

SPECIAL TOPICS IN COMPUTING AND ICT RESEARCH

Strengthening the Role of ICT in Development

Editors

Janet Aisbett
Greg Gibbon
Anthony J. Rodrigues
Joseph Kizza Migga
Ravi Nath
Gerald R Renardel

Volume IV

FOUNTAIN PUBLISHERS
Kampala

Fountain Publishers
P.O. Box 488
Kampala
E-mail: sales@fountainpublishers.co.ug;
publishing@fountainpublishers.co.ug
Website: www.fountainpublishers.co.ug

Distributed in Europe and Commonwealth countries outside Africa by:
African Books Collective, P.O. Box 721, Oxford OX1 9EN, UK.
Tel: 44(0) 1869 349110, Fax: +44(0)1869 349110
E-mail: orders@africanbookscollective.com
Website: www.africanbookscollective.com

Distributed in North America by:
Michigan State University Press
1405 South Harrison Road
25 Manly Miles Building
East Lansing, MI 48823-5245
Tel: +1 517 3559543, 800 6782120
Fax: +1 517 4322611
E-mail: msupress@msu.edu
Website: www.msupress.msu.edu

© Makerere University 2008

All rights reserved. No part of this publication may be reprinted or reproduced or utilised in any form or by any means, electronic, mechanical or other means now known or hereafter invented, including copying and recording, or in any information storage or retrieval system, without permission in writing from the publishers.

ISBN 978-9970-02-871-2

Contents

<i>Preface</i>	<i>vii</i>
<i>Keynote speaker biographies</i>	<i>viii</i>
<i>Author Biographies</i>	<i>xii</i>
<i>Acknowledgements</i>	<i>xviii</i>

Part 1: Computer Science

1. The Status of Software Usability in Uganda <i>Florence Tushabe, Venansius Baryamureeba, Paul Bagyenda, Cyprian Ogwang and Peter Jehopio</i>	1
2. Game Theoretic Multi-agent Systems Scheduler for Parallel Machines <i>Elisha T. O. Opiyo, Erick Ayienga, Katherine Getao, William Okello-Odongo, Bernard Manderick and Ann Nowé</i>	12
3. Dynamic Resource Allocation: A Comparison of First Price Sealed Bid and Vickrey Auctions <i>Elisha T. O. Opiyo, Erick Ayienga, Katherine Getao, William Okello-Odongo, Bernard Manderick, Ann Nowé</i>	21
4. Bootstrapping Machine Translation for the Language Pair English–Kiswahili <i>Guy De Pauw, Peter Waiganjo, Wagacha Gilles and Maurice De Schryver</i>	30
5. Towards Domain Independent Named Entity Recognition <i>Fredrick Edward Kitoogo, Venansius Baryamureeba & Guy De Pauw</i>	38
6. Towards Full Comprehension of Swahili Natural Language Statements for Database Querying <i>Lawrence Muchemi</i>	50
7. Generalized Association Rule Mining Using Genetic Algorithms <i>Peter P. Wakabi-Waiswa, Venansius Baryamureeba, K. Sarukesi</i>	59
8. Collaborative Filtering: A Comparison of Graph-based Semi- supervised Learning Methods and Memory-based Methods <i>Rasna R. Walia</i>	70

Part 2: Information Systems

9. A Simulation Modelling-based Investigation of the Impact of IT Infrastructure on Business Process Reengineering <i>Innocent Baguma, and Joseph K. Ssewanyana</i>	85
---	----

10.	A Flexible Approach for user Evaluation of Biomedical Ontologies <i>Gilbert Maiga and Ddembe Williams</i>	100
11.	Facilitating Grass-root NGO’s Adoption of Information Systems and Information Communication Technologies <i>Kathy Lynch, Stacey Lynch and Festus Bazaira</i>	113
12.	Career Guidance Using Expert System Approach <i>Ojenge Winston and Muchemi Lawrence</i>	123
13.	Exploring the Implementation of Blended Learning in a Developing Country: A Case Study of Uganda <i>Philip O. Ayoo and Jude T. Lubega</i>	132
14.	Use of “Flight Simulator” for Training IT Managers in Investment Decision Making <i>Paul Ssemaluulu and Ddembe Williams</i>	144
15.	System Dynamics Modeling in Healthcare: The Ugandan Immunisation System <i>Agnes Semwanga Rwashana and Ddembe Wileese Williams</i>	160
16.	Geometrical Spatial Integration Model for Geo-information Management <i>Ismail Wadembere and Patrick J. Ogao</i>	177

Part 3: ICT Sustaible Development

17.	The Role of ICT in Creation and Sustenance of Trust in SME Clusters and Inter-organizational Systems <i>Agnes Mindila, Dorothy McCormick, Anthony Rodrigues and Ronald Waweru Mwangi</i>	194
18.	Data Security Lapses in Developed Societies: Public and Private Sector Cases <i>Anthony J Rodrigues</i>	203
19.	ICT-enabled Services: A Critical Analysis of the Opportunities and Challenges in Uganda <i>Venansius Baryamureeba</i>	215
20.	The localisation into isiXhosa of the iLanga Telephone System <i>Lorenzo Dalvit, Mosiuoa Tsietsi, Alfredo Terzoli, Pamela Maseko Msindisi Sam and Thandeka Mapi</i>	236

21.	Agro-meteorological Knowledge Management System for Small Scale Farmers <i>Omondi Lwande and Muchemi Lawrence</i>	243
22.	Impact of Funded Research Projects on University Quality: A System Dynamics Approach <i>Benedict Oyo, Ddembe Williams and Erik Barendsen</i>	256
23.	Tertiary ICT Education in the East African Community (EAC): Human Capital and Curriculum Development <i>Rodrigues J. Anthony and Wafula J. Muliaro</i>	273

Part 4: Data Communication and Computer Networks

24.	Using Mobile Technology & Services for Sustainable Development in Africa <i>Fisseha Mekuria</i>	290
25.	Implementing Security in Wireless Sensor Networks <i>Joseph Migga Kizza</i>	296
26.	Delay-optimized Path in Peer-to-Peer Networks <i>Francis Lowu and Venansius Baryamureeba</i>	312
27.	Service Oriented Telecom Architectures for Enabling Innovative Mobile Services <i>F. Mekuria and W. Mwebaza</i>	321

Part 5: Information Technology

28.	Comparison of Rwandan and Russian ICT Educational Policies <i>Irina Zlotnikova and Voronezh State University of Education (Russia)</i>	332
29.	A Generic E-Learning Framework: A Case Study Faculty of Computing and Information Technology (CIT) Makerere University <i>Jude T. Lubega and Fiona Ssozi Mugarura</i>	340
30.	Integrating E-Supervision in Higher Educational Learning <i>Jude T. Lubega and Michael Niyitegeka</i>	351
31.	Mobile Research Supervision Initiative (MRSI) at Makerere University: Lessons to Learn <i>Paul B. Muyinda, Kathy Lynch and Jude T. Lubega</i>	359

32.	Assessing Implementation of Blended learning in Constrained Low Bandwidth Environment <i>Nazir Ahmad Subail</i>	377
33.	Information Communication Technologies Acceptance and use Among University Communities in Uganda: A Model for Hybrid Library Services end-users <i>Prisca K. G. Tibenderana and Patrick J. Ogao</i>	391
34.	Value Difference Metric for Student Knowledge Level initialization in a Learner Model-based Adaptive e-Learning System <i>Robert O. Oboko, Peter W. Wagacha, Euphraith M. Masinde, Elijah Omwenga and Arno Libotton</i>	411

Preface

This is the fourth conference in the SREC series. While the original theme of ICT for sustainable development is still strong, for the last two years the conference has been held as the International Conference on Computing and ICT Research (ICCIR). This name reflects the broad nature of ICT research that, in the end, contributes to sustainable development.

A brief look through these proceedings will convince you of the lively mix of applied and basic research across the disciplines. Keynote speakers and contributing authors alike address both generic developments and applications of state-of-the-art research to specific African problems.

Thus, in the Computer Science stream, basic research contributions range across semi-supervised learning, dynamic resource allocation, automatic auction mechanisms, and more. In the Information Systems stream, researchers look at appropriate selection of biomedical ontologies and at geometric integration in GIS. In Data Communications and Computer Networks, basic research ranges from network path optimisation, through telecommunication architectures for special mobile services, to various wireless security issues.

On the applied research side, we see investigations into ICT adoption amongst grassroots NGOs; library users; and stakeholders in blended learning. There are studies into applications of ICT to support interaction amongst SME clusters, between lecturers and students, and amongst stakeholders in small scale farming. There are simulation models, for example of the Ugandan immunisation systems and of IT effects on business processes; and work on developing interfaces in local languages. We are taken by key note speakers into innovative centres for IT and Development, and for Language Technology, and made to think in the context of African development about ICT educational policies, mobile services and more. There is a lot of exciting research happening.

The Faculty of Computing and Information Technology at Makerere University again hosts ICCIR, with the generous support of the Netherlands Organization for International Cooperation in Higher Education. The conference starts off with PhD colloquium sessions which give trainee researchers the valuable opportunity to expose their research to kindly but searching critique. The main conference has 11 key note speakers and the 28 contributed papers that each survived review by two independent reviewers.

Sincere thanks are due to those reviewers, and to all those on the organising and technical committees who worked hard to bring together the many strands needed to create a successful conference. Special thanks go to those reviewers who provided detailed helpful feedback to the authors, because this both strengthens the conference and helps the researchers, whether or not their paper is accepted. As in previous years, the publishing team at Fountain Publishers, Kampala, were professional and prompt, and the Conference Secretariat worked hard and long. The biggest thanks, though, go to Jude Lubega who, as Acting Dean of Research, has been tireless in his efforts to keep everything on track.

Janet Aisbett

Keynote Speaker Biographies

Venansius BARYAMUREEBA holds a PhD in Computer Science and is a Professor of Computer Science. He currently serves as the Dean of the Faculty of Computing and IT at Makerere University, Uganda. He is a member of Senate and Council at Makerere University. He is also a member of Senate at Busitema University. He is the Chairman and Managing Director of ICT Consults Ltd, one of the prestigious ICT consultancy firms in Africa. He is also currently involved in initiatives of growing and sustaining ICT/ computing human capital in Sub-Saharan Africa region. In 2007 Prof. Baryamureeba received the ‘Top ICT Educator/ Academic’ Award in Africa for 2007 and was appointed Judge for the African ICT Achievers Awards 2008. Professor Baryamureeba has been considered for inclusion in the upcoming 2009 Edition of “Who’s Who in the World”.

Arvi HURSKAINEN was born in Finland. After graduating from the University of Helsinki he worked for nine years in various parts of Tanzania in teaching and research tasks. This work gave him an opportunity to learn Swahili to the extent that in later years it came to play a major role in his career. Since 1980 he has been working at the University of Helsinki, with language technology as major teaching and research duty. Over the years, in cooperation with the University of Dar-es-Salaam, he has carried out field-work in Swahili-speaking areas. As an outcome of this work, there are two language resources of Swahili accessible in the Internet, i.e. Computer Archives of Swahili Language and Folklore that contains speech with transcription, and Helsinki Corpus of Swahili with 12 million analyzed words. Prof. Hurskainen has worked on various topics of language technology, including morphology, syntax, disambiguation, machine translation of text, information retrieval, and automatic dictionary compilation. All these components are included in the comprehensive system termed Swahili Language Manager. An early outcome of this work is the spelling checker of Swahili, currently included in the MS Office 2007 suite. He has also been the Editor of Nordic Journal of African Studies in 1992-2006

Joseph Migga KIZZA received a BSc in Mathematics and Computer Science from Makerere University, Kampala, Uganda, a MSc in Computer Science from California State University, USA, an M.A. in Mathematics from the University of Toledo, Ohio, USA and a PhD in Computer Science from the University of Nebraska-Lincoln, Nebraska, USA. Dr Kizza is currently a professor of Computer Science and Director of the Center for Information Security and Assurance at the University of Tennessee - Chattanooga, Tennessee, USA. His research interests are in Social Computing and Computer Network Security in which he has so far published numerous conference and journal articles and more than seven books. He was appointed a UNESCO expert in Information Technology in 1994. He is the chief editor for the Journal of Computing and Information Research.

Victor W. MBARIKA is on faculty in the College of Business at Southern University and A&M College and is founding director of the University's International Center for Information Technology and Development. He is an information and communications technology (ICT) consultant with various governmental and private agencies. He holds a B.Sc. in Management Information Systems from the U.S. International University, a M.S. in MIS from the University of Illinois at Chicago, and a PhD in MIS from Auburn University. With over 100 academic publications, Dr. Mbarika has authored three books and his research on ICT diffusion in developing countries and on multimedia learning has appeared (or is forthcoming) in 35 peer reviewed journals, 5 book chapters and over 60 national and international conferences publications. His scholarly publications have appeared in journals such as IEEE Transactions, IEEE Spectrum, Communications of the ACM, and the Journal of the Association for Information Systems. His work has also appeared in major news outlets such as Reuters and The Associated Press (AP). He has received several research and teaching grants and fellowships from the National Science Foundation, NASA, KPMG, Southern Regional Education Board and the Louisiana Board of Regents. He is Editor-in-Chief of The African Journal of Information Systems and he serves as a senior board member of several other journals such as IEEE Transactions on IT in Biomedicine. Dr. Mbarika is Founder of the Cameroon Computer and Network Center (CCNC). He is a member of the Association of Information Systems (AIS), the Institute of Electrical and Electronics Engineers (IEEE), and the Information Resources Management Association (IRMA). He holds over 15 research, teaching and service excellence awards.

Fisseha MEKURIA is presently a Visiting Professor of Wireless Communications Systems at the Faculty of Computing & IT, Makerere University, Uganda. He received a PhD from Linkoping University, Sweden, in 1993, where he worked for several years as postdoctoral Research Fellow. He has over 20 years experience in ICT and telecommunications research, education and industry R&D. He has worked as senior research scientist at Ericsson Mobile Communications R&D center in Sweden. During this time he developed 12 granted US & EUPO patents in the areas of wireless and mobile communications. He has designed a curriculum for Mobile Computing and Software Development. Prof. Mekuria's leading role in establishing a College of Telecommunications & IT in Addis Ababa, and the paper "Educating the Architects of the Mobile Economy", earned him a nomination for the prestigious IEEE Continuing Education Award. His research interests are Next Generation Wireless & Mobile Communications Systems and Services, and ICT for sustainable development.

Ravi NATH is the director of the Joe Ricketts Center and the holder of the Jack and Joan McGraw Endowed Chair in Information Technology Management in the College of Business Administration at Creighton University. Previously, he held the positions of the Associate Dean of Graduate Programs at Creighton University. At the University of Memphis he served as the Associate Dean for Academic Programs

and Director of the PhD program in Management Information Systems. Dr. Nath has a Masters degree from Wichita State University and a PhD degree from Texas Tech University. He spent one year lecturing and conducting research in Zimbabwe as a Fulbright Scholar. Dr. Nath has published nearly 100 research papers in the area of Electronic Commerce and Information Systems in various national and international journals. He is the Past-President of the Southwest Decision Sciences Institute. Also, he has served on the Board of Directors of Decision Sciences Institute and was a Vice-President of the Institute. Dr. Nath received the University of Memphis' Distinguished Teaching Award in 1995. He is also a recipient of the Southwest Decision Sciences Institute's 1996 Distinguished Service Award. Currently, Dr. Nath serves on the Board of Directors of several for-profit and non-profit organizations.

Dilip PATEL holds the chair of Information Systems at London South Bank University. He is the head of the Centre for Information Management and E-Business, which consist of four research groups: E-Business and The Digital Society, Health Informatics, Information Management and Modelling and Knowledge Based Systems. Professor Dilip Patel is currently on the editorial board for two international journals: The International Journal of Cognitive Informatics and Natural Intelligence (IJCiNi) and the International Journal of Information Technology and Web Engineering. He is also Editor-in- Chief of The International Journal of Computing and ICT Research.

Vir V. PHOHA is a professor of Computer Science in the College of Engineering and Science at Louisiana Tech University, US. He holds a W. W. Chew Endowed Professorship at Louisiana Tech and is directing the Center for Secure Cyberspace. Dr. Phoha holds an M.S. and a PhD in Computer Science from Texas Tech University. Professor Phoha has done fundamental and applied work in anomaly detection in network systems, in particular in the detection of rare events. He has eight patent applications and many reports of inventions. He is author of over 90 publications and author/editor of three books: (1) Internet Security Dictionary, Springer-Verlag (2002) and (2) Foundations of Wavelet Networks and Applications, CRC Press/Chapman Hall (2002), and Quantitative Measure for Discrete Event Supervisory Control, Springer (2005). He directs the Anomaly Detection and Mitigation Laboratory at Louisiana Tech University. He has won various distinctions, including outstanding research faculty and faculty circle of excellence award at Northeastern State University, Oklahoma, and as a student was awarded the President's Gold Medal for academic distinction.

Anthony J. RODRIGUES is Professor of Computer Science at the School of Computing and Informatics, University of Nairobi. Research interests include Scientific Computing, Approximation Theory, Error Analysis and Systems Modeling. He has also been involved in the design and development of sustainable

integrated management information systems at various universities and is currently studying the developmental impact of various ICT policies and strategies (or the lack thereof) in the region.

Irina Ya. ZLOTNIKOVA holds a PhD in Theory and Methodology of Computer Science Education (Doctor of Pedagogical Sciences, Moscow, Russia, 2005), a PhD in Solid-State Electronics, Nano- and Microelectronics (Candidate of Technical Sciences, Voronezh, Russia, 1995) and a specialist degree in Radiophysics and Electronics (Voronezh, Russia, 1988). She has been employed by the Voronezh State University of Education since 1991, since 2007 as a full Professor in the Department of Information Technology in Education. In 2005-2006, during a one-year sabbatical, she worked in Kigali Institute of Science and Technology (Kigali, Rwanda). She is joining the Information Technology Department, Makerere University, as a Professor in August 2008. Professor Zlotnikova has taught a variety of undergraduate and post-graduate computer-based courses and has experience teaching in distance mode with the Voronezh Institute of Distant Education and the African Virtual University. Professor Zlotnikova is an expert in the areas of Software Engineering for Educational Purposes, including Management, Development of Educational Web-resources, Theory and Methodology of Computer Science Education, Computer-Based Distance Learning and E-learning, and International ICT Policy. She has more than 70 professional publications, including 3 monographs. She has supervised Masters and Doctoral Dissertations in the above-mentioned areas.

Author Biographies

Eric AYINGA is a member of academic staff at the School of Computing and Informatics, University of Nairobi. His interests include wireless networks, network quality of service assurance, grid computing, multi-agent systems, game theory and reinforcement learning.

Philip AYO is the Electronic Database Specialist at the Inter-University Council for East Africa (IUCEA), a regional organization that coordinates university education in East Africa. Here, he oversees the development of appropriate ICT capacity at the IUCEA and member universities. He holds a Master of Philosophy degree in Information Science and a Bachelor of Science in Information Technology. Currently, he is pursuing a PhD degree in Information Systems at Makerere University, Uganda. His research interests are in the areas of online learning and ICT policy development, where he has published and presented papers in conferences. Before joining the IUCEA he was teaching at the Department of Information and Media Technology, Moi University. He is a senior member of the Uganda Computer Society.

Erik BARENSEN is a senior lecturer and director of education in the Institute of Computing and Information Sciences, Radboud University Nijmegen, the Netherlands. He holds a PhD in mathematics and computing science. His research interests include foundations of computing (lambda calculus, rewriting, logic, type theory), static analysis, methodology, and didactics of science. He is involved in several local and national activities concerning curriculum development, quality assurance, teacher education, secondary education, and seminars for education executives.

Festus BAZIRA founded VOLSET 1997 and has been its Director since this time. Festus was educated in Uganda and Kenya, and has qualifications and experience in teaching, community health, first-aid, and drug management. He is responsible for seeking funding, the every-day operation of VOLSET, and the many volunteers that spend time with VOLSET.

Lorenzo DALVIT is a lecturer in ICT Education at Rhodes University and the Research and ICT coordinator of the SANTED programme within the Rhodes School of Languages (African Studies Section). He has worked for several years as a researcher on multilingualism and ICT for the Telkom Centres of Excellence of Rhodes and Fort Hare Universities in South Africa. Dalvit's areas of academic interest are ICT for development and localisation into African languages. As part of his multi-disciplinary interests, he is currently completing a PhD in Education. He has published in academic journals and presented at conferences within various disciplines both in South Africa and internationally. He is active in post-graduate supervision and in establishing collaborations with universities in other parts of Africa.

Guy DE PAUW is a senior researcher at the CNTS – Language Technology Group (University of Antwerp, Belgium) and a visiting lecturer at the School of Computing and Informatics (University of Nairobi). His research focuses on the application of machine-learning techniques to natural language processing (NLP) for African languages. A recurring theme in his research is how we can port current state-of-the-art NLP techniques that are developed with Indo-European languages in mind, to Bantu languages. Recent publications concentrate on morphological and syntactic processing of Swahili and the development of resources for and components of a machine translation system for the language pair English – Swahili. He is also one of the founders of the AfLaT organization (aflat.org), which aims to bring together researchers on language technology for African languages.

Katherine GETAO is the former Director of the School of Computing and Informatics, University of Nairobi. Her interests include artificial intelligence and natural language applications.

Fredrick KITOOGO is a holder of a Bachelor of Statistics degree from Makerere University, Uganda and a MSc in Computer Science degree from the National University of Science and Technology (NUST), Bulawayo, Zimbabwe. Kitoogo is a PhD student in the Department of Computer Science, Faculty of Computing and IT, Makerere University and he is currently writing his thesis on Improved Use of Machine Learning Techniques in Named Entity Recognition. His main areas of research interest are Machine Learning, Natural Language Processing and Business Intelligence & Data Warehousing.

Kathy LYNCH is an Associate Professor in ICT Research and Development at the University of the Sunshine Coast, Australia. She has a PhD in Education and a Masters in Business Systems from Monash University, Australia. Her tertiary teaching background is primarily in Information Systems. Kathy is a Research Associate at Makerere University, Faculty of Computing and Information Technology, and the Editor-in-Chief of the International Journal of Information, Knowledge and Management. Her research interests revolve around the intersection of applied information technology and society; in particular mobile computing, information systems, emerging interfaces and education.

Stacey LYNCH has a Bachelor Science Honours (Microbiology) from Monash University, Australia. She worked as a research scientist for Pfizer Animal Health, Melbourne within the research and development group for animal viral vaccines. In 2007 Stacey commenced a PhD at Melbourne University, Veterinary Science, looking at the pathogenesis and vaccine development for a horse viral pathogen. Stacey has a broad interest in virology and has experience in data collection, data systems and working within GxP guidelines.

Gilbert MAIGA holds the following academic qualifications: MSc Computer Science (Makerere University), MSc Animal Physiology (UON), Bachelor of Veterinary Science (Makerere University), and Diploma in Computer Systems

Design (WLC), Certificate in Computer Programming (WLC). Networking Basics Certificate (CISCO). He is a PhD student in the Department of Information Systems, Faculty of Computing and IT, Makerere University in August 2005, where he is also an assistant lecturer. His MSc research concerned information systems integration in HIV-AIDS management, and his current area of research concerns integration of clinical and biological information in bioinformatics.

Bernard MANDERICK is the Co-Director of Computational and Modeling Laboratory (COMO), Free University of Brussels, Belgium. His interests include artificial intelligence, game theory, machine learning and evolutionary computational modeling.

Lawrence MUCHEMI is a lecturer at the University of Nairobi in the School of Computing and Informatics. He has previously worked at Jomo Kenyatta University of Agriculture and Technology and was head of the Computer Science Department at Africa Nazarene University. Muchemi graduated from Moi University in 1994 with a Bachelors degree in Engineering and later with a Master of Philosophy specializing in Artificial Intelligence systems. He has maintained long life interest in the research and development of systems employing engineering and computer science theories and practice. Of particular interest is natural language processing for African Languages. He has conducted research and published in Natural Language Queries, Natural Language Synthesis, Expert Systems and Information Systems. He is currently working towards a PhD degree in Computer Science on Natural Language Queries for Kiswahili.

Paul MUYINDA is lecturer of Information Technology in the Department of Distance Education, Institute of Adult and Continuing Education, Makerere University. He is also charged with the responsibility of pedagogically integrating ICTs into the distance learner student support system. He holds a Bachelor in Statistics from Makerere University obtained in 1995 and a Master of Computer Science and Applications from Shanghai University, People's Republic of China, obtained in 1998. Muyinda is a Certified Netware Administrator (CNA). His Master's Dissertation was titled "A National Passport IntranetWare: A Case of Uganda National Passport Office". He is currently a PhD (Information Systems) student at the Faculty of Computing and IT, Makerere University with research interest in e-learning. His thesis is titled "A Framework for Instantiating Learning Objects Applications for Mobile and Fixed Communication Systems".

Michael NIYITEGEKA is an Assistant Lecturer in the Department of Information Systems, Faculty of Computing & IT, Makerere University. He is specifically interested in Technology Management as a research area and is currently working on his PhD proposal in the area of ICT and SME business development.

Ann NOWÉ is the Co-Direcrtor of Computational and Modeling Laboratory (COMO), Free University of Brussels, Belgium. Her interests include artificial intelligence, machine learning and multi-agent systems.

Patrick OGAO has been involved in visualization research since 1996. He has a PhD (Utrecht University, the Netherlands) and MSc (ITC Enschede, The Netherlands) focusing on visualization in information systems and dynamic geo-phenomena. Prior to joining Makerere University, he worked with the research group in scientific visualization and computer graphics in the department of computing and mathematics of the University of Groningen, The Netherlands. He is currently a visiting senior lecturer and acting Head of Information Systems Department in the Faculty of Computing and Information Technology, at Makerere University. He is the Long Term Expert (on secondment by the University of Groningen) in a NUFFIC/NPT project in ICT training capacity building in four public universities in Uganda. His research interests are in visualization applications in bio-informatics, geo-informatics, and software engineering and in developing strategies for developing countries

Winston OJENGE holds a Bachelor of Technology Education (Electrical Technology) from Moi University. He lectured at Bumbe Technical Training Institute where he rose to Acting Principal, before joining the Kenya Polytechnic as a Senior Lecturer in Telecommunications and Control systems in 2005. His MSc specialization at the University of Nairobi, School of Computing and Informatics, is Artificial Intelligence. His research interest is in the application of Machine Learning algorithms to Remote Sensing, in particular to crop yield prediction. He is currently working on the PhD proposal.

William OKELLO-ODONGO is the Director of the School of Computing and Informatics, University of Nairobi. His interests include distributed systems and high performance computing.

Elisha T. O. OPIYO is a member of academic staff at the School of Computing and Informatics, University of Nairobi. His interests include scheduling, resource allocation, grid computing, multi-agent systems, game theory and reinforcement learning.

Benedict OYO is a PhD student in Information Systems at Makerere University, Faculty of Computing and Information Technology, Uganda. His research area is Decision Support Models for Quality Management in Higher Education using System Dynamics Modelling and Action Research. Benedict has an MSc in Computer Science from Makerere University. His current research is jointly supervised by the Faculty of Computing and Information Technology, Makerere University, and the Institute for Computing and Information Sciences, Radboud University Nijmegen

Agnes RWASHANA Semwanga is an Assistant Lecturer in the Information Systems Department, Faculty of Computing and Information Technology, Makerere University. She is currently undertaking PhD research (final year) and her main research interests include Simulation and Modeling, System Dynamics and Health Informatics.

Paul SSEMALUULU is a PhD student in the Department of Information Systems, Faculty of Computing and IT, Makerere University. His main area of research is the Application of System Dynamics Modeling to IT Investment Evaluation. Ssemaluulu holds a FTC in Telecommunications from Uganda Polytechnic, Kyambogo; a BBA and MSc Computer Science Degree from Makerere University. Ssemaluulu is employed in the Faculty of Computing and IT, Makerere University as a Research Coordinator. He is also employed as part-time lecturer in the Department of Information Technology. He has a strong technical and management background, having trained with such international companies like Phillips, AT&T, ITS-France, Nepostel and Fujitsu.

Nazir Ahmad SUHAIL is a PhD (Information Technology) student at the Faculty of Computing and Information Technology, Makerere University, Kampala-Uganda in August 2005. He is holder of BSc, MSc (Mathematics) degrees from The University of The Punjab, Lahore-Pakistan and a PGD(Computer Science) from Makerere University. He is also lecturing at Kampala University in the Departments of Computer Science and Mathematics and has previously taught at the Department of Mathematics Makerere University.

Prisca TIBENDERANA is a part time PhD student in Information Systems at the Faculty of Computing and Information Technology, Makerere University, Kampala. The study is funded by Innovation at Makerere Committee (i@mak.com). Since 2007, she has been a Deputy Librarian at Kampala International University. She started her career studying librarianship in Nigeria and was previously a Deputy Librarian in Charge of Book Bank System at Makerere University. As Senior Librarian and Head of Periodical Section, she introduced online journals, spearheaded and coordinated the Programme for the Enhancement of Research Information (PERI) in Uganda. An advocate of digital technologies in libraries, she has written a number of articles on the subject.

Ismail WADEMBERE holds a BSc Surveying degree from Makerere University and a MSc in planning - Geographical Information Systems (GIS) degree from the University of Science, Malaysia. Wadembere is a PhD student in the Department of Information Systems, Faculty of Computing and IT, Makerere University. His main areas of research interest are GIS, Service-Oriented developments and Online Information Systems. Wadembere is employed in Kyambogo University, where he handles GIS, Computing for Surveying and Mapping.

Joseph Muliaro WAFULA holds a BSc (Kenyatta University), M.Sc (University of Nairobi) and M.Phil (University of Cambridge UK). He is currently a lecturer at the Institute of Computer Science and Information Technology at Jomo Kenyatta University of Agriculture and Technology, Kenya and a Ph.D student researching ICT policies and e-strategies.

Rasna R. WALIA is a faculty member at the School of Computing and Informatics, University of Nairobi, where she is involved in research and teaching of both undergraduate and graduate students of Computer Science and Bioinformatics. She holds an MSc degree in Intelligent Systems from University College London, UK. Her project dealt with the field of Collaborative Filtering and the application of methods such as Graph-based Semi-Supervised Learning and Memory-based methods to the Collaborative Filtering Domain. She also holds a Bachelors degree in Computer Science from the University of Nairobi, Kenya. Her areas of interest include Bioinformatics and Machine Learning. She is currently supervising two final year undergraduate projects in the area of Genetic Algorithms and Ant Colony Optimization.

Peter Patrick WAKABI Waiswa earned a Bachelor's degree in Statistics and Applied Economics and a Postgraduate Diploma in Computer Science from Makerere University and he holds a MSc. in Computer Science from National University of Science & Technology from the National University of Science and Technology (Zimbabwe). He is also a Certified Network Administrator (CNA). Wakabi is a PhD student in the Department of Computer Science, Faculty of Computing and IT, Makerere University and he is currently writing his thesis on Association Rule Mining Using Evolutionary Computing. His main areas of research interest are Evolutionary Computing, Algorithms and Complexity, Artificial Intelligence, Data Mining and Knowledge Discovery. Peter is currently employed by Uganda Electricity Transmission Company Limited as the Head of IT where his main duty is the strategic positioning of the company using ICTs. He previously worked as a lecturer at ICS Makerere University and Uganda Christian University; as Commonwealth IT Technical Adviser to the NCC in Tanzania; and as an IT Consultant with PricewaterhouseCoopers, among other positions.

Acknowledgements

The chief editor would like to recognize and give special thanks to the following:

- The Conference Chair and the Chairpersons of the conference technical committees;
- The different researchers around the world who reviewed the papers anonymously;
- Members of the conference organising committee and conference technical committees;
- The Conference Secretariat for supporting the conference activities and putting together this book;
- Makerere University Management;
- The Netherlands Organization for International Cooperation in Higher Education for providing financial support for the Conference and publication of this book; and
- Fountain Publishers, for a job well done.

We are above all grateful to those who submitted their papers for peer review.

Editors

Janet Aisbett

Greg Gibbon

Anthony J. Rodrigues

Joseph Kizza Migga

Ravi Nath

Gerald R Renardel

PART 1



Computer Science

1

The Status of Software Usability in Uganda

Florence Tushabe, Venansius Baryamureeba,
Paul Bagyenda, Cyprian Ogwang and Peter Jehopio

This research investigates the current software usability practices in developing countries, particularly in Uganda. It documents the contemporary user needs, attitudes and challenges that are being experienced in regard to software usage and adoption. The performance of software in circulation is judged based upon indicators like efficiency, effectiveness and ease of use. In addition, local and internationally produced software are compared in terms of functional performance and user satisfaction. This study proposes some recommendations about how to further improve software usability within the present circumstances, as well as some predictions about the future trends.

1. Introduction

Previous studies showed that as much as 68% of software acquired by government bodies of developing countries remained either unused or unsuccessfully utilized [Gib 1998]. This is partially caused by indigenous software developers not being well equipped to transform their innovative ideas into usable products, capable of meeting the local demands. Software usability plays a critical part of any countries growth because software is the central nervous system of Information and Communication Technology.

Software usability is the quality of being convenient and of practical use [Merriam-Webster 2002] or a measure of the quality of a user's experience with a product or system [Nielsen 1993]. According to the International Standards Organisation [ISO/IEC TR 9126-2 2003], usability is the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a given context of use.

Usability is often ignored within the software development process and this results in the production of software that does not address the real needs of the user community [Centennial Software 2005; Bhagwati et. al. 2004; Mankiw and Swagel 2006]. The developed software therefore remains less relevant to society and so does the corresponding economic stimulus. This can partly be overcome by involving users within the software development process [Ogwang, 2006].

This research highlights the reasons why software usability remains relatively low in developing countries. The goal is to identify what can be put in place to promote production of locally designed and socially relevant products demanded

for by the private and public sectors. It involved going to the software end-users to inquire about their software experiences, practices and needs. The findings, which are documented in this report, will be useful to the trainers / mentors of software education and incubation centers especially during the process of updating the curriculum to include indigenous solutions to the revealed problems.

This article is organised as follows: Section 2 summarizes the previous software usability practices in developing countries, Section 3 and 4 briefly describe the purpose and description of this survey. The results obtained from the investigation are given in Section 5 while Sections 6 and 7 propose some recommendations and concluding remarks.

2. Software Usability in Developing Countries

Developing countries first became consumers of software in the 1950s and 1960s [Heeks 1998]. To a lesser degree they also started producing software around this time. However, most of their national policies and the investment strategies have been overly-oriented to the hardware of information and communications. This has been a setback to software usability.

Recent studies have shown that the local software-industries in most low-income developing countries are dominated by subsidiaries of large multi-national computer/consultancy companies [Duncombe and Heeks 2001]. The developers consist of mostly one- and two-person software firms with low turnover [Heeks 1998]. Their work ranges from developing custom-built software that meets the needs of users in the small but growing market of smaller enterprises, to customizing existing software packages or simply trading imported software packages. The four main approaches to software development practiced in developing countries in the order of priority are: custom-built, re-engineered, customisation, and off-the-shelf packages [Duncombe and Heeks 2001].

A study conducted in 2001 [Duncombe and Heeks 2001] showed that computer applications that are being developed by local software firms/individuals in developing countries are of poor quality and produced very slowly at a high cost. They are usually characterised by a low degree of integration and effectiveness, the consumption is very limited, very undemanding, and very much unrelated to any kind of market mechanism. Moreover they are heavily skewed towards certain sectors like finance management, military, and criminal investigations.

Sadly, most of their software applications and systems remain essentially in demonstration form only, without being tested and validated in a real user environment [Ogwang, 2006]. This is attributed to the lack of market demand for what is being produced. Potential user organisations either do not perceive a need for the particular application or have a preference for manual rather than computerised systems. The technical education and extension institutions are weak while local supply capabilities and access to international know-how restricted. All these constraints, and more, mean that software diffusion into the local community is low.

This study will review the current status of software usability in a developing country. It will report the changes that have emerged over time especially in regard to software integration in the work environment, given from the users' point of view.

3. Purpose of the Survey

The major aim of the study was to gather information about software usage in Uganda. It sought to answer the following questions:

- a. How optimally is a given software being used? Which tasks are users opting to perform manually instead of by computer and why?
- b. What are the current software needs of the public/private sector?
- c. What type of software is most common or lacking in the market?
- d. How is the locally developed software performing in comparison to imported ones? What are their strengths and weaknesses?
- e. What are the most popular factors that are considered during procurement/acquisition of software? And why?
- f. To what extent has adoption of software impacted job performance in Uganda?
- g. What are the challenges that software users are facing?

4. Description of the Survey

The survey targeted users of customized and job-specific software. This excluded users of general non domain-specific packages like MS Word, Excel etc. The respondents included the end users (operators), managers and stakeholders of businesses/organisations. The geographical location was Kampala city whose specific area is enclosed in the pink boundary shown in Figure 1, and Entebbe city, particularly all major offices found on portal, circular and airport roads up to Entebbe International airport.

co-operation from some organisations like Bank of Baroda and embassies who said that they do not participate in surveys.

4.3 Data Analysis

Descriptive statistics and logistic regression were moved to analyse quantitative data in order to establish current status and likely future trend of software usage in Uganda. Qualitative data were analysed by applying coding and expert opinion using MS Excel and SPSS packages for data entry and analysis.

5. Results

The results presented below represent the views of respondents sampled from government bodies (27%), local private organizations (52%), international private organizations (17%), diplomatic missions (1%) and others (3%). The majority of respondents turned out to be working in administration, accounts and the operations departments of their organizations. Responses from companies dealing in a variety of fields like legal services, education, IT firms, tourism etc are documented here. We believe that this is a fair representation of the total population of domain-specific software users in Uganda and here are our findings:

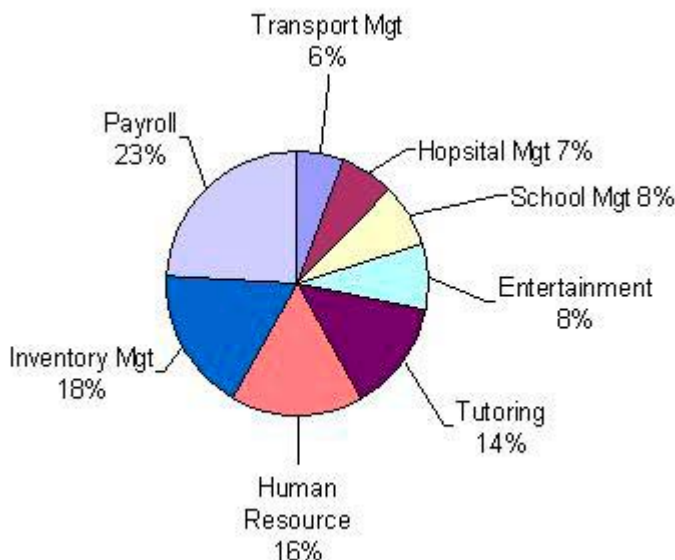
1. Most Ugandans are happy to have adopted the use of software in their organisations. 80% of respondents said their expectations were fully met through the adoption of computers in their organizations.
2. Computer usage is dominated by males. The respondents reported that an average of 73% of computer users in their organisation is male and 27% female.
3. Software acquired by organizations in Uganda are mostly custom built (48%) then freeware (27%) or re-engineered (25%). Of these, 62% are proprietary and 38% open source.
4. There is a good return-on-investment for software users. All respondents said that adoption of software registered only positive impact on their work in terms of performance improvement and profits. Their experience shows that software speeds up tasks by an average of 74%, improves quality by 72% and quantity of output by 66%. It reduces errors by 64%, increases customer satisfaction by 61%, employee morale by 54% and finally reduces costs (or increases profits) by 52%.
5. Most users of customized software enjoy working with it. Over 70% said that they find the software pleasant to use, with a nice visual appearance and responding the way they expect. 72% said that accomplishing ordinary tasks can be very fast by someone who has learnt the basics of the software while 65% found error recovery easier. 56% reported that the software is usually the one in control rather than the user.
6. Only 12% of customized software in organizations is locally developed while 59% is imported and 29% is freeware / open source. 27% said that they use pirated software.

Table 1: Local Vs International performance

Source of Software	Employee Morale	Error Reduction	Customer Satisfaction	Quantity Increase
Local	56	59	65	61
International	57	66	66	69
	Cost Reduction	Speed	Profit Increase	Quality Improvement
Local	51	71	47	68
International	57	77	59	74

- Local and internationally produced software are performing similarly. The general trend is that locally developed software performs slightly worse but this becomes insignificant when sample space imbalances are factored in. Table 1 shows the average percentage improvement of the indicators used to judge software performance as recorded by users.
- The most popular customized software in circulation are payroll / financial ones. Figure 2 shows the detailed sectoral distribution of software in Uganda today.

Figure 2: Distribution of software in circulation



- Procurement of software is largely dependant on the task that the software performs (87%) or ease of use (83%). Other factors are the reputation of the software vendor (57%), how popular the software is (54%), likely training or technical support by the supplier (43%) and donor / higher authority prescription (26%).

10. Software is not being fully utilized. The average functional usage of general software is at 47%. Security software is the most utilized at 66% followed by text processors (65%), internet software (63%), spreadsheets (59%), readers (46%), media players (43%), presentation software (42%), databases (40%), java (24%), and project management software at 23%.
11. Functional usage of customized software is even lower at an average of 42%. Payroll software seems to be the most functionally explored with 56% followed by inventory / assets management (51%), human resource management (49%), tutoring software (11%), entertainment (39%), school management (37%), hospital management (35%) and transport / fleet management at 30%.
12. Technical challenges to computer usage are mainly the use of commands (29%) followed by security features (25%), help facilities (22%), navigation features (18%), menus (17%), dialog boxes (17%) and icons at 16%.

Software users remain ignorant of how to use software for common tasks like networking and remote login, report creation, calculations, insertion of special characters and use of help and back-up facilities. Many users prefer to perform some activities manually, and the reasons for that should be established. These include customer care and accounting tasks - stock keeping, auditing, balancing of accounts, coordinating payment records, preparation of receipts, delivery notes, quotations and invoices.

The respondents said that users are not playing a sufficient role within the software development process. They want software that is lighter (during start-up or loading tasks) and compatible with other platforms or applications. They say that software should be made more user-friendly by reducing the menu items / options and increasing the icons and prompts. The help facilities are also not being quite utilized. Some say they'd like to see more details and others say it is "too long and complicated to understand".

More training and sensitization was been frequently suggested as one method of improving software usability. Users find it expensive to sponsor support trips to vendors based abroad. They feel that maintenance costs are biting them hard. Also, employers of software developers would like to see fresh graduates competent in hard core programming like ASP.

5.1 Discussion

The fact that software users are quite happy with the benefits of adoption of customized software means that its demand will continue to rise. The software in circulation is performing well and largely to the satisfaction of users. Moreover, locally produced software is performing as well as internationally produces ones. This is good news for indigenous software developers especially since only 30% of customized software in circulation is freeware. Local developers have the benefit of producing mass-off-the-shelf software for sale with high chances of it being bought.

They are however presented with a challenge of out-competing imported services. Local developers must prove that their products are equal or even better than others, if they are to enjoy a decent cut of the market share.

Another issue which has to be tackled is how to increase user knowledge of the complete functionalities of a given software. This causes a lot of unnecessary software duplication. One possible explanation why functional utilization of customized software is worse than that of general software is that fewer consultation options are available in case of any queries or problems. This problem can be reduced when local solutions (suppliers) are more readily available.

Based on the findings compiled, the future trend as predicted by logistic regression will be as follows:

1. There will be an increase in demand for Internet based systems in all fields especially for human resource and payroll systems. Figure 3 and 4 below illustrate the technological shifts / future trends. The demand for the type of software moving in the positive direction will rise.
2. The use of more advanced security features and menus will rise while that of commands and icons will decline. It is traditional and very wide spread to find passwords as the only security check for almost all of the systems in place. This trend should and will shift to more advanced security measures like voice, face or fingerprint recognition. The continued rejection of commands is in line with the general trend that users find them too difficult to use.

Fig. 3: Future trend for general software

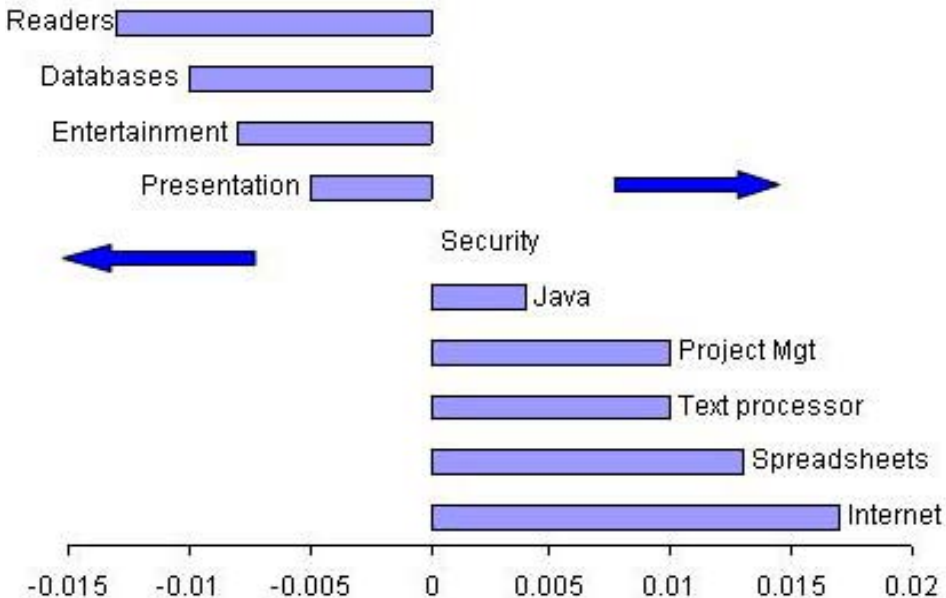
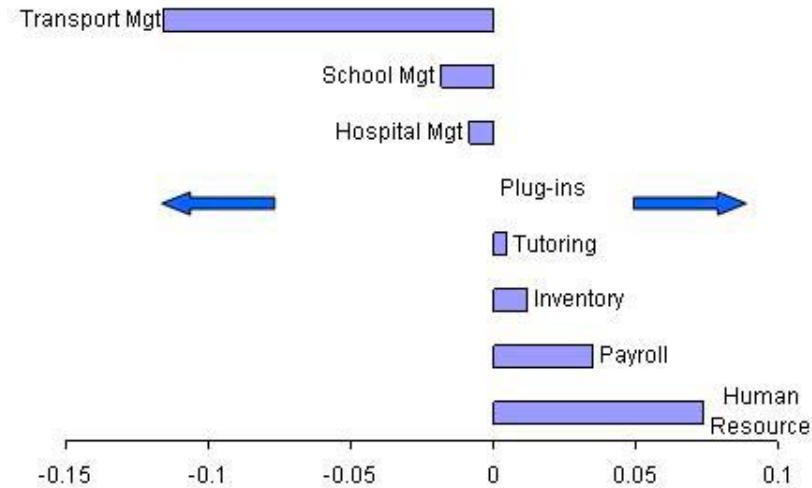


Fig. 4: Future trend for customized software



6. Recommendations

In light of the findings of this survey, we recommend the following:

1. Increase the presence of locally developed software. The 12% prevalence is too low and not cost effective for users. This can be achieved by:
 - Creating a regulatory environment (policies, laws) that favors local software.
 - Directing government departments and academic institutions to buy local software where both local and non local solutions exist as a way of promoting local talent.
 - Setting up software incubation centers that provide practical training solutions to software engineers. Training in open source platforms should be encouraged.
 - Marketing and publicity of local innovations. This is even more important since half of all respondents said they purchase software based on how popular it is.
2. Develop software that meets the current needs of the public / private sector, which are:
 - Internet-based systems (includes e-commerce, e-business, e-banking, e-governance etc).
 - Geared away from presentations or entertainment but towards human resource, customer care, financial / payroll, inventory and tutoring software.
 - Future software innovations should focus on improving quality of products, customer care, employee morale and error management.

- Enhancement of security and menu features. This can be implemented by exploiting biometrics for security as well as localization (translation into local languages) for easier interpretation and navigation.
 - Venturing into virgin areas like cartoons and children's programs, animations and local games / sports. Software that targets farmers and markets, environmental and health simulations etc also have a high potential of meeting many peoples needs.
3. Develop software requested for by respondents. They would like to see more automation in the following areas:
- Product design and assembly, detection of mechanical faults and plant maintenance. This can be for computers, cars and factory machinery. Users also want software that will manage devices to perform physical tasks like loading and offloading of commodities, packing goods, automatically opening doors, cleaning etc.
 - Financial management. Companies would like automatic reminders for re-ordering of materials or other purchases and services like payment of utilities and wages. Many parents complain that after paying fees in the bank, the school will still demand payment slips. They want an instant reflection of payment records to all bank partners like all schools / universities.
 - Monitoring and supervision of resources. These include devices that output something of economic interest and yet do not have record keeping facilities. Same applies to employee / worker movements in and out of office.
 - In-built software tasks. Many users change their minds after sending an email and they would like to un-send a sent email. Some want software that can photocopy and fax pages in addition to current printing tasks. Others would like instant messengers for machines on LANs. One person said that the timing on his computer is not the same as actual time.
 - Others - computerised bill-boards and car meters that automatically record the amount of fuel remaining in the tank.
4. Increase software usage and accessibility.
- Address the gender issues that are creating the huge imbalance of female to male computer interaction (currently at 3:7).
 - Encourage the use of open source software to make it more affordable and accessible.
5. Improve functional software utilization which is currently below average (47%).
- Make software development multidisciplinary i.e. involve fine artists, social scientists, gender experts, software engineers etc to address the different components in the software lifecycle.

- Shorten the distance between software developers / suppliers and end-users. Physically distant suppliers become too expensive to consult and this acts as a hindrance to full software utilization.
- Improve training of both software engineers and end users with regular refresher courses. Mandatory training and certification of users who successfully finish the training can be adopted.

7. Conclusion

The software industry in Uganda, and most likely, other developing countries has grown tremendously. The quality of software being developed is now of international standards and a concerted effort is needed to change people's mindset and perception that local software development is poorer. Users are ready to embrace the software revolution. They are happy with the products and are no longer complaining about huge costs involved. They know what they want and understand the benefits of software usage but still lack the knowledge and necessary expertise to optimally exploit the benefits. Both users and developers have to be supported in their pursuit to eradicate poverty in the world by 2015. The recommendations outlined in this report can be a starting point.

References

- BHAGWATI, J., A. PANAGARIYA, AND T. SRINIVASAN. 2004. The Muddles over Outsourcing, *Journal of Economic Perspectives* (Fall).
- CENTENNIAL-SOFTWARE 2005. Tracking Software Usage: Managing The Cost. Available on 15/02/2008: www.centennial-software.com
- DUNCOMBE, R., and HEEKS, R. 2001. Enterprise Development and Information and Communications in Developing Countries: Supporting "ICT Flyers". Available 15/02/2008 at: <http://www.unbotswana.org/bw/undp/docs/bhdr2002/Supporting%20ICT%20flyers.pdf>
- GIB, T. 1998. Principles of software engineering. *Addison-Wesley*. Boston
- GRUNDEY, M., R. HEEKS. 1998. Romania's Hardware and Software Industry: Building IT Policy and Capabilities in a Transitional Economy. *Development Informatics Working Paper* No. 2, IDPM Publication, Manchester (ISBN - 1 90251801 2).
- ICC / ESOMAR. 2007. International code on market and social research, *International chamber of commerce*
- ISO/IEC TR 9126-2. 2003. Software Engineering - Product Quality. External Metrics. ISO, Geneva.
- MANKIW, G., AND P. SWAGEL. 2006. The Politics and Economics of Offshore Outsourcing. Working Paper, American Enterprise Institute.
- MERRIAM-WEBSTER (2002). Webster's third new international dictionary unabridged. Springfield, MA, Merriam-Webster, Inc.
- NIELSEN, J. 1993. Usability Engineering, AP Professional. Boston, Mass.
- OGWANG, C. 2006. The Impact on Usability Of Users' Involvement in Software Development Projects. *Masters thesis of Makerere University*.

2

Game Theoretic Multi-Agent Systems Scheduler for Parallel Machines

Elisha T. O. Opiyo, Erick Ayienga, Katherine Getao,
William Okello-Odongo, Bernard Manderick and Ann Nowé

This paper considers the scheduling of n independent jobs on m non-identical machines using the ideas from game theory and multi-agent systems. The values of n and m are fixed at 30 and 6 respectively giving a schedule space with a dimension of approximately 10^{23} schedules. The agents are used to represent the jobs and they select machines on which the jobs should be processed, resulting into schedules. The schedules that are generated are evaluated using the makespan which is the total time taken for all the jobs to be processed. The makespan of the schedules that are generated vary when the agents that represent the jobs change the way they make their selection decisions. The agent selection policies that are investigated in this paper include pure random choice, potential game strategy and dispersion game strategy. The results that are obtained show that the random choice strategy and the potential game strategy generate the empirical best schedules by chance. The dispersion game strategy however is shown to converge very quickly to a stable schedule type whose best makespan value is between 3.1 to 3.4 times larger than the empirical best schedule. The main contributions in this paper include generating schedules in a concrete schedule space using ideas from game theory and multi-agent systems and the results that are obtained.

1. Introduction

Scheduling of a given number, n of independent jobs on a given number, m of non identical parallel machines is an area that has attracted much interest for several years especially from Operations Research [Graham et al. 1977; Sing & Sheik 1998; Cicirello 2003; Nowicki and Smatnicki 2005; T'Kind and Billaut 2006; Johannes 2006; Trietsch and Baker 2008]. The scheduling problem has also attracted the interest of researchers from Artificial Intelligence [Jurgen 1998; Jennings and Jackson 1995; Sycara et al. 1995; Opiyo et al. 2005; Opiyo et al. 2006; Opiyo et al. 2007]. Scheduling independent jobs on parallel non identical machines seek algorithms that optimize, through minimization, such factors as the total time it takes for all the jobs to be completed, also called the *makespan* [Ibarra and Kim 1977]. The scheduling problem has been well structured for the job shop related applications. Graham et al. [1977] introduced a three component format for specifying the problems in which the machine properties are specified followed by the constraints then the optimization details. In that format the general class

of the scheduling problem handled in this paper is $Q || C_{\max}$, denoting uniform machines, with no constraints specified and minimizing the total completion time for all the jobs. In this paper the values of n and m are fixed at 30 and 6 respectively giving a schedule space with a dimension of approximately 10^{23} schedules. The schedules are generated using ideas from Game theory and Multiagent systems. This differs from other approaches based on Operations Research (OR). The issue with the OR approaches is that most solutions are limited to each class of the scheduling problem that is solved. This makes it necessary to seek the invention of algorithms or heuristics for different problem classes. For example algorithms for $1 || C_{\max}$ are not guaranteed to solve the $3 || C_{\max}$ or the $Q || C_{\max}$ problems. The agent-based approaches are different. The schedules are generated according to the agent behaviour. This associates the qualities of schedules that are produced with the behaviour of the agents. This shifts the burden of the scheduling problem from the invention of algorithms to determining the agent behaviour that would lead to good schedules. The main advantage of using the agent-based approach is that in the extreme case that the problem class is unfamiliar the agents can learn the behaviour that leads to good schedules on their own, see Galstyan et al. [2005]. The scalability problem is also less restricting with the agent based approach. The agent based approach is also more appropriate for the emerging distributed systems environments such as the Internet, cellular networks and grid computing [Foster et al. 2001; Araujo et al. 2007; Berman et al. 2003].

1.1 The Scheduling Problem as a Multiagent Systems Game

In this section the way that the scheduling process is modeled as the agent game is outlined. The agents are entities that can sense and react to changes in their environments autonomously. The multiagent system is a system that consists of the agents [Wooldrige 2002]. Game theory is the study of interactions in contexts where the participants make the choices that are conflicting. A game is a structure that consists of a set of the agents, a set of the agent actions or choices and a set of the agent payoffs associated with their actions [Wooldrige 2002; Shoham and Leyton-Brown 2008]. The scheduling task under consideration involves assigning n independent jobs that are to be processed on m non identical machines so that the total processing time for all the jobs is minimized. A game can be formulated in which the agents represent the jobs, the machines or both [Porter 2004; Heydenreich et al. 2006; Angel et al. 2006]. In this work the agents represent the jobs. The agents act by choosing the machines on which the jobs that they represent are to be processed. The agent payoffs are based on the *makespan*. A schedule with a smaller *makespan* is the one that is more preferred. The makespan is however, a group payoff rather than an individual payoff. The types of games that are of interest in this paper include dispersion games [Trond et al. 2002], potential games [Monderer and Shapley 1996], and random choice games. Random choice games are those in which the agents make choices at random without considering any other matters.

The agents are labeled 0 – 29 and the machines are labeled 0 – 5. The machines are shown to the left. Each machine has the jobs scheduled on it labeled 0 – 29 on its right. The core of the allocation algorithm is outlined as follows:

Initialization: AgentList(0-29) each with processing time, MachineList(0-5) each with a speed

While (rounds < some pre-determined value- 100,000: 1,000,000)

Agents select machines (random || potential || dispersion game strategy)

Shortest processing time (SPT) heuristic is used to arrange agents on local machines

Note schedule statistics; adjust rounds EndWhile

The remaining sections of this paper include related work where similar research is highlighted; the experimental set up that outlines the simulation experiments and the results that were found; the discussion of the findings, and the conclusion.

2. Related Work

This work was motivated by the work of Heydenreich et al. [2006] and Porter [2004] in their work that is part of the broader area of algorithmic game theory [Marios et al. 2006; Nisan et al. 2007; Shoham and Leyton-Brown 2008]. Porter's model had n jobs that were represented by the agents, a centre that applied a scheduling mechanism and *one* machine processor. This was of the type $1 || C_{\max}$ using the notation of Graham et al. [1977]. This model was extended by Heydenreich et al. [2006] to the general type $P || C_{\max}$ in which several jobs run on several *identical* parallel machines. The jobs are represented by the agents. The agents select the machines on which they want the jobs to be processed. This work differs from their work in that the problem class that is considered in this work is of the type $Q || C_{\max}$, where the machines are not identical. It also differs in the approach in that these researches used analytical techniques to investigate the loss of efficiency from a central to a distributed environment using the price of anarchy. This work on the other hand conducts investigations in a concrete space, uses an empirical approach, and investigates the quality of schedules that can be obtained by varying the agent behaviour.

3. The Experiment and the Results

This section gives an outline of the way that the experiments were conducted. The aim of the experiment was to determine the agent behaviour that generates the schedules that are good, that is with minimum *makespan*. The important properties of jobs include the release time, the deadline, the processing time and the weight. The important machine property is its speed. There were 30 jobs labeled 0 to 29. There were 6 machines labeled 0 to 5. The machine speeds and the job processing times were fixed using integer values. It was assumed that the scheduling process begins when all the jobs are available. The release time was therefore fixed to 0 in all cases. It was also assumed that the jobs were equally weighted. The schedules were produced in rounds. One round gave one schedule. After each round, the schedule, the round of the schedule and the *makespan* of the schedule were noted.

If the schedule turned out to be better than the best schedule then its details were remembered, replacing those of the incumbent best schedule. If the schedule turned out to be the worse than the worst schedule then its details replaced those of the worst incumbent schedule. The rounds ranged from 100, 000 to 1 million.

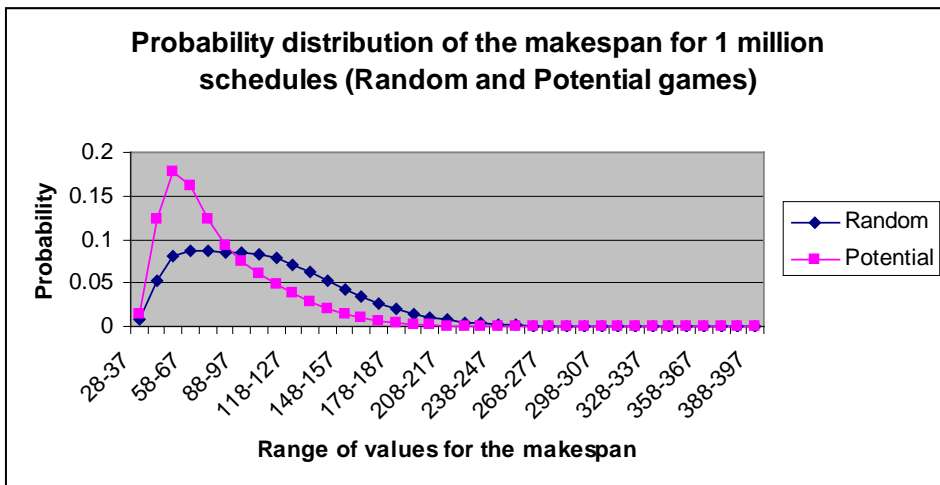
The empirical best *makespan* was 28 while the empirical worst *makespan* was 400. These values only hold for the fixed integer values that were used for the processing times of the jobs and the speeds of the machines. The results are summarized in Table 1, and Figure 2.

Table 1: Summary of the results

Attribute	Random Choice	Potential Game	Dispersion Game
Best makespan	28 - 30	28 - 30	30 -60
Worst makespan	350 - 480	270 - 380	100 - 220
Effort for best-	65%	75%	Under 5 %
Effort for worst -	15%	10%	Under 5%
Schedule quality-	Just below 8 %	7% - 10 %	8% constant
Quality of best	98%	82 %	35 %
Quality of worst	5%	5%	10%
Probability distr.	90 % in 158 - 167	90% in 118 - 127	99 % in 88-97

Legend: best-: best makespan; worst-: worst makespan; quality-: quality index distr.: distribution; 158-167, 118-127, 88-97: intervals for makespan

Fig. 2: Distribution of the *makespan* based on Random Choice and Dispersion game strategies



4. Discussion

The results on the best makespan obtainable using various agent selection schemes indicate that only the random choice and the potential game strategies give

a makespan equal to or close to 28. The makespan of 28 is the empirical best makespan for the given integer values of the job processing times and the machine speeds [Table 1]. The results from dispersion game strategy are not as good since the best makespan lies between 30 and 60. This is explained by the fact that the random effect that is present in both random choice and potential games causes the free exploration of the schedule space and makes it possible for them to stumble on schedules with the best makespan. The dispersion game strategy scheme on the other hand converges much faster to the load balanced configurations that restrict the makespan to a similar range of 88 – 97 [Table 1]. This makes the best and the worst schedules to occur around this range for the dispersion game strategy. On whether there is some predictable amount of search effort required to get the best schedule the results indicate that no reliable predictable rule is immediately evident for random choice and potential game strategies. From the results it is only indicative that more rounds are needed to guarantee getting a good schedule when using the random choice or the potential game selection policies. The dispersion game gives a very different result. The number of rounds, for getting both the best and the worst schedules is extremely low. This is explained by the convergence to a balanced load on all machines that occurs very fast mostly by the third schedule generated or third round. The results indicate that highest quality and lowest quality schedules are generated using the random and potential game strategies. Dispersion game strategy, however, produces schedules that are moderate, though between 3.1 and 3.4 larger than the empirical best schedule. These schedules are not so good and not so bad. The distribution of schedules in the space was also investigated [Figure 2 and Table 1]. The random choice strategy scheme gives insight into the possible schedule distribution in the schedule space based on the quality of the makespan. This is because the random choice strategy generates schedules at random and independently and therefore draws random samples from the schedule space. The distributions obtained from the potential game strategy and dispersion game strategy only depict the way the agent behaviour affects quality schedules. The potential game is designed such that the agents will tend to repeat the actions that previously led to the best schedules. This is visible in Figure 2, where the displacement is to the left, or towards the good schedules, those with smaller makespan. The dispersion game however led to the schedules with the makespan mostly in the range of 88 – 97 [Table 1].

5. Conclusion

This work has taken an empirical approach to handling the scheduling problem involving independent jobs on parallel non identical machines. The long term objective is to find a way of producing schedules that are good based on some measure and in this case, the *makespan*. It has been demonstrated that the ideas from game theory and multi-agent systems can be used to generate schedules in a concrete space. It has also been shown that the quality of schedules is affected by the agent behaviour as it selects a machine. It has also been demonstrated that

there is a schedule space in which the performance of load balanced solutions is consistently 3.1 to 3.4 times worse than the empirical best. In real life terms the scheduling scenarios occur in more complex contexts than the 30 jobs and 6 machines scenario. The schedule spaces in such scenarios are much larger. It means that there is much room, and by some chance, to operate with schedules that are close to the best and plenty of room, also by some chance, to operate with schedules that are bad. The findings in this work point to further investigations on the desired agent behaviour since none of the selection schemes seem satisfactory. The dispersion game converges very quickly to a solution but such a solution can be far from the empirical best schedules. This work has investigated the quality of schedules where some given integer values of the job processing times and the machine speeds are used, however, it may be extended to other selected integer values. It may also be extended to investigate contexts in which the jobs and the machines are dynamic. These are left as the issues for further work. Another area for further work is to investigate if it is necessary to use any heuristics at all where the agents are used to produce schedules as they play the allocation game for the scheduling problems of the type $Q || C_{\max}$.

Acknowledgments

Our work is supported by the Flemish Inter-University Council: Institutional University Cooperation Programme (VLIR-IUC) of Belgium. We thank them for this support. We are also grateful to the VLIR Secretariat at the University of Nairobi, and the Free University of Brussels (VUB), for their support. We also appreciate the contributions of the anonymous reviewers in improving this paper.

References

- ANGEL ERIC, EVRIPIDIS BAMPIS, FANNY PASCUAL. 2006. *Algorithmic Game Theory and Scheduling*. IBISC, University of Every, France. CIRM, May 2006. vs .. www.cirm.univ-rs.fr/videos/2006/exposes/14/Monday%20Afternoon/Bampis.ppt - last visited February, 2008.
- ARAUJO, RICARDO MATSUMURA DE; LAMB, LUIS C. 2007. An information theoretic analysis of memory bounds in a distributed resource allocation mechanism. In: *Proceedings of the 20th International Joint Conference on Artificial Intelligence (IJCAI 2007)*, 2007, Hyderabad, India.
- BERMAN FRAN, FOX GEOFFREY C., HEY ANTHONY J. G. 2003. Editors. *Grid computing: making the global infrastructure a reality*. Wiley publishers.
- CICIRELLO VINCENT A. 2003. Weighted tardiness scheduling with sequence-dependent setups: A benchmark library. *Technical report, Intelligent Coordination and Logistics Laboratory*, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, February 2003.
- FOSTER IAN, CARL KESSELMAN, STEVEN TUECKE. 2001. *The Anatomy of the grid: enabling scalable virtual organizations*, reprint in Berman et al. (2003) pp. 171 – 191.

- GALSTYAN ARAM, KARL CZAJKOWSKI, AND KRISTINA LERMAN. 2005. Resource Allocation in the Grid with Learning Agents. In the *Journal of Grid Computing*, Volume 3, pages 91-100, Springer 2005.
- GRAHAM R. L., E. L. LAWLER, J. K. LENSTRA, A.H. G. RINNODY KAN. 1977. Optimization and Approximation in Deterministic Sequencing and Scheduling: A Survey. Stichting Mathematisch Centrum. Mathematische Besliskunde; BW 82/77.
- HEYDENREICH BIRGIT, RUDOLF MÜLLER, MARC UETZ. 2006. Decentralization and Mechanism Design for Online Machine Scheduling, Algorithm Theory- SWAT 2006, Volume 4059, *Lecture Notes in Computer Science*, pp 136-147. Springer.
- HEYDENREICH BIRGIT & MÜLLER RUDOLF & UETZ MARC. 2006b. Games and Mechanism Design in Machine Scheduling – An Introduction, *Technical Report 022*, Maastricht : METEOR, Maastricht Research School of Economics of Technology and Organization, 2006.
- IBARRA OSCAR H., CHUL E. KIM. 1977. Heuristic Algorithms for Scheduling Independent Tasks on Nonidentical Processors. In *Journal of the Association for Computing Machinery*, Vol 24, No. 2, April 1977, pp 280-289.
- JENNINGS N. R., A. J. JACKSON. 1995. Agent-based Meeting Scheduling: A Design and Implementation. *IEE Electronics Letters*, Vol. 31, No. 5, The Institution of Electrical Engineering, pp. 350–352, 1995.
- JOHANNES BERIT. 2006. Scheduling parallel jobs to minimize the makespan. In *Journal of Scheduling*, 9 pp. 433 – 452, Springer.
- JURGEN DORN. 1998. *Towards reusable intelligent scheduling software*. In *Knowledge-Based Systems*, pp. 101-112, 1999, citeseer.ist.psu.edu/456393.html.
- MARIOS MAVRONICOLAS, VICKY PAPADOPOULOU, PAUL SPIRAKIS. 2006. Algorithmic Game Theory and Applications. Wiley-Interscience, John Wiley & Sons publication.
- MONDERER DOV, LLOYD S. SHAPLEY. 1996. Potential Games. In *Games And Economic Behavior* 14, 124–143 (1996) Article No. 0044.
- NISAN NOAM, TIM ROUGHGARDEN, ÉVA TARDOS, VIJAY V. VIZIRANI. 2007. Algorithmic Game Theory. Cambridge University Press, 2007.
- OPIYO ELISHA T. O., ERICK AYIENGA, KATHERINE GETAO, BERNARD MANDERICK, OKELLO-ODONGO, ANN NOWÉ. 2005. Computing research challenges and opportunities with grid computers. In proceedings of The first International workshop on sustainable research excellence in computing for developing countries. 9-12 August 2005, Makerere, Uganda.
- NOWICKI EUGENIUCZ, CZESLAW SMATNICKI. 2005. An Advanced Tabu Search Algorithm for the Job Shop Problem. In *Journal of Scheduling*, 8, pp. 145-159, 2005.
- OPIYO ELISHA T. O., ERICK AYIENGA, KATHERINE GETAO, BERNARD MANDERICK, OKELLO-ODONGO, ANN NOWÉ. 2006. Multi-agent systems scheduler and open grid computing standards. In Proceedings of The secondt International workshop on sustainable research excellence in computing for developing countries, 9-12 August 2006, Makerere, Uganda.

- OPIYO ELISHA T. O., ERICK AYIENGA, KATHERINE GETAO, BERNARD MANDERICK, OKELLO-ODONGO, ANN NOWÉ. 2007. Modeling a Multi-Agent Systems Scheduler for Grid Computing Using the Gaia Methodology. In 1st International Conference in Computer Science and Informatics, 5th – 7th February 2007, Nairobi, Kenya, School of Computing & Informatics, University of Nairobi, Kenya.
- PORTER RYAN. 2004. Mechanism Design for Online Real-Time Scheduling. Proc. 5th ACM Conf. Electronic Commerce, ACM Press, 2004, 61-70.
- SANDHOLM W. H. 2001. Potential Games with Continuous Player Sets. In Journal of Economic Theory, Volume 97, pp. 81 – 108, 2000.
- SHOHAM YOAV, KEVIN LEYTON-BROWN. 2008. Multiagent Systems; Algorithmic, Game-Theoretic and Logical Foundations. Cambridge University Press, 2008.
- SINGH JAIN ANANT, MEERAN SHEIK. 1998. A state-of-the-art review of job-shop scheduling techniques. A Technical report, Department of Applied Physics, Electronic and Mechanical Engineering, University of Dundee, Dundee, Scotland, 1998.”, citeseer.ist.psu.edu/jain98stateart.html. Last visited January 2008.
- SYCARA KATIA, DAJUN ZENG, KAZUO MIYASHITA. 1995. Using CBR to acquire user scheduling preferences. In Proceedings of the Eleventh IEEE Conference on Artificial Intelligence Applications, 1995. www.cs.cmu.edu/~sycara/cabins-papers/caia95-final.ps.gz
- T’KINDT VINCENT, JAN-CHARLES BILLAUT. 2006. Multi-criteria scheduling: Theory, models and algorithms. Springer-Verlag.
- TRIETSCH DAN, KENNETH R. BAKER. 2008. Minimizing the number of tardy jobs with stochastically ordered processing times. In Journal of Scheduling, 11, pp. 71 – 73, Springer.
- TROND GRENAGER, ROB POWERS, YOAV SHOHAM. 2002 Dispersion Games: General Definitions and Some Specific Results. In proc. AAAI02, pages 398-403. AAAI Press.
- WOOLDRIDGE MICHAEL. 2002. An introduction to multiagent systems. John Willey & Sons, Ltd, 2002.
- YURI M. ERMOLIEV, SJUR DIDRIK FLAM. 1997. Learning in Potential Games. INTERIM REPORT IR-97-022 / June, International Institute for Applied Systems Analysis _ A-2361 Laxenburg _ Austria.

3

Dynamic Resource Allocation: A Comparison of First Price Sealed Bid and Vickrey Auctions

Elisha T. O. Opiyo, Erick Ayienga, Katherine Getao,
William Okello-Odongo, Bernard Manderick and Ann Nowé

Resource allocation involves assigning the resources to the resource users. Where the resources and the resource users do not change with time, the resource allocation problem can be solved as a scheduling problem. Where the resources and the resource users change with time then some different allocation mechanisms are needed. In this paper an environment is considered in which the resources emerge and the resource requests also emerge unpredictably as time goes. The resources emerge with different capacities and in the same way the resource requests also emerge from the users with different demands as time goes. In this dynamic environment the resource allocation performance of the first price sealed bid and the Vickrey auctions are explored and compared. The system allocation performance is measured using the number of the emerging resource provisions and the resource requests that get matched. The simulation results show that there is no performance difference between the two mechanisms apart from the fact that the Auctioneer's earnings are higher for the first price sealed bid auction.

1. Introduction

Resource allocation is the process of assigning the resources to the resource users. In a static environment the number of the resources and the number of the resource users do not change. The resource allocation problem can, therefore, be solved as a scheduling problem. The resources can, however, change with time in different ways. The resource capacities can be changing with time. The resources can also change by being available at different times with different capacities. Consider a situation in which there are three computers each with the different types of packages that are used through online access by some ten people in a small organization. At any moment there may be a number of requests for the packages. If there are only three requests then there may be no problem as the requests can be immediately satisfied if no additional conditions are imposed. At some other time there could be, say, five requests. This requires making a decision on whom to allocate the computers immediately and who should wait. With time the number of requests change and the number of computers may also change. This is a dynamic resource allocation scenario. The computers are the resources

and the workers are the resource users. This paper presents an exploration of the performance of the allocation process when it is managed by the first price sealed bid, or the Vickrey auctions. It is assumed that the resources do not fail and a resource user is not transferred from one resource to another for the sake of continuity (i.e. no pre-emption).

Auctions involve buying and selling items by offering them for bids, handling the bids, and then allocating the item to the one whose bid is the highest. The goods change hands from the auctioneers who manage the bids and represent the resource owners to the bidders who represent the resource consumers. The *first price sealed bid auctions* are those in which the proposed bids are private in that the participants do not see each other's bids. The highest bidder is allocated the item and that bidder pays the amount that was in the bid. The *Vickrey auctions* are also private but the highest bidder gets the item but pays the amount that was proposed by the second highest bidder [Wooldrige 2002]. *Grid computers* are virtual computers that are constructed out of flexible, secure, coordinated resource sharing among dynamic collections of individuals, institutions and resources [Forster et al. 2001]. Although they involve many computing elements that are on the network their individual users think that they are each using a single computer. *Pre-emption* is a notion that is used to describe a situation where a job that is being executed is interrupted and continued at a later time on the same or a different machine. The job may also simply be transferred from one machine to another. *Schedules* are plans for performing work or achieving an objective, specifying the order and allotted time for each activity; or a program of events or appointments expected in a given time [Houghton 2000]. *Agents* are autonomous entities in the environment that can sense and react to the environmental changes. *Multi-agent systems* are those systems that are made up of several agents [Wooldrige et al. 2000].

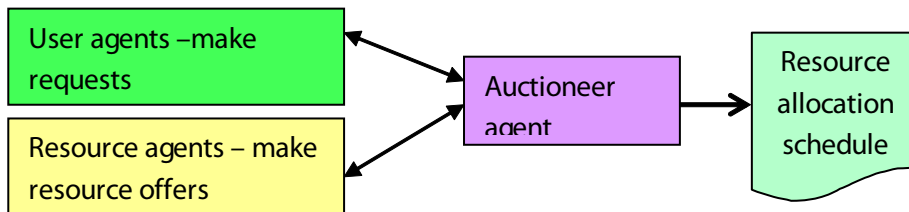
Scheduling has been identified as one of the grid computing challenging issues [Buyya 2002]. Foster et al. [2001] also identify resource sharing as important activity in grid computing. The resources should be assigned to users so that the users can work smoothly and harmoniously. In this paper the dynamic resource allocation process is considered. The resources and the users emerge and leave on their own in an unpredictable manner as time goes on. This is the basis of the dynamism. The grid computing problem can be structured as consisting of two sets one set consists of the resource users and another set consists of the resource providers requiring matching dynamically. Section 3 gives more details on this problem. This paper reports the use of the auction-based techniques for the dynamic resource allocation of the resources in a simplified environment. The dynamism is retained but pre-emption and resource failures are not allowed. The performance measurements are based on the time that the users and the resources take before they are served. Ideally a user wants to get a resource immediately while a resource seeks to be utilized immediately.

This paper continues with a presentation of the simplified resource allocation model and then gives the outline of the allocation algorithm. Related work is considered and the experimental set up is outlined. The summarized results are given, and discussed, and finally the conclusion is made with further areas of research pointed out.

2. The Resource Allocation Model

In this section the auction-based allocation model and an outline of the allocation algorithm are presented. The resource allocation model that is investigated and is reported in this paper is a simplified version of an agent-based grid scheduling model of Opiyo et al. [2007]. In that model, the auctioneer, bank, registrar and service agents manage the resource allocation process as resource agents offer their facilities while the user agents make the requests for the resources. The simplified model that is used in this paper is found in Figure 1.

Fig. 1: Auction-based resource allocation framework



The core of the resource allocation algorithm is as follows:

WHILE (the system is running)

Form a list of resource users and providers as they emerge

At some pre-determined time intervals {

For each resource compile a list of eligible users

Allocate the resource to most eligible user using auction (sealed bid/Vickrey)

Remove scheduled resources and users from their respective lists

Note important statistics

Release the current allocation schedule

}

ENDWHILE

3. Related Work

In this section the scheduling problem as related to the grid computing context is outlined. The way this problem has been approached is considered. The grid computing resource allocation problem can be stated in the same way as the general scheduling problem. Consider the set of user tasks $\{T_1, T_2, T_3, \dots, T_n\}$ and a set of resources $\{R_1, R_2, R_3, \dots, R_m\}$. The problem is to find a way to assign resources to the tasks such that some objectives are optimized. In grid computing, however, due to the dynamism, uncertainties, wide distribution

and resource autonomies the optimization factor is not pursued much. The grid computing problem has been approached using the techniques that are economic on the one hand and non-economic on the other [Buyya 2000]. The economic approaches use commodity market-based and auction-based techniques [Buyya 2000; Ferguson et al. 1999; Rajkumar et al. 1997; Wolski et al. 2000; Buyya et al. 2005]. The commodity-based approaches involve buyers and sellers where the prices are publicly agreed upon and may be demand driven. The auction-based methods involve the auctioneers and the bidders. The allocations are done by the auctioneer awarding the resources to the most qualified bidders according to the rules of the respective auctions. The non-economic approaches to solving the resource allocation problem are responsible for the early grid computing platforms such as Condor. These non-economic allocation methods rely on scheduling techniques that emerged out of distributed operating systems [Berman et al. 2003c; Zoltan and Lazlo 1999; Ammar et al. 1999; Gabrielle et al. 2001; Cao et al. 2001; Tomasz et al. 2003; Walfredo et al. 2004; Adnan 2004; Bucur 2004]. These researchers were mostly concerned with delivering complete grid computing architectures using various scheduling techniques that are based on queues, service discovery or brokering. A non-economic approach that is close to this work is that of Getao and Opiyo [Getao and Opiyo 2007], in which data structures are used in managing the resource allocations process. This paper presents the work that differs from the other previous work in that the auction-based approach is used in the resource allocation. The focus in this paper is on the performance related issues that would arise in the grid computing context. The rate at which the emerging resource requests and resource offers are matched by the dynamic allocation mechanism is the performance matter that is being investigated. The other previous auction-based work has concentrated on the allocation of particular commodities such as the bandwidth but not concerned with the grid computing [Rosu et al. 1997; Ramanathan et al. 1999; Chandra et al. 2002; Ghosh et al. 2004; Georgiadis et al. 2006]. The other previous work involving auctions has also been concerned with the centralization and the decentralization issues. It has involved investigating the nature of the auction mechanisms that are strategy-proof and encourage truthful revelations [Nisan & Ronen 1999; Nisan & Ronen 2000]. The pioneering work of Chunlin and Layuan [2008] concentrates more on the layered modeling of global optimization in grid computing with focus on the quality of service.

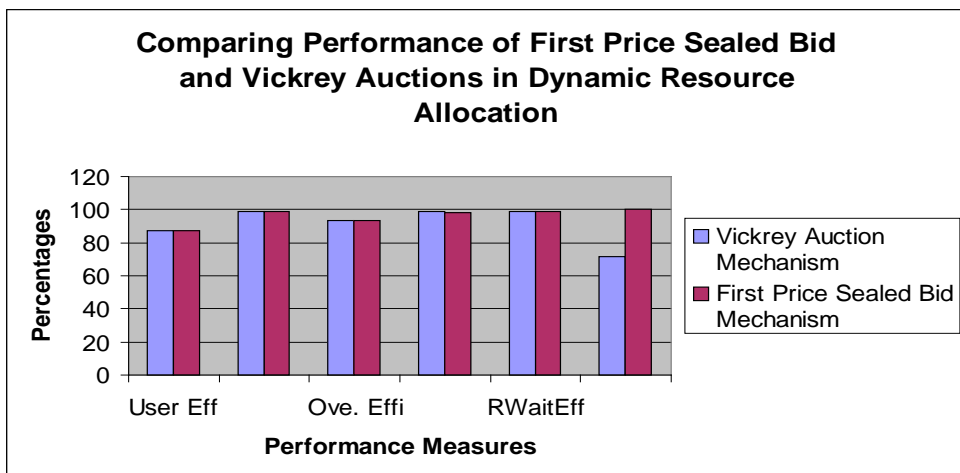
4. The Experiment and the Results

The objective is to investigate rate at which the resource users and the resource providers are matched by the system's dynamic allocation mechanism as they emerge. Some questions are used to guide the investigation. When a user is looking for a resource how long would the user wait before the resource request is met? Similarly, how long does it take a resource that has been availed for use before it begins to be used? What is the overall waiting time for the system? What does the auctioneer gain? The variables whose values are recorded include user efficiency,

resource efficiency, overall efficiency, user waiting time efficiency, resource waiting time efficiency and auctioneer's efficiency. The *user efficiency* is the ratio of the total users whose requests are met to the total number of users expressed as a percentage. The *resource efficiency* is the ratio of the total resource providers whose offers are utilized to the total number of resource providers expressed as a percentage. The *overall efficiency* is the ratio of the total resource users and providers whose needs are met to the total number of users and providers expressed as a percentage. This variable acts as the systems performance metric. The user's waiting time efficiency is computed as: $(Total\ Length\ of\ Time - Cumulative\ Waiting\ Time\ for\ Users) * 100 / Total\ Length\ of\ Time$. The resource provider's waiting time efficiency is computed as: $(Total\ Length\ of\ Time - Cumulative\ Waiting\ Time\ for\ Resources) * 100 / Total\ Length\ of\ Time$. The *auctioneer's efficiency* is computed as the ratio of total amount earned to potential amount that can be earned expressed as a percentage.

The average values were obtained from all the sessions beginning with 10,000 rounds to 100,000 rounds for each dynamic allocation mechanism. The first price sealed bid auction allocation mechanism yielded the efficiencies above 85% for the user, resource, system, user waiting and resource waiting. The auctioneer efficiency was 100%. The Vickrey auction allocation mechanism yielded the efficiencies above 83% for the user, resource, system, user waiting and resource waiting. The auctioneer efficiency was about 70%. Figure 2 gives a comparison of the performances.

Fig. 2: Comparing performances of the auction mechanisms



Legend: User Eff - User Efficiency; Res. Eff- Resource Efficiency; Ove. Effi - Overall Efficiency; UWaitEff - user waiting efficiency; RWaitEff - Resource waiting efficiency; Auc. Eff - Auctioneer Efficiency

5. Discussion

The resource users are cleared at an efficiency rate above 83%. The resource provider's efficiency is above 95% for both mechanisms. The non attainment of

100% is explained by the fact that there will be some requests or offers that cannot be met. Some of the resources may not be used due to the non availability of the suitable resource users. Sometimes some resource users are always bidding lower than the other competing users. The overall performance in both cases is over 80%. In both mechanisms the mean waiting time efficiencies are above 95% that means that the emerging resource requests and resource offers do not wait for a long time before their needs are met. This is an indication of the potential that exists in the use of the auction-based allocation mechanisms in a dynamic environment such as the grid computing.

Figure 2 shows how the performances of the two mechanisms compare. The percentage difference between the corresponding values of the performance measurements are below 3 % except for the auctioneer's earning efficiency. The auctioneer's earning is more for the first price sealed bid allocation mechanism because in it the bidder pays the amount that is on the bid. This enables the auctioneer to collect all the possible revenue. For the Vickrey auction, on the other hand, the bidder pays the value of the second highest bidder which is most likely lower than the winner's bid price. The auctioneer therefore earns less than the possible total revenue from the winning bids. So apart from the auctioneer's earnings, all other performance parameters seem to be similar. The similarity of other performance values is explained by the similarity of the two allocation processes. Both are sealed bids. Both have the same allocation process. In both cases to get an eligible bidder the list of bidders must be in descending order of the bidding amounts. The process of picking the second highest bidder when the auctioneer is collecting payment in case of the Vickrey allocation mechanism does not seem to introduce any waiting overhead. This is because the process of picking payment from the highest bidder and from the second highest bidder need not take different lengths of time.

6. Conclusion

The results from these experiments point to the fact that the use of the first price sealed bid and the Vickrey allocation mechanisms in the grid computing environment is promising. Either of them may be used in managing the grid computing resource allocation tasks. Since both mechanisms deliver similar performance capabilities except for the auctioneer's revenue, the circumstances may determine which one to use. Where the resource owners insist on maximizing their revenues the first price sealed bid auction allocation mechanism may be selected. However, if truthful bidding is emphasized then the Vickrey auction allocation mechanism can be used. The auction-based resource allocation mechanisms such as the ones that have been compared can only be implemented where the resource requests are always accompanied by some numeric offers. The numeric offers should represent the valuations that are attached to the resources by those making the requests. These mechanisms can also work only under the circumstances where the resource requests and resource offers are periodically processed particularly

to allow the auctioneers to perform the allocations. In computing environments there are periodic checks that enable the events to be recognized for servicing. As an item of further work, investigations can be conducted to find out the effect on performance of pre-emption, uncertainties of the resource providers and of the resource consumers. There is also need to investigate the effect of the length of the time that the auctioneer takes before the allocations are conducted.

Acknowledgments

We are very grateful to the Flemish Inter-University Council: Institutional University Cooperation Programme (VLIR-IUC) of Belgium. They have supported our work. We also thank the VLIR Secretariat at the University of Nairobi, and the Free University of Brussels (VUB), for their support. We are thankful to the anonymous reviewers for helping in the improvement of this paper.

References

- ADNAN FIDA . 2004. *Scheduling Policies for Enterprise Wide Peer-to-Peer Computing*. In Proceedings of the 2003-2004 Grad Symposium. CS Dept, University of Saskatchewan, April 2004. [Citeseer.ist.psu.edu/709611.html](http://citeseer.ist.psu.edu/709611.html) – last visited February, 2008.
- AMMAR H. A., PRASANNA V. K., RAGHAVENDRA C.S. 1999. A Unified Resource Scheduling Framework for Heterogeneous Computing Environments. In *Proc. HCW'99, 8th Heterogeneous Computing Workshop*, pages 156–168, San Juan, Puerto Rico, April 1999.
- BERMAN F., FOX G. C., HEY A. J. G. 2003. (Editors). *Grid computing: making the global infrastructure a reality*. Wiley publishers.
- BERMAN, F., FOX, G.C., HEY, A.J.G. 2003b. The Grid: Past, Present, Future. In Berman, F., Fox, G.C., Hey, A.J.G.,(eds): *Grid Computing: Making the Global Infrastructure a Reality*. Wiley publishers, 9 – 50.
- BERMAN F., WOLSKI R., CASANOVA H., CIRNE W., DAIL H., FAERMAN M., FIGUEIRA S., HAYES J., OBERTELLI G., SCHOPF J., SHAO G., SMALLEN S., SPRING ., SU A., DMITRII Z. 2003c. Adaptive Computing on the Grid Using AppLeS. In *IEEE Transactions On Parallel And Distributed Systems*, VOL. 14, NO. 4, APRIL 2003.
- BUCUR ANCA. 2004. Performance Analysis of Processor Co-Allocation in Multicuster Systems. *A PhD Thesis*, Delft University of Technology.
- BUYAYA, R. 2002. Economic-Based Distributed Resource Management and Scheduling for Grid Computing. *PhD Thesis*, Monash University Melbourne Australia.
- BUYAYA R., D. ABRAMSON, AND S. VENUGOPAL, 2005. The Grid Economy. In *Proc. IEEE*, vol. 93, no. 3, March. 2005.
- CAO J, KERBYSON D. J., NUDD G. R. 2001. Use of Agent-Based Service Discovery for Resource Management in Metacomputing Environment. In *Proc. 7th Int. Euro-Par Conf.*, Manchester, UK, LNCS 2150, 882-886, 2001.
- CHANDRA A., W. GONG, AND P. SHENOY. 2002. Dynamic resource allocation for shared data centers using online measurements. *Technical Report TR02-30*, Department of Computer Science, University of Massachusetts, 2002. <http://citeseer.ist.psu.edu/>

[chandra03dynamic.html](#)

- CHUNLIN LI, LI LAYAN. 2008. Multi-level scheduling for global optimization in grid computing. In *Computers and Electrical Engineering*, Volume 34, pp. 202-22, Elsevier 2008.
- FERGUSON D., C. NIKOLAOU, J. SAIRAMESH, AND Y. YEMINI. 1996. Economic models for allocating resources in computer systems. In Scott Clearwater, editor, *Market-Based Control: A Paradigm for Distributed Resource Allocation*, Scott Clearwater. In World Scientific, Hong Kong, 1996. <http://citeseer.ist.psu.edu/ferguson96economic.html>.
- FOSTER, I., CARL, K., STEVEN, T. 2001. The Anatomy of the Grid: Enabling Scalable Virtual Organizations .2001. Reprint in Berman, F., Fox, G.C., Hey, A.J.G.,(eds): *Grid Computing: Making the Global Infrastructure a Reality*. Wiley publishers (2003) 171 – 191.
- GABRIELLE A. ANGULO D., FOSTER I., LANFERMANN G.,LIU C., RADKE T ., SEIDEL E., SHALF J. 2001. *The Cactus Worm: Experiments with Dynamic Resource Discovery and Allocation in a Grid Environment*. www.cs.uchicago.edu/files/tr_authentic/TR-2001-28.pdf - last visited February, 2008.
- GETAO KATHERINE, OPIYO E. T. O. 2007. A Grid Service Matrix Model for Resource Scheduling in Grid Computing Environments. In *1st International Conference in Computer Science and Informatics*, 5th – 7th February 2007, Nairobi, Kenya, School of Computing & Informatics, University of Nairobi, Kenya.
- GEORGIADIS L., NEELY M. J., TASSIULAS L . 2006. Resource Allocation and Cross-Layer Control in Wireless Networks. In *Foundations and Trends in Networking* Vol. 1, No 1 (2006) 1–144.
- GHOSH S., RAJKUMAR R., HANSEN J., LEHOCZKY J. 2004. *Integrated Resource Management and Scheduling with Multi-Resource Constraints*. www.cs.cmu.edu/People/hansen/papers/radar-sched.pdf - last visited February, 2008.
- HOUGHTON. 2000. *The American Heritage® Dictionary of the English Language*, Fourth Edition, 2000, Houghton Mifflin Company. education. yahoo.com/reference/dictionary- last visited February 2008.
- NG C., D. C. PARKES, AND M. SELTZER. 2003. Strategyproof Computing: Systems Infrastructures for Self-Interested Parties. *Technical report, Harvard University*, 2003. <http://citeseer.ist.psu.edu/article/ng03strategyproof.html>
- NISANN. RONEN A. 2000. Computationally Feasible VCG Mechanisms. In *ACM Conference on Electronic Commerce 2000*: 242-252
- NISAN N., RONEN A . 1999. Algorithmic Mechanism Design. In *STOC 1999*: 129-140.
- RAJKUMAR R., C. LEE, J. LEHOCZKY, AND D. SIEWIOREK. 1997. A Resource Allocation Model for QoS Management. In *Proceedings of the IEEE Real-Time Systems Symposium*, December 1997. <http://citeseer.ist.psu.edu/rajikumar97resource.html>
- RAMANATHAN P., K. M. SIVALINGAM, P. AGRAWAL AND S. KISHORE. 1999. Dynamic resource allocation schemes during handoff for mobile multimedia wireless networks. In *IEEE Journal on Selected Areas in Communications*, 17(7), July 1999. <http://citeseer.ist.psu.edu/ramanathan99dynamic.html>
- ROSU D., K. SCHWAN, S. YALAMANCHILI, AND R. JHA. 1997. On Adaptive Resource Allocation for Complex Real-Time Applications. 18th *IEEE Real-Time Systems Symposium*, Dec., 1997. <http://citeseer.ist.psu.edu/article/rosu97adaptive.html>

- SYCARA K., S. ROTH, N. SADEH M. FOX. 1991. Distributed constrained heuristic search. In *IEEE Transactions on Systems, Man, and Cybernetics*, 21:1446--1461, 1991.
- TOMASZ W., KURZYNIEC D, DRZEWIECKI D, SUNDERAM V. 2003. Resource Monitoring and Management in Metacomputing Environments. In *Proceedings 10th European PVM/MPI User's Group Meeting*, Venice, Italy, pp. 629-635, September 2003.
- WALFREDO C., BRASILEIRO F., COSTA L., PARANHOS D., SANTOS-NETO E., ANDRADE N. 2004. *Scheduling in Bag-of-Task Grids: The PAUÁ Case*. <http://citeseer.ist.psu.edu/cis?q=Luiz+Biondi+Neto>, walfredo.dsc.ufcg.edu.br/ papers /paula.pdf - last visited February 2008.
- WOOLDRIDGE MICHAEL. 2002. *An introduction to multiagent systems*. John Willey & Sons, Ltd.
- WOLSKI R., PLANK J. S., BREVIK J., BRYAN T. 2000. Analyzing Market-based Resource Allocation Strategies for the Computational Grid. In *International Journal of High Performance Computing Applications*, 2001, 258-281.
- WOOLDRIDGE M., NICHOLAS R. JENNINGS, DAVID KINNY. 2000. The Gaia Methodology for Agent-Oriented Analysis and Design, in *Autonomous Agents and Multi-Agent Systems*, 3, 285-312. Kluwer Academic Publishers.
- ZOLTAN J., LAZLO K. 1999. A Jini-based prototype Metacomputing framework. *Lecture Notes in Computer Science*, Vol. 1900, P. 1171--, 2001.

4

Bootstrapping Machine Translation for the Language Pair English – Kiswahili

Guy De Pauw, Peter Waiganjo, Wagacha Gilles and Maurice De Schryver

In recent years, research in Machine Translation has greatly benefited from the increasing availability of parallel corpora. Processing the same text in two different languages yields useful information on how words and phrases are translated from a source language into a target language. To investigate this, a parallel corpus is typically aligned by linking linguistic tokens in the source language to the corresponding units in the target language. An aligned parallel corpus therefore facilitates the automatic development of a machine translation system. In this paper, we describe data collection and annotation efforts and preliminary experiments with a parallel corpus English - Kiswahili.

1. Introduction

Language technology applications such as speech recognition and machine translation can provide an invaluable impetus in bridging the digital divide. For a language like Kiswahili, digital resources have become increasingly important in everyday life both in urban and rural areas, particularly thanks to the increasing number of web-enabled mobile phone users in the language area. Most research in the field of language technology for African languages is however still firmly rooted in the knowledge-based paradigm, in which language applications and tools are built on the basis of manually compiled rules. This approach makes development of language technology applications expensive, since it requires a significant amount of expert knowledge and manual effort. The need for a cheaper and faster alternative for developing African language technology applications is therefore high.

The data-driven, corpus-based approach envisioned in this paper establishes such an alternative, so far not yet extensively investigated for African languages. The main advantage of this approach is its language independence: all that is needed is (linguistically annotated) language data, which is fairly cheap to compile. Given this data, existing state-of-the-art algorithms and resources can consequently be reused to quickly develop robust language applications and tools.

Most African languages are however resource-scarce, meaning that digital text resources are few. An increasing number of publications however are showing that carefully selected procedures can indeed bootstrap language technology for Kiswahili [De Pauw et al. 2006] and even smaller local Kenyan languages [De Pauw and Wagacha 2007; De Pauw et al. 2007a; De Pauw et al. 2007b].

In this paper we outline on-going research on the development of a data-driven machine translation system for the language pair English – Kiswahili. We first provide a short survey of the different approaches to machine translation (Section 2). We then concentrate on the required data collection and annotation efforts (Section 3) and describe some preliminary experimental results with automatic sentence and word alignment tools (Section 4). We conclude with a discussion of the current limitations to the approach and provide pointers for future research.

2. Machine Translation

The main task of Machine Translation (MT) can be defined as having a computer take a text input in one language, the Source language (SL), decode its meaning and re-encode it producing as output a similar-meaning text in another language, the Target language (TL). The idea of building an application to automatically convert text from one language to an equivalent text-meaning in a second language traces its roots back to cold war intelligence efforts in the 1950's and 60's for Russian-English text translations. Since then a large number of MT systems have been developed with varying degrees of success. For an excellent overview of the history of MT, we refer the reader to [Hutchins 1986].

The original dream of creating a fully automatic MT system has long since been abandoned and most research in the field currently concentrates on minimizing human pre- and post-processing effort. A human translator is thus considered to work alongside the MT system to produce faster and more consistent translations.

The Internet brought in an interesting new dimension to the purpose of MT. In the mid 1990s, free online translation services began to surface with an increasing number of MT vendors. The most famous example is AltaVista's BabelFish, offering on-line versions of Systran to translate English, French, German, Spanish and other Indo-European languages. Currently Google.inc is also offering translation services. While these systems provide far from perfect output, they can often give readers a sense of what is being talked about on a web page in a language (and often even character set) foreign to them.

There are roughly three types of approaches to machine translation:

1. **Rule-based methods** perform translation using extensive lexicons with morphological, syntactic and semantic information, and large sets of manually compiled rules. These systems are very labor intensive to develop.
2. **Statistical methods** entail the collection and statistical analysis of bilingual text corpora, i.e. parallel corpora. The technique tries to find the highest probability translation of a sentence or phrase among the exponential number of choices.
3. **Example-based methods** are similar to statistical methods in that they are parallel corpus driven. An Example-Based Machine Translator (EBMT) scans for patterns in both languages and relates them in a translation memory.

Most MT systems currently under development are based on methods (2) and/or (3). Research in these fields has greatly benefited from the increasing availability of parallel corpora, which are needed to bootstrap these approaches. Such a parallel corpus is typically aligned by linking, either automatically or manually, linguistic tokens in the source language to the corresponding units in the target language. Processing this data enables the development of fast and effective MT systems in both directions with a minimum of human involvement. In the next section we describe data collection and preprocessing efforts on the Sawa Corpus, a parallel corpus English – Kiswahili.

3. Data Collection And Annotation

While digital data is increasingly becoming available for Kiswahili on the Internet, sourcing useful bilingual data is far from trivial. At this stage in the development of the MT system, it is paramount to use faithfully translated material, as this benefits further automated processing. The corpus-based MT approaches we wish to employ, require word alignment to be performed on the texts, during which the words in the source language are linked to the corresponding words in the target language (also see Figure 1). But before we can do this, we need to perform sentence-alignment, during which we establish an unambiguous mapping between the sentences in the source text and the sentences in the target text. While some data is inherently sentence-aligned, other texts require significant preprocessing before word alignment can be performed.

The Sawa Corpus currently consists of a reasonable amount of data (roughly half a million words in each language), although this is not comparable to the resources available to Indo-European language pairs, such as the Hansard corpus [Roukos et al. 1997] (2.87 million sentence pairs). Table 1 gives an overview of the data available in the Sawa Corpus. For each segment it lists the number of sentences and words.

Table 1: Overview of the Data in the Sawa Corpus

	English Sentences	Kiswahili Sentences	English Words	Kiswahili Words
New Testament	7.9k		189.2k	151.1k
Quran	6.2k		165.5k	124.3k
Declaration of HR	0.2k		1.8k	1.8k
Kamusi.org	5.6k		35.5k	26.7k
Movie Subtitles	9.0k		72.2k	58.4k
Investment Reports	3.2k	3.1k	52.9k	54.9k
Local Translator	1.5k	1.6k	25.0k	25.7k
Full Corpus Total	33.6k	33.6k	542.1k	442.9k

We found digitally available Kiswahili versions of the New Testament and the *uran* for which we sourced the English counterparts. While religious material has a specific register and may not constitute ideal training material for an open-ended MT system, it does have the advantage of being inherently aligned on the verse level, facilitating further sentence alignment. Another typical bilingual text is the UN Declaration of Human Rights, which is available in many of the world's languages, including Kiswahili. This text was manually sentence-aligned.

The downloadable version of the on-line dictionary English-Kiswahili [Benjamin 2008] contains individual example sentences associated with the dictionary entries. These can be extracted and used as parallel data in the Sawa corpus. Since at a later point, we also wish to study the specific linguistic aspects of spoken language, we opted to have some movie subtitles manually translated. These can be extracted from DVDs and while the language is compressed to fit on screen and constitutes scripted language, they nevertheless provide a good sample of spoken language. It is inherently sentence-aligned, thanks to the technical time-coding information and also opens up possibilities for MT systems with other language pairs, since a commercial DVD typically contains subtitles for a large number of other languages as well.

The rest of the material consists of paragraph-aligned data, which was manually sentence-aligned. We obtained a substantial amount of data from a local Kenyan translator. Finally, we also included Kenyan investment reports. These are yearly reports from local companies and are presented in both English and Kiswahili. A major difficulty was extracting the data from these documents. The company reports are presented in colorful brochures in PDF format, meaning automatic text exports require manual post-processing and paragraph alignment. They nevertheless provide a valuable resource, since they come from a fairly specific domain and are a good sample of the type of text the projected MT system may need to process in a practical setting.

The reader may note that there is a very diverse range of texts within the Sawa corpus, ranging from movie subtitles to religious texts. While it certainly benefits the evaluation to use data from texts in one specific language register, we have chosen to maintain variety in the language data at this point. Upon evaluating the decoder at a later stage, we will however investigate the bias introduced by the specific language registers in the corpus.

All of the data in the corpus was subsequently tokenized, which involves automatically cleaning up the texts, conversion to UTF-8 and splitting punctuation from word forms. The next step involved scanning for sentence boundaries in the paragraph-aligned text, to facilitate the automatic sentence alignment method described in Section 4.

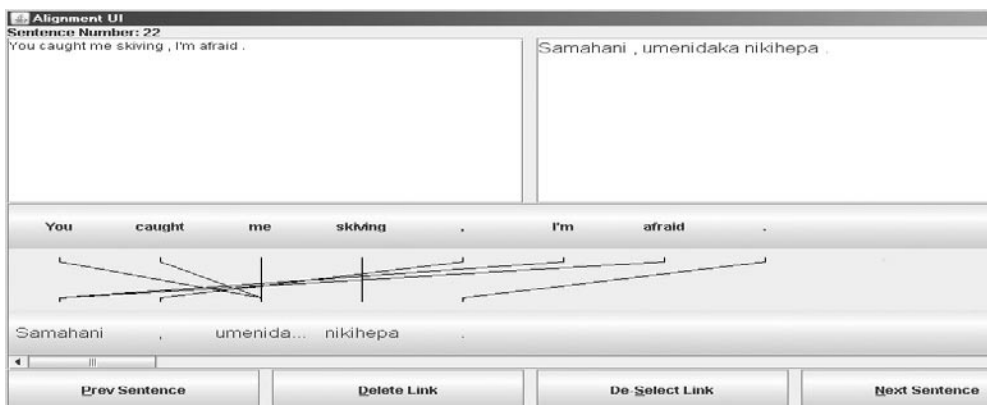
While not necessary for further processing, we also performed manual word-alignment annotation. This task can be done automatically, but it is useful to have a gold-standard reference against which we can evaluate the automated method. Monitoring the accuracy of the automatic word-alignment method against the

human reference, will allow us to tweak parameters to arrive at the optimal settings for this language pair.

We used the UMIACS word alignment interface [Hwa and Madnani 2004] for this purpose and asked the annotators to link the words between the two sentences (Figure 1). Given the linguistic difference between English and Kiswahili, this is by no means a trivial task. Particularly the morphological richness of Kiswahili means that there is a lot of convergence from words in English to words in Kiswahili. This alignment was done on some of the manual translations of movie subtitles, giving us a gold-standard word-alignment reference of about 5,000 words. Each annotator’s work was cross-checked by another annotator to improve correctness and consistency.

4. Proposed Methods

Fig. 1: Manual word alignment using the UMIACS interface



There are a number of packages available to process parallel corpora. To preprocess the paragraph-aligned texts, we used Microsoft’s bilingual sentence aligner [Moore 2002]. The output of the sentence alignment was consequently manually corrected. We found that 95% of the sentences were correctly aligned with most errors being made on sentences that were not present in English, i.e. instances where the translator decided to add an extra clarifying sentence to the direct translation from English. This also explains why there are more Kiswahili words in the paragraph aligned texts than in English, while the situation is reversed for the sentence aligned data.

For word-alignment, the state-of-the-art method is GIZA++ [Och and Ney 2003], which implements the word alignment methods IBM1 to IBM5 and HMM. While this method has a strong Indo-European bias, it is nevertheless interesting to see how far we can get with the default approach used in statistical MT. We evaluate by looking at the word alignments proposed by GIZA++ and compare them to the manually word-aligned section of the Sawa Corpus. We can quantify the evaluation by calculating precision and recall and their harmonic mean, the F-

score (Table 2). The former expresses how many links are correct, divided by the total number of links suggested by GIZA++. The latter is calculated by dividing the number of correct links, by the total number of links in the manual annotation. While the results presented in Table 2 are encouraging, it is clear that extra linguistic data sources and a more elaborate exploration of the experimental parameters of GIZA++ is needed.

Table 2: Precision, Recall and F-score for the word-alignment task using GIZA++

Precision	Recall	F _($\beta=1$)
39.4%	44.5%	41.79%

5. Discussion

In this paper we presented parallel corpus collection work that will enable the construction of a machine translation system for the language pair English – Kiswahili. We are confident that we have a critical amount of data that will enable good word alignment that can subsequently be used as a model for an MT decoding system, such as the Moses package [Koehn et al. 2007]. While the currently reported scores are not yet state-of-the-art, we are confident that further experimentation and the addition of more bilingual data as well as the introduction of extra linguistic features will raise the accuracy level of the proposed MT system.

The most straightforward addition is the introduction of part-of-speech tags as an extra layer of linguistic description, which can be used in word alignment model IBM5. The current word alignment method tries to link word forms, but knowing that for instance a word in the source language is a noun, will facilitate linking it to a corresponding noun in the target language, rather than considering a verb as a possible match. Both for English [Ratnaparkhi 1996] and Kiswahili [De Pauw et al. 2006], we have highly accurate part-of-speech taggers available. Another extra information source that we have so far ignored is a digital dictionary as a seed for the word alignment. The kamusiproject.org electronic dictionary will be included in further word-alignment experiments and will undoubtedly improve the quality of the output.

Once we have a stable word alignment module, we will further conduct learning curve experiments, in which we train the system with gradually increasing amounts of data. This will provide us with information on how much more data we need to achieve state-of-the-art performance. This additional data can be automatically found by parallel web mining, for which a few systems have recently become available [Resnik and Smith 2003]. Furthermore, we will also look into the use of comparable corpora, i.e. bilingual texts that are not straight translations, but deal with the same subject matter. These have been found to work as additional material within a parallel corpus [McEnery and Xiao 2007] and may further help improve the development of a robust, open-ended and bidirectional machine translation system for the language pair English - Kiswahili.

Acknowledgments

The research presented in this paper was made possible through the support of the VLIR-IUC-UON program. The first author is funded as a Postdoctoral Fellow of the Research Foundation - Flanders (FWO). We are greatly indebted to Dr. James Omboga Zaja for contributing some of his translated data, to Mahmoud Shokrollahi-Far for his advice on the Quran and to Anne Kimani, Chris Wangai Njoka and Naomi Maajabu for their annotation efforts.

References

- BENJAMIN, M. 2008. *The Kamusi Project*. Available at: <http://www.kamusiproject.org> (Accessed: 2 June 2008).
- DE PAUW, G., DE SCHRYVER, G.-M. AND WAGACHA, P.W. 2006. Data-driven part-of-speech tagging of Kiswahili. In *Proceedings of Text, Speech and Dialogue, 9th International Conference* Springer Verlag, Berlin, Germany, 197-204.
- DE PAUW, G. AND WAGACHA, P.W. 2007. Bootstrapping morphological analysis of Gikuyu using unsupervised maximum entropy learning. In *Proceedings of the eighth INTERSPEECH conference*, Antwerp, Belgium.
- DE PAUW, G., WAGACHA, P.W. AND ABADE, D.A. 2007a. Unsupervised Induction of Dholuo Word Classes using Maximum Entropy Learning. In *Proceedings of the 1st International Conference in Computer Science and ICT (COSCIT 2007)* University of Nairobi, Nairobi, Kenya.
- DE PAUW, G., WAGACHA, P.W. AND DE SCHRYVER, G.-M. 2007b. Automatic diacritic restoration for resource-scarce languages. In *Proceedings of Text, Speech and Dialogue, 10th International Conference* Springer Verlag, Berlin, Germany, 170-179.
- HUTCHINS, W.J. 1986. *Machine translation: past, present, future*. Ellis, Chichester.
- HWA, R. AND MADNANI, N. 2004. The UMIACS Word Alignment Interface. Available at: <http://www.umiacs.umd.edu/~nmadnani/alignment/forclip.htm> (Accessed: 2 June 2008).
- KOEHN, P., HOANG, H., BIRCH, A., CALLISON-BURCH, C., FEDERICO, M., BERTOLDI, N., COWAN, B., SHEN, W., MORAN, C., ZENS, R., DYER, C., BOJAR, O., CONSTANTIN, A. AND HERBST, E. 2007. Moses: Open Source Toolkit for Statistical Machine Translation. In *Annual Meeting of the Association for Computational Linguistics (ACL), demonstration session*, Prague, Czech Republic.
- MCENERY, A.M. AND XIAO, R.Z. 2007. Parallel and comparable corpora: What are they up to? In *Incorporating Corpora: Translation and the Linguist. Translating Europe. Multilingual Matters*, Clevedon, UK.
- MOORE, R.C. 2002. Fast and Accurate Sentence Alignment of Bilingual Corpora. In *Proceedings of the 5th Conference of the Association for Machine Translation in the Americas on Machine Translation: From Research to Real Users* Springer-Verlag, London, UK, 135-144.
- OCH, F.J. AND NEY, H. 2003. A systematic comparison of various statistical alignment models. *Computational Linguistics* 29, 19-51.

- RATNAPARKHI, A. 1996. A Maximum Entropy Model for Part-of-Speech Tagging. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, E. BRILL AND K. CHURCH Eds. Association for Computational Linguistics, Somerset, New Jersey, 133-142.
- RESNIK, P. AND SMITH, N.A. 2003. The Web as a parallel corpus. *Computational Linguistics* 29, 349-380.
- ROUKOS, S., GRAFF, D. AND MELAMED, D. 1997. *Hansard French/English*. Available at: <http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC95T20> (Accessed: 2 June 2008).

5

Towards Domain Independent Named Entity Recognition

Fredrick Edward Kitoogo, Venansius Baryamureeba and Guy De Pauw

Named entity recognition is a preprocessing tool to many natural language processing tasks, such as text summarization, speech translation, and document categorization. Many systems for named entity recognition have been developed over the past years with substantial success save for the problem of being domain specific and making it difficult to use the different systems across domains. This work attempts to surmount the problem by proposing the use of domain independent features with a maximum entropy model and a multiobjective genetic algorithm (MOGA) to select the best features. The methods used in this work are backed up by experiments of which the classifications are evaluated using two diverse domains. Conclusions are finally drawn and the outlook for future work is considered.

1. Introduction

The rising need of automatic recognition and classification of information from different literature sources has given rise to the development of many named entity recognition systems. The systems that have emerged over the years have however been tailor-made to work with specific domains such as the newswire in the CoNLL 2003 shared task [Tjong Kim Sang and De Meulder 2003]. The systems are designed with respect to unique semantics, domain specificity, document genre and syntax which render it very difficult for adaptation across domains.

The focus of this work is placed on recognition and classification of text from diverse domains. We propose the following approach to handle the task: first, we identify the different diverse domain corpora from which the recognition is going to be done, second, we make out the entity types that are common to the different domains; in this work we adopted the entity types used in the CoNLL 2003 shared task [Tjong Kim Sang and De Meulder 2003] and those used in the MUC-7 Named Entity Task, third, we introduce the idea of domain independent features (features derived from different domain data sets) on top of local and global features. We use a maximum entropy model together with a MOGA to choose the best features performance on all domain data sets. Performance improvement with the use of a domain independent gazetteer automatically generated using the method used in Nadeau et al. [2006] will also be investigated.

The recognition and classification of entities in the approach above will be achieved in a two step process; (i) Named Entity Detection [NED], where the named entities (NEs) in the target text are tagged using the illustrious BIO model

used by Tjong Kim Sang [2002] and originally proposed by Ramshaw and Marcus [1995], where a tag shows that a word is at the beginning of a NE (B), inside a NE (I) or outside a NE (O) and (ii) Named Entity

Classification [NEC] where the previously discovered named entities are then classified into predefined class types such as person name, location name, and organization name.

The novelty in our approach is that domain independence will be engineered by not only utilizing local context of a word in a sentence and other occurrences of the word in the same document, but also makes use of occurrences of each word within other domain documents to extract features (domain independent features) and the use of a multiobjective algorithm to choose out the most optimal domain independent features.

Evaluation of the approach will be realized by comparing the classification performance without using the domain independence techniques with that when the domain independence techniques have been applied over the various domain data sets. The improvement in performance using the latter demonstrates the feasibility of the approach.

1.1. Organization

The organization of this paper is as follows: after this introduction, a brief of related work is given in Section 2; The Domain Independent Named Entity Recognition System (DINERS) which includes the Data, Algorithms and Features used is illustrated in Section 3. In Section 4 the different experiments and results obtained are examined and discussed. Finally we conclude in Section 5 with a summary of the most important observations and an outlook on future work.

1.2. Related Work

The goal of the NER aims at automatically and robustly annotating named entities in large volumes of text. NER systems are required to offer good performance by being able to adapt to different domains and document genre's without much (or any) tuning. Many attempts have been made by existing NER researchers to develop systems that can successfully be tuned to new domains and applications using both hand-crafted and semi-automatic methods, however, there have been few successes in developing systems that are robust enough to automatically adapt across domains. The adaptability is mainly mired by lack of ontology and rule bottlenecks [Bontcheva et al. 2002].

Our work builds on the state-of-the-art approaches by other researchers such as Giouli et al. [2006], where in achieving domain independence their main focus was on building a homogenous, reusable and adaptable linguistic resource to different domains and languages.

