ISSN 1913-8989 (Print)
ISSN 1913-8997 (Online)

Computer and Information Science

CANADIAN CENTER OF SCIENCE AND EDUCATION ®

Vol. 16, No. 4 November 2023

COMPUTER AND INFORMATION SCIENCE

An International Peer-reviewed and Open Access Journal for Computer and Information Science

Computer and Information Science is an international, double-blind peer-reviewed, open-access journal. CIS is published by the Canadian Center of Science and Education in both print and online versions. CIS aims to promote excellence through dissemination of high-quality research findings, specialist knowledge, and discussion of professional issues that reflect the diversity of this field.

The scopes of the journal include:

Programming Theory
Information and Information Systems
Computer Systems and Organisations
Machines, Languages, and Computation

The journal is included in:

DBLP (2008-2019)
ERA
Google Scholar
The Index of Information Systems Journals
Ulrich's
WorldCat

Open-Access Policy

We follow the Gold Open Access way in journal publishing. This means that our journals provide immediate open access for readers to all articles on the publisher‟s website. The readers, therefore, are allowed to read, download, copy, distribute, print, search, link to the full texts or use them for any other lawful purpose. The operations of the journals are alternatively financed by article processing charges paid by authors or by their institutions or funding agencies.

Copyright Policy

Copyrights for articles are retained by the authors, with first publication rights granted to the journal/publisher. Authors have rights to reuse, republish, archive, and distribute their own articles after publication. The journal/publisher is not responsible for subsequent uses of the work.

Submission Policy

Submission of an article implies that the work described has not been published previously (except in the form of an abstract or as part of a published lecture or academic thesis), that it is not under consideration for publication elsewhere, that its publication is approved by all authors and tacitly or explicitly by the authorities responsible where the work was carried out. However, we accept submissions that have previously appeared on preprint servers (for example: arXiv, bioRxiv, Nature Precedings, Philica, Social Science Research Network, and Vixra); have previously been presented at conferences; or have previously appeared in other “non-journal” venues (for example: blogs or posters). Authors are responsible for updating the archived preprint with the journal reference (including DOI) and a link to the published articles on the appropriate journal website upon publication.

iThenticate® Professional Plagiarism Prevention

The publisher and journals have a zero-tolerance plagiarism policy. We check the issue using the plagiarism prevention tool and a reviewer check. All submissions will be checked by iThenticate before being sent to reviewers.

CIS accepts both Online and Email submission. The online system makes readers to submit and track the status of their manuscripts conveniently. For any questions, please contact cis@ccsenet.org.

Computer and Information Science; Vol. 16, No. 4; 2023
ISSN 1913-8989 E-ISSN 1913-8997
Published by Canadian Center of Science and Education

EDITORIAL TEAM

Computer and Information Science; Vol. 16, No. 4; 2023
ISSN 1913-8989 E-ISSN 1913-8997
Published by Canadian Center of Science and Education

CONTENTS

The Evolution of Information Security Strategies: A Comprehensive Investigation of INFOSEC Risk Assessment in the Contemporary Information Era 1
Abdullah Al Hayajneh, Hasnain Nizam Thakur, Kutub Thakur

Electronic Health System Integration Framework for Secure M-Health Services: A Case of University of Nairobi Hospital 21
Samuel Nandasaba, Gregory Wanyembi, Geoffrey Mariga Wambugu

Malware Investigation and Analysis for Cyber Threat Intelligence: A Case Study of Flubot Malware 47
Uchenna J. Nzenwata, Frank Uchendu, Haruna Ismail, Eluwa M. Jumoke, Himikaiye O. Johnson

Proposal of a Visualization System for a Hierarchical Clustering Algorithm: The Visualize Proximity Matrix 65
Sulaiman Abdullah Alateyah

Applying AI in the Healthcare Sector: Difficulties 78
Abdussalam Garba, Muhammad Baballe Ahmad, Mukhtar Ibrahim Bello

Reviewer Acknowledgements for Computer and Information Science, Vol. 16, No. 4 84
Chris Lee

Computer and Information Science; Vol. 16, No. 4; 2023
ISSN 1913-8989 E-ISSN 1913-8997
Published by Canadian Center of Science and Education

The Evolution of Information Security Strategies: A Comprehensive Investigation of INFOSEC Risk Assessment in the Contemporary Information Era

Abdullah Al Hayajneh[1], Hasnain Nizam Thakur[1], Kutub Thakur[1]

[1] Professional Security Studies Department, New Jersey City University, Jersey City, NJ, USA

Correspondence: Abdullah Al Hayajneh, Professional Security Studies Department, New Jersey City University, Jersey City, NJ, USA.

Received: September 14, 2023 Accepted: October 21, 2023 Online Published: November 20, 2023

doi:10.5539/cis.v16n4p1 URL: https://doi.org/10.5539/cis.v16n4p1

Abstract

In the contemporary era marked by the extensive utilization of data, information systems have been extensively embraced by global organizations and also hold a pivotal position in national defense and various other domains. The growing interconnectedness between individuals and diverse information systems has resulted in an intensified emphasis on the evaluation of potential risks. The mitigation of these dangers extends beyond simple technological solutions and includes established standards, legal structures, and policies, adopting a complete approach based on safety engineering concepts. This study aims to develop a robust framework for the harmonization of Information Technology Security Standards. It will explore prevalent techniques for conducting risk assessments and differentiate between quantitative and qualitative approaches to evaluation. Moreover, this study illustrates the combination of quantitative and qualitative evaluation methodologies, providing a comprehensive framework for the analysis and design of risk assessment. In addition, this study advances our understanding of INFOSEC risk assessment and contributes to the advancement of more efficient information security strategies by sharing global perspectives, addressing challenges in classification, clarifying the incorporation of Information Security Management Systems (ISMS), and highlighting the significance of Artificial Intelligence in the domain of Information Security (INFOSEC).

Keywords: information Security Systems (INFOSEC), Information Security, Risk Assessment, Quantitative and Qualitative approaches, Artificial Intelligence

1. Introduction

1.1 Information Risk

The necessity for information security arises from the inherent risks associated with the utilization of technology in the management of information. From a comprehensive perspective, information is susceptible to unauthorized disclosure, which has the potential to compromise its confidentiality; it is vulnerable to unauthorized alterations, posing risks to its integrity; and it is also susceptible to destruction or loss, which could undermine its availability. The financial impact on the proprietor resulting from the loss of a valuable information asset, whether explicitly apparent or not, is a consistent outcome. Such financial repercussions can manifest both directly, influencing the inherent worth of the information asset itself, and indirectly, giving rise to diverse consequences encompassing service disruptions, reputational harm, erosion of competitive advantage, legal accountabilities, and other contributing factors (Blakley et al., 2001). The application of risk assessment principles to the realm of information security has been evident since the 1960s. Initially, the focus of risk assessment was primarily directed toward matters of confidentiality. However, during the transition period spanning from the late 1980s to the mid-1990s, the safeguarding of computers and network security emerged as prominent concerns within the ambit of risk assessment (Blakley et al., 2001), (Behnia et al., 2012). Notably, experts introduced additional dimensions such as integrity and availability, thereby enhancing the comprehensiveness of information security considerations. The swift and profound evolution of the internet and mobile communication technologies has engendered a ubiquitous challenge for global denizens of the digital realm – that of effectively preserving the confidentiality of personal information. A pivotal milestone in the domain of international information security standards occurred in 2013 with the revision of the "Information Technology – Security Techniques – Code of Practice for Information

Security Management," a publication endorsed by ISO/IEC (Behnia et al., 2012).

In the pursuit of enhancing the precision and effectiveness of risk assessment, the development of supplementary methodologies has been undertaken to facilitate the evaluation of security risks. A notable illustration of such advancements is exemplified by COBRA (Consultative, Objective and Bi-functional Risk Analysis), devised by British C & A Systems Security Ltd. in the year 1991 (Schmidt, 2023). Within this framework, the enterprise leverages collected questionnaire data to appraise the security status of an organization within the context of the risk assessment report. Another noteworthy tool, CRAMM (CCTA Risk Analysis and Management Method), stands as an expansive and adaptable mechanism tailored for the strategic substantiation of prioritized countermeasures at a managerial echelon. Notably, the optimal deployment of CRAMM necessitates the engagement of proficient and seasoned practitioners to ensure efficacious outcomes (Schmidt, 2023), (Fredriksen et al., 2002). Furthermore, the arena of risk assessment has witnessed the emergence of CORA (Cost-of-Risk Analysis), introduced by International Security Technology, Inc. (ICT). CORA employs meticulously gathered data pertaining to the spectrum of threats, functions, assets, and vulnerabilities inherent in said functions and assets. Subsequently, these factors are harmonized with the corresponding threat profiles, thereby facilitating the quantitative estimation of potential consequences ensuing from identified risks (Behnia et al., 2012), (Schmidt, 2023), (Fredriksen et al., 2002).

Modern risk assessment goes beyond what was previously limited to a technological study report. Its scope has expanded to include a detailed analysis of many different aspects, such as technology infrastructure, technical examination of systems and physical hardware, human resource management, and a nuanced characterization of the benefits and drawbacks of current methods. The modern paradigm of information security risk assessment demands a holistic approach that goes beyond narrow focal points. However, it is important to recognize that results from similar situations may show variances that might be ascribed to the various criteria created by various national committees.

This paper focuses on the following key areas:

a) A meticulous examination of current global research efforts related to risk assessment alongside a succinct introduction to prevalent evaluative methodologies.

b) An investigation into the classification and valuation of essential risk components such as assets, threats, and vulnerabilities, along with a comprehensive explanation of the challenges inherent to risk assessment and the resulting evolving requirements.

c) A clear explanation of the standards governing Information Security Management Systems (ISMS), along with an exploration of how these standards are put into practice and the presentation of a comprehensive framework detailing the stages of evaluation.

1.2 Risk Assessment of Information System

The field of information security encompasses the protection of the confidentiality, integrity, and availability of information. Furthermore, it covers additional properties such as authenticity, accountability, non-repudiation, and reliability (ISO/IEC JTC1 SC 27, 2013). Within the framework of risk assessment talks, there is a tendency for financial concerns, assets, and threats to assume a position of priority. Examining a scenario from the perspective of risk assessment: let us contemplate a circumstance in which I experience a financial loss of one hundred dollars as a consequence of my own carelessness, leading to a lack of finances for the purpose of having dinner. Within this particular framework, the term '100 dollars' serves as a representation of an asset, while 'negligence' denotes a state of vulnerability. Additionally, 'stealing' is characterized as a potential danger, 'loss of money' is conceptualized as a manifestation of risk, and the subsequent inability to afford supper is regarded as the resultant effect. Significantly, vulnerability and threat are identified as causal elements, whereas risk and effect are regarded as end results. The process of conducting a security risk assessment entails the examination of factors such as Confidentiality, Integrity, and Availability (CIA), in addition to other safety considerations, during the various stages of information processing, transmission, and storage within a system. The objective of this analytical procedure is to assess the level of security by incorporating the elements of vulnerability, threat, risk, and the subsequent consequences (ISO/IEC JTC1 SC 27, 2013). The primary objective of information security management is to mitigate risk to a degree that is financially feasible and in accordance with established norms through the implementation of comprehensive and cohesive risk control methodologies.

1.3 Meaning of Risk Assessment

Risk assessment serves as the fundamental basis for ensuring the security of information systems. The ability of an executor to effectively assess, manage, and mitigate risks is contingent upon their accurate and comprehensive

understanding of the various hazards involved (Stoneburner et al., 2002). Moreover, it is impractical to strive for absolute safety or eradicate all risks inside a system without taking into account associated costs. Information security necessitates being directed by social demands and prioritizing numerous crucial aspects. This entails doing a thorough risk assessment followed by implementing appropriate control measures. Organizations use risk assessment to determine the extent of the potential threat and the risk associated with an IT system throughout its Software Development Life Cycle SDLC. Risk is a function of the likelihood of a given threat-source's exercising a particular potential vulnerability and the resulting impact of that adverse event on the organization (Stoneburner et al., 2002). To determine the likelihood of a future adverse event, threats to an IT system must be analyzed in conjunction with the potential vulnerabilities and the controls in place for the IT system. Impact refers to the magnitude of harm that could be caused by a threat's exercise of vulnerability. The level of impact is governed by the potential mission impacts and, in turn, produces a relative value for the IT assets and resources affected (e.g., the criticality and sensitivity of the IT system components and data) (Stoneburner et al., 2002).

The risk assessment technique consists of nine fundamental components, including: The first step involves system characterization, followed by the identification of threats in the second step. The third step entails identifying vulnerabilities, while the fourth step involves doing a control analysis. The fifth step focuses on determining the likelihood of occurrence, and the sixth step involves conducting an impact analysis. The seventh step entails determining the overall risk, followed by the eighth step, which involves making control recommendations. Finally, the ninth step involves documenting the results obtained from the previous steps. Steps 2, 3, 4, and 6 have the potential to be executed concurrently subsequent to the completion of Step 1. Figure 1 illustrates the sequential progression of these processes, together with the corresponding inputs and outputs associated with each step (Stoneburner et al., 2002)

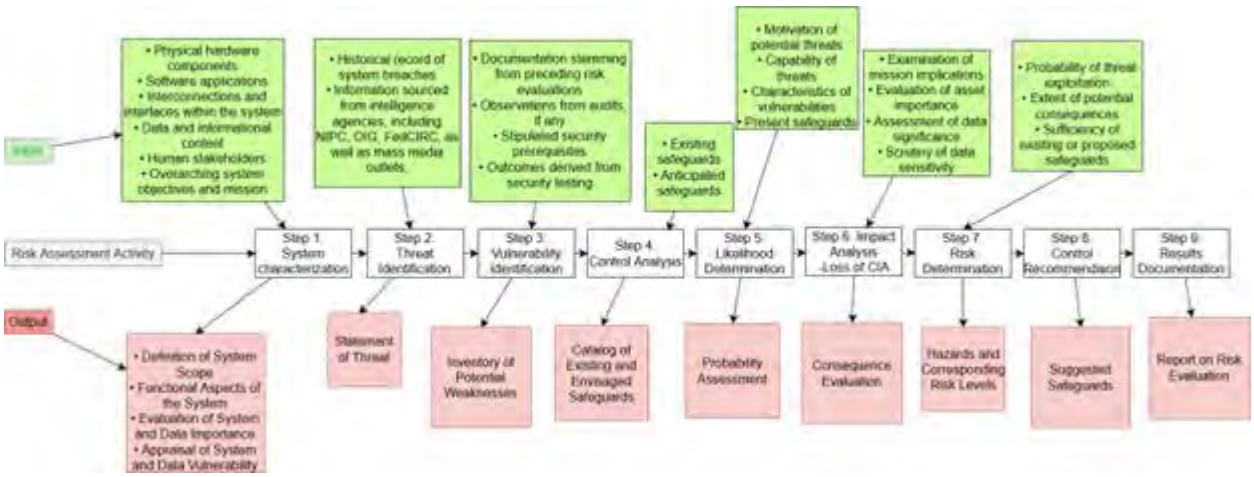


Figure 1. Flowchart Illustrating Risk Assessment Methodology

1.4 Foundation for Harmonization of Information Technology Security Standards

Many countries have made significant efforts to develop standards for information technology (IT) security. In the past ten years, the concepts and standards that govern the assessment of security goods have reached a level of maturity in the European Community (EC), United States (US), and Canada (Task, 1993). The widespread prevalence of products in the global market highlights the necessity for a standard that achieves broad acceptance and relevance among vendors in international contexts. Manufacturers face the challenge of the infeasibility of constructing and subjecting items to review systems in several countries, each with their own unique requirements. The main goal of this project is to develop a strategy that promotes the coordination of standards, thus facilitating harmonization (Task, 1993). The United States Department of Defense (DoD) introduced the initial safety assessment standard in the field of information technology, commonly referred to as "TC SEC" or the "Orange Book" (Latham, 1986), in 1983. The primary objective of the development of this standard was to assess the security of operating systems. This program represents a significant endeavor within the field of Information Technology Security. Other examples of evaluation criteria for information technology security are the Information Technology Security Evaluation Criteria (ITSEC) (Gehrke et al., 1992), the Canadian Trusted Computer Product Evaluation Criteria (CTCPEC) (Bacic, 1990), and the Federal Criteria for Information

Technology Security (Rannenberg, 1993). The complete Trusted Computer System Evaluation Criteria (CTCPEC) was developed in April 1992 as a merged framework that combines the Trusted Computer System Evaluation Criteria (TCSEC) and the European Information Technology Security Evaluation Criteria (ITSEC). Its purpose is to provide a complete structure for evaluating information technology (IT) products (Rannenberg, 1993). In addition, the International Organization for Standardization (ISO) introduced ISO/IEC 13335 in 1996, followed by the development of an information system management system centered on risk management in the year 2000 (Gehrke et al., 1992), (Bacic, 1990), (Rannenberg, 1993). The diagram below (Figure 02) depicts the comprehensive framework for the management of information security (Wawrzyniak, 2006).

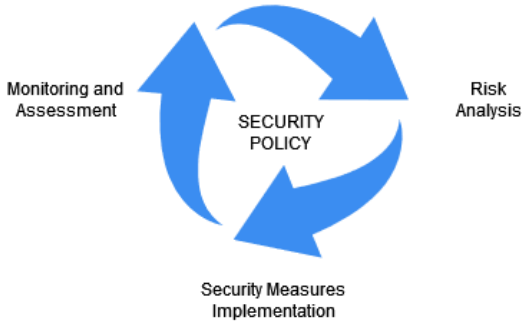


Figure 2. The overall structure for the process of managing information security (Wawrzyniak, 2006)

2. Common Methods of Risk Assessment

The main objective of doing risk assessment is to identify potential threats and consequences faced by information systems, and thereafter develop solutions to address the gap between the system's security level and organizational needs. Nevertheless, it is crucial to recognize that various assessment procedures can result in disparate evaluation results. Hence, the careful incorporation of comprehensive system-specific information is imperative when determining suitable approaches. In a general sense, risk assessment methodologies can be classified into three main categories: those that rely on quantitative analysis, those that are based on qualitative analysis, and those that integrate both quantitative and qualitative elements (Szczepankiewicz et al., 2006). To be more specific, Quantitative and qualitative methods are two primary categories of methodologies that are employed for the examination of risks to which assets within companies are exposed. Table I presents the primary pros and cons associated with IT risk assessment approaches. There exist various categories of IT risk analysis approaches, as outlined by the literature (Szczepankiewicz et al., 2006)

- Quantitative procedures involve the assessment of the amount of risk by employing numerical measurements. In this particular framework, the quantification of asset worth is based on measurable quantities, the measurement of threat frequency is expressed in terms of occurrence instances, and vulnerability is assessed by taking into account the probability of experiencing loss (Rot, 2008). The result of these approaches manifests in the form of measurable indicators. There are other quantitative approaches that can be identified, including Annual Loss Expected, Courtney's method, Fisher's method, and the ISRAM model (Rot, 2008).
- Qualitative methods, in contrast to quantitative approaches, do not depend on numerical data but instead provide outcomes in the form of descriptive narratives and corresponding recommendations. In the context of these methodologies, the process of risk assessment is intricately linked with:
    - The qualitative assessment of asset value, along with the development of qualitative scales indicating the occurrence frequency of threat incidents and the vulnerability linked to a specific threat, or alternatively: - The formulation of threat scenarios by predicting the essential elements that contribute to risk (Szczepankiewicz et al., 2006), (Rot, 2008).

