



**UNIVERSITY OF NAIROBI**  
**SCHOOL OF COMPUTING AND INFORMATICS**

---

**An Algorithm for Identity Theft Mitigation: Keypoint Signature  
Verification**

**BY**

**Caroline Wambui Mwangi**

**P53/73007/2014**

**Supervisor**

**Dr. Elisha Abade**

**SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR  
THE DEGREE OF MASTER OF SCIENCE IN DISTRIBUTED COMPUTING  
TECHNOLOGY AT THE UNIVERSITY OF NAIROBI**

**July, 2016**



## **DEDICATION**

To my parents Rev. Dr. Humphrey Mwangi, Mrs Catherine Muthoni Mwangi, Mr Patrick Ndirangu and Mrs Lucy Ndirangu.

## **ACKNOWLEDGEMENT**

Firstly, it is with God's strength that I have achieved this far in my studies. I would like to specially thank my Supervisor, Dr Elisha Abade for his wise counsel and great insight to this study. To the members of the panel for each and every input I received each time I presented, prospered the project to completion in so many ways.

I would also like to extend my regards to Edinah Mose for editing this work. Not forgetting the entire fraternity of the School of Computing and Informatics who in one way or another helped me as a student at the University of Nairobi.

My appreciation also goes to members of the various SACCOs that greatly helped in gathering of information that was used to make this project a total success.

I would also like to extend my gratitude to Dr. Reuben Njuguna, my mentor, leader and role model. His constant support in my office work even when my schedule was very tight went a long way in ensuring a smooth flow of both my office and school work. I will forever be grateful for having worked under him during the period of my studies.

Lastly, and not the least I appreciate my friends Newton Juma, Yeddah, Diana Iropia, Nasserian, Victor Mbithi, Peter Mbatha, Victor Nyamota, Nelson Analo, Kevin Mwangi and Tamba who have been of great help and encouragement in this academic journey.

## TABLE OF CONTENTS

DECLARATION .....	i
DEDICATION .....	iii
ACKNOWLEDGEMENT .....	iv
TABLE OF CONTENTS.....	v
LIST OF FIGURES .....	ix
LIST OF TABLES .....	x
ABBREVIATIONS AND ACRONYMS .....	xi
ABSTRACT.....	xii
CHAPTER ONE .....	1
INTRODUCTION .....	1
1.1 Background of the Study.....	1
1.2 Statement of the Problem .....	3
1.3 Research Objective.....	4
1.4 Research Questions .....	4
1.5 Hypothesis.....	4
1.6 Justification of the Study.....	4
1.7 Scope of the Study.....	5
1.8 Limitation .....	5
CHAPTER TWO .....	6
LITERATURE REVIEW AND CONCEPTUAL FRAMEWORK .....	6
2.1 Introduction .....	6
2.2 Identity Theft.....	6
2.3 Techniques of Identity Theft .....	7
2.3.1 Physical Theft Techniques.....	7

2.3.2 Technology-based Techniques .....	9
2.3.3 Social Engineering Techniques .....	11
2.4 Signature Verification .....	12
2.4.1 Graph Matching .....	12
2.4.2 Hidden Markov Model (HMM).....	12
2.4.3 Neural Networks.....	13
2.4.4 Binary Robust Invariant Scalable Keypoints.....	13
2.5 Existing Identity Theft Mitigation Strategies .....	13
2.6 Mitigation of Identity Theft .....	14
2.7 Theoretical/Conceptual Framework.....	15
CHAPTER THREE .....	16
RESEARCH METHODOLOGY.....	16
3.1 Introduction .....	16
3.2 Research Design.....	16
3.4 Research Setting.....	17
3.5 Study Population .....	17
3.6 Sampling.....	17
3.7 Data Collection Method .....	18
3.8 Research Evaluation.....	18
3.9 Research Validation .....	19
3.10 Data Analysis .....	20
3.11 Human Expert Comparison.....	20
CHAPTER FOUR.....	21
SYSTEM DESIGN .....	21
4.1 Introduction .....	21
4.2 Design of the Proposed Prototype.....	21
4.3 Requirement Specification of the Proposed Prototype .....	23

4.3.1 Participants .....	23
4.3.2 Use Case Diagram .....	23
4.3.3 Functional Requirements .....	24
4.3.4 Non Functional Requirements .....	25
4.4 Developing the Prototype .....	25
4.4.1 Design Decision.....	25
4.4.2 Program Development .....	25
4.4.3 Program Development Steps .....	26
4.5 System Testing .....	26
CHAPTER FIVE .....	27
DATA PRESENTATION, ANALYSIS AND DISCUSSION.....	27
5.1 Introduction .....	27
5.2 Data Presentation.....	27
5.3 Qualitative Data Analysis.....	31
5.3.1 SACCOs Representation in Kenya.....	31
5.3.2 Loan Application Requirements .....	31
5.3.3 Guarantor’s Requirements .....	32
5.3.4 Forgery Details .....	33
5.4 Image Matching Process .....	35
5.4.1 Vector Rasterization .....	35
5.4.2 Ascertain the Reference Object .....	35
5.4.3 Similarity Value Calculation .....	35
5.4.4 Preliminary Matching .....	35
5.5 Proposed Algorithm .....	36
5.6 Pseudo Code.....	38
5.7 Discussion .....	38
5.7.1 Introduction .....	38

5.7.2 Experiment Results.....	39
5.7.3 Discussion of results.....	42
5.7.4 Comparison with Human Expert.....	43
5.7.5 Evaluation of the System.....	44
CHAPTER SIX.....	47
FINDINGS, CONCLUSION AND FURTHER RESEARCH.....	47
6.1 Summary of Findings of the Research Questions.....	47
6.2 Comparison with other Existing Applications.....	48
6.3 Contribution to Previous Work.....	48
6.4 Conclusion.....	49
6.5 Further Research.....	49
REFERENCES.....	50
APPENDIX.....	54
Appendix 1: Project Schedule.....	54
Appendix 2: Questionnaire.....	55
Appendix 3: SACCO Request Letter.....	58
Appendix 4: System Testing Review.....	59
Appendix 4: Code.....	60
Appendix 5: User Manual.....	64
Appendix 6: Research Permit.....	67



## LIST OF FIGURES

Figure 1: Conceptual Framework .....	15
Figure 2: Constraint Propagation .....	17
Figure 3 : A diagram of loan process .....	22
Figure 4: Loan Management system that requires a guarantor for loan applications .....	22
Figure 5: Use case for SACCO Member .....	23
Figure 6: Use case for Loan Officer .....	24
Figure 7 A representations of the responses from the SACCOs.....	31
Figure 8: Mode of loan application.....	32
Figure 9: How many guarantors are required .....	32
Figure 10: Guarantor's requirement.....	33
Figure 11: Forgery Existence.....	33
Figure 12: How many have had their details forged.....	34
Figure 13: Red Flags.....	34
Figure 14: Matching Process .....	36
Figure 15:Member no 3001's Signature.....	39
Figure 16:Member no 3005's Signature.....	39
Figure 17: Bar Chart Indicating Comparison.....	44
Figure 18: Bar Chart System Evaluation .....	45

## LIST OF TABLES

Table 1: Evaluation Matrix .....	19
Table 2: Demographics Data .....	27
Table 3: SACCO .....	28
Table 4: Red Flags .....	29
Table 5: Steps the SACCO has taken .....	30
Table 6: Point variants of known signatures .....	39
Table 7: Point variants of test signature 16.jpg and known signatures.....	40
Table 8: Point variants of known signatures.....	40
Table 9: Point variants of test signature 20.jpg and known signatures.....	41
Table 10: Point variants of known signatures.....	41
Table 11: Point variants of test signature 24.jpg and known signatures.....	42
Table 12: Results from Prototype .....	42
Table 13: Results from Human Expert .....	43
Table 14: Comparison of Human Expert verses the Prototype.....	43
Table 15: Prototype Evaluation .....	45

## **ABBREVIATIONS AND ACRONYMS**

ID-	Identification
ATM-	Automated Teller Machine
DNS-	Domain Name System
HMM-	Hidden Markov Model
BRISK-	Binary Robust Invariant Scalable Key point
AGM-	Annual General Meetings
SACCO-	Savings and Credit Co-operative
PDA-	Personal Digital Assistance
PC-	Personal Computer
SCE-	Signature Code-Euler
DDM-	Directional Difference Matching
TP-	True Positive
FP-	False Positive
TN-	True Negative
FN-	False Negative
FAR-	False Acceptance Ratio
FRR-	False Recognition Ratio

## **ABSTRACT**

The existence of identity theft in society has become a major concern due to the effects it causes to those that are affected by it, more especially in the financial sector. Thus this thesis establishes the existence of identity theft issues in the financial sector loan sections and proposes an algorithm that addresses the mitigation processes of identity theft by having the signatures on the loan forms verified using the implementation of the proposed algorithm, then the results are compared with the human experts verification that are done on a daily basis. From the qualitative data collected from the four SACCOs presented indicate the 93% of the respondents knew that forgery of one's signature in the SACCO exists and from the 93%, 95% of them had been victims of identity theft and 50% of them knew it after deductions were been made from their accounts. The algorithm was implemented in a prototype that was used to test the signatures that were corrected from various individuals that belonged to various SACCOs. The prototype had successfully verified 80.1% of the test signatures and as expected the highest results from the four Human experts verification of forged signature was 8.3% indicating that they had indicated more signatures as originals. The prototype thus recorded an accuracy of 91.4% and a precision of 60.0%.

### **Key words:**

Algorithm, Identity Theft, Mitigation, Handwritten Signature, Signature Verification

# CHAPTER ONE

## INTRODUCTION

Signature verification is the most common natural way of personal verification. It is termed as one of biometric aspects in comparison to finger printing and facial recognition. Biometric systems identify individuals based on their distinguishing characteristics (Zimmerman et al., 2003). The rise in identity theft in financial institutions that provide loans which require guarantors' authorising signature, gives the need to provide convenient and credible measures when loans are guaranteed in order to draw and maintain customer loyalty. Signature verification is based on two modes; the off-line mode and the on-line mode. In this study we dwell on the off-line mode of signature verification since the SACCO members are given loan forms on which the guarantors would also need to sign on in order to facilitate the loan.

### 1.1 Background of the Study

The art of forgery is as old as the letters of the alphabet. Forgery was practised since ancient times in every country where writing existed and paper was used for financial transaction (Koppenhaver, 2007). Forgeries are classified as follows (Karounia, et al., 2010):

1. Random forgery; the forger does not have the shape of the writer signature but comes up with a scribble of his own.
2. Unskilled forgery; the forger knows the name of the original signer but not how his/her signature looks like.
3. Skilled Forgery; the forger has unrestricted access to genuine signature model, practices it and eventually comes up with a forged sample.

Handwritten Signatures are accepted forms of verification in the process of loan applications. These signatures can be manipulated by forging and using the members' details without their knowledge.

Identity theft dates way back even before the immerging of internet technology or technology in general. But with technology, identity theft has become a common crime and even easier and safer to perform without being caught, thus making it one of the most charted white collar crime (CIPPIC, 2007).

On Identity-Theft-Scenarios.com, Identity theft was once a physical crime. The first criminals who stole identities actually murdered their victims. Once the victim's corpse is disposed, the criminal would then acquire the identity of the victim, ID numbers and other private information

(Identity-Theft-Scenarios.com, 2015). By then, the inspiration was never a real financial gain but it was a way to acquire a new beginning. When the telephone was invented, the identity thieves graduated to using this device to acquire an individual's information like date of birth, addresses, bank accounts since at the end of the call there was a promise of financial rewards or other rewards. (Identity-Theft-Scenarios.com, 2015) This was the first ever gadget that was used in making identity theft easy and it is still used today especially in Africa where we have many who are not informed on the dangers of giving out personal information to someone you do not know about. Then the use of paper shredders was encouraged since people would fish in the trash bins and get thrown away bills or documents that contained personal information (Identity-Theft-Scenarios.com, 2015). Just as this was thriving, the internet boomed with varying breakthroughs of gathering personal information.

This existing paradigm creates openings for loopholes in information security of both criminal and civil nature (Barske et al., 2010). For instance, identity crime can be achieved in various ways and where an opportunity has presented itself. This may be by assuming your identity to gain employment, gain a loan or open a bank or credit accounts. In this research, the focus is on financial sectors, specifically SACCOs, as institutions through which identity theft has been used to acquire loan/financial credits wrongfully at the expense of the guarantors, institution or agency.

In the world today, the most commonly used means of document authentication of self or another person is a handwritten signature. These documents would be bank cheques, log books, forms like: Opening bank accounts, loan application forms and Visa application forms. All these may either prove what you possess, who you are or what you know. Financial sectors specifically SACCOs offer members loans and make flexible interest percentages of repayment. This has led to forgery of signatures so as to get loans. When acquiring a loan from an existing financial institution, a guarantor's signature is a necessary requirement in the application form, since it is termed as the authentication stamp of the person guaranteeing the loan in-case someone defaulted on payments. This constitutes the falsification of a guarantor's details and signatures. Signatures have their own uniqueness even when forged they can never be an exact of the owner's. Therefore, a research on the key-point features of a handwritten signature on an applicant's form is used and an algorithm that extracts these features is applied thus enabling the verification process of this signature to confirm that the owner of the signature is indeed the one guaranteeing the loan.

In Kenya, so many people have turned to these institutions as a mode of saving and banking since it is easily accessible and can deliver more efficiently to low income earners than the banks. There is difficulty in distinguishing between common fraud which would be the use of someone's credit card or ATM card to conduct financial transaction. The actual act of identity theft takes place when a criminal fiddles with information that directly affects the identity of a person which would be name, address, signature, membership number, date of birth, telephone number, driving licence number among others (CIPPIC, 2007). This information is then used by the criminal to charade as the victim, gradually taking over his/her identity. With this information, the identity thief may open a bank account, obtain loans, secure employment or even begin a new life in other countries (CIPPIC, 2007).

In one of the SACCOs, a scenario of a forgery case was presented where the one taking the loan applied for a loan and had the guarantors' signatures forged. As the loan was processed, there was no evidence of any existing identity theft. So the client was given his/her loan and for the first few months the SACCO member paid the monthly amount as it was required. Then on the sixth month the client stopped repaying the loan. At this instance of default, the SACCO gets in touch with the people who had guaranteed the loan and it was at this moment that the guarantors become aware of the particular loan. It was hard for the loan officials to understand how they would not have known about the loan until the loan request form was retrieved and the signatures on the loan forms re-compared to the originals of the guarantors and the mismatch was then detected.