On the other hand Jiang and Zhai [2006] present several strategies for exploiting the domain structure in training data to learn a more robust named entity recognizer that can perform well on a new domain. They improvise a way to automatically

rank features based on their how generalizable they are across domains. They then train a classifier with strong emphasis on the most generalizable features.

Nadeau et al. [2006] use an un-supervised strategy for domain independence by creating a system that can recognize named-entities in a given document without prior training by using automatically generated gazetteers and later resolving ambiguity.

In their work, Bontcheva et al. [2002] presented an approach to domain independent named entity recognition which is portable and was built to be usable in many different applications, on many different kinds of text and for many different purposes. They specifically showed how the system is applied for annotating (semi)-automatically digital library content and also for indexing such content by entities and events.

Most of the approaches explored employed the use of local content to design features; our approach extended to exploit external content of words in different domain text and used a strategy of MOGA that is used to empirically make a choice of the most useful features.

In this work we have used some methods which are orthogonal to those used in the related work, implying that combination of the methods is capable of improving performance of domain independence in NER.

2. The Domain Independent Named Entity Recognition System (Diners)

The Domain Independent Named Entity Recognition System (DINERS), an innovation by the authors, was developed with the capability of a machine learning approach to named entity recognition to be ported from one domain to another with very limited re-engineering. The steps followed to achieve the domain independence functionality as shown in Figure 1 below and the following detail.

1.1. The Data

In this work we use English texts from the following domains:-

2.0.1. CoNLL 2003

The English set of the CoNLL 2003 data was used for this work; the CoNLL 2003 data was originally taken from the Reuters Corpus. This corpus consists of Reuters news stories. This data set represents the general domain of news stories.

The data used was already subjected to some linguistic preprocessing and files were already formatted into four columns separated by a single space. Each word has been put on a separate line and there is an empty line after each sentence. The first item on each line is a word, the second a part-of-speech (POS) tag, the third a syntactic chunk tag and the fourth the named entity tag. The chunk tags and the named entity tags have the format I-TYPE which means that the word is inside a phrase of type TYPE. Only if two phrases of the same type immediately follow each other, the first word of the second phrase will have tag B-TYPE to show that it starts a new phrase. A word with tag O is not part of a phrase. An illustration follows in Table I.

2.0.2. Courts of Judicature (Uganda)

This data was extracted from court case judgments from the Uganda Courts of Judicature mainly divided into three categories; Supreme Court, Court of Appeal and High Court. Two files of each type were used and divided into three zones, namely the Title Line (TL), Litigant Line (LL), Presiding Judicial Officer(s) Line (PJL) and Date Line (DL).

2.1. Data Processing

The different data sets were received in different formats; some were already tagged for Part of Speech (PoS), chunk tags and named entity tags, and others were not, those not tagged were manually tagged for all the three categories by annotators from AfLaT [2008].

The different data sets were further prepared by extracting useful features from which the actual learning by the machine learning algorithm is done; the final training and testing files are built in the format shown in Table II.

2.2. Entity Types to recognize

The NE guidelines developed for the CoNLL 2003 shared task [Tjong Kim Sang and De Meulder 2003] are used as the basis for the decision of the classes of NEs used in this work. This work is however limited to Person Names, Organization Names and Location Names

Fig. 1: The Domain Independent Named Entity Recognition System(DINERS) Flowchart

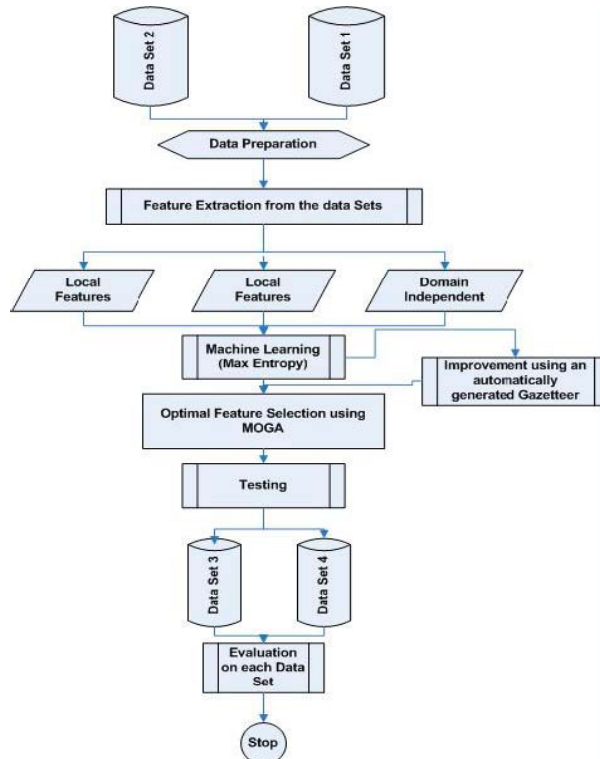


Table II. Format of the final Testing and Training files

Name Classes along with the Features				
Class	Feature1	Feature2	Feature N
Class1	binaryValue11	binaryValue21	binaryValue2N
..
..
Classn	binaryValue1n	binaryValue2n	binaryValuenN

Table I. Format of the CoNLL 2003 Data Set

Words along with the corresponding tags			
Word	PoS Tag	Chunk Tag	Feature N
U.N.	NNP	I-NP	I-ORG
official	NN	-NP	O
Ekeus	NNP	-NP	I-PER
heads	VBZ	-VP	O
for	IN	-PP	O
Baghdad	NNP	-NP	I-LOC
.	.	.	O

2.3. Feature Derivation

Features are different types of knowledge sources which are used to make tagging decisions; they are implemented a binary valued functions which query the history information in the training files and the future in the testing files to determine whether or not they fire.

2.3.1. Useful Lists

Useful lists will be generated some from without the training data while others will be derived from the training data, the following lists are derived using automated methods:-

- Frequent Word (FW) - Words that appear in more than five documents
- Rare Words (RW) - Words that appear less than five times in a document
- Useful Unigrams for each name class [UNI] - For each of the Named Entity classes (single words that come before a name); such as Honorable, Mr., Justice etc.
- Useful Bigrams for each name class [BIG] - For each of the Named Entity classes (two words that come before a name); such as His Worship, His Lordship

- Useful Name Class Suffixes (UNC) - For each of the Named Entity classes, a list of tokens that terminate them are determined; e.g. for the ORG Class we can have tokens such as Company, Limited, Inc, etc.
- Function Words for each name class [FUN] - Lower case words that occur in a name class; e.g. of the
- Gazetteer List - an important reference for information about known names

Amongst all the features some will be used for the detection stage while others will be for the classification phase. The detailed list of features is shown in Table III.

Table III. Features for Named Entity Detection and classification
Named Entity Detection Features

Feature	Local	Global	Domain Independent
Rare Words - Token word found in the list of Rare Words	✓		
Frequent Words - Token word found in the list of Rare Words	✓		
Contextual	✓		
Anchor word	✓		
([-1,...+1]) Window	✓		
Orthographic			
InitCap of Anchor Word	✓		
Capitalization of whole Anchor word	✓		
InitCap of [-1,...+1] Window	✓		
Capitalization of [-1,...+1] Window	✓		
Position of Anchor Word in Sentence	✓		
Contains digits	✓		
Contains Dollar Sign	✓		
Hyphenated	✓		
Proportion of capitalization of token word in the document	✓	✓	
Proportion of capitalization of token word in another domain documents		✓	✓
Proportion of InitCaps of token word in the document			
Proportion of InitCaps of token word in another domain document			✓
Part of Speech Features			
PoS of token word	✓		
PoS of token [-2,...+2] Window	✓		
Class Suffix			
Suffix to Anchor word has UNC in the same document	✓		
occurrence of the token word appears with the same UNC in the same document		✓	
occurrence of the token word appears with the same UNC in another domain document			✓
Unigrams			
Anchor Word has a unigram appearing in the UNI list	✓		
occurrence of the token word appears with the same UNI in the same document		✓	
occurrence of the token word appears with the same UNI in another domain document			✓
Bigrams			
Anchor Word has a bigram appearing in the BIG list	✓		
occurrence of the token word appears with the same BIG in the same document		✓	
occurrence of the token word appears with the same BIG in another domain document			✓
Function Word - Anchor words have FUN in inside it	✓		
Lexicon Features			
Lexicon feature of Anchor Word	✓		
Lexicon feature of [-1,...+1] Window Word	✓		
Anchor word is in the Gazetteer List	✓		
Anchor word is in the Trigger List	✓		
Anchor word is in the Name List	✓		

2.1. The Classification Algorithm

The algorithm used in this work is a Maximum Entropy (ME) model based on the Kullback-Leibler divergence described by Le [2004]. The model uses the form shown in Equation 1 below.

$$p(y|x) = \frac{1}{Z(x)} \exp \left[\sum_{i=1}^k \lambda_i f_i(x, y) \right] \quad (1)$$

Where; $p(y|x)$ denotes the conditional probability of predicting an outcome y given the context x , $f_i(x, y)$ are the features with the associated weighting parameters λ_i 's. k is the number of features and $Z(x)$ is a normalization factor to ensure that $\sum_y p(y|x) = 1$.

The ME model represents evidence with contextual predicates in the form in Equation 2.

$$f_{cp,y'}(x, y) = \begin{cases} 1, & \text{if } y = y' \text{ and } cp(x) = true \\ 0, & \text{otherwise.} \end{cases} \quad (2)$$

Where cp is the contextual predicate that maps a pair of outcome and context to $\{true, false\}$.

A typical example is given in Equation 3 below.

$$f_j(x, y) = \begin{cases} 1, & \text{if } word(y) = Fred \text{ and } y = I - PER \text{ (or B - PER)} \\ 0, & \text{otherwise.} \end{cases} \quad (3)$$

$word(x) = Fred$ is an example of a contextual predicate. Where I - PER means the text lies within a person name or B - PER meaning the text begins a person name.

The choice of the estimation and weighting method (Limited-Memory Variable Metric (L-BFGS)) was made using the backing of Malouf [2002], in which it was empirically discovered that the Limited-Memory Variable Metric (L-BFGS) algorithm of Benson and More' [2001] outperformed the generalized Iterative Scaling, improved Iterative Scaling function optimization algorithms such as conjugate gradient and variable metric methods .

The maximum entropy classifier is used to classify each word as one of the following: the beginning of a NE (B tag), a word inside a NE (I tag), and a word outside the tag (O tag).

During experiments, it is possible that the classifier produces a sequence of inadmissible classes (e.g., PER-B followed by LOC-L). To eliminate such sequences, we adopt from Chieu and Ng [2003] the use of a transition probability between word classes $P(c_i - c_j)$ to be equal to 1 if the sequence is admissible, and 0 otherwise. The probability of the classes c_1, \dots, c_n assigned to the words in a sequence s in a document D given other documents OD_j is defined in Equation 4 that follows:-

$$p(c_1, \dots, c_n | s, D, OD_j) = \prod_{i=1}^n P(c_i | s, D) * P(c_i | c_{i-1}) \quad (4)$$

Where $P(c_i | s, D, OD_j)$ is determined by the maximum entropy classifier. A Viterbi algorithm is then used to select the sequence of word classes with the highest probability.

2.2. The Multiobjective Genetic Algorithm (MOGA)

The multiobjective algorithm will ideally be used to select out of a pool of all the designed domain independent features an optimal set. In this work we adopted the MOGA that was used in the work of Kitoogo and Barya [2007].

2.3. Named Entity Recognition

2.3.1. Named Entity Detection

This stage is for identifying a word or sequences of words that make up the names of specific entities using the capitalization feature and a tagger to spot proper nouns. Disambiguation will be performed to solve the problem of mistaken capitalized words for named entities using the method applied in the work of Nadeau et al. [2006]. The major disambiguation tasks at this stage will be; Entity - Noun and Entity - Boundary.

Table IV. Data Sets used in the Experiments

Data Sets along with the Instances	
Data Set	Instances
Uganda Judiciary	
High Court 1	5,006
High Court 2	5,457
Court of Appeal 1	3,070
Court of Appeal 2	4,451
Supreme Court 1	3,156
<i>TOTAL</i>	21,140
CoNLL 2003	
English (Train Set)	51,577
English (Test Set A)	46,665
English (Test Set B)	204,566
<i>TOTAL</i>	302,808

Table V. Gazetteers

Name lists along with the Instances	
Data Set	Instances
Magistrates	135
Registrars	17
Judges	42

2.3.2. Named Entity Classification

At this stage, the already recognized entities are categorized into predefined classes a maximum entropy classifier using the identified feature then thereafter perform disambiguation (Entity - Entity) as in Nadeau et al. [2006].

3. Experiments

The experiments conducted in this work have been guided by the following framework and are by no means fully exhaustive; they are ideally for exploring the feasibility of the domain independence concept and DINERS specifically.

3.1. Implementation

3.1.1. Toolkit

As an implementation basis, we have used MaxEnt (Maximum Entropy Modeling Toolkit for Python and C++) toolkit developed by Le [2004], a software toolkit designed to ease the development of systems working on maximum entropy modeling [Wang et al. 2006; Hendrickx and Bosch 2004; Hendrickx and Bosch 2005]. MaxEnt offers routines for conditional maximum entropy modeling, parameter estimation and smoothing amongst others.

3.1.2. Data

We used five real-world data sets from the judgments from the Uganda Courts of Judicature and three real world data sets from the CoNLL 2003 shared task [Tjong Kim Sang and De Meulder 2003]. The two data groups (Judicial data sets and the CoNLL data sets) were integrated to form two grand data sets which were used

for both training and testing interchangeably. Details of the data sets are given in Table IV.

3.1.3. Gazetteer

A list of Magistrates, Registrars and that of Judges (Justices) of the Uganda Courts of Judicature were used as the gazetteers in these experiments. A summary of the name lists is given in Table V.

Table VI. Baseline Results

Results - CoNLL 2003 used as the Training Data Set			
NAMED ENTITY	PRECISION (%)	RECALL (%)	$F_{\beta=1}$ (%)
PERSON	91.35	91.87	91.61
LOCATION	90.23	90.05	90.14
ORGANIZATION	81.55	80.20	80.87
OVERALL	90.44	91.18	90.81
Results - Judicial Text used as the Training Data Set			
NAMED ENTITY	PRECISION (%)	RECALL (%)	$F_{\beta=1}$ (%)
PERSON	70.12	71.30	70.71
LOCATION	68.93	68.21	68.57
ORGANIZATION	66.04	65.84	65.94
OVERALL	69.55	69.16	69.35

3.2. Experimental Methodology

The two domain data sets (CoNLL 2003 and Judicial Text) are each used as the training set as well as the testing set interchangeably; i.e first the CoNLL 2003 is used as the training set and the model is tested on the Judicial text then vice-versa. The model is trained with 100 iterations of the L-BFGS method.

The first experiments do not have domain independent features employed, while in the second experiments the domain independent features have been used together with a MOGA to test for performance improvement. The experiments are further run to test the effect of the usage of gazetteers for the PERSON entity.

The performance of the DINERS is evaluated using standard precision (P), recall (R), and F-score, where F-score is defined as $2PR/(P + R)$.

For comparison of performance between the use of the different feature options ($p < 0.01$), we employed the McNemar's significance tests [Dietterich 1998].

3.3. Results

The baseline results without domain independent features are shown in Table VI. Table VII in turn shows the results when domain independent features have been used. Clearly the use of domain independent features has a positive significant impact on both precision and recall rates across all the entities in both cases, i.e. even if the training and test sets are interchanged.

The experiments reveal that there is a drop in the overall f-score performance for 90.81% to 69.35% and 92.04% to 70.27% respectively for both the baseline

results and domain independent features case when the data sets are swapped for training and testing. Another clear finding is that the CoNLL 2003 data set yields significantly better results when used as a training set.

As shown in Table VIII, the use of gazetteers indicates an improvement from 94.15% to 94.90% and 76.31 to 77.03 respectively, the difference in both cases is not statistically significant.

Table VII. Results - Application of Domain Independent Features

Results - CoNLL 2003 used as the Training Data Set			
NAMED ENTITY	PRECISION (%)	RECALL (%)	$F_{\beta=1}$ (%)
PERSON	94.08	94.22	94.15
LOCATION	91.69	91.06	91.37
ORGANIZATION	83.78	84.05	83.91
OVERALL	91.90	92.19	92.04
Results - Judicial Text used as the Training Data Set			
NAMED ENTITY	PRECISION (%)	RECALL (%)	$F_{\beta=1}$ (%)
PERSON	75.89	76.74	76.31
LOCATION	69.65	69.33	69.49
ORGANIZATION	67.25	66.43	66.84
OVERALL	70.26	70.28	70.27

Table VIII. Results - Gazetteers used with Domain Independent Features

Results - CoNLL 2003 used as the Training Data Set			
NAMED ENTITY	PRECISION (%)	RECALL (%)	$F_{\beta=1}$ (%)
PERSON	95.03	94.78	94.90
Results - Judicial Text used as the Training Data Set			
NAMED ENTITY	PRECISION (%)	RECALL (%)	$F_{\beta=1}$ (%)
PERSON	76.07	78.01	77.03

4. Conclusions

This work began with the identification of the different corpora, from which named entities were to be recognized, it was discovered that technical domains already annotated with named entities such as person name, location name, and organization name are not available and this necessitated manual annotation of some of the data used for evaluation.

In this work one of the initial steps of the identifying, manual annotation, and processing of judicial text was a leap towards an attempt to building a Judicial Corpora for named entity recognition. The judicial data set together with the CoNLL 2003 data set were used for evaluating the Domain Independent Recognition System (DINERS) developed in this research.

Our DINERS demonstrates that the use and empirically crafted combination of domain independent feature sets using optimization techniques such as MOGA yields improved performance. It can further be concluded from this work that non-technical data offers better training than technical data sets. The use of gazetteers for domain independence does not significantly improve performance.

Although the results reported in this work are affirmative, many research directions in this arena are yet to be explored. For future work, the system may be experimented for portability with other technical domains; the range of the domain independent features could also be widened. The use document sectioning (Title, Date, Body, etc.) will also be explored for possible performance improvement.

This work targeted generic named entities of Person, Location and Organization because it was difficult to identify technical entities that are shared between different technical domains, however for future work the work will be extended to a wider spectrum of technical domains to identify and work towards shared technical entities.

Acknowledgments

The authors wish to thank AfLaT [2008] for providing some of the data that was used for the experiments and Naomi Maajabu at AfLaT [2008] for the manual annotation of some of the judicial text data that have supported this work.

References

- AFLAT. 2008. African Language Technology (AfLat). <http://aflat.org/>
- BENSON, J.S. AND MORE, J.J. 2001. A limited memory variable metric method for bound constrained minimization. Preprint
- ANL/ACSP909-0901, Argonne National Laboratory.
- BONTCHEVA, K., MAYNARD, D., CUNNINGHAM, H. AND SAGGION, H. 2002. Using Human Language Technology for Automatic Annotation and Indexing of Digital Library Content. In *Proceedings of the 6th European Conference on Research and Advanced Technology for Digital Libraries*, 613 – 625.
- CHIEU, H. L. AND NG, H. T. 2003. Named entity recognition with a maximum entropy approach. In *Proceedings of the seventh conference on Natural language learning at HLT-NAACL 2003* 4, 160 – 163.
- CHINCHOR, N. 1998. MUC-7 Named Entity Task Definition (Version 3.5). In *Proceedings of the Seventh Message Understanding Conference*.
- DIETTERICH, T. G. 1998. Approximate statistical tests for comparing supervised classification learning algorithms. *Neural Comput.* 10(7), 1895 – 1923.

- ETZIONI, O., CAFARELLA, M., DOWNEY, D., POPESCU, A.M., SHAKED, T., SODERLAND, S., WELD, D. S., AND YATES, A. 2005 Unsupervised Named-Entity Extraction from the Web: An Experimental Study. *Artificial Intelligence*, 165, 91 -134.
- GIOULI, V., KONSTANDINIDIS, A., DESYPRI, E., AND PAPAGEORGIOU, X. 2006. Multi-domain, Multi-lingual Named Entity Recognition: revisiting and grounding the resources issue. In *Proceedings of the 5th Language Resources and Evaluation Conference (LREC) 2006*, Genoa, Italy.
- HENDRICKX, I. AND VAN DEN BOSCH, A. 2004. Maximum-entropy parameter estimation for the k-nn modified value-difference kernel. In *Proceedings of the 16th Belgian-Dutch Conference on Artificial Intelligence*, 19 – 26.
- HENDRICKX, I. AND VAN DEN BOSCH, A. 2005. Algorithms with Instance-based Classification. In *Proceedings of the 16th European Conference on Machine Learning*, Porto, Portugal, October 3-7 2005, 158 – 169.
- JIANG, J. AND ZHAIC, X. 2006. Exploiting Domain Structure for Named Entity Recognition. In *Proceedings of the main conference on Human Language Technology Conference of the North American Chapter of the Association of Computational Linguistics*, 74 – 81.
- KITTOGO, F.E., AND BARYAMUREEBA, V. 2007. A Methodology for Feature Selection in Named Entity Recognition, *International Journal of Computing and ICT Research in Computing and ICT Research (IJCIR)*, 1(1), 18 – 26.
- LE, Z. (2004). Maximum Entropy Modeling Toolkit for Python and C++ . *Natural Language Processing Lab*, Northeastern University, China.
- MALOUF, R. 2002. A comparison of algorithms for maximum entropy parameter estimation. In *Proceedings of the International Conference on Computational Linguistics*, 1 – 7.
- NADEAU, D., TURNEY, P., AND MATWIN, S. 2006. Unsupervised Named-Entity Recognition: Generating Gazetteers and Resolving Ambiguity. In *Proceedings of 19th Canadian Conference on Artificial Intelligence*.
- RAMSHAW, L. AND MARCUS, M. 1995. Text chunking using transformation-based learning. In *Proceedings of the Third ACL Workshop on Very Large Corpora*, 82 – 94.
- TJONG KIM SANG, E.F. 2002. Introduction to the conll-2002 shared task: Language independent named entity recognition. In *Proceedings of CoNLL-2002*, 155 – 158.
- TJONG KIM SANG, E.F. AND DE MEULDER, F. 2003. Introduction to the CoNLL-2003 shared task: Language-independent named entity recognition. In *Proceedings of CoNLL-2003*, 142 – 147.
- WANG, M., SAGAE, K., AND MITAMURA, T. 2006. A fast, accurate deterministic parser for Chinese. In *Proceedings of the 21st International Conference on Computational Linguistics and the 44th annual meeting of the ACL*, 425 – 432.

6

Towards Full Comprehension of Swahili Natural Language Statements for Database Querying

Lawrence Muchemi

Natural language access to databases is a research area shrouded by many unresolved issues. This paper presents a methodology of comprehending Swahili NL statements with an aim of forming corresponding SQL statements. It presents a Swahili grammar based information extraction approach which is thought of being generic enough to cover many Bantu languages. The proposed methodology uses overlapping layers which integrate lexical semantics and syntactic knowledge. The framework under which the proposed model works is also presented. Evaluation was done through simulation using field data on corresponding flowcharts. The results show a methodology that is promising.

1. Introduction

The quest for accessing information from databases using natural language has attracted researchers in natural language processing for many years. Among many reasons for the unsuccessful wide scale usage is erroneous choice of approaches where researchers concentrated mainly on traditional syntactic and semantic techniques [Muchemi and Narin'yan 2007]. Efforts have now shifted to interlingua approach [Luckhardt 1987 and Androutsopoulos 1995]. The problem requires syntactic and semantic knowledge contained in a natural language statement to be intelligently combined with database schema knowledge. For resource scarce languages the problem is acute because of the need to perform syntactic and semantic parsing before conversion algorithms are applied. In general database access problems should have a deeper understanding of meaning of terms within a sentence as opposed to deeper syntactic understanding. The successful solution to this problem will help in accessing huge data repositories within many organizations' and governments' databases by users who prefer use of natural language.

In this paper a methodology for comprehending Swahili queries is presented. The wider frame work for achieving the conversion to structured query language (SQL) is also explained. The scope for natural language (NL) comprehension reported in this paper is limited to the 'select' type of SQL statements without involving table joins. The approach used in this work borrows concepts from information extraction techniques as reported in Jurafsky and Martin [2003] and Kitani et al [1994] and transfer approach reported in Luckhardt [1984]. Like in information extraction, pattern search identifies terms and maps them to noun

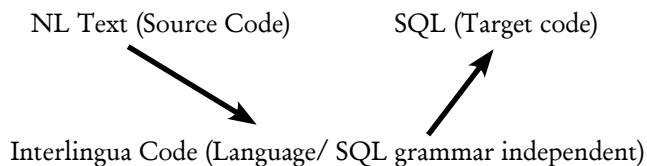
entities. A term is a word form used in communicative setting to represent a concept in a domain and may consist of one or more words [Sewangi 2001]. Templates are then used to map extracted pieces of information to structured semi-processed statements which are eventually converted to SQL code. Arrays or frames may be used to hold the entities and their meanings.

The rest of the paper first gives a synopsis of the general frameworks used for database access in natural language and shows criteria for the selection of approach. This is followed by an outline of a survey that investigates into Swahili natural language inputs. The conclusions are incorporated in the model presented.

2. General Frameworks for Data Base Access in Natural Language(NL)

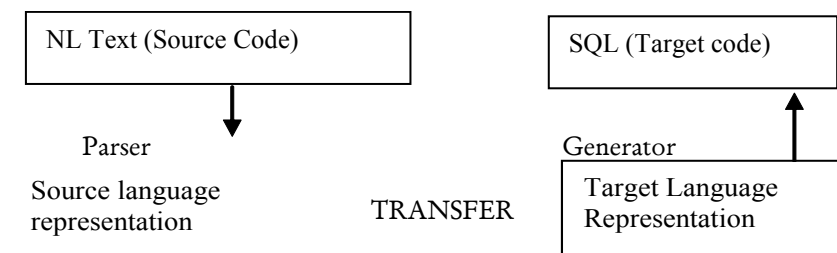
NL processing for generation of SQL statements has evolved from pattern matching systems to semantic systems and to a combination of semantics and syntactic processing [Androustopoulos, 1996]. Perhaps what is attracting researchers to a great extent today is intermediate representation language, also referred to as interlingua systems [Jung and Lee 2002]. The two dominant approaches are direct interlingua approach and transfer models. The direct interlingua which may be roughly modeled as illustrated in figure 2.1 is attractive due to its simplicity. The assumption made in this approach is that natural language may be modeled into an interlingua. An SQL code generator would then be applied to produce the expected SQL codes.

Fig. 2.1. Direct Interlingua Approach



This model has never been achieved in its pure form [Luckhardt, 1987]. Many systems use the transfer model. The transfer model, illustrated in figure 2.2 below uses two different types of intermediate code one closely resembling the source language while the other resembles the target language. This approach has experienced better success and has been used in several systems such as SUSY [Luckhardt, 1984] among others.

Fig. 2.2. Transfer Approach



Recent works in machine translation especially grammatical frameworks [Ranta 2004] have brought forth ways that inspire looking at the problem as a translation problem. However it remains to be established whether the heavy reliance on grammar formalism has a negative impact on the prospects of this approach. Rule-based machine translation relies heavily on grammar rules which are a great disadvantage when parsing languages where speakers put more emphasis on semantics as opposed to rules of grammar. This is supported by the fact that human communications and understanding is semantic driven as opposed to syntactic [Muchemi and Getao 2007].

This paper has adopted the transfer approach because of past experiences cited above. To understand pertinent issues for Swahili inputs, that any NL-SQL mapping system would have to address, a study was conducted and the results and analysis of collected data is contained in the sections here below.

3. Methodology

Swahili language is spoken by inhabitants of Eastern and central Africa and has over 100 million speakers. Only limited research work in computational linguistics has been done for Swahili and this brings about a challenge in availability of resources and relevant Swahili computational linguistics documentation. The methodology therefore involved collecting data from the field and analyzing it. The purpose was to identify patterns and other useful information that may be used in developing NL-SQL conversion algorithms. The survey methods, results and analysis are briefly presented in subsequent sections. The conclusions drawn from the initial review together with other reviewed techniques were used in developing the model presented in later sections of this paper.

A. Investigating Swahili NL inputs

Purposive sampling method as described in Mugenda and Mugenda [2003] was used in selecting the domain and respondents. Poultry farmers are likely beneficiaries of products modeled on findings of this research and were therefore selected. Fifty farmers in a selected district were given questionnaires. Each questionnaire had twenty five information request areas which required the respondent to pose questions to a system acting as a veterinary doctor. Approximately one thousand statements were studied and the following challenges and possible solutions were identified:

- a) There is a challenge in distinguishing bona fide questions and mere statements of facts. Human beings can decipher meanings from intonations or by guessing. A system will however need a mechanism for determining whether an input is a question or a mere statement before proceeding.

For example: Na weza kuzuia kuhara? {Can stop diarrhea?}

From the analysis it was found that questions containing the following word categories would qualify as resolvable queries:

- Viwakilishi viulizi {special pronouns} such as yupi, upi, lipi, ipi,,/ kipi /kupi, wapi, {which/what/where} and their plurals.
- Vivumishi viulizi {special adjectives} such as gani {which}, -pi, -ngapi {How many}
- Vielezi viulizi {special adverbs} such as lini {when}
- In addition when certain key verbs begin a statement, the solution is possible. For example Nipe {Give}, Orodhesha {list}, Futa {delete}, Ondoa {remove} etc

Hence in the preprocessing stage we would require presence of words in these categories to filter out genuine queries. In addition discourse processing would be necessary for integrating pieces of information. This can be done using existing models such as that described in Kitani et al [1994] among others.

- b) Swahili speakers are geographically widely dispersed with varying first languages. This results in many local dialects affecting how Swahili speakers write Swahili text.

Examples of Swahili statements for the sentence “Give me that book”

1. Nipe kitabu hicho Standard Swahili (Kiugunja) dialect
2. Nifee gitafu hisho Swahili text affected by Kikuyu dialect
3. Nipeako kitabu hicho ... Swahili text affected by Luhya dialect
4. Pea mimi gitavu hicho..... Swahili text affected by Kalenjin dialect

The study revealed that term structures in statements used by speakers from different language backgrounds remain constant with variations mainly in lexicon. The term structures are similar to those used in standard Swahili. This research therefore adopted the use of these standard structures. The patterns of standard Swahili terms are discussed fully in Wamitila [2006] and Kamusi-TUKI [2004]. A methodology for computationally identifying terms in a corpus or a given set of words is a challenge addressed in Sewangi, [2001]. This research adopts the methodology as presented but in addition proposes a pre-processing stage for handling lexical errors.

- c) An observation from the survey shows that all attempts to access an information source are predominantly anchored on a key verb within the sentence. This verb carries the very essence of seeking interaction with the database. It is then paramount for any successful information extraction model for database to possess the ability to identify this verb.
- d) During the analysis it was observed that it is possible to restrict most questions to six possible templates. This assists the system to easily identify key structural items for subsequent processing to SQL pseudo code. The identified templates have a relationship with SQL statements structures and are given below:
- Key verb + one Projection
 - Key Verb + one Projection + one Condition

- Key Verb + one Projection + many Conditions
- Key Verb + many Projections
- Key Verb + many Projections + one Condition
- Key Verb + many Projections + many Conditions

The terms ‘key verb’ in the above structures refer to the main verb which forms the essence of the user seeking an interaction with the system. Usually this verb is a request e.g. Give, List etc. It is necessary to have a statement begin with this key verb so that we can easily pick out projections and conditions. In situations where the key verb is not explicitly stated or appears at the middle of a sentence, the model should assign an appropriate verb or rephrase the statement appropriately. An assumption here is that most statements can be rephrased and the original semantics maintained. The term ‘projection’ used in the templates above, imply nouns which can be mapped onto field names within a selected domain and database schema. ‘Conditions’ refer to restrictions on the output if desired.

e) Challenge in the use of unrestrained NL text as input

One major challenge with unrestrained text is that questions can be paraphrased in many different ways. In the example given above the same question could be reworded in many other ways not necessarily starting with the key verb ‘give’. For example,

“Mwanafunzi mwenye alama ya juu zaidi ni nani?”...”The student having the highest grade is called who?”

In such situations it is necessary to have a procedure for identifying the essence of interaction. Information contained within a sentence can be used to assign appropriate key verbs. For example ‘ni nani’ (who) in the above example indicates that a name is being sought, hence we assign a key verb and noun; Give Name.

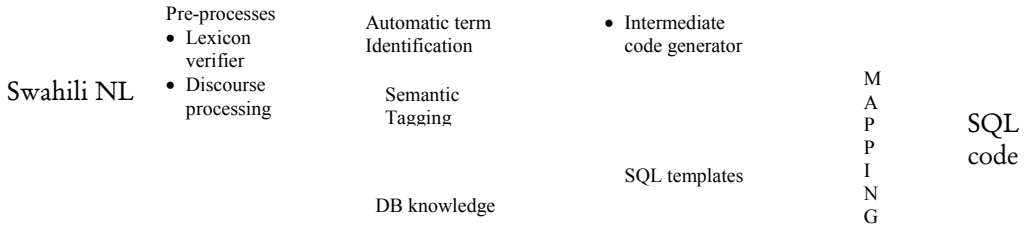
Nouns that would form the projection (Table name and column name) part of an SQL statement are then identified. This is followed by those that would form the condition part of the statement if present. Presence of some word categories signify that a condition is being spelt out. For example adjectives such as ‘Mwenye, kwenye/penye, ambapo (whom, where, given)’ signify a condition. Nouns coming after this adjective, form part of the condition. This procedure of reorganizing a statement so that it corresponds to one of the six identified templates would guarantee a solution. An algorithm for paraphrasing based on the above steps has so far been developed.

Model Architecture

The following is a brief description of the step by step processing proposed in the model. The input is unrestricted Swahili statement which undergoes pre-processing stage that verifies that the input is a genuine query, the lexicon is recognizable and there is no need for discourse processing. Terms are then generated and assembled into a suitable intermediate code. Generating intermediate code requires the use of

domain specific knowledge, such as semantics and templates proposed in section 3.1 above. The process proceeds by integrating this intermediate code with the specified database schema knowledge through mapping process and generates the SQL code. The process is illustrated in figure 3.1 here below.

Figure 3.1. The Transfer Approach Frame Work for Swahili



Steps in the generation of SQL scripts

Preprocessing

To illustrate the processes of each stage of the above model, we consider a sample statement:

“Nipe jina la mwanafunzi mwenye gredi ya juu zaidi?” “.....“Give me the name of the student with the highest grade?”

The model accepts the input as a string delivered from an interface and ensures that key words are recognizable. If not recognizable, the user is prompted to clarify. Preprocessing also involves verifying whether a statement is a resolvable query. The above statement begins with the word ‘give’, hence the statement is a resolvable query using the criteria described in section 3.1. If the statement contains pronouns and co-referential words, these are resolved at this stage. The output of this stage is a stream of verified words. The words are labeled to indicate relative position within the sentence. This forms the input of the term identification stage.

Automatic Term Identification

Term identification is the process of locating possible term candidates in a domain specific text. This can be done manually or automatically with the help of a computer. Automatic implementation involves term-patterns matching with words in the corpus or text. The model described here proposes application of automatic term identification algorithms such as those proposed in Sewangi [2001] at this stage. A tool such as the Swahili shallow syntactic parser described in Arvi [1999] may applied in identifying word categories.

Examples of term-patterns obtained through such algorithms would be:

N(noun)	Example	Jina
V(Verb)	Example	Nipe
V+N	Example	Nipe jina
N+gen connective+N	Example	Gredi ya juu

There are over 88 possible term patterns identified in Sewangi [2001] and therefore this large number of patterns would be difficult to fully present in this paper. However these patterns are used in identifying domain terms within the model.

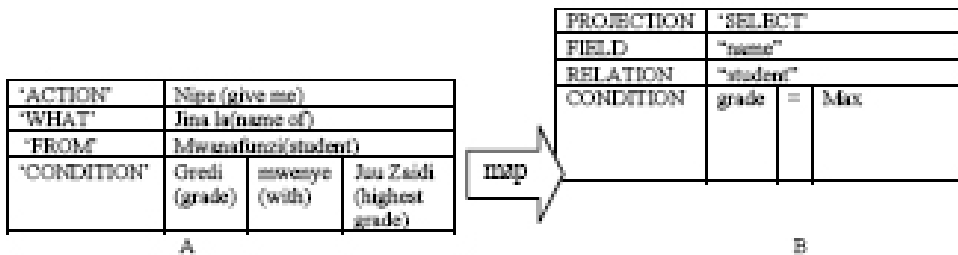
Semantic Tagging

The terms identified in the preceding stages are tagged with relevant semantic tags. These include terms referring to table names, column names, conditions etc. Information from the database schema and the specific domain is used at this stage for providing the meanings. For example, '*jina la mwanafunzi*' (*name of student*) gives an indication that column name is 'name', while table name is 'student'. Knowledge representation can be achieved through the use of frames or arrays.

Intermediate Code Generation and SQL Mapping

In the generation of intermediate code, we store the identified and semantically tagged terms in the slots of a frame-based structure shown in fig 3.2 (A) below. This can be viewed as an implementation of expectation driven processing procedure discussed in Turban et al. [2006]. Semantic tagging assists in the placement of terms to their most likely positions within the frame. It is important that all words in the original statement are used in the frame. The frame appears as shown here below:

Fig 3.2 Mapping Process



The semantically tagged terms are initially fit into a frame which represents source language representation. From a selection of SQL templates, the model selects the most appropriate template and maps the given information as shown in fig 3.2 (B) above. The generated table has a structure closer to SQL form and hence it can be viewed as a representation of the target language. This is followed by generation of the appropriate SQL code.

4. Discussions

As described, the methodology proposed here is an integration of many independent researches such as discourse processing found in Kitani [1994], automatic term identification found in Sewangi [2001], expectation-driven reasoning in frame-

structures [Turban et al. 2006] among others. The methodology also proposes new approaches in query verification as well as paraphrasing algorithm.

Research for this work is on-going. The algorithms for paraphrasing and mapping are complete and were initially tested. Randomly selected sample of 50 questions was used to give an indication of level of success. The statements were applied to flow charts based on the algorithms and the initial results show that up to 60% of these questions yielded the expected SQL queries. Long statements cannot be effectively handled by the proposed algorithm and this is still a challenge. Due to the heavy reliance on automatic term generation which relies on up to 88 patterns, there is over generation of terms leading to inefficiencies. Machine learning may improve the efficiency of the model by for instance storing successful cases and some research in this direction will be undertaken. Though not entirely successful, the initial results serve as a good motivation for further research.

5. Conclusions

This paper has demonstrated a methodology of converting Swahili NL statements to SQL code. It has illustrated the conceptual framework and detailed steps of how this can be achieved. The method is envisaged to be robust enough to handle varied usage and dialects among Swahili speakers. This has been a concept demonstration and practical evaluations would be required. However, test runs on flow charts yield high levels of successful conversion rates of up to 60%. Further work is required to refine the algorithms for better success rates.

References

- ANDROUTSOPOULOS, I., RITCHIE, G., AND THANISCH, P. 1995. Natural Language Interfaces to Databases - An Introduction. *SiteSeer- Scientific Literature Digital Library*. Penn State University, USA.
- ANDROUTSOPOULOS, I. 1996. A Principled Framework for Constructing Natural Language Interfaces to Temporal Databases. *PhD Thesis*. University of Edinburgh
- ARVI, H. 1999. Swahili Language Manager- SALAMA. *Nordic Journal of African Studies* Vol. 8(2), 139-157.
- JUNG, H., AND LEE, G. 2002. Multilingual question answering with high portability on relational databases. In *Proceedings of the 2002 conference on multilingual summarization and question answering. Association for Computational Linguistics, Morristown, NJ, USA. Vol.19, 1-8*.
- JURAFSKY, D., AND MARTIN, J. 2003. *Readings in Speech and Language Processing*. Pearson Education, Singapore, India.
- KAMUSI-TUKI. 2004. Kamusi ya Kiswahili Sanifu. Taasisi ya Uchunguzi Wa Kiswahili, Dar es salaam 2ND Ed. Oxford Press, Nairobi, Kenya.
- KITANI T., ERIGUCHI Y., AND HARA M. 1994. Pattern Matching and Discourse Processing in Information Extraction from Japanese Text. *Journal of Artificial Intelligence Research*. 2(1994), 89-110.

- LUCKHARDT, H. 1987. *Readings in Der Transfer in der maschinellen Sprachübersetzung*, Tübingen, Niemeyer.
- LUCKHARDT, H. (1984). *Readings in Erste Überlegungen zur Verwendung des Sublanguage-Konzepts in SUSY. Multilingua*(1984) 3-3.
- MUCHEMI, L., AND NARIN'YANI, A. 2007. Semantic Based NL Front-end For Db Access: Framework Adaptation For Swahili Language. In *Proceedings of the 1st International Conference in Computer Science and Informatics*, Nairobi, Kenya, Feb. 2007, UoN-ISBN 9966-7284-0-6, Nairobi, Kenya 151-156.
- MUCHEMI, L. AND GETAO, K. 2007. Enhancing Citizen-government Communication through Natural Language Querying. . In *Proceedings of the 1st International Conference in Computer Science and Informatics* ,Nairobi, Kenya, Feb. 2007, UoN-ISBN 9966-7284-0-6, Nairobi, Kenya 161-167.
- MUGENDA, A., AND MUGENDA, O. 2003. *Readings in Research Methods: Quantitative and Qualitative Approaches*. African Centre for Technology Studies, Nairobi, Kenya
- RANTA, A. 2004. Grammatical Framework: A type Theoretical Grammatical Formalism. *Journal of Functional Programming* 14(2):145-189.
- SEWANGI, S. 2001 Computer- Assisted Extraction of Phrases in Specific Domains- The Case of Kiswahili. PhD Thesis, University of Helsinki Finland
- TURBAN, E., ARONSON, J. AND LIANG, T. 2006. *Readings in Decision Support Systems and Intelligent Systems*. 7th Ed. Prentice-Hall . New Delhi, India.
- WAMITILA, K. 2006. *Chemchemi Ya Marudio* 2nd Ed., Vide Muwa, Nairobi, Kenya.

7

Generalized Association Rule Mining Using Genetic Algorithms

Peter P. Wakabi-Waiswa, Venansius Baryamureeba and K. Sarukesi

We formulate a general Association rule mining model for extracting useful information from very large databases. An interactive Association rule mining system is designed using a combination of genetic algorithms and a modified a-priori based algorithm. The association rule mining problem is modeled as a multi-objective combinatorial problem which is solved using genetic algorithms. The combination of genetic algorithms with a-priori query optimization make association rule mining yield fast results. In this paper we use the same combination to extend it to a much more general context allowing efficient mining of very large databases for many different kinds of patterns. Given a large database of transactions, where each transaction consists of a set of items, and a taxonomy (is-a hierarchy) on the items, we find associations between items at any level of the taxonomy. We show how the idea can be used either in a general purpose mining system or in a next generation of conventional query optimizers.

1. Introduction

Association rule mining (ARM) is one of the core data mining techniques. The major aim of ARM is to find the set of all subsets of items or attributes that frequently occur in many database records or transactions, and additionally, to extract rules on how a subset of items influences the presence of another subset. The ARM problem was introduced in 1993 by Agrawal et al. [1993] who developed the Apriori algorithm for solving the ARM problem. Most of the existing algorithms are improvements to Apriori algorithm [Ghosh and Nath, 2004; Zhao and Bhowmick, 2003].

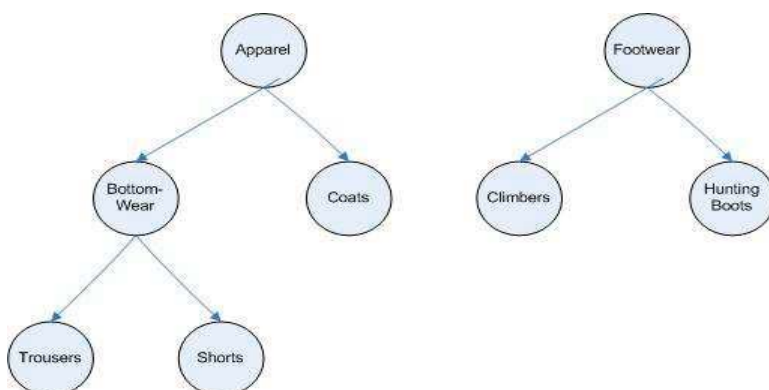
Given a set of transactions, where each transaction is a set of items, an association rule is an expression of the form $X \Rightarrow Y$ where X and Y are sets of items. This rule translates to stating that transactions in the database which contain the items in X tend to contain the items in Y . An example of such a rule might be that "20% of transactions that contain beer also contain diapers; 3% of all transactions contain both these items". Here 20% is called the confidence of the rule, and 3% the support of the rule. The support of the rule $X \Rightarrow Y$ is the percentage of transactions that contain both X and Y . The problem of mining association rules is to find all rules that satisfy a user-specified minimum support and minimum confidence.

In ARM, frequent pattern mining (FPM) is the most time-consuming part. The traditional association rule mining algorithms that extract all frequent generalized patterns are not very efficient due to the fact that a large portion of the frequent patterns are redundant. This subjects users to a drawback in that when faced with real problems they tend to get a lot of uninteresting or redundant rules along with the interesting rules. During the mining task, in most cases, taxonomies (is-a hierarchies) do exist over the items being mined.

Therefore association rules can be extracted in a generalized form conveying knowledge in a compact manner with the use of taxonomies. According to Srikant and Agrawal (1995) [10], given a large database of transactions, where each transaction consists of a set of items, and a taxonomy (is-a hierarchy) on the items, generalized association mining involves finding associations between items at any level of the taxonomy. Figure 1 shows an example of a taxonomy. The taxonomy says that trousers is-a bottom wear, shorts is-a bottom wear, bottom wear is-a Apparel. The motivation for generalized association mining is that users are interested in generating rules that span different levels of the taxonomy. In generalised association rules, application specific knowledge in the form of taxonomies over items are used to discover more interesting rules.

In this paper, we formulate a mechanism for finding for generalized association rules. This is achieved as a combination of an a-priori based algorithm for finding frequent itemsets with multi-objective evolutionary algorithms (MOEA) with emphasis on genetic algorithms (GA). The main motivation for using GAs is that they perform a global search and cope better with attribute interaction than the greedy rule induction algorithms often used in data mining tasks. We emphasize generalized association rule mining because users may be interested in generating rules that span different levels of the taxonomy. Multi-objective optimisation with evolutionary algorithms is well discussed by (Fonseca and Fleming, 1998) [4], (Khabzaoui et al. 2005) [8] and (Freitas, 2003 [5]).

Fig.1: Example of is a Taxonomy



$L = \{i_1, \dots, i_m\}$ denotes a set of literals, also referred to as itemsets. T denotes

a directed acyclic graph (DAG) on L and it is a set of taxonomies. An edge in T represents an “is-a” relationship, and T represents a set of taxonomies. The taxonomy is modelled as a DAG rather than a forest to allow for multiple taxonomies. Lower case letters denote items and upper case letters denote itemsets. x^{\wedge} denotes an ancestor of x (and x a descendant of x^{\wedge}) if there is an edge from x^{\wedge} to x in the transitive-closure of T . Note that a node is not an ancestor of itself, since the graph is acyclic. D denotes a set of transactions, where each transaction T is a set of items such that $T \subseteq L$. A transaction T is said to support an item $x \in L$ if x is in T or x is an ancestor of some item in T . We say that a transaction T supports $x \in L$ if T supports every item in X .

A generalized association rule is an implication of the form $X \Rightarrow Y$, where $X \subset L$, $Y \subset L$, $X \cap Y = \emptyset$, and no item in Y is an ancestor of any item in X . The rule $X \Rightarrow Y$, holds in the transaction set D with confidence c if $c\%$ of transactions in D that support X also support Y . The rule $X \Rightarrow Y$, has support s in the transaction set D if $s\%$ of transactions in D support $X \cup Y$. The reason for the condition that no item in Y should be an ancestor of any item in X is that a rule of the form $x \Rightarrow ancestor(x)$ is trivially true with 100% confidence, and hence redundant Agrawal (1995) [10]. We call these rules generalized association rules because both X and Y can contain items from any level of taxonomy T .

The rest of this paper is organised as follows. In Section 2 we provide an overview of work related to mining generalised association rules. In Section 3 we discuss the proposed algorithm. In Section 4, we present an overview of multi-objective optimisation and rule mining problems. In Section 5 the analysis of the results are presented. Section 6 is the conclusion.

2. Related Literature

According to Srikant and Agrawal (1995) [10], the traditional ARM approach does not consider the presence of taxonomies and it restricts the items in association rules to the leaf-level items in the taxonomy [Ghosh and Nath, 2004; Zaki, 2001]. It is valuable, however, to find rules across different levels of the taxonomy due to the following reasons:

1. Rules at lower levels may not have minimum support; thus many significant associations may not be discovered if we restrict rules to items at the leaves of the taxonomy
2. Taxonomies can be used to prune uninteresting or redundant rules.

Tsur et al. [1997] formulated Query Flocks, an algorithm for generalising the market-basket problem. They achieved this by parameterizing the query with a filter condition to eliminate values of the parameters that are uninteresting. The query flocks algorithm is a generalisation of the A-Priori technique.

Dehuri and Mall [2006] developed the multi-objective evolutionary algorithm called improved niched Pareto genetic algorithm (INPGA). It is an algorithm for mining highly predictive and comprehensible classification rules from large databases. INPGA strongly emphasizes predictive accuracy and comprehensibility of the rules.

In ARM, frequent pattern mining (FPM) is the most time-consuming part. The traditional ARM algorithms that extract all frequent patterns are not very efficient due to the fact that a large portion of the frequent patterns are redundant. This presents users with a lot of uninteresting or redundant rules. Association rules, however, can be extracted in a generalized form conveying knowledge in a compact manner with the use of taxonomies.

3. The Generalization of Association Rule Mining Algorithm (GARM)

The problem of discovering generalised association rules is handled in three steps:

1. Find all itemsets whose support is greater than the user-specified minimum support as proposed by Hipp et al. [1998].
2. Extract the rules from frequent itemsets
3. Find the generalized association rules by pruning the taxonomy using genetic algorithms. We apply comprehensibility, surprise, interestingness and confidence measures to form a coherent set of four complementary measures to extract interesting rules.

3.1. Frequent Itemset Generation

To generate the frequent itemset (FI), a lattice is formed that contains all the $2^{|L|}$ subsets of the itemset L . The FI's are generated by traversing the lattice in such a way that all frequent itemsets are found but as few infrequent ones as possible are visited. This is achieved by algorithms that use the downward closure property. The downward closure property states that: All subsets of frequent itemsets must also be frequent.

Using the support of a candidate itemset, a candidate may be pruned or added to the frequent itemsets. The strategy of the algorithm is to join every two frequent

$(k-1)$ -itemsets which have a $(k-2)$ -prefix in common. Such a join results in a candidate k -itemset. After the support of a candidate is counted, it is pruned or added to the set of frequent itemsets depending on its value. This approach starts with the 1-itemsets as the set of candidate 1-itemsets. All frequent candidates are added for further candidate generation. This ensures that all frequent itemsets are visited and reduces the number of the visited infrequent itemsets.

3.2. Lattice Traversal

There are several mechanisms for traversing the tree (lattice traversal) including breadth first search (BFS) as used in Apriori based algorithms Agrawal [1993]. In BFS the frequent (k-1) itemsets are discovered when generating the candidate k-itemsets. These algorithms exhibit good performance because they prune the candidates that have infrequent subset before counting their supports. The depth-first search (DFS) lattice traversal based algorithms on the other hand are restricted to mine only frequent k-itemsets with $k > 3$. They also do not guarantee that the infrequent subsets of a candidate C are predetermined at the time the support of C has to be determined. This limits them in that candidates having infrequent subsets cannot be pruned hence causing operating overheads. Especially when mining generalized rules this problem becomes impeding because the pruning is required for optimization purposes.

3.3. Support Counting

Counting actual occurrences of candidates as done by Apriori relies on a hash tree structure Agrawal [1993]. Counting candidates that occur infrequently is not computationally demanding. But with growing candidate sizes, this approach causes performance degradation because the number of levels of the hash tree increases. In this work, support is counted by intersecting transaction identifiers is done in Partition and Eclat, Hipp *et. al.*[1998].

3.4. Prune the Taxonomy Using Genetic Algorithms

We use a genetic algorithm to prune the taxonomy of the generated frequent itemsets (FI). A set of three complementary metrics (confidence, comprehensibility and J-Measure) is used as criteria for pruning the FIs.

4. Design of the Genetic Algorithm

In this section the design of the genetic algorithm for generalised rule induction is presented.

4.1. Knowledge Representation

A crucial issue in the design of an individual representation is to decide whether the candidate solution represented by an individual will be a rule set or just a single rule Freitas [2003]. Genetic algorithms (GAs) for rule discovery can be divided into two broad approaches, the Michigan approach and the Pittsburgh approach [Dehuri *et al.* 2006]. The biggest distinguishing feature between the two is that in the Michigan approach (also referred to as Learning Classifier Systems) an individual is a single rule, whereas in the Pittsburgh approach each individual represents an entire set of rules.

In the context of this research the use of the term Michigan approach will denote any approach where each GA individual encodes a single prediction rule. The choice between these two approaches strongly depends on which

kind of rule is to be discovered. This is related to which kind of data mining task being addressed. Suppose the task is classification. Then evaluate the quality of the rule set as a whole, rather than the quality of a single rule. In other words, the interaction among the rules is important. In this case, the Pittsburgh approach seems more natural [Frietas 2002].

On the other hand, the Michigan approach might be more natural in other kinds of data mining tasks. An example is a task where the goal is to find a small set of high-quality prediction rules, and each rule is often evaluated independently of other rules. The Pittsburgh approach directly takes into account rule interaction when computing the fitness function of an individual. However, this approach leads to syntactically-longer individuals, which tends to make fitness computation more computationally expensive. In addition, it may require some modifications to standard genetic operators to cope with relatively complex individuals.

By contrast, in the Michigan approach the individuals are simpler and syntactically shorter. This tends to reduce the time taken to compute the fitness function and to simplify the design of genetic operators. However, this advantage comes with a cost. First of all, since the fitness function evaluates the quality of each rule separately, now it is not easy to compute the quality of the rule set as a whole - i.e. taking rule interactions into account. Another problem is that, since we want to discover a set of rules, rather than a single rule, we cannot allow the GA population to converge to a single individual which is what usually happens in standard GAs. This introduces the need for some kind of niching method. The need for niching in the Michigan approach may be avoided by running the GA several times, each time discovering a different rule. The drawback of this approach is that it tends to be computationally expensive.

We have, therefore, used a modified Michigan encoding/decoding scheme which associates two bits to each attribute. If these two bits are 00 then the attribute next to these two bits appears in the antecedent part and if it is 11 then the attribute appears in the consequent part. And the other two combinations, 01 and 10 will indicate the absence of the attribute in either of these parts. So the rule $ACF \sqcap BE$ will look like 00A 11B 00C 01D 11E 00F. In this way we can handle variable length rules with more storage efficiency, adding only an overhead of $2k$ bits, where k is the number of attributes in the database. The decoding is performed as follows:

$$DV = mnv + (mxv - mnv) * \left(\sum (2^{i-1} * i^{th} \text{bitvalue}) / (2^n - 1) \right) \quad (1.1)$$

where DV is the decoded value; $1 \leq i \leq n$ and n is the number of bits used for encoding; mnv and mxv are minimum and maximum values of the attribute; and bitvalue is the value of the bit in position i . For brevity, this encoding scheme will not deal with relational operators and as such the rules generated from this formula will not include relational operators.

Due to the fact that there may be a large number of attributes in the database, we propose to use multi-point crossover operator. There are some difficulties to use the standard multi-objective GAs for association rule mining problems. In case of rule mining problems, we need to store a set of better rules found from the database. Applying the standard genetic operations only, the final population may not contain some rules that are better and were generated at some intermediate generations. The better rules generated at intermediate stages should be kept. For this task, an external population is used. In this population no genetic operation is performed. It will simply contain only the non-dominated chromosomes of the previous generation. At the end of first generation, it will contain the non-dominated chromosomes of the first generation. After the next generation, it will contain those chromosomes, which are non-dominated among the current population as well as among the non-dominated solutions till the previous generation.

The scheme applied here for encoding/decoding the rules to/from binary chromosomes is that the different values of the attributes are encoded and the attribute names are not. For encoding a categorical valued attribute, the market basket encoding scheme is used. For a real valued attribute their binary representation can be used as the encoded value. The range of values of that attribute will control the number of bits used for it.

The archive size is fixed, i.e., whenever the number of non-dominated individuals is less than the predefined archive size, the archive is filled up by dominated individuals. Additionally, the clustering technique used does not loose boundary points.

4.2. Individual encoding

In this paper an individual corresponds to a single association rule of the form where A is the rule antecedent, consisting of a conjunction of conditions on the values of the predicting attributes, and B is the rule consequent, which is constituted by a conjunction of conditions of the values of the predicted attributes. We encode the rule antecedent and the rule consequent separately. Both the antecedent and the consequent are separately encoded as a variable-length list of rule conditions. The rule's conditions can contain both propositional logic and first-order logic, in which the pair of operands is compatible, as defined in the Attribute-Compatibility Table (ACT) for the data being mined [Freitas 2003].

The genetic material of an individual's rule antecedent is randomly generated when the initial population is created and is thereafter subject to the action of crossover and mutation. In contrast, the production of the genetic material of an individual's rule consequent is treated in a special way, due to its strategic importance in determining the quality of the rule represented by the individual. The basic idea is to delay the generation of the rule consequent until fitness-computation time, when the rule consequent is generated in such a way

that the rule's predictive accuracy is maximized. A more elaborated description of this idea is as follows.

First of all, crossover and mutation operators are applied only to an individual's rule antecedent. Once these genetic operators have been applied, the just-produced rule antecedent of an offspring individual is matched against all the tuples of the database, to compute statistics to be used for calculating the individual's fitness. These statistics contain the frequency of tuples satisfying the rule antecedent, distributed per goal attribute and per goal attribute value.

After computing the statistics to determine the rule's fitness, the rule's consequent part is generated by selecting the frequency that maximizes the predictive accuracy of the rule. The main reason for determining the consequent part after the antecedent part is that it maximizes the individual's predictive accuracy for a given rule antecedent; and it leads to a gain in computational efficiency, since by matching a single rule antecedent against the database just once we are computing, in a single database pass, the predictive accuracy associated with several possible rules. This is significantly more efficient than matching the entire rule (antecedent and consequent) against the database for each possible rule. Note also that this is important because predictive accuracy computation usually by far dominates the processing time of a GA mining a large database.

4.3. Fitness Assignment

The fitness function used in this paper consists of three metrics including. We combine these metrics into an objective fitness function. The complementary set of measures include confidence defined in equation (1), comprehensibility defined in equation (2) and J-Measure defined in equation (4).

The following expression is used to quantify the comprehensibility of an association rule

$$\text{Comprehensibility} = \log(1 + |Y|) + \log(1 + |X \cup Y|) \quad (1.2)$$

where, $|Y|$ and $|X \cup Y|$ are the number of attributes involved in the consequent part and the total rule, respectively.

The confidence factor or predictive accuracy of the rule is given in the following equation

$$\text{Confidence} = \sigma(X \cup Y) / \sigma(X) \quad (1.3)$$

The J-Measure is a good indicator of the information content of the generated rules. In rule inference we are interested in the distribution of the rule "implication" variable Y , and especially its two events y and complement. The purpose is to measure the difference between the priori distribution $f(y)$, i.e. $f(Y = y)$ and $f(Y \neq y)$, and the posteriori distribution $f(Y | X)$. The J-Measure gives the average mutual information between the events y and $f(Y = y)$. The J-Measure shows how dissimilar our a priori and posteriori beliefs are about Y meaning that useful rules imply a high degree of dissimilarity. The J-Measure is calculated as:

$$JM = f(y|x) \log_2 \left(\frac{f(y|x)}{f(y)} \right) + (1 - f(y|x)) \log_2 \left(\frac{(1 - f(y|x))}{1 - f(y)} \right) \quad (1.4)$$

where JM is the J-Measure.

The fitness function is calculated as the arithmetic weighted average confidence, comprehensibility and J-Measure. The fitness function ($f(x)$) is given by:

$$f(x) = \frac{W_1 * \text{Comprehensibility} + W_2 * (J - \text{Measure}) + W_3 * \text{Confidence}}{W_1 + W_2 + W_3} \quad (1.5)$$

where W_1, W_2, W_3 are user-defined weights.

4.4. Environmental Selection

The number of individuals contained in the archive is constant over time, and the truncation method prevents boundary solutions being removed. During environmental selection, the first step is to copy all non-dominated individuals, i.e., those which have a fitness lower than one, from archive and population to the archive of the next generation. If the non-dominated front fits exactly into the archive the environmental selection step is complete. In case the archive is too small, the best dominated individuals in the previous archive and population are copied to the new archive. Otherwise, truncate the archive.

5. Experiments

The experimental design we employed in this paper was to generate a table of frequent itemsets and then generate transactions by picking itemsets from I and inserting them into the transaction. The process starts with building taxonomy over the items. The taxonomy is created by assigning children to the roots, and then to the nodes at the next depth until the frequent itemsets are exhausted.

The data used were data from supermarket customer purchases. There are 25,000 items and the taxonomy has 3 levels with 55 roots. The total number of transactions in the database is 350,000. The number of generalised rules discovered in more than one and a half times more that in cumulate. This shows a good performance of the designed algorithm. We evaluated the performance of this algorithm and compared it with an implementation of Cumulate [10]. Default values of the parameters are: Population size = 40, Mutation rate = 0.5, Crossover rate = 0.8, Selection in Pareto Archive (elitism) = 0.5. The stopping criterion used is the non evolution of the archive during 10 generations, once the minimal number of generations has been over passed.

5.1 Results and Discussion

Minimum Support was varied from 2.5% to 0.5%. GARM was faster than Cumulate by approximately 1.5, with the performance gap increasing as the

minimum support decreased. The number of transactions was varied from 1,000 to 10,000. GARM performed better than Cumulate though they tend to converge as the number of transactions grows.

6. Conclusion And Future Works

In this paper we have combined a-priori query technique with a genetic algorithm to deal with generalisation of the association rule mining problem. The results show that our proposed model can attain considerable performance improvement in terms of minimum support and number of items. The intended future improvements include consideration of various other characteristics that a good generalisation algorithm should have. These characteristics include a limitless number of roots and/or levels in the taxonomy, depth-ratio, and number of transactions. We also hope subject the algorithm to larger sample sizes and different data types.

References

- AGRAWAL R., IMIELINSKI T., SWAMI A., (1993). Mining Association Rules Between Sets of Items in Large Databases. Proc. of the 1993 ACM SIGMOD Conf. on Management of Data
- AGRAWAL R., SRIKANT R.(1994). Fast Algorithms for Mining Association Rules. Proc. of the 20th Int'l Conf. on Very Large Data Bases
- DEHURI S., JAGADEV K. , GHOSH A. AND MALL R. (2006). Multi-objective Genetic Algorithm for Association Rule Mining Using a Homogeneous Dedicated Cluster of Workstations. American Journal of Applied Sciences 3 (11): 2086–2095, 2006 ISSN 1546-9239, 2006 Science Publications
- FONSECA M. C. AND FLEMING J. P. (1998) Multi-objective Optimization and Multiple Constraint Handling with Evolutionary Algorithms-Part I: A Unified Formulation. IEEE Transactions on Systems, Man and Cybernetics - Part A: Systems and Humans, 28(1):26-37
- FREITAS A. A., (2003) A Survey of Evolutionary Algorithms for Data Mining and Knowledge Discovery. Advances in evolutionary computing: theory and applications, Pp 819–845
- GHOSH A. AND NATH B.,(2004). Multi-objective rule mining using genetic algorithms. Information Sciences 163 pp 123- 133
- HIPP J., MYKA A., WIRTH R., DIERIGER, AND GUNTZER. (1998). A new Algorithm for Faster Mining of Generalised Association Rules. Proceeding of the 2nd European Symposium on Principles of Data Mining and Knowledge Discovery (PKDD '98), pages 75–82, Nantes, France, September 1998. Springer.
- KHABZAOU M., DHAENES C., AND TALBI E. (2005). Parallel Genetic Algorithms for Multi-Objective rule mining. MIC2005. The 6th Metaheuristics International Conference, Vienna, Austria. SRIKANT R. AND AGRAWAL R. (1995). Mining Generalized Association Rules. Proceedings of the 21st VLDB Conference Zurich, Switzerland, 1995

- TSUR D., ULLMAN J. D., ABITEBOUL S., CLIFTON C., MOTWAMI R., AND NESTOROV S. (1997). Query Flocks: A Generalization of Association Rule Mining.
- ZAKI, M.J. (2001). Generating non-redundant association rules. In Proceedings of the Sixth ACM-SIGKDD International Conference on Knowledge Discovery and Data Mining, New York, NY: ACM, 34-43.
- ZHAO Q. AND BHOWMICK S. S. (2003). Association Rule Mining: A Survey. Technical Report, CAIS, Nanyang Technological University, Singapore, No. 2003116 , 2003.

8

Collaborative Filtering: A Comparison of Graph-Based Semi-Supervised Learning Methods and Memory-Based Methods

Rasna R. Walia

Collaborative filtering is a method of making predictions about the interests of a user based on interest similarity to other users and consequently recommending the predicted items. There is a widespread use of collaborative filtering systems in commercial websites, such as Amazon.com, which has popularized item-based methods. There are also many music and video sites such as iLike and Everyone's a Critic (EaC) that implement collaborative filtering systems. This trend is growing in product-based sites. This paper discusses the implementation of graph-based semi-supervised learning methods and memory-based methods to the collaborative filtering scenario and compares these methods to baseline methods such as techniques based on weighted average. This work compares the predictive accuracy of these methods on the MovieLens data set. The metrics used for evaluation measure the accuracy of generated predictions based on already known, held-out ratings that constitute the test set. Preliminary results indicate that graph-based semi-supervised learning methods perform better than baseline methods. However, some of the memory-based methods outperform the graph-based semi-supervised learning methods as well as the baseline methods.

1. Introduction

Collaborative filtering is basically a method of matching people with similar interests for the purpose of making recommendations. The basic assumption of collaborative filtering is that those who agreed in the past tend to agree in the future. An example of a collaborative filtering system in use is Amazon.com, where new books are recommended to users based on what they have previously bought as well as their similarity to other users.

The task of collaborative filtering is split into **prediction** and **recommendation**. Collaborative prediction refers to the task of predicting preferences of users based on their preferences so far, and how they relate to the preferences of other users. On the other hand, collaborative recommendation is the task of specifying a set of items that a user might like or find useful. In this work, focus was on the collaborative prediction task and the testing of different algorithms for their accuracy.

Collaborative filtering systems are also distinguished based on whether they use **implicit** or **explicit** votes. Explicit voting refers to a user expressing his preference

for an item, usually on a discrete numerical scale. On the other hand, implicit voting arises from interpreting user behaviour to estimate a vote or preference. Implicit votes are often based on things like browsing data, purchase history or other types of information access patterns. Throughout this work, explicit votes were used to form predictions, because of their simplicity and natural interpretation.

Collaborative filtering algorithms suggest new items or predict their usefulness for a particular user based on the user's previous likings and the opinions of other like-minded users. Typically, there is a list of m users $U = \{u_1, u_2, \dots, u_m\}$ and a list of n items $I = \{i_1, i_2, \dots, i_n\}$. Each user u_b is associated with a list of items I_{ub} on which he has expressed a liking, between a range of 0-5, with 1 meaning least preferred, 5 meaning most preferred and 0 meaning the item has not been rated by the user. The entire $m \times n$ user-item data is represented as a ratings matrix, R . Each entry r_{ij} in R represents the rating of the i^{th} user on the j^{th} item. An example ratings matrix with 5 users and 10 movies is given below:

$$\begin{bmatrix} 0 & 0 & 5 & 0 & 0 & 0 & 2 & 0 & 4 & 3 \\ 0 & 3 & 0 & 4 & 2 & 0 & 0 & 5 & 0 & 0 \\ 1 & 0 & 2 & 0 & 0 & 3 & 0 & 0 & 1 & 2 \\ 0 & 5 & 1 & 0 & 4 & 0 & 1 & 0 & 0 & 0 \\ 2 & 4 & 0 & 0 & 5 & 0 & 0 & 4 & 0 & 2 \end{bmatrix}$$

This paper contains three contributions:

- The application of graph-based semi-supervised learning methods within the collaborative filtering domain.
- Implementation of several methods on a non-trivial sized data set i.e. the MovieLens data set, with 100,000 ratings from 943 users on 1,682 movies.
- Alteration to memory-based methods: User-based methods were applied to items and a significant improvement in the accuracy of the predictions was observed.

This paper is organized as follows: Section 2 briefly discusses related work. Section 3 outlines the experimental methods employed, including a description of the data set and error measures used. The results are presented in Section 4. Section 5 presents a discussion of the results obtained and Section 6 outlines avenues for future work.

2. Related Work

A number of collaborative filtering techniques have been proposed, of which the most popular ones are those based on the correlation criteria and matrix factorization.

Breese et al. [1998] carried out an empirical analysis of the predictive accuracy of several memory-based algorithms like techniques based on the correlation coefficient and vector-based similarity as well as model-based methods like the Naïve Bayes formulation and Bayesian Networks. They evaluated their algorithms on the MS Web data set, Nielsen data set and EachMovie data set. A key observation from their work is that the performance of the different methods is greatly influenced by the nature of the data set and the availability of votes with which to make predictions.

Results generated by memory-based methods are generally quite accurate. A major drawback of these methods is that they are computationally very expensive because the similarity between each pair of users needs to be computed before predictions can be made on the desired items. These algorithms also cannot detect cases of item synonymy. Another major disadvantage of memory-based methods is that they do not construct any explicit statistical models so nothing is really “learnt” from the available user profiles.

A major computation in memory-based methods is the calculation of similarities between users. Algorithms where predictions are made based on the similarity computation among items have been implemented by Sarwar et al. [2001]. These methods are known as item-based collaborative filtering algorithms. Unlike the user-based methods, the item-based approach looks into the set of items the target user has rated and computes how similar they are to the target item i and then selects k most similar items. Once the most similar items have been found, the prediction is computed by taking a weighted average of the target user’s ratings on the similar items. In cases where large datasets are being worked on, these methods provide better quality recommendations than user-based methods [Sarwar et al. 2001]. Item-based methods allow similarity computations among items to be performed in advance, leading to faster recommendations for a user.

In many cases, the collaborative prediction task is viewed as a classification problem [Marlin 2004]. Algorithms like the k -nearest neighbour classifier among others have been implemented in Marlin’s work. Semi-supervised learning is also considered a classification method. It targets situations where labelled data are scarce and unlabelled data are in abundance. Semi-supervised learning on a graph has been studied from different perspectives in Belkin and Niyogi [2004], Herbster et al. [2005] and Herbster and Pontil [2006]. A common theme in all these papers is the use of the graph Laplacian, L .

One of the characteristics of the data used within the collaborative filtering framework is sparsity. The user-movie matrix, where rows represent the users and columns represent the movies, has relatively few actual ratings and has many missing values which are represented as zeros. Dimensionality reduction methods and matrix decomposition techniques such as singular value decomposition (SVD) based prediction algorithms can overcome the sparsity problem by utilizing the latent relationships that are captured in the low-rank approximations of the user-movie matrix. [Sarwar et al. 2000], [Kleeman et al. 2006]

Several model-based methods such as Bayesian networks [Breese et al. 1998] and clustering [Connor and Herlocker 2001], [George and Merugu 2005] have been applied to the collaborative filtering domain. These different methods provide item recommendations by first building a model of user ratings. The Bayesian network model formulates a probabilistic model for the collaborative filtering problem while the clustering model treats it as a classification problem. The clustering model works by similar users and/or similar items in the same class and estimating the probability that a particular user belongs to certain class C and from there, computes the conditional probability of ratings.

In this work, graph-based semi-supervised learning algorithms and memory-based algorithms have been implemented and compared with each other because they work in a similar manner. Both methods operate over the entire user database to generate the required prediction for a particular movie for the given active user.

2. Experimental Method

1.1 Methods Employed

There are generally two broad classes of collaborative filtering algorithms. **Memory-based** algorithms operate over the entire user database to make predictions while **model-based** algorithms use the user database to learn a model which is then used for prediction. The emphasis of this work is on memory-based methods.

Semi-supervised learning is a class of machine learning techniques that makes use of both labelled and unlabelled data for training. Typically, such learning problems are characterized by a small amount of labelled data with a large amount of unlabelled data. It has been found that the use of labelled data in conjunction with unlabelled data can produce considerable improvement in learning accuracy. The graph-based learning algorithms presented in this paper are all categorized as semi-supervised learning methods.

The problem of collaborative filtering lends itself well to the semi-supervised learning framework since an analogy can be drawn between labelled data and user ratings for movies as well as the unknown ratings and unlabelled data.

3.1.1 *Graph-Based Semi-Supervised Learning Algorithms*

Two graph-based semi-supervised learning algorithms were used in this work:

- **Minimum Norm Interpolation:** This method uses a Laplacian kernel to predict labels on the unlabelled data.
- **Harmonic Energy Minimizing Functions:** This method uses the graph Laplacian and labels on the labelled data to predict the labels on the unlabelled data.

3.1.1.1 Hilbert Space of Functions on a Graph

Functions defined on the graph are represented by a Hilbert space associated with the graph Laplacian. Let $G = (V, E)$ be an undirected graph with a vertex set $V = \{1, \dots, n\}$, edge set $E(G) = E \subseteq \{(i, j) : i < j\}_{i, j \in V}$ and $n \times n$ adjacency matrix $\mathbf{A} = (A_{i,j} : i, j \in V)$ such that $A_{i,j} = A_{j,i} = 1$ if $(i, j) \in E$ and zero otherwise. The graph Laplacian \mathbf{L} is the $n \times n$ matrix defined as $\mathbf{L} = \mathbf{D} - \mathbf{A}$, where $\mathbf{D} = \text{diag}(d_1, \dots, d_n)$ and d_i is the degree of vertex i . There are l labelled points and u unlabelled points. Usually, $l \ll u$ and $n = l + u$ is the total number of points in the graph.

3.1.1.2 Minimum Norm Interpolation

Let $R(G)$ be the linear space of real-valued functions defined as $\mathbf{g} = (g_1, \dots, g_n)^T$, where ‘‘T’’ denotes transposition. A linear subspace $H(G)$ of $R(G)$ is defined which is orthogonal to the eigenvectors of \mathbf{L} with zero eigenvalue, that is,

$$H(G) = \{\mathbf{g} : \mathbf{g}^T \mathbf{u}_i = 0, i = 1, \dots, r\}$$

Since G is connected \mathbf{L} has only one eigenvector with eigenvalue zero (the constant vector) and therefore

$$\text{Equation 1: } H(G) = \{\mathbf{g} : \sum_{i=1}^n g_i = 0\}$$

Within this framework, the aim is to learn a classification function $\mathbf{g} \in H(G)$ on the basis of a set of labelled vertices. \mathbf{g} is obtained as the minimal norm interpolant in $H(G)$ to the labelled vertices, i.e., the unique solution to the problem:

$$\text{Equation 2: } \min_{\mathbf{g} \in H(G)} \{\|\mathbf{g}\| : g_i = y_i, i = 1, \dots, l\}$$

The reproducing kernel of $H(G)$ is the pseudo-inverse of the Laplacian $\mathbf{K} = \mathbf{L}^+$. With the representer theorem, the coordinates of \mathbf{g} are expressed as:

$$\text{Equation 3: } g_i = \sum_{j=1}^l K_j c_j$$

The solution of Equation 3 is given by $\mathbf{c} = \tilde{\mathbf{K}}^+ \mathbf{y}$ where $\tilde{\mathbf{K}} = (\tilde{K}_j)_{i,j=1}^l$.

In this work, the following symmetric positive definite graph kernel has been used:

$$K_c^b = \mathbf{L}^+ + b\mathbf{I}^T + c\mathbf{I}, (0 < b, 0 \leq c)$$

This method assumes that all labels are centred around zero and hence, all ratings in the data set are shifted relative to one of the following: User average, Movie average, Weighted average, User median, Movie median.

3.1.1.3 Harmonic Energy Minimizing Functions

There are l labelled points and u unlabelled points. L denotes the labelled set and U denotes the unlabelled set of points. In this formulation, the labels on the labelled

data lie between two intervals a and b such that $y_i = y_{L(i)} \in [a, b]$ for $i = 1, \dots, l$. This constraint is denoted by $y|_{L=y_l}$. To satisfy the requirement that unlabelled points that are nearby in the graph have similar labels, the *energy* is defined to be:

$$\text{Equation 4: } E(y) = \sum_{i,j \in E(G)} (y_i - y_j)^2$$

so that low energy corresponds to a slowly varying function over the graph.

The graph Laplacian matrix and known labels are used to calculate the labels of unlabelled data points. The harmonic property means that the value at each unlabelled node is the average of the neighbouring nodes. The harmonic energy minimizing function f is computed with matrix methods. The Laplacian matrix L is partitioned into blocks for labelled and unlabelled nodes as follows:

$$L = \begin{bmatrix} L_l & L_u \\ L_u & L_u \end{bmatrix}$$

Let $f = \begin{bmatrix} f_l \\ f_u \end{bmatrix}$ where $f_l = y_l$ and f_u denotes the mean values on the unlabelled data points. The solution is given by:

$$\text{Equation 5: } f_u = -L_u^{-1} L_u f_l$$

This method also requires labels to be centered on zero and so ratings are shifted in the user-movie matrix relative to one of the following: User average, Movie average, Weighted average, User median, Movie median.

The basic idea in both the methods is that the learnt function should be *smooth* with respect to the graph. The smoothness in the minimum norm interpolation method and harmonic energy minimizing functions method is ensured by Equation 1 and Equation 4, respectively.

3.1.2 Memory-Based Algorithms

Memory-based algorithms operate over the entire user database to make predictions. Given an unknown test rating (of a test item by a test user) to be estimated, memory-based collaborative filtering methods first measure similarities between the test user and other user **user-based methods**. After that, the unknown rating is predicted by taking a weighted average of the known ratings of the test item by similar users. In this work, different user-based methods were implemented, as described by Breese et al. [1998] and they are distinguished mainly by the method used to calculate the ‘‘weight’’.

Prediction Computation

The basic task in collaborative filtering is to predict the votes of a particular user, the active user, from a database of user votes, by finding other users similar to the

active user. The user database consists of a set of votes $v_{b,j}$, corresponding to the vote for user b on movie j . I_b is the set of movies on which user b has voted. The mean vote of user b is defined as:

$$\bar{v}_b = \frac{1}{|I_b|} \sum_{j \in I_b} v_{b,j}$$

The votes of the active user, a , are predicted based on some partial information regarding the active user and a set of weights calculated from the user database. To calculate the predicted vote of the active user for item j , a weighted sum of votes of the other users in the database is used.

$$\text{Equation 6: } p_{a,j} = \bar{v}_a + \hat{e} \sum_{b=1}^n w(a,b)(v_{b,j} - \bar{v}_b)$$

where n is the number of users in the collaborative filtering database with non-zero weights and \hat{e} is a normalizing factor such that the absolute values of the weights sum to unity. $w(a,b)$ is either the distance, correlation or similarity between each user b and the active user a .

An important term in Equation 6 is $w(a,b)$, leading to the generation of different predictions depending on how it is calculated. The table below summarizes different methods used for calculating $w(a,b)$.

Table 1: Different Weight Calculating Methods

$w(a,b)$	Method
Distance	Euclidean, Manhattan
Correlation	Pearson Correlation
Weight/Similarity	Vector Similarity, Default Voting, Inverse User Frequency

In this work, the correlation and weight/similarity methods were used to calculate $w(a,b)$.

3.1.2.2 Pearson Correlation

The correlation between users a and b is:

$$w(a,b) = \frac{\sum_j (v_{a,j} - \bar{v}_a)(v_{b,j} - \bar{v}_b)}{\sqrt{\sum_j (v_{a,j} - \bar{v}_a)^2 (v_{b,j} - \bar{v}_b)^2}}$$

where the summations over j are over the movies for which both users a and b have recorded votes.

3.1.2.4 Vector Similarity

In this formulation, each user is treated as a vector of votes in n -dimensional space, where n is the number of votes in the user database. The weight between two users is computed by calculating the cosine of the angle formed by the vectors. The weight is now calculated as:

$$w(a,b) = \sum_j \frac{v_{aj}}{\sqrt{\sum_{k \in I_a} v_{ak}^2}} \frac{v_{bj}}{\sqrt{\sum_{k \in I_b} v_{bk}^2}}$$

where the squared terms in the denominator serve to normalize votes so that users that vote on movies will not *a priori* be more similar to other users. Again, the summations over j are over the items for which both users a and b have recorded votes.

3.1.2.4 Default Voting

Default voting is an extension to the correlation algorithm. Correlation, as a similarity measurement, does not work very well on sparse data sets. When two users have few movies in common, their weights tend to be over-emphasized. Default voting deals with this problem by adding a number of imaginary items that both have rated in common in order to smooth the votes. A default vote value d is assumed as a vote for movies which do not have explicit votes. The same default vote value d is taken for some number of additional items k that neither user has voted on. This has the effect of assuming that there are some additional number of unspecified movies that neither user has voted on, but they would generally agree on. In most cases, the value of d is selected to reflect a neutral or somewhat negative preference for the unobserved movies. The weight is now calculated as:

$$w(a,b) = \frac{(n+k)(\sum_j v_{aj}v_{bj} + kd^2) - (\sum_j v_{aj} + kd)(\sum_j v_{bj} + kd)}{\sqrt{((n+k)(\sum_j v_{aj}^2 + kd^2) - (\sum_j v_{aj} + kd)^2)((n+k)(\sum_j v_{bj}^2 + kd^2) - (\sum_j v_{bj} + kd)^2)}}$$

where the summations j are now over the union of items that either user a or b has voted on $I_a \cup I_b$ and $n = |I_a \cup I_b|$.

3.1.2.5 Extension

In this work, user-based methods were also applied to items. The user-based methods were modified to replace values pertaining to ‘users’ with corresponding ‘movie’ values. These methods are referred to as **user-based methods on items** in this work. To predict the vote for user a on item j Equation 6 is modified slightly to become:

Equation 7:
$$p_{aj} = \bar{v}_j + \hat{e} \sum_{i=1}^n w(j,i)(v_{ai} - \bar{v}_i)$$

where n is the number of items in the collaborative filtering database with non-zero weights and $\frac{1}{\sum w(j,i)}$ is a normalizing factor such that the absolute values of the weights sum to unity. $w(j,i)$ is either the distance, correlation or similarity between each item i and the active item j . Two user similarity computation methods, vector similarity and default voting, were employed to calculate the similarities between movies.

3.1.3 *Baseline Methods*

In order to determine the relative performance of the various algorithms implemented for the collaborative filtering problem in this work, the most basic prediction methods were used. These methods involve predicting an unknown rating by returning one of the following statistics as the predicted value: (a) User Average; (b) Movie Average; (c) Weighted Average.

3.2 Data

The different methods were evaluated on the **MovieLens** dataset. The fastest of methods generally ran for ten minutes on this dataset and the slowest took up to thirty minutes. The dataset consists of 100,000 ratings (in the range of 1 - 5) from 943 users on 1,682 movies. A null vote i.e. a zero entry for a movie, means that the movie has not been watched by the user. Each user has rated at least 20 movies. Previous work carried out on the MovieLens dataset in relation to collaborative filtering includes:

- Singular value decomposition based prediction by Sarwar et al. [2000].
- Item-based collaborative filtering methods by Sarwar et al. [2001].
- Different matrix factorization techniques such as maximum margin matrix factorization, incremental SVD and repeated matrix reconstruction by Kleeman et al. [2006].

This work presents a novel application of graph-based semi-supervised learning methods in the collaborative filtering domain. A similar graph-based semi-supervised approach has been used to address the sentiment analysis task of rating inference for unlabelled documents. [Goldberg and Merugu 2005]

3.3 Error Measures

Statistical accuracy metrics were used to evaluate the quality of predictions of the different algorithms that were implemented. Also known as predictive accuracy metrics, they measure how close the predicted ratings are to the true user ratings in the test set.

3.3.1 *Mean Absolute Error*

Mean absolute error (MAE) measures the average absolute deviation between a predicted rating and the user's true rating. If the number of predicted votes in the test set for the active user is m_a , then the mean absolute error for the user is:

$$S_a = \frac{1}{m_a} \sum_{j \in P_a} |p_{a,j} - v_{a,j}|$$

where $p_{a,j}$ is the predicted rating for user a on movie j and $v_{a,j}$ is the actual rating of user a on movie j . These scores are then averaged over all the users in the test set. The lower the MAE, the more accurate is the prediction.

3.3.2 Root Mean Squared Error

Root mean squared error (RMSE) is a slight variation to the MAE. It squares the error before summing it up. This results in more emphasis on large errors.

$$RMSE = \sqrt{\frac{\sum e^e}{N}}$$

where e is the error at each point and N is the number of points tested. The lower the RMSE, the more accurate is the prediction.

3.4 Experimental Protocol

Experiments using the different methods were carried out in a similar manner. The data sets used in the experiments were partitioned into training sets and test sets. The methods were trained on the training sets and their performance evaluated on the test sets. In all cases, error rates taken over the set of held-out ratings used for testing and not the set of observed ratings used for training are reported. The RMSE and MAE values presented in the experiments are average error rates across multiple test sets.

4. Results

To compare the predictive accuracy of the different methods, two large datasets provided by MovieLens were used. One pair is ua.base and ua.test and the other pair is ub.base and ub.test. These datasets split the main dataset into a training and test set with exactly 10 ratings per user in the test set. The sets ua.test and ub.test are disjoint. The training sets (*.base) have 90,570 ratings and the test sets (*.test) have 9,430 ratings.

The results presented in Table 2 are for the following methods:

- **Baseline Methods:** User Average, Movie Average, Weighted Average
- **Standard Methods - User-Based Methods:** Pearson Correlation, Vector Similarity and Default Voting
- **Variation of Standard Methods - User-Based Methods on Items:** Vector Similarity and Default Voting
- **Graph-based Semi-supervised Learning Methods:** Minimum Norm Interpolation and Harmonic Energy Minimizing Functions

Table 2: Predictive Accuracy of the Different Methods

Methods	Data Set			
	RMSE		MAE	
	ua.test	ub.test	ua.test	ub.test
User Average	1.043	1.06	0.833	0.849
Movie Average	1.043	1.05	0.836	0.841
Weighted Average	1.009	1.016	0.812	0.819
Pearson Correlation	0.971	0.987	0.7639	0.7771
Vector Similarity	0.973	0.989	0.7642	0.7771
Default Voting	0.989	1.002	0.777	0.787
Vector Similarity (on items)	0.955	0.972	0.751	0.766
Default Voting (on items)	0.583	0.591	0.447	0.455
Minimum Norm Interpolation	0.98	0.988	0.785	0.791
Harmonic Energy Minimizing Functions	0.989	0.998	0.793	0.8

5. Discussion

The MovieLens data set used in the experiments contains 100,000 ratings from 943 users on 1,682 movies. Out of a possible 1,586,126 ratings, only 100,000 ratings are present in the data set. Due to the sparse nature of the data set, it was expected that the graph-based semi-supervised learning methods would perform better than the other methods employed. The reason for this was that these graph-based semi-supervised learning methods use both labelled and unlabelled data to build better classifiers.

However, the observation was that the performance of the graph-based semi-supervised learning methods were almost the same as that of the memory-based methods. The performance of the minimum norm interpolation method is approximately the same on both datasets. The harmonic energy minimizing functions method returns a slightly higher rate of error than the minimum norm interpolation method. Promising results were achieved when applying user-based methods to items. Baseline methods, such as using the weighted average as the predicted value, gave fairly good predictions.

Generally, the graph-based semi-supervised learning methods perform better than the baseline methods. However, memory-based methods such as those that use Pearson correlation, vector similarity and default voting to calculate weights and generate predictions performed better than the graph-based methods.

6. Future Work

There is a growing need for the use of collaborative filtering systems because of the increasing volume of customer data available on the Web and the growth of product-based websites. New technologies are continuously being exploited to improve the predictive accuracy of collaborative filtering systems.

This study shows that graph-based semi-supervised learning methods is one of the techniques that can be used in the collaborative filtering domain. Further research is needed to understand under what conditions certain methods work well and others do not. Future work includes implementation and observation of the performance of the methods presented in this paper to larger datasets, such as the Netflix dataset, which contains over 100 million ratings from approximately 480,000 users on 17,700 movies. Further analysis and experiments also need to be carried out to look for ways to improve the performance of the various methods presented in this paper.

In addition to the above, a comparative study of the methods presented in this paper with other methods such as singular value decomposition, non-negative matrix factorization, clustering algorithms, Naive Bayes classifier and Bayesian networks needs to be carried out. It is believed that these model-based methods actually “learn” something from user profiles and item characteristics and are therefore expected to give better prediction accuracy than memory-based methods.

Acknowledgements

Sincere thanks to Dr. Mark Herbster (m.herbster@cs.ucl.ac.uk) of the Department of Computer Science, University College London, for introducing me to the field of collaborative filtering and developing my interest in it.

I would also like to thank the GroupLens Research Project at the University of Minnesota for making the MovieLens dataset available for use.

References

- BELKIN, M. AND NIYOGI, P. [2004]. Semi-Supervised Learning on Riemannian Manifolds. *In Machine Learning*, 56, pages 209 – 239.
- BREESE, J.S., HECKERMAN, D. AND KADIE, C. 1998. Empirical Analysis of Predictive Algorithms for Collaborative Filtering. *In Proceedings of the Fourteenth Annual Conference on Uncertainty in Artificial Intelligence*, pages 43 – 52, July 1998.
- CONNOR, M. AND HERLOCKER, J. [2001]. Clustering Items for Collaborative Filtering. *In SIGIR-2001 Workshop on Recommender Systems*.
- GEORGE, T. AND MERUGU, S. [2005]. A Scalable Collaborative Filtering Framework Based on Co-Clustering. *In Proceedings of the Fifth IEEE International Conference on Data Mining*, pages 625 – 628.
- GOLDBERG, A.B. AND ZHU, X. [2006]. Seeing stars when there aren’t many stars: Graph-based Semi-supervised Learning for Sentiment Categorization. *HLT-NAACL 2006 Workshop on Textgraphs: Graph-based Algorithms for Natural Language Processing*.
- HERBSTER, M., PONTIL, M. AND WAINER, L. [2005]. Online Learning over Graphs. *In ICML 2005*, pages 305 – 312.
- HERBSTER, M. AND PONTIL, M. [2006]. Prediction on a Graph with a Perceptron. *In NIPS 2006*.

- KLEEMAN, A., HENDERSEN, N. AND DENUIT, S. Matrix Factorization for Collaborative Prediction. *ICME*.
- MARLIN, B. [2004]. Collaborative Filtering: A Machine Learning Perspective. *Master's Thesis, University of Toronto*.
- SARWAR, B.M., KARYPIS, G., KONSTAN, J.A. AND RIEDL, J.T. [2000]. Application of Dimensionality Reduction in Recommender Systems – A Case Study. *In ACM WebKDD Workshop, 2000*.
- SARWAR, B.M., KARYPIS, G., KONSTAN, J.A. AND RIEDL, J.T. [2001]. Item-Based Collaborative Filtering Recommendation Algorithms. *In Proceedings of the Tenth International World Wide Web Conference*, pages 285 – 295, May 2001.
- ZHU, X., LAFFERTY, J. AND GHAHRAMANI, Z. [2003]. Combining Active Learning and Semi-Supervised Learning using Gaussian Fields and Harmonic Functions. *In ICML, 2003*.

PART 2



Information Systems

9

A Simulation Modelling-Based Investigation of the Impact of IT Infrastructure on Business Process Reengineering

Innocent Baguma and Joseph K. Ssewanyana

The literature on business process reengineering (BPR) has promoted business process reengineering as an essential element in helping organizations to achieve their goals. The literature has further claimed that Information Technology (IT) is a major enabler of business process reengineering. Despite all these claims, there has been hardly any effort to demonstrate the interrelationship between business process and information technology infrastructure, and which modeling technique can be used to demonstrate the relationship. This paper investigates how simulation modeling can demonstrate the impact of IT infrastructure on business process. The dynamic synthesis methodology is used with data collected from five commercial banks. The same collected data is used to validate the model. The results indicate that network infrastructure and data integration were significantly related to the process improvement thrust, such that network infrastructure is critical in improving business processes and enhancing customer service.

1. Introduction

Since the 1990s there has been a growing research interest in business process reengineering (BPR), also referred to as business process innovation (Davenport, 1993; Hammer, 1990 and Davenport and Short, 1990). The BPR-paradigm, defined as the radical redesign of business processes enabled by information technology (IT) in order to achieve dramatic improvements in their performance (Hammer, 1990), has been described abundantly in managerial journals (Earl, 1994; Hall, 1993; Hammer, 1990, Davenport and Short, 1990). Several management concepts, principles, guidelines, checklists and step approaches have been proposed to demonstrate how to make BPR work. In spite of all these, many BPR projects have failed. Mistakes are realized after the redesigned processes are implemented, and it is too late, costly, and difficult to reverse the situation (Hlupic and Robinson, 1998). To minimize the impact of these failures, simulation modeling has been used to identify the bottlenecks in the processes and to think of various alternative solutions before implementation. Simulation models can provide graphical display that can interactively be edited and animated to demonstrate the dynamics of the processes.

Although the technological environment has been changing since the industrial revolution, there has been little or no effort to show the imminent interrelationship between Business process and Information Technology infrastructure. Many have argued that IT infrastructure has little or no impact on business process both intra and inter-organization

(Attaran; 2004, Lee, 2004; Meel, 1995; Davenport, 1993). Furthermore, despite the many simulation environments that allow organizations to model both elements, there has been no attempt to model and identify the impacts and changes that one variable will have on the other (Painter, 1996). This paper investigates how simulation modeling can demonstrate the impact of IT infrastructure on Business Process Reengineering. The paper starts by discussing business process reengineering and IT infrastructure, followed by simulation modeling and system dynamics, conceptual model and hypothesis, methodology, results and discussion, and lastly model validation before the conclusion.

2. Business Process Reengineering and IT Infrastructure

There are several literatures on business process re-engineering that support business processes as an essential element for organizations to achieve their goals (Ray and Serrano, 2003). How the business processes are designed and deployed determine the profitability and performance of the organization. BPR makes the business process efficient, effective and flexible. It is designed to meet and often exceed customer expectations. For an organization to maximize benefits from business process re-engineering, information technology plays a big role.

The IT infrastructure has been defined as the extent to which data and applications through communication networks can be shared and accessed for organizational use (Broadbent et al, 1999; Wyse and Higgins, 1993). The main purpose of IT infrastructure is to provide consistent and quick information support to the organization so as to respond to the dynamic changes in the market.

The IT infrastructure and BPR are interdependent in the sense that deciding the information requirements for the new business processes determines the IT infrastructure constituents, and recognition of IT capabilities provides alternatives for BPR (Ross, 1998; Venkatraman, 1993). Building a responsive IT infrastructure is highly dependent on an appropriate determination of business process information needs, which is based on the various activities in the business processes.

3. Simulation Modeling and Systems Dynamics

The increasing popularity of simulation has resulted in its widespread use for modeling and analysis of systems in various application areas, such as manufacturing, transport, logistics, communication networks, health and military. Shannon (1975) has defined simulation as the process of designing a model of a real system and conducting experiments with this model for the purpose, either of understanding the behavior of the system or evaluating various strategies (within the limits imposed by the criterion or set of criteria) for the operation of the system.

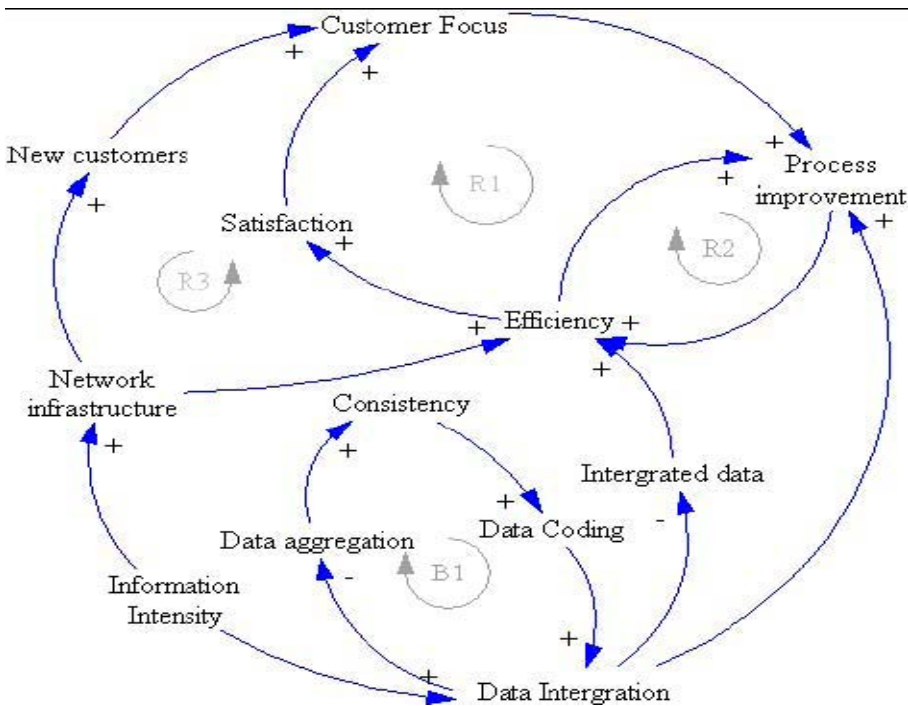
Simulation models provide a potentially powerful tool for conducting controlled experiments by systematically varying specific parameters and rerunning the model (Tumay, 1995).

System Dynamics (SD) enables a company to look beyond individual issues towards a broader perspective. SD uses computer simulations to create knowledge based on the data and information available. Computer simulations performed at different time intervals, generate results which can help in forecasting, solving problems and developing policies. Some of the benefits of simulation are the possibilities of exploring new policies, operating procedures and many others, without disrupting on going operations, and the ability to gain insights into system variables.

4. Conceptual Model and Hypothesis Development

Figure 1 presents the conceptual model of the underlying hypothesis proposed in this paper. It shows the interaction and relationship of variables using cause and effect analysis that determine the behavior of variables over time. It contains four dominant feedback loops of which 3 are reinforcing (R) and the other one balancing loop (B)

Fig 1: Conceptual Model



The model has four major variables to assess the impact of IT infrastructure on business process reengineering. Firstly, the network infrastructure that focuses on the integrative capability to coordinate different and disparate functions and activities. It helps in sharing IT resources within and across the boundaries of a firm. The measures include flexibility, connectivity, efficiency, accessibility, effectiveness and design. Secondly, the process improvement thrust which is the driving force of data integration, information intensity and network infrastructure on business process redesign. The measures include change, efficiency, coordination, flexibility, accuracy, response time and reliability.

Thirdly, data integration which is the standardization of data definition codes, and formats that allow organizations to easily manage their activities and compare their performance with similar and comparable functional units (Goodhue et al, 1992). The introduction of a shared-access database often introduces new ways of working, as many sequential processes can be handled in parallel (Barua and Lee, 1997). The measures are data coding, data consistency, accuracy and timeliness. And lastly, customer focus that looks at activities directed towards meeting and often exceeding customer expectations. It involves tracking of customer complaints, and the causes of their dissatisfaction, and subsequently corrects the root cause of customer dissatisfaction so they can provide innovative products and services (Davenport, 1993). The measures are quality of product or service, response rate, reliability, and innovativeness.

5. Hypothesis Development

The justification for each variable and the derived hypothesis are presented below. Information intensity of the industry and type of service are two moderator variables included in the study. The type of service, at firm-divisional level, is categorized into portfolio banking and general banking and currency. Information intensity of the industry is defined as the extent to which products and services of the divisions are dependent on information (Glazer, 1991). Complex services, such as the long-term mortgage and credit facilities extension, require a lot more information than those of simple services, such as deposit and withdraw of money. Moreover, in the operation of complex services, the contents of information also increase, as customer requirements become quite specific. In some cases, increasing information contents in the services enable customers to order for customized services, thus creating the need to capture, store, and manipulate customer-related information.

Research Hypothesis 1

The strengths of network infrastructure in regard to business lie in its integrative capability to co-ordinate different and disparate functions and activities (Davenport 1993). For example, electronic mail, video conferencing and computer to computer links such as Electronic Data Interchange increases not only the support to organizational processes by increasing the range and depth of information about

business activities but also make it feasible for team members to co-ordinate their activities synchronously (across time zones) and geographically (across remote locations) more easily than ever before (Rockart and Short, 1991).

H1. *Higher level of networks infrastructure is associated with higher level of process improvement thrust.*

Research Hypothesis 2

Data integration solutions are very common in corporate and governmental organizations, where they continually acquire, merge, and transport data. When central management recognizes the interdependence of multiple units within the same function, it often enforces standards of data definition and data coding to co-ordinate activities for higher performance (Davenport, 1993).

It is presumed among many researchers that data integration is beneficial to a point of integrating disparate business processes, but beyond that data integration may not be useful (Goodhue *et al.*, 1992). However, several studies argue that as the firms begin to become familiar with data integration methodologies, businesses will move toward a better fit between business planning and IS planning. For example, Teng *et al* (1995) reported that the key for aligning business planning with IS planning is related to the ability of the firm to plan for integrated databases. Furthermore, Richardson *et al.* (1990) in a case study with Texaco and Star Enterprises found the importance of data integration in businesses. They argued that data needs should be managed as corporate asset and data planning should be integral part of business process improvement.

H2. *Higher level of data integration is associated with higher level of Process improvement thrust*

Research Hypotheses 3 and 4

Some organizations generate a lot of information in their operations and are likely to make use of sophisticated IT infrastructure to meet customer demands in products and services (Glazer, 1991). Tasks and activities which were information intensive are affected by the use of IT, and often need high level of IT support to satisfy customer demands in products and services. Moreover, in higher information intensive environments, companies need a higher level of IS support for customizing products or services to meet customer demands (Glazer, 1991).

H3. *The greater the degree of information intensity of the industry, the stronger the effect of networks infrastructure on process improvement thrust*

H4. *The greater the degree of information intensity of the industry, the stronger the effect of data integration on process improvement thrust.*

Research Hypothesis 5

Many practitioners believe that the critical advantage of IS networking lies in its integrative capability. By integrating different core tasks and departments, a company can take advantage of efficiency and effectiveness (Davenport, 1993). The activities that were done serially can be done in parallel, reducing the cycle time by as much as 80 percent (Stalk and Hout, 1990). According to Boynton (1993), the creation of information architecture serves as a flexible structure to access and handle the rapidly changing demands of businesses. Systems built around information architecture provide enough flexibility to adapt to new products and production processes.

Railing and Housel (1990) in his study with TWRs space and defense sector found that the integrated communication networks offered advantages through quick response, cost reduction and fast decision making. Likewise, Kriebel and Strong (1984) in a survey of 48 firms found that the leading edge firms made more use of integrated systems as compared to the low-end firms. In a study of the insurance firms, Venkatraman and Zaheer (1990) found positive effects of electronic integration in the property and casualty market. Furthermore, Hall and McCauley (1987) reported that well integrated communication systems were related to product cycle time reduction, customer satisfaction and proactive management response.

H5. *A higher level of network infrastructure is associated with a higher level of customer Focus*

Research Hypothesis 6

There is a need for accurate, precise, and timely information when dealing with customers. Businesses aim at finding out customers requirements but turning these requirements into products and services requires a change in management philosophy. Scheter (1992) noted that to satisfy the timeliness constraint, inspecting and testing of products or processes must be done promptly. Accumulation of old data is often useless and mislead about the current problems. An integrated database support offers flexibility in collecting relevant data and analyzing problems quickly. At the same time, data integration reduces data redundancies and inconsistencies in customer information.

H6. *The greater the degree of information intensity of the industry, the stronger the effect of data integration on customer focus.*

6. Methodology

The methodology applied a systems dynamics approach based on the dynamic synthesis methodology (DSM) as the framework for analysis, model building and simulation (Williams 2002), to assess the impact of IT infrastructure on business processes. This was an iterative research process that covered: problem statement,

field study, system dynamics model building, case studies, simulation experiments and model use, and theory extension, as described and illustrated by Williams (2002).

Dynamic Synthesis Methodology (DSM) refers to the integration of theoretical concepts and structuring of parts and elements of a process overtime in such a manner to form a formal functional entity by synthesis as a philosophy of science. Synthesis is an attempt to fuse the findings of various branches of science into coherent view, in order to explain why things operate the way they do (Williams, 2002).

The study combined a case study research method and system dynamic modeling approach as they provide the qualitative information that is used to understand the problem domain in the more detail.

The study was conducted in five commercial banks that were purposively selected due to their willingness to participate in the study, and were at the same time interconnected on the Bankom network. The principal data collection instrument was the questionnaire which was administered to 100 purposively selected IT personnel and middle level managers. The questionnaire was followed up with interviews. The Cronbach Alpha model of internal consistency was applied to establish the reliability of the instrument. An alpha of 0.9194 was generated which confirmed reliability of the instrument. On the other hand, to determine the validity of the instrument and its items, a Content Validity Index (CVI) was used and a CVI of 0.92 was generated which confirmed that the properties of measurement scales and the items of the instrument were reliable and acceptable since the CVI was greater than 0.5 (Grant and Davis, 1997; Rubio et al, 2003). The questionnaire administration was followed by interviews.

The causal loop diagram (CLD) in Figure 2 was created using the Vensim PLE software. The CLD is a tool that is used to identify, analyze and communicate feedback loop structures. It helps to understand the cause and effect variables, and how changes manifest in the problem whether positive or negative. The Stock and Flow diagram in Figure 3 was developed with the STELLA (version 8.1) software. It was constructed from the causal loop diagram.

Fig 2: Causal loop diagram

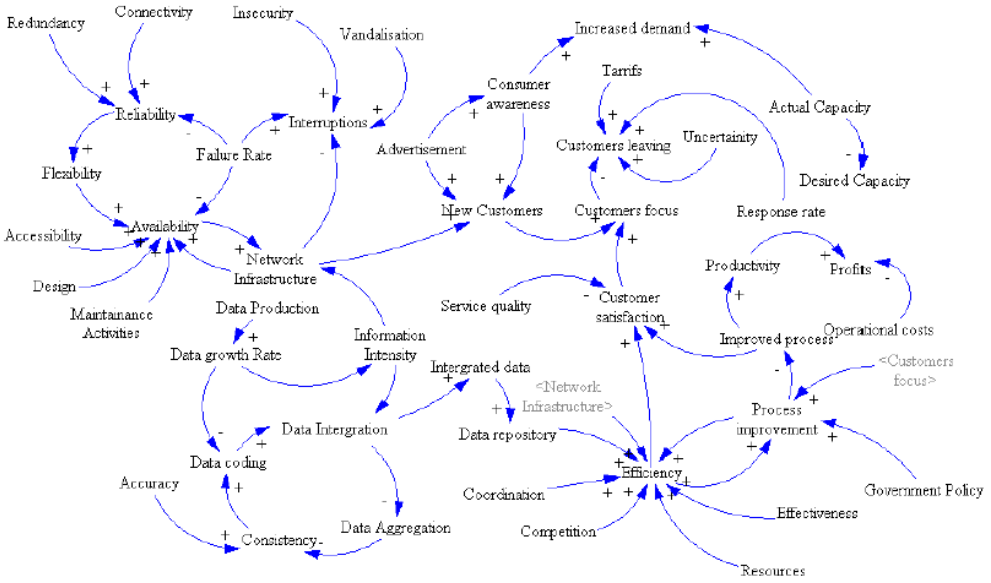
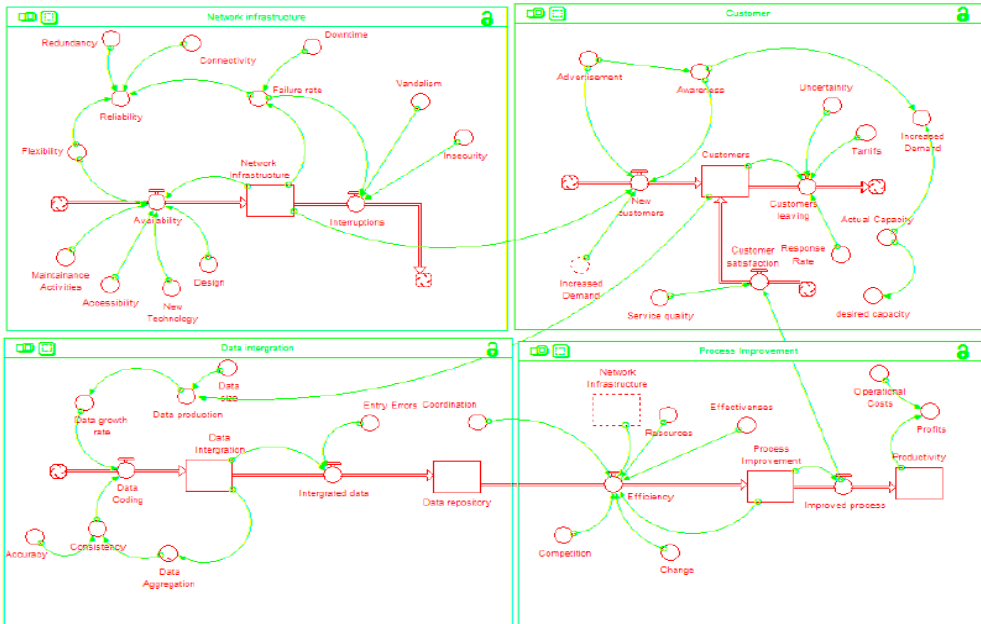


Fig 3: Stock and flow diagram



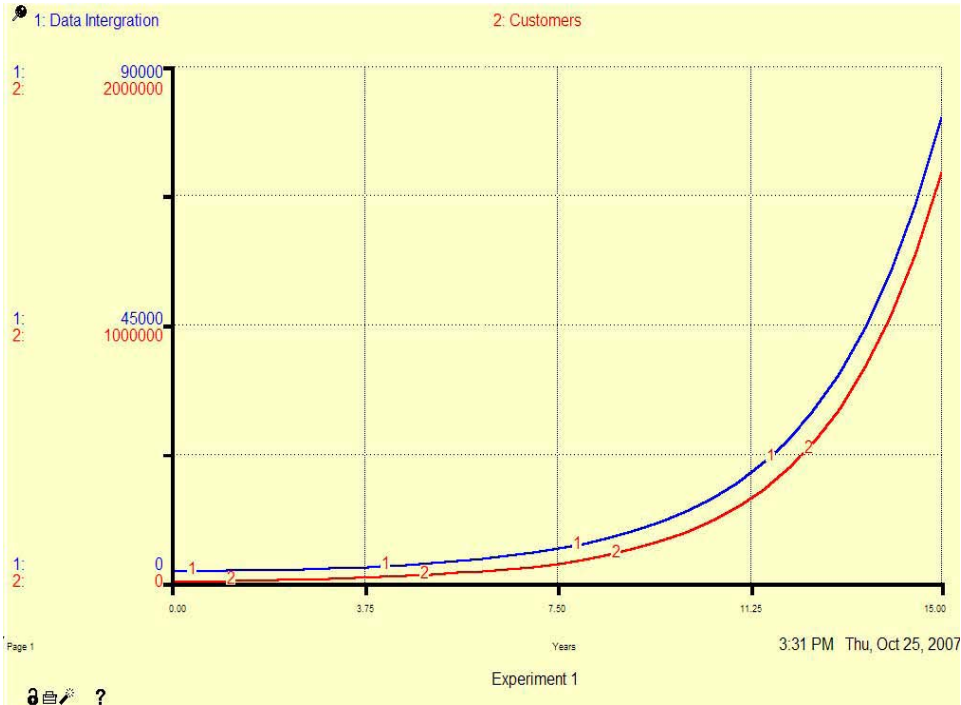
7. Results and Discussions

The results from the simulation are discussed below together with the validation of the model.

Simulation results

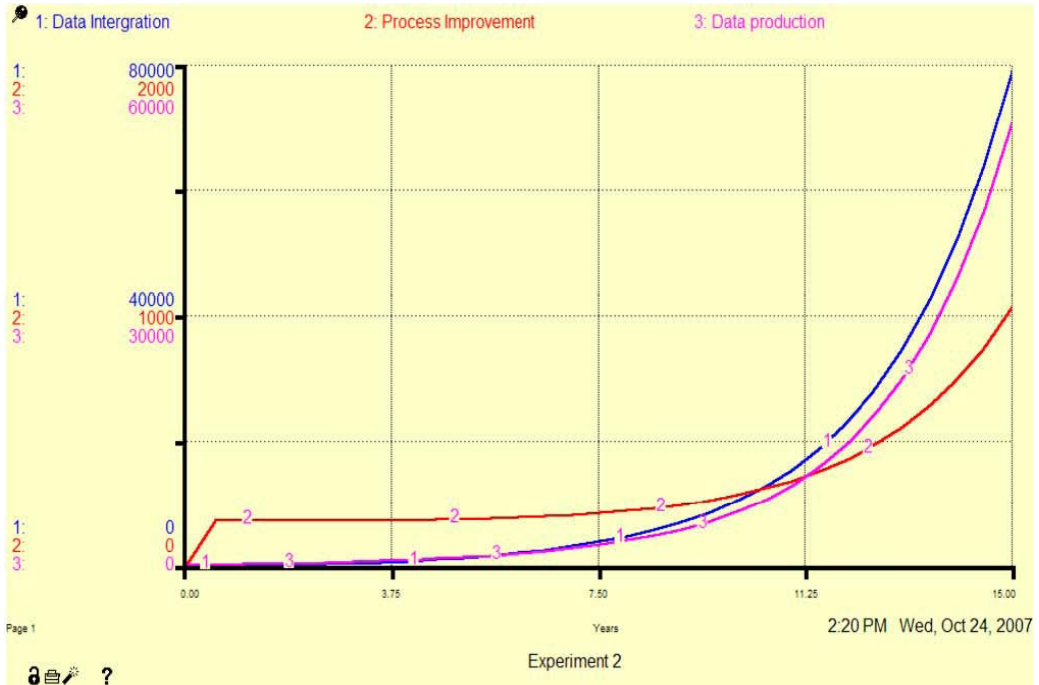
Six key variables were used in generating the simulation results that were defined in the dynamic hypothesis. The figures 4, 5, and 6 present the simulation outputs based on the model.

Fig 4: Simulation Experiment 1



The first simulation of the model examined data integration and customer focus. One observes a behavior defined by one of the key propositions in the dynamic hypothesis that a higher level of data integration (curve 1) leads to a higher level of customer focus (curve 2). In the first 4 years the number of customers is low at about 20500 customers, likewise not much data is integrated, in the later years as the number of customers grows as a result of improved service delivery because of the introduction of new technologies, integrated data from both the new customers and old ones steadily rises in the subsequent years. It can be observed that towards the 15th year customer growth continues to rise because the integration of data makes service provision much more effective and efficient and this leads to customer satisfaction and continued customer inflows. Also this is because the need for accurate, precise, and meaningful information is indispensable in customer focus by integrating core tasks and departments, banks can take advantage of efficiency and effectiveness. The creation of information architecture serves as a flexible structure to access and handle the rapidly changing demands of customers, because networks offer advantages through quick response, and cost reduction, this in turn lead to an increase in customer focus.