Failure Mode and Effects Analysis/Failure Mode, Effects, and Criticality Analysis (FMEA/FMECA), the Risk Management Framework utilized by the Microsoft Corporate Security Group, the National Institute of Standards and Technology Special Publication 800-30 (NIST SP 800-30), and the Computerized Risk Analysis and Management Methodology (CRAMM) encompass various instances of quantitative methodologies (Rot, 2008).

The selection of risk metrics is contingent upon the magnitude of a particular hazard (Szyjewski, 2004). The metrics encompass a range of evaluations, starting from basic assessments that classify risk into high, medium, or low categories to more accurate measures that calculate the probability of a specific occurrence taking place.

When assessing the inherent risk associated with information security in an Information System, it is customary to do a qualitative analysis. This technique commonly relies on the fundamental principles of information security, specifically confidentiality, integrity, and availability. Each bestowed element can undergo a comprehensive examination of risk on an individual basis. In order to enhance the efficiency of the analysis process, a pre-established scale of information value (low, medium, high) has been implemented. The evaluation of risk magnitude can be demarcated by the utilization of a categorical scale containing designations such as very low, low, medium, high, and very high. A thorough assessment of risk and an understanding of its probability of occurrence are essential for gaining a comprehensive understanding of its impact on the overall operating efficiency of the Information System (Rot, 2008), (Szyjewski, 2004).

Table 1. The Advantages and Disadvantages of Quantitative and Qualitative Approaches in IT Risk Analysis (Rot, 2008)

Risk Analysis	Quantitative methods	Qualitative methods
Selected benefits	● The utilization of quantitative methods enables the determination of the repercussions resulting from incidents, hence facilitating the assessment of costs and benefits when selecting protective measures. ● They provide a more precise representation of the level of risk.	● This feature facilitates the prioritization of risks in a systematic manner. ● This approach enables the identification of high-risk locations quickly and cost-effectively. ● The process of analysis is characterized by its relative ease and affordability.
Selected drawbacks	● The effectiveness of quantitative measures is contingent upon the extent and precision with which the measuring scale is defined. ● The findings of the analysis may lack precision and potentially lead to confusion. ● It is necessary to enhance conventional methodologies by incorporating qualitative descriptions, such as comments and interpretations. ● The utilization of these methodologies in analysis typically incurs more costs, necessitates a higher level of expertise, and requires the use of specialized equipment.	● The method lacks the capability to ascertain probability and outcomes through numerical metrics. ● The process of evaluating costs and benefits becomes more challenging when making decisions on the selection of protective measures. ● The obtained outcomes possess a broad scope and are characterized by their general nature, as well as an approximate quality.

2.1 Methods of Quantitative Assessment

When employing quantitative approaches, analysts are faced with the challenge of accurately evaluating the essential quantities necessary for calculation. The quantification of risk can be expressed through various scales or directly within the financial domain as the projected magnitude of losses associated with a specific type of risk over a defined time period (Szczepankiewicz et al., 2006).

Quantitative risk analysis comprises a wide range of methods and procedures used for evaluation purposes, which include (Volkan Evrin, 2021), ("Risk Assessment and Analysis Methods," ISACA, n.d.):

- Heuristic methods refer to procedures employed to estimate contingency, which are based on either experience or expertise.
- The three-point estimate method involves the utilization of optimistic, likely, and pessimistic values in order to achieve an ideal estimation.
- Decision tree analysis is a visual tool used to illustrate the potential consequences of different choices or alternatives. It provides a diagrammatic picture of the implications associated with each alternative.
- The utilization of monetary evaluation to establish contingency reserves in a project or business process budget is commonly referred to as Expected Monetary Value (EMV).
- The Monte Carlo analysis method involves the utilization of estimated values for different outcomes in order to calculate business costs and project completion deadlines.
- Sensitivity analysis involves the identification of the most influential risks associated with a project.

Fault Tree Analysis (FTA) and Failure Modes and Effects Analysis (FMEA) are two methodologies utilized in engineering and risk management to systematically identify and analyze the various factors that contribute to system failure. These methodologies employ structured diagrams to visually represent the elements responsible for the occurrence of failures within a given system ("Risk Assessment and Analysis Methods," ISACA, n.d.).

Quantitative risk assessment encompasses essential variables, such as Single Loss Expectancy (SLE), which denotes the anticipated loss resulting from a singular occurrence. The Annual Rate of Occurrence (ARO) denotes the frequency at which the incident is anticipated to transpire during a one-year time frame. The integration of System Loss Expectancy (SLE) and Annual Rate of Occurrence (ARO) results in the derivation of the Annual Loss Expectancy (ALE), which serves as a basis for determining the economic feasibility of countermeasure investments. The Asset Loss Expectancy (ALE) can be calculated by multiplying the Single Loss Expectancy (SLE) by the Annual Rate of Occurrence (ARO). The aforementioned value functions as the resultant of risk assessment in quantitative analysis (Volkan Evrin, 2021).

The availability of trustworthy data is crucial for the accurate and effective execution of the IT risk assessment process, but it is uncommon for the team responsible for this activity to have access to such data. Furthermore, accurately measuring the extent of losses pertaining to individual and organizational assets might pose challenges. One of the primary focal points revolves around the potential vulnerability of sensitive data. The establishment of value requires a clear and precise demarcation of information, which is crucial for the effective implementation of various business activities. Moreover, it is of utmost importance to understand the value of information in maintaining the operational integrity of different segments within an organization, which in turn has implications for the overall enterprise. The underlying correlation employed in IT risk assessment is delineated as follows (Szczepankiewicz et al., 2006), (Galach, 2004).

$$\mathbf{R = P \times W \text{ and } P = F \times V}$$

In this context, the variables are denoted as follows: R represents the risk value, P represents the probability or expected number of incidents that may occur, resulting in the loss of assets value within a specified time period. W refers to the value of loss, which represents the anticipated decrease in the value of assets resulting from a single incident. F, on the other hand, denotes the frequency at which threats are expected to occur. The susceptibility of an information system, or its individual elements, to a danger, refers to the likelihood that a certain vulnerability would be exploited by that threat. The justification for this rests in the understanding that the evaluation of IT risk is typically expressed through the anticipated value of a loss, a metric derived from the description of three essential variables (Szczepankiewicz et al., 2006). The quantification of resources, which includes crucial factors such as knowledge necessary for efficient organizational operation, is attributed with value. The quantification of possible threats targeting resources, such as processed information, is determined by calculating the frequency of their occurrences within a specified timeframe. The usual practice is to establish a one-year duration for the purpose of determining frequency. The quantification of the susceptibility of an information technology (IT) system or its individual components to threats is achieved by an evaluation of the probability of experiencing losses as a result of prospective occurrences (Rot, 2008). The ALE model (Annual Loss Expected) is widely recognized as the prevailing and commonly utilized quantitative approach for risk assessment. The underlying idea of this model is based on the concept of expected loss, which is determined by multiplying the probability of bad occurrences impacting information technology (IT) systems with the monetary losses associated with these events. The subsequent models offer comprehensive information regarding this method (Rot, 2008), (Wawrzyniak, 2007), (Szczepankiewicz et al., 2006).

ALE = (Probability of event) x (value of loss)

$$ALE = \sum_{i=1}^{n} I(O_i)F_i$$

Here, let $\{O_1, O_2, ..., O_n\}$ represent the collection of negative effects coming from an event. The value of $I(O_i)$ denotes the loss indicated by the event. The variable F_i represents the frequency of the event i.

The calculation of annual predicted losses in companies is based on the summation of anticipated yearly losses, which serves as a fundamental basis. A variety of alternative approaches for the assessment and evaluation of IT risk arise from the aforementioned methodology. These adjustments are tailored to address specific needs and conditions that are inherent to a given organizational context. One strategy that deserves consideration is Courtney's method, which was formulated by Robert Courtney. Similar to the ALE (Annualized Loss Expectancy) methodology, this method assesses potential losses by calculating the product of the value of losses associated with a hazard and an indicator that measures the likelihood of its happening. Courtney asserts that the crux of risk assessment relies on the utilization of the following formula (Rot, 2008), (Ryba et al., 2009).

$$\mathbf{R = P \times C}$$

R represents the risk value. The variable P represents the probability of the occurrence of a specific number of times throughout a year, pertaining to an event that results in a loss for the organization. The term "C-loss" refers to the financial loss experienced by an organization as a direct consequence of a singular incident that leads to such loss.

$$\boldsymbol{ALE} = \frac{\mathbf{10}^{\boldsymbol{f+i-3}}}{\mathbf{3}}$$

The variable "f" represents an index that quantifies the assessed frequency of the event that leads to a loss. The index is used to determine the extent of loss resulting from the occurrence of an event (Rot, 2008).

Courtney's method presents a comprehensive framework consisting of six primary vulnerability types. These categories comprise accidental data disclosure, unintended data modification, inadvertent data deletion, intentional data disclosure, deliberate data change, and deliberate data removal. The approach in question has been officially acknowledged by national bodies in the United States as a recognized method for performing risk assessments (Ryba et al., 2009).

Fisher's method, which builds upon Courtney's approach, functions as a complete framework for the development of security solutions in the field of Information Systems. The presence of a well-established information security policy within the company is essential for the efficient application of this concept. This methodology comprises various stages within the process of managing risk in Information Systems (Ryba et al., 2009), (Rot, 2008).

- Phase 1 of the research process entails the systematic collection of data. This phase involves identifying and categorizing the various resources inside Information Systems. Once these resources have been identified, relevant information about them is gathered. This information will then be subjected to additional analysis in later phases (Rot, 2008).
- Phase 2 of the study involves the process of identifying threats. This entails the mapping of the six threat categories indicated in Courtney's technique onto eleven Fisher control points. These points comprise several acts including the gathering, transmission, transformation of form, transit, reception, processing, migration, removal, and utilization of data (Rot, 2008).
- Phase 3 of the project involves conducting a risk assessment. This assessment aims to evaluate the level of risk using Courtney's method formula, which is represented by the equation R = P × C. In this equation, P represents the probability of an event causing loss occurring a specific number of times per year, while C represents the loss incurred by the organization as a result of a single occurrence of the event causing loss (Rot, 2008).
- Phase 4 of the project entails the design of control mechanisms, wherein the task at hand involves the careful selection of appropriate control mechanisms for each risk that has been identified. The strategies under consideration encompass preventive, detective, and remedial aspects (Rot, 2008).
- Phase 5 - Economic Viability The evaluation of mechanisms entails doing a comprehensive business assessment of the identified mechanisms. This assessment involves the utilization of the Return on Investment (ROI) indicator, which was previously described, and is expressed through a specific formula (Rot, 2008).

$$\boldsymbol{ROI = Operational\ profit\ in\ a\ given\ period \div value\ of\ invested\ capital}$$

This approach considers the operational gain as the representation of the degree of risk for certain control mechanisms, while the measure of invested capital is viewed as the evaluated cost of these mechanisms (Ryba et al., 2009). The ISRAM model (Information Security Risk Analysis Method) is presented as the subsequent approach in the paper. The foundation of this model is based on the ALE (Annual Loss Expected) approach, which predominantly utilizes survey research as its primary methodology. The assessment of risk in information technology is achieved by employing the following formula (Wawrzyniak, 2007), (Karabacak et al., 2005).

$$RISK = \left(\frac{\sum_m T_1\left(\sum_i w_i p_i\right)}{m}\right)\left(\frac{\sum_m T_2\left(\sum_j w_j p_j\right)}{n}\right)$$

Here, the variable "i" denotes the number of survey questions pertaining to the assessment of the probability of

incident occurrences. The variable "j" represents the quantity of survey questions that pertain to assessing the effects. The variables "m" and "n" represent the number of participants involved in the survey. The weights assigned to questions "i" and "j" are denoted as wi and wj, respectively. The variables pi and pj are used to denote the values that correspond to the selected responses for "i" and "j" respectively. T1 represents the tabulated probability associated with the occurrence of specific events. The term "T2" refers to a table that encompasses the unfavorable consequences that result from the occurrence of events. The Courtney's approach, which is commonly employed in risk assessment, has been reconstructed through the utilization of the Exposure Analysis Matrix. The foundation of this approach is the underlying notion that the magnitude of risks is contingent upon the number of individuals capable of causing harm, hence necessitating a risk analysis that involves the categorization of specific occupational cohorts inside the organization. Parker's approach incorporates Courtney's method but expands upon it by incorporating qualitative analysis of risk. Additionally, Parker's method formalizes the assessment of the impact of human factors on risk, setting it apart from other methods (Szyjewski, 2004).

2.2 Methods of Qualitative Assessment

The main purpose of qualitative risk analysis is to identify risks that require thorough consideration, and to develop appropriate controls and actions based on the implications and impact of the risks on objectives (Behnia et al., 2012). In the realm of qualitative risk analysis, there exist two well acknowledged and easily implementable approaches for risk appraisal (Kuzminykh et al., 2021). The Keep It Super Simple (KISS) approach is well-suited for projects that have a narrow focus or restricted magnitude. Its purpose is to minimize needless complexities and facilitate uncomplicated review by teams with limited experience in risk assessment. This particular methodology involves assessing risks using a fundamental scale, commonly classified as very high, high, medium, low, and very low ("Risk Assessment and Analysis Methods," ISACA, n.d.), (Kuzminykh et al., 2021). The utilization of this approach is more appropriate for complex and substantial challenges, as well as for teams who possess expertise in risk assessment. The methodology employed in this study involves the assessment of both the probability of risk occurrence and the subsequent consequences or impacts. Probability is a measure that measures the likelihood of a risk being realized, whereas impact pertains to the potential consequences that can affect several aspects, including time, cost, scope, and quality. The assessment of probability and impact is conducted using a predetermined scale, such as a range of 1 to 10 or 1 to 5. The risk score is then derived by multiplying these two evaluations (Behnia et al., 2012), ("Risk Assessment and Analysis Methods," ISACA, n.d.). Qualitative risk analysis is a widely applicable approach that may be utilized to effectively identify risk areas associated with regular business operations across different industries. This methodology evaluates the extent to which employees' concerns over their job are consistent with the identified risk domains. The utilization of the quantitative technique is employed in tandem to investigate relevant risk scenarios, hence providing detailed insights for making well-informed decisions ("Risk Assessment and Analysis Methods," ISACA, n.d.). Quantitative risk analysis provides enhanced objectivity and exact facts in the context of crucial decision-making or complex activities, as opposed to its qualitative equivalent. However, it is crucial to recognize that quantitative analysis is still susceptible to estimation or inference, leading prudent risk managers to take into account supplementary considerations during the decision-making procedure (Kuzminykh et al., 2021). Although qualitative risk analysis is commonly favored for its simplicity in execution, there may be situations that need the utilization of a quantitative approach. Following the completion of qualitative research, the subsequent step involves doing quantitative analysis. However, in cases where qualitative analysis provides sufficient insights, it may not be necessary to perform a quantitative analysis for every individual risk ("Risk Assessment and Analysis Methods," ISACA, n.d.), (Kuzminykh et al., 2021), (Peixoto et al., 2022).

There are several qualitative procedures that can be utilized for risk analysis. This discourse will focus on three specific methods: FMEA/FMECA methodologies, NIST 800-30, and CRAMM methodologies. The origins of FMEA (Failure Mode and Effects Analysis) and FMECA (Failure Mode and Effects Criticality Analysis) may be traced back to the 1950s, during which they were developed for the purpose of evaluating the reliability of weaponry. These methods are still utilized, particularly in sectors such as aviation, space exploration, and electronics (Bialas, 2006). The FMEA/FMECA process essentially entails a thorough examination of the potential consequences of each fault on the overall functionality of the system, as well as the classification of probable flaws based on their severity levels. The FMECA methodology enhances the analysis by evaluating the severity of defects and their possible impact on the system's functionality. Although these methods have proven to be effective, they require a significant amount of manual effort and skill from practitioners. Additionally, specialized tools that integrate parts of knowledge engineering and fuzzy logic are necessary for their implementation (Bialas, 2006). The NIST SP 800-30 methodology outlines a comprehensive framework for conducting IT risk assessment, which encompasses nine fundamental phases (Stoneburner et al., 2002), (Ryba et al., 2009).

a) **Selection of Evaluated Systems:** The task at hand involves the identification of the systems that are subject to assessment.

b) **Scope Definition and Information Collection:** The process of delineating the parameters of evaluation and collecting the requisite data.

c) **Threat Identification:** Identifying potential risks associated with the assessed systems.

d) **Susceptibility Identification:** The task at hand involves the identification of vulnerabilities inside the systems that have been analyzed.

e) **Analysis of Control and Protection Mechanisms:** This study aims to assess the existing or proposed systems of control and protection.

f) **Determination of Susceptibility Probabilities:** The probabilities of susceptibility incidence can be specified by identifying threat sources and assigning likelihood levels categorized as low, medium, or high.

g) **Analysis of Incident Impact:** This study aims to conduct an analysis and assessment of the effects of incidents on the system, data, and organization. These incidents will be grouped into three levels: high, medium, and low.

h) **Risk Level Determination:** The Risk Level Matrix is utilized to determine risk levels by multiplying the probabilities of incident occurrence, which are assigned weights of 1.0 for high, 0.5 for medium, and 0.1 for low, with the corresponding impact strengths of incidents, which are assigned weights of 100 for high, 50 for medium, and 10 for low. The matrix presented below serves to determine the comprehensive level of risk associated with each identified danger, which is classified into three categories: high (with a product range of (50,100]), medium (with a product range of (10,50]), or low (with a product range of [1,10]).

i) **Elaboration of Control and Protection Recommendations:** The objective of this study is to propose control methods and alternative alternatives that can effectively mitigate risks to an acceptable level.

