UNIVERSITY OF NAIROBI

COLLEGE OF BIOLOGICAL AND PHYSICAL SCIENCES

SCHOOL OF COMPUTING AND INFORMATICS

ENHANCED WEIGHTED TRUST EVALUATION SCHEME FOR DETECTION OF MALICIOUS NODES IN WIRELESS SENSOR NETWORKS.

KORIATA PATRICK TUYAA

A Research Project Report Submitted in Partial Fulfillment of the Requirements of the Degree of Master of Science in Distributed Computing Technology of the University of Nairobi.

November 2016.

DECLARATION

I certify that this research project report to the best of my knowledge, is my original authorial work except as acknowledged therein and has not been submitted for any other degree or professional qualification award in this or any other University.

Signature: _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ Date: _

Koriata Patrick Tuyaa

P53/79371/2015

This research report has been submitted as partial fulfillment of the requirements for the degree of Master of Science in Distributed Computing Technology of the University of Nairobi with my approval as the University supervisor.

Signature: _ Date: _

Prof. William Okelo-Odongo

DEDICATION

I dedicate this research project to my parents; I owe special gratitude to my late dad Kasuka ole Koriata and my mum Naiyiari Koriata who have always been there for me in every step of my life. Forever greatly indebted to you for the sacrifices made and your invaluable counsel.

Above all, I am greatly grateful to the Almighty God who has been my source of wisdom, strength and inspiration throughout this project work. His grace and mercies have really been abundant in my life.

ACKNOWLEDGEMENT

My sincere appreciations go to my thesis advisor and supervisor. Prof. William Okelo-Odongo who has been very kind and insightful throughout this thesis despite his busy schedule. His counsel, guidance, encouragement and expert opinions were indispensable and greatly vital towards the completion of this project.

I also give thanks to all members of the thesis panel for their criticisms, corrections and comments which were indeed of great help towards the successful completion of this project work. I am also grateful to my colleagues at the School of Computing and Informatics for extending support to me in one way or another.

ABSTRACT

Wireless Sensor Networks (WSNs) present myriad application opportunities for several applications such as precision agriculture, environmental and habitat monitoring, traffic control, industrial process monitoring and control, home automation and mission-critical surveillance applications such as military surveillance, healthcare (elderly, home monitoring) applications, disaster relief and management, fire detection applications among others.

Since WSNs are used in mission-critical tasks, security is an essential requirement. Sensor nodes can easily be compromised by an adversary due to unique constraints inherent in WSNs such as limited sensor node energy, limited computation and communication capabilities and the hostile deployment environments. These unique challenges render existing traditional security schemes used in traditional networks inadequate and inefficient. An adversary may take control of some sensor nodes and use them to inject false data with the aim of misleading the network's operator (Byzantine attack). It is therefore critical to detect and isolate malicious nodes so as to prevent attacks that can be launched from these nodes and more importantly avoid being misled by falsified information introduced by the adversary via them. This research gives emphasis on improving Weighted Trust Evaluation (WTE) as a technique for detecting and isolating the malicious nodes. Extensive simulation is performed using MATLAB in which the results show the proposed WTE based algorithm has the ability to detect and isolate malicious nodes, both the malicious sensor nodes and the malicious cluster heads (forwarding nodes) in WSNs at a reasonable detection rate and short response time whilst achieving good scalability.

Table of Contents

TABLE OF FIGURES

ACRONYMS

WSN	Wireless Sensor Network
SWSN	Surveillance Wireless Sensor Network
FN	Forwarding Node
SN	Sensor Node
AP	Access Point
BS	Base Station
WTE	Weighted Trust Evaluation
STL	Stop Transmit and Listen
SWATT	Software based Attestation for Embedded Devices
SPRT	Sequential Probability Ratio Testing
DoS	Denial-of Service
PDoS	Path based DoS
UML	Unified Modeling Language
TCP	Transmission Control Protocol
SYN	Synchronize Message
RSSI	Received Signal Strength Indicator
GPS	Geographical Positioning System

CHAPTER ONE: INTRODUCTION

1.1) Background.

Wireless sensor network (WSN) consists of a large number of spatially distributed autonomous sensor nodes working cooperatively to monitor the surrounding physical phenomena or environmental conditions (monitored target) and then communicate the gathered data to the main central location through wireless links. A sensor node, also known as mote is defined as a small, low-powered, wireless device, capable of gathering sensory information, perform limited data processing and transmit the gathered information to other nodes in the network via optical communication (laser), radio frequencies (RF) or infrared transmission media. A sensor node senses physical phenomena like light, temperature, humidity, pressure, chemical concentrations and any other phenomenon capable of causing the transducer respond to it. Once the phenomena is sensed, the data collected (measurement) is converted into signals for further processing to reveal some characteristics pertaining the phenomenon from the target area (Hussain, et al., April, 2013).

WSNs have a myriad of application areas including environmental and habitat applications, healthcare applications, military applications, agricultural monitoring applications and commercial applications like vehicle tracking, industrial processes control, inventory control and traffic flow surveillance. A number of these applications areas are mission-critical; for example battlefield surveillance applications, healthcare (elderly people, home patient monitoring), and disaster relief management as well as fire detection applications among others. The fault-tolerance, rapid deployment and self-organization characteristics of WSNs make them ideal for military's C4ISRT systems: "command, control, communications, computing, intelligence, surveillance, reconnaissance and targeting" (SOHRABY, et al., 2007). The scenario outlined below describes an area in which border surveillance WSN can be very useful for the military.

The porous Kenya-Somali border continues to pose a security threat to Kenyans; dangerous Al-Shabaab militants, illegal immigrants and smugglers are using it as a back door to the country. The porous border also contributes to the flow of illegal weapons and ammunition. These illegal entry points are made up of dirt tracks crisscrossing the bushes through the dusty and thickets-filled no man's land into and out of Kenya. They are situated far from the official checkpoints and border crossing points hence there are no government officials and security searches. These routes are commonly referred to as "Panya" or "Rat Routes". Somalia-based Islamist militant group Al Shabaab have taken advantage of this to conduct a series of cross-border raids including abductions of tourists and aid workers inside Kenya, Westgate Mall's attack in September 2013, storming Garissa University and killing 147 students in 2015 among other attacks in which the group has claim responsibility.

The Kenyan Authorities have stepped up efforts to beef up border security. Kenya Defense Forces (KDF) has invaded Somali with the aim of crushing the militant. The government is also constructing a wall along some sections of the border and it intends to install security cameras and expands the border patrol units. The Kenya-Somali border is roughly 780 kilometers and the Ministry of Interior admits that putting in place proper controls along the entire borderline is a serious and complex challenge. The construction and maintenance of the Kenya-Somali separation wall is an expensive endeavor, it may cost as much as US$2 million, per kilometer (Cannon, May 2016). There are other issues like the ongoing Kenya-Somali maritime border dispute that has been referred to the Hague-based International Court of Justice (ICJ). Somali is concerned that Kenya may use the construction of the border wall to draw the boundary to its favor (Cannon, May 2016).

In light of the above challenges and difficulties, a Surveillance Wireless Sensor Network (SWSN) would be appropriate to detect unauthorized intrusions and analyze enemy movements at the border locations. SWSNs can be employed in monitoring (gathering information) and protection of critical areas like borders, any precious asset, private properties or even rails. They detect intrusions and alert the military or the responsible personnel of targets of interest such as trespassers or moving vehicles in hostile environments or within a predefined area.

The hostile environment in which WSN are deployed in, the wireless medium and the constrained resources (limited energy, processing capability, and storage capacity) on the tiny sensor devices used pose a challenge in designing and implementing WSN security. (CHELLI, 2015).

Most wireless sensor network protocols, owing to the constrained resources inherent in the sensor node, assume a high level of trust between the communicating sensor nodes so as to eliminate the authentication overhead. This creates the danger of adversaries injecting malicious nodes to the sensor network or manipulate the operation of existing ones. The adversary may take control of some sensor nodes and use them to inject false data with the goal of misleading the network operator. Consequently, there is a risk of attackers launching an array of attacks on the sensor networks (Karuppiah & Rajaram, 2014.). According to (Alam & Debashis, 2014) the most dangerous attack in WSN is the insertion of a malicious node as it can destroy the whole network.

Several schemes for malicious nodes detection and isolation in WSN have been proposed. This research explores and improves one of them, the Weighted Trust Evaluation (WTE) Scheme. WTE is a lightweight algorithm use in a three-layer hierarchical network architecture consisting of low-powered Sensor Nodes (SN) having limited capabilities, higher-powered Forwarding Nodes (FN) which collect data from the lower layer (SNs) and the Base Stations (BS) or Access Points (AP) layer that route information between the wireless sensor network (WSN) and the wired infrastructure. Weighted Trust Evaluation Scheme is based on several assumptions i.e. both Forwarding Nodes (FNs) and Base station (BS) are trusted and won't be compromised and that the number of normal working nodes exceeds the compromised nodes. (Sumathi & Venkatesan, 2014) Once an adversary gains control over the BS then it leads to create any possible attacks in the network. The threat of Forwarding Nodes being compromised is not considered, a compromised FN gives an adversary control of all the sensor nodes under it.

Another scheme that has been used in the detection and isolation of malicious nodes is Stop Transmit and Listen (STL). STL employs non-transmission times to detect malicious nodes; nodes transmitting during these times exhibit malicious behavior. STL has a drawback in that when the whole network or a major portion of it stopped their transmission at a time (during the non-transmission time) and then resume transmission, congestion and unwanted delay in the network operations arises.

In this research, we propose an enhanced WTE based detection algorithm that aims to address the drawback of the WTE scheme by employing STL. The STL will come in handy to address the threat of the compromised forwarding nodes and since there are few, issues of congestions and delays in the network are avoided.

1.2) Problem Statement

The border surveillance wireless sensor networks (WSNs) are deployed in unattended and hostile environments. This among other issues such as unreliable wireless medium used and the constrained resources (limited energy, processing ability, and storage capacity) on the tiny sensor devices pose a challenge in designing security mechanisms for the WSN. In order to eliminate authentication overhead, most WSN protocols assume a high level of trust among the communicating nodes. However, this creates the danger of adversaries introducing malicious nodes to the sensor network or manipulate existing ones and then subsequently use them to propagate a wide range of attacks.

Detection and isolation of malicious or malfunctioning nodes in border surveillance WSN is a major security issue. It is crucial that these nodes be detected and excluded in the sensor network to avoid catastrophic decision being made as a result of falsified information injected by the adversary as well as prevent an array of attacks that can emanate from malicious nodes. Attacks emanating from malicious nodes are the most dangerous attacks. These necessitate that their detection and isolation be given top priority as malicious nodes can send erroneous or falsified report (Byzantine problem) to the base station leading to a disastrous decision; such as, in a battlefield surveillance WSN a misleading report about the enemy operations may result to extra casualties.

1.3) Objectives

The following are the aims of the research:

i. Investigate wireless sensor networks security design issues and challenges and the various attacks that adversaries can launch via malicious nodes.
ii. Design and implement a prototype of an enhanced malicious node detection scheme by amalgamating the Weighted Trust Evaluation Scheme and Stop Transmit and Listen (STL) scheme.
iii. Evaluate malicious node detection and isolation by analyzing the response time, detection ratio and the misdetection ratio of the above-proposed scheme.

1.4) Research questions

i) What are the wireless sensor networks security design issues and challenges?
ii) What are the various attacks that adversaries can launch via malicious nodes?
iii) Can the proposed scheme correctly detect and isolate malicious nodes from the WSN?
iv) What are the response timings, detection and misdetection ratio of the scheme under various simulation parameters?

1.5) Scope

The research identifies malicious node detection and isolation as a major security concern in Wireless Sensor Networks (WSNs). Data accurateness and timeliness are two major attributes that are central to the correct workings of the envisioned scheme. The goal is to detect and isolate the nodes that are transmitting incorrect data whilst taking into consideration the non-transmissions times of so as to treat them as malfunctioning (malicious).

The research uses simulation to test and evaluate the envisioned model. Simulation is a suitable tool to study WSN as setting up, deploying and operating a test bed for real experiments is difficult and expensive (López, et al., 2005). MATLAB/Simulink is chosen as the appropriate tool to model and simulate the envisioned scheme.

MATLAB and its backend software, Simulink, is a high performance tool with rich computational and visualization features. It offers a platform of easy programming capability where users can easily develop own custom functions. MATLAB/Simulink provide a communication toolbox to set up and build a complete Wireless Sensor Network system model (Nayyar & Singh, 2015) **.**

1.6) Justification

Malicious and faulty nodes in Wireless Sensor Networks can send falsified reports or erroneous data to the base station. This is catastrophic as mission-critical applications like military surveillance and health care WSN rely on accurate and timely data for reliable functioning of the network; for example, in a battlefield surveillance WSN a distorted misleading report pertaining the enemy operations may result to extra casualties

Ensuring that malicious nodes are detected and isolated as early as possible in the SWSN ensures accurate data is send leading to right decisions being made. This would go a long way in ensuring that the border surveillance WSN is reliable and can be depended upon to alert the personnel manning the border of terrorists, illegal immigrants and smugglers crossing the borders at non-designated points.

The study also adds to the body of knowledge in the field of surveillance WSN. The envisioned malicious node detection and isolation scheme once deployed can be useful for further research in WSN malicious node detection.

CHAPTER TWO: LITERATURE REVIEW

2.1) Introduction

This section analyses and reviews published work in the area of security of Wireless Sensor Networks (WSNs) and malicious node detection and isolation schemes. It commences with a brief introduction of WSN, an investigation on the major design challenges and issues in the WSNs security, the security goals of WSNs, attacks emanating from malicious nodes and then an analysis of malicious nodes detection schemes.

2.2) Wireless Sensor Networks

Wireless sensor network (WSN) consists of a large number of spatially distributed autonomous sensor nodes operating collaboratively to monitor the surrounding physical or environmental conditions (monitored target) and then communicate the gathered sensory data to the main central location through wireless links. A sensor node (mote) is a small, low-powered, wireless device, with limited computation and communication capabilities, capable of gathering sensory information, perform limited data processing and transmit the gathered information to other nodes in the network via optical communication (laser), radio frequencies (RF) or infrared transmission media. (Hussain, et al., April, 2013).

A sensor node comprises of a sensor, memory, processor, mobilizer, communication system, power units and position finding system. Each sensor node is made up of three subsystems namely:

- Sensor subsystem that senses the physical phenomena or environmental conditions.
- Processing subsystem that performs local computations operations on the sensed data.
- Communication subsystem that is responsible for message transmission and exchanges among neighboring sensors.

Sensors can monitor several phenomena such as humidity, temperature, lighting conditions, pressure, vehicular movement, noise level, chemical concentrations, soil makeup, and other properties. There are several types of sensors which include infrared, seismic, thermal, magnetic, acoustic, visual and radar based on the sensing mechanism employed by them (Ali , 2012). Once the phenomena is sensed, the data collected (measurement) is converted into signals for further processing to reveal some characteristics pertaining the phenomenon from the target area (Hussain, et al., April, 2013)

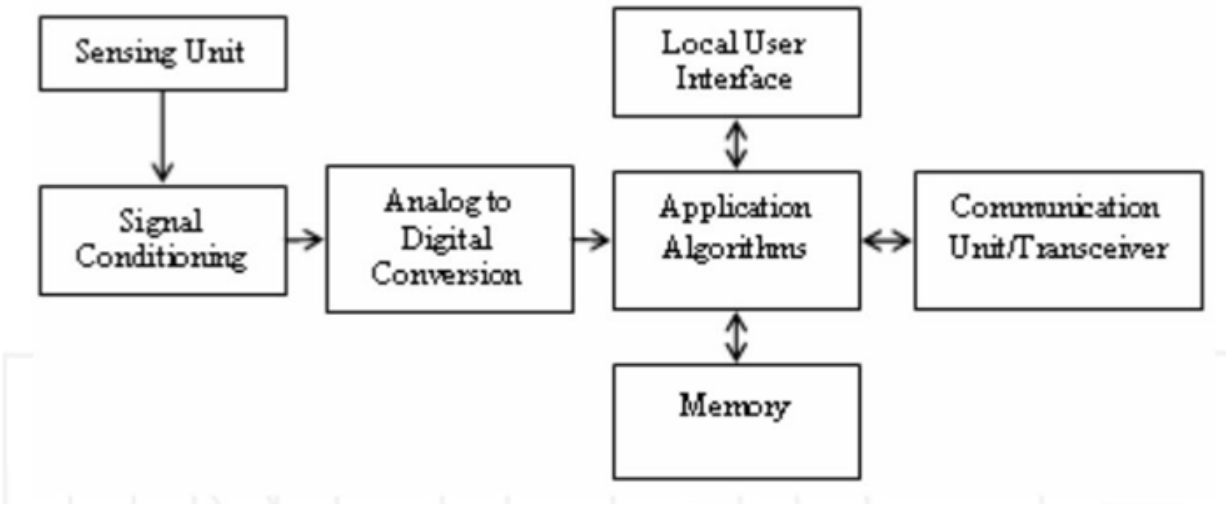


Figure 1: Sensor node basic architectural components (Ali , 2012)

WSN have great potential for deployment in mission-critical applications like battlefield surveillance applications, healthcare (elderly people, home-patient monitoring), disaster relief as well as fire detection applications among others. Since WSNs are employed in mission-critical tasks, security is an essential requirement. However, sensor networks pose unique challenges and as such existing traditional security schemes used in traditional networks are inadequate (PERRIG, et al., June 2004). Limited sensor node energy, computation and communication capabilities and the hostile deployment environments bring a challenge of employing efficient security solutions in WSN.

2.2.1) Surveillance Wireless Sensor Network

Surveillance Wireless Sensor Networks (SWSN) are deployed along the border or perimeter areas to monitor the real-world phenomena of interest in detail and detect unauthorized intrusions by hostile elements. The sensor nodes can either by deployed randomly via aerial deployment or deterministically where the exact locations of the sensor nodes are pre-determined. A SWSN can be employed in a broad range of places ranging from country borders for military surveillance, wildlife parks to monitor endangered animal species, embassies, and factories.