Fig 5: Simulation Experiment 2

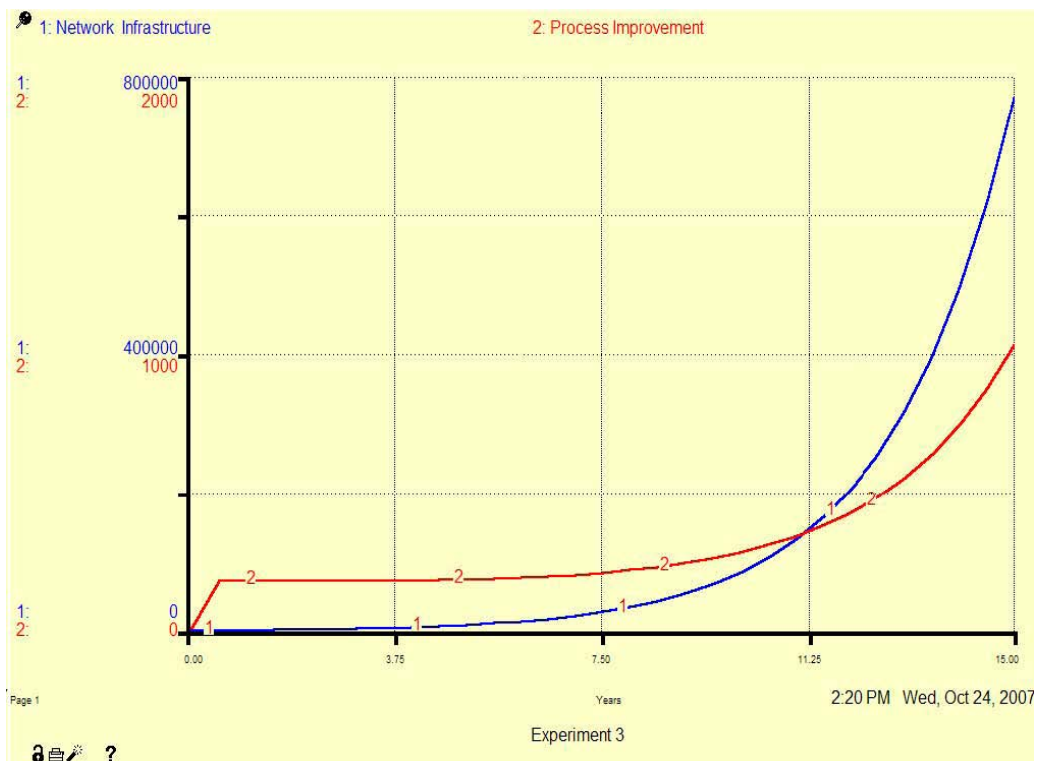


The second simulation examined data integration, process improvement and data production. The simulation output shows, a rise in data production (curve 3) as a result of the increasing number of customers. Due to the need to access services from different points, there is a need for data Integration (curve 1) which makes it possible to have shared access to data bases. One also observes that process improvement (curve 2) will rise slowly and this is fueled by the fact that data that was integrated can be accessed easily from shared databases, thus reducing delays in service provision. In year 1, process improvement increases sharply because available capacity exceeds required capacity; this is due to the low number of customers. As the customers' numbers grow, process improvement stabilizes rising at almost the same rates as compared to data production and data integration in the second half of the year 2 and this continues till the 12th year. In the 13th year, process improvement increases at a decreasing rate. The rate at which data production and data integration grow is higher as compared to process improvement. From the simulation experiment, a behaviors defined by one of the key propositions that a higher level of data Integration (curve 1) is associated with a highest level of process improvement (curve 2) can be deduced. This can also be supported by the fact that data Integration reduces data redundancies and inconsistencies in customer information.

The third simulation examines the behaviour of network infrastructure and process improvement. The output from the simulation shows that an increase in network infrastructure is associated with a higher level of process improvement.

This can be explained by the fact that management recognises the interdependence of multiple units and the need to co-ordinate different and disparate functions and activities such as electronic mail, peer to peer links such as those integral to Electronic Data Interchange by putting in place the necessary infrastructure. This not only increases support to organisational processes by increasing the range and depth of information about business activities but also makes it feasible for team members to co-ordinate their activities (across time zones) and geographically (across remote location) more easily than ever before. In the first six years network growth is slow because banks roll out a few equipment as they test for customer acceptance. In the later years network growth rises steadily because new access points are created to cater for the increasing number of customers, to reduce congestion in certain service centres so as to improve on service delivery. This is later followed by continuous process improvement.

Fig 6: Experiment 3



The results from the simulation experiments have indicated that a strong network infrastructure, customer focus, and data integration can lead to process improvement thrust. The results provide support to the contention that an integrated technology environment is one of the important considerations in business improvement initiatives. The findings regarding network infrastructure dimensions were consistent with earlier studies on BPR (Bhatt G, 2000). Secondly

it was observed that data integration is equally important in determining overall customer focus. Specifically creation of databases appears to rank slightly high when integrated database support offers flexibility in collecting relevant data and analyzing problems quickly. These results are consistent with Porter (1985) argument that with use of IT, organizations can provide a higher value to customers.

8. Model Validation

The model was validated with empirical analysis, and the findings had similarities. Overall, the model shows a good fit to the data as evidenced by the significant F-values. The t-values accompanying the individual coefficients are obtained using SPSS package. The statistical significance of all individual relationships provides strong empirical support for the developed model. Networks infrastructure was found to be significantly related with process improvement thrust ($f=4.09$, $\text{sig}=0.032$) (p is less than 0.05), where as the relationship between network infrastructure and customer focus was not found to be significant ($f=2.691$, $\text{sig}=0.105$) (p less than 0.01), (Hypothesis 1 and 3). Data integration was also found to be significantly related with process improvement thrust ($f=16.334$, $\text{sig}=0.000$) (p is less than 0.01) and process improvement ($f=9.234$, $\text{sig}=0.03$) (p is less than 0.05) (Hypothesis 2 and 4). The effect of information intensity of the industry was not found to be significantly moderating the relationship between networks infrastructure and process improvement thrust, Similarly information intensity had little effect on customer focus and data integration, also Information intensity had little effect on process improvement thrust and customer focus (Hypothesis 5, and 6).

Fig 7: Model Validation Results

Hypothesis	Relationship	Coefficient	t-Value	P-value	Conclusion	Adjusted R2 (%)	F	Sig
H1	Network Infrastructure → Process Improvement	0.255**	3.99	<0.05	Fail to reject H1	5.2	4.09	0.032
H2	Data integration → Process Improvement	0.412**	6.711	<0.01	Fail to reject H2	15.9	16.334	0.000
H3	Network Infrastructure → Customer focus	0.192	5.979	>0.01	Fail to reject H3	2.3	2.691	0.105
H4	Data integration → Customer focus	0.318**	8.003	<0.01	Fail to reject H4	9	9.234	0.003
H5	Information intensity → Network Infrastructure → Process Improvement	-0.18	4.209	<0.05	Fail to reject H5	7.1	3.69	0.030
H6	Information intensity → Data Integration → Network Infrastructure → Process Improvement	-0.124	5.434	<0.01	Fail to reject H6	16	8.262	0.000
H7	Information intensity → Infrastructure → Customer focus	-0.057	5.412	<0.01	Fail to reject H7	2.1	1.762	0.179
H8	Information intensity → Data Integration → Customer focus	-0.057	5.826	<0.01	Fail to reject H8	8	4.553	0.014

9. Conclusion

Despite the continuing debate on the effect of IT on BPR, the strategic potential of IT in BPR cannot be denied, even though incompatible islands of IT may be a hindrance to BPR. The role of network infrastructure is critical to improve business processes and enhance customer services, by enabling sharing of real time information throughout the organization. Moreover a network infrastructure

enables coordination between people, regardless of their physical locations and background. The IT infrastructure is an area where bank management should focus a large proportion of their resources, as there is growing evidence that customers associate quality of service with the bank's possession of a good IT infrastructure.

The current banking environment is ripe for the use of simulation. The pressure to control costs is higher than ever, so, there is a critical need for powerful tools which can help bank executives and administrators make good decisions on how to achieve objectives of reducing costs while maintaining high quality service. The highly stochastic nature of banking processes, as well as the complexity of subsystem interactions, makes simulation the decision-support tool of choice for analyzing the delivery of banking services.

10. Limitations and Future Research

In spite of the good findings, the study had limitations which were influenced first and foremost by resource constraints in terms of time and money. Secondly, the banks were reluctant to divulge detailed information about the utilization of IT facilities and functions in their institutions. Thirdly, the survey was conducted in only Kampala, extending the study to other areas would provide a better understanding of the way the hypothesized factors influence BPR and how managers can effectively apply Information Technology. More work need to be done to obtain more data to support the findings. Potential areas of research include analyzing the relationship of effective redundancy with factors such as reliability, cost of operations and maintenance, life-cycle costs for equipment and profitability of banking institutions.

Future studies can examine different dimensions of IT infrastructure in depth, such as comparing the role of IT expertise, and organizational capabilities in BPR. Researchers may also be interested in exploring the different contextual factors which could unduly influence the strength of the effect of IT infrastructure on BPR.

References

- ATTARAN, M. (2004). "Exploring the Relationship Between Information Technology and Business Process Reengineering," *Information and Management*, Vol.41, No.5,: 585-596
- BARUA A., LEE B. (1997), An Economic Analysis of the Introduction of an Electronic Data Interchange systems. *Information Systems Research*. Vol 8 No. 4: 98 – 422.
- BHATT, G. (2000), "Exploring The Relationship Between Information Technology, Infrastructure And Business Process Reengineering", *Business Process Management Journal*, 6(2):139-63.
- BOYNTON A.C. (1993), Achieving Dynamic stability through Information Technology, *California. Management Review* 58 – 77.

- BROADBENT, M, WEILL, P., ST. CLAIR, D (1999), "The Implications Of Information Technology Infrastructure For Business Process Redesign", *MIS Quarterly*, bf 23(2):159-82.
- DAVENPORT, P. H. (1993). *Process Innovation: Reengineering Work Through Information Technology*. Harvard Business School Press, Boston.
- DAVENPORT, T.H. AND SHORT, J.E. (1990), "The New Industrial Engineering: Information Technology And Business Process Redesign", *Sloan Management Review*, 31(4):11-27
- EARL, M. (1994) The New And The Old Of Business Process Redesign, *Journal of Strategic*
- EL SAWY O.A. (2001) *Redesigning Enterprise Processes for e-Business*. McGraw-Hill Irwin.
- GIAGLIS, G.M., PAUL, R.J., AND SERRANO, A. (1999), Reconciliation of Business and Systems Modelling Via Discrete Event Simulation. Farrington, P.A., Nembhard, H.B., Sturrock, D.T., and Evans, G.W. (Eds). *Proceedings of the 1999 Winter Simulation Conference*, Phoenix, Arizona,
- GLAZER, R (1991), "Marketing In An Information Intensive Environment: Strategic Implications Of Knowledge As An Asset", *Journal of Marketing*, 55:1-19.
- GOODHUE, D.L, WYBO, M.D, KIRSCH, L.J. (1992), "The Impact Of Data Integration On The Cost And Benefits", *MIS Quarterly*, 16(3):293-311.
- GRANT J.S., DAVIS L.L. (1997), Selection and use of content expert for instrument development, *Research in Nursing and Health*, 20, 269-174.
- HALL, W.A, MCCAULEY, R.E (1987), "Planning And Managing Corporate Network Utility", *MIS Quarterly*, 11(4):437-47
- HAMMER, M. (1990, July-August). "Reengineering Work: Don't Automate, Obliterate," *Harvard Business Review*, pp. 104-112.
- HLUPIC V, ROBINSON S (1998), Business Process Modeling and Analysis using Discrete-event Simulation. *Proceedings of the 1998 Winter simulation conference*, <http://portal.acm.org/citation.cfm?id=293172.293489> Accessed on 12/11/2007
- KRIEBEL C.H., STRING D.M. (1984), A survey of the MIS and Telecommunications Activities of major business firms, *MIS Quarterly* 8(3): 171-7
- LEE, I. (2004), Evaluating Business Process Integrated Information Technology Investment. *Business Process Management Journal*, 10 (2):214-233
- MADNICK, S.E. (1995), "Integration Technology: The Reinvention Of The Linkage Between Information Systems And Computer Science", *Decision Support Systems*, 13:373-80.
- MEEL, J.W. VAN AND SOL, H.G. (1996) Business Engineering: Dynamic Modelling Instruments for a Dynamic World, *Simulation and Gaming*, 27(4):440-461
- PAINTER, M.K., FERNADES, R., PADMANABAN, N. AND MAYER, R.J. (1996) A Methodology for Integrating Business Process and Information Infrastructure Models. In Charnes, J.M., Morrice, D.J., Brunner, D.T. and Swain, J.J. (Eds.), *Proceedings of the 1996 Winter Simulation Conference*, San Diego, California, December, pp. 1305-1312
- PORTER, M (1985), "Technology And Competitive Advantage", *Journal of Business Strategy*, pp.60-78.

- RAILING L., HOUSEL T., (1990), A Network Infrastructure to control costs and enable fast response: The TRW process. *MIS Quarterly* 14(4):405- 19
- RAY J.P., SERRANO A. (2003), The process of process reengineering: Simulation for business process and information systems design. *Proceedings of the 35th conference on Winter simulation: Driving innovation*, New Orleans, Louisiana, USA.
- RICHARDSON G.L., JACKSON B.M., DICKSON G.W. (1990), A Principles Based Enterprise Architecture: Lessons from Texaco and Star Enterprise. *MIS Quarterly*, 14(4):385-403
- ROCKART, J., SHORT, J. (1991), "The Networked Organization And The Management Interdependence", in Scott Morton, M.S (Eds), *The Corporation of the 1990s: Information Technology and Organization Transformation*, Oxford University Press, New York, NY.
- ROSS, J. (1998), "IT infrastructure management", *Paper Presented At The 98' Annual Conference Of Information Systems Management At The London Business School Centre for Research in Information Management*, London.
- RUBIO, D. M., BERG-WEGER, M., TEBB, S. S., LEE, E. S., & RAUCH, S. (2003). Objectifying content validity: Conducting a content validity study in social work research. *Social Work Research*, 27, 94-104.
- SCHEER, A.W. 1994. *Business Process Engineering, Reference Models for Industrial Enterprises*. Springer-Verlag
- SCHETER E.S. (1992), *Managing the world class quality*, Marcel Dekker, New York, NY.
- SHANNON, R.E.(1975) *Systems Simulation: The Art And The Science*, Prentice-Hall New Jersey
- STALK G, HOUT T.M. (1990). *Competing Against Time*. Free Press, New York, NY.
- TENG, T., GROVER, V., JEONG, S., KETTINGER, W. (1995), "The Implementation Of Business Process Reengineering", *Journal of Management Information Systems*, 12(1):109-43.
- TUMAY, K. (1995), Business Process Simulation. *Proceedings of the WSC 95 – Winter Simulation Conference*, Washington DC, USA. Ed. By Alexopoulos A, Karg K., Lelegdon W.R., and Godsman D., SCS, 55 -60
- VENKATRAMAN, N. (1991), "IT-Induced Business Reconfiguration", in Scott Morton, M.S (Eds), *The Corporation of the 1990s: Information Technology and Organization Transformation*, Oxford University Press, New York, NY.
- WILLIAMS, D.(2004). Dynamics Synthesis Methodology: A Theoretical Framework for Research in The Requirements process Modelling Analysis. *Proceedings of the 1st European Conference on Research Methods for business Management studies*. Cambridge, UK. April 29-30, 467-490. ISBN:0-9540488-3-0.
- WILLIAMS, D., (2002). *An Application of System Dynamics to Requirements Engineering Process Modeling*, Unpublished PhD, London South Bank University
- WYSE, J.E. AND HIGGINS, C.A. (1993), "MIS Integration: A Framework For Management", *Journal of Systems Management*, February, pp. 32-7.

A Flexible Approach for User Evaluation of Biomedical Ontologies

Gilbert Maiga and Ddembe Williams

There has been an emergence of various ontologies describing data from either clinical or biological domains. Associated with this has been the development of integration systems, attempting to combine such biomedical ontologies using various strategies to overcome issues of scope and differing levels of granularity. However, such ontology based integration systems still find little use in distributed computing environments. This is attributed to both the lack of knowledge about user needs for such systems and the absence of a general framework to assess their relative suitability for specific applications. This paper aims to bridge this gap by proposing such a reference framework. The framework draws on existing information systems and ontology evaluation approaches in relating user objectives to ontology characteristics. The paper suggests that such a framework bridges the gap by emphasizing the dynamics of a biomedical environment. The framework therefore includes feedbacks from the evaluation process to the user characteristics of the integrated systems. This reference framework was validated in a study using structured interviews and a survey. The results indicate that the framework is sufficiently flexible for evaluating ontology based biomedical integrated systems, taking into account the conflicting needs of different users interested in accessing complex libraries of biomedical data.

1. Introduction

Biomedical ontology based integrated systems (BOIS) bring together disparate sources of heterogeneous and rapidly changing biological and clinical data [Kohler *et al.* 2003 ; Kumar *et al.* 2006; Yugyung *et al.* 2006; Rey-Perez *et al.* 2006; Sioutos *et al.* 2006]. Ontologies represent domain knowledge in a generic way. They provide a commonly agreed understanding of a domain, which may be reused, shared, and operationalized across applications and groups. They are used to create reusable models for integrating biomedical data in order to achieve interoperability between such sources. However, the widespread use of BOIS in distributed computing environments remains hampered by: 1) the lack of a general framework to assess their suitability for specific applications; 2) the lack of knowledge about user needs for such BOIS since ontologies are subjective knowledge artifacts in terms of time, place and cultural environment, as reflected in their design [Alani and Brewster 2006]. This underlines the difficulty in articulating specific properties to use in ranking ontologies since selection can depend on personal preferences of user requirements [*ibid.*].

Ontology integration builds new knowledge structures by assembling, extending, specializing and adapting existing ontologies. This enables purpose-built ontologies to emerge [Pinto and Martins 2000]. The challenge for evaluating ontologies has become an important research and development endeavor for assessing and selecting the appropriate biomedical integration ontology for a given system, especially in a distributed computing environment like the semantic web. Comparing the effectiveness of integrated ontology based systems remains challenging due to lack of standard frameworks for evaluating them, given that they differ in function, and in their expected input and outputs [Natalya and Musen 2002]. Existing frameworks evaluate ontologies largely on the basis of the taxonomical structure. These are therefore not suited for assessing the functional and processual objects present in biomedical ontologies. Furthermore, there is a lack of knowledge about the metrics and properties [scope, inputs, processes, outputs and structure] users require when judging the general quality or suitability of an integrated ontology [Alani and Brewster 2006; Kalfoglou and Schorlmer 2003; Lambrix and Tan 2006]. As a result, the effective reuse and wide adoption of ontology-based biomedical integration systems by industry is not yet realized [Gangemi *et al.* 2005; Kalfoglou and Hu 2006]. The EON [2006] workshop helped to bring into focus this lack of standard evaluation approaches and metrics for determining quality in the increasing number of ontologies made available for the semantic web.

In this study, a mixed method research strategy combining quantitative and qualitative deductive approaches was applied. A deductive descriptive approach helped to identify general properties and ontological theories that support knowledge integration in biological and clinical integration systems. The properties that users require in an ontology for biomedical integration systems were investigated using structured interviews and a survey. This helped to determine the scope, inputs, processes, outputs and evaluation criteria important to users of an integrated biomedical system. The results were used to derive a flexible framework for user centered assessment of systems that integrate clinical and biological ontologies. The framework is seen as a dynamic system, with feedbacks from the evaluation process guiding improvement of the integration systems.

The rest of this paper is organized as follows. Section 2 discusses work related to information systems evaluation and biomedicine. Section 3 explores existing approaches to ontology evaluation, multicriteria evaluations and systems theory as key underpinning concepts for deriving the framework. Conclusions from the study are drawn in section 4.

2. Biomedical and Information Systems Evaluation

2.1 Integrating Biomedical Information with Ontologies

Ontology-based attempts to bridge the gap between clinical and biological information in order to achieve interoperability include work on ONTOFUSION

[Rey-Perez *et al.*, 2006], SEMEDA [Kohler *et al.* 2003], ASLER [Yugyung *et al.*, 2006], colon carcinoma [Kumar *et al.* 2006], and cancer [Sioutos *et al.*, 2006].

While the ONTOFUSION tool provides semantic level integration of genomic and clinical databases [Rey-Perez *et al.* 2006], it does not bridge across levels of granularity for biomedical data. Using principles of the basic formal ontology (BFO), Kumar *et al.* [2006] describe a framework for integrating medical and biological information in order to draw inferences across levels of granularity using the Gene Ontology. Yugyung *et al.* [2006] describe a method for medical ontology integration using an incremental approach of semantic enrichment, refinement and integration (ALSER) that depends on measures of similarity between ontology models. A terminology and description logic based framework for integrating molecular and clinical cancer-related information given by Sioutos *et al.* [2006]. It however leaves integrating with external sources largely unresolved. Despite these efforts, integrating biomedical data objects across structure, function and processes using ontologies remains challenging with lack of a single unifying approach (frame of reference) to ontology integration, against which users can assess an ontology based integration system.

Evaluation of ontology integration systems is a criterion based technical judgment guiding the construction process and any refinement steps for both the integrated and resulting ontologies [Pinto and Martins 2000]. Ontologies may be assessed by user ratings and reviews, and for general ontological properties. Evaluation is done during design, development and prior to use [Kalfoglou and Hu 2006]. Evaluation is important if ontologies are to be widely adopted for use in distributed computing environments [*ibid.*]. The lack of a unifying evaluation framework for integrated systems remains an obstacle for ontology reuse and may hinder their adoption by industry and the wider web community [Alani and Brewster 2006].

Little work exists on evaluating biomedical ontology integration systems (BOIS). To evaluate both the integrated and resulting ontologies, Pinto and Martins [2000] recommend both technical and user assessment of the candidate ontologies by domain experts and ontologists respectively. These authors also recommend selection of candidate ontologies using strict (hard) and desirable (soft) requirements. Soft requirements provide flexibility, as they can be adapted to integration processes that take into account particular features during the choice of one ontology.

Evaluation of the resultant ontology can therefore be done according to criteria used for any ontology to meet assessment criteria of completeness, conciseness, consistency, expandability, robustness [Gomez-Perez and Pazos 1995], and clarity, coherence, extendibility, minimal encoding bias, minimal ontological commitment [Gruber 1995]. This work and the proposed evaluation framework is largely motivated by the need to contribute to the evaluation of biomedical integration systems so they can gain wide use in distributed computing environments like the semantic web.

2.2 Information Systems Evaluation

The theory of evaluation is rooted in the twin ideas of accountability and social enquiry [Alkin and Christie 2004]. Thus Cronholm and Goldkuhl [2003] divide information system (IS) evaluation approaches into formal rational, interpretive or criteria based, using respectively goal based, goal free and criteria based evaluation strategies. Formal rational evaluations are largely quantitative processes, usually concerned with technical and economic aspects, employing goal based strategies. These strategies focus on intended services and outcomes, to achieve goals which can be phrased in quantitative or qualitative terms.

Interpretive approaches view IS as social systems with embedded information technology. In the same light, Walsham [1993] argues that IS evaluation should consider not only the purpose for conducting the evaluation and associated factors, but also social context and process, and stress the need to consider evaluation as a learning process for all involved. Goal free strategies are appropriate in an interpretive approach, performed with limited evaluator involvement. In contrast criteria based evaluation use selected general qualities for evaluation where scores from multiple criteria are combined into an overall weighted sum [Walsham 1993].

The most appropriate evaluation approach depends largely on its context [Cronholm and Goldkuhl 2003]. Whatever the approach adopted, Alkin and Christie [2004] argue that evaluation models must consider *methods*, *valuing* and *use*. *Methods* deal with how the evaluation is to be done, and so focus on knowledge construction. *Valuing* concerns the role of the evaluator in assigning criteria while *use* focuses on the purpose of the evaluation. These factors can be seen as the *process, people and purpose* that Ballantine *et al.* [2000] identify as drivers of IS evaluation.

IS evaluation is considered as a multi-stage process occurring at different points, in different ways, during the product life-cycle. However, IS evaluation literature and practice have been closely linked with the technical project development cycle. For example, Beynon-Davis *et al.* [2004] propose a model for IS evaluation that distinguishes between the strategic, formative, summative approaches, and post mortem analysis [done if the project has to undergo partial abandonment - which is not relevant to the comparative integration system evaluation being investigated here]. Strategic evaluation is conducted as part of the planning process for an intended IS. Formative evaluation is a continuous, iterative informal process aimed at providing systematic feedback to designers and implementers, influencing the process of development and the final IS [Kumar 1990; Remenyi and Smith 1999; Walsham 1993]. Summative evaluation, usually done at the end of the project is concerned with assessing the worth of a project or program outcome in light of initially specified success criteria [*ibid*].

BOIS can also be seen as a type of IS, with systemic features that can be variously evaluated against user requirements. What distinguishes them from general IS is arguably the dynamics of the environment in which they are used. The vast amounts

of biomedical data require ontology integration systems that can capture and represent new structure, processes and functional objects that emerge with data in the two domains. In a comparative evaluation of biomedical ontology integration systems as IS, the evaluator is the proposed user. Neither an interpretative nor a formal rational approach seems as appropriate as a criteria-based approach in which the user's requirements motivate the evaluation criteria. Such an approach can be undertaken in a strategic, formative or summative evaluation.

3. Deriving the Evaluation Framework

This section explores existing approaches to ontology evaluation and systems theory as key underpinning concepts for deriving the framework.

3.1 Ontology Evaluation Approaches in Biomedicine

The pre-modeling, modeling and post release stages of the ontology development life cycle have been used to categorize ontology evaluation [Hartmann et. al. 2004]. Based on type and purpose, Brank et al. [2005] provide a taxonomy of ontology evaluation approaches as follows:

- 1) Approaches that evaluate an ontology by comparing it to a golden standard ontology or other representation of the problem domain for which an appropriate ontology is needed [Gomez-Perez 1994; Guarino, 1998; Hovy 2001].
- 2) Task based approaches that use the ontology in an application and evaluate the quality of results [Porzel and Malaka 2004].
- 3) Data or corpus driven approaches. These evaluate the congruence of an ontology with a given corpus to determine how appropriate it is for the representation of knowledge of the domain represented by the texts [Brewster et al. 2004].
- 4) Assessment by humans to show how well the ontology meets a set of predefined criteria, standards, requirements as in Ontometric [Lozano-Tello and Gomez-Perez 2004] and the peer-review based approach [Supekar 2005].

3.2. Multicriteria Approaches to Ontology Evaluation

Multiple criteria approaches deal with the problem of selecting a good ontology from a given set based on defining several decision criteria or attributes. For each criterion, the ontology is evaluated and given a numerical score and an overall score computed as a weighted sum of its per-criterion scores [Brank et al. 2005; Lozano-Tello and Gomez-Perez 2004]. While this approach requires a lot of manual involvement by human experts, it allows a combination of criteria at many levels [ibid].

Ontometric [Lozano-Tello and Gomez-Perez 2004] is a multicriteria decision making method for use by knowledge engineers to select an appropriate ontology among various alternatives or to decide the suitability of a particular ontology for a

project using the dimensions of: cost, content represented, language, methodology and software development environment used. [Hartmann et al. 2004]. The selection criteria is based a four step analytic hierarchy process (AHP) [Saaty 1980] namely: 1) Decide upon the criteria for selection; 2) Rate the relative importance of these criteria using pair-wise comparisons; 3) Rate the potential choice relative to the other on the basis of each selection criterion, achieved by pair wise comparisons of the choices; 4) Combine the ratings derived in steps 2 and 3 to obtain an overall relative rating for each potential choice.

3.3. Contributions from Systems Theory

General systems theory (GST) is here adopted to explain the emergent properties from the complex and dynamic nature of data in biomedical integration systems [Huang et al. 2004]. GST is “elements in standing relationship, the joining and integrating of the web of relationships to create emergent properties of the whole that are not found in any analysis of the parts” [Von Bertalanffy 1962]. GST explains structure and properties of systems in terms of relationships from which new properties of wholes emerge that are not found among those of elements, and the corresponding behavior of the whole cannot be explained in terms of the behavior of the parts. Concepts from the process of self organization (SO) may be used to extend systems theory. In SO, the internal organization of a system increases in complexity without being guided or managed by an outside source and displays emergent properties which do not exist if the lower level is removed [Gershenson 2006]. In SO the environment is unpredictable and the elements interact to achieve dynamically a global function [*ibid*].

Engineered systems self organize by adaptation, anticipation, robustness or a combination of these features [Gershenson 2006]. Self-organizing systems, rather than being a type of systems, are a perspective for studying, understanding, designing, controlling, and building systems; the crucial factor being the observer, who has to describe the process at an appropriate level and aspects, and to define the purpose of the system; SO can therefore be everywhere, it just needs to be observed [Heylighen and Gershenson 2003]. Organization is seen as structure that has a purpose. The observer has to focus their viewpoint, set the purpose of the system to see the attractor as an organized state at the right level and in the right aspect in order to observe self-organization - a perspective used to design, build, and control artificial systems. A key characteristic of an artificial self-organizing system is that structure and function of the system emerge from interactions between the elements [*ibid*]. It is this perspective of self organizing systems that is adopted for this study to explain the emergent properties faced by frameworks for evaluating BOIS.

3.4. Field Study

A mixed method research strategy combining quantitative and qualitative deductive survey was used in the study to identify the dimensions and requirements for a

meta-ontology of biomedical integration systems. Existing literature and document analysis were used to determine scope and identify common structures, functions, inputs, outputs for such a meta-ontology. Similarly, ontological theories that support knowledge integration in biological and clinical integration systems were described. The meta-ontology acts as the frame of reference for the evaluation framework. The dimensions of the reference evaluation framework were tested for potential user agreement using a field study.

The study population consisted of 450 (four hundred fifty) health care workers. Structured interviews and a survey using questionnaires were conducted to collect data and clarify user needs relevant to biomedical integrated systems. In the survey, potential users of such a system were asked for their level of agreement with characteristics of a meta-ontology for such a system, and its evaluation criteria namely: the scope, purpose, requirements, granularity, inputs, processes, outputs.

A high level [$>70\%$] of respondents agreed that a biomedical integration system should:

- 1) Integrate data from diverse sources and yet useful to molecular biologists, clinicians, healthcare workers, legal practitioners, information systems developers and policy makers
- 2) Support heterogenous data integration across levels of granularity
- 3) Support functions and processes that link and relate both existing and future genetic to clinical patient data objects, determine the fit between task/user needs to existing biomedical integration models and provide feedbacks for continuous improvements of such models.
- 4) Provide for a generic meta-model that define structures (concepts) that are common to both patient clinical and genetic data.

The results of the field study were incorporated into the reference framework. A new framework for evaluating ontology based integrated systems was the result. Inputs, processes, outputs and feedbacks for such a framework were identified from the field study. It relates user needs to structure, functions and processes of an ontology and seeks to determine the effectiveness of a biomedical integration system using completeness to measure system quality.

5.5. The Framework

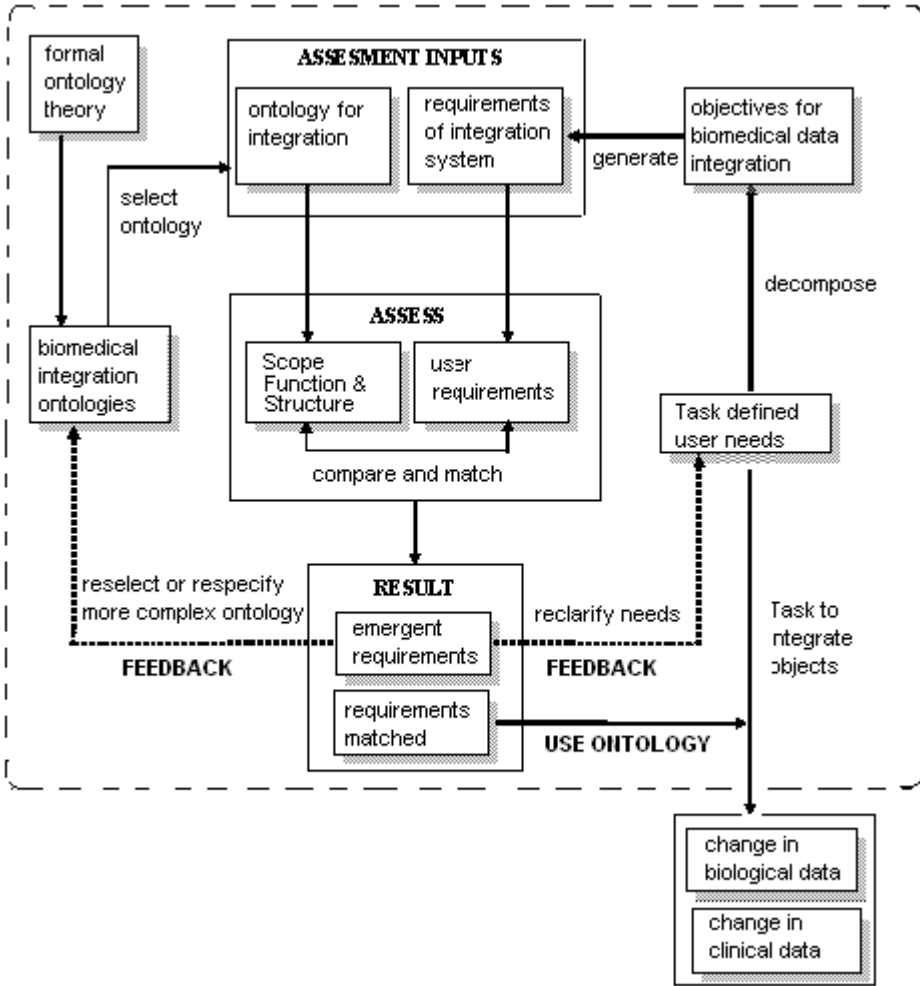
Existing ontology based evaluation models are attempts to define standards for structuring, representing and integrating existing knowledge in a domain of discourse using specified criteria. Such criteria are often combined and used for user assessment and selection of an ontology appropriate for a given task [Lozano-Tello and Gomez-Perez, 2004; Supekar 2005]. In biomedicine, where new data objects emerge and new user requirements emerge with the rapidly increasing data, flexible models that enable users to iteratively search through an ontology library using multiple criteria are more likely to result into: (1) selection of an

appropriate ontology for a given task or (2) re-specification of new requirements for an ontology to fit the task. This framework is an attempt to evaluate biomedical ontology integration systems as information systems with the evaluator as the proposed user. A criteria-based approach in which the user's requirements motivate the assessment criteria in a formative evaluation [Kumar 1990; Remenyi and Smith 1999; Walsham 1993] framework is found appropriate as it aims to provide systematic feedback to designers and implementers, influencing the process of development and the final IS.

This framework (in figure1) extends Ontometric [Lozano-Tello and Gomez-Perez 2004], a multicriteria decision making method for selecting the most appropriate ontology among various alternatives or for deciding the suitability of a particular ontology for a project. Ontometric, based on the AHP framework [Saaty 1980] allows ontology engineers and domain experts the flexibility to select the hierarchy for the decision criteria to be used in evaluations. It however offers no specific features that support evaluating systems in dynamic environments like biomedicine that require reusing, extending and modifying existing ontologies to accommodate new types of data objects and avoid the huge effort of starting or building entirely new ontologies [Alani and Brewster 2006]. The evaluation framework as proposed in this study (figure1) mitigates this shortcoming by adopting a systems approach so that domain or task evaluator (user) needs are considered when pruning, extending or modifying an existing ontology.

Inputs for the ontology based biomedical evaluation framework are defined in terms of scope, function, structure, clinical and biological objects and their relationships. User requirements for the integration system are the other type of requirements for evaluation. Biomedical ontology integration systems (Ontologies) also have taxonomic structures (defined using objects and relationships between them), function and processes. The processing requirements for the evaluation framework are: determining evaluation criteria from user needs; visualization and identification of relationships between clinical and biological concepts; comparison of the overlap between user needs and the ontology model; determining the fit between a task and user integration requirements; providing feedback to improve upon and enable existing models meet user task requirements. The evaluation process determines the relevance of an integration ontology for a give task. The overlap between user needs and the ontology model determines the fit between the two.

Fig 1: A User Centered Biomedical Ontology Evaluation Framework



The evaluation framework seeks to establish how well the biomedical integration system conforms to task derived user requirements i.e. completeness. On evaluation, the scope, structure (objects and relations), processes and functions of a given ontology may meet user requirements leading to its selection for use in integration of clinical and biological data objects. Alternatively, a proportion of total user requirements defined may not be met and are here referred to as emergent requirements (ER). Emergent requirements are used as feedbacks for re-examination and reclarification of the scope of user needs, and for reselection or re-specification of the ontology model. This leads to new ontology integration systems with more complex structure and function. General systems theory and that of self organization help to explain these feedbacks caused by these emergent properties of the framework, leading to complexity.

System quality, as used to evaluate an ontology in this framework is defined as “conformance to requirements” [Crosby 1979; Garvin 1984]. A need can be decomposed into a set of requirements upon features and characteristics of a product or service, and if all these requirements are conformed to, then that need has been satisfied. The evaluation criterion seeks to establish how well the biomedical integration system ontology conforms to requirements. This is completeness for the system, and is here expressed as:

$$\text{Completeness} = \frac{\text{All Requirements met}}{\text{Total system requirements defined}} \dots\dots\dots\text{Equation 1}$$

In the framework, this relationship expresses how well the ontology based structures satisfy task requirements. Following evaluation, a proportion of total requirements defined are not met. These emergent requirements [ER] require re-examination of user needs and /or ontology structure leading to a redefinition of the integration requirements.

$$\text{ER} = [\text{Total system requirements defined} - \text{All Requirements met}] \dots\dots\text{Equation 2}$$

The framework guides users (ontology developers or domain experts) to select an appropriate ontology integration system for a given task through the following steps:

1. Determine the selection criteria (in terms of scope, function, structure, clinical and biological objects and the relationships between them) for an ontology based on the requirements of a given integration task
2. Rate the relative importance of these selection criteria to the integration task
3. Select and Visualize ontology to assess (scope, structure, function, processes).
4. Assess ontology by matching it to identified selection criteria for the task.
5. Identify any unmatched (emergent) requirements
6. Use emergent requirements to help rescope/reselect user or ontology or Select matched ontology for use in the integration process.

4. Conclusions

The literature shows the important need for evaluation of ontologies use in integrating data in distributed computing environments. The paper identifies the challenges faced by existing frameworks in evaluating such ontology based biomedical integration systems in a dynamic environment. This framework as given in this paper extends the OntoMetric evaluation method [Lozano-

Tello and Gomez 2004] using systems theory. Systems theory is then adopted as an underpinning theory in a user driven criteria based formative evaluation framework for biomedical ontology integration systems. Steps for the utility of such a framework are also given. The framework emphasizes the properties users tend to look for when judging the suitability of a biomedical ontology integration system. Interviews and questionnaires in a descriptive survey were used to validate the framework. The framework enables users to assess and select a suitable ontology for their particular use case when accessing ever-increasing libraries of biomedical data. The novelty of this framework lies in the ability to relate ontology structure and user derived objectives in order to derive requirements for evaluating ontology based integrated biological and clinical information systems in environments where the amounts of data are ever increasing and user needs are changing.

References

- ALANI, H. AND BREWSTER, C. 2006. Metrics for Ranking Ontologies. In *Proceedings of the 4th International Conference on Evaluation of Ontologies for the Web*. EON Workshop Edinburgh International Conference Center. Edinburgh UK. May 22nd 2006.
- ALKIN, M.C. AND CHRISTIE, C.A. 2004. An evaluation theory tree. In M.C. Alkin Ed., *Evaluation Roots*. Thousand Oaks, CA: Sage. Chapter 2, 12-54.
- AVISON, D., HORTON, J., POWELL, P., AND NANDHAKUMAR, J. 1995. Incorporating Evaluation in the Information Systems Development Process. In *Proceedings of the third evaluation of Information Technology Conference*. Henly, UK.
- BALLANTINE, J., LEVY, M., MARTIN, M., MUNRO, I., AND POWELL, P. 2000. An Ethical Perspective on Information Systems Evaluation. *International Journal of Agile Management Systems*. 23, 233- 241.
- BEYNON-DAVIS, P., OWENS, I. AND WILLIAMS, D.M. 2004. Information systems Evaluation and the Information Systems Development Process. *The Journal of Enterprise Information Management*. 17, 276-282.
- BRANK, J., GROBELNIK, M. AND MLADENIC, D. 2005. A Survey of Ontology Evaluation Techniques. *At the conference on Data Mining and Data Warehouses SIKDD*.
- BREWSTER, C., ALANI, H., DASMAHAPATRA, S., AND WILKA, Y. 2004. Data Driven Ontology Evaluation. *Proceedings of the Language Resources and Evaluation Conference LREC 2004*, Lisbon, Portugal.
- CRONHOLM, S. AND GOLDKUHL, G. 2003. Strategies for Information Systems Evaluation: Six Generic Types. *Electronic Journal of Information Systems Evaluation*. 62, 65 -74.
- CROSBY, P.B., 1988. *Quality is Free: The Art of Making Quality Certain*, New York: McGraw-Hill, 1988.
- DAVIDSON, S., OVERTON, C. AND BUNAN, P. 1995. Challenges in Integrating Biological Data Sources. *Journal of Computational Biology*. 24.
- DING, Y. AND FOO, S. 2002 Ontology research and development, Part 2. A Review of Ontology mapping and evolving. *Journal of Information Science*, 28, 375 - 388.

- EON. 2006. Evaluation of Ontologies for the Web, *4th International EON Workshop* located at the 15th International World Wide Web Conference, Edinburgh International Conference Center, Edinburgh, United Kingdom, Available: <http://km.aifb.uni-karlsruhe.de/ws/eon2006>. Accessed May 25th, 2006.
- GANGEMI, A., CATENACCI, C., CIARAMITA, M. AND LEHMANN, J. 2005. A theoretical framework for ontology evaluation and validation. In *Proceedings of SWAP 2005*.
- GARVIN, D.A., 1984. What does 'Product Quality' Really Mean?, *Sloan Management Review*, fall, 1984.
- GERSHENSON, C. AND HEYLIGHEN, F. 2003. When can we call a system self-organizing? In *Advances in Artificial Life, 7th European Conference*, and ECAL 2003 LNAI 2801
- GERSHENSON, C. 2006. A General Methodology for Designing Self-Organizing Systems. *ECCO working paper*. Available: <http://uk.arxiv.org/abs/nlin.AO/0505009>
- GOMEZ-PEREZ, A. 1994. Some ideas and Examples to Evaluate Ontologies. Technical Report KSL-94-65. Knowledge System Laboratory. Stanford University. Also in Proceedings of the 11th Conference on Artificial Intelligence for Applications, CAIA94.
- GRUBER, T. 1995. Towards Principles for the Design of Ontologies Used for Knowledge Sharing. *International Journal of Human-Computer studies*, 43 5/6:907-928.
- GUARINO, N. AND WELTY, C. 2002. Evaluating Ontological Decisions with OntoClean. *Communications of the ACM*. 45, No. 2.
- GÓMEZ-PÉREZ, A. 2004. Ontology Evaluation. *Handbook on Ontologies*. 251-274.
- HEYLIGHEN, F. AND C. GERSHENSON. 2003. The meaning of self-organization in computing. *IEEE Intelligent Systems*, 72-75.
- HONGZHAN, H., CATHY, H., ANASTASIA, N., ZHANGZHI H., WINONA, C., AND BARKER, B. 2004. The iProClass integrated database for protein functional analysis, *Computational Biology and Chemistry*, 28. 2004:87-96.
- HOVY, E. 2001. Comparing sets of semantic relations in ontologies. In Rebecca Green, Carol A Bean, and Sung Hyon Myaengeds., *Semantics of Relationships, chapter 6*. Dordrecht, NL: Kluwer.
- HUANG, H., HU, Z., SUZEK, B.E. AND WU, C.H. 2004. The PIR integrated protein databases and data retrieval System. *Data science journal*, 3:163.
- KALFOGLOU, Y. AND SCHORLMER, M. 2003. Ontology Mapping: The State Of The Art. *The Knowledge Engineering Review*, 181: 1-31.
- KALFOGLOU, Y. AND SCHORLMER, M. 2006. Issues with evaluating and using publicly available ontologies. In *Proceedings 4th International EON Workshop, Evaluating Ontologies for the Web*, Edinburgh, Scotland.
- KOHLER, J., PHILIPI, S. AND LANGE, M. 2003 SEMEDA: Ontology based semantic integration of biological databases. *Bioinformatics*, 19, 2420 -2427.
- KUMAR, K.1990 Post Implementation Evaluation of Computer based Information Systems CBIS: current practices, *Communications of ACM*, 33, 203-12

- KUMAR, A., LINAYIP, Y., SMITH, B. AND GRENON, P. 2006 Bridging the gap between medical and bioinformatics: An ontological case study in colon carcinoma. *Computers in Biology and Medicine* 36, 694–711
- LAMBRIX, P. AND TAN, H. 2006. SAMBO. A system for aligning and merging biomedical ontologies *Web Semantics: Science, Services and Agents on the World Wide Web*, 4 196 – 206
- LOZANO-TELLO, A. AND GOMEZ-PEREZ, A. 2004. ONTOMETRIC: A Method to Choose the Appropriate Ontology”. *Journal of Database Management*. Special Issue on Ontological analysis, Evaluation.
- NATALYA F. N. AND. MUSEN, M.A. 2002. Evaluating Ontology-Mapping Tools: Requirements and Experience. Available: <http://seer.nj.nec.com>. Accessed 6 may 2006.
- PINTO, H. S., AND MARTINS, J. P. 2000. Reusing Ontologies. In *Proceedings of AAAI 2000 Spring Symposium Series, Workshop on Bringing Knowledge to Business Processes*, AAAI Press.
- REY-PEREZ, D., MAOJO, V., GARCIA-RESAL, M., ALONSO-CALVO, R., BILLHARDT, H., MARTIN-SÁNCHEZ, F., AND SOUSA, A. 2006. ONTOFUSION: Ontology-based integration of genomic and clinical databases. *Computers in Biology and Medicine*, 36, 712-730.
- SMITH, B. 2003. Ontology. *Blackwell Guide to the Philosophy of Computing and information*, Oxford: Blackwell, 55- 166.
- SIOUTOS, N., CORONADO, S., HABER, M., HARTEL, F., SHAIU, W., AND WRIGHT, W. C. 2006. NCI thesaurus: A semantic model integrating cancer-related clinical and molecular information. *Journal of Biomedical Informatics*.
- USHOLD, M. AND M. GRUNINGER. 1996. Ontologies: Principles, methods and applications, in *The Knowledge Engineering Review*, 112, 93–55.
- VON BERTALANFFY, L. 1962. General system theory - A Critical Review. *General Systems* 7, 1-20.
- WALSHAM, G. 1993. Interpreting Information systems in organizations. Wiley, Chichester.
- YUGYUNG, L., SUPEKAR, K. AND GELLER, J. 2006 Ontology Integration: Experience with medical Terminologies. *Computers in Biology and Medicine*, 36, 893 - 919.

Facilitating Grass-root NGO's Adoption of Information Systems and Information Communication Technologies

Kathy Lynch, Stacey Lynch and Festus Bazairra

Increasingly, information and communication technologies (ICTs) are being employed to share, generate and analyse information. In response to their decreasing cost, an increase in accessibility, and an increase in the adoption of modern ICTs across the globe, non-government organizations (NGOs) in developing countries are not only welcoming these technologies into their communities, but are using them to connect to resources outside their communities. On the lowest rung of a country's organizational ladder, are grass-root NGOs; these small organizations are often the life line for isolated communities, and it is through them, developing countries are making progress – both economically and socially, towards advancing a country's development. This paper presents a pilot project that explored the adoption of information technology by a small grass-root NGO in Uganda. We found the adoption of an easy to use information system together with skills training in the use of the information system and communication technologies opened up opportunities, pride and professionalism in the NGO, with a flow on effect to the people within the communities they serve. The paper concludes with several guidelines that could be followed when facilitating the adoption of ICT by organisations in developing countries.

1. Introduction

The research presented in this paper is a pilot study that explores the use of Information and Communication Technologies (ICTs) by one grass-root Non Government Organisation (NGO) in Uganda; The Voluntary Services Trust Team (VOLSET). The paper reports on an instance where VOLSET embraced the development of a purpose-built information system used to systematically record Voluntary Counselling and Testing (VCT) data. The transferring of VCT data from an archaic, inadequate, paper-based system into an electronic application where data can be manipulated, analysed, graphed and reported, is critical as valid evidence is needed to further advance the benefits of VCT across the Ugandan nation.

The study evolved from two researcher's participation in a five week volunteer program in Uganda. The researchers, one education/information technology academic, the other a microbiologist, spent time with VOLSET during November/

December 2005. The study design is retrospective, and uses an ethnographic approach influenced by participatory action research.

The paper commences with an overview of NGOs and VOLSET in particular, the state of ICTs in developing countries and Uganda in particular, the study design, observations, interventions, and concludes with a discussion of the outcomes of the research, including draft guidelines for the development of information systems for NGOs in developing countries who are on (or are moving towards) the path of ICT adoption.

1.1 Non-government Organisations (NGOs)

Non-government organisations have traditionally contributed significantly to a country's future, however, to continue this work effectively in today's technological age, they need not only the technology but the skills to use and gain value from applications from which the technologies have enable. Furthermore, they require timely access to 'markets' to enable the sharing and dissemination of the information obtained. Often NGOs are at a 'grass-root' level; these organisations give assistance to communities at a local level by the local or common 'man'. Grass-root NGOs are a life line for many people in developing countries in improving skills, knowledge, and support from outside the confines of their small, isolated communities.

Uganda is a country with many NGOs that sustain life within the region. In 2003 (more recent figures could not be found) there were 3,499 registered NGOs in Uganda [The Uganda National NGO Forum, 2003], plus at least another 30% unregistered faith-based. These organisations are responsible for distributing approximately 25% of international aid funding [Ndegwa, 1996] , (more recent figures could not be found, however it is presumed to be higher in 2007-2008). The total amount of foreign aid entering Uganda each year approximately 11% of its GDP and for the fiscal year 2005-06 was predicted to reach about \$989 million [Worldpress.org].

1.2.1 Voluntary Counselling and Testing (VCT)

VCT refers to the process of giving people professional counselling before and after HIV testing. The process helps people prepare for and understand their test results. Over the past 20 years, VCT programs have helped millions of people learn their HIV status, nonetheless there are more than 80% of people living with HIV in low and middle-income countries who do not know that they are infected [World Health Organization]. In 2003, Uganda had approximately 6.5% prevalence of HIV in its adult population [Coutinho 2003] - a decline from 22% in 1992. One of the factors claimed to have brought about this decline is large-scale VCT. This testing has primarily been provided by the 2000 NGOs active in HIV work in Uganda (http://www.aidsuganda.org/pdf/role_of_VCT.pdf). Although this decline is impressive, one of the challenges that lie ahead is not only spreading VCT across Uganda but scaling up best practice of VCT [Coutinho 2003]. Furthermore,

work conducted by (or through) such government organisation as the Population Secretariat (POPSEC), the Ministry of Health, the Ministry of Gender, Labor and Social Development, aim at strengthening leadership at national and district level in creating a supportive environment for family planning and reproductive health [Babikwa 2005]. It is claimed by Coutinho [2003] that VCT reinforces positive behaviour whether HIV positive or negative, and the more people who go through VCT, the less stigma will be attached to those who live with HIV/AIDS, and that early VCT can lead to a delay in HIV deaths [Coutinho 2003].

1.2.2 *VOLSET: One Ugandan NGO*

The Voluntary Services Trust Team (VOLSET) was established in 1997 by a small group of concerned Nakaseke (a village) Ugandans in response to the government's concern with the HIV/AIDS pandemic, and an increase in the number of needy children orphaned by HIV/AIDS and poverty. The small group made it their mission to "alleviate human suffering caused by disease and poverty" [VOLSET]. VOLSET is a registered non-political, non-partisan, non-religious, non-governmental organization, operating in the districts of Mukono and Luwero/Nakaseke. **Over the last ten years, VOLSET's mission has changed little, though it now specifically includes HIV:** "to alleviate human suffering caused by HIV/AIDS, disease and poverty in rural and less served communities" [VOLSET]. VOLSET's office is located in Ntenjeru Town: a town with few shops, no running water, dusty roads, limited transport, and a very intermittent electricity supply. However, VOLSET does own (through donations) a mobile phone, two computers, has a website (<http://volset.org/>) and an e-mail address.

VOLSET operates within a grassroots community, and is managed by a small dedicated group of village Ugandans. Year-round, year after year volunteers come from all over the globe to work with VOLSET for one month to three years. These volunteers are primarily from developed countries, and bring skills which they willingly share with VOLSET and villagers they interact with to build capacity and educate the villagers in an attempt to break the poverty circle. It is through the learning from and sharing these skills that enable VOLSET to grow, to be productive, and recognized as leader in the community in which it serves.

One of the programs carried out by VOLSET is the VCT program. VOLSET members and volunteers counsel and HIV test villagers in their districts around and on Lake Victoria. The use of mobile phones the Web e-mail and computer applications are changing the way VOLSET (and other NGOs) engage with their communities, and doing business.

1.3 Information and Communication Technologies (ICTs) in Developing Countries

In many developing countries, traditional information and communication infrastructure such as telephone landlines and reliable electricity supply, are not in abundance across the nation, or accessed by the general populace as a whole.

However, modern information and communication technologies have emerged, and are becoming increasingly popular for the 'common man' therefore, enabling them to become connected to the world outside their village. In particular, mobile phone telephony, basic computer applications, and the Internet have taken on an acceptance and ownership level not seen before. The adoption of ICTs in countries such as Uganda, compared to developed countries where they have become the norm, ubiquitous and pervasive, have brought the power of communication, immediacy and the world to these people, and have enabled them to share information, and communicate with others beyond their communities.

Billions of dollars every year are given to African nations to assist them in developing their nation's skills. Funding for the use, training, implementation and development of ICTs is seen as a pathway out of poverty as they are seen as change agents that are capable of empowering the populace with capacity building, self efficiency, and education. Information technology has been identified as one of the rapidly growing areas that has the potential to 'leap-frog' developing countries, into the globalised economy [United Kingdom, Parliamentary Office of Science and Technology, 2006]. The study reports that access to ICTs in developing countries is on the increase; for example; in developing countries the total (fixed and mobile) telephone access increased from 2% in 1991 to 31% in 2004: and Internet usage has also grown rapidly from 0.03% in 1994 to 6.7% in 2004 of developing country inhabitants accessing the Internet. Uganda is no different, and though it is one of the poorest nations in the world, ICT implementation is high on its lists of national priorities [Uganda National Council for Science and Technology, 2002].

The number of Ugandan's subscribing to a mobile phone service has increased from 500,000 in December 2002 to just over 3 million in March 2007 [Uganda Communications Commission, 2007]. Prior to 1996, Uganda's communication infrastructure was among the least developed in Africa, furthermore, even today approximately 70% of the communication services are concentrated in urban areas, leaving the rural areas with the least access to these vital communication service [Uganda Communications Commission, 2007]. Internet service providers (ISP) have also been on the incline- from two in 1996, to 17 in 2002. Uganda's communication systems favour those living in and around the capital city of Kampala, with at least 25 Internet cafes in Kampala (the capital), with more being established every year. Mobile phones and the Internet are not the only ICTs in existence, however, they are the ones most reported in the literature.

Regardless of the insurgence of ICTs into developing countries, their benefits are not fully realised as many countries have inadequate infrastructure and human capacity to support ICT [Uganda National Council for Science and Technology, 2002].

2. The Study

Studies exploring the role ICTs play in developing countries have been conducted by international originations (for example, the Canadian International Development

Agency (CIDA), and studies specifically in Uganda have been conducted by the International Development Research Centre (IDRC)). However, the participants of their studies have commonly come from large, well financed urban Ugandan NGOs; but, little is known about the adoption of ICTs by grass-root, rural-based Ugandan NGOs. The study conducted and presented in this paper, looks at the adoption of ICTs by one rural, grass-root Ugandan NGO. The study itself is a pilot, however the results of the pilot have informed a wider project that is in the early stages of design and implementation.

The study emerged from a five week volunteer program with VOLSET by two volunteers who also happened to be researchers; one a researcher in education and information technology, the other, a microbiologist. The study design is retrospective, and uses an ethnographic approach influenced by participatory action research (PAR). Furthermore, data has been collected since November 2005, and is continuing. As the study falls in to the PAR domain there is no research question per se nor hypothesis, however the impetus of the pilot study was exploring the adoption of information and communication technologies by grass-root NGOs in Uganda.

2.1 Research Methodology

The study follows the paradigm of ethnographic research. This methodology is described as “.. the ethnographer participating, overtly or covertly, in people’s lives for an extended period of time, watching what happens, listening to what is said, asking questions” [Hammersley & Atkinson 1995: 1]. More recently, Johnson (2000) defines ethnography as “a descriptive account of social life and culture in a particular social system based on detailed observations of what people actually do” [Johnson 2000: 111]. The study explored the social structures and social dynamics surrounding the NGO and the villagers they interacted with, thus enabling a view of real-life situations and behaviours, which hopefully contributes to an empirically stronger theory which in turn can contribute to grounded theory.

Ethnographic research is holistic, with no element examined or observed in isolation. Consequently the use of ICTs by VOLSET needs to be looked at in the frame of the environment in which it exists: that is isolated, rural, poor transport, poor health services, dusty roads, high unemployment, subsistent village life, non-existent or very limited power supply, and limited financial resources. Together with the adoption of ICT the researchers were also interested in the transference of ICT related skills from the volunteers to VOLSET and the community. This interpretation and how it relates to a plan to bring about a different desired state for the culture is common in ethnographic research [Garson 2008].

Participatory action research is embedded in a study as the researchers were also two of the volunteers who worked with false set to five weeks and have continued to work with VOLSET from their own country. PAR can be defined as “collective, self-reflective enquiry undertaken by participants in social situations in order improve the rationality and justice of their own social...practices” [Kemmis

& McTaggart 1988: 5], it embraces the principles of participation and reflection, and empowerment and emancipation of groups seeking to improve their social situation. This is a common stimulus for PAR, and therefore it is suitable method for the research.

2.2 Methods

The data collection methods for the research were primarily qualitative, and include the personal journals of the researchers, email communications, observations, and interviews.

2.3 The Participant: VOLSET

As previously described, VOLSET is a small grass root NGO working in the towns of Ntenjeru and Nakaseke, Uganda. Its first foray into the use of ICT was in 1998 when it acquired a mobile phone for which it uses to communicate with government authorities and other similar NGOs. In the same year, a VOLSET volunteer developed the website and established an e-mail address. In 2003 they were donated a computer, a scanner and a printer. The computer was used to keep a record of volunteer information, typing letters for VOLSET and the community, the scanner was used to scan school report cards for sponsored children and as a photocopier. During 2005 the first digital camera was donated, which was used to capture images for the Web site, and of the sponsored children for emailing to their sponsors.

3. Reflection, Planning, Action, and Observation

The PAR cycle revolves around four elements: reflection, planning, action, and observation. The interplay between VOLSET and the research is described through the lens of these elements.

Reflection: To reflect one needs to have an understanding of what they have been doing, where they want to go, how to get there, and the influences along the way. Since VOLSET's establishment in 1997, and its first volunteer intake, there have been many opportunities for VOLSET to reflect its mission and accomplishments. The directors of VOLSET are the first to admit that their accomplishments would not have been achieved without the knowledge, time and skill transference from volunteers to the VOLSET community. This is not to say that it is a one-way reflection. Personal communications with volunteers show that they themselves have reflected on the skills and knowledge in particular that members of VOLSET and its community have transferred to them and how it has made an impact on their own lives. This co-construction of knowledge through collaboration is significant, and is visualised when one reflects.

Planning: Strategic planning and long-term planning is not a strength of ANON_ NGO, neither is it for many grassroots NGOs in developing countries. However, day to day and short-term planning is always through the lens of the mission of VOLSET. Much planning depends primarily on the resources available at any

given time, in particular money / cash. Volunteers come from across the globe, some directly via the website, others by word of mouth, and others through intermediaries such as Global Volunteer Network [Global Volunteer Network]. With every volunteer comes a plan, a plan that is devised what the volunteer joins VOLSET in Uganda. The plan is a nexus between what the volunteer brings to VOLSET- that is their skills, and what VOLSET needs 'today'. Furthermore, the volunteer may have had an agenda, aim or preference of their own, such as teaching English, or counselling or agriculture. VOLSET and the volunteers co-operate together to develop a plan for projects that are rewarding and beneficial to the VOLSET community and to the volunteer.

Action: Volunteers involvement is always focused within the bounds of VOLSET's mission bringing about actions and interventions. These interventions could be as simple as the payment of patient transfer to hospital or health clinic, the donation of a mobile phone with extended battery life, contribution to 'air time' for mobile phone communications or Internet connection, or more strategic as presented in the example below The addition of volunteers swell the human resources of VOLSET enabling it to physically be in more places, furthermore, with more people comes the opportunity to build the capacity of those whom they share their skills.

Observation: The last element of PAR before the cycle repeats itself in full or only in part, is observation. The organisation continually refers to their mission and on the many projects that they could explore to fulfill their mission. During the rollout of any particular project, leaders of VOLSET assessed its benefit to the community, financial commitment, use of human resources use of human resources (VOLSET members and volunteers) and its chance for success. It is observation by VOLSET members and volunteers alike that reflection, planning and action are legitimized. The following section presents an overview of one intervention related to the use of ICTs by VOLSET.

3.1 VOLSET, ICTs and one Intervention

The researchers spent five weeks in late 2005 with VOLSET. The professional skills that they brought with them were in education, ICT and microbiology. Upon arrival in Ntenjeru, they gathered information, observed, and participated in planning the flexible schedule for the following five weeks. ICTs played a significant part in their influence and contribution during their stay - and after they left. Of the number of projects that were conducted with and without ICTs, the one outlined below is the most significant and has actioned a significant change: the development of information systems to capture VCT data.

3.1.1 Pre-intervention

VOLSET regularly runs education seminars in mainland and island villages around and on Lake Victoria. They talk about the importance of healthy living, education, and VCT. An attendance list is recorded in a small notebook and used in

annual reports to illustrate the number of people attending these sessions. During these seminars another meeting time is organized where VOLSET specifically conducts VCT. VCT data has been collected by VOLSET since March 2004. When a person attends a VCT session, VOLSET writes the information into a small book. The data collected includes demographic data such as name, gender, age, village, occupation, and the result of their last HIV test. The dataset, however is often more incomplete than complete. For example, 132 people participated in VOLSET's VCT program between August 2004 and October 2005, during these consultations basic demographic data such as gender and age were only collected in 40% of cases, with marital status collected in 60% of cases. Using this current method of data collection, the information obtained during VCT was often ad hoc, inconsistent, incomplete, and difficult to retrieve for analysis and reporting.

3.1.2 *The Intervention*

The ICT intervention was multi-faceted with each element relating to each other in some way.

Education seminars. The first intervention was re-designing the attendance list collected at the education seminars. The purpose of this list was to record not who attended, but how many people attended, and their gender, therefore, the requirement for attendees to write their name were removed, and the following data were collected instead: village name, date, how many people attended, and their gender. The data are still collected in small exercise books, however the cover now includes a summary indicating that dates and the villages where the data collection was conducted. The summary data enables easy retrieval of the primary source data when and if required. Upon return to the VOLSET office where there is a computer and (intermittent) power, volunteers, senior school students or a member of the community entered the data into a spreadsheet.

VCT data. The second intervention was more detailed in that it involved the development of a detailed spreadsheet with appropriate data fields, embedded macros, and selection boxes. After consultation with VOLSET, additional questions were included such as number of children, and date of last HIV test. VCT participants are now given a small business card sized word-processed note stating the date and that they have participated in VCT, therefore giving them a record (and a reminder) of their VCT. In-the-field data collection booklet or sheets were designed to discourage incomplete responses, and to ensure consistent and comparable data were captured. Upon return to the VOLSET office volunteers, senior school students or members of the community entered the data into a spreadsheet. The dataset has now reached over 1800 with greater than 95% usable responses in regards to collection of basic demographic data between November 2005 and February 2008. The data collected over the last two years can be used by VOLSET to find ways to improve their service to communities, for example by determining what age groups are under represented in current VCT and understand what extent of home care is required by HIV positive community members.

The benefits from using the spreadsheets in particular the VCT spreadsheet is that the data can now be retrieved and manipulated for graphing and analysis showing a true picture of VCT and education, and thus can be embedded into reports. These reports can be used to attract funding or in annual reports that portray the work in progress of VOLSET.

Basic computer literacy. As a subsequence to the above there is a requirement to enter data into a spreadsheet, and an awareness in the community of the skill transference from volunteers to VOLSET members. These brought forth the need for rudimentary computer literacy training in such areas as data entry and word processing. Though not an information system, these skills are crucial to the adoption and advancement of ICT's, therefore face-to-face, one-on-one training was given by the volunteers, and a manual was written for village Ugandans on 'An introduction to word processing'. The face-to-face training together with the word processing manual have enabled the up skilling of a number of villagers and schoolchildren in basic computer literacy.

4. Discussion and Future Work

The project outlined in this paper is continuing and evolving as VOLSET develops a greater understanding of the power of good, reliable data, as well as further developing their ICT skills. The cycle of developing the spreadsheet, and the associate paper collection instrument, is evolving as VOLSET volunteers continue to work in the field and beyond. Not only do the volunteers see the benefit of what they are doing, and continue to refine the instruments, VOLSET itself is developing more effective process and systems to advance their cause.

It is through such projects as this, that the vision of those that lead Uganda can be realized as VCT becomes more common throughout the far reaching villages of Uganda, and the people become more aware and take action to overcome the HIV pandemic. Putting aside the benefits of the implementation of a computer based information system to capture, manipulate and report data; the people of VOLSET's exposure to computers and other ICTs, other than as typing and talking tools, is a change that cannot be ignored.

As a consequence to the work, the volunteers have set up a web site in their own country (<http://www.friends-of-volset.org>) to enable a wider dissemination of the work and respective outcomes. Future work, though is dependant on funding, involves extending the pilot study. The full study will be robust (methodological, conceptually and in execution) and explore opportunities and instances of the adoption of ITCs by grass-root NGOs in Uganda and, if possible to cross-validate the findings with other developing countries. The outcome of the full study is the development of guidelines that can be viewed by policy makers to hasten the adoption of ICTs, the skills required to benefit from them, and services and systems that take advantage of them.

References

- BABIKWA, D. 2005. *The POLICY Project in Uganda: Evaluation Report*. Prepared for the U.S Agency for International Development. <http://www.policyproject.com/countries.cfm?country=Uganda>
- COUTINHO, A. 2003. *A situation analysis of HIV/AIDS in Uganda and the role of VCT*. Retrieved October 2007, from http://www.aidsuganda.org/pdf/role_of_VCT.pdf
- GARSON, D (2008) *Ethnographic research*. Retrieved February, 2008 from <http://www2.chass.ncsu.edu/garson/pA765/ethno.htm>
- GLOBAL VOLUNTEER NETWORK. <http://www.volunteer.org.nz/>
- HAMMERSLEY, M. and ATKINSON, P. 1995. *Ethnography: Principles in practice, Second Ed.* London: Routledge.
- JOHNSON, A.G. 2000. *The Blackwell Dictionary of Sociology, Second ed.* Oxford, UK: Blackwell.
- KEMMIS, S., and MCTAGGART, R. (Eds.) 1988. *The action research reader* (3rd ed.), Geelong: Deakin University Press.
- UGANDA COMMUNICATIONS COMMISSION. 2007. <http://www.ucc.co.ug/marketInfo/marketstatistics.php>
- THE UGANDA NATIONAL NGO FORUM. 2003 <http://www.ngoforum.or.ug/>
- UGANDA NATIONAL COUNCIL FOR SCIENCE AND TECHNOLOGY (UNCST). 2002. *Status of ICT in Uganda: SCAN-ICT Preliminary baseline report*. <http://www.uneca.org/aisi/ScanGhana/Documents/3.%20SCAN%20ICT%20UGANDA.pdf>
- UNITED KINGDOM, PARLIAMENTARY OFFICE OF SCIENCE AND TECHNOLOGY. 2006. *ICTs in developing countries, 2006*. (March, 2006 Number 261). Retrieved September 2007 from <http://www.parliament.uk/documents/upload/postpn261.pdf>
- VOLSET. <http://www.volset.org> (earlier web site can be found at <http://volsetuganda.tripod.com/>)
- WORLD HEALTH ORGANIZATION. <http://www.who.int/hiv/topics/vct/en/>.
- WORLD PRESS.ORG. <http://www.worldpress.org/Africa/2074.cfm>

12

Career Guidance Using Expert System Approach

Ojenge Winston and Muchemi Lawrence

The lack of quality career guidance in African high schools, caused by shortage of human and time resources that the process demands, has led to the choice of unsuitable careers resulting in widespread poor performance at the workplace. The focus of this paper is the use of Expert Systems that store the highly specialized knowledge of career counseling and dispenses it to the needy youth in a cheap and quick process. Initial results of the evaluation of the model, has shown a career diagnosis accuracy of 71%.

1. Introduction

A study, done as part of this research, and detailed later in the paper, concluded that there is a mismatch between the careers most high school graduates choose and the ones of their natural interest.

Appearing in a report by Osoro et al [2000], are results of a separate study that concluded that one of the main reasons why people are later dissatisfied in their vocations is uninformed choice of vocation immediately after school. Interviews with Ministry of Education (Kenya) officials and career guidance teachers confirmed that students are simply given the careers booklet with university courses, their prerequisite subjects and cut-off points, instead of career guidance and counseling. This is due to limited skilled human and time resources. This problem may be minimized by the use of an Expert System that would store the specialized knowledge of the career counselors and dispense this knowledge in under thirty minutes to a high school graduate.

There is, clearly lack of a quality and cost-effective Career Guidance Model. This paper presents results and analysis of the studies mentioned above, the process of identifying desirable activities in the career guidance process and the integration of those activities into a career guidance model using Expert Systems, for easy and quick querying by a high school graduate. Based on the proposed model, an application may be constructed to provide instantaneous career guidance to high school graduates.

2. Related Approaches

2.1. Career Guidance Activities

Most reputable career guidance models involve the following activities; student visits to prospective colleges to get a first hand experience of training tasks; job simulation to enable the student get a first hand experience of tasks in the job environment; exposing students to resource persons and industry people. White [2002] reports a model that integrates a personality analysis, besides the above-mentioned activities. The premise for the use of the personality analysis to diagnose one's natural career is that human performance in a job situation is dependent on;

- a) One's ability to comprehend certain unique cognitive concepts demanded by a job situation.
- b) One's natural disposition toward certain tasks and environments and not others [Naralmbaiah and Yan Wah, 2004]

Point a), above is about scholastic aptitude ability and is addressed by most education system's college entrance examinations.

Natural disposition toward certain tasks, objects and environments is defined as personality. It implies that if one's personality type could be determined, then their natural tasks and working environments could also be determined [Keirsey and Bates 1998].

2.2. Personality Analysis

The main personality analysis models, as presented by Quenk [1999] are the following; The Enneagram, the Big Five, the Astrology Sign, the Aura Color models and the Myers Briggs Typology Indicator (MBTI). Other personality analysis models are derivatives of the above-mentioned ones. A summary of the approaches in each of these models is briefly presented here below;

The Enneagram model is based on theology. It identifies which of the 9 personality types, one belongs to. It also describes the nature of the conflicting forces of each personality and the cycle of change that is constant in human nature.

The Astrology Sign model analyses personality by determining which planet rules one's sign of zodiac. It then indicates one's associating quadruplicate (one of four attributes). It then reveals which of the four elements (air, earth, water and fire) one most relates to in order to show how that element affects one's sign's predisposition.

The Aura Color model is based on color psychology. It proposes that each person has a natural inclination toward certain colors and a natural dislike towards others, and that one's favorite color points to their personality.

The above-mentioned models are not founded on solid science, hence are not reliable.

The Big 5 model considered as one of the most accurate models [Quenk 1999], is based on cognitive science. It defines 5 distinct personality traits, namely; level of extroversion; level of orderliness; level of emotional stability; level of accommodation; level of inquisitiveness.

Each of the above parameters is measurable in five degrees. A user's set of responses to appropriately designed questions may, therefore, be 1 of 125. This becomes its main weakness, as it introduces too many groupings (125 different ways to respond) which are hard to model.

The Myers-Briggs Typology Indicator (MBTI) measures 4 significant human attributes; Personality, Outlook, Temperament and Lifestyle to identify one's personality type.

Under Personality, one can either be extroverted or introverted. Under Outlook, one can either be sensory (uses any or a combination of the five senses to comprehend situations and make decisions) or intuitive (imagine beyond sensory experience). Under Temperament, one can either be thinking or feeling. Lastly, under Lifestyle, one can either be judgmental or perceptive.

Appropriate questions are presented to a test-taker. The taker's responses to these questions are used to evaluate his/her personality type, which clearly is one out of a possible 16 (four attributes, each having two options implies 16 possible sets of responses).

MBTI is therefore known as the 16-Personality Type model.

Each of the 16 personality types has been competently linked, after decades of research, to a set of related job categories [Keirse 1995]. The study reported here adopted this model due to ease of grouping and its firm grounding in psychology.

3. Methodology

The approach used in the research reported in this paper involved field surveys from which observations and analysis were done. The conclusions were used in developing a model which is presented in later sections of this paper. In order to assess the level of effectiveness, an experiment was designed and results evaluated. These are also presented.

3.1. Surveys

In a survey carried out in 2007, to estimate the level of professionals' satisfaction with the tasks and nature of their careers, 314 professionals were interviewed. They were over 35 years of age and drawn from 16 different professions around Nairobi city. The results are detailed in Table 1 and Fig. 1 here below. In a second survey, also done as a part of this research in 2007, Ministry of Education (Kenya) officials and career guidance teachers in 22 randomly-sampled high schools across the country were interviewed. It was in an attempt to determine what career guidance practices are carried out in Kenyan high schools. The results of the survey are detailed in Table 2 below.

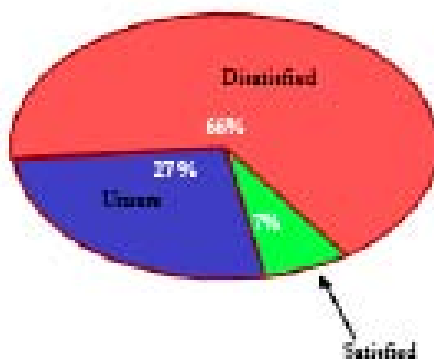
3.2. Survey Results and Analyses

3.2.1. Extent of Satisfaction with Vocation

Table 1 - Extent of Satisfaction with Vocation

	OPINION	NUMBER & PERCENTAGE
1.	Those who are sure they chose wrong	207 (66%)
2.	Those who have not really thought about it	84 (27%)
3.	Those who are sure they chose right	23 (7%)

Fig. 1: Extent of satisfaction with Vocation



As can be seen in Table 1 and Fig. 1 above, only 7% of survey participants were satisfied with the tasks and nature of their vocation, with 66% being unsatisfied. This survey concluded that there is a mismatch between the careers most high school graduates choose and the ones of their natural interest.

3.2.2. Career Guidance Practices in Kenya

Table 2 below shows the various categories of schools and the means of career guidance being practiced currently.

Table 2 - Career Guidance Practices in Kenyan High Schools

	TYPE OF HIGH SCHOOL	PRACTICES
1	Private Academies	Trips to university fairs; Individual attention to students; Career debates among students
2	High Cost Public Schools	Counselor discusses with interested student. No full time counselor as each counselor has teaching load too
3	Other Public Schools (nearly 90% of all Kenyan high school students)	No real counseling done, only instruction on use of the careers booklet, which contains a list of college courses and their cut-off points

Approximately 90% of public high school students in Kenya are not provided any reasonable career guidance. This is due to limited skilled human and time resources. This research concentrated on public schools because they form a bulk of the student population. This problem may be minimized by use of Expert Systems, which would store the specialized knowledge of the career counselors and dispense this knowledge as required by potential users.

3.3. Model Design

From the description of the activities reviewed in Section 2.1 above, we propose that a suitable system should have the following features;

- a) Personality Analysis module
- b) Decision-making classes, class discussions; Visits to prospective colleges so that the students get first hand experience of training tasks in each course; Job simulation to enable student make informed decisions. These can be grouped into a simulation module which we here call the auxiliary module.
- c) Scholastic Aptitude Testing (SAT), such as college-entrance examinations, to evaluate one’s cognitive ability. This enables determination of the level of college course the student should join.
Specifically whether at certificate, middle-college diploma or degree level.

We propose a model that consists of three sections; the first being the Personality Analysis Expert System, with the last being the College-Entrance Criteria Expert System.

The middle section contains a simulation of activity b) above. To enable one to choose a specific job category from the small set recommended by the Personality Analysis, interactive decision-making and class discussion sessions can be modeled by the use of multimedia or other techniques. The proposed Career Guidance model would appear as illustrated in Fig. 2.

Fig. 2: Proposed Career Guidance Model

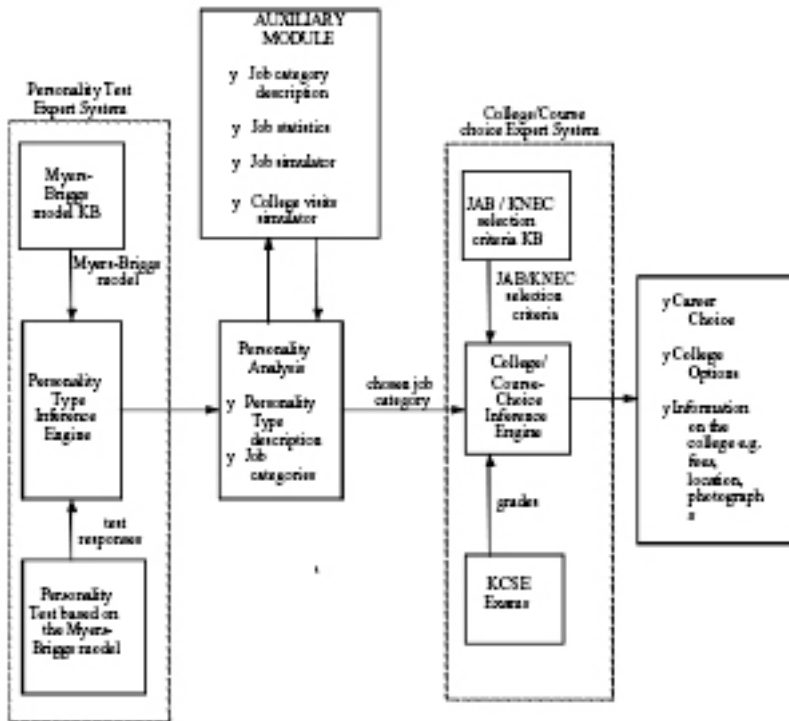


Figure 2 above shows the architecture of the proposed model. It indicates how various activities are integrated. A description of the working of the core modules is presented briefly here below.

3.3.1. The Personality Analysis module

This module hosts the knowledge and rules needed to perform personality analysis. The knowledge and rules are based on the Myers-Briggs Typology Indicator (MBTI) model. The following is short description of the working of this model;

According to Keirsey [1995], there are four possible human natures Artisan, Idealist, Guardian and Rational. Parameters for classifying these human natures are indicated below;

1. Personality - Extrovert (E) or Introvert (I)
2. Outlook - Sensory (S) or Intuitive (N)
3. Temperament - Thinking (T) or Feeling (F)
4. Lifestyle - Judgmental (J) or Perceptive (P)

Note: The letters used against a name indicate abbreviations used later on in the paper or in the prototype.

As shown in Table 3 below, an Idealist nature is one whose Outlook is Intuitive (N) with a Temperament of Feeling (F). If his/her Lifestyle is Judgmental (J), then he/she is an Idealist who is a Mentor. If his Personality is Extrovert (E), then he/she is a Mentor of personality type ENFJ and is fit for the careers of Teacher, Counselor or Home Economist. Artisan, Guardian and Rational natures can be similarly analyzed into their more detailed derivatives, to determine the appropriate careers.

Table 3 - Personality Analysis and job suitability Chart

OUTLOOK	HUMAN NATURE	TEMPERAMENT/ LIFESTYLE	PERSONALITY	
Introspective (Intuitive), N	Idealist (Feeling), NF	Mentor (Judgmental), NF J	Teacher/Counselor/Home Economist, ENFJ Counselor/Teacher/Psychiatrist, INFJ	
		Advocate (Perceptive), NFP	Childcare/Secretary/Counselor, NFP Psychologist/Editor/Artist, INFP	
	Rational (Thinking), NT	Coordinator (Judgmental), NTJ	Manager/Lawyer/Operations Research, ENTJ Sales/Journalist/Photographer, INTJ	
		Engineer (Perceptive), NTP	Engineer/Scientist/Social-Scientist, ENTP Scientist/Engineer/Writer, INTP	
		Guardian (Judgmental), SJ	Administrator (Thinking), SJT	Financial Manager/Manager/Supervisor, ESTJ Accountant/Engineer/Technician, ISTJ
			Conservator (Feeling), SJF	Counselor/Childcare/Social Worker, ESFJ Health-worker/Teacher/Librarian, ISFJ
Artisan (Perceptive), SP	Operator (Thinking), SPT		Builder/Plumber/Carpenter, ESTP Craft Worker/Statistician/Technician, ISTP	
	Entertainer (Feeling), SPF		Actor/Musician/Teacher, ESFP Painter/Outdoor/Stock-clerk, ISFP	

Responses to a set of questions given by the user, as shown on the interface of the prototype illustrated in Fig. 4, determine which of the possible sixteen personality types the respondent belongs to. Using the personality type, the model would determine which career one is best for [Keirsey 1995].

3.3.2. College Entrance Expert System

Many colleges have some criteria based on grades scored in selected subjects in specified examinations, for admitting students. The rules in this module should be derived from the criteria each specific country, state or region uses. For the prototype described here below, the Kenya national examination council (KNEC) grading system, Kenya certificate of secondary examination (KCSE) and public university joint entrance criteria, Joint admission board (JAB) were modeled.

3.4. The Prototype and Evaluation of the Model

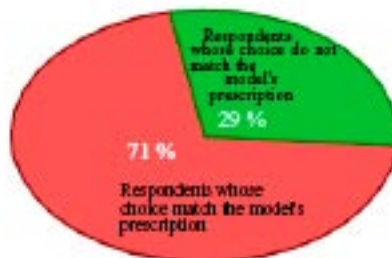
In order to assess the effectiveness of the model presented above, a prototype implementing the designs described was developed. A typical screen showing the user interface and the type of questions is shown in Fig. 3;

Fig. 3: A Typical Interface of the Developed Prototype

The implementation was done in Visual Studio and MS Access. The choice of the tools was purely on ease of coding.

The prototype was tested on 104 professionals, all over 35 years of age, from diverse occupational backgrounds. Of the 104 professionals, 74 respondents representing 71% of the sample had the occupations they desired match their personality types. The types were automatically determined by the prototype depending on the answers given by the respondents. A summary of the results is shown in Fig. 5

Fig. 4: Respondents Whose Desired Occupations Match Personality Types



The assumption made during evaluation is that at over 35 years of age, most literate professionals have been exposed to different environments and people and therefore they know their strengths and weaknesses, likes and dislikes. They also know almost all possible career opportunities that are available.

4. Conclusion

This paper has demonstrated a model based on expert system concepts that can achieve effective career guidance objectives. It was also established that a significant percentage of professionals in Kenya are dissatisfied with the tasks and nature of their current vocations. Career counselors play an insignificant role in career choice by public high school students and that most Kenyan high school students do not get quality career guidance as qualified counselors and time for guidance is lacking.

Through prototype evaluation, we have shown that if public high schools in Kenya implemented a system based on this model, then it is possible to achieve a career satisfaction level of up to 71% from the current 7%. This is a clear improvement.

References

- KEIRSY, D. 1995, *Portraits of Temperament*. Del Mar, California: Prometheus Nemesis Book Company; 3rd edition
- KEIRSEY, D. AND BATES, M. 1998 *Please Understand Me II: Character and Temperament Types*. Del Mar, California: Prometheus Nemesis Book Company
- LEONDES, C. 2002. *Knowledge Based Systems*, English Language Edition, Eisevier Science and Technology Books
- NARALMBAIAH, G. AND YAN WAH L. 2004. Who Should Work With Whom? *Building Effective Software Teams*. Communications of the ACM Archive 47, 79-82
- OSORO, B. ET AL. 2000. *Career decision-making of high school students in Kenya*. An International Journal for the Advancement of Counseling
- QUENK, N. 1999. *Essentials of Myers-Briggs Indicator Assessment* (Essentials of Psychological Assessment). Wiley
- WHITE, P. 2002. College Selection Program, Department of College Guidance, Summer Country Day School.

Exploring the Implementation of Blended Learning in a Developing Country: A Case Study of Uganda

Philip O. Ayoo and Jude T. Lubega

This paper explores on how the blended learning method has been implemented in the developing country such as Uganda. Many higher education institutions in Uganda have acquired a new form of learning using technology that they combine with the traditional learning methods. Implementation of this form of learning using technology has been challenged by several factors. Many of the challenges are attributed to the conditions that exist in a developing world such as lack of appropriate technology. In spite of the several challenges blended learning has been implemented in relation to the existing internal procedures to suit the developing world context. The study strategy aimed to identify if the procedures initiated for implementing blended learning in the developing world were suitable. A case study approach was adopted in order to inspect these implementation procedures used within a higher education setting. The results from the study indicate that despite the lack of the appropriate technologies, infrastructure, pedagogical knowledge, quality assurance procedures and capital to invest, hybrid policies/frameworks are implemented. For this reason, a lot of higher education institutions are falling short of the implementation policies adopted. Therefore the paper stipulates that it is not simply adopting what the developed world has done but designing what suits the contextual setting that the particular country poses.

1. Introduction

Traditional learning methods do not meet the contemporary needs of our information society any more (Gütl *et al*, 2004). Traditionally learning was carried out in a designated place, at a particular time and by a known tutor. To exemplify this, Jay Cross observes (in Downes, 2007), “For sixty years, we’ve thought of learning as residing in the formal models exemplified by schools, universities, and training programs. Common to these top-down formats is a curriculum that rests on the beliefs and worldview of the authorities in charge.”

Advances in technology all over the world have contributed to the paradigm shift in several areas such as education. Today learning can occur ubiquitously due to the technological factors that have broken the geographical barriers to learning. Detecon (2002) observes that the potential of e-technology for the developing countries is breathtaking, and the significance of this for e-learning

is huge. Especially in Africa, the current status requires innovative ways to support education for all (UNESCO-IITE Report, 2003). Thus, technology-based solutions have been increasingly established to overcome these problems. The need to invent more innovative ways of providing learning is being driven by the changing conditions and unique contexts in which learners find themselves. Devlin *et al.*, (2001) note as follows:

“Half the world’s population is under age 20; we have essentially no hope of educating these young people by conventional means. New learning modules can provide access to high-quality learning opportunities for a wide variety of students, the vast majority of whom will never have the luxury of a residence-based higher education.” (Devlin, et al. (2001) in Larson and Wasserstein (2003))

According to Downes (2007), learning will increasingly be available not so much in learning institutions but in any given environment in which learners find themselves. Many developed countries have adopted the use of technologies to enhance learning in almost all educational sectors. This has further increased the need for lifelong learning especially in the adults. Developing countries have also appreciated the need for this form of learning that can be carried out ubiquitously.

According to Gütl, *et al.* (2004), the main objectives of an innovative solution for an e-learning system are:

- i) personalised retrieval, management and presentation of relevant and timely information for learning activities;
- ii) support of various learning and teaching paradigms; and
- iii) improved knowledge of the users’ behaviour.

Innovative learning environments are not so much dependent on the use of ICT, but rely more on the organization of the learning situation and the ability of teachers to use technology to support pedagogical learning objectives that change and transform traditional learning activities (Grünbaum, *et al.*, 2004). According to Dalgarno (1996), the changes that have occurred in accepted approaches to teaching and learning in recent years have been underpinned by shifts in psychological and pedagogical theory, culminating in moves towards a constructivist view of learning.

The future direction of e-learning has been defined as “blended learning” (Mitchell, 2001). Looking at Uganda in particular as a developing country, emphasis on the use of technology during education has been encouraged by the government. Many educational institutions have adapted the use of technologies for enhancing the traditional classroom teaching. How the technology has been implemented varies from one institution to another depending on the needs and resources. The growing demand for post-secondary education and the teacher-student ratio in Uganda has also encouraged the adoption of blended learning. This form of flexible learning is well sought by many institutions to lead them in attempting the major challenges to delivering learning. Due to the wide implementation of blended

learning in Uganda today, it is significant that strategies, policies, frameworks be explored and reviewed for effective use.

2. Brief Background

According to Brennan (2004), there have been six major waves of technological innovation in learning, and this is applicable to both developed and developing countries. Writing can be considered as the first technological innovation, with phonetic alphabets, papyrus and paper. Printing was the second, with moveable type. The third was broadcast media such as film, radio and television. The fourth was a range of mass media storage devices including audiocassette, videotape and CD. The fifth was the mass produced computer with CD-ROM. The sixth is the current networked, web-based e-learning revolution. With each of these innovations, new forms of blended learning arose, and the various stages overlap depending on the particular application context.

Blended learning therefore is not new, as it has been driven by a series of technological innovations in learning for many centuries (Brennan, 2004). What has given the term 'blended learning' a new impetus is the exciting and powerful array of web-based options that have cascaded on to the market over the last few years. Institutions that formerly relied purely on students gathering in campus-based classrooms are suddenly able (and many seem eager) to offer their programming on the Internet. Similarly, institutions accustomed to large-scale distance delivery via print or television are now being asked to provide more flexible, interactive, and responsive net-based alternatives.

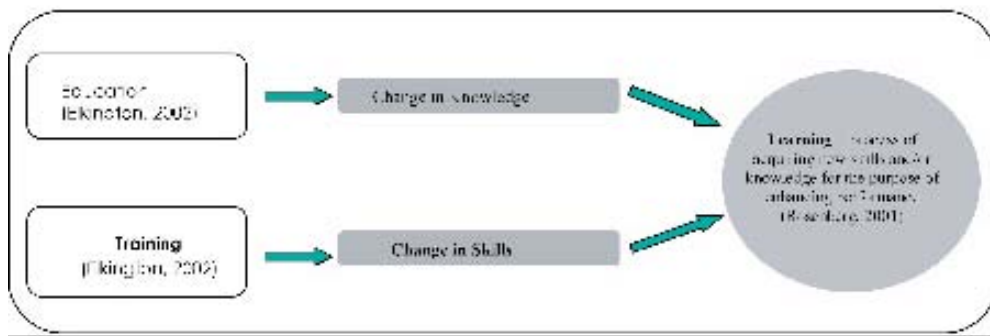
In Uganda, e-learning is now becoming popular in providing college education in most institutions of learning and in the workplace. Distance education, especially of the correspondence type, has been at the vanguard of these developments, but campus-based students are also mixing and matching their classroom and online learning in all sorts of often unanticipated ways.

The transformation of the education sector in order to embrace the new learning paradigms is being driven by a number of broad economic, technological, and social trends that have accelerated in recent years. One of the key ones is the significant increase in the demand for higher education in both developed and developing countries. LaRocque and Latham (2003) estimate that the overall demand for higher education is expected to grow from 48 million enrolments in 1990 to 159 million in 2025 – an annual growth rate of 3.5 percent. From an economic perspective, the International Data Corporation (IDC, 2004) has estimated that world revenues in the corporate e-learning market will surpass \$21.1 billion by 2008, compared with a \$7.9 billion market in 2004 (van Dam, 2005). According to this estimate, the market for online courses offered at universities will continue to grow about 25% to 30% a year. For this reason, it is important to explore how developing countries are embracing the new learning environment, with a view to establishing the inhibiting factors and suggesting probable solutions.

3. Putting Blended Learning Into Context

To understand the concept of blended learning, it is good to understand the concept of learning. van Dam (2005) believes that formal learning occurs to a large extent through education and training. Elkington (2002), as cited in van Dam (2005), defines training as a change in skills, and education as a change in knowledge. van Dam (2005), therefore, concludes that learning is the process of acquiring new skills and knowledge with the goal of improving performance. He summarizes the relationship between education, training and learning in **figure 1** below.

Fig. 1: The Relationship Between Education, Training and Learning (Adapted from van Dam, 2005)



With the introduction of modern technology to deliver learning, several terms have been introduced, including distance learning, e-learning, online learning, computer-based learning, and web-based learning, etc. Kruse and Keil (2000), as cited by van Dam (2005), acknowledge in their book *Technology-Based Training*, that there is much overlap and convergence in the definitions with regard to technology and learning. Over the past few years, training professionals have become more pragmatic in their approach to technology-based media by using it to augment traditional forms of training delivery, such as classroom instruction and text-based materials (Brennan, 2004). This trend has led to the rise of the term “blended learning.”

In defining “blended learning”, most authors concentrate on the word blended. To blend means to “form a harmonious compound, become one” (Concise Oxford Dictionary 8th edition, 1990). Blended learning, therefore, refers to the use of two or more distinct methods of training (Brennan 2004). This may include combinations such as blending classroom instruction with online instruction, blending online instruction with access to a coach or faculty member, blending simulations with structured courses, blending on-the-job training with brown bag informal sessions, or blending managerial coaching with e-learning activities (Brennan, 2004).

Dr. Margaret Driscoll identifies four different concepts in defining blended learning (Wikipedia). The first defines blended learning as meaning “to combine or mix modes of web-based technology (e.g. live virtual classroom, self-paced instruction, collaborative learning, streaming video, audio, and text) to accomplish an educational goal” (Driscoll, 2002). The second definition describes blended learning as meaning “to combine various pedagogical approaches (e.g. constructivism, behaviorism, cognitivism) to produce an optimal learning outcome with or without instructional technology” (Driscoll, 2002). The third definition from Driscoll (2002) defines blended learning as meaning “to combine any form of instructional technology (e.g. videotape, CD-ROM, web-based training, film) with face-to-face instructor-led training.” The fourth concept defines blended learning as meaning “to mix or combine instructional technology with actual job tasks in order to create a harmonious effect of learning and working” (Driscoll, 2002). In summing this up, Valiathan (2002) describes blended learning as a method that mixes various learning events including traditional face-to-face classroom activities, live e-learning and self-paced learning. This learning method includes a variety of tools that create flexible rich learning environments that simulate and maximize the learner’s potential to learn.

Research has indicated that providing a learning process with a variety of methods through which learners can acquire knowledge improves their learning potentials (Dean et. al., 2001; Lubega and Williams, 2003; DeLacey and Leonard, 2002). This is the reason why this learning method has been adopted in many institutions of learning all over the world. Many researchers have experimented on single learning approaches and made conclusions to reflect that there is no difference in the learning output between the methods.

Blended learning in the developing countries has been carried out in several ways. These include use of the traditional classroom teaching, CD-ROMs, TV, Radio, and Managed Learning Environments (MLE). The blend has often been done by combining any of the above learning methods depending on their availability.

According to Singh and Reed (2001), blended learning focuses on optimizing achievement of learning objectives by applying the “right” learning technologies to match the “right” personal learning style to transfer the “right” skills to the “right” person at the “right” time. Embedded in this definition are the following principles:

- i) We are focusing on the learning objective rather than the method of delivery
- ii) Many different personal learning styles need to be supported to reach broad audiences
- iii) Each of us brings different knowledge into the learning experience
- iv) In many cases, the most effective learning strategy is “just-what-I-need, just-in-time”

The experience of pioneers in blended learning shows that putting these principles into practice can result in radical improvements in the effectiveness, reach and cost-effectiveness of learning programmes relative to traditional approaches (Singh and Reed, 2001).

4. Situation in Uganda

Worldwide, classroom training has served as the primary means for delivering behavioral skills training for years. Nevertheless, its effectiveness at delivering lasting instruction on its own is now being questioned (Snipes, 2005). In Uganda, like in many of Africa countries, higher education is being challenged by new opportunities relating to technologies that are improving the ways in which knowledge is produced, managed, disseminated, accessed and controlled. In order to open up and widen access to education to the public, a number of education institutions in Uganda are currently establishing a distance education (DE) component in their programmes.

The education system in Uganda covers eight years of primary (basic) education, four years of ordinary level secondary education, two years of advanced level secondary education, and two-to-five years of tertiary education. Alternative tracks branch off from ordinary level secondary to technical colleges and primary teachers colleges. Tertiary education covers post-advanced level secondary education, which includes universities and other institutions of higher learning such as polytechnics.

Uganda, as a partner of the Education for All (EFA) coalition, launched Universal Primary Education (UPE) in 1997. This resulted in the increase of primary school enrolment figures from 2.7 million pupils in 1996 to 5.3 million in 1997, and to 7.1 million in 2005 (Ministry of Education and Sports, 2005). On the other hand, most of the growth in tertiary education in the last decade in Uganda has been in the universities sub-sector, accounting for 65% of total tertiary enrolment (Ministry of Education and Sports, 2005). The remaining 35% are shared by other tertiary institutions.

By 1995, Uganda's gross enrolment ratio at tertiary level was 2%, way below the 3.2% sub-Saharan average. However, between 1995 and 2003 tertiary enrolment increased by 230% (Balihuta, 2001). The annual average rate of increase in tertiary enrolments has been 46% per annum in the last decade (World Bank, 2000). This raised the gross enrolment ratio to an estimated 2.8%. The projected demand for higher education with improved access to secondary education is expected to reach 126,396; making up only 3.1% gross enrolment ratio (Ministry of Education and Sports, 2003).

Of the four public universities – Makerere University, Kyambogo University, Mbarara University of Science and Technology, and Gulu University – only the first two offer distance education programs. Both Makerere University and Kyambogo University are dual-mode institutions. Although the Uganda Management Institute (UMI) is not ranked as a university, it is a degree-awarding

higher institution of learning that also has a DE component. Nsamizi Training Institute for Social Development (NTI) is yet another tertiary institution offering at least one program through DE.

Of the twelve private universities in Uganda, only Uganda Martyrs University (UMU), and Uganda Christian University (UCU) are known to run distance education programmes. Kampala International University too has mooted plans to start computer-based DE programs in the near future.

Because nearly all the DE programs in Uganda are within dual mode institutions, the main focus when introducing ICTs in these institutions has not been directed at ICTs for enhancing DE delivery; rather the emphasis has been on ICTs for enhancing face-to-face delivery. For instance, Makerere University and Kyambogo University have made evident strides in introducing ICTs in the management and teaching/learning processes but no infrastructure developed or policy formulated puts a primary emphasis on DE. The proposed single-mode Open University of Uganda is the only institution of higher learning in Uganda whose policy documents and proposed structures put primary emphasis on ICTs for DE. The Open University of Uganda proposes to use radio, video conferencing, satellite-based and computer-mediated instructional technologies as the main media of instruction.

In order to implement technology-enhanced learning, the question of existing ICT infrastructure becomes important. Prior to 1996, Uganda's communication infrastructure was among the least developed, not only in the World, but also in Africa (Wasukira and Naigambi, 2002). Further more, 70% of the communication services were concentrated in urban areas, leaving the rural areas with the least access to communication services. However, as a result of the liberalization policies adopted by the Uganda Government during the 1990s, the infrastructure situation has now changed drastically. As a result of the liberalization, the telecommunications sector has experienced tremendous growth, which is typified by:

- i) New market entrants have brought in a whole range of satellite-based technology
- ii) Two national operators have been licensed
- iii) Four mobile phone operators have been licensed
- iv) Cellular mobile phone services marked a turning point in business trends

There is growing evidence of success in infrastructure development and growing consensus that new initiatives that build on these successes are needed, and this bode well for the future (UNCST, 2002). New opportunities exist to strengthen infrastructure and to allow delivery of services in society. For instance, the laying of an optic fibre cable network is ongoing connecting the main towns, and infrastructure rollout programmes to cover the rest of the country with fibre-optic broadband bandwidth is now being extended to the major business districts outside the capital.

A comprehensive National IT Policy is in place and being implemented. There is a rural communications development policy to cater for bridging the digital-divide with the rural communities. The vision is to provide “bandwidth-for-all by 2010” and improve the infrastructure to extend to the rural communities. A national ICT master plan to cover e-education and e-health is also underway.

In spite of this, however, the level of infrastructure and services are still way below the average compared with other economies in the World (Wasukira and Naigambi, 2002). Moreover, most of the developments are still concentrated in urban areas, benefiting a small percentage of Ugandans.

5. Strategies to Enhance Blended Learning in Uganda

In order to successfully mainstream blended learning into the culture of higher education institutions in Uganda, some strategies can be recommended.

5.1. Culture

In order to make available and put to use the potential of the technology, there are two major obstacles: infrastructure and culture (Detecon, 2002). There are certain cultural traits that further or hinder a country’s participation in the global village life. The culture needs to be open or, better still, positive towards learning and knowledge as well as the delivery technology (Detecon, 2002). People need to have access to the system and the freedom to use it (regarding information and communication). The following questions need to be answered: What is the cultural attitude towards learning in the target population? Does it differ by the learner’s age or gender? Are there taboos regarding learning? What is the cultural attitude towards the technological aspects of e-learning? What are the socially/culturally accepted/preferred learning styles?

5.2. Communications Infrastructure

For e-learning to succeed in the developing world, it needs to build on another important pillar: the existence of infrastructure, along with some degree of connectivity (Sehrt, 2005). One must put in place, or make available, electrical power, the network, hardware and software to deliver e-learning, and the content and services ready for use (Detecon, 2002).

Uganda is one of the nations falling on the other side of the digital divide (i.e. the have-nots), with telecommunications infrastructure among the least developed. No countrywide telecommunications network has been established in Uganda, and given the current low level of infrastructure development for ICTs, it is obvious that the majority of Ugandans will not be able to have access to DE through these technologies. In Uganda, up to 80% of the digital telephone lines and modern switching equipment is located in the capital city, Kampala, with other areas having largely old and unsuitable lines. Resulting problems include high usage costs due to the low rates at which data can be transmitted or received, high percentage of failure due to poor quality of the lines, and limitations on the

applications that can be used (in some areas, for instance, it is not possible to access the web).

In order to support and deliver blended solutions, it is crucial to address the following issues from a technical perspective:

- i) What learning delivery infrastructure is currently in place?
- ii) What is the desired technical environment given the stated goals?
- iii) What access does the target audience have to technology?
- iv) What information (if any) should be recorded/tracked to address the agreed goals?
- v) What authoring environment/features (if any) are required to address the agreed goals?
- vi) What technical standards should be defined for training design and development?

5.3. Human Infrastructure

The introduction of the new technologies has to be combined with the building of the skills necessary to operate and maintain them. Consequently, to succeed, blended learning initiatives require a supportive human infrastructure. Computer literacy is an imperative precondition for learners to benefit from technology-based learning (Sehrt, 2005). e-Learning can only build on a set of basic computer literacy skills. Hence one of the first tasks is to assess the existing capability to implement and support blended learning. The following issues must be considered:

- i) What skill-sets are required to address the learning goals?
- ii) What new skills will be required? How will these skills be developed?
- iii) Who will design, develop and deliver the solutions?

5.4. Personal Prerequisites

On the personal level, e-learning needs “connectivity” in terms of language and literacy to succeed: either the learners are able to understand the language the content is presented in or the content has to be produced in the learners’ native language. The same holds true for literacy: either the learners are able to read and write and use the computer or the computer’s interface has to be adjusted according to the literacy level.

The lack of content in native languages is a serious impediment to Internet use in many countries. Even in the developed world, the preponderance of English online poses a serious obstacle to universal access.

5.5. Content

When we are considering e-learning we are of course talking about technology. However the best technology in e-learning is irrelevant if the content is not available (Detecon, 2002). The content needs to be in the adequate formats in order to enable success. Strategies must be in place to develop learning content

that address the specific needs of the learners depending on the individual learning contexts.

5.6. Quality Assurance (QA) Processes and Procedures

Naturally, one must ensure that the blended learning solutions adhere to best practice both in terms of development methodologies and delivery techniques. The following questions may guide this review:

- i) How are learning materials currently designed and developed?
- ii) What QA processes are in place?
- iii) Are any formal quality procedures, such as ISO 9001 in place?
- iv) What guidelines and methodologies are in place for designing and developing learning?
- v) Are materials used for accreditation/certification against professional standards/qualifications?
- vi) What is your policy for accessibility and usability?
- vii) How are materials maintained and updated?

5.7. Regulatory Situation

There is need to establish the policies that are in place to guide the use of various technologies for learning. The following questions need to be answered:

- i) What is the regulatory situation regarding information, communication and technology?
- ii) How complicated (if possible) is it to set up/obtain access to electronic services?
- iii) Is there censorship in place regarding information and communication?

5.8. Evaluation

Before implementing blended solutions, it is crucial that you determine how success will be measured. Such measures should be built into the plan and a process established for continual evaluation. How should the solution be evaluated in order to learn lessons for the future? There is need to determine:

- i) How are learning initiatives currently evaluated? How often?
- ii) Who is involved/consulted in the evaluation process?
- iii) How are plans revised in light of new developments?

6. Conclusion

The continuous advances in technology enable the realization of a more distributed structure of knowledge transfer (Iahad, 2004). This becomes critically important for developing countries that lack the resources and infrastructure for implementing cutting-edge education practices. The two main themes of technology in education for developing countries focus either on aspects of technological support for

traditional methods and localized processes, or on the investigation of how such technologies may assist distance learning. Commonly such efforts are threefold, relating to content delivery, assessment and provision of feedback. **As noted by Nashua (2003)** blended learning has too often been viewed as some kind of hastily mixed ‘learning stew,’ where learning institutions create a potluck combination of instructor-led classes, self-study courses, library resources, and various seminars and expect it to work. The successes of such models are never as great as they could be if these institutions used a more methodological approach to developing blended learning programs.

Learning institutions with greatest success in blended learning are those that take a structured, methodical approach to creating defined curricula (Nashua, 2003). These institutions analyze the strengths and weaknesses of various learning activities in the context of a blended learning environment and build the curricula accordingly. They also know in advance how results and associated value will be measured. This paper has attempted to establish the opportunities that exist for proper implementation of blended learning in a developing country, taking the case of Uganda. Obstacles that hamper use of technology to enhance learning have been identified, and strategies to overcome these have been suggested.

References

- BRENNAN, M. (2004) Blended Learning and Business Change.
http://www.clomedia.com/content/templates/clo_fairfield.asp?articleid=349&zoneid=13
 (retrieved 08/10/2007)
- CROSS, J. (2004) Foreword to *The Handbook of Blended Learning: Global perspectives, local designs*, by Curt Bonk and Charlie Graham.
<http://metatime.blogspot.com/#111778520060079866> (retrieved 26/09/06)
- DETECON (2002) e-Learning in developing countries. Opinion Paper. http://www1.worldbank.org/education/stuttgart_conference/download/6_2.pdf (retrieved 10/10/2007)
- DOWNES, S. (2006). Learning Networks and Connective Knowledge. Discussion Paper #92. Instructional Technology Forum.
<http://it.coe.uga.edu/itforum/paper92/paper92.html> (retrieved 08/10/2007)
- DEVLIN, M. ET AL. (2001) Introduction to the Internet and the University. Forum 2001. Epic’s Blended Learning Consultancy.
http://www.epic.co.uk/content/what_we_do/consultancy_index.htm (retrieved 08/10/2007).
- HARRIMAN, G. (2007) Blended Learning. e-Learning Resources.
http://www.grayharriman.com/blended_learning.htm (retrieved 01/10/2007)
- IAHAD, N. (2004) e-Learning in Developing Countries: Suggesting a Methodology for Enabling Computer-Aided Assessment. Proceedings of the IEEE International Conference on Advanced Learning Technologies (ICALT’04). Vol. 00, pp. 983 – 987. <http://portal.acm.org/citation.cfm?id=1018423.1020190&coll=&dl=&CFID=15151515&CFTOKEN=6184618> (retrieved 08/10/2007).

- LARSON, R. and WASSERSTEIN, J. (2003) LINC - Learning International Network Consortium. A new distance learning initiative in education in developed and developing countries.
- MITCHELL, L. (2001) New training methods offer personalized e-learning. InfoWorld. <http://www.itworld.com/Man/3889/IWD010416tcelearning/> (retrieved 01/10/2007).
- NAKABUGO, M. et al. (2007) Large Class Teaching in Uganda: Implications for E-Learning. <http://www.education.up.ac.za/de/eng/deta/presentations/word/Large%20Class%20Teaching%20in%20Uganda%20Implications%20for%20E-Learning%20Nakbugo,%20Hurd,%20Ezati.pdf> (retrieved 08/10/2007).
- NASHUA, N. (2003) SkillSoft seminar series expands on blended learning: focuses on best practices for structuring successful programs and proving results. Free Seminars in 40 U.S. and Canadian Cities Incorporates Research and Case Studies Supporting Value of Formal and Informal Learning. http://www.skillsoft.com/about/press_room/press_releases/sept_24_03_seminar.asp (retrieved 08/10/2007).
- SEHRT, M. (2005) e-Learning in the developing countries: digital divide into digital opportunities. UN Chronicle Online Edition. <http://www.un.org/Pubs/chronicle/2003/issue4/0403p45.asp> (retrieved 08/10/2007).
- SINGH, H. AND REED, C. (2001) A white paper: achieving success with blended learning. 2001 ASTD State of the Industry Report, American Society for Training & Development. <http://www.centra.com/download/whitepapers/blendedlearning.pdf> (retrieved 03/10/2007).
- SNIPES, J. (2005) Blended Learning: Reinforcing Results. http://clomedia.com/content/templates/clo_article.asp?articleid=1070&zoneid=25 (retrieved 08/10/2007).
- UNCST (2002) Status of ICT in Uganda: Scan-ICT preliminary baseline study Uganda country report. www.uneca.org/aisi/ScanGhana/Documents/3.%20SCAN%20ICT%20UGANDA.pdf (retrieved 10/10/2007).
- WASUKIRA, E. AND NAIGAMBI, W. (2002) Report on the Usage of ICT in Local Governments in Uganda. <http://unpan1.un.org/intradoc/groups/public/documents/CAFRAD/UNPAN006316.pdf> (retrieved 10/10/2007).

Use of “Flight Simulator” for Training IT Managers in Investment Decision Making

Paul Ssemaluulu and Ddembe Williams

In this paper it has been suggested that the findings of research on managerial decision-making may be relevant in IT, but that research needs to be carried out on the factors that influence the decisions of IT managers in order to develop an effective “flight Simulator”. While this is useful as a first step, the paper argues that a systems thinking/ systems dynamics approach is necessary for developing flight simulators as learning environments for managers. The advantage of SD is not merely that it captures the complexity of decision-making processes; it also offers a way of exploring the impact of change on decision-making over time. The paper describes the model, identifies a number of propositions that can be derived from it and suggests how these may be tested empirically through a flight simulator.

1. Introduction

Improving the information Technology (IT) investment evaluation process has become an important area of research and professional practice [Young 2002; Cohen et al. 2000]. Investment evaluation is a complex and dynamic decision making process as there are many interacting factors that impact on the choices made by the decision maker. There is some evidence to suggest that the introduction of system dynamics-based flight simulator as learning environment can be effectively used to promote understanding and improve learning of complex application domains in which management decision making take place. Improving managerial learning process effectiveness is arguably one of the most important aspects of research and development in education [Scriven 1960; Erwin 1991; Levin 2000; Jin 2002]. Despite technological benefits of the emerging tools to facilitate and leverage a broad range of cognitive learning processes [Pfahl et al. 2001], and the significant progress made in educational research [Levin 2000], seldom has flight simulators been used as a form of computer based instruction [Spector et al. 1998] which is claimed to foster effective learning, and enhanced problem-solving to managers [Young 2002].

In research and development education, Brown [1992] further extended in Levin [2000] calls for contemporary educational intervention that synthesises discourse of design experiments and other classroom-based studies. Levin [2000] further argues “*demonstrations studies and design experiments have their place in the developmental stages of educational intervention research*” [pp 201]. The advances

in instructional technologies and explosion of simulation tools are changing how managers work and make decisions on the day-to-day basis. Current improvements in the world economy, both industry and business are searching ways to improve their competitiveness, hence the need for experienced and well-trained decision support analysts. Various researchers have claimed that advanced technologies, like system dynamics simulation tools [Forrester 1961; Fisher and Zahara 1997; Sterman 2000; Spector et al. 2001] might fundamentally change how the learning process occurs [Draper and Swanson 1990; Isaac and Senge 1992].

In the paper a systems thinking/systems dynamics-based flight simulator approach is used to model the impact of change on IT decision-making. The paper further examines how this flight simulator could be used to improve decision-making and proposes a programme of future research.

1.1. Research Issues

Business and industry environments in which managers make decision are constantly changing, there by increasing demand for technology to increase productivity to support effective and timely decisions. This has led to developing decision support systems one of the most difficult and challenging tasks in many organizations. Higher education is not immune to these developments. In response to the demand for decision support analysts, higher education have designed curriculum for learning and teaching decision sciences [Williams 2002]. System dynamics-based flight simulators can be viewed as opportunities for postgraduate decision support education to leverage learning effectiveness [Dunlap and Gabringer 1996]. The focus of classroom experiments is the simulation of a specific decision making process and to learn from it.

The use of simulation methods in classroom settings is growing, however, Cohen et al. [2000] emphasise that "there is an increasing need to evaluate claims concerning the advantages and effectiveness of these newer approaches against more traditional methods" [pg 79]. The advantage of the use of a flight simulator in decision support systems is that the subject's humanity is left intact in that students are given a realistic situation in which to act in whatever way they think appropriate. Another advantage of the simulation approach, is the inclusion of time dimension in flight simulators that allow the students to take an active role in interacting with the learning environment and the experimenter the opportunity of observing a social process in action with its feedback loops, multidimensional causal connections that may help in improving our understanding of the effectiveness of flight simulator learning environment [Palys 1978]. Feedback from "flight simulators" can stimulate changes in the students' mental models about managerial decision support. Such learning involves new understanding, reframing of a decision situation and leads to new goals and new decision rules, not just new decisions.

These issues raise the following research questions (RQ) that will be investigated:

- RQ1: How sensitive to the characteristics of the learning environment are decision rules that managers use?
- RQ2: How can these flight simulator-learning processes be captured in models?
- RQ3: What types of feedback experiences and training might mitigate learning environment and develop our systems thinking capabilities in decision making context?
- RQ4: How can IT Managers improve their decision making process?

2. Related Literature

Flight Simulator

In investment evaluation, design experiments have been used [Levin 2000; Brown 1992] as experimental research strategy in classroom intervention [Cohen et al. 2000], however system dynamics based flight simulator as learning environment have not been used before. Sterman [2000] calls for the creation of managerial practice fields “flight simulator”. The underlying concept of rehearsals and practices can be central to successful team performance in training of aircraft pilots. Although flight simulator environments have focused on military aspects of planning and training, however they have not been used in higher education setting to teach and training managers learning decision support systems [Cohen et al. 2000]. In the System Dynamics (SD) literature, researchers have proposed the use of either case study or purely problem simulation. Researchers have not stated the advantages of combining the two research methods neither have the relative merits of quantitative (simulation modelling and qualitative (case study) been examined and debated [Williams 2000].

IT Managers

The term ‘IT Manager’ is used throughout the paper to describe an occupational group whose members ‘plan and oversee multiple projects and project managers and work with CIO’s and senior managers to determine systems development strategy and standards’ [Juliussen and Juliussen 1996]. The term ‘IS manager’ is sometimes used instead of, or in addition to, IT manager in the IT/management literature. The difference is one of emphasis. IS managers are more concerned with the management of information; IT managers with the management of the technical resources used to gather, process and disseminate information to different levels of the organisation [Ward et al. 1990]. This is justified on grounds that knowledge of the technical aspects of systems management is vital if organisations are to keep pace with the rapid rate of technological change and make sound decisions about IT investments.