## **1.2 Statement of the Problem**

The copying of signatures so as to imprint them on a form for one's financial gain is a crime. In the scenario presented in section 1.1 above, the SACCO member applied for a loan like another he/she had taken before. And since the same guarantors had guaranteed before there was no cause for alarm until the loan was defaulted. If this was to be detected during the loan processing then it would save people from incurring unnecessary charges that they did not have a hand in. The detection of the red flags would immensely give members more confidence in providing guarantorship to other members without fearing the aspect of the same person reapplying for a loan without their knowledge.

## **1.3 Research Objective**

### **General Object**

To analyse features on the signatures that are imprinted on the loan forms by the guarantors and propose an algorithm that will enable the verification of this signatures and thus enhance efficiency in handling of the identity theft cases.

### **Specific Objectives**

1. To evaluate identity theft tactics currently applied in the financial sectors during loan applications.
2. To establish the red flags that can be used as identification in order to point out a forgery case.
3. Propose an algorithm that will enable identity theft mitigation for institutions that use signature verification
4. Evaluate and analyse the algorithm's efficiency

## **1.4 Research Questions**

1. How have the financial sectors been handling the identity theft cases?
2. What is the occurrence rate of these cases?
3. How does the identity thief get the necessary individual's information?
4. Will the algorithm enable quick and easy detection of the forgery?
5. How efficient will it be for the financial sectors to take action on a forgery?
6. How does the algorithm perform in signature verification?

## **1.5 Hypothesis**

1. The staff working in the selected financial sectors (SACCOs) know about Identity theft.
2. They have a record of all members' signatures in soft or hard copies.
3. The guarantor's signature is required so that one can get a loan.

## **1.6 Justification of the Study**

When someone's signature has been used to acquire a loan, the identity thief mostly defaults on the loan and the guarantors who were listed on the application form end up carrying the burden of repaying the loan. For the financial sectors, integrity issues on how they conduct their activities and financial responsibility is also doubted. Thus this algorithm outlines the process of



events that will ensure that the detection of this forgery is done before the loan is processed and thus mitigation takes place at a stage in which a crime will be prevented hence ensuring that the involved victims are not affected.

### **1.7 Scope of the Study**

Due to time limitations, a sample collection was conducted on the existing financial sectors by narrowing down to SACCOs only. There was also the aspect of time and financial constraints, thus the study was limited to Nairobi County.

### **1.8 Limitation**

1. The openness of the financial sectors on how they manage their information.
2. From exiting financial sectors, Savings and Credit Co-operatives (SACCOs) were more readily accessible and willing to offer information that would help in driving the research forward. Thus the research covered only a fraction of the existing financial sectors in Kenya.
3. Time limitation on conducting a full research on the known identity theft cases in more financial institutions other than SACCOs
4. The recognition of Identity theft as a crime in the institutions' constitution and how to deal with it.
5. The changing nature of the identity theft crime and the fact that it is still concealed.

## CHAPTER TWO

### LITERATURE REVIEW AND CONCEPTUAL FRAMEWORK

#### 2.1 Introduction

In the current financial environment, the number of SACCO registered and those starting have rapidly increased to bridge the existing gap between the rich and the poor. In any financial environment the economic health is always threatened since there will be someone who will want to make an easy gain. And the use of Identity theft crime becomes a swift mechanism. Cases of identity theft can be solved if they are discovered earlier or during the build-up of the crime. Thus the research sought to ensure that, in all aspects or undertakings of an institution in Kenya where identity theft is prone to be encountered it is most appropriate to counter attack this beforehand. The research focused on the existence of identity theft in various financial institutions in Kenya.

#### 2.2 Identity Theft

In identity theft, there exist prominent differences which are determined either by the regions or to a certain percentage according to age. Data available suggests that depending on the form of identity theft, all persons, in spite of social or economic background are potentially vulnerable to identity theft.

The USA Congress in 1998 passed the Identity Theft Assumption and defence Act (the US public law 105-138), stating that an identity thief is anyone who:

*["Knowingly transfers or uses, without lawful authority, any name or number that may be used, alone or in conjunction with any other information, to identify a specific individual with the intent to commit, or to aid or abet, any unlawful activity that constitutes a violation of federal law, or that constitutes a felony under any applicable state or local law."]*

Thus, identity theft is the thievery of personal information such as a name, date of birth, passport number or credit card information. Any activity in which personal information is shared or made readily accessible to others creates a chance for identity theft to take place.

In identity theft every human being in spite of his/her social or economic background is a potential victim (Graeme & McNally, 2005). There are three stages that have been identified which are:

- Acquisition- This is the gaining of information through the web.

- The use of someone's identity for a financial scam for the benefit of self.
- Discovery- The time taken for full realization of the crime directly relates to the amount lost or gained by a victim (Graeme & McNally, 2005). At this stage, the criminal justice system may or may not be involved thus at this point; there is need for extensive research.

Identity theft has been referred to as the crime of the new millennium (Hoar, 2001). The increase in the use of Information Technology facilities like the Internet has led to the rise of crimes and many people are susceptible to various criminal activities. In this study, Identity Theft is considered as a standalone crime reference to the United States as defined in the Identity Theft and Assumption Deterrence Act (1998) and belongs to federal crimes (Ogla, 2007).

Identity theft can be accomplished anonymously and easily through a variety of ways as is discussed in the next section. Its experience and consequences to the victims can be shattering. This is readily evident with the recent change of ATM machines and ATM cards. (Maxwell, 2012). This may also occur through credit or debit (visa) card transactions. Various individuals have gradually discovered transactions in their accounts that they had not authorised from the use of their credit or debit (visa) cards. The identity thief images the information while the customer is knowingly purchasing, then uses the same information to perform malicious activities.

## **2.3 Techniques of Identity Theft**

The techniques employed keep evolving and may vary according to the information corrected from the victim. In this research, the techniques are categorised into various types depending on the various different approaches used to solicit information from an individual.

### **2.3.1 Physical Theft Techniques**

This involves theft of sources of personal information: Any item that stores data like a cell phone, iPad or any item which may carry important documents like a wallet and a purse. Then there is dumpster diving, where identity thieves will trash ones household garbage, business trash in order to pick up any paper bills or pieces of paper which may be containing the information they require. The most important thing for this technique is for a company or even at home to ensure that the paper disposed is either shredded or disposed in a way that one cannot pick it up and gain important information. Like in Kenya this would be the ATM slips, copies of one ID or passport and so on.

Change of address and mail theft is also used to get the necessary information. When the identity thieves target ones mail, they will redirect it to another address to which it will go to. This has proved sufficient mostly in the US and Canada where there have been cases of this nature. Through mail redirect, the identity thief gets the necessary information the easiest way possible and abundant time to commit the crime before the victims can even make sense of what is already taking place. Mail theft is easily achieved by getting it from the mail boxes and recycling bins. For instance in Kenya, people still share post boxes making it a very easy technique from which one can acquire the necessary information.

Another technique is where a company engages a person in what seems as a legitimate business. This may be through wire transfers in an auction or a reshipping deal of some items. This scheme becomes a way of acquiring one's personal information, credit card and auction fraud. The victim is usually contacted through chat rooms, over the internet or bulletin boards for job applications.

In the process of digitization, governments are making public records accessible online through other electronic means. While their aim is to reduce costs, make service more readily available and boost openness and accountability, it has significantly become a benefit to identity thieves who gain access to personal information and defraud unsuspecting individuals.

There are also cases where personal information of a deceased person can be accessed from newspapers. This is generally referred to as Tombstone theft. Obituaries mostly provide a person's full name and date of birth. Uninformed funeral homes may also pass out information. All this enables the identity thief to acquire loans or even withdraw from their account like a case in Atlanta, where 80 deceased persons' information was sold for \$600 each and then used to acquire car loan totalling to \$1.5 million (CIPPIC, 2007).

Skimming and personal information trafficking is another means characterised under the techniques of physical identity theft. Skimming is not only limited to debit, credit and calling cards since there are so many other cards that are now in existence and use magnetic strips to store information. For example airline boarding passes contain loads of information on an individual as indicated on the Guardian Unlimited. Personal information trafficking is achieved by "Carder Networks" and other underground networks. The information available in these sources will be as a result of insider abuse or remote exploitations of computer vulnerabilities to access clients databases (CIPPIC, 2007).

An insider job is also a major contributing factor to how one's personal information gets into the wrong hands. The security of the information is only as good as the integrity of the employees.

Once the identity thief has some information he would need to dig deeper to acquire more sufficient information this is known as identity consolidation or “identity breeding”. For example if an identity thief has someone’s ID, he or she can use this to get a sim card replacement then use this sim card to siphon money from unsuspecting persons by calling them, requesting for information to offer rewards, job promises as long as they get some money from the victim, sending false M-Pesa messages and luring the receiver to send back the money when there was nothing sent.

### **2.3.2 Technology-based Techniques**

Phishing is a technique that involves the use of a social engineering by camouflaging as a trustworthy organization in an e-mail message or through their web platforms (Oppliger & Gajek, 2005). Email messages become the main channels of gathering the necessary information from the victims. The phishing messages have become very sophisticated that it is very hard to determine their legitimacy. They may appear as if they have been sent from a bank, indicating an issue with your account that would need immediate attention. These messages are written with a somewhat similar message as the organization would use and the identity thieves also ensure that colours and logos are of the same brand. This is known as spoofing (Oppliger & Gajek, 2005). Spoofed websites are used to entice victims who are then manipulated to accomplish the Phishing technique (Vishesh, 2007). This can be accomplished in two ways, mapping legitimate domain names to legitimate IP addresses with the aim of compromising computer host files and Domain name system poisoning. Exploitation of the DNS is done so as to gain control of the existing website and change the numerical address associated with the textual domain name. This results to any victims visiting the actual address been referred to the spoofed site, but the address bar on the victims browser never changes from the obvious. This is also similar to DNS Cache Poisoning whereby the addressed is changed locally on the actual machine accessing the website instead of the DNS server.

The other technology-based technique applied is the Spyware programs (Post, 2003). They are known for slowing down or even crashing the system and may also cause unwanted advertising and perpetual pop-up messages. This may seem unimportant but in the unknown the spyware tracks the activities of the computer user or enable access to the content of the hard-disk drive. This is because the spyware programmes are able to send information via the internet to the creator of the spyware (Post, 2003). It usually consists of the core functionality which appeals to users and entices them to install and use the spyware and functionality for information gathering.

Internet searches and Google hacking can also be a great source of information (Billig, et al., 2008). Searches from legitimate websites could give vital information of employees or management members. Google hacking consists of using Google search engine to find “hidden” documents on a website (Billig, et al., 2008). Many organization have no idea just how much information one can get from their site if it is not properly managed and configured. This would include payroll details, contacts, ID numbers, NHIF details among others. Apart from Google hacking the other type of hacking that can be used involves exploiting known security holes and vulnerabilities in softwares such as Microsoft Windows (Billig, et al., 2008). For this to take effect corrupt data and a set of instructions are sent to the softwares running on a targeted computer. The corrupted data confuses the software and it start executing the new instructions sent by the hacker. Also, information for a large group of people can be captured by hacking into and stealing data from financial or government databases of business. For example In China, there exists lack of individual privacy concerns and mechanisms for public protection thus making it a potential ground for identity theft (Shao-Bo, et al., 2008).

With the rise of wireless technology, many home users are now getting connected. Identity thieves will visit neighbourhoods detecting these Wi-Fi wireless networks. Wireless equipped laptops, PDAs and other softwares are used to detect the unsecured wireless networks. Once a connection is established then the identity thief can go ahead and acquire the information he/she needs from the device.

When organizations are doing an upgrade of their servers or PCs or a home owner needs to sell out an old PC, they will just sell the out or sometime donate them to schools and so on. The hard-disk drive may have “mother lodes” of the former owner. For servers if they had been used to store up the organizations database, then it would have this information that would be retrieved from it. So when discarding this devices proper measures should be taken since just deleting them from plain site does not necessary mean the using the computer forensics softwares of data recovery or the other existing softwares that are commonly used for data recovery cannot be applied to reveal the still present that is locked away on the hard-disk drive (SecurityFocus, 2003). Students in a Massachusetts Institute of Technology proved that wrongly disposed computers equipment posed high risks. They purchased used computer hard drives and then scanned them for personal information. They recovered medical e-mails that contained personal information and credit card numbers and so much more, thus proving the gold-mine in improperly disposed of computer equipment.

### **2.3.3 Social Engineering Techniques**

This involves the natural aspects of a person in trusting someone else especially those that are close to them. Remarkable efforts have been made in the past years by governments, business and academic research community in understanding the Identity theft issues and also developing solutions to deal with the crime from a social, technological, law enforcement and legislative, business and management angles (Shao-Bo, et al., 2008). One of the most effective ways of achieving identity theft is through “Social engineering”, it involves: An Identity theft criminal contacting the victim directly and convincing them to disclose passwords or other information by posing as agents representing an organisation or a particular individual (Shao-Bo et al., 2008). This is a common aspect in Kenya. We have had public notices made on the newspapers concerning employees who have left an organisation (22<sup>nd</sup> Sept Monday, DailyNation, 2014). The victims can also be targeted via the internet by the use of social engineering like the use of email messages, phone text messages, or intercepting and capturing financial or identity information while transaction is in progress. This trust can be exploited by identity thieves through various ways as indicated in this section to acquire the information they need (CIPPIC, 2007):

1. Pre-texting: This involves “smooth talking” the victim into trusting you and eventually persuading them to give off vital information
2. Obtaining credit reports: Identity thieves may pose as legitimate business people such as landlords, potential employer and used car dealers but with the interest of getting their credit reports.
3. Bogus Employment and Visa Processing Schemes: This is the most common way identity thieves use to get ones information especially in Kenya. Instances of people complaining on the assurance given when filling forms and copies of their documents they had given out only for the said to just vanish and the person completely dupes them and robs their personal information and some amount of cash too.
4. Through Contests and Surveys: unwitting victims may give off personal information while under the impression that they are joining up a contest or participating in surveys. This can happen through written submissions to contents or draws for prizes.

## **2.4 Signature Verification**

The necessity of ensuring that only the required persons get the required services, products and authorization has led to the automatic personal authentication. Biometrics on facial recognition, signature, handwriting, palm prints and voice are used to establish the identity of an individual. Signature verification is the most common form of verification when it comes to document verifications.