Once the sensor nodes are deployed to a region of interest; they organize themselves forming an operational sensor network and then start sensing the target area for intrusions such as tank vibrations, troop movements or sniper gun noise. The sensed event is relayed to the sink node via the cluster heads (forwarding nodes). In order to lessen the communication overhead, forwarding nodes perform data aggregation/compression on the sensed data before its transmission to the base station to provide situational awareness so that an appropriate action can be taken.

The main objective of border SWSN is the detection of enemy intrusions and alerting the military or the responsible personnel of targets of interest such as trespassers or moving vehicles in hostile environments or within a predefined area. Dense sensor nodes deployment is done in the border location to ensure robustness.

Security is an essential requirement in SWSNs used in mission-critical tasks such as military surveillance. Sensor nodes can easily be compromised by the attacker due to constraints like limited sensor node energy, limited computation and communication capabilities and the hostile deployment environments. The adversary may inject false data using the compromised nodes thus misleading the network operator; this has catastrophic consequences. In this research we investigate malicious node detection schemes with special interest in weighted trust evaluation scheme.

2.3) Challenges in Designing Wireless Sensor Network Security Schemes

The following are the various design issues and challenges within Wireless Sensor Network's platform that make the employment of existing security mechanisms inadequate and inefficient.

2.3.1) Very Limited Resources

The acute resource scarcity of sensor nodes poses significant challenges to resource-intensive security mechanisms. These mechanisms require certain amounts of resources such as energy, data memory and code space to function well but these resources are constrained in a tiny sensor node. The hardware constraints demand that the security algorithms used be extremely efficient in terms of memory, computational complexity and bandwidth, (Padmavathi & Shanmugapriya, 2009) .

Energy which is the most treasured resource for sensor networks also happens to be the biggest constraint as it limits its capabilities and must therefore be conserved or used effectively by the security mechanisms in place. Since the internal batteries of sensor nodes deployed in the field (hazardous environments) cannot be replaced or recharged easily; battery charge must be conserved as much as possible so as to extend the lifetime of the node and the sensor network in general. (SHARMA & TRIPATHI, April 2015). Communication is a power-intensive task and the security mechanisms used are required to be energy-efficient.

Clearly, security mechanisms employed in a sensor network must strive to be communication efficient in order to achieve energy usage minimization. Effective security mechanisms are also required to limit the security algorithm's size since the sensor node has limited memory and low storage capacity.

2.3.2) Unreliable Communication

Due to the inherent broadcast nature of the wireless communication medium employed in WSNs; packets may be distorted as a result of channel errors leading to conflicts, packets may also be dropped at highly congested nodes and an adversary can easily launch a Denial-of Service (DoS) attack.

The multi-hop routing, network congestion and node processing can result to greater latency in the sensor network resulting to synchronization issues among sensor nodes. These issues can hinder sensor network security especially where the security mechanism is based on cryptographic key distribution and critical event reports. (CHELLI, 2015)

2.3.3) Unattended Operations

The sensor nodes may be left unguarded for a long period of time in the field; this though depends on the application function of the sensor network in consideration. There are three major cautions to these unattended sensor nodes (Padmavathi & Shanmugapriya, 2009):

• **Exposure to Physical Attacks:** Sensor nodes may be deployed in a hostile environment exposed to adversaries and bad weather conditions. The probability that a sensor node suffers a physical attack like capture or destruction by an attacker in such an environment is therefore high.

• **Managed Remotely:** Sensor network remote management makes it nearly impossible to detect physical node tampering and manipulation by the adversaries.

• **Lack of a Central Management Point:** In order increase sensor network vitality, a wireless sensor network need be a distributed network devoid of a central management point. However, an incorrect or poor design will make the sensor network organization inefficient, difficult and fragile.

2.3.4) Hostile Environments

Sensor nodes in extremely hostile deployment environments are susceptible to destruction or capture by the adversaries as they are exposed to them. Attackers can capture a sensor node, disassemble it, and extract valuable information such as cryptographic keys from it.

2.4) Security Goals for Wireless Sensor Networks

The main objectives of Wireless Sensor Networks (WSNs) security are as follows:

2.4.1) Data Confidentiality

Confidentiality refers to the ability to conceal vital messages' content from being disclosed to unauthorized party or protect the messages against unintended access. Sensor nodes may exchange or pass highly sensitive information such as cryptographic key distribution and it must therefore remain confidential. This means that it is very crucial to build a secure communication channel in a sensor network. Data encryption should also be used to secure the data being transmitted across the sensor network.

2.4.2) Data Integrity

Data integrity is referred as the ability to assert that the message was not altered, tampered with or improperly modified in transit by an adversary. It is essential to guarantee data reliability.

The sensor network integrity will be compromised when (Padmavathi & Shanmugapriya, 2009):

- A malicious node in the network injects incorrect and misleading data.
- Unstable and turbulent conditions resulting from the wireless communication channel causing data damage or loss. (Akykildiz, et al., 2002)

2.4.3) Data Authenticity

Authentication ensures the reliability of the received message through source identity verification. An attacker can alter the data packet or even modify the whole packet stream by introducing extra bogus packets. Data authentication is therefore needed so that the recipient node can confirm that the data actually originates from the claimed sender (correct source).

2.4.4) Data Availability

Availability seeks to ensure that the required network services are functioning at a desired level of performance and work promptly in normal situations as well as in the event of attacks or environmental mishaps. It implies that the sensor node has the ability to access and utilize the available resources and that the network is operational and ready for use to transmit messages.

2.4.5) Data Freshness

This ensures that the transmitted messages are current and old content (expired packets) are not replayed by an adversary to either mislead the network or keep the network resources busy thereby reducing the sensor network vitality. It is essential especially in shared-key design strategies that require the keys be changed over time. (CHELLI, 2015)

2.4.6) Secure Localization

Sensors may get displaced during their deployment, after a certain length of time or after a critical displacement incident. WSN operations depends on its ability to automatically and accurately locate each sensor node in the network after the displacement. (CHELLI, 2015).

2.4.7) Self-Organization

WSN being an ad-hoc network and lacking a fixed infrastructure for network management requires that each node be independent and versatile so as to be able to self-organize and self-heal depending on the various situations, topology and deployment strategy. This inherent feature of the sensor network is a great challenge to WSN security. If self-organization is absent in a wireless sensor network, an attack or the risky deployment environment may have dire consequences. (Padmavathi & Shanmugapriya, 2009)

2.4.8) Time Synchronization

Time synchronization is required by many WSN applications, it is essential in multi-hop communication, conservation of node energy (periodic time sleep) and node localization. Sensor nodes may wish to determine the network latency of a packet as it transits between a pair of sensor nodes (sender-receiver) (Padmavathi & Shanmugapriya, 2009). Collaborative time synchronization may be needed by wireless sensor network for tracking applications.

2.5) Attacks Launch From Malicious Sensor Nodes

Since the wireless sensor networks are set up in hostile environments, sensor nodes can be compromised easily by the adversary due to the resource constraints such as limited memory space, battery lifetime and computing capability. Detection and isolation of these compromised nodes is crucial to avoid being deceived and misguided by falsified data injected by the attacker.

An adversary can easily launch a range of attacks against the wireless sensor network through the compromised (malicious) nodes Some of the attacks that can emanate from malicious nodes include sinkhole attacks, black hole attack, wormhole attack, Sybil attack, HELLO flooding attacks and Denial-of-Service attacks. (Atakli, et al., 2008)

2.5.1) Denial-of-Service attacks

Denial of Service (DoS) attack refers to an explicit attempt by the adversary to deny the victim (legitimate user) use or access to all or part of their network resources (Soomro, et al., 2008). In a DoS attack an adversary may destroy or disrupt a network, the attacker can also overload the network with bogus requests thereby diminishing the network's ability to provide a service (Virmani, et al., 2014). These attacks make the sensor node depletes the battery power and degrade the overall sensor network performance.

DoS attacks span across all the layers of the sensor network protocol stack as discussed below:

2.5.1.1) Denial-of-Service Attacks in Physical Layer

The most common attack at the physical layer is jamming; it aims at disrupting normal sensor network operations. The adversary may continuously relay radio signals or transmit high-energy signals that disrupt or block sensor nodes' communication by increasing the noise to signal ratio in the wireless communication medium. This results in a denial of transmission service attack.

The defense strategies employed include detect and sleep or detect and reroute around jammed regions (Raymond & Midkiff, 2008). For the physical protection of nodes tamper-proof packaging of nodes, redundant or camouflaged (hidden) nodes are deployed in the field.

2.5.1.2) Denial-of-Service Attacks in Data Link Layer

Link layer protocols are responsible for coordination of access (channel arbitration) of the shared wireless medium channels by neighboring sensor nodes as well as provide abstraction to the upper layers. Link layer attackers breach predefined protocol rules and behaviors. These attacks may include inducing packet collisions by disrupting a packet, repeated packet transmissions thereby draining sensor node energy or abuse of a cooperative Media Access Control (MAC) layer priority mechanism resulting to unfairness (Woo, & Culler, 2001) (CHELLI, 2015).

Link-layer authentication which allows for communication with trusted parties only, antireplay protection which ensure packets are only sent and received once, error correcting codes (used for resisting collision), detect attack and sleep and broadcast attack protection are used as the attack mitigation strategies.

2.5.1.3) Denial-of-Service Attacks in Network Layer

Denial of service attacks in the network layer of sensor networks are intended at disruption of sensor node routing information interfering with the whole operation of the network.. (Raymond & Midkiff, 2008). Routing-disruption attacks can result in a DoS attack as a malicious node that disrupt the network's routing protocol can advertise itself as part of a favorable route only to drop the received packets.

An adversary may also alter, spoof, replay routing information with the aim of traffic disruption in WSNs. These attacks can be mitigated by using egress filtering and authorization (Pathan, 2010).

2.5.1.6) Denial-of-Service Attacks in Transport Layer

The attacker may make numerous SYN connection requests to overflow buffer state; until all the resources needed for the connection are depleted or reach the upper limit. This in effect causes severe resource constraint or resource exhaustion for a legitimate node (Raymond & Midkiff, 2008).

The SYN flooding attack is mitigated through SYN cookies in which the client maintains state and the server is relieved of keeping the connection requests queue but otherwise encode the information it should keep in the TCP sent as the response to the SYN.

2.5.1.7) Denial-of-Service Attacks in Application Layer

This layer is susceptible to a variety of attacks such as repudiation, overwhelm, Path based DoS (PDoS), data corruption and malicious code. The attacker can launch an overwhelm attack by overwhelming sensor nodes causing the network to transmit huge volumes of traffic. This drains the node energy as well as consumes the network bandwidth. (CHELLI, 2015) (Parno, et al., 2005). In PDoS attack a huge number of bogus packets are transmitted through a particular path to the base station, keeping the nodes in the path busy hence starving legitimate traffic and draining the sensor network resources (Pathan, 2010).

Use of replay protection, data aggregation and various authentication mechanisms could be effective defense strategies against these types of attacks.

2.5.2) Black Hole attack

A malicious node take advantage of routing protocol's packet route discovery process vulnerabilities to advertise itself to other nodes in the sensor network as having the shortest valid route to the packets destination node (Y-C & Perrig, 2004). The attack modifies the routing protocol so as to channel traffic through a particular node (malicious node) controlled by the adversary.

In the route finding process, the source node relays RREQ (Route Request) packets to intermediate forwarding nodes to find the best valid route to the intended packet destination node. Since malicious nodes do not consult the routing table, they reply immediately to the source node. (Das, et al., 2002) . The source node then assumes that the route finding process is over, ignores other nodes' RREP (Route Reply) messages and selects the route through the malicious node as the best route to transmit the data packets to the intended destination. The malicious node is able to accomplish this by allocating a high sequence number to the RREP packet. The source node starts forwarding its packets to the black hole trusting that they will be relayed to the destination. The adversary controlling the black hole may now discard these packets instead of relaying them to the destination node as stipulated by the protocol.

The countermeasures against black hole attack are implicit acknowledgments that ensures received packets are forwarded exactly as they were sent to the node, multipath routing that routes the packets over multiple paths increasing the probability of them reaching the destination,

2.5.3) HELLO Flood attack

A laptop-class adversary with a higher radio transmission power and range relays routing protocol HELLO packets to a number of other sensor nodes within a WSN making them assume the attacker is their neighbor (Padmavathi & Shanmugapriya, 2009). The hello packets recipient sensor nodes are influenced that the compromised node (adversary) is within their radio range. These node during data transmission to the base station may forward packets to the adversary since they assume it is their neighbor and are eventually spoofed by the adversary.
Hello flood attack could be mitigated using pairwise authentication of nodes or by employing geographic routing protocols. Pairwise authentication enable sensor nodes verify bi-directionality of a link before they can construct routes to forward traffic received over the link. Geographic routing protocols like Geographic and Energy-Aware Routing allow nodes discard hello messages received from nodes not within their communication range in terms of locations, which nodes broadcast to each other (Raymond & Midkiff, 2008).

2.5.4) Sink hole attack

In a sinkhole attack, the attacker's main goal is to allure the traffic from nodes in its close proximity (neighboring nodes) through a compromised sensor node. These attacks make the compromised attacking node look enticing and ideal to be used by the surrounding neighboring nodes to forward traffic. (Padmavathi & Shanmugapriya, 2009)

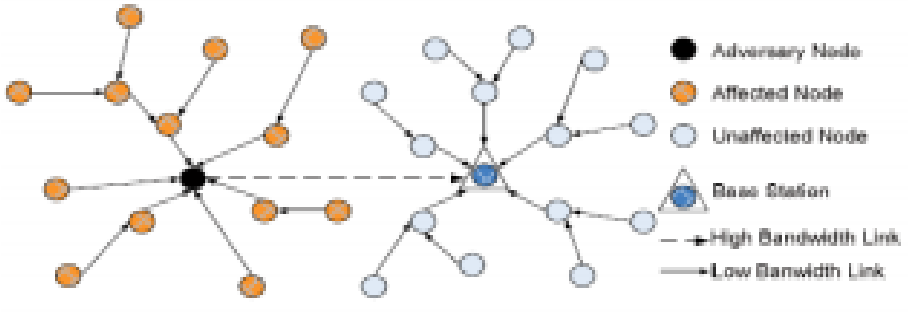


Figure 2: Sinkhole Attack (Alajmi, July 2014)

Some of the countermeasures against sinkhole attack include Hop Count Monitoring scheme, calculation of RSSI (received signal strength indicator) value at the receiver and the use of Mobile Agent based approach where a mobile agent delivers information of one node to others in the network.

2.5.5) Sybil attack

Sybil attack is an identity-based attack in which an attacker infects a single node with malicious code that duplicates the node; presenting multiple identities in multiple locations to other nodes in the sensor network. The multiple identities of node degrades the integrity of data as well as straining the network's resources. The Sybil attack decreases the efficiency of fault tolerant schemes like multipath routing, distributed storage and topology maintenance (Padmavathi & Shanmugapriya, 2009).

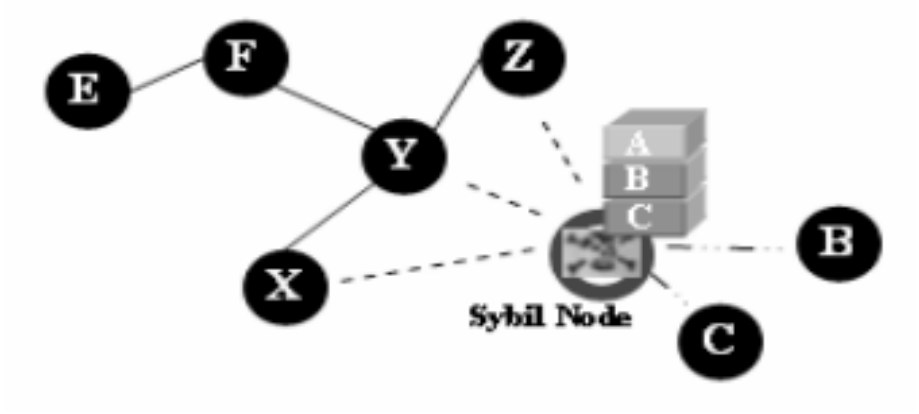


Figure 3: Sybil Attack (Alajmi, July 2014)

Authentication and encryption schemes can protect a sensor network from Sybil attacks. Trusted certification in which a centralized authority issues exactly one identity (certificate) can also be used to counter Sybil attack.

2.5.6) Worm Hole attacks

This is an attack in which the packets or their individual bits are captured at one part of the sensor network, tunneled over a low latency link to another location and are then replayed at their destination location (Hu, et al., 2003). This is usually accomplished by two distant colluding nodes which create an impression that the two locations involved are directly

connected even though they are genuinely distant (Virmani, et al., 2014).

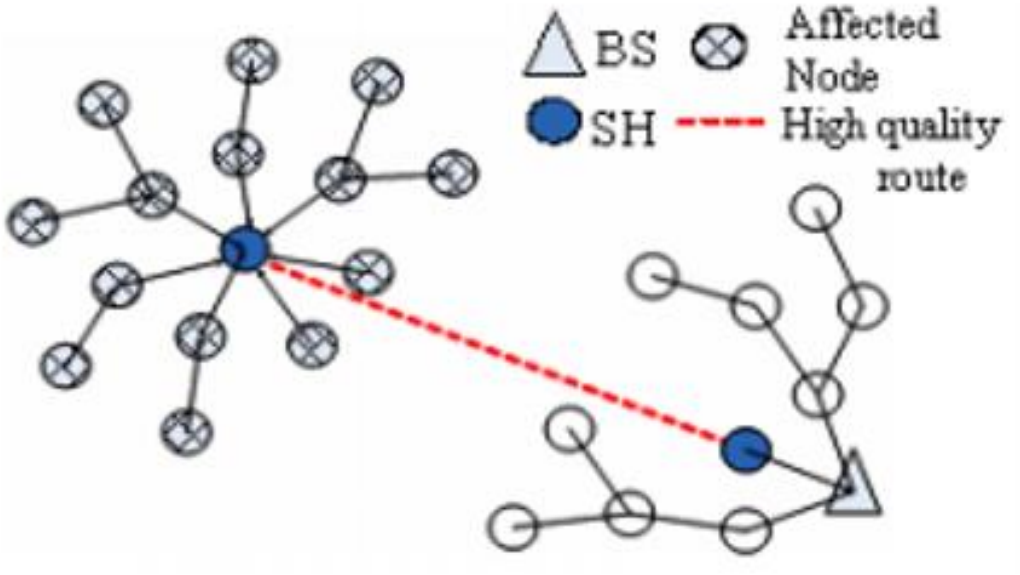


Figure 4: Wormhole attack (Abdullah, et al., 2015)

A wormhole attack involve two distant malevolent nodes conspiring to understate their inter-node distance by relaying packets through an out-of-bound channel which is only available to the adversary.