Why are IT Managers Different?

- i. The rapid pace of technological change and the demands imposed on IT managers make the decision-making environment more fraught and complex than for many managers. The high turnover rates at the senior levels of the IT organization and the widespread criticism of IT managers' ability to handle management and business issues with which they have been confronted suggests that organizational change may place enormous strain on decision-making competence [Flynn and Williams 2000]
- ii. Psychological factors that cause managers to convince themselves that things do not look so bad and that continuation will eventually lead to success [Brockner 1992]. These factors include the manager's previous experience with similar projects, the degree to which the manager feels personally responsible for the outcome of the project, as well as psychological and cognitive biases that can affect the way in which information concerning the project is perceived or proposed [Keil 1995 Pg: 423].
- iii. Another type of bias can lead to "throwing good money after bad" in an effort to turn around a failing project (Garland 1990)

Investment Decision Making

The decisions that managers make regarding IS investment most times impact on the cost and eventual success of the IS systems that are put in place. In addition to this, the eventual cost of the IS investment may also depend on the decisions that these managers make. When making these decisions, it is important for them to be well informed with quality information, so as to come out with informed decisions. Accampo [1989] as quoted by Murphy and Simon [2002] states that "quantitative techniques can be hard to apply to activities in which information is the key commodity". Murphy and Simon [2002] contend that "many of the measures found in the IS literature that are used to evaluate system success are intangible" and that "traditional methods of project evaluation fall short if these measures cannot be quantified in monetary terms". In their work on system success, Delone and McLean [1992] have observed that System quality and information quality are related to system development; and system use and user satisfaction are relevant to implementation. In this study, system use and user/customer satisfaction reflect expectations of the customer. Delone and McLean [2002] state that information quality and system quality, are the most important quality components to measure the success of an IS in an organisation. On the other hand, Reicks [200]) states that "most people want access to the right information, as they recognize that sharing the right information with the right people at the right time, can empower these individuals". This helps them make the right decisions. Khalifa and Liu [2004] conclude that perceived benefits (IS success) are measured by expectations, ease of use and perceived usefulness.

Several authors have highlighted problems associated with IS investment which makes it unique. Some of these are:

Complexity of IS projects- It has been found that Information systems themselves are complex social objects inseparable from the organisational context within which they are situated and the infrastructure supporting them, and are products of history and human agency [Symons 1994]. This means that they affect the people who use them and are in addition affected by these very people. Many IS projects are designed to improve the operation of business activities that are dynamic, complex, non-linear systems which cannot be readily understood by using static modelling approaches. The dynamic systems are characterised by interactions of closed chains (or feedback loops) that, when combined, define the structure of the system and hence how it behaves over time [Kennedy 2001].

This affects correctness of output and makes it difficult to estimate the exact expenditures and therefore benefits [Marquez and Blanchar 2004].

IS investments are dominated by multiple view points which are not captured using current evaluation methods [Marquez and Blanchar 2004].

The role of soft factors such as motivation and perceptions has also not been appreciated in IS investment literature [Caulfield and Maj 2002]. If you omit soft variables you run the risk of failing to capture something essential to driving human affair

Poor information quality impairs decision making and promotes inefficiency. One reason for this is that managers depend on this data to order for equipment, select suppliers and contractors.

3. Factors Affecting It Managers Decision Making

The existing body of literature identifies factors that affect IT managers' decision making. Also the literature point to the significant differences the way managers make decisions. It is now possible to turn to the factors identified in the management literature as important in decision-making and consider how these may influence IT managers.

Organisational factors are likely to have a significant influence on the decision-making behaviour of IT managers. It could be argued that IT managers employed in organisations that depend on IT are likely to take decisions about IT that differ in nature and complexity to those of managers working in organisations that are less IT intensive or have a culture that is not supportive of IT or technical innovation. IT managers in organisations that are dependent on IT may play a more significant role in strategic decision-making and enjoy more influence in the decision-making process at all levels than IT managers in organisations that regard IT as a backroom activity. Although this may suggest that IT managers are in an ascendant position, the trend towards devolving responsibility for IT to the business areas and increasing user responsibility for IT and control over the IT process may undermine this trend. Whatever the precise status of IT managers, the greater involvement of users and increasingly complex IT structures

in organisations have almost certainly increased the complexity of the decision-making process. The ability of the organisation to learn and to ensure that learning is passed to organisational members may be another organisational factor that influences the decision-making behaviour of IT managers. Clearly, if the organisation has created mechanisms that distil learning from past events, the IT manager will be able to utilise this in the decision-making process.

Another key factor that influences decision-making is change in the external environment [Cooke and Slack 1984]. In recent years, major changes in the economic and political environment in which firms operate, the introduction of new "enabling" technologies, changes in social attitudes and demographic patterns combined with the general growth of competitiveness have resulted in dramatic changes in the strategic direction of many organizations [Doherty and Nyham 1996]. It could be argued of course, that the factors that influence IT managers are the same as those that influence other managers. There is some evidence to show that this is not the case. Studies of personality differences between IT professionals and general managers, for example, suggest that they are more intuitive and analytical and are less influenced by emotion than general managers [Lyons 1985]. This may mean that they are less swayed by sentiment or practical considerations when making decisions. An understanding of these differences is clearly important if decision-making in IT is to be improved, particularly since IT professionals are under increasing pressure to "hybridise" and become knowledgeable about the business [Earl 1989; British computer Society 1990; Earl and Skyme 1990; Jones 1993; Flynn and Williams 2000].

The types of decisions taken by managers obviously reflect the nature and purposes of the business [Cooke and Slack, 1984]. The structure of an organization may influence both the ease and effectiveness of decision making.

Decisions regarding the strategic direction of the organization influence the business rules that guide decision-making at lower levels of the organization. Thus if senior managers change strategic direction, decision-making at every level of the organization will be affected [Flynn and Williams 2000]. The speed with which change in strategy can be implemented is dependent on management information about change and the availability of systems to disseminate that information. It seems reasonable to assume that managers who have access to high quality information and IT used to support decision-making are likely to make more competent decisions than those who do not have such facilities [Flynn and Williams 2000].

One of the ways in which organisations can store and distil learning about important events is through the use of IS/IT. A factor that may have a very important bearing on the decisions IT managers make may therefore be access to information and IT support. IT managers who lack access to information are likely to make ill-informed and poor decisions. The fact that IT managers are frequently involved in the design, development and implementation of Management Information Systems (MIS) and Decision Support Systems (DSS) presumably

means that they require an understanding of the decision-making process and the factors that affect decision-making. It is ironic that so little is known about the decision-making behaviour of the managers responsible for delivering the systems that help other managers make decisions on which the survival and well being of the organisation may rest.

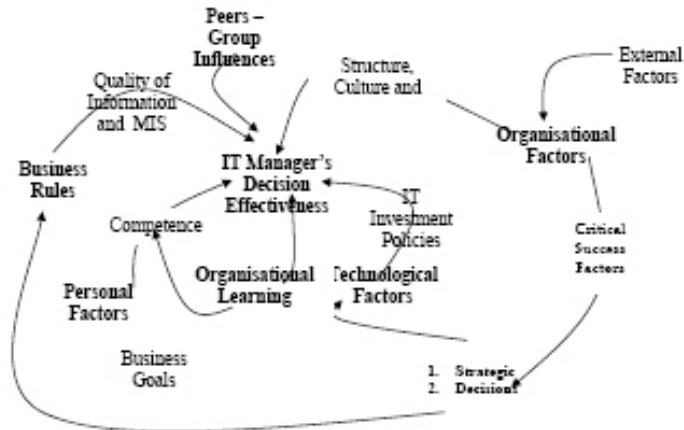
The idea that the IT department should be a centre of innovation suggests that IT managers are at the centre of change processes in organisations. Indeed they are frequently referred to in the literature as 'agents' or managers' of change.[Earl and Skyrme 1990; BCS 1990]. The question of how change may affect IT managers and how it might influence decision-making, however, has not been the subject of detailed investigation. Is there any reason to suppose that they behave differently from other managers? It could be argued that IT managers are more accustomed to dealing with change and are therefore more likely to deal with it effectively. This may be true of very experienced managers and managers in organisations that have been established long enough to have accumulated experience in dealing with change. However, the rapid pace of technological change and the demands imposed on IT managers make the decision-making environment more fraught and complex than for many other managers. The high turnover rates at the senior levels of the IT organisation and the widespread criticism of IT manager's ability to handle the management and business issues with which they have recently been confronted suggests that organisational change may place enormous strains on decision-making competence.

4. Model of the Factors that Influence Decision-making

The model over page (see Figure 1) provides a starting point for understanding the range of factors that influence decision-making in IT. At the centre of the model is the IT manager. The decisions he/she makes is depicted as being influenced by the factors described in the previous section. Personal factors are shown to be important as are the influence of peers and the groups to which the manager belongs. Personal and peer/group factors are subject to the influence of organisational factors - the structure, culture and political ethos of the organisation, its goals, quality of information available, etc. "Business rules" are shaped by the organisational context and strategic objectives, while investment policies are influenced by technological factors and are also shown to have an impact on the IT decision-maker.

The model also suggests that organisational learning may have a direct impact on the individual IT manager's ability to learn and to make effective decisions.

Fig. 1: Factors that Influence Decision-Making of IT Managers (based on Williams & Flynn, 2000)



The Model is useful insofar as it depicts the main factors that influence the decision-making of IT managers and indicates some of the possible interrelationships between them. However, it does not depict the process, delays, complexity or dynamic nature of decision-making in IT. If decision-making competence is to be improved, it is necessary to go beyond merely listing factors that influence decision-making and consider the dynamic nature of the decision-making process. Pidd [1996] contends that if organisations are to deal with the complexity of systems and decision-making within systems, there is a need for a new way of thinking about decision-making. Systems thinking/systems dynamics offers a vehicle for conceptualising the dynamics of the decision-making process [Morecroft and Senge 1990; Sterman 1994]. The next section explains the nature of the systems thinking/systems dynamics approach and highlights its potential in understanding the decision-making behaviour of IT managers.

- i) It has been shown that IS managers frequently lack a full understanding of their organizations business and are often not involved in the senior management decision-making of the company [Andresen 2000]. Senior management who do understand the business and have to make the decisions are usually not comfortable with the emerging information technologies. [Andresen 2000].
- ii) When considering new IS investments, senior management seldom have feedback from previous investments to provide comfort for their earlier decisions. [Ballantine and Stray 1998].
- iii) Complexity. Although significant numbers of IS projects are routinely completed successfully, a recent study on the state of IS in the UK carried out by Oxford University and Computer weekly reported that a mere 16 % of IS projects were considered successful [Sauer and Cuthbertson 2003].

- iv) Visualisation. IS project outcomes are effectively invisible. This visualisation problem is a source of many IS project failures.
- v) Traditional analysis focuses on the separation of individual pieces of what is being studied. Systems thinking, in contrast, focuses on how the thing being studied interacts with the other constituents of the system—a set of elements that interact to produce behavior of which it is a part. This results in sometimes strikingly different conclusions than those generated by traditional forms of analysis, especially when what is being studied is dynamically complex or has a great deal of feedback; like IS investment projects have.
- vi) *Management Support*- Top Management support refers to the senior executives favorable attitude toward and explicit support for the IS [Doll and Torkzadeh 1991]. When top management is highly supportive of IS, greater resources are likely to be allocated to develop and support information systems [Sabherwal et al. 2006]. This improves the decision making capabilities of IT managers.

From the diagram, by enabling the visualization of the factors at play in the investment process, complexity is reduced. In case complexity is increased, understanding would be impaired, but if reduced, understanding would be enhanced. With increased understanding, there is increased management support which leads effective management decision making. This now calls for newer approaches in Investment decision making.

Integrating emerging technologies to design flight simulator-like learning experiences is of high pedagogical value and andrological. The technological and pedagogical implications of designing such state-of-the-art learning environment can extend the pedagogical objectives to include concepts and dimensions that are usually difficult to communicate [Pfahl et al. 2000; Wang and Newlin 2002; Jin 2002], if not impossible for a manager that has not had an exposure of mathematics or computing, to reach through traditional educational methods. The flight simulators developed should not be viewed as an answer to the problems, but rather as a vehicle for exploring many of the problems reported in the literature [Levin 2000]. Several tentative propositions can now be drawn from the literature and suggests how they may be tested.

- P1: There is a time lag between the a IT manager gaining understanding of the decision environment technical potential and the manager's understanding of their own information requirements and choices at the time of decision making process.

The flight simulators will make a useful starting point for developing a theory of the manager's learning decision making process and managerial decision support training.

- P2: The quality and availability of information has a major impact on the IT Manager's capacity to respond to choices and satisfaction with the decision-making process.

P3: IT Managers, who have access to high quality information, and use flight simulators to support decision-making, are likely to make more effective management decisions than those who do not have flight simulators at the time of decision making process.

Managers in Information Systems are faced with increasingly dynamic, complex, and uncertain environment in which to make decisions. Modern Science does not reflect on all reality, but only on that part that is ordered, isolatable, predictable and controllable [Heylingen 1997; Cilliers 2005; Gershenson and Heylingen 2005]. A different set of tools is required to navigate this increasingly complex environment [Gershenson, 2007]. Rapid technological advances, information explosion, and the widening gap between the developed and underdeveloped countries of the world all contribute to today's complex environment [Daellenbach 1994].

Fundamentally, a system is complex when we cannot understand it through simple cause-and-effect relationships or other standard methods of analysis. Every year there are more people, more computers, more devices, more cars, more medicines, more regulations, more problems [Gershenson and Heylingen 200]). This results into more complex information systems.

A distinguished researcher in role-playing, gaming and computer simulation, van Ments [1978], proposed a design process for using students and suggests that, in educational setting, one may use simulation as situation from which others may learn. The underlying system dynamics model's most distinctive feature is that it can deal with mathematical concepts and user-friendly technology that support decision making process. As suggested previously, the inadequacies of current decision support models make it difficult to capture the complex relationships and feedback loops that characterise the decision-making processes. This "fly by wire" concept of the learning process has been used in organisations to facilitate learning [Senge 1994] but it has not been used as a basis for developing manager' decision-making effectiveness. The research will apply Dynamic Synthesis Methodology, grounded on well-tested and developed theoretical anchors and builds on an existing epistemological of science in the acquisition of knowledge (Williams 2001), as a basis for theory building and extension in the areas of problem-solving field.

5. Flight Simulator for Investment: A Case Study

The use of dynamic synthesis methodology to develop a "decision explorer" as flight simulator learning environment will make significant contributions to the literature and to knowledge as to how managers learning difficult concepts like those common in investment evaluation process. The concept of decision explorer has not been explored in the literature nor has it been applied in educational or training context. The results of such a study have also significant implication for student life long learning process [Yin, 2002] and manager who may wish to learn or train using the decision explorer.

One advantage of using system dynamics models over more traditional models is the inclusion of feedback effects. This makes the model's output more realistic

and exposes complexity that may be hidden in other modeling techniques. Thus, decision makers can more easily identify the interrelationships in the environment of the model. Feedback occurs in another level as well. As managers practice with the simulation, their understanding of the environment changes. They learn from the feedback the model gives them.

Sterman emphasizes thus “Learning is a feedback process. We make decisions that alter the real world; we receive information feedback about the real world and using the new information, we revise our understanding of the world and the decisions we make to bring the state of the system closer to our goals” [Sterman, 1994 Pg 291].

Building a Flight Simulator

Building a flight simulator is accomplished in four steps, that is; charting the trends or patterns of behavior; finding the variables; using system thinking tools to build causal loop diagrams and finally using stock and flow diagrams to build a dynamic model. Simulations are run on the model to test policies and decision rules. When the results are satisfactory, an interface is added to allow non-technical users interaction with the flight simulator. By doing this we are empowering the managers and promoting team learning.

Reference Modes

In System Dynamics research, reference modes constitute perceived qualitative mental models about key variables that are problematic and central to investigation. They would essentially constitute what the manager perceives as the behavior of the phenomenon over time. These are plotted and we then move on to the next stage of identifying the key variables.

Identifying Key Variables

In identifying the key variables, we are trying to establish the system’s dynamics. This reveals the factors that make it change over time. When managers are exposed to this kind of reasoning through participation, they soon learn to identify the most important issues instead of moving around in circles.

Building a Causal Loop Diagram

According to Maani and Cavana [2003, Page 28], “a causal loop is a conceptual tool which reveals a dynamic process in which the chain effect(s) of a cause is/are traced, through a set of related variables, back to the original cause (effect). Once a group of variables has been linked together in a connected path, this forms a causal loop”

Building a Dynamic Model

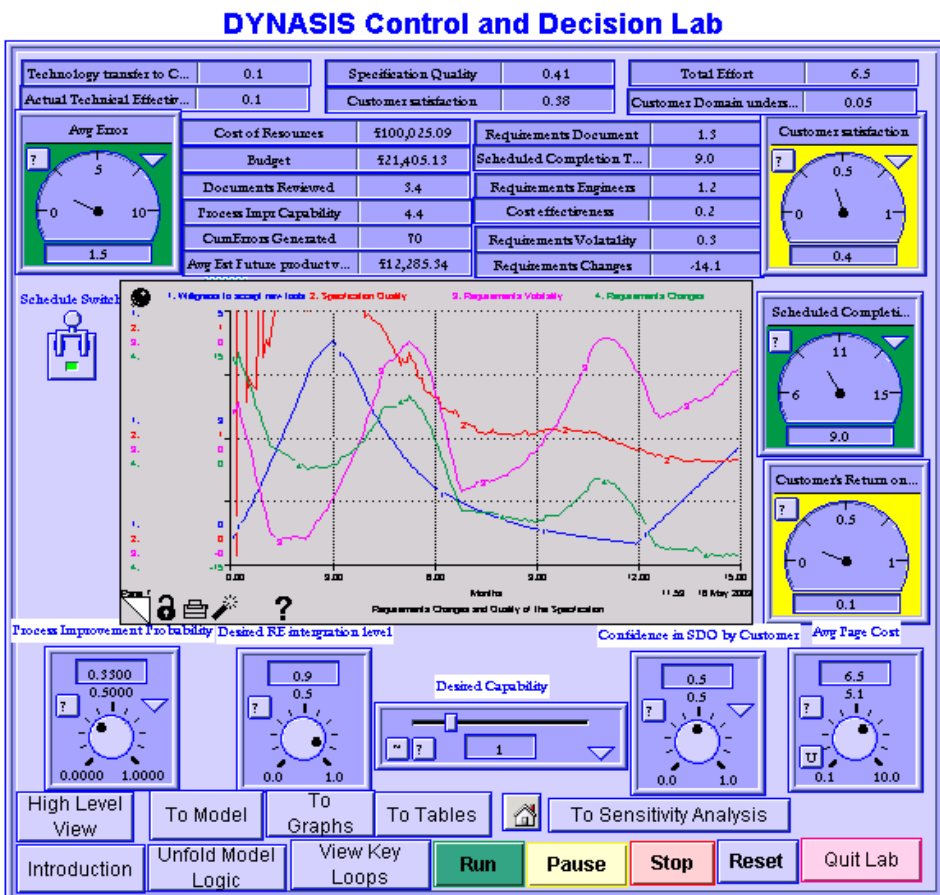
This is done by converting the causal loop diagram into stock and flow diagrams, which is a formal quantitative model of the problem in question. The mathematical relationship between and among variables is then defined.

Design of a graphical user Interface

Designing an intuitive interface for IT managerial decision making requires significant research effort. For example in an earlier papers Williams, [2000]; [2001a/b], and [2002] proposed and developed a prototype research flight simulator resulting from the application of Dynamic Synthesis Methodology to the modelling and analysis of requirements engineering process. This RE process system dynamics model, whose interface is illustrated in Figure 2 below predicts the performance of the RE process in terms of cost, quality of specifications and schedule of the RE projects.

As illustrated in Figure 2, a requirements engineer or manager who uses DYNASIS can evaluate the completeness, consistency and accuracy of the model in predicting RE process performance against known performance measures as well as gain requirements engineering planning and control expertise through training with tool [Williams, Hall and Kennedy, 2000; Williams, 2001a/b]. Such a prototype can be a basis for developing flight simulators for IT managerial decision making

Fig. 2: Main Interface for DYNASIS Tool



6. Conclusion and Future Work

The paper indicates that the high turnover rates at the senior levels of the IT organization the managers' ability to handle both management and business issues are a big strain on IT managers. It has been suggested that the findings of research on managerial decision-making may be relevant in IT but that research needs to be carried out on the factors that influence the decisions of IT managers in order to develop an effective "flight Simulator". The paper argues that a systems thinking/ systems dynamics approach is necessary for developing flight simulators as learning environment for managers. The advantage of SD is not merely that it captures the complexity of decision-making processes; it also offers a way of exploring the impact of change on decision-making over time. The paper describes the model, identifies a number of propositions that can be derived from it and suggests how these may be tested empirically through a flight simulator.

The practical applications of the model are promising for both theorists and practising managers. The direction of future research will focus on validating the model and to develop a range of flight simulators for IT and other managers to improve their decision-making processes and decision-making effectiveness.

References

- ANDRESEN, J. (2000). A Framework for Measuring IT Innovation Benefits. *Doctoral Dissertation*. Technical University of Denmark, Copenhagen
- BALLANTINE, J. and Stray, S. (1998). Financial Appraisal and the IS/IT Investment Decision Making Process. *Journal of Information Technology* 13, 3-14.
- BCS. (1990). The BCS (British Computer Society) Report 'Hybrid Managers - From Potential to Reality'.
- BROWN, A. AND DOWLING, P (1998) Doing Research/ Reading Research: A Mode of Integration in Education. Master Classes in education Series, Oxford
- BROWN, A. L (1992) Design Experiments: Theoretical and methodological challenges in creating complex interventions in classroom setting. *Journal of the Learning Sciences*, 2, 141-178.
- CAULFIELD, C.G., AND MAJ, S.P. (2002). A Case for System Dynamics. *Global Journal of Engineering Education*. 6: (1)79
- COHEN, L. AND MANION, L. (1994) Research methods in Education, fourth edition, London: Routledge.
- COHEN, L, MANION, L AND MORRISON, K (2000) Research Methods in Education, 5th Edition, Routedledge- Falmer London.
- CILLIERS, P. (2005). Complexity, Deconstruction and Relativism. *Theory, Culture and Society*. 22, 255-267.

- COOKE, S. AND SLACK, N. (1984). Making Management Decisions. Prentice Hall.
- COUGAR, J.D. AND ZAWACKI, R.A. (1980). *Motivating and managing computer personnel*. New York: John Wiley.
- DAELLENBACH, H.G. (1994). *Systems and Decision Making: Management Science Approach*. John Wiley and Sons. New York.
- DELONE, W.H., AND MCLEAN E.R. (1992). Information Systems Success: The Quest for the Dependent Variable . *Information Systems Research*. **3(1)**, 60-95.
- DELONE, W.H., AND MCLEAN E.R. (2002). Information Systems Success Revisited. *Proceedings of the 35th Hawaii International Conference on System Sciences*,. **Jan 7-10**.
- DOLL, W.J. AND TORKZADEH, G. (1991). The Measurement of End-User Computing Satisfaction: Theoretical and Methodological Issues. *MIS Quarterly*. **15(1)**, 5-10.
- DRAPER, F AND SWANSON, M, (1990) Learner-directed systems education: A successful example, *System Dynamics Review*, Vol. 6 (2), 209-213.
- DUNLAP J. C., & GABRINGER, R. S. (1996). *Rich environments for Active learning in the higher education*. In Wilson B. G. (Eds.), *Constructivist Learning Environments: Case studies in Instructional Design* (pp. 151-174). New Jersey: Educational Technology Publications.
- EARL, M.J. AND SKYRME, D.J. (1990). Hybrid Managers: What should you do? *Computer Bulletin*:19-21. May.
- ERWIN, T (1991) *Assessing Student Learning and Development*. San Fransisco, Jossey-Bass Publishers
- FISHER, D. and Zahara, R (1997) Seamless Integration of Systems Dynamics into High School Mathematics: Algebra, Calculus, Modelling Courses. A Systems Approach to learning and Education into 21st Century, Proc. Of International Systems Dynamics Conference, System Dynamics Society Istanbul pp. 825-829.
- FLYNN, V. AND WILLIAMS, D. (2000). Improving the Decision-Making Behavior of IT Managers: A Competence Approach. In Eds. Ron Sanchez and Eine heene, *Advances in Business Strategy*. 6(B), 67-83. Jai Press Inc. ISBN:0-7623-0676-9.
- FORRESTER (1961) *Industrial Dynamics*, Productivity Press.
- FRIEDMAN, A. (1989), *Computer Systems Development: History, Organisation and Implementation*, John Wiley & Sons, Chichester.
- GARLAND, H. (1990). Throwing Good Money After Bad: The Effect of Sunk Costs on the Decision to Escalate Commitment to an Ongoing Project. *Journal of Applied Psychology*. **75(6)**, 721-727.
- GERSHENSON, C. (2007). *Design and Control of Self-organizing Systems*. PhD Dissertation- Vrije University. Brussels-Belgium.
- GERSHENSON, C. AND HEYLIGHEN, F. (2005). How Can we Think the Complex? Managing the Complex. **1**, 47-62
- HEYLIGHEN, F. (1997). Publications on Complex, Evolving Systems. A Citation-Based Survey. *Complexity*. **2(5)**, 31-36.
- ISAACS W., SENGE P. (1992), *Overcoming limits to learning in computer-based learning*

- environments, *European Journal of Operational Research*, **59**, 183-196.
- JIN, QUN (2002) Design of a virtual community based interactive learning environment, *International Journal of Information Sciences*, 140, 171-191.
- KEIL, M. (1995). Pulling the Plug: Software project Management and the problem of Project Escalation. *MIS Quarterly*. 19(4), 421-447.
- KENNEDY, M. (2001). The role of System Dynamics Models in improving the Information Systems Investment Appraisal in respect of Process Improvement Projects. *Proceedings of Nineteenth International System Dynamics Conference*. Atlanta, Georgia, USA.
- KHALIFA, M. AND LIU, V. (2004). State of research on Information system Satisfaction. *Journal of Information Technology Theory and Application*.
- LEVIN, J. R (1999) What to do about educational research's credibility Gaps? *Issues in Education*, 5 (2), 177-229.
- LYONS, M.L. (1985). The Data Processing Psyche. *Datamation*. 31, 103-110.
- MAANI, K.E. AND CAVANA, R.Y. (2003). *Systems Thinking and Modeling: Understanding Change and Complexity*. Prentice Hall, Auckland. New Zealand.
- MARQUEZ, A.C., AND BLANCHAR, C. (2004). A Decision support System for Evaluating Operations Investments in High-Technology Systems. *Decision Support Systems*. DESCUP-11036.
- MORECROFT, J.D. (1999) System Dynamics for Policy Strategy and Management Education. *Journal of the Operational Research Society Special Issue* . 50 (4).
- MURPHY, K.E. AND SIMON, S.J. (2002). Intangible Benefits Evaluation in ERP Projects. *Information Systems Journal*,12, 301-320.
- PFAHL, D., KLEM, M. AND RUHE, G. (2001) A CBT Module with Integrated Simulation Component for Software Project Management Education and Training, *Journal of Systems and Software* **59**, 283-298
- van MENTS M. (1978) Role-playing: playing a part or mirror to meaning? *Sagest Journal* 8(3), 83- 92.
- PALYS, T.S. (1978) Simulation methods and Social Psychology. *Journal of the Theory of Social Behaviour*, 8, 341- 68.
- REICKS, R. (2001). How to Ensure that Knowledge is Power. *The Journal of Business Strategy*. 22(4), 32-35.
- SABHERWAL, R.; JEYARAJ, A. AND CHOWA, C. (2006). Information System Success: Individual and Organizational Determinants. *Management Science*. 52(12), 1849-1864.
- SAUER, C. AND CUTHBERTSON, C. (2003). The State of IT Project Management in the UK. *Report of the Working Group of The Royal Academy of Engineering and The British Computer Society* .Templeton College, Oxford University. November.
- SCRIVEN, M (1960) The philosophy of science in educational research, *Review of Educational research*, 30, 422-429.
- SENGE, P. M. (1994). *The Fifth Discipline: The Art and Practice of the Learning Organization*.

New York: Doubleday/Currency.

- SPECTOR, J. M.; CHRISTENSEN, D. L; SIOUTINE, A. V AND MCCORMACK, D (2001) Models and simulations for learning in Complex domains: using causal loop diagrams for assessment and evaluations, *Computer in Human Behaviour Journal*, 17, 517-545
- STERMAN, J.D. (1994). Learning in and About Complex systems. *Systems Dynamics Review*. 10, 291-330.
- STERMAN, J (2000) *Business Dynamics: McGraw-Hill*
- SYMONS, V.J. (1994). *Evaluation of Information Systems Investments: Towards Multiple Perspectives*. Chapman and Hall, London. ISBN 0-412-41540-2.
- WANG, A AND NEWLIN, M. H (2002) Predictors of web-student performance: the role of self-efficacy and reasons for taking an on-line class, *Computers in Human Behavior* 18, 151-163
- WILLIAMS, D. (2001/a). Towards a Systems Engineering Process Methodology for Requirements Engineering Process Modeling. International Council on Systems Engineering (INCOSE), Symposium, Daventry, UK.
- WILLIAMS, D. (2001/b) An Application of System Dynamics to Requirements Engineering Process Modelling, Un Published PhD Thesis, School of Computing, Information Systems and Mathematics, South Bank University
- WILLIAMS, D. W (2000) *Dynamic Synthesis: A Methodology for Process Modelling and Analysis*, South Bank Technical Report
- WILLIAMS, D. W (2002) *Profiling System for Employability Curriculum Development for Decision Sciences Postgraduate Higher education*, South Bank Technical Report.
- WILLIAMS, D.W; HALL, T. AND KENNEDY (2000). A Framework for Improving the Requirements Engineering Process Management. *Software Quality Journal*. 8(2), 133-147. February.
- YIN, R. K, (1984) *Case Study Research: Design and Methods*, Sage Beverly Hills California.
- YIN, R.K. (2002). *Case Study Research, Design and Methods*. 3rd ed. Newbury Park, Sage Publications.
- ZAWACKI, R. (1992) Motivating the IS people of the future. *Information Systems Management*, 9(2), 73-75.

System Dynamics Modeling in Healthcare: The Ugandan Immunisation System

Agnes Semwanga Rwashana and Ddembe Wileese Williams

The paper develops a system dynamics simulation model to understand the dynamics of immunisation with the aim of aiding decision making process by proposing policies that would enhance immunisation utilization. The model is designed with the intent to show how particular variables influence immunisation demand and coverage rather than predict immunisation coverage rates. The paper applies system dynamics modeling and field study research methods to capture the complex and dynamic nature of the immunization process, to enhance the understanding of the immunization health care problems and to generate insights that may increase the immunization coverage effectiveness. The model is divided into four sectors interacting with one another. The paper suggests key leverage points which could substantially improve immunisation demand, effectiveness of the health system as well as vaccine management.

1. Introduction

Simulation modeling has become an important area for strategy implementation, policy analysis and design in many application areas including health care management. Some of the health care management systems include billing, administrative, staffing, patient management, hospital management, pharmacy, consulting systems, disease screening, and emergency care among others. Computer models are used extensively in many areas of systems management to provide insight into the working of a system. This is particularly useful when the system is complex and / or when experimentation is not possible. Simulation modeling has been applied in a number of health care situations [Bayer et al., 2007; Cody et al., 2007; Haslett, 2007; Chu et al., 2003; Fitzpatrick et al., 1993; Dexter et al., 1999; Dittus et al., 1996] and clearly has the potential to play a role in health care decision making at all levels. The availability of a variety of simulation software has significantly expanded the role of simulation modeling in healthcare research, policy making and operational decisions [Fone et al., 2002; McAleer et al., 1995].

Simulation models provide quantitative information that provides a clear understanding of the problem thus enhancing decision making. Computer simulation provides a powerful tool that can be used to model and understand complex systems from the macro systems level to the micro genetic level [Hethcote et al., 2004]. The availability of a variety of simulation software has significantly

expanded the role of simulation in research, policy making and operational decisions [Dexter *et al.*, 1999; Dittus *et al.*, 1996; Fitzpatrick *et al.*, 1993; McAleer *et al.*, 1995].

1.1. Background

Preventable childhood diseases such as measles, polio and premature deaths still occur particularly in the developing countries due to low immunisation coverage [WHO, 1999]. According to WHO/UNICEF [2006], the global coverage for measles stands at 77 per cent, 28 million infants worldwide had not been vaccinated with DTP3 in 2005 with 75% of these belonging to developing countries. In a study to evaluate new tendencies and strategies in international immunisation, Martin and Marshall [2002] suggest that “*failure to immunize the world’s children with life saving vaccines results in more than 3 million premature deaths annually*”. In resource constrained countries, despite numerous immunisation campaigns over media, health visits and improved health services, immunisation coverage are still low [WHO/UNICEF, 2006]. In Uganda, a nationwide survey showed that 46% of the children (12-23 months) had received all the recommended vaccines [UBOS, 2007].

Various approaches have been applied to understand immunisation coverage problems, however, there are still acknowledged deficiencies in these approaches and this has given rise to research efforts for alternative solutions including the need to adopt new technologies to address some of these problems. The immunisation system like many health care systems is very complex with many stakeholder (clients, doctors, nurses, managers, policy implementers) viewpoints that are difficult to manage. Due to increased pressure to reduce costs, increase efficiency and raise standards, health policy makers internationally have introduced a wide range of changes to healthcare in the quest for improved performance thus requiring System Dynamics techniques that have the ability to enhance the generation of knowledge in health services. Health care services like any other business involve a lot of transactions such as importation and delivery of medicines; construction of hospitals and clinics; hiring and deploying staff; processing and payments of staff salaries, purchase and delivery of food for patients; collecting and analyzing disease spread data, and educating the community [Wasukira *et al.*, 2003]. Some of the issues that are faced by the health services include deficiencies in service delivery, growing gaps in facility and equipment upkeep, inequity of access to basic health services by the communities, inefficient allocation of scarce resources and lack of coordination among key stakeholders [Fraser and McGrath, 2000].

System Dynamics is highly applicable to the dynamic simulation of the immunisation system since it has a number of key attributes involving strong feedback loops thus requiring the systematic approach. Creation of a complex and sufficiently rich description of the immunisation system helps to analyze and study the dynamics associated with policy implementation in immunisation coverage.

1.2. Research Aim and Scope

The authors initiated this study with aim of understanding the immunisation health care problems, generate insights that may increase the immunisation coverage effectiveness, develop and simulate a healthcare policy design model by applying the System Dynamics (SD) modelling approach. This research covered the activities, strategies and policies associated with immunisation coverage. Modelling techniques as well key activities employed in health care modeling processes were investigated. Explorations of the relationships, dynamics and processes of the current immunisation system were done to establish how these factors influence insufficient demand and poor immunisation coverage. Vaccine management, health care service provision and the programme management at both national, district and village levels were examined. Specific in-depth investigations of the issues pertaining to immunisation demand were carried out in Mukono district.

2. Research Methodology

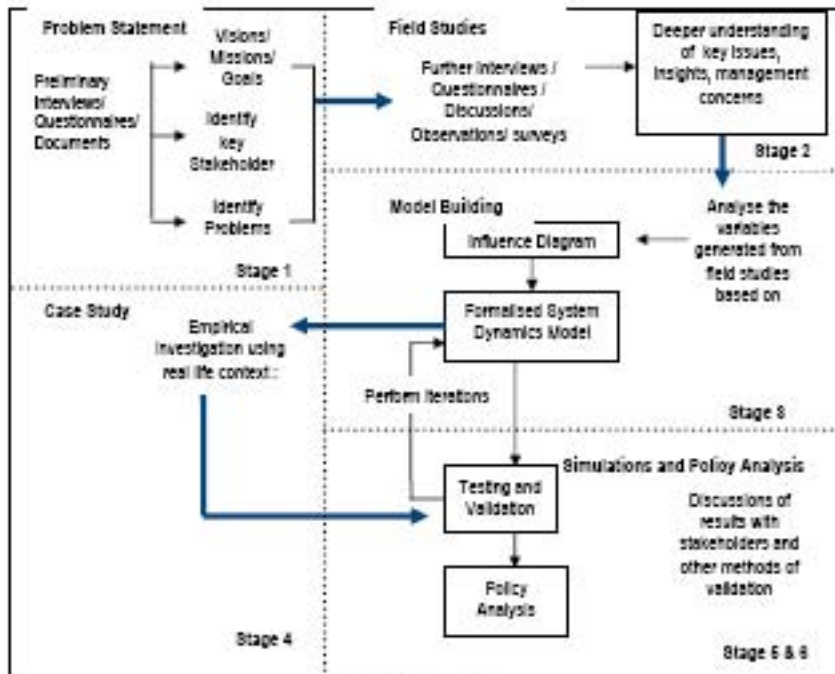
The study employed the Dynamic Synthesis Methodology (DSM) which combines two powerful research strategies; System Dynamics and case study research methods [Williams, 2000]. Combining simulation and case study methods as proposed by the Dynamic Synthesis Methodology is beneficial in that the strength of the case study enables the collection of data in its natural setting. Case study enables the collection of on-site information of the current system, owners and user requirements and specifications used to develop the generic model.

The System Dynamics methodology illuminates three principal effects which are systemic feedback loops, systemic delays and unintended consequences. The SD methodology is more suitable to solve the problems presented by the immunisation system as illustrated below:-

- Models benefit decision makers by identifying data that should be available but is not being collected.
- The immunisation system presents exogenous shocks (come from sources outside the system they affect) such as changes in demand for immunisation and the emergence of epidemics such as measles.
- The immunisation system contains feedback loops, communication paths and methods that impact behaviour such as the following illustration : immunisation knowledge enhances utilisation for immunisation services which in turn results in more knowledge.
- The immunisation system has systemic delays within its structure which develop over time in response to internal and external influences. Examples of such delays are those arising from delivery of health services and cold chain maintenance especially to the rural communities as well as delays in uptake of immunisation.
- The different policy changes, feedback loops and behavioural changes in the immunisation system result in either/both intended and unintended consequences which can be well investigated using the SD methodology.

The vast problems faced by the nation's immunisation system policy can be interpreted in terms of the information, actions and consequences which the system dynamics view point presents. Figure 1 presents the research design framework used to conduct this research. This framework consists of six stages and was developed based on the Dynamic Synthesis Methodology by Williams [2000].

Fig. 1: The Research Design



In the first stage, preliminary interviews were done to identify problems and key stakeholders. Field studies were used to determine the full range of activities and challenges associated with immunisation coverage. Factors affecting immunisation coverage as well as national immunisation policies used for immunisation coverage were critically analysed. The model was built based on the field study results which provide a descriptive model on which the SD conceptual feedback structure is developed. The feedback structure model was developed with the help of Causal Loop Diagrams. Preliminary causal loop diagrams generated from the field studies were presented to various stakeholders for their comments and feedback on understanding of the immunisation coverage problems. Causal loop diagrams were converted to stock and flow diagrams which form a quantitative model of the problem in question. Mathematical relationships between and among variables which enable the simulation of the model were defined after which simulations of the key variables were run.

3. The Immunisation Policy Analysis Model

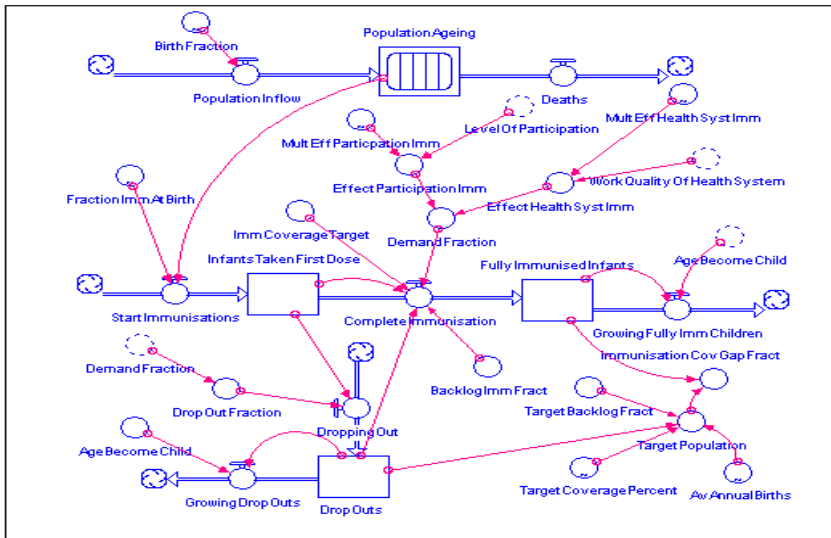
Immunisation coverage focuses on the supply and delivery of immunisation services, while immunisation demand is reflected through the parent’s acceptance and the child’s attendance according to the immunisation schedule. The immunisation simulation model was constructed so that it could be used to investigate deeper dynamic issues that are of concern to management and provide varying scenarios with different structures and policies. The immunisation model was built based on the analysed data that was collected through literature and field studies.

Demand is interplay of factors arising from the supply of immunisation services (healthcare factors) as well as population characteristics (socio-economic factors and background factors). The immunisation model was divided into the four sectors (population, immunisation demand, immunisation operations and vaccine stock management) and simulated over a 15-year period to evaluate potential policies development and debate in health care management.

3.1. The Population Sector

The population sector shows the dynamics that are involved in the various population categories (see Figure2). The population sector has a Population Ageing sub-model with four stocks which represent the four different age groups namely infants (below 1 year), children (above1to 14years), reproductive age group (15-49years) and adults above 50years. Infants (below1year) are particularly separated from the children because this is the age group that receives immunisation.

Fig. 2: The Population Sector



The population categories that are generated in this sector are *InfantsTakenFirstDose*(infants who have received the initial dose (at birth)), *FullyImmunisedInf*

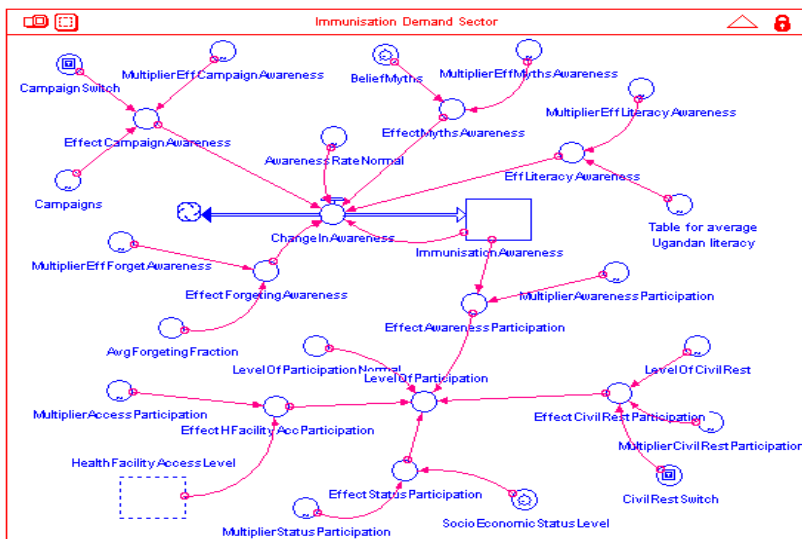
ants (fully immunised infants who have completed the immunisation schedule by taking all the required doses) and *Drop Outs* (infants who take initial doses but do not complete the immunisation schedule or take all the required doses of immunisation).

3.2. The Immunisation Demand Sector

The immunisation demand sector captures and models the dynamics that are associated with level of parents' participation in immunisation activities which affects the demand (see Figure 3). The output of the immunisation demand is the level of participation in immunisation activities and the key factors that were found to affect the level of participation are:

- Immunisation awareness which can be obtained through effective campaigns (media-radio, newspapers, films, television) coupled with high level of literacy provides knowledge concerning the importance of immunisation which increases the level of participation. Loss of immunisation awareness occurs after sometime and this can be accelerated by belief in myths.
- Accessibility to health facilities which makes it easier for the parents to take the children for immunisation without incurring extra costs.
- A high socio-economic status level increases the parents' availability thus enabling them to take off sometime and take the children for immunisation.
- Areas that have civil unrest reduce the parents' level of participation.
- Reminders sent to the parents of the immunisation schedule were found to tremendously increase the level of participation towards immunisation activities.

Fig. 3: The Immunisation Demand Sector

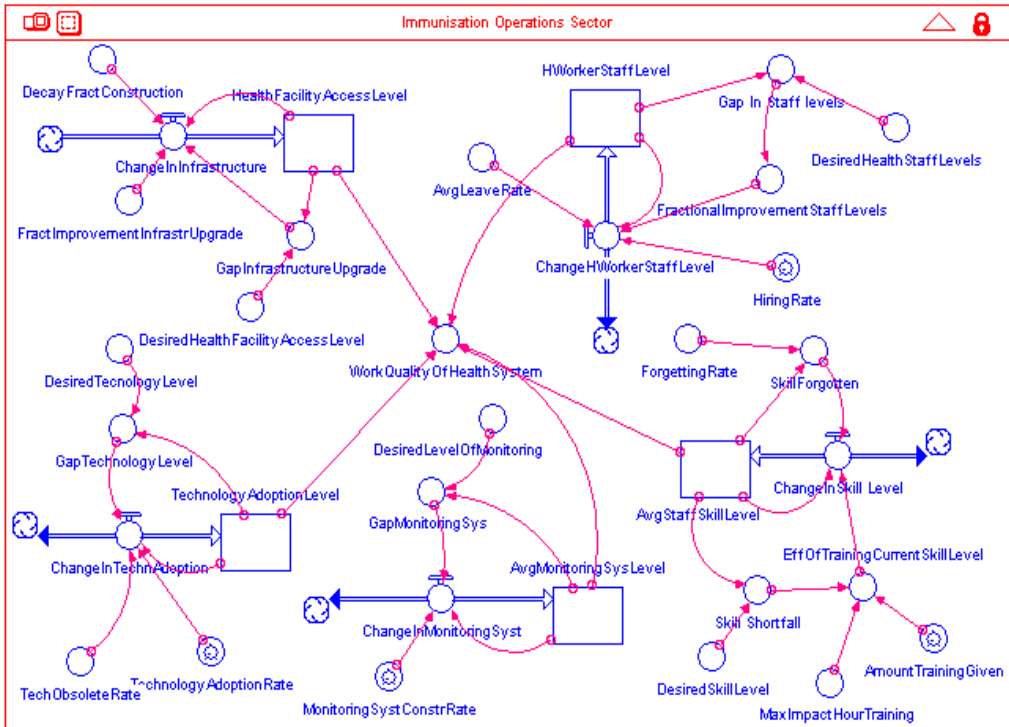


The above are soft variables which can not be quantified but are represented as level on scale ranging 0-1 with 1 as the highest level and 0 as the lowest level. The level of participation is an aggregate of the effects of awareness on participation (*EffectAwarenessParticipation*), civil rest (lack of civil wars) (*EffectCivilRestParticipation*) on participation, socio-economic status (*EffectStatusParticipation*), access to health facility (immunisation services) (*EffectHFacilityAccParticipation*). Each of these effects is derived by first assessing a multiplier effect for each of the variables. The multiplier effect are graph representations of the behaviour of the level of participation based on the changes of that variable.

3.3. The Immunisation Operations Sector

The effectiveness of health systems is vital to the immunisation coverage. The immunisation operations sector presents the dynamics involved in the provision of immunisation services (see Figure 4).

Fig.4: The Immunisation Operations Sector



Some of the key issues in the quality of health system include the following:

- Health facility infrastructure level (*HealthFacilityInfrustrLevel*) refers to the fraction of parishes that have hospitals, clinics or dispensaries where immunisation services are being provided. The construction of new health facilities is influenced by the gap between the actual quality of service and that which is desired.

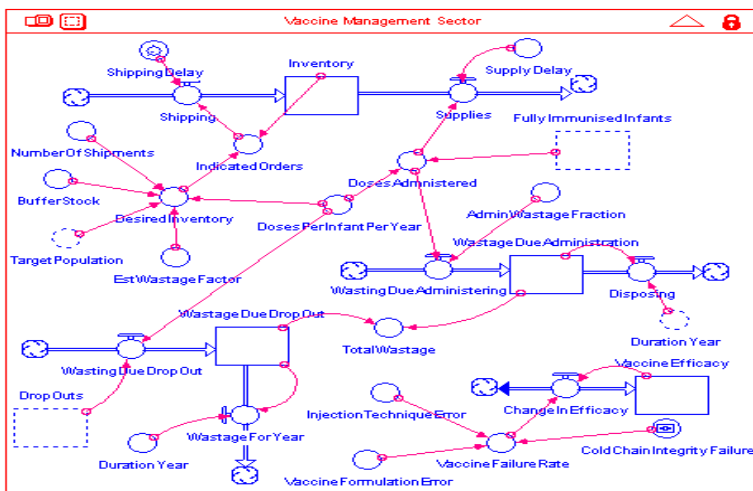
- Health worker staff level (*HWorkerStaffLevel*) refers to the fraction of the number of health workers (nurses) involved in the government immunisation programme compared out of the required level.
- Health worker skill level (*AverageStaffSkill*) refers to the level of staff skill in the various health centres. Change in skill levels can be brought about through the amount of training provided coupled with the impact of the training. Increased skill levels provided efficiency in the health centres which improves the quality of service.
- Technology adoption levels (*TechnologyAdoptionLevel*) refers to the fraction of health centres that has adopted the use of technology in their operations. Change in technology adoption is influence by various factors which this model does not take care of.
- Level of monitoring systems (*LevelMonitoringSys*) refers to the levels of monitoring of immunisation activities. The level of monitoring systems in health centres and hospitals greatly affects the quality of health system.

There are various factors such as payment of salaries and allowances for health workers which are left out of this model since these are done by the public service and not necessarily handled by the immunisation system.

3.4. Vaccine Management Sector

Vaccine management is the key to the success of immunisation programmes. There is need to monitor the management of vaccines as well as have them replenished regularly in order to prevent over and under stocking which may lead to expiry of vaccines and low coverage respectively. This sector represents the dynamics involved in the management of stocks with the aim of minimising wastage and avoiding stock out situations. Vaccine wastage is the proportion of vaccine supplied but not administered to the recipients (see Figure 5).

Fig.5: The Vaccine Management Sector



Barenzi *et al.* (2000) predicted the vaccine wastage rate for period 2000-2004 to be 60%. Factors associated with vaccine wastage are :-

- Administrative factors which arise during the delivery of vaccines. Vaccines are not fully utilised when being administered (e.g. only 16 or 17 doses are obtained from a 20 dose vial).
- Break down of cold chain management which reduces the potency of vaccines
- Stock management problems which result in expiry of vaccines resulting from short expiry dates provided by the manufacturers. Vaccines such as the measles vaccine with a shelf life of 5 months have to be distributed as soon as possible by following the Earliest Expiry First Out principle (EEFO).

Vaccines are administered to the infants to boost their immunity thus enabling them to have a higher survival probability even when they are infected. Vaccine efficacy is the strength of the vaccine and is vital in providing the immunity for a period of time. Change in vaccine efficacy is affected by the vaccine failure rate (*VaccineFailureRate*) is an aggregate of the factors below which are quantified by considering levels of 0-1 :-

- Injection technique errors arising due to lack of skill in administering the injection.
- Cold chain integrity errors arising due to poor maintenance of cold chain from the manufacture to the end user.
- Vaccine formulation errors arising due to poor formulation of the vaccine while mixing it.

The model was populated with data from various census and health reports of Uganda from the year 2001-2007.

3.5. Reference to other model structures

This section highlights some of the models that were referred to during model development. Subramanyam and Sekhar [1987] in a vaccine optimal distribution system, used inputs based on population, growth rates, birth rates, infant mortality and calculated the target population based on the growth rate alone and does not take into account the drop outs who return for missed doses which this new model considers. The developed model uses vaccine failure rates resulting from vaccine formulation, cold chain integrity and injection techniques similar to those used by Tobias and Roberts [2000] in a prediction and prevention of Measles Epidemics model. The model developed by Maani and Stephenson [2000] highlights provider incentives, strategy fragmentation and parents' role as the key factors impacting immunisation rates. In the new developed model the key factors associated impacting immunisation rates are parents' role, quality of healthcare system, vaccine management, programme funding and social factors which affect the community.

4. Results And Discussion

Policy experiments refer to how a manager uses information about the system in the formulation and design of policies [Maani and Cavana, 2000]. With the help of screens the model is developed into a tool that can easily be used by policy makers of the immunisation system. This section provides snapshots of the output and a few scenarios of the model simulations.

4.1. Scenario 1 : Adjusting Conditions of Immunisation Demand

This scenario allowed a user to adjust the conditions of the mother (level of immunisation awareness, socio-economic status and (health centre penetration) distance to the health centers under normal conditions of the healthcare system.

Fig. 6: Screen Shot of the Model

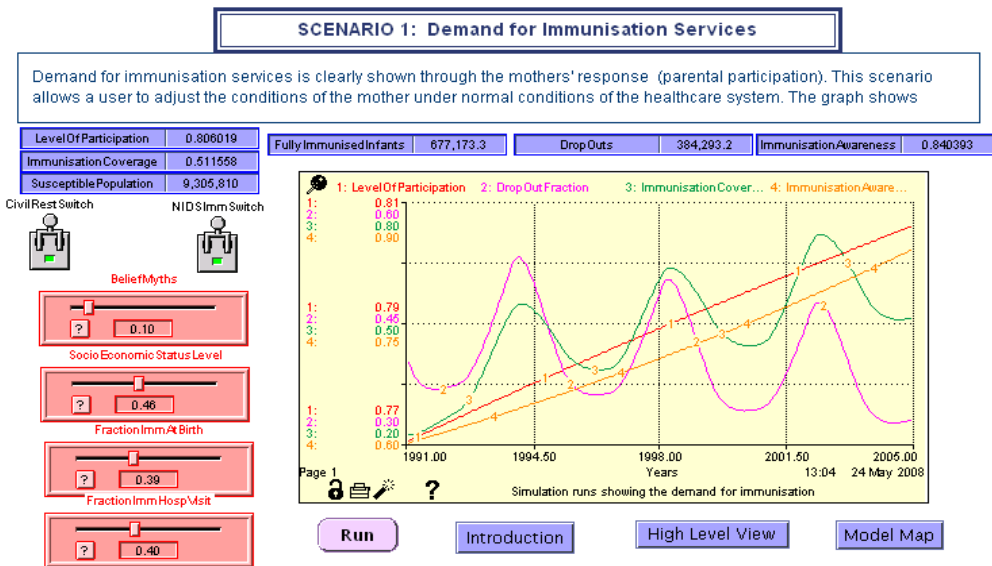
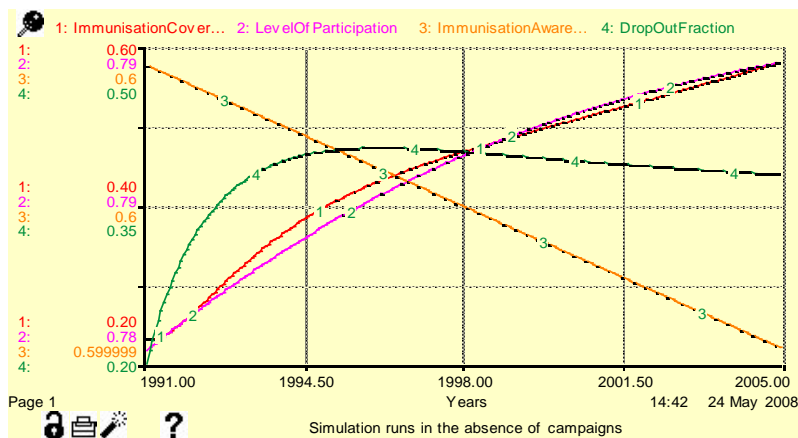


Figure 7 demonstrates the behavior of the model in the absence of campaigns. The immunisation coverage rates move gradually over time and do not reach high levels compared to Figure 6, where the rates rise close to 0.80, which implies that the absence of campaigns negatively affects awareness which affects participation in immunisation activities.

Fig. 7: Simulation Runs In The Absence Of Campaigns



Simulation runs showing the effect of Socio Economic Status of the people on immunisation response are presented in Table I.

Table 1. Variation of Socio-Economic Status

	Variable	Normal rates	Increase by 0.1	Decrease by 0.1
Inputs	Socio Economic Status	0.46	0.56	0.36
Outputs	Immunisation coverage	0.518	0.523	0.511
	Level of participation	0.818	0.834	0.797
	Fully immunised infants	682,518	685,496	678,796
	Drop outs	370,006	354,968	388,784

Results of the simulation in demonstrate that an increase in socio-economic status of the mothers results in an increase in immunisation coverage rates, the level of participation in immunisation activities and the number of fully immunised infants while a decrease results in lowered rates. An increase in socio-economic status results in lowered drop out rates, while a decrease in socio economic status depicts an increase in drop out rates. Table 2 presents results of simulation runs in the presence of civil wars.

Table 2. Effect Of Civil Wars on Immunisation Demand

Variable	Normal rates	Presence of civil wars
Immunisation coverage rate	0.518	0.49
Level of participation	0.818	0.75
Fully immunised infants	682,518	585,333
Drop outs	370,006	377,544

The presence of civil wars significantly lowers the immunisation coverage rates, the level of participation and the number of fully immunised infants thus increasing the drop outs as shown in Table 2. This demonstrates that there should be peaceful environment that enhances parental participation in immunisation. Results of the

simulation show that in order to increase the number of fully immunised infants as well as reduce the number of drop outs, the mothers should have immunisation awareness, good social economic status, absence of wars and have a health centre in the community.

4.2. Scenario 2: Adjusting Conditions of Healthcare Service

This scenario allowed a user to adjust the conditions of the healthcare service (staff levels, skill levels, health facilities, stock availability) under normal conditions of the mothers conditions. Results of the simulation showed that in order to increase the number of fully immunised infants as well as reduce the number of drop outs, the quality of the healthcare service provision should be maintained at a high quality. The health workers should be motivated and each health centre should have a sufficient number to avoid overloading which lowers the motivation. Results of the simulation show that in order to increase the number of fully immunised infants, the healthcare services have to be improved and maintained bearing in mind the growing population.

Fig. 8: Simulation Runs at the Normal Growth Rate in the Healthcare System

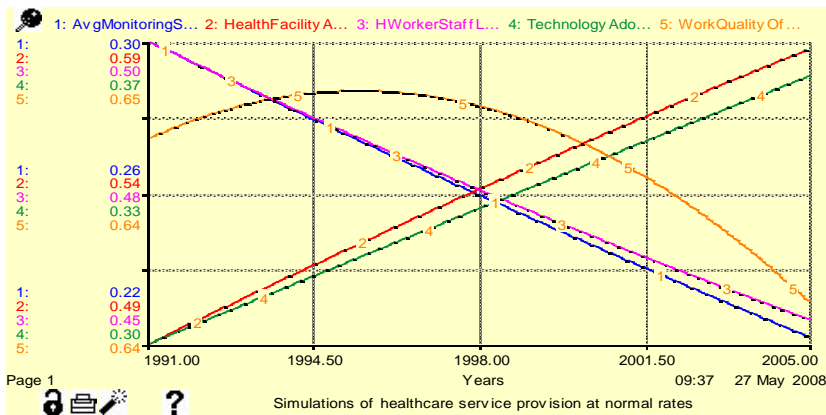


Figure 8 demonstrates a decline in the goal seeking behavior of the level of staff and level of monitoring systems which implies that the corrective action being taken is not sufficient to counteract the current trend. This contributes to the gradual decline in the work quality of the health service provision.

Fig. 9: Simulation Runs with Increase in the Healthcare System

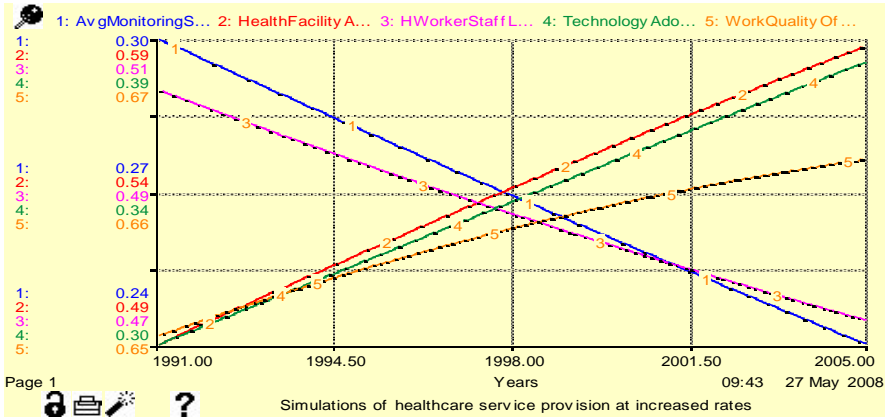


Figure 9 demonstrates that the quality of health service must be increased to achieve greater coverage and more fully immunised infants. For example, an increase in the health service components (technology adoption rate, hiring rate, monitoring systems rate) increases the fully immunised infants by over 4,000.

A comparison between Figure 8 and 9 illustrates that when the healthcare system is not upgraded (hiring new health workers, building more hospitals, adoption of new technologies and increasing the skill level of health workers), the work quality of the health system deteriorates and immunisation coverage rates drop.

Results of the simulation in Figure 9 illustrate that in order to increase the number of fully immunised infants and as reduce the number of drop outs, the quality of health care service provision should be maintained at a high level. The number of health workers should be motivated and each health centre should have a sufficient number to avoid overloading which lowers the motivation.

Table 3. Healthcare System Scenarios

	Variable	Normal growth rate	Decreased growth rate	Increased growth rate
Inputs	Technology adoption rate	0.05	0.04	0.06
	Hiring rate of health workers	0.02	0.01	0.03
	Monitoring systems growth rate	0.03	0.02	0.04
	Amount of training	0.08	0.07	0.09
Outputs	Work quality of health systems	0.643	0.640	0.657
	Immunisation coverage rates	0.464	0.463	0.470
	Fully Immunised infants	648,385	647,427	652,298
	Susceptible population	9,389,132	10,226,954	9,379,880

A comparison of Figure 8 and 9 shows when the healthcare system is not upgraded (hiring new health workers, building more hospitals, adoption of new technologies and increasing the skill level of health workers), the work quality of the health system deteriorates and immunisation coverage rates drop.

5. Proposed Intervention Strategies

Maani and Cavana define leverage points as actions or interventions that can have lasting impact on the system in terms of reversing a trend or breaking a vicious cycle. Leverage points can have substantial effect on the system by providing a fundamental and long term changes to the system as opposed to dealing with symptoms of the problems. This section highlights the key leverage points which could substantially improve immunisation demand, effectiveness of the health system, vaccine management as well as reduce immunisation costs. The proposed interventions include both short term and long term strategies as explained below:

1. *Level of participation.* The level of participation of mothers (caretakers) is key as far as immunisation coverage is concerned. The level of participation can be enhanced by considering short term and long term strategies. In the model (Figure 7), the absence of campaigns, the level of immunisation awareness declines which eventually results in lowered levels of participation in immunisation activities after some delay. This demonstrates the long term effect such a policy decision would have. Table II also demonstrates that the presence of wars in a region would tremendously impact the level of participation which is reflected by the fully immunised infants and the infants dropping out of immunisation. Some of the short term strategies include continued immunisation campaigns, increasing accessibility to health care services through the use of mobile clinics as demonstrated in Figure 9 and Table III. Some of the long term strategies include increasing the literacy levels, improving the livelihood of the people by lowering the poverty level (Table I) and minimising civil wars in the communities (Table II).
2. *Upgrade of immunisation healthcare system.* The model shows that the current level of access of healthcare services is not sufficient to meet the growing population. The level of accessibility of healthcare services currently at 49% should be increased. The nurse to patient ratio at 1: 3,065 should be increased to meet the required standard of 1:125 nurse to patient ratio (Musoke *et al.*, 2006). The level of monitoring of health services should be increased through the adoption of technologies to improve the quality of immunisation services. The model reveals that there is need to increase the hiring rate of the healthcare staff as well as address issues pertaining to the rate of attrition which results in a decline in the staff levels (Table III and Figure 9). The model reveals that access to health services would improve participation in immunisation activity. The model demonstrates that an increase in infrastructure upgrade would increase immunisation coverage.
3. *Design of relevant health information systems.* The research suggests that in order to have well suited immunisation community based health initiative

in developing countries, there is need to develop community health monitoring information systems that have procedures that track individuals by recording events and needs as services are extended to the communities (Rwashana and Williams, 2007). The success of a health information system depends on whether it meets the requirements of the stakeholders, which necessitates a deeper understanding of the organisational environment such as the one provided by the System Dynamics. A broad integrated view of the immunisation system provides an analysis of the complex interactions between behavioural, technical, policy and cultural issues which fosters communication among the stakeholders by linking up the non-technical view and technical view of the designers and programmers of information systems there by capturing the requisite information.

4. Management of resources. The immunisation system uses various resources to provide services to the communities. There is need to develop automated resource ordering and tracking systems in order to minimise stock out situations as well as overstocking which may lead to losses due to expiry of drugs.

6. Conclusion

The model focuses on the dynamic aspects that may be potentially within control by the stakeholders involved in the immunisation process. The model helps to generate insight to policy development as far as immunisation and healthcare is concerned and generates deeper insight of the issues that are associated with immunisation coverage. The research demonstrates the benefits of using systems thinking methods in facilitating the understanding of healthcare system.

References

- BAYER, S., KOBERLE-GAISER, M. AND BARLOW, J.(2007). Planning for Adaptability in Healthcare Infrastructure. *Proceedings of the 25th International System Dynamics, 29 Jul-2 Aug 07, Boston, USA*. System Dynamics Society. ISBN 978-0-9745329-7-4.
- CHU, S.C., KIN, C.K. AND LAM, S.S. (2003). Hospital lift system simulator: A performance evaluator-predictor. *European Journal of Operational Research*, **146**: 1:156-180.
- CODY, J., CAVANA, R.Y. AND PEARSON, D. (2007). Social Structure and health : Preliminary work on a System Dynamics Concept Model. *Proceedings of the 25th International System Dynamics, 29 Jul-2 Aug 07, Boston, USA*. System Dynamics Society. ISBN 978-0-9745329-7-4.
- DEXTER,F., MACARIO,A., TRAUB, R.D., HOPWOOD, M. AND LUBARSKY,D. A.(1999). An operating room scheduling strategy to maximize the use of operating room block time:computer simulation of patient scheduling and survey of patients' preferences for surgical waiting time. *Anaesthetic Analog*, **89**:7-20.
- DITTUS, R. S., KLEIN, R. W., DEBROTA, D. J., DAME, M. A., AND FITZGERALD, J.F. (1996). Medical Resident Work Schedules: Design and Evaluation by Simulation Modeling. *Management Science*, **42**(6)891-906.

- FITZPATRICK, K.E., BAKER, J.R. AND DAVE, D.S. (1993). An application of computer simulation to improve scheduling of hospital operating room facilities in the United States. *Int Journal of Computer Applications Tech*, 6:205-224.
- FONE, D., HOLLINGHURST, S., TEMPLE, M., ROUND, A., LESTER, N., WEIGHTMAN, A., ROBERTS, K., COYLE, E., BEVAN, G. AND PALMER, S.(2003). Systematic review of the use and value of computer simulation modeling in population health and health care delivery. *Journal of Public Health Medicine*, 25(4):325-335.
- FRASER, H.S.F. AND MCGRATH, J.D. (2000) . Information technology and telemedicine in sub-Saharan Africa. *British Medical Journal MJ* 321 : 465-466.
- HASLETT, T. (2007). From Hospital Plans to Computer Simulation : A Case study of Alfred Centre. *Proceedings of the 25th International System Dynamics, 29 Jul-2 Aug 07, Boston, USA. System Dynamics Society*. ISBN 978-0-9745329-7-4.
- HETHCOTE, H.W., HORBY, P. AND MCINTYRE P. (2004). Using computer simulationsto compare pertussis vaccination strategies in Australia. *Vaccine Journal*,22(17-18):2181-91. Pub-Med.
- MAANI, E.K. AND CAVANA, R.Y. (2000). *Systems Thinking Modelling : Understanding Change and Complexity*. Prentice Hall.
- MAANI,K.E.AND STEPHENSON, P. (2000). Evaluation of Immunisation strategies in New Zealand-A Systems thinking approach. 1st International Conference on System Dynamics in Management.
- MARTIN, J.F. AND MARSHALL, J. (2002). New tendencies and strategies in international immunisation : GAVI and The Vaccine Fund. *Vaccine* 21:587-592.
- MBARIKA, V.W.A. (2004). Is Telemedicine the Panacea for Sub-Saharan Africa's Medical Nightmare ? *Communication for the ACM*, 47(7): 21-24.
- MCALEER, W.E., TURNER, J.A., LISMORE, D., AND NAQVI, I.A.(1995). Simulation of a Hospital's Theatre Suite. *Journal of Management in Medicine*, 9(5): 14-26.
- MUSOKE G., AKAMPUMUZA G., ADAIR M., IUTUNG P., BURROWS R.,NSAKANYA R., WE-INSTEIN M., FARTHING C., CHANG H. AND OKONGO B.(2006).Using Lay Providers as HIV Medics to Bridge the Human Resource Gap in ART Service Delivery: Experiences from Uganda and Zambia. *XVI International AIDS Conference* 13-18 August 2006. Toronto, Canada
- RWASHANA A.S. AND WILLIAMS D.W. (2007). Enhancing Immunisation Coverage through Health Information Systems : A System Dynamics Approach. *Stud Health Techno Inform*. 130: 247-56.
- SUBRAMANYAM, K. AND SEKHAR, S.C.(1987). Improving Vaccine Supplies to Rural India. *Socio-Economic Planning Science*, 21(2)131-138.
- TOBIAS, M.I. AND ROBERTS, M.G.(2000). Predicting and preventing Measles epidemics in New Zealand : application of a mathematical model. *Epidemiological Infections*. 124: 271-287.
- UBOS (2007). 2006 Uganda Demographic and Health Survey.
- WASUKIRA, E., SOMERWEI, F AND WENDT, D. (2003). ICT for development : ICT in Health Seminar. *I-network Uganda*. 2:4

- WILLIAMS, D., (2000). *Dynamic Synthesis: A Theoretical Framework for Research in Requirements Engineering Process Management*. *Operational Research Society*, ISBN: 0 903440202.
- WHO/UNICEF (2006). Global Immunisation Coverage. http://www.who.int/immunisation/_monitoring/data/en [accessed on 17 March, 2007]
- WHO (1999). Measles. Progress towards global control and regional elimination. *Weekly Epidemiology*, 74(50): 429-434.

Geometrical Spatial Integration Model for Geo-Information Management

Ismail Wadembere and Patrick J. Ogao

Geospatial information systems provide avenues to fulfill users' quest for both absolute and relative locations of features/objects. To achieve this, the users need to have access to different geospatial data sets from various sources. This calls for integration of data from different geospatial sources. But, there are shortcomings as different data sets may not map exactly onto one another, one of reasons being the difference in features' geometry. This paper presents a conceptual model for geospatial data integration that can identify and measure differences and adjust spatial geometries of geospatial features to form meaningful objects which can be used for geo-spatial analysis, modeling, and easy geo-information management.

1. Introduction

Geospatial information systems especially Geographical Information Systems (GIS) have enabled users to analyze, model, and visualize geospatial (space and earth related) data. This has made GIS a useful media for exchange of geo-spatial information among individuals and organizations and a basis (both as science and tool) for space and location based decision-making.

However, it takes a lot of time, effort, resources, and skills to create and store geo-spatial data and because of that, in most situations many practitioners do not have all the data they need in one database (FGDC, 2007). To meet this need, many individuals and organizations use geo-spatial data from different sources. This has made geo-spatial data integration difficult to ignore as there is need to take advantage of different spatial data in different locations to reduce on the cost and time involved in data production.

Geospatial data integration is not straight forward as these data sets are collected by many individuals/organizations and sometimes using different methods, standards, data models, and information technologies. As result, the data may vary in spatial geometries causing heterogeneity, which can be zonal/horizontal or layered/vertical fragmentation (Kampshoff, 2006). In this paper, we handle zonal heterogeneity which is a result of geometrical difference in thematically similar data sets that can be evidenced by incomplete or inaccurate mapping of the data sets.

This paper looks at the need for geospatial data integration which motivates investigation into the different data integration approaches, spatial representation,

geometrical integration and their limitations. We use the outcomes of that research to triangulate different methods as we move towards designing specifications for a Geometrical Spatial Integration Model (GSIM). We end this paper by a look at proposed future work and conclusion.

2. The Need for Geo-spatial Data Integration

Currently, we are seeing developments in technologies which are creating services which need geospatial data comparison (Najar et al., 2006) and integration (Sester et al., 2007). This is evident with increased geo-spatial developments like Google Earth, Microsoft's MapPoint.Net, Oracle Locator and Spatial, and ESRI's ArcGIS; new platforms, vendors, users, and conferences and publications on location-based services (Zhao and Motorola, 2002), (Chang et al., 2007), (Sester et al., 2007), (Busgeeth and Rivett, 2004). These developments and changes although seemingly diverse, are related as they all use and take advantage of the geospatial information available (Sagayaraj et al., 2006). That is why we are seeing increased importance of location in information systems and services (Strader et al., 2004).

Increasingly, technologies like the Internet, web services, image capture and communication, low cost spatial data collection tools, grid computing power, and On-Line Transaction Processing (OLTP) application (Skog, 1998), (Sharma, 2005), (Sohir, 2005) among others utilize geo-spatial data. In the process, technologies are positively influencing the need for spatial data integration.

The need for geospatial integration is further fuelled by Location Based Services and Location-Tracking trends; for example in the United States the move to location-aware mobile phones have been driven by the federal government's enhanced 911(E911) mandate (FCC, 2006). Also, researchers are experimenting with ideas such as leaving messages for other people attached to certain places, all enabled by on-line location services (Sharpe, 2002). All these examples underline the emergent need for geospatial data integration to improve on sharing and information management (Shi and Yeh, 1999).

Further use of geospatial data and need for integration can be seen in Enterprise Location Services and Software (ELSS); which concern the use of location-specific information within information systems. Several technologies combine to make ELSS including the Service-Oriented Architectures (SOA) and the automated location determination technologies like GPS, cellular networks, and radio frequency. Also, spatially-enabled IT Infrastructure, Open Standards like Geographical Markup language (GML), Spatial Data infrastructures (SDI), 3G System (Zhao and Motorola, 2002), etc are becoming very common (Sharma, 2005, Sohir, 2005).

All these technologies provide avenues for sharing geo-information from different sources. But for many organizations, integration is still a major challenge in GIS projects.

3. Data Integration Approaches

The need for integration arise from the nature of geo-spatial data and information - huge data sets, different formats, and the need to share information between organizations and individuals (Evans, 1997), (Erdi and Sava, 2005). However, most data sources, databases and information systems are not designed for integration. Thus, whenever integrated access to different source systems is desired, the sources and their data that do not fit together have to be coalesced using additional adaptation and reconciliation functionality. The goal is always to provide a homogeneous, unified view on data from different sources.

Integration tasks may depend on various factors such as those (Ziegler and Dittrich, 2004) list: “the architectural view of system; the content and functionality of the component systems; the kind of information that is managed by component systems (spatial data, alphanumeric data, multimedia data; structured, semi-structured, unstructured data); requirements concerning autonomy of component systems; intended use of the integrated system (read-only or write access); performance requirements; the available resources (time, money, human resources, know-how, etc.); and the source and type of method used to collect the data”.

Additionally, several kinds of heterogeneity typically have to be considered. These include differences in (i) hardware and operating systems, (ii) data management software, (iii) data models, schemas, and data semantics, (iv) middleware, (v) user interfaces, and (vi) business rules and integrity constraints (Ziegler and Dittrich, 2004).

Different researchers ((Ziegler and Dittrich, 2004), (Nebert, 2003), (Sohir, 2005), (Musinguzi et al., 2004), (Evans, 1997), (Nebert, 2003), (Erdi and Sava, 2005), (Kilpelainen, 1997), (Friis-Christensen et al., 2005), (Skog, 1998), (Sharma, 2005), (Sester et al., 2007), and (Bell, 2005), and so on) present different ways of addressing the integration problem which can be done at presented system layers, data storage, open systems, and setting sharing channels. These include:-

Manual Integration: users directly interact with all relevant information systems and manually integrate selected data. The users have to deal with different user interfaces and query languages that require detailed knowledge on location of data, logical data representation, and data semantics.

Peer-to-peer (P2P) Integration: a decentralized approach to integration between distributed and autonomous peers where data can be mutually shared and integrated. P2P integration constitutes, depending on the provided integration functionality, either a uniform data access approach or a data access interface for subsequent manual or application-based integration.

Common User Interface: user is supplied with a common user interface (e.g. web browser) that provides a uniform look and feel when accessing data from different relevant information systems. System are separated but homogenization and integration of data is done by the tools like search engines.

Integration by Applications: integration applications access various data sources and return integrated results to the user. This is mostly suitable for systems with a few components since applications become many as the number of system interfaces and data formats to homogenize and integrate grows. Examples include workflow management systems that allow implementation of business processes where each single step is executed by a different application.

Integration by Middleware: Middleware provides reusable functionality that is generally used to solve dedicated aspects of the integration problem e.g. SQL-middleware. Applications are needed in the form of middleware tools and are usually combined to build integrated systems and they are relieved from implementing common integration functionality.

Uniform Data Access: logical integration of data is accomplished at the data access level. Global applications are provided with a unified global view of physically distributed data, though only virtual data is available on this level. However, global provision of physically integrated data can be time-consuming since data access, homogenization, and integration have to be done at runtime. Examples include mediated query systems, portals for Internet/intranet, federated database systems, etc.

Open systems: geospatial data integration is done using distributed models that are based on open systems technologies like OpenGIS, XML, GML, etc.

Geo-information Sharing Infrastructures: communities agree on how to share data like SDI, Geospatial Data Clearinghouse, Proprietary protocols, GIS-network integration, CRUD matrix (which stands for Create, Read, Update, and Delete) where modeling tool is used in data-oriented methodologies to show processes and their associated data. CRUD is a spatial data integration approach based on comparing organizations spatial data in a matrix. It shows data sets on the vertical axis and their associated departments or users on the horizontal axis. The matrix provides information on data ownership, maintenance, use, and geographical distribution.

Common Data Storage: Here, physical data integration is performed by transferring data to new data storage; local sources can either be retired or remain operational. This approach provides fast data access. Examples include federated database systems, data warehouses (data from several operational sources like OLTP are extracted, transformed, and loaded (ETL)), storing geo-data with varying standards in same geodatabase using multiple representation, and attribute versioning.

However, these approaches do not meet all the requirements, as some times, there is need to change and adjust geometries of spatial features (Peng, 2005), so that data with varying geometries can fit exactly on each other to allow the process of integration to succeed. Thus, in this paper we base our approach on common data storage to develop a conceptual GSIM. The model is to support the

determination of differences between the geometries of geo-spatial datasets; then to adjust these geometries so that data can be integrated before they are adapted/adopted in sharing avenues and used in spatial analysis and modeling.

4. Spatial Representation and Geometrical Integration

Geo-Spatial data is data that occupies cartographic (map-able) space that usually has specific location according to some geographic referencing system (e.g. x-y) or address. This data is represented using two major data structure: raster and vector, which help to store spatial characteristics of features. Vector data are based on features and have x and y coordinates. Raster data are based on pixels, with a grid like system of rows and columns (Burrough, 1986), (Nievergelt and Widmayer, 2006).

In developing GSIM, our focus is vector data. Vectors are scalable drawings created from mathematical calculations and stored in the computer as mathematically defined geometrical primitives (points, lines/curves, and polygons) making them easy to create and edit. Vectors can be reduced and enlarged with little loss of quality because the equations can easily be recalculated at new scales (Chrisman, 1997).

A geometrical point is a one-dimensional feature (has length only and no width). It is a spot (location) that has no physical or actual spatial dimension. Points indicate specific location of features, which are usually not shown in true size, especially for things that are too small to depict properly at a given scale. Points also show location of intangible (non physical entities) features e.g. address or location of occurrence like traffic accidents. A lines has a beginning and an end and are linear features; either real (e.g. road) or administrative (e.g. boundary). Sometimes the thickness of the line indicates a measure (such as amount of traffic on a road) or type of road (such as primary verses secondary). A polygon is an enclosed area, a two dimensional feature with at least 3 sides (and therefore with an inside area and perimeter) e.g. plot of land - (Burrough, 1986).

Spatial objects can be grouped into layers, also called overlays, coverages, or themes. One layer may represent a single entity type or a group of conceptually related entity types e.g. a layer may have only river or may have rivers, lakes, and swamps. Options depend on the system as well as the database model and some spatial databases have been built by combining all entities into one layer (Goodchild, 1992).

Elements of reality modeled in a GIS database have identities: (i) the element in reality – entity “a phenomenon of interest in reality that is not further subdivided into phenomena of the same kind”, (ii) the element as it is represented in the database – object “a digital representation of all or part of an entity”, and (iii) a third identity important in cartographic applications, namely the symbol that is used to depict the object/entity as a feature on a map or other graphic display (Tomlin, 1990).

Geometrical integration incorporates schema integration and solving semantic conflicts of the datasets (Kampshoff, 2005). The interest here is the schema integration at the model level as the spatial feature can be broken into primitives (such as point, line/arc, and polygon). The primitives are expected to be the same for similar thematic datasets (datasets representing the same feature), but it is not always the case. In some situations, we have spatial geometrical heterogeneity existing which can be zonal/horizontal or layered/vertical fragmentation (Kampshoff, 2005).

Zonal heterogeneity is the result of capturing thematically similar data independently by different individuals with no common framework being followed. Examples of problems include:- same points with varying coordinates, arcs which are supposed to be the same intersecting, neighboring areas overlapping each other, etc. These lead to geometrical and topological inconsistencies like creating sliver polygons, dangling arcs and nodes. Layer fragmentation occurs when a base-layer on which the integrated multi-layer datasets depend is replaced by an updated version but the rest of the new layers are collected using the old one. This is termed as systematic tessellation (Masuyama, 2006) which is the difference due to use of different geodetic models when capturing layers.

In this research, we look at zonal heterogeneity, where the differences have to be determined and adjusted in order to have physical data integration at layer levels. (Casado, 2006) and (Kampshoff, 2005) present several approaches to geometrical integration which include:-

Similarity or Helmert Transformation approach which deal with whole layer being treated as solid image where the geometry adjustment is done to change the whole layer. Although the angles between features are preserved, the distances are not. This approach is not applicable in our situation as we do not consider layer as one solid object but we are interested in changing the geometry of a single element on the layer.

Affine Transformations approach to geometrical correction preserve neither distances nor angles although it can change the geometrical of the whole layer. Again this approach can not accomplish our need of just adjusting specific features.

Rubber sheeting method allows local distortions and preserves topological properties by distributing random errors through out the data set. This is good to overcome the problem of having other feature to appear more changed than others, but it is not an individual feature change method and can not directly be applied to change the shape of specific primitive as it will distribute the error to others which do not have any problem.

Complex Conformal Mapping approach preserve angles and multiplies the length of short segments by a certain factor as it tries to keep the features on the layer undistorted. This approach is for whole layers and may not be applied to a single element of the layer's geometry.

Other geometric correction algorithms like one developed by (Lee et al., 2002) are for satellite image correction as the parameters considered in the correction are instrument characteristics, the satellite motion, earth rotation, terrain effects, and the ground control points (GCP) which are not applicable in vector GIS.

Stochastical geometrical integration tries to balance the geometrical error available in the layer and in the process does not reproduce the values for data points as the random error is filtered in these locations.

Deterministic geometrical integration models can be applied when the random error is of negligible magnitude. That means that if the separation between features is very big, this approach can not be applied.

Kampshoff (2005) combines the last two approaches to come up with an improved geometrical integration model. Since this approach needs several points from a true map which have known coordinates in order to carry out adjustment, it can not be applied in our situation where we may need only to move a feature without using known points as they are may not available .

5. Triangulation of Methods For GSIM

For established fields of science like Information Systems (IS), academicians often take different approaches in order to solve different problems, and expand on the researches. There are many reasons for this including (King and Lyytinen, 2005), (Benbasat and Zmud, 1999), (Niehaves, 2005), (Wangler and Backlund, 2005), (Erdi and Sava, 2005), (Aronson, 1998), (Bernard et al., 2003), (Mingers, 2001) :-

- a. The need to achieve a relation between IS and other fields.
- b. Different approaches have been used by researchers as it gives them the opportunity to be creative.
- c. The question of whether existing research approaches provide a suitable basis to start new research.
- d. New researchers in IS with little or no knowledge about IS research approaches and paradigms end up creating new approaches.
- e. The rapid change in Information Technology (IT) has led IS researchers to adopt or create new approaches which can cope with IT changes.
- f. The changing needs of users such as the use of geo-spatial/location data in many IS
- g. The emergence of open standards has forced researchers to change approaches.

Diverse academic disciplines, approaches, and research communities contribute to IS research (Niehaves, 2005). It is against this background of diversity that we develop GSIM using multi-methods so that it can benefit from such a variety of approaches and contributions. In the process, it helps us to see IS-related phenomena from different points of view (Mingers, 2001), (Weber, 2004), thus eventually benefitting the GSIM development. Among the issues that we considered was to develop GSIM using different approaches/paradigms basing on independent

components/units/object/sub-models, which can function both as stand alone or integrated with others to accomplish a task.

This means that GSIM is not tied to one format (Bernard et al., 2003) and changes are implemented by modifying specific component concerned. Emphasis will be placed on developing the GSIM that can be implemented using messaging, UML, GML, and XML manipulation for identifying data components and triggering different units/objects. The development of the GSIM will also utilize spatial validation, spatial matching, spatial interpretation, reverse engineering, and system thinking (Aronson, 1998), and GIS spatial functions such as overlay, merge, connectivity, neighborhood, and operations.

6. Specifications and Design of GSIM

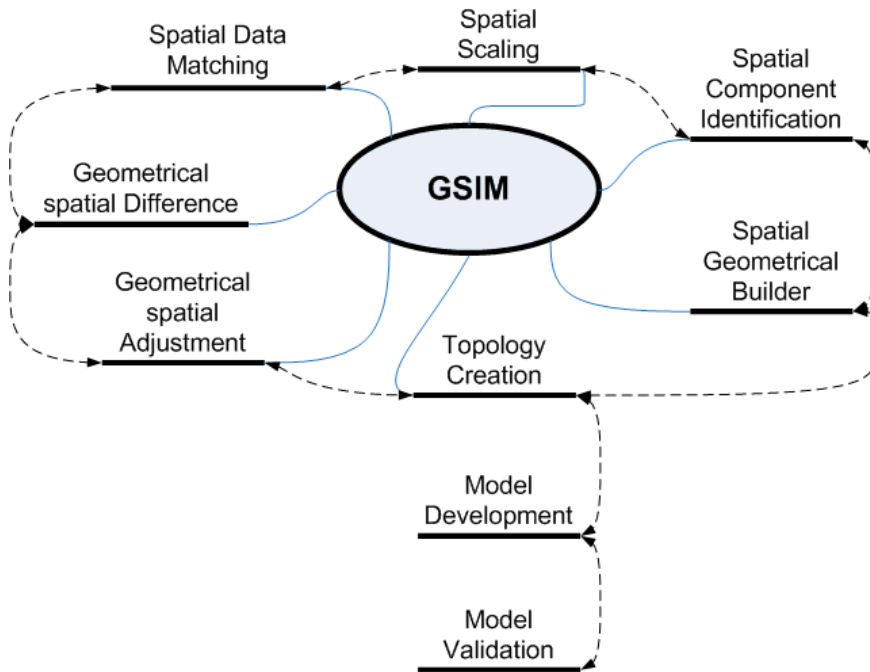
Assume a user has two geo-spatial data sets from different sources and s/he is interested in using the two data sets, thus s/he has to integrate the two data sets. Several issues may emerge that need to be considered during the design of GSIM:-

- a) geo-spatial data may contain bias e.g. some features being over emphasized and some under represented
- b) different data sets representing the same entity may extend over a different area
- c) data sets may just be incomplete i.e. do not have all the required elements for certain analysis and modeling to be meaningful
- d) sources may contain closely related and overlapping data
- e) data may be stored in multiple data models and schemas
- f) topology requirements like overlap, containment, connections, and adjacency may be missing
- g) data sources may be having differing query processing capabilities

Questions we asked ourselves in order to define a problem that formed the basis for designing GSIM were: - a) When are two spatial data sets different? b) Is the difference due to geometry? c) Which of the data sets is correct or better? d) What parameters of geometry are different? e) Can the geometrical parameter be changed? f) How much of the difference is due to geometrical difference? g) How can geometrical parameters be adjusted? h) How can a combined spatial geometry be built? i) How can data sets be combined?

The above questions were filtered to get the specific components of GSIM (see **Figure 1**): spatial scaling, spatial data matching, best spatial data set, spatial components identification, geometrical spatial difference, geometrical spatial adjustment, geometrical spatial builder, topology creation, and final model development.

Fig 1: Geometrical Spatial Integration Model Components



Spatial Scaling

Two data sets are checked to see if they occupy the same aerial extent, if not the data sets are adjusted so that they occupy the same area. This may involve a change of scale because of shrinkage or enlargement of a particular data set. If data sets representing the same area have different scales, then scale change has to be done so that the data sets cover the same area extent. The intension here is not to handle whole scale of the data within the object although scale change can take place so that the datasets cover the same area extent. The scale change, which is handled by the GSIM, affects individual geometries of the spatial elements for example changing the shape of polygon say a plot of land.

Spatial Data Matching

The main purpose for this object is to determine if two spatial data sets are different and if the difference is due to geometry. Here, we also determine which geometrical parameters are causing the difference. This involves spatial object matching to predict errors due to inexact matching, which end up producing for example sliver lines/polygons.

We employ the ideas of spatial data matching and spatial interpretation in determining the spatial data sets with similar geographies and comparing spatial data sets at object level, will take advantage of the computational model developed by (Masuyama, 2006) . The end product at this level is an algorithm that can determine spatial geometry differences and identify which components of the geometry are causing the mismatch. The main considerations are scale, resolution, and actual errors in a data set.

Best Spatial Data Set

This compares different data sets using known data sets (especially base data set) or known controls. The purpose is to identify correct data sets or to assess which data sets should be used as the basis and whose integration parameters will be used during adjustment.

Spatial Component Identification

As the various components of geo-spatial data sets have to be handled differently during special adjustment, the role of this component is to breakdown the different primary component of the data set into individual layers i.e. point layer, ploylines layer, and polygon layer.

Geometrical Spatial Difference

This is used for geometrical quantitative difference determination. This involves the process of extracting the geo-data's spatial geometry and determining the geometry relationships of components depending on the particular geo-data source.

Spatial validation is utilized to determine elements and is based on a) Logical cartographic consistency such as overlap, containment, adjacency, and connectivity, b) Closed polygons, c) One label for each polygon, d) No duplicate arcs and e) No overshoot arcs (dangles). GIS spatial functions like overlay, merge, connectivity, neighborhood, operations are also utilized to develop an algorithm for geometrical quantitative spatial difference determination. Current computational models which are based on geometric features like one suggested by (Masuyama, 2006) will be employed.

The outcome this should be able to handle:-

- *For line features*, the whole line feature is picked at a time and comparison takes place in the following parameters:- length of the ploylines, location (using coordinates) of the nodes and vertices, overlap, number of nodes, thickness, and continuity (connectivity).
- *For the point features*, an imaginary fine grid is over laid on the two data sets so that the values in each grid can be compared. The grid size is according to the scale of the data set i.e. the grid square is determined by the scale at which the comparison is being done. The difference is determined and comparison in precision, resolution, and actual data values is done. If the difference is zero, we assume that the two cells are the same and if the resultant is not zero (could be positive or negative), then the two are different.
- *For polygons*, whole polygons are compared, the shape, corner/edge coordinates, overlaps, label points, adjacency, containment, over size, under size, neighborhood, not closed, and does not have a label.

Geospatial Ontology and Geometrical Spatial Adjustment

After identifying the components causing the differences and determining how much of the difference exist; this unit/object breaks down geometries of data to individual primitives (points, arcs and polygons). This is done as geometrical differences (not semantics differences as described by (Budak Arpinar et al., 2006)) within individual features among thematically similar data sets can still remain present even after overlaying and merging. So the adjustment should be done on the primary elements – the primitives of individual features.

This approach is being adopted as spatial data has been conceptually defined as geometric data consisting of points, arcs and polygons (Lu et al., 2007) and these conceptual objects (geometric data) are used to represent GIS objects/features which should be robust and efficient in terms of storage, retrieval, processing and good graphical reconstructing.

However, to handle the different primitives at the same time brings in challenges in terms of requiring different algorithms to adjust the different primitives (Masuyama, 2006). This is so as it can be observed by taking a close look at the different primitives. It is evident they have different geometries and situation could be made simpler if only one type of primitive can be handled during geometry adjustment. In this research we look at disintegrating the lines/arcs and polygons into points as a move towards providing only one primitive type (points) on which geometry adjustment can take place.

Spatial validation and Spatial Reverse Engineering (SRE) are employed here to help in disintegrating geo-data from any source into primitive objects/elements. This is done by considering the strength and goals of SRE (i) to analyze a data set, (ii) to identify the data's components, (iii) to identify the interrelationships of the data components, and (iv) to create representations of the data in another form or at a higher level of abstraction. SRE helps to return to a less complex or more primitive state or stage of the spatial data geometry. The resultant algorithm has to accomplish *geometrical point correction* according to the geometrical spatial difference of features which could have comprised of *polylines, points and polygons*.

Geometrical Spatial Builder

This unit/object looks at the different spatial primitives that have been adjusted and builds them into objects which have spatial meaning; these are later combined during topology creation and building.

Topology Creation and Building

As GIS data should have proper topology to enable users to carry out modeling and analysis, this unit/object combines the spatial objects generated using geometrical spatial builder into proper polygons, ploylines, and points and creates relationships between them to obtain meaningful spatial data sets.

Final Model Development

This is the combining of data sets after the above units/objects have played their roles. After developing the different units, the resultant GSIM is achieved using an iterative process where the design rules for each possible sequence of the unit/object are taken into consideration and deployed according to given user needs/requirements. The resultant GSIM and also the different units/objects can be deployed using approaches like novel service retrieval (Klein and Bernstein, 2004) that capture service semantics using process models and apply a pattern-matching algorithm to find the services with the behavior the user wants.

Model Validation and Evaluation

To make sure the model can be utilized; GSIM will be evaluated with the help of topological requirements testing. This will be carried out both theoretically and experimentally using the following criteria: the maturity and capabilities of each object, whether or how well each object can be used in conjunction with others, and how the whole model can fulfill geometrical integration requirements. A goal-based evaluation will be performed to decide if our goals are met. We will look for logical cartographic consistency involving testing for overlap, containment, adjacency and connectivity. This requires all polygons be closed, there be no duplicate arcs and no overshoot arcs (dangles) and only one label for each polygon, and so on. GIS spatial functions like overlay, merge, connectivity, neighborhood, and operations will also be utilized in testing.

7. Future Work

The conceptual GSIM will be implemented and tested to determine if it meets the need of the geospatial data integration. We will examine how GSIM can take advantage of the expanding geo-technologies like GML. Subsequently we plan to investigate if it can be adapted and used in spatial analysis and modeling which can feed into sharing avenues like SDI and general geo-information management.

8. Conclusion

The main function of geospatial information systems is to provide location specific information which is easy to understand and can be used in decision-making. In this paper, we have presented a conceptual Geometrical Spatial Integration Model (GSIM) which when implemented can identify and compare differences and adjust geometries of primitives (points, lines, and polygons) so that they fit exactly to come up with complete and meaningful spatial features. With this model in place, geospatial practitioners may be able to identify sliver primitives and features in data sets which can be eliminated, making integration easier, and leading to increased use of different data sets created and kept by different individuals. This may facilitate natural resources managers/users access to geo-information, and lead to improved analysis and modeling results, reduced cost and time involved in geo-spatial data production, and informed location-based decision-making.

References

- ARONSON, D. (1998) Systems thinking. *R\&D Innovator. (now Innovative Leader)*, 6.
- BELL, C. A. (2005) Attribute-Level Versioning: A Relational Mechanism for Version Storage and Retrieval. *School of Engineering*. Richmond, Virginia (USA), Virginia Commonwealth University.
- BUDAK ARPINAR, I., SHETH, A., RAMAKRISHNAN, C., LYNN USERY, E., AZAMI, M. & KWAN, M. P. (2006) Geospatial Ontology Development and Semantic Analytics. *Transactions in GIS*, 10, 551-575.
- BURROUGH, P. A. (1986) *Principles of geographical information systems for land resources assessment*, Oxford, Oxford University Press.
- BUSGEETH, K. & RIVETT, U. (2004) The use of a spatial information system in the management of HIV/AIDS in South Africa. *International Journal of Health Geographics* 3, 13.
- CASADO, M. L. (2006) Some Basic Mathematical Constraints for the Geometric Conflation Problem. IN
- CAETANO, M. A. P., M. (Ed.) *7th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences*.
- CHANG, C.-L., CHANG, Y.-H., CHIANG, M., CHUANG, T.-R., DENG, D., HO, S. & LIN, F.-T. (2007) Experience in Building GML-based Interoperable Geo-spatial Systems. Institute of Information Science, Academia Sinica, Taipei, Taiwan.
- CHRISMAN, N. R. (1997) *Exploring geographic information systems*, New York, John Wiley and Sons. Inc.
- ERDI, A. & SAVA, D. S. (2005) Institutional Policies on Geographical Information System (GIS) Studies in Turkey. *Pharaohs to Geoinformatics, FIG Working Week 2005 and GSDI-8*. Cairo, Egypt.
- EVANS, D. J. (1997) Infrastructures for Sharing Geographic Information among Environmental agencies. Massachusetts Institute of Technology.
- FCC (2006) Federal Communications Commission. Public Safety and Homeland Security Bureau, USA
- FGDC (2007) The Federal Geographic Data Committee.
- FRIIS-CHRISTENSEN, A., NYTUN, J. P., JENSEN, C. S. & SKOGAN, D. (2005) A Conceptual Schema Language for the Management of Multiple Representations of Geographic Entities. *Transactions in GIS*, 9, 345-380.
- GOODCHILD, M. F. (1992) Geographical information science. *International Journal of Geographical Information Science*, 6, 31-45.
- KAMPSHOFF, S. (2005) Mathematical Models for Geometrical Integration. *First International Workshop on Next Generation 3D City Models*. Bonn, Germany, Geodetic Institute, RWTH Aachen Univeristy, Germany.
- KAMPSHOFF, S. (2006) Mathematical Models for Geometrical Integration. Geodetic Institute, RWTH Aachen Univeristy, Germany.

- KILPELAINEN, T. (1997) Multiple Representation and Generalization of Geo-databases for Topographic Maps. Finnish Geodetic Institute.
- KLEIN, M. & BERNSTEIN, A. (2004) Toward high-precision service retrieval. *Internet Computing, IEEE*, 8, 30-36.
- LEE, Y. Y., WU, A. M. & WU, F. (2002) An Algorithm for Geometric Correction of High Resolution Image Based on Physical Modeling. *The 23rd Asian Conference on Remote Sensing*.
- LU, C.-T., SANTOS, R. D., SRIPADA, L. & KOU, Y. (2007) Advances in GML for Geospatial Applications. *Geoinformatica*, 11, 131-157.
- MASUYAMA, A. (2006) Methods for detecting apparent differences between spatial tessellations at different time points. *International Journal of Geographical Information Science*, 20, 633-648.
- MINGERS, J. (2001) Combining IS research methods: towards a pluralist methodology. *Information Systems Research*, 12, 240-259.
- MUSINGUZI, M., BAX, G. & TICKODRI-TOGBOA, S. S. (2004) Opportunities and Challenges for SDI development in development countries - A Case study of Uganda. *Proc. 12th Int. Conf. on Geoinformatics - Geospatial Information Research: Bridging the Pacific and Atlantic*. University of Gävle, Sweden, Geoinformatics.
- NAJAR, C., RAJABIFARD, A., WILLIAMSON, I. & GIGER, C. (2006) A Framework for Comparing Spatial Data Infrastructures An Australian-Swiss Case Study. *GSDI-9 Conference Proceedings*. Santiago, Chile.
- NEBERT, D. (2003) *Issues of Data Standardization*, Global Spatial Data Infrastructure (GSDI).
- NIEHAVES, B. (2005) Epistemological perspectives on multi-method information system Research” , . European Research Center for Information Systems, University of Münster, Leonardo-Campus 3, 48140 Münster, Germany.
- NIEVERGELT, J. & WIDMAYER, P. (2006) Spatial data structures: Concepts and design choices. *Handbook of Computational Geometry*, 725-764.
- PENG, Z. R. (2005) A proposed framework for feature-level geospatial data sharing: a case study for transportation network data. *International Journal of Geographical Information Science*, 19, 459-481.
- SAGAYARAJ, F., P. THAMBIDURAI, B.S. BHARADWAJ, G.N. BALAGEI & N. HEMANT (2006) An Improved and Efficient Storage Technique for GIS Geometric Primitives Based on Minimum Bounding Rectangles. *9th Annual International Conference, Map India 2006*. India.
- SESTER, M., GÖSSELN, G. V. & KIELER, B. (2007) Identification and adjustment of corresponding objects in data sets of different origin. *10th AGILE International Conference on Geographic Information Science 2007*. Aalborg University, Denmark.
- SHARMA, J. (2005) Spatial Database Architecture - Influence of IT Trends on Spatial Information Management. *Directions Magazine*. Directions Magazine.
- SHARPE, B. (2002) Trends in Information Technology. *The Appliance Studio Ltd*.

- SHI, X. & YEH, A. G. O. (1999) The Integration of Case-based Systems and GIS in Development Control. *Environment and Planning B*, 26, 345-364.
- SKOG, J. (1998) Trends in applying spatial Technologies. *13th ESRI European User Conference*. France.
- SOHIR, M. H. (2005) The Role of ESA in Building the Egyptian Spatial Data Infrastructure (ESDI) Towards the Electronic Government (E-gov.). Spatial Portals and e-Government. *From Pharaohs to Geoinformatics, FIG Working Week 2005 and GSDI-8* Cairo, Egypt.
- STRADER, T., TARASEWICH, P. & NICKERSON, R. C. (2004) The State of Wireless Information Systems and Mobile Commerce Research *Information Systems and e-Business Management*, 2, 287-292,.
- TOMLIN, C. D. (1990) *Geographic information systems and cartographic modeling*, S. 1.]: Englewood Cliffs.
- WANGLER, B. & BACKLUND, A. (2005) Information Systems Engineering: What is it? *CAiSE Workshops*, 2, 427-437.
- WEBER, R. (2004) The Rhetoric of Positivism versus Interpretivism. *MIS Quarterly*, 28, iii-xii.
- ZHAO, Y. & MOTOROLA, I. (2002) Standardization of mobile phone positioning for 3G systems. *Communications Magazine, IEEE*, 40, 108-116.
- ZIEGLER, P. & DITTRICH, K. R. (2004) Three decades of data integration - all problems solved? *Database Technology Research Group*. Winterthurerstrasse 190, CH-8057 Zurich, Switzerland, Department of Informatics, university of Zurich.

PART 3



ICT Sustainable Development

17

The Role of ICT in Creation and Sustenance of Trust in SME Clusters and Inter-organizational Systems

*Agnes Mindila, Dorothy McCormick,
Anthony Rodrigues and Ronald Waweru Mwangi*

Research has shown that it is through clustering and interorganizational relationships that SMEs can be innovative, therefore creating development opportunities. Literature stresses that the success of these clusters and inter-organizational systems can be traced to the system of social and economic relationships based on trust among other factors. It is therefore a great concern for researchers to define processes that would enhance and contribute to trust. On the other hand research has shown that ICT and networked computing provide geographical clusters and interorganizational systems with new development opportunities, drawing new possible trajectories of evolution. One specific area where ICT provides new development opportunities is by reinforcing intra-cluster and inter-organizational relationships. Many social scientists have been skeptical about the contribution of ICT towards trust arguing that trust is only built where people have face-face communication, have demographic similarity, have prior history of relationship and are co-located. However it has been noted that Inter-organizational systems exist that are enabled by ICT, they are located at different physical locations and have no prior relationships. These in itself creates a gap for researchers to be able to explain the existence of such relationships. By considering different dimensions of trust this paper posits that ICT contributes to creation and sustenance of trust, which is a key component for the success of inter-organizational systems. The paper seeks to provide a framework upon which ICT is analytically evaluated for its contribution towards creation and sustenance of trust.

1. Introduction

Economic development literature identifies innovation systems as a key vehicle that brings about development. This is through its contribution to economic growth by technological change and by contributing to institutional change by way of providing new institutional frameworks and practices. Innovation no longer depends on how firms perform on individual basis but on how they work in unison, hence the emergence of SME clusters and the concepts of inter-organizational relationships (IOR) and inter-organizational systems (IOS). Literature stresses that the success of these clusters, IOR and IOS can be traced to the system of social and economic relationships based on trust.

On the other hand, research has shown that ICT and networked computing provide geographical clusters, IOS and IOR with new development opportunities

drawing new possible trajectories of evolution (Carbonara, 2005). Besides it reinforces intra-cluster and inter-organizational relationships. This then means that establishment of inter-organizational level trust is a big concern for researchers in ICT and IOS alike. A stream of research on trust goes on in social science and on the other hand a stream of research on trust goes on in ICT. They seem to be parallel and not synergizing. This paper attempts to synergize these two streams of research and provide a common ground. It's important to investigate the contribution of ICT towards trust, which is key in the success of these inter-organizational systems. The rest of the paper is organized as follows:

Section 2 of the paper introduces the concept of SME clusters and IOS. Section 3 introduces trust and its dimensions. Section 4 suggests a conceptual framework with which ICT can be analyzed for its role in contributing towards trust creation and sustenance. Lastly section 5 makes conclusions and recommendations on future developments and research.

2. SME Clusters and IOS

Emerging evidence suggests that use of collaborative inter-organizational relationships is an important source of innovation (Marshall, 2004). SMEs can afford to continually be innovative, not because they have resources to invest in research and development but because of collaborations and inter-firm networks. Innovativeness in this context is looked at in two ways, one, innovativeness that enables SMEs to come up with new products and processes hence increasing productivity and contributing to economic growth. Secondly innovativeness that enables institutional entrepreneurs to be strategically positioned and engage in bricolage, which eventually results in institutional change (Campbell, 2004). It is then the unique combination of economic growth and institutional change that brings about economic development.

Both kinds of innovativeness are key towards economic development and are all dependent on clusters and inter-organizational systems for the case of the SMEs. Marshall, 2004 emphasizes that more enduring relations entail obligation and trust and that economic action is embedded in a rich structure of ongoing networks of relationships. Networks, partnerships and consortiums succeed in part through the social glue that holds them together, the glue entails norms of trust and appropriate behavior (Reich and Kaarst-Brown, 2003). Literature on IOS has emphasized trust to be an important factor in the development and success of IOS (Ibrahim and Ribbers, 2006).

Currently experts disagree about whether face-face interaction is the only means for formation of trust and collaboration. Social science, policy and public management research is silent on the question of what would happen if potential network actors could be digitally linked (Fountain, 2001), not just with networks but with information systems that carry out duties of control, monitoring and enforcement. Technology researchers celebrate the ability of ICT to make distance and time constraints virtually meaningless. They also argue that easier

communication, enforcement, control and co-ordination lead to enhanced trust hence enhanced inter-organizational relationships. Speaking on technology, Bellussi, 2004 states “if ICT adoption and implementation are aimed at reinforcing intra-cluster and inter-organizational relationships then they contribute towards integration of the economic actors operating along global supply chains”. This integration is only possible if there is trust. How to measure the contribution of ICT on networks and social capital of which trust is part, has been the greatest challenge for researchers (Hurriagotia, 2004).

It’s the importance of trust in these SME clusters and IOS that leads to section 3, which aims at examining trust and its different dimensions with the aim of finding the role of ICT in any of the dimensions available.

3. Trust and Its Dimensions

Gefen D, 2002 argues that trust itself is a multidimensional construct. Each of these dimensions is important in the achievement of the desired behavior. Many researchers, (Gefen D., 2002, Areniurs and Antro, 2002, Erden and Ozen, 2003) have classified trust into two major categories each with its dimensions.

- (i) Institutional trust: Situational Normality, Structural assurances and facilitating conditions.
- (ii) Inter-organizational trust: perceived ability, perceived integrity and perceived benevolence.

Institutional trust was suggested by Zucker, 1986 when he suggested that institutional trust is the most important mode by which trust is created in an impersonal economic environment where the sense of a community with common values is lacking. This is a perspective that has been widely adopted by information systems researchers perhaps because ICT brings together organizations with no familiarity and similarity.

Institutional trust has been defined by various researchers (Mcknight et al, 2002) as “the organizations belief that favorable conditions are in place that are beneficial to outcome success”.

Institutional trust is defined to have three dimensions, which are structural assurances, situational normality and facilitating conditions. Structural assurances are beliefs that favorable outcomes are likely because of contextual structures such as contracts, regulations and guarantees whereas situational normality beliefs that success is anticipated because the situation is normal (Pavlou P.A., Tan Y.H., and Gefen D., 2003). Whereas facilitating conditions refer to common beliefs, goals and standards.

Interorganizational trust on the other hand has been defined as “one organizations belief that the other party in the exchange relationship will behave in accordance with the trustors confident expectations” (Pavlou P.A., Tan Y.H., and Gefen D., 2003). Four dimensions of interorganizational trust have been identified namely:

- Competence
- Credibility

- Benevolence
- Openness

Competence refers to the capacity to fulfill promises and contracts. Credibility refers to predictability, reliability and honesty towards fulfilling contractual obligations. Benevolence refers to the expectation that a party will act fairly and will not take unfair advantage of the trustor, even given chance (Ridings C.M., Gefen D., and Arinze B., 2002, Stoutland S.E., 2001). Openness refers to the perception of honesty of communications and completeness of information (Ibrahim M. and Ribbers P., 2006).

It is important at this point to note that while trust plays a major role in the inter-organizational systems supported by networked computing and ICT environment, it seldom complies with the traditional view of dyadic trust that the academic literature in social sciences focuses on but rather the institutional trust.

Many researchers like Pavlou P.A., 2002 and (Pavlou P.A., Tan Y.H., and Gefen D., 2003) have shown that institution based trust positively influences inter-organizational trust which positively influences inter-organizational relationship continuity whether in a digital environment or not.

This paper posits that even though ICT may have less or may not have a direct role in creating and sustaining Inter-organizational trust per se, it surely has a direct influence on institutional trust which positively influences inter-organizational trust. This then leads to section 4, which conceptualizes the 'how' of this hypothesis.

4. Conceptual Framework of the Role of ICT in Contributing to Trust in Inter-organizational Systems.

Several researchers have contributed to the work on institutional trust and how it relates to inter-organizational trust. This section will examine some of these works and draw important resolutions.

Hung Y.C et al, 2004 posit that individuals have three possible routes by which to form trust namely peripheral route, central route or habitual route. That the peripheral route is used due to reasons such as limited personal knowledge, time constraints or lack of traditional social cues. In such a situation individuals are not able to engage deliberate cognitive information processing and have to rely on peripheral cues available in context. The peripheral cues that form initial trust include dispositional trust, third party as conduits of trust, category based trust, role based trust and rule based trust.

The central route is achieved through interactions and accumulated personal knowledge. Personal attributes are deliberately assessed. The accumulated knowledge becomes a basis for initially calibrating and then upgrading trust related expectations.

Habitual route is where the accumulated personal knowledge based on prior successful trust transactions contribute to a habitual trust attitude.

Pavlou, 2002, states that institutional trust positively contributes to interorganizational trust. Where he defines institutional trust as obtained through Structural assurances, Situational normality and facilitating conditions.

The commonness in the research done on institutional trust is the fact that it can be achieved through third party or bilateral relationships between various firms. This is done through three major ways namely:

- Structural assurances
- Situational normality
- Facilitating conditions.

Another way mentioned by Hung et al is category based.

This section sheds light on how ICT can be used in inter-organization systems to foster structural assurances, situational normality, facilitating conditions and category based trust.

4.1 Situational Normality

Situational normality is where it is believed that success is anticipated because the situation is normal. It is suggested that this situation is achieved when the following factors are guaranteed.

- Authorization
- Integrity
- Non-repudiation
- Confidentiality.

The factors above can be achieved by way of various technologies and information systems. By employing technologies such as encryption where special mathematical algorithms are used to transform digital data into a scrambled code before they are transmitted and decode the data when received and firewalls authorization and confidentiality can be achieved in IOS.

Development of Information systems controls will ensure integrity in IOS by way of ensuring proper data entry, processing techniques, storage methods and information output. Input controls would include encryption, validation procedures, error signals and error corrections among others. Processing controls will include programmed edit checks and controlled totals. Storage controls include logs and time stamps, secured databases, disaster recovery capabilities, backup files, security codes and proper database management procedures. Output controls would include control totals, report distribution lists, encryption, control totals and end user feed back. This information system controls will ensure integrity in IOS.

Information system audit trails would contribute greatly to non-repudiation in IOS. These trails provide a way of documenting all transactions done from the first step to the very last step. They help auditors check for errors or fraud.

Cyber forensic, sniffer software and investigative services also provide integrity within IOS.

By use of the above information technologies and systems IOS can achieve situational normality in the sense that authorization, integrity, non-repudiation and confidentiality is attained.

4.2. Structural Assurances.

Structural assurances are beliefs that favorable outcomes are likely because of contextual structures such as contracts, regulations and guarantees. This is especially when looking at IOS as sources of technological change, which is one aspect that brings about economic development. This paper would want to take the other perspective of economic development occasioned by institutional change. This then implies that learning of new ideas becomes a key issue or what we call brocolage. The structural assurances required for this kind of set up are those that intentionally foster knowledge transfer between organizations. The structures put in place should be such that the ability to extract knowledge from individuals within the organization and to structure the knowledge foe use by partners is made easier. It also means assurance that the actual transfer of the knowledge will occur and that the partners will be able to absorb the knowledge provided by other organizations. Two key issues become very important when addressing structural assurances that will achieve the above description. These are:

- Transparency-level of openness of a firm to its partner and the opportunity it provides to its partner to learn
- Receptivity-the capacity or ability of the collaborating partners to absorb the knowledge provided by another. (Davidson M. and Olfman L., 2004)

This implies that collaborating organizations should have high levels of transparency and receptivity. In order to achieve this organizations have to make a deliberate effort to set up structures that will help them achieve these two important factors. This implies creating structural assurances that help them achieve transparency and receptivity.

ICT provides a fertile ground for easily achieving this two factors. By employing knowledge management systems that embeds techniques, technologies and reward systems for getting individuals share what they know and to make use of the accumulated IOS knowledge. By providing accessibility to people, documentation and organization partners through groupware software, e-mail, telephone, shared databases, instant messaging organizations increase their level of transparency.

By providing help desk systems and web portals firms are provided with the opportunity to learn by their partners. An organization that deliberately puts the above technologies and information systems in place provides structural assurances which is assurance that actual transfer of knowledge will occur.

4.3. Facilitating Conditions

This refers to shared standards, common beliefs about behavior and goals, and relationship values. We focus on shared standards. Organizations can achieve shared standards by considering adopting similar standards in terms of data

management. Organizations can also ensure shared standards by supporting interoperable Information technology platforms.