A handwritten signature is as a result of a rapid movement (Plamondon & Lorette, 1989). Thus the stroke features of the signature stays the same when written out on a periodic frame over time.

Signatures are made up of special characters therefore, most of the time they can be unreadable (Ozgunduz, et al., 2005). The verification process determines if the signature is a forgery while the process of recognition is used to find the identification of the signature owner. There exist various aspects, techniques and models for on-line and off-line signature verification.

### **2.4.1 Graph Matching**

It uses a multi-layer grammatical face model. It increases the robustness of recognition under varying lighting conditions. It has been used in face recognition fields where it applies a high-level semantic understanding of the face, enabling the users of the process to perform an intelligent recognition process driven by the status of the face i.e. change in expression and positions (Abuhaiba, 2007). Graph matching is used in other image similarity problems. The advantage of this model is the ability to accommodate spatial attributed relations and support supervised and unsupervised learning from training data (Abuhaiba, 2007).

### **2.4.2 Hidden Markov Model (HMM)**

It is a well-established probabilistic model essentially based on a notion of system state (Ehab, et al., 2010). There is a Markov chain modelling the system's transactions between a set of internal states, in which each state is connected with a specific probability allocation over the set of likely outcomes. The output of HMM is a sequence of outcomes where each outcome is sampled according to the probability distribution of the underlying state (Ehab, et al., 2010). Thus the HMM model is used for the learning and verification process (Justino & Yacoubi, 2000) The learning phase generates an HMM  $\lambda = \{A, B, \pi\}$  model that adequately characterizes each author's signature while the verification process is made up of a forward algorithm that determines the logarithm of the probability of the observed sequence (Justino & Yacoubi, 2000).



### **2.4.3 Neural Networks**

Neural networks have been a fundamental part of computerised pattern recognition tasks for more than half a century (Alan, et al., 2008). The main reasons they are applied is power and the ease of use. The neural networks (Alan, et al., 2008) firstly extract a feature set representing the signature from several samples of differing signers. Then the neural network learns the relationship between a signature and its class. Once this relation is learned, the network is then presented with test signatures that can be classified as belonging to a particular signer.

### **2.4.4 Binary Robust Invariant Scalable Keypoints**

Binary Robust Invariant Scalable Keypoint (BRISK) detector finds salient image regions such that they are repeatedly detected despite change of viewpoint making it robust to all possible image transformations. Thus tackling the classic computer vision problem of detecting, describing and matching image keypoints for cases without sufficient or prior knowledge of the position or camera pose (Leutenegger et al., 2003). The core phases in Robust Invariant Scalable keypoint (BRISK) are: characteristic recognition, descriptor composition and comprehensive keypoint harmonization. The modularity of this method enables the use of BRISK detector in combination with any other key-point descriptor and vice versa (Leutenegger, et al., 2003). It works by; applying the BRISK detector for characteristic recognition to estimates the true scale of each keypoint in the continuous scale-space where the scale-space pyramid layers consist of  $n$  octaves,  $c_i$  and  $n$  intra-octaves,  $d_i$  for  $i = \{0, 1, \dots, n-1\}$  and typically  $n=4$ . Given these set of key-points, it is then composed as a set of binary string by linking the results of the simple brightness comparison test in the descriptor composition phase and finally comprehensive key-point harmonization by simply computing their hamming distance (Calonder, et al., 2010) which is the number of bits differing in two descriptors makes it a measure of existing dissimilarity in recognition of the signatures.

## **2.5 Existing Identity Theft Mitigation Strategies**

The influence of ICT on society goes far beyond establishing basic information infrastructure. It has made the communication platforms more reachable to all worldwide. Emails have replaced the post letters, an online presence for a business, organization or institution has become critically important than printed publicity while mobile phones have become a growing norm in the society more than any other personal means on communication. With all this technology aspects in place, Identity theft has been given a fair playing field.

Applications such as e-government, e-commerce, e-education, e-health and e-environment have become enablers for development, as they provide efficient channel to deliver a wide range of basic services thus enabling developing countries to easily adapt to the digital era. In light of this, identity theft becomes a main threat to further deployment of e-government used in our legal systems and e-business used in our institutions.

Enhancing cyber security and protecting critical information infrastructures are essential to each nation's security and economic well-being (Gercke, 2012). Making the internet platforms safer has thus become a critical component to the development of any new services and government policies. Thus, they have created a global identity theft mitigation process; particular governments in various countries have come up with mitigation strategies for identity theft according to their specific requirements while specific organizations and institutions have mitigation strategies that are according to their exiting red flags.

## **2.6 Mitigation of Identity Theft**

Fighting Identity theft calls for actions from multiple parties. It requires the involvement of the technological, legal and law enforcement, economical and managerial solutions. The process of solving the identity theft problem involves the coordination of multiple parties. This is because the verification of an individual's identity engrosses multiple steps, methods and parties.

The most appropriate method for identity theft management would be the implementation of comprehensive and integrated identity fraud enterprise management algorithms (Jamieson et al., 2007). The designed stages of such algorithms should be functional, complementary, integrate knowledge innovations that fine-tune policy statements, processes and procedures. It should also be grounded from theories. In California, all companies are required by the Security Breach Information Act to inform customers when data has been breached or lost (Jamieson, et al., 2007). The implementation of identity fraud enterprise management framework provides significant economic benefits.

## 2.7 Theoretical/Conceptual Framework

This research incorporates the conceptual framework. It was viewed in this research as the best approach of the study. It outlined the research conception, prospects, philosophy and theories that support and inform the research questions by setting a distinct outline on the flow of the research.

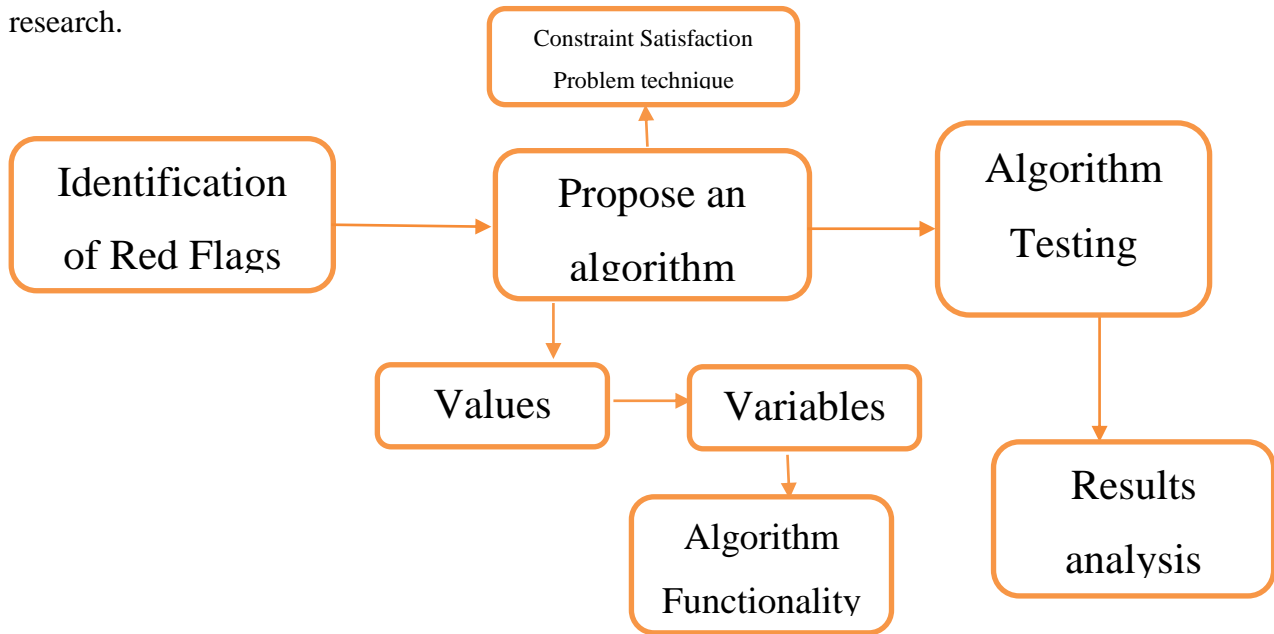


Figure 1: Conceptual Framework

## **CHAPTER THREE**

### **RESEARCH METHODOLOGY**

#### **3.1 Introduction**

In this chapter the research design, data collection methods, analysis and validation are discussed. The research design entails the chosen life cycle of the research in which it was ran, the data collection methods outlines the best of the existing methods that were used to gather the required information for the success of this research and the analysis method/tools that were applied to analyse the collected information. Finally validation of the data was established.

#### **3.2 Research Design**

For this study, the qualitative research design was applied to enable the collection of data facts and explaining of the phenomena more deeply and exhaustively. The study critically examined the knowledge of Identity theft among Kenyan Citizens and investigated the existing logical processes that are stipulated in the constitution for Identity theft reconstruction purposes.

Qualitative research design sequence and methods applied in the research are: Assessment of a problem statement, formulation of research questions, selection of a population sample, collection of data and analysis and lastly but not least, the presentation of the findings and conclusions (Mugenda & Mugenda, 1999).

This research design also entails giving empowerment to the persons been evaluated by having question that allow them to give their own opinions and voice out their ideologies or concern on the topic of discussion. This assures the respondents, thus making them more willing to participate in the undertakings. It also enriches the data collected making conclusions made unbiased and full informative thus producing almost accurate outcomes in the data analysis processes.

For the algorithm design, the Constraint Satisfaction Problems technique of developing an algorithm was implemented since the algorithm works with random generated key points from the scanned image. In Constraint Satisfaction Problems technique, the constraints are the key component in expressing a problem and are determined by the way the variables and the set of values are chosen (Bacchus, 2010). Constraints are a logical relation among a set of variables (Wagner & Urli, 2013) and they limit probable values that variables can obtain for example (Bacchus, 2010) all-different( $X_1, X_2, X_3$ ). This constraint says that  $X_1, X_2$  and  $X_3$  should acquire differing values thus with a set of values  $\{1,2,3\}$  created for each of the set should be  $X_1=1, X_2=2$  and  $X_3=3$  (Bacchus, 2010). It also symbolizes some partial information regarding

the variables of interest (Wagner & Urli, 2013). The Constraint Satisfaction Problems technique involves 3 components (Bacchus, 2010; Wagner & Urli, 2013):

1. A set of variables: Randomly generated key point variables from the scanned image
2. A set of values for each of the variables
3. A set of constraint restricting the values between various collections of variables.

An assignment of a value from its set to every variable satisfies all of the constraints. Constraint propagation enables forward checking which controls the future conflicts from the value set formed and enables earlier pruning as shown in figure 2 (Wagner & Urli, 2013).

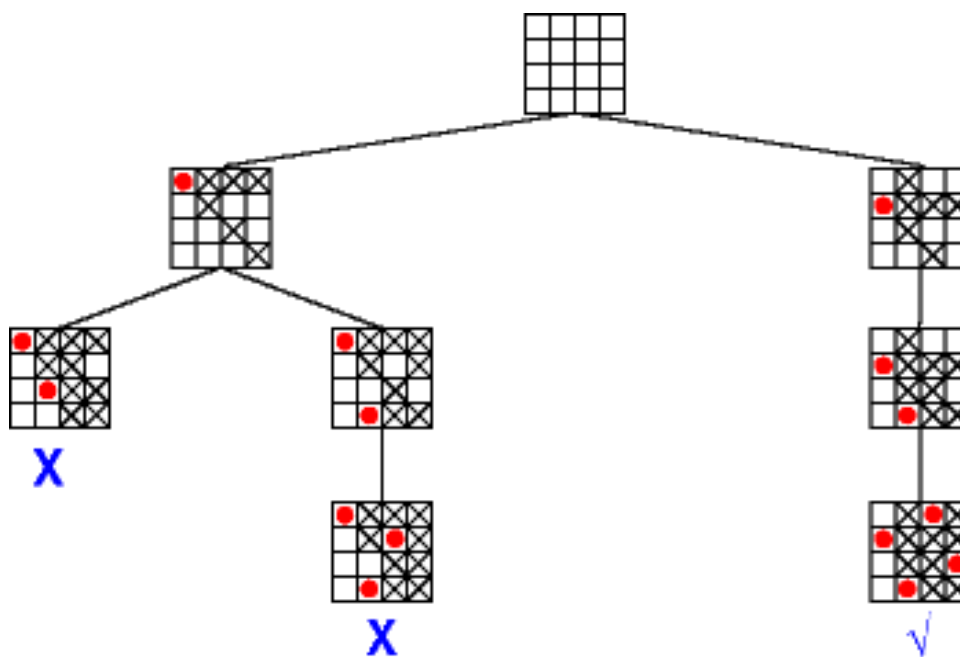


Figure 2: Constraint Propagation

### 3.4 Research Setting

The proposed county of research was Nairobi County.

### 3.5 Study Population

The population under study involved the members of four participating SACCOs and the staff members in these SACCOs as well.

### 3.6 Sampling

In this research, the data gathering process was guided by a sample of persons from various SACCOs. Convenience sampling provides the most convenient and cost effective method of selecting the study group (Battaglia, 2008). This is because it eases the cost of locating elements

of the population, the geographical distribution of the sample and obtaining the data from selected elements (Battaglia, 2008). It involved visiting SACCO offices or SACCO AGM meetings for those that agreed to participate in this research to seek participation of the individuals in this SACCOs. This ensured that the target sample of research was only members of a particular SACCOs. With this information a hypothesis was established on the knowledge and actions taken on identity theft in SACCOs.

### **3.7 Data Collection Method**

Since this research was more societal based, the experimental research best applied. This helped in the testing of the effectiveness of the prototype that that was developed from this research and how efficient it proved in dealing with identity theft crimes. The experimental research provided a method of investigation to derive basic relationships among phenomena under controlled conditions or to identify the conditions underlying the occurrence of a given phenomenon (Ross & Morrison, 2001).

### **3.8 Research Evaluation**

In this phase, the Evaluation method was applied so as to enable the measure of the True positives against the False positives and True Negatives against the False Negatives (Witten, et al., 2011). The true positives (TP) showed the number of negatives identified and true negatives (TN) are correct classifications. A false positive (FP) is when the outcome is incorrectly predicted as yes (or positive) when it is actually no (negative). A false negative (FN) is when the outcome is incorrectly predicted as a negative when it is actually positive (Witten, et al., 2011). Thus, the TP is where the signature accessed, is the original signature of the writer while the FP is the forged signature of a writer but it is viewed as an original due to the similarities in the stroke aspects of the signature. The TN is where the forged signature is identified through the algorithm verification procedure while the FN is where an original signature of a writer is termed as a forgery yet it is his/her actual signature and this is summarised in the table below (Witten, et al., 2011).