Some defense strategies against wormhole attacks are packet leashes which ensures that packets are not accepted “too far” from their source. Geographical leashes uses GPS information embedded into the packet being send whereas temporal leashes uses the nodes’ clocks timestamps added to the packet.

2.6) Malicious Nodes Detection Techniques

Several schemes for malicious node detection and isolation in WSNs have been proposed.

(Sung & Choi, 2013) Proposed a Dual Threshold technique for malicious node detection that employs two thresholds to minimize false alarm rate as well as improve the detection accuracy. All deployed sensor nodes do have transmission ranges, ’tr’, and any other sensor node in close proximity i.e. within the node transmission range is considered its neighbor. Each individual

sensor node maintains its neighbors' trust values to designate their trustworthiness. The sensor node makes a localized decision based on its own readings and those of its neighbors taking into account their trust values. Trust values lie between 0 and 1. If Tik=0 means node Ni does not trust Nk at all. A node also has its own trust value, once Tii=0 means the node is faulty.

(Curiac, et al., 2007) Proposed Auto regression Technique which is a mechanism that relies on past/present sensor node values. The sensor node present value is compared with an estimated value computed from its own previous values by an autoregressive predictor placed at the base station. The two values are compared to check if node behavior is normal or abnormal. If the variance between these two values is higher than a set threshold, the node is regarded malicious.

(Yang, et al., 2007) Proposed SoftWare-based ATTestation (SWATT) mechanism to authenticate the embedded device (sensor nodes) memory contents and detect any falsification or maliciously altered or inserted code in memory. The verifier send to the embedded device a randomly generated MAC key, which then calculates Message Authentication Code (MAC) value on the whole memory using the received key and returns the MAC value. The verifier uses the checksum to verify the memory contents. If the memory has been maliciously altered by the adversary then the checksum is false.

(Bao, et al., 2011) Proposed a Trust-Based Intrusion Detection approach which considers a composite trust metric derived from both social trust and quality of service (QoS) trust to identify malicious nodes in the wireless sensor network. The cluster head apply intrusion detection in the sensor nodes to assess the trust worthiness and maliciousness of the nodes in its cluster. This is achieved by statistically examining peer-to-peer trust evaluation results gathered from the different sensor nodes (Sumathi & Venkatesan, 2014).

(Nidharshini & Janani, December 2012.) Proposed a Sequential Probability Ratio Testing (SPRT) to detect duplicate nodes made by an adversary in the WSN. The attacker can easily capture and make replicas of unattended nodes and then use them to take control of the entire network. The base station is responsible for identifying compromised nodes by computing the speed of observed sample nodes and decides which nodes' speed exceeds the decided threshold speed, these ones are regarded malicious.

2.6.1) Weighted Trust Evaluation Scheme.

Weighted-Trust Evaluation (WTE) based scheme is a light-weighted algorithm used to detect and subsequently isolate compromised (malicious) nodes by monitoring their reported data in a hierarchical WSN architecture. (Zhao, et al., March 2013) (Atakli, et al., 2008) Employed and demonstrated this method using a three-layer hierarchical sensor network. The components of the three-layer hierarchical network architecture are:

a) Low-power Sensor Nodes (SN) whose functionalities are limited. SN is in the lowest tier and does not offer multi-hop routing capacity as in a traditional flat sensor network. SNs report the data to its Forwarding Node.
b) Higher-power Forwarding Nodes (FN) which collect data from the lower layer (SNs), verify its correctness, aggregate and forward it to other FNs or to the upper layer (Base Station).
c) Base Stations (BS) or Access Points (AP) which verifies data reported by the FNs as well as routing data between the wireless sensor network and the wired infrastructure.

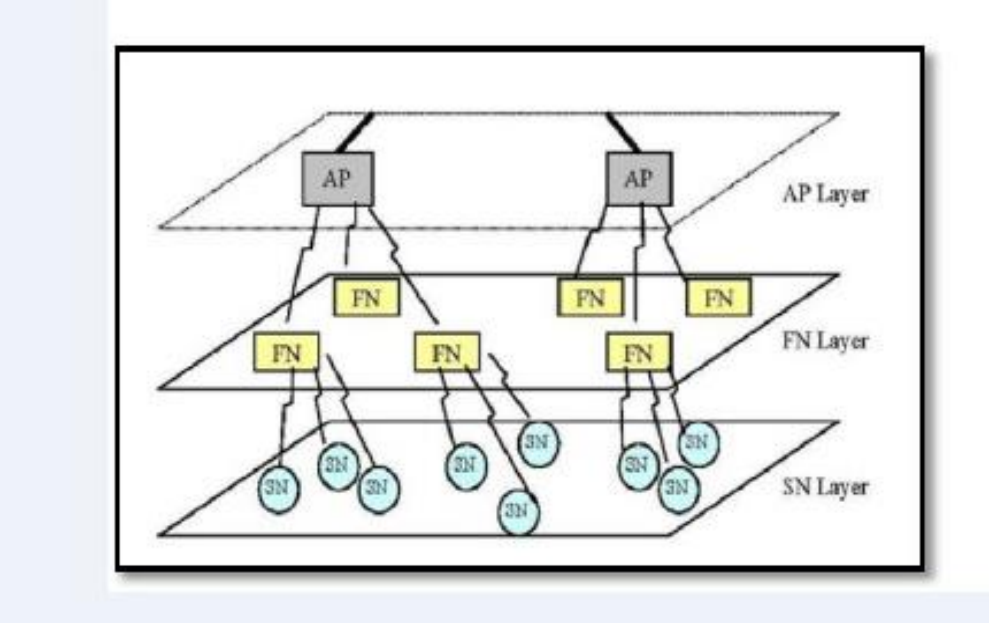


Figure 5: Architecture of the hierarchical WSN (Atakli, et al., 2008).

This scheme is based on two assumptions; first, the FNs and Base station are trusted nodes that cannot be compromised by an attacker since once an adversary seize control of the BS then they can launch any possible attack in the sensor network (Sumathi & Venkatesan, 2014) (Hu, et al., 2009) (Atakli, et al., 2008). Another critical assumption is that the normal nodes (working in proper condition) in the sensor network exceeds in number the compromised nodes. Otherwise, the scheme may misidentify normal node as compromised nodes increasing false positives. The proposed enhanced WTE intends to detect and isolate malicious FNs in the sensor network instead of assuming they won't be compromised by adversaries. This aims to cautions all the SNs under a FN which the attacker can control and manipulate once it take control of a particular FN.

2.6.1.1) Malicious Nodes Detection

A compromised sensor node provides falsified information that may wrongly mislead the senor network. This problem is referred as the Byzantine problem. A compromised/malicious sensor node can continuously forward wrong information to the upper layers. The aggregator (AP or FN) in the upper layer may compute an incorrect aggregation result due to the misleading information emanating from the malicious nodes. This may have disastrous effects to the decision making process.

WTE scheme models malicious node detection and isolation in 2 steps;

First, an initial weight Wn is assigned to every sensor node (SN) in the sensor network. The Forwarding Node (FN) gathers all the reported data from all the SNs under it and computes an aggregated result taking into account each SN weight.

$$E = \sum_{n=1}^{N} Wn \; x \; U \; n \;\; / \sum_{n=1}^{N} Wn$$

Where:
E = FN aggregate result.

Wn = SN assigned weight (Ranging between 0 and 1).

Un = SN output information (Un is usually dependent on the sensor network application. The output value may be “true” or “false” or continuous numbers like in a case of temperature readings).

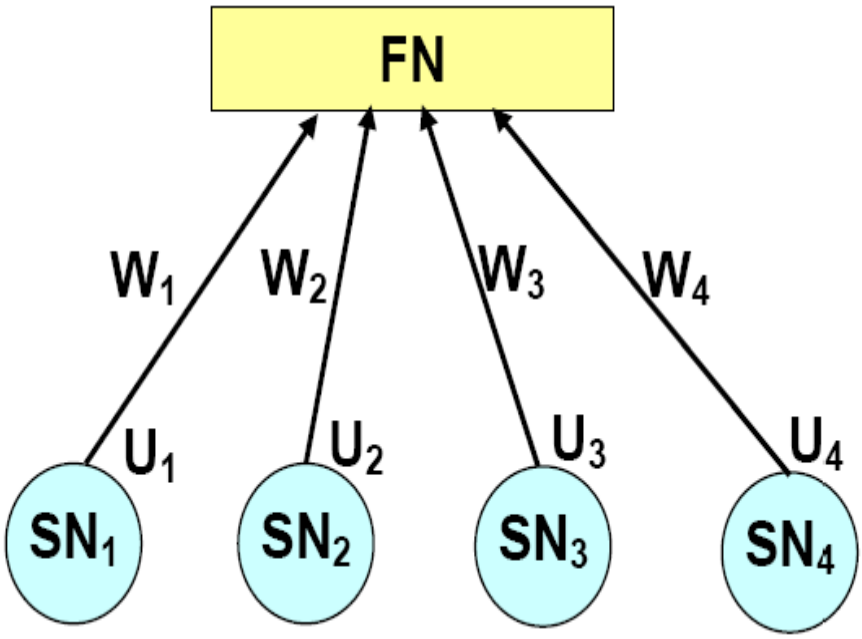


Figure 6: Weight-based hierarchical wireless sensor network (Atakli, et al., 2008).

Each SN weight is updated based on the accuracy of the reported information. The SN weight is updated for two reasons. First, if a compromised sensor node continuously forward data that is inconsistent with the final aggregate decision, its weight is likely to be reduced by a set weight penalty. If the weight decreases below a given threshold, then it is identified as a malicious node. Second, the SN weight determine how much a sensor node report contribute to the final aggregate decision. This is meant to lower the effect of incorrect reports from malicious sensor node.

2.6.1.2) Weight Value Recovery

The SN weight is decreased by a certain penalty value once it is detected to be reporting falsified data. However, the false report may be a result of a temporary communication channel interruption and the SN is neither malicious nor faulty. The weight values for such SNs needs to be recovered after the disturbance rather than keeping these values low permanently. The SNs that behave correctly thereafter longer than a set recovery time have their weight value increased.

2.6.2) Stop Transmit and Listen (STL)

The STL scheme employs non-transmission time slots to detect malicious nodes. Each sensor node have an inbuilt time limit to stop their data transmissions and listen for traffic. Once the nodes have been deployed and they have started sensing the target phenomena, the sensed data is sent to the base station. After every few seconds or after a set transmission time, each sensor node halt their data transmission process and listens for malicious traffic. If a sensor node transmits data during the non-transmission time (listening time) , it is caught by its neighbor nodes in the sensor network and it is regarded malicious as it exhibits malicious behavior. If a malicious node doesn't transmit data during a non-transmission time slot, it will still be caught in other frequent non-transmission times. The malevolent behavior of a malicious node is broadcasted across the entire sensor network. (Sathyamoorthi, et al., 2014). Then every other sensor network node desist from either forwarding data to the detected malicious node or accepting from it.

This technique has some weaknesses such that when the whole network or a major portion of it stopped their transmission at a time (during non-transmitting time) and then resume transmission, congestion and unwanted delay in the network operations arises (Sumathi & Venkatesan, 2014).

2.7) Literature Summary

Several schemes for malicious node detection and isolation in sensor networks have been fronted. They include Software based ATTestation (SWATT), Weighted Trust Evaluation Scheme**,** Stop Transmit and Listen (STL), Auto regression technique, Sequential Probability Ratio Testing and Trust Based Approach.

Weighted Trust Evaluation Scheme is based on several assumptions i.e. both Forwarding Nodes (FNs) and Base station (BS) are trusted and that the number of normal nodes exceeds the compromised nodes. (Sumathi & Venkatesan, 2014) Once an adversary gains control over the BS then it leads to create any possible attacks in the network. The threat of Forwarding Nodes being compromised by the adversary is not considered.

STL has a drawback in that when the whole network or a major portion of it stopped their transmission at a time (during non-transmission times) and then resume transmission, congestion and unwanted delay in the network operations arises. STL can be improved by separating the network into several groups/clusters, each having their transmission in a separate non-overlapping time interval. The cluster based approach is used to overcome the delays problem which results to delays in the wireless sensor network.

This research project proposes a scheme that enhances Weighted Trust Evaluation Scheme by amalgamating it with STL in a hierarchical WSN. Weighted Trust Evaluation Scheme is employed in the lower SN layer to detect and isolate malicious sensor nodes and the STL is used to detect malicious Forwarding Nodes (FNs). Instead of assuming that the FNs won't be compromised, this research considers malicious FNs detection in WSN as a weighty matter since a malicious FN is a threat to all the SNs under it.

2.8) Conceptual Framework

The research evaluates the performance of the enhanced WTE for detection and isolation of malicious node in WSN via three identified metrics namely response time, detection rate and misdetection ratio.

Response time is used to show how quick the enhanced WTE based scheme detects malicious nodes present in a sensor network. It is the average number of cycles required by the scheme to correctly detect malicious nodes. A short response time that ensures malicious nodes are isolated as early as possible in the WSN is desirable so as to lessen the disastrous effects of these nodes on the overall operation of the sensor network.

Detection ratio (DR) refers to the ratio of malicious nodes detected by the scheme to the total number of malicious sensor nodes present in the WSN. Detection ratio is used to indicate the effectiveness of our enhanced WTE scheme. The DR should be high to ensure that all malicious nodes are detected and isolated in the sensor network. This is key in eliminating misleading report emanating from malicious sensor nodes present in the WSN.

The third metric is Misdetection ratio, which refers to the ratio of misdetected nodes to the total number of all detections made by the scheme; this includes malicious nodes correctly detected and all misdetected nodes. Misdetected nodes belong to two classes: malicious nodes considered normal by the scheme and normal nodes considered malicious. The misdetection ratio of the scheme should be as low as possible so as to reduce the false positives reported.

The aim entailed designing and developing a sensor network malicious node detection scheme with high detection rate, short response time and low misdetection ratio.

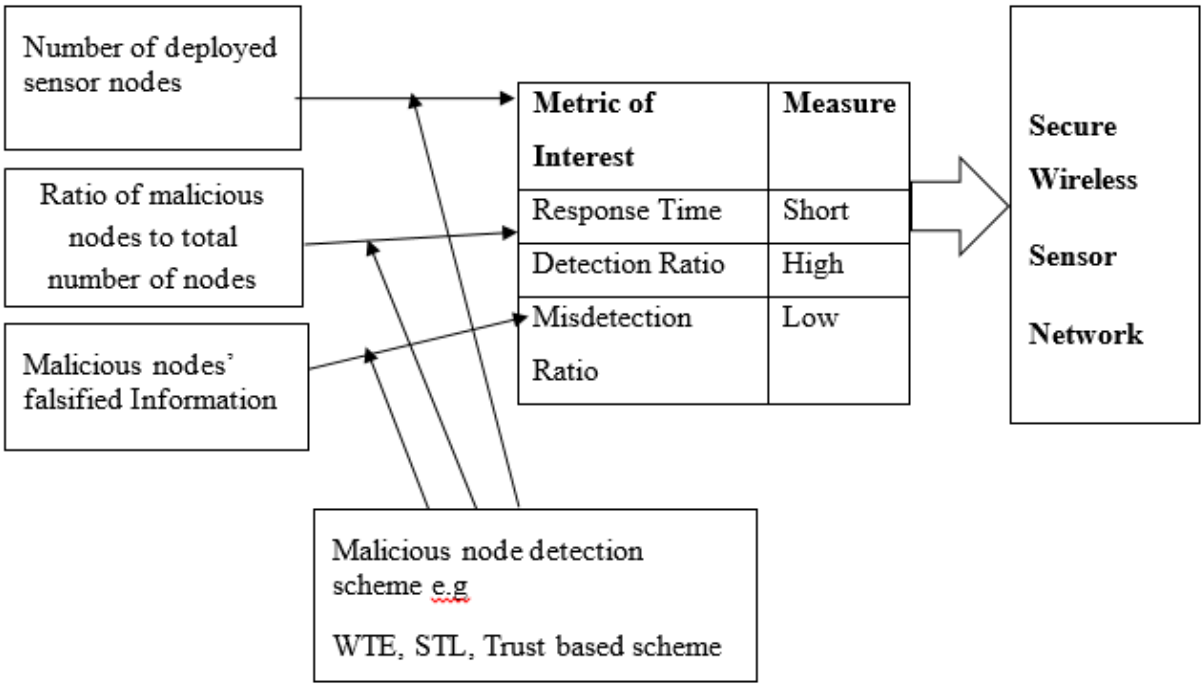


Figure 7: Conceptual Framework

The ratio of malicious nodes present in the network to the total number of deployed sensor nodes affects the detection and misdetection ratios in that when malicious nodes are the majority compared to the normal nodes, the number of misdetected nodes increases. Malicious nodes inject falsified data to mislead the sensor network.

CHAPTER THREE: METHODOLOGY

3.1) Introduction

This chapter centers on the methodology for designing and evaluating an enhanced Weight Trust Evaluation scheme (WTE) for Surveillance Wireless Sensor Network (SWSN). The main objective of the study was to design and evaluate a malicious node detection scheme (algorithm) using a SWSN model. The model design was achieved using prototyping and evaluated through simulation methodology. Modeling and simulation tools for WSN already exist and hence a tool was not developed. We used MATLAB/Simulink as the modeling and simulation tool for this project.

3.2) Simulation and Modeling

Simulation is the technique of representing or imitating the operations and behaviors of real-world processes, events or systems over time by means of something suitably analogous, studying the representation (model) to arrive at some outcome/results.