The CRAMM technique, known as the CCTA's Risk Analysis and Management technique, has been officially recognized by the CCTA (UK Government Central Computer and Telecommunications Agency) as a standard for risk analysis and management. This methodology is structured around a three-stage procedure (Ryba et al., 2009).

a) **Resource Identification and Evaluation:** The process of identifying and evaluating resources.

b) **Threat and Susceptibility Evaluation:** Evaluating potential risks and weaknesses.

c) **Control and Protection Mechanism Selection and Recommendation:** The process of selecting and recommending control and protection systems.

The objective of IT risk analysis is to assess the probability of occurrences that may disturb the optimal operation of resources. The process entails the classification of identified resources into distinct asset groups, followed by the creation of inventories of significant threats associated with each asset group. This ultimately leads to the assessment of the risk level for each group, utilizing a five-level scale.

The aforementioned approach incorporates specialized software as an essential component that facilitates the aforementioned processes (Rot, 2008).

2.3 Combination of Quantitative and Qualitative Evaluation

The key advantages of quantitative risk assessment stem from its utilization of empirical and measurable data. This approach enables the provision of accurate outcomes pertaining to risk appraisal and the determination of the optimal investment necessary for effective risk treatment, hence ensuring the profitability of the company ("Risk Assessment and Analysis Methods," ISACA, n.d.). One instance of a quantitative methodology employed for doing cost-benefit analysis is the Annual Loss Expected (ALE) calculation. This approach assists companies in assessing the projected financial loss linked to a particular asset or investment, as a result of associated risks, within a span of one year ("Risk Assessment and Analysis Methods," ISACA, n.d.).

The determination of Annualized Loss Expectancy (ALE) for an investment in a virtualized system encompasses the subsequent procedures ("Risk Assessment and Analysis Methods," ISACA, n.d.):

The monetary worth of the virtualization system's hardware is estimated to be $1 million, as determined by the Single Loss Expectancy (SLE) for hardware ("Risk Assessment and Analysis Methods," ISACA, n.d.).

The estimated monetary worth of virtualized system management software is $250,000, as determined by the

Single Loss Expectancy (SLE) for software ("Risk Assessment and Analysis Methods," ISACA, n.d.).

According to the vendor data, there is an occurrence of system catastrophic failure, caused by either software or hardware issues, once per 10 years, resulting in an Annual Rate of Occurrence (ARO) of 0.1 ("Risk Assessment and Analysis Methods," ISACA, n.d.).

The annualized loss expectancy (ALE) for the homework assignment is calculated by multiplying the asset value of $1 million by the annualized rate of occurrence (ARO) of 0.1, resulting in an ALE of $100,000 ("Risk Assessment and Analysis Methods," ISACA, n.d.).

The annual license expense (ALE) for software (SW) can be calculated by multiplying the initial cost of $250,000 by the annual maintenance fee rate of 0.1, resulting in an ALE of $25,000 ("Risk Assessment and Analysis Methods," ISACA, n.d.).

In the given context, the organization is confronted with a recurring risk of incurring a financial loss up to $100,000 per annum due to hardware failure, and a loss of $25,000 per annum due to software failure in the event of virtualization system malfunction. Any control mechanism that is applied, such as backup systems, disaster recovery plans, or fault tolerance systems, which incurs costs lower than the specified amounts, would result in a profitable outcome ("Risk Assessment and Analysis Methods," ISACA, n.d.).

Some risk evaluations require the consideration of complex parameters. Further illustrations can be obtained by employing the subsequent "sequential deconstruction of quantitative risk analysis" ("Risk Assessment and Analysis Methods," ISACA, n.d.).

- Perform a comprehensive evaluation of potential hazards and weaknesses to ascertain relevant factors of risk. Determine the Exposure Factor (EF), which is the percentage of asset loss attributed to the identified threat. The Single Loss Expectancy (SLE) can be calculated by taking into account the risk variables and the values of the assets that are at risk. The SLE is obtained by multiplying the value of the asset by the Exposure Factor (EF). When considering adjustments, it is important to consider historical records of incidents and the prevalent institutional culture as influential elements.
- Estimate the Annual Rate of Occurrence (ARO) associated with each risk determinant. Determine the necessary steps to mitigate each risk factor. Utilize a numerical scale spanning from 1 to 10 in order to quantitatively assess the level of severity, where a rating of 10 represents the highest degree of severity. This scale functions as a corrective factor for evaluating risk, taking into account the company's specific risk profile.
- Determine the Annualized Loss Expectancy (ALE) associated with each risk factor. It is imperative to acknowledge that, despite the implementation of countermeasures, the Annual Rate of Occurrence (ARO) that contributes to the Annual Loss Expectancy (ALE) may not be completely eliminated in every case. The computation of the adjusted Annual Loss Expectancy (ALE) involves multiplying the ALE, as indicated in the table, by the adjustment factor and then by the size correction factor.
- Perform a thorough assessment of the expenses and advantages by conducting a comparative analysis of the Annual Loss Expectancy (ALE) before and after the installation of countermeasures. The Internal Rate of Return (IRR) can be employed as a foundational metric for computing the Return on Investment (ROI) within the context of a cost/benefit analysis methodology.
- In conclusion, it is imperative to concisely summarize the results in order to facilitate the assessment of performance by management.

The utilization of combined approaches has the potential to optimize process efficiency and facilitate the achievement of required security levels. When it comes to the risk assessment process, the decision between quantitative and qualitative methodologies may often be made with reasonable ease. The execution of qualitative risk assessment is expeditious since it does not heavily rely on mathematical calculations or measurements, hence facilitating its straightforward implementation. Organizations derive advantages from the presence of seasoned personnel who possess knowledge of assets and processes. However, it is crucial to acknowledge the inherent biases that may arise when evaluating the likelihood and consequences of certain events (Rot, 2008). In general, a balanced integration of qualitative and quantitative research methods, along by thorough preparation of assessment procedures and suitable modeling techniques, may be the most advantageous approach for a successful risk assessment procedure (Rot, 2008), ("Risk Assessment and Analysis Methods," ISACA, n.d.).

3. Analysis and Design of Risk Assessment

The Information Security Management System (ISMS) is a comprehensive framework consisting of a set of

well-defined policies designed to effectively manage information security and address IT-related risks (Cranor et al., 2005). The aforementioned methodology has increasingly garnered attention as a highly efficient strategy for tackling intricate issues within the field of information security, thus garnering broader recognition and approval on an international level. The effectiveness and functioning of this system are inherently dependent on the careful implementation of the risk assessment phase, which is a crucial step in its formation and smooth operation. It is crucial to highlight that the domain of ISMS functions within a context marked by dynamic aspects, and these constituents inject a certain degree of unpredictability into the equation (Abbas et al., 2011).

One notable element concerns the dynamic nature of security requirements inside an organization. The aforementioned requirements are subject to change as a result of the inherent obsolescence of current security methods and the need to address developing vulnerabilities. The rapid advancements in technology have given rise to a set of vulnerabilities that businesses must actively confront, marking the onset of a new era of problems (Abbas et al., 2011).

The second dynamic component underscores the potential for the implementation of an Information Security Management System (ISMS) to result in unanticipated external effects that impact other interconnected systems. The prediction of these externalities is inherently unpredictable and difficult to ascertain in advance. The complex interaction among different systems might give rise to cascading consequences that may only manifest themselves once the Information Security Management System (ISMS) has been implemented and is actively engaging with its surrounding environment (Abbas et al., 2011).

The evaluative procedures for security issues included within Information Security Management Systems (ISMS) are important to the concept of the third dimension. The aforementioned mechanisms are intrinsically interconnected with the technical environment of their respective era. As technological advancements progress, novel risks and susceptibilities arise, hence rendering the assessment standards employed in Information Security Management Systems (ISMS) outdated. This highlights the importance of employing a flexible strategy that can adjust to the changing landscape of threats, guaranteeing the resilience and efficacy of the Information Security Management System (ISMS) in the presence of continuously increasing difficulties (Abbas et al., 2011).

In light of the aforementioned context, the incorporation of a perpetual security assessment mechanism within the Information Security Management System (ISMS) emerges as not merely a choice, but a vital necessity. The mechanism, deeply embedded in the organizational structure, functions as a dynamic feedback loop that consistently assesses the effectiveness and pertinence of the Information Security Management System (ISMS). The implementation of this measure guarantees the maintenance of a resilient and flexible system, which is capable of effectively addressing emerging threats and vulnerabilities. As a result, it significantly contributes to the attainment of the overall objectives related to information security (Cranor et al., 2005). In an environment marked by swift technological progress and ever-changing risk elements, the implementation of a proactive and flexible security evaluation framework is crucial for maintaining the integrity and efficacy of an Information Security Management System (ISMS) (Cranor et al., 2005), (Abbas et al., 2011).

3.1 Information Security Standards

Within the field of information security, a multitude of standards have been developed. The integration of security standards within corporations and enterprises not only boosts the efficacy of security measures but also optimizes their development process. In order to achieve a collective agreement regarding the extent of safeguarding offered by these standards, it is imperative to incorporate security measures in a systematic fashion. The following is a comprehensive review of the primary standards (Asosheh et al., 2013):

The ISO/IEC 27000-series comprises a collection of information security standards that are collaboratively issued by the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC). At now, this series encompasses a total of eleven well-established standards, while more standards are still in the developmental process. The primary criteria in this series are:

- ISO 13335, also known as ISO/IEC 13335, is a standard named "Information technology — Security techniques — Management of information and communications technology security." This standard is divided into two portions. The ISO 27005 standard has superseded specific portions of the old ISO 13335 standard.
- ISO 17799, often referred to as BS7799, offers a complete examination of security considerations. The subject matter involves a set of control criteria that are especially complex.

- ISO 27001 is a standard that delineates the necessary elements for establishing, executing, overseeing, evaluating, sustaining, and improving an Information Security Management System. The standard is based on a methodology that emphasizes the importance of process-oriented practices.
- ISO 27002 provides a set of control objectives and best practice procedures that can be selected and implemented to mitigate risks and achieve information security objectives. This standard includes 11 clauses that consist of a total of 39 primary security categories and 133 individual controls.
- ISO 27003 provides practical recommendations for the deployment of Information Security Management Systems (ISMS) in accordance with ISO/IEC 27001.
- ISO 27004 offers guidance on the development and implementation of metrics for assessing the efficiency of Information Security Management Systems (ISMS), control goals, and controls specified in ISO/IEC 27001.
- ISO 27005 provides guidelines for the effective management of risks associated with information security.
- ISO 27006 provides a comprehensive framework that outlines the necessary criteria and offers guidance for entities involved in the auditing process.
- ISO 27007 provides guidance on the implementation of audits for Information Security Management Systems (ISMS) and outlines the necessary skills and qualifications for ISMS auditors.
- ISO 27008 provides recommendations pertaining to the evaluation and execution of controls in order to assess their implementation and operational effectiveness.
- There are further standards that are part of the ISO27k series. The ISO27k series encompasses standards 27000–27019 and 27030–27044, which jointly provide a range of security management characteristics that are in line with the requirements of ISO 27001.

The BS7799 Security Standard was initially developed by the British Standards Institution (BSI). The 1995 version, together with its succeeding iterations in 1999 and 2000, are widely acknowledged and referred to as BS Version 1, BS Version 2, and BS Version 3, respectively (Broderick, 2006).

The Payment Card Industry Data Security Standard (PCIDSS) is an internationally recognized information security standard developed by the Payment Card Industry Security Standards Council. Its purpose is to assist enterprises involved in card payment processing in their efforts to prevent credit card fraud. This applies to all organizations that are responsible for managing cardholder information associated with card brands that display the corresponding logos (Susanto12 et al., 2011).

The Information Technology Infrastructure Library (ITIL) emerged in the 1980s as a direct response to the insufficient quality of IT service. ITIL comprises a comprehensive framework of principles and practices that pertain to the management of Information Technology Services (ITSM), including aspects connected to security. The primary objective of ITIL is the management of IT services, with an emphasis on the viewpoint of service providers (Tofan, 2011).

COBIT, which stands for Control Objectives for Information and Related Technology, presents a structured approach for addressing and managing the risks associated with information technology inside business operations. COBIT, developed by the IT Governance Institute (ITGI) under the auspices of the Information Systems Audit and Control Association (ISACA), serves as a framework for governing and managing IT to ensure its alignment with business objectives (Broderick, 2006).

Finally, the origins of the Government Access to Secure Systems Program (GASSP) can be traced back to the year 1992, with its establishment being supported by the support of the United States government, the Information Security Institute, and several other institutions. Following a series of developmental stages, the framework was designated as the 'Generally Accepted System Security Principles' (GAISP) and later transformed into the 'Generally Accepted Information Security Principles' (GAISP) to align with the expanding range of its objectives. The establishment of GASSP received support from a wide range of stakeholders, and the resulting document encompasses a comprehensive structure. The GAISP iteration is a crucial reference material that provides developers with direction based on concepts supported by respected organizations such as the OECD and ISF. It can be likened to a comprehensive cookbook, with detailed instructions and recommendations (Siponen & Willison, 2009), (Asosheh et al., 2013).

4. The Significance of Artificial Intelligence in Information Security (INFOSEC)

The preservation of information and system security is a significant challenge in the digital communication landscape. The inherent significance of information highlights the crucial need to guarantee data security in communication systems. The utilisation of inherent properties of artificial neural networks, such as their adaptive learning capabilities, can facilitate the possibility for security advancements (Karapilafis, 2015). Moreover, artificial intelligence (AI) functions across multiple modalities, including learning, comprehension, and decision-making. The evolution of AI is apparent in the development of autonomous intelligence, as demonstrated by the breakthroughs made in the field of self-driving vehicles. Prominent business companies such as Google, IBM, Juniper Networks, and Balbix have utilised artificial intelligence (AI) in the field of information security. Google utilises artificial intelligence (AI) to enhance the performance of Gmail, while IBM applies AI to identify potential attacks. Additionally, Balbix's Breach Control Platform harnesses AI to proactively anticipate and mitigate potential hazards. By efficiently integrating artificial intelligence (AI) with security measures, a more robust cybersecurity posture is established, leading to the successful mitigation of many threats, including ransomware. The combination of AI capabilities with security requirements is a significant shift, marking the beginning of a new age in digital defence ("Artificial Intelligence and its Application," Onyango).

4.1 Methodological Approaches for Integrating Artificial Intelligence into INFOSEC

Artificial Intelligence (AI) is a distinct field of study within the discipline of computer science that focuses on the development of computer systems possessing sophisticated cognitive ability to perform a wide range of activities. The area of information security management can be categorised into three distinct classifications: Artificial Narrow Intelligence (ANI), Artificial General Intelligence (AGI), and Artificial Super Intelligence (ASI) ("Artificial Intelligence and its Application," Onyango).

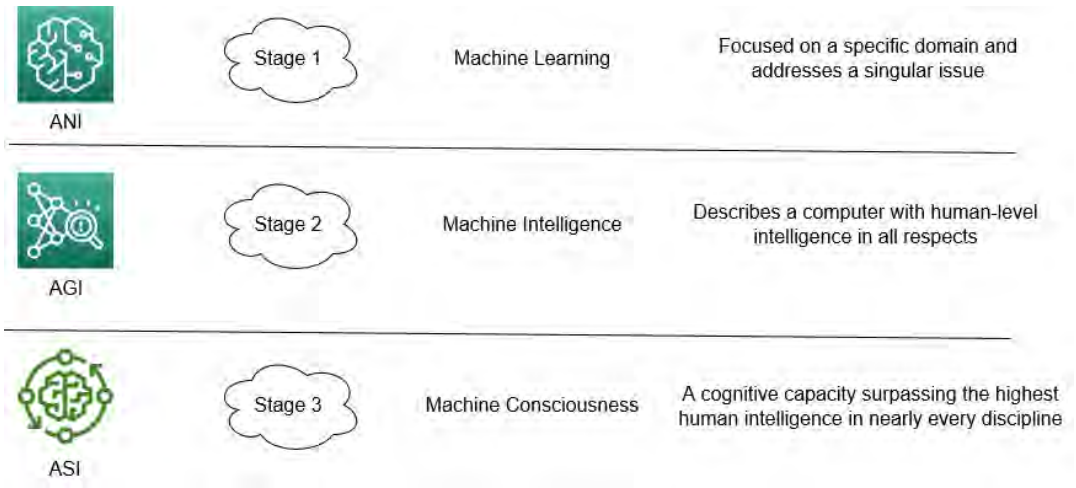


Figure 3. Artificial Intelligence across its three classifications (Granata, 2023)

Artificial Narrow Intelligence (ANI) demonstrates a high level of competence in narrow activities within the field of artificial intelligence ("Artificial Intelligence and its Application," Onyango). It is particularly good in managing information, as seen in its application in tasks such as retrieving smartphone data through virtual assistants. In contrast, Artificial General Intelligence (AGI) encompasses a wide range of reasoning capabilities and flexibility, similar to the cognitive capacities observed in humans ("Artificial Intelligence and its Application," Onyango), (Tolani & Tolani, 2019). This is exemplified by healthcare-focused robots such as pillo. In the realm of artificial intelligence, Artificial Superintelligence (ASI) represents the pinnacle of AI effectiveness, surpassing human cognitive abilities by actively engaging with intricate concepts, accurately predicting errors, and devising solutions. This is exemplified by the multifaceted humanoid robot "Alpha 2," which contributes to enhancing information security and fortifying system resilience ("Artificial Intelligence and its Application," Onyango), (Tolani & Tolani, 2019), (Sundu & Ozdemir, 2020).

4.2 Security Risks in the Realm of Artificial Intelligence (AI)

Cyberattacks can be classified into discrete areas, including integrity, confidentiality, authenticity, and non-repudiation considerations. Based on the aforementioned concerns, it is apparent that AI security risks might arise in three basic aspects. Such as, ("Artificial Intelligence and its Application," Onyango).