Actual Condition (Truth)			
(-ve) Forgery	(+ve) Geniune		
<b>FP</b>	<b>TP</b>	(+ve) Geniune	<b>Output of the system</b>
<b>TN</b>	<b>FN</b>	(-ve) Forgery	

Table 1: Evaluation Matrix

The Original Signature rate (Sensitivity):

$$\frac{TP}{TP+FN} \quad \text{Equation: 3.1}$$

The Forged Signature rate (Specifity):

$$\frac{TN}{FP+TN} \quad \text{Equation: 3.2}$$

Accuracy:

$$\left( \frac{TP + TN}{TP+TN+FP+FN} \right) 100\% \quad \text{Equation: 3.3}$$

Precision:

$$\left( \frac{TP}{TP+FP} \right) 100\% \quad \text{Equation: 3.4}$$

### 3.9 Research Validation

Validity is largely determined by the presence or absence of systematic error in data (Mugenda & Mugenda, 1999). In this case, we justify the results gathered and how effective they are to the new acquired information. The gathered data was theorised in the aim of confirming the gathered information, thus making it a continuing process of building assurance in the effectiveness of the acquired information.

### **3.10 Data Analysis**

Collected data was cleaned in order to determine incompleteness or unreasonable data and then improved the quality through correction of detected errors and omissions. Then data was entered for analysis using the Microsoft Excel package.

### **3.11 Human Expert Comparison**

Since the SACCOs' have loan officers who offer the loans to various individuals, it was necessary to compare the way in which this loan officers verify signatures and the way in which the systems perform the same duty. Thus the experiment involved the participation of four loan officers each examining the signatures used in the system as well.



## **CHAPTER FOUR**

### **SYSTEM DESIGN**

#### **4.1 Introduction**

In this chapter a presentation of the proposed prototype is made based on the understanding of the current signature verification process in the Financial Institutions specifically SACCOs through the data collected, processing and analysis within the SACCO as well as from the literature review.

System design is concerned with establishing how to deliver the functionality that was specified in analysis while at the same time, meeting non-functional requirements that may sometimes conflict with each other (Simon & Ray, 2002). This section focuses on making high-level decisions concerning the general structure of the system. In section 4.2 we look into the design of the prototype, Requirement specification of the proposed prototype 4.3, Developing the prototype 4.4, Program Development 4.5 and finally System Testing 4.6

#### **4.2 Design of the Proposed Prototype**

It is the gaps found in the literature review that motivated the design of a prototype for Signature verification. The particular aspect that this research addressed is the fraud that occurs as people seek to acquire a loan and why guarantorship is the security of the loan. Figure 3 below shows a Loan Management system that requires a guarantor for loan applications. This prototype closes this gap by ensuring that the guarantors' details are clearly confirmed in order to erase any unfriendly activity. An offline option was chosen for the system since the forms are still filled on paper thus making it quite efficient.



Figure 3 : A diagram of loan process

The screenshot shows a web-based Loan Management system. The user is logged in as "Abbas Harrell - 1880". The "Loan" tab is selected and circled in red. The interface displays various loan details and a table of loan history.

Loan No.	Loan Product	Short Description	Request Date	Effective Date	Repay Start Date	Requested Loan	Approved Loan	Loan Repayment	Interest Rate (%)	Interest Method
2090	BUS	Business Support	13/03/2011	01/03/2011	31/03/2011	600000.00	600000.00	12500.00	0.01000000	FRT
2081	BUS	Business Support	30/04/2010	01/11/2009	30/11/2009	120000.00	120000.00	10000.00	0.01000000	FLT
2063	DEV	Developmental	26/01/2009	01/12/2008	31/12/2008	60000.00	60000.00	5000.00	0.01000000	FLT
2062	BUS	Developmental	21/01/2009	01/12/2008	31/12/2008	2000.00	2000.00	66.70	0.05000000	FLT
1880	DEV	Loan	12/12/2008	20/12/2008	20/12/2008	139000.00	139000.00	3861.00	0.01000000	FLT

Figure 4: Loan Management system that requires a guarantor for loan applications

### 4.3 Requirement Specification of the Proposed Prototype

This is the serious part in design as it formally describes the requirements necessary to prevent incidence of uncertainties during the development process of a particular system. This section therefore presents the concerned participants with the assigned functional and non-functional requirements of the proposed system.

#### 4.3.1 Participants

One important aspect of any software development process is identification of the relevant participants as it helps shape the requirement process. This section therefore sought to identify the key participants of the proposed signature verification prototype. The identification process was guided by literature review knowledge in Chapter 2 and the Loan application system presented in figure 4 in section 4.2 above. The following are the major participants.

1. SACCO Members: These are the main stakeholders of the system. Without the members there would be no picking of loans. Thus there would be no need of verifying anyone since there would be no loan granted. Members request for a loan and are given forms to fill in for that purpose.
2. Loan Office/Lender Underwriter as indicated in the process: These are the persons that actually check the loan forms and ones eligibility of getting the loan. They are charged with the verification of the guarantors and are the ones that use the prototype created.

#### 4.3.2 Use Case Diagram

This section models the users' characteristics so as easily understand their roles and variations. It is used to show how the system participants enable the systems functionality.

##### 1. Use case Diagram for SACCO Member

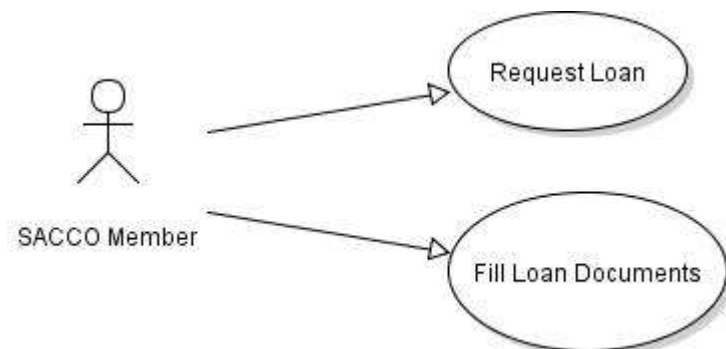


Figure 5: Use case for SACCO Member

## 2. Use case Diagram for the Loan Officer

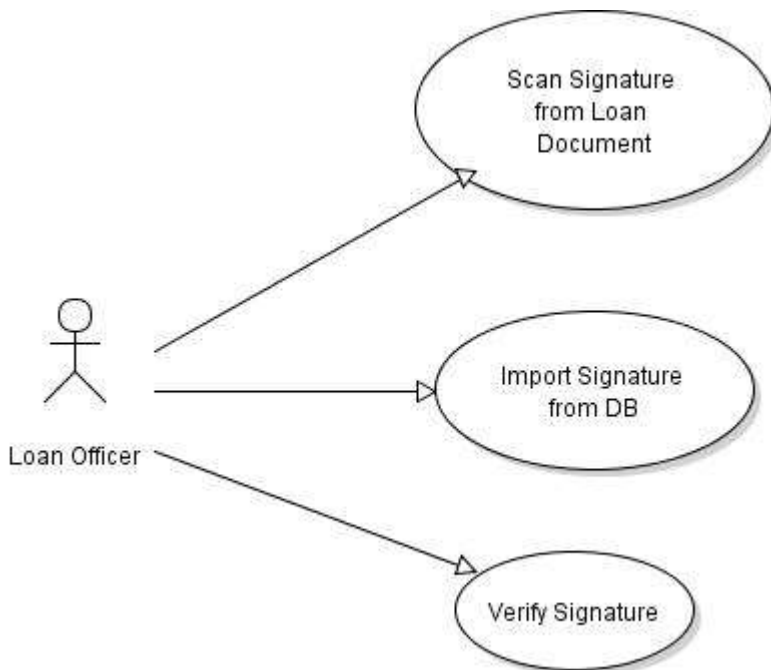


Figure 6: Use case for Loan Officer

### 4.3.3 Functional Requirements

These are defined as specific statements of service that define how a system responds to particular inputs as well as its behavior in given situations. This section therefore presents the statement of service which matches to the requirements analysis found in literature review.

**F1:** The database that provides the records of all the existing and new members of the SACCO

**Motivation:** The acquiring of the members original signature that is recorded when the member registers the very first time with the SACCO.

**F2:** The proposed signature verification prototype provides the ultimate process for ensuring that the guarantor does not become a victim of fraud.

**Motivation:** Usability and reliability to verify and produce output of the signatures as required.

#### **4.3.4 Non Functional Requirements**

These are system quality related statements illustrating the limitations on the services being rendered by the system. This section outlines these statements to help understand the development plan for the proposed prototype.

**Availability:** The prototype is offline thus works in all cases of application. The database can be picked from any server and my database can be used by the application.

**Usability:** The proposed signature verification prototype is easy to use and does not require any special computer skills

**Usefulness:** The proposed signature verification prototype aims to make signature verification easier and efficient.

#### **4.4 Developing the Prototype**

In this section, the development process is described on the functionality of the proposed prototype.

##### **4.4.1 Design Decision**

While designing the proposed prototype, a decision was made on the choice of suitable languages, database and development Tool kit for the given context. The following were applied:

1. Java was the language of choice for the coding of the interface and functionality implementation.
2. Net Beans was used to bring all the functionality of the prototype together.
3. SQLite was the choice of the Database. This was the fact that due to the time limitation we needed something that was server less unlike the jdbc:Derby

##### **4.4.2 Program Development**

The program development was done using a methodology called Joint Application Development (JAD). It was developed for designing a computer based system (Laign, n.d.). It dramatically shortens the life time of a project, improves the quality of the final product by focusing on the up-front portion of the development lifecycle thus eliminating the occurrence of errors that are expensive to correct later on (Laign, n.d.).

JAD was deemed an appropriate approach based on its suitable characteristics outlined below:

1. The involvement of the clients in the design and development of the project.
2. Its approach is used to lead to shorter development times and greater client satisfaction.

3. A series of interviews are conducted with the participants throughout to acquire the system requirements thus ensuring the application will be accepted after development due to their involvement.
4. It is very focused and conducted in a dedicated environment.

#### **4.4.3 Program Development Steps**

The system development process begun with development of an initial prototype based on the requirements previously gathered through interviews. The prototype was then introduced to four loan officers who were the target users of the project. The aim was to see how the prototype would be refined through participatory approach into a final product that met their needs through an iterative approach. The system underwent a series of three iterations with several recommendations being made for each iterative process in order for the system to work according to their requirements.

##### **1. Iteration 1:**

The users suggested the addition of two more file extensions instead of just one. The prototype would only allow a Jpeg file extension image while loading up the scanned image. This was because the Jpeg is viewed as the most common file extension. But, while scanning the saving options allow for more extensions options thus the .png and .bmp were included in the image file name extension filter.

##### **2. Iteration 2:**

After the incorporation of the recommendations in iteration 1 and presenting the system again, the users recommended the need to have a progress bar during the verification process. This would not leave the user wondering if it is processing or not.

#### **4.5 System Testing**

The users, a total of 4, were taken through a training process. They were specifically informed on the need to observe the image file extensions that the scanned images can be in. They filled in a review form indicating how they felt about the prototype and its use in a review sheet sampled in Appendix 4. One user stated:

*“What a useful tool this is. It is quite efficient and will help reduce the time taken in the comparison process.”*

## CHAPTER FIVE

### DATA PRESENTATION, ANALYSIS AND DISCUSSION

#### 5.1 Introduction

In this chapter, we presents the raw data, analyse it and present the findings of the study that focuses on signature forgery and discusses the proposed algorithm, its application and from this a prototype is created that is used in the implementation of the proposed algorithm. Data presentation and analysis includes the data from evaluation of identity theft tactics currently applied in the financial sectors during loan applications, how the red flags are established in order to point out forgery cases and the confidence levels of members that the SACCOs are indeed doing something to have a control of the occurring forgeries. To clearly present the findings, this chapter outlines the raw data demonstration, demographic data, quantitative data analysis and graphic statistical analysis to comprehensively respond to the research questions.

#### 5.2 Data Presentation

100 questionnaires were issued to willing participants from various SACCOs within Nairobi County, out of which 85 questionnaires were returned and 7 partially filled, making 87.6% collection. The demographic data of the respondents is presented in the Table below. The data shows that 50.59% of the respondents were females while 49.41% males, with all the respondents being below age 55. The number of years a member has been in the SACCO was less than 1 year 20%, 35.29% 1-9 years and 44.71% for 10-20 years.

Gender	Male		Female		Total
		42		43	
Age	Below 35 years		Between 35-55 years		Above 55 years
		47		38	
How long have you been in the SACCO	Less than 1 year	1-9 years	10-20 years	21-30 years	Above 30 years
	17	30	38	0	0

Table 2: Demographics Data

Data on picking a loan from a SACCO and what the guarantors would need in order to guarantee your loan is presented in the table below. The figures represent the responses given by the

respondents based on the Likert scale that was provided. For instance, when the respondents were asked what you would need to do in order to get a loan from your SACCO, 85 respondents said they would need to fill a form to get a loan while no respondent said they would apply for it online either via Mobile request or a website. This points out that in many SACCOs form filling is still the mode of loan application used. Other responses to the rest of the questions are as presented in table 3 below.

What would you need to do in order to get a loan from your SACCO	Fill a form	Apply online either via (Mobile request or a website)
	<b>85</b>	<b>0</b>
Have you ever applied for a loan	Yes	No
	<b>68</b>	<b>17</b>
How many guarantors would you need so as to get a loan	1-5	5 and above
	<b>53</b>	<b>15</b>
What would the guarantors need to do to guarantee your loan	Fill a guarantor's form	Sign your loan form
	<b>8</b>	<b>60</b>

Table 3: SACCO

Data on identifying the red flags was corrected through the guarantor's details falsification section and is presented in Table 4. The figures represent the responses given by the respondents based on the Likert scale that was provided. For instance, when the respondents were asked if they knew if guarantor signature forgery exists, 63 respondents said they have never known of its existence while 5 respondents said they actually have known of its existence. This points out that signature forgery is happening, but to many members it is unknown but to those that know, they know the impact of its existence. Other respondents to the rest of the questions are as presented in Table 4.