A simulation model is a conceptual representation of a process, system or phenomenon under study. A model is simpler (not so complex) but a close representation of reality. Modeling is the process of generating a model that represents the process or system of interest. Simulations allow evaluation of a model to be able to make predictions about a real system or to optimize system performance.

Simulations are suitable to study properties of a model of a real-life system that would otherwise be too complex, expensive to test in the real world, not accessible, too large/small, too dangerous, too fast/slow, or unacceptable to engage. Simulation also offers a flexible way of controlling experimental factors, allowing the changing of the model's inputs and studying the effect to the output.

Simulation is a suitable tool to study WSNs as setting up, deploying and operating a test bed for real experiments is difficult and expensive. The simulation methodology was adopted in evaluating and analyzing our malicious node detection and isolation scheme under different conditions i.e. number of deployed sensor nodes and the altering the ratios of the malicious and normal nodes with respect to total number of nodes in the sensor network while studying the outputs. The simulation and modeling of our scheme is detailed in chapter four and five of this report.

There are various tools use for simulation and modeling based on the area of application. Some of the simulation tools for wireless sensor networks are NS-2 (Network Simulator-2), Mannasim , OMNeT++, OPNET, WSN Localization Simulator**,** SENSE, JSim, MATLAB/Simulink including others. Tools with a graphical user interface (GUI) are more preferred and widely used since they enable the designer and audience to experience the simulated model as it would appear in real life.

3.3) Modeling Technique

The research made use of the agent based modeling (ABM) technique. ABM approach models systems composed of discrete, autonomous and interacting agents. It models the individual agents' behaviors and how these behaviors influence other agents. The complex WSN system is composed of different elements (sensor nodes) that behave differently at different times making agent based modeling ideal.

Although the detection and isolation of malicious nodes is a continuous process, agent based modeling allowed us to represent the system and processes as operations being performed at different points in time and at different devices. Events triggered other events in the model and from these triggers and responses data was generated. The generated data formed the backbone of this research.

3.4) Modeling and Simulation Tool

This research used MATLAB/Simulink simulation tool for the modeling and simulation of our enhanced WTE scheme. MATLAB and its backend software, Simulink, is a high performance tool with rich computational and visualization features. It offers a platform of easy programming capability where users can easily develop own custom functions. MATLAB/SIMULINK provide a communication toolbox to setup and build a complete Wireless Sensor Network system model (Nayyar & Singh, 2015) **.**

3.5) Software Development Methodology

Software development methodology/software development life cycle refers to a set of principles, procedures, tools and techniques that are used to structure, plan, organize, control and manage the whole process of building and implementing software systems projects.

The following are some of the software development methodologies;

Waterfall Methodology: This is a linear-sequential, structured and phased software development approach flowing from requirement gathering and analysis, design, implementation, testing, to deployment. Each of these phases has a set of well-defined activities.

Iterative and Incremental development: In this methodology, the system functionalities are developed through a cyclic/iterative approach with a gradual increase in the system features per iteration resulting to a stable iteration release. The cycle is repeated taking advantage of what has been learned in the previous cycles and user feedback given until the desired software functionalities have been fully developed.

Prototyping: In prototyping, a few aspects of the system are simulated and may be different from the real system. It typically involves determining the system's basic inputs and output, building, reviewing and enhancing the prototype. Suitable for demonstration in which building the actual system would take long or expensive. It helps to quickly gauge the accuracy of the system early enough in the project.

Spiral: The spiral methodology combines waterfall and prototyping methodologies and is well suited for large complex projects.

Unified Rational Process: It emphasizes on following software development best practices. Requirements are managed through use cases, visual modeling is done using Unified Modeling Language (UML) and the designed is accomplished through component based architectures. The software is developed incrementally and iteratively. Good for software projects done in teams.

This project employed the prototyping methodology to design, build and test the malicious node detection and isolation scheme. Prototyping was suitable for this project due to the following factors; the acquisition of the real sensor hardware and related software is expensive, practical deployment of the real sensor nodes in the test bed/field is time consuming and since the sensor nodes are fragile and susceptible to failure they need to be well protected. Prototyping offers the best alternative for the experimentation of the scheme under different conditions using the available software tools (simulation and modeling tool) before it can be deployed in the real world fields.

The specific steps carried out in the prototyping of our scheme include:

I. **Requirements identification** : The requirements for the developed algorithm are non-functional (performance) requirements and they are;
    a. Short response time - the scheme should possess the capacity to identify the malicious nodes in the wireless sensor network as fast as possible.
    b. High detection ratio – the scheme should be able to identify a very high number of malicious sensor nodes present in the wireless sensor network
    c. Low misdetection ratio – the scheme should not misdetect nodes or the number of misdetect nodes should be very low.

II. **Design of the algorithm:** We used system model diagrams, control flow diagrams, flow charts and pseudo codes to show how the algorithm works and to depict the flow of our algorithm's logic. Details of the algorithm's design are found in chapter four.

III. **Algorithm Implementation:** The algorithm design was converted into a computer program using the MATLAB language. The MATLAB platform was chosen due to its advantage of quick prototyping and fast computational engine.

IV. **Testing the developed algorithm:** Testing involved verifying that the initial model requirements specification have been met by the output of the implemented algorithm. The main features tested include generation of short response timings, high detection rate and low misdetection ratio.

3.6) Evaluation and Analysis

The algorithm evaluation and analysis was achieved via the simulation methodology. Extensive simulations were carried out in the MATLAB environment. A number of sensor nodes 'n' were deployed randomly in a square field to form a heterogeneous wireless sensor network. Some nodes were set malicious whereas others acted as normal nodes. The simulations were then run a number of times and the results observed.

Evaluation and analysis were conducted to determine the effectiveness our scheme in terms of how quick the malicious nodes are identified (response time), the number of detected and misdetected nodes.

The simulation tool identified above is capable of producing reports in graphical as well as tabular form. The simulation performed generated data in form of graphical images and reports as well. This data generated is based on events at various points in the WSN model. The generated data contained the response timings, detection rate as well as misdetection ratios in the WSN model.

CHAPTER FOUR: DESIGN AND IMPLEMENTATION

4.1) Introduction

In the previous chapters; the tools, procedures and methods for implementing our proposed model were identified. In this chapter, we will discuss these tools (MATLAB/Simulink) in detail and how they were used to achieve the implementation of the solution.

4.2) Overview of MATLAB/Simulink

MATrix LABoratory (MATLAB) is a multi-paradigm tool for technical computing, data analysis and visualization, algorithm and interactive development. It is both a computational and application development platform. MATLAB is well suited for performing computationally intensive tasks, solving mathematical problems, modelling, simulation and control theory-related applications.

Simulink is a block diagram environment fully integrated with MATLAB .It supports model-based design, simulation, automatic generation of code, testing and verification of dynamic and embedded systems including image, signal and video processing, communications and control systems. It also provides customizable block libraries and an interactive graphical environment.

The customizable block libraries and the responsive graphical environment enable the users perform design, simulation, implementation, and testing of a range of time-varying systems such as controls, communications, signal processing, image and video processing.

MATLAB offers the advantage of quick prototyping, fast computational engine, integration with other programming languages and ability to work with different data sources.

4.2.1) MATLAB Data Input/Output

MATLAB achieves data Input/output (data I/O) through functionalities such as data import functions, support for JDBC (Java Database Connectivity) and ODBC (Open Database Connectivity) compliant databases, interface with data providers such as Hyperfeed, Yahoo etc. and also interfacing with Excel.

Several commands are used to open, read files and display output in MATLAB namely:

i. Textread - Read text file data and write to multiple outputs(store it in an array)
ii. Fopen - Open a file or obtain open files' information.
iii. Fscanf - reads formatted data from a file and returns it as an array of values.
iv. Fprintf - Write formatted data to file.
v. Load - Load variables from file into workspace
vi. Xlsread - Read MS Excel files (.xls)
vii. Textscan - Read text file data, convert it and write to a cell array.

4.2.2) MATLAB Data Analysis

MATLAB has a library of functions that are very useful for data analysis. The workflow for data analysis in MATLAB is data access, pre-process, analyze, visualize and then sharing of the results. MATLAB and related data analysis tools perform analysis and give insights, identify trends, estimate uncertainty, create visualizations and models as well as publish reports. Some of the tool boxes and algorithm that are of help in analysis include:

i. Statistical analysis Toolbox

It is used to analyze historical data as well as modeling data, simulating and developing statistical algorithms. It can be used to perform probability distributions, descriptive statistics, hypothesis tests, statistical plots and linear as well as non-linear modelling.

ii. Curve Fitting Toolbox

These are routines used for preprocessing data, creating, analyzing ,comparing, and managing models. It can be used to determine the goodness of fit, analyze fit and perform Fourier series fit.

iii. Optimization Toolbox

It contains functions used to find parameters for minimization or maximization of objectives while satisfying their constraints. It includes quadratic and linear programming, as well as nonlinear optimization and nonlinear least squares solvers. These solvers can be applied in finding optimal solutions to discrete and continuous problems, performing tradeoff analyses as well as integrating these optimization solutions into algorithms and applications.

4.2.3) Modeling in MATLAB

Modeling is the creation of virtual representation of real world systems. Simulation is valuable in testing the model to see how it behaves under a wide range of conditions. Modeling and simulation is useful in the design process as it can be used to identify errors early in the design process.

The common representations used in MATLAB for system models include block diagrams, schematics and state charts. They can be used to model communication systems, signal processing algorithms, control software and mechatronics systems.

4.3) System Model

Our research considers a wireless sensor network (WSN) with n sensor nodes randomly distributed in a region R. A subset of the n nodes are powerful forwarding nodes. The nodes form clusters and the powerful nodes act as cluster heads/forwarding nodes forwarding data to the base station. Sensor nodes in close neighborhood (members of one cluster) register similar readings else they are deemed malevolent. Each node j collects data samples about its local environment and transmits the data to the forwarding node which act as the intermediate to the base station. The communication path over which the sensed values are propagated from the source node j to the forwarding and then to the base station is assumed to be error-free so the data reaches to the base station without modification enroute. We also assume that the bandwidth of the wireless channel used in transmission is not limited so contention issues are reduced.

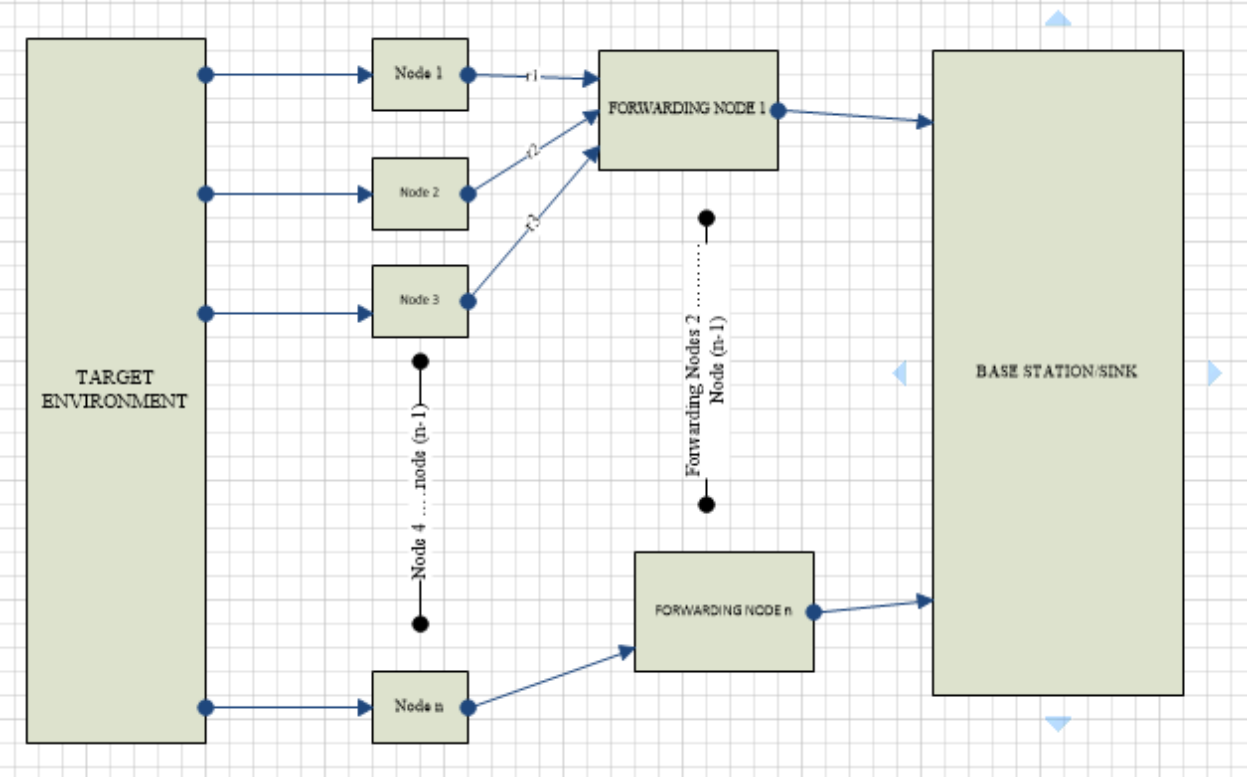


Figure 8: System Conceptual Model

4.4) Enhanced Weighted Trust Evaluation Scheme.

A heterogeneous wireless sensor network made up of sensor nodes with different energy levels and processing power is assumed. The deployed sensor nodes are assumed to form two sets in the ratio of p: 1-p where p is the percentage of higher energy sensor nodes. The higher energy (powerful) subset are elected as the forwarding nodes (cluster heads). The forwarding nodes broadcast its presence to all the normal sensor nodes. Normal sensor nodes choose the cluster to belong based on the broadcasted signal strength. It is assumed that the stronger the signal, the closer the forwarding node. The normal sensor node ends up choosing the forwarding node with the shortest distance from it as its cluster head.

All the cluster sensor nodes members forward their sensed data to the forwarding nodes whereas the forwarding nodes forward the aggregated value to the sink node for further processing and decision making.

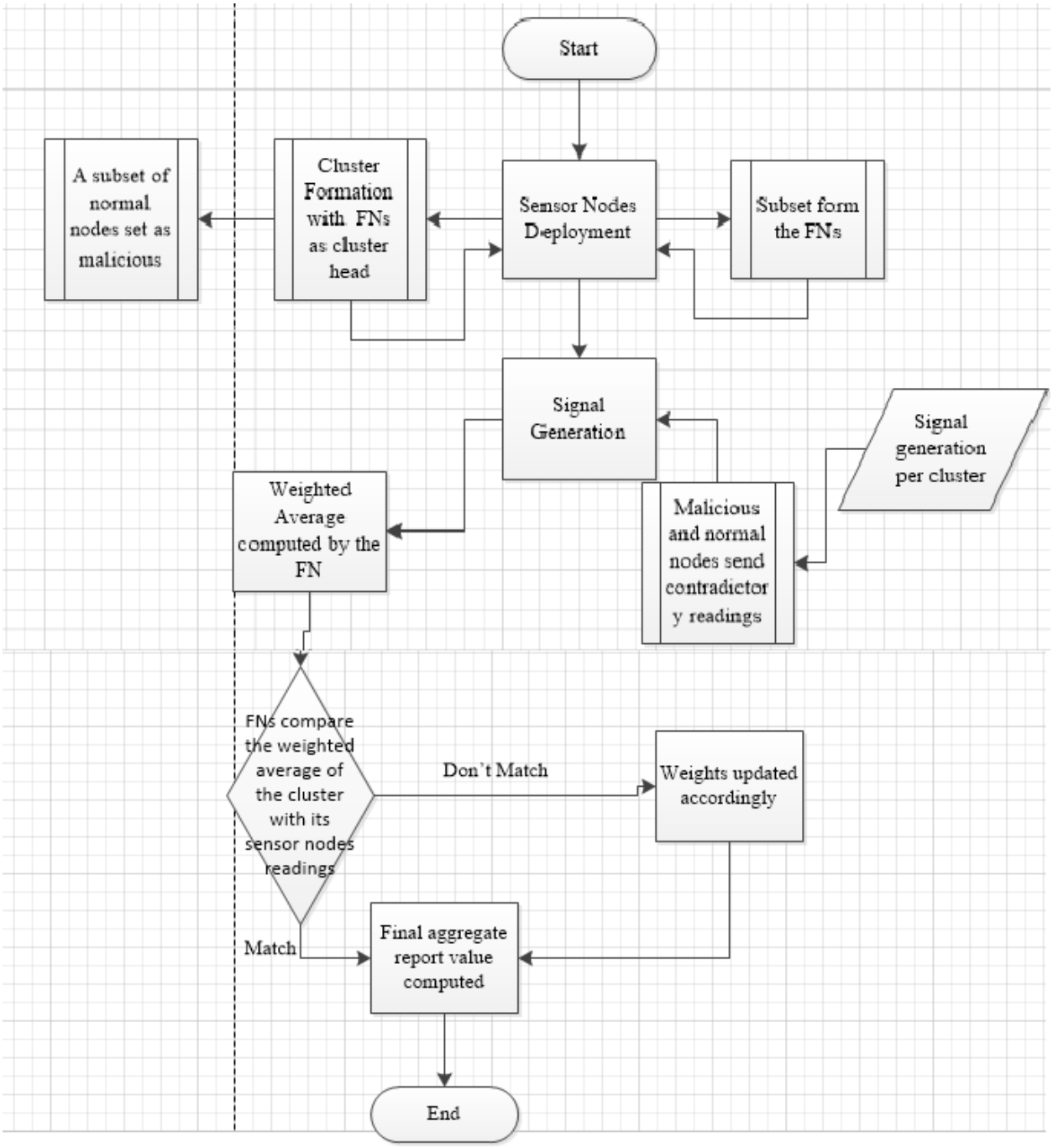


Figure 9: Enhanced Weighted Trust Evaluation Scheme - Control Flow Diagram

4.4.1) Enhanced Weighted Trust Evaluation Algorithm

The algorithm comprises of two phases:

4.4.1.1) Deployment and selection phase

Step 1: n sensor nodes deployed.