- Within the realm of cybersecurity, espionage refers to the act of obtaining intelligence related to a specific target's information system. This serves as a preliminary step before executing complex cyberattacks. An entity with antagonistic intentions possesses the capability to utilize approaches driven by artificial intelligence in order to thoroughly examine an information management system, consequently deriving detailed insights by leveraging inherent properties such as datasheets ("Artificial Intelligence and its Application," Onyango).
- Sabotage refers to the intentional obstruction of an artificial intelligence (AI) system's ability to function effectively. This can be achieved by manipulating the system's models or overwhelming it with demands that beyond its processing capacity ("Artificial Intelligence and its Application," Onyango).
- Similarly, fraud entails the purposeful manipulation of roles through the misclassification and contamination of data, including the deliberate insertion of fabricated data or orchestrated interactions during system training in order to exert influence on decision-making processes ("Artificial Intelligence and its Application," Onyango).

The harmful application of artificial intelligence (AI) poses a complex set of difficulties to the field of information security, which can be categorized into three main areas: digital threats, physical threats, and political threats ("Artificial Intelligence and its Application," Onyango).

4.2.1 Digital Security

Threats can manifest through the manipulation of individuals within a social context, a phenomenon commonly referred to as social engineering. These threats can be categorized and visualized in the accompanying diagram ("Artificial Intelligence and its Application," Onyango).

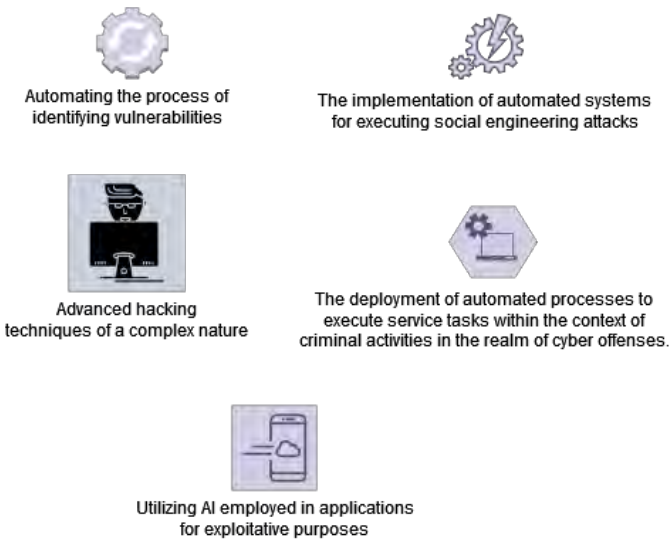


Figure 4. Digital Security ("Artificial Intelligence and its Application," Onyango)

Automated social engineering attacks encompass the utilisation of Natural Language Processing (NLP) to emulate the writing style of a certain target, hence enabling AI systems to collect online data pertaining to the subject and afterwards generate malevolent content, such as detrimental links, emails, and webpages ("Artificial Intelligence and its Application," Onyango). The process of identifying system vulnerabilities using artificial intelligence (AI) is dependent on the analysis of past trends in order to detect potential holes that can be exploited by malicious actors in a covert manner ("Artificial Intelligence and its Application," Onyango). Artificial intelligence (AI) has the potential to enhance the efficiency of hacking activities by employing a prioritisation system that focuses on selecting target victims according to their specific vulnerabilities. Moreover, artificial intelligence has the capability to automate operations that cause disruptions in the flow of data, namely in the realm of cybercrimes, such as payment processing. In the field of information security, the utilisation of data poisoning is employed as a means to establish backdoors or breach the security protocols that are integral to artificial intelligence (AI) systems ("Artificial Intelligence and its Application," Onyango).

4.2.2 Physical Security

Security threats can manifest via compromising the physical integrity of a machine, for as through the utilisation of

weaponized hard drives ("Artificial Intelligence and its Application," Onyango).

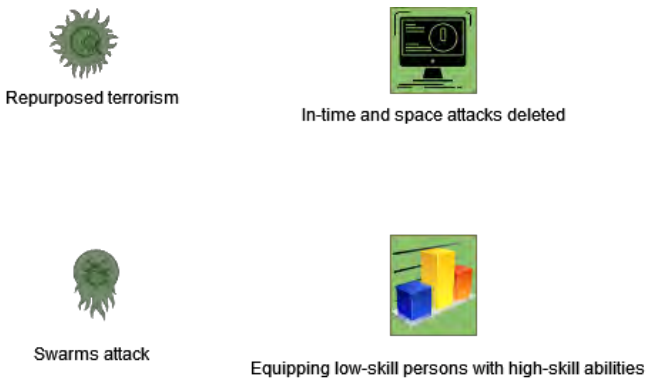


Figure 5. Physical Security ("Artificial Intelligence and its Application," Onyango)

Terrorist repurposing refers to the unauthorised utilisation of commercially available artificial intelligence (AI) devices, such as drones, with the intention of undermining the security of individuals or organisations, hence constituting a possible threat to data storage facilities. The integration of wireless and remote communication functionalities in artificial intelligence (AI) systems facilitates the persistent and automatic targeting of data centres. Furthermore, the use of distributed networks in the field of artificial intelligence (AI) results in the development of autonomous robotic systems that possess the ability to carry out highly coordinated attacks on a significant magnitude. Moreover, artificial intelligence (AI) grants malicious actors with advanced proficiencies, presenting different avenues for carrying out assaults by optimising algorithms, navigating systems, and detecting weaknesses ("Artificial Intelligence and its Application," Onyango).

4.2.3 Political Security

Threats possess a significant impact on society by manifesting itself in various forms, such as profiling, surveillance, and the utilisation of automated disinformation operations ("Artificial Intelligence and its Application," Onyango). AI generates photos and movies with provocative material that is hard to check, enabling fake news.

Figure 6. Political Security [33]

Furthermore, AI allows governments to collect and exploit data from individuals and organisations without consent. Third, social network AI recognises harmful influencers that spread personalised falsehoods. Finally, AI's sophisticated algorithms alter data to affect user behaviour, as shown in bot-driven denial-of-information attacks that flood information channels and render data inaccurate ("Artificial Intelligence and its Application," Onyango).

4.3 Benefits and Areas of AI Appliactions in Infosec Management

Artificial Intelligence (AI) assumes a diverse position in augmenting cybersecurity practises. Firstly, it facilitates the development of complete inventories of IT assets, providing accurate insights on the utilisation of users, devices, and applications inside information systems, hence streamlining inventory management (Fozilovich, 2022). Additionally, artificial intelligence (AI)-powered solutions provide timely and up-to-date data about worldwide and organization-specific risks, allowing organisations to efficiently determine the order of importance for implementing security measures. Furthermore, artificial intelligence (AI) plays a crucial role in the assessment of security tools and procedures, contributing to the maintenance of a robust security posture via the identification

of vulnerabilities. In addition, artificial intelligence (AI) has the capability to forecast prospective breaches, therefore assisting organisations in the proactive deployment of resources to avoid or mitigate vulnerabilities. In addition, artificial intelligence (AI) technologies play a crucial role in expediting and providing comprehensive responses to security events by effectively detecting underlying causes and implementing measures to avoid future occurrences. Finally, artificial intelligence (AI) plays a crucial role in improving the comprehensibility of cybersecurity systems, facilitating the effective communication of studies and suggestions to diverse parties engaged in the oversight and protection of data. Artificial intelligence (AI) plays a crucial role in enhancing the cyber resilience and security protocols used by organisations ("Artificial Intelligence and its Application," Onyango).

4.4 Summary of AI's Role in Information Security

The efficacy of conventional security procedures in protecting information systems from cyber assaults has shown a growing susceptibility. The prevalence of hackers successfully circumventing security firewalls has grown more frequent, hence highlighting the need for a more sophisticated and strategic approach to the field of cybersecurity (Tolani & Tolani, 2019). Artificial Intelligence (AI) presents a diverse array of efficacious security solutions capable of detecting atypical behaviours inside data management systems. By using machine learning techniques, AI can discern data abnormalities and forecast possible security threats (Everitt et al., 2017). Artificial intelligence (AI) solutions provide the capability to effectively identify and isolate tainted data, hence aiding organisations in the eradication of system vulnerabilities and the prevention of malware assaults. This technology functions across three distinct tiers in the realm of information security management, including prevention and mitigation, detection, and reaction. Artificial intelligence (AI) plays a crucial role in bolstering information security via its ability to better decision-making processes, monitor system activity, and automate jobs, hence enabling fast responses to potential attacks. The flexibility of artificial intelligence (AI) in the domains of data management and security makes it a viable instrument for the development of more secure cyber environments in the future ("Artificial Intelligence and its Application," Onyango), (Everitt et al., 2017).

5. Protective Measures

The concept of security countermeasures spans multiple elements within the domain of information security management (Taylor, 2015). The concept comprises four fundamental elements, namely deterrent, prevention, detection, and cures. Deterrent countermeasures primarily emphasize non-technical approaches, such as the implementation of security regulations and the provision of awareness training, in order to dissuade potential security events. On the other hand, preventive measures cover a range of technologies such as firewalls, intrusion detection systems, encryption, and access controls that are designed to proactively mitigate security breaches (Farahmand et al., 2005). Detection measures encompass the diligent surveillance of suspected vulnerabilities via a range of mechanisms. Remedies, in the context of security breaches, encompass corrective measures implemented subsequent to the occurrence of such breaches. The installation of appropriate countermeasures is necessary in order to lessen the information security risk faced by a company. The task for identifying suitable solutions to mitigate security threats lies with management, albeit requiring diligent managerial oversight (Taylor, 2015), (Goodhue & Straub, 1991), (Taylor & Brice Jr, 2012).

The mere existence of security countermeasures does not inherently ensure a decrease in risks associated with information security. Despite the use of such precautions, the probability of security events continues to be substantial. For example, the efficacy of security policies is compromised when staff do not possess a comprehensive understanding of them. In a similar vein, the inadequacy or incorrect administration of access restrictions can result in the failure to effectively protect computer-based systems. The efficacy of security countermeasures is frequently impacted by human factors, such as errors in installation and configuration (Mattord et al., 2014), (Anderson, 2001).

In order to address these difficulties, it is imperative for responsible management to assess not just the magnitude of risks, but also the feasibility and effectiveness of prospective mitigation strategies (Taylor, 2015), (Farahmand et al., 2005). The alignment between a sensible cost-benefit evaluation and cost-effectiveness may not always be guaranteed, thus requiring a deliberate allocation of resources. As a result, it is imperative for management to make well-informed judgments regarding the prioritization of security threats and the selection of countermeasures that will effectively mitigate risk. Nevertheless, despite previous endeavors to ascertain the most advantageous amounts of expenditure in security countermeasures, a conclusive consensus has yet to materialize. The ultimate determination to enact security countermeasures is a multifaceted process, shaped by limitations in resources and a comprehensive evaluation of hazards within the company (Taylor & Brice Jr, 2012). Given the constraints of limited budgets, companies are required to carefully evaluate concerns pertaining to information security risk management

prior to making decisions linked to security. This process entails the identification, prioritization, assessment, and selection of suitable countermeasures to successfully mitigate risks. It recognizes that security decisions can be susceptible to irrational biases, potentially leaving businesses open to information security threats (Yue et al., 2007).

6. Conclusion

Risk assessment plays a crucial role in the framework of Information Security Management. It is imperative for organizations to implement a methodical and meticulously organized approach in evaluating the risks associated with information security pertaining to their assets. Moreover, the evaluation of risks related to information systems has become a crucial topic due to the growing need for resilient information systems and the heightened importance of data protection. In this particular context, risk assessment plays a fundamental role in ensuring the operational integrity of information systems, serving as an essential component within the broader framework of system architecture. The present study commences by providing a comprehensive overview of essential concepts related to information security. Subsequently, it proceeds to examine the historical development and progression of these concepts across time. Through a thorough examination of the present context, this investigation highlights the notable importance of risk assessment in the wider scope of information security. However, it is important to highlight that there are other possible areas that require additional investigation and improvement in the field of risk assessment. One potential area of study involves a comprehensive examination of the underlying complexities present in control methodologies. Moreover, the utilization of risk assessment models and managerial frameworks in various industrial sectors presents potential for uncovering unique intricacies that are distinctive to each sector. These possible pathways present an opportunity to explore and enhance the field of risk assessment, ultimately contributing to the overall improvement of information security standards.

Acknowledgments

Not applicable

Authors contributions

Not applicable

Funding

Not applicable

Competing interests

Not applicable

Informed consent

Obtained.

Ethics approval

The Publication Ethics Committee of the Canadian Center of Science and Education.

The journal’s policies adhere to the Core Practices established by the Committee on Publication Ethics (COPE).

Provenance and peer review

Not commissioned; externally double-blind peer reviewed.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Data sharing statement

No additional data are available.

References

Abbas, H., Magnusson, C., Yngstrom, L., & Hemani, A. (2011). Addressing dynamic issues in information

security management. *Information Management & Computer Security*, *19*(1), 5-24. https://doi.org/10.1108/09685221111115836

Anderson, R. (2001, December). Why information security is hard-an economic perspective. In *Seventeenth Annual Computer Security Applications Conference* (pp. 358-365). IEEE.

Asosheh, A., Hajinazari, P., & Khodkari, H. (2013, April). A practical implementation of ISMS. In *7th International Conference on e-Commerce in Developing Countries: with focus on e-Security* (pp. 1-17). IEEE. https://doi.org/10.1109/ECDC.2013.6556730

Bacic, E. M. (1990, December). The Canadian trusted computer product evaluation criteria. In *[1990] Proceedings of the Sixth Annual Computer Security Applications Conference* (pp. 188-196). IEEE.

Behnia, A., Abd Rashid, R., & Chaudhry, J. A. (2012). A survey of information security risk analysis methods. *SmartCR*, *2*(1), 79-94. https://doi.org/10.6029/smartcr.2012.01.007

Bialas, A. (2006). Security of information and services in modern institution and company. *WNT, Warsaw*.

Blakley, B., McDermott, E., & Geer, D. (2001, September). Information security is information risk management. In *Proceedings of the 2001 workshop on New security paradigms* (pp. 97-104). https://doi.org/10.1145/508171.508187

Broderick, J. S. (2006). ISMS, security standards and security regulations. *Information Security Technical Report, 11*(1), 26-31. https://doi.org/10.1016/j.istr.2005.12.001

Cranor, L. F., & Garfinkel, S. (2005). *Security and usability: designing secure systems that people can use*. " O'Reilly Media, Inc.".

Everitt, T., Goertzel, B., & Potapov, A. (2017). Artificial general intelligence. *Lecture Notes in Artificial Intelligence. Heidelberg: Springer*. https://doi.org/10.1007/978-3-319-63703-7

Farahmand, F., Navathe, S. B., Sharp, G. P., & Enslow, P. H. (2005). A management perspective on risk of security threats to information systems. *Information Technology and Management*, *6*, 203-225. https://doi.org/10.1007/s10799-005-5880-5

Fozilovich, Y. O. (2022). Artificial Intelligence and its Application in Information Security Management. *Central Asian Journal of Theoretical and Applied Science*, *3*(4), 90-97.

Fredriksen, R., Kristiansen, M., Gran, B. A., Stølen, K., Opperud, T. A., & Dimitrakos, T. (2002). The CORAS framework for a model-based risk management process. In *Computer Safety, Reliability and Security: 21st International Conference, SAFECOMP 2002 Catania, Italy, September 10–13, 2002 Proceedings 21* (pp. 94-105). Springer Berlin Heidelberg. https://doi.org/10.1007/3-540-45732-1_11

Galach, A. (2004). Instruction of IT system security management. *Osrodek Doradztwa i Doskonalenia Kadr Publishing House, Gdansk*.

Gehrke, M., Pfitzmann, A., & Rannenberg, K. (1992, September). Information Technology Security Evaluation Criteria (ITSEC)-a Contribution to Vulnerability? In *IFIP Congress (2)* (pp. 579-587).

Goodhue, D. L., & Straub, D. W. (1991). Security concerns of system users: A study of perceptions of the adequacy of security. *Information & Management*, *20*(1), 13-27. https://doi.org/10.1016/0378-7206(91)90024-V

Granata, P. (2023, March 8). The three types of Artificial Intelligence: A glimpse into the future. *DeltalogiX.* Retrieved from https://deltalogix.blog/en/2023/03/08/artificial-intelligence-a-look-at-its-three-types-and-their-possible-future-implications/

Joint Technical Committee ISO/IEC JTC1. Subcommittee SC 27. (2013). *Information Technology--Security Techniques--Information Security Management Systems--Requirements*. ISO/IEC.

Karabacak, B., & Sogukpinar, I. (2005). ISRAM: information security risk analysis method. *Computers & Security*, *24*(2), 147-159. https://doi.org/10.1016/j.cose.2004.07.004

Karapilafis, G. (2015). Implementation of Artificial Intelligence in INFOSEC tasks and applications. *Journal of Applied Mathematics and Bioinformatics*, *5*(3), 113.

Kuzminykh, I., Ghita, B., Sokolov, V., & Bakhshi, T. (2021). Information security risk assessment. *Encyclopedia*, *1*(3), 602-617. https://doi.org/10.3390/encyclopedia1030050

Latham, D. C. (1986). Department of defense trusted computer system evaluation criteria. *Department of Defense*, *198*.

Mattord, H. J., Levy, Y., & Furnell, S. (2014). Factors for measuring password-based authentication practices. *Journal of Information Privacy and Security*, *10*(2), 71-94. https://doi.org/10.1080/15536548.2014.924812

Onyango, O. O. *Artificial Intelligence and its Application to Information Security Management.*

Peixoto, U. I., Casal-Ribeiro, M., Medeiros-Leal, W. M., Novoa-Pabon, A., Pinho, M., & Santos, R. (2022). Scientific and Fisher's Knowledge-Based Ecological Risk Assessment: Combining Approaches to Determine the Vulnerability of Fisheries Stocks. *Sustainability*, *14*(22), 14870. https://doi.org/10.3390/su142214870

Rannenberg, K. (1993, August). Recent Development in Information Technology Security Evaluation-The Need for Evaluation Criteria for Multilateral Security. In *Security and control of information technology in society* (pp. 113-128).

Rot, A. (2008). IT risk assessment: Quantitative and qualitative approach. *Resource*, *283*(March), 284.