Do you know if guarantor signature forgery exists	Yes	No
	<b>63</b>	<b>5</b>
Have you ever had your signature	Yes	No



and details forged by someone you once guaranteed	60		3	
How did you know about it	During the loan application process	Through a call from the SACCO for confirmation	After the person defaulted on the loan	Didn't know until there was a deduction from your account Not Applicable
	0	10	30	20
Did the SACCO assist after raising the issue	Yes		No	
	60		0	

Table 4: Red Flags

Data on the steps the SACCO takes are presented in Table 5 and it shows the responses of the 60 loan application respondents on the 10 questions that sought to identify the natures of stagnation they experienced in their SACCOs. The responses were arranged into Likert scale from strongly agree (SA), agree (A), Disagree (D), Strongly Disagree (SD) and Not Applicable (NA). The figures indicate how many of the respondents supported strongly agree (SA), agree (A), Disagree (D), Strongly Disagree (SD) and Not Applicable (NA) on each of the 10 questions. For instance, in question one that asked if the SACCO called the guarantor's beforehand would make sure forgery does not happen, only 10 agreed to a very great extent, 48 agreed to a great extent while 5 disagreed and no respondents strongly disagreed indicating that something more need to be done to assure them security. This shows that a larger percentage of the SACCO members really need something done about it. The responses to other questions are outlined in the table 5 below.

Statements on Improvements and observations	SA	A	D	SD
I think if they called guarantors before giving the loan would help in making sure am protected	10	45	5	0

If I had knowledge of the impending action I would not have initially guaranteed him/her	60	0	0	0
I had total trust in my fellow member	60	0	0	0
I would have guaranteed the loan if he/she had approached me like before	50	10	0	0
I will not stop guaranteeing other follow members because of one or two untrust worthy persons	3	7	50	0
After I discovered I had my signature forged and informed the SACCO they went out of their way to assist me	33	13	14	0
The SACCO is doing enough to sensitize people on Guarantors details falsification	13	37	10	0
The SACCO has shown effort in managing these cases	10	50	0	0
You were confident that when you informed them of the issue they quickly understood what had taken place	33	13	14	0
The issue was resolved amicably	33	13	14	6

Table 1: Steps the SACCO has taken

### 5.3 Qualitative Data Analysis

#### 5.3.1 SACCOs Representation in Kenya

Having requested people from various places to fill in the questionnaires, various SACCOs were represented in this research. Though it would be great to use their actual names it may be better representing the number of SACCOs in a graph but number them. Convenience sampling was applied meaning data from different SACCOs was corrected and the number of responses that were received from a particular SACCO represented in data presentation. This representation is shown in figure 7 below.

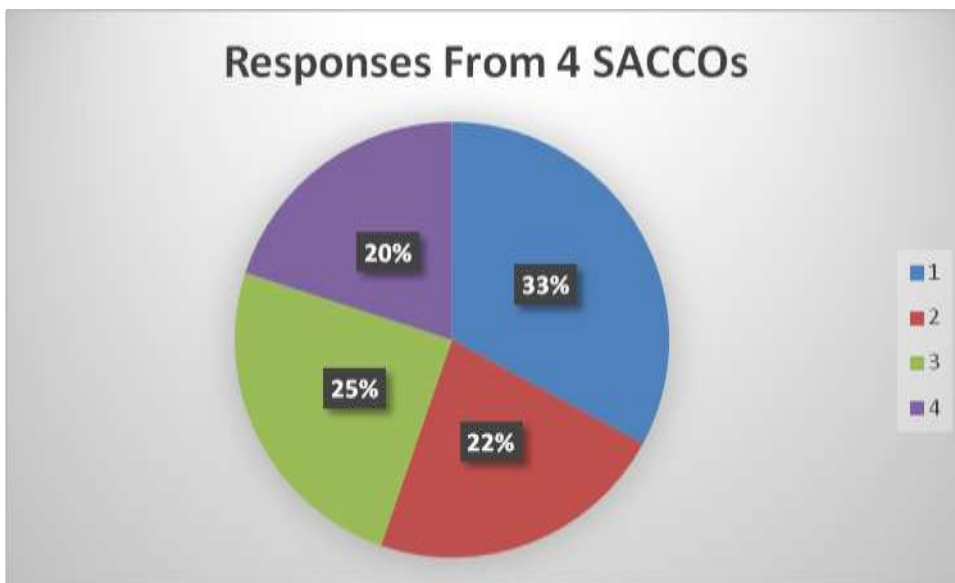


Figure 7: A representations of the responses from the SACCOs

#### 5.3.2 Loan Application Requirements

So as to find out the mode of loan application processes in SACCOs, questions that would enable the researcher have a clear vision of what the SACCOs actually need their members to do in order to get a loan were applied in the questionnaire. Unlike some banks that have turned to mobile banking, SACCOs have not yet adapted this mode during loan applications but have become flexible during repayment where they have adapted exiting mobile payment modes. The figures 8 and 9 below represent the Mode of loan application and how many guarantors are required for loan application respectively. From these results the main process of loan application was through form filling while in other loan applications at least 1 guarantor was required.

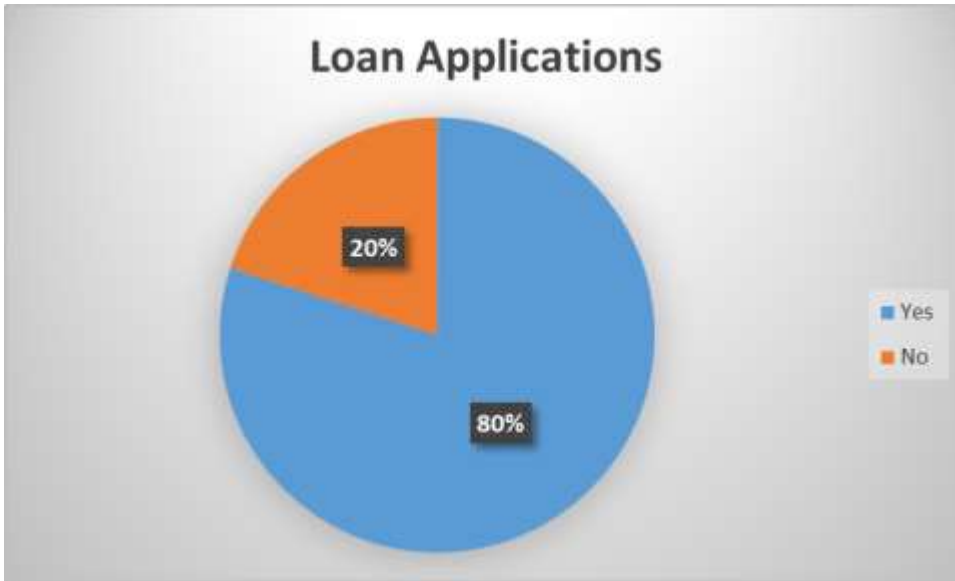


Figure 8: Mode of loan application

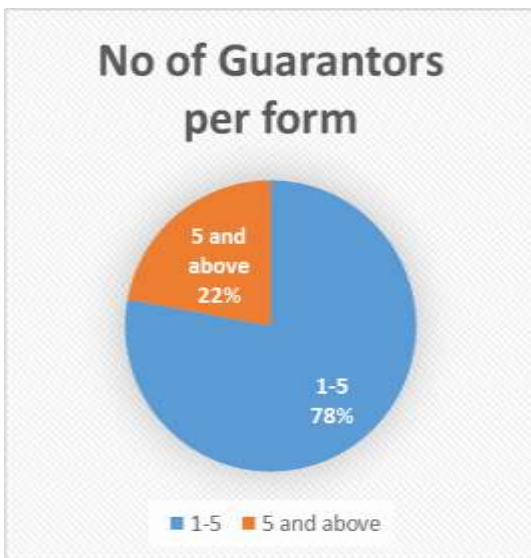


Figure 9: How many guarantors are required

### 5.3.3 Guarantor's Requirements

To find out what was needed from the guarantors; a question on if they would either require to fill a separate form or sign a loan application form was used. The common answer provided was that the guarantor would need to sign the actual loan application form. This would mean he/she would have to give details of his membership and also a signature on the applicant's form thus giving away his/her critical information that can be used again if the applicant required to apply again. This is represented in figure 10 below.



Figure 10: Guarantor's requirement

### 5.3.4 Forgery Details

In this section of the research questionnaire, the aim was to investigate if members knew that forgery existed, if any of the respondents have encountered an instance where they have had their data forged and at what stage of the loan application was it discovered. The figures 11, 12 and 13 below represent this information. The research indicated that 93% of the 68 respondents knew that forgery of one's information existed out of which that 95% of the 63 respondents have actually had their information used against them and with 50% of them knowing it only after the person had defaulted the loan and money was already being deducted from their salaries/accounts.



Figure 11: Forgery Existence

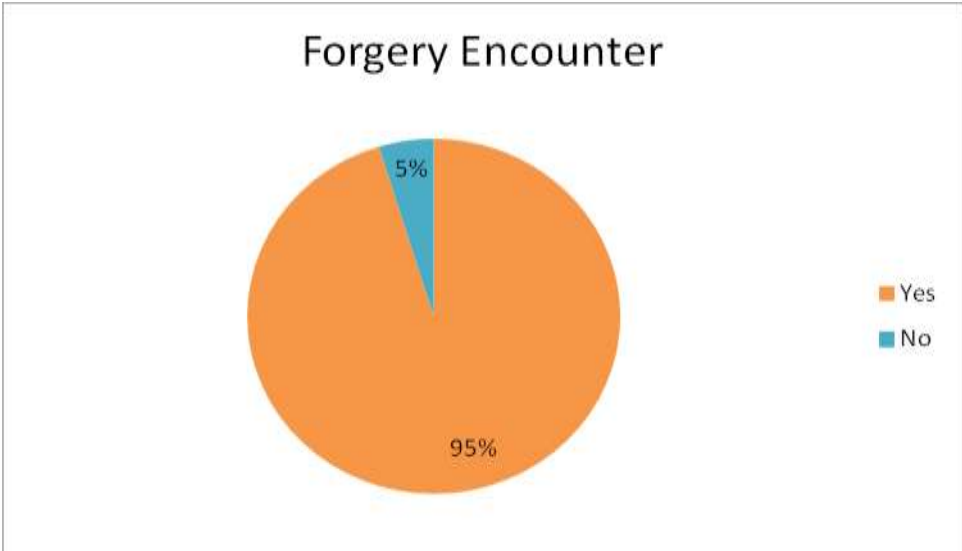


Figure 12: How many have had their details forged

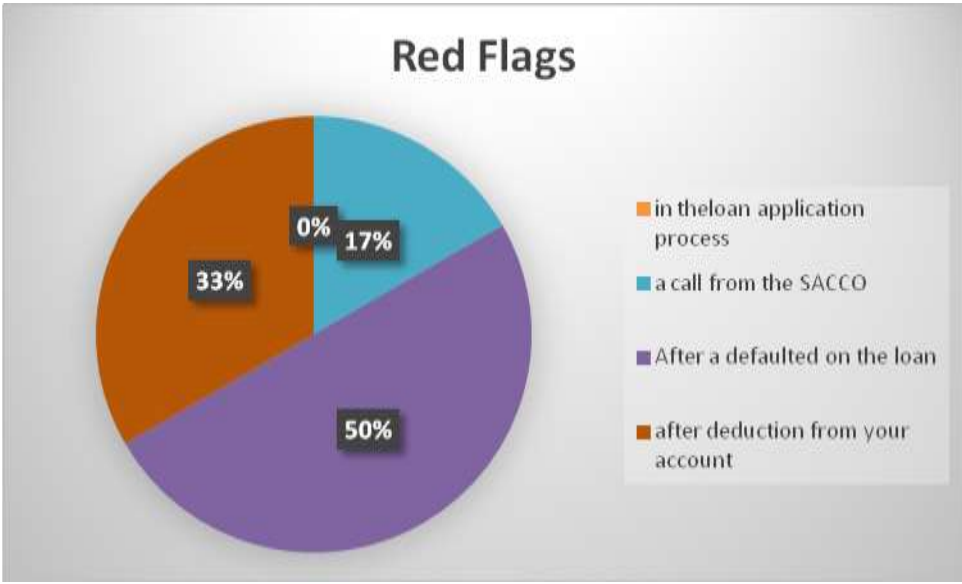


Figure 13: Red Flags

## **5.4 Image Matching Process**

There exists different images that are to be processed and values attained. Thus the following matching process established by Guo Li and et al. (Li, et al., 2010) is used and it is presented in the figure 14.

### **5.4.1 Vector Rasterization**

The image is described in vector graphics form thus in this stage it is changed to a raster form so that it is read as a bitmap file format.

### **5.4.2 Ascertain the Reference Object**

The image inputs are referenced to check for the critical points required to do a similarity value calculation.

### **5.4.3 Similarity Value Calculation**

The Euler number calculation is adopted so as to establish values of the objects which have been identified during Ascertain the reference object stage.

### **5.4.4 Preliminary Matching**

If the similarity values of the objects are within the scope of threshold, the preliminary conclusion are thus established which will imply that the images are homograph entities.

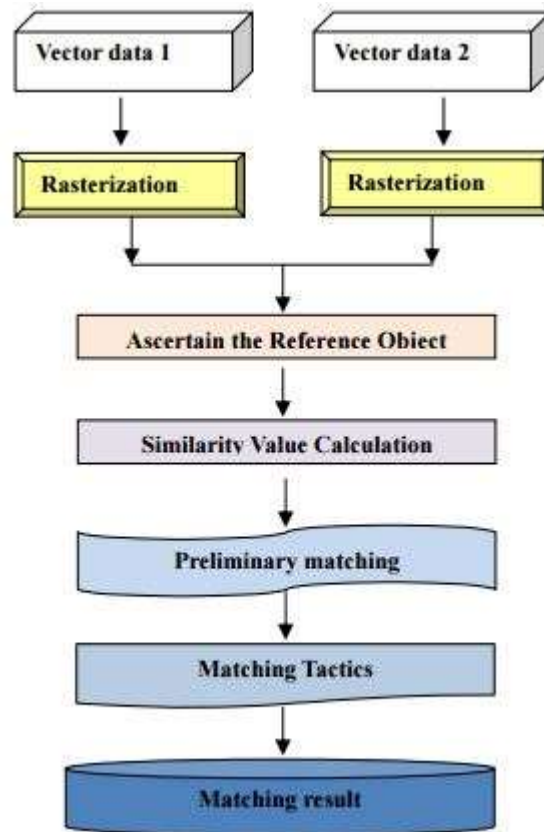


Figure 14: Matching Process

### 5.5 Proposed Algorithm

The algorithm works for both the database image and the scanned image. This is because features are extracted from both images (173x116 px) and then analysed and matched. The algorithm can be summarized as follows:

1. Signature feature extraction: Key point, these points are extracted by scanning the image for pixel continuous forming a line or a curve, from the scanned image and database image and from this, the critical points are created.
2. Critical points are selected within a rectangular space. From this area a random selection of points is done where the signature has unique characteristics. This can be a corner or a full-stop depending on how your signature appears.
3. From the vectors collected in the plots of the two images  $X_1$  and  $X_2$   $\{a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8, a_9 \dots a_n\}$  and  $\{b_1, b_2, b_3, b_4, b_5, b_6, b_7, b_8, b_9 \dots b_n\}$  a two dimensional matrix is created from it.  $(a_1, a_2, a_3, a_4)$  and  $(a_5, a_6, a_7, a_8)$ .