4.4. Category Based Membership

Membership in a certain social or organizational category is said to contribute to institutional trust. With the invention of social software like blogs, individuals across partnering firms can be able to share information in their area of interest.

The framework in Figure 1 summarizes the contribution of ICT to institutional trust.

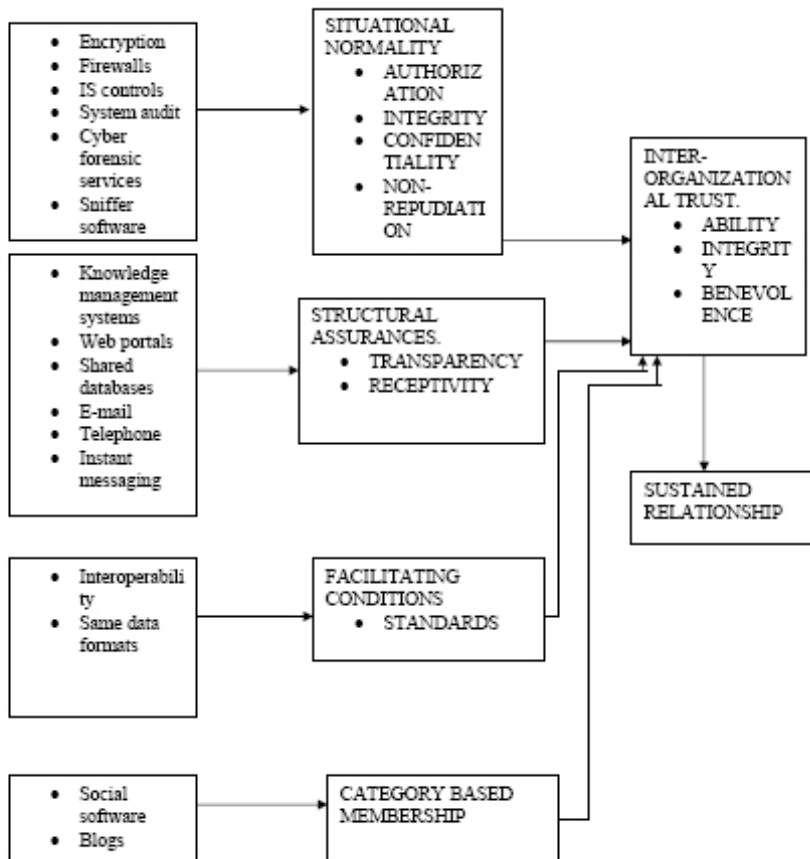
5.0. Conclusion and Further Research

Trust affects performance and it is critical in organization co-operation, coordination and control. As Information system researchers we are faced with the challenge of developing new models for creation and maintenance of trust better suited to the dispensation of new ICTS'. This paper has focused on creation of trust for the sake of causing brocolage to occur. This brocolage is what brings about institutional change that is a key ingredient to economic development in SMES'.

Two dimensions of trust have been identified i.e. institutional trust and inter-organizational trust. The paper dwelt on institutional trust and illustrated how ICTS' can be used to contribute to structural assurances, situational normality, facilitating conditions and membership category, which are key ingredients that make up institutional trust.

Other researchers have done work on how institutional trust leads to inter-organizational trust. Further research on this work will endeavor to look at how ICTS' directly contributes to inter-organizational trust. This will be a leap forward in trying to inform managers of information systems on how strategically ICTS' can be utilized to foster trust in IOS.

Fig. 1: Framework: contribution of ICT to institutional trust



References

- ARENIURS P. AND ANTRO E. (2002) *International Social capital and its effects on resource acquisition in born global firms*. Working paper series. 91-wp-2002-004.
- BELLUSSI F., (2002) *Industrial districts: networks without technologies?* Proceedings of the DRUD summer conference “Industrial Dynamics of the new and old economy-who is embracing whom?”
- CAMPELL J.K. (2004) *Institutional change and globalization* . Princeton: Princeton University Press.
- CARBONARA N., (2005) *Information and Communication Technology and Geographical clusters: Opportunities and Spread*. Technovation 25 pp 213-222
- DAVIDSON M. AND OLDFMAN L. (2004) *The Impact of Information Technology use on interorganizational Learning in an IT Outsourcing Collaboration*. Proceedings of the 37th Hawaii International Conference on System sciences
- ERDEN F. AND OZEN J. (2003) *Cognitive and affective dimensions of trust in developing team performance*. Team Performance Management Vol 9 Number 5-6. pp131-135.

- FOUNTAIN J.E., (2001). *Building the virtual state Information Technology and Institutional Change*. Brookings Institution press. Washington D.C
- GEFEN D. (2002) *Reflections on the dimensions of trust and trustworthiness among online consumers*. ACM vol 33 issue 3. pp38-53.
- HUNG Y.C., DENNIS A.R AND ROBERT L. (2004) *Trust in Virtual Teams:Towards an Intergrative Model of Trust Formation*. Proceedings of the 37th Hawaii International Conference on System sciences
- HURRIAGAGOITIA J.M.Z., (2004) *Analysis and measurement of interactions in European innovation systems. A sectoral approach*. Economics of Technological and institutional change. Doctoral training program ETIC 2004. Maastricht, October 11th-24th 2004.
- IBRAHIM M. AND RIBBERS P. (2006) *Trust, Dependence And Global Interorganizational Systems*. Proceedings of the 37th Hawaii International Conference on System sciences
- MARSHALL C., (2004). *The dynamic nature of innovation partnering: a longitudinal study of collaborative interorganizational relationships*. European Journal of innovation Management. Vol 7 No.2 pp128-140
- MCKNIGHT ET AL. (2002) *Developing and Validating trust measures foe e-commerce: An integrative typology*. Information Systems Research. 13 (3) pp 334-359
- O'BRIEN J. (6th ed) (2004)*Management Information Systems. Managing Technology in the Business Enterprise*. New york: McGraw Hill
- PAVLOU P.A. (2002) *Institution based trust in interorganizational exchange relationship. The role of online B2B market places on trust formation*. Journal of strategic information systems 11. pp215-243
- PAVLOU P.A., TAN Y.H AND GEFEN D. (2003) *The transitional Role of Institutional Trust in Online Interorganizational Relationships*. Proceedings of the 37th Hawaii International Conference on System sciences
- RIDINGS C.M., GEFEN D. AND ARINZE B., (2002) *Some antecedents and effects of trust in virtual communities*. Journal of strategic Information systems 11. pp271-295.
- STOUTLAND S.E., (2001) *The multiple Dimensions of Trust in resident/Police relations in Boston*. Journal of Research in Crime and Delinquency vol 38. No.3 pp 226-256.

18

Data Security Lapses in Developed Societies: Public and Private Sector Cases

Anthony J Rodrigues

The loss of twenty five million (25) child-benefit records was disclosed on 20th November 2007 to the House of Commons. The scale of the security lapse at the Treasury's tax agency, affected almost half the British population. Two computer discs containing the names, addresses and dates of birth of nine and a half (9.5) million parents and fifteen and a half (15.5) million children, together with their national-insurance and child-benefit numbers and bank-account details, went missing. The records would be a treasure trove for identity fraudsters. Such large scale institutional data loss is not an isolated incident. Whereas personal data loss by or stolen from credit-card companies, online retailers, government departments and banks is well known, there now appears to be no point in fraudsters stealing one person's details when records can be had in their thousands and millions from processes and computer systems that are not secure. Moreover such theft is not limited to public sector or the UK. The policy implications and regulatory issues and possible sanctions are discussed.

1. Background

The UK government, Chancellor of the Exchequer, told parliament on 20th November 2007 that two computer discs containing the personal details of 25 million British individuals and 7 million families had gone missing. A junior official at HM Revenue and Customs (HMRC), the tax-collection department which also pays out benefits to families, had breached security guidelines by putting the discs in the internal-government post run by TNT, a private courier. It was only when they failed to arrive at their destination, the National Audit Office (NAO), Parliament's spending watchdog that the security breach came to light. The Chancellor of the Exchequer was informed of the loss on 10th November 2007; the public heard about it ten days later.

The discs included names, addresses, bank-account details, dates of birth and names of spouses and children. The information on the discs was protected by a password but was not encrypted; the mail had not even been registered; and the NAO did not want the addresses and bank-account details anyway. While the fate of the discs is unknown, they contain just the sort of information sought after by identity thieves, who could use it to procure fake documents, commit fraud and empty bank accounts.

This is the latest in a series of such losses, Economist 22nd November 2007:

- o In September 2007 HMRC, lost a laptop containing personal data on 400 people, and

- o In October 2007 it lost another disk in the post, containing pension records for 15,000 people.

The lost CDs rank alongside the earlier theft of US public sector data in :

- o 2005 from the Bank of America of tapes containing information on 1million American government employees and
- o 2006 from the home of an employee of the Department of Veterans Affairs in America on 26.5 million people.

There have been dozens of smaller cases around the world in which personal data have been lost by, or stolen from, credit-card companies, online retailers, government departments and banks. Fear of identity fraud means that many people now routinely shred receipts and sensitive documents before binning them. However, there is no point in stealing one person's details, when records can be had in their thousands and millions from processes and computer systems that not secure. Large databases have become central to the operation of governments, health systems, banks and other large companies. While nobody really wants to turn back the clock and revert to paper records what can be done?

2. Policy Implications

With regard to computer security, there are two issues:

- o security depends on a combination of technology and policy; and
- o no system is ever totally secure.

It is safer to assume that there will be breaches, and work out how to minimise the damage. That means:

- o storing, and moving around, as little data as possible;
- o anonymising records, linking to personal details stored in a separate database; and
- o using encryption to protect data in transit.

None of this was done by the HMRC. It was asked to supply anonymised data to the National Audit Office, but provided reams of unencrypted personal information instead.

Regulation has a role to play, too. Many European countries and 35 American states have rules that require companies and government departments to disclose breaches of information security to anyone affected. In many cases they are also legally liable for any loss. This gives them an incentive to store as little data as possible and to look after it properly. Britain, alas, has some of the most toothless data-protection rules in the developed world: the government recently rejected a plan to make reporting of breaches compulsory. According to one estimate, setting up new bank accounts for everyone affected by the lost benefits records could cost around £300m (\$600m). Data debased even when sensible regulation is in place is, however, of no use if the rules are ignored, as in the British lost benefits records case. Data-protection watchdogs need to be able to carry out spot checks

to ensure that everything is being done by the book. But where other regulators, such as European antitrust investigators, have the power to carry out unexpected raids, Britain's data-protection watchdog does not. The government has pledged to change this, but as things stand, all it can do is issue HMRC with a warning not to make the same mistakes again.

Large databases have their uses, doing away with paperwork and speeding things up. But the centralisation of so much data also has drawbacks. In its enthusiasm for huge technology projects, such as its plan for a national identity card see section 6, the British government has failed to take such dangers sufficiently seriously. For example departments face no penalties for ignoring procedures and losing data. If organisations were confronted by the risks of building large databases and forced to balance them against the benefits they provide, they would design them differently and monitor their use more closely. Sometimes, they would decide that they were not worth building at all.

3. Impact

The Opposition allege that cost-saving imposed by senior officials lay behind the problem. The prime minister's personal responsibility may be involved if it can be shown that HMRC's failings are linked to an administrative merger and job cuts that he pushed through as chancellor. The lost benefit records have already caused the resignation of the chairman of HMRC. The Chancellor will be next in line if the security breach leads to actual fraud. But the chief loser is likely to be the Prime Minister.

Careless use of personal data highlighted the malaise at the heart of the UK government's IT infrastructure. Will it change the internal culture? The loss of two CDs has effected the government's £14bn-a-year IT programme, Cross (2007):

- o The NHS IT agency Connecting for Health warned hospitals not to post discs containing unencrypted personal data to the central NHS Tracing Service, run by a private contractor in the Midlands. Media not meeting security standards "will be destroyed upon receipt", it warned, and
- o Ministers announced a five-month delay to ContactPoint, a database with details about every child in the UK.

Data sharing between departments about individuals can have benefits; what is needed is a culture within government where both the power and the responsibility for implementing those benefits is understood throughout. The first priority for IT chiefs was to report to the prime minister on the analysis of systems and procedures by 10th December 2007. The reports would feed in to a review by the government's intelligence chief. Yet all these efforts make one big assumption: that as long as "systems and procedures" are properly followed, everything can continue as before.

An interim report specifically into the HMRC see Poynter 2007, Chairman of PricewaterhouseCoopers, published on 17 December 2007 by the Chancellor, *inter alia*, criticised HMRC's complex management structure, saying "roles and

responsibilities do not relate to accountability”. He also admitted that a letter from HMRC apologising to families over the loss of child benefit had itself contained too much personal information, some of which was liable to have gone astray. In addition he promised to bring in tougher laws to prosecute the willful misuse of data, so those found guilty may be imprisoned rather simply fined, as at present.

4. Systemic Failure?

The UK government on 17th December 2007 revealed another disturbing loss of citizens’ data when the Department for Transport (DfT) Secretary disclosed that a computer disc had gone missing that contained the records of more than 3 million driving test applicants, including their names, home addresses, email addresses and telephone numbers. The lost data covers every applicant for a driving theory test between September 2004 and April 2007. The disc was lost from a “secure store” in Iowa in the US midwest by private firm Pearson Driving Assessments in May 2007, but this was not relayed either to Westminster or to the police. The DfT Secretary who had only been informed in November 2007 apologised for the latest data embarrassment, but stressed the missing disc did not contain personal bank account, national insurance or credit card numbers. Whitehall officials argued that most of the data is available in telephone directories, Wintour (2007).

Nevertheless, the loss adds to the challenge the government faces in persuading the public that the benefits of data-sharing in the government outweigh the risks. This incident emerged in a trawl of government procedures by the cabinet secretary set up in the wake of the loss of 25 million child benefit records by HMRC reported in November 2007. An interim report from the cabinet secretary published on 17 December 2007 revealed that four government departments - the Treasury, the Department for Work and Pensions, the Ministry of Defence and the Crown Prosecution Service - had felt forced to tighten data handling procedures in the wake of the child benefit fiasco. The DfT Secretary criticised Pearson Driving Assessments, a contractor to the Driving Standards Agency, for failing to inform the police of the disc’s loss and disclosed that the drive had been formatted to fit Pearson configurations and as such was not readily usable or accessible by third parties. The Information Commissioner on learning of the loss advised the DfT Secretary that there was no need to send notification to individuals, Wintour 2007.

The opposition in Parliament attacked ministers, pointing out that only a year ago they had promised to get a grip on data security in their department. “This latest debacle shows a systemic failure within the government’s data protection controls. The importance of confidentiality for the DfT’s agencies is particularly acute given the value of vehicle and driver details to car criminals and rogue drivers.” Another opposition party parliamentarian said “If data and valuable information is consistently lost or stolen or abused, the public completely lose confidence in government ... it is very difficult to see how we can be confident of the government proceeding with much more ambitious initiatives, not just the

compulsory ID cards scheme, but the DNA database and the NHS spine.” The shadow justice secretary also questioned why the information had ever been sent overseas.

As if on cue, Watt 2008, the UK government on 19th February 2008 faced fresh embarrassment over lost personal data, after admitting that a disc containing the DNA details of thousands of suspected foreign criminals was mislaid for a year. Whitehall sources said the disc, containing the details of 2,000 suspects wanted by the Dutch authorities, lay on a desk at the Crown Prosecution Service (CPS) for a year after being “ignored” by officials. The CPS insisted there had been no breach of security, because the disc never left its premises after being sent to Britain by the Dutch authorities in January 2007. But the government was accused by the opposition of endangering the British public because eleven (11) of the suspects allegedly committed crimes in Britain, including serious assault and non-payment of fines, before the CPS looked at the information. The Dutch authorities sent DNA details of 2,000 people believed to have been involved in unsolved crimes, asking the CPS to match the data against British DNA records. Whitehall sources admitted that the disc lay unnoticed on the desk of a CPS official until last month. “It was just ignored, it sat on someone’s desk,” one source said. “It did not get lost. It did not leave the building. It was not a data security issue”.

The checks were carried out last month when the disc was spotted, and found that the records of fifteen (15) of the Dutch suspects were on the British database. The shadow home secretary, said: “It beggars belief that after all the disasters that have happened in the government’s handling of data it can still lose a disc with the details of serious criminals who are a danger to the British public. This shows it is incapable of learning the lessons of how to handle vital data.” In 2007 HMRC lost two discs containing the details of half the population as discussed above.

The CPS was asked by the attorney general’s office to begin an inquiry into why a disc holding DNA details of thousands of suspected foreign criminals was left unchecked for a year, the Prime Minister told the Commons. In a heated exchange the Conservative leader accused the government of “catastrophic failures” for putting the safety of the British public at risk. “Why is it that this government is so incompetent when it comes to protecting information from criminals?” referring to recent scandals over lost data that have dogged the government. Earlier the Liberal Democrat leader, said on the BBC that “very grave” implications and would deal “yet another hammer blow” to public confidence in the government’s handling of information. “Of course the responsibility for this lies with the home secretary, but I think there is a more systemic political failure of this government to get to grips with the database culture it has created,” he said. “It has created a database state which it doesn’t seem to know how to handle in human terms.”

This latest embarrassment for Whitehall came after a series of blunders involving missing data. Apart from the most high-profile loss of two computer discs by HM Revenue and Customs containing 25 million people’s child benefit details equivalent to almost half the population in 2007 discussed earlier in section

1 above, six thousand (6000) personal details were lost in January 2008 after a Royal Navy laptop was stolen and the defence secretary admitted there have been two more similar thefts since 2005.

5. Internal Culture

The government appears to be overwhelmed by the scale of what it is trying to do with IT. “It’s indicative of a lack of expertise,” said Helen Margetts, professor of society and the internet at Oxford Internet Institute and the co-author of a study that is devastatingly critical of the government’s IT programme, see Cross 2007. The study of IT projects in seven leading countries found that governments that place big IT contracts in the hands of a few big contractors are the ones most likely to experience failures see Dunleavy, P., et al. (2006).

The UK was unique in the extent to which it outsourced projects so that large IT companies had the government over a barrel. The study found that the UK had “the most concentrated government IT market in the world, with:

- o a near-monopolistic lead supplier (Electronic Data Systems, or EDS),
- o huge contract sizes,
- o poorly understood use of private finance initiative (PFI) contracts for inappropriate IT projects, and
- o virtually no in-house capacity to manage (let alone develop) IT systems.

IT contractors “drove a coach and horses” through budgets, the report alleges. “It became expected practice to pitch prices for initially completed tranches of work relatively low, in the confident expectation that later revisions and extensions would create negotiated contracts of between four and six times the initial competed contract price.” It’s only a short step from there to demanding £5,000 to “strip” sensitive personal information from the child benefit data - a task that, with a properly designed database, would take no extra effort. That HMRC is charged so much extra for elementary tweaks shows how weak the government’s grasp of its own IT has become.

The subsequent loss of the discs then exposes a culture - not just a one-off error - where unencrypted personal data is regularly sent back and forth between public bodies on physical media, rather than via secure networks. “They just weren’t considered important,” Margetts was reported to have said, see Cross 2007. In theory, the government has been trying to raise its game for more than two years.

6. Transformational Government Strategy

The three central aims of the Transformational Government Strategy, see Cabinet Office 2005 are first:

- o create a new “IT profession in government” which is to be achieved in part by hiring people with IT qualifications for the civil service fast stream, where they can expect to rise to the top.

The other two strands of Transformational Government create a radically new IT infrastructure, based on public bodies that :

- o share systems, and the same institutions
- o share data on every individual in the country.

In December 2006, the former chairman of HMRC and current Prime Minister's adviser on transforming public services published further radical plans, Varney Report 2006. They are based on a passionate commitment to sharing data and systems. It was pointed out that 'today, each citizen has at least five widely used identity numbers. In future, one accurate and robust registration should underpin all services. There is absolutely no objection to public services sharing basic administrative information like names, numbers, addresses and dates of birth.' This vision of transformed government meshes with a barrage of IT-based initiatives aimed at tackling specific political priorities which include:

- o ID card, which will be enabled by linking at least three existing identity databases;
- o ContactPoint, the newly renamed index of information relating to every child in the country - subsequently delayed for five months after the lost benefits records case;
- o e-Borders, the immigration system supposed to give immigration officers and airline check-in staff overseas the ability to check passengers' credentials against government records;
- o NHS care records service, creating a shared electronic health record for every individual in England (Scotland and Wales have parallel schemes).

But there have been repeated warnings that the schemes are not feasible. Professor Ross Anderson, a security expert at Cambridge University, suggested that the proposed children's database was intrusive and possibly illegal; the government brushed it off - as it did with warnings about internet security and the centralisation of personal health records as reported in Cross 2007. The two lost discs have put those dismissals under fresh scrutiny.

Despite these warnings the home secretary on the 6th March 2008 gave a broad indication that millions of public-sector teachers, nurses and care staff in "positions of trust" were to be given a fast-track incentive to sign up early for an ID card to get the "critical mass" needed for the controversial scheme, as reported in Guardian 7th May 2008. She promised to speed up the introduction of the scheme by allowing people to get a card without waiting to renew their passport, enabling its full introduction by 2017, two years ahead of previous plans. She also claimed that the next official cost estimate for the scheme, to be published in May, would confirm that the £5.4bn price tag has been reduced by £1bn. The revised timetable confirms that groups are being targeted for early registration on the national identity database - the key part of the scheme - to secure its popular acceptance.

ID cards become compulsory in 2008 for foreign nationals from outside Europe living in Britain, and in 2009 for 200,000 airport workers and Olympic security

staff. Incentives will encourage early voluntary participation from 2010 for young people, including students, and millions of public sector staff in “positions of trust” currently subject to criminal record vetting. Such staff that “chose to have an identity card” would be able to use it to fast-track checks on their status as part of their job. A trial by the Criminal Records Bureau and the Identity and Passport Service has shown criminal record checks could be cut from four weeks to four days.

Young people who registered for an ID card would find it easier to enroll on a course, apply for a student loan or open a bank account. The scheme’s critics said the revised timetable also put back by two years, from late 2009 to 2011, the date from which the bulk of the population will have their personal “biometric” - electronic fingerprint - details automatically entered on the national identity database when they renew their passport. This is likely to come after MPs vote after the next general election on whether the scheme is to be made compulsory. People will then be given a choice.

An identity card alone will cost about £30, while combining it with a passport will cost at least £93. Attempts to reduce the eventual cost of signing up to the ID card register are ongoing. A government poll, published on 6th March 2008, claimed 59% support but that included a third of voters who believe the cards will be free. The home secretary acknowledged that the ID cards debate evoked strong feelings but said that rather than thinking of the state as an opponent of civil liberties, its role was to “defend our interests, to offer reassurance and trust, and to work in the most effective way possible to ease and to enable our lives”. But the shadow home secretary said that the ID card scheme was now being implemented by spin and stealth. “The home secretary’s claim that she is offering people a choice is misleading - identity cards will still be compulsory. The government has lost the argument. They have lost 25 million personal records. And they have lost the public’s trust. It’s time they faced up to these stark truths - and ditched ID cards for good.” Others said it was a marketing exercise: “Whether you volunteer or are coerced on to the ID database, there’s no way back. You’ll be monitored for life. That is why the government is targeting students and young people, to get them on before they realise what is happening” reports *The Guardian* 7th March 2008.

The social need to ensure public health and safety, national security and crime control, economic prosperity and environmental sustainability, will all guarantee the presence of governments and their active involvement in its citizen’s lives, visibly or “behind the scenes”. What the public do expect however, is improved efficiency in the results obtained from expenditures incurred. The result of a better-educated public is that, in terms of suggestions for improving government (other than exclusively by think tanks drawn from Universities) more and more people feel entitled and empowered to express their views. Governments that therefore rule with the consent of their publics cannot afford to ignore or dismiss their constituents’ proposals. Some critics claim that what occurs may be termed

“impression management” wherein a show is made of being technology savvy but only minor adjustments are made in the internal institutional IT and support culture.

Nevertheless, the public does demand more (services) for less (taxes), and the possibility of shifting party allegiances exist if citizens are unhappy with current government policy. Given this reality, electronic productivity seems to be by far the best way for governments to achieve the trade-offs between increasing expectations and diminishing resources provided that effective safeguards are in place. Nowadays electronic technologies play an increasingly large role in shaping the mind-set of citizens, and it is natural that they will want that mind-set reflected in social governance.

7. Private Sector Immune?

To dwell only on public sector cases would be misleading as the private sector also has its problems with data as the following two cases show. Norwich Union Life an insurance company in the UK was fined £1.26m for failing to protect customers’ personal details after fraudsters were able to steal £3.3m from policyholders, by the Financial Services Authority (FSA) reported Osbourne 2007. The FSA said the insurer had failed to take reasonable care to look after policyholders’ details and protect itself and them from the risk of paying out policies to fraudsters. The failures were highlighted by a number of frauds carried out against customers in mid-2006, when criminals used publicly available data such as names and addresses and dates of birth to impersonate Norwich Union customers. In a series of telephone calls to Norwich Union Life the fraudsters obtained confidential information and were able to change customers’ details so that policies were paid out to the wrong accounts. More than 632 policies were targeted, and 74, worth a total of around £3.3m, were surrendered to fraudsters. Nine of those surrendered belonged to directors of Norwich Union’s parent company Aviva. The FSA said that although the other policyholders had not had their money stolen, Norwich Union’s call centre staff had disclosed confidential information about the customers, in some cases divulging their full bank account details. As a result, it said the company’s 6.8 million customers had been exposed to “an additional, ongoing and unquantifiable risk of identity theft”.

During an investigation the watchdog found the insurer had failed to properly assess the risks posed to its business by financial crime, and by not doing so had left its customers more likely to fall victim to crimes such as identity theft. The FSA’s director of enforcement said: “Norwich Union Life let down its customers by not taking reasonable steps to keep their personal and financial information safe and secure.” It is vital that firms have robust systems and controls in place to make sure that customers’ details do not fall into the wrong hands. Firms must also frequently review their controls to tackle the growing threat of identity theft. This fine is a clear message that the FSA takes information security seriously and requires that firms do so too.

When setting the fine, the FSA said it had taken into account measures taken by the insurer to mitigate the seriousness of its failing, including reinstating the value of stolen policies and appointing a third party to review its caller identification procedures. However, it said it still deserved a significant fine because it had not taken the same action to inform and protect victims who were not members of the company as it had to protect those working within Aviva, and had posed a significant risk to the watchdog's objective of reducing financial crime. The chief executive of Norwich Union Life, said: "We are sorry that this situation arose and apologised to the affected customers when this happened. We have extensive procedures in place to protect our customers, but in this instance weaknesses were exploited and we were the target of organised fraud. Whilst the number of customers affected is very small compared to the number of policies we manage overall, any breach in customer confidentiality is clearly unacceptable." He added that steps had been taken to improve the company's security systems.

This is not the first time the FSA has fined a financial service provider for failing to protect its customers' personal information. In the past two years BNPP Private bank, Capita Financial Administrators and Nationwide building society have all been hit with fines, with Nationwide paying £980,000 after the theft of a laptop carrying account details.

More recently on 8th April 2008 it was reported see Bowers 2008, that HSBC was facing the possibility of a regulatory inquiry and heavy fines after it was forced yesterday to confirm it had lost an unencrypted disk containing life insurance policy details for 370,000 customers. Britain's biggest bank confirmed a newspaper report revealing the disk had disappeared in the post more than seven weeks ago. Customers have not been told their details have gone astray and questions have been raised about the speed and thoroughness of HSBC's investigation into the data loss.

HSBC said the disk contained names, dates of birth and policy numbers as well as the amount of cover taken out and smoking habits of policyholders. A spokesman stressed that no other information, such as personal bank account details or medical details, was on the disk. It was password-protected but not encrypted. "There is nothing else that could in any way compromise a customer and there is no reason to suppose that the disk has fallen into the wrong hands," HSBC said in a statement. "Nonetheless, HSBC would like to apologise to its customers for any concern this may cause them. Each customer will be contacted shortly and a thorough investigation into this matter is under way." The bank is not recommending any action be taken by customers whose details have gone astray. The disk went missing in February 2008, three months after the Chairman of HMRC resigned over the earlier loss in the post of two child benefit information disks discussed in section 1, which contained data on every child in Britain and the bank and national insurance details of their parents. A spokesman for HSBC said an internal inquiry began as soon as the bank was told the disk had not arrived at its destination. But Royal Mail said it had not been contacted about the lost disk.

The City regulator, the FSA was informed six weeks later, HSBC admitted. The bank said the data loss had come about because of a “unique set of circumstances”, blaming the failure of a secure encrypted digital link between HSBC’s Southampton office and reinsurer Swiss Re in Folkestone, Kent. It said the information had been required urgently by Swiss Re and, when the secure link failed, it had been burned on to a disk and put in the Royal Mail’s business post. No courier service was involved said HSBC. The FSA has signaled its determination to clamp down on firms not taking reasonable care of customer data. It has imposed heavy fines on companies failing to own up to breaches in data security. Last December 2007 the FSA fined Norwich Union £1.26m for lax security which allowed fraudsters to target life insurance policyholders. The regulator criticised the insurer for failing to address deficiencies swiftly as discussed earlier in this section.

8. Conclusion

Over the past decade, the public sector’s ability to manage big IT projects has repeatedly been called into question. The usual response is that the government’s record is no worse than private industries, or its counterparts in the rest of the world. However, repeated examples - and the Dunleavy et al 2006 study - suggest there is simply too little IT expertise within the government. And yet far from scaling up efforts to meet the challenge of a new IT infrastructure, the government is scaling back IT, one symptom is the changing of the name of the old e-government unit. A similar nomenclature trend in Canada was discussed by Riley 2003. It is important for government to continue using the nomenclature of e-government. This subject matter is growing and harmony, coherency and consistency in any technology mediated evolutions being undertaken is paramount so that it can be understood by the public.

Government and governance are both about getting the consent and cooperation of the governed. But whereas government is the formal apparatus for this objective, governance is the outcome as experienced by those on the receiving end. The challenges for good e-government (and e-governance) are establishment resistance to change in the status quo, and public cynicism. Nevertheless, as technology developments are currently shaping public expectations citizens will continue to expect more as new state of the art technologies evolve.

The big issue for government is whether it will be able to reverse its internal culture, where IT expertise is not treated as a “core activity”, while it implements sweeping policies which rely on huge IT projects that embrace the population. This could prove to be the biggest task it has ever faced

References

- Bowers S. 2008 HSBC loses disk with policy details of 370,000 customers 8th April 2008 The Guardian News & Media Ltd UK.
- Cabinet Office, November 2005. Transformational Government: Enabled by Technology,

<http://www.cio.gov.uk/documents/pdf/transgov/transgov-strategy.pdf> last accessed 10th April 2008

CROSS M. 2007 Who lost our data expertise? 29th November 2007 The Guardian News & Media Ltd UK.

DUNLEAVY, P., MARGETTS, H., BASTOW, S. and TINKLER, J. (2006) *Digital Era Governance: IT Corporations, the State and e-Government* (Oxford University Press: Oxford).

ECONOMIST 2007, Lost in the post, 22nd November 2007 .The Economist Newspaper & The Economist Group U.K

GUARDIAN 2008 National ID 7th March 2008 The Guardian News & Media Ltd UK.

OSBORNE H. Norwich Union fined over security lapse. 17th December 2007 The Guardian News & Media Ltd UK.

OECD, March 2003a "The E-Government Imperative: Main Findings" in OECD Observer.

POYNTER K. December 2007 Data Security at HMRC Interim Report 14th December 2008 http://www.hmtreasury.gov.uk/independent_reviews/poynter_review/poynter_review_index.cfm last accessed 10th April 2008

RILEY T.B, 2003 E-Government Vs. E-Governance: Examining The Differences In A Changing Public Sector Climate 2003, International Tracking Survey Report '03 No.4

VARNEY D. 2006 Service Transformation: A better service for citizens and businesses, a better deal for the tax payer. HM Treasury 2006

http://www.hmtreasury.gov.uk/pre_budget_report/prebud_pbr06/other_docs/prebud_pbr06_varney.cfm last accessed 10th April 2008

WATT N., 2008, Lost DNA data. 20th February 2008 The Guardian News & Media Ltd UK.

WINTOUR P. 2007 Drivers Data lost, 18th December 2007 The Guardian News & Media Ltd UK.

19

ICT-enabled Services: A Critical Analysis of the Opportunities and Challenges in Uganda

Venansius Baryamureeba

In this paper we analyze the potential for ICT enabled services by Ugandan firms to both clients within Uganda and outside Uganda. We highlight the opportunities and challenges of providing ICT-enabled services by Ugandan firms to companies, organizations, and government departments within Uganda and outside Uganda. We also suggest ways of addressing a common challenge of managing and meeting client expectations in ICT-enabled sourcing services and highlight key recommendations. Categories and Subject Descriptors: J.1 [Computer Applications]: Administrative Data Processing-Business, education, marketing; K.4.2. [Computers and Society]—Social Issues—employment; K.4.3. [Computers and Society]—Organizational Impacts—employment.

1. Introduction

Information and Communications Technologies (ICTs), broadly defined, facilitate by electronic means the creation, storage, management and dissemination of information. ICT is both a vehicle for communication and a means of processing information. ICTs are part of the economic infrastructure that supports global production, trade, investment and capital flows. ICTs are means by which individuals, institutions and organizations network, undertake activities, and participate in the development process at local, national and global levels. Specifically the ICT industry is the main driver of the economies of countries like India and China. In Europe, according to the EU Commission's annual progress report on i2010 [EU Annual report 2007], ICT drives 50% of EU growth. Denmark, The Netherlands, Finland, Sweden, The UK and Belgium -- all have higher broadband penetration rates than the USA and Japan. Technology is fuelling innovation and productivity, and there are signs of fundamental change in markets and user behavior, as countries move towards a knowledge-based economy.

By critically analyzing the economic structure of different countries, it is clear that a higher average of real income is always associated with a high proportion of the working population engaged in the service or public utility sectors. In many large economies, the service sector is the largest in terms of employment, mostly due to a massive increase in productivity growth and progressively higher income elasticity in the primary and secondary sectors. More economically advanced

countries such as the United States of America (USA), Germany, and The United Kingdom, have followed the movement from Agriculture as the largest source of employment, to industry and finally to services. The structural transformation of employment has occurred even more markedly from Agriculture to services in the later developed and some of the currently developing countries. The service sector is becoming a dominant feature of the economic landscape in these countries as its contributions towards Gross Domestic Product (GDP) steadily increase and the contributions of the agricultural sector to GDP decrease.

ICT enabled services is a form of outsourced service which has emerged due to the involvement of ICT in various fields such as banking and finance, telecommunications, insurance, etc. ICT-enabled services include ICT-intensive business processes, projects, and tasks that use ICT as an enabler for designing services, coordinating service deployment, and delivering services. Some of the examples of ICT enabled services are medical transcription, back-office accounting, insurance claim, credit card processing and many more.

Business Process Outsourcing (BPO) is the contracting of a specific business task, such as payroll, to a third-party service provider. Usually, BPO is implemented as a cost-saving measure for tasks that a company requires but does not depend upon to maintain its position in the marketplace. Organizations are increasingly delegating their ICT intensive business activities to external service providers, taking advantage of the rapid evolution of the global telecommunications infrastructure. Studies [Roland Berger 2004] show that various service processes are increasingly being shifted outside Western Europe, USA and Canada. BPO is often divided into two categories: back office outsourcing, which includes internal business functions such as billing or purchasing, and front office outsourcing, which includes customer-related services such as marketing or technical support. The most common examples of BPO are call centers, human resources, accounting and payroll outsourcing.

2. Offshoring

Offshoring of services reflects new ICT that allows companies to reconfigure their value chains with a view to increasing their competitiveness. Offshoring generally refers to the practice by companies, organizations or government entities, of replacing goods or services previously produced domestically with goods and services produced abroad. Advances in ICT and developments in the management of business processes, coupled with a large pool of educated workers in other countries, allow companies to move services work outside the developed countries as part of a larger trend towards global interdependence. The business processes being offshored range from routine to non critical tasks, which are resource intensive and operational, to strategic processes that directly impart revenues. Offshoring places a premium on geographical proximity, language skills, customer service orientation and familiarity with developed-country cultures.

Service offshoring refers to the decision of firms to source services from a foreign country instead of their home country. Service jobs at the risk of being offshored can be defined through the following four conditions [Bernhard 2007]: (i) people in those jobs ‘are likely to make intensive use of ICTs in order to produce their output’, (ii) “their output can be traded/transmitted with the help of ICTs (ICT-enabled trade in services)”, (iii) “the work has a high information or ‘knowledge’ content” and (iv) “the work does not necessarily require face-to-face contact”. Jobs that require a face to face interaction or physical presence like to stock shelves, provide nursing care, and install a network cannot be offshored. Therefore, in developed countries, by and large architects and town planners, judges and magistrates, notaries and other legal professionals, estate agents, commercial sales representatives and technical representatives are not at risk of being offshored. Finance, accounting, software development, ICT support, human resource services, telephone call center services together with other back office operations are currently being undertaken offshore (lower-wage locations such as India, The Philippines, Eastern Europe and part of Africa). Easy of control, flexibility and organizational feasibility are the main drivers when selecting business models for Offshoring.

2.1. Offshoring Potential of Companies Based in USA

While the potential for companies to reduce costs by offshoring back office operations is dramatic, companies can potentially increase these savings by over 50 percent by selectively integrating transformation and process improvement efforts into their globalization initiatives, according to research from The Hackett Group [Tekrati Inc. 2007]. According to Hackett’s research, the Fortune 500 could generate over \$91 billion annually, or about \$182 million on average per company, by strategically moving back office processes overseas without first improving them. Hackett’s research finds that the key to generating these savings is careful planning and analysis to determine which processes to offshore and whether or not to integrate transformation, and a staged approach to offshoring over a five to ten year period. In some cases, Hackett’s research shows that the potential exists for companies to capture the majority of the available cost reduction through process optimization alone.

The Hackett Group [Tekrati Inc. 2007] recommends that companies analyze multiple factors as they evaluate their globalization and transformation decisions, including the current performance levels of each process, the risk of transformation and globalization, the complexity of the process, and the net present value of the strategic alternatives. In examining process complexity, companies should look well below the functional level, at individual processes and the learning curve required to becoming proficient at them. The longer the learning curve, the more likely the process will benefit from transformation prior to offshoring.

There is high potential for offshoring jobs in packaged software and ICT services in the USA [Farrell 2006]. In the USA, companies consider a host of

factors beyond labor cost when deciding where to place an activity including each potential location's risk profile, infrastructure, domestic market, non-labor costs, business and living environment, and the availability of vendors [Farrell 2006]. Interestingly many programming jobs have moved offshore and at the same time more positions of systems analysts and software engineers have been created in the USA [Farrell 2006]. Thus, offshoring is creating a win-win situation.

2.2. Offshoring Potential of Companies Based in EU

A study by Roland Berger [2004] on the offshoring strategies of Europe's largest companies shows that various service processes are increasingly being shifted abroad. So far, the focus is on back office services, but most service processes are potential future candidates for offshoring. While they lag behind their USA rivals, European companies especially from the United Kingdom – see offshoring as a way to reduce costs and improve their competitiveness. Most offshoring projects remain on the European continent, but India – as the most frequently mentioned single destination tends to attract larger offshoring projects in terms of the number of jobs offshored. More than 80% of companies with experience in offshoring are satisfied with the results, reporting cost savings in the range of 20%-to-40%.

Furthermore, Table 1 shows industry breakdown of the occupational employment data for Belgium that can be used to show industries that have the highest share of jobs at risk of being offshored. This clearly shows that there is great potential for offshoring jobs in Belgium. The share of jobs at risk of being offshored due to ICT-developments is lower for Belgium than the European Union or the USA as of 2007 [Bernhard 2007]. Therefore there is high potential for offshoring jobs from the European Union and the USA.

Table 1: ICT-enabled off shoring potential by industry (NACE 2-digit) for Belgium (average and values for 1993 and 2005); Adopted from Bernhard [Bernhard 2007]

NACE- Code share	Industry	Avg 1993	Share 2005	Share 93-05
High potential				
66	Insurance and pension funding, except compulsory social security	81.6%	78.5%	87.5%
65	Financial intermediation, except insurance and pension funding	78.0%	76.5%	78.6%
72	Computer and related activities	76.3%	76.8%	78.8%
67	Activities auxiliary to financial intermediation	74.9%	62.3%	70.9%
30	Manufacture of office machinery and computers	51.8%	64.8%	38.0%
73	Research and development	46.8%	37.7%	48.8%
74	Other business activities	46.1%	48.0%	48.3%
70	Real estate activities	39.0%	41.6%	45.9%
51	Wholesale trade & commission trade, except of motor vehicles & motorcycles	30.7%	30.7%	31.0%

23	Manufacture of coke, refined petroleum products and nuclear fuel	30.1%	38.6%	31.7%
Medium potential				
32	Manufacture of radio, television and communication equipment and apparatus	28.7%	24.9%	38.2%
24	Manufacture of chemicals and chemical products	26.9%	24.5%	26.2%
33	Manufacture of medical, precision and optical instruments, watches and clocks	26.0%	23.6%	31.3%
71	Renting of machinery and equipment and of personal and household goods	25.1%	29.3%	19.3%
91	Activities of membership organizations n.e.c.	22.3%	20.5%	25.1%
40	Electricity, gas, steam and hot water supply	20.5%	18.8%	24.9%
16	Manufacture of tobacco products	18.8%	22.1%	12.6%
63	Supporting and auxiliary transport activities; activities of travel agencies	18.3%	14.8%	20.6%
31	Manufacture of electrical machinery and apparatus n.e.c.	17.8%	11.4%	22.8%
29	Manufacture of machinery and equipment n.e.c.	16.8%	14.8%	16.7%
41	Collection, purification and distribution of water	16.3%	26.7%	18.9%
22	Publishing, printing and reproduction of recorded media	15.7%	16.7%	17.4%
64	Post and telecommunications	14.6%	5.4%	19.4%
25	Manufacture of rubber and plastic products	14.0%	13.4%	12.7%
92	Recreational, cultural and sporting activities	13.8%	13.2%	13.3%
62	Air transport	13.5%	11.3%	15.8%
35	Manufacture of other transport equipment	13.1%	14.8%	12.9%
21	Manufacture of pulp, paper and paper products	12.8%	13.8%	12.6%
15	Manufacture of food products and beverages	12.2%	11.6%	16.3%
27	Manufacture of basic metals	11.8%	9.8%	14.4%
90	Sewage and refuse disposal, sanitation and similar activities	11.4%	5.6%	13.1%
26	Manufacture of other non-metallic mineral products	10.8%	10.1%	13.7%
50	Sale, maintenance and repair of motor vehicles and motorcycles	10.6%	10.4%	12.9%
19	Tanning and dressing of leather; manufacture of luggage, handbags, and footwear	10.5%	10.4%	14.9%
Low potential				
28	Manufacture of fabricated metal products, except machinery and equipment	9.8%	8.7%	10.9%
75	Public administration and defence; compulsory social security	9.0%	7.9%	9.7%
36	Manufacture of furniture; manufacturing n.e.c.	8.6%	7.3%	11.9%
18	Manufacture of wearing apparel; dressing and dyeing of fur	8.2%	8.1%	11.4%
34	Manufacture of motor vehicles, trailers and semi-trailers	7.9%	7.3%	11.3%
17	Manufacture of textiles	7.8%	5.4%	11.6%
61	Water transport	7.4%	6.4%	8.4%
52	Retail trade, except of motor vehicles and motorcycles	6.9%	4.9%	9.3%
85	Health and social work	6.9%	7.0%	7.5%
20	Manufacture of wood and of products of wood and cork, except furniture	5.9%	6.3%	5.8%

45	Construction	5.6%	4.1%	7.5%
60	Land transport; transport via pipelines	5.0%	3.2%	6.6%
80	Education	4.6%	3.5%	6.0%
2	Forestry, logging and related service activities	3.3%	1.9%	8.5%
55	Hotels and restaurants	2.5%	2.4%	3.9%
93	Other service activities	1.4%	0.5%	1.7%
1	Agriculture, hunting and related service activities	1.2%	1.1%	1.8%

2.3. Impact of Offshoring and International Outsourcing on USA and EU Economies

Offshoring and international outsourcing activities are creating a new international division of labor, which should be beneficial for consumers worldwide. By allowing firms to fragment and relocate their activities according to the best conditions, ultimately improves their ability to compete internationally, increasing their share in world markets and eventually creating jobs. In developed countries, it is expected that in the long run increased imports from abroad will be matched by increased exports, i.e. as economic growth accelerates overseas and trading partners grow richer, the conditions are created for a multiplier effect to boost further growth in employment and jobs in the home economy.

The different theoretical expectations about how Offshoring and international outsourcing will impact the USA economy include [GAO 2005]:

- The average USA standard of living. One school of thought is of the view that Offshoring and international outsourcing is likely to undermine USA technology leadership and harm USA long term living standards. On the other hand, economic theory also generally predicts that Offshoring and international outsourcing will benefit USA living standards in the long run.
- Employment and job loss. There is fear that there will be an increase in unemployment in the USA as a result of job loss to offshoring and international outsourcing destinations.
- Distribution of income. There is an argument that Offshoring and international outsourcing could increase income inequality while others argue that changes in the income distribution are driven primarily by factors such as technological change.
- Security and consumer privacy. There is growing concern about the impact of services Offshoring and international outsourcing on the security of the national defense systems and critical infrastructure such as utilities and communication networks as well as privacy and security of consumers' financial and medical information.

Companies in the European Union are delocalizing i.e. shifting economic activities towards foreign sites, including scaling down activities or not expanding at home in favour of expansion abroad [EU Annual report 2007]. This is done through

either offshoring, which implies the firms retain the ownership of the whole of the production process while relocating parts of their activities abroad via setting up of subsidiaries; or (international) outsourcing, which implies contracting out parts of the production process to the external suppliers located in another country. Each company needs to adjust its (industrial) structure to better match its comparative advantage. Hence, inevitably some sectors as well as occupations are likely to expand while others will be reduced, reflecting changing comparative advantages. However, the fact that the phenomenon increasingly extends to all parts of the value added chain raises the fear that the EU may not be able to sustain comparative advantage in a sufficiently broad range of activities and will eventually fail to prevent industries from relocating entirely.

Furthermore, EU's rate of population growth is falling while the inhabitants are ageing [Murray 2008]. The ageing population in Europe presents challenges for public health (concerns over possible bankruptcy of Medicare and related programs) as well as for economic development (shrinking and ageing of labor force, possible bankruptcy of social security systems). Thus offshoring and international outsourcing are likely to produce the wealth that will sustain the ageing population otherwise immigration of skilled labor from outside Europe may slow down ageing population and delocalization.

Within the EU, there is also growing concern that international outsourcing and offshoring will destroy jobs [EU Annual report 2007]. Sectors like clothing and textiles, leather and footwear, ship building and basic metals where labor represents a large share of the total costs have high risk of being offshored or internationally outsourced. Despite evidence on delocalization of highly skilled jobs particularly in the services sector, the majority of delocalization affecting developed economies impacts on the lower skilled workers. On the other hand, the potential benefits of offshoring and international outsourcing include increased competitiveness and lower prices, positive feedback cycles from target countries and labor redeployment on better paid jobs [Hewitt 2005].

3. Challenges And Opportunities

Cost reduction remains the overarching objective for Offshoring and international outsourcing by large firms in developed countries. Cost is normally followed by quality of staff and local services in the list of hard facts benchmarked for the selection of target countries [Hewitt 2005]. Location criteria are also influenced by soft factors such as following competitors, promotion by target countries and internal lobbying by a foreign affiliate. Target countries need to provide the right package in terms of cost, skills availability, infrastructure, and time zone.

The foundations of effective participation in export-related sectors of globalized commerce directly or indirectly related to ICTs lie in the current state of factors such as geographic location, political stability, ICT infrastructure, human resource availability, and legislation. The greatest deficiencies in these areas are normally found in the number of appropriately-educated workers, the quality of the national

electricity grid, relatively low level of Internet penetration, and problems with the enforcement of intellectual property protection [Hewitt 2005].

The Government of Uganda can play an important role in creating the environment for ICT and ICT-enabled services and innovations, by not only putting in place appropriate policy, regulation and investment incentives, but also by its own intelligent and widespread use of ICT as a tool of efficiency and transparency. By understanding the role of ICT in improving private sector competitiveness and economic growth in Uganda, the Government can take the necessary steps to create the appropriate enabling environment and address the skills gap in the economy. Uganda is a reasonably priced offshore services provider [Kartano et al. 2004] due mainly to cheap human resource availability; is familiar with the Common Wealth Culture i.e. a former colony of Britain and currently a member of the Common Wealth. English is the official language.

3.1. Geographical Location

Uganda's standard time is three (3) hours ahead of the Greenwich Mean Time (GMT). When its morning in U.S. its afternoon in Uganda and this still offers Uganda a competitive advantage over India and China. As a result of the time difference with the USA, Uganda can benefit from daytime and overnight work whereas with EU the similarity of time zone, i.e. 1 to 3 hours difference with Uganda can benefit from daytime work in addition to overnight work. Also Uganda is approximately 8 hours away by air from EU capitals. The time difference places Uganda in a favorable position to provide ICT enabled services to EU especially in areas where EU countries may need to interact during office hours with any public and private employees in Uganda.

3.2. Political Stability

Since 1986 there has not been a civil war apart from the terrorist activities by the Lord's Resistance army in the Northern part of Uganda. There is currently a mediation team working on finding a lasting solution to the Northern Uganda conflict. Also Since 1986, Uganda has had regular elections at all levels. There was a return to multiparty system of governance as a result of the referendum that was held in July 2005 and both the Presidential and parliamentary elections of March 2006 were held under a multiparty system of governance. By and large there is freedom of speech in Uganda and according to a survey by Germany firm, The German Bertelsmann Foundation Transformation index put Uganda at number 8 out of 38 African countries [New Vision 2008]. This firm analyses the quality of democracy, market economics and political management in 145 developing countries across the world.

3.3. Legislation

3.3.1. Policy and Regulatory Frameworks

The Ministry of ICT of the Republic of Uganda coordinates all government efforts and regulatory bodies in the ICT sector. Some of the bodies being coordinated by the Ministry of ICT are Uganda Communications Commission (UCC), The National Information Technology Authority –Uganda (NITA-U) and the Broadcasting Council. The ICT sector in Uganda has been influenced by various policies, statutes, laws, acts and regulations, passed and enacted in the recent past. These have brought about liberalization in various social / economic sectors leading to an impressive economic performance.

The Uganda Communications Commission (UCC) was established by the Uganda Communications Act of 1997.

UCC's mission is to facilitate sustainable development of communication services that are universally accessible through effective regulation. UCC regulates and promotes the developments in the communications industry in Uganda. The communications sector is governed by the following legislations: Uganda Communications Act Cap 106 Laws of Uganda and UCC Regulations [UCC 2008]. UCC regulations include Radio Communications Regulations, Equipment Type Approval Regulations, Postal Regulations, Practice and Procedure Regulations, Tariff and Accounting Regulations, Fair Competition Regulations, Licensing Regulations, Universal Service Regulations and Interconnection Regulations. Furthermore, UCC has put in place policies that drive the private sector [UCC 2008] and they include: Telecom Sector Policy, Postal Policy, National ICT Policy Framework and Rural Communications Policy. UCC has the responsibility of developing the rural telecommunications environment through the Rural Communications Development Policy enacted in July 2001 whose purpose is to support the development of communications infrastructure in rural Uganda and to ensure that people in rural areas have reasonable and affordable access to communications services. By June 2007, The Telecommunication policy review had been undertaken and post duopoly policy developed by UCC.

Other ICT related legislations in place include:

- a. The Press and Journalist Statute 1995 (Cap 105), The Statute extended Article 29(1) (*Freedom of expression*) of the Constitution of the Republic of Uganda to the print media. It also created the Media Council, the National Institute of Journalists of Uganda and a Disciplinary Committee within the Media Council. The Council is responsible for regulating eligibility for media ownership and requires journalists to register with the National Institute of Journalists of Uganda.
- b. The Electronic Media Statute 1996 (Cap 104), Laws of Uganda provides for the establishment of the Broadcasting Council that licenses radio, television and video rental licenses. The purchase, use and sale of television

sets is also subject to licensing by the Council. As of September 2007 [UCC 2007] 153 radio stations and 25 TV stations had been licensed.

- c. The Access to Information Act, 2004, that makes provision for access by individuals and/or corporations to information of interest relating to a public company.

Proposed legislations that are due for consideration include:

- a. National Information Technology Authority—Uganda (NITA-U) Bill. The proposed Act establishes an Authority as a coordinating, monitoring and supervisory body to promote National IT development in support of the principles of modernization within the context of poverty eradication; to provide a national IT-policy framework and technical support for IT related services for Government-wide use; and for other matters incidental to the foregoing.
- b. The Electronic Transactions Bill, 2003 proposes an Act to provide for the use, security, facilitation and regulation of electronic communications and transactions; and to encourage the use of e-government services. It intended to inter alia promote e-government and e-commerce in Uganda.
- c. The Electronic Signatures Bill, 2003 proposes an Act to make provision for and to regulate the use of electronic signatures. It attempts to do this through provisions relating to the establishment of the Public Key Infrastructure for authenticity and security of documents.
- d. The Computer Misuse Bill, 2003 proposes an Act to make provision for the safety and security of electronic transactions and information systems; to prevent unlawful access, abuse or misuse of information systems including computers and to make provision for securing the conduct of electronic transaction in a trustworthy electronic environment.

The Commercial Justice Reform Programme (CJRP), focuses on improving the regulatory and service delivery framework for businesses in Uganda. The key areas for reform in the CJRP have included Courts, Commercial Registries (that is the Company Registry and Land Registry), law reform and strengthening the legal profession. The sector seeks to promote alternative dispute resolution mechanisms to ease access for those enterprises that face barriers due to financial and other constraints. A key initiative has been the establishment of a Centre for Arbitration and Dispute Resolution (CADER) with enabling legislation – The Arbitration and Conciliation Act - to allow disputants to seek arbitration, mediation or negotiation as an alternative to adjudication in the commercial court.

1.3.2. ICT Associations

Several industry associations have been formed in Uganda mainly focusing on policy advocacy and creating a competitive environment. The Private Sector ICT Association (PICTA) seeks to create a forum for constructive dialogue with government and attempt to influence policy input for fair taxation and fair play

in government processes. The Uganda Software Services Association (USSA) was established with a specific focus on promoting software development and services in Uganda. USSA has advocacy and advisory functions and it intends to disseminate information to its members about policies, markets, and other relevant issues. The Private Sector Foundation Uganda (PSFU) is an umbrella private sector body comprised of business associations and corporate members. Its focus is on policy advocacy and building competitive capacity for partners. Uganda National Academy for Computing and Internet Technologies (UNACIT) is an Umbrella organization that brings together all the computing professionals in higher education in Uganda. The organization is charged with quality assurance on all computing and ICT programmes in universities in Uganda. The organization assumes the role of implementing national certifications in all areas of Computing and ICT to avoid recognition of certifications being awarded by fake institutions. UNACIT establishes working linkages with the National Council for Higher Education-Uganda and any other regulatory bodies or professional bodies in the discipline of computing among others.

Government should pay serious attention to developing a better enabling environment for ICT business through training, incentive schemes, incubation centres, access to venture capital and improved selection criteria to enable local participation in government procurement contracts. The present climate tends to support larger corporations and multinationals, leaving the local entrepreneurs at a disadvantage [UIA 2007].

3.4. ICT Infrastructure

3.4.1. Telecommunications Infrastructure

SEACOM [SEACOM 2008] will provide high capacity bandwidth linking businesses and communities in Southern and East Africa, Europe and South Asia. When it is fully functional in 2009, SEACOM will be a service provider of international fibre optic bandwidth along the East Coast of Africa linking Southern and East Africa, Europe and South Asia. SEACOM has an enormous capacity of 1.2TB/s, to enable high definition TV, peer to peer networks, IPTV, and surging Internet demand. SEACOM has been structured to meet the policy objectives of Governments and new Partnership for Africa's Development (NEPAD). SEACOM will be the first to launch services with a planned Ready for Service date of June 2009.

The East African Submarine Cable System, EASSy [EASSy 2008] an undersea fiber optic cable is a fully integrated multi-technology network. Construction commenced on 14th March 2008 and its expected to be completed in the last quarter of 2009. It will run 10,500 kilometres from the continent's southern tip to the Horn of Africa, connecting Mozambique, Madagascar, Tanzania, Kenya, Somalia, Djibouti, and Sudan to the rest of the world. Thirteen (13) landlocked countries will also be linked to the system via their terrestrial backbone networks to be developed by the participating companies in Uganda, Botswana, Burundi, The

Central African Republic, The Democratic Republic of Congo, Chad, Ethiopia, Lesotho, Malawi, Rwanda, Swaziland, Zambia and Zimbabwe. MTN (U) Ltd and Uganda Telecom Ltd are the only two companies in Uganda participating in this venture.

There are other similar initiatives to SEACOM and EASSy in the region such as TEAMS which is between the Government of Kenya and the private sector, and Uhurunet which is a partnership of governments and private sector from potentially 23 countries in the region under NEPAD umbrella.

The work on the national backbone commenced early 2007 and by June 2007, optical fibre laying in Kampala, between Kampala and Jinja and between Kampala and Entebbe had been completed [Mulira 2007].

In September 2007, Uganda had 26 licensed telecom operators [UCC 2007], namely MTN Uganda Ltd, Uganda Telecom Ltd, Celtel Uganda Ltd, WARID Telecom Uganda Ltd, HiTs (U) Ltd, Infocom, Africa Online Ltd, Afsat Communications (U) Ltd, I-Tel Ltd, Datanet.Com LLC, Anupam GlobalSoft (U) Ltd, TMP Uganda Ltd, Satellite Communications network Ltd, International Telecom Ltd, Talk TeleKom Solutions Ltd, Roke Investment International Ltd, Mo Telecom International Ltd, Yo Uganda Ltd, Fast Com Ltd, Kampala Siti Cable Ltd, Nomad Com Ltd, Multi Choice Uganda Ltd, Link (U) Wireless Ltd, Bukasa Telecom International Ltd, RCS Ltd, and Kanodiko Systems Ltd.

Uganda has a population of approximately 28 million as of September 2007 [UCC 2007] 4,195,278 were mobile cellular subscribers. With more mobile operators licensed this subscription is likely to go up dramatically.

3.4.2. Energy

Uganda has an acute shortage of electricity, which is negatively affecting the nation's economy and the well-being of its citizens. About 97% of Uganda's population does not have access to electricity. Uganda's power shortage is the single greatest obstacle to the country's economic growth. The demand for electricity is growing rapidly.

Nalubaale Power Station, often known by its old name, Owen Falls Dam, is a hydroelectric power station across the River Nile near to its source at Lake Victoria in Uganda. The rating of the Nalubaale power station is 180MW. The extension to the Owen Falls, Kiira power station has space for five hydroelectric turbine generators with three installed as of 2003. Each unit at the extension has a capacity of 40 megawatts.

The Bujagali Hydropower Project [Bujagali 2008] aimed at addressing the power shortage is a 250MW power-generating facility proposed by Bujagali Energy Limited, a company jointly owned by affiliates of Sithe Global Power, LLC and the Aga Khan Fund for Economic Development. It is sponsored by Industrial Promotion Services (Kenya) Limited and SG Bujagali Holdings Ltd, an affiliate of Sithe Global Power, LLC (USA). The power plant is

Table 2: Academic Programmes Taught at CIT as Well as the New Ones Proposed to Start in Next Academic Year 2008/09, adopted from CIT Annual Report [2007].

Department	Programmes Offered	Short Courses	Proposed Programmes
Computer Science	Diploma in Computer Science & IT (DCSIT) B.Sc. Computer Science (B.Sc.CS) PGD Computer Science (PGDCS) M.Sc. Computer Science (M.Sc. CS) PhD Computer Science (PhD.CS)	Certificate in Computer Applications [CCA] International Computer Driving Licence (ICDL)	NONE
Information Technology	Bachelor of Information Technology (BIT) PGD Information Technology (PGDIT) Master of Information Technology (MIT) PhD Information Technology (PhD. IT)	NONE	NONE
Networks	PGD in Data Communication & Software Engineering (PGDDCSE) PGD ICT Policy & Regulations (PGD ICT PR) M.Sc. Data Communication and Software Engineering (M.Sc. DCSE) PhD Software Engineering	Cisco Certified Network Associate [CCNA] Cisco Certified Network Professional [CCNP] IT Essentials I & II Microsoft IT Academy [MCDBA, MCSE, MCSA, MOS & MCSD]	Fundamentals of Wireless Local Area Networks (WLANS) B.Sc. Computer Engineering B.Sc. Software Engineering M.Sc. Mobile Computing
Information Systems	PGD Information Systems (PGDIS) M.Sc. Information Systems (M.Sc.IS) PhD Information Systems (PhD. IS)	Oracle Certified Associate (OCA)	Oracle Associate Professional (OCP) Bachelor of Information Systems

planned to be situated 1,100m above sea level at Bujagali Falls, about 8km north of Lake Victoria, the source of the Nile. On August 21 2007, Uganda's President Yoweri Museveni and the Aga Khan, Prince Karim al-Hussaini, spiritual leader of the Ismaili Muslims, laid the foundation stone for the Bujagali hydropower dam on the River Nile in a show of commitment to address Uganda's continuing energy crisis. An associated power transmission system, the Bujagali Interconnection Project, is a separate project sponsored by Uganda Electricity Transmission Company Limited (UETCL) that will distribute the hydro electricity to the Ugandan power grid. More projects for power generation are planned by World Bank and other development partners.

Table 3: Graduate & Undergraduate students per Programme as of October 2007, adopted from CIT Annual Report [2007].

Programme	1 st year	2 nd year	3 rd year	4 th year		
PhD by Research only	4	2	4	1	11	
PhD by Coursework & Research	PhD. CS	1	6	4	-	11
	PhD. IT	3	5	-	-	8
	PhD. SE	1	1	1	-	3
	PhD. IS	5	7	7	-	19
Subtotals					52	
M.Sc. CS	15	9	-	-	24	
MIT	35	28	-	-	63	
M.Sc. IS	35	27	-	-	62	
M.Sc. DCSE	39	25	-	-	64	
Subtotals					213	
PGDCS	1	-	-	-	1	
PGDIT	5	-	-	-	5	
PGDDCSE	3	-	-	-	3	
PGDICT PR	-	-	-	-	-	
PGDIS	2	-	-	-	2	
Subtotals					11	
B.Sc. CS	564	303	266	-	1133	
BIT	1038	774	939	-	2751	
DCSIT	121	44	-	-	165	
Subtotals					4049	
Grand Total					4325	

Table 4: Students who have trained on the short courses in 2007, adopted from CIT Annual Report [2007].

Course Name	Total
CCNA	2,421
CCNP	320
IT Essentials (A+, Server+, N+)	312
Microsoft IT Academy [MCDBA, MCSE, MCSA, MOS & MCSD]	21
Oracle Certified Associate (OCA)	34
ICDL	59
CCA	2203
Total	5,370

3.5. Human Resource Availability

The single most important thing that any country will need in order to prosper in the future is a large, well educated workforce. That is why most companies are operating in India and China and will continue to shift their offshore offices to such countries. Fortunately, Uganda has highly skilled human resources that are reasonably priced in comparison with other countries in the region.

Makerere University Faculty of Computing and IT (CIT) [CIT 2008] has positioned itself as the main ICT training centre in Africa. CIT has the largest teaching and laboratory infrastructure on the African continent totaling about 15,000 sq metres of space. In positioning itself to address the ICT human resource gaps in the country, Makerere University Faculty of Computing and IT (CIT) is already running most of the computing programmes and is also planning to establish new computing programmes as indicated in Table 2:

As of December 2007, CIT had a total of 4335 students (4049 undergraduate and 276 postgraduate) of which 331 are international students as indicated in Table 3. An addition of 5,370 were trained on short courses in 2007. All undergraduate students take Cisco Certified Network Associate (CCNA) and there are excluded from the CNNA total of 2421 in Table 4.

Makerere University Faculty of Computing and IT has been designated the National ICT incubation centre with the mandate to undertake software and business incubation among others.

Within Makerere University there are other faculties running ICT related programmes such as Bachelor of Library and Information Science and Bachelor of Business Computing. Makerere University has about 30,000 undergraduate students and over 5000 postgraduate students [Mak 2008]. Uganda has a total of 5 public Universities and over 15 private Universities all of which run programmes in the field of ICT. As per the tables above, Makerere University alone produces over 1000 ICT graduates per year. The rate of unemployment is as high as 70% of the graduates. So it goes without saying that Uganda has sufficient and qualified human resources needed to support ICT-enabled services.

3.6. Managing and Meeting Client Expectations

Due to variations in working culture it is important that measures are put in place to ensure that client expectations are met in ICT-enabled services. This requires that the whole process must be managed by someone whom the client has greater expectations and trust in order to build client confidence. For new markets of ICT enabled services like Uganda the following are recommended:

3.6.1. *Hiring Expatriates*

The local companies need to initially hire expatriates from countries that have high potential to offshore work. These expatriates could be recommended by the company/ firm intending to outsource work to a company in Uganda. The expatriate (s) would manage the process and ensure that deadlines are adhered to and the work is of high quality and meets clients' expectations. The local employees would also have the opportunity to understudy the expatriate and after a given period be able to take over from the expatriate.

3.6.2. *International Placements*

The other option to be considered in ensuring that customer expectations are met is to place local Ugandan staff with firms/ companies in developed countries for a given period of time so that they can acquire on job skills in managing outsourcing projects. After this placement period they can return to Uganda and be key persons in ensuring that clients' expectations are met.

3.6.3. *Regional Placements*

Regionally, we are lucky that countries like Mauritius are already success stories in the outsourcing industry. Under arrangements at government, firm or individual level Ugandans could acquire management and other skills from firms in regions that are already successful in the business. Thereafter, they can return and set up their own outsourcing businesses in Uganda.

3.6.4. *Providing Incentives to African Diaspora*

A good number of the African Diaspora work with multinational ICT companies like Microsoft, IBM, Lenovo, Google, Cisco Systems, Yahoo, etc. to name but a few. They have accumulated a lot of skills that could enable ICT enabled services be a core component of the services industry in Uganda. Therefore, incentives must be put in place to encourage them to either relocate to Uganda or to enable them set up offshore offices in Uganda or provide services like telemarketing, technology transfer and investment capital at a distance.

1.6.5. *Providing Incentives to International Firms*

Incentives could be provided to international firms experienced in doing outsourcing work to open their offshore offices in Uganda. This would give an opportunity to Ugandans to train on the job within the country. Thereafter, Ugandans can start their own outsourcing businesses.

3.6.6. *Setting up a National Incubation Facility*

The government of Uganda should seriously consider setting up an incubation facility that will produce up to speed skilled workers that can effectively engage in the outsourcing industry.

4. Recommendations and Conclusion

4.1. Recommendations

- a. ICT enabled services require a wide range of skills not only in computing disciplines (computer science, computer engineering, software engineering, information systems, information technology) but also in areas such as business administration, accounting, inventory control, marketing and customer service. We recommend that all graduates in all disciplines before they graduate or immediately after graduation be taken through mandatory training in ICT areas such as; use of computers based on the international computer driving license, use of common software, internet and communication, and computer/computing ethics. On the other hand specialized training that produces high level specialists in computing disciplines (computer science, computer engineering, software engineering, information systems, and information technology) to design and create new technology and information systems should continue to be carried out and more should be planned. Also specialized training for technicians to install, configure, and maintain hardware, software, and networks through certifications such as CCNA, MCSE, OCA and OCP should continue to be carried out and more should be planned. In addition, industries which the government regards as especially important, such as call centers, will require specialized training that the state may help to provide: in the case of call centers, this currently involves (English, French, Spanish, Germany etc) language immersion courses and courses on call center services. The Centre of Excellence in Computing and ICT at Makerere University with its affiliated incubations centres in areas such as call centre and software development should be replicated at all institutions of higher learning.
- b. A national educational and skills inventory should be undertaken, which will in return serve as an aid in national initiatives aimed at improving competitiveness of the country. In particular, a far higher potential for commercial success is in the provision of software services – especially contract programming – and the provision of front-office call center services and back-office clerical work to foreign clients who either have offices in Uganda, or who can be provided with services through telecommunications networks (“offshore outsourcing”). The success of Uganda in these areas depends critically on the availability of a sufficient number of highly-trained workers. Thus a systematic inventory of the skills and numbers of workers must be carried out to make the actual situation clearer, and to plan effectively for the creation of a large number of new workers to meet future demand.

- c. As much as possible academic programs should follow international curricula benchmarks or certifications as is the case at Makerere University Faculty of Computing and IT [Wanyama and Baryamureeba 2007; Baryamureeba 2007] if the graduates are to remain competitive. This concept is being implemented at all the public Universities in Uganda on all their computing programs. The private Universities in Uganda have also started aligning their curricula with international standards. All higher education institutions are listed on the National Council for Higher Education website [NCHE 2008]. The National Council for Higher Education is a statutory agency, a watchdog for quality and relevant higher education established under “The Universities and Other Tertiary Institutions Act, 2001” for; (i) regulating and guiding the establishment and management of institutions of higher learning, and; (ii) regulating the quality of higher education, equating of higher education qualifications and to advise government on higher education issues.
- d. Uganda should avoid being trapped in “low-end” services provision characterized by low salaries, poor working conditions, and low investment in employee training. Emphasis should be placed at all times on trying to position the country’s citizens and businesses as providers of more skilled services, and to favor those foreign businesses who are oriented towards the use of such services, and the training of Ugandan workers to supply them.
- e. There is need to continue subsidizing computers and computer software. The extension of regular telephony and Internet connectivity through the use of new long-range wireless connection technology in metropolitan and rural areas should likewise be seriously considered. The lack of widespread Internet connectivity and frequent failures of the national electrical grid (which also negatively affects the penetration of computers) make it likely that ICT and ICT-enabled businesses will for the foreseeable future have to be concentrated in urban areas like Kampala. Also, once fully equipped the Centre of Excellence in Computing and ICT at Makerere University is expected to make available over 8000 computers to be used in overnight outsourcing work.
- f. The Republic of Uganda must not only have laws and signed treaties that protect intellectual property and personal data, it must also enforce them strictly, and be internationally seen to be doing so if it is to have reasonable hopes of participating fully in some of the most promising areas of ICT enabled services.
- g. The software and outsourced office services sectors can involve either local companies, or foreign companies that are attracted to the country to carry out their operations. In the early stages of their operations, local businesses almost inevitably face serious problems in finding adequate financial resources and in obtaining information about foreign markets and competitors, have serious problems in making their offerings known to an international audience, and have equally serious problems convincing that audience to have confidence in

their products and services. The following actions should therefore be taken [Hewitt 2005]:

- Financing – Uganda government should do whatever it can to make financial resources and financial mechanisms accessible for local companies. This can include providing grants or subsidies for companies in strategic areas, having Bank of Uganda extend credit to these companies on a portfolio basis (to distribute the risks of investing in startup businesses), helping to find local or foreign investors, and stimulating the formation of local venture capital funds, whose own portfolio-based approach helps private sector investors to overcome their reluctance to invest in startups.
- Investigation of markets and competition – both the Uganda government and local businesses should search out international trade data, and data on the internal demand for software and outsourced office services in the countries of potential clients. They should also investigate the products, services, prices, and marketing strategies of their competitors in developed and developing countries.
- International marketing and relationship formation – local businesses should undertake international marketing programs and searches for possible partners in foreign countries. The Uganda government can help to introduce local companies and their products and services to potential foreign clients and partners through the standard mechanisms of trade fair sponsorship, help with visits of potential clients to the country, trade missions, and consular activities.
- Confidence – local businesses should make all possible efforts to evaluate and implement internationally-recognized quality certification methodologies (e.g., Carnegie-Mellon CMM and ISO standards for software, Customer Operations Performance Center [COPC] for call centers), to have workers qualify for various professional certifications (e.g., Microsoft, Cisco, and Oracle certifications), and to publicize their qualifications in the international marketplace. The Uganda government should stimulate and strongly support quality control efforts in strategic sectors.

There are at least two extremely important types of activities that can assist in most or all of the areas mentioned in (g) above:

- Strengthen existing professional associations and form new ones in strategic areas where they do not exist. The high costs of international market research and marketing, and the implementation of quality certification methodologies, can be shared among the members of a strategic sector through sector-specific professional associations. Participation in such associations can also benefit their members by giving them increased influence in lobbying for favorable legislation, and benefit the government by giving it a single source to talk to when considering new legislation.

- Mobilization of the Diaspora – the government and private sector associations should cooperate to identify, contact and organize members of the Ugandan community living abroad who might be especially disposed towards contributing financial support, technology, entrepreneurship, and market intelligence, and who might, in some cases, actually be clients for the local products and services offered.
- h. It is pertinent that Uganda aggressively tackles corruption and bad governance at all levels to ensure long term political stability and investor confidence.
- i. Provide investment incentives. In the case of foreign businesses, standard incentives include selective tax reductions, economical telecommunications, loose restrictions on foreign ownership, repatriation of earnings, manpower training initiatives, assistance with locating and renting appropriate facilities, preferential customs clearance procedures, and labor laws that do not unduly favor local workers at the cost of foreign business owners. In case of local investors, incentives should include financial resources, International marketing and relationship formation using government machinery among others.

4.2. Concluding Remarks

The opportunities lie in the skilled but low priced human resources availability, sufficient language skills, familiarity with mainly The Common Wealth culture and low cost of living. The short term challenges being addressed are international bandwidth and energy. However, thorough planning for energy needs to be undertaken to cater for future growth. Uganda must ensure continued political stability and accountability. All in all we have shown that despite the few challenges which are being addressed, Uganda has great potential to be the number one ICT-enabled services provider on the African continent.

References

- BARYAMUREEBA, V. 2007. ICT as an Engine for Uganda's Economic Growth: The Role of and Opportunities for Makerere University. *International Journal of Computing and ICT Research*. 1, 1, 47-57.
- BERNHARD, M. 2007. Potential ICT-Enabled offshoring of Service Jobs in Belgium, Working Paper 6-07, Federal Planning Bureau, Brussels.
- BUJAGALI, 2008. http://www.bujagali-energy.com/bujagali_projectOverview1.htm. Accessed on 11th march 2008.
- CIT ANNUAL REPORT, 2007. Makerere University Faculty of Computing and IT Annual Report.
- CIT, 2008. <http://www.cit.ac.ug>. Accessed on 11th March 2008.
- EASSy, 2008. <http://eassy.org/>. Accessed on 11th March 2008.
- EU ANNUAL REPORT, 2007. http://ec.europa.eu/information_society/eeurope/i2010/annual_report/index_en.htm, dated March 30, 2007 and accessed on 1st November 2007.

- FARRELL, D. 2006. U.S. offshoring: Small steps to make it win-win, The Berkeley Electronic Press, www.bepress.com/ev. Accessed on 25th February 2008.
- GAO 2005. Offshoring of services: An overview of the issues, United States Government Accountability Office, Report to Congressional Committee.
- HEWITT, J. 2005. Analysis of the Export Potential of ICT and ICT-Enabled products and Services in the Dominican Republic. Reproduced by Chemonics International Inc. under the Contract No. PCE-I-830-98-00015-0.
- KARTANO, K., TINA, J., MUSISI, C. 2004 ICT Investment opportunities in East Africa: Country specific market analysis. International Finance Corporation, Country Report for Uganda.
- MAK, 2008. <http://www.mak.ac.ug>. Accessed on 12th March 2008.
- MULIRA, H.M. 2007. Ministerial Policy Statement. Financial Year 2007/2008.
- MURRAY, A. 2008. Growing old gracefully: How to ease population ageing in Europe. Published by Centre for European Reform, 14 Great College Street London, ISBN 978 1 901229 79 0. http://www.cer.org.uk/pdf/e790_17jan08.pdf. Last accessed on 4th June 2008.
- NCHE, 2008. <http://www.unche.or.ug/>. Accessed on 12th March 2008.
- NEW VISION. 2008. The New Vision, Vol. 23, No 061, pp. 1-3, 25th March 2008.
- ROLAND BERGER. 2004. Service Offshoring takes off in Europe—In search of improved competitiveness. United Nations Conference on Trade and Development.
- TEKRATI INC. 2007. <http://services.tekrati.com/research/8918/> Accessed on 25th February 2008.
- UCC, 2007. Status of the Communications Market. Uganda Communications Commission. <http://www.ucc.co.ug/marketReviewSept2007.pdf>. Accessed on 9th March 2008.
- UCC, 2008. <http://www.ucc.co.ug/>
- UIA, 2007. Uganda Investment Authority. <http://www.ugandainvest.com/uia.php?uhpl=incincentives&&uhpl1=Incentives>. Last accessed on 4th June 2008.
- SEACOM, 2008. <http://seacom.mu/index2.asp>. Accessed on 11th March 2008.
- WANYAMA, T. AND BARYAMUREEBA, V. 2007. The Role of Academia in Fostering Private Sector Competitiveness in ICT Development. *International Journal of Computing and ICT Research*. 1,1, 58-67.

The Localisation into isiXhosa of the iLanga Telephone System

*Lorenzo Dalvit, Mosiuoa Tsietsi, Alfredo Terzoli,
Pamela Maseko Msindisi Sam and Thandeka Mapi*

In this paper we describe the localisation of iLanga, a VoIP PBX system, into isiXhosa. In South Africa, for social and political reasons, making systems available in the African languages is becoming more and more important. On the one hand, it gives access to the use of technology to people who are not fluent in English. On the other hand, it has a strong symbolic value. Although a variety of software is already available in a number of African languages, localisation in such language is still in its early stage. The process presented interesting challenges both from the technical and the linguistic point of view.

1. Introduction

In South Africa, like in other parts of Africa, English is the dominant language of Science and Technology. In most cases, proficiency in English is therefore required in order to access Information and Communication Technology (ICT). Various sources [Department of Education and Department of Communication 2001] acknowledge the potential role of English as a “gatekeeper” to the use of ICT by members of disadvantaged communities and emphasise the need to promote technological discourse in the African languages.

We start our paper with an overview of the presence of African languages in ICT and on localisations in the African languages. We then focus on the localisation process itself and the specific challenges to localisation in the African languages. After that, we describe our experience in localising iLanga, a VoIP PBX system developed in the Department of Computer Science at Rhodes University. We end with some feedback and some reflections on the possible scenarios for the implementation of the localised system.

2. Localisation in the African Languages

2.1. African Languages And ICT

The *localisation* (i.e. translation of software interfaces) in the South African indigenous languages is a comparatively recent but fast growing phenomenon. Microsoft Windows XP has been released in Setswana, isiZulu and Afrikaans [Ashword 2006]. Up-coming Microsoft Vista is expected to include more African

languages. The popular search engine Google is also available in a number of South African languages.

Wikipedia [Wikipedia 2007] features a number of articles in various South African indigenous languages. A growing number of resources in the African languages are available on the Web [De Schryver 2003]. These are mainly on-line dictionaries, often maintained by speakers of other languages.

Translate.org.za [Translate 20007] is a South African NGO committed to the localisation of open-source software into all South African official languages. Software localised so far includes OpenOffice, Mozilla Firefox, Gnome and KDE as well as a number of other applications such as. Tuxpaint, for instance [see <http://downloads.translate.org.za>]. This makes it possible to run a computer almost entirely in a language such as isiXhosa, for instance.

Translate.org.za developed an alternative methodology for the localisation of open-source software called `translate@thon`. This is basically a translation marathon, during which volunteer translators work collaboratively on-line using PO-based On-line Translation and Localisation Engine (Pootle). This is an application developed by Translate.org.za themselves, which allows translators to work on the same project collaboratively and provides suggestions from a database to ensure consistency. The main Pootle server is maintained by the Telkom Centre of Excellence of Rhodes University and is available at <http:pootle.translate.org.za>.

2.2. The localisation process

The two main components of localisation are translation and modification [Keniston 1999]. Translation is the linguistic component of localisation. Ideally, it consists of five phases: translation into the target language, back-translation into the original language, comparison between the two versions, adjustments to the version in the target language and integration into the interface [4]. The localised items are usually stored and indexed into a separate file.

Modification refers to more “structural” changes, such as scrolling patterns (e.g. for Arabic), character set (e.g. for Russian or Chinese) and box sizes. It can involve cultural aspects (e.g. colours or icons) and linguistic aspects (e.g. dictionary search patterns).

2.3. Challenges

Localisation can be a lengthy and expensive process. The localisation of an operating system, which can involve up to half a million words, can cost up to one US Dollar per word and take one to two years [Keniston 1999].

Another problem is finding appropriate terms which are at the same time effective and compliant to a common standard. This is particularly true for African languages, which often have very different dialects and varieties. Moreover, since localisation in such languages is a comparatively new phenomenon, various interventions are usually piecemeal and lack coordination.

Related to this, there is the issue of lack of appropriate terminology [Heugh 2002]. This affects the use of African languages in Science and Technology in general. A considerable effort is put by the South African Government into developing new terminology and promoting its use [National Language Services 2007]. In spite of this, many speakers of an African language prefer to use English borrowings [Heugh 2002], such as *ipassword* (password) or *imeyile* (e-mail) in isiXhosa.

3. Background to the Project

3.1. A Multi-disciplinary Collaboration

In recent years, a common interest for localisation issues has brought together Translate.org.za, the Telkom Centres of Excellence (CoE) of Rhodes and Fort Hare University and the South Africa-Norwegian Tertiary Education Development (SANTED) multilingualism programme at Rhodes.

The SANTED programme is specifically tasked with the implementation of the Rhodes Language Policy. Rhodes is an historically English-medium institution, but is situated in a predominantly isiXhosa-speaking area. Its new language policy involves mainly the promotion of isiXhosa in all domains. Although the CoE of both Rhodes and Fort Hare focus mainly on telecommunication, their leaderships have long realised the importance of language issue in promoting access to ICT. For this reason both Centres have sponsored research on localisation in the past.

3.2. An Asterisk-based Telephone System

The telecommunication landscape has been evolving at a fast pace, spurred on by the exciting possibility of delivering a range of new multimedia services to consumers largely through the adoption and development of open source telephony applications, among which Asterisk is chief [Meggelen et al 2005]. Asterisk is a softswitch that is able capable of handling voice and data communication between heterogeneous endpoints in a seamless fashion. It provides a framework for producing a variety of useful telephony services such as IVR (Interactive Voice Response), voicemail storage and retrieval and multi-party conferencing.

Since 2004, postgraduate students sponsored by the CoE at Rhodes University have been developing an integrated telephony solution called iLanga which is based on Asterisk [Penton and Terzoli 2004]. The initial installation was complimented in subsequent years with an aesthetically pleasing, dynamic interface built with Flash, Action Script and PHP [Hitchcock 2006]. Its main objective was to provide an alternative method through which users could interact with the system. Many of the functions that are provided on the user telephones, such as accessing voicemail messages and initiating calls, are also available via the interface.

iLanga is a complete, cost effective, telephony solution based exclusively on open source telephony and non-telephony software, which embeds asterisk at the core. Since 2005, it has been connected to the University's PBX and tested with postgraduate students and staff in the Department of Computer Science. During

2006, a similar system was deployed in Dwesa, a rural community in the Eastern Cape, South Africa, as part of an ICT-for-Development project discussed at a previous edition of the conference [Dalvit et al 2007]. The system is now available to a wider and linguistically (as well as culturally) diverse community of users. This prompted its localisation into isiXhosa, the most widely spoken African language in the area.

4. Localisation of iLanga

4.1. A Bilingual Interface

iLanga currently allows a user to choose which language they want to use in their interaction with the system, either English or isiXhosa [Dalvit et al 2006]. The system interface has two components: a Web-based interface and a telephone-based audio interface.

The Web-based interface comprises a total of sixty items, divided into six screen shots. The initial translation was performed by approximately a dozen isiXhosa-speaking students of Computer Science using the translate@thon method. The initial phase took approximately six hours. The translations were later reviewed by a language expert before being integrated into the system (see Figure 1).

Fig.1: iLanga Main Page



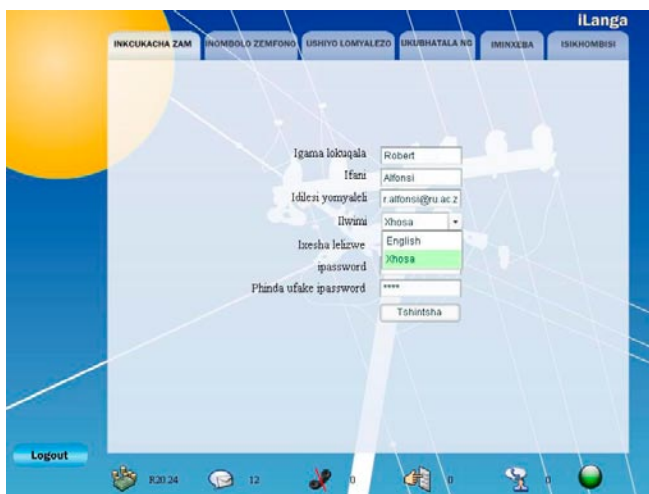
To interact with users via phone, Asterisk uses a set of pre-recorded voice prompts, stored in .gsm format. A complete language pack consists of 436 prompts, equivalent to approximately 2500 words. Obviously, some of these are more frequently used than others and some others refer to features (e.g. video conference) which are not yet implemented in iLanga. For this reasons we focused primarily on frequently used services, such as voicemail for instance.

The majority of the prompts were translated by a team of postgraduate students from various disciplines. These ranged from Computer Science to African Languages and Communication. On a smaller scale, the translate@thon method was again used.

Prompts were recorded and converted into .gsm format using standard open-source tools [Audacity 2007][Sox 2007]. The translated prompts were included in a separate directory identified by the relevant language code (i.e. “xh” for isiXhosa). The system automatically used prompts from the right directory based on the value of the “language” variable set through the user interface.

This configuration option is shown in Figure 2. The same does not apply to the audio portion of the service, which replays prompts completely in English. Fortunately, asterisk does provide configuration options for multi-language support so that users can listen to audio prompts in their language of preference (see Figure 2).

Fig 2: Language Configuration in iLanga



4.2. Challenges

To maintain the cohesion between the iLanga interface and the audio system, it is necessary to ensure that any language preference changes made on the interface are implemented in the audio system. iLanga stores the preference in a MySQL database, thus the Asterisk configuration on iLanga must be extended to interact with the database to serve the appropriate audio content on the user’s telephones.

Localisation of Asterisk systems is available for most European languages, such as French and Spanish. African languages have been marginalised in this respect, and currently there is no isiXhosa language package available for users of Asterisk systems to download and install. The compilation of an isiXhosa package, containing the voice prompts and working configuration settings, is one of the envisaged spin-offs of our project.

The main technical challenge was to get the system to respect the language-specific order of the prompt for isiXhosa instead of the default one for English. For example, in isiXhosa the adjective follows the noun it refers to, while in English the adjective precede the noun. In the initial stages, we relied on the example configurations for other languages (i.e. Italian) which follow a similar word order. A completely independent configuration, specific to isiXhosa, is currently being developed.

5. Summary

Localisation of software in African languages is important for both practical and symbolic reasons. This presents specific challenges, such as lack of resources and terminology for these languages. During the localisation of the iLanga VoIP PBX, a collaborative approach called translateathon was used for the translation component. The modification component presented a few technical challenges, mainly related to the structure of the languages involved and the differences compared to English. While the Web interface of the system is complete, the localisation of the voice prompts is limited mainly to the voicemail system, but suitable for demonstration purposes. By the time of the conference, we are confident we will be able to run a complete demonstration of the system. On a conclusive note it should be emphasised that, to our knowledge, this is the first VoIP PBX system ever to be localised in an African language. We hope it will constitute a valuable experience to inform future similar efforts.

Acknowledgements

We would like to thank the Telkom Centre of Excellence in Distributed Multimedia and the South African-Norwegian Tertiary Education Development Programme of Rhodes University for sponsoring this work.

References

- ASHWORD, W. (2006 April). Windows goes African. *ITWEB*. Available: <http://www.itweb.co.za/sections/software/2006/0604211033.asp?S=MICROSOFT%20WATCH&A=MSW&O=FRGN>
- DALVIT, L., ALFONSI, R., MINI, B., MURRAY, S., & TERZOLI, A. (2006). "The localisation of iLanga, a next generation PBX". Poster presented at the *Southern African Telecommunication Networks and Applications Conference (SATNAC)*, 3 - 6 September 2006, Cape Town.
- DALVIT, L., MUYINGI, H., TERZOLI, A., & THINYANE, M. (2007). "The deployment of an e-commerce platform and related projects in a rural area in South Africa". *International Journal of Computing and ICT Research*, 1(1): 7-19. Available: <http://www.ijcir.org/volume1-number1/index.php> (last accessed 30 November 2007).
- Department of Education and Department of Communication. (2001). *Strategy for Information and Communication Technology in Education*. Pretoria: Government Printer.
- HITCHCOCK, J. (2006). *Decorating Asterisk: Experiments in Service Creation for a Multi-Protocol Telephony Environment Using Open Source Tools*. Master's Thesis, Rhodes University, South Africa
- Gtranslator project details. Available: <http://freshmeat.net/projects/gtranslator/>

HEUGH, K. "The Case Against Bilingual and Multilingual Education in South Africa: Laying Bare the Myths". *Perspectives in Education*, vol. 20, no. 1, pp. 1-19.

KENISTON, K. "Language, power and software" in *Cultural Attitudes towards Technology*. C. Ess, Ed. NY: Suny Press. 1999.

MEGGELEN, J.V, SMITH, J. AND MADSEN, L (2005) Asterisk, the Future of Telephony. National Language Services. Available: http://www.dac.gov.za/about_us/cd_nat_language/home.htm

PENTON, J., AND TERZOLI, A (2004). iLanga: A Next Generation VOIP-based, TDM-enabled PBX. In *SA TNAC Conference Proceedings (2004)*

Pootle <http://pootle.translate.org.za>

Translate project. Available: <http://www.translate.org.za>

21

Agro-meteorological Knowledge Management System for Small Scale Farmers

Omondi Lwande and Muchemi Lawrence

A lot has been done in the use of Information and Communication Technology (ICT) in business yet very little in solving agricultural problems in developing countries. The problem addressed in this paper is the integration of widely available knowledge, processing and providing effective delivery for countries with low ICT infrastructure. Data was collected from small-scale farmers, meteorologists, agricultural extension officers and a research station and analyzed. From this, a suitable knowledge based model was developed. In order to evaluate the model, a prototype was developed and used in tests with sample data. The results reveal that the best way of disseminating agro-meteorological information to small scale farmers is using an integrated knowledge management system and mobile phone technology.

1.0 Introduction

A common problem in developing countries is the lack of integrated means of processing and delivering agrometeorological information to small scale farmers. Even with improved agricultural technology and improved level of farm inputs the agricultural sectors of these countries operate below their potential level owing to the challenges imposed by the marked weather and climate variability. These countries lack a formal system that integrates data from different sources such as agricultural research institutes, meteorological stations, agricultural advisory agencies and the actual ecological requirements. To understand the magnitude of the problem and suggest possible solutions, a study was carried out in Kenya. Kenya has many districts where small scale farming is carried out. Purposive sampling method was used in picking up the district where this research was done [Mugenda and Mugenda 2003]. Simple random sampling technique was used in picking out the respondents during the survey. The survey showed that Information has to be delivered to farmers in a purposeful and meaningful manner that can enable them make informed decisions. It also revealed that majority of the farmers are not able to access meteorological information and even where available processing requires human expertise which is rarely available. The mode of delivery mainly involves mass media especially the radio and is not effective.

This paper therefore addresses the challenges brought about by the above short comings. We specifically address the problem of lack of appropriate system that collects, organizes and disseminates agro-meteorological information to small

scale farmers in a meaningful format. We present a model of knowledge based system that integrates the collection, processing and dissemination of vital agro-meteorological information. It integrates an expert system, and comprehensive knowledge dissemination techniques involving the mobile phone and the internet. A prototype based on the model was developed. This paper also presents the evaluation results of a prototype based on the above model. Through this prototype, we demonstrate that small scale farmers in developing countries can quickly obtain expert information on ecological requirements, pests, diseases and meteorological information by taking advantage of the fast growing mobile phone technology.

2.0 Related Approaches

The challenges of integrating agro-meteorological data in developing countries are wide and varied. Many models have been brought forward but of all the reviewed models, none could comprehensively address the unique challenges in the developing countries. The unique challenges include a large population of rural people with limited access to technology, among many other challenges addressed later in this paper. Most proposed models address agriculture or weather related issues separately. Further they do not treat the aspect of reasoning with the knowledge resident in people and giving focused information a priority. Activities of mainstream research have concentrated on the following basic categories: field data collection and storage, data processing and knowledge dissemination. Majority of the reviewed models propose the use of remote sensory equipment in collecting field data. The upper air atmospheric observation is done by remote sensors carried in special balloon such as that reported by Kenya Meteorological Department [KMD, 2008]. The ground observation is done using remote sensors as reported in Khalid and Steve et al. [2005] among others. Developing countries are unable to easily implement the automatic ground observation techniques due to inadequate level of funding. The spread of mobile telephones seems an attractive option for collecting weather data in the field and sending it to a centralized database. After processing, the knowledge may be disseminated through short message service as reported by Hansen and Jensen. [2002]. This information should be combined with ecological requirements, pest and diseases information as well as crop field observation to produce meaningful information. Rule based systems have been proposed for processing such information as reported in Okinyokun and Anyiam [2001].

3.0 Methodology

In order to understand the underlying issues, the methodology for the research involved a survey before the model design phase. The survey was conducted to reveal the meteorological, ecological, crop health and field information needs for farmers. It was also meant to study the most suitable way of processing the available data and disseminating knowledge to small scale farmers and the most

suitable model that would be necessary to implement the processes. The research was divided into survey and data analysis phase, design and implementation phase and prototype evaluation phase. These together with the results are described here below.

3.1 Field Survey

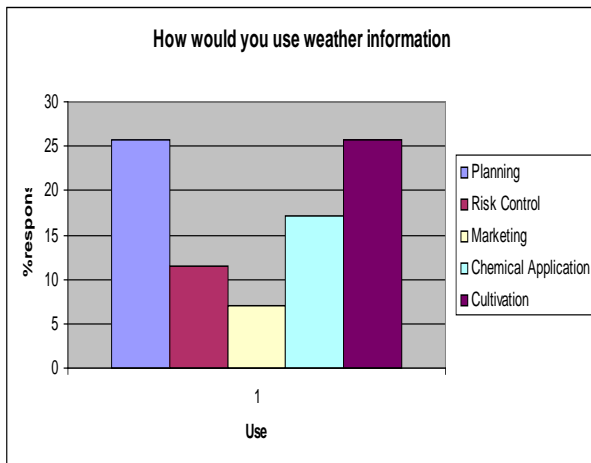
Purposive sampling method was used in picking up the district where this research was done. The district selected was a convenient sample intended to represent a broad range of small scale farmers in Kenya. Simple random sampling technique was used in picking out the respondents during the survey. Structured questionnaire was developed and pre-tested on twenty respondents. The questionnaire was then modified prior to distribution. Fifty respondents were identified during a parents' day in a school in Kiambu district of central Kenya. All the respondents were small scale farmers. The sample was convenient because it represented farmers from all areas of the district. Each questionnaire returned was re-checked and when information was incomplete, attempts were made to trace the respondents who had filled it to request completion. The researchers also observed the actual procedure of data collection, processing, storage and dissemination of weather information to consumers at one location being served by a particular weather station, the Kabete meteorological station. Interview questions were used for three categories of respondents namely small scale farmers, meteorologists and agricultural extension officers. Face to face interviews were conducted to get the views of six meteorological officers at the meteorological station. Six agricultural officers at the crop field stations were also interviewed to know data collection, processing and dissemination methods. Documents and tools used for data collection, recording, storage and processing were collected from the meteorological station, a university campus Crop Research station and the national agricultural research institute, Kenya Agricultural Research Institute (KARI). These assisted in knowing sources of information as well as storage and dissemination methods.

3.2 Survey Results and Analysis

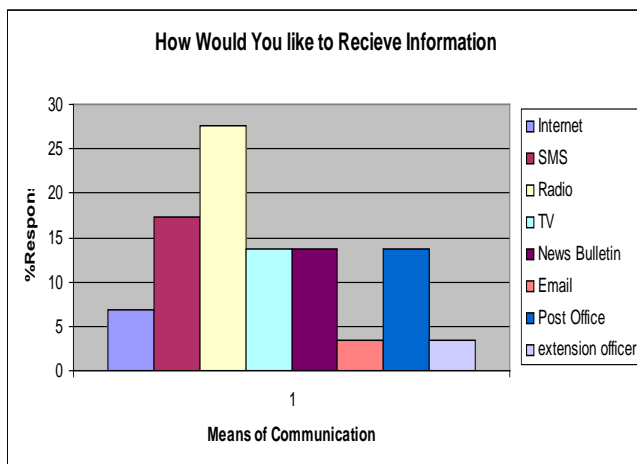
From the field survey, majority of the farmers preferred using weather information for planning planting and cultivation activities. This represented over 50% of the respondents as shown in figure 1.1 A below. 27% preferred receiving weather information through radio compared with 17% SMS as shown in figure 1.1 B below. However, respondents showed reservations with radio because bulletins are done at specific hours when probably the user is not available. This hinders free access to information when needed.

Fig 1.1: Usage of information (fig. A) and preferred communication methods (fig. B)

A



B

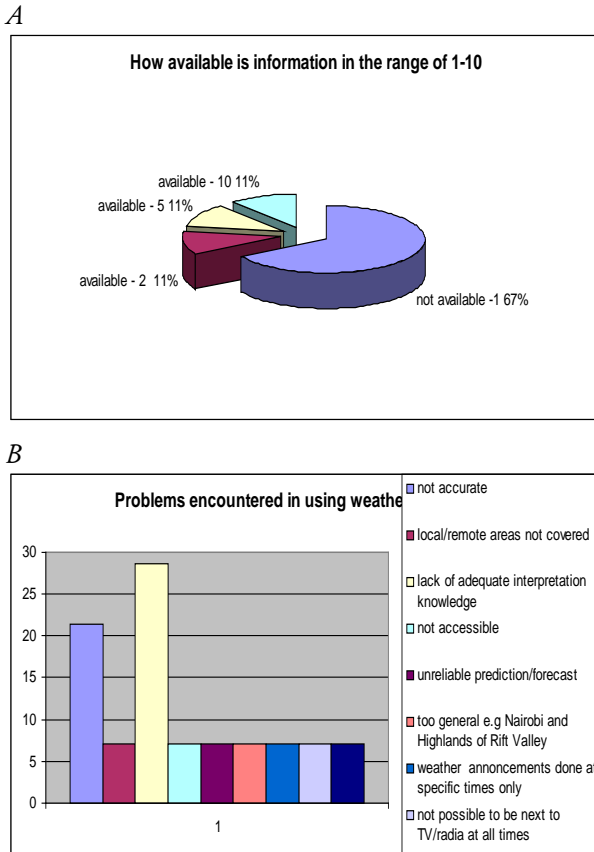


From the above results, one can see that farmers need timely information to make various decisions such as planning for planting, cultivation and chemical application. A good number of respondents appreciate the use of the mobile phone for receiving information. This gives a good motivation for the use of mobile phones in the model.

From Fig 1.2 A below, 67% of respondents claimed weather information was completely not available. From fig 1.2 B below 28% lacked ability to interpret

weather information, 21 % of the respondents claimed weather information was not reliable.

Fig 1.2: Information Availability (fig. A) and problems encountered in using information (fig. B)

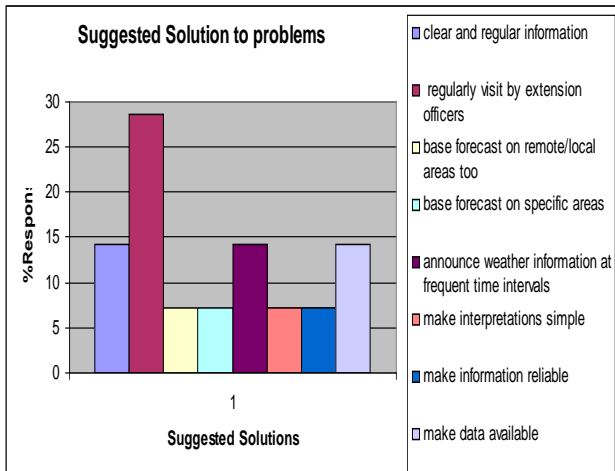


From the above results, it can be observed that information is not readily available and even in cases where it is available, it is perceived to be not accurate. Farmers cannot easily relate the given weather information and planning of the various cultivation activities. This gives the impetus for the design of an expert system with reasoning capabilities. It is also important to give the farmers control of interaction with information sources so that they can perceive better availability of information. In order to increase the perception of accuracy, it would be important to include an explanation element to such an expert system.

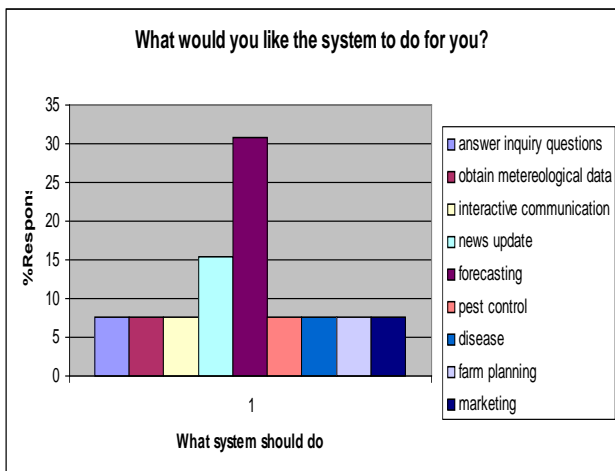
From fig 1.3 A below, 28% of respondents preferred regular visits by agricultural extension officers while 14% recommended clear information. The questionnaires also revealed that the agricultural extension officers were hardly available. From fig 1.3 B below, 30% of respondents wanted advisory system to help in weather forecast while 15% preferred news update.

Fig. 1.3: Suggested solutions (fig. A) and what they would like the system to do (fig. B)

A



B



From the above results, it can be observed that there is need to supplement the services offered by the agricultural extension officers. In general it can be conclude that an expert system incorporating agricultural and weather related data would be a suitable solution.

4.0 Model Design

From documents collected from the field and interviews with experts, rules were coded and formed the inference engine of the expert system. A ruled based system is usually preferred because of ease of understanding and interpretation [Okinyokun and Anyiam 2001]. The rules were grouped in four categories namely crop suitability, pest/disease forecast, yield prediction and advance warning. An example of the type of rules used in all categories is as shown below.

Example of Yield prediction rule

IF crop STATE is POOR and crop STAGE IS EMERGENCE and effects is INSUFFICIENT RAIN and Temperature > OPTIMUM and rainfall < OPTIMUM

Then yield is LOW

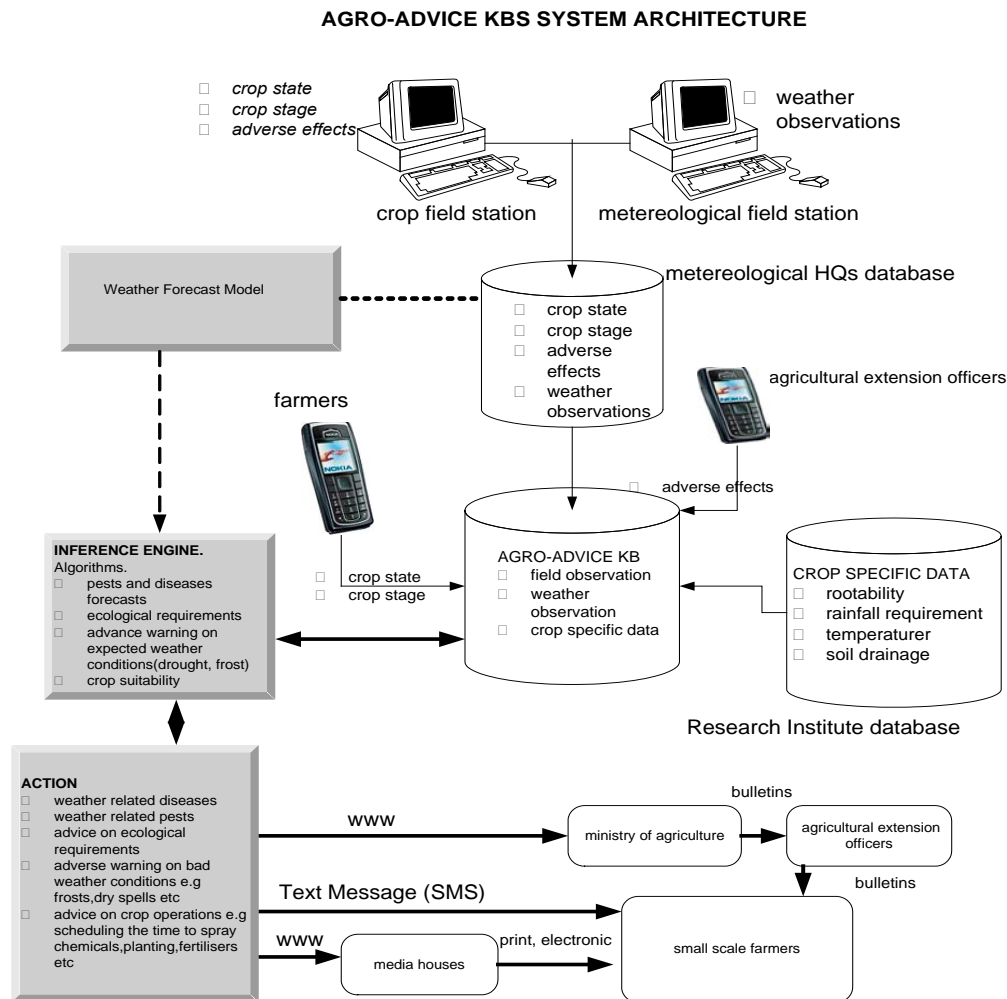
This rule shows that the yield is low if majority of reported observations by farmers, agricultural extension officers and meteorologists in the field show that the crop state is poor and crops are at emergence stage and there is insufficient rain and temperature is above crop optimum temperature and the rainfall is below optimum. All the other rules for advance warning and pest/disease forecast follow a similar pattern.

Relational Database approach was used to design the knowledge base. A similar approach was used by Okinyokun, and Anyiam [2001] to design a prototype of knowledge base system for weather monitoring and forecasting with satisfactory results. The model developed here envisages the use of two relational databases and a knowledge base. The databases would simulate the weather and agricultural research databases as currently in use by the respective institutions. The knowledge base contains crop specific information and ecological requirements. It also contains knowledge about pests and diseases. Data is entered in the knowledge base through a web based interface and text messaging using a mobile phone (SMS). Short message is convenient for small scale farmers who cannot access internet facilities. A web based interface where a user could login and post weather and crop data was used. Web technology has been proposed by Maracchi et al. [2002] in designing an internet application for agro-meteorological products in Europe. The same approach has been used by Kurnik and Susnik [2006] to come up with Slovenian agro-meteorological information system (SAGMIS). Dissemination is achieved through mobile phone and internet. As shown by the survey and analysis shown above, mobile phone would be a preferred means of information dissemination. The web interface also enables widely distributed meteorologists to login and post observations.

The Model for the Agro-meteorological System

The above design considerations and conclusions from the survey if put in context would yield to the model shown in figure 2 here below.

Fig.2: Architecture of the Model



The above model shows the flow of information from various sources such as the farmers, the agricultural research institutes, meteorological stations and agricultural extension officers. The knowledge from these sources is brought together in the Knowledge Base (KB). This is then processed by the inference engine with some the algorithms as shown in the diagram. The system can perform various actions as shown. Small scale farmers can then interact with the system through short message services (SMS). The farmers can also obtain information through

mass media. The government through respective department can also obtain the information through the internet. This enables the department to produce the necessary bulletins.

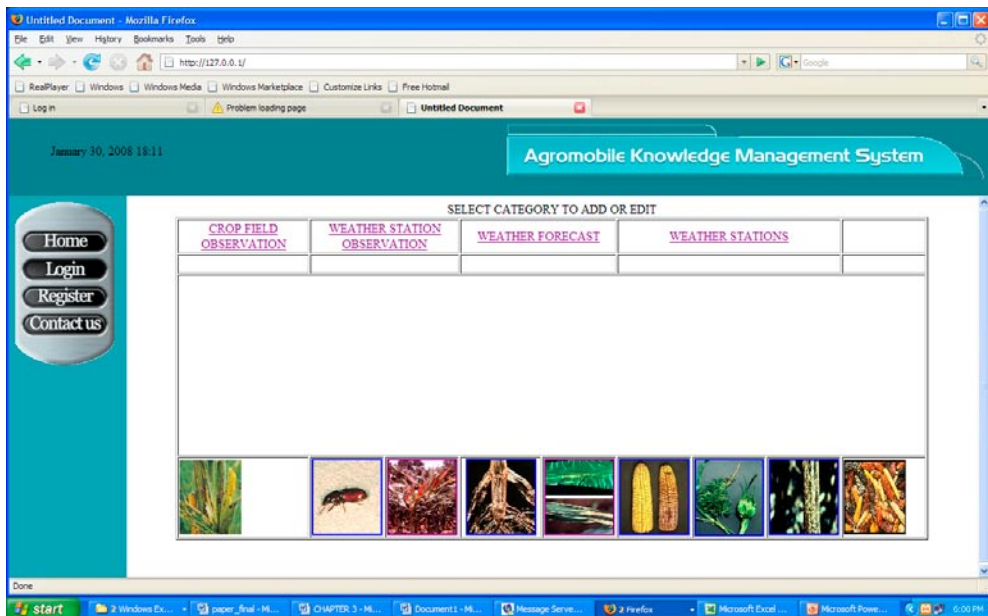
Implementation of the Model

Microsoft SQL 2005 was used for implementation of the storage component of the knowledge base. Two databases were created mainly the meteorological and agricultural research institute databases. Storage tables were created in each database and relationships established. The tool was preferred because it had appropriate features for relational database development including security and relationship facilities. A combination of PHP and HTML programming languages were used to implement the inference engine, SMS modules and user interface. Codes for inference engine rules were written using PHP. SMS scripts were also implemented using PHP. Data entry forms were implemented using PHP. The two applications were platform independent and therefore the most suitable for the model. The web based interface could allow users to login and post, edit and if permission allowed edit records in the database. On the same note, small scale farmers with little computer literacy could query the system through mobile phones. Ozeki SMS server manager was used as an interface between the inference engine and the knowledge base through storage of incoming and outgoing messages. This was a messaging platform that enabled access to mobile messaging services. A mobile phone was attached to a computer where the system was hosted to act as a gateway. Users could query by sending the messages to the gateway phone which in turn could query the knowledge base and reply by sending back the results

The prototype

The prototype has a web and SMS user interface component. Web components can be used by meteorologist and agricultural research specialists

Fig 2: The Interface used by Meteorologist



Meteorologist can login and post crop weather station data like temperature, rainfall, humidity, wind speed and sunshine. Users can also add crop field observation like crop state, stage and adverse effects. Meteorologists can also add weather forecast data. Meteorologist can also add weather station information like location, altitude, latitude, longitude and station type. For this research, SQL 2005 database was used for storage. It was preferred because it has good data manipulation mechanisms. HTML was used to implement the user interface. It was chosen because it easily implements internet applications and can be interfaced with SQL with ease. PHP was used for implementing the algorithms and other reasoning procedures and the module developed acted as the inference component of the system. PHP interfaces easily with HTML and SQL and therefore the reason for its choice. The Knowledge Base was implemented in PHP as a rule based system.

Evaluation of the Model

The goal of evaluation was to assess quality of information given by the expert system. Two simple experiments were carried out. The first was for assessing the quality of advice on crop suitability given certain climatic and ecological conditions. The second set of experiment was for predicting crop yield given some input conditions such as soil type and condition. The quality is bench marked against a human expert's response, in this case agricultural extension officer, given the same input data.

a) Crop Suitability Test

Six sets of past weather observation were entered into the system using the mobile interface. The results are summarized in table 1 here below.

Table 1: crop suitability evaluation results

Data set	System response	Agricultural Knowledge	Expert	test
Set 1	suitable	suitable		true
Set 2	suitable	Not suitable		false
Set 3	suitable	suitable		true
Set 4	suitable	suitable		true
Set 5	Not suitable	suitable		false
Set 6	Not suitable	Not suitable		true

Four out of six trials representing 66% of the tests yielded outputs that agreed with the human expert. This result shows a perceived accuracy level that is satisfactory. The actual accuracy level may be measured by planting the actual crops and determining how well the plants perform. If assumed that the two accuracy levels are insignificantly the same, then we can conclude a high level of performance.

b) Yield predictions test

Eleven sets of past crop field observation data were entered into the system. For each set, a yield prediction query was sent through mobile phone interface and results compared with expert prediction. Results are summarized in the table 2 here below.

Table 2: Yield prediction evaluation results

Data set	crop	Expected yield by agro-meteorologists	Expected yield by the system	test
Set 1	maize	high	high	true
Set 2	maize	Moderate	moderate	true
Set 3	maize	Moderate	moderate	true
Set 4	maize	Moderate	moderate	true
Set 5	maize	Moderate	moderate	true
Set 6	maize	High	moderate	true
Set 7	maize	Moderate	moderate	false
Set 8	maize	High	moderate	true
Set 9	maize	Moderate	high	false
Set 10	maize	High	true	false
Set 11	maize	Moderate	moderate	true

Eight out of eleven trials were true meaning that they tallied with the expected results from a human expert. This showed that the system has perceived accuracy level of 72%. From the results presented in Fig. 1.1 28 respondents out of a total

50 (representing 56%) indicated that the information they got was not accurate or unreliable. This shows an accuracy level of only 44%. The system has a mean accuracy level of 69% based on the two sets of experiments. This is an improvement and indicates that it's a worthwhile exercise to implement the actual system.

Conclusion

This paper has demonstrated a model that can be used to implement an agro-meteorological system that can assist farmers to obtain timely information. The information is useful in planning various farming activities. We have demonstrated how communication issues can be handled between government agencies knowledge bases, field officers in agriculture and meteorology departments and small scale farmers whose computer literacy level is assumed low. Knowledge contained in agro- meteorological repositories is difficult to interpret especially for small scale farmers. This paper has demonstrated how Expert Systems can overcome these challenges.

The paper has demonstrated a prototype that is suitable for capturing, organizing, processing and disseminating knowledge to small scale farmers, government departments and media houses. The prototype demonstrated how expert knowledge can be formulated and implemented in the form of rules to provide reasoning mechanism for yield prediction, crop suitability check, pest and disease prediction as well as scheduled warning based on weather condition. Through evaluation tests conducted on the prototype, it was showed that a system developed out of the designed model would enjoy better accuracy levels of up to 69% as opposed to 44% experienced by current means of processing and dissemination.