Ryba, M., Poniewierski, A., Sulwinski, J., & Górnisiewicz, M. (2009). The methodology for managing the abuse of IT systems. *Information Security Journal: A Global Perspective*, *18*(3), 107-115. https://doi.org/10.1080/19393550902791457

Schmidt, M. (2023). Information security risk management terminology and key concepts. *Risk management*, *25*(1), 2. https://doi.org/10.1057/s41283-022-00108-8

Siponen, M., & Willison, R. (2009). Information security management standards: Problems and solutions. *Information & management*, *46*(5), 267-270. https://doi.org/10.1016/j.im.2008.12.007

Stoneburner, G., Goguen, A., & Feringa, A. (2002). Risk management guide for information technology systems. *Nist special publication*, *800*(30), 800-30. https://doi.org/10.6028/NIST.SP.800-30

Sundu, M., & Ozdemir, S. (2020). The effect of artificial intelligence on management process: challenges and opportunities. *Challenges and Opportunities for SMEs in Industry 4.0*, 22-41. https://doi.org/10.4018/978-1-7998-2577-7.ch003

Susanto12, H., Almunawar, M. N., & Tuan, Y. C. (2011). Information security management system standards: A comparative study of the big five. *International Journal of Electrical Computer Sciences IJECSIJENS, 11*(5), 23-29.

Szczepankiewicz, E. I., & Szczepankiewicz, P. (2006). Risk analysis in IT environment for the Purpose of Operational Risk Management. Part 3–Strategies of Dealing the Operational Risk. *Monitor Rachunkowosci i Finansow*, *8.*

Szyjewski, Z. (2004). Methodologies of IT projects management. *Placet, Warsaw*.

Task, J. (1993). *Foundations for the Harmonization of Information Technology Security Standards.*

Taylor, R. G. (2015). Potential problems with information security risk assessments. *Information Security Journal: A Global Perspective*, *24*(4-6), 177-184. https://doi.org/10.1080/19393555.2015.1092620

Taylor, R. G., & Brice Jr, J. (2012). Fact or fiction? A study of managerial perceptions applied to an analysis of organizational security risk. *Journal of Organizational Culture, Communications and Conflict*, *16*(1), 1.

Tofan, D. C. (2011). Information security standards. *Journal of Mobile, Embedded and Distributed Systems*, *3*(3), 128-135.

Tolani, M. G., & Tolani, H. G. (2019). The use of artificial intelligence in cyber defense. *International Studies Journal of Engineering and Technology (IRJET)*, *6*(7), 3084-3087.

Volkan Evrin, C. I. S. A., & CRISC, C. (2021). *Risk Assessment and Analysis Methods: Analysis and Quantitative Risk Assessment and Analysis Methods: Qualitative and Quantitative*. (n.d.). ISACA. Retrieved from https://www.isaca.org/resources/isaca-journal/issues/2021/volume-2/risk-assessment-and-analysis-methods

Wawrzyniak, D. (2006). Information security risk assessment model for risk management. In *Trust and Privacy in Digital Business: Third International Conference, TrustBus 2006, Kraków, Poland, September 4-8, 2006. Proceedings 3* (pp. 21-30). Springer Berlin Heidelberg. https://doi.org/10.1007/11824633_3

Wawrzyniak, D. (2007). *Models of IT risk assessment–classical approach and possibilities of its development.*

Yue, W. T., Çakanyıldırım, M., Ryu, Y. U., & Liu, D. (2007). Network externalities, layered protection and IT security risk management. *Decision Support Systems*, *44*(1), 1-16. https://doi.org/10.1016/j.dss.2006.08.009

Computer and Information Science; Vol. 16, No. 4; 2023
ISSN 1913-8989 E-ISSN 1913-8997
Published by Canadian Center of Science and Education

Electronic Health System Integration Framework for Secure M-Health Services: A Case of University of Nairobi Hospital

Samuel Nandasaba[1], Gregory Wanyembi[1], & Geoffrey Mariga Wambugu[2]

[1] Mt. Kenya University, Kenya

[2] Murang'a University, Kenya

Correspondence: Samuel Nandasaba, Mt. Kenya University, Kenya.

Received: September 1, 2023 Accepted: October 5, 2023 Online Published: November 20, 2023

doi:10.5539/cis.v16n4p21 URL: https://doi.org/10.5539/cis.v16n4p21

Abstract

The purpose of this article sought to design a secure framework that can be used in M-Health systems development. The researcher used the integrated information theory as a framework for enforcing system security as a holistic approach. To actualize this study, objectives that were meant to guide in carrying out the research were: To evaluate the significance of Confidentiality, Integrity and availability on the security of M-health systems and to develop a framework for secure integration of M-Health systems. The researcher used University of Nairobi Hospital because of ease of accessibility and financial resources available to conduct the research. The study adopted a cross section survey design methodology that included a sample size of 44 ICT personnel and users of University Health System at the University of Nairobi Hospital. Data collection methods were observation, conducting interviews and filling questionnaires that were administered to the target population in the University Hospital. The target population were handed the questionnaires and had them filled. The filled in questionnaires were then picked later from the respondents. SPSS version 23 was used for data analysis, then presented in frequency tables, bar charts, pie charts and standard deviation.

Keywords: availability, confidentiality, integrated information theory, integrity and M-Health

1. Introduction

With the advancement of wireless information technologies and applications, a rapid rise has been recorded in the use of smartphones, tablets, and other electronic gadgets in the health sector. Researchers have developed frameworks such as (Maranda, 2016),

(Gejibo, 2015), (Nkosi, 2014), (Leon et al., 2012), and (Elkhodr, 2012). In addressing the M-health information services, these frameworks are faced with security challenges, the major being confidentiality, availability, and integrity, this is negatively affecting the usage of the frameworks in sorting out the security risks. (Vimalachandran et al., 2018), because of the effects on encouraging good standards of patient care, maintaining CIA data in EHR systems has grown to be a significant issue. This research therefore aims to offer an intervention by proposing an integration framework of EHIS into M-health with much focus on the security aspect to enhance M-health applications security. Iwaya et al (2020), it has become clear that security and privacy are the most difficult parts of healthcare information systems, and it is vital to properly comprehend and handle the security concerns of M-Health.

2. Objectives

Research Objectives

The research was guided by the following objectives:

Main Objective

The main objective of this study was to design a framework for electronic health system Integration for secure M-health services at the university of Nairobi hospital.

Specific Objectives

The following two specific objectives served as the study's guide in order to accomplish the overall goal.

1. To evaluate the significance of a comprehensive information security of M-health systems at the

University of Nairobi Hospital.

2. To develop a framework for secure integration of M-Health systems at the University of Nairobi Hospital.

Research Questions

1. How can comprehensive information security be realized in an M-health system at the University of Nairobi Hospital?
2. How can a secure integration framework of M-Health system at the University of Nairobi Hospital be developed?

2.1 Justification

Cyber security has been a key challenge to computer-based systems in the last few years. It is important to note that systems security should be considered as a part of the system development in any information system and not an implementation requirement. Therefore, the ability to analyze the key security demands of the system and integrate them within the computer-based system as it is being built is critical. Furthermore, for health information systems, preservation of privacy of a patient's health data is one of the key tenets of any health system/facility. Thus, a comprehensive security model must be enforced in any kind of health system for it to be considered effective and beneficial.

The study focuses on the influence of electronic health system integration framework for secure m-health services: a case of university of Nairobi hospital. The study will adopt a cross section survey design.

Mobile Health applications development plays a crucial role in today's lives with the increasing number of tablets and smartphone users. Mobile technology is growing exponentially and therefore organizations and government are making use of their power to collect, collate, transmit and present data in a timely manner hence overcoming limitations that are in manual systems and the University of Nairobi hospital is not an exception.

Fast growth of mobile technology has made it possible for electronic systems to share data more often, therefore providing decision makers with useful information and improving their capacity to have very important decisions about health matters. While mobile phone technology has shown tremendous potential to transform health-care distribution, there is little guidance to keep university of Nairobi hospital developers updated about the development of secure frameworks for M-Health systems.

2.2 Literature Review

2.2.1 Theoretical Framework

This section looked at the various theories that were used to inform the study on security features of an E- health system. The study was founded on one theory, the Integrated Systems Theory. Specifically, literatures pertaining to health system information security in health care systems were reviewed.

2.2.2 The integrated Systems Theory

This theory was proposed by (Hong et al.,2003), as an interdisciplinary theory dealing with any structure of nature, culture, and multiple empirical disciplines, as well as a paradigm with which a phenomenon can be studied from a systematic perspective (Capra,1997). Integrated System Theory entails enforcement of information security policies, management and assessment of risks, information Auditing and internal controls. Consequently, information security is covered comprehensively in terms of many aspects by integrated system theory. It describes organizational actions in terms of information security management and techniques, as well as including alternatives.

In order to fully comprehend information security management, explain information security management techniques, and anticipate management outcomes, integrated systems theory is crucial for the study. Consequently, the theory offers a solid foundation for evaluating the level of information security controls implemented at the University of Nairobi hospital. Internal control is the prevention, detection, and correlation of system-related activities in order to prevent unauthorized and illegal access.

Controls can also be referred to as administrative, operational, and technical measures that safeguard the system's availability, integrity, and confidentiality.

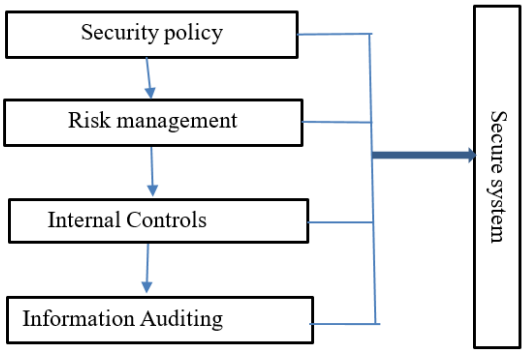


Figure 2. Integrated system Theory (Adapted from Hong et al., 2003)

The nature of this theory makes it difficult to adapt to highly dynamic surroundings, and it also takes a top-down approach that may not be consistent with reality.

We need to consider M-health as a service delivery system; as argued above in Chapter 2, Integrated System Theory entails enforcement of information security policies, management and assessment of risks, information Auditing and internal controls. Consequently, information security is covered comprehensively in terms of many aspects by integrated system theory. It describes organizational actions in terms of information security management and techniques, as well as including alternatives. The relevance of the theory in the management of M-health services is deficient in the fact that the theory fails to holistically look at the overall IT infrastructure holistically but instead lays its overall emphasis on information security management.

The big question that this research needs to ask is therefore, how would a holistic integrated information security delivery system look like? What are the optimal IT infrastructure and what security principles would govern the measures that would be put in place to guarantee the security of services for which the system is built?

According to IBM, IT infrastructure configurations vary based on organizational requirements and objectives, however some objectives apply to all businesses. Business high-performance storage, a low-latency network, security, an efficient wide area network (WAN), virtualization, and zero downtime are all features of the ideal infrastructure. This means that to guarantee the security of services supported by these systems, vulnerabilities to each of the components of these systems must be adequately analyzed and addressed.

In our case, M-health systems are critical not only for provision of accurate medical information on patients, but the preservation of this data and a guarantee that only those with the authority to access it can access it whenever they need it. It therefore implies that the collection, storage and processing functions must be secure.

High-performance storage systems include a data recovery system and data archiving and backup capabilities.

Low-latency networks use enterprise-level infrastructure elements to minimize data flow delays.

Secure infrastructures systems that regulate data availability and information access. Regardless of where the data is kept, it may protect a company from hacks and breaches while maintaining customer trust.

WANs prioritize traffic on the network and adjust the bandwidth allocation for certain applications as necessary.

Virtualization increases uptime, enhances disaster recovery, and saves energy while providing quicker server provisioning.

Zero downtime to keep costs low and earnings high, this strategy tries to minimize system outages and business operations disruptions.

There is therefore need to divide the security of M-health services into two

Information security: (principles include confidentiality, availability, integrity)

MHI security (M-health infrastructure security): mhealth infrastructure includes enterprise servers, data storage servers, Mainframes, mobile devices and software and operating systems

Further, in having a holistic view of secure M-health services, we need to consider that any effective Service Delivery system strategy is comprised of five key components (IBM,2016):

- Service level management
- Financial management for IT services

- Capacity management
- Availability management
- IT service continuity management

Therefore, any effective M-heath system must be able to integrate all these functionalities and requirements into the system and its operations.

3. ICT and Healthcare Systems

Patient Care Information Systems (PCIS) deployment in healthcare organizations has not been successful. This has been so because of a number of challenges that may be faced when the systems are being implemented or thereafter (Berg, 2001). Healthcare Information and Communication Technologies (ICT) are complex operational technologies whose applications, purposes, disadvantages, and ramifications are not well defined, nor are the advantages of usage guaranteed. However, there are some compelling theories about how IT can be used in healthcare to increase efficiency, consistency, and connectivity in order to promote acceptance and guide effective deployment of e-healthcare systems. ICT use ideas are discussed among a group of partners that include medical practitioners, representatives of healthcare organisations, legislative and regulatory authorities, as well as ICT suppliers and consultants. This interactions between a group and systems form and decide the consequences of healthcare ICT technologies. As a result, understanding the social development and interpretive mechanisms by which healthcare ICT technologies are created and shared is important for forecasting consequences of ICT implementation and informing policymakers of the threats presented to patient information contained within these digital networks.

Information security.

Information systems (IS) are highly depended on by organizations. Consequently, these firms employ technical controls to lessen information security risks (Gundu & Flowerday, 2013). Risk identification.

This refers to the process where potential risks of a project and their characteristics are listed. The results are usually recorded in a risk register.

3.1 Risk Management.

Risk refers to the possibility of something adverse happening. Risk management is therefore about transforming organizational culture to accept risk and facilitate risk discussion when doing business activities or making any strategic investment on various projects.

3.2 Risk Mitigation

It is a strategy in preparation for and lessen the effects of threats faced by a system.

3.3 Existing Frameworks for M-Health Systems.

There are researches that have been done on the frameworks for M-health and have been published in various referred journals as discussed below.

3.3.1 A Framework for Assessing M-Health Challenges in South Africa.

A qualitative study conducted in South Africa to review the benefits and challenges of M-health in community-based health services. There were four key system dimensions that were identified and assessed. These were;

- Government stewardship
- Organisational
- Technological
- Financial

According to the report, prospects for effective M-health adoption in South Africa include a high prevalence of cell phones, a positive policy framework for M-health, successful use of M-health for community-based health programs in a variety of initiatives, and a well-developed ICT industry. However, there were some shortcomings in other main aspects of the health system, such as corporate culture and potential for using health knowledge for management, as well as a lack of access and usage of ICT in primary health care. The complexities of ensuring interoperability and convergence of information systems, as well as ensuring information safety, is among the technical challenges. There was also the issue of sufficient financing for large-scale M-Health usage in a resource-constrained world. The limitations of this study was that it never dealt on technological Security of

M-health systems.

The framework that was developed is as shown in the figure 3;

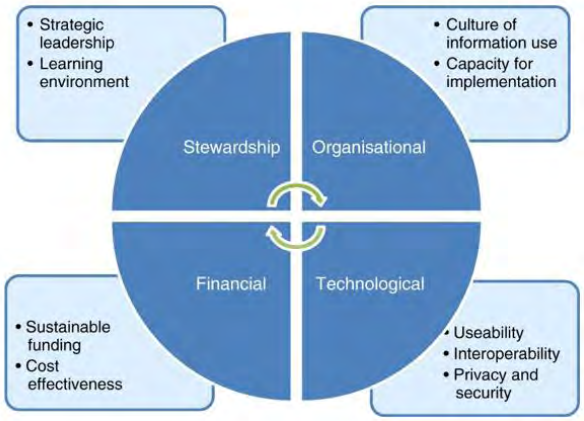


Figure 3. A framework for assessing M-Health challenges in South Africa (Leon et al., 2012)

The figure shows four health systems dimensions that should be addressed when assessing the complexities of scaling up M-Health from a health systems perspective.

3.3.2 M-Health Decision Making Framework for Community Based Services

(Maranda ,2016), the researcher encompassed additional security procedures using encryption, integrity of the data and the security keys. In the encryption perspective, encrypted data was sent to the server .The researcher used built-in libraries for encryption of string data. The messages were transferred in XML format to the server

(Maranda,2016) implemented integrity of the data using Digital signatures. The Digital signature ensured that the message that was sent was exactly what was received. The signature depended on the encryption which assured authentication.

The Digital signature solely relied on a private key algorithm. This meant that the message owner was the only one who knew it but the public key was known. The researcher used checksum algorithms and a checksum function to transform the input and produced a numerical output of smaller size.

He suggested the use of security keys. In this case he used 128bit strings which were delivered to the server before requesting sensitive data. The security keys were encoded with chosen cryptographic algorithm and was unified with entire application but quite independent from the device.

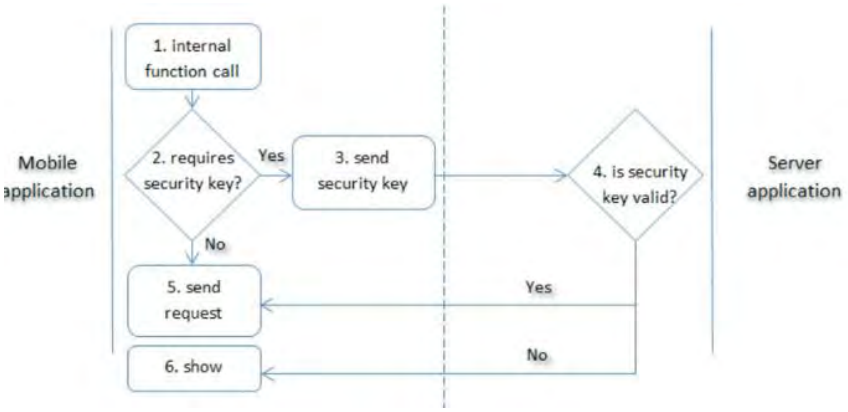


Figure 4. Data transfer using encryption and digital signatures. (Maranda, 2016)

In this study, the researcher introduced three core components that must be taken into account while developing and deploying mobile applications. These components were: data transfer, storage, and access. All these components must be viewed as equally important all through all phases of development. The Secure Development Strategy detailed the assumptions and frameworks that ought to be enforced within the application context to provide mobile protection. The most important feature of Secure Development Strategy is that it embraced all crucial aspects by describing the concepts and grouping them. The geo-location and ADID

(Application Device Identifier) for data access, the encryption of sensitive data in database files, and the encryption of requests transmitted over the internet with digital signatures and security keys were not addressed by the researcher in this study. The Secure Development Strategy's objectives were to limit the number of potential risk points in the program rather than to completely protect it from assault by preventing potential attackers from encrypting important data. With the use of a developed security architecture called iSec, the Secure Development Strategy's pillars were really put into practice.