4. Matching of the signature (Vatsa M., 2004), SCE based matching using the Directional Difference Matching (DDM). Construction of the comparison matrix is done with its components as binary numbers. The matrix comparison of input SCE from the scanned image and SCE from the database image is done. The following equation for comparison is applied (Vatsa M., 2004):

$$[X_1, X_2] = Y \pm \varepsilon$$

Y in this instance represents the Euler number/Different from input Euler code and  $\varepsilon$  is the tolerant error. The value of X lies between  $X_1$  and  $X_2$  then it is indicated as 1 in the comparison matrix otherwise a 0 is adopted. To match the signatures, the numbers of 1's and 0's in the matrix are counted. The 1's refer to a match while the 0's refer to a mismatch.

## 5.6 Pseudo Code

1. Original and test image inputs  $w=173, h=116$
2. Set Pixel array total number of elements =  $w*h$
2. Input X, Y
3. Plot X,Y in the window=  $X+(Y*width)$
4. Location in 1d pixel array: Location =  $X+ (Y*img.w)$
5. Create new pixels entries from the neighbour on the left side of the image:  $getpixels ((x-1) + y*img.width)$
6. Matrix: Plot four symmetric pixels
  - a.  $Setpixels(x-1,y)$
  - b.  $Setpixels(x+1,y+1)$
  - c.  $Setpixels(Round(x_c+x), Round(y_c-y),1)$
  - d.  $Setpixels(Round(x_c-x), Round(y_c+y),1)$
7. Checks if matrix exceeds 4. If so then stop. Else repeat steps 4-7
8. Compute a matrix differentiation from the two matrixes to get Y  
 $[i]-[j]= Y \pm \epsilon$
9. Set verification limit to 1-10.
10. Return verification status

## 5.7 Discussion

### 5.7.1 Introduction

The results are made in comparison with the human expertise on signature verification using all the processes explained in section 4.4. A total of 532 signatures were used. 66 of which were database signatures and while 466 signatures were test signatures. For each class of know signatures contained both a training and a test signature. The overall performance of the signature verification process was measured in terms of accuracy in which it would determine the genuine or forged signature in a particular test.

Figure 15 and Figure 16 below are original signatures from two members with member number 3001 and 3005 respectively from the database.

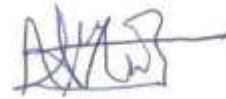


Figure 15: Member no 3001's Signature

Figure 16: Member no 3005's Signature

### 5.7.2 Experiment Results

In this section a sample of the tested signatures are presented. Signatures from the same known writer were tested against each other as shown in Table 6 below. The critical points picked had a threshold that did not exceed the set value. Then in Table 7, an unknown signature was used to test its originality. It was classified as a forgery signature as the results are shown below. From this experiment the test signature used can be classified as a skilled forgery because from a glance one would say it may be one and the same as the owners.







Signatures	Critical Points	Time (Nano)	Verification
  13.jpg   14.jpg	Set1(12,13)	1678	Signatures match
  13.jpg   15.jpg	Set2(12,12)	1666	Signatures match
  14.jpg   15.jpg	Set3(13,12)	1550	Signatures match

Table 6: Point variants of known signatures







Signatures	Critical Points	Time (Nano)	Verification
    13.jpg   16.jpg	Set1(12,24)	1389	Signature Mismatch
    14.jpg   16.jpg	Set2(13,25)	1354	Signature Mismatch
    15.jpg   16.jpg	Set3(14,25)	1411	Signature Mismatch

Table 7: Point variants of test signature 16.jpg and known signatures

The next experiment shows signatures in Table 8 which are from the same know writer. And as in the experiment above we test them to verify them against each other. Then in Table 9 the known signatures in Table 8 were tested against an unknown signature and from the results it was determined as a forgery.







Signatures	Critical Points	Time (Nano)	Verification
    17.jpg   19.jpg	Set1(63,60)	1469	Signature Match
    18.jpg   19.jpg	Set2(65,60)	1472	Signature Match
    17.jpg   18.jpg	Set3(63,65)	1464	Signature Match

Table 8: Point variants of known signatures







Signatures	Critical Points	Time (Nano)	Verification
    17.jpg   20.jpg	Set1(63,43)	1657	Signature Mismatch
    18.jpg   20.jpg	Set2(65,43)	1645	Signature Mismatch
    19.jpg   20.jpg	Set3(60,43)	1650	Signature Mismatch

Table 9: Point variants of test signature 20.jpg and known signatures

The next experiment shows signatures in Table 10 which are from the same know writer. And as in the experiment above we test them to verify them against each other. Then in Table 11 the known signatures in Table 10 were tested against an unknown signature and from the results it was determined as a forgery.







Signatures	Critical Points	Time (Nano)	Verification
    21.jpg   23.jpg	Set1(29,30)	1469	Signature Match
    22.jpg   23.jpg	Set2(30,30)	1472	Signature Match
    21.jpg   22.jpg	Set3(29,30)	1464	Signature Match

Table 10: Point variants of known signatures







Signatures	Critical Points	Time (Nano)	Verification
    21.jpg   24.jpg	Set1(29,44)	1264	Signature Mismatch
    22.jpg   24.jpg	Set2(28,44)	1260	Signature Mismatch
    23.jpg   24.jpg	Set3(30,44)	1262	Signature Mismatch

Table 11: Point variants of test signature 24.jpg and known signatures

### 5.7.3 Discussion of results

In this section the proposed evaluation techniques of the signature verification are applied. The matrix on the true positive, true negative, false positive and false negative is established in order to establish the preciseness at which the forged signatures are determined.

Out of the 532 signatures the prototype achieved the following results: The percentage of accuracy in this test was 91.4% while the percentage of precision was 60.0%.

<b>TP</b>	11.3%	<b>FP</b>	7.5%
<b>TN</b>	80.1%	<b>FN</b>	1.1%

Table 12: Results from Prototype

There were 66 members that were recorded in the database and each member was asked to provide ten of her/his signature. Thus in total there were 66 original signatures that were recorded in the database which thus were termed as the actual member's signature while 466 were the actual test signatures. Then as earlier stated 532 signatures were used as the test and training signatures. Out of these signatures 80.1% of the signatures were clearly identified as forgeries while 11.3% out of the test signatures were identified as originals. From this test it indicated that a signature may vary depending on the pen used and the circumstance while signing among others.

### 5.7.4 Comparison with Human Expert

The human experts from four different SACCOs were asked to have a look at the sets of signatures and compare them to the original signatures. They were shown the original signatures with which they were to compare with the 532 test signatures. The results are as presented in the table below:

Human Expert	TP	TN	FP	FN
1	28.9%	4.7%	61.7%	4.7%
2	30.3%	8.3%	56.9%	4.5%
3	30.6%	6.0%	59.0%	4.3%
4	34.6%	7.7%	53.2%	4.5%

Table 13: Results from Human Expert

As this experiment with the human experts was being conducted some factors were indicated as situations at which one may not even check the signature as long as the guarantor is a member and there is an imprint at the signature section. Also the fact that we are human there may be something that is over looked in the process. In comparison to the results, the signatures verified by the human eye had more forgeries compared to the test signatures totalled. The signatures contained 466 test signatures while the other 66 signatures were the copies of the same signatures in the database. Thus from the results above the human experts indicated a very high percentage: Human Expert 1 had 61.7% of the false positive signatures (forgeries termed as originals) and a low on the true negatives (forgeries), Human Expert 1 had 4.7% while as indicated by the prototype above the vice versa should have been the case. Thus the percentage comparison of the human experts and the results from the prototype were as follows:

<b>Comparison</b>				
	<b>TP</b>	<b>TN</b>	<b>FP</b>	<b>FN</b>
<b>Human Expert 1</b>	28.9%	4.7%	61.7%	4.7%
<b>Human expert 2</b>	30.3%	8.3%	56.9%	4.5%
<b>Human expert 3</b>	30.6%	6.0%	59.0%	4.3%
<b>Human expert 4</b>	34.6%	7.7%	53.2%	4.5%
<b>Prototype</b>	11.3%	80.1%	7.5%	1.1%

Table 14: Comparison of Human Expert versus the Prototype

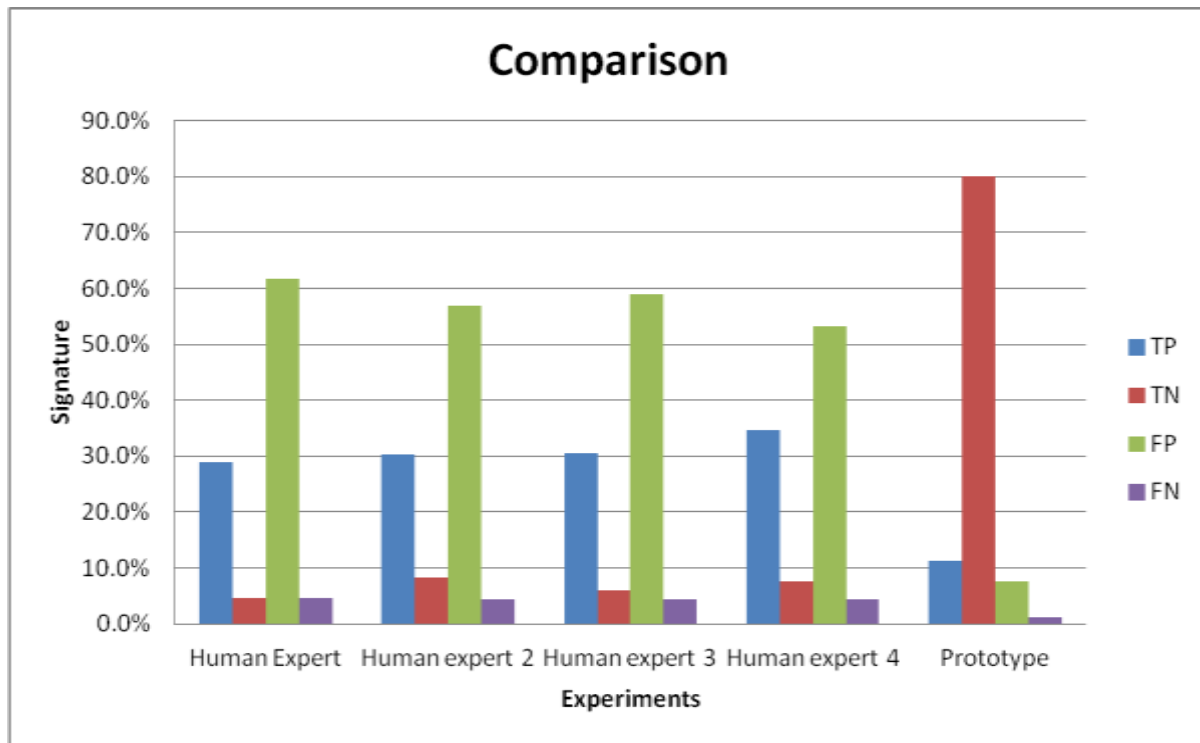


Figure 17: Bar Chart Indicating Comparison

### 5.7.5 Evaluation of the System

Evaluation of the system performance was done through statistical analysis of experimented results and compared to the human expert experiment. Statistical results in terms of precision (3.4) and accuracy (3.3) were calculated.

Accuracy:

$$\left( \frac{TP + TN}{TP + TN + FP + FN} \right) 100\% \quad \text{Equation: 3.3}$$

Precision:

$$\left( \frac{TP}{TP + FP} \right) 100\% \quad \text{Equation: 3.4}$$



Experiments	TP	TN	FP	FN	Accuracy	Precision
<b>Prototype</b>	60	426	40	6	91.4%	60.0%
<b>Human Expert 1</b>	139	12	315	12	31.6%	30.6%
<b>Human Expert 2</b>	147	30	289	10	37.2%	33.7%
<b>Human Expert 3</b>	148	18	300	12	34.7%	33.0%
<b>Human Expert 4</b>	168	28	270	13	40.9%	38.4%

Table 15: Prototype Evaluation

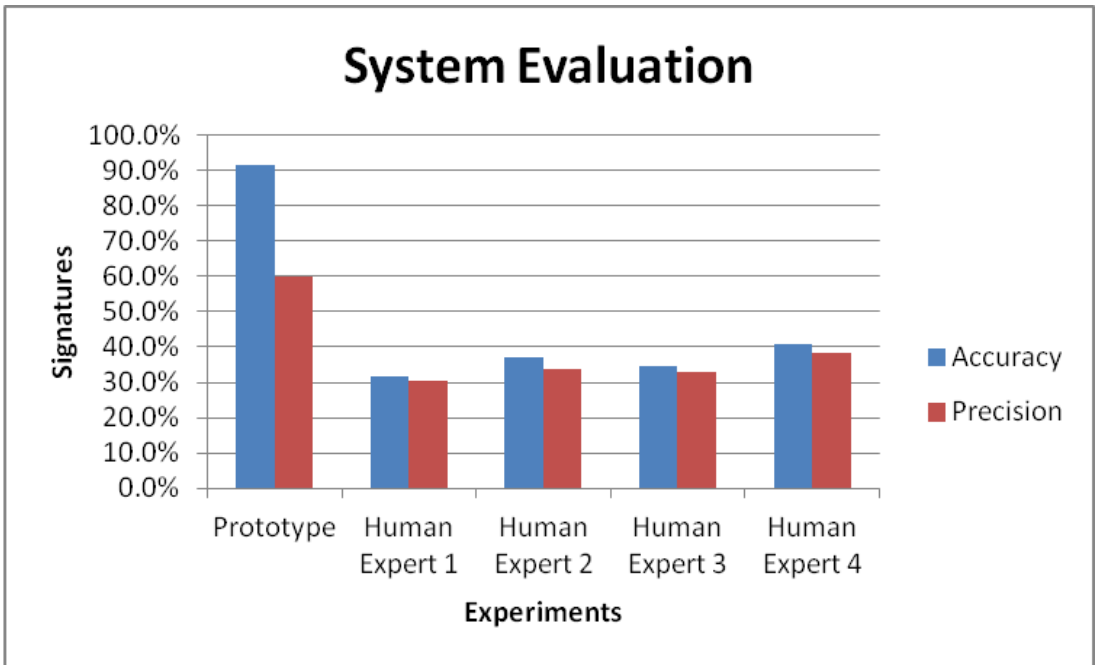


Figure 18: Bar Chart System Evaluation

### 5.7.5.1 False Acceptance Ratio

This is the evaluation of the probability that the system may erroneously accept a forged signature as an original signature. It is calculated by dividing the number of false signature acceptances (FP) by the number of total test signatures (Thakkar, 2016).

$$\left[ \frac{FP}{\text{Total test signatures}} \right] \times 100 \quad \text{Equation: 3.5}$$

Thus making the percentage FAR of the system 7.5%

### 5.7.5.2 False Rejection Ratio

This is the evaluation of the probability that the system may erroneously reject the original signatures and mark them as forged signatures. It is calculated by dividing the number of false rejections (FN) by number of identified know original signatures (Thakkar, 2016).

$$\left[ \frac{\text{FN}}{\text{Number of Known |Original signatures}} \right] \times 100 \quad \text{Equation: 3.6}$$

Thus making the percentage FRR of the system 9.1%

Thakkar further indicates that for a good working system it should have a low FAR and a high FRR which would thus ensure that no unauthorized transactions are allowed (Thakkar, 2016).