Step 2: Select a subset (p) of the deployed nodes as the powerful forwarding nodes.

Step 3: The forwarding nodes broadcasts a hello message (an advertisement message) to all normal sensor nodes.

Step 4: The normal sensor nodes that have selected a particular forwarding node as their cluster head send an acknowledgement message to it and they become cluster members.

Normal sensor nodes decide on the cluster to belong based on its proximity to the cluster head since it is assumed that the nearest forwarding node (FN) broadcasted the strongest signal.

4.4.1.2) Data computation and transmission phase

Step 1: Cluster member(s) transmit sensed data to the forwarding node (FN).

Step 2: FN gathers the data forwarded by the normal sensor nodes under it.

Step 3: FN perform an aggregation of the data collected taking into account the weights assigned to the normal sensor nodes.

Step 4: The aggregate value is compared to the individual values of the normal sensor nodes.

Step 5: The weights of the cluster members whose values are not in sync with the aggregate value are gradually reduced till their values is below the minimum weight threshold set.

Step 6: When the sensor node weight is below the minimum weight threshold, they are detected as malicious and isolated from the sensor network.

Step 7: The forwarding nodes forward the aggregate data value to the base station during the transmission times.

Step 8: The forwarding nodes stop transmitting and listen for malicious traffic in the network during the non-transmission times.

Step 9: The forwarding nodes transmitting during non-transmission times are detected as malicious.

The normal forwarding nodes only the send data to the base station during the transmission times. During the non-transmission times, they listen for any malicious traffic and are caught transmitting during these time slots are identified as malicious.

4.5) Malicious sensor node modeling

We consider a border monitoring WSN where the field or region is filled with IR (Infrared) sensors to detect any human presence. The region where the human presence is actually sensed is called an 'event region' whereas the other region is known as 'non-event region'. In case of human intrusion, the normal nodes in an event region send '1' directly to the FN indicating alarm. The other nodes (malicious nodes) send no alarm i.e. '0' to the FN. The malicious nodes in the non-event region send 1 (alarm) to the FN and the normal ones send a 0 (no alarm).

Let's consider each sensor node 'nj' in the network field reporting reading 'rj' such that rj= 1 for an event condition and 0 for no event condition. The aggregated value (E) gives the weighted average of the signal sensed by the deployed sensor nodes. If a sensor node is compromised by the adversary, it will send incorrect data to the FN making it transmit wrong data to the base station enabling the attackers achieve their aim of misleading the sensor network operator.

This malicious node detection algorithm is illustrated below:

1) Each sensor node nj sends a reading, rj to the Forwarding Node (FN). The normal sensor nodes send 1 (alarm) whereas the malicious ones send a 0 in case of an event and vice versa.

2) Each FN computes the aggregate value, E:

$$E=\sum_{j=1}^{J} Wn \; x \; rj \; / \sum_{j=1}^{J} Wn$$

Where Wn= Weight assigned to the node

3) Each FN computes the percentage of nodes (Pe) that have reported an event and those that didn't (Pn). Aimed at achieving majority voting in a cluster.

Pe = No. of nodes that reported an event/ Total number of nodes

Pn = No. of nodes that did not reported an event/ Total number of nodes

4) If Pe >= Tu (upper threshold) then majority of the nodes sent an alarm signal, an event has occurred and the weights are updated accordingly.
5) If Pn <= Tl; Tl being lower threshold, then an event hasn't occurred and the weights are updated accordingly.
6) For steps 5 and 6 the interchange of the percentages Pn >= Tu and Pe <= Tl also applies
7) Determine the nodes with Wn = 0 as malicious.

The weight of the sensor node is gradually reduced by the penalty factor if it sends reading not in sync with the aggregate value of the forwarding node. The weight assigned to a node is updated to Wn = 0 if its weight is reduced below the set minimum weight threshold.

4.6) Sensor Node Weight Updates

The sensor nodes are assigned a weight value (Wn) which represent its reliability or the confidence level. This helps to monitor their behavior as they report their readings as well as modifying their contribution to the final report of the forwarding node. The weight value (Wn) is between 0 and 1. Initially it is set to 1, Wn=1, and it is updated each time the sensor node reports a wrong value i.e. its reading does not correspond to the aggregate value. The node weight is set to 0 if its weight is reduced below the set minimum weight threshold, detected as malicious and isolated from the network. Every time that a sensor node is reporting a false value, its weight is reduced by a penalty value.

4.6.1) Weight Reduction Flowchart

The flowchart below depicts the weight reduction procedure. The weight of the node, Nj is reduced by the penalty factor, Pf, if it reports a false value. The initial condition St = N is done to ensure that only the normal nodes are considered and malicious are isolated from the network. Once the node's weight is reduced below the minimum weight threshold, St = M and thereafter its readings are ignored. The procedure reduces the weight of the node 'Wj' each time it sends false data till it is declared malicious.

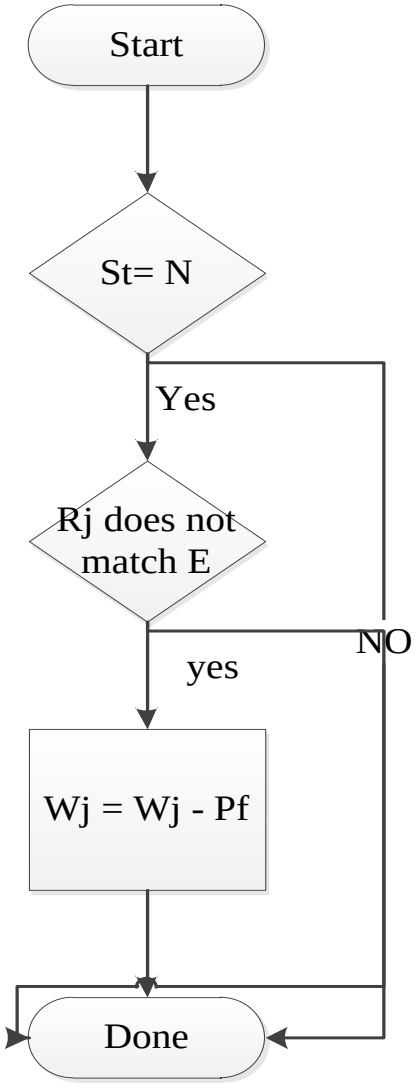


Figure 10: Weight Reduction Flowchart

4.7.) Simulation Setup

Extensive simulations of the proposed scheme described in the previous chapters are carried out in MATLAB. Heterogeneous wireless sensor network of 100 sensor nodes deployed randomly between [0,0] and [100,100] in a square area with field dimensions of 100*100 m is considered.

Parameter	Values
No of sensor nodes, n	100
Percentage of the powerful nodes subset, p	0.2
Percentage of malicious nodes to total nodes deployed, m	0.2
Weight penalty factor	0.2
Minimum weight threshold	0.6
Sink Location	[50, 100]
Network Field Dimensions	100*100 m

Table 1: Simulation parameters

At the initial setup, the sensor nodes are of three types; normal sensor nodes, forwarding nodes and the sink node. The forwarding nodes are p percent of the total number of nodes (n) deployed in the field. In the network of n=100 nodes considered, the powerful forwarding nodes would be p*n whereas the remaining (1-p) nodes are normal nodes. This translates to (0.2 * 100) =20 forwarding nodes and ((1-0.2)* 100) =80 normal sensor nodes.

Both the normal and forwarding nodes are randomly deployed whereas the sink node is placed outside the sensing area [50, 150].

The following colors were used to represent the sensor nodes:

a) ‘g’ - green to denote a normal sensor node.
b) ‘b’ - blue to denote a forwarding sensor node.
c) ‘m’ - magenta to denote the sink node.
d) ‘r’ - red to denote the malicious ordinary sensor node.
e) ‘k’ - black to denote the malicious forwarding sensor node.

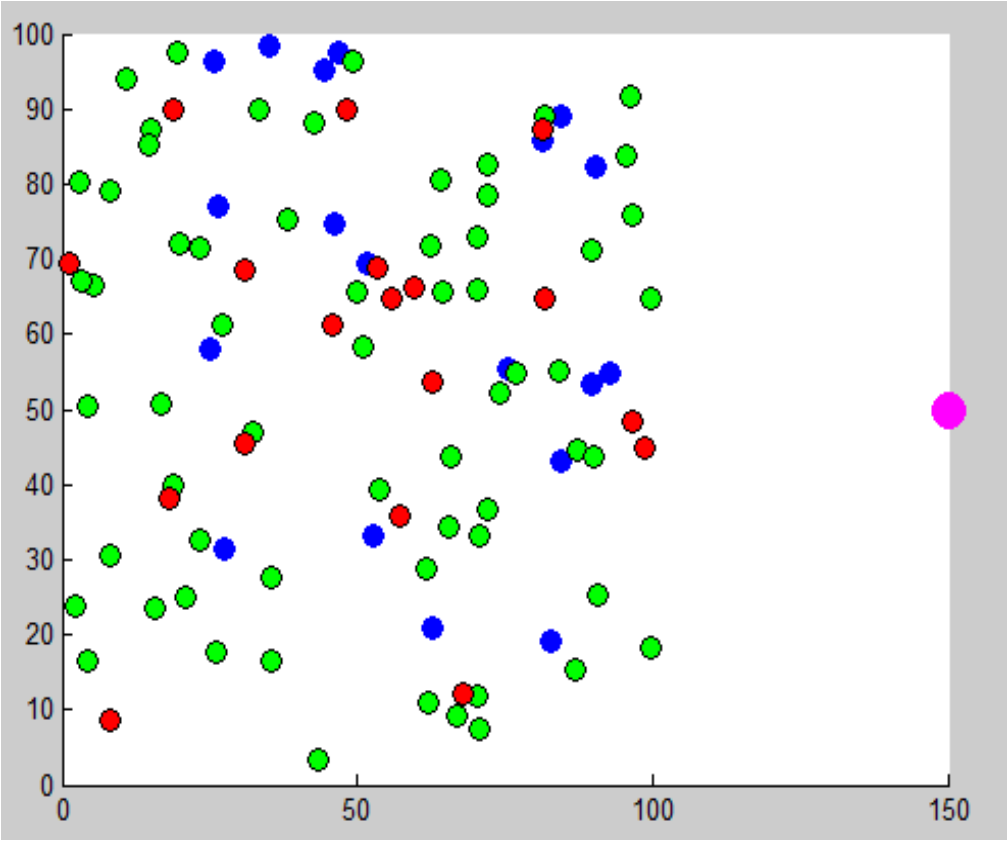


Figure 11. Random deployment of sensor nodes

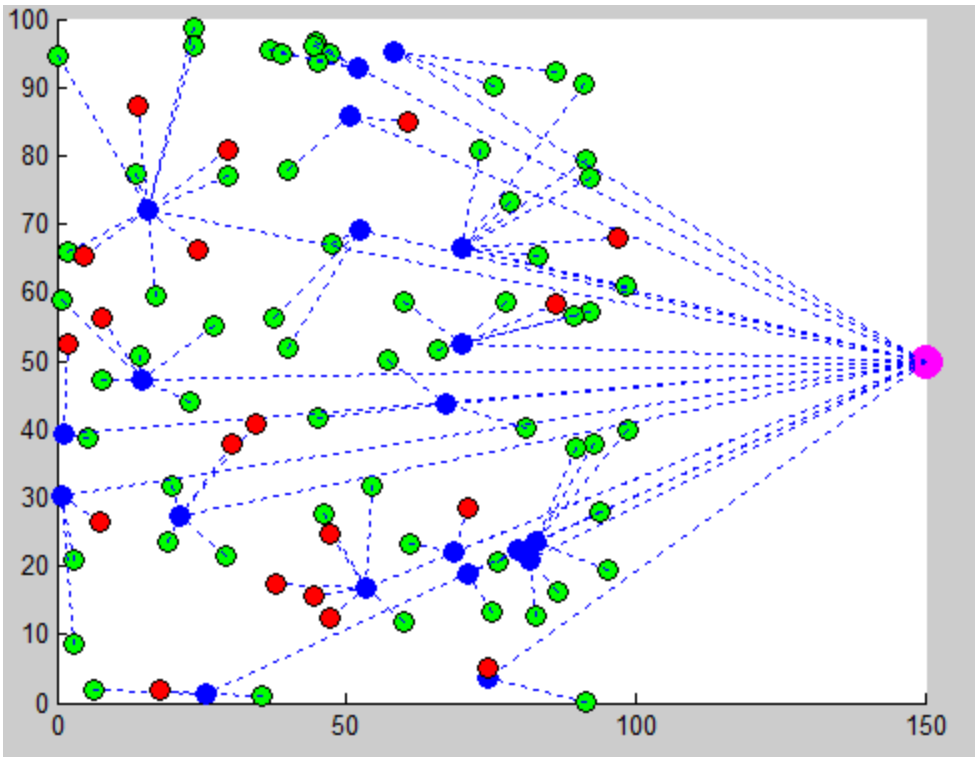


Figure 12: Sensor Network Data Transfer

The following assumptions were considered during the design and evaluation of the WSN model:

i. The sink and the forwarding nodes possess powerful processing power and have unlimited supply of energy.
ii. The normal sensor nodes are of limited processing power and limited supply of energy.
iii. The deployed sensor nodes are not mobile and are distributed randomly.

CHAPTER FIVE: SIMULATION RESULTS AND DISCUSSION

5.1) Introduction

In order to carry out performance evaluation of the enhanced weighted trust evaluation scheme, the following evaluation specific objectives were set.

I. To find out the average number of cycles required to correctly detect malicious nodes present in the wireless sensor network.
II. To find out the number of correctly detected malicious sensor nodes with respect to the total number of malicious nodes in the sensor network under various simulation parameters.
III. To find out the ratio of misdetected sensor nodes to the total number of detections made by the scheme under the various simulation parameters.
IV. To show that the scheme has a short response time, high detection rate as well as a low misdetection ratio.

5.2) Evaluation Metrics

1) Response Time

Response time (RT) refers to the average number of cycles required by the scheme to correctly detect malicious nodes in the sensor network. Response time is used to show how quick the enhanced WTE based scheme detects malicious nodes present in a sensor network.

2) Detection Ratio

Detection Ratio (DR) refers to the ratio of malicious nodes detected by the scheme to the total number of malicious sensor nodes present in the WSN. It is used as a scheme effectiveness indicator. A high DR is key in ensuring misleading report emanating from malicious sensor nodes present in the WSN are eliminated..

$$DR = \frac{\text{Number of correctly detected malicious nodes}}{\text{Total number of malicious nodes}}$$

3) Misdetection Ratio

Misdetection ratio (MR) refers to the ratio of misdetected nodes to the total number of all detections made by the scheme; this includes malicious nodes correctly detected and all misdetected nodes. Misdetected nodes belong to two classes: malicious nodes considered normal as well as normal nodes considered malicious. The misdetection ratio of the scheme should be as low as possible so as to reduce false positives.

$$MR = \frac{\text{Number of misdetected nodes}}{\text{Total number of detections}}$$

5.3) Simulation of Enhanced Weighted Trust Evaluation

5.3.1) Sensors Nodes Deployment

Once the sensor nodes are deployed, they form clusters and elect the forwarding node nearest to them as the cluster head and then they start forwarding the sensed data to it. The forwarding node does data aggregation and forward the aggregate value to the base station. In our simulation, the normal sensor nodes are represented by the green color, forwarding nodes are in blue whereas the sink node is magenta in color as shown in the figure below.

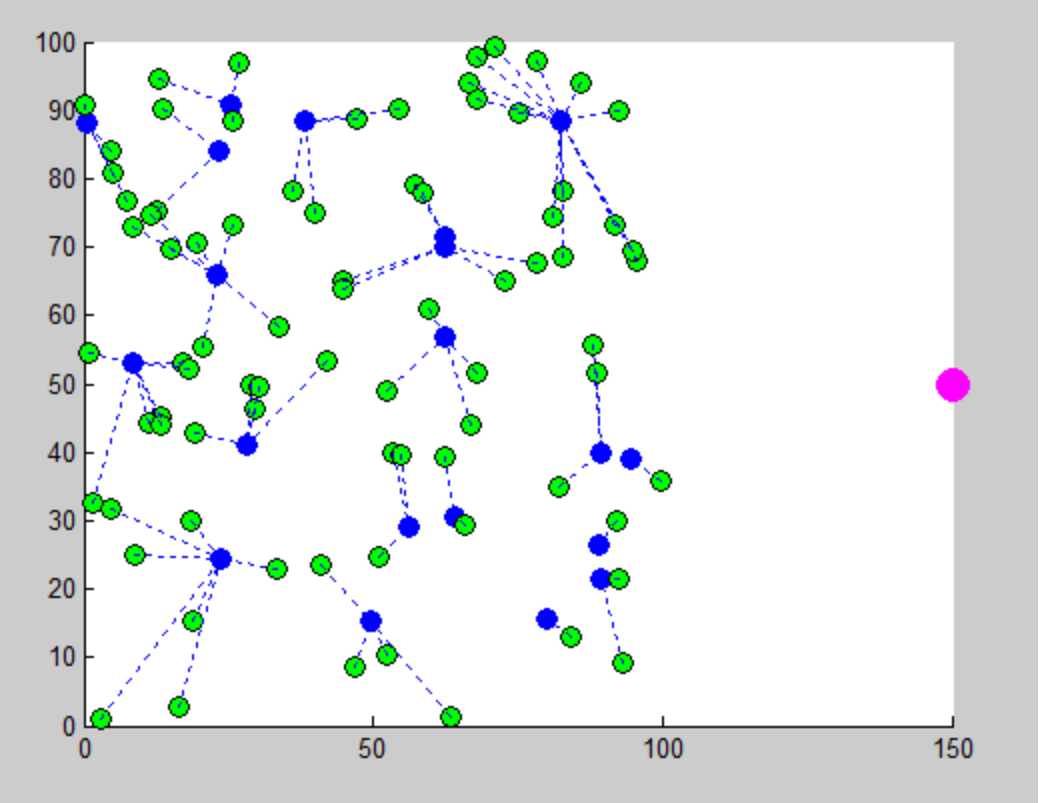


Figure 13: Sensor nodes and their cluster heads

5.4) Detection of Malicious nodes

In our simulation, n=100 sensor nodes are deployed randomly in an area 100 * 100 and a subset (p of n) of them are powerful nodes which act as the cluster heads; forwarding the aggregated data to the base station. Enhanced Weighted Trust Evaluation Scheme is implemented in the forwarding nodes. A sensor network of IR sensors to detect any human presence is simulated.