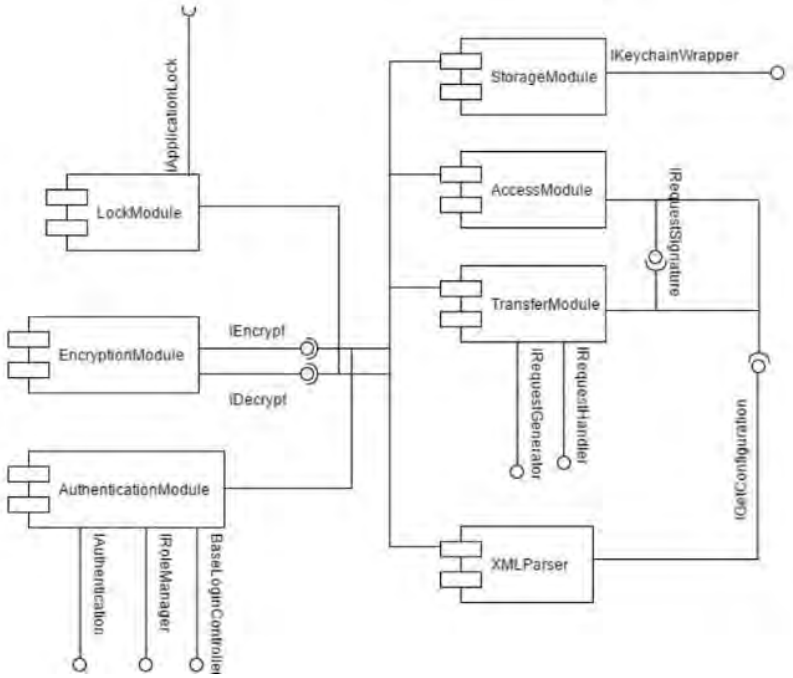


Figure 5. Components of the iSec framework. (Maranda, 2016).

3.3.3 Mobile Application Security System Framework

For mobile application networks, Floyd (2006) developed creative layering of security mechanisms that offered a dispersed security solution. Code signature variation, remote hashing, a mobile application generator, application time to live, resident monitor applications, distributed application monitoring, code obfuscation, and hashing algorithms were all provided by the researcher in a way that maximizes benefits while minimizing overhead. The researcher offered a distributed method that could keep safeguarding even if hosts and programs were deleted, corrupted, or destroyed. The integrity and security of the application system were strengthened by the security measures described in this study.

The limitations of these research is that the researcher didn't address methods for detecting and preventing denial of service (DoS) attacks and a secured interprogram communications. Diverse application kinds must collaborate through interprogram communications in order to carry out an essential task.

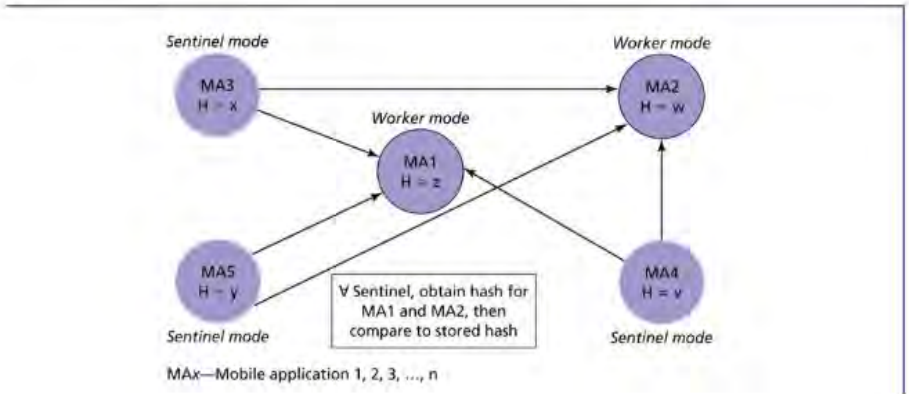


Figure 6. Analyzed hashes obtained for the conflicting parties. (Floyd, 2006)

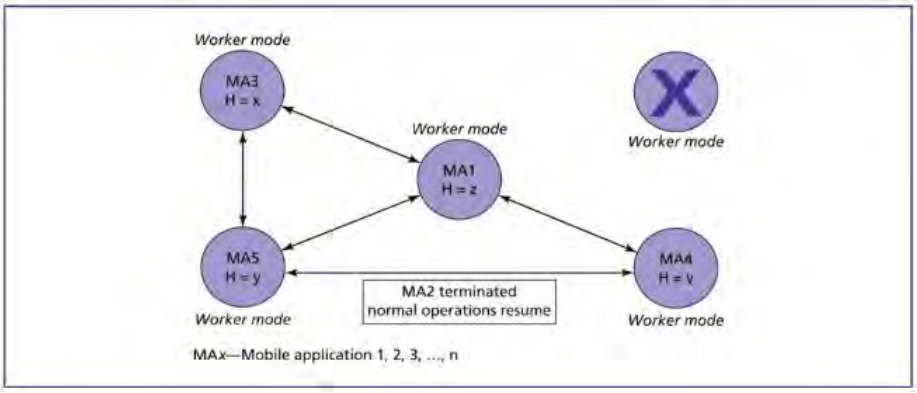


Figure 7. Corrupt application identified by consensus vote (Floyd, 2006)

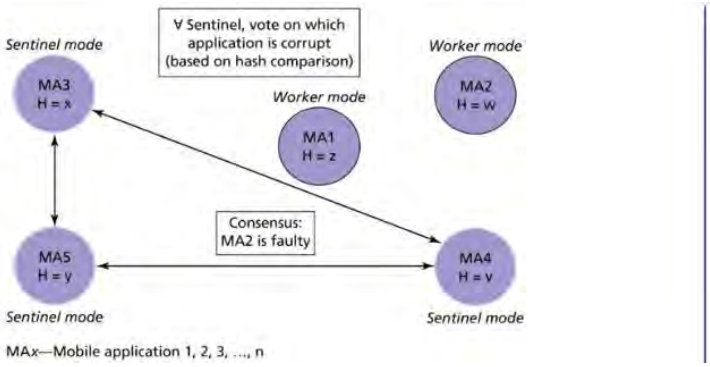


Figure 8. Termination of corrupt application from the network. (Floyd, 2006)

3.3.4 A Framework to Improve Mobile Banking Security

In this case, the researcher recommended providing banking customers with a smartphone device so they could safely access their business or personal accounts from any location at any time.

To resolve these security issues, a confidence negotiation approach was suggested in this report. As the underlying protocol, trust negotiation was paired with Transport Layer Security (TLS). This technological mix aimed to improve the current security of M-banking applications.

The proposed framework verified the requesters and their devices. These gave users the chance to register their mobile devices and provided a mechanism for financial organizations to confirm that the device was being used. By automating the authentication process, this strategy improved two factor authentication.

The limitation of this study is that it did not factor in the location verification method to the m-banking system.

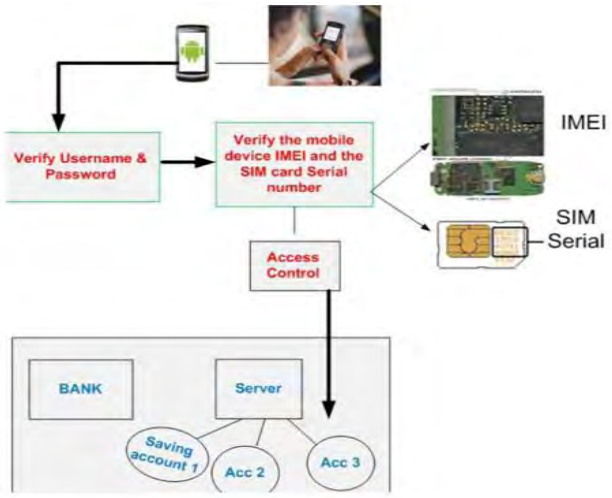


Figure 9. Sec ure M-banking model ((Source, Elkhodr (2012))

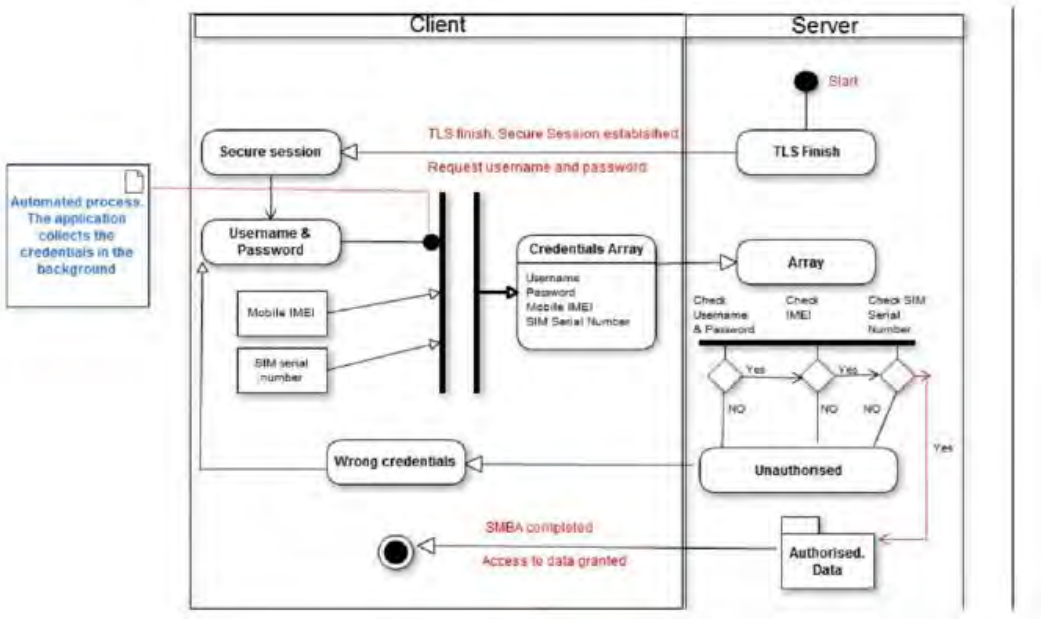


Figure 10. Secure mobile banking approach framework activity diagram. (Elkhodr, 2012)

3.3.5 An Enhanced Mobile Health Applications Cloud Computing Framework

The researcher described the difficulties that mobiles encounter when providing Secure Multimedia-based Health Services due to computation and power supply constraints in this report.

In this case, the researcher postulates that mobile devices are not able to perform complex multimedia and security Algorithms due to the fact that they run on small batteries and also have inadequate computational capacity, as a result, researcher devised a cloud computing platform to support mobile devices while running heavier multimedia and encryption algorithms in the distribution of mobile health services. In this study, the suggested framework makes use of a cloud computing protocol management approach to offer mobile devices security as a service (SaaS) and multimedia sensor data processing. The researcher in this study hypothesized that security and multimedia operations may be carried out in the cloud, enabling mobile health service providers to subscribe and expand the features of their mobile health applications beyond the limitations of currently accessible mobile devices.

In this research, the security of mobile health systems data was not addressed by the researcher hence the need to carry out more research on this topic.

The initials of the diagram

NI- is a non-intrusive sensor that is used to gather the required sensor signals, which are then fed to an embedded digital signal processor (DSP) in a mobile device.

SIPS-is the session initiation protocol signaling

SIP-EP-is the session initiation protocol event packet, this connect the IMS client to the call session control functions (CSFC).

The CSFCs are SIP proxy servers, supporting IMS signaling and session control functions. XDMS- is the database management system which controls and organizes data created by the health monitoring services.

The Application server hosts the ongoing mobile service and sends and receives data from the IMS client. The application server also functions as a branch of the Home Subscriber Server (HSS), which is the primary repository of mobile-related user data. The IMS system monitor acts as the recipient and interpreter of the sensed physiological information and therefore relays back the necessary decision and action to be taken.

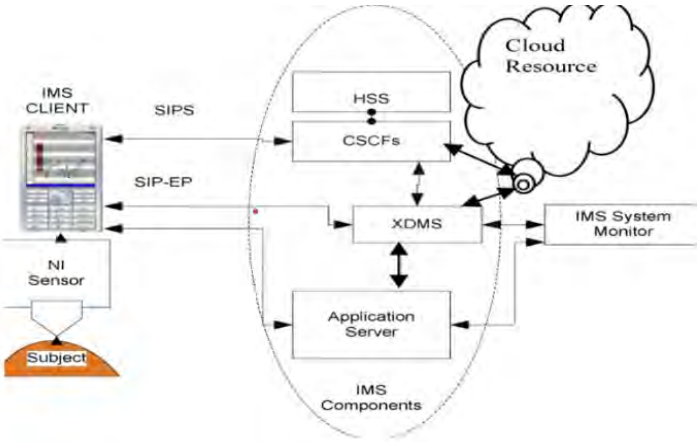


Figure 11. A framework of IP multimedia subsystem standard –based mobile health monitoring with cloud support. (Nkosi&Mekuria, 2010)

3.3.6 Secure Mobile Data Collection System Framework

Gejibo (2015) developed a secure mobile data collection system framework with a set of modular security features that were built to meet MDCS standards and design criteria. The framework included simple interaction interfaces. The framework was created to be flexible, scalable, and adaptable to various MDCS security settings while simultaneously being secure by default. The framework's primary goal was to offer an all-encompassing safe solution for user identification, secure mobile and cloud storage, and secure communication.

Authenticator: a security module that dealt with account recovery, remote server authentication, and user authentication on mobile devices. With a default concrete implementation, it offered the authentication services through straightforward interfaces.

This particular module received the user authentication delegation from the MDCS client. As a result, whenever an attempt was made to access the MDCS (mobile data collection systems) client, the Authenticator module was invoked. The Authenticator is adaptable and may be set up to offer further capabilities like single sign-on and device authentication. The requirement that a phone may be shared by several collectors who should not have access to each other's acquired data was the major justification for the module's existence.

2. Secure Storage: Security module in charge of managing and protecting the mobile device's MDCS application resources. With a default concrete implementation that handles encryption, decryption, cleaning up leftover data once the user logs out, and a recovery strategy in case the application crashes or the battery runs out, the secure storage is available via straightforward APIs.

3. Secure Communication: is a security component in charge of creating a secure tunnel between the client and the server. A popular protocol for protecting HTTP messages is Hypertext Transfer Protocol Secure (HTTPS).

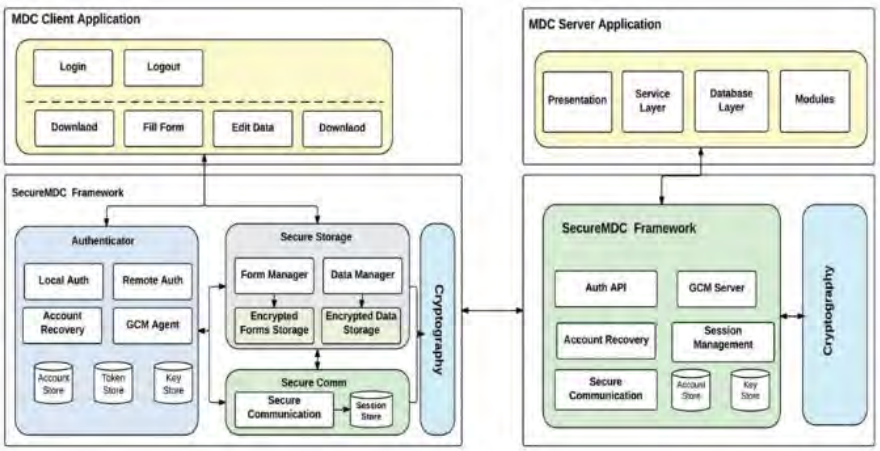


Figure 12. showing the modules and the explanation(Gejibo,2015)

3.3.7 SecourHealth Framework

Simplicio etal (2015) in his SecourHealth, developed a minimal security framework with a focus on very sensitive data collection systems. In this he identified various pillars which included:

I. Lack of connectivity and tolerance delays- Users were able to function offline when necessary and authenticate themselves on any device they had previously registered.

II. Loss or theft device protection- When a device is stolen while a legitimate user's session is still active, this module has a method that restricts the attacker's ability to receive information from the server.

III. Data transfer between a mobile device and sever in secure manner- In this module, even in the absence of an underlying secure connection, all data sent between a server and a mobile device was encrypted and authorized.

IV. By incorporating this security framework into the GeoHealth system and the Android-based "Family Health Program" application, which were both used by the government to collect health data in Sao Paulo city, Marcos (2015) put this security architecture into action.

3.4 Conceptual Framework

Figure 12 shows the conceptual framework for this research. In order to assess your electronic health record security, credibility, and availability requirements, you must first thoroughly consider your practice's health IT climate. This could include the technology your profession uses for both therapeutic and institutional purposes, as well as when and how those technologies are physically used and located within your practice. Consider the circumstances that could result in unwanted entry, use, leak, interruption, alteration, or loss of electronic health records as you assess the health IT climate. These circumstances are likely to be specific to the practice and can take the form of technology problems (e.g., a lack of securely installed computing equipment), procedural issues (e.g., a lack of a security incident management plan), or staff issues (e.g., lack of comprehensive information security training).

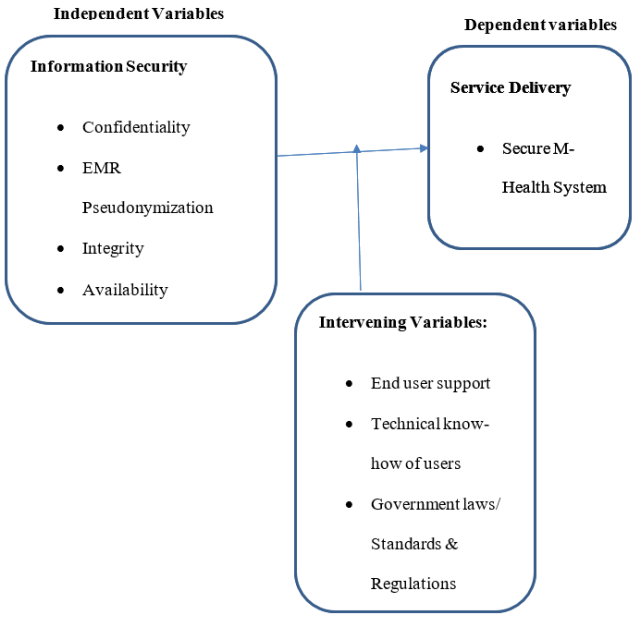


Figure 13. Conceptual framework, (Researcher, 2023)

3.4.1 Confidentiality

The security of information contained inside networks from unwanted or unintended access is referred to as confidentiality.