References

- HANSEN, J., AND JENSEN, L.E. 2002. *Readings in E-Advise for Adoption of Agricultural Advisory Services to e-Business*, Danish Institute of agricultural Information Systems, Denmark.
- JOCELYN, S., AND BURSTEIN, F. 2002. A Framework for Case-Based Fuzzy Multi-criteria: Decision Support for Tropical Cyclone Forecasting. In *Proceedings of the 36th Hawaii International Conference on System Sciences*, Hawaii, U.S.A., 2002.
- KMD. 2008. *Readings in Kenya meteorological department website*. KMD, Nairobi, Kenya
- KHALID J., AND STEVE M. 2005 Knowledge Based System for Weather Information Processing and Forecasting. In *Proceedings of the World Congress of Computers in Agriculture and Natural Resources*, Iguacu Falls, Brazil, 662-668
- KURNIK, B., AND SUSNIK A. 2006. Slovenian agro-meteorological information system (SAgMIS). In Copernicus Online Service and Information System, <http://www.cosis.net>
- MALHOTRA, Y. 2000. Knowledge Management & New Organization Forms: A Framework for Business Model Innovation, *Information Resources Management Journal*, Jan- Mar 2000, 13(1), 5-14

- MARRACHI, G. et al 2002. Internet applications for Agrometeorological Products-Experiences from Europe. In *Proceedings of the Expert Group Meeting on Internet Applications for Agrometeorological Products*, Washington D.C, USA, 6-9 May 2002.
- MUGENDA, A., AND MUGENDA, A. 2003. *Readings in Research Methods: Quantitative and Qualitative Approaches*, African Centre for Technology Studies, Nairobi, Kenya
- OKINYOKUN, C. AND ANYIAM, T. 2001. A Prototype of Knowledge Base System for Weather Monitoring and Forecasting. *International Journal of The Computer, The Internet and Management*, Vol. 9, No.1, 2001

22

Impact of Funded Research Projects on University Quality: A System Dynamics Approach

Benedict Oyo, Ddembe Williams and Erik Barendsen

Whereas there are diverse university funding sources including: funding/research grants from government, research contracts from private sector and research councils, tuition fees from students, and internally generated funds through short courses and consultancies, it is widely acknowledged that universities worldwide are facing growing difficulties in financing their activities. It is further claimed that in the times of financial constraints in universities, teaching is favourably emphasised, rendering research and development deficient. This paper explores the impact of funded research projects on perceived quality of university using systems thinking and system dynamics simulation modelling. A conceptual system dynamics model based on current literature and supported by research practice in the Ugandan universities is used to demonstrate the effects of research funding on: new research resources, staff development, quality of graduate outcomes, and demand for graduate training. The authors envisage that the model developed may provide the basis of testing various acknowledged theories on the problem of funding university research projects.

1. Introduction

It has become common for university's core activities including: teaching, research and professional services to compete for dominance. At the centre of this issue is the World Bank's [2000, p.15] claim that "the wealth of nations is no longer measured solely in terms of natural resources or level of industrialisation, but also on higher education's productivity in which knowledge is the new capital and information is the currency of exchange". In similar arguments, Williams and Van Dyke [2007] assert that due to globalisation, today's students, employers, and academics demand for indicators of international academic standing of universities in order to make decisions about where to study, whom to employ, or where to seek professional expertise. Given such expectations lies the ingenuity required in managing institutions of higher education while maintaining their effectiveness. Kennedy and Clare [1999] report that institutional managers have a crucial role to play in relation to quality in the following ways:

- Finding ways of using the institution's resources to better effect and generate more resources.

- Being accountable to the wider society, through use of effective means of assuring academic standards.
- Developing improved systems of strategic planning and institutional management.

In addressing these roles, the institutional managers must involve other stakeholders including government, quality agencies, individual academics, students, employers, parents and the society, short of which undermines the institutional ability to foresee future needs and ultimately future competitiveness.

While the body of literature qualitatively acknowledges that university research performance or productivity generates new resources [Williams and Van Dyke 2007; Geuna and Martin 2003; Vahdatzad and Mojtahedzadeh 2000; Kennedy and Clare 1999], strengthens graduate training [Williams and Van Dyke 2007; Howard and Sharobeam 2002], enhances university reputation [Vahdatzad and Mojtahedzadeh 2000; Watty 2003], and facilitates staff development, corresponding quantitative studies are lacking. In this paper, a quantitative system dynamics model is developed and used to test relationship between research projects funding, projects execution, staff development, demand for graduate training, and quality and quantity of graduate students' research.

The debate on how quality is defined, measured, and managed is one of the hottest issues in higher education studies [vanderHeyden 2001; Watty 2003; Teferra and Altbach 2004]. This paper contributes to this debate on the premise that: although universities excel because of quality, they don't necessarily survive because of quality especially in the times of financial constraints. Therefore, balancing between excellence or quality and survival is among the greatest challenges universities face today. On the other hand, research productivity intervention seems positioned at the centre of this challenge and hence an area for further inquiry. On the basis of these arguments, the next section of this paper, gives the rationale for focusing on research and quality, thereafter, a review of the Ugandan research practice is discussed, and finally the model and its ensuing simulation results are presented. Throughout this paper, university and higher education (HE) are synonymous. In the same spirit, the university quality implies: the overall quality of research in terms of quality of staff, and quality of students' research.

2. Rationale for Investigating Research Productivity In Universities

Continuous reforms, functions, and expectations in HE fuel concerns over its ability to provide quality products that rhyme or surpass business needs. A long-standing tradition of universities is the high value placed on academic freedom and institutional autonomy, and therefore full responsibility of its quality. The traditional elitist perception of HE: small size, homogeneous student body, autonomous operation, and relative political detachment [Mehrez and Mizrahi 2000] has eroded as a growing percentage of the population in all countries enters higher education. Distance education, lifelong learning, mobile learning, online programmes, internationalisation, and evening classes are recent flavours in HE.

Amidst the growing complexity of HE activities and means of education delivery, research productivity is core if quality and competitiveness are to be maintained.

The academic culture of universities generally, is tipped strongly toward research to such a level that Nossal [2006] terms it “the cult of research intensity”. To this end, developed countries such as Australia [Williams and Van Dyke 2007], UK [Kennedy and Clare 1999], and US [Howard and Sharobeam 2002], attach university funding to research performance. A more comprehensive study was undertaken by Geuna and Martin [2003] in which they compared and contrasted approaches to funding and evaluation of university research in 14 countries in Europe and the Asia-Pacific region. Their findings indicate that, in some countries, the allocation of research funds to universities now depends wholly or partly on the results of evaluation of their previous research results.

From the viewpoint of the benefits of research to a university, Vahdatzad and Mojtahedzadeh [2000] use the experience of Yazd research productivity and its future research strategy to assert that:

- Whereas the equality of schooling and its continuous improvement are critical for any educational institution, it is the quantity and quality of researchers that provides the necessary capacity for further growth.
- Research makes education more effective; it also brings money for capacity expansion. A university without the research capability has little success in attracting high quality and energetic faculty members.
- It is in fact the publication of the research that best satisfy the faculties and is the research outcomes that give the faculty members the opportunity for their tenure.

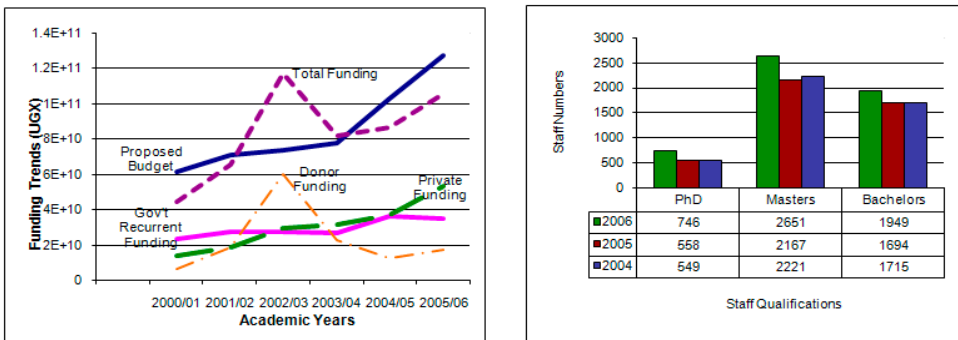
It is therefore prudent for any university to invest in research and development while simultaneously maintaining and improving the quality of teaching. Failure to balance teaching and research creates no difference between a university and a high school or knowledge creation through research and knowledge consumption through reading and teaching.

3. Research and Publications in the Ugandan Universities

Research is recognised as a parallel top priority for universities alongside teaching. The basic indicators of research in universities are: level of funding including subscription to journals and periodicals, percentage of graduate training versus graduation rates, ratio of staff with PhD, and publications rate [Geuna and Martin 2003; Teferra and Altbach 2004; Williams and Van Dyke 2007]. The focus of this section therefore is guided by these indicators in the context of Ugandan universities. First, on the issue of research funding, the National Council for Higher Education’s 2006 edition of the state of higher education in Uganda reports that much of research done in Ugandan universities is funded by donors. The report specifically shows that: public and private universities respectively spend 1.1% and 0.4% of their budgets on research. This practice is in part depicted by funding in

Makerere University (Figure 1a) in which funding trends over the academic years 2000/01 to 2005/06 was dominated by deficits yet Makerere is highly supported by donors on addition to government funding as a public university [Wabwire 2007].

Fig. 1: Funding Trends for Makerere University and Academic Staff Distribution in the Ugandan Tertiary Institutions



a) Funding trends for Makerere University
(Source: Wabwire, 2007)

b) Distribution of Academic Staff in Ugandan Tertiary Institutions
(Source: National Council for Higher Education, 2006)

By reflecting on Figure 1a, this paper contends that as university funding becomes more constrained, investment in activities with more short term visibility such as increased enrolment as confirmed by rising private funding at undergraduate level takes course. As a result, enrolment growth outstrips the university’s ability and commitment to research, rendering research productivity stagnant. Similar trends have been reported for universities elsewhere, e.g., National Council for Higher Education [2006] for the case of Ugandan universities, Vahdatzad and Mojtahedzadeh [2000] for the University of Yazd, while Mehrez and Mizrahi [2000] have sighted Israeli universities, George and McAllister [1995] for case of Canadian universities, and certainly in much of the developing countries [Tilak 2000; Salmi 1992].

On the issue of academic staff qualifications, the desired qualification of PhD for university academic staff is a dream far from reality not even in the next decade, judging by values in Figure 1b. This argument is in exception of Makerere University whose percentage of staff with PhD is increasing steadily. The portion of university staff by qualifications from the general tertiary institutions in Figure1b are respectively in the range of 92%-93% and 80-81% for PhD and Masters. Given these large numbers of staff with Maters and Bachelors, emphasis on staff development is not an option anymore but an obligation for university management. A less expensive way to achieve this is through prospecting of research projects with staff training component as discussed in section 4.2.

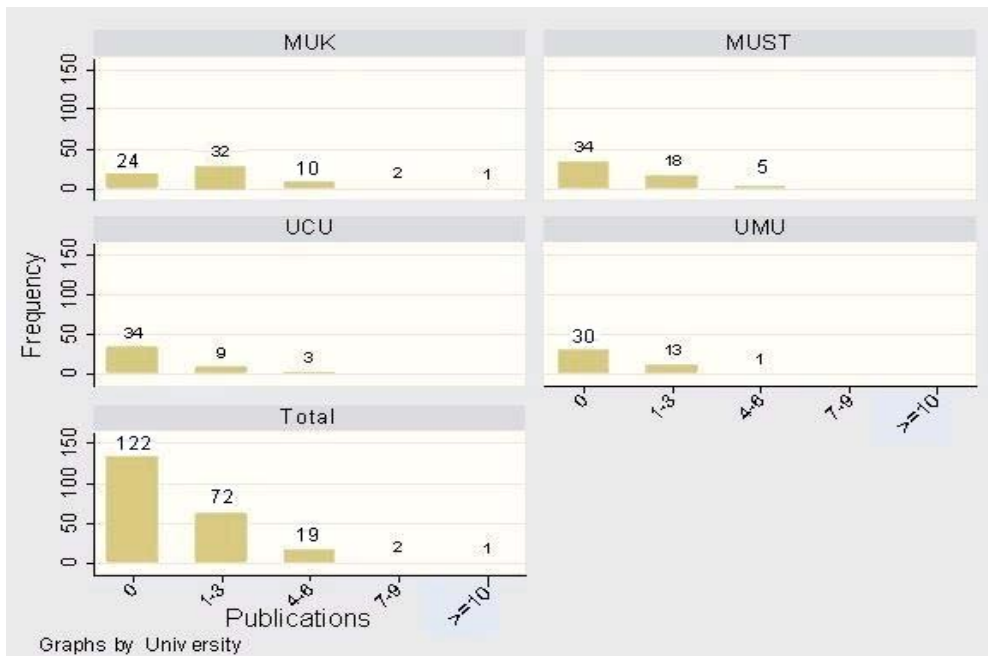
By exploring the diversity in training, it is observed that the level of graduate training (Masters and PhD) falls in the range of 2 percent to 12 percent of total

enrolment in per university with graduate training component. The focus however, shouldn't be on graduate training alone but its impact/quality in terms of 'thesis publication index' [Badri and Abdulla 2004] of Masters and PhD degrees. It is further observed that none of the universities has a thesis publication benchmark for Masters and PhD degrees, except for Makerere University that introduced journal publication requirement for PhD starting from 2007/2008 academic year.

In a recent survey of publication trends, out of 216 academic staff from two leading public and private universities, we find that 44 percent submitted at least one publication (conference paper, journal, book chapter or book) in the period 2004-2007. Narrowing down to the individual universities changes the general picture significantly as depicted in Figure 2. For instance out of 69 academic staff of Makerere, 65 percent had at least one publication in the same period. While these findings may not represent the actual publications output, they can be used as threshold for future publications analysis. Figure 2 summarises publications on a per university basis. The acronyms in Figure 2 are:

- MUK - Makerere University
- MUST - Mbarara University
- UCU - Uganda Christian University
- UMU - Uganda Martyrs University

Fig.2: Publications by Selected Ugandan Universities



Judging by the publications grouping in Figure 2, it is clear that a large section of academic staff in leading Ugandan universities do not consistently publish.

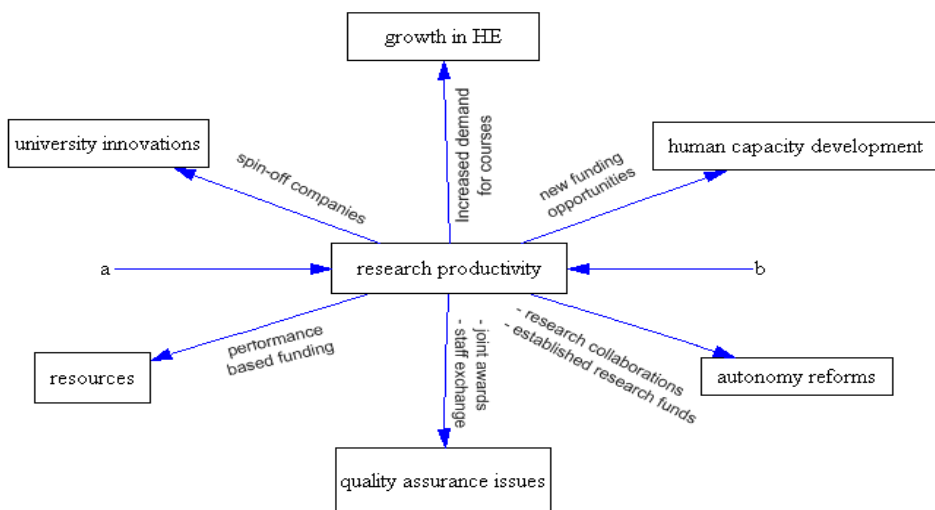
In terms of qualifications of the surveyed academic staff, 13.5% hold PhD, 67% hold Masters and 19.5% have a Bachelors. These percentages are comparable with staff numbers by qualifications in Figure 1b and hence the survey distribution is justified.

In view of the research practice and hence challenges as discussed, it is necessary to investigate the long term effect these challenges pause on quality as well as test policies aimed at offsetting them. This line of argument is further explored in this paper, first by deriving the dynamic hypothesis for research productivity and quality from the results of the survey and relevant literature, then building the model.

4. Factors Affecting Research Productivity

At the recent World Bank's regional higher education conference on strategic choices for higher education reform held in Malaysia from 3-5 December 2007 (<http://go.worldbank.org/BA1EX32U10>) several strategic issues were explored. These include: growth in HE, universities innovations systems, human capital development, research development and commercialisation, quality assurance issues, resource mobilisation, autonomy reforms, and globalisation in HE. By internalising these university strategic choices, it is possible to place research productivity at the centre and derive the rest of the strategies from it, as articulated by Figure 3.

Fig.3: Placing Research Productivity at the Centre of University Strategies



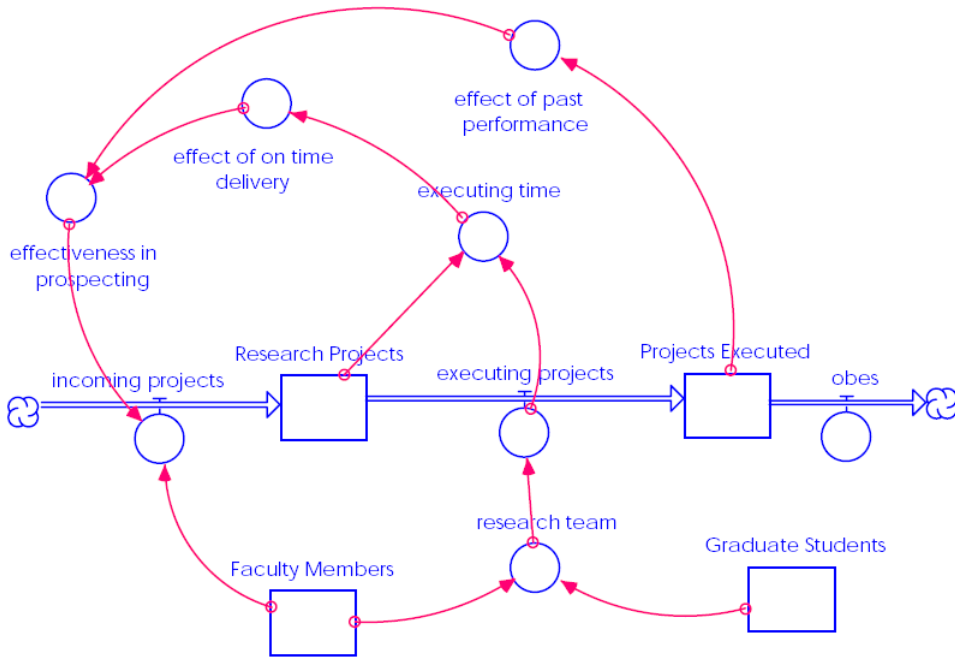
Arrows **a** and **b** in Figure 3 represent the factors that influence research productivity. Since these factors are fundamental to the contribution of this paper, they are explored in two phases, first from the literature and from Ugandan universities practice.

By examining the available literature, research productivity which also reflects outcomes in research in terms of publications, scientific applications, and income, arises from:

1. Level of research funding. Indeed, funding is acknowledged as the main determinant of research productivity across much of the literature [Mehrez and Mizrahi 2000; Martin 2003; Teferra and Altbach 2004; Frølich and Klitkou 2006].
2. Percentage of staff with PhD. A PhD qualification is perceived as the ability of the PhD graduate to do research. This is mainly supported by the fact that publication volumes are highly dominated by staff with PhD.
3. Volume and quality of graduate training. Universities world wide have benchmarks for the number of publications for PhD and in some cases Masters Award [Badri and Abdulla 2004].
4. Basic resources like computers, Internet connectivity, and journal subscriptions deter research undertaking by faculty academics [Mizrahi 2000; Martin 2003; Teferra and Altbach, 2004].
5. Efficiency of research groups. Large scale research is most successful in a collaborative environment that is natured by research groups.
6. Staff salaries. Low salaries persuade faculty academics to pursue private goals outside the university to meet their overlapping needs at the expense of active research (Vahdatzad and Mojtahedzadeh, 2000).
7. Past research reputation. The answer to the question on: “how organizational structure of the university and other research organizations influence rate of innovation and hence research productivity” lies in research reputation of the university. Numerous publications depict a general decline in research productivity in historical universities in the developing countries [Salmi 1992; Tilak 2000; Teferra and Altbach 2004] yet the same universities are most prominent in research performance in their respective countries because of previous research reputation.

Vahdatzad and Mojtahedzadeh [2000] integrate some of these factors in their structure for growth of research and development in a university clarifies the core research process in a university. This structure as given in Figure 4, illustrates that: faculty members invest time in prospecting new projects and their effectiveness depends on previous success. Projects are executed by the research team that is made up of faculty members and graduate students.

Fig. 4: A Structure for the Growth of Research and Development in a University (Source: Vahdatzad and Mojtahedzadeh, 2000)



Following from the factors that affect research productivity and insights from Figure 4, a survey to confirm the manifestations of these factors in the Ugandan context was made. Two leading Ugandan public universities (Makerere and Mbarara) and two private universities (Uganda Martyrs and Uganda Christian) were considered in this survey, in which a total of 216 academic staff were involved. In short, the results of the survey indicate low publications output as already discussed in section 3. From a more informative perspective, the following patterns were confirmed:

- The trends in publications (conference papers, book chapter, journal or book) are highly correlated with staff qualifications. Over 95% of Staff with PhD had at least one publication in the period 2004-2007, while Masters was 20% and Bachelors only 2%.
- 84% of the total publications resulted from projects funded either by government, donors or universities. The remaining 16% resulted from graduate training.
- Facilitation of academic staff for conference participation is allocated in the university budget according university management, although most faculty academics seemed not aware of such provisions.
- Research groups either do not exist in some faculties/academic units or are non functional because of funding constraints, and group commitment issues due to the culture of part-time teaching in other tertiary institutions and/or consultancy engagements.

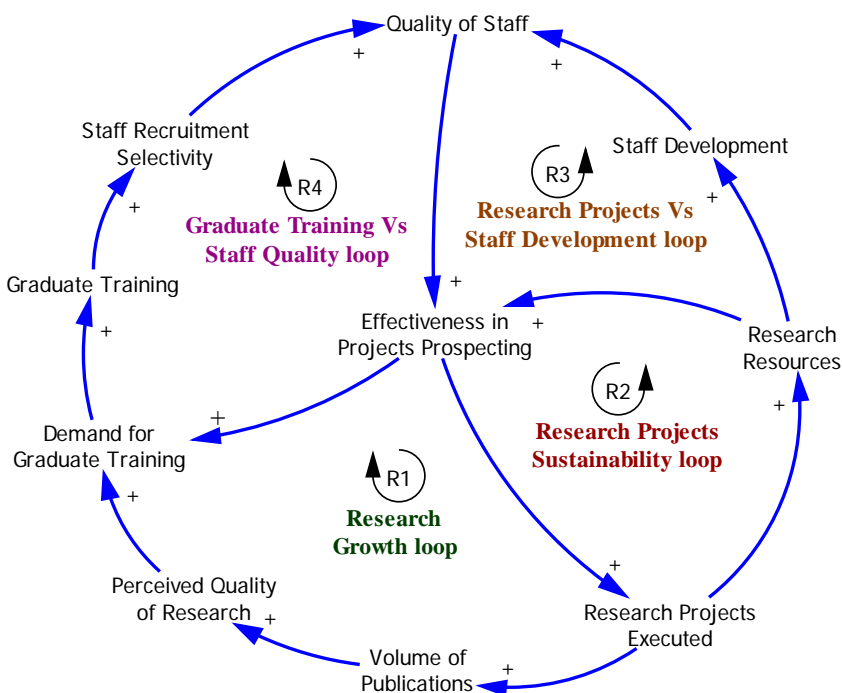
- Staff development is more prevalent in public universities. At the lower end, 8% of academic staff without PhDs are facilitated by the university for further training. The actual percentage of staff on training however, is higher than the 8% as other funding sources for training exist
- Recruitment of academic staff with higher qualifications (PhD or Masters) is favourable in faculties that offer graduate training.
- Graduate training is generally low. Graduate enrolment trends vary from desired enrolment, to no enrolments irrespective of available places for graduate training.

On the basis of these findings and relevant literature, the dynamics hypothesis, which is presented in the next section is conceptualised.

4.1. Dynamic Hypothesis

The problem being modelled is the relationship between funded research projects and quality in HE. Based on the fundamental systems premise that a system’s structure generates its behaviour, the interactions between: graduate training, volume of research, research projects execution, research resources, staff development, and quality of staff as determinants of the relationship under investigation is shown in Figure 5. The four dominant reinforcing feedback loops (R1 to R4) in Figure 5 therefore provide the theoretical analysis of the factors that underpin research productivity and quality in HE.

Fig. 5: The Dynamic Hypothesis for Research Productivity and Quality System



The dynamics in Figure 5 reflect two categories of feedback loops:

- Indicators of research productivity loops (R2 and R4)
- Effects of research productivity loops (R1 and R3)

Effects of research productivity loops (R1 and R3)

Both loops R1 and R3 associate effectiveness in projects prospecting and research projects executed with research productivity in terms of research resources and publications. Specifically, R1 articulates that: an increase in the effectiveness in projects prospecting increases the number of executed research projects. Further increase in executed research projects increases volume of publications which results into increases in perceived quality of research and ultimately demand for graduate training. Finally, an increase in graduate training increases research recruitment selectivity which in turn improves the quality of staff. The loop ends by suggesting that an increase in quality of staff increases effectiveness in projects prospecting. The loop R3 is similarly explored.

Indicators of research productivity and quality loops (R2 and R4)

The reinforcing loops R2 and R4 summarise indicators of research productivity and quality. R2 demonstrates how research projects can be sustained while R4 depicts the dependency of staff quality on graduate training.

The policies implied by the dynamic hypothesis are further expanded into stocks and flows diagrams for the model, which is presented next.

4.2. The Model

The model presented in this paper, is developed by converting the dynamic hypothesis (Figure 5) into corresponding stocks and flows (Appendix A). The model's parameters are based on literature on university research productivity (Vahdatzad and Mojtahedzadeh, 2000; Mehrez and Mizrahi, 2000; Badri and Abdulla, 2004; Williams and Dyke, 2007), and research practice in leading Ugandan universities. In keeping the model simple but without compromising relevance, the following assumptions have been made:

1. University academic staff with PhD qualifications and graduate students constitute research groups. Every research group in turn manages at least one research project at a time, and the maximum number of research groups an academic unit can have is fixed.
2. Benefits from projects executed include: new research resources, more faculty/university research funds, and more highly qualified staff through research projects with PhD training component.
3. The planned graduate students capacity is fixed since other factors beyond the scope of this paper affect the overall students' capacity that includes undergraduate and graduate students.
4. Tuition from students doesn't contribute to the research funds since it is evident that universities generally operate under budget deficits and thus research thrives only when it is funded through a separate channel.

5. A minimum number of publication(s) is mandatory for graduate students prior to their graduation and a publication is a journal or its equivalent
6. Graduate students imply Masters and PhDs only

Although communicating model structure begins with the dynamic hypothesis through influence diagrams and ends with stocks and flows, since the goal of this research is analysis of simulation results, the influence diagrams are not discussed. Rather, priority is placed on the stocks and flows from which the simulations ensue. As such, the stocks and flows diagram in appendix A, makes a point that the impact of funded research on university quality is determined by:

(a) Demand for Graduate Training

The number of applicants per place, which is established from previous admission trends of graduate students, is a major indicator of demand for graduate training. Due to the volatility of demand for graduate training, a bi-flow is used to model it. The actual demand for graduate training (denoted by T_d) in equation (1) is determined by combining: perceived quality of research (Q_r), number of applicants per place (A_p), desired intake (D_i), and effectiveness in projects prospecting (P_p) as follows:

$$T_d = A_p * D_i * (1 + Q_r + P_p) \quad (1)$$

The effectiveness in projects prospecting implies that graduate training costs are compensated by involvement in research projects, while the perceived quality of research is preferred over the job market trends since the latter's effect is short term and hence unreliable for predicting future behaviour. As indicated in Appendix A, the demand for graduate courses accumulates or depletes the "potential students" stock, which in turn determines the admission rate of graduate students.

(b) Students' research quantity and quality

The outcome of graduate training is reflected by the quality and quantity of students' publications. With respect to quantity, "students' publications" stock is used to determine volume of students' publications over time. As indicated in Appendix A, "students' publications" stock is accumulated by "students' publications" rate. This rate is obtained from "upper quartile publications per student", "staff commitment to students' supervision" and "graduation rate". The upper quartile publications are determined from historical publications trends of graduate students. It is assumed that a university sets a minimum number of journal publications or its equivalent as graduation requirements for the graduate degrees. The "staff commitment to students' supervision" is further determined as:

$$MEAN(\text{staff training indicator, perceived quality of research}) \quad (2)$$

Equation (2) is logically conceived following the survey on determinants of research productivity from four leading Ugandan universities. Since upper quartile publications are considered as "ideal publications per thesis" (refer to Appendix

A), average publications per thesis \leq ideal publications per thesis. Furthermore, “expected students’ publications” is derived from ideal publications per thesis, and subsequently “quality of students’ research” which is computed as the ratio between “research students’ publications” and “expected students’ publications” lies in the range [0, 1].

(c) Quality of Academic Staff

The model computes quality of academic staff as the mean value of the percentage of staff with PhD and research productivity. These two parameters are considered most influential determinants of quality of staff given the prevailing scenario of Ugandan universities as previously discussed.

(d) Staff Development

Staff development is associated with staff training. Staff training depends on demand for training (number of staff without PhD qualifications) and availability of training funds. With regard to these factors, staff on training is modelled as an oven stock that keeps staff with Masters and Bachelors qualifications also known as assistant lectures and teaching assistants respectively for two and three years of their training. After the respective training periods, assistant lecturers become lecturers (PhD holders) and teaching assistants become assistant lecturers and subsequently, the quality of staff increases. On the other hand, the “PhD project staff training” conveyor in Appendix A, depicts a scenario in which executing faculty research projects have a component of PhD training, thus facilitating more staff training at PhD level on addition to the general allocations for staff training.

(e) Research Project Initiation and Execution

Another set of interrelated stocks and flows in model (Appendix A) is found in the funded research projects loop. These transcend incoming projects rate, to executed projects stock through executing projects stock. Starting from executed projects stock, research resources capacity stock is derived. Subsequently, research growth factor is computed as:

$$\frac{INIT(Research\ Resources\ Capacity)}{Research\ Resources\ Capacity} \quad (3)$$

Considering the result of equation (3) to be \square , and since “Research Resources Capacity” is a stock that accumulates with number of projects executed, then \square decreases over time. Research resource availability which is determined from \square as $e^{-\square}$, undergoes exponential growth in the range [0, 1]. The “research project publications” stock and “projects income” stock are similarly derived from “projects executed” stock.

By all measures, much discussion of model structure has negligible effect on model conceptualisation unless supporting simulation results are given. The next section connects structure and behaviour of the model.

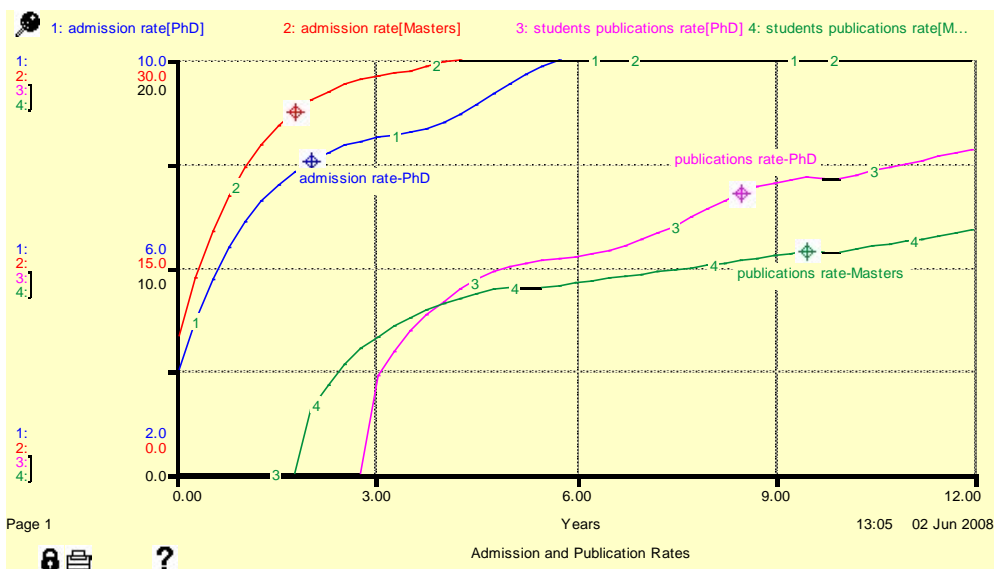
4.3. Simulation Results and Analysis

As part of the model’s behavioural validation, several simulations runs were explored. Some of these are discussed in the following subsections.

4.3.1. Students Admission and Publication Trends

Figure 6 shows behavioural trends for graduate students’ publications in correspondence with their admissions. As depicted by Figure 6, there is a general rise in graduate students’ admission before stabilising to a constant value, which is the desired admission number. This admission behaviour trends correspond with the dynamics of demand for graduate training that depends on the perceived quality of research and effectiveness in project prospecting. As these parameters improve (compare with graph1-Figure 9), both admission numbers for Masters and PhD students’ rise until the maximum/desired value is attained. As for the publications rates, the initial values before the first admitted students graduate are zero (refer to graph3 and graph4-Figure 9).

Fig. 6: Behavioural Trends for Graduate Students Admission and Publication



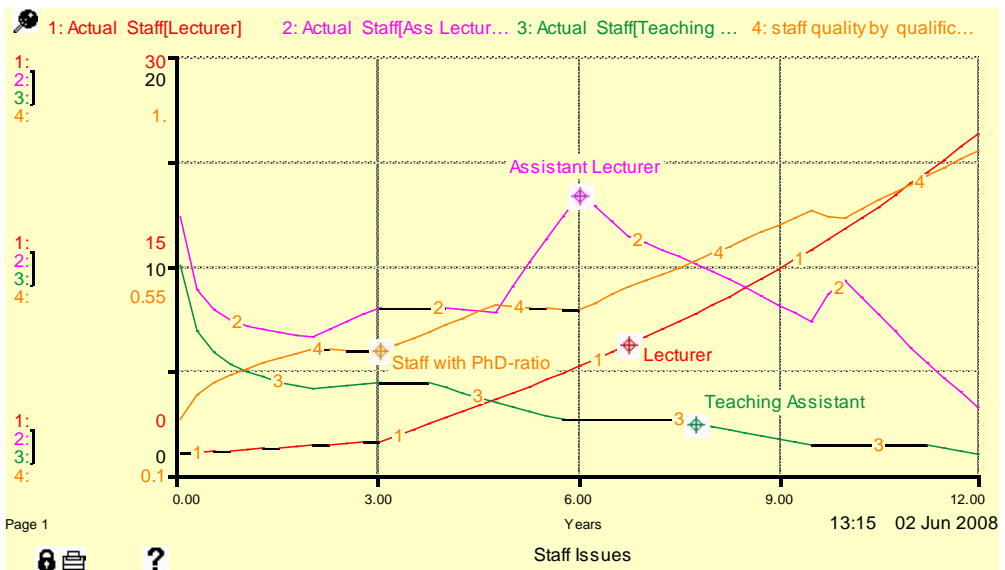
In contrast with graduation rate (arising from admission rate), it is observed that publications per thesis for PhD are higher than PhD graduation rates while the reverse is true for Masters. This observation is justified by the upper quartile publications index values as already highlighted, which are higher for PhD thesis. For this reason, it is imperative that academic units prospect research projects with PhD training component since the outcome of such training directly benefit both the academic units (in terms of publications) and the students.

4.3.2. Staff Capacity Trends

The behaviour in Figure 7 demonstrates the effect of staff development and graduate training on composition of academic staff and hence their quality in terms of the ratio of staff with PhD qualifications (lecturers). As indicated, the number of lecturers increases steadily while that of assistant lecturers and teaching assistants generally decrease over the simulation period. These behaviours are caused by two logical factors in the model:

- The staff recruitment ratio for assistant lecturers and teaching assistants decrease as Masters and PhD students graduate. This is modelled using 'if-then-elseif-then' conditions;
- When funds for staff training are available, assistant lecturers and teaching assistants in service, benefit from this opportunity such that after training, assistant lecturers become lecturers while teaching assistants become assistant lecturers and ultimately lecturers after another training period.

Fig. 7: Dynamic Behaviour for Academic Staff Capacity



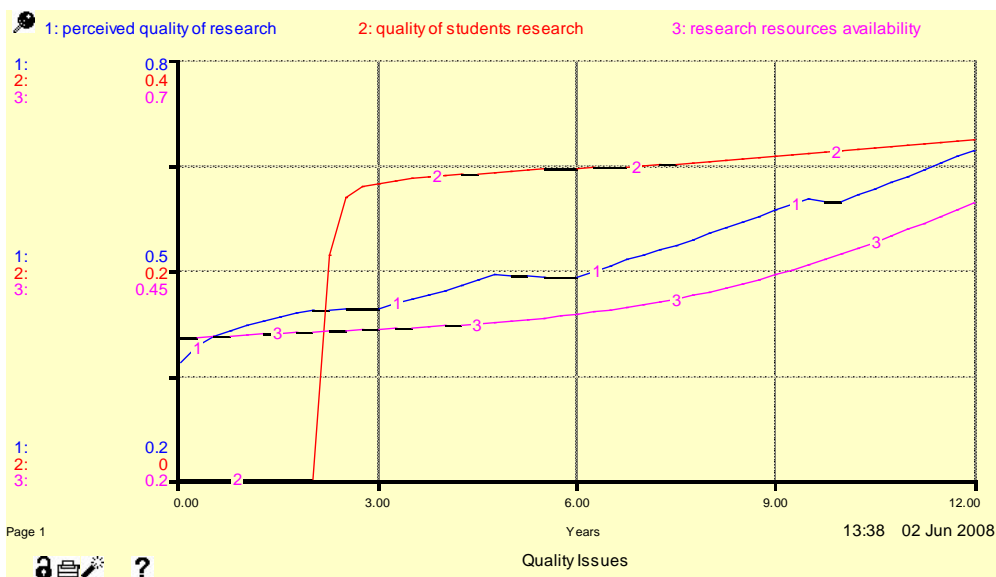
Graph 4 in Figure 7 depicts a scenario in which the percentage of staff with PhD increases over time. The stabilising behaviour of graph 4 in the sixth and tenth year reflect recruitment of staff (e.g., assistant lecturers) without PhD qualification. Whereas the actual staff capacity arises from the established staff to students' ratios, which include undergraduate students, the current model uses a staff growth constant as an equivalent for staff demand, which is subsequently used to compute staff recruitment rate. The ensuing behaviour in Figure 7 therefore affirms the assumption that recurrent staff recruitment ratio (teaching assistants, assistant lecturers, and lecturers) improves as graduate students complete their studies,

since a pool of highly qualified graduates will then be available, i.e., as more PhD graduates are available, the recruitment pattern is scaled upwards.

4.3.3. Quality Issues

Following from the dynamic hypothesis, behavioural trends for perceived quality of research, quality of students' research, and research resources availability, are reflected in graphs 1-3 in Figure 8. Specifically, graph 3 undergoes exponential growth that depends on projects executed (compare with graph 3 in Figure 9). Quality of students' research (graph 2-Figure 8) stabilises at 0.4, which is far less than the ideal value of unity. This arises because the gap between ideal publications per thesis and average or indicated publications per thesis is significant. Finally, the perceived quality of research (graph 1- Figure 8) depicts a relatively non chaotic behaviour since its determinants (graph 2 and graph 3) are non chaotic.

Fig. 8: Behaviour Trends for Quality Factors



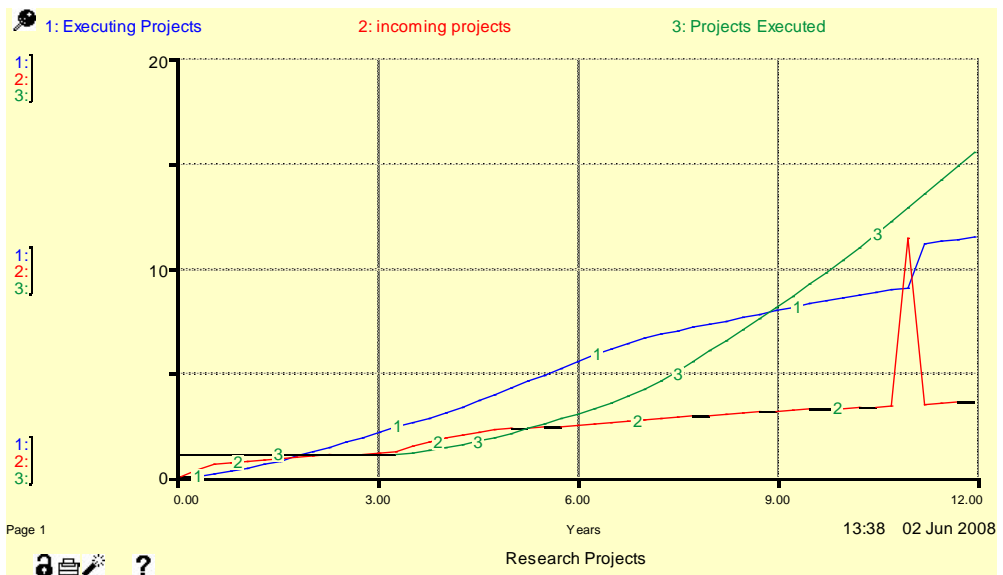
The behaviours in Figure 8 generally depict rising quality judging by the shapes of the graphs. However, since the values underpinning the graphs are in the range $[0, 1]$, insights from these graphs do not arise from the shapes but the range of values that determine these shapes as well the deviations of the final simulated value from unity. In this context, graph 1 and graph 3 have a greater effect on overall quality than graph 2.

4.3.4. Research Projects Execution Trends

The behaviours in Figure 9 depict the theory that successful research project execution depends on: quality of staff, graduate students enrolment, research funding, and effectiveness of research groups [Badri and Abdulla 2004; Vahdatzad and Mojtahedzadeh 2000]. In comparison, executing projects (graph 1-Figure 9)

increases less sharply than projects executed although both are stocks. This is supported by the structural logic in which the former out-flows into the latter and is further depleted by project failures. From another perspective, since part of the income from projects executed is used to fund new projects, this scenario has been modelled as an oven (refer to Appendix A) and is reflected by “step growth” in executing projects as depicted by the spike in the 11th year. This spike therefore, is an example of project sustainability, whereby a new project is funded using funds saved from previous projects.

Fig. 9: Research Projects Prospecting and Execution



By comparison, graph2 in Figure 9 shows the steady flow of incoming projects occurring after the third year. This corresponds with behavioural trends in Figure 6 where graduate students’ admission normalises about the same time. Similarly, the increase in gradient of graph3 (executing projects) after the sixth year matches the average period of three years to execute a research project, given that incoming projects (graph2) also stabilise after the third year.

5. Conclusion

As stated initially, this paper has explored the effects of funded research projects on staff development, quality of graduate outcomes, demand for graduate training, new research resources, and funding of future research. Subsequently, a conceptual model was implemented using system dynamics approach and ensuing simulation results discussed. The simulation results have shown that integrating the factors that influence research productivity increases quality of graduate outcomes while simultaneously strengthening demand for graduate training as well as future research projects prospecting. The paper has further, demonstrated the relevance of theory in the development of quantitative system dynamics models for quality

in higher education. Specifically, the model described in this paper extends Vahdatzad and Mojtahedzadeh's [2000] qualitative model on university research and development.

If funding mechanism influences strategies universities make, and since there is no guarantee that research funding may improve in the near future as in the case of Uganda where pressure for expansion is in conflict with limited resources, then the impact of funded research projects on quality is hard to delineate. Nevertheless, the theory that income from research projects can lead to research sustainability as has been modelled and discussed in this paper, is worth exploring in practice.

23

Tertiary ICT Education in the East African Community (EAC): Human Capital and Curriculum Development

Rodrigues J Anthony and Wafula J Muliaro

The liberalization of education has fostered an unprecedented demand especially at the tertiary level. In several deliberations, human capital ultimately emerges as the decisive factor governing the success of any endeavor of scale and substance. In the Information and Communications Technology (ICT) era, as indeed in any other era, the needs of all people cannot be completely fulfilled. Even now the power of ICT has been fully harnessed to apply to a relatively small portion of humanity even though nations, institutions, and individuals are continually thinking of how to use this technology to reach the marginalised. The principal difficulties include the galvanizing of people and aligning different cultures to forge ahead with initiatives that converge on endeavors to satisfy human aspirations, be they physical, emotional, mental, or even spiritual. The development of human capital with the requisite skills, understanding, motivation, vision, focus and commitment without vexing restrictions or constraints are essential attributes for eventual success. The area for human resource training has been foremost in the minds of the planners. Higher funding levels, the expansion of public institutions of higher learning, and the recognition of the role of the private tertiary education sector to enable a more effective and fast-tracked training of the requisite manpower have been most obvious steps. In the drive to achieve manpower training effectiveness, private education sector has been encouraged, resulting in timely approvals to provide university level courses, especially in capital intensive and or marketable fields like Computer science, Information Technology, Information Systems, Software Engineering, Communication, and Business Administration. This paper analyses twelve tertiary institutions (private and public) of EAC with a focus on both human resource and curriculum development and presents a snapshot of ICT education status in which the identified opportunities, risks and challenges are discussed. The paper also suggests solutions that could be considered at sub-regional level.

1. Introduction

Human capital has ultimately emerged as the decisive factor governing the success of any endeavor of scale and meaning. In the Information and Communications Technology (ICT) era, or indeed any era, the needs of all people cannot be completely fulfilled. Countries, corporations, and individuals are constantly thinking of using this technology to carry them forward. The principal difficulties include the galvanizing of people and aligning different cultures to forge ahead with initiatives to converge on a meaningful endeavor to satisfy the human

aspiration, be they physical, mental, or spiritual. The development of a pool of relevant skill workers with the requisite skills, understanding, motivation, vision, focus, commitment, and without vexing restrictions are essential attributes for eventual success.

The area for human resource training has been foremost in the minds of the planners [Jonhendro et al, 2002]. The injection of more money, the opening of more places in institutions of higher learning, and the opening up and recognition of the private tertiary education sector to enable a more effective and fast-paced training of the requisite manpower has been most obvious. In the big push to achieve manpower training effectiveness, the private education sector has been given a boost, resulting in many of them being quickly given the approvals to provide university level courses, especially in the emphasized priority fields like computing, communication, and business.

There is need to consciously promote the creation of knowledge workers, builders, and innovators. This human capital must, in the course of their development, be imbued with an open mindset, a positive attitude of problem resolution, a keen eye for current trends, drive and energy for embarking on new initiatives, and a commitment to personal excellence.

ICT education is a young field and not in itself a unique discipline [Almstrum et.al. 2004]. As it matures, the field is developing its own unique perspectives. It is for this reason the new approaches to its curriculum development are being considered and proposed in this paper. Clement [2004] reports that pedagogical methods resulting from research in one field can be revised and transferred to another.

This paper analyses twelve tertiary institutions (private and public) of EAC with a focus on both human resource and curriculum development and presents a snapshot of ICT education status upon which the identified opportunities, risks and challenges are discussed. The paper also suggests solutions that could be considered at sub-regional level including setting up an Information Management Centre, the Education Technology Center, and Center for Instructional and Technology.

The paper also seeks to establish if the general lack of strategic direction (policy instruments) and leadership in ICT education could be a contributor to the observed status.

2. Human Capital

Although terrorism is a major threat for many nations the greatest danger facing most of them could be the challenge of human capital development. Human capital theory suggests that the human capital of a nation or organizations is an important factor in its economic success, implying that human capital must therefore be of great interest to politicians, economists and training and development specialists [Elkin 2004] The latter study on OECD shows that human resource development as a factor enhances labour market flexibility, facilitates structural adjustment in economy and cause prosperity. On the same footing, Teitel (2006) observed that

the rapid accumulation of human capital, development of technical institutions and public policies in support of enterprise development and innovation, led to the emergence of advanced technical capabilities in a number of semi-industrialized countries. The significant human capital developed in the ICT sector in Asia brought about the so-called miraculous Asian economies with high growth rates (Anonymous, 2006). For example, the application of ICT in the activities of Malaysia's economy contributes significantly to its productivity growth, and the growth of the economy.

Ziliak (2007) argues that there is a strong tie between educational attainment and economic status, hence intensive investments in human capital are needed over the course of individuals' lives to address persistent poverty especially in developing countries. Jonhendro et al [2002] state in their paper that financial success and prosperity usually come with the persistent persuasion of all people to adopt and use any innovation that will increase productivity. In the Information and Communications Technology (ICT) era, or indeed any era, the needs of all people cannot be completely fulfilled. Countries, corporations, and individuals are constantly thinking of using ICT to carry them forward. The principal difficulties include the galvanizing of people and aligning different cultures to forge ahead with initiatives to converge on a meaningful endeavor to satisfy the human aspiration, be they physical, mental, or spiritual. The development of a pool of relevant skill workers with the requisite skills, understanding, motivation, vision, focus, commitment, and without vexing restrictions are essential attributes for eventual success.

Thus, there is need to consciously promote the creation of knowledge workers, builders, and innovators. This human capital must, in the course of their development, be imbued with an open mindset, a positive attitude of problem resolution, a keen eye for current trends, drive and energy for embarking on new initiatives, and a commitment to personal excellence. The question is, does our EAC institutions ICT education curriculum instill and guarantee output of such characters?

For all practical purposes it is sufficient to say that a knowledge worker exemplifies two organic intertwining attributes: attitude and respect to a vocation and the command of basic knowledge and skills. It is insufficient just to have only one of the attributes. A knowledge worker must have some basic knowledge in some fields and aims at using that existing knowledge to help discover new information or how to do things that have not been done before. A knowledge worker has to be innovative rather than just being productive hence need to inquire, analyze, conceptualize, reason, synthesize, construct, and be willing to impart new knowledge, skills, or wisdom to others.

In order for worker to have all these attributes, a predisposing learning environment must be created in their educational development. This is the role of our education institutions. It is common to get the feeling of lecturers that our present students are a lot of dependent, unoriginal, non-innovative, and selfish

individuals who lack the ability to do something in order to improve. Tendencies to copy, plagiarize, regurgitate, and like to be spoon-fed are some of the words used whenever assessments of students' work and attitude are made. With a view to produce competent knowledge workers, the curriculum will have to be structured with emphasis on communication skills, critical thinking skills, creative skills, problem solving skills, social skills, and life skills. Collaborative learning can be introduced into the curriculum and learning process to mimic life situations. Cooperative and collaborative learning help students to develop the requisite skills in teamwork, while life skills prepare students to cope in an increasingly stressful and changing environment. Such skills prepare students to face the challenges of the work place.

3. Curriculum

ICT education is not in itself a unique discipline [Almstrum et. al. 2004]. As it matures, the field is developing its own unique perspectives. It is for this reason the new approaches to curriculum development are being considered.

ICT education is challenged by several demands such as globalization that raises the importance of certain attitudes and transferable skills and social competences of graduates. The EAC sub-region through Inter University Council of East Africa (IUCEA), just like the Bologna Process that is aimed at creation of a common European Higher Education Area (EHEA) by 2010 [Heitmann, 2005], can do the following:

- Facilitate mobility of students and staff and professional mobility of graduates,
- Promote internationalization and global competitiveness,
- Raise quality of ICT education and contribute to economic development and growth,
- Enhance the EAC Integration.

According to Heitmann (2005), the paradigm shift to outcomes orientation and student learning have recently fostered the use of systematic and comprehensive approaches. Heitmann (2005) observes that the experience of undertaking a major curriculum revision points to the fact that the "curriculum as planned" is not yet the "curriculum as implemented" and for sure differs later on from the "curriculum as experienced by students and staff". The successful implementation of a comprehensively and systematically planned new or revised curriculum requires to a certain extent an organizational development and a change of action and behaviour of the persons involved.

Jonhendro et al [2002] argue that it is imperative that education institutions reconsider the necessary ingredients to make learning effective, relevant, and fun for the students. Hence, there is need for EAC sub-region to create a knowledge ecosystem, comprising students, lecturers, administrators, parents, partner institutions, employers, and the entire community at large. Such ecosystem can

provide the necessary enabling environment for establishing and promoting Learning Communities (LC).

Scharff and Brown [2004] reported that effective Learning Communities could lead to higher academic achievement, better retention rates, diminished faculty isolation, and increased curricular integration. As an import to ICT education, Learning Communities can provide a means for tying together courses to help students better understand the connections between ICT education and other fields, as well as across different areas of computing. An interesting pedagogical development in higher education has been the advent of the Learning Community (LC). For instance, the connection between computer science and logic is as significant as the connection between computer science and mathematics. Courses that frequently make use of and refer to logic concepts include Functional Programming, Formal Languages, Automata Theory, Artificial Intelligence, Databases, and Software Engineering. Examples of connections between logic and computer science include pre- and post-conditions, invariants, recursion, symbolic computation, and formal specification and verification. Many mathematical concepts are also basic concepts in logic. These fundamental concepts are very important for computer science students, but tend to be difficult for novices to learn. The opportunity to emphasize the connections between computing and logic suggests the power of including Learning Communities in the computer science curriculum, both as a curricular enhancement and to address issues important to computer science educators.

The society expects college graduates to be able to think critically, solve complex problems, act in a principled manner, be dependable, read, write and speak effectively, have respect for others, be able to adapt to change, and engage in life-long learning. However, according to Scharff and Brown [2004] students undertaking the traditional curriculum generally do not meet expectations, most importantly in three key cognitive areas: using abstract symbols, epistemology, and principled, ethical reasoning – as well as the ability to work cooperatively in teams with different people. They further point out that it is unlikely that the traditional curriculum can lead to the needed changes. They offer two reasons for this. One reason is the relative isolation of topics that can occur during instruction, which deprives students of important opportunities to develop their abilities. Regardless of how well a discipline relates to other topics, instructors often present content without reference to related disciplines. This approach leads to difficulties when students need to apply the skills and concepts they have learned, which they will almost certainly need to do while studying a variety of other disciplines. It also has the effect that students do not develop their abilities. Secondly, the traditional curriculum fails to bring about the needed changes on the grounds that education is often seen as merely fact acquisition, a view frequently connected with the traditional lecture. The goal of education is to improve minds, enabling them to acquire abilities and skills to do things they could not do previously.

Learning Communities (LC) as a possible curricular alternative deserves consideration and has been recommended by Scharff and Brown [2004]. One typical mission for LCs is to develop a sense of community. Other goals and practices are to:

- (1) Incorporate and value diversity,
- (2) Share a culture,
- (3) Foster internal communication,
- (4) Promote caring, trust, and teamwork,
- (5) Involve maintenance processes and governance structures that encourage participation and sharing of leadership tasks,
- (6) Foster the development of young people, and
- (7) Have links with the outside world

LCs commonly follows one of three models below (Scharff and Brown, 2004):

- The integrated model, where courses are integrated into one interdisciplinary course, or
- The linked course or cluster model, where the curriculum materials in two or more stand-alone courses establish a series of connections between the interwoven courses, or
- The student cohorts/integrative seminar model

LCs that follow the linked course model allow the faculty involved to plan their respective curricula so that students will have planned and supervised opportunities to discover a variety of connections or interrelationships. Developing and elucidating two or three such connections or relationships is generally all that is needed to put to rest the notion of the stand-alone course. Because two – or perhaps three – subjects are involved, the faculty must cooperate to plan the curricula together. As one faculty member develops a particular topic, his or her colleague can refer to, or in some other way make use of, that topic. Thus, not only are connections between the subjects made explicit, but also much of what is learned in each subject turns out to reinforce or help illuminate what is learned in the other.

Beyond the content-related goals of each course, the LC becomes a learning environment that should lead students to as many of the following objectives as possible:

- Acquire a deeper understanding of course materials by making connections between courses and disciplines, in part by learning to transfer skills, concepts, and ideas learned in one discipline to appropriate uses in another discipline;
- Learn to find similarities in increasingly disparate subject areas in order to find better solutions of a broader range of problems;
- Import and export knowledge from one course to another;

- Experience increased interaction with other students and faculty, as well as a rich learning-centered community inside and outside classroom;
- Participate in active and collaborative learning; and
- Explore and begin to understand diverse perspectives.

Indeed, LCs provide many benefits for students: increased academic achievement, better retention rates, greater satisfaction with college life, reduced time to degree, improved quality of thinking and communicating, better understanding of self and others, and greater ability to bridge the gap between the academic and social worlds. In addition, there are benefits for faculty: diminished isolation, shared purpose and cooperation among faculty colleagues, increased curricular integration, a fresh approach to one's discipline, and increased satisfaction with their students' learning.

A difficult issue for assessment under any circumstance is determining whether students have learned material well enough to apply it practically. Therefore, a LC that combines a computing course with a course in another area, for example business that would give the ICT professors insights into whether their students had learned a skill or concept well enough to apply it in "real life" situations in the other course is recommended.

LCs within the ICT education curriculum could allow courses to provide mutual support for one another. For instance, even when there are prerequisites, computer science topics are often taught as if they are relatively independent of one another. It's a good idea to link computer science courses such as discrete mathematics with data structures, algorithms and data structures with databases, databases with software engineering, and data structures with operating systems.

Introducing LCs in the curriculum can also serve as a tool for programme recruitment. Faculty wishing to attract good students to their disciplines must offer courses that will attract those students. A well thought out LC with an attractive title can have that effect. Students intimidated when faced with studying a difficult subject might find that subject less frightening if it were linked with another subject that they perceive as more accessible. At present, interdisciplinary pairings appear to be particularly interesting and important in attracting students to computer science, as many students apparently pursue computer science because of the influence of computers in other fields. In addition, the interdisciplinary approach might help attract women in ICT computer science and education in general. This paper raises fundamental questions such as why has LC not been considered in any of the EAC institutions of higher learning ICT education despite the numerous benefits and solutions it provides to most of the existing challenges? What policies need to be put in place at sub-regional level that can create an enabling environment for an appropriate ICT education curriculum to be developed? How can EAC institutions of higher learning establish a roadmap to acquisition of ICT human capital needed to spur economic development?

4. Analysis of EAC Institutions of Higher Learning

This section presents the analyses of twelve tertiary institutions (private and public) of East African Community (EAC) namely: Daystar University (private), Gulu University, Hubert Kairuki Memorial University (private), Islamic University in Uganda (private), JKUAT, Kyambogo University, Makerere University, Moi University, Mzumbe University (private), Tumaini University (private), Uganda Martyrs University (private), and University of Nairobi. It focuses on both human resource and curriculum development and presents a snapshot of ICT education status as published by IUCEA Report [2007] and discussed in this paper.

4.1. Challenges, Risks and Opportunities

4.1.1. Resource Development

EAC Sub-region institutions of higher learning require more academic, technical and administrative staff if quality ICT education is to be offered. Most ICT education departments in East Africa sub-region are faced with several challenges in maintaining a high quality team of academics. According to the IUCEA report [2007], the challenges include: high turnover due to industry demand for ICT professionals; high remuneration of ICT professionals in the market; rapid growth of ICT departments nationally and worldwide leading to aggressive headhunting; and, emerging professional aspect of ICT.

Identified barriers to recruitment in ICT education departments included: high expectations in terms of qualifications and experience, unattractive terms of service and poor job advertisement strategy. In order to create an enabling environment for the academics, there is need for sound and inspiring leadership; competitive terms of service; opportunities for advancement; recognition; and a collegial atmosphere. Adoption of a remuneration package that is cognizant of market realities could help surmount this barrier.

Human resource development in EA sub-region requires mobilization of internal resources so as to reduce dependence on development partners. Hence universities need to develop ICT policies in line with their vision. Also adoption or use of standard capacity building framework as in the case of University of Nairobi is encouraged. This would minimize the tendency of waiting to seize opportunities as they emerge and promote deliberate planning for human resource development.

A business/customer oriented approach need to be emphasized in the human resource development as a shift from the traditional ways. Employment of ICT Education staff on contract has shown a negative impact on long-term commitment.

A gender imbalance in ICT education exist and a strategy needs to be put in place to correct the anomaly. Having more female staff is one of the strategies suggested. An interdisciplinary approach promoted by adoption of LCs can help attract women both students and staff into ICT education.

There is need for promoting ICT4D network of Academic in each University as an ICT development watchdogs' mechanism, and extended to sub-regional levels. One of the objectives of AISI that states that every man and woman, school child, village, government office and business can access information and knowledge by 2010 depends on the strategy of fostering a new generation of men and women in Africa able to use ICTs to leverage the development of their nations, which ICT education has a major role to play. Shortage of human capital and appropriate training programs especially in ICT education in EAC sub-region could hamper the development of an inclusive and effective Information Society despite having put infrastructure in place.

E-literacy level varied immensely from country to country, institution to institution, and this has been aggravated by lack of clear and systematic mechanism of sharing the best practices and lessons learned by those who have solved such problems. There appear to be a disconnect between national ICT policy makers and the education sector, which complicates and create a challenge when it comes to buying-in the national ICT vision by the education sector.

Staff development in some of the institutions was guided by established training needs and informed experience that required investment on individuals, departments and faculties, so that the staff and the institution could acquire and utilize new techniques and technologies in order to better meet new performance demands. Endeavours are being made to motivate staff, by giving them opportunities to do research work that can increase their remuneration.

Introduction of short courses to enable staff to upgrade their ICT skills is suggested as a means of improving e-literacy. Also introduction of ICT loan schemes to enable university staff acquire computers and other ICT related items at an affordable rate and payment in installments was among strategies some universities have adopted towards enabling staff acquire necessary facilities for effective ICT education teaching.

4.1.2. Curriculum Development

Curriculum development (CD) involves identifying learning needs, assessing the audience, developing goals and objectives, selecting and organizing content, selecting methods and aids, facilitating learning process, and developing action plan. It is a continuous process. There should be regular reviews to ensure quality. Feedback from industry, former students and all key stakeholders is important in reviews.

Institutional collaboration is encouraged in CD. On the question of e-learning, it was noted that most institutions are taking a blended e-learning approach, where conventional teaching methods are combined with e-learning to reduce the risk of failing and increase chances of successful transfer to full e-learning.

The EAC institutions suggested Learning outcomes approach to curriculum development without specifying any particular one. We propose that the concept of LCs be embraced.

Open source software (OSS) is not yet popular in most EAC sub-region institutions despite have a promising future. OSS has the potential to give education Institutions flexibility that cannot be found with proprietary software. More so, OSS Learning Management System (LMS) have the potential to serve the institutions in terms of knowledge expansion and collaboration in developing e-learning courses and curriculum.

There is a need for developing standards at sub-regional level upon which content, skills, products can be benchmarked. This would enable the creation of individual, institutional, national and even sub-regional identity. The sub-regional standards would preside over curriculum developed among its institutions in tune with international standards.

Integration of ICT departments and Education departments need to be explored. Currently they are structured separately, which is hindering further developments in ICT education. The EAC universities each needs to identify areas of focus so as to create highly specialized institutions and avoid duplication and unnecessary competition.

The EAC sub-region society as a whole embraces its cultural and linguistic diversity and supports those who are disadvantaged. A society of strong growth and competitiveness is a society of shared values, mutual respect and humanity. In order to move towards such a society, there is need to review the curricula and fine-tune them so that they respond to the dynamically changing world.

5. Discussion

The discussion made attempts to explore the link between the identified challenges, risks and opportunities and the following recommendations found in the IUCEA Report [2007].

1. An EAC sub-regional ICT in education policy be development.
2. Funding for the human capital development and infrastructure be sourced.
3. Capacity for University Staff in virtual learning systems be built.
4. Learner support systems be developed.
5. Content be development
6. A framework for harmonization and development of an EAC sub-regional quality assurance be developed that will take into consideration:
 - Identification of current subject domains in relation to ICT education body of knowledge
 - Development of sub-regional benchmarks and standards for ICT education
 - Development of key reference points for quality assurance
7. Establishment of a journal for ICT research called *EAC Journal for ICT Development*.
8. Development of an IUCEA portal for dissemination of information.

9. Promotion of OSS for specialized applications.

10. Support ICT R&D incubation and entrepreneurship.

On analyzing the twelve universities of EAC sub-region, it was found that there are some issues that were addressed by all the institutions that we consider critical and deserve attention. These are:

- Sound and inspiring leadership
- Development of ICT policies in line with vision
- Promotion of deliberate planning for human resource development•
Development of human capital and appropriate training programs towards an inclusive and effective Information Society.
- Learning outcomes approach to curriculum development.
- Integration of ICT departments and Education departments.

The way forward from the IUCEA Report gives general recommendation that if implemented would tackle most of the challenges raised. However, we would like to argue further on one major issue: A curriculum development approach that would promote ICT human capital development.

As stated earlier, there is need to consciously promote the creation of knowledge workers, builders, and innovators. This human capital must, in the course of their development, be imbued with an open mindset, a positive attitude of problem resolution, a keen eye for current trends, drive and energy for embarking on new initiatives, and a commitment to personal excellence. The concern is that all the EAC institutions ICT education curriculum in the IUCEA report [2007] do not guarantee output of such characters since they are products of the traditional curriculum development approach. We challenge the EAC ICT education sector to create the necessary predisposing learning environment that can output the desired human capital.

According to Heitmann [2005], the paradigm shift to outcomes orientation and student learning have recently fostered the use of systematic and comprehensive approaches. A successful implementation of a comprehensively and systematically planned new or revised curriculum requires to a certain extent an organizational development and a change of action and behaviour of the persons involved. Hence, there is need for EAC sub-region to create a knowledge ecosystem, comprising students, lecturers, administrators, parents, partner institutions, employers, and the entire community at large. Such ecosystem can provide the necessary enabling environment for establishing and promoting Learning Communities (LC).

Scharff and Brown [2004] reported that effective Learning Communities could lead to higher academic achievement, better retention rates, diminished faculty isolation, and increased curricular integration. As an import to ICT education, Learning Communities can provide a means for tying together courses to help students better understand the connections between ICT education and other fields, as well as across different areas of computing. LCs that follow the linked course model allow the faculty involved to plan their respective curricula so that

students will have planned and supervised opportunities to discover a variety of connections or interrelationships. Developing and elucidating two or three such connections or relationships is generally all that is needed to put to rest the notion of the stand-alone course. Because two – or perhaps three – subjects are involved, the faculty must cooperate to plan the curricula together. As one faculty member develops a particular topic, his or her colleague can refer to, or in some other way make use of, that topic. Thus, not only are connections between the subjects made explicit, but also much of what is learned in each subject turns out to reinforce or help illuminate what is learned in the other.

LCs within the ICT education curriculum could allow courses to provide mutual support for one another. For instance, even when there are prerequisites, computer science topics are often taught as if they are relatively independent of one another.

We observe that LCs have not been considered in any of the EAC institutions of higher learning ICT education despite their great potential for numerous benefits and solutions to most of the existing challenges faced by EAC institutions of higher learning.

6. Conclusion

The EAC institutions need to focus on creation of knowledge workers, builders, and innovators. Such human capital can only be attained if suitable curriculum is put in place. A curriculum that produces characters with open mindset, positive attitude of problem resolution, keen on current trends, and with commitment to personal excellence.

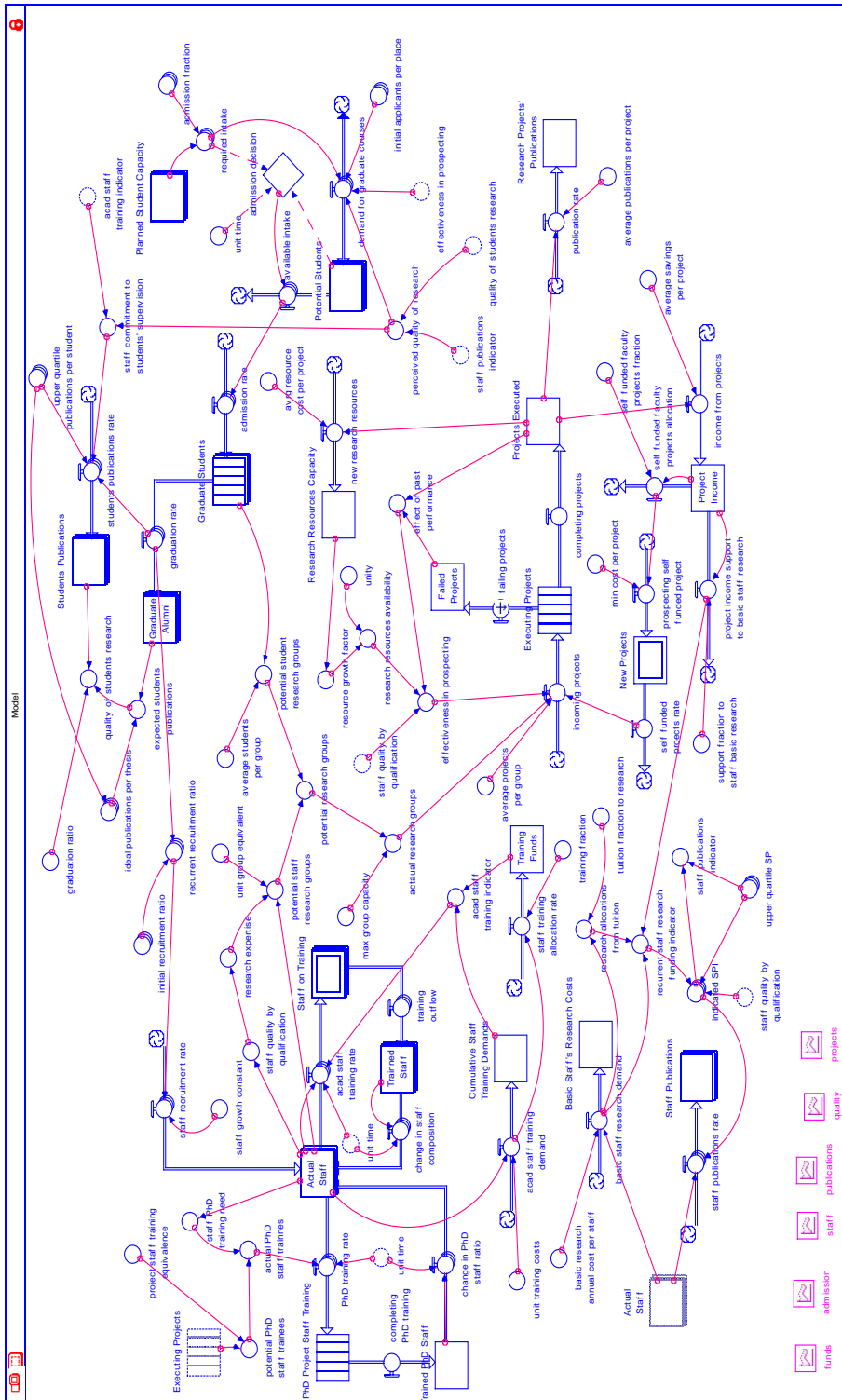
The EAC institutions ICT education curriculum stated in the IUCEA report [2007] cannot guarantee output of desired characters due to the fact that they are products of the traditional curriculum development approach. We challenge the EAC ICT education sector to embrace a new paradigm shift in curriculum that have room for LCs as well as create the necessary predisposing learning environment that can output the desired human capital.

References

- ALMSTRUM V.L., HAZZAN O. AND GINAT H (2004). Special Issue on Import/Export Relationships to Computer Science Education Research. *Computer Science Education* Vol. 14, No. 4, pp. 251–256.
- ANONYMOUS. (2006). The impact of ICT and human capital on achieving knowledge-based economy: applications in Malaysia's economy. *World Review of Science, Technology, and Sustainable Development*; January 2006, Vol. 3 Issue: Number 3 p270–283
- CLEMENT J.M. A Call for Action (Research): Applying Science Education Research to Computer Science Instruction. *Computer Science Education* 2004, Vol. 14, No. 4, pp. 343–364.
- ELKIN G.(2004), New Zealand human capital development and structural reform .2:ISSN 1360-3736

- HEITMANN G. (2005). Challenges of engineering education and curriculum development in the context of the Bologna process. *European Journal of Engineering Education* Vol. 30, No. 4, December 2005, 447-458
- GO'RNIK-KOČIKOWSKA K (2006). Ethics of global ICT society
- IUCEA Report on ICT Education in EAC. Accessed on April 14, 2006
- SCHARFF C. AND BROWN H. (2004). Thinking Through Computing: The Power of Learning Communities. *Computer Science Education* Vol. 14, No. 4, pp. 297-320
- TEITELS (2006). On semi-industrialized countries and the acquisition of advanced technological capabilities. *Econ. Innov. New Techn.*, 2006, Vol. 15(2), March, pp. 171-194.
- WEBBER C.F., (2003). New technologies and educative leadership. *Journal of Educational Administration* Vol. 41 No. 2, 2003 pp. 119-123
- ZILIAK J.P.(2007). Human Capital and the Challenge of Persistent Poverty in Appalachia. USA.
- GEORGE, P. J., AND MCALLISTER, J. A. 1995. The expanding role of the state in Canadian universities: can university autonomy and accountability be reconciled? *Higher Education in Management* 7, 309-321.
- GEUNA, A. AND MARTIN, B.R. 2003. University Research Evaluation and Funding: An International Comparison. *Minerva* 41, 227-304
- FRØLICH, N.AND KLITKOU, A. 2006. Strategic management of higher education institutions: performance funding and research output. Paper presented at the Conference on Indicators on Science, Technology and Innovation, Session of Indicators for strategic management of higher education institutions, November 15-17, in Lugano, Switzerland
- KASOZI, A.B.K., 2006. A proposed quality assurance framework for institutions of higher learning in Uganda. *Paper presented at a Seminar on quality assurance in tertiary education*, Servres, France, 3-5 June.
- KASOZI, A.B.K. 2003. *University Education in Uganda: Challenges and Opportunities for Reform*. Fountain Publishers, Kampala.
- KENNEDY, M. AND CLARE, C. 1999. Some Issues in Building System Dynamics Models for improving the Resource Management Process in Higher Education. *Paper presented at the seventeenth International System Dynamics Conference*, Wellington, New Zealand July 30-23.
- KENNEDY, M. 1998. Some Issues in System Dynamics Model Building to Support Quality Monitoring in Higher Education. *Paper presented at the Sixteenth International System Dynamics Conference*, Quebec, Canada. July 20-23
- MEHREZ A. AND MIZRAHI S. 2000. Quality requirements in rapidly growing higher education systems: the Israeli example. *Higher Education Policy* 13, 157-171
- National Council for Higher Education-Uganda (2006). *The State of Higher Education*. Kampala: National Council for Higher Education
- NOSSAL, K.R. 2006. A question of balance: The cult of research intensity and the professing of political science in Canada, *Canadian Journal of Political Science*, 39:4

- SALMI, J. 1992. The Higher Education Crisis in Developing Countries: Issues, Problems, Constraints and Reforms, *International Review of Education* 38(1), 19-33.
- SHAROBEAM, M.H., AND HOWARD, K. 2002. Teaching Demands versus Research Productivity. *Journal of College Science Teaching*, 31, 436-441
- TEFERRA, D. AND ALTBACH, P.G. 2004. African higher education: Challenges for the 21st century. *Higher Education* 47, 21-50.
- TILAK, J.B.G. (2000), Higher Education in Developing Countries. *Minerva* 38, 233-240
- VANDERHEYDEN, M. A. 2001. Measures of quality in higher education – An American Perspective. In Steven M. J., Creamer, D.G., Alexander, M.D (Eds), *International Perspectives on Quality in Higher Education*. Educational Policy Institute of Virginia Tech Department of Educational Leadership, pp. 14-20.
- VAHDATZAD, M.A., AND MOJTAHEDZADEH, M.T 2000. Some issues in the strategic management in a fast growing academic Institution: The case of University of Yazd. *Paper presented at 18th International Conference of the System Dynamics Society*, August 6 - 10, Bergen, Norway
- WATTY, K. 2003. When will Academics Learn about Quality? *Quality in Higher Education*, 9(3)
- WILLIAMS, R., AND VAN DYKE, N. 2007. Measuring the international standing of universities with an application to Australian universities. *Higher Education* 53, 819-841.
- WORLD BANK 2000. *Higher Education in Developing Countries: Peril and Promise*. Naylor Design, Inc. Washington DC **Appendix A: Stocks and Flows Diagram for Research and Quality Issues**



PART 4



Data Communication and Computer Networks

Using Mobile Technology and Services for Sustainable Development in Africa

Fisseha Mekuria

Due to the explosive growth of mobile and wireless technologies and services in Africa, a number of questions are being raised by academicians, technology professionals, industry and policy makers on how it is going to affect the future development of society in Africa. The recent World Wide Web consortium (W3C) and World bank organized workshop in Brazil [1,2], have discussed issues on the effects of mobile technology for social development. This was also partly motivated by the increased internet and web presence in the mobile area, and the possibility of using mobile subscriber terminals as a tool for delivering public information and services. A number of mobile projects for health, governance, banking and learning have been identified in the boundary region of mobile and web technologies. One important conclusion that came out from the discussion was that, the current few mobile SMS based services, appearing on the scene are just a proof of concept and announce the beginning of the thousands of mobile data based services that will engulf the mobile economies of the future. The main reasons for this are: 1) The ubiquity of the mobile device and the high increase in mobile subscriber base, soaring above the 3.5 billion value, and 2) The improved capabilities of mobile devices, 3) The flexibility of time and geographic independence that is possible in accessing information using mobile technologies. This paper will try to address some issues of mobile technology and sustainable economic development in a developing country context.

1. Introduction

Recently developing countries have shown the largest rate of increase in mobile technology deployment, services and subscriber base. Therefore it is time to devise models and harness the mobile technology for a sustainable economic development. Learning from the history of the developed nations is one way to meet these challenges. The history of developed nations teaches us that educational institutions were the core players when addressing technological and societal challenges appearing in history. Such framework has resulted in a sustainable development model which is well documented [IDEON Research Park, 2003; Clarysse and Wright, 2005; Mekuria, 2007]. In this paper the role of educational institutes in building the capacity for affordable and sustainable mobile technology and service development will be discussed. The frame work consists of research, educational and entrepreneurial components. In the final analysis the keynote will

argue that the harnessing and development of a local service and content provider sector is of paramount importance, in order for developing countries to benefit from the spinoff effects of the explosive growth in mobile technology & services. Other issues such as models for the reduction of the total cost of ownership for a mobile device and services will be shown to result in effective implementation and development of mobile web & social networking services in developing countries.

To maintain the mobile technology and services growth in Africa and migrate towards more relevant data services: such as mobile-health, mobile-learning, mobile payment,..., we need to look beyond the voice & SMS services that are now prevalent. The challenges to be addressed can be classified into infrastructure and human resource to develop the necessary localized services and content. This paper tries to address the latter challenge, by developing an MSc program titled: “Mobile Computing & Application Software development” (short MCASD) to be delivered at the University of Makerere. The program will address the issue of the development of a localized mobile content and service provider sector in a formal and competitive way. Such a sector will ensure that developing countries can benefit in the form of employment creation and economic development as a spinoff of the explosive growth of mobile technology and services. The program is also designed to work as a University-Industry collaborative initiative. To make the program successful three areas are identified that will be addressed in the paper: 1) The MCASD curriculum, 2) The Mobile Laboratory for Research, usability testing & prototyping, and 3) A research group in usability & localization for Mobile Computing, Technology & services.

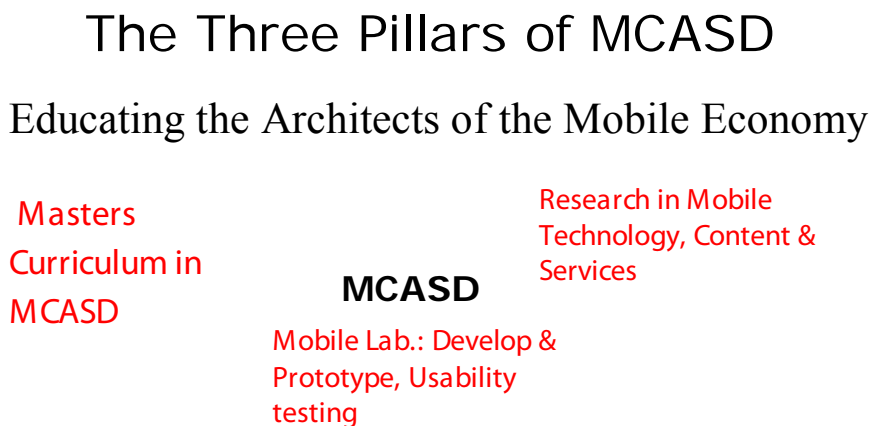
1.1 The MCASD Educational Program

The masters in mobile computing and applications software development, MCASD program at Makerere University is designed to give a service oriented understanding of the wireless technology, standards and architecture of the mobile network. It will also address the mobile services eco-system from different stake holder’s perspective. The program will also equip students with mobile programming skills to design innovative, and locally relevant mobile content and services. In general the program will focus on producing graduate students with competence to address the following issues:

- 1- To Bridge the mobile services and local content development knowledge gap between African subscribers and the mobile communications & service industry.
- 2- To perform relevant research in the areas of mobile technology and services. To cater for this a research group focused in the areas of “Mobile Computing & Communications” technology & Services, in collaboration with international partners and industry is being formed at CIT, Makerere university.
- 3- To support the mobile telecom industry and regulatory authorities with research to extend connectivity and mobile services to rural areas by performing locally relevant research in mobile communication services with community based usability lab studies.

- 4- To play an important role in harnessing the explosive mobile communications services growth in emerging markets of Africa to the benefit of the society and economic development.
- 5- To maintain the mobile market growth with active collaboration of the telecom network and service provider industry and improving the user experience by usability studies.
- 6- To establish a mobile software and applications programming Lab. to support the MCASD program, perform research in mobile software development and service scalability & usability studies.
- 7- Support graduating students' entrepreneurship, and promote the creation of a mobile web content & service provider sector in Uganda and the Eastern Africa region.

Fig 1: A Conceptual Block Diagram of the MCASD Program Functional Units.



2. Research in Mobile Technology, Content and Services

The second pillar of the MCASD program is the research that Makerere University, faculty of Computing and IT, as an academic & research institute has initiated to support the long term sustainability of the MCASD program. It is built in collaboration with international research faculty, the local mobile network & service industry, and the involvement of the public policy and regulatory authorities. Students' final year thesis work will be guided with high quality research and prototyping lab support, to produce relevant and localized mobile content and services.

To make the research component in MCASD feasible and to support new mobile services such as mobile search, mobile web services, mobile blogging and mobile social networking, a research group is formed which will spearhead the MCASD research areas. Such a research group will spearhead the development of

new innovative mobile services, which will guarantee successful launch, is based on real need of local mobile users, and will enrich the lifestyle of the community where it is deployed. Using the mobile Lab which is discussed in the next section, the research group will also be involved in mobile service usability and localization research to support the successful diffusion of mobile services to underserved rural communities in Africa.

3. Mobile Lab for Developing, Prototyping & Usability Testing of Innovative Mobile Services

The MCASD program embodies the important component of mobile Lab to teach graduate students the tools for mobile applications programming, perform research in mobile service usability and carry out practical field research on mobile technology and service usability. Once the lab trials are completed and a prototype mobile service is developed, the mobile laboratory can be easily transported to the community where the service is to be deployed. Such a field usability testing will give crucial feedback to the developed mobile service, and will help promote sustainable diffusion of technology and services specifically for rural and urban underserved communities.

The MCASD mobile Lab will also be used to promote graduate students entrepreneurial activities in mobile software development, innovative mobile service based startups and the emergence of a mobile content and services provider sector. The Lab also will be a resource tool for operators and service providers to test new mobile services before launch, for testing of usability and localization. Another wish expressed from the public sector policy and regulatory authorities and consumer groups [3], is the need to generate consumer information material on capabilities of mobile terminals and which services are possible with affordable terminals of customers. This is especially important in developing regions where the literacy level and exposure to technical gadgets is minimal.

4. Developing Regions and the Challenges for Mobile Development

As next generation mobile technologies are being deployed in developing regions, one important challenge that the MCASD program aims to perform research is the provision of innovative mobile data services in an environment where the customer base is composed of greater than 95 % prepaid customers and belong to the low ARPU segment of the market. Issues such as affordability and total cost of mobile innovative services will be studied together with usability, language & cultural context localization.

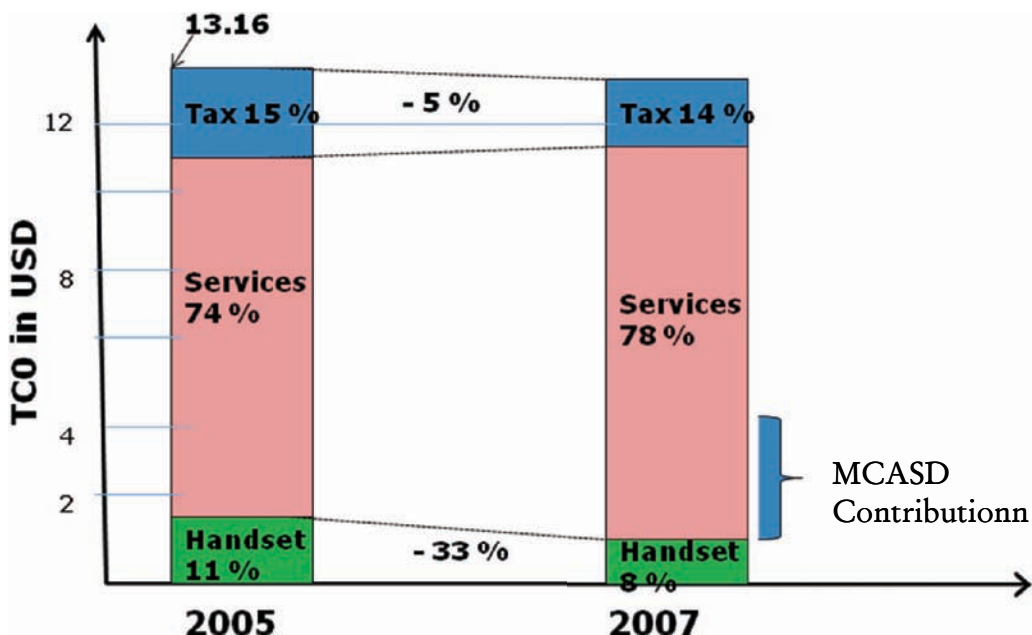
If the mobile industry and service provider sector is to develop towards the provision of relevant innovative mobile data services in developing regions, it has to look above the present complacent position and look towards solving the issues associated with affordability, usability and localization. As shown above in figure 2, research in the area of affordability and reduction of TCO is required to make sure that the innovative services developed by students and the mobile service provider sector can be used by all income groups in developing regions[6]. Initial

indications support the notion that the MCASD program has a potential to result in reduction of the TCO by 20-30 % in the long term, on innovative mobile services for developing countries [5].

5. Educational Program for Promoting Sustainable Development

The objective of the MCASD masters program is to produce graduates who will be the architects of the mobile economy and lead the sustainable development of mobile technology and services. At Makerere university we have students from the various communities in Uganda and students from all Eastern Africa countries. The sustainable development aspect of the program is developed through two steps, 1) Students going back to their communities with a developed mobile content or service, for the purpose of usability testing, localization and relevancy check. This will give students a chance to interact with the rural community and provides an educational social networking experience for both parts.

Fig. 2: Average total cost of ownership for low income group. (Expanding horizons Nokia 1/2008),



2) The incubation support to start a mobile content and service provider sector is the next sustainable development aspect where employment creation within communities and in the region in general can be achieved. The program will prepare students with the following competency areas:

- The principles of mobile /wireless communications. The mobile terminal hardware and software set-up, the API and the radio interface to the base station. Mobile operating systems and user interface design. Manipulate and modify installed mobile applications and trouble shoot.
- Be able to use the tools for designing mobile optimized web-sites, mobile content development, and mobile web services. Mobile programming tools, design of Innovative mobile applications and test on a mobile

platform. Modify existing applications to achieve localization of mobile services with desired usability characteristic.

- Have state of the art competence in the areas of mobile communications & services and be readily available for employment and consulting assignments by the mobile communications industry. To provide graduating students competence in the areas of mobile content, application software, and web services development, and to be able to pursue entrepreneurial path and start a mobile content & service provider company.

Finally a software incubation center is being set-up at Makerere university, faculty of computing and IT which will be used to support, ICT incubation and startups of graduating students. The center is aim to support and be a green house for the creation of a local mobile service and content provider sector [4].

6. Conclusion

In this paper the role of educational institutes in building the capacity for affordable and sustainable mobile technology and service development is presented. An educational, research and incubation program with a local center of excellence for generating locally relevant innovative mobile services, in a developing region context is described. The success of such a program will in the author's opinion, guarantee successful and sustainable diffusion of mobile technology and services into the communities, and the needed continued growth of the market for the mobile industry. At the same time the program has a potential to create, a conducive environment in developing countries to build a local mobile content and services sector. This in turn will bring benefits in the form of employment creation for sustainable social and economic development. In the final analysis the affordability issue for the new services should be studied for mobile total cost of ownership reduction. Initial indications suggest a 20-30% reduction in TCO is expected due to the effects of the MCASD program.

References:

www.w3c.org/m4sd/

IDEON Research Park: Swedens First Research Park, Source of Innovation, 2003. WWW.ideon.se.

B. CLARYSSE, M. WRIGHT, "Spinning out new Ventures: A typology of incubation strategies from European research institutions". Elsevier Journal of Business Venturing Vol. 20, pp. 183-216,2005.

F. MEKURIA "ICT Parks, ICT Business Incubation and sustainable development", Proceedings of ICTe Africa-2007 NEPAD symposium, Nairobi, Kenya.

F. MEKURIA, "Educating the Architects of the Mobile Economy: A program for Mobile Computing and Applications Software Development", Proceedings of World wide web consortium, W3C, and World Bank Workshop on Mobile Technology for Social Development.

"Affordability Key in bringing Digital Inclusion.", Expanding horizons: www.nokia.com/expandinghorizons, 1/2008, pp.12-14.

25

Implementing Security in Wireless Sensor Networks

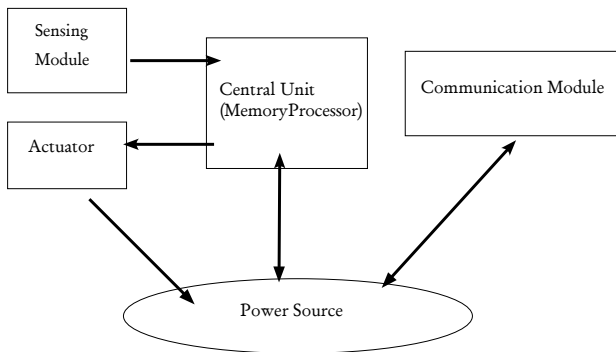
Joseph Migga Kizza

Wireless sensor networks (WSN) consist of a large number of small sensor nodes, usually spread out in hard accessible areas and communicating wirelessly. A sensor node combines the abilities to sense, compute, and communicate to other nodes. The large numbers of nodes with minimum capacity devices operating in constraining and demanding real-world environments impede communication within and outside the network, making the process of implementing security in wireless networks exceptionally difficult and expensive. We discuss these difficulties and what is being done to overcome them in order to meet the ever-growing and popular wireless sensor networks. We conclude by considering several potential future directions for security solutions.

1. Introduction

The rapid development of wireless technology in the last few years has created new interest in low-cost wireless sensor networks. Wireless sensor networks (WSN) or just sensor networks are grids or networks made of spatially distributed autonomous but cooperating tiny devices called sensors all of which have sensing capabilities that are used to detect, monitor and track physical or environmental conditions, such as temperature, sound, vibration, pressure, motion or pollutants, at different locations [2]. A sensor, similar to that in Figure 1, is a small device that produces a measurable response to a change in a physical condition. Sensor nodes can be independently used to measure a physical quantity and to convert it into a signal that can be read by an observer or by an instrument [2]. The network may consist of just a few or thousands of tiny, mostly immobile randomly deployed nodes, covering a small or large geographical area. Sensor networks do not require predetermined positioning when they are randomly deployed making them viable for inaccessible terrains where they can quickly self organize and form a network on the fly.

Fig 1: A Wireless Sensor Node



The use of sensors to monitor physical or environmental conditions is not new. Sensors have been used in both mechanical and electrical systems for a long time. However, what is new and exciting is that the new sensor nodes are now fitted with onboard tiny processors forming a new class of sensors that have the ability to partially process the collected data before sending it to the fusing node or base station. The sensor nodes now also have sophisticated protocols that help in reducing the costs of communications among sensors and can implement complex power saving modes of operations depending on the environment and the state of the network [6]. The accuracy of the data gathered has also greatly improved.

These recent advances have opened up the potential for WSN. According to David Culler et al [5], wireless sensor networks could advance many scientific pursuits while providing a vehicle for enhancing various forms of productivity, including manufacturing, agriculture, construction, and transportation. In the military, they are good for command and control, intelligence and surveillance. In health, they are beneficial in monitoring patients, in commercial application they can be used in managing inventory, monitoring production lines and product quality and monitoring areas prone to disasters [7]. New technologies are creating more powerful and yet smaller devices. This miniaturization trend is leading us to ubiquitous computing capacities that are exponentially faster and cheaper with each passing day. With these developments, researchers are refocusing and developing techniques that use this miniaturization process to build radios and exceptionally small mechanical structures like sense fields and forces in physical environments that could only be imagined just a few years ago. Culler et al believe that these inexpensive, low-power communication devices can be deployed throughout a physical space, providing dense sensing close to physical phenomena, processing and communicating this information, and coordinating actions with other nodes including a base station [5].

However, as wireless sensor networks with vast potential of applications unfold and their role in dealing with sensitive data increases, the security of these networks have become one of the most pressing issues in further development of these networks. This paper gives a general discussion of these limitations and how these limitations contribute to the security problems faced by the sensor network.

We survey several interesting security approaches aimed at enhancing security and we conclude by considering several potential future directions for security solutions.

2. The Growth of Sensor Networks

WSN have evolved slowly from simple point-to-point networks with simple interface protocols providing for sensing and control information and analog signal providing a single dimension of measurement to the current large number and sophisticated wireless sensor nodes networks. The development of the microprocessor boasted the sensor node with increased onboard intelligence providing it with different computing capabilities. The sensor node is now a microprocessor chip with a sensor on board. The increased intelligence in the node and the development of digital standards such as RS-232, RS-422, and RS-485 gave impetus to the creation of numerous sensor networking schemes [3]. In addition, the popularization of the microcontrollers and the development BITBUS, a field bus developed by Intel to interconnect stand-alone control units and terminals thus making them able to interchange data telegrams further improved the sensor node communications capabilities, thus, bringing the dream of sensor networks closer[3].

Another outstanding development that further made the road to fully functioning sensor networks possible was the development of the Manufacturing Automation Protocol (MAP), by General Motors to reduce the cost of integrating various networking schemes into a plant wide system. As Jay Warrior observes, this further resulted in the development of the Manufacturing Messaging Specification (MMS), a specification that made it possible for the networked nodes to exchange real-time data and supervisory control information [3]. With the development of other communication protocols that allowed simultaneous analog and digital communications for smart instruments, the sensor network, as we know it today, was born. Currently there is a whole spectrum of different sensor network protocols for the many different types of sensor networks in use today.

3. Design Factors in Sensor Networks

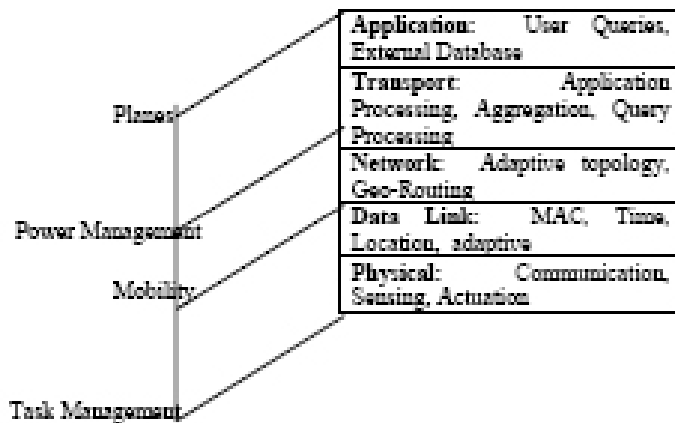
Several factors influence the design philosophy of sensor networks. Among these factors are first whether the nodes are stationary or moving and whether the network is deterministic or self-organizing. Most sensor network applications use stationary nodes. However, there is a good number of applications that use mobile nodes. In this case, therefore, the network is bound to use more energy because of the need to track the moving node and the increase in bandwidth requirements for periodic reporting which increases traffic. In a deterministic topology, the positions of the nodes and the routes in the network are pre-determined and the nodes are manually placed. In self-organizing topology, node positions are random and the routes are also random and unreliable. Routing in these networks, therefore, becomes the main design concern. Also since self-organizing sensor networks

demand a lot of energy, direct routing is not desirable and multi-hop routing is more energy efficient. However, multi-hop requires considerable routing management techniques. In addition to routing and energy, other factors also influence the design philosophy of sensor networks [7]:

3.1 Routing

Communication in wireless sensor networks, like in most network communication, is based on a protocol stack with several layers as seen in Figure 2. This stack also combines power and routing awareness, integrates data with networking protocols, communicates power efficiently through the wireless medium, and promotes cooperation between nodes [7]. To achieve all these, the stack consists of five layers and three management planes and these are: physical layer, data link layer, transport layer, application layer, power management plane, mobility management plane, and task management plane. This stack is different from those of the traditional networks made of non-sensor nodes like the TCP/IP and the ISO's OSI.

Fig 2: Sensor Network Protocol Stack



Physical Layer - is responsible for several tasks including frequency selection, carrier frequency generation, signal detection, modulation, and data encryption.

- Data Link Layer - is responsible for a number of tasks including multiplexing of data streams, data frame detection, medium access and error control.
- Network Layer - is responsible for network routing. Routing in sensor networks, unlike in the traditional networks, is influenced by the following [7]:

- o Power efficiency as an important consideration.
- o Sensor networks being mostly data-centric
- o data aggregation being useful only when it does not hinder the collaborative efforts of the sensor nodes, and
- o an ideal sensor network having attribute-based addressing and location awareness.
- Transport Layer - not yet in place. Because, unlike traditional networks, protocols like TCP where the end-to-end communication schemes are possible, here there is no global addressing. The development of global addressing schemes is still a challenge.
- Application Layer - also not available. Although there are many application areas for sensor networks, application layer protocols are yet to be developed.

Based on the above discussion, therefore, sensor networks are largely still multi-hop wireless networks whose nodes can be either a host or a router, forwarding packets to other nodes in the network. In many sensor networks the information collected from a large number of sensors is either lightly processed locally at the node, or transmitted to the base station or other sensors for further processing using one of the three routing techniques: one-to-many, many-to-one, and one-to-one/point-to-point. However, the two most fundamental communication primitives are broadcast (one-to-many) and point-to-point (one-to-one).

Broadcast Communication

The broadcast packet routing technique is used extensively in wireless sensor networks due to the large number of sensor nodes in any wireless sensor network. Broadcasting, as a means of node communication is highly desirable in this environment because of the limited signal range for each node. In broadcast mode the node that initiates the broadcast of a packet is referred to as the source or sender node and all others as receivers. The receivers of broadcast packet then forward the packet to their nearest adjacent neighbors which causes the packet to move throughout the network enabling all network nodes to eventually receive the broadcast packet.

Point-to-Point Communication

Though not common, point-to-point routing is still important for many applications in wireless sensor networks including games based on wireless sensor networks and data-centric storage where nodes store information about the detected events using the geographic location as the key. Point-to-point routing can also be used to send data from the detection node to the storage node [8].

3.1.1 Routing Protocols

There are several routing protocols in use today for sensor networks including data-centric, hierarchical and location-based [15].

Data-centric routing. Because the sensor network may have thousands of nodes which are randomly deployed it is inconceivable to have network-wide external addressing and network layer managed routing protocols found in traditional networks. If each node were to send out data to other nodes in the region, there would result a considerable redundancy of data and an inefficient use of scarce energy. For these reasons, data-centric routing techniques are more resource efficient. In data-centric routing the sink node, desiring data, sends out an attribute-based query to the surrounding nodes in the region. The attributes in the query specify the desired properties of the data. The sink then waits for the data [15]. Common data-centric routing protocols include sensor protocols for information via Negotiation (SPIN) and directed diffusion [15].

Hierarchical Routing. Hierarchical routing involves multi-hop communication and the aggregation and fusion of data within clusters of nodes in order to decrease the number of transmitted messages to the sink nodes which leads to conservation of energy. There are several hierarchical protocols in use today including LEACH, PEGASIS, TEEN and APTEEN[16].

Location-based routing. In location-based routing, each node maintains a location list consisting of location information for a number of nodes in a region of a sensor network. Each node periodically updates its location list by receiving information about locations and location lists of all its direct neighbors. It also, in turn, sends its location and location list to all its adjacent nodes. This keeps the location list of all nodes in the region current and up to date.

3.2 Power Consumption

Most sensor networks are entirely self-organizing and operate with extremely limited energy and computational resources. Because most nodes may be either in inaccessible environments, replenishing them with new power may be almost impossible. The life of a sensor node, therefore, may be in question and it may not be able to transmit critical data when desired. The functionality of the network, therefore, depends on the consumption rate of energy by node units.

3.3 Fault Tolerance

If a sensor network is to face anyone sensor node failure, we would like the network to be able to sustain all its functionalities. That is to say that the sensor network should be as reliable as possible and continue to function as much as possible.

3.4 Scalability

We want to have a network such that the addition of more nodes to the network does not have any diverse effects to the functionality of the network.

3.5 Production Costs

Wireless sensor networks most often use large numbers of sensor nodes. The unit cost of each individual sensor node plays a crucial role in determining the overall costs of the entire sensor network. We would like a well functioning network having a least per unit cost for individual nodes.

3.6 Nature of Hardware Deployed

A sensor node consists of four basic parts: the sensing unit, the processing unit, the transceiver unit, and the power unit. All these units must be packaged in a very small, match-box-sized package. In addition, all these units and the overall sensor node must consume very low power.

3.7 Topology of Sensor Networks

Because a normal sensor network may contain thousands of sensor nodes deployed randomly throughout the field of observation, the wireless sensor network resulting may have uneven densities depending on how the nodes were deployed. Nodes can be deployed by dropping them from a plane, carefully placing them, or dropped by artillery. Also not every deployed sensor may work as expected. So the topology of the resulting network may determine the functionality of the wireless sensor network.

3.8 Transmission Media

In a wireless sensor network, the nodes are linked by a wireless medium. The medium could be by radio like RF and Bluetooth, infrared or optical waves. Both infrared and optical links require no obstructions like objects in the line of sight. The functionality of the network may depend on these media.

4. Security in Sensor Networks

Modern wireless sensor networks many times consist of hundreds to thousands of inexpensive wireless nodes, each with some computational power and sensing capability and usually operating in a random unsupervised environment. The sensors in the network act as “sources” as it detects environmental events either continuous or intermittently whenever the occurrence of the event triggers the signal detection process. The data picked up is either lightly processed locally by the node and then sent off or just sent off to the “sink” node or a base station. This kind of environment presents several security challenges.

4.1 Security Challenges

The most pressing of these challenges include:

4.1.1 Aggregation

Data aggregation in sensor networks is the process of gathering data from different sensor “source” nodes and expressing it in a summary form before it is sent off to a “sink” node or to a base station. There are two types of data aggregation: in-stream aggregation, which occurs over a single stream, generally over a time window, and multi-stream aggregation, which occurs across the values of multiple streams, either at the same time or over a time window. Data aggregation is essential in sensor networks because as it combines data from different “source” nodes, it eliminates redundancy thus minimizing the number of transmissions

and hence saving energy. In fact significant energy gains are possible with data aggregation. The gains are greatest when the number of sources is large, and when the sources are located relatively close to each other and far from the sink [23]. However, as sensor network applications expand to include increasingly sensitive measurements of everyday life, preserving data accuracy, efficiency, and privacy becomes an increasingly important concern as this is difficult to do with many current data aggregation techniques.

4.1.2 Node Capture /Node deployment

Node compromise is a situation where a sensor node can be completely captured and manipulated by the adversary [13]. The conditions for node compromise are made possible as a result of sensor nodes in a wireless sensor network being random deployed many times in inaccessible or hostile environments. Usually these nodes are also unsupervised and unattended. In this kind of environments, nodes are undefendable and easy to compromise or totally capture by an adversary. There are several ways to capture a sensor node. One approach is the physical capture where an adversary can physically capture the node because of the node being in a hostile or unprotected environment. In another approach, is the software-based capture when an attacker uses software like a virus to capture a node?

4.1.3 Energy Consumption

Sensor networks are entirely self-organizing and operate with extremely limited energy and computational resources. Conservation of energy by sensor nodes results in minimizing their transmit power in order to maintain acceptable connectivity. This may prevent the network from maintaining the security solution needed to transmit and protect critical data.

Computational challenges: because of node energy conservation, sensor nodes have very limited computational power which presents a challenge in programming security functionalities like good encryptions.

4.1.4 Large Numbers of nodes/Communication challenges

Because modern wireless sensor networks consist of hundreds to thousands of inexpensive wireless nodes, this large number of nodes presents a challenge of guaranteeing a secure, reliable, sometimes ad hoc communication among sensor nodes or groups of sensor nodes which sometimes can be mobile units. For example, since sensor nodes are typically battery-driven, large numbers of them in a network make it a challenge to find and replace or recharge batteries.

4.2 Sensor Network Vulnerabilities and Attacks

Because of these limitations and the high dependency on the physical environment of deployment, sensor networks pose unique challenges that traditional security techniques like secrecy, authentication, privacy, cryptography, robustness to denial-of-service attacks used in traditional networks cannot be applied directly [9]. This means that existing security mechanism fit for traditional networks

cannot be used whole sale to sensor networks. Yet there are no comprehensive security mechanisms and best practices for sensor networks. One of reasons why traditional network security mechanisms and best practices fail with sensor networks is because many of these security mechanisms and best practices are taken and viewed as standalone. To achieve any semblance of desired security in a sensor network; these security mechanisms and best practices must be a part of and be embedded into every design aspect of the sensor network including the communication protocols and deployment topologies. For example we cannot talk about the security of a sensor network if that network lacks secure routing protocols. Secure routing protocols are essential security entities in sensor networks because a compromised routing protocol compromises the network nodes and a single compromised network sensor node completely compromises the entire network. However, current sensor network routing protocols suffer from many security vulnerabilities as we will see shortly.

We have established that sensor networks have a number of issues that separate them from traditional networks. Among these are the vulnerability of sensor nodes to physical compromise, significant power and processing constraints, aggregation of node outputs, and compromising individual nodes. Physical vulnerability includes physical node access and compromise or local eavesdropping. Power and processing constraints prevent sensor networks from running good security encryptions. And aggregation of output may grossly obscure the effects of a malicious attack from spreading throughout the network.

Senor network adversaries target and exploit these weaknesses and other network loopholes embedded within these limitations. Let us look at some of these next.

4.2.1 *Attacks*

There are several attack types including: eavesdropping, disruption, hijacking and rushing [12, 22]:

Eavesdropping. Here, the attacker (eavesdropper) aims to determine the aggregate data that is being output by either the node or the sensor network. The attacker captures the message from the network traffic either by listening to for some time to the network traffic transmitted by the nodes, or directly compromising the nodes . There are two types of eavesdropping:

- **Passive:** The attacker's presence on the network remains unknown to the sensor nodes and uses only the broadcast medium to eavesdrop on all messages.
- **Active:** The attacker actively attempts to discern information by sending queries to sensors or aggregation points, or by attacking sensor nodes.

Disruption. The intent of the attacker here is to disrupt the sensor's working. It is usually done in two ways:

- **Semantically:** where the attacker injects messages, corrupts data, or changes values in order to render the aggregated data corrupt or useless. Examples of this type of attacks includes [17]:

- **Routing loop:** where an attacker injects malicious routing information that causes other nodes to form a routing loop causing all packets injected into this loop to go round in circles and eventually resulting into wasting precious communication and battery resources
- **General DoS attacks:** where an attacker injects malicious information or alter the routing setup messages which end up preventing the routing protocol from functioning correctly.
- Sybil attack: where a malicious node influenced by an attacker creates multiple fake identities to perform desired attacks on the network.
- **Blackhole attack:** where a malicious node influenced by the attacker advertises a short distance to all destinations, thus attracting traffic destined to those destinations into the backhole.
- **Wormhole attack:** where two nodes are caused to use an out-of-band channel to forward traffic between each other, enabling them to mount several other attacks along the way.
- **Physically:** where the attacker tries to upset sensor readings by directly manipulating the environment. For example, generating heat in the vicinity of sensors will result in erroneous values being reported.

Hijacking. In this case the attacker attempts to alter the aggregated output of an application on several network sensor nodes.

Rushing attack: According to YihChun Hu et al [22], in an on-demand protocol, a node needing a route to a destination floods the network with ROUTE REQUEST packets in an attempt to find a route to the destination. To limit the overhead of this flood, each node typically forwards only one ROUTE REQUEST originating from any Route Discovery. In fact all existing on-demand routing protocols, such as AODV, DSR, LAR, AODV and other, only forward the REQUEST that arrives first from each Route Discovery. In the rushing attack, the attacker exploits this property of the operation of Route Discovery. The rushing attack is a very powerful attack that results in denial-of-service and it is easy to perform by an attacker.

4.3 Securing Sensor Networks

The choice of a good security mechanism for wireless sensor networks depends on network application and environmental conditions. It also depends on other factors like sensor node processor performance, memory capacity and energy. While in traditional networks, standard security requirements, such as availability, confidentiality, integrity, authentication, and non-repudiation, are sufficient for security, in sensor networks, special security requirements such as message freshness, intrusion detection, intrusion tolerance, are necessary in addition.

4.3.1 Necessary Conditions for A Secure Sensor Network

Any security solution to sensor networks must preserve the confidentiality, integrity, authentication, and non-replay of data within the network [14, 17].

Data Confidentiality. Confidentiality of data in a sensor network is achievable only if those with access to network data are authorized to do so. Under no circumstances should sensor readings leak outside the network. The standard approach for preventing this from happening is to use encryption. This requires the use of a secret key that only intended receivers possess.

Data Integrity. The integrity of data in any network means that data in that network is genuine, undiluted without authorization. This implies that data between the sender and the receiver is unaltered in transit by an adversary.

Data Authentication. The process of authentication of both network data and users is very important in preserving network data integrity and preventing unauthorized access to the network. Without authenticating mechanisms in place, an attacker can easily access the network and inject dangerous messages without the receivers of the new altered data knowing and making sure that the data being used originates from a malicious source.

Data Freshness/Non-replay. Adrian Perrig et al [14] define sensor network data freshness to mean recent data which for a sensor network would ensure that no adversary replayed old messages. There are two types of freshness: weak freshness, which provides partial message ordering, but carries no delay information, and strong freshness, which provides a total order on a request-response pair, and allows for delay estimation. Weak freshness is required by sensor measurements, while strong freshness is useful for time synchronization within the network [14].

4.3.2 *Security Mechanisms and Best Practices for Sensor Networks*

We cannot ensure the confidentiality, integrity, authentication, and freshness of data in sensor networks without paying attention to the following issues particular to sensor networks:

- **Data aggregation.** Aggregation is generally consensus-based compromise where missing readings from one or a few nodes may not significantly affect the overall system. Data aggregation is used in sensor networks to reduce energy consumption. With aggregation, however, raw data items from sensor nodes are invisible to the base station throwing in doubt the authenticity of the aggregated data. Without securing the data aggregation process, a compromised sensor node may forge an aggregation value and mislead the base station into trusting a false reading.
- **Antijamming** - Attackers can cause denial of service by jamming the base station or any other sensor node in the network. Attackers can also jam sensor radio frequencies. Protocols and services must be in place to stop this from happening.
- **Access control** - is a process of granting user the access right to the sensor network resources. It is essential to have an effective and efficient access control mechanisms especially via a base station to authenticate user requests to get access to the network resources.
- **Key management** - key management is crucial in supporting the basic security tenants like authentication and encryption in sensor networks.

As the number of applications for sensor networks grows, an effective key management scheme is required.

- *Link layer encryption* - Most widely used encryption schemes in sensor networks today involve the use of pre-distribution of key broadcasts by sensor nodes to thousands of sensors for pair wise exchange of information. But this scheme does not square well with known sensor network security problems like node compromise, low network connectivity, and a large communication overhead. However, a link-layer key management scheme can mitigate these problems and therefore be more efficient.
- *Data replication* - is the process of storing the same data on several sensor network nodes which created enough redundancy which in turn improves on reliability and availability and hence security.
- *Resilience to node capture*. One of the most challenging issues facing sensor networks is that of node capture because online traditional networks which can get high physical security, sensor networks can be deployed in environments with limited physical security if any.

Although we have outlined the difficulties in making a sensor network secure due to inherent limitations, it is however, possible to design security protocols that specific for a particular security issue. This is the direction current sensor network security research is taking [14].

There are several cryptographic approaches being used to secure sensor networks.

Cryptography. One of the first tasks in setting up a sensor network is to establish cryptographic system with secure keys for secure communication. It is important to be able to encrypt and authenticate messages sent between sensor nodes. However, doing this requires prior agreement between the communicating nodes on keys for performing encryption and authentication. Due to resource constraints in sensor nodes including limited computational power, many key agreement schemes like trusted-server, public-key, and key pre-distribution used in traditional networks are just not applicable in sensor networks. Also pre-distribution of secret keys for all pairs of nodes is not viable due to the large amount of memory this requires when the network size is large. Although over the years, efforts have been made to propose several approaches to do this, the inherent limited computational power of sensor nodes and the huge numbers of network nodes are making public-key cryptographic primitives too expensive in terms of system overhead in key-establishment [10]. Modern research has tried to handle the key establishment and management problem network-wide by use of a shared unique symmetric key between pairs of nodes. However, this also does not scale well as the number of nodes grows [10]. Another approach to establish keys that seem more appropriate for sensor networks is via pre-distribution, where (secret) key information is distributed to all sensor nodes prior to deployment [11].

4.4 Trends in Sensor Network Security Research

There are several active research tracks in sensor security and we will discuss a few here.

4.4.1 *Key Management*

Because of sensor nodes deployment and other sensor network limitations like limited computation capabilities, it is not possible to use key management as usually used in traditional networks where there may be a relationship in key sharing among members of the network. Because of these difficulties in sensor networks, if there were to be a single shared key, a compromise of just one node, may be through capture, would lay the entire network bare. A new framework of key exchange is needed. Eschenauer and Gligor [18] first proposed a framework of key exchange where a sensor randomly chooses m keys from the key pool with n keys before the deployment. After the node is deployed, it then contacts all its immediate neighbors to see if it shares any key with them. What must be noted in this solution is the non-involvement of the base station in this key management framework. Several extensions of this framework have been developed including [19]:

- The q -composite random key pre-distribution framework – where two nodes share a common key hashed from q common keys. This approach adds more strength to the above approach. Because now an intruder would need to capture communication from more nodes in order to be able to compute a shared key.
- Multi-key reinforcement framework – where a message from a node is partitioned into several fragments and each fragment is routed through a separate secure path. Its advantages are balanced by its high overhead.
- Random-pairwise framework - where in the pre-deployment phase, N unique identities are generated for each network node. Each node identity is matched up with other m randomly selected distinct node identities and a unique pairwise key is generated for each pair of nodes. The new key and the pair of node identities are stored on both key rings. After deployment, the nodes then broadcast their identities to their neighbors.

Other frameworks include a localized encryption and authentication protocol (LEAP) by Zhu et al [20]. Under LEAP, it is observed that there are different types of messages in a sensor networks. This lead to the use of four keys: individual, group, cluster, and pairwise key [19].

4.4.2 *Confidentiality, Authentication and Freshness*

It is common knowledge to all security professionals that the use of strong cryptographic techniques strengthens the security of communication. In sensor networks, like in traditional networks, this is also the case. During authentication

in sensor networks, the sending node, using a shared key with the receiving node, computes a MAC (Message Authentication Code) on the message about to be transmitted using a known hash function. The receiving node, upon receipt of the message, applies the shared key and the same hash function to the message to generate a new MAC. If this MAC agrees with the sender node's MAC, then the message has not been tampered with and the receiving node knows that the message has been sent by the sending node since it is only this sending node that shares the key with the receiving node. Several studies including [14] SPINS which has two building blocks: Secure Network Encryption Protocol (SNED) which provides data confidentiality, two part data authentication, and data freshness; and micro Timed, Efficient, Streaming, Loss-tolerant Authentication (μ TESLA) which provides authentication to node streaming broadcasts. In addition to SPINS, TinySec [21] which also supports message confidentiality, integrity, and authentication in wireless sensor networks also uses this approach. There are several other works on message confidentiality, authentication, and integrity including Perrig et al [14].