A percentage of the sensor nodes, m, are set malicious. For our simulation since n =100, the number of forwarding nodes, fno = (p*n) = (0.2 * 100) = 20. The malicious nodes are of two sets; the malicious ordinary sensor nodes and the malicious forwarding nodes. The malicious sensor nodes numbers are be given by;

Malicious forwarding nodes = m*(p*n).

Whereas;

Malicious ordinary sensor nodes= m*(n-(p*n))

For our first simulation, this translates to:

Malicious forwarding nodes =m*(p*n).

= 0.2 *(0.2 *100)

= 0.2 * 20

= 4

Malicious ordinary sensor nodes= m*(n-(p*n))

=0.2 * (100-(0.2*100))

=0.2 *(100-20)

=0.2 * 80

=16

Total number of malicious nodes = m*n

= 0.2 * 100

= 20

The network is as shown in the figure below:

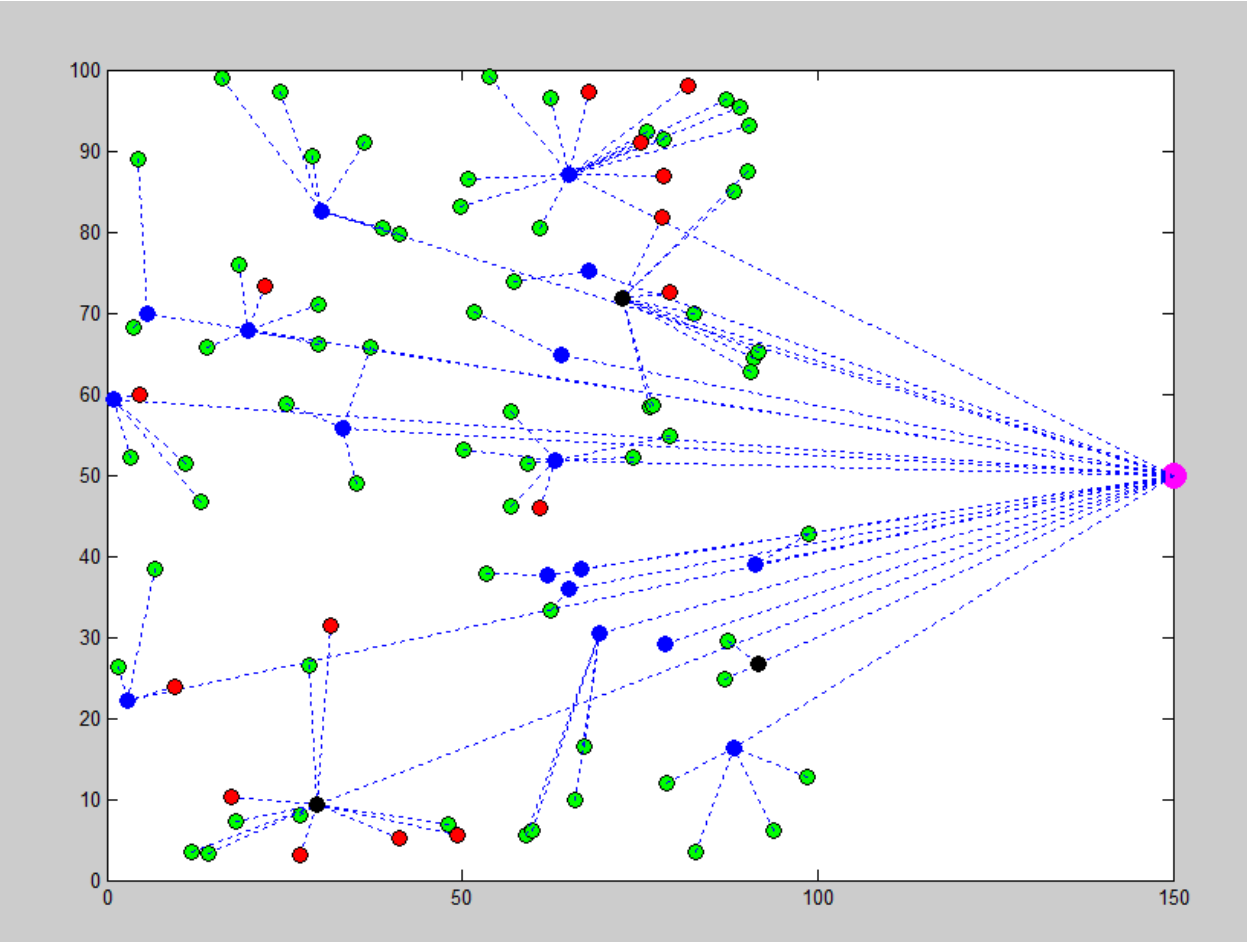


Figure 14: Simulated Sensor Network

The nodes in green color represent normal sensor nodes, nodes in blue are the normal forwarding sensor nodes, nodes in red represents malicious ordinary sensor nodes, nodes in black are the malicious forwarding nodes whereas the dotted blue line represents the flow of traffic from one node to another.

In our simulations, we assume that the normal sensor nodes in a cluster record and forward similar readings representing the actual happenings in the field, the malicious nodes however distort the data in order to mislead the decision made at the base station. The malicious ordinary sensor nodes sense and forward data that contradicts that of normal nodes whereas the malicious forwarding nodes transmit during non-transmission times thereby building up illegal traffic.

5.4.1 Sample Sensed Data

The sample simulated sensed data in our sensor network is as represented in the figure below.

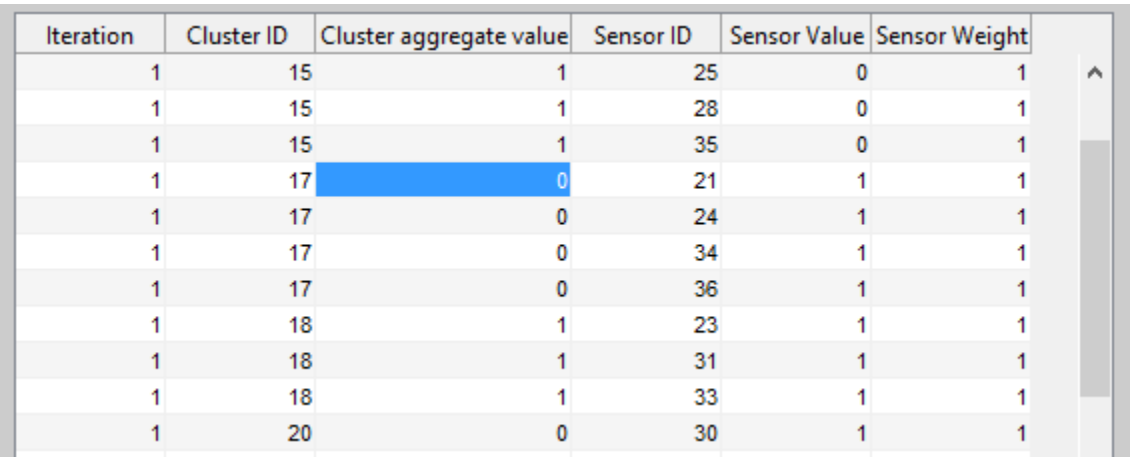
Iteration	Cluster ID	Sensor ID	Sensor Value
1	14	90	0
1	16	44	0
1	16	62	0
1	16	67	0
1	16	71	0
1	17	35	1
1	18	74	1
1	19	81	1
1	20	28	1
1	20	58	0
1	20	70	0
1	20	77	0
1	20	83	0
2	1	26	1
2	1	54	0
2	1	69	0
2	1	87	0
2	2	41	0
2	2	52	0
2	2	53	0
2	2	85	0
2	2	89	0
2	2	91	0
2	3	37	1
2	3	56	1

Figure 15: Simulated sample sense data

The figure above shows the iterations during the simulation run, the ids of the clusters involved, the ids of their member nodes and the data sensed by the individual sensor nodes. In iteration 1, cluster node 16; all the member nodes (44, 62, 67 and 71) sense and send a value of 0 to the forwarding node (FN) implying no alert whereas nodes in cluster 20 for the same iteration contradict each other. Sensor nodes 58, 70, 77 and 83 sense 0 meaning no alert whereas the malicious sensor node 28 sense and forward 1 implying an alert (presence of an intruder).

5.4.2 Sensor Node Weight Update

Each ordinary sensor node in the sensor network is assigned a weight of 1 which represent its reliability/confidence level. In our simulations, the malicious sensor node's weights is reduced per transmission by a weight penalty, pf = 0.2, until their weight reaches a set minimum weight threshold (0.6) from the initial value of 1. The weights are only reduced if and only if the malicious sensors are sending malicious traffic i.e. data that is not consistent with what other sensor nodes under the same FN (same cluster) are sending. When the weight of the sensor node reaches 0.6 it is considered malicious and isolated from the network; its weight updated to 0 meaning it is not reliable at all. The weight reductions are as shown in the figures below.



Iteration	Cluster ID	Cluster aggregate value	Sensor ID	Sensor Value	Sensor Weight
1	15	1	25	0	1
1	15	1	28	0	1
1	15	1	35	0	1
1	17	0	21	1	1
1	17	0	24	1	1
1	17	0	34	1	1
1	17	0	36	1	1
1	18	1	23	1	1
1	18	1	31	1	1
1	18	1	33	1	1
1	20	0	30	1	1

Figure 16: Weight reductions (Iteration 1)

The figure above shows that all the sensor nodes are assigned an initial weight value of 1. Sensor nodes 25, 28 and 35 of cluster 15 sensed 0 (no alert) whereas the majority of other nodes in the cluster as depicted by the cluster aggregate value sensed a 1 (alert). Sensor nodes 21, 24, 34 and 36 of cluster 17 sensed a 1 whereas other sensor nodes sensed a 0. These sensor nodes are considered malicious and their weights are subsequently reduced.

Iteration	Cluster ID	Cluster aggregate value	Sensor ID	Sensor Value	Sensor Weight
2	15	0	25	1	0.8000
2	15	0	28	1	0.8000
2	15	0	35	1	0.8000
2	17	0	21	1	0.8000
2	17	0	24	1	0.8000
2	17	0	34	1	0.8000
2	17	0	36	1	0.8000
2	18	0	23	0	1
2	18	0	31	0	1
2	18	0	33	0	1
2	20	1	30	0	0.8000

Figure 17: Weight reductions (Iteration 2)

The figure above shows that the weight (a representation of reliability) of the malicious ordinary sensor nodes has been reduced by a weight penalty of 0.2 to a weight of 0.8 from the initial value of 1.

Iteration	Cluster ID	Cluster aggregate value	Sensor ID	Sensor Value	Sensor Weight
3	15	0	25	1	0.6000
3	15	0	28	1	0.6000
3	15	0	35	1	0.6000
3	17	0	21	1	0.6000
3	17	0	24	1	0.6000
3	17	0	34	1	0.6000
3	17	0	36	1	0.6000
3	18	1	23	1	1
3	18	1	31	1	1
3	18	1	33	1	1
3	20	0	30	1	0.6000

Figure 18: Weight reductions (Iteration 3)

In the third iteration, their weights are reduced much further to 0.6 and they are subsequently isolated from the network. The sensed data from the malicious nodes will no longer contribute to the final decision/report of the cluster head since their weights have been reduced to zero as shown below.

Iteration	Cluster ID	Cluster aggregate value	Sensor ID	Sensor Value	Sensor Weight
4	15	0	25	1	0
4	15	0	28	1	0
4	15	0	35	1	0
4	17	1	21	1	0
4	17	1	24	1	0
4	17	1	34	1	0
4	17	1	36	1	0
4	18	0	23	0	1
4	18	0	31	0	1
4	18	0	33	0	1
4	20	1	30	1	0

Figure 19: Malicious nodes isolation from the network

5.5 Evaluation of Enhanced Weighted Trust Evaluation Scheme

5.5.1 Response Time

As discussed earlier, response time (RT) refers to the average number of cycles required to correctly detect a malicious node in the sensor network. A node is considered malicious in our scheme if its weight is reduced below a set minimum weight threshold. In our simulation we have set the minimum weight threshold as 0.6. Since the penalty factor by which the weight of each sensor node is reduced by is 0.2, it means that it takes an average of three iterations to detect the malicious sensor node assuming that it send wrong data continuously.

In one of our simulation runs, sensor node 32,33,66,29,23,27,21,22,28,31,35,34,24,26,30 and 36 are set malicious. Results shows that it takes the scheme an average of 3 cycles to correctly detect and isolate the malevolent nodes from the sensor network and their weights are set to 0.

Sensor ID	Iteration	Sensor Weight
32	3	0
33	3	0
66	3	0
29	3	0
23	3	0
27	3	0
21	3	0
22	3	0
28	3	0
31	3	0
35	3	0
34	3	0
24	3	0
26	3	0
30	3	0
36	3	0

Figure 20: Malicious Nodes Response Time

5.5.1.1 Effect of Minimum Weight Threshold and Penalty Factor on Response Time

The minimum weight threshold and weight penalty factor set have a direct effect on the response time of the scheme. A node is declared malicious when its weight reaches a certain pre-defined minimum weight threshold and the response time is concerned with the number of iterations the node goes through before it is detected. The penalty factor has a direct bearing on response time since the sensor node weight is gradually reduced by the set penalty factor each iteration that it sends wrong data.

When the minimum weight threshold is set to a lower value say 0.2 and the weight penalty factor remains 0.2 as the results below show; the response time is long. The number of cycles required to detect the malicious node increases to 5.

Sensor ID	Iteration	Sensor Weight
29	5	0
23	5	0
31	5	0
27	5	0
33	5	0
36	5	0
30	5	0
22	5	0
35	5	0
25	5	0
28	5	0
32	5	0
24	5	0
46	5	0
26	5	0

Figure 21: Malicious Nodes Response Time (Small minimum weight threshold)

When the minimum weight threshold is set to a higher value say 0.8 and the weight penalty factor remains 0.2 as the results below show; the response time is short. The number of cycles required to detect the malicious node reduces to 2.

Sensor ID	Iteration	Sensor Weight
21	2	0
23	2	0
26	2	0
29	2	0
32	2	0
36	2	0
28	2	0
31	2	0
22	2	0
30	2	0
33	2	0
24	2	0
25	2	0
35	2	0
27	2	0
34	2	0

Figure 22: Malicious Nodes Response Time (Large minimum weight threshold)

Changes in the penalty factor value also affect the response time. Increasing the weight penalty to a higher value of 0.6 from 0.2 and keeping the minimum weight threshold at 0.6. The results indicates that the response time reduces from 3 cycles to 2 cycles.

Sensor ID	Iteration	Sensor Weight
22	2	0
24	2	0
32	2	0
23	2	0
28	2	0
31	2	0
34	2	0
29	2	0
33	2	0
36	2	0
64	2	0
72	2	0
86	2	0
30	2	0
35	2	0
26	2	0

Figure 23: Malicious Nodes Response Time (Large penalty factor)

In general, assuming a constant penalty factor; as the minimum weight threshold decreases the response time increases. Also assuming a minimum weight threshold as size of the penalty factor increases, the response time decreases.

5.5.2 Detection Ratio

Detection Ratio (DR) is given by the ratio between the number of malicious nodes correctly detected by the scheme and the total number of malicious nodes present in the sensor network (set at the beginning of simulation).

In our simulation, the number of malicious nodes (m) in the network is given as a percentage of the total number of sensor nodes in the network. Ordinary sensor nodes are detected malicious in our scheme if they sent data that is not consistent with what the majority of the sensor nodes are sensing and reporting to the cluster head. Forwarding sensor nodes in the scheme are identified as malicious when they send data to the base station during non-transmission times.

In one of our simulation run the percentage of malicious nodes, is set to 0.2 (m=0.2). This means that:

Malicious nodes = m *n

= 0.2 * 100

=20

Where n = number of deployed sensor nodes.

The total number of malicious sensor nodes is 20 but there are two sets of malicious sensor nodes in the network i.e. malicious ordinary sensor nodes and malicious forwarding nodes.

Malicious forwarding nodes = m *(p*n)

=0.2 *(0.2 * 100)

=4

Where p= percentage of forwarding nodes in the sensor network.