–Privacy is an individual's right to choose when, how, and to what degree knowledge about them is transmitted to others" (Brands, 2003, pp2). Privacy involves the right of the individual to be left alone, to withdraw from the influence of his environment (Innab, 2018). Additionally confidentiality relates to disclosure or nondisclosure of information.

Patients agree that with their permission their medical data can be shared with other institutions such as insurance companies to facilitate the payment of their dues. Therefore, the patients only understand that the shared individual medical data can only be used for the intended purpose.

Furthermore, outside the hospital, patients have expectations that individuals will not be given access to confidential information or institutions not permitted to keep such content. The content's genuine users won't abuse this access for uses other than those for which it was intended in the first place.

3.4.2 Integrity

The property whose electronic health records has not been tampered with or lost in an improper way is referred to as integrity. It refers to the accuracy and continuity of data stored about a person, entity, or event, in this case, the patient (Charitoudi & Blyth, 2013). Integrity of data encompasses documentation accuracy throughout the entire health record. It entails patient identification, information governance, record correction and validation of authorship. Additionally, the accuracy of the data provided at the time of capture has a significant impact on the quality of the data in the EHR.

The quality of a patient's healthcare may be significantly impacted by inaccurate health information. As health information becomes more computerized and the extent of organizational interchange of health information expands into Health Information Exchanges, maintaining the accuracy and completeness of health data is essential (HIEs) (Kellerman & Spencer, 2013).

(Lucas, 2013) postulates that the accuracy, reliability, and completeness of the demographic information related to or associated with a specific patient is known as patient identity integrity.

Availability

Availability refers to the property where electronic health information can be accessed and used when demanded by an authorized person. System availability looks at the period of up-time for operations and is a measure of how often the system is alive and well. It is always denoted as (up-time)/(up-time + downtime) with numerous variants (Ahmed & Mousa, 2016). Up-time and downtime refer to dichotomized conditions. (Charitoudi & Blyth, 2013) states that Up-time is the ability to perform assigned duty as downtime denotes not being able to perform the given task.

When a system is up and running and ready for use, it is said to be available. A system may go offline for a variety of reasons, ranging from scheduled maintenance downtime to catastrophic failure (Innab, 2018). The goal of high availability solutions is to reduce this downtime and/or the amount of time it takes to recover from an outage (Zdravkova, 2015). How much downtime may be permitted will influence the solution's comprehensiveness, complexity, and cost.

On the extreme end of the scale, high availability can literally refer to a disaster response plan that can get an organization back up and running as soon as possible. For small systems, this may be as straightforward as an uninterruptible power source and a strict backup strategy. The peak of consistent availability, exemplified by comprehensive workload-sharing solutions distributed across several sites, is at the other end of the spectrum. There are differing degrees of availability between these two extremes (Nganji & Nggada, 2011). Computing systems availability has been described using various concepts: high availability computing; fault tolerant systems; system redundancy (Lizasoain et al., 2015). The idea behind all this is to ensure that no matter what happens,

users must be able to access the systems for the data and information that they require. In the case of e-health, system failure can be initiated at five levels.

Mobile device failure (hardware/software)

Network outrages Or server failure (hardware or software)

Wherever any of this happens, it leads to delays in access of the systems or failure by the users to perform the functions for which they are supposed to. Thus, system designers must do their best to develop contingency plans that will ensure continued access of these vital systems.

3.4.3 Pseudonymization

A pseudonym can help protect privacy. By using a method known as pseudonymization, sensitive data can be secured while still providing people access to less important components. To handle sensitive data, this method replaces crucial data elements with pseudonyms. This method prevents immediate access to the information.

3.4.4 Standards and Regulations

Any electronic health records system that ensures the privacy and security of patient data must adhere to standards. The Health Insurance Portability and Accountability Act (HIPAA) of 1996 is the regulation that is most frequently used in the US. Health information is protected by federal law known as HIPAA, which also guarantees that patients can access their own medical records. The people in charge of safeguarding this information now have additional obligations. HIPAA sets security standards that are comprised of four areas as stated by Maiwald (2005). These four sections include technical security services, physical safeguards, technical security mechanisms and finally administrative procedures. The main objective of HIPAA is to keep customers' and employees' personal health information private, secure and confidential. The personal Health information must also be maintained in a manner that ensures high integrity and high availability in the event of an emergency. Another standard in use is the CEN/ISOEN3606-PartIV in Europe, which includes privacy and security directive. The CEN created this standard in 2008, which was then modified in 2010 with ISO approval. This standard's primary goal was to offer universal guidelines for creating interoperable electronic health systems. Efficiency, cost savings, and risk avoidance are the practical justifications for implementing standards.

3.4.5 Common Data Security Architecture

This is an open and extensible software framework and API specification that usually addresses communication and data security requirements. It was originally developed by Intel Architecture Lab (IAL).The main objectives of Common Data Security Architecture were:

Encourage interoperable ,horizontal standards

Offer essential components of security capability to the industry.

The Common Data security Architecture has three layers which include

System security services-this offers language interface adapter

Common security services manager-this is a cryptographic service provider, which performs bulk encrypting, digesting and digital signature in addition there is trust policy modules which implement policies defined by authorities and level of trust required to perform certain actions.

Security add-in modules- this layer provides modules that offer basic components in cryptographic algorithms and storage.

4. Methodology

4.1 Introduction

The section outlined the method to use in the study focusing on the significance of electronic health system integration framework for secure m-health services: a case of university of Nairobi hospital. The study adopted a cross section survey design. It outlined research design, the population of study, sample size, and data collection methods from the study participants and the tools that will be used for data analysis.

4.2 Study Design

According to (Wausi, etal, 2009), a research design is an account of the logical steps used to connect the research questions and procedure to data collection, analysis, and interpretation in a coherent manner. According to (Tarus, et al., 2015), the researcher points out that in a descriptive study, the researcher can use results obtained from the sample to make a generalization about the entire population. This study adopted a cross section survey design, the study collected both qualitative and quantitative data from various respondents by conducting interviews with participants or giving a questionnaire to the intended audience.

Study Population and sample size determination

A systematic random sampling technique was applied in this study. Equal opportunity was given to each person to participate in the study. The population for this study was staff using the University Health systems and the system developers at the University of Nairobi Hospital. The study's sample size was 44 system users and ICT personnel. The researcher was convinced that because they are more familiar with how the system works, the ICT staff members had the necessary knowledge of University Health systems. The two objectives of the study was used to come up with research questions. The study site for this research was chosen because they have had an E-Health system for the last five years which enabled collection of the required data successful.

Less than 100 of the employees listed on the University of Nairobi Health Services website meet the criteria for inclusion. The sample size of the staff who were willing to participate were determined by using Yamane Taro's

sample size calculation formula (Yamane, 1967).

n=N/(1+N(e)2)

Where:

n is the sample size of target population required for the study

N is the total population size of target population

e is the level of precision (error estimate) which is 0.05

n=N/(1+N(e)2)= n=100/(1+100(0.05)2)= 44 participants

44 people were therefore be contacted to participate in this study.

4.2.1 Pre-test

To guarantee that all study parameters are tested from the target group, the researcher did pre-test for the created questionnaire. Pre-testing was done by the researcher distributing questionnaires to a few random participants at the UoN Hospital. If queries arose, then the researcher was able to make necessary changes to the tool to ensure that the required information was captured during the study.

4.2.2 Data Collection

The primary method of data collection for this project was through questionnaires. The questionnaires were left and picked later at an arranged time by the respondents. To ensure a high response rate and to help when respondents sought clarifications, there was follow-ups via email, phone calls, and visits as needed. The questionnaire was administered to ICT staff and users who have roles in the University Health System (UHS). In addition to questionnaires the researcher used observation and structured interviews in order to gain more information on the Security of University Health System.

4.2.3 Data Management

Once data was collected, the questionnaires were checked by researcher to ensure that none was incomplete. Once this was done, the researcher stored these questionnaires in a lockable cabinet where they were safe. Data entry was then done followed by data analysis and finally the researcher once again stored the questionnaires in a lockable cabinet.

4.2.3.1 Data Analysis and Presentation

The information from the respondents' completed questionnaires was coded and entered into a computer statistics tool. Data analysis and the presentation of the findings were done using SPSS version 23.0, a statistical package for the social sciences. Correlation and Regression data analysis techniques were used.

4.2.3.2 Correlation Analysis Technique

In statistics, correlation means that there is a relationship between various events. In statistics, the term "correlation" refers to the relationship between various occurrences.

For the purpose of conducting a reliable correlation study, detailed observations of two variables are required, which gives us a benefit in terms of acquiring results. In order to examine the level at which the variables under study were related, Karl Pearson’s as a measure of coefficient of correlation was used. The Pearson product-moment correlation coefficient denoted as r was primarily used to gauge the level of association between two variables and ranges from +1 to -1. Lack of relationship between variables is denoted by zero value. Where there is a positive relationship, a figure greater than zero is in indicated whereas a negative relationship is indicated by a figure less than zero.

Simple metrics: Findings from research can be categorized easily. The results can be between -1.00 and 1.00. There can only be three possible overall conclusions from the analysis.

4.2.3.3 Regression Analysis technique

A multiple regression analysis was undertaken to further gauge the association among the independent variables on secure service delivery at the University of Nairobi Hospital. To aid this SPSS V 21.0 was used to facilitate the outcomes of the multiple regressions for the study. Predict a research in the near and long term, Understand service security levels. Review and comprehend how various factors affect each of these things. The extent to which changes in the dependent variable (secure service) was influenced by all the four independent variables (Information integrity, Pseudonymization, information confidentiality and information availability) was explained by the coefficient of determination.

Yi=f(xi,β)+ɵi

Where Y is the dependent variable which was secure M-health system

Xi was the independent variable of Confidentiality.

Xii was the independent variable of pseudonymization.

Xiii independent variable integrity

Xiv independent variable availability

Thus

Y=(Xi+Xii+Xiii+Xiv,β)+ɵi

In both techniques, results were presented using tables, frequency charts and graphs, and the findings will be presented using tables, graphs, bar charts , pie charts, mean and also standard deviation.

4.2.4 Ethical Considerations

Approval to conduct research was sought from the Mount Kenya University ethical review committee as well as National Commission for Science, Technology and Innovation (NACOSTI) clearance certificate before the commencement of the study.

5. Research Analysis, Findings

5.1 Results

The main objective of this study sought to design a model for electronic health system Integration framework for secure M-Health information systems.

This objective was achieved and managed to evaluate and investigate the existing frameworks for electronic health Integration. The review of the current form of framework revealed that mobile based Health Information systems are unreliable and do not enable professional health workers access to patients' data at any given time.

The results from this project revealed that over 70% believed introduction of Confidentiality, Integrity and availability on the security of M-health systems would make University Health systems processes convenient.

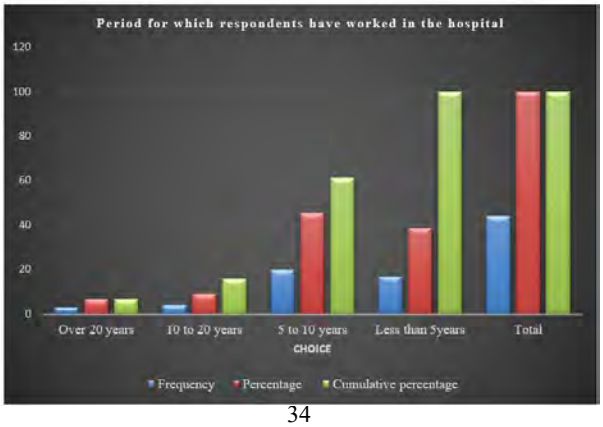
The second objective was to develop a framework for secure integration of M-health systems.

The framework was developed, built and tested. University Health System framework was found to be working well, consisting of entities for security measures. The developed framework was validated. Evaluation of its applicability and usability revealed that it can reduce the vulnerability and improve security level of university health Systems, thus making seamless intervention where M-Health security concern is raised.

The first question is to find out period for which the respondents have worked in the hospital.

Table 1. Period which respondents have worked in the Hospital

Choice	Frequency	Percentage	Cumulative percentage
Over 20 years	3	6.82	6.82
10 to 20 years	4	9.10	15.92
5 to 10 years	20	45.46	61.38
Less than 5years	17	38.62	100.00
Total	44	100.00	



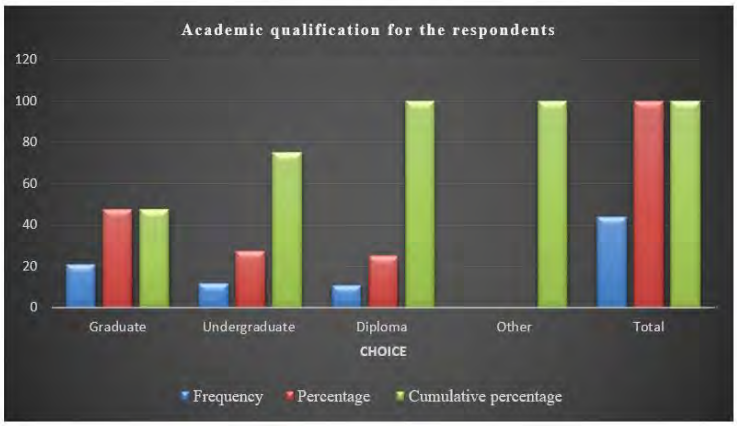
Based on 44 respondents, 6.82% of them indicated to have worked in hospital for over 20years. 9.10 % had worked for between 10 to 20 years. Another 45.45% had worked in hospital between 5 to 10 years and the last 38.62% had worked in hospital a period less than 5 years.

From the study the researcher noted a good number of staff had relative experience at the hospital over 5years.

The second question was to find out the academic qualification for the respondents.

Table 2. Academic qualification for the respondents

Choice	Frequency	Percentage	Cumulative percentage
Graduate	21	47.73	47.73
Undergraduate	12	27.27	75.00
Diploma	11	25.00	100.00
Other			
Total	**44**	**100**	



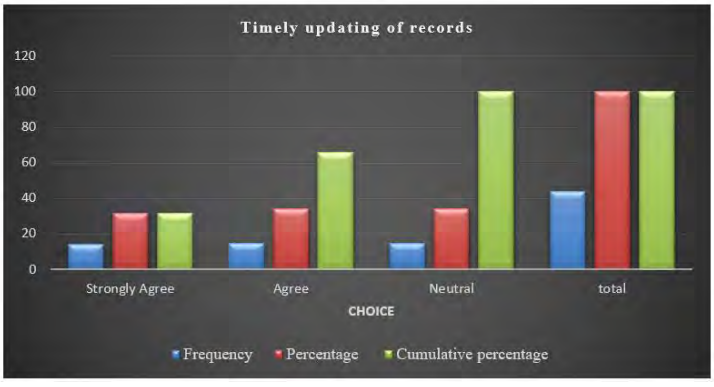
Out of 44 respondents, 47.73% of respondents hold Graduate. 27.27% holds Undergraduate and lastly 25.00% diploma holders.

From the table above, evidently, the vast majority of respondents' i.e., 75.00%, have undertaken education with research component in it and understand the research activities well.

The questions intend was to know the level of Hospital ensuring that all actors add the medical records as soon they are through with the patient to ensure completeness and reliability.

Table 3. Timely addition of medical records in ensuring completeness

Choice	Frequency	Percentage	Cumulative percentage
Strongly Agree	14	31.80	31.80
Agree	15	34.10	65.90
Neutral	15	34.10	100.00
total	**44**	**100.00**	**100.00**



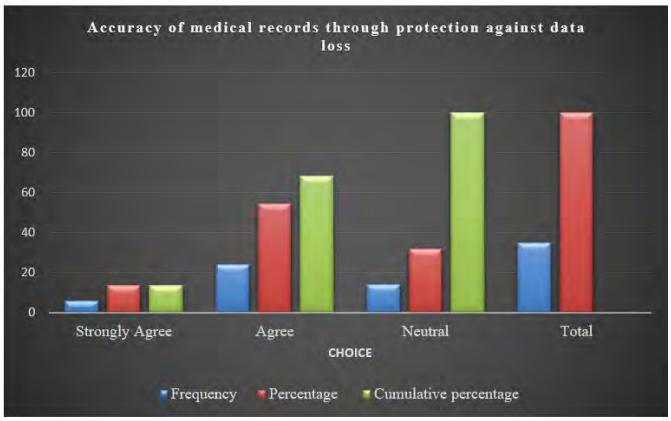
From a total of 44 responses, 31.80% of the respondents strongly agreed that timely addition of medical records in ensuring completeness is observed, 34.10% agreed. A similar 34.10% were neutral.

This indicates that if Electronic Health System Integration Framework for Secure M-Health Service, 65.90% of staff would find it helpful.

Researcher asked respondents if the hospital has ensured accuracy of medical records though protection of information against loss.

Table 4. Response on accuracy of medical records through protection of information against loss

Choice	Frequency	Percentage	Cumulative percentage
Strongly Agree	6	13.64	13.64
Agree	24	54.54	68.18
Neutral	14	31.82	100.00
Total	35	100.00	

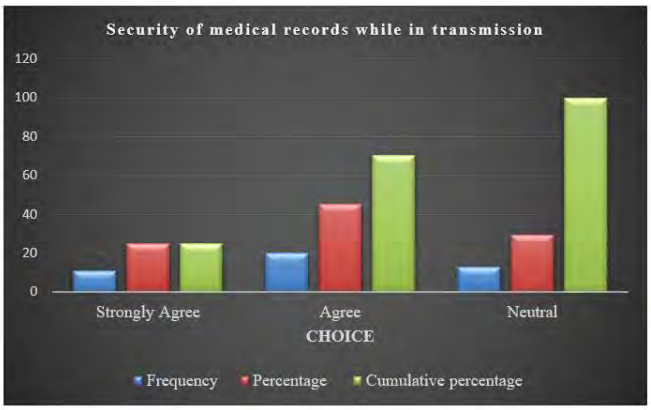


Out of 44 responses, 13.64% of the respondents strongly agreed on accuracy of medical records through protection of information against loss. Another 54.54% agreed on the same. 31.82% of the respondents were not sure. So, we can conclude that protection of information against loss of medical records was key.