4.4.3 Resilience

While sensor networks, because of their size and deployment, are ideal for information gathering and environmental monitoring, node compromise poses a very serious security problem in these networks. While existing ad-hoc security solutions can address a few security problems, on a limited number of nodes in a network, many of these solutions cannot scale up when the numbers of nodes in the network grows. Also when the node number is high and typically these nodes are unattended, they are prone to node compromise.

To overcome this problem, Yang et al [21] have proposed a novel local-based key management solution through two techniques in which they bind symmetric secret keys to geographic locations and then assign those location-bound keys to sensor nodes based on the nodes' deployed locations. There are two approaches to this scheme: location-binding keys and location-based keys. In both of these approaches the network terrain is divided into a grid where each cell on the grid is associated with multiple keys. Each node in a grid stores one key for each of its local neighboring cells and a few randomly selected remote cells. Any genuine real event must be validated by multiple keys bound to the specific location of that event. This requirement rules out any bogus event which might be a result of an attacker obtaining multiple keys from some compromised nodes because such event cannot combine all necessary keys to make the event genuine.

5. Conclusion

Although sensors have been used for a long time and many times these sensors have been used to create a network, new technologies developed in the last five years have created a new vibrancy and exuberance in the field of sensor and sensor

networks in general. More interesting are the developments of technologies and protocols that have helped create a marriage between digital and analog sensors to work together and exchange data across packet networks. This has resulted into a multitude of applications. The large number of new applications for wireless sensor networks has lead to unprecedented growth of wireless sensor networks. Along with this growth though is the rising concern for security of these networks as crucial data is being moved within and outside these networks. We have discussed these concerns in this paper and outlined working protocols and best practices in use today. Finally we have introduced and started a brief dialog into new research trends in wireless sensor networks.

References

- ANKAR SURE, S. S. IYENGAR, EUNGCHUN CHO. *Ecoinformatics Using Wireless sensor Networks: An Overview*. *Ecological Informatics*. Vol 1, 2006, pp 287-293.
- ADRIAN PERRIG. "Secure Routing in Sensor Networks". <http://www.cylab.cmu.edu/default.aspx?id=1985>
- ADRIAN PERRIG, JOHN STANKOVIC, and DAVID WAGNER. "Security in Wireless Sensor Networks", *Communication of the ACM*, June 2004/Vol. 47, No. 6, pp 53-57.
- ADRIAN PERRIG, ROBERT SZEWCZYK, VICTOR WEN, DAVID CULLER, J. D. TYGAR. "SPINS: Security Protocols for Sensor Networks", <http://www.ece.cmu.edu/~adrian/projects/mc2001/mc2001.pdf>
- BHASKAR KRISHNAMACHARI, DEBORAH ESTRIN, AND STEPHEN WICKER. "The Impact of Data Aggregation in Wireless Sensor Networks", http://lecs.cs.ucla.edu/Publications/papers/krishnamacharib_aggregation.pdf.
- C. KARLOF, N. SASTRY, AND D. WAGNER. "TinySec: A Link Layer Security Architecture for Wireless Sensor Networks". In *ACM SenSys*, 2004
- DAVID MALAN, THADDEUS FULFORDJONES, MATT WELSH, and STEVE MOULTON. "CodeBlue: An Ad Hoc Sensor Network Infrastructure for Emergency Medical Care", <http://www.eecs.harvard.edu/~mdw/papers/codeblue-bsn04.pdf>
- DAVID CULLER, DEBORAL ESTRIN, and MANI SRIVASTAVA. "Overview of Sensor Networks". *Computer*. August 2004, pp 41-50. <http://www.archrock.com/downloads/resources/IEEE-overview-2004.pdf>
- ERIC SABBAAH, ADNAN MAJEED, KYOUNG-DON KANG, KE LIU, and NAEL ABU-GHAZALEH. "An Application-Driven Perspective on Wireless Sensor Network Security", *Q2SWinet'06*, October 2, 2006.
- IAN F. AKYILDIZ, WEILIAN SU, YOGESH SANKARASUBRAMANIAM, ARDAL CAYIRCI. "A Survey on Sensor Networks", *IEEE Communications Magazine*, August 2002, pp 102 – 114.
- JAY WARRIOR. "Smart Sensor Networks of the Future". *DA Systems*. http://archives.sensorsmag.com/articles/0397/net_mar/main.shtml
- JOHN A. STANKOVIC. "Research Challenges for Wireless Sensor Networks". <http://doi.acm.org/10.1145/1121776.1121780>.

JORGE ORTIZ, DAEKYEONG MOON, CHRIS R. BAKER. "Location Service for Point-to-Point Routing in Wireless Sensor Networks".

http://www.cs.berkeley.edu/~jortiz/papers/cs268_ortiz_moon_baker.pdf

KAY R'OMER AND FRIEDEMANN MATTEN. "The Design Space of Wireless Sensor Networks". <http://www.vs.inf.ethz.ch/publ/papers/wsn-designspace.pdf>

Wikipedia. <http://en.wikipedia.org/wiki/Sensor>

L. ESCHENAUER AND V.D. GLIGOR. "A Key-management Scheme for Distributed Sensor Networks". In ACM Conference on Computer and Communications Security, 2002.

MADHUKARANAND, ZACHARY VIVES, INSUP LEE. "Quantifying Eavesdropping Vulnerability in Sensor Networks", ACM International Conference Proceeding Series; Vol. 96. Proceedings of the 2nd international workshop on Data management for sensor networks. <http://delivery.acm.org/10.1145/1090000/1080887/p3-anand.pdf?key1=1080887&key2=9635208021&coll=ACM&dl=ACM&CFID=63535215&CFTOKEN=62146412>

NALIN VIMALKUMAR SUBRAMANIAN. "Survey on Energy-Aware Routing and Routing Protocols for Sensor Networks." http://archives.cs.iastate.edu/documents/disk0/00/00/05/19/00000519-00/energy_sr.pdf

PRADIP DE, YONGHE LIU, AND SAJAL K. DAS. "Modeling Node Compromise Spread in Wireless Sensor Networks Using Epidemic Theory",

Proceedings of the 2006 International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM'06). <http://delivery.acm.org/10.1145/1140000/1139411/25930237.pdf?key1=1139411&key2=1877208021&coll=ACM&dl=ACM&CFID=63535215&CFTOKEN=62146412>

S. ZHU, S. SETIA, and S. JAJODIA. "LEAP: Efficient Security Mechanisms for Large-scale Distributed Sensor Networks". In the 10th ACM Conference on Computer and Communication Security (CCS '03), 2003.

SEAPAHNMEGUERDICHIAN, FARINAZ KOUSHANFAR, MIODRAG POTKONJAK, MANIB. SRIVASTAVA. "Coverage Problems in Wireless Ad-hoc Sensor Networks". http://www.ece.rice.edu/~fk1/papers/Infocom_coverage_01.pdf

WENLIANG DU, JING DENG, YUNGHSIANG S. HAN, PRAMOD K. VARSHNEY, JONATHAN KATZ and ARAM KHALILI. "A Pairwise Key Predistribution Scheme for Wireless Sensor Networks", ACM Transactions on Information and System Security, Vol. 8, No. 2, May 2005, Pages 228-258

YIHCHUNHU, ADRIAN PERRIG, and DAVID B. JOHNSON. "Rushing Attacks and Defense in Wireless AdHoc Network Routing Protocols", <http://www.monarch.cs.rice.edu/monarch-papers/wise03.pdf>

26

Delay-Optimized Path in Peer-to-Peer Networks

Francis Lowu and Venansius Baryamureeba

Peer-to-Peer (P2P) networks are popular in sharing multimedia resources such as video, audio and images. Its popularity has attracted very many users thus increasing data streaming along a congested path. The increased sharing of the resources in the network has caused delays making the network resource sharing inefficient. This paper presents a delay-optimized path algorithm that increases data flow within the network among the peers and utilizes the network resources for effective and efficient streaming of data. We consider a distributed search for peer nodes with multi-peer participation in both searching and initiating resource transmission within the network. The resource streaming across the peer network is controlled by the resource function. Single peer initiated searching is also considered where a single peer node initiates the resource search process. Our analysis shows that the nature of growth of the network is linear.

1. Introduction

In a P2P network, resource streaming is based on active nodes in a decentralized and distributed network. Sharing of resources such as images, video, audio and speech among others in the network booms. Peers stream packets to each node as long as the peer nodes in the network request for the resources. Decentralized lookup services that store key value pairs are good in scaling peer-to-peer networks [Stoica et al. 2001]. Other peers may also need to advertise some resources available to them for searching in the network by other peer nodes. Peer-to-Peer network's resource allocations are not normally distributed. Each peer in the network has its own resources interests and multimedia requirements for each user is different. Enterprises such as businesses, educational institutions and government organizations have collections of computing resources that require P2P architecture in order to use these resources efficiently [Adnan Fida 2004]. The transmission in bit/sec of the multimedia streams to the recipient has a varying quality of service provision for each multimedia packet along a given path. In their study Valduries and Pacitti [2004] say that, Peer-to-Peer computing offers new opportunities for building highly distributed data systems. Unlike client-server computing, P2P can operate without central coordination and offer important advantages such as a very dynamic environment where peers can join and leave the network at any time; direct and fast communication between peers, and scale up to large

number of peers. However, most deployed P2P systems have severe limitations: file-level sharing, read-only access, simple search and poor scaling. Keong et al. [2004] say that Internet today, computing and communications environments are significantly more complex and chaotic than classical distributed systems, lacking any centralized organization or hierarchical control. Peers need effortless ways to discover, consume and provide services so as to take advantage of new services as they become available [Elenius et al. 2004].

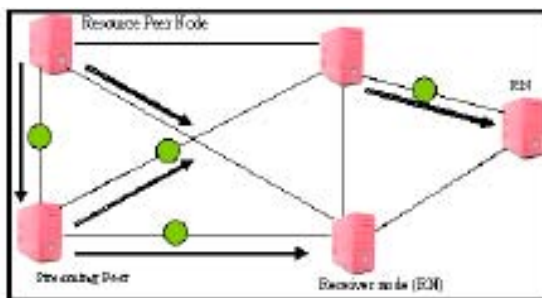
1.1 Overview

The rest of the paper is arranged as follows: in Section 2 we present the problem formulation, the network model and work of other researchers. Section 3 discusses the network maintenance for joining and departing peers. Section 4 presents the design of the algorithm. In Section 5 we present results and discussion of our findings and lastly Section 6 presents the concluding remarks.

2. Problem Formulation and Network Model

Peer-to-peer network involve multiple participating nodes streaming resources to different peer nodes. The scaling nature of the network makes participation in the streaming of resources not normally distributed. More peers may want to join the network at different times and this requires a way of catering for both the existing and joining or leaving members. We consider two cases in the network: a single-peer initiated streaming, where a single peer streams resources to multiple peer nodes and secondly a multi-peer initiated streaming where many peer nodes stream either the same or different resources along the network paths with an optimal delay path and high throughput. The network is modeled as an undirected graph $G(V,E)$, with V being the set of network vertices or peers and E is the link or the streaming path to the network nodes. For the case of the single-peer streaming there is a source S and the receiving node R such that the overall participation in streaming and receiving resources within the network is such that $V = (SUR)$, where $|(SUR)| = n$, where n is the number of nodes. While for the case of the multi-peer streaming there is no specified source node since at least two peers can stream network resources along any path to the recipients at the same time. Our objective is having an optimal-delay path for peer-to-peer resource streaming in the network.

Fig. 1: Single-Peer-Initiated Streaming



a. Single-Peer-Initiated Streaming

We consider that a single peer streams packets to all active nodes in the network. We further consider that for all links $(i, j) \in E$ there exists an optimal delay path $p(i, j)$ such that, the delay $d(i, j)$ is minimal. Resources from the peers is directed to the node that request or search for it as shown in Figure 1. Each peer node can act as a source while others are at the receiving end, a centralized resource streaming, however this happens if other nodes do not have resources to stream to the expectation of the receiver nodes. The peer node ready to transmit packets is called the initiator node, it sends messages to the other nodes to inform them that its ready to stream resources via their addresses.

b. Multi-Peer-Initiated Streaming

In this type of packet streaming; peers are distributed and each node has a link in an undirected graph $G(V, E)$. We formulate the problem by considering that each peer node is in the neighborhood of another peer node in a defined magnitude γ , which also determines the rate of resource streaming along a link. A peer p_k for $k = 1, 2, \dots, n$ will have resources to share with the other peers in the network; this peer will identify its neighbor as $p_i = \gamma(p_j)$, for $i \in j$ and where γ is the magnitude of p_i from p_j on a link (i, j) . Then peer p_i will forward a request for a given resource from peer p_j and visa-versa. The presence of resource at a peer node will be identified as Boolean $(0, 1)$, and a requested resource will be labeled as either at level 0 or level $1 \in (0, 1)$. A resource rx is at level 0 if it can not be retrieved from a given peer node by its neighbors or any other peer in the network. For example a resource $rx_0(p_i) = p_j$ indicates that a resource rx can not be retrieved from peer node p_i by a peer p_j . The magnitude γ will vary for each link in the network, and it will be key in determining the delay along that link.

%Each path in the network will be subject to a delay counter c between peer nodes. A delay counter $cd(i, j)$ determines how much delay a packet will take along a link (i, j) .

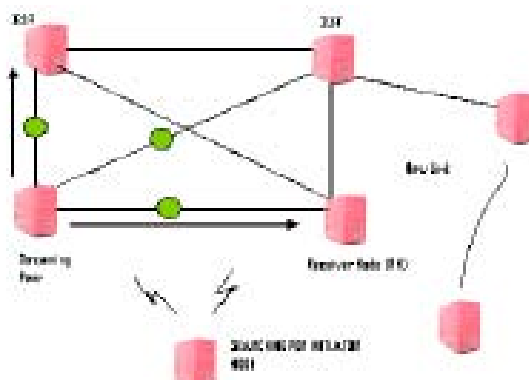
We require that the delay be minimized along the path streaming packets to all destination node-receiving resources from nodes with resource $\{1\}$ level. The

optimal delay path is such that $\text{delay}(i,j) = \min(\text{rx}1p(i,j))$, $\forall(i,j) \in E$ also we require that $\text{delay}(i,j) = \min \sum \text{rx}1p(i,j)$, $\forall(i,j) \in E$. Under the condition that the resource function, holds for all the nodes in the network.

$$\text{ResourceFunction}(i,j) \begin{cases} \text{rx} = 1 & \text{If resources are available} \\ \text{rx} = 0 & \text{Otherwise} \end{cases}$$

The **Figure 2** below shows how multi-peer streaming of resources in the network is arranged. Every peer node can be an initiator node in this type of streaming.

Fig. 2: Multi-Peer-Initiated Streaming



3. Maintaining the Network Model

The topological graph of the network has to be maintained in order to scale, since peers may require leaving or joining the network to share resources such as images, video, audio and speech. Peers that join the network have to identify a peer node already in the network in its nearest neighborhood. The following steps are essential for a peer joining the network. The node that accepts other peers to join the network is termed the integrating peer node.

a. Joining Peer

Initially we assume that there is only one peer node, say, and an arbitrary peer p_i . A peer p_j interested in the resources of p_i requests to join the network. Since only p_i has resources, then only resource level-1 will apply for the arbitrary peer p_i to join. If the peer joins the network, the magnitude γ of the peer p_j from peer p_i is noted and the delay function $d(i,j)$ is computed subject to γ for the available resource rx at peer p_i . This implies that p_j becomes the neighbor of p_i . Any other peer p_k will contact any of p_i or p_j to join the network. Suppose it contacts peer

p_j , then it will become the immediate neighbor of that peer node. A new node that has been integrated will have an address that depicts its presence in the network. Peers will need to identify their network requirements immediately they start sending and receiving packets.

b. Departing Peer

The peer node leaving the network sends a departure message to all other peer nodes in the network. With the departure message, it embeds a resource level 0 in the message to inform the other peers that it does not have any resource to share. This makes all its immediate neighbors to terminate the entire request to the departing node, such that $\gamma = 0$ along the links connecting to the departing peer. Nodes can be lost during the process of joining the network. For such cases no departure message is received from that node by either the integrating peer node or any other node in the network.

4. Delay-optimization Algorithm

The delay optimization algorithm is presented in this section. To create the peer-to-peer graph, there is an initiator node that has resources it may share with other peers. A peer gets interest in the resources of the initiator node and sends a request to share the resources. The initiator node realizes that it has no immediate neighbor, on getting the request from the joining peer. If the neighborhood node is found, it's labeled a peer node and a resource-sharing link is immediately formed between the peers.

For each link formed a link delay requirement is registered and its magnitude is assigned. The magnitude is assigned only if the new peer has resources to share that make the link a busy path for resources streaming. Each link in the network will have a delay bound Delta which depends on the resource requirement along that link.

Input: Number of peer nodes n

Output: An optimized delay path

begin algorithm

```

Initiator node =  $p_o$ 
  for all  $n$  nodes  $p_k$ ,  $k = 1, 2, \dots, n$ 
     $p_k =$  (true or false)
  if  $p_k =$  true, find next node  $p_o$ 
    link ( $p_o, p_k$ )
  else  $p_k =$  false
    Delta = link delay bound
    while ( $p_o, p_k$ ) is linked do
      magnitude( $p_o, p_k$ ) =  $\gamma$ 

```



```

      ( $p_o, p_k$ ) = linkDelReq
if linkDelReq ≤ Delta and resource level = 1
  Stream resource
    delay(i,j) ≤  $\Sigma rx1p(i,j)$  ≤ Delta
     $\leq rx1p(i,j)$  = optimal
  else
    resource level = 0
  locate new node =  $p_{\{k-1\}}$ 
  repeat
    return optimal delay
end

```

5. Results

5.1 Discussion

Quality of service (QoS) constraints such as delay is common in peer-to-peer networks and need to be addressed. In our study we talk about the optimal delay path in peer networks and how we can achieve it to increase resource streaming among the peers. We present two types of streaming: the single-peer-initiated streaming where a single peer will be allowed to initiate resource searching in the network. This streaming is considered a broadcast and only the initiator node will be responsible for allowing new peers to join as others in the network search for the resources. This is to help in reducing the congestion encountered if all peers are at the same level of streaming resources hence minimal delay on the network paths. We also present the multi-peer initiated streaming where every node is free to stream and request for resources in the network. However only nodes with resources to share are eligible for search. A peer node that does not have resources to share cannot be an initiator node and therefore no new member can join from that peer node. This increases reliability and reduces congestion due to crowding at the peer node and along the streaming link.

5.2 Correctness

Peers in the entire network can at least share or search for resources, subject to the resource function. Each peer joins the peer network as a sole node, this scales the network additively. Since no link can exist without a peer in the network, this study assumes that the initial state of the algorithm holds. A peer is either in the network or outside, if not joining the network. We denote by true a peer already in the network or false a peer not in sight. The algorithm further expects that for all peers ($pk, pk-1$) in the network, there is a binding link that scales the network. In line 7, the while loop holds only if ($pk, pk-1$) holds as a link: whose magnitude can then be computed. It is easy to see that the algorithm is linear, from line 2, the while loop iterates once for every entry of a peer in the network, requires $(n+1)$ number of times to execute. Whereas the for loop requires n times to read through each input values (joining nodes).

5.3 Contribution

Considering the network model discussed, the paper made the following contribution towards the study.

1. The network model designed gives a good picture of how members join and leave the network without disorganizing packet streaming.
2. We have designed an algorithm that optimizes delay along the peer links and we find that the nature of growth of the network is linear.
3. The paper clearly gives how a peer can easily find a resource at the other peer node subject to the resource function: from the time the peer joins and leaves the network.
4. The searching of resources at the streaming (source) node is found to be proportional to the searching peer. We define the constant of proportionality as the magnitude, γ which varies depending on the placement of the nodes within the network.

5.4 Related Work

Rongmei et al. [2004] proposed the improved search in unstructured peer-to-peer (P2P) overlay networks by building a partial index of shared data. The index maintains two types of information: the top interests of peers and globally unpopular data, both characterized by data properties. A variety of P2P systems for sharing digital information are currently available and most of them perform searching by exact key matching. Panos et al. [2003] focused on similarity searching and describe Fuzzy peer, a generic broadcast-based P2P system that supports a wide range of fuzzy queries. As a case they studied and presented an image retrieval application implemented on top of Fuzzy peer where users provide sample images whose sets of features are propagated through the peers. Tzanetakis et al. [2004] describe a robust, scalable P2P system that provide flexible search semantics based on attribute value pairs and supports automatic extraction of musical features and content-based similarity retrieval in P2P networks. Srdjan et al. [2005] present P2P sensor networks overlay on 3G mobile networks. Each sensor network acts as one peer node and is represented by its gateway in the P2P network. Peers communicate and provide collaboration and information on the executions. Zihui et al. [2003] develop a mathematical model to explore and illustrate the fundamental performance issues of peer-to-peer file sharing. Their model showed flexibility in different characteristics of the P2P system. Much work has been published in the peer-to-peer optimization topology, [Aaron et al. 2004] however little work has been done on the resource optimization between neighbor peers. They propose traffic based learning protocol that learns new connections between neighbor peers. Nazanin and Reza [2006] examined the key issues and tradeoffs in incorporating swarm-like content delivery into mesh-based peer-to-peer streaming of live content. Ashen and John [2003] propose a rank-based selection mechanism for peer-to-peer media streaming systems. Their mechanism provides a differentiation of services with incentives for cooperation.

6. Conclusion and Future Work

Communication in peer-to-peer networks is scaling very fast with the changing Internet and support of wireless technology. Peer nodes require to stream and search for resources with minimum delay constraint. The paper has presented a delay optimization algorithm that caters for the minimum delay path along a streaming link in the network, for a scaling peer group. Where each link is given a link magnitude that helps in identifying the QoS requirements of any given link in the network. Further the link delay requirements are noted for each link with resource level, such that its easy to measure the delay on the link that attempts to stream resources. We further looked at the way peers behave to minimize congestion during streaming. Our proof of the correctness of the algorithm supports the objective of the study and it shows that the members scale linearly, with a time complexity of $O(n)$, which is consistent with the precondition and the post condition of the algorithm. In future we intend to consider how an insecure peer node in the network affects P2P multimedia streaming. And how that peer node on attack can increase delay and affect searching. Our interest will be in securing a peer node with resources to share.

References

- AARON, H., EGEMEN, T., MINH, T. (2004) Fast Learning of Optimal Connections in a Peer-to-Peer Network. Dept of Computer Science and Software Engineering, University of Melbourne, Victoria.
- ADNAN, F. 2004. Scheduling Policies for Enterprise Wide Peer-to-Peer Computing Department of Computer Sciences University of Saskatchewan 57 Campus Drive Saskatoon, SK, Canada.
- AHSAN, H., JOHN, C. (2003) Incentive Mechanism for Peer-to-Peer Media Streaming. School of Information Management and Systems. University of California, Berkeley.
- ELENIUS, D., MAGNUS, I. (2004). Ontology-based Service Discovery in P2P Networks Department of Computer and Information Science Linköping University 581 83 Linköping, Sweden.
- KEONG, L., JON, C., MARCELO, P., RAVI, S., and STEVEN, L. (2004). A Survey and Comparison of Peer-to-Peer Overlay Network Schemes. *IEEE Communications Survey and Tutorial*, March.
- NAZANIN, M., REZA, R.(2006). Understanding Mesh-based Peer-to-Peer Streaming. Dept of Computer and Information Science. University of Oregon.
- PANOS, K., WEE, S., NG BENG, C., KIAN-LEE, T. (2003). Answering Similarity Queries in Peer-to-Peer Networks Department of Computer Science National University of Singapore 3 Science Drive 2, 117543 Singapore.
- RONGMEI, Z., AND CHARLIE, H. (2004). Assisted Peer-to-Peer Search with Partial Indexing Purdue University West Lafayette, IN 47907.
- ROWSTRON, A., DRUSCHEL, P. (2001). Pastry: Scalable, Distributed object location and routing for large-scale peer-to-peer systems. Lecture notes in Computer Science.

- SRDJAN, K., CLEARY, D., DARYL, P. (2005). P2P Mobile Sensor Networks. R & D, LMI Ericsson, Ireland. {\it Proceedings of the 38th Hawaii International Conference on System Sciences - 2005.}
- STOICA,I.,MORRIS,R.,KARGER,D.,KAASHOEK,M.,BALAKRISHNAN,H.(2001).Chord: A Scalable Peer-to-Peer Lookup Service for Internet Applications. {\it Technical Report TR-819,MIT, March.}
- TZANETAKIS, G., JUN, GAO., STEENKISTE, P. (2004). A Scalable Peer-to-Peer System for Music Content and Information Retrieval. Computer Science Carnegie Mellon University 5000 Forbes Avenue Pittsburgh, PA 15218.
- VALDURIEZ, P., PACITTI, E. (2004). Data Management in Large-scale P2P Systems Atlasgroup, INRIA and LINA, University of Nantes – France.
- ZIHUI, GE., DANIEL, R.F., JAISWAL, S., KUROSE, J., DON, T. (2003). Modeling Peer-Peer File Sharing Systems Department of Computer Science University of Massachusetts, Amherst

Service Oriented Telecom Architectures for Enabling Innovative Mobile Services

F. Mekuria and W. Mwebaza

Future networks are characterized by convergence at different technological and service levels. Convergence of fixed and mobile network is a trend that is used to develop service oriented communication networks. This paper discusses service oriented telecom architecture for convergent networking and enablement of new innovative mobile services. Analysis of different existing migration strategies for service orientation results in a near similar IP-Multimedia Subsystem (IMS)-based architecture for the case of developing countries. The main migration and networking guidance parameters used are: capital and operational cost reduction, simplicity of introducing & managing new mobile services, unified pricing & billing for revenue generation, smooth migration of legacy network, and adherence to international telecom standards.

1. Introduction

The convergence of mobile and fixed (or wireline) telecommunications services, terminal devices, and networks is inevitable. Its ecosystem is comprised of several key components beyond network operators and service providers. There are several layers to this ecosystem and there are many specialized companies within the layers, but the key players are standard development and certification organizations, network equipment vendors, handset vendors, application developers, and consumer electronics companies. These all players have their own interest in and contribution to fixed-mobile convergence (FMC). Therefore, the main drivers of FMC-SOTA are: fixed mobile substitution (FMS), operator/service providers, customers/end-users, equipment vendors, and technologies.

Nowadays, mobile telephone use is dominating and the revenue keeps increasing. On the other hand, fixed telephone use is reducing and its revenue generation is also decreasing, especially for voice service. If this trend keeps up, there is a high probability that traffic will entirely shift to mobile communication which is known as FMS. Thus, FMS is one of the major driving forces that push operators to converge the two networks either to keep or increase their revenue.

Important technologies and key architectures have been identified for FMC based service oriented architecture (FMC-SOTA) in developing countries [5].

Convergence of the current packet and circuit networks and services over both fixed and mobile networks is becoming very crucial in order to satisfy customer

needs, to reduce cost and introduce new innovative mobile services by mobile operators. FMC services can provide one or more of the following characteristics: seamlessness, user flexibility in access methods, converged customer premise equipment (CPE), personalization, and the best of services taken from mobile and fixed worlds.

In order to ensure the end user a continuity of experience, a total integration of the different access and core networks from the fixed and mobile sides needs to be achieved. Ovum [1] indicates four main stages of FMC development:

Stage-1: price bundling of fixed and mobile services. Bundles are delivered to customers at discounts, often via a single bill. The services and the networks which provide them remain separate.

Stage-2: Service convergence/orientation in which a few common services are offered over both fixed and mobile networks e.g. a single voice mailbox or linked email.

Stage-3: device convergence in which a common device offers both fixed and mobile network access with seamless handover.

Stage-4: network convergence in which fixed and mobile services run over a common IP transport network and use a common platform for service creation and control - the IMS.

Some countries are taking the lead in moving along the path of staged development [1,14]. Network convergence is the most important stage such that it involves the deployment of a common services platform that allows for the development of integrated services. Users can access these services from a range of devices over both fixed and mobile networks. This leads to “anywhere anytime” communication service architectures where the subscriber is always best connected in terms of bandwidth, quality of service and price.

2. Fmc Networking and Service Oriented Telecom Architecture

FMC includes convergence of services, networks, and end user terminals. Enabling standards to support FMC and the service orientation of the telecom architecture.

The standards focused are Third Generation Partnership Project (3GPP) and European Telecommunications Standards Institute (ETSI) Telecom & Internet Converged Services and Protocols for Advanced Networks (TISPAN). 3GPP is working on GSM networks evolution to IMS and ETSI TISPAN is developing standard for fixed networks migration to NGN taking IMS requirements into consideration. The two standard bodies are working together to build common IMS architecture for both fixed and wireless networks. IMS is built on layered architecture with open interfaces and it can provide IP multimedia services independent of the underlying access technologies.

2.1. IP Multimedia Subsystem (IMS)

IMS is being developed by 3GPP in collaboration with IETF, ITU-T, and ETSI. Initially, the main aim of IMS was to bring together the Internet and the cellular world, leading to service oriented telecom architectures (IMS-SOTA). To this end,

IMS uses existing Internet protocols such as SIP (Session Initiation Protocol), AAA (Authentication, Authorization and Accounting protocol), and COPS (Common Open Policy Service) to meet the requirements of reliable and real-time data communication. Further, IMS include easy quality of service (QoS) management, mobility management, service control and integration.

The core network elements or nodes of IMS communicate with each other using specific protocols; each of these interfaces are identified using a reference point label. A detailed list of all the interfaces and their operation is available in 3GPP TS 23.002 [6]. Also another important characteristic of the IMS architecture is that it exclusively uses IPv6; it requires network elements such as NAT-PT (Network Address Translation - Protocol Translation) and IMS-ALG (IMS Application Level Gateway) to interoperate with the traditional Internet (which mostly uses IPv4).

Few applications and services that can be developed on the IMS platform include: presence related services, instant messaging, push-to-talk services. The services that can be developed on the IMS platform are limited only by the creativity of the application developers and requirements of the customer.

3. Fmc-sota: Case of Dev. Regions

Fixed-mobile convergence is getting attention from telecommunications industry and operators in developing regions. There are service providers and vendors that are interested in the potential benefits convergence offers [14]. Work is also underway in this area in some standard organizations like 3GPP and ETSI-TISPAN. Migration to FMC may be evolutionary or revolutionary. It is the evolutionary one that is appropriate particularly for developing countries. This is because it enables them to invest their limited capital on future-proof technologies by leveraging the previous investment on legacy networks. However, directly moving to the complete convergent platform, which is revolutionary, is not recommended. This is because it needs high investment cost, affect the huge investment on legacy networks, result in less customer experience for multimedia services, and so on.

FMC enabling technologies and/or architectures should meet the following requirements, among others.

- Based-on standards in order to have the following advantages of open interfaces.
 - Easy service creation, modification and provisioning
 - Telecom operators can buy equipment from any supplier. This enables operators to have the best-breed product with reasonable cost.
- Enable telecom operators to implement future-proof technologies step-by-step. This enables operators to:
 - Protect investment on legacy infrastructure.
 - Gradually build customer base for the future IP multimedia services.
- Flexible and scalable network architecture.
- Simplified network architecture that enables telecom operators to gain capital and operational expenditures (CAPEX and OPEX) reduction.

- High revenue generation that guarantee return on investment.
- Guarantee end-to-end quality of service.
- Inter-working with packet- and circuit-switched networks.

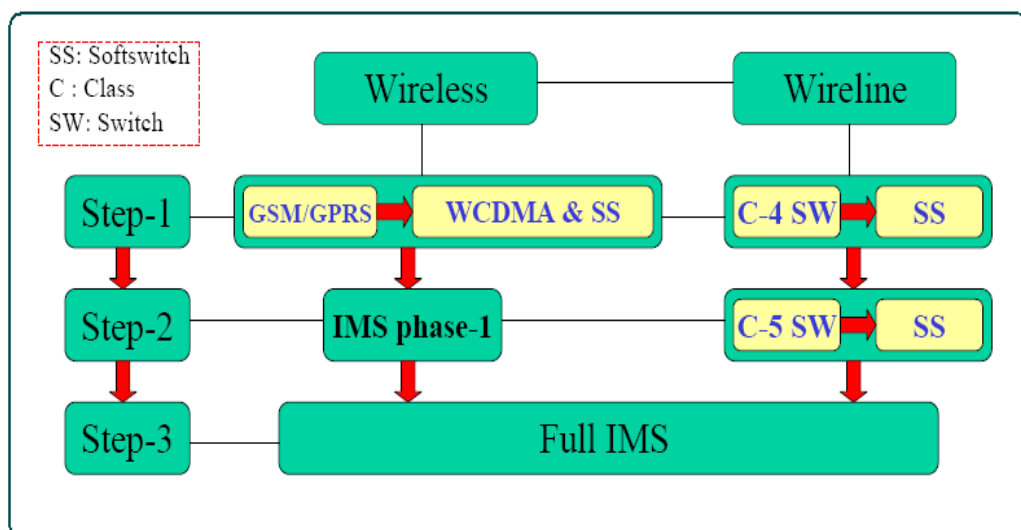
IP multimedia subsystem (IMS) can provide multimedia services over common IP backbone irrespective of the underlying access technologies and it is the one that meets the above requirements. Therefore, a migration path to IMS-based FMC architecture could take the following path [13].

Step-1: Network consolidation and building a foundation for IMS infrastructure

Step-2: Upgrading the consolidated platforms or injecting new IMS elements

Step-3: Full IMS deployment

Fig 1: Steps for migration to FMC/IMS [13,14]



Telecom operators should have a planned migration path to deploy a complete fixed-mobile converged architecture. The migration process must be self-financed by revenue from new services, and cost reductions from network consolidation and optimization. A three steps smooth migration path have been identified for developing countries, to fixed-mobile converged network architecture (as shown in figure-1 above).

Each step is applied to the specific case of fixed and mobile networks. Each of them can also be justified by new services and/or advantages brought from consolidation. Each will be discussed in the coming sections.

4.1. Step-1: Network Consolidation

This step involves network consolidation and building a foundation that will allow telecom operators to create flexible IMS architecture and hence to capture Capital and Operational Expenditure (CAPEX and OPEX) efficiencies. Through consolidation, it is possible to gain scale and reduce the number of network elements making the network more IMS-ready.

The existing mobile networks in developing countries provide basic voice service based on GSM 900/1800 MHz. It also is capable to provide short messaging service (SMS) and GPRS services. During this step, mobile network consolidation can be achieved by introducing wide-band code division multiple access (WCDMA) in the base station subsystem (BSS) and softswitch in the core network. 3GPP Release-99 [7, 8] basically introduces a new type of radio access network called UTRAN (UMTS Terrestrial Radio Access Network), which enhances the downlink & uplink speeds to about 2Mbps. 3GPP Release-4 [16] is defined to facilitate the evolution of the Release-99 circuit switched core network architecture into an IP transport development with a separation of control (MSC server) and bearer (media gateway) planes easing the introduction of packet switched backbone technology in the circuit switched domain. This is a pre-IMS phase so that the cellular network will be ready for initial IMS implementation.

The current trend in the telecommunications industry is the migration from voice-optimized switched circuit networks to multi-service packet networks. This migration will provide benefits in terms of economies of scale and revenue from new services. One of the greatest challenges in this migration is creating a packet-based infrastructure that will preserve the ubiquity, quality, and reliability of voice services while allowing the greatest flexibility for use of the new packet technologies. In practice, this means the retaining of current telephony-related services or the transparent offering of existing public switched telephone network (PSTN)-bound services on new packet-based deployments [10].

Telephony soft-switch solutions can be deployed to support today's legacy call-oriented traffic over an IP-based infrastructure, with telecom-grade scalability and quality of service. Beyond voice services, IP-based networks open up opportunities for integrated voice applications and services – such as Video Telephony and presence aware telephony – that would be impossible (or prohibitively expensive) using traditional Time Division Multiplexed (TDM) technology [18].

The fixed network is providing basic telephony service and supplementary services based on class-based TDM switches and signaling system #7 (SS7). During this step, it is recommended to replace class-4 switches, which are national tandem switches, by softswitches. Class-4 switches represent a smaller investment for the service provider and with less functionality than a Class-5 switch, means a simpler and therefore lower risk point at which to move to packet voice.

4.2. Step-2: Upgrading the Consolidated Platforms or Injecting New IMS Elements

By enabling and upgrading the consolidated platforms created in step one and/or injecting new IMS elements into the network, network operators can enhance the existing services and/or launch new IMS powered applications and services. This step basically involves the initial deployment of IMS in mobile networks, and providing voice over packet up to the access level and other high revenue generating value added services by upgrading the softswitch in fixed network. It is the first step towards using IP session control mechanisms in the packet switched domain, which is IMS, enabling flexible session negotiation for multimedia applications in mobile network.

In mobile network evolution of step one above, the GPRS main components (GGSN and SGSN) that brings packet data services to the mobile network, the UTRAN in Release-99, and the softswitch (MSC server and media gateway) in Release-4 have been introduced. This resulted in a good mobile network foundation for IMS implementation. The basic introduction of the IMS network arrives in 3GPP Release-5 set of standards, while enhancements such as PSTN and IPv4 inter-working are added in Release-6, which will be discussed in step-3. IMS is an overlay control network that reuses the packet domain of the Release-99/Release-4 networks [11].

In this second step of mobile network evolution, new IMS elements will be introduced in the already established network foundation. This involves the implementation of 3GPP's Release-5 [2, 3]. Some major release-5 items [11] are as follows.

- IP-based multimedia services (i.e. IP multimedia core network subsystem).
- Reliable QOS for packet switch domain (end-to-end QOS).
- Introduction of IP UTRAN.
- Introduction of High Speed Downlink Packet Access (HSDPA).
- CAMEL phase 4.
- Wideband adaptive multi-rate (AMR) coding.

The introduction of softswitch to legacy fixed network has been discussed in evolution step one. In first step, it was recommended that softswitch replaces the legacy class-4 or national transit switches and gradually offloads voice trunking traffic to the packet network. This also brings the advantage of network consolidation due to the distributed media gateways for trunking functionality while all can be controlled by centralized media gateway controller. Further more, the transport infrastructure in the core network may also be optimized due to the usage of IP transport.

In step-2 of fixed network evolution, the replacement of class-5 switches by softswitches is recommended to provide IP-based services to the end user. Developing countries should gradually replace these legacy class-5 switches. This can be done in parallel with step-1, which places trunking gateways alongside existing class-5 switches to provide access to the packet network. Access gateways (AGs) can be used to supplement or replace existing TDM access nodes for voice connections to the packet network. At the end of life the remainder of TDM exchange equipment can be replaced by media gateways.

4.3. Step-3: Full IMS deployment

In the third and final step of the roadmap follows a move to full IMS deployment, utilizing the IMS enablers to create new multimedia services and applications not previously possible.

Bear-in-mind that in the previous step of mobile network evolution, the major achievements are 3GPP Release-5 enhancements like HSDPA and initial IMS deployment. In step-3, introduction of 3GPP Release-6 to mobile network is recommended. 3GPP Rel'6 is an enhancement of earlier 3GPP releases that

continues UMTS momentum by enabling greater speeds, capacity improvements and new applications. It includes numerous new features such as:

- High Speed Uplink Packet Access (HSUPA).
- The second phase of IP Multimedia Subsystem (IMS).
- Inter-working with Wireless Local Area Networks (WLAN).
- Multimedia Broadcast Multicast Service (MBMS).

During this third step, IMS also serves wireline access. ETSI-TISPAN [12] drives the application of IMS for wireline access in its efforts on the Next Generation Network (NGN), in collaboration with 3GPP. The service definitions for NGN multimedia telephony will also be reused and adopted for wireless accesses by 3GPP. When all legacy equipment has been removed and SIP based IMS is fully introduced, the target network has been reached.

3GPP release-6 and beyond are basically characterized by enhancing or improving features of previous releases and also focus on inter-working with other networks such as PSTN and WLAN. ETSI TISPAN is also working on fixed network migration to NGN based on IMS architecture. Its first release provides basic architecture built on this concept [11]. Other consecutive releases that will add new features and/or enhance features of the previous release (release 1) are under way.

IMS is a layered architecture where service enablers and common functions can be reused for multiple applications. The architecture in IMS also specifies interoperability and roaming, and provides bearer control, charging and security.

For end-users, IMS-based services enable person-to-person and person-to-content communications in a variety of modes – including voice, text, pictures and video, or any combination of these – in a highly personalized and controlled way. IMS provides services like presence, video telephony, and instant messaging that will bring quicker return on broadband investments.

4. Conclusion

Service oriented and converged telecom architectures are necessary to enable the network to support the thousands of mobile services expected by customers in the future [5,9]. Standard bodies and different stakeholders in the telecom industry are making effort and contribution to achieve fully converged network architecture to make this possible. This is reflected in the convergence of 3GPP IMS and ETSI TISPAN NGN. The IMS standards are promising network development for fully converged fixed and mobile/wireless networks and service orientation.

Convergence enables telecom operators to use the mobility advantage of wireless networks and the Broadband infrastructure of the fixed network in a seamless fashion, to provide a platform for service orientation and enablement of innovative mobile services. The convergence criteria used were protection of investment on legacy infrastructure, Service orientation: ease of developing, activating & billing of new services, lower network OPEX cost . Other criteria were adherence to international telecom standards & Inter-working with packet- and circuit-switched networks which guarantee smooth functionality.

Acknowledgements

The authors would like to thank colleagues at the CIT for valuable comments and support, and the various standards Organizations for providing us their documents.

References

- 3GPP, "Service Requirements for the IP Multimedia Core Network Subsystem", Release 5, Technical Specifications Group Services and System Aspects, TS 22.228 V5.7.0, March 2006.
- 3GPP, "IP Multimedia Subsystem Stage 2", Release 5, Technical Specifications Group Services and System aspects, TS 23.228 V5.13.0, November 2004.
- 3GPP2, "IP Multimedia Domain System Requirements", 3GPP2 S.P0058.
- 3GPP, "Network architecture", Technical Specification Group Services and System Aspects, 3rd Generation Partnership Project (3GPP), TS 23.002 V7.0.0, December 2005. http://www.3gpp.org/ftp/specs/latest/Rel-7/23_series/
- 3GPP, "3rd Generation Mobile System Release 1999 Specifications", TS 21.101. <http://www.3gpp.org/ftp/Specs/>
- 3GPP, "Overview of 3GPP Release 99", ETSI 2004. www.3gpp.org
- 3GPP, "IP Multimedia Subsystem (IMS); Stage 2", Technical Specification Group Services and System Aspects, 3rd Generation Partnership Project (3GPP), TS 23.228 V7.2.0, Sep 2005. http://www.3gpp.org/ftp/specs/latest/Rel-7/23_series/
- ENDALE SHIFERAW "Strategies for Fixed mobile convergence: The case of ETC" MSc Thesis, GSTIT, Nov, 2006, Addis Ababa, Ethiopia.
- ETSI ES 282 007: "Telecommunications and Internet converged Services and Protocols for Advanced Networking (TISPAN); IP Multimedia Subsystem (IMS); Functional Architecture", ETSI ES 282 007 V1.1.1, March 2006.
- Future Communications Networks & Services for Emerging Markets", F. Mekuria & E. Sutherland, Proc. of IEEE international conference on "Innovations in IT"-IIT2006, Nov. 19-21, 2006, Dubai, UAE.
- F. MEKURIA, "Educating the Architectures of the Mobile Economy.", Proc. of W3C Workshop, on the Role of Mobile Technologies for Social Development, Sao Paulo, Brazil, June 2-3, 2008.
- F. MEKURIA, E. SHIFERAW, "Fixed Mobile Convergent Networking for Service Oriented Telecom Architecture." Proc. of NEPAD ICT Africa Conference,
- OVUM LIMITED, "Review of the Regulatory Framework for Fixed-Mobile Convergence in Hong Kong", April 2006.
- PAUL G. A. SIJZEN, JOHN P. L. SEGERS, LOUISE F. A. SPERGEL, AND JACK KOZIK, "Building the Bridge: Devising an Architecture to Migrate Voice-Band Calls to Packet Transport and Multimedia Services", Bell Labs Technical Journal, July–September 2000.

PART 5



Information Technology

28

Comparison of Rwandan and Russian ICT Educational Policies

*Irina Zlotnikova and Voronezh State University of Education
(Voronezh, Russia)*

The paper contains results of the comparison between African (namely Rwandan) and Russian ICT educational policies. First the author presents a review of ICT educational strategy in Rwanda, including The Vision 2020 Plan, National Information and Communications Infrastructure Plans and international ICT projects in Rwanda. Then the author gives a review of ICT educational strategy in Russia, including Informatization of Education Process, Program of Computerization of Rural Schools and some non-governmental initiatives. Several parameters, like the level of ICT literacy, the degree of integration into the international educational system, technical base of ICT, the degree of syllabi unification, the syllabi contents, the degree of students' motivation to ICT learning, have been compared. The author draws the conclusion that the Rwandan ICT educational policy being completely realized should lead to prosperity and educational superiority of the country. The Russian ICT educational policy has been quite effective for a long time but today has to be adjusted to the contemporary world.

1. Introduction

The paper contains results of the comparison between African (namely Rwandan) and Russian ICT educational policies. The author has an experience of 17 years of working in Voronezh State University of Education (Voronezh, Russia) and 1 year of working in Kigali Institute of Science and Technology (Kigali, Rwanda, 2/2006 – 2/2007).

This analysis does not pretend to be complete and reflects the subjective opinion of the author, especially regarding to African education. But the author hopes to develop her experience working in Makerere University (Kampala, Uganda). The author realizes also that the comparison between Rwandan and Russian ICT educational policy could be not quite correct since our countries are very different.

The paper addresses the following questions:

1. What is common in Rwandan and Russian ICT education and what is the difference between them?
2. What are positive features of Rwandan and Russian ICT education and ICT educational policies?

3. How ICT educational policies can be improved to satisfy social requirements and conditions of the contemporary world?

2. ICT Educational Strategy in Rwanda

Since Rwanda has limited natural resources, comparatively high density of population (340 people per square kilometer) and high population growth rate (5.5%), the only way to prosperity is a knowledge-based economy. The Rwandan government develops the consistent ICT strategy.

2.1 ICT Components of the Vision 2020 Plan

The Rwandan government has released a plan for Rwanda's social and economic development, with the ultimate goal of being a prosperous nation by 2020. The Vision 2020 Plan is centered on "a prosperous knowledge-based economy." ICT are mentioned as extremely important components of the development of the educational system in Rwanda.

According to this plan by 2020 Rwanda will have scientists, engineers and technicians to meet the needs of the economy, and that they develop this through secondary and tertiary education, including an increase in the number of professional and polytechnic universities.

Meeting this plan Rwandan higher education institutions (for example, Kigali Institute of Science and Technology/KIST which is a key player in ICT education in Rwanda) develop their own Program 2020 Envisioning.

2.2 National Information and Communications Infrastructure Plans

In 2001 the Rwandan government released the 1st National Information and Communications Infrastructure (NICI) plan, "An Integrated ICT-led Socio-Economic Development Policy and Plan for Rwanda" to address a plan for the ICT development of the nation between 2001 and 2005. There are subsequent 2nd, 3rd, and 4th phase plans guiding ICT policy through 2020. The 1st NICI plan (2001-2005) aimed to develop an approach to the knowledge-based economy around ICT. The 2nd phase of the plan (2006-2010) is centered on strengthening this new economic base. The 3rd plan (2011-2015) looks to sustain the development and compete in the global market for information-based services and products. The 4th plan (2016-2020) will finally explain how the new processes and development can bring Rwanda to middle-income status.

2.3 International ICT projects in Rwanda

Rwanda is a participant of the imfundo project (means "education" in the Ndebele language). Imfundo is the team within the United Kingdom Department for International Development (DFID) which creates partnerships between several DFIDs, large HiTech companies and developing countries to deliver information technology-based activities supporting gender equality and universal primary education in sub-Saharan Africa. The aim of the project is to raise attainment through distance learning and professional development courses in ICT for teachers.

Some students have been studying ICT courses through the African Virtual University which is allowing students to learn online, while being taught by lecturers from other countries.

In October 2006 NEPAD (New Partnership for Africa Development) e-Africa Commission launched a project to further develop ICT in Rwandan schools. The project will link up schools across Africa. The project will include primary and secondary level and is intended to grow so that eventually it will incorporate all Rwandan secondary schools.

Global Education Alliance is a new venture of the Global Education Initiative (GEI). It is a multi-stakeholder partnership for education created with UNESCO, private companies and the Education For All Fast Track Initiative (EFA-FTI). Rwanda is the first country to pilot the multi-stakeholder alliance.

Rwanda has been testing the One Laptop per Child (OLPC) technology developed by Massachusetts Institute of Technology (MIT) Media Lab in effort to enhance the quality of learning. A non-profit association OLPC, independent of MIT, promotes the inexpensive laptop as a learning tool.

3. ICT Educational Strategy in Russia

The Soviet and Russian government always paid special attention to science and technology education. In the second half of XXth century Computer Science (and ICT as its part) has become one of the most important scientific disciplines. Outstanding Soviet scientists addressed to the government with the question about compulsory studying of Computer Science at secondary schools and universities.

3.1 Informatization of Education Process

Informatization of Education is the long-term process originating since 1985 when Computer Science has been included to the school curricula as a compulsory subject, alongside with mathematics, physics and other scientific disciplines. Simultaneously the schools have been supplied with computer facilities. First computers were developed specially for schools by Soviet research institutes and enterprises. They had no world analogues. But since 1990s the schools have been supplied with IBM PCs. In 2006 in schools one computer was shared by 13-45 pupils, in vocational schools – by 4-15 students. Up to 69 % of schools had access to the Internet.

In 1990s Computer Science has become compulsory for all higher education institutions and for all specialties. It has been included to the State Educational Standards. Today all higher education institutions have faculties and departments of Computer Science/ICT. There are also research institutes which study the multiple-sided problem of Informatization of Education.

3.2 Computerization of Rural Schools Program

In 2000 the Russian government has paid attention to the fact that graduates of rural schools are in the worst position than graduates of city schools. While city schools were equipped with modern computer facilities, rural schools had no computers, with exception, maybe, one out-of-date computer in the director's office.

In 2001 every rural school has been supplied with at least one modern computer. At the next stage of the program, in 2004, each rural school had not less than 4-5 computers and one computer in the school library. Computers had an Internet access through fiber optic cables.

Today almost all rural schools have at least one computer class. Normally, it is used very intensively for teaching Computer Science. There is no time left for other subjects. Teachers of non-computer subjects have no possibility to teach their disciplines using computer facilities. As a result the graduates of rural schools have an idea about computers, they have seen computers and touched them, but they are not computer/ICT literate in full sense of this expression.

Other problem of rural schools is lack of qualified teachers (especially of Computer Science). It is the same problem, like in Africa and all over the world.

Teachers of non-computer subjects are mostly ICT illiterate. They have to be prepared to using ICT in teaching their subjects.

3.3 Federation of Internet Education (FIE)

The problem of teachers' ICT literacy is solved by many governmental and non-governmental projects. But the project of Federation of Internet Education (sponsored by Yukos Company) was especially impressive. According to the general plan of this project started in 1999-2000, well-equipped learning centers have been created in every region/district of Russia. The author has had an experience of working in one of these centers, namely Voronezh Regional Center of FIE.

The teachers of different subjects were trained intensively during two weeks. They studied how to create their own digital educational resources and how to use existing ones. Every center graduated 1000 to 2000 teachers per year. Unfortunately, in 2004 Yukos Company has had serious financial problems which led to its closing. As a result Federation of Internet Education doesn't exist anymore.

3.4 Intel® Teach to the Future

The international educational project Intel® "Teach to the Future" is partially sponsored by Intel Company and partially by higher education institutions where it is applied. Today more than 4 millions teachers from 35 countries participate in it. This project has no official support of Russian government but it corresponds to the Russian ICT educational policy.

This projects aims to help teachers to seize ICT as far as innovative methods of teaching, to use it in their professional activity then teaching and preparing to lessons. It is targeted to expanding of advanced technologies into education.

4. Comparison between Rwandan and Russian ICT Educational Policies

4.1 Higher Education System as a whole

Today in Rwanda there are 20 higher education institutions, 6 of these are governmental ones. Most of them are situated in Kigali, the capital of Rwanda. The

oldest institution, National University of Rwanda (NUR), is situated in Butare, the leading academic center of the country. Rwandan education has a support not only from the national government as it was mentioned above, but also from the international community which feels partially responsible for the genocide of 1994.

In Russia there are 685 governmental higher education institutions, all of these having state accreditation. Besides, 619 non-governmental higher education institutions have been licensed for educational activities. Thus, the number of higher education institutions is 1,304. In 2003–2004, the total number of students of higher education institutions was 5,947,500.

In the Soviet Union, education of all levels was free for anybody who could pass entrance exams; students were provided with small scholarships and free housing. It provided access to higher education to all gifted people. More than 20% of Russians of age 30–59 hold five/six-year degrees (this number is twice as high as that of the United States).

Institutions had to be funded entirely from federal and regional budgets. After the collapse of the Soviet Union institutions found themselves unable to provide adequate teachers' salaries, students' scholarships, and maintenance for their facilities. To solve these problems many state institutions started to open commercial positions. The number of those positions has been growing steadily since then as far as tuition fees.

4.2 Basic Computer Literacy

The literacy rate in Rwanda is currently 49%. The literacy rate in Russia is 99.4%. As for computer/ICT literacy, in Rwanda, according to the author's experience, most school graduates entering the higher education institution never have seen computers before.

In Soviet Union the Computer Science had been proclaimed as a compulsory school subject (1985). Since then, every school graduate is presumed to be ICT literate, but, of course, there is some exclusion, especially in rural schools. But, again according to author's personal experience, last few years almost all students entering Voronezh State University of Education have certain ICT skills and know how to use computers in their future professional life. Many of them have their own computers (for work and entertainment).

4.3 Integration into the International Educational System

The Rwandan educational system is completely integrated to the international education system and can be considered as a part of it. It presumes the same academic degrees like in UK, US and many other countries – Bachelor's, Master's and Doctor's ones. If students want, they have many possibilities to continue their education abroad. Most of them possess good communicative skills in English and/or French, since the teaching language in secondary schools and higher education institutions is English or French.

The author should mention that the Russian (former Soviet) educational system has been isolated for a long time and is still not completely integrated into the international educational system. The system of academic degrees in Russia differed from the international one. It creates certain problems for Russian professionals who want to develop their experience of working or learning abroad. Last years Russia has become a part of the Bologna Process. Recent graduates have Bachelors' and Masters' degrees. But those degrees didn't exist during Soviet and post-Soviet period.

Russia has more academic graduates than any other country in Europe, especially doctorates. There are two doctor's degrees. First of these, Candidate of Sciences is roughly equivalent to the Ph.D. in the United States, UK and other countries. Second doctoral degree is Doctor of Sciences. The average time between obtaining Candidate of Sciences and Doctor of Sciences degrees is roughly 10 years, and most of new Doctors are 40 and more years old. Only 1 of 4 Candidates reaches this grade.

4.4 Technical Base of ICT Education

In Rwandan universities one could observe the comparatively poor technical base (too few computers, lack of licensed software) and too large students groups especially when conducting ICT labs. There are 3-4 or more students working at the same computer. It is easy to imagine the student who studies computers but never has touched the keyboard.

In Russia as a result of the Informatization of Education process there is a quite good technical base of higher education including contemporary models of computers and mostly licensed software. It is granted by federal and regional budgets as far as commercial positions and other forms of non-budget financing. When conducting ICT labs, students' groups are divided into smaller subgroups (10-12 students), according to the principle "one student – one computer". For tutorials the amount should not exceed 25 students in the group.

4.5 Educational Standards and Syllabi Content

The students of computer-based specialities in Rwanda are taught according to the international standards and using syllabi of remarkable universities from all over the world. These syllabi have been found in the Internet and got adapted to the local situation. The members of the academic staff are involved into the process of module specification. But state educational standards which would be compulsory for all higher education institutions don't exist.

Russian education, including tertiary one, last few years is in process of total unification and standardization. The main goal of this process is to achieve a better quality of education as a whole. As it was mentioned above there are hundreds of tertiary education institutions in Russia. But for some reasons not all of them can provide an equally high level of education quality. One of these reasons is that in the past the lecturers could teach what and how they wanted according to

their own syllabi. So it was necessary to develop a structure and content for all existing curricula to provide a standard high level of education quality. The syllabi are elaborated in accordance with the State Educational Standards which regulate almost 80% of their content. But the other 20% are elaborated by the university itself. As a result any student of given specialty in any university wherever it is situated will be taught according to the same syllabus.

The contents of Computer Science syllabi in Rwanda and Russia are different. For example, in Rwanda the students of both computer and non-computer specialities learn C as a first programming language. Then they study C++ . There is no Programming subject at secondary schools with exclusion of most prestigious ones.

As a first programming language in Russia teachers in schools mostly choose Pascal/Delphi and Basic. There is also special School Algorithmic Language which is close to the natural Russian language, like Basic is close to natural English language. There were also some attempts to teach Logo and its local versions as a first programming language. C is a professional programmer's tool unlike Basic which is a programming language for beginners. But it's very easy to learn and apply Basic for solving elementary problems.

4.6 Students' Motivation to Learning

In Rwandan universities the system of credits and ratings stimulates students to work actively during all the semester. Students can gain 50 marks for working during the semester (CATs and assignments) and 50 marks for the examination maximum, so their final marks depend not only on their efforts right before and at the examination. Rwandese students are highly motivated to learning because they believe that their future career and their life as a whole depend on results of their learning in the university. They respect education itself and want to be educated people.

It is necessary to say that Russian students en masse are poorly motivated to learn during the semester, since the main tool of assessment is the mark at the examination. There are just 4 marks - "5=excellent", "4=good", "3=satisfactory", "2=unsatisfactory". The student can do nothing during the semester, then make a short-term effort before the examination and get the positive mark. But now Russian teachers try to change this situation stimulating students to work intensively during the semester. One of the most effective tools of in-semester assessment, control and monitoring is computer testing. It is widely applied in Voronezh State University of Education. Today for some part of young Russian people education has no value, since they can see that educated people sometimes have lower salaries and live worse than people who have no education.

5. Conclusion

The government of Rwanda realizes that ICT education is a key tool in development of the national economy. The Rwandan ICT educational policy being completely realized should lead to prosperity and educational superiority of the country.

Thus Rwanda will have scientists, engineers and technicians to meet all needs of the economy. Then there will be no necessity to send students to study abroad or to invite expatriate ICT professionals.

The Russian ICT educational policy has been quite effective for a long time. As a result of emphasis on science and technology education, Russian ICT as far as medical, mathematical, scientific, space and aviation research is generally of a high order. The effectiveness of Russian ICT educational policy can be proven with the fact that the author has seen many Rwandese ICT professionals who were graduated by Russian/Soviet universities as far as Russian citizens who work as programmers or ICT teachers. But now Russian ICT educational policy has to be adjusted to the contemporary world.

6. Acknowledgements

I wish to thank Professor V.K. Srivastava and Dr Donart Ngarambe from Kigali Institute of Science and Technology (Kigali, Rwanda) who kindly explained me basic features of African ICT Educational Policy while I have been working there; Professor Alexander V. Mogilyov from Voronezh State University of Education (Voronezh, Russia) with whom I discussed details of Russian ICT educational policy and Professor Yaroslav A. Vagramenko from Moscow State Open University of Education (Moscow, Russia), the leading expert in Informatization of Education Process.

References

- ZLOTNIKOVA, I. (2004). *Theoretical Bases of Distant Informational Training of the Pedagogical Staff*. Voronezh State University of Education Press, Voronezh, Russia.
- ZLOTNIKOVA, I. (2004). *Practical Engineering of the Distance Education System for Informational Training of the Pedagogical Staff Based on Web Technologies*. Voronezh State University of Education Press, Voronezh, Russia.
- ZLOTNIKOVA, I. (2006). *Analysis and Comparison of Russian and African Teaching Experience*. In *Informatization of Education and Educational Management*, L. A. BACHURINA, Ed. Voronezh Teachers' Training Institute Press, Voronezh, Russia, 81-84.
- ZLOTNIKOVA, I. (2006). *Teaching Teachers – How to Construct Your ICT Lesson Effectively*. In *ICT in Education (Volume 1)*, V.N. Malyshev, Ed. Lipetsk State University of Education Press, Lipetsk, Russia, 136-143.
- ZLOTNIKOVA, I., ZOLOTAREVA, O. (2008). *Using Adaptive Tests at the Pedagogical University*. In *ICT in Education (Volume 3)*, S.L. BLYUMIN, T.P. FOMINA AND I.N. FROLOV, Eds. Lipetsk State University of Education Press, Lipetsk, Russia, 208-213.
- ZLOTNIKOVA, I., KOZLOVA, O. (2008). *Computer Testing: Practical Experience*. In *ICT in Education (Volume 3)*, S.L. BLYUMIN, T.P. FOMINA AND I.N. FROLOV, Eds. Lipetsk State University of Education Press, Lipetsk, Russia, 213-219.

29

A Generic E-Learning Framework: A Case Study Faculty of Computing and Information Technology (CIT) Makerere University

Jude T. Lubega and Fiona Ssozi Mugarura

Advancement in Information Technology (IT) has led to better facilities and services for learning within educational organizations. Many higher educational institutions of learning have adopted technology in their learning methods in order to enhance the traditional learning method and to offer distance/purely online learning. This enhancement is in the form of management, administration, facilitation, delivery and support of the modules/courses delivered using the electronic media. Makerere University's strategy for ICT implementation strives for use of technology during learning hence its adoption by several faculties within their teaching methods. Makerere University Faculty of Computing and Information Technology (CIT) developed a generic e-learning framework to guide it in implementing an effective blended learning process (traditional classroom and e-learning). CIT's goal is to implement a complete e-learning program gradually by first initiating a blended learning approach. The lessons learnt from the blended learning approach would enhance implementation of a complete e-learning program. This paper describes the generic e-learning framework developed by CIT in the implementation of the blended learning approach. The framework constitutes several components that are derived from the organizational environmental factors. These components include organizational policy, funding, stakeholders, services and infrastructure. The components' constituting factors are inter networked to allow a specific learning method (blended, purely online or distance learning) to be implemented. The framework in operation is illustrated by a typical case study of blended learning in CIT.

1. Introduction

Advances in IT and new developments in learning provide opportunities to create well-designed, engaging, interactive, affordable, efficient, easily accessible, flexible, meaningful, distributed, and facilitated learning environments [Khan, 2001]. Many organizations such as Makerere University Faculty of Computing and Information Technology (CIT) have utilized the opportunity that technology brings to offer ubiquitous learning. CIT's future plan is to gradually offer a complete online program in any of its degree courses. However, e-learning implementation can only be adopted in phases until CIT is in position to deliver a complete online program. The focus of CIT is on creating an effective foundation for pedagogical needs and appropriateness of the implemented e-learning method. Therefore, the teaching,

learning and assessment have to be linked to the learning outcomes, module/course delivery and assessment objectives. In order to attain this objective a generic e-learning framework was developed by CIT. This framework was meant to guide the entire process of implementing e-learning in the faculty and gradually to the entire university. The UNESCO report [2003] suggests that the African education status requires innovative ways to support it in achieving the millennium goals. These innovative features can be achieved through the integration of technologies that can process information and transmit it for the purpose of learning and educational development [UNESCO, 2004]. Many African countries have lagged behind in the implementation of technologies for enhancement of services such as education. Many hindering factors such as access to funds, electricity and human capacity slow down the implementation of ICT within education.

It was estimated that the overall demand for higher education is expected to grow from 48 million enrolments in 1990 to 159 million in 2025 [LaRocque and Latham, 2003]. Therefore there is a need for other alternative educational methods that can be used for knowledge dissemination. In the past two years CIT has seen a tremendous increase in students applying for its courses. However not all applicants are admitted to undertake courses at CIT which leaves the demand for admission high. In order to cater for the masses of applicants, CIT has devised means of adopting e-learning in its teaching methods. It is believed that with the introduction of e-learning, education will be delivered ubiquitously to whoever needs it cheaply and easily. Research has shown that whenever there is an abrupt increase in the demand for learning but without adequate resources, inefficiency and ineffective in service provision is the result [Adelman & Taylor, 2000]. This is true for CIT where the demand for its courses surpasses what can be offered to the population. In such circumstances many African institutions have not bothered to find a solution through research. World Bank [2000] also notes that there is lack of ICT research in Africa in the areas of effective educational uses and potential impacts on the quality of African education. On this note the paper addresses how CIT has undertaken the gradual implementation of ICT within its courses.

2. Learning Process

Learning can be defined as a process of knowledge construction that involves different stakeholders (learners, tutors and instructional designers) undertaking several activities like courseware development, assessments, application and feedback generation. The different stakeholders aim at achieving different objectives depending on the activities carried out within the learning process. Constructivism theory states that, learners play an active role and take on responsibility to construct their own knowledge and meaning [Fosnot, 1996; Steffe and Gale, 1995; Honebein *et. al.*, 1993]. During learning learners act in different ways based on their own understanding hence making meaning out of the learning process. On many occasions learners prefer different delivery methods for content presentation because it affects how they act on it. The constructivist theory has

been used in the design of constructivist e-learning environments. Within these e-learning environments, learners take on responsibility for their learning process by managing the learning activities and collaboration with others. The constructivist e-learning environments are capable of offering personalized content, track learning activities and offer dynamic feedback to the learners.

For any e-learning process there are several factors that are necessary for its effective adoption. Several e-learning frameworks been developed to guide organization in adopting e-learning [Khan, 2001; Jennifer, 2005]. Many of the frameworks are developed not to suit the developing world context hence being difficult to adopt in Africa. However there are key components that are always considered including courseware instructional design. There are different instructional design methods that have been adopted during courseware development including the Learning Object (LO).

Several organizations in the world such as SCORM [2003], IEEE [2003], and IMS [2003] have greatly contributed to the standardizing of content development and e-learning delivery. The learning objects concept is part of the many standards that have been set by these organizations to enhance courseware instruction. Wiley [2003], states that a learning object is any digital or non digital chunk of content that has been split into reusable small and manageable chunks of content. Cisco Systems [Cisco 2001] was a pioneer in the development of the learning objects and content objects architecture. This architecture has been further development to suite the higher education [Sun and Williams, 2004]. This higher education learning object concept contains overview, learning objectives, information objects, practical objects, assessment objects and summary. With such structure embedded into learning courseware, students can effectively undertake a self study session which is encouraged by e-learning.

3. A Generic E-learning Framework

A generic e-learning framework has been developed in relation to the institutional setting and existing ICT policy. This framework considers the environmental constituents as the main dynamic factors to the e-learning process. The learning environment can be influenced by several things that include available resources, infrastructure, services and literacy. Many frameworks adopted in developing countries ignore some important constituents hence affecting the entire e-learning process. The developed generic e-learning framework contains main constituents such as organizational policy, funding, stakeholders, services and infrastructure. When these constituents are interlinked together they are able to effectively deliver an effective e-learning process. Figure one describes the generic e-learning framework that CIT developed to effectively help in the implementation of blended learning.

The framework signified the importance of the existing environmental set up when adopting a new technology to enhance e-learning practices. Many times it is assumed that with technology everything is possible but it can be noted from

the framework that other constituents such as stakeholders, services, policies and funding resources greatly impact on the e-learning process. A description of the different constituents and how it interlinks with others is made in the following sections.

3.1. Organizational Policies

This refers to a set of guidelines set up by the organization to guide in administering of the organization. These guidelines stipulate what should be done, how it can be done and the end result. The policies that affect e-learning implementation include course accreditation, teaching and ICT policies. When the faculty decides to offer an online course, the University has to agree on the accreditation of the course based on the existing policy for accreditation, the delivery methods, appropriateness of the courseware and target group of learners. The University's policy on teaching using ICT encourages the faculty in implementing e-learning. Therefore these policies have a direct impact how e-learning is implemented within an organization. The policies also outline clearly how the monitoring and tracking of the e-learning implementation is carried out within the organization.

Fig. 1: Organizational Policies



3.2. Funding

This refers to both the money spent (expenses) on the implementation of e-learning and the one generated from the revenues (training fees, tuition, etc...). Money can

be spent on several items such tutor recruitment, human capacity development, rent for e-learning platform, purchase of hardware and software. The funding directly and indirectly influences the type of e-learning that will be provided based on initial input (expenditure). The funding will determine what stakeholders to be recruited (staff) for teaching and content creation. It also influences the quality of infrastructure to be installed for delivering the e-learning process. Funding also determines the type and quality of services that can be offered. The fees paid by the students should be well articulated to indicate the reasons why they are paying a specific amount as compared to another. Full time students will end up paying more fees than part time students.

3.3. Stakeholders

These are people that participate in the e-learning process through different activities. The stakeholders include administrators, tutors and students.

- **Administrators:** these are people who carry out the administering of the entire e-learning process. Administering includes the support offered to manage the module/courses that are offered using this ICT. The administrators may include network administrators, mentors, departmental heads, e-learning managers, policy makers etc... These are key stakeholders in the planning and play a major role during the e-learning process implementation.
- **Tutors:** these are teachers who teach, design or provide learning content that is delivered electronically to the students. The tutors develop instructional content that conforms with SCORM standards which they pass on to students through an e-learning environment. Tutors also streamline, monitor, engage and offer constructional feedback during the e-learning process. Without the tutors to develop the content, engage and monitor students' effective e-learning can hardly occur. Therefore the tutors have to be well equipped with the pedagogical and instructional design principles/techniques.
- **Students:** these are the individuals who take a contract to learn using the electronic means. The students recruited to the e-learning process come from different backgrounds and have different motives for undertaking the learning. Knowing the common background and motives of the students enhances the content development process. Life long learners have different learning motives as compared to students learning in order to acquire employment. The students' learning levels have to be determined too before they undertake the e-learning process.

3.1. Services

This is the assistance that is offered by the stakeholders using the available ICT to enhance the e-learning process. This assistance offered through ICT is in different ways as described below.

- **Administration:** this assistance refers to the support that is offered in the administration and management of ICT that is used in enhancing of the e-learning process. Adoption of ICT within administration can greatly improve on the managing of other services such as communication between students, staff members etc. Using ICT to administer the flow of learning resources within the e-learning environment is very vital for the effectiveness of the process. Having a timely flow of information between the stakeholders is very important for the e-learning process. For example administration can assist in managing timetable information, course announcements, and tutor information.
- **Development:** this assistance involves designing and implementing both infrastructure and learning content for the e-learning process. Design of the network that will facilitate the e-learning process is very vital. For example poor implementation may lead to an ineffective internet connection hence leading to poor e-learning process. The design of learning content that conforms to the pedagogy is very crucial in this service hence leading to the constructive content. Curriculum and staff development is another vital service that is very essential within the e-learning process. The staff members have to undergo training very often in order to constantly have updated skills on offering an effective e-learning process.
- **Support:** this assistance refers to any help that is offered to the stakeholders during the e-learning process. Support may be offered in different forms, to the students they may receive feedback for improving learning and the tutors may receive bonus funding to motivate them during content development. The tutors may receive other forms of support such as training that is essential for attaining the e-learning pedagogical skills. There is also a need to sensitize all stakeholders about the e-learning environment before they can embark on using it. Students need to be sensitized on how to use the e-learning environment and its effectiveness to their learning process.
- **Delivery:** this assistance involves the dissemination of information and learning content to the stakeholders using ICT. The delivery methods may vary from one stakeholder to another. Tutors may pass on the necessary information to the students through emails and discussion boards. It is very vital to decide on the delivery method that will be implemented for the learning method you want to embark on (is it a blended, purely online or distance learning approach). Some tutors may prefer to deliver the content through the use of CDs or e-learning platform. Different delivery methods have merits and demerits depending on the circumstances in which it is being utilized for learning.
- **Access:** this assistance involves offering stakeholders the appropriate avenues through which they can interact with the information offered in the e-learning process. The access that is required by the tutors may be

different from what is required by the students. Is e-learning offered at a 24 hour basis or the e-learning process occurs at only designated times of the day? Can students access the e-learning environment ubiquitously or restricted to the university circumference? Some tutors prefer to avail the online assessment at specific periods for the sake of monitoring.

- **Instruction:** this assistance involves selection and enforcing a particular instructional method for designing the e-learning content, delivery and communication of information. The instructional methods used in e-learning are quite many and determine how the e-learning process will occur. Instructional methods such as use of the higher educational learning object concept during content design, use of discussion boards for disseminating specific information amongst students (contribution is a must), and design of assessments (multiple choice, short answer questions, essay question) are very crucial. We should note that not all designed content is suitable for the e-learning process and may vary depending on the type of learners (life long, mature or others).

3.1. Infrastructure

These are the physical or no physical items that enable the e-learning process to occur. The infrastructure can be ICT related or not but should be able to support the e-learning project that the organization is embarking on. The infrastructure may contain the following:

- **Hardware:** these are the physical ICT components that are integrated together in order to offer an effective e-learning process. Some of the current hardware is designed specifically for multi-media processing which includes e-learning. E-learning that involves video conferencing may need specialized ICT as compared to a normal desktop computer. Therefore hardware has to critically be considered during the implementation of the e-learning process.
- **Software:** these are the types of software that can be selected for use in the e-learning process. There exists several e-learning environment software on market which offer different functionality. Some of the e-learning environment software are free to use and others are for rentable. Selection of e-learning environment software for use will depend on organizational policy and probably the funding available. Examples of the e-learning environment software include Blackboard, WebCT and Chisimba.
- **Networks:** this is the connection carried out on the different hardware so that they are linked to each other for the purpose of sharing information during the e-learning process. Different networks can be designed for the e-learning purpose (local or wide area network). The university can decide that the e-learning environment is only accessible within its premises (local network) or ubiquitously (wide area network). Such policies have to be debated before the initialization of the e-learning process.

- **Maintenance or Help desk:** this is the centre where stakeholders can seek advice related to e-learning process. This is very essential because of the limited direct contacts between the students and the tutors. If the students require anything related to their e-learning process, there is always some one to help. For example if the network is off, lost password, labs being occupied; stakeholders would need clarification from whoever is concerned. The help can undertake this task of linking the stakeholders with other administrative units.

4. A Case Study of Implementation of the Generic E-learning Framework in Cit

In order to illustrate the proper working of the framework, CIT was used as a case study for implementing blended learning. Initially CIT established that the University had an ICT policy that required all faculties to implement within their teaching and learning (Makerere ICT Policy). Therefore CIT's determination to implement a blended learning process was in line with the University ICT Policy. CIT preferred to implement the blended learning process on diploma courses that were already being offered in the faculty. This decision was in line with the accreditation policy of the university hence no violation was expected to occur.

Having established that the appropriate policies were not being violated, CIT undertook the next step to find out if there was funding for the implementation of blended learning. It was found that CIT's administration was in full of support of the implementation and had set up a fund to cater for expenses incurred. The fund would be used for training staff in pedagogical skills and buying more multi-media equipment. It was also noted that the University was renting an e-learning environment (Blackboard) that could be used for the e-learning process. Therefore no expenses were incurred on renting the e-learning environment. However for sustainability CIT recommended that we should adopt a free e-learning environment for the future e-learning implementation. There was no direct funding from the students to pay for the access to the e-learning content

The stake holders identified within CIT were the tutors who teach the diploma courses, the students on the diploma courses, e-learning expert, departmental administrators and head. These stakeholders were initially sensitized on the benefits of implementing blended learning. The students were informed of the new delivery method that required them to spend one hour in a physical class and the other two hours studying ubiquitously per week through the e-learning environment. Students were made aware that all the time they spend on the e-learning environment was being tracked and recorded for future reference. This encouraged them to better utilize the available two hours per week of self study. The tutors were offered a pedagogical training course and requested to utilize the skills learnt to develop courseware for the diploma courses they were teaching. This was done appropriately and the courseware reviewed as per the selected instructional design concept. The courseware was uploaded on the e-learning environment and was made available to the students.

CIT had an establishment of adequate computer infrastructure that could easily be used for offering an effective e-learning process. The infrastructure included several computer laboratories fully connected to the internet, software necessary for the e-learning and a department for systems administration responsible for help desk roles.

The content utilized during the e-learning process was designed to suite the higher education learning object concept. Figure 2 and 3 illustrate the diploma courses developed and details of one course fundamentals of software systems design respectively. It was noted that the students were excited to utilize technology for self study through an e-learning environment. The students were encouraged to utilize other tools within the e-learning environment for self study, communication and assessment. It was further observed that students who missed the classes could easily access the content they missed in class hence failing to lag within their studies. Students were also found to develop maturity in being responsible for their learning by constantly requesting for help in several scenarios for example being unable to access the e-learning environment outside the campus.

Despite having access to the content ubiquitously, some students did not take the initiative to undertake self study. It was observed from the course statistics that some students did not bother accessing the online content. However the students had a 24 hour access to the computer laboratories for study time. There was no excuse for the students who did not access the content other than being simply lazy. Therefore after one month of use, the students were notified that it was a must for every one to utilize the e-learning content. Whoever did not access the content would loose marks on their final score. This encouraged the lazy students to access the content.

Fig. 2: Diploma Courses Developed for E-learning

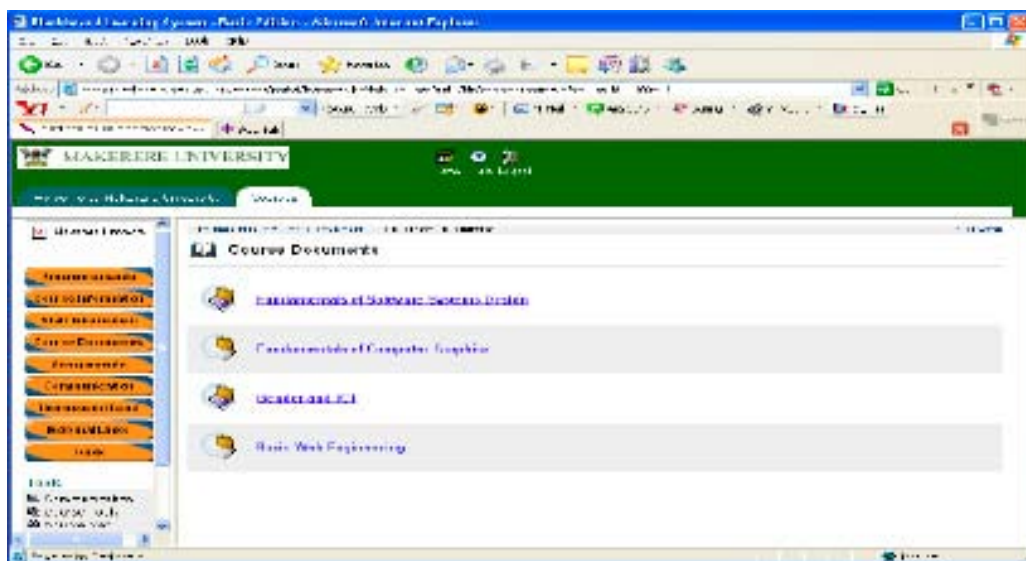
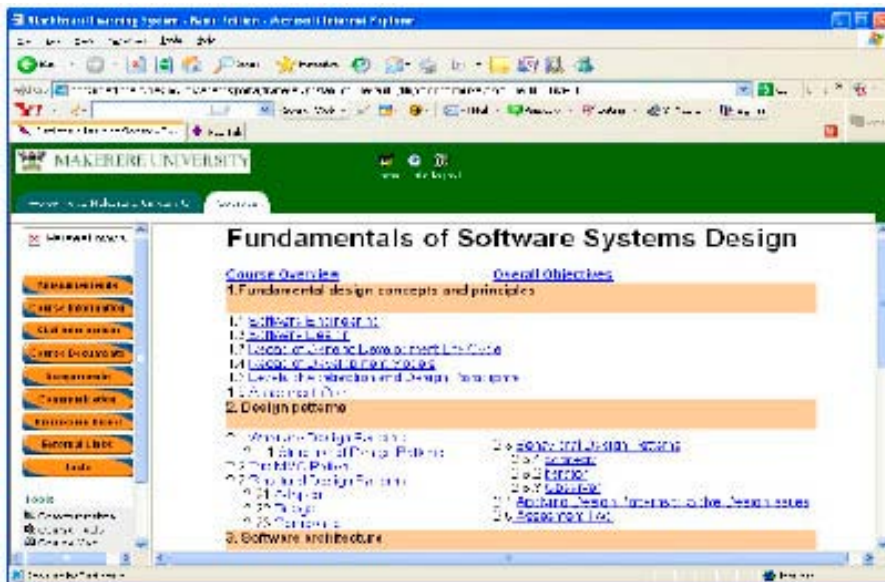


Fig. 3: Courseware Developed for Fundamentals of Software Systems

5. Conclusion

A generic e-learning framework has been described that was utilized by CIT in the implementation of blended learning for its diploma courses. The framework contains different components that are useful during the implementation of e-learning within an organization. The components included organizational policies, funding, stakeholders, services and infrastructure. Each component contains other constituents that are necessary for effective implementation of learning.

The framework was partially implemented in CIT's diploma courses as a pilot study for blended learning. It was noted that the service component was very important because it contained the pedagogical skills training for tutors and students. Without a well developed courseware, effective knowledge construction can hardly occur during e-learning. It was important to train tutors on how to develop content based on the higher educational learning object concept. The results showed that students were excited on using technology for their self study in the absence of the tutors. By the end of the course students felt that they had gained maturity in being responsible for their learning process. It was also observed that there is resistance in the undertaking a self study by the students but the situation can be changed through motivation and enforcement. Several students were initially adamant on utilizing technology for self study but their perception was changed through enforcing compulsory access to the content. CIT has plans to implement blended learning in all its undergraduate courses based on the developed generic e-learning framework.

References

- ADELMAN, H.S. & TAYLOR, L. 2000. Looking at school health and school reform policy through the lens of addressing barriers to learning. *Children's Services: Social Policy, Research, and Practice*, 3, 117-132.
- CISCO SYSTEMS, 2001. Reusable Learning Object Strategy – Designing Information and Learning Objects through Concept, Fact, Procedure, Process, and Principle Templates, *Cisco Systems Inc*, http://business.cisco.com/servletwl3/FileDownloader/iqprd/86575/86575_kbns.pdf
- CRAMER, S., KRASINKI, S., CRUTCHFIELD, M. D., SACKMARY, B., & SCALIA, L. 2000. Using collaboration and the web to implement the new CEC standards. *Teaching Exceptional Children*, 32(5), 12-19.
- FOSNOT, C. 1996. Constructivism: *A Psychological theory of learning*. In C. FOSNOT (Ed.): *Constructivism: Theory, perspectives, and practice*. New York: Teachers College Press, 8-33.
- HONEBEIN, P.C., DUFFY T. AND FISHMAN B. 1993. Constructivism and the Design of Learning Environment: Context and Authentic Activities for Learning, in T.M. DUFFY, J. LOWYCK AND D. JONASSEN, Eds. *Design Environments for Constructivist Learning*, Springer-Verlag, NY, 87-108.
- IEEE, 2003. Standard for Learning Object Metadata, *Learning Technology Standards Committee (LTSC)*, retrieved on 29/10/07 from <http://grouper.ieee.org/LTSC/wg12/>
- IMS, [2003]. Open Specifications for Interoperable Learning Technology, retrieved on 29/10/07 from <http://www.imsglobal.org/>
- JENNIFER DV. 2005. E-learning strategy: *A Framework for Success*. Learning Circuits. Accessed on 10/01/2008 from <http://www.learningcircuits.org/2005/aug2005/devries.htm>
- KHAN, B. 2001. *A framework for e-learning*. *LTI Newslines: Learning and Training Innovations*, retrieved on 29/10/07 from <http://www.elearningmag.com/ltimagazine/article/articleDetail.jsp?id=5163>
- MOHD, H.H. 2001. Integrating e-Learning into the workplace. *The Internet and Higher Education*, 4(3-4), pp.301-310.
- MOOIJ, T. 2007. Design of educational and ICT conditions to integrate differences in learning: Contextual learning theory and a first transformation step in early education. *Computers in Human Behavior*, 23(3), pp. 1499-1530.
- ROSSETT, A., Ed. 2002. *The ASTD E-Learning Handbook*. New York: McGraw-Hill.
- SCORM. 2003. Best Practices Guide for Content Developers, *Learning Systems Architecture Lad*, <http://www.lsal.cmu.edu/lsal/expertise/projects/developersguide/developersguide/>
- STEFFE, L.P., & GALE, J., Eds. 1995. *Constructivism in Education*. HILLSDALE, NJ: ERLBAUM,
- SUN, L. AND WILLIAMS S. 2004. An Instructional Design Model for Constructivist Learning, Eds. *Media 2004 World Conference on Multimedia, Hypermedia and Telecommunication*, Lugano, Switzerland, ISBN: 1-880094-53-3, 2476-2483.
- WENTLING, T.L., WAIGHT, C., GALLAHER, J., LA FLEUR, J., WANG, & C., KANFER, A. 2000. E-learning-A review of literature.
- WILEY, D.A. 2003. *The Instructional Use of Learning Objects*, Association for Instructional Technology, <http://www.reusability.org/read/>
- WORLD BANK. 2002. *Enhancing learning Opportunities in Africa*, in Working paper series of Africa Region Human Development

Integrating E-Supervision in Higher Educational Learning

Jude T. Lubega and Michael Niyitegeka

Technology has contributed a lot of improvements in pedagogical aspects of education. To many educators it is believed that technology will some day control the entire activities within education. This approach which is referred to as e-learning is transforming how education is being delivered. However the e-learning approach depends on the proper integration of technology, with the existing traditional means. The technology enabled approach should be positioned to enhance and not replace the existing traditional mean. There are a lot of ways through which technology has been adopted to enhance learning and these could include content dissemination, interaction, assessment, evaluation, communication and supervision. In this paper e-supervision is discussed as one method that technology is enhancing within higher educational institutions of learning. The paper discusses a pedagogical model for e-supervision that is facilitated by the available technology. This model indicates that there are several methods that are being adopted to enhance the traditional supervision. The methods include use of e-mails, discussion boards, forums, telephony, chat rooms, wiki, blogs and e-research group. The methods can be effective in enhancing supervision but would need a strong foundation in setting up a technological infrastructure, social atmosphere, communication, solidarity, time schedule, prompt of response and respect for members.

1. Introduction

There has been great emphasis on research within higher institutions of learning for all graduate programs. The quality of research from graduate students is greatly attributed to both the supervisor's and supervisee's efforts to do all activities that are offered in the specific research being undertaken. In many developed countries research has been considered as the driving force for the economy and therefore taken to be an essential component for all. There are several ways through which research can be supervised using traditional means; however technology has introduced other forms of supervision. These new forms of supervision methods are currently practiced informally by different supervisors. Supervision is a form of mentorship that develops within some one after a few times of practice [Persson and Brew, 2002]. Therefore it is a result of constant giving of constructive guidance to some one being guided and there after measure the outcome. Research has indicated that supervision is one of the major influences on the research student outcomes [Seagram et. al., 1998; Latona and Browne, 2001]. This is because it directly contributes to what students do in relation to how they have been guided. Many times the supervisors are assumed to be knowledgeable in the specific research areas hence being used as reference points by the students.

The supervisory role helps to serve as a reflective practice through which the supervisee can question and modify their activities. Therefore the supervisor has to present a positive experience which question and lead to a modification of the supervisee's activities. Both the supervisor and supervisee's experience can be undertaken through traditional and modern (technological) means. Today technology has encouraged both the supervisor and supervisee to keep in touch, manage their activities, operate effectively and share experience despite the great barriers that may exist. However before the experience can take place and be shared, both the supervisor and supervisee have to establish a personal relationship which forms the foundation for the experience [Nelson, *et. al.*, 2001; Pearson, 2000]. With the founded relationship a teaching-learning alliance is easily created between the stakeholders and it is this relationship that contracts them to do their respective roles. It is the role of the supervisor to make sure that the supervision bond is created irrespective of the existing circumstances. This bond can be further facilitated by the constant open communication between the supervisor and supervisee. For an effective supervision process, supervisor and supervisee need to agree on the activities schedule and goals. When such an important step has been taken, there is less ambiguity surrounding the supervision process [Nelson *and* Friedlander, 2001]. For any effective supervision process, key factors have to be considered. A multi model for pedagogical supervision is proposed for effective supervision through technology.

2. Supervisory Process

Supervision process involves several people undertaking different tasks together at different levels to accomplish a specific goal and in guidance of each other. The supervision process involves usually the client (research funding body), the supervisee and the supervisor. These people have different roles they undertake during the supervision process which directly or indirectly affect the final goal. It is been noted by Storm and Todd [1997] that supervision will involve the supervisors safe guarding the welfare of their clients, mentoring supervisees and protecting their professional practice. Therefore the tasks undertaken by the supervisor should not conflict to hinder any of the three stakeholders. In so many scenarios supervision is initiated with the consent of the supervisor who requires commitment from the supervisee and supervision process. This could be due to the expertise possessed, supervisory model undertaken, competence to supervisee, availability of the supervisor, honesty and integrity possessed by the consenting persons. .

During the supervision process, personal relationship between the supervisee and supervisor creates the foundation on which the supervision process can be undertaken [Ladny *et. al.*, 2001; Pearson, 2001; Watkins, 1995]. When the relationship fails during the supervision process it is always advisable to change supervisors in order to protect research interests. Initiating this relationship should be done by the supervisor who clearly outlines what this bond will hold and its

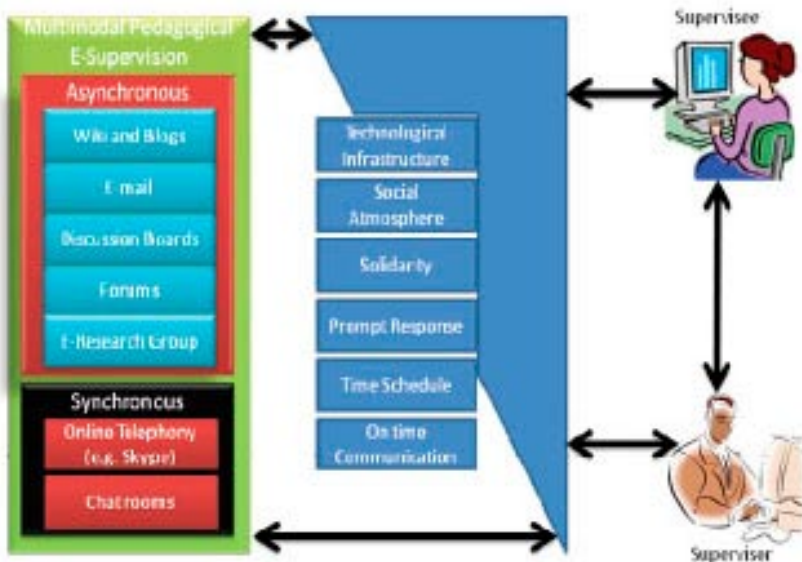
boundaries [Pearson, 2001; Bernard and Goodyear, 1998]. When the supervisee notices these efforts, they take the initiative to abide with the guiding principles of the supervision process. The supervisee will always feel comfortable when their learning interests are catered for irrespective of the existing circumstances. Despite the efforts from both supervisor and supervisee to undertake their tasks more diligently, both have always fallen short of their expectations. Carlason and Erickson [1999] note that both supervisor and supervisee need to reflect on what they have discussed on attaining the research goals. During supervision the norm of saying that you should do what I say but not what I think is not right in attaining the stipulated research goals. All stakeholders need to be offered an opportunity to discuss the requirements for the research.

A more prosperous supervision process will even involve the supervisor deepening the relationship by encouraging a friendship form of relationship [Bernard and Goodyear, 1998] that contains genuineness, respect and empathy. Such a relationship will motivate both supervisor and supervisee to carry out their tasks without any fear for each other. Many times as the supervisor gets to know about the supervisee's characters, attitudes and the desire to be supervised, the relationship blossoms and the working pace increases drastically. Open and free communication is so vital for such a process that involves two people carrying out tasks at different levels. The communication should always aim at reaching a consensus on the research objectives, goals and tasks [Vespia *et. al.*, 2002; Nelson *et. al.*, 2001]. When both supervisee and supervisor develop the attitude of trust between each other, they tend to accept each other's experiences and the desire to learn from each other. This creates a deeper feeling for the supervisee that they can be helped on areas where they have challenges, vulnerable and need more understanding. Having an attitude of openness to the new knowledge being offered by the supervisor creates an avenue for effective supervision. Irrespective of human factors such as gender, age, nationalities both supervisor and supervisee need to understand their roles during supervision.

The quality of postgraduate research supervision has been questioned due to the failure by both supervisors and supervisee to undertake their roles explicitly. Research results have been questioned, discarded, work plagiarised and probably the entire process of supervision mismanaged. This has encouraged many of the higher degrees institutions to occasionally undertake research supervision audits in order to maintain the quality of the research outputs. With the advancement in technology, the roles for the supervisor and supervisee during supervision have drastically changed. Certain roles can now be shared by both stakeholders hence making the entire supervision process dynamic.

Supervision today can be undertaken through two major ways; traditional face to face and e-supervision. In many respects today both supervision ways have been utilised to increase on the chances of providing effective guidance. However e-supervision has been increasingly adopted within higher institutions of learning for several purposes. It is been adopted due to the increasing large

numbers of students, increased number of distant collaborations, availability of technology and appreciation of technology by the educational society. In order to have e-supervision effectively adopted, there is need of full commitment from all stakeholders concerned. This commitment can only be offered when the stakeholders are aware of what needs to be undertaken. In many occasions



e-supervision is carried out unknowingly hence not receiving the expected and required commitment. It is from this view that a model for pedagogical integration of e-supervision is proposed for higher educational institutions.

3. A Multi-model for Pedagogical Integration of E-supervision

The Multi-Model for Pedagogical E-Supervision (MMPES) has several components that are interlinked to each other to achieve the goal of electronic supervision as described in figure 1. The model contains three main components, the users (supervisor and supervisee), prior factors and e-supervision methods (synchronous and asynchronous).

3.1. Users

The *users* within the MMPES include supervisor and supervisee who are continuously interacting with each other in several forms. These two types of people initiate a form of agreed upon interaction that they use during the supervision process. The *supervisor* has the role of initiating most of the interaction methods that could be used during the supervision process. Many times the supervisor contacts the supervisee with suggestions on how the supervision process could be carried out. The supervisor has the role of formulating and maintaining the supervisory bond, creating an orientation to the supervision process, resolve conflicts and support the supervisee in any way. The *supervisee* has the role of making sure that the interactions within the supervision process are effective at all times, learning

without resistance, create Collaboration Avenue, monitor progress and publish research findings. The supervisee oversees the supervision schedule so that they take more responsibility of the process. With the different roles undertaken supervision can be enhanced through the existing prior factors.

3.2. Prior Factors

Prior factors are the requirements that need to be in place for e-supervision interaction to occur in an effective format. In order to have an effective e-supervision there is need for prior planning that involves setting up a technological infrastructure, social atmosphere, communication, solidarity, time schedule, prompt of response and respect for members.

Technological infrastructure: refers to the hardware and software that are necessary for the e-supervision to occur. Before e-supervision can take place both the supervisor and supervisee should have access to a computer fully connected to the internet and containing an appropriate operating system such Microsoft Windows. Sometimes all stakeholders may utilize specific application software such as Microsoft Office, Adobe Editor for making additions, deletions to the documents.

Social Atmosphere: refers to the acceptance of use of technology for learning within the community where e-supervision is being introduced. In many societies use of technology to enhance learning has not yet been appreciated for fear of replacing the traditional teachers. Therefore for e-supervision to effectively occur the society should have good feelings about its contribution to education. Technology appreciation by the educational society can be demonstrated through its adequate adoption for both teaching and learning.

Solidarity: refers to the harmony that has been created between the supervisor and supervisee on the entire process of e-supervision. There is need of harmonizing the entire supervision activities, virtual meeting schedules, response format plus other guidelines. Without a shared vision on how to achieve the stipulated research goals, both supervisor and supervisee can hardly embark on this challenging process of supervision. Therefore it is very important for them to have a work plan that will be adhered to during e-supervision.

Prompt Response: refers to replying to your counter part in the shortest time possible when contacted. Many times communication between people at a distance fails due to late replies. During e-supervision communication between the supervisor and supervisee has to be immediate to avoid being misunderstood. When communication is not done on time, participants loose trust in the entire process.

Time Schedule: refers to the routine like forms of planned supervision meeting that are agreed upon by both supervisor and supervisee. Before e-supervision can be effectively undertaken by both the supervisor and supervisee, the need to plan when and at what time they will meet for the supervision is a critical success factor. If any of the members abuses this time schedule or violates its existence, the impactions are loss of trust and commitment

3.3. E-supervision Methods

There are several methods that have been identified to help in the e-supervision process and can be categorised under synchronous and asynchronous.

Synchronous category includes all methods where both the supervisor and supervisee interact in real-time during the e-supervision. There is immediate contact and response between the supervisor and supervisee. This category has methods such as chat room and online telephone (e.g. sky pee). *Chatroom* are electronic tools for real-time communication where by two or more people interactively write to each other and receive immediate response. These tools have restrictive content editors that allow a specific number of lines to be typed at a time and never store the content after the chat. *Online Telephone* is an online communication that involves two participating members talking to each other through online tools such as sky pee. This is a cheap form of communication through the internet and requires participating members to share user ids for identification.

Asynchronous category is one that involves the supervisor and supervisee communicating to each other and do not receive immediate reply. Such communication faces challenges of slowness to respond, ignoring of communication as well as members being impatient. This method has categories such as e-mails, wikis/blogs, discussion boards, forums and e-research groups. The aim of this method is for the members to contribute to the e-supervision and receive a reply at a later stage. This method is very common today and is thought to be effective especially when both supervisor and supervisee are distant apart.

A blend of the components within the model contributes to a successful e-supervision process and can effectively influence research performance. With many students today undertaking ICT related modules, it is quite easy to utilize the form of e-supervision. However the educational level at which e-supervision should be administered needs further research before a conclusion can be made. Although the youth utilize synchronous communications as a tool for their specific activities, does it indicate its significance in transmitting knowledge to them?

4. Motivation for the Mppes

This model emanates from a case scenario at the Faculty of Computing and IT, Makerere University. The faculty has seen tremendous growth in terms of graduate students and this implies increased demand on the part of the supervisors.

The categories of students who seek graduate studies in most cases are working. This implies that they are engaged through out the greater part of the day undertaking their normal duties and will tend to be free after the normal working hours. This however implies that the supervisor has to remain at work for extra hours which has a lot of demands not only on the supervisor but on the educational system as a whole. Majority of the supervisors are teaching on the graduate programmes and hence most of the supervisory time is spent in class. This scenario implies that the research students may not have adequate supervision if they are to use traditional face-to-face supervision. Unless alternatives are made

possible it means that majority of the research student may not progress through the research stage within the required time.

The faculty of Computing and IT starting October 2007 implemented this model partially through encouraging both the supervisor and supervisee to use the existing technologies. As a pilot, students with similar research area like E-services were grouped together and were required to form a group mail of which the supervisor is a member by default. The students were encouraged to submit their work to the group mail and fellow students comment on their colleagues work. Because they were in the same research area there is sharing of information especially literature. The supervisor watches the comments or the guidance being given and will comment to clarify, refocus the research, and provide the general guidance. This has definitely yielded results in that students learn from each other and also support each other to complete within the specified time. To a large extent the e-mail approach has been predominantly used. Only in a limited number of instances the chat facilities has been used for remediation/feedback from both supervisor and supervisee. It is important to note that this was a blended approach which had both the face-to-face and e-supervision. The group was expected to meet the supervisor once a week at an agreed time and the rest of the interaction was done through e-mail.

There were over-riding assumptions with this approach; each of the group members will be expected to contribute when ever there is a submission or else the team spirit will fade, access to the internet and e-mail are assumed to be possible to ensure timely feedback by colleagues.

In general there was appreciation from the group of students who participated in this pilot model testing. This is because it provided the students with alternatives and enable self mentorship which was a new learning opportunity. It also enabled group effort to be realized. As a result this particular group has thoroughly progressed in their research.

5. Conclusion and Future Work

With the current trend that indicates higher number of graduate students enrolling for postgraduate education, institutions are faced with a challenge of educating them (teaching and supervising). Technology has contributed to several opportunities for teaching and learning which include e-supervision. The multi model for pedagogical e-supervision illustrates that there exists several electronic methods that can be adopted when supervising graduate students. It is assumed that many of these graduate students are mature and can satisfy the prior requirements for this form of supervision. With this form of electronic supervision, supervisors and supervisee are able to develop a strong attachment to each other through the constant collaboration and communication. Technology tends to create a neutral platform for all stakeholders involved in the supervision process. The supervisee and supervisor learn how to respect, appreciate and share knowledge between each other as a sign of commitment to the supervision process.

References

- BERNARD, J. AND GOODYEAR, R. 1998. *Fundamentals of Clinical Supervision* (2nd ed). Needham Heights: Allyn & Bacon.
- CAMPBELL, J. 2000. *Becoming an effective supervisor: A workbook for counselors and psychotherapists*. Philadelphia: Taylor & Francis.
- LATONA, K. AND BROWNE, M. 2001. Factors Associated with Completion of Research Higher Degrees. Australian Government Publishing Services.
- NELSON, M.L., & FRIEDLANDER, M.L. 2001. A close look at conflictual supervisory relationships: The trainee's perspective. *Journal of Counseling Psychology*, 48, 384-395.
- NELSON, M.L., FRIEDLANDER, M.L., GRAY, L.A., LADANY, N., & WALKER, J.A. 2001. Towards relationship-centered supervision: *Journal of Counseling Psychology*, 48, 407- 409.
- PEARSON, B. & PIAZZA, N. 1997. Classification of dual relationships in the helping professions. *Counselor Education and Supervision*, 37, 89-99.
- PEARSON, Q. (2001). A case in clinical supervision: A framework for putting theory into practice. *Journal of Mental Health Counselling*, 23, 1-7.
- PEARSON, M. AND BREW, A. 2002. Research Training and Supervision Development. *Studies in Higher Education*, 27 (1), 135-150.
- SEAGRAM, B. C., GOULD, J. AND PYKE, S. W. 1998. An Investigation of Gender and other variables on time to Completion of Doctoral Degrees. *Research in Higher Education*, 39 (3), 319-335.
- VESPIA, K. M., HECKMAN-STONE, C. & DELWORTH, U. 2002. Describing and facilitating effective supervision behavior in counseling trainees. *Psychotherapy: Theory, Research, Practice, Training*, 39, 56-65.
- WATKINS, C.E. 1995. Psychotherapy supervision in the 1990s: Some observations and reflections. *American Journal of Psychotherapy*, 4, 568-581.

31

Mobile Research Supervision Initiative (MRSI) at Makerere University: Lessons to Learn

Paul B. Muyinda, Kathy Lynch and Jude T. Lubega

The rapid technological development in mobile phones coupled with their rapid diffusion into all walks of life has instigated various sectors to exploit them for various purposes aimed at enhancing organizational efficiency and flexibility. The communication, business, financial, banking and education sectors have developed and continue to develop applications for mobile phones. Our research sought to evaluate an initiative introduced at Makerere University's Department of Distance Education (Mobile Research Supervision Initiative – MRSI), for guiding (using mobile phones) distance learners who were completing their final year field research projects. The study adopted a mixed methods research approach, employing a self administered semi-structured questionnaire/survey, in-depth interviews and document review methods for data collection. Results from the evaluation have indicated that the mobile research supervision: i) created a virtual community of practice amongst research students and their supervisors, ii) motivated lonely distance learners in the field, iii) created a customer care feeling amongst students, and iv) bred intimacy amongst the alumni and their institution. The results of this study could be used to inform the development of mobile learning policy and pedagogy for developing countries. **Keywords:** Distance learning, m-learning, e-learning, mobile research supervision

1. Introduction

Mobile phones' diffusion, capability and portability are making them the number one companion for the human race. Other than their orthodox purpose of communication, mobile phones are also being used in the entertainment, business and education sectors. The education sector is using them to extend learning support to students in what is now called mobile learning (m-learning). Many institutions are implementing various educational applications for mobile phones. At the Department of Distance Education, Makerere University, mobile phones are being used to extend support (academic and administrative) to the distance learners. In particular, the Department of Distance Education introduced the Mobile Research Supervision Initiative (MRSI) on 1st August 2005, to support its students completing their final year field research project. Most of the support is embedded in SMS messages. SMS messages are sent out to research students to guide them in various aspects of research including: field data collection, making face to face meeting appointments with supervisors, providing pointers to useful

literature sources, motivating the 'lonely' student in the field, pacing the students, encouraging students to collaborate amongst themselves, informing students of deadlines and providing a two way collaboration mechanism between the students and their supervisors. The mobile phone therefore has made the supervisor available to the research students at anytime in any place. This study sought to evaluate the extent to which the Department of Distance Education and the students are benefiting from the MRSI and draw lessons for improvement.

The Department of Distance Education was established at Makerere University in 1991. Its programmes cater for students who study through the distance education mode. The Department's programmes are run on a collaborative venture with other academic faculties in the University which run conventional/internal programmes [Aguti and Fraser 2006]. Hence Makerere University is a dual mode University. The Department of Distance Education carries out student support activities to students scattered in many parts of Uganda mainly via the radio [Kajumbula 2006]. As a mass media, the radio cannot provide synchronous or asynchronous one to one customized communication. Hence the Department introduced mobile learning to ensure one to one student to staff, staff to student, student to student and staff to staff support. The MRSI aims at enhancing collaboration among research students and their supervisors. Mobile research supervision involves the supervisor guiding the research student (using a mobile phone) as he/she carries out his/her final year field research project. It ensures that guidance is provided to the student at anytime in any place in the field.

2. Genesis of the Mobile Research Supervision Initiative (MRSI)

As a partial fulfilment for the requirements for the award of a bachelors degree from Makerere University, undergraduate students in their final year of study are required to undertake an independent field study in an area contributing to knowledge in any of the Department's specified research themes. As is the practice everywhere, each research student is allocated a supervisor, who is supposed to collaborate (through face to face or otherwise) with the research student on a regular basis. However, the face-to-face collaboration between the supervisor and the student was found to be inadequate. Often, research students were frustrated by the 'non-availability' of their supervisors when they came to meet them at the main campus, meaning that the students' face to face interactions with their supervisors were only limited to a few occasions. In some instances, the research paper unnecessarily prolonged the students' stay at the University beyond their programme's stipulated minimum duration. In other instances, the research paper proved to be too 'difficult' for some students to comprehend hence leading to failure in completing their programmes. These and other factors have been argued as the reason for poor throughput (60%) of distance learners at Makerere University [Otto and Wrightson 2005]. According to a number of research students, the perceived 'difficulty' in the research course was attributed to lack of adequate research supervision.

In a bid to make the research process a rewarding and exciting experience for the students and increase the supervisors' availability (physical and virtual) to their research students, the Department of Distance Education encouraged supervisors to integrate readily accessible ICTs into the research supervision process. The mobile phone has been highlighted in particular; because an earlier study [Kajumbula 2006] had revealed that 97% of distance learners at Makerere University had access to it. The initiative was named the Mobile Research Supervision Initiative (MRSI).

It was hoped that the additional voice and text collaboration on the mobile phone would help minimise the perceived negative impact of lack of face-to-face interaction and encourage the development of 'a personal touch' between the research students and their supervisors. This initiative, which falls in the wider context of m-learning, began in the 2005/06 academic year. Since then, the Department had not taken stock of its effectiveness and had not established the students' perception to it. We evaluated the initiative to establish its outcomes to the Department of Distance Education and students. An evaluation is "the systematic investigation of the worth or merit of some object" [Commonwealth of Learning 2007, p.37]. Evaluations are carried out to improve evidence-based decision making. This evaluation project was aimed at producing evidence to guide open, distance and e-learning (ODEL) practitioners and policy makers in deciding whether or not to continue investing in mobile research supervision and other mobile learning initiatives. It answered the general question – 'Is it worthwhile to continue with mobile research supervision and should we invest more resources and effort into it or not?' Specifically, the study has answered the following questions:

- i. What research supervision aspects were most adequately handled through mobile research supervision?
- ii. What influence did mobile research supervision have on students' throughput?
- iii. What were the constraints faced by the research students during mobile research supervision?
- iv. What were the cost implications of mobile research supervision to the student, supervisor and Department?
- v. What other academic and administrative components could be delivered via mobile phones to distance learners at Makerere University?

3. Evaluation Framework

The evaluation study was underpinned by two major theories of learning, namely: the social constructivist learning theory [Brown and Campione 1996] and the conversational theory [Pask 1975]. The social constructivist learning theory required the research students to act and reflect within their social environment. The messages sent and calls made to the research students required them to act by solving a given research problem and reflect on the derived solution(s) with the aim of enriching their experiential knowledge. The conversation theory suggests that successful learning requires continuous two-way conversations and interactions between the teacher and learner and amongst the learners themselves [Pask 1975].

The MRSI embraced these two theories because text and voice messages sent to a supervisee's mobile phone could be saved for future action and reflection, while the learners could use their mobile phones to converse/collaborate with fellow research students or with their supervisors. The MRSI is an abstraction from m-learning which in turn is a subset of e-learning.

M-learning is attracting considerable research among the academia and industry. While evaluating the effectiveness of SMS communication to distance learners at Makerere University, Kajumbula [2006, p.6] concluded thus, "... students' access to mobile phone technology is very high and therefore the mobile phone presents a very attractive option to easing communication between the students and the Department". This conclusion implied the need to explore the mobile phone as an ICT for delivering a multitude of student support services. Fagerberg and Rekkedal [2004] see the mobile phone as a tool which increases flexibility in distance learning. With its collaborative features and portability, the mobile phone is capable of delivering learning at any time in any place.

However, simply jumping onto the bandwagon of technological innovations is not a panacea to improving student support. Different features of the technology must be understood by all the stakeholders before deploying the technology for learning or student support [Keegan 2005]. Pedagogical issues must be born in mind [Muyinda 2007]. Costs of the technological innovation have to be evaluated vis-à-vis the anticipated benefits [Graham, 2004]. The limitations of the technology must also be known in advance and mitigated. For instance, while considering delivering content on mobile phones with small screens and low memory, Zhang [2003] proposed to use Wireless Access Protocol (WAP) gateways, agent profilers and caching proxies to deliver content to users of the mobile phone in a combination of push and pull mechanisms. Some researchers [Sharples *et al.* 2005] are actually wondering whether the use of mobile phones in mobile learning is any different from the conventional e-learning. In other words, they see no difference between mobile learning and e-learning. To them, what is important is the careful integration of a new learning technology and not its hypnotization. Based on their view, one would be given a license to simply integrate mobile phones into the research supervision process without any due consideration. This would be unacceptable to Muyinda [2007] and Graham [2004] who prefer a careful and considered integration of new technology into the learning process.

4. Evaluation Methodology

A mixed methods approach employing quantitative and qualitative paradigms was employed. Distance learners who had completed their research project (herein referred to as research students), research project supervisors (herein referred to as supervisors) and administrators in the Department of Distance Education at Makerere University (hereafter referred to as administrators) formed the target population of the evaluation study. Using simple random sampling, 100 research

students were selected from the Research Report Submission Register found in the Research Office at the Department of Distance Education. Phone calls were made to the respondents to recruit the research students for their voluntarily participation in the study. The phone calls to respondents also helped us to determine the most convenient mode of delivering the research instrument. Of the 100 research students selected, 86 agreed to participate in the research and provided postal mail addresses or physical locations. However 71 responses were returned representing a response rate of 84%. A semi-structured self-administered questionnaire was mailed or delivered physically to the research students. The questionnaire asked standard sets of questions including mobile phone ownership, mobile phone connectivity issues and use of mobile phone in research supervision. Such questions provided answers to questions from the measurement and understanding perspectives. Learners were given freedom, in the instrument, to express their opinions and views about the conduct and experience of mobile research supervision. Seventy-one responses were returned.

In order to obtain views, opinions, experiences and understanding about the MRSI, in-depth interviews were conducted with 5 supervisors and 2 administrators. These were purposively chosen as key informants. Since the in-depth interviews were semi-structured, we were able to obtain detailed answers to non-standard questions that were posed.

The data gathered was analyzed using descriptive statistics, and are presented in tables and text forms in the following section.

5. Results

5.1. Mobility Characteristics of Research Students

5.1.1. Employment Status of Research Students

One of the objectives of introducing the distance education learning mode at Makerere University was to extend university education to adults who found it difficult to leave their jobs or families to attend fulltime university education. We sought to establish the employment status of the research students and their distribution by region. The regional distribution variable was necessary for establishing the proximity of the research students to Makerere University's main campus. Employment status would provide a lead into the ability of the research students to pay for their own mobile communications costs. Results are presented in Table 1.

Table 1. Research students' region of aboard and employment status

Region	Education Sector (Freq)	Business/ Financial Sector (Freq)	Self Employed (Freq)	Unemployed (Freq)	Total (Freq)
Western	1	2	1	2	6
Central	13	8	3	20	44
Eastern	3	3	0	5	11
Northern	2	0	0	8	10
Total	19	13	4	35	71

Source: Primary data

From Table 1, it is evident that 35 of the 71 (or 50%) of the research students were unemployed. Sixty seven percent (or 24 out of 36) of those that were employed were working in the central part of Uganda with the northern part of the country contributing only 2 out of 36 working research student. The majority (44 out of 71 or 62%) of the research students were from the central region of Uganda. This region is in close proximity to Makerere University's main campus; therefore, the research students could easily travel to the main campus to physically meet their supervisors. Of those who were employed, 53% (19 out of 36) were employed in the education sector, 36% (13 out of 36) in the business/financial sector, and 11% (4 out of 36) were self employed. It is not surprising that the education and financial sectors featured most because respondents were drawn from the Department of Distance Education which offers the Bachelor of Education (external) and Bachelor of Commerce (external) degree programmes.

An assumption is made that the 50% unemployed research students, may not be able to afford mobile phones and their associated costs. This assumption was thwarted by the pervasiveness of the mobile phones among research students as presented in Section 5.1.2 below.

5.1.2. Mobile Phone Pervasiveness

Here, we present findings related to mobile phone ownership, research collaboration and connectivity issues. These findings are important for bringing out the plausibility of mobile research supervision among distance learners in Uganda.

Mobile Phone Ownership

Respondents were asked to indicate whether they had a personal mobile phone or not in order to determine the proliferation of this device among research students. The results are indicated in Table 2 below and in the text that follow.

Table 2 Mobile phone ownership

Mobile Phone Ownership	Frequency	Percent
Owned a Mobile Phone	68	96
Did Not own Mobile Phone	3	4
Total	71	100

Source: Field data

The data presented in Table 2 contradicts the assumption made previously in regard to unemployed research students not being able to own and maintain a mobile phone, as 96% of the research students, regardless of employment status, owned mobile phones. Also all supervisors and administrators interviewed owned a mobile phone. This indicates a high penetration of the mobile phone among the students, supervisors and administrators.

Mobile Phone Research Collaboration

A follow up question asked research students who had a personal mobile phone to indicate whether they had used it to collaborate with their supervisors and amongst themselves during the research process. The data collected indicates that the mobile phone enabled 93% (n=63) of the research students to collaborate with their colleagues and supervisors on various aspect of research using their mobile phones while 7% (n=5) did not use their mobile phones for research collaboration.