This also implies that some several tests would need to be redone.

## **CHAPTER SIX**

### **FINDINGS, CONCLUSION AND FURTHER RESEARCH**

#### **6.1 Summary of Findings of the Research Questions**

##### **Question 1: How have the Financial Sectors been Handling the Identity Theft Cases?**

From the study, the SACCOs used in the study as a sample of the financial sectors in Kenya, indicated that identity theft is not noted until late in the loan process which is after the person granted the loan has defaulted it.

##### **Question 2: Occurrence Rate of these Cases?**

When a qualitative analysis that was done, the occurrence rate was established at 95% of the total respondents which is also the total number of members that had encountered forgery of their signatures while only 5% had not encountered it but knew of its existences.

##### **Question 3: How does the Identity Thief Get the Necessary Individual's Information?**

From the qualitative analysis 80% of the persons who participated indicated that forms are supposed to be filled in order to make a loan application.

##### **Question 4: Does the Algorithm Enable Quick and Easy Detection of the Forgery?**

With accuracy achievement of 91.4% and a precision of 60.0% the system in the test environment did well compared to the human experts that are usually assigned to grant loans to persons that have made applications.

##### **Question 5: Does the Use of Signature Verification Improve the Security of the Loan Process?**

From the study, the materialization of the signature verification application enhanced the impact on quality of identity theft mitigation. This arises from the qualitative data collected from participants given that most had experienced forgery of their signatures. However, from the research it was noted that more needs to be done to shift everything to the digital platform. Thus, for a greater difference to be manifested, starting to use an online loan application system will enable better and faster monitoring. Quantitative results that were conducted show that the prototype accuracy and precision as compared to the current implemented Human expert mode of verification is more advocated for.

##### **Question 6: What is the Usefulness of the Signature Verification in Identity Theft Mitigation?**

Two main pointers arose from the study namely: enhancing online applications for loans and the efficiency of real time programming. Utilization of the signature verification application as an

encouragement tool in the improvement of security and the building of trust in the clients also emerged. This was noted by a user as an easier means of verifying the signatures from all the users that used the system unlike when verifying it with just the normal eye.

## **6.2 Comparison with other Existing Applications**

Signature verification algorithm was based on static and dynamic features of online signature data (Vatsa M., 2004). The texture and topological features are extracted from the signature image while a digital table captures in real-time the pressure values, breakpoints and the time taken to create a signature. (Vatsa M., 2004) Euler numbers were used to analyze the textural and topological features of the signature. With this application they system had an accuracy of 98.18%. Patil G. et al classification of offline handwritten signature applied wavelets and a pattern recognition neural network (Patil & Hegadi, 2014). The system implemented Discrete Daubechies Wavelet transform to extract wavelet coefficients in three directions namely horizontal, vertical, diagonal and a pattern recognition neural network classifier is designed where the training algorithm is a Quasi-Newton algorithm and the classification is done (Patil & Hegadi, 2014). Their system had a false acceptance rate quite high and indicated that some modifications in the form of increasing the number of hidden layers and their nodes along with the more efficient training algorithms are highly desirable and have the potential of better accuracies (Patil & Hegadi, 2014). Anand et al. (2014) enhanced signature verification and recognition using Matlab. They indicate feature extraction as a main necessity for successful results in a system. They implement neural network for the verification of signatures. Various features are used in the creation of the feature sets used, they include: eccentricity, skewness, solidity, entropy, euler number just but to mention a few. For this project euler number application to acquire the features from the scanned signature images was applied and a handwriting analysis done for the line or curve in the images that was then compared to the data collected from the images of the known signatures stored in the database. Its accuracy was 91.4% and precision at 60.0%.

## **6.3 Contribution to Previous Work**

Previous studies on the use of signature verification like Jarad et al. (2014), offline handwritten signature verification system using a supervised neural network approach aims at limiting the computer singularity in deciding whether the signatures are forged or not but inturn allows the signature verification personnel to participate in the decision making process by adding a label which indicates the amount of similarity between the signatures that are been analysed (Jarad, et

al., 2014). Other works suggest the use of Associative Memory Net (Dash, et al., 2012) and Contourlet transform (Pourshahabi, et al., 2009).

This research proposes an image based verification application that implements the measure of the boundary from which a point is calculated and plotted on the image then a rectangle plotting done from each pixel, values are vectorised and then a matrix which gathers points from these vectors is developed in order to aid in the comparison using matrix comparison methods.

## **6.4 Conclusion**

The general objective of this research was to analyse features on a signature that are imprinted on loan forms by the guarantors and propose an algorithm and develop a prototype that enables the verification of this signatures, therefore, enhance efficiency in the handling of the identity theft cases. The algorithm proposed used the Euler number as the mode of attaining the key-points from the 2d image of the signature that were retrieved from the database and scanned from the forms. The evaluation of the identity tactics and establishment of red flags was conducted through the filling of questionnaires and this helped in actualising the research's necessity. From the project Signature verification with the application of the algorithm was termed more accurate and timely unlike when conducted by a human expert as shown in the results above. The human expert eye sought to outline the most visible set of strokes that may seem to look similar in one way or the other while for the system sought out wholes found in the image pixels and used this to determine the signatures validity. Due to the pressure one uses to imprint their signature on a paper the ink placement largely varies.

## **6.5 Further Research**

This research focused on the offline version of input since the forms are still filled manually. An online version of the system can be created in which one can sign on a tablet and the signed signature is registered and verified real time through the online process. If the manual form is not altered, the scanning of the signatures on the forms can be made easier by having it done via a scanning mobile application. For the algorithm, since the image is a scan, at times it may appear dirty, crumbled or blurry depending on the scan done, thus a rasterization may be necessary in order to map it from picture geometry unto pixels, this will not entail much since the algorithm applied does not impose a specific way to work out the colour of those pixels and neither would rasterization require one to do so.

## REFERENCES

- ABUHAIBA, I. S. I., 2007. Offline Signature Verification Using Graph Matching. *Turk J Elec Engine.*, 15(1), pp. 89-104.
- Alan, M., Jarrod, T. & Wayne, R., 2008. Neural Network-based Handwritten Signature. *JOURNAL OF COMPUTERS*, 3(8), pp. 9-22.
- Anand, H. & D.L, B., 2014. Enhanced signature verification and recognition using Matlab. *International Journal of Innovative research in Advanced Engineering (IJIRAE)*, 1(4), pp. 2349-2163.
- Bacchus, F., 2010. *Computer Lecture notes*. [Online] Available at: <http://www.cs.toronto.edu/~fbacchus/Presentations/CSP-BasicIntro.pdf> [Accessed 16 July 2015].
- Barske, D., Stander, A. & Jordaan, J., 2010. A Digital Forensic Readiness framework for South African SME's. *Information Security for South Africa (ISSA)*, (DOI) 10.1109(ISSA.2010.5588281).
- Battaglia, M. P., 2008. Nonprobability Sampling. In: *Encyclopedia of survey Research methods*. s.l.:SAGE Publications, pp. 523-526.
- Billig, J., Danilchenko, Y. & Frank, C. E., 2008. *Evaluation of Google Hacking*. Kennesaw, InfoSecCD Conference'08.
- Calonder, M., Lepetit, V., Strecha, C. & Fua., P., 2010. *BRIEF: Binary Robust Independent Elementary Features*. s.l., The European Conference on Computer Vision.
- CIPPIC, 2007. *Identity Theft: Introduction and Background*, Ottawa: CIPPIC Working Paper No. 1.
- CIPPIC, 2007. *Techniques of Identity Theft*, Ottawa: CIPPIC Working Paper No.2.

Dash, T., Nayak, T. & Chattopadhyay, S., 2012. Offline Handwritten Signature Verification using Associative Memory Net. *International Journal of Advanced Research in computer Engineering & Technology*, 1(4).

Ehab, E., Vladimiro, S. & Mogens, N., 2010. HMM-based Trust Model. In: J. D. G. Pierpaolo Degano, ed. *Formal Aspects in Security and Trust*. Netherlands: Springer Berlin Heidelberg, pp. 21-35.

Gercke, M., 2012. *Understanding cybercrime phenomena, challenges and legal response*, s.l.: ITU Telecommunication Development Sector.

Graeme, N. R. & McNally, M. M., 2005. *Identity theft Literature Review*, s.l.: U.S. Department of Justice.

Hoar, S. B., 2001. Identity Theft: The Crime of the new Millennium. *HeinOnline journals*, 1423(80).

Identity-Theft-Scenarios.com, 2015. *History of identity Theft*, s.l.: Identity-Theft-Scenarios.com.

Jamieson, R. J., Winchester, D. W. & Smith, S., 2007. *Development of a Conceptual Framework for Managing Identity Fraud*. Hawaii, IEEE Computer Society.

Jarad, M., Al-Najdawi, N. & Tedmori, S., 2014. Offline handwritten signature Verification System using a Supervised Neural Network approach. *IEEE Computer Society*, 6(ISBN:987-1-4799-3999-2), pp. 189-195.

Justino, E. J. R. & Yacoubi, A. E., 2000. An Off-Line Signature Verification System Using HMM and Graphometric Features. *2000,4th IAPR International on Document Analysis Systems*.

Karounia, A., Dayab, B. & Bahlakb, S., 2010. Offline signature recognition using neural networks approach. *Elsevier Ltd*, Volume Procedia Computer Science 3, pp. 155-161.

Koppenhaver, K. M., 2007. History of Forgery, Forensic Document examination. *Humana Press*.

Laign, S., n.d. *Rapid Applicaton Development*. Southern California: CS 470 Fall I.

Leutenegger, S., Chli, M. & Siegwart, R. Y., 2003. BRISK: Binary Robust Invariant Scalable Keypoints. *ETH Zurich*, pp. 1-8.

Li, G., Zhiping, L., Bin, Z. & Yaoge, W., 2010. The study for Matching Algorithms and Matching Tactics about area Vector Data Based on Spatial Directional Similarity. *The Joint International Conference on Theory, Data Handling and Modelling in GeoSpatial Information Science*, 38(II), p. 395.

Maxwell, M., 2012. *Police warn over card skimming syndicate*, Nairobi: Star newspaper.

22<sup>nd</sup> September Monday, DailyNation, 2014. *Public Notice Adverts*. Nairobi: Monday Daily Nation.

Mugenda, O. M. & Mugenda, A. G., 1999. *Research Methods: Quantitative and Qualitative Approaches*. 2nd ed. s.l.:Acts Press.

Mwangi, E. K., 2008. *Offline Handwritten signature verification using SIFT Features*, Kampala: Academia.edu publishing.

Ogla, A., 2007. *ID Theft: A Computer Forensics' Investigation Framework*. Australia, 5th Australian Digital Forensics Conference Paper.

Oppliger, R. & Gajek, S., 2005. Effective Protection Against Phishing and Web Spoofing. In: J. Dittmann, S. Katzenbeisser & A. U. (, eds. s.l.:IFIP International Federation for Information Processing, pp. 32-41.

Ozgunduz, E., Karsligil, E. & Senturk, T., 2005. *Off-line Signature Verification and Recognition by Support Vector Machine*. s.l., European Signal processing Conference.

Patil, P. G. & Hegadi, R. S., 2014. Classification of offline Handwritten Signatures using Wavelets and a Pattern Recognition Neural Network. *International journal of Computer Applications*, Volume Recent Advances in Information Technology, pp. 0975-8887.

Plamondon, R. & Lorette, G., 1989. Automatic signature verification and writer identification—the state of the art. *Pattern Recognition*, Volume 22, p. 107–131.

Post, A., 2003. *The Dangers of Spyware*, s.l.: Symantec Security Response.



Pourshahabi, M. R., Sigari, M. H. & Pourreza, H. R., 2009. Offline Handwritten signature identification and verification using contourlet transform. *IEEE Computer Society*, 2(9), pp. 670-673.

Ross, S. M. & Morrison, G. R., 2001. *Experimental Research Methods*. s.l., Association Educational Communication and technology .

SecurityFocus, 2003. *Discarded computer hard drives prove a trove of personal info*, s.l.: Security Focus.

Shao-Bo, J., Shawn, S.-C. & Quing-Fei, M., 2008. Systems Plan for Combating Identity Theft-A Theoretical Framework. *J Service Science and management*, Volume 1, pp. 143-152.

Simon, B. & Ray, M. S. a. F., 2002. System Design. In: G. Conor & D. Sarah, eds. *Object-Oriented Systems Analysis and Design Using UML*. Berkshire: McGraw-Hill Education, pp. 321-342.