Malicious ordinary sensor nodes = m *(n- (p*n))

=0.2 * (100 – (0.2* 100))

= 16

In our simulation we use the red color to represent the malicious ordinary sensor nodes and the black color to represent malicious forwarding nodes. Simulation results in one of the runs are as shown in the figure below:

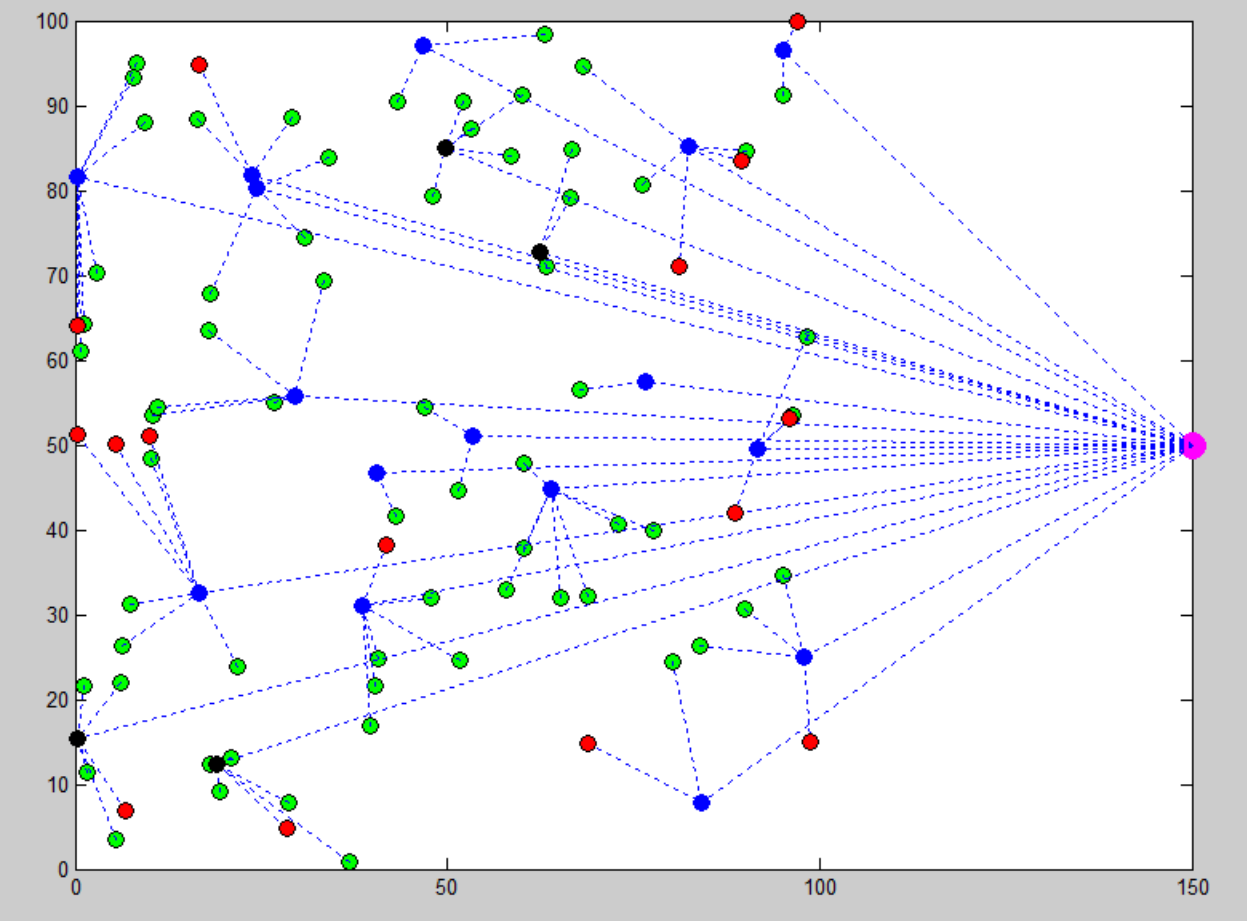
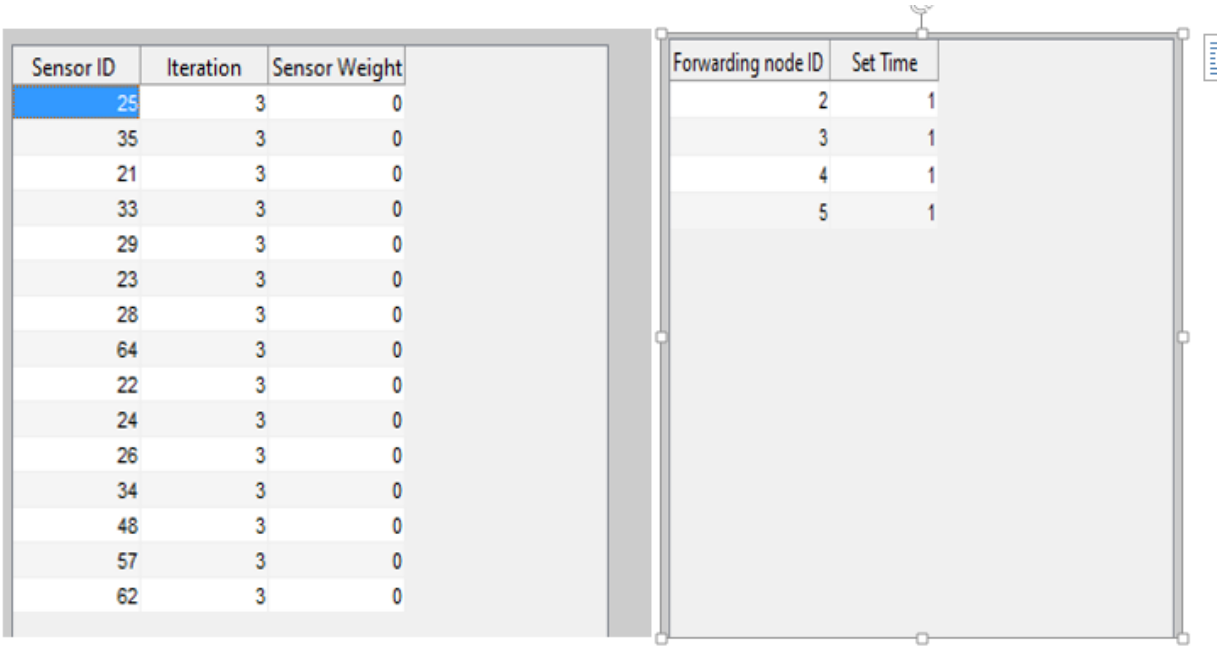


Figure 24: Malicious nodes

The number of detected malicious ordinary sensor nodes is 15 out of the 16 that had been set as malicious whereas all the malicious forwarding nodes are detected by the scheme.



Sensor ID	Iteration	Sensor Weight
25	3	0
35	3	0
21	3	0
33	3	0
29	3	0
23	3	0
28	3	0
64	3	0
22	3	0
24	3	0
26	3	0
34	3	0
48	3	0
57	3	0
62	3	0

Forwarding node ID	Set Time
2	1
3	1
4	1
5	1

Figure 25: Number of Detected Malicious Nodes

DR = Number of malicious nodes correctly detected
Total number of malicious nodes in the network.

DR = (15 + 4) / 20

= 0.95

The ideal value of detection ratio is 1. Some of the ordinary sensor nodes may not be detected by the scheme in cases where the number of the malicious sensor nodes in a cluster outnumber the normal sensor nodes. This is attributed to the fact that a node is established to be malicious if its reported value is not consistent with the reporting of the majority of sensor nodes in the cluster implying that if the malicious nodes are the majority, then their value will be regarded as the rightly sensed data.

5.5.2.1 Effect of the Number of Malicious Nodes to Detection Ratio

The detection ratio is affected by the total number of malicious nodes in the network in that when the majority of the sensor nodes are malicious, their values tilt the aggregate value of the cluster head towards the values sensed by the malicious nodes at the expense of the correct values reported by the normal nodes.

Iteration No	Cluster ID	Cluster Members	Normal Nodes	Sensed Value	Malicious Nodes	Malicious Sensed Value	Cluster Aggregate Value
1	11	7	2	0	5	1	1
1	12	2	1	0	1	1	1
1	13	3	2	0	1	1	0
1	14	4	3	0	1	1	0
1	15	8	2	0	6	1	1
1	16	3	1	0	2	1	1
1	17	4	3	0	1	1	0
1	18	4	3	1	1	0	1
1	19	9	3	1	6	0	0
1	20	7	1	1	6	0	0
2	1	3	1	0	2	1	1
2	2	2	1	1	1	0	0
2	3	1	0	1	1	0	0
2	4	2	1	1	1	0	1
2	5	3	0	1	3	0	0
2	6	3	2	1	1	0	1
2	7	0	0	0	0	0	0
2	8	6	3	0	3	1	0
2	9	5	1	1	4	0	0
2	10	4	2	1	2	0	0
2	11	7	2	1	5	0	0
2	12	2	1	0	1	1	1
2	13	3	2	1	1	0	1
2	14	4	3	1	1	0	1
2	15	8	2	0	6	1	1

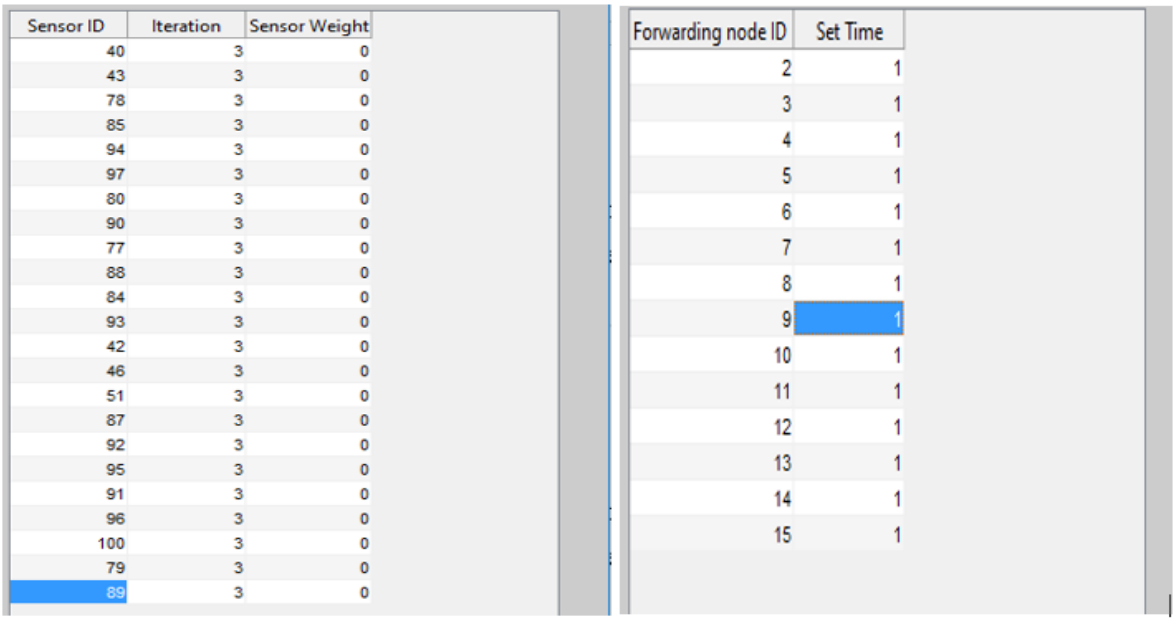
Figure 26: Majority Malicious Nodes in a Cluster

The effect of the majority of malicious of malicious sensor nodes affecting the aggregate value and subsequently the forwarding node report is illustrated in the above figure. Cluster 11, in iteration 1 has 7 sensor nodes as its members, 2 of them are normal nodes whereas the rest are malicious. The normal nodes report 0 whereas the malicious nodes report a 1 (an alert) and the cluster aggregate value is 1. This is because the malicious nodes outnumbers the normal nodes. The same effect can be seen in cluster 15, 19 and 20 in the same iteration.

The percentage of malicious nodes in the WSN can be increased and simulation can be used to illustrate its effect to the detection ratio. The results of one of the simulation runs in which the percentage of malicious nodes ‘m’ is set to 0.7 are shown below.

Malicious nodes = 0.7 * 100

= 70



Sensor ID	Iteration	Sensor Weight
40	3	0
43	3	0
78	3	0
85	3	0
94	3	0
97	3	0
80	3	0
90	3	0
77	3	0
88	3	0
84	3	0
93	3	0
42	3	0
46	3	0
51	3	0
87	3	0
92	3	0
95	3	0
91	3	0
96	3	0
100	3	0
79	3	0
89	3	0

Forwarding node ID	Set Time
2	1
3	1
4	1
5	1
6	1
7	1
8	1
9	1
10	1
11	1
12	1
13	1
14	1
15	1

Figure 27: Effect of Number of Malicious Nodes on Detection ratio (0.7)

The number of detected malicious ordinary sensor nodes is 23 out of the 56 that had been set as malicious whereas all the malicious forwarding nodes are detected by the scheme.

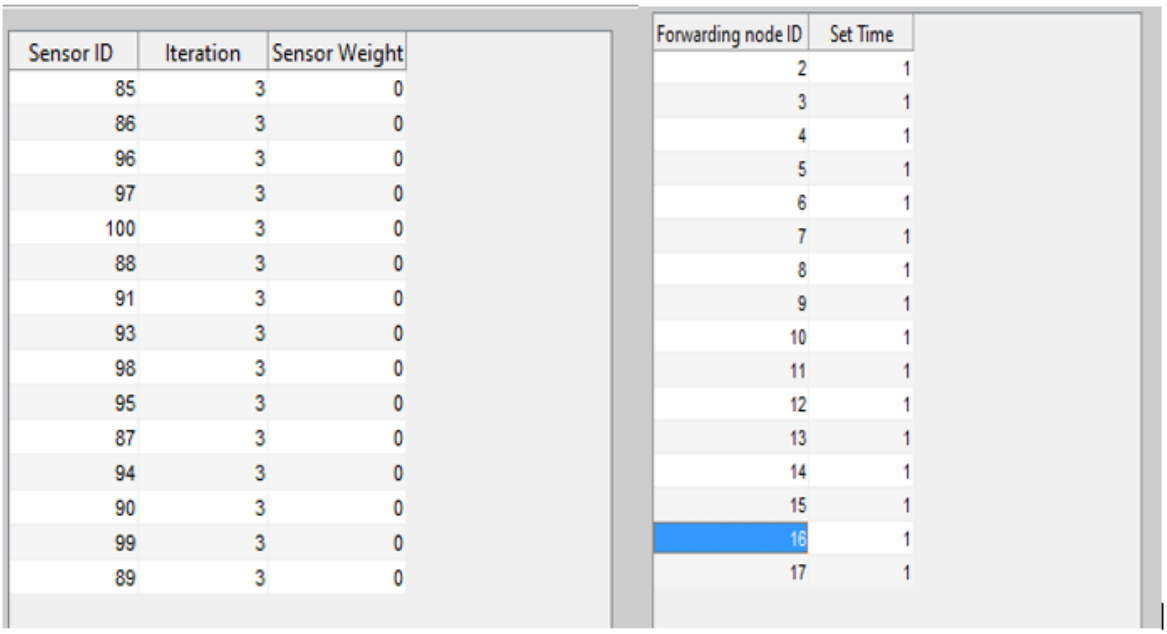
DR = (23 + 16) / 80

= 0.557

On increasing further the percentage of malicious node ‘m’ to 0.8, the following results are gotten.

Malicious nodes = 0.8 * 100

= 80



Sensor ID	Iteration	Sensor Weight
85	3	0
86	3	0
96	3	0
97	3	0
100	3	0
88	3	0
91	3	0
93	3	0
98	3	0
95	3	0
87	3	0
94	3	0
90	3	0
99	3	0
89	3	0

Forwarding node ID	Set Time
2	1
3	1
4	1
5	1
6	1
7	1
8	1
9	1
10	1
11	1
12	1
13	1
14	1
15	1
16	1
17	1

Figure 28: Effect of Number of Malicious nodes on Detection ratio (0.8)

The number of detected malicious ordinary sensor nodes is 15 out of the 64 that had been set as malicious whereas all the malicious forwarding nodes are detected by the scheme.

DR = (15 + 16) / 80

= 0.388

Malicious nodes	10	20	30	40	50	60	70	80	90
Detected Malicious nodes	10	19	24	31	33	35	39	31	25
Detection Ratio	1	0.95	0.8	0.78	0.66	0.58	0.56	0.39	0.28

Table 2: Malicious nodes (Both SNs and FNs) and Detection Ratio

The results in the table above are from a sensor network in which the number of deployed sensor nodes is one hundred (n =100).

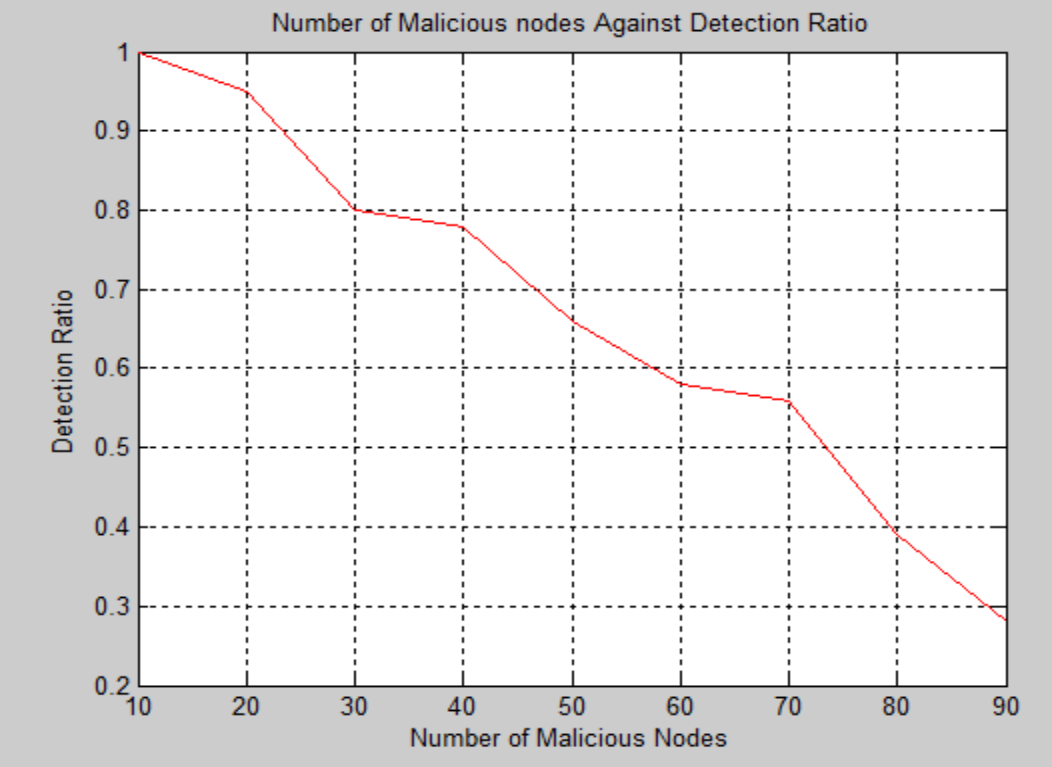


Figure 29: Number of Malicious Nodes against Detection Ratio

The graph above illustrates that as the number of malicious nodes increases the detection ratio decreases.