Mobile Phone Connectivity Issues

We wanted to establish whether a research student could send/receive a text message or make/receive a call wherever he/she was, at anytime and whether he/she could afford the cost of mobile communication. Sixty nine percent (69%) of the research students never had any disruptions in network connectivity, 30% sometimes had disruptions while only 1% always had mobile network connectivity problems. The less disruption in mobile network connectivity could be attributed to the fact that the majority of respondents (67%) were located in the central region which has a high concentration of mobile telecommunication companies.

There were a significant number of research students (71%) who got disruptions in communication due to lack or insufficient airtime credit on their phones. Twenty nine percent (29%) never ran short of airtime credit. This implies that universities wishing to embrace m-learning for student support services will have to provide a budget for phone airtime credit or work out a mechanism for toll free communication.

Other phone connectivity issues considered related to phone battery life. Research students reported having no disruptions to communication arising from phone battery power. Phone batteries were charged using multiple sources of energy, including: the national power grid (UMEME - 97% - n = 67), generators (7% - n = 5), solar power (3% - n = 2) and cars (3% - n = 2). The batteries were

charged either at home (82% - $n = 60$), place of work (18% $n = 13$), commercial charging vendor (4% - $n = 3$), by a neighbor/friend (6% $n = 4$), or a combination of power sources (18% - 13). This indicates that the mobile phone could be used in rural (with no grid power) and urban (with grid power) areas.

5.2. Mobile Phones and Research Supervision: Research Aspects that Were Effectively Handled Through the MRSI

We sought to establish, the aspect of research supervision that were adequately handled through mobile phones. We began by establishing whether, on being assigned to a particular supervisor, a researcher sought that supervisor's mobile phone contact, the reason for seeking or not seeking the phone contact and mechanism of collaboration during the research process.

Ninety seven percent (97% - $n = 69$) of the research students sought to get their supervisors' phone contacts in order to: i) fix appointments for face to face discussions (96% - $n = 68$), seek guidance in the field and report progress (90% - $n = 64$) and iii) communicate with supervisors to create intimacy that could breed better academic working relationship (56% $n = 40$). By enabling the fixing of face to face meetings appointments with supervisors, the MRSI improved efficiency of arranging these meetings hence saving research student's time and money that could be spent in failed meetings.

By communicating with their supervisors, research students broke the phobia that usually affects them in student-teacher relationships. This enabled them to freely express problems and challenges encountered in the field. The mobile phone also reduced the 'loneliness' that is usually associated with distance learning.

Three percent (3%) of the research students did not seek phone contacts of their supervisors because they could physically get access to them at any time they wished. The small percentage of research students being physically able to meet their supervisors at any time is an indication of limited physical access to the supervisors. It confirms the students' outcry of limited face to face supervision. This calls for technologies that could increase supervisors' 'availability' to the research student. This evaluation research has revealed that the mobile phone is one of such technologies. After all, 90% ($n = 64$) of the research students were at liberty to call their supervisors, who in turn received their calls positively (87% - $n = 62$). In some instances, research students (7% - $n = 5$) reported that their supervisors had discouraged calls from them. When asked why they did not encourage calls from research students, one supervisor said, "... sometimes students call at awkward hours of the day. Besides, it is costly to the student". Supervisors felt it right and fitting for them to use University resources to reach out to the research students in the field. Fifty two percent (52% - $n = 37$) of the research students benefited from this generosity.

When asked to tick off only one option which was closest to the feeling experienced as soon as they received a call or text message from their supervisors or administrators in the Department, research students generally felt that they were cared for. Table 4 provides the different feelings.

Table 4. Single most feeling when a student received a call or text message from the supervisor/Department

Feeling	Frequency	Percent
Felt good and gratified	10	28
Felt sense of care by supervisor	6	18
Was encouraged/motivated to accomplish my research	6	18
Felt sense of responsibility by the department	6	18
Was excited and surprised	3	9
I was tensed up	3	9
Total	34	100

Source: Primary Data

Table 4 gives an indication that mobile research supervision, created a customer care feeling among the research students. Research students who were called or sent a text message expressed gratitude to the responsible and motivating supervisor/department action. Others were simply surprised because they never expected such a thing to happen. Consequently, the call/text message tensed up some students, who felt they were so insignificant to receive such a 'first class' treatment from their university or supervisor. Being excited, surprised and tensed up by the call/text message received from their supervisors is an indication of lack of prior knowledge about the MRSI on the part of the research students. The administrators in an interview confirmed that research students were not briefed about this mode of supervision.

When asked to write down some of the SMS messages received on their mobile phones, research students gave a number of messages some of which are reproduced below. It should be noted that the 150 character limitations of text messages dictated the use of abbreviations and symbols in some messages.

... come with Chapter one we shall meet at CCE Common Room at 2:00pm. Come for feedback from your work. We are meeting the supervisor at the Institute. The supervisor is not around. Please find out if supervisor is in office. How far have you reached? Can we meet at campus? We are meeting at the Institute at 10am – supervisor. Please come 4 feedback on sat 23 June 2007 at 9:00am. Deadline 4 handing in research reports is July 15 2007. Come for supervision meeting from 9-5pm. Have you typed the questionnaire, lets discuss them 2moro. Will not be available, meet 2gthor next week same time. Come for your research questionnaire. Research supervision meeting is on sat 17th feb 2007 starting at 9:00am-prepare to travel. Kindly confirm whether you have completed your research or not, I am updating my workload. Please hurry the supervisor is about to leave. .SoP is that problem whose answer calls for research....

Source: Message sent on research students' mobile phones

An analysis of the messages sent to research students reveals that they sought to: i) fix face to face research supervision meeting appointments, ii) motivate the 'lonely' research student in the field, iii) pace and motivate the research student, iv) provide a two way communication mechanism, v) establish and monitor research students' progress, vi) inform research students about deadlines and important dates, and vii) provide research students with some useful research tips.

In an asynchronous way, research students responded to the SMS messages or calls sent or made to them by their supervisors and fellow research students. Again we reproduce some of the messages in their raw form in order to emphasise the need to learn a new form of vocabulary for SMS messaging in m-learning.

.... I will be waiting 4 the instructions as u promised. I was ur rsch student and I am requestg u 2 B my referee. Am sorry am late, but about to arrive. I am unable to meet u now, I have to get a loan 4 S/fees from the association & the meeting is at 3:00. By GODS name, do me favour again to RSCH thru somebody my body stil betrays me.or give alternative. I will not manage to see u have tests the whole day please give another day. Ok i will pass de message 2 her 2day evening. Thank you 4 accepting 2 be my supervisor and offering all your time 2 ensure that I complete my research be4 the deadline. God bles u. Gd evening sir! was calling you abt my research report. you told me lst wk dat you would meet me this week. the deadline for submitting is friday. thanks. My report is at your desk for signing. May I come over sir? Ur rsch student is on bed rest 4 two months due backache can't travel a long distance. My marks 4 research are out bt my name is not on d list, I need ur help sir. Am very sorry 2 sms u sir, bt I called, u've left d fon at home. My supervisor help me 2 b put on graduatn list, i checkt but my research marks r not out. I would like to know ua schedule of duties so that I come to meet u for de research report i have submitted. Am sorry I dnt inform u bt due 2 my fixed schedule i had 2 change supervisor. Very sorry 4 not informing u.. THAXS FR DE MSG AM SORRY I HV BN AWAY I'M CONFIRMING THAT I WILL HAND IN MY DRAFT REPORT AT DE END OF FEB 2007 YR STUDENT. Mai RESEARCH was given 2 da RECEPTIONIST. Sir, sorry 4 not responding in time, I'm upcountry & I do request u kindly 2 allow me bring my work mid-next month. I will be so grateful upon ur consideration. Nice time sir. I am about to complete my RSCH, I had some problems. Thank u. Thanks 4 ur service. Thanks 4 the message. I finished & graduated in oct. 2006. I'm finalizing with the last bit of the research a hope to travel any time this week to present it. Am realy very sorry sir! I got an accident dt almost put me 2 death but now i've just dropped mi

report in ur office, kindly act as a parent & tel me wn 2 mt. I still have adv.a/c to retake next sem.my resch work wil be handed over to yu soon for oct gradtn. Thks so much....

Source: Messages sent on supervisors' phones

The SMS messages above indicate that there was a two way asynchronous collaboration between research students and their supervisors and amongst students themselves. The two way collaboration created intimacy between research students and their supervisors and amongst research students themselves. Research students upheld the relationships created even after the supervision. Messages seeking for help from supervisors even after a student had graduated are evident. "I was ur rsch student and I am requestg u 2 B my referee", read one of the messages. The collaboration also paced research students and helped supervisors to monitor their progress. "THAXS FR DE MSG AM SORRY I HV BN AWAY. I'M CONFIRMING THAT I WILL HAND IN MY DRAFT REPORT AT DE END OF FEB 2007 YR STUDENT", read another message. The prompting messages also brought hope to those who had almost given up with research due to one reason or another. "Am realy very sorry sir! I got an accident dt almost put me 2 death but now i've just dropped mi report in ur off, kindly act as a parent & tel me wn 2 mt". Here, the research student was responding to a call made to him by his supervisor trying to find out his ware about, as the student had taken over 12 months in the field without the supervisor hearing from him. The collaboration text messages and calls kept the supervisor in the know about what was happening to his/her research students out there in the field. However, space limitations of the supervisors' mobile phones only allowed a given number of messages to come through. Also, supervisors reported a cognitive overload especially when many messages required instant response. This called for the need to ration messages.

5.3. Influence of Mobile Research Supervision on Students' Throughput

A research project paper is a partial fulfilment for the award of a bachelor's degree at Makerere University. This implies that the longer one takes to accomplish the research paper the longer he/she stays on any given programme or even not finish the programme.

We used the mean comparison test to cross tabulate the ownership of a mobile phone versus the duration (in months) it took a student to accomplish his/her research. The mean comparison test results are presented in Tables 5. We established, as is indicated in Table 5, that those who owned a mobile phone and used it for research collaboration on average used the same time (about 5 months) to complete their research paper as those who owned a mobile phone but did not use it for research collaboration. However those who did not own a mobile phone and never collaborated with their supervisors and colleagues using a mobile phone were likely to spend 2 more months doing their research paper (used 7 months) than those who owned mobile phones and used them for research collaboration.

These results indicate that mobile research supervision contributed less to student throughput. We conclude from these findings that there are other intervening variables that determine students' throughput.

Table 5 Mobile phone research collaboration and research project completion duration

	Completion Duration		
	Mean (Months)	N	Std. Deviation (Months)
Used Mobile Phone for Research Collaboration			
Owned a mobile phone and used it for research collaboration	4.556	63	3.686
Owned a mobile phone but did not use it for research collaboration	4.800	5	4.087
Did not own a mobile phone and did not have research any research collaboration on any mobile phone	7.333	3	5.774
Total	4.690	71	3.776

Source: Primary Data

5.4. Constraints Faced by Research Students and Supervisors During Mobile Research Supervision

In an open ended question, research students were asked to list the challenges faced during the mobile research supervision process. Table 6 provides a tally of the challenges mentioned. The challenges are ranked using frequency counts from the most frequent (ranked as number 1) through to the least frequent (ranked as number 11).

In Table 6, the most important challenge reported was the rigor of the research process. As a way of maintaining quality, supervisors were strict on many issues including: the research methodology and writing style. "I do not entertain mediocrity", said one of the supervisors during an interview with him. Evidence of the reported rigor is attested to in the responses the research students gave to the question – What major research supervision challenges did you encounter during your research?

Table 6. Challenges faced during mobile research supervision

Challenge	Research Students faced with the Challenge (Percent)	Ranking as per most frequent
Rigor of the research process	45 (n = 32)	1
Busy schedule of the research students	42 (n = 30)	2
Non-availability of the supervisor at certain times	40 (n = 28)	3
Costly research process	33 (n = 23)	4
Limited time allocated to the research project paper	32 (n = 22)	5
Limited guidance	20 (n = 14)	6
Long distance traveled to come and meet the supervisor	18 (n = 13)	7
Need for word processed manuscripts	17 (n = 12)	8
Non response/respondents being suspicious	11 (n = 8)	9
Using e-mail and Internet to communicate	9 (n = 6)	10
Electricity load shedding	2 (n = 1)	11

Source: Primary Data

... I disliked a lot of criticism to my work. I was unable to put up a clear research problem for five consecutive times. The supervisor made me rewrite the report several times due to mistakes I had made. My supervisor wanted me to be as knowledgeable as him in research method, something I was not able to cope up with in the little time given. Not knowing exactly what the supervisor wanted me to do in certain instances. Crossing my work all together ...

Source: Primary Data

The responses above point to the fact that the supervisors placed strict emphasis on quality of the research project report. Therefore the MRSI did not compromise research quality standards.

Supervisors reported supervising, in some instances up to 30 students. Towards the closing in of the deadline for research report submission, supervisors reported having an information overload on their mobile phones as students simultaneously sent SMS messages to their mobile phones. This created a problem of increased cognitive load to the supervisor especially when multitasking.

Electricity load shedding caused the least challenge (2%) due to there being a number of alternative power sources which could be used to charge mobile phones (see Section 5.1.2) and power secretarial bureau computers that research students used for processing their work. Other important challenges included:

busy schedule of the research students ($n=30$), non-availability of the supervisor at certain times ($n=28$), costly research process ($n=23$) and limited time allocated to the research project paper ($n=22$).

The above challenges point to the need to have a robust mechanism to aid research students and their supervisors keep in constant touch. This could motivate research students to carry on with their research, an exercise some students have tagged to be 'difficult'.

Research students were asked in an open ended question to give ICT based solutions for solving the challenges mentioned in Section 5.4 above. Table 7 summarises the responses given and uses frequency counts to rank them from most important (as number 1) to the least important (as number 8).

Table 7 ICT based suggestion for improving research supervision process at Makerere University

ICT based Suggestion	Number Research students making the suggestion (Percent)	Ranking as per most frequent
Encourage use of Internet/e-mails in research supervision	92 ($n=65$)	1
Improve on the mobile/internet communications infrastructure	87 ($n=62$)	2
Train research students in use of ICTs	81 ($n=58$)	3
Entrench mobile phones in student support services	81 ($n=58$)	3
Use mobile phones to give guidelines to research students	79 ($n=56$)	4
Send regular SMS reminders to research students	71 ($n=50$)	5
Establish subsidized call centers at regional offices	52 ($n=37$)	6
Partner with phone dealers so as to provide low cost phones to students	50 ($n=36$)	7
Messages, calls and e-mails should be sent/made to research students in good time	50 ($n=36$)	7
Distance education should forge a partnership with secondary schools for sharing ISP fees	40 ($n=28$)	8

Source: Primary Data

The students of today, even in developing countries such as Uganda, are quickly embracing the digital age requirements. Many of the research students (65 out of 71) wished to have online supervision through Internet discussion fora and e-mails. Generally speaking, Table 7 brings out a high enthusiasm for mobile and online supervision. However, there is recognition that the infrastructure and capacity to enable this preferred mode of supervision is not yet mature and needs to be developed.

Asked why those with access to Internet services did not utilise it during the research supervision, 75% of the research students said that their supervisors did not encourage them to use Internet and e-mail. This implies that uptake of mobile and online supervision can be deeply entrenched if the University encouraged staff to adopt its use through provision of the necessary facilities.

5.5. Cost Implication of Mobile Research Supervision

Our research did not go into the detail of determining the cost/benefit ratio of mobile research supervision. However, views were collected from research students, supervisors and administrators regarding their mobile research supervision cost experiences. Supervisors and administrators singled out airtime credit as the major cost of maintaining a mobile phone. This was confirmed by the responses generated from the question which required research students to state whether they had received any disruption in collaboration due to lack of or insufficient airtime credit. Lack or insufficient supply of airtime credit sometimes disrupted communication of 62% of the research students, 8% were always constrained while 30% communicated without limitation of phone airtime credit. For the supervisors, they only communicated (using their mobile) to research students on goodwill since the Department did not provide them with any airtime credit for supervision purposes. Some supervisors used the Department's landline phone to communicate to research students. This, according to supervisors and administrators interviewed, was neither sustainable nor convenient. The supervisors proposed that a proposal be developed and presented to mobile telecommunications service providers for them to consider giving a concession in communications charges to distance learners as one of their social responsibility package.

5.6. Other Academic and Administrative Services that can be Delivered via Mobile Phones to Distance Learners at Makerere University

Other than the text and verbal collaboration, supervisors and administrators indicated the need to use mobile phones for delivery of content to learners and providing links to useful study materials. The distance education administrators felt that the mobile phone could be used to remind students about registration, examination and face to face dates, in addition to giving them motivational messages.

6. Discussion

This evaluation study was aimed at answering the question, "is it worthwhile to continue with mobile research supervision and should we invest more resources and effort into it or not?" Our specific intentions were geared at establishing; i) the supervision aspects that could be adequately handled through mobile research supervision, ii) determine the influence of mobile research supervision on students throughput, iii) establish the constraints faced by research students, iv) find the cost implication of mobile research supervision and v) find other academic and administrative components that could be supported via the mobile phone. The

answers would be useful to educators and policy makers wishing to create an m-learning pedagogy and formulate m-learning policies.

The results indicate that over 96% (n=68) of the research students who participated in the study, owned mobile phones and were using them for research collaboration (93%). This pervasiveness has made the mobile phone a handy and cost effective collaboration device. The phones owned by students and the general public in Uganda are of various generations (1G, 1.5G, 2G, 2.5G, 3G, etc.) and capabilities. However, the cross cutting advantage of these varied phones is their ability to broker voice and at least 150 character text messages. With continuous improvements in mobile technology and decline in prices of mobile devices, current low end mobile phone holders will soon be able to acquire high end mobile phones capable of accessing the World Wide Web and send and receive e-mails. With more learners preferring online and mobile supervision (92%), a new breed of learners has emerged. Such breed of learners consider their mobile phone as the number one source of information. Learning, just like their mobile devices, is becoming personalized and a new form of pedagogy referred to by Fisher and Baird [2006-2007] as the 'mobile digital pedagogy' has to be put in place, together with personalized technologies to support the student to collaborate with fellow research students and his/her supervisor.

The study has also revealed that mobile research supervision is capable of removing the loneliness associated with distance learning and increase flexibility in distance learning. This is consistent with Fagerberg and Rekkedal [2004] who see m-learning as having come to re-introduce the flexibility that had been phased out by online learning. "The introduction of the desktop computer (and other learning technologies), which required the student to study at a certain place, often also at a certain time, reduced flexibility of distance learning", [Fagerberg and Rekkedal 2004, p.3]. With the mobile phone, it was possible at anytime in any place, for the research students to schedule face-to-face meeting appointments with their supervisors, seek guidance from their supervisors, communicate progress to their supervisors and create intimate working relationships. These aspects of supervision amplified Fisher and Baird's [2006-2007] need to personalize learning.

Because mobile research supervision personalised learning, the study has revealed that the additional voice and text collaboration on the mobile phone helped to minimise the perceived negative impact of lack of face-to-face interaction and encouraged the development of 'a personal touch' between the research students and their supervisors. This increased students' throughput. On average, research students who owned a mobile phone and used it for collaboration completed their research project paper in the statutory time frame (5 months) while those who did not collaborate using their mobile phones took longer (7 months and above) to complete. This confirms Otto and Wrightson's [2005] exposition that effective student support increases students' throughput in distance learning.

7. Conclusion and Future Work

The evaluation study has showed that the mobile phone is a very handy device for facilitating collaborative/cooperative learning. It is therefore worthwhile to invest more time and resources in developing m-learning solutions geared towards personalising learning and providing learning support to learners at anytime in any place. Research needs to be carried out to determine the actual unit cost of implementing m-learning and how this cost could be shared between the learner and the institution. It has been espoused that the research paper did not significantly contribute to the high attrition rate of distance learners. Hence there is need to undertake a study to determine the causes of high attrition rates at Makerere University. Also, m-learning policies and pedagogy for learners in developing countries need to be developed. The research has also pointed to the need to extend the use of the mobile phone beyond SMS communication to learning content access, delivery and use on mobile phones.

References

- AGUTI, N. J. AND FRASER, J. W. 2006. Integration of Information Communication Technologies (ICTs) in The Distance Education Bachelor of Education Programme, Makerere University, Uganda. *Turkish Online Journal of Distance Education -TOJDE*, 7(3), Article 7.
- BROWN, A., AND CAMPIONE, J. 1996. Psychological theory and design of innovative learning environments: on procedures, principles, and systems. In *Innovations in learning: new environments for education* (pp. 289-325), L. SCHAUBLE AND R. GLASER Eds. Mahwah, NJ, Erlbaum.
- COMMONWEALTH OF LEARNING 2007. Unit 4 – Research and evaluation? PREST-A1 Course Module. In *Research Foundation for ODL Practitioners (A1+A2) course materials*. Online course, 24th September 2007 – 18th November 2007.
- FAGERBERG, T. AND REKKEDAL, T. 2004. Enhancing the flexibility of distance education - designing and trying out a learning environment for mobile distance learners. *A Paper presented at the 21st World Conference on Open and Distance Education, Lifelong Learning in the Networked World*. Hong Kong, 18-21 February 2004. Retrieved from <http://learning.ericsson.net/mllearning2/files/icde2004finalpaper.doc> on February 9, 2008.
- FISHER, M., AND BAIRD, E. D. 2006-2007. Making mLearning Work: Utilizing Mobile Technology for Active Exploration, Collaboration, Assessment, and Reflection in Higher Education. *Journal of Educational Technology Systems*, 35(1), 3-30.
- GRAHAM, G. 2004. E-learning: a philosophical enquiry. *Education + Training*, 46 (6/7), 308-14.
- KAJUMBULA, R. 2006. The effectiveness of mobile short messaging service (SMS) technologies in the support of selected distance education students of Makerere University, Uganda. *A Paper presented at the Forth Pan-Commonwealth Forum on Open Learning*, Ocho Rios, Jamaica, 2006.

- KEEGAN, D. 2005. The incorporation of mobile learning into mainstream education and training. *In Proceedings of the 4th World Conference on MLearning (MLearning: 2005)*, Cape Town, SA, 25-28 October 2005.
- MUYINDA, B. P. 2007. Mlearning: pedagogical, technical and organizational hypes and realities. *Campus-Wide Information System*, 24 (2), 97-104
- OTTO, A. AND WRIGHTSON, T. 2005. *Makerere University Distance Learning Programmes Evaluation Report*. Kampala, Makerere University.
- PASK, G. 1975. Minds and media in education and entertainment: some theoretical comments illustrated by the design and operation of a system for exteriorizing and manipulating individual theses. *In Progress in cybernetics and systems research (Vol. 4, pp. 3850)*, R. TRAPPL AND G. PASK Eds. Washington and London, Hemisphere.
- SHARPLES, M., TAYLOR, J., VAVOULA, G. 2005. Towards a Theory of Mobile Learning. *In Proceedings of the 4th World Conference on Mlearning (M-Learning 2005)*, Cape Town, SA, 25-28 October 2005.
- ZHANG, D. 2003. Delivery of personalized and adaptive content to mobile devices: a framework and enabling technology. *Communications of the Association for Information Systems*, 12, 183-202.

Assessing Implementation of blended Learning in Constrained Low Bandwidth Environment

Nazir Ahmad Subail

In order to meet the exponentially growing demand of higher education, implementation of e-learning is an ideal choice for the universities. And the introduction of the concept of blended learning model is a natural start for the implementation of e-learning solutions for the universities in developing countries which is pursued by many organizations in rich economies. When organizations are in the process of implementation of new type learning, a number of factors come into play. Some factors are about the technology, others about the prospective users, still others about the local context of use and the associated costs [Wilson et al. 2002]. This paper sought to study the readiness of organizations seeking to implement blended learning in constrained low bandwidth environment with an emphasis on university settings and proposes an e-readiness model. The study employed three types of instruments; Institutional Readiness Questionnaire, Lecturers' Perception Questionnaire, and Students' Perception Questionnaire.

1. Introduction

In order to meet the exponentially growing demand [Goddard 1998] of higher education, implementation of e-learning is an ideal choice for the universities. And the introduction of the concept of blended learning model is a natural start for the implementation of e-learning solutions for the universities in developing countries which is pursued by many organizations in rich economies. Blended learning model combines best experiences in face-to-face learning and e-learning thoughtfully to optimize the learning process. Blended learning has been successful due to combining successful elements of both traditional and technology-enhanced learning and that it enhances the learners' performance over single-delivery mode [Zenger and Uehlein 2001; Barbian 2002].

Least Developed Countries (LDCs) fall under the category of constrained low bandwidth environment. Constrained low bandwidth environment includes: insufficient bandwidth, inadequate telecommunication infrastructure, irregular power supply, high cost of technology, among others. The bandwidth available to a user in LDCs is much lower than that in developed world [aidworld 2006]. An average university in LDCs pays 50 times more for their bandwidth than a typical university in other parts of the world [Steiner et al. 2005].

The issue of low bandwidth can be addressed through bandwidth optimization. This approach puts emphasis on how to explore ways to control and manage many bandwidth hungry Internet applications that consume this institutional resource [infobrief 2003]. Universities can use the available conserved institutional bandwidth to optimize the media performance.

However, addressing only the issue of low bandwidth environment doesn't suffice to complete of the complex process of blended learning implementation composed of many inter-related factors. According to Wilson et al. [2000], "Some factors are about the technology, others about the prospective users, still others about the local context of use and the associated costs". To address the issues of complexities that abound during the implementation process of blended learning, researchers have called for the need of a framework to guide the practicing managers. For example, Bates [2000] argues that an over reaching framework is needed to ensure that the use of all technology needs are embedded within a framework. Others researchers [Hadjerrouit 2007; Uys et al. 2004] support the idea of Bates and suggest that blended learning needs to be implemented within a clear and strategically developed framework in order to have the desired impact.

But prior to the development of the framework, it is imperative to assess the preparedness of the organizations planning to adopt blended learning strategy for their course delivery. This in turn would lead to the successful design of a strategy based on the realities, opportunities and constraints identified in the e-readiness model [e-readiness Guide 2002]. E- Readiness as defined by Borotis and Poulymenakou [2004] is "the mental or physical preparedness of an organization for some e-learning experience or action". Also it is equally important that various stakeholders; students, lecturers and administrative staff [Sorenson and Reiner 2003] are "e-ready so that a coherent achievable strategy, tailored to meet their needs, may be implemented" [infodev 2001; Sife et al. 2007].

Based on this background, this paper sought to study the readiness of organizations seeking to implement blended learning in constrained low bandwidth environment. And assess the awareness and perceptions of academic staff and students with an emphasis on university settings and proposes an e-readiness model.

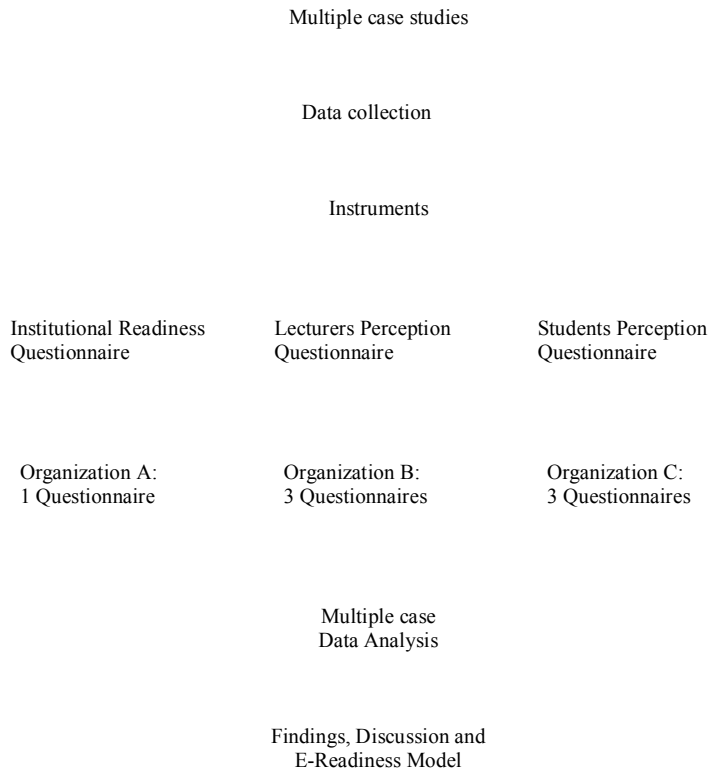
2. Methodology

By using multiple case study approach this study developed three survey questionnaires; Institutional Readiness Questionnaire, Lecturers' Perception Questionnaire, and Students' Perception Questionnaire to explore factors relating to e-readiness in constrained low bandwidth environment.

3. Research Design

Figure 1 below shows the research design for the study

Fig. 1: Research design framework



4. Participants

For the survey we selected one public and two private universities, namely; Makerere University, Kampala University and Kampala International University all located in Kampala, Uganda. The reason for selecting Makerere University is that this is the most popular Government funded University in Uganda with a population of about 30,000 students which would provide better understanding of the context. Other two universities are private but they have upcountry branches. Hence these three universities make a good representative sample of Ugandan universities.

5. Data Collection Instruments

A combination of three data collection instruments; Institutional Readiness Questionnaire, Lecturers' Perception Questionnaire, and Students' Perception Questionnaire were designed for the study after identifying three key stakeholders; institutions, lecturers, and students. The questionnaires contained open ended and

closed ended questions. Mukazi [2005] argues that identification of stakeholders and to secure their support is key to the success for any technology enhanced initiative in an organization in the context of LDCs.

6. Procedures

After having the survey instruments ready which were paper based, researchers distributed 10 Institutional Readiness (survey) Questionnaire to various Deans/ Heads of department in three Universities in Kampala, Uganda. A total of 12 Lecturers Perception (survey) Questionnaires were also distributed to lecturers by the researchers. In addition to that 47 Students Perception (survey) Questionnaires were distributed among students in two universities. We distributed the instrument to the students belonging to Faculties of Computing and Information Technology in two universities at the time when they were sitting free in their lecture rooms. While for other Faculties questionnaires were distributed and collected through the lecturers of the respective faculties.

The study took place during the period November 2007 –January 2008.

7. Data Analysis and Results

7.1 Institutional Readiness Questionnaire (IRQ)

The provision of appropriate technological infrastructure, administrative, technical, and logistic support is crucial for the successful implementation of blended learning in an organization. Out of 10 IRQs which we distributed to various Dean/ Head of departments of various universities, only 8 responded. The universities include; Makerere university, Kampala university, and Kampala International university in Uganda.

The Institutional Readiness Questionnaire was [synthesized from Badri 2002]. The main areas covered were on Infrastructure, Human resource, National policy and Implementation of blended learning initiatives as shown in Tables I, II, and III.

Table I. Statistics for the items related to Technological Infrastructure

Table 9. Statistics for the items related to human resources

Item	Human resources	No. of Respondents	Percentage of respondents' choice	
			Good	Bad
Q.9	How is the availability of IT professionals for E-transactions in your department?	8	75%	25%
Q.10	What is the level of computer literacy in your department?	8	25%	75%

Q.11	Is there any ICT training facility in your department or university?	8	Yes 100%	No 0.00%
Q.12, 13	What is the attitude of the academic staff, and students towards blended learning implementation in your department?	8	Positive 75%	Not known or not good 25%

Table 10. Statistics for the items relating to National policy & blended learning implementation initiative

Item	Na National policy & blended learning i implementation initiative	No. of Respondents	Percentage of respondents' choice	
Q.15	Is there any Government authority in-charge of ICT in the country?	8	Yes 100%	No 0.00%
Q.16	Is implementation of blended learning readiness a national priority?	8	Yes 62.5%	Don't Know 37.5%
Q17	Is there any partnership between NGOs and Government that exist in the country to improve ICT readiness?	8	Yes 100%	Don't know 0.00%

7.2 Lecturers' Perception Questionnaire (LPQ)

Lecturers' Perception Questionnaires were distributed to lecturers in various faculties in two universities, namely, Makerere University and Kampala University in Kampala, Uganda. The Questionnaire was [synthesized from Kamsin 2005] and had two sections; Background Information of respondents, and their Perception to assess their self-reported readiness for blended learning.

7.2.1 Background Information

The survey included 4 (33.33%) female, and 8 (66.66%) male lecturers. The majority of the lecturers are above 30 years of age as shown in Table IV. The survey reveal that most lecturers have high level of teaching experience; 5 (41.7%) of them have teaching experience ranging from 1-5 years, 2 (16.6%) have 5-10 years while 5 (41.7%) have teaching experience more than 10 years.

Table 11. The respondents' age distribution

Age range	Frequency	percentage
20-30 years	4	33.33%
30-40 years	4	33.33%
40 years and above	4	33.33%

7.2.2 Lecturers' perception to asses their self reported readiness for blended learning

This section concerned how the participant lecturers perceived their readiness for blended learning. Table V illustrates the survey results.

Table 12. Statistics of the items related to Lecturers' perception

Item	Lecturers' perception	No. of Respondents	Percentage of respondents' choice		
Q.7	Where do you access a computer & Internet when necessary?	8	Office / Faculty 33.33%	Internet Café 33.33%	Home (Computer only) 33.33%
Q.8	Do you have basic computer skills?	8	Yes 100%	No 0.00%	
Q.9	What technical problem do you face when accessing the Internet?	8	Low speed 90%	Unreliable Internet service 10%	
Q.10	How is power supply in your department?	8	Fair 66.33%	Irregular 33.33%	
Q.11	Do you have blended learning teaching experience? If yes state the number of years	8	1-10 years 33.33%	< 1year 33.33%	No 33.33%
Q.12	What type of Instructional mode do you prefer?	8	Traditional face-to-face learning 0.00%	Fully online learning 0.00%	Blended learning 100%
Q.13	Do you think implementation of bended learning can enhance learning skills?	8	Yes 100%	No 0.00%	
Q.14	Do you think that blended learning embraces numerous benefits such as; Cost effectiveness, Quality, Interactivity, and Flexibility?	8	Yes 100%	No 0.00%	
Q.15	Can implementation of blended learning make any contribution to education system of Uganda and LDcs at large?	8	Yes 100%	No 0.00%	

7.3 Students’ Perception Questionnaire (SPQ)

Students’ Perception Questionnaire was also [synthesized from Kamsin 2005]. Our survey included the students from various universities belonging to the following faculties: Faculty of Computer Science and Information Technology, School of Education, Faculty of Arts, Faculty of Social Sciences, Faculty of Business Management, Faculty of Art and Interior Designing, which made up the largest number of students being surveyed. More samples had been taken from the Faculty of Computer Science and Information Technology because assumptions are made that their degree is more related to the e-learning system [Kamsin 2005]. The Questionnaire also had two sections like Lecturers’ perception Questionnaire.

7.3.1 Background Information

As revealed by the survey shown in table VI majority of the students were in the age range of 20-30 years. The respondents include; 15 (32.60%) females and 32 (67.40%) males offering courses from Certificate to Doctorate degree. However, there was no student in the sample from Masters Program.

Table 13. Students’ age range

Age range	Frequency	Percentage
< 20 years	10	21.28%
20-30 years	34	72.34%
>30 years	3	6.38%

7.3.2 Students perception to assess their self reported readiness for blended learning

This section describes the students’ perception to assess their readiness for blended learning. The survey results are shown in Table14.

Table 14. Statistics for the items relating to students’ perception

Item	Students’ perception	No. of respondents	Percentage of respondents’ choice		
			Yes	Little Knowledge	No knowledge
Q7	Do you have enough knowledge to use a computer?	47	23.40%	70.22%	6.38%
Q.8	Do you have enough knowledge as how to use Internet?	47	29.79%	48.93%	21.28%
Q.9	Do you like to undergo some formal training in computer applications?	47	91.49%	8.51%	0.00%

Q.10	Do you have some idea about e-learning or blended learning?	47	Yes 19.16%	Not sure 40.42%	No idea 40.42%
Q.11	Do you have any experience of using e-learning or blended learning?	47	Used many many times 10.64%	Used once 19.15%	No 70.21%
Q.12	Do you think Implementation of blended learning can enhance learning skills?	47	Yes 65.96%	Not sure 29.78%	No 4.26%
Q.13	Do you have access to your own personal computer and Internet connection?	47	PC & Internet 4.26%	Only PC 29.78%	No 65.96%
Q.14	What type of learning mode do you support?	47	Traditional face-to-face learning 21.28%	Fully online learning 12.76%	Blended learning 65.96%
Q.15	Do you have uninterrupted power supply in your computer lab at the university?	47	Yes 44.68%	No 55.32%	
Q.16	Do you have uninterrupted power supply at your residence?	47	Yes 44.68%	No 55.32%	
Q.17	Can implementation of blended learning make any contribution to education system of Uganda and least developed countries at large?	47	Yes 65.96%	Not sure 29.78%	No 4.26%

8. Discussion

8.1 Institutional Readiness Questionnaire

In Institutional Readiness Questionnaire the major focus was on areas such as; Infrastructure, Human resource, National policy, and Implementation of blended learning initiatives.

Availability of necessary technological infrastructure is the basic requirement for the implementation of blended learning strategy in any organization, among others. The access to an adequate network infrastructure includes: speed (bandwidth) of Internet, service and support, reliable and affordable Internet connection, hardware and software [Canaria report 2002].

We found that many of blended learning advantages such as to acquire knowledge just-in-time, anytime and anywhere are not guaranteed in developing countries. Most of the students do not have access to Internet and personal computers due to prevailing economic conditions and are thus restricted to come to the lab which could be far away from their homes or workplaces for access to a computer. We also found that it is difficult to attain real-time communication with our students due to insufficient bandwidth and irregular power supply.

It is also confirmed that many faculties and departments either do not have computers, or have very few as compared to the student population, except some faculties such as Faculties of Computing and Faculties of Science. According to survey results, Technical support is non-existent in some faculties but fair in some faculties. Regarding students computer literacy rate, we found that it is fair only among students offering computer science related courses but very low among students in other faculties. Survey results show that attitude towards implementation of blended learning is positive in most of the faculties; even some faculties' say they are eagerly waiting to use it when implemented.

National policy or Institutional policy plays an important role to introduce e-learning. Policy is meant to create an enabling environment and culture at institutional levels through: telecommunications regulations, ethics, mandatory courses and increasing access [Canaria report 2002]. Survey results confirm that there is a Government Ministry responsible for all matters relating ICT in the country and some partnership exists between Government of Uganda and other foreign organizations.

8.2 Lecturers' Perception Questionnaire

As revealed by the results, one third of the lecturers were females and two third were males. Similarly we found that one third of the lecturers were in the age range of 20-30 and two third were of 30 years and above. All the lecturers have a high level of teaching experience and many of them have taught for more than ten years. In addition to that many lecturers stated that they have experience in teaching online courses.

Regarding the respondents basic computer knowledge, the findings were quite encouraging. All the respondents have indicated that they have basic computer knowledge and they access computers and Internet from their office or faculty or Internet café, when necessary. But none have computer with Internet connection at home, although very few have PCs at home. Low speed Internet and unreliable Internet services are among the main technical problems revealed by the lecturers. Non-existence of power supply is another technical problem some faculties are facing as revealed by the survey.

All the lecturers prefer blended learning over traditional face-to-face learning or fully online learning and agreed that blended learning can enhance the learning skills. Following benefits of blended learning are agreed by the respondents: Cost effectiveness, Quality, Interactivity, Flexibility, More direct control, Ease of

use, Self paced, Time saving, and High level of participation and engagement. All lecturers agree that implementation of blended learning will contribute to education system of Uganda.

8.3 Students' Perception Questionnaire

Survey results indicate that majority of the respondents fall in the age range of 20-30 years, offering Certificate, Diploma, Degree, and PhD programs. One third of the students were female in the sample. A smaller percentage of students have computer skills and Internet knowledge, majority has very little knowledge and about one quarter doesn't have any knowledge. It is interesting to note a bigger percentage of students have shown interest to go for computer training. Although some say it is expensive and time consuming.

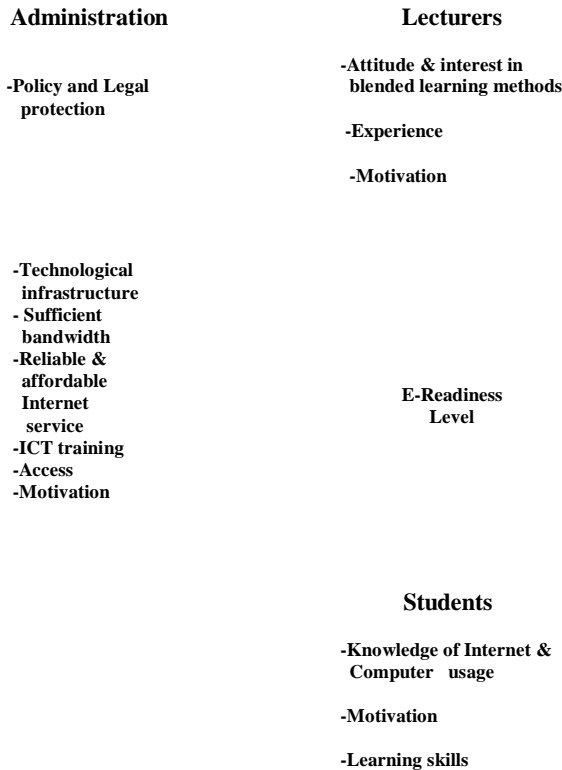
The survey results show that idea of blended learning or e-learning is new for most of the students. But majority of the respondents prefer blended learning method over face -to-face or fully online learning and agreed that blended learning mode can hence the learning skills and can make contributions to the education system in Uganda and LDCs at large.

Survey results also reveal that majority of the students do not have access to PCs and Internet connection, neither had they capacity to purchase it. Although a smaller percentage has access to PCs. Irregular power supply is another technical problem revealed by the students' survey.

9. E-readiness Model

Machado [2007] argues that e-readiness formula for an organization may be achieved by "equation of ICTs plus e-skilled people". Carlos further explains that e-readiness is conceptualized by two factors; 'the ability of Higher Education Institutions and the capacity of stakeholders (Lecturers and students). Figure 2 below provides the e-readiness Model for this study and shows the relationship among three constituent parts of the model; Administration, lecturers and students.

Fig. 2: E-Readiness Model



10. Conclusion

By utilizing blended learning solutions effectively and thoughtfully, universities ability to transfer knowledge to the potential learners can be enhanced in developing countries. Before that, it is necessary to understand the issues that may confront during the complex process of integrating blended learning with multimedia tools, as blended learning relies heavily on technologies. Also, collaboration with other government and organization is important to overcome the high cost of technology. Technical support in form of experienced personnel with technical skill is also needed to implement blended learning in an institution.

However, according to survey, the implementation of blended learning is tough in the universities in LDCs as they are facing a number of challenges such as; low bandwidth environment, irregular power supply, inadequate telecommunication infrastructure, high cost of technology, and accessibility. This study has provided the researcher an opportunity to understand the concern of stakeholders which in turn will lead to successful design of blended implementation framework adaptable

in constrained low bandwidth environment. And the framework will be based upon realities and constraints that abound the context.

Currently, most universities in constrained low bandwidth environment have not strategically implemented online learning or blended learning in their programme offerings and therefore this study, besides the scientific, has its highly pragmatic justification [Begičević, et al. 2006]. This study made a significant contribution theoretically and practically to the existing literature on blended learning in general and its implementation in constrained low bandwidth environment in particular.

The study is based on realistic and idealistic grounds and addresses questions of practicality which leads to instrumental consequences and increased usefulness of knowledge created [JMPE 2004]. The research will offer particular value when the organizations are considering larger change management initiatives within an organization to be addressed through a combination of training using a variety of modalities and media [Shaw et al. 2006]. It is important to understand the concern of stakeholders when designing

The study has served to confirm that ‘ability’, and ‘capacity’ are key factors for the successful implementation of an e-readiness measurement tool [Machado 2007] for HEIs in constrained low bandwidth environment. The ability and capacity are measured through assessing ‘Institutional Readiness’ and ‘Perceptions of Lecturers and Students, respectively.

Acknowledgments

This research was conducted in collaboration with three universities in Uganda; Makerere University, Kampala University, and Kampala International University. Many thanks to the administrators, academic staff, and students of these institutions who contributed to this research by providing useful information.

References

- AIDWORLD 2006. Low Bandwidth Guidelines: Why Should Websites is optimized for Bandwidth? URL:<http://www.aidworld.org/Wiki.jsp?page=ContactUs> (Retrieved on 18 May 2007)
- BOROTIS, S. AND POULYMENAKOU, A. 2004. e-Learning Readiness Components: Key Issues to Consider Before Adopting e-Learning Interventions. (Retrieved on Nov. 2, 2005) http://www.eltrun.gr/papers/eLReadiness_ELEARN2004.pdf
- BATES, A.W. T. 2000. Managing Technological Change: Strategies for College and University Leaders, San Francisco: Jossey-Bass. ISBN 0-7879-4681-8. ... [edutechwiki.unige.ch/en/Bates managing technological change model](http://edutechwiki.unige.ch/en/Bates_managing_technological_change_model) .
- BEGECIVIC, N., BLAENKA, D. 2006. Validation of theoretical model for decision making about elearning implementation: Faculty of Organization and Informatics University Zagreb, Croatia URL:www.projekti.hr/articles/ValidatioTheoreticalModelBegicevicDivjakKOMPLETNO (Accessed on 13 February , 2007)
- BARBIAN, J. 2002. Blended Works: Here’s Proof! *Online Learning*, 6, 6, 26-31.

- BADRI N. K. 2002. Higher Distance/Virtual Education in The Anglophone Caribbean (A Report on the Evolution, Present Status and Future of Higher Virtual Education in the Anglophone Caribbean) October/November 2002.
- CANARIA REPORT 2002. Framework for Rural and Remote Readiness in Telehealth. June 2002. URL: www.readinessguide.org. (Accessed on 8 March, 2008)
- E-READINESS GUIDE 2002. How to develop and implement a national e-readiness action plan in developing countries GeoSINC International, Release Date: April 2002 URL: <http://www.apdip.net/documents/evaluation/e-readiness/geosinc01042002.pdf>
- GODARD, A. 1998. Facing up to market forces. Times Higher Education Supplement (Retrieved on August 2, 2007)
- HADJERROUIT, S. 2007. A Blended Learning Model in Java Programming: A Design-Based Research Approach on Hybrid Learning. Proceedings of the 2007 Computer Science and IT Education Conference
- INFODEV 2001. E-Readiness as a Tool for ICT Development. (Retrieved on Feb. 24, 2005) URL: <http://www.infodev.org/library>.
- INFOBRIEF 2003. Optimizing internet bandwidth in developing countries higher education, INASP infobrief, 1 July 2003. available at: ... www.emeraldinsight.com/.../viewContentItem.do?
- JMPE 2004. Journal of Multi disciplinary Evaluation, Number 1, October 2004
- KAMSIN, A. 2005. Is E-Learning the Solution and Substitute for Conventional Learning?. In International Journal of The Computer, the Internet and Management Vol. 13.No.3 (September-December, 2005) pp 79-89
- MACHADO, C. 2007. Developing an e-readiness model for higher education institutions: results of a focus group study .In British Journal of Educational Technology Vol 38 No 1 2007. pp 72–825.
- MUZAKI, F. 2005. A Framework for Adaptation of Online Learning. In Towards enhancing Learning With Information and Communication Technology in Universities: A Research Dissertation Submitted to the Graduate School in Partial Fulfillment of the Requirements for the Award of the Degree of Master of Science in Computer Science of Makerere University, Kampala-Uganda
- SHAW, S., NICHOLA, S. I. 2006. Effectively Implementing a Blended Learning Approach; Maximizing Advantages and Eliminating Disadvantages; An Eedo Knowledgeware White Paper.
- SIFE, A. S., LWOGA, E.T., SANGA, C. 2007. New technologies for teaching and learning: Challenges for higher. In International Journal of Education and Development using Information and Communication Technology (IJEDICT), 2007, Vol. 3, Issue 2, pp. 57-67. learning institutions in developing countries
- STEINER, R., TIRIVAYI, N., JENSEN, M., GAKIO, K. 2005. African tertiary institution connectivity survey . Nairobi: African Virtual University. URL: <http://www.avu.org/documents/Partneship9> (Retrieved on 24 Feb 2006)
- SORENSEN, D. L. & REINER, C. 2003. Charting the uncharted seas of online student

ratings of instruction. *New Direction for Teaching and Learning*, Vol. 96, 1-21.

UYS, P. M., NLEYA, P., MOLELU, G. B. 2004. Technological Innovation and Management Strategies for Higher Education in Africa: Harmonizing Reality and Idealism. In *Educational Media International* Volume 41, Number 1 / 2004 Pp: 67 - 80

WILSON, B., SHERRY, L., DOBROVOLNY, J., BATTY, M., AND RYDER, M. 2002. Adoption of learning technologies in schools and universities. In H. H. Adelsberger, B. Collis, and J. M. Pawlowski (Eds.), *Handbook on information technologies for education and training*. New York: Springer-verlag URL:http://www.learningtechnologies.co.uk/magazine/article_full.cfm?articleid=156 and [issueid=18](http://www.learningtechnologies.co.uk/magazine/article_full.cfm?issueid=18) and [section=1](http://www.learningtechnologies.co.uk/magazine/article_full.cfm?section=1) (Retrieved on 16 September, 2006)

ZENGER, J. , & UEHLEIN, C. 2001. Why Blended Will Win. *T+D*, 55, 8, 54-62 Aug 2001.

33

Information Communication Technologies Acceptance and Use Among University Communities in Uganda: A Model for Hybrid Library Services End-Users

Prisca K. G. Tibenderana and Patrick J. Ogao

Since their inception during the first century, library institutions have used manual operations to provide services to patrons. Information seekers in these libraries spent a lot of time perusing through card catalogues and searching rows of stacks for material that may have been checked out by someone else. The tradition tool and techniques were inadequate and time consuming. The introduction of the new technologies has changed this concept and what we have now are hybrid libraries based on computer networks and physical facilities. This investigation is an effort to examine alternative measures of intention in revalidating and expanding (Venkatesh et al. 2003) UTAUT model in the context of hybrid library services using university communities in a less developed country, Uganda.

1. Introduction

University libraries play a critical role in the emerging dynamic environments of higher education.

Emerging Issues of Hybrid Libraries

The introduction of computers and other telecommunication technologies in libraries during the last twenty years has altogether changed the concept of a library and the library profession, since the information seeker is no longer confined to the walls of the library, (Kwak et. al., 2002; Bevis, 2003; Rosenberg, 2005; Vinitha, 2006). As ascertained by (Kwak et. al., 2002), majority of university libraries are now hybrid libraries, depending on both electronic and print media based on network and physical facilities. And because of this, many university libraries are part of campus-wide networks, with Internet access (Alan, 1996; Kiondo 1997; Martey 2004; Rosenberg, 1998 & 2005). In this way, university libraries worldwide are able to offer their patrons remote services and access to the vast networked resources. Unlike in the traditional library where users are required to have the ability to read, in an e-library environment users require basic ICT and information searching skills.

Many university libraries embarked on automating their operations and services using the new technologies. To enable automation processes, a library purchases and installs application software. The library is then equipped with a network server and a number of computers (PC) distributed in a local area network (LAN) and in some cases also on a world wide web (www) within different parts of the

library and faculties and administrative buildings of the university. According to (Alan 1996), “the system is then used to create bibliographic databases, control acquisitions, cataloguing and serials, effect bar-coded circulation, book reservation and recall system, current awareness services (CAS), selective dissemination of information (SDI), online literature searches of international databases through CD-ROMs and via Internet and support interlibrary loan services and electronic document delivery services (EDDS). Card catalogues are replaced with computer terminals.” Indeed, Internet and CD-ROM technology have made it possible to store much more information as well as provide immediate access to resources throughout the world.

Observations made by (Leedy 1993) that in the past, information seekers spent much time pouring through card catalogue, and searching rows of stacks for material that may have been checked out by someone else. This process was time consuming. Information seekers knew a librarian as a source of assistance when the catalogue and guides were not useful. In addition, an information seeker often found vital information in a book located near the one he/she has used before, because library materials (books, journals) were organized by discipline. With the use of computers and other communication networks, one can (in theory) get the information required electronically from wherever it is located. The main role of a librarian is now to assist end-users in searching techniques and the use of technology.

The trend towards the delivery of information services in university libraries using ICT gathered momentum on the African Continent after about 2001 (Rosenburg 2005). This is particularly true in the case of online information access, CD-ROM databases and serial publications (Kiondo, 1997 & Martey, 2004). Electronic library services introduced in universities in Uganda at the turn of the last century include but not limited to: End-Users Training Programmes, Printing, Photocopying, Internet web browsing, E-mail, Online full text journals articles, Electronic books, Online Public Access Catalogue (OPAC), Bibliographic databases, Online indexing and abstracting services, CD-ROM services, Library websites, EDDS, Electronic Document Reserve, Electronic Reference and Information, Bar Coded Circulation, Current Awareness, SDI, Electronic Book Reservation and Recall, Electronic Interlibrary Loan, Digitization and Microfilming Services.

According to Information System research, for such services to be used they need to be accepted. The study, therefore, aims at establishing levels of acceptance and use of the technologies, as the major motivation for this investigation.

Rationale for the Study

Despite the fact that technology acceptance has been studied for over two decades now resulting in several evaluation models which have been validated in different contexts and different cultures, this study found that non of the models has been validated within electronic library services context in a country South of the Sahara.

- a) The result of this is that no end-users acceptance and usage levels of e-library services have been studied not only in Uganda but also elsewhere in the world.
- b) By addressing a direction highlighted in (Venkatesh et al. 2003), to examine alternative measures of intention and behavior in revalidating and expanding the research to other contexts, this investigation studied end-users' acceptance and use of e-library services. Accordingly, the investigation modified UTAUT model and developed an end-users evaluation model using e-library services context in universities in Uganda an African country South of the Sahara.
- c) Lastly, (Venkatesh et al. 2003) found that gender and age have received very little attention in technology acceptance research, yet Venkatesh's results indicate that these variables moderate most of UTAUT four key relationships. This study further validated the findings using e-library services in universities in Uganda a less developed country in an African setting (Anandarajan, et. al., 2000 & Oshlyansky et. al. 2007).

Research Questions

In order to study the underlining problems stated above and based on UTAUT (Venkatesh et al., 2003) six assumptions were postulated:

Ha₁. End-users in universities accept to use e-library services.

Ha₂. The UTAUT constructs demonstrate an effect on behaviour intent to use e-library services.

Ha₃. The research model constructs will account for a significant percentage of the variance on the end-users intention to accept and use electronic library services.

Ha₄. Relevancy does demonstrate an effect on behavioral intention to use e-library services.

Ha₅. Relevance moderated by awareness demonstrates an effect on behavior intention to use e-library services.

Ha₆. Usage and intention to use e-library services account for a significant percentage of variation of perceived benefits.

Conceptual Model: The Unified Theory of Acceptance and Use of Technology (UTAUT)

The theoretical basis for this investigation steams from the study by (Venkatesh, et. al. 2003). The Unified Theory of Acceptance and Use of Technology Model (UTAUT) provided the motivation for this research because it enables the studying of acceptance and use visa verse none acceptance and none use of technology. UTAUT indicates that behavior intention to use a technology, in this case the hybrid library, is influenced by people's perceptions of performance expectance, effort expectance, social factors and facilitating conditions which are moderated

by gender, age, experience and voluntariness. The tool used in UTAUT model utilizes communal cognitive elements which better outline and dictate user acceptance of Information Systems. These elements are highly rated constructs because they include culture, which has a great effect on the highly rated constructs (Oshlyansky, 2007). Besides, (Venkatesh et al. 2003), model allows expansion or deduction in order to provide alternative measures of intention and behavior in revalidating and expanding the model to other contexts. The UTAUT Model is unification from eight user acceptance and use of technology models developed by (Venkatesh et al. 2003). The historical background for its four constructs is presented in the UTAUT (Venketesh at al. 2003) and the researchers find the background well documented and will use them as presented therein. Moran et. al., (2006) too give a fairly detailed background of the constructs.

The current investigation therefore, extended the model to study end-users acceptance and use of e-library services context. Universities in Uganda provided a good environment because of the recently introduced e-library services to end-user communities. The study was structured using UTAUT (Venketesh at al. 2003) constructs with modifications. The two additional constructs were got from external user holistic studies of library services (Ranganathan, 1931 & Nicholson, 2004). Introduced was the independent variable of *relevance* which means “suited given requirements or pertinent to specific requirements” and a moderating variable of *awareness* that means “having knowledge of some fact about something” and were considered vital for end-users services oriented technologies.

2. Related Works

Technology Acceptance Model (TAM)

Technology acceptance is mainly about how people accept and use the introduced technology. Quite often people think that introducing new technology results in service acceptance and use. However, several research findings dispute the claim, showing that there are several other factors that affect technology acceptance and use [Carlsson, et al. 2006]. According to Louho, Kallioja and Oittinen (2006) technology acceptance models are used to explain how users will accept and use a specific technology in future.

Developed by Davis (1989) for his doctoral research, the Technology Acceptance Model (TAM) was based on Theory of Reasoned Action (TRA) Fishbein & Ajzen, (1975). Davis together with Bogozzi and Warshaw (1989) jointly published an article in MIS Quarterly, where TAM theory was presented. According to Davis (1989) perceived usefulness (PU) and perceived ease of use (PEOU) both affect people's intention to use, thereby, contributing to either usage or non-use. Davis (1989) indicates that usefulness was more significantly affected by usage than ease of use. Davis (1989) concludes that perceived usefulness had a stronger correlation with user acceptance of technology. TAM's major strengths are that it provides factors which lead to IS acceptance, provides room for extensions and elaborations

better than other competing models (Taylor & Todd, 2001). Some observed shortcomings by users of TAM are its failure to determine barriers that hinder technology adoption (Taylor & Todd, 2001) and possibly its simplicity, which has led to its being over-used at the expense of designing other models. TAM's acceptance in IS research is documented in Lee, Kozar and Larsen (2003).

The Unified Theory of Acceptance and Use of Technology (UTAUT)

A comparison of the determinants found in major acceptance and use models is presented by Moran (2006); Lee, et al. (2003); Vanketesh et al. (2003) and Venketesh & Davis (2000). Vanketesh, et al. (2003) in particular review the user acceptance literature systematically by comparing eight previous models and the predictive factors specified by each and by so doing developed a new model the "Unified Theory of Acceptance and Use of Technology model" (UTAUT).

Determinants of acceptance of technology in UTAUT are: *performance expectancy*, *effort expectancy*, *social influence* and *facilitating conditions*: where *performance expectancy* means the degree to which a user believes that using a technology will provide gains in job/study/teaching/research performance; *effort expectancy* means the degree of ease in using the system; *social influence* means the degree to which an individual perceived that it is important that others believe that they should use the new system; *facilitating conditions* on the other hand means the degree to which individuals believe that there is organizational and technical support for using the system. The four determinants of user acceptance in UTAUT were moderated by gender, age, experience and voluntariness (Venkatesh et al. 2003).

Since UTAUT was designed, it has attracted many scholars in IS research. Amongst the first was Garfield (2005) who used its tools to analyze the acceptance of Computers in Bentley College. A study by (Pu-Li and Kishore 2005) weblog systems to validate UTAUT constructs and conclude by advising researchers to be cautious when interpreting results using UTAUT scales. The study by (Louho, Kallioja and Oittinen 2006) discusses factors that affect the use of hybrid media applications using UTAUT as the conceptual model. By studying College Students' acceptance of Tablet Personal Computers and modified UTAUT, (Moran 2006) introduced "self efficacy" and "anxiety" determinants because of their significance in other technology acceptance models.

Results of Moran show a high correlation between attitude toward technology use and anxiety. Calrsson et al. (2006) studied the adoption of wireless mobile communication in Europe using UTAUT, while (Anderson and Schwager 2006) examined the application of UTAUT to wireless LAN technology in smaller enterprises in the USA. Cody-Allen and Kishore (2005) extended UTAUT by adding e-quality, trust and satisfaction constructs to develop an E-Business Quality Model. Heerik et. al. (2006) used cooperation, empathy, assertion, self control, responsibility, trust and competence to evaluate social abilities in the elderly people within an experimental setup. After the experiment, participants were

interviewed using a questionnaire related to that used by (Venkatesh et al. 2003). The investigators used user data collected on human-robot interactions in nursing home for the elderly, and the experiences they went through were utilized to develop guidelines to support human-robot user studies in elderly institutions. Scholars who have also used UTAUT Model include (Zhang 2006) & Yang, et al. (2005). Engebretsen (2005), tested UTAUT constructs in health research project to study the acceptance of EpiHandy in Uganda and South Africa. Results of this study show that Health Workers in Uganda accept the EpiHandy more than those in South Africa.

Technology Acceptance or Use in Libraries

Jung (1997) used a comprehensive tool to measure digital library users' satisfaction and developed two models; one for objective and the other for subjective measurements. Hill, et al. (1997) used multiple methods to evaluate the use of Alexandria Digital Library. Battenfield, (1999) presents strategies for evaluating the usability of Digital Libraries (DL). During the implementation of Perseus Digital Library Project, using a longitudinal and multifaceted view, (Marchionini, 2000) suggested diverse measurements for evaluation of DL. Saracevic, (2000), presents a conceptual evaluation model for digital libraries. Borgman, et. al., (2000) identifies the effectiveness of DL based on the relationship between DL and students' creativity. Lilly, (2001) evaluated a virtue library collection. In their longitudinal survey of opinions, Kwak, et al. (2002] developed a framework for evaluating university libraries in digital environments. Lagier, (2002) used citations, to evaluate usage and usability of online databases. Baruchson-Arbib (2002] carried out a survey on the use of electronic information sources by Israeli College Students. Bar-Ilan, (2003) carried out a survey on the use of electronic databases and electronic journals accessed through the web by academic staff in Israeli universities. Nicholson, (2004) developed a matrix conceptual framework for the holistic measurement and cumulative evaluation of library services (Table 1). Here, *aboutness* refers to the location of the information within the system (library in this case). According to Schamber (1994) the information view of relevance introduces the concept of aboutness which is based on a content match between the query and the documents being sought. *Usability* refers to how well the system can be used without you having any hardship. *Knowledge status* refers to how well one is aware of what is available, and is linked to the introduced concept of awareness. *Value of works* refers to the value that the material has to the user, which is largely influenced by the relevance of the works. Nicholson therefore views a user's use of library services as being affected by their awareness and by the relevance to them of the library's offerings

Table 1: Measurement Matrix. *Source: Nicholson (2004)*

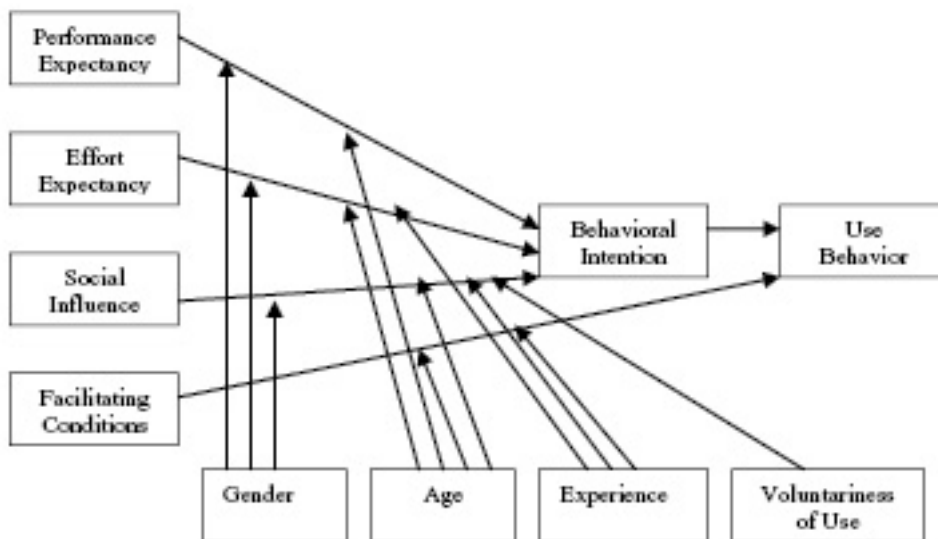
Measurement	Topic	
	<i>System</i>	<i>Use</i>
Perspective <i>Internal (Library System)</i>	Procedures Standards	Recorded interactions with interface & Materials-bibliomining
<i>External (User)</i>	Aboutness Usability	Knowledge status Value of works

3. Methodology

The study data were collected from a randomly selected cross-sectional survey sample of 494 library end-users during the months of June, July and August 2007. Stratified random sampling was important as the first step in fulfilling the study objective. The goal was to achieve the desired representation from various subgroups in the population. The variables considered within the stratification were: public visa via private in which, fifty percent of the sampled universities were government-funded (Mak., Kyambogo, Mbarara & Gulu,) and the remaining fifty percent were privately funded (UMU, IUIU, UCU & KIU); sampling those which had introduced ICT in their libraries. In addition, three faculties were targeted due to the homogeneous characteristics of teaching and subject disciplines across the eight universities. With a population of 28,423 library end-users in the three faculties across the sample universities, a number much greater than 10,000 (Fisher et al. 1983), the z-statistic of 1.22 was used to get the sample size of 346 subjects. However, not all respondents would be willing to respond. Assuming a response rate of 70%, the number of questionnaires administered was calculated as $(100/70) \times 346 = 494$. Eighty nine percent of the target population was undergraduates, this stratum was given greater sample representation of fifty percent. Thirty percent represented postgraduate student samples and twenty percent were administered to faculty staff. After establishing the number of samples in each stratum, random sampling was used when administering the questionnaire to respondents.

The survey instrument was a self-administered questionnaire, which was distributed in hard copy format to respondents by the main investigator and the field research coordinator. For proper follow up on the questionnaires, eighteen research assistances were recruited, trained and assigned to collect the questionnaires from the eight universities. Makerere University that had the greatest number of respondents (47%) was assigned four and the other seven universities, two for each. The field research coordinator, together with the main researcher traveled from one university to the other to ensure systematic and good data collection

Fig. 1: The UTAUT Model



Source: Venkatesh et al., (2003)

Table 2: Distribution of Valid Responses by University

University	Mak	MUST	KYA	GU	UMU	IUIU	UCU	KIU	Total
Frequency	210	36	37	40	34	15	46	27	445
Percentage	47.19	8.09	8.31	8.99	7.64	3.37	10.34	6.07	100

Table 4. The Average Shared Variances and the Correlation Coefficients of Latent Variables.

Model (N= 445)								
	AVC	PE	RE	SI	FC	BI	UB	EB
PE	0.85	0.92						
RE	0.60	0.35	0.77					
SI	0.76	0.35	0.38	0.87				
FC	0.66	0.36	0.44	0.74	0.81			
BI	0.82	0.21	0.38	0.50	0.52	0.90		
UB	0.63	0.37	0.47	0.24	0.28	0.23	0.79	
EB	0.73	0.34	0.36	0.63	0.67	0.55	0.42	0.85

Table 5: T-ratio Test and R² for the Study Model Dependent Constructs.

Dependent Construct	R ²	t-test statistic	Probability Value
Behavioural Intentions	0.30	42.99	< 0.001
Usage behaviour	0.09	86.58	< 0.01
Expected benefits	0.18	44.72	< 0.01

Survey Instrument

The questionnaire was based on the pre-existing tool developed by (Venkatesh et al. 2003), and used by (Anderson & Schwager 2004 and Moran 2006), with modifications to make it relevant to e-library end-users. The tool was pre-tested in two workshops on 20th and 21st April 2007 and results from the pre-test study were presented in a Joint Conference on Digital Libraries in Vancouver, Canada on 20th June 2007 where experts in the area made great inputs in the study instrument. The survey participants responded to questions with a five-point likert scale where a one represents strongly agree and a five represents strongly disagree. All the questions and the wording were critically scrutinized and approved by the two academic supervisors and the reliability of the constructs were determined after the data was gathered, entered, cleaned and ready for analysis.

Data Analysis Software

Besides using PLS Graph for data analysis, SPSS and STATA softwares were found appropriate at different stages for this study because of their ability to model latent variables under both normal and non-normal conditions. This study adopted the Partial Least Squares (PLS) statistical analysis method developed within PLS Graph (Version, 2.91.03.04) software. PLS-Graph is statistical technique software for generating covariance matrix based on structural equation modeling and it has been used by many researchers in technology acceptance and adoption studies (Moran, 2006; Venkatesh et al., 2003; Compeau & Higgins, 1995 & others). The software was used to determine the validity of various construct indicators.

4. Data Analysis

Data was collected from library end-users in three faculties (Education, Science & Social Sciences) across the eight universities. A total of a randomly selected 494 respondents were each given a print copy of the questionnaire. At the end of the field work (end of August 2007), 475 questionnaires had been filled and were collected by research assistants from respondents. Of the 475 collected questionnaires, 3.4% were from unwanted sources, e.g. university administrative staff, Faculty of Medicine etc. In addition, of the 475 returned questionnaires, 0.7% was partially filled and 2.5% had poorly designed responses, for examples one of the respondents ticked the rating scale of 3 throughout the questionnaire. All together, of the returned responses, 7% were categorized invalid and were not included in the study analysis. After entering the survey data, the database

recorded 445 valid responses, indicating valid response rate of 90 percent, which means that a great proportion of the targeted groups provided valid responses. The unreturned questionnaires account for only 3% of the total administered questionnaires.

Demographic Characteristics

The data for valid responses across the eight universities is summarized in Table 2. Of the valid responses, 25.8% were from the faculty of education, 49.2% were affiliated to faculty of science, and 25% reported themselves as social scientists. Respondents were a well-educated class of people since 51% were under taking various undergraduate courses, 22% were postgraduate students and 27% were faculty-teaching staff.

As far as gender is concerned 37% respondents were female and 63% were male. Respondents were between 20 and 60 years of age. Further more, 57% had personal computers or laptops as compared to 43% who had no computers. However, 54% survey participants had their first encounter with a computer during the past five years and 46% had been using computers for six to ten years, a great indication of their acceptance and use of computers. While 51.5% respondents had very good electronic information searching skills, 49% had the required searching skills only to some extent.

Availability of ICTs in Libraries

Availability of the ICTs facilities is a key pre-condition towards learning, adopting and benefiting from electronic library services. Accordingly, 50% respondents across all universities indicated the availability of computers, World Wide Web, Local Area Network, Network Servers and CD-ROM Reader/Writer in their libraries. The availability of the above-mentioned facilities is a guiding measure to this study towards evaluating and comparing different study groups of interest in terms of end user's perception of acceptance, usage and expected benefits from the e-library services. It should be noted that 51.7% and 46.9% of the respondents were not sure about the existence of bar code readers and microfilm readers respectively. Awareness can, therefore, be regarded as an impediment to acceptance and use of a new technology (Brown et. al. 2006) and is one of the measures to be evaluated in this study.

Electronic Library Services

Electronic library services offered by the universities as reported by most respondents include: Internet Web browsing services 85.1%, E-mail services 88.5%, full text journals articles 65.8%, Printing Services 65.5%, Photographic Services 84.1%, CD-ROM Services 55.1%, Library Website 55.9% and photocopying services 84.1%. Electronic library services where 50% respondents were not sure if they were offered include: E-Document Reserve Services 55.5%, Bar Coded Circulation Services 60.3%, Selective Dissemination of Information 53.3%, Digitization Services 50.8% and Microfilming 55.5%.

The Research Model Construction

Reliability of Constructs

The internal consistency of the variables within the construct was determined using Cronbach's alpha. Generally, according to Zhang, Li, & Sun (2006), Moran (2006) and Venkatesh et. al., (2003), the reliability coefficients greater or equal to 0.6 are considered acceptable in technology acceptance theory and practice. The reliability of individual construct was evaluated using STATA (version 9), one of the recommended scientific software for analysis of all kinds of data. Results from the analysis are displayed in Table 3.

Table 3: Scale Reliabilities of Constructs

Construct	No. of Questions	Reliability of the Group
Awareness of the e-library services	5	0.58
Performance Expectancy	5	0.77
Relevance of e-library services	6	0.62
Social Influence	4	0.67
Facilitating Conditions	7	0.60
Behavioural Intentions	5	0.83
Usage behaviour	8	0.81
Expected Benefits from using e-library services	6	0.79

The Research Model Validity

The study adopted conceptual model generated by PLS-Graph to measure direct effects and Generalized Linear model (which is the most common method used in information systems research for moderator analysis) was used to measure the direct and interaction effects between the constructs moderated by other variables. The research model under this study comprises of eight latent variables, which cannot be measured directly but can be measured through other variables called construct indicators. The latent variables used in this study are "Performance Expectancy, Relevancy and Social Influence" which determine behavioral intentions; two determinants of use behavior are "Behavioral Intentions and Facilitating Conditions"; one determinant of expected benefits is "use behavior", and moderator variables of gender, age, experience and awareness. UTAUT model included four moderator variables; gender, age, experience, and voluntariness. However, the research model has replaced voluntariness with awareness to suit the requirements of the context being studied. This is justified by the fact that the study focuses on e-library end users in underdeveloped culture and the issue of awareness is of great importance in both cases (Oshlyansky, 2007; Brown, et al. 2006; Anandarajan, et al. 2000 &. 2002).

All indicators for respective constructs were evaluated for internal consistency (IC) and only indicators which had an IC greater than 0.7 were retained (Moran, 2006; Venkatesh et al., 2003; Chin et al. 2003) & others. Also generated by the PLS-Graph software were the weights, factor loadings, reflective indicators used to generate latent variables and the research conceptual model without moderators. PLS is supported in publications by Moran (2006; Jarvis et al 2003; Chin et al. 2003; Venkatesh et al., 2003) and others.

The factor loadings and R-squared were used to evaluate the complete model analysis and goodness of fit. The goodness of fit values (R-squared and Adjusted R-squared) measures how well the model parameter estimates are able to predict the model performance and also generate the sample covariance matrix. The technique works in such a way that the postulated model is taken as true and adjusts the parameter estimates while minimizing the covariance difference between population parameters and the sample estimates.

Reliability of Variables and Constructs

Reliability in technology acceptance model refers to the degree that the variables are stable and consistent with what they are intended to measure (Moran, 2006; Singleton & Straits, 2004). Cronbach's alpha coefficient is commonly used to measure the internal consistency or reliability of the variables in question. The coefficient generally increases when the correlations between the items (inter-test correlation) increase. The alpha coefficient also depends on the size of the average inter-item covariance, while unidimensionality depends on the pattern of the inter-item covariance. With the exception of one construct (i.e. Awareness), all other constructs had internal consistency or reliability coefficients greater than 0.6, a level that is generally considered good and acceptable in theory and practice (Moran, 2006; Venkatesh et al., 2003; Fornell & Larcker 1981; and Compeau & Higgins, 1995).

Constructs Validity

Construct validity refers to the degree to which a test measures an intended hypothetical construct (Cronbach, 1951 and Straub et al, 2004). The two sub categories of construct validity are Convergent Validity and Discriminate Validity. The important thing to recognize is that two work together. If it can be demonstrated that there is evidence for both convergent and discriminate validity, then by definition there would be evidence for construct validity. One alone is not sufficient for establishing construct validity. Besides, validity means that a measurement that we take represents what we think it should.

Convergent validity is the degree to which similar constructs are related; whereas discriminant validity is the degree to which different constructs are different from each other. Construct validity was evaluated by performing the principal component analysis and factor analysis as discussed by (Venkatesh et al. 2003 and Straub et al. 1989 & 2004). To determine convergent validity, the

main researcher together with the academic supervisors reviewed the construct items. Moreover, the study instrument was pre-tested to ascertain the information flow and identify any missing gaps as described by (Straub et al 2004]. The survey indicators were based on statements from (Moran 2006; Anderson & Schwager 2004 and Venkatesh et al. 2003).

The dimensionality of the forty-one indicators which made up the seven constructs of the survey instruments for the model were subjected to factor analysis resulting into removal of fourteen items. Each construct indicator was evaluated for internal consistency factor loading (ICFL). Any indicator with ICFL less than 0.7 (Compeau & Higgins, 1995b) was dropped and the model was re-estimated. In PLS Graph, the indicators for each construct are evaluated for their internal consistency within the model, if any ICFLs are less than 0.7, they are removed from the model and the ICFLs are recalculated. If there are ICFLs less than 0.7, the lowest ICFL is removed and then the model is recalculated. The process is continued until no ICFL for any construct indicator is less than 0.7 (Gefan & Straub, 2005 and Compeau & Higgins, 1995). The iterative process was done for all constructs including the moderator latent variable (awareness). During the process of estimation, the loading coefficients get lower and others increase indicating the convergence and divergence of the different indicators used to measure the latent constructs. The retained indicators were then weighted and the weights were used to determine the latent variables. Considering the model, the retained constructs together with the indicators are: performance expectancy (PE) with two indicators retained, relevance (RE) with three indicators, social influence (SO) with four indicators, facilitating conditions (FO) with five indicators, behavioral intentions (BI) with four indicators, use behavior (UB) with six indicators, and expected benefits with three indicators respectively. Awareness (AW) as a moderator latent variable has two indicators.

Under the study, the discriminate validity of the model constructs was evaluated. The criteria for measuring discriminate validity are to measure the average shared variance (AVC) between the constructs and their indicators (Maron, 2006; Venkatesh et al., 2003 and Fornell & Larcker, 1981). Discriminate validity is adequate when constructs have the shared variance greater than 0.5. This implies that any single construct is different from the other constructs in the model, and also indicates that at least fifty percent of the measurement variance was captured by the construct (Maron, 2006 and Chin, 1998). The discriminate validity is further confirmed if the diagonal elements are significantly higher than the off diagonal values in the corresponding rows and columns, and the diagonal elements are the square root of the shared variance between the constructs and their indicators. The average shared variances and the correlation coefficients of latent variables are summarized in Table 4. Since all constructs have average shared variances (AVC) greater than 0.5 and the diagonal elements (in bold) greater than correlation values in the respective corresponding rows or columns, the instrument therefore demonstrates acceptance levels of efficiency and truthfulness (validity and reliability) and successful discriminate validity (Moran, 2006 and Chin 1998).

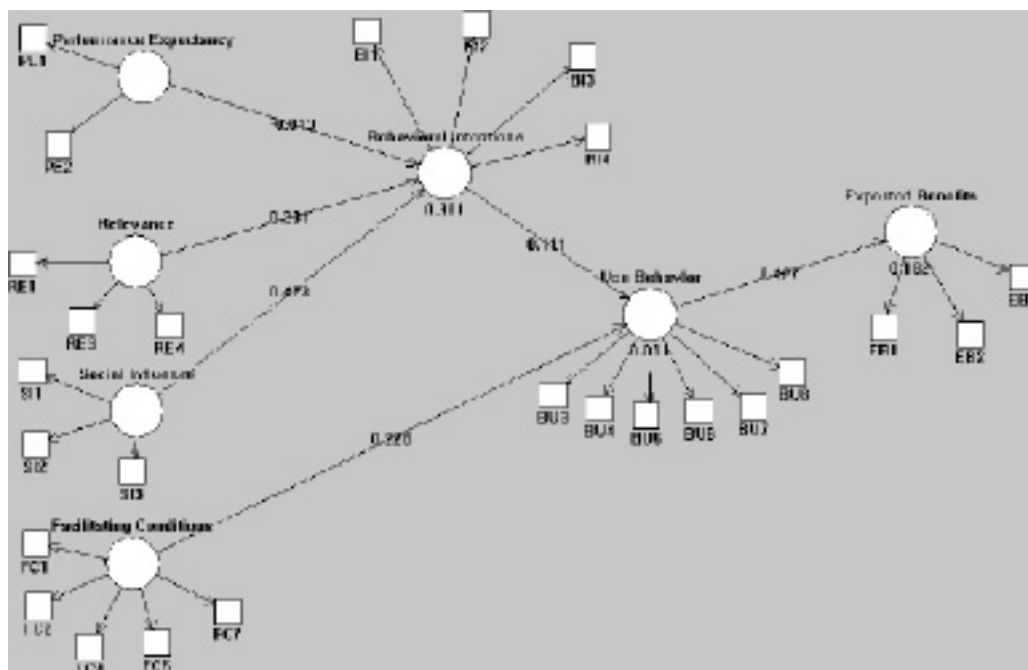
5. Testing the Hypotheses

The hypotheses analysis was based on the research model path results in Figure 3. Based on the analysis, the study was able to respond to the set research questions and hypotheses.

H_{1a}: University end-users accept and use electronic library services.

Results from this study support this statement in that at the eight universities studied, end-users have relatively high inclination to behavioral intention to accept and use electronic library services at 30 percent, followed by relatively significantly lower usage behaviour of the technology services at 9 percent and they moderately expect the benefits at 18 percent giving the overall prediction of 57 percent as presented by the PLS-Graph analysis of the conceptual model in Figure 3. Moreover, all the dependent constructs; behavioral intentions use behavior and expected benefits indicate a positive inclination towards the acceptance and usage of electronic library services. The path coefficients for all the constructs except performance expectancy indicate that end-users in this study had a positive inclination towards acceptance and usage of electronic library services. The above observations support the alternative hypothesis that University end-users accept and use electronic library services.

Fig 3: The PLS-Graph Structural Model



H_{a2}: The UTAUT model constructs will demonstrate an effect on users' acceptance of electronic library services.

This hypothesis is specific and is partially supported by the positive path coefficients between all the model constructs towards the dependent constructs; behavioural intent, usage behaviour and expected benefits with the exception of performance expectancy which demonstrates a negative effect on behavioural intent to use electronic library services. The support of the hypothesis is partial because performance expectancy is a component of the modified model. Social influence and performance expectancy constructs from the modified UTAUT model positively and negatively contribute 42 percent and 01 percent towards behavioural intent. The researcher included relevance construct, which partially contributes 23 percent towards behavioural intent. It is also interesting to note that the influence from the social setting of the end-users has the higher (nearly double) contribution to behavioral intent at 42 percent as compared to a combination of relevance and performance expectance at 24 percent. This finding is not different from Anandarajan, et al. (2000), Brown, (2006) and Oshlyansky, Cairns, & Thimbleby, (2007) whose investigations were carried out in almost similar cultures. This is quite different from Venkatesh et al. (2003), whose findings show performance expectancy construct one of the strongest contribution to behaviour intention to use a technology.

Ha3: The research model constructs will account for a significant percentage of the variance on end-users' intention to accept and use electronic library services.

The model supports this hypothesis as it is constructed from the study which supports 30 percent of the behavioural intention to accept and use electronic library services, followed by 0.09 percent of usage behaviour and 18 percent of the expected benefits to use electronic resources. Overall, the model predicts 57 percent of the intention to accept and use electronic library services.

Ha4: Relevancy does demonstrate an effect on behavioural intention to use e-library services.

This hypothesis is supported by the 23 percent contribution to behavioral intentions by the relevance construct. This is as significant as the social influence construct, which has the highest contribution 42 percent towards behavioural intent in this study. The researcher to this study included this construct for the purposes of establishing end-users' opinion on how relevant electronic resources are towards their learning, teaching and research.

Ha5: Relevancy moderated by awareness does demonstrate an effect on behavioral intention to use electronic e-library services.

This hypothesis is not supported by the data because the effect of relevancy construct being moderated by the awareness on behavioural intent is positive

(beta coefficient = 0.02) but insignificant ($p > 0.1$). Under the GLM, the relations of the different SOUTAUT factors were analyzed. Regression was made based on the structural form of the model which describes the nature of the variables as fixed factors, covariates and the possible measures of interaction effects. The GLM results are used to measure and compare the direct effects of both fixed factors and covariates, and effects due to interactions.

H a6: Usage behaviour will account for a significant percentage of variation on perceived benefits.

Hypothesis six is supported by the path coefficient which acts as correlation between usage behaviour and expected benefits. Usage behaviour positively contributes 43 percent towards expected benefits. It should be observed that this contribution is the highest in the overall study model setting. It is also worth noting that the researcher to this study included the expected benefits construct in. The purpose was to ascertain any possible benefits the end-users expected as a result of accepting and using the electronic library services.

Significance of the Study Model Dependent Constructs

The significance measure was used to determine the level of inclination the dependent constructs have towards the acceptance and usage of electronic library services. This was done using the t-ratio test statistic on the dependent constructs (behavioral intentions use behaviour and expected benefits) and the results are summarized in Table 5.

It can be observed from the results in Table 6, that all the dependent constructs have a significant positive inclination to end-users' acceptance and usage of electronic library services. The significance of inclination is determined by the probability values of the t-test statistic (less than 1 percent and 0.1 percent) and the contribution are determined by the values of squared multiple variances, (R^2).

6. Conclusion

A model for studying levels of end-users' acceptance and use of ICT services has been developed. The constructed research model, a modification of UTAUT eliminated

"*effort expectancy*" and "*volunteriness*" variables found irrelevant in e-library context and replaced the independent variable with "*relevance*" and the moderator variable with "*awareness*" which were found appropriate to this study context. An associated instrument has been developed by specializing the items to e-library services context, taking into account the changed variables. Each of the research model constructs was analyzed and results compared with studies in other contexts. The study findings shade light on levels of end-users' accept and use of hybrid library services in universities especially those in less developed countries. The study also confirms the efficiency and robustness of the UTAUT model to determine acceptance and use of a technology.

Acknowledgement

The authors are greatly indebted to the following individuals for the various support they provided during the course of carrying out the research and the writing of this paper: Prof. Janet Aisabett, Dr. P. H. Schwager, Mr. J. Wokadala, the cross-sectional survey participants and our families; to Dr. W. W. Chin of University of Houston, the developer of PLS-Graph software, who provided it free of charge and. Last but not the least, to i@mak.com the sponsors of this research for the generous financial support

References

- Alan, F. (1996). Dhaka University Library Automation (DULA): An Appraisal Report and Proposals for Action. Dhaka. University of Dhaka.
- Anandarajan, M., Igbaria, M. & Anakwe, U. P. (2000). Technology Acceptance in the Banking Industry: A prospective from A less developed country. *Information Technology & People*. Vol 13, No 4 pp 298-312.
- Anandarajan, M., M. Igbaria, M. & Anakwe, U. P.(2000). IT Acceptance in a Less- Developed Country: A motivational Factor Perspective. *International Journal of Information Management*. Vol 22 pp 47-65.
- Anderson, J. E. & Schwager, P. H. (2004). SME Adoption of Wireless LAN Technology: Applying the UTAUT Model. In Proceedings of the 7thConference of the Southern Association for Information Systems. Online At: <http://sais.aisnet.org/sais2004/Anderson%20&%20Schwager.pdf>. Retrieved on 20 Sept. 2006.
- Baruchson-Arbib S. and F. Shor (2002). The Use of Electronic Information Sources by Israeli College Students. *The Journal of Academic Librarianship*. 28 (4): 255-257.
- Bar-Ilan, J., B. C. Peritz, Y. Wolman. (2003). A Survey on the use of electronic databases and electronic journals accessed through the web by the academic staff of Israeli universities. *Journal of Academic Librarianship*. 29(6):346-361.
- Bevis, D. M. & Graham, J. B. (2003). The Evolution of an Integrated Electronic Journal Collection. *Journal of Academic Librarianship*. 29(2): 115- 119.
- Borgman, C. L. (2000). *From Gutenberg to the Global Information Infrastructure: Access to Information in the Networked World*. Cambridge, MA: The MIT Press.
- Brown I., Rijk K. D., Patel K., Twum- Ampofo, Y., & Belle, J. P. (2006). T-Commerce: an Investigation of Non-Adoption in South Africa.
- Buttenfield, B. (1999). Usability Evaluation of Digital Libraries. In *Digital Libraries: Philosophies, Technical Design Considerations, and Example Scenarios*. Ed. D. Stern. New York: Haworth Press.
- Carlsson, C., J. Carlsson, K. Hyvonen, J. Puhakainen, and P. Walden, (2006) Adoption of Mobile Device/Services – Searching for Answers with the UTAUT. In *Proceeding of the 39th Hawaii InternationalConference on System Science*. S. pp 1-10
- Chin, W. W. (1998). “The Partial Least Squares Approach for Structural Equation Modeling” In *Modern Methods Methods for Business Research*, George A. Marcoulides. Ed. LawrenceErbium Associates, NY. 295-333

- Chin, W. W., Marcolin, B. L. & Newsted, P. R.(2003). A Partial Least Squares Latent Variable Modeling Approach for Measuring Interaction Effects: Results from a Monte Carlo Simulation Study and an Electronic-Mail Emotion/Adoption. *Study Information Systems Research*, 14 (2): 189-217
- Cody-Allen, E. & Kishore, R. (2006). An Extension of the UTAUT Model with E-Quality, Trust & Satisfaction Constructs. Online at: Retrieved on 28 Sept. 2006.
- Compeau, D. R. & Higgins, C. A. (1995a). Application of Social Cognitive Theory to Training for Computer Skills. *Information Systems Research*, 6 (2): 118-143. Compeau, D. R. and Higgins, C. A. (1995b).
- Computer Self-Efficacy: Development of a Measure and Initial Test. *MIS Quarterly*, 19, 189-211.
- Crawford, W. and Gorman, M. (1995). *Future Libraries: Dreams, Madness & Reality*. Chicago, American Library Association.
- Cronbach, L. (1951). Coefficient Alpha and the Internal Structure of tests. *Psychometrika*, 16, 297- 334.
- Davis, F.D. (1989). Perceived Usefulness, Perceived Ease of Use and User Acceptance of Information Technology. *MIS Quarterly* 13 (3): 319-340.
- Davis, F.D., Bagozzi, R.P. and Warshaw, P.R. (1989). User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. *Management Science*. 35(8): 982-1002.
- Engebretsen, T. (2005). Acceptance of Information Technology by Health Research Projects in Low- Income Countries: Intention to Use and Acceptance of Using EpiHandy (IUAUE). Master's Thesis. University of Bergen, Norway.
- Fishbein, M. & Ajzen, I. (1975). *Belief, Attitude, Intention and Behavior. An Introduction to Theory and Research*. Addison-Wesley, Reading, MA.
- Fisher, A., Laing, J. & Stoeckel, J. (1983). Guidelines for overcoming design problems in family planning operations research. *Medline*. 16 (2): 100-105
- Fornell, C., and Larcker, D. F. (1981). Evaluating Structural Equation Models With Unobservable Variables and Measurement Error. *Algebra and Statistics," Journal of Marketing Research*, Vol. 1, No. 3 pp 382-388.
- Garfield, M. J. (2005). Acceptance of Ubiquitous Computing. *Information Systems Management*. Vol. 22 No. pp 424-31.
- Gefen, D., and Straub, D. (2005). A Practical Guide to Factorial Validity Using PLS- Graph: Tutorial & Annotated Example. *Communications of the Association for Information Systems*. Vol. 16 pp 91-109.
- Heerik, M., B. Krose, B. Wielinga and V. Evers (2006). *Human-Robot User Studies In Eldercare: Lessons Learned*. Institute for Information Engineering, University of Amsterdam, Oudweg, The Netherlands. Online document at: <http://www.digital-life.org/jos/index.php?option=com>. Retrieved on 28 Sept. 2006.
- Hill, L. R., Dolin, J. Frew, R. B. Kemp, M. Larsgaard, D. R. Montello, R. M. & Simpson, J. (1997). *User Evaluation: Summary of the methodologies and results for the Alexandria Digital Library*, University of California at Santa
- Barbara. *Proceedings of the ASIS Annual Meeting*. Vol. 34 pp 225-243.

- Jarvis, C. B., Mackenzie, S. B., & Podsakoff, P. M. (2003). A Critical Review of Construct Indicators and Measurements and Measurement Model Misspecification in Marketing and Consumer Research. *Journal of Consumer Research*, 30, 199-218
- Jung, J. K. (1997). Measuring User Success in the Digital Library Environment. Ph. D. Dissertation, Drexel Univ.
- Kiondo, Elizabeth, (1997). The Challenge of CD- ROM and Document Delivery Services at the University of Dar es Salaam Library, Tanzania. *African Journal of Library, Archives and Information Science*, Vol. 7 No. 1 pp 19 – 31.
- Kwak, B. H., Jun, Woodchun, Gruenwold, L. and Hong, Suk-Ki (2002). A Study on the Evaluation Model for University Libraries in Digital Environments. In: *Research and Advances Technology for Digital Technology: Proceedings of 6th European Conference, ECDL. Rome, Italy. Sept. 16 – 18.*
- Lagier, J. (2002). Measuring Usage and Usability of Online Databases at Harnell College: An Evaluation of Selected Electronic Resources. Ph. D. Dissertation, Nova Southeastern University. Online document available at: <http://www.harnell.edu/faculty/jlagier/dissertation.pdf>. Retrieved on 28 Sept. 2006
- Lee, Y., K.A. Kozar, and K.R.T. Larsen, (2003). The Technology Acceptance Model: Past, Present, and Future. *Communications of the Association for Information Systems*. 12, 752-780
- Leedy, P. (1993). *Practical Research, Planning and Design*: Columbus, Ohio, Merrill. pp 17 - 42.
- Lilly, E.B. (2001). Creating accessible Web sites: an introduction. *The Electronic Library*. 19, 6 pp 397 – 405.
- Louho, R., M. Kallioja and P. Oittinen. (2006). Factors Affecting the Use of Hybrid Media Applications. *Graphic Arts in Finland*. 35:3.
- Marchonini, G. (2000). Evaluating Digital Libraries: A Longitudinal and Multifaceted View. *LibraryTrends*. Vol. 49 No. 2 pp 303-333
- Martey, A. (2004). ICT in Distance Education in Ghana. *Library High Tech News*. 5, 16 – 18.
- Moran, M. J. (2006). College Student's Acceptance of Tablet PCs and Application of the Unified Theory of Acceptance Technology (UTAUT) Model. Ph. D. Thesis. Papella University. Mennesota, USA. Online at: [http://www.homepages.dsu.edu/moranm/Research/Mark-](http://www.homepages.dsu.edu/moranm/Research/Mark-Moran_proposal_for_research_v3.51-with_IRB.doc)
- [Moran_proposal_for_research_v3.51- with_IRB.doc](http://www.homepages.dsu.edu/moranm/Research/Mark-Moran_proposal_for_research_v3.51-with_IRB.doc). Retrieved on 26 Sept. 2006.
- Nicholson, S. (2004). A Conceptual Framework for the Holistic Measurement and Cumulative Evaluation of Library Services. *Journal of Documentation*. 60: 164-182. Also an Online Document. Retrieved on 15 August 2006.
- Oshlyansky, L., Cairns, P. & Thimbleby, H. (2007). Validating the Unified Theory of Acceptance and Use of Technology (UTAUT) tool cross- culturally. A Web document. Retrieved on 15 December 2007 at:<http://www.users.cs.york.ac.uk/~pcairns/papers/oshlyansky-Hc107.pfg>

- Pu-Li, J. and R. Kishore, (2006). How Robust is the UTAUT Instrument? A Multigroup Invariance Analysis in the Context of Acceptance and Use of Online Community Weblog Systems. An Online Document at: <http://portal.acm.org/poplogin.cfm?dl>. Retrieved on 28 Sept. 2006.
- Ranganathan, S. R. 1931. *The Five Laws of Library Services*. Madras, The Madras Library Association.
- Rosenberg, D. (1998). IT and university libraries in Africa. *Internet Research: Electronic Networking Applications and Policy*, 8 (1), 5 – 13.
- Rosenberg, D. (2005). Towards the digital library: Findings of an investigation to establish the current status of status of university libraries in Africa. Retrieved on 21 May 2005 from: <http://www.inasp.info>.
- Saracevic, T. (2000). Digital Library Evaluation: Toward an Evaluation of Concept. *Library Trends*. Vol. 49 No. 2 pp 350-369
- Schamber, L. (1994). Relevance of Information Behavior. *Annual Review of Information Science and Technology* Vol. 29 pp 3-48.
- Singleton, L. & Straits, B. C. (2004). *Approaches to Social Research* (4th Ed.). Oxford, England: Oxford University Press.
- Straub, D., Boudreau, M. C., & Gefen, D. (2004). Validation Guidelines for IS Positivist Research. *Communications of the Association for Information Systems*, Vol 13 pp 380-347.
- Straub, D., (1989). Validating Instruments in MIS Research. *MIS Quarterly*. 13(12), 146-247.
- Taylor, R. S. (1986). *Value-added processes in information systems*: Norwood, NJ: Ablex.
- Venkatesh, V., (2000). Determinants of Perceived Ease of Use: Integrating Perceived Behavioral Control, Computer Computer Anxiety and Enjoyment into the Technology Acceptance Model. *Information Systems Research* Vol. 11 No. 4 pp 342-365.
- VVenkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly*, 27(3), 425-478.
- Vinitha, K., Kanthimathi, S. & Devi, K. Tharani (2006). Impact of Information & Communication Technology on Library and its Services. In *Proceedings of DRTC –ICT Conference on Digital Learning Environment*. pp 11 – 13 January.
- Yang, H. I. and Yang, H. L. (2005). The Role of Personality Traits The UTAUT Model Under Online Stocking. *Contemporary Management Research*, Vol. 1. Also an online document.
- Zhang, P., Li and Sun, H. (2006). Affective Quality and Cognitive Absorbtion. Extending Technology Acceptance Research. In *Proceedings International Conference on System Sciences*. Kauai, Hawaii.

34

Value Difference Metric for Student Knowledge Level initialization in a Learner Model-based Adaptive e-Learning System

Robert O. Oboko, Peter W. Wagacha, Euphraith M. Masinde,
Elijah Omwenga and Arno Libotton

Web-based learning systems give students the freedom to determine what to study based on each individual learner's learning goals. These systems support learners in constructing their own knowledge for solving problems at hand. However, in the absence of instructors, learners often need to be supported as they learn in ways that are tailored to suit a specific learner. Adaptive web-based learning systems fit in such situations. In order for an adaptive learning system to be able to provide learning support, it needs to build a model of each individual learner and then to use the attribute values for each learner as stored in the model to determining the kind of learning support that is suitable for each learner. Examples of such attributes are learner knowledge level, learning styles and learner errors committed by learners during learning. There are two important issues about the use of learner models. Firstly, how to initialize the attributes in the learner models and secondly, how to update the attribute values of the learner model as learners interact with the learning system. With regard to initialization of learner models, one of the approaches used is to input into a machine learning algorithm attribute values of learners who are already using the system and who are similar (hence called neighbors) to the learner whose model is being initialized. The algorithm will use these values to predict initial values for the attributes of a new learner. Similarity among learners is often expressed as the distance from one learner to another. This distance is often determined using a heterogeneous function of Euclidean and Overlap measures (HOEM). This paper reports the results of an investigation on how HOEM compares to two different variations of Value Difference Metric (VDM) combined with the Euclidean measure (HVDM) using different numbers of neighbors. An adaptive web-based learning system teaching object oriented programming was used. HOEM was found to be more accurate than the two variations of HVDM. Key words: learner modeling, initialization, web-based learning, nearest neighbors, overlap measure, knowledge level, object oriented programming

1. Introduction

Initialization of learner models is important because it seems unreasonable to consider every student to be starting up with the same knowledge about a given

domain being learned. The system can lose credibility with the students [De Bra, 2000] because of using an inaccurate learner model to generate inappropriate advice for students. Instance-based learning algorithms typically have been used for initializing models of learners in intelligent tutoring as well as web-based intelligent tutoring systems [Virvou & Tsiriga, 2004]. They typically handle numeric attributes well using the Euclidean distance measure but they do not handle nominal attributes as well because mostly, the overlap measure is used for nominal attributes, which loses some information in the process of comparing the values of two feature vectors [Virvou & Tsiriga, 2004; Wilson & Martinez, 1997]. The Value Difference Metric (VDM) was designed to find reasonable distances between nominal attribute values. This paper reports results of investigation on the use of VDM for nominal attributes in the initialization of learner models in a web-based intelligent tutoring system. The focus is on the accuracy in the prediction of initial student scores for different concepts of an object oriented programming course and comparison of a heterogeneous function using Euclidean and VDM distance measures with a heterogeneous function using the Euclidean and overlap distance measures.

2. Context of the Study

The experiment was run for using an object oriented programming course delivered by an adaptive web-based learning system. The second year students at the School of Computing and Informatics of the University of Nairobi were the participants. These students were studying object oriented programming for the first time. They had already studied a programming methodology course which is a pre-requisite for this course.

The students studied the course on their own during laboratory sessions, with minimal support from the course lecturer and the researchers. The experiment was run between February and March 2008. The students started studying the course at different times and the models of those who started later on were initialized using the attribute values of those who were ahead of them in studying the course. The attribute of interest in the learner model, and which is the subject of the initialization process is the learner's estimated knowledge level.

The course was organized into goals. Each goal had one or more concepts, with the content of the concepts of a goal being considered sufficient to achieve the learning goal by the student. In turn, each concept was organized into a number of modules which contained the actual content of the course.

3. Review of Related Literature

Virvou and Tsiriga [2004] used a heterogeneous function with the overlap measure for nominal attributes and the Euclidean distance measure for numerical attributes. The overlap measure compares two values and returns 1 if they are different and 0 if they are the same [Mitchell, 1997]. The Euclidean distance between two training examples (considering all the numerical attributes in the feature vector) is square

root of the sum of the square of the difference between pairs of numerical attribute values [Mitchell, 1997].

Before learning commences, a set of domain independent attributes e.g. program of study and indirectly domain dependent attributes e.g. number of programming languages already learnt by the student are used to form the initial student model [Virvou and Tsiriga, 2004]. The attributes in this model are used to find the distance between two students. These attributes are those considered important to the domain being learned. In this study, these attributes were arrived at after studying related literature [Burkhardt et al, 2002; Schmidt, 2005; Wiedenbeck, 2005; Ventura, 2005].

Figure 1 shows a snapshot of the page for capturing domain independent and indirectly domain dependent information in the learning system for this study.

The feature vector of a new student is compared to the vectors of already registered students, who have had time to learn using the system. The total distance obtained between two student vectors is used to determine the contribution of the score of the already learning student on a given concept during the estimation of the initial knowledge level of the new student for the same concept. The distance weighted k-NN algorithm is used for the estimation and the weight is the inverse of the distance [Mitchell, 1997], meaning that students who are furthest from the new student contribute the least in the initialization of the new student's knowledge level for the different concepts.

The overlap measure looks for the equality or otherwise of values, without taking into consideration the class of the training example (i.e. categorization of the student) [Wilson & Martinez, 1997] or how close the nominal values are from each other [Virvou & Tsiriga, 2004]. For example, the attribute "preference for use of hypermedia during learning" may have possible values as low, moderate and high. The overlap measure would consider the difference between low and moderate as being the same as the difference between low and high just because the values are not the same in both cases. The metric used in this experiment to estimate the distance between two nominal attribute values i.e. VDM takes into consideration not only the similarity of the nominal attribute values but also the similarity/correlation between the classes of the training examples (i.e. the classes of the two students) being considered [Wilson & Martinez, 1997].

Fig 1: Page used by students to enter their personal information.



With VDM, the distance between two values x and y of a nominal attribute a is defined by the formula:

$$VDMa(x, y) = \sum_{c=1}^C |N_{a,x,c} / N_{a,y}|^q$$

where

$N_{a,x}$ is the number of instances in a training set T

$N_{a,x,c}$ is the number of instances in a training set T that have x as the value of attribute a and output class c

C is the number of output classes in the problem domain

q is a constant, usually 1 or 2.

In the RISE system [Domingoes, 1995], q was equal to 1. On the other hand, in the empirical study [Wilson & Martinez, 1997], values of both 1 and 2 were used for q . The experiment was conducted over 15 databases from the machine learning database repository of the University of California, Irvine, all with nominal- and numeric-valued attributes. As a result, a heterogeneous distance function was required. Euclidean and VDM distance measures were used. $Q = 2$ was found to be more robust than $Q = 1$.

4. Methodology

There were two aims of the experiment: To evaluate the effectiveness of student model initialization and to compare VDM with the overlap measure.

The students took an adaptive pre-test for each concept. The pre-test was adaptive so that students did not have to take all questions in the pre-test for a given concept, especially if they are starting out [Yau, 2004]. The overall pre-test score across all the concepts was used to place a new student into one of novice, intermediate and advanced classes.

Then the heterogeneous distance functions (HOEM and VDM) and distance-weighted k-NN algorithm were used to estimate the initial knowledge level for each student, based on the models of the other students of the same class who had interacted with the system for some time. Different numbers of neighbors were also considered for the experiment. Where there were no other students in this class already, then the pre-test scores of the student were used as the student's initial scores.

The initialization scores were expressed approximately to the possible scores from the pre-test (0, 4, 8, 13, 18, 24 and 30) and therefore, comparing them to see if they were similar to the pre-test scores for any given student and for any given concept was possible.

Then, the differences between the pre-test scores and the initial scores are summarized for each of the different scenarios. The difference between the mean of the differences for two groups were tested for significance using the t-test.

The scores initialized using HVDM with $q = 2$ were used for proposing the concepts the system though the student was ready to study. This is because this variation of HVDM had been found to be more robust than HVDM with $q=1$ [Wilson & Martinez, 1997]. In evaluating the overall effectiveness of the initialization by this version of HVDM, a comparison was made between the initial values of students using the k-NN algorithm and those of the pre-tests.

5. Comparisons with Related Work

In the work of Virvou and Tsiriga [2004], the initialization of student models is based on default assumptions, which are attached to each class, called a stereotype. This means that students of the class will have common initial knowledge level scores. They have not given the students a pre-test because a student might score in a pre-test as a result of a 'simple guess or slip', leading to reduction in the accuracy of the student model. pp 292.

In this experiment, we have provided two pre-test questions per level/class instead of one question to reduce chances of a 'slip' or 'guess'. Answers to both questions are used to determine a student's class. If both are answered correctly, then the student is provided with questions of the next level. If one of the questions is answered incorrectly, the student gets another chance to answer it. But if both are wrong (student is in the previous class), or the second question is wrong on second attempt (student is in current class), then the class is confirmed and no

more questions are provided (adaptive pre-testing). This reduces the time spent in the pre-test, an issue raised by Virvou and Tsiriga [2004], but at the same time leading to a reasonable estimation of the student's initial knowledge [Guzman & Conejo, 2002].

In this study, instead of using default scores (when the student first registers his/her category), the student's pre-test scores are used for initialization since they are already individualized and are real scores. When there are other students already of the same class who have interacted with the learning system for some time and their initial scores have been updated, then their scores are subjected to the distance-weighted k-NN to estimate the initial concept scores for the new student.

In the same study of Virvou and Tsiriga[2004], the heterogeneous distance function used has the Euclidean measure for numeric attributes and the overlap measure for nominal attributes. The overlap measure only compares the similarity of the values without considering the class of the training example, yet this information can be used for better estimation of distances between feature vector values as happens with VDM [Wilson & Martinez, 1997]. In our study, VDM was used.

6. Results of the Experiment

The results reported in this paper are part of an ongoing experiment at the School of Computing and Informatics, started in February 2008 and running into March 2008. Three different machine learning algorithms were compared in terms of their accuracy at initializing the concept knowledge scores of a student, based on the scores obtained by similar students (neighbours) who were of the same knowledge category. The knowledge categories were Novice, Intermediate and Advanced. The three different machine learning algorithms used were HOEM, HVDM with the value $Q=1$ and HVDM with the value $Q=2$.

The first part of the experiment, whose results are being reported in this paper, had 8 concepts which students were supposed to learn. Each concept had a pre-test of six questions. The six questions were organized into three categories namely Novice, Intermediate and Advanced. Each category had two questions. The Advanced level questions had a score of 6 each, the questions of the Intermediate level had a score of 5 each and the questions of the Novice level had a score of 4 each. Therefore, the possible scores a student could obtain in a test were as in Table 1:

Note: The scores were not assumed to be on an ordinal scale. There were no such scores as 1, 2, 3, 5, 15 or 27. Scores obtained by a student were always one of the 7 possible scores i.e. 0, 4, 8, 13, 18, 24 or 30. Scores initialized were also rounded to be one of the 7. Scores were considered close to each other if they were next to each other in the order of the 7 possible scores given above. For example, 4 is close to 0 and 8, 13 is close to both 8 and 18.

Table 1: Possible scores in a pre-test for a student

Scenario	Level1 Questions Right	Level2 Questions Right	Level3 Questions Right	Level1 Score	Level2 Score	Level3 Score	T o t a l Score
1	0	0	0	0	0	0	0
2	1	0	0	4	0	0	4
3	2	0	0	8	0	0	8
4	2	1	0	8	5	0	13
5	2	2	0	8	10	0	18
6	2	2	1	8	10	6	24
7	2	2	2	8	10	12	30

The accuracy of an algorithm for a given student concept pre-test and for a given number of neighbours was obtained by comparing the score obtained by the student in the pre-test and the initialized score of the student for the concept according to the algorithm (HOEM, HVDM with $Q=1$ or HVDM with $Q=2$) and the number of neighbours considered during the initialization process (1, 3, 5, 7, ... i.e. odd numbers).

Two measures of accuracy were considered. One of the measures considered the exactness of the two sets of scores i.e. those initialized and those obtained by the student in the pre-tests for the different concepts. Therefore 4 and 4 are considered the exact and 0 and 4 are considered completely different. The exactness is indicated by an entry of 1 in the column of exactness in table 2 below. Inexactness of the scores was indicated by an entry of 0 in the exactness column in Table 2. The second measure expressed the goodness of the classifier's estimation of the initial scores for the student. If an initialized score was either exact or next to the pretest score such as 4 to 8 and 0, then the estimation was considered good. This is indicated by an entry of 1 in the column of goodness of the estimate in Table 2 below. Otherwise, the comparison was considered not good and an entry of 0 was made in the 'goodness' column in Table 2.

Table 2 below shows a segment of the summary of the comparison of the initialized scores with their corresponding pre-test scores using HOEM. There were similar summaries for HVDM with $Q=1$ and HVDM with $Q=2$.

From the three summaries (for the different algorithms) similar to the results shown in Table 2 above, it was possible to compare the three different algorithms. The results obtained when 1 nearest neighbour and the best result obtained when other different numbers of neighbours were considered are tabulated in Table 3:

7. Discussion

Metrics for expressing the distance between two neighbours, especially as used with the k-nearest neighbour algorithm, are many. It has been proved using different datasets that HOEM does not do as well as HVDM with both $Q=1$ and $Q=2$ [Wilson & Martinez, 1997]. The research being reported on in this paper was carried to see if this proof applied to data about students learning a course using an

adaptive web-based course. Other researchers who have used HOEM in adaptive web-based learning environments [Virvou & Tsiriga, 2004] believed that other algorithms for estimating distance between students could do better in initializing student models than HOEM. This was because HOEM only considered the difference between two nominal attribute values and did not consider the class of the entities described by the nominal attributes.

From the results obtained from this research and presented in Table 3, HOEM seems to be doing better than HVDM with $Q=1$ and $Q=2$ because it has on average attained higher 'good' classification accuracy. However, it is important to note that only 17 students used adaptive support to learn the first 8 concepts of the course offered through our adaptive web-based learning system. The other students who participated served as the control group and did not use adaptive learning support. If there were more students learning using the system, perhaps studying a full course, the results might have been different. By the end of the experiment, we expect to have more information that can be used to compare the accuracy of initialization of student models using the three algorithms. HVDM is expected to perform better, especially in the exact comparison of initialized and pre-test scores.

HVDM with $Q=1$ has also attained a higher average accuracy than HVDM with $Q=2$. This, however, is expected to change [Wilson & Martinez, 1997], especially in the exact comparison of scores.

When initialization was done using more than 1 neighbour, the average accuracy seems to improve as the number of neighbours increases. This can be confirmed from the example shown in Table 2 on the initialization of concept scores for student 10. The second phase of learning using the system is going on and the students who started using adaptive learning support in this phase are expected to have a higher number of neighbours participating in the concept score initialization process. Therefore, it will be possible to confirm this conclusion.

Table 3: Comparison of 'Good' estimates of classifier accuracy of the three different algorithms HOEM, HVDM with $Q=1$ and HVDM with $Q=2$.

Student Number	Neighbours					
	K = 1 (1 Neighbour)			Best results from other K (other numbers of neighbours)		
	HOEM (%)	HVDM1 (%)	HVDM2 (%)	HOEM (%)	HVDM1 (%)	HVDM2 (%)
10	60	60	60	80	80	
13	25	25	25			
14	17	17	13			
16	29	29	29	43	29	29
17	53	57	57			
18	100	100	100			

19	50	43	43	17	43	43
21	25	0	0	0	0	0
24	50	100	75	25	100	75
25	60	80	60	60	80	60
26	29	29	29			
27	67	29	43	83	43	43
28	60	40	60			
	13	13	13	7	7	6
	625	609	594	308	375	250
	48.1	46.8	45.7	44	53.6	41.7
	100	100	100	83	100	75
	17	0	0	0	0	0

8. Future Work

In order to test the algorithms further, experiments should be carried out using a bigger sample than has been the case in this study, where only 43 students participated. This will make it possible for the researchers to monitor better the trends in accuracy as the number of neighbours increases.

The classifiers also need to be tested in a learning environment that is arts based so that the performance of the classifiers can be compared to that of science-based courses such as object oriented programming, which is the course students have studied during this research.

A different type of students should also be considered for the experiment. It is important to test distance learning systems such as web-based learning systems with mature students who are probably working. Such students normally have limited time for making contact with the course instructors. These students are more likely to favour systems which support flexible learning in which students have more independence to plan their own learning [Oboko et al, 2007] compared to full time younger students who have more time to be in college and therefore more opportunities for getting learning support from the course instructors.

References

- BURKHARDT, J., DE' TIENNE, F. AND WIEDENBECK, S.: 2002, Object-Oriented Program Comprehension: Effect of Expertise, Task and Phase. *Empirical Software Engineering*, 7, 115–156, 2002
- DE BRA, P.: 2000, Pros and cons of adaptive hypermedia in web-based education. *Journal of CyberPsychology and Behavior* 3(1), 71-77.
- DOMINGOS, P.: 1995, Rule Induction and Instance-Based Learning: A Unified Approach. *The 1995 International Joint Conference on Artificial Intelligence (IJCAI-95)*

- GUZMAN, E. AND CONEJO, R.: 2002, Simultaneous evaluation of multiple topics in SIETTE. In: S. A. Cerri, G. Gouarderes and F. Paraguacu (eds.): *Proceedings of the Sixth International Conference on Intelligent Tutoring Systems, Lecture Notes in Computer Science*, Vol. 2363. Springer-Verlag, Berlin, Heidelberg, pp. 739-748
- MITCHELL, T.: 1997, *Machine Learning*. McGraw-Hill, New York.
- SCHMIDT, C.: 2005, *Cognitively Flexible Hypertext in an Object Oriented Course: Effects of Case-based Instructional Support on Student Learning*. Unpublished thesis, Wichita state University
- WIEDENBECK, S.: 2005, Factors affecting the success of non-majors in learning to program. *Proceedings of the 2005 international workshop on Computing education research* Seattle, WA, USA Pages: 13 - 24
- OBOKO, R, WAGACHA, P, LIBBOTON, A., OMWENGA, E. & ODOTTE, Z. (2007). Non-obstrusive determination of Learning Styles in Adaptive Web-based Learning. *Proceedings of the Moi University 3rd Annual International Conference*. Eldoret, Kenya
- VENTURA, R.P.: 2005, Identifying Predictors of Success for an Objects-First CS1. *Computer Science Education* Vol. 15, No. 3, September 2005, pp. 223 – 243
- VIRVOU, M. AND TSIRIGA, V.: 2004, A Framework for the Initialization of Student Models in Web-based Intelligent Tutoring Systems. *User Modeling and User-Adapted Interaction* 14: 289-316, 2004
- WILSON, D.R. & MARTINEZ, T.R.: 1997, Improved Heterogenous Distance Functions. *Journal of Artificial Intelligence Research* 6 (1997) 1-34
- YAU, J.Y. 2004, *Learning Objects and the Teaching of Java Programming*. Unpublished thesis, University of Warwick