Thakkar, D., 2016. *Bayometric: False Acceptance Rate (FAR) and False Recognition Rate (FRR)*. [Online]  
Available at: <https://www.bayometric.com/false-acceptance-rate-far-false-recognition-rate-frr/>  
[Accessed 2 June 2016].

Vatsa M., S. R. M. P. N. A., 2004. Signature Verification Using static and Dynamic Features. In: N. P. e. al., ed. Berlin Heidelberg: Springer-Verlag, pp. 350-355.

Vishesh, T., 2007. *Phishing and Pharming- The Deadly Duo*, s.l.: SANS Institute.

Wagner, M. & Urli, T., 2013. *Computer Lecture notes*. [Online]  
[Accessed 16 July 2015].

Witten, I. H., Frank, E. & Hall, M. A., 2011. *Data Mining Practical machine Learning Tools and Techniques*. 3rd ed. s.l.:Morgan Kaufmann Publishers .

Zimmerman, T. G. et al., 2003. *Retail Applications of Signature Verification*, Almaden: IBM Research.

## APPENDIX

### Appendix 1: Project Schedule

#### MSC Projects Timetable (May/June 2015)

	Activity	Start date	End date
1	Consultations and picking of project titles	11th May 2015	12th June 2015
	Supervisor Allocation		
2	Preparing the proposal	15th June 2015	10th July 2015
3	Milestone one presentation	13th July 2015	31 <sup>st</sup> July 2015
4	Working towards Milestone Two	July 2015	Oct 2015
5	Milestone Two Presentations	19th October 2015	23 October 2015
6	Working towards milestone three	Oct 2015	Nov, 2015
7	Milestone Three Presentations	23th Nov 2015	27th Nov 2015

## **Appendix 2: Questionnaire**

### **SACCO MEMBERS QUESTIONNAIRE**

#### Introduction

This questionnaire is prepared to collect information on the knowledge of Identity theft and the red flags that can be used as indicators. The information given is for academic purposes and not any other business and it will be kept confidential. Your name is not required, fill by putting a tick (√) on only one option in an item and kindly do so in every question.

#### **Section 1: Demographic Information**

1. What is your gender?

Female

Male

2. How old are you?

Below- 35 years

Between 35-55 years

Above 55 years

3. What is the name of the SACCO you are in?

---

SACCO

4. How long have you been in the SACCO?

1-9 years

10-20 years

21-30 years

30 years and above

#### **Section II: SACCO**

5. What would you need to do in order to get a loan from your SACCO?

Fill a form

Apply online either via (Mobile request or a website)

6. Have you ever applied for a loan?

yes

no

7. How many guarantors in your SACCO would you need to get a loan?

1-5

5 and above

8. What would the guarantors need to do to guarantee your loan?

---

Fill a guarantor's form

Sign your loan form

**Section III: Guarantor Details Falsification.**

9. Do you know if guarantor signature forgery exists?

yes

no

10. Have you ever had your signature and details forged by someone you once guaranteed?

yes

no

11. How did you know about it?

During the loan application process

Through a call from the SACCO for confirmation

After the person defaulted the loan

Didn't know until there was a deduction from your account

12. Did the SACCO assist after raising the issue?

yes

no

10. The table below contains statements regarding things that would be done. Kindly state the extent to which you feel the SACCO can improve to make sure, you as a guarantor you are protected, by putting a tick (✓) in the appropriate column. Use the key below in your responses:

**SA – Strongly Agree A – Agree D – Disagree SD – Strongly Disagree**

<b>Statement</b>	<b>SA</b>	<b>A</b>	<b>D</b>	<b>SD</b>
1. I think if they called guarantors before giving the loan would help in making sure am protected				
2. If I had knowledge of the impending action I would not have initially guaranteed him/her				
3. I had total trust in my fellow member				
4. I would have guaranteed the loan if he/she had approached me like before				
5. I will not stop guaranteeing other follow members because of one or two untrustworthy persons				
6. After I discovered my signature had been forged and informed the SACCO they went out of their way to assist me				
7. The SACCO is doing enough to sensitize people on Guarantors details falsification				
8. The SACCO has shown effort in managing these cases				
9. You were confident that when you informed them of the issue they quickly understood what had taken place				
10. The issue was resolved amicably				

Thank you for taking you time to fill this form.

### **Appendix 3: SACCO Request Letter**

Caroline Mwangi  
P.O.Box 247-00621  
Village market,  
Nairobi, Kenya

28/07/2015

Dear Sir/Madam,

#### **RE: REQUEST TO RESEARCH AT YOUR SACCO SOCIETY**

I am a student at the University of Nairobi, Reg no. P53/73007/2014 undertaking my masters course. I am currently in my project year and I am required to work hand in hand with SACCOs in my project. My title is ‘An Algorithm for Identity Theft Mitigation: Keypoint Signature Verification’. I would require information on how you identify signature forgeries, the recorded signature forgeries that you have experienced in the past and how you have dealt with them in order to resolve these incidents.

With this brief introduction to my project I will be seeking your help in making my project a total success. Your assistance will be highly appreciated. To contact me kindly use this details: phone number 0724987362.

I will be looking forward to hear from you. Thank you in advance.

Yours Sincerely

Caroline Mwangi

#### **Appendix 4: System Testing Review**

The table below contains statements regarding usability of the prototype that would be done. Kindly indicate how your experience was, by putting a tick (√) in the appropriate column. Use the key below in your responses:

**In a scale of 1-4 One been Best.**

<b>Statement</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
1. Was it easily accessible?				
2. How did you find the user interphase?				
3. Does it reduce time taken in verification?				

## Appendix 4: Code

```
package signaturesdk.verification;

import java.util.LinkedList ;
import signaturesdk.beans.PairList;
import signaturesdk.features.utils.Copier;
import java.awt.BasicStroke;
import java.awt.Color;
import java.awt.Cursor;
import java.awt.Graphics;
import java.awt.Graphics2D;
import java.awt.RenderingHints;
import java.awt.event.MouseEvent;
import java.awt.event.MouseListener;
import java.awt.event.MouseMotionListener;
import java.awt.image.RenderedImage;
import java.util.Arrays;
import signaturesdk.acquisition.SignatureJlabel;

public class EulerNumber {
    private double[][] SCEMatrix;
    //private int[][] wrapPath;
    private int criticalPointsOrg, criticalPointsTest;
    private int e;
    //private int T, costSCB; // Sakoe-Chiba band attribs.
    private int numOptimalDistance;
    private SignatureJlabel inputSignlabel, testSignlabel;
    private RenderedImage image, newImage; // Local copy of the image.
    private boolean displayLine; // Should we show the sampling line?
    private int samples;
    //private double npS1,npS2; // Used to get pixels from the image.
    private int x0,y0,x1,y1; // The points that define the sampling line.
    private int thisX,thisY; // Last mouse cursor position.
```



```
//private boolean firstSelected; // True if the first point is already selected.  
//private boolean secondSelected;
```

```
public EulerNumber () {  
    // the default value  
    numOptimalDistance = Integer.MAX_VALUE;  
}
```

```
public void draw() {  
    loadPixels();  
    // Since we are going to access the image's pixels too  
    img.loadPixels();  
    for (int y = 0; y < height; y++) {  
        for (int x = 0; x < width; x++) {  
            int loc = x + y*width;
```

```
// The functions red(), green(), and blue() pull out the 3 color components from a pixel.
```

```
    float r = red(img.pixels[loc]);  
    float g = green(img.pixels[loc]);  
    float b = blue(img.pixels[loc]);
```

```
    // Set the display pixel to the image pixel  
    pixels[loc] = color(r,g,b);  
    }  
    }  
    updatePixels();  
}
```

```
public void setLine(int x0, int y0, int xT, int yT)  
{  
    this.x0 = x0;
```

```

this.y0 = y0;
this.x1 = x1;
this.y1 = y1;
npS1= calcSamples();
npS2= calcsamples();
}

public void setRect(double x, double y, double w, double h) {
    inputSignlabel.setBounds(100, 100, 500, 300);
    testSignlabel.setBounds(100,100,500,300);
}

public void addPoints(float[] xValues, float[] yValues, float[] yErrorBars, int shape, String
label) {
    if (xValues==null || xValues.length ==0) {
        xValues = new float[yValues.length];
        for (int i=0; i<yValues.length; i++)
            xValues[i] = i;
    }
    //allPlotObjects.add(new PlotObject(xValues, yValues, yErrorBars, shape,
currentLineWidth, currentColor, currentColor2, label));
    //if (plotDrawn) updateImage();
}

private double computeSCEMatrix(double[][] distanceMatrix) {
    this.SCEMatrix = new double[this.criticalPointsOrg][this.criticalPointsTest];
    this.SCEMatrix[0][0] = distanceMatrix[0][0];

    // fill the first row
    for (int i = 1; i < this.criticalPointsOrg; i++)
        this.SCEMatrix[i][0] = distanceMatrix[i][0]
            + this.SCEMatrix[i - 1][0];

    // initialize the first column
    for (int i = 1; i < criticalPointsTest; i++)

```

```

        this.SCEMatrix[0][i] = distanceMatrix[0][i]
            + this.SCEMatrix[0][i - 1];

// fill the others
for (int i = 1; i < criticalPointsOrg; i++)
    for (int j = 1; j < criticalPointsTest; j++)
        this.SCEMatrix[i][j] = distanceMatrix[i][j]
            + Math.min(this.SCEMatrix[i - 1][j], //insetion
                Math.min(
                    this.SCEMatrix[i - 1][j - 1], // match
                    this.SCEMatrix[i][j - 1]) //deletion
            );

    return this.SCEMatrix[criticalPointsOrg-1][criticalPointsTest-1];
//return computePath();

    }
}

```

## Appendix 5: User Manual

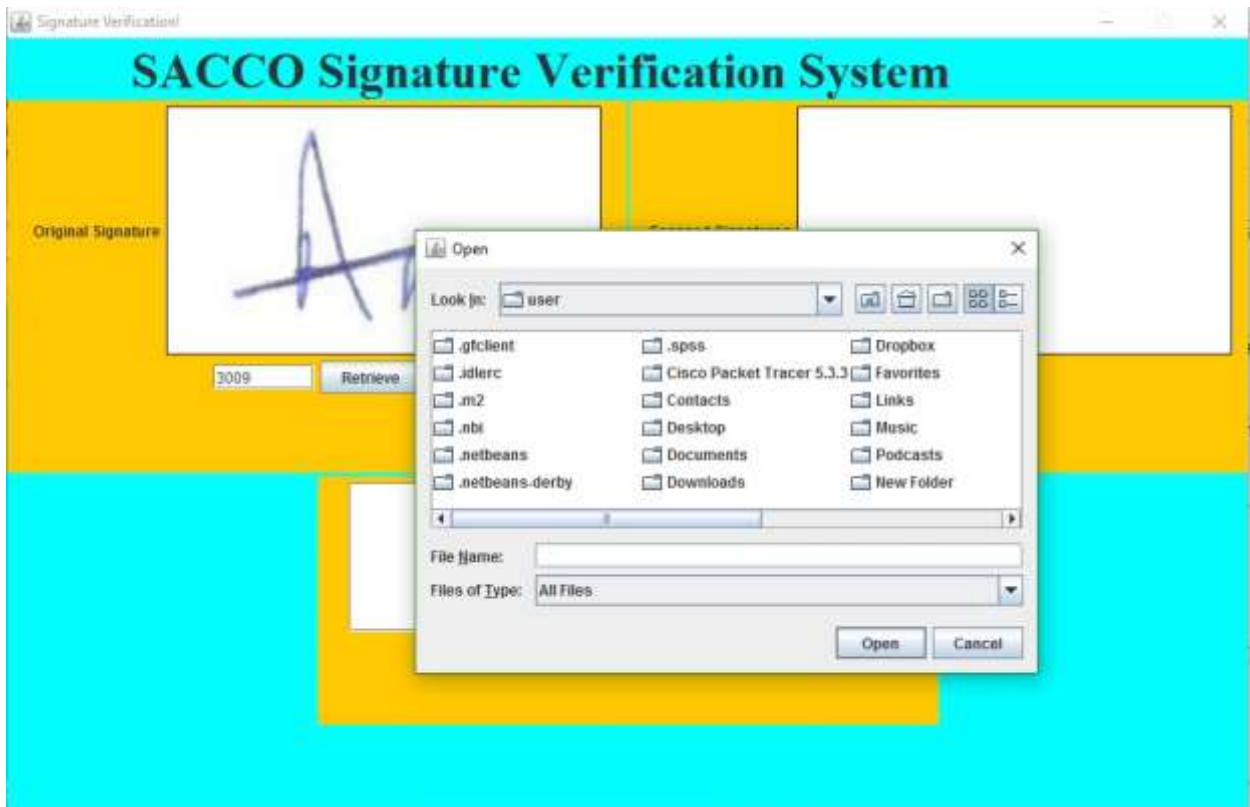
### 1. Open application



### 2. The retrieve the original signature from the Database by inputting the membership number



3. Then input the scanned signature through Open Signature button.



6. Then now click on verify for signature verification.



## Appendix 6: Research Permit



### NATIONAL COMMISSION FOR SCIENCE, TECHNOLOGY AND INNOVATION

Telephone: +254-20-2213471,  
2241349,3310571,2219420  
Fax: +254-20-318245,318249  
Email: dg@nacosti.go.ke  
Website: www.nacosti.go.ke  
when replying please quote

9<sup>th</sup> Floor, Utalii House  
Uhuru Highway  
P.O. Box 30623-00100  
NAIROBI-KENYA

Ref. No

Date:

**NACOSTI/P/16/99647/10532**

**11<sup>th</sup> May, 2016**

Caroline Wambui Mwangi  
University of Nairobi  
P.O. Box 30197-00100  
**NAIROBI.**

#### **RE: RESEARCH AUTHORIZATION**

Following your application for authority to carry out research on "*An algorithm for identity theft mitigation: Keypoint signature verification,*" I am pleased to inform you that you have been authorized to undertake research in **Nairobi County** for the period ending **10<sup>th</sup> May, 2017.**

You are advised to report to **the County Commissioner and the County Director of Education, Nairobi County** before embarking on the research project.

On completion of the research, you are expected to submit **two hard copies and one soft copy in pdf** of the research report/thesis to our office.

**DR. STEPHEN K. KIBIRU, PhD.  
FOR: DIRECTOR-GENERAL/CEO**

Copy to:

The County Commissioner  
Nairobi County.

The County Director of Education  
Nairobi County.