For the same set of simulations; considering malicious forwarding nodes (FNs) only, the results are as follows:

Malicious FNs	2	4	6	8	10	12	14	16	18
Detected Malicious FNs	2	4	6	8	10	12	14	16	18
Detection Ratio	1	1	1	1	1	1	1	1	1

Table 3: Malicious FNs nodes and Detection Ratio

The difference in the detection ratios when both malicious SNs and FNs are considered vis a vis when only malicious FNs are considered is due to the effect of false positives that is attributed to cases where the number of malicious nodes exceeds legitimate nodes in a cluster under an FN. This influences the sensor nodes aggregate data value.

This effect does not affect the detection of malicious FNs since the scheme relies on trapping malicious behavior during non-transmission times to capture malicious FNs as opposed to the use of sensor node reported value and the cluster head aggregate value.

5.5.3 Misdetection Ratio

Misdetection ratio (MR) is given by the ratio between misdetected nodes and the total number of detections made by the scheme.

In one of our simulation run, sensor node 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36 are set malicious. The scheme detected the following nodes as malicious; 23, 21, 27, 30, 31, 35, 32, 22, 24, 36, 28, 65, 66, 34, 26 whereas node 25, 29 and 33 were detected as normal. From the results above we can see that normal nodes 65 and 66 were misdirected as malicious and malicious node 33 was misdetected as normal. The total number of misdetections was 5

$$MR = \frac{\text{Number of misdetected nodes}}{\text{Total number of detections}}$$

$$= 5/18$$

$$= 0.278$$

The misdetections can be attributed to the clusters in which the number of malicious nodes exceeds the number of normal working nodes. Node 33 is in cluster 1 alone hence its value is the one treated as the aggregate value at the forwarding node, whereas sensor node 25, 29, 65 and 66 are all in cluster 13. Sensor node 25 and 29 send a 1 (an alert) whereas sensor node 65 and 66 sent 0 indicating no alert; when the aggregate is performed the value is 1 and therefore the normal nodes 65 and 66 are regarded as malicious.

CHAPTER SIX: CONCLUSION

6.1 Introduction

The research project had several objectives set and were to be achieved over the course of the project implementation. The objectives were; first, Investigate wireless sensor networks security design issues and challenges and the various attacks that adversaries can launch via malicious nodes. Secondly, design and implement a prototype of an enhanced malicious node detection scheme by amalgamating the Weighted Trust Evaluation Scheme and Stop Transmit and Listen (STL) scheme and thirdly evaluate malicious node detection and isolation by analyzing the response time, detection ratio and the misdetection ratio of the above-proposed scheme.

The research project delved into detailed wireless sensor network security design issues and challenges such as limited energy and computational capabilities, unreliable wireless communication medium and the hostile deployment environment. These design issues and challenges render the employment of existing security mechanisms inadequate and inefficient. This coupled with the fact that owing to the constrained resources inherent in the sensor node, most wireless sensor network protocols tend to assume a high level of trust between the communicating sensor nodes so as to eliminate the authentication overhead creates the danger of adversaries introducing malicious nodes to the sensor network or manipulate existing ones and subsequently using them to propagate a wide range of attacks such as sinkhole attack, Sybil attack, black hole attacks, wormhole attack, HELLO flooding attacks and Denial-of-Service attacks. Detection and exclusion of such malicious nodes is crucial.

The second objective was met via the design and development of the scheme. An enhanced weighted trust evaluation scheme was designed using system model diagrams, control flow diagrams, flow charts and pseudo codes to show how the algorithm works and to depict the flow of our algorithm's logic. The algorithm design was then developed and implemented using the MATLAB language. The MATLAB platform was chosen due to its advantage of quick prototyping and fast computational engine.

The third objective which involved testing and evaluating the performance of our scheme was also accomplished. The evaluation involved verifying that the initial model requirements specification have been met by the output of the implemented algorithm. The main features tested include generation of short response timings, high detection rate and low misdetection ratio. It was also demonstrated that the ratio of malicious nodes to total number of sensor nodes in WSN has an effect on the detection ratio of the scheme. Results obtained were plotted graphically and it was found out that the ratio of malicious nodes to the total nodes deployed directly affect the detection ratio in that as the number of malicious nodes in the network increases, the detection ratio decreases. The algorithm can be thus be said to be suitable to detect malicious nodes in WSNs in which the ratio of malicious nodes to the total number of nodes is less than 0.5.

Simulation results obtained also indicated that varying the minimum weight threshold and the weight penalty factor has an effect to the response time of the scheme. These results show that the value of the pre-defined minimum weight threshold and the penalty factor have a direct effect on the response time. As the weight threshold is decreased, the response time increases and vice versa. Also as the penalty factor decreases, assuming that the predefined minimum weight threshold is kept constant, the response time tends to be high since the weight reduction tends to be low.

Generally in this research project, we discussed wireless sensor networks (WSN) and the detection and isolation of malicious sensor nodes in a bit to secure the WSN from attacks that can be propagated by the adversary via the malicious nodes. We proposed an enhanced Weighted Trust Evaluation (WTE) based detection algorithm to detect and isolate malicious nodes in wireless sensor network. The fundamental operation of the algorithm is that a weight representing the confidence level of a sensor node is assigned to every sensor node and also the forwarding nodes are assigned transmission time-slots. The weights of sensor nodes reporting wrong data to mislead the network are gradually decreased. They are detected as malicious and isolated from the network when their weights reach a pre-defined minimum allowed weight threshold. Malicious forwarding nodes are detected in the WSN when they send data to the base station during non-transmission times, the traffic is regarded illegal.

Extensive simulation is performed using MATLAB. Simulation results show that our WTE based algorithm is able to detect and isolate malicious nodes in WSNs. The solution can be applied to a flexible number of sensor nodes that operate under a cluster head, it thus achieve good scalability with a reasonable detection rate and short response time.

6.2 Challenges and Assumptions

Several challenges and limitations were faced while undertaking the project, at the same time a number of assumptions were made.

The simulation tool chosen, MATLAB, though it offers the advantage of quick prototyping, fast computational engine, rich computation and visualization features it is limited in that it lacks built-in routines for wireless sensor networks (WSN). This necessitated that we built our own for the project.

The issue of false positives in some clusters where compromised nodes outnumber the legitimate nodes. In such cases, the normal nodes were treated as malicious and malicious ones treated as normal. This leads to an increase in misdetection ratio.

The assumptions made include:

i) The access point (sink) is cannot be compromised by an adversary otherwise the attacker can launch any possible attack against the WSN upon taking control of the access point (AP).
ii) The communication path over which the sensed values are propagated from the source sensor to the forwarding node and then to the base station is considered to be error-free so the data reaches to the base station without modification enroute.
iii) The bandwidth of the wireless channel used in transmission is not limited so contention issues are reduced.

6.3 Future Work

There is a lot that we desired to have accomplished in this project but due to the aforementioned limitations and challenges we could not achieve them. First, an insecure access point (sink) can be a gateway to an array of attacks once an adversary takes control of it. We have assumed in our work that it cannot be compromised, future work should look into ways of securing the sink node from being compromised or other nodes being able to detect that it has been compromised. Another area of improvement would be detection and isolation of malicious nodes even in clusters in which the number of compromised nodes exceeds the number of normal nodes. This would be a key improvement to reduce the misdetection ratio.

REFERENCES

Alam, D. . S. & Debashis, 2014. ANALYSIS OF SECURITY THREATS IN WIRELESS SENSOR NETWORK. *International Journal of Wireless & Mobile Networks (IJWMN),* Volume 6.

Ali , Q. . I., 2012. Simulation Framework of Wireless Sensor Network (WSN) Using MATLAB/SIMULINK Software. In: s.l.:s.n., pp. 263-264.

Das, R., Purkayastha, D. B. S. & Das, D. P., 2002. Security Measures for Black Hole Attack in MANET: An Approach. *Proceedings of Communications and Computer.*

Karuppiah, A. B. & Rajaram, S., 2014.. False Misbehavior Elimination of Packet Dropping Attackers during Military Surveillance using WSN. *Advances in Military Technology*, 9(1).

Abdullah, M. I., Rahman, M. M. & Roy, M. C., 2015. Detecting Sinkhole Attacks in Wireless Sensor Network using Hop Count. *I. J. Computer Network and Information Security*, p. 51.

Akykildiz, . I. . F., Su, W., Sankarasubramaniam, Y. & Cayirci, E., 2002. A Survey on Sensor Networks. *IEEE Communication Magazine.*

Alajmi, N., July 2014. Wireless Sensor Networks Attacks and Solutions. *International Journal of Computer Science and Information Security (IJCSIS)* , 12(7).

Atakli, I. M. et al., 2008. Malicious Node Detection in Wireless Sensor Networks. *The Symposium on Simulation of Systems Security (SSSS'08), Ottawa, Canada,* p. 838.

Bao, F., Chen, I.-R., Chang, M. & Cho, J.-H., 2011. *Trust-Based Intrusion Detection in Wireless Sensor Networks.* Kyoto, Japan, s.n.

Cannon, B. J., May 2016. Terrorists, Geopolitics and Kenya's Proposed Border Wall with Somalia. *Journal of Terrorism Research,* 7(2), pp. 27-28.

CHELLI, K., 2015. Security Issues in Wireless Sensor Networks:Attacks and Countermeasures. *Proceedings of the World Congress on Engineering 2015,* Volume 1, pp. 1-6.

Curiac, D.-I.et al., 2007. *Malicious Node Detection in Wireless Sensor Networks Using an Autoregression Technique.* Athens, Greece, s.n.

Hu, H. et al., 2009. Weighted trust evaluation-based malicious node detection for wireless sensor networks. *Int. J. Information and Computer Security,* 3(2), p. 148.

Hussain, M. Z., Singh, . M. P. & Singh, R. K., April, 2013. Analysis of Lifetime of Wireless Sensor Network. *nternational Journal of Advanced Science and Technology,* Volume 53, p. 1.

Hu, Y.-C., Perrig, A. & Johnson, D. B., 2003. *Packet Leashes: A Defense against Wormhole Attacks in Wireless Networks.* s.l., s.n.

López, E. E. et al., 2005. *Simulation Tools for Wireless Sensor Networks.* Cartagena, Spain., s.n., pp. 5-9.

Nayyar, A. & Singh, R., 2015. A Comprehensive Review of Simulation Tools for Wireless Sensor Networks (WSNs). *Journal of Wireless Networking and Communications,* Issue 2167-7328, pp. 110-113.

Nidharshini, T. & Janani, V., December 2012.. Detection of Duplicate Nodes in Wireless Sensor Networks Using Sequential Probability Ratio Testing. *International Journal of Advanced Research in Computer and Communication Engineering,* 1(10).

Padmavathi, . D. . G. & Shanmugapriya, M. D., 2009. A Survey of Attacks, Security Mechanisms and Challenges in Wireless Sensor Networks. *International Journal of Computer Science and Information Security,,* Volume 4.

Parno, B., Perrig, A. & Gligor , V., 2005. Distributed Detection of Node Replication Attacks in Sensor Networks. *Proceedings of the IEEE Symposium on Security and Privacy (S&P'05).*

Pathan, A.-S. K., 2010. DENIAL OF SERVICE IN WIRELESS SENSOR NETWORKS: ISSUES AND CHALLENGES. In: *Advances in Communications and Media Research.* s.l.: Nova Science Publishers, Inc..

PERRIG, A., STANKOVIC, J. & WAGNER, D., June 2004. SECURITY IN WIRELESS SENSOR NETWORKS. *COMMUNICATIONS OF THE ACM,* 47(6).

Raymond, D. R. & Midkiff, S. F., 2008. Denial-of-Service in Wireless Sensor Networks: Attacks and Defenses. *IEEE Pervasive Computing,* Volume 7.

Sathyamoorthi, T., Vijayachakaravarthy, D., Divya, R. & Nandhini, M., 2014. A SIMPLE AND EFFECTIVE SCHEME TO FIND MALICIOUS NODE IN WIRELESS SENSOR NETWORK. *International Journal of Research in Engineering and Technology,* 03(02).

SHARMA, R. & TRIPATHI, N., April 2015. Comprehensive Review on Wireless Sensor Networks. *ORIENTAL JOURNAL OF COMPUTER SCIENCE & TECHNOLOGY,* 8(1), pp. 59-64.

SOHRABY, K., MINOLI, D. & ZNATI, T., 2007. *WIRELESS SENSOR NETWORKS: Technology, Protocols, and Applications.* Hoboken, New Jersey.: John Wiley & Sons, Inc..

Soomro, S. A., Soomro, S. A., Memon, A. G. & Baqi, . A., 2008. Denial of Service Attacks in Wireless Ad-hoc Networks. *Journal of Information & Communication Technology,* Volume 04, pp. 01-10.

Sumathi , K. & Venkatesan, D. M., 2014. A Survey on Detecting Compromised Nodes in Wireless Sensor Networks. *(IJCSIT) International Journal of Computer Science and Information Technologies,* Volume 5, pp. 7720-7722.

Sung, . Y. L. & Choi, Y.-H., 2013. Malicious Node Detection Using a Dual Threshold in Wireless Sensor Networks. *Journal of Sensor and Actuator Networks.*

Virmani, D., Soni, A., Chandel, S. & Hemrajani, M., 2014. Routing Attacks in Wireless Sensor Networks: A Survey. *Bhagwan Parshuram Institute of Technology, India.*

Woo, , A. & Culler, . D., 2001. A Transmission Control Scheme for Media Access in Sensor Networks. *Proceedings of the Seventh Annual ACM/IEEE International Conference on Mobile Computing and Networking, MobiCom, Rome, Italy.*

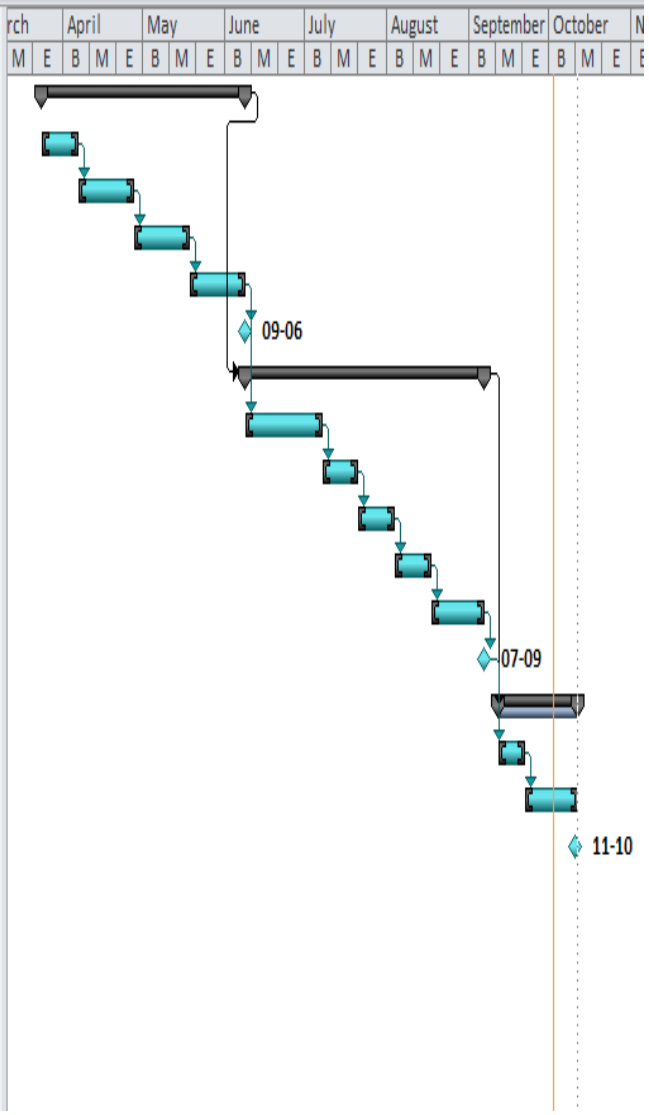
Yang, Y., Wang, X., Zhu, S. & Cao, G., 2007. *Distributed Software-based Attestation for Node Compromise Detection in Sensor Networks.* Pennsylvania, s.n.

Y-C , H. & Perrig, A., 2004. A Survey of Secure Wireless Ad Hoc Routing. *IEEE Security and Privacy*.

Zhao, S., Tepe, K., Seskar, I. & Raychaudhuri, D., March 2013. Routing Protocols for Self-Organizing Hierarchical Ad-Hoc Wireless Networks. *Proceedings of the IEEE Sarnoff Symposium, Trenton, NJ,*.

APPENDIX I: PROJECT SCHEDULE

	Task Name	Duration	Start	Finish
1	**Proposal Preparation and Presentation**	**55 days**	**Thu 24-03-16**	**Thu 09-06-16**
2	Topic Selection	10 days	Thu 24-03-16	Wed 06-04-16
3	Literature Selection and Review	15 days	Thu 07-04-16	Wed 27-04-16
4	Proposal Writing	15 days	Thu 28-04-16	Wed 18-05-16
5	Proposal Review	15 days	Thu 19-05-16	Wed 08-06-16
6	Presentation	0 days	Thu 09-06-16	Thu 09-06-16
7	**Undertaking Research**	**64 days**	**Thu 09-06-16**	**Wed 07-09-16**
8	Simulation and Modeling Tool Learning	21 days	Thu 09-06-16	Thu 07-07-16
9	Model Development	9 days	Fri 08-07-16	Wed 20-07-16
10	Model testing	10 days	Thu 21-07-16	Wed 03-08-16
11	Data Collection from the Model	10 days	Thu 04-08-16	Wed 17-08-16
12	Data Analysis and Documentation	14 days	Thu 18-08-16	Tue 06-09-16
13	Results and Model Presenation	0 days	Wed 07-09-16	Wed 07-09-16
14	**Project Report Preparation**	**22 days**	**Mon 12-09-16**	**Tue 11-10-16**
15	Report Writing	8 days	Mon 12-09-16	Wed 21-09-16
16	Report Review	14 days	Thu 22-09-16	Tue 11-10-16
17	Report Presentation	0 days	Tue 11-10-16	Tue 11-10-16



APPENDIX II: PROJECT BUDGET

	Expense	**Amount**	
	Internet Charges	12,000.00	
	Transport Charges	8,000.00	
	Stationery & Printing	10,000.00	
	Report Binding	6,000.00	
	Miscellaneous	10,000.00	
	Total	**46,000.00**	