The researcher asked the respondents whether the hospital ensured that medical records were protected against distortion while in transmission through electronic media.

Table 5: Security on medical records while in transmission

Choice	Frequency	Percentage	Cumulative percentage
Strongly Agree	11	25.00	25.00
Agree	20	45.45	70.45
Not Sure	13	29.55	100.00

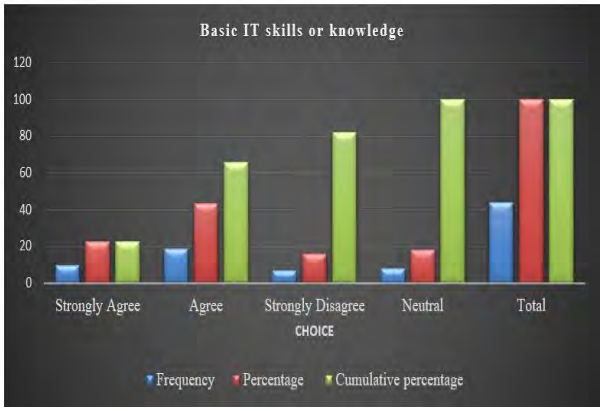


From the above question, 44 responded. Out of which 25.00% strongly agreed presence of security on medical records while on transmission. 45.45% agreed too. 25.81%, while 29.55% are not sure. This means that 70.45% respondents believed in the availability of secure environment in medical records transmission.

The researcher asked respondents for thought about the Hospital in ensuring that employees have basic IT knowledge to key in accurate data.

Table 6. Responses on whether the Hospital ensured its employees have basic IT knowledge to key in accurate data

Choice	Frequency	Percentage	Cumulative percentage
Strongly Agree	10	22.72	22.72
Agree	19	43.20	65.92
Strongly Disagree	7	15.90	81.82
Not Sure	8	18.18	100.00
Total	44	100.00	

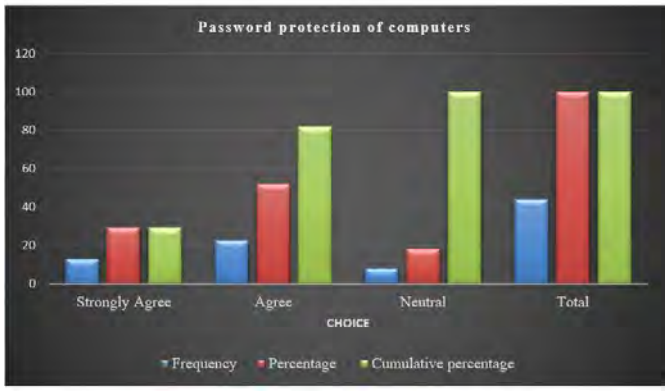


There were 44 responses, 22.372 strongly agreed. 43.20% Agreed, 15.90% Strongly Disagreed, while 18.18% were not sure. Guided by the responses, the researcher came to the conclusion that most of respondents were satisfied that the hospital ensured employees have basic IT knowledge thus using the proposed system was viable.

The researcher asked respondents whether Passwords have been put in computers for protection of data.

Table 7. Protection of data through computer passwords

Choice	Frequency	Percentage	Cumulative percentage
Strongly Agree	13	29.55	29.55
Agree	23	52.27	81.82
Not sure	8	18.18	100.00
Total	44	100.00	

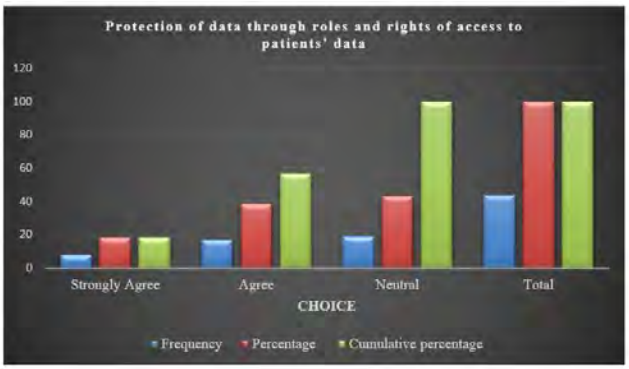


Most respondents agreed with the computer protection that has been put in place.

The researcher asked respondents whether The Hospital has ensured that the system has various users with different roles to avoid unauthorized access of patients' data.

Table 8. Protection of data through roles and rights of access to patients' data

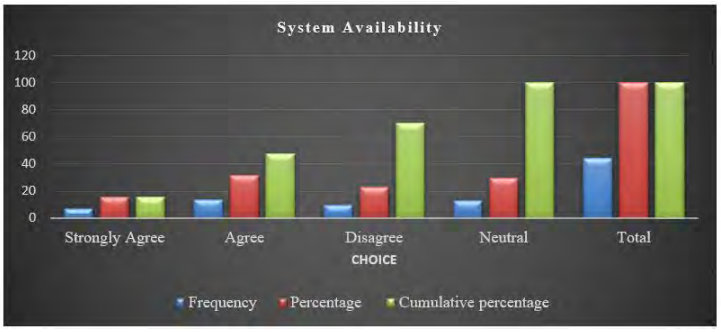
Choice	Frequency	Percentage	Cumulative percentage
Strongly Agree	8	18.18	18.18
Agree	17	38.64	56.82
Not sure	19	43.18	100.00
Total	44	100.00	



The researcher asked respondents if The University Health system is always up and running

Table 9. Availability of system

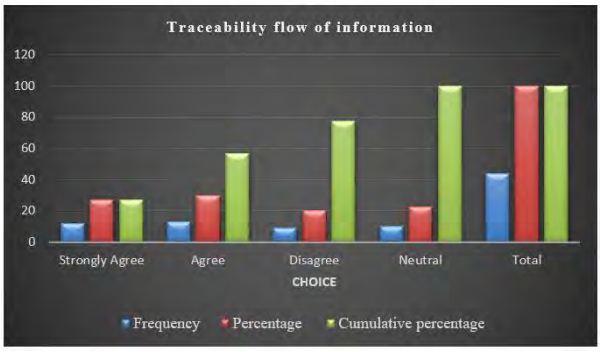
Choice	Frequency	Percentage	Cumulative percentage
Strongly Agree	7	15.91	15.91
Agree	14	31.82	47.73
Disagree	10	22.72	70.45
Not sure	13	29.55	100.00
Total	44	100.00	



The researcher asked respondents whether the flow of information in the University Health system is traceable through logging and documentation

Table 10. Traceability flow of information

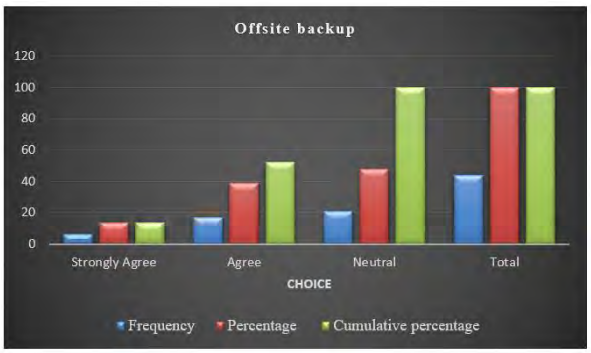
Choice	Frequency	Percentage	Cumulative percentage
Strongly Agree	12	27.27	27.27
Agree	13	29.55	56.82
Disagree	9	20.45	77.27
Not sure	10	22.73	100.00
Total	44	100.00	



The researcher asked respondents if there is an offsite backup of the patient data.

Table 11. Offsite backup availability

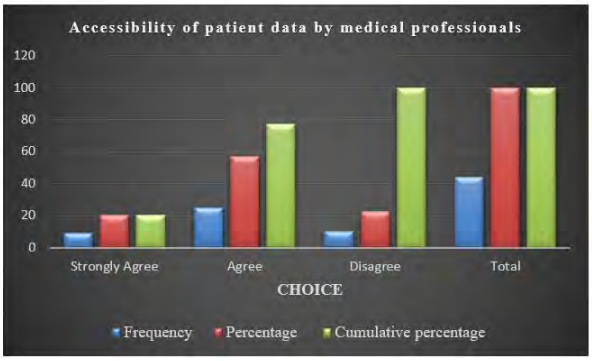
Choice	Frequency	Percentage	Cumulative percentage
Strongly Agree	6	13.64	13.64
Agree	17	38.64	52.28
Not sure	21	47.72	100.00
Total	44	100.00	



The researcher asked respondents if Healthcare professionals have access to patients' information when needed.

Table 12. Accessibility of information by medical professionals

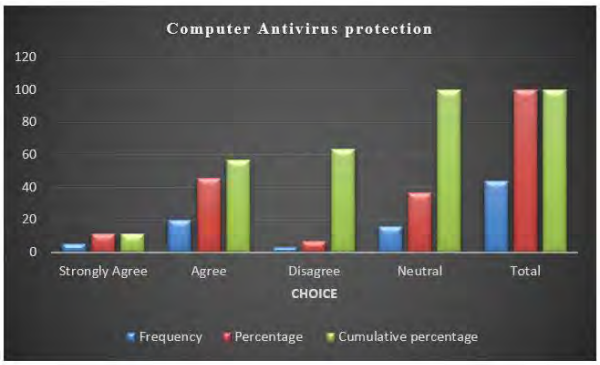
Choice	Frequency	Percentage	Cumulative percentage
Strongly Agree	9	20.45	20.45
Agree	25	56.82	77.27
Disagree	10	22.73	100.00
Total	44	100.00	



The researcher asked respondents if the computer being used has an updated Antivirus

Table 13. Use of Anti-Virus

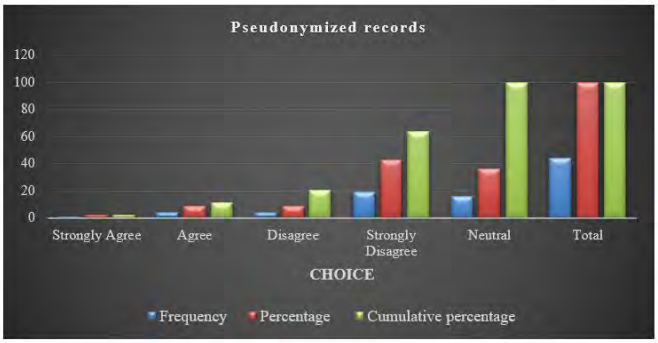
Choice	Frequency	Percentage	Cumulative percentage
Strongly Agree	9	20.45	20.45
Agree	25	56.82	77.27
Disagree	10	22.73	100.00
Total	44	100.00	



The researcher asked respondents if Patient records are always Pseudonymized.

Table 14. Patient records Pseudonymized

Choice	Frequency	Percentage	Cumulative percentage
Strongly Agree	1	2.27	2.27
Agree	4	9.10	11.37
Disagree	4	9.10	20.47
Strongly Disagree	19	43.18	63.65
Neutral	16	36.35	100.00
Total	44	100.00	



5.2 Relationship between Information Security on Secure Service Delivery

5.2.1 Correlation Analysis

In order to examine the level at which the variables under study were related, Karl Pearson's as a measure of coefficient of correlation was used. The Pearson product-moment correlation coefficient denoted as r is primarily used to gauge the level of association between two variables and ranges from +1 to -1. Lack of relationship between variables is denoted by zero value. Where there is a positive relationship, a figure greater than zero is in indicated whereas a negative relationship is indicated by a figure less than zero. The findings are as at:

Table 15. Correlation Analysis

	Secure Service delivery	Information Integrity	Information confidentiality	Information availability
Secure Service delivery(r) (p) Sig. (2 tailed)	1.000			
Information integrity	0.679	1.000		
(p) (2 tailed)	0.009			
Information confidentiality	0.612	0.326	1.000	
(r)	0.013	0.021		
Information availability	0.574	0.254	0.076	1.000
(p) Sig. (2 tailed)	0.026	0.123	0.046	

The results show that information integrity and service delivery have a strong beneficial association (r = .679, P-value < 0.009). This implies that Information integrity influences service delivery in University of Nairobi Hospital. The results also revealed a strong favorable association between information confidentiality and service delivery (r =.612, P-value < 0.013). Hence, suggesting that information confidentiality influences service delivery in University of Nairobi Hospital.

The results showed a strong correlation between information accessibility and service delivery (r = .574, P-value < 0.0426) thus, depicting that information availability influences service delivery in University of Nairobi Hospital.

5.2.2 Regression Analysis

A multiple regression analysis was undertaken to further gauge the association among the

independent variables on service delivery in University of Nairobi Hospital. To aid this, SPSS V 21.0 was used to facilitate the outcomes of the multiple regressions for the study.

The extent to which changes in the dependent variable (Secure service delivery) is influenced by all the three independent variables (Information integrity, information confidentiality, and information availability) is explained by the coefficient of determination.

Table 16. Model Summary

Model	R	R Square	Adjusted Square	R Std. Error of the Estimate
1	.896[a]	.802	.775	0.0131

a. Predictors: (Constant), Information integrity, information confidentiality, and information availability

b. Dependent Variable: Secure Service Delivery

Table 16 shows model summary of regressed variable of the study. The three independent

variables in the study explain 80% effect of level of information security as applied by the

University of Nairobi hospital and how it affects service delivery as represented by R Squared (Coefficient of determinant). This therefore means 20% are other factors not studied in this research that influence secure service delivery.

Table 17. Anova (Analysis of Variance)

Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	6.942	3	2.314	6.51	.001[a]
	Residual	11.005	31	0.355		
	Total	17.947	34			

a. Predictors: (Constant), Information integrity, information confidentiality, and information availability

b. Dependent Variable: Service Delivery

The model summary also shows that the regression model significantly and accurately predicts the dependent variable. The F test shows the statistical significance of the employed regression model. The regression model generally statistically and significantly predicts the outcome variable that is a good fit for the data, as shown by the P=0.001, which is less than 0.05.

Table 18. Coefficient of Determination

	Unstandardized Coefficients			Standardized	Coefficients
	B	Std. Error	Beta	t	Sig.
Constant	7.232	0.451		16.035	0
Information integrity	0.802	0.243	0.126	3.3	0.0011
Information confidentiality	0.769	0.261	0.146	2.946	0.0036
Information availability	0.473	0.213	0.045	2.2	0.0274

The overall equation model for service delivery, Information integrity, information confidentiality, and information availability was as follows:

Ybt = 7.232 + 0.802X1+ 0.769X2+ 0.473X3

According to the model, the service delivery will always be 7.232 if all of the predictor values are zero. The model predicts that service delivery will rise by 0.802 units whenever the value processed through information integrity changes by one unit. In addition, if information confidentiality changes by one unit the service delivery increases by 0.769.

The study's results also showed that service delivery will rise by 0.473 when information availability changes by one unit.

The t-test was used to determine the significance of each variable, which had a 0.05 base value. The result indicates the information confidentiality and information availability have a value of 0.0036 and 0.0274 against the service delivery in the model respectively.

This demonstrates how there is a connection between service delivery, information confidentiality and information availability is significant. The relationship between service delivery and

Information integrity recorded at rate of 0.0011 which is significant since it's less than p- value (P.0.05).

The result indicates the information confidentiality and information availability have a value of 0.0036 and 0.0274 against the service delivery in the model respectively. This shows that the relationship between service delivery, information confidentiality and information availability is significant. The relationship between service delivery and Information integrity recorded at rate of 0.0011 which is significant since it's less than p-value (P.0.05).

Similar to the study findings, Brundtland, (2001) noted that health care delivery has been noted as one of the services deliveries that require high involvement of the consumer in the

consumption process (Peprah, 2014). The customer is involved throughout the entire service delivery process. Information security breaches can result in clinical diagnoses that are incorrect, which can have serious consequences for patients.

5.4 Developed Model

Table 15 in the example below displays the real data, but after the pseudonymization process, the sensitive material is concealed, and the de-identified data is still significant and useful for research.

Table 19. Basic Pseudonymization technique

Healthcare ID	DoB	Name	Medication	Condition
2001567898761234	12/12/2022	Babu Loliondo	Insulin	CD(Conduct Disorder)
2008123456785000	10/05/2022	John Pombe	Dapotum	MH(Malignant hyperthermia)
2001567898761234	10/10/2022	Babu Loliondo	Thalitone	CKD

Table 20. Pseudonymized data (Health care ID & Name)

[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
0102	12/12/2022	A12	Insulin	CD(Conduct Disorder)
452	10/05/2022	B02	Dapotum	MH(Malignant hyperthermia)
2712	10/10/2022	N17	Thalitone	CKD

The data for the concealed connective index is kept on another computer or in a secure location that is inaccessible to regular users.

Table 21. Health ID, healthcare identifier Pseudonym

Healthcare ID	Healthcare Identifier Pseudonym
2001567898761234	0102
2008123456785000	452
2001567898761234	2712

Table 22. Name and name Pseudonym

Name	Name Pseudonym
Babu Loliondo	A12
John Pombe	B02
Babu Loliondo	N17

The distinction between encryption and pseudonymization is that sensitive information and relationships are exposed when encryption or password authorization is used. Pseudonymization, on the other hand, exposes relationships while concealing critical information. Data patterns must be preserved for linking or analysis, and personal data that will be shared—internally or with a partner—must be concealed while being used—are the two key conditions for Pseudonymization.

As a result, risk exposure will be lower, and any possible effects of internal and external security breaches will be lessened. Pseudonymization successfully makes stolen data unusable for identity theft and other types of fraud. By employing de-identified data to identify accounts, process account papers, and record accounts, this makes secure outsourcing and offshore possible. The hospital can save money while greatly decreasing the security concerns associated with hiring third parties.

De-identified data can be used by system integrators, developers, and system administrators for the health software industry to estimate E-Health projects that deal with sensitive health data, design and test new systems that draw on existing operations for sensitive health data, and maintain E-Health systems that manipulate sensitive data.

6. Conclusion and recommendation

6.1 Introduction

This section summarizes the recommendations and conclusions which were arrived at after analysis of the data. It also gives suggestions for further research reference to the general objectives of the study.

6.2 Summary of Findings

The results from this project revealed that over 70% believed introduction of Confidentiality, Integrity and availability on the security of M-health systems would make University Health systems processes fairly convenient.

It was established that a good number, over 60% of staff had relative experience at the hospital over 5years, and at least 75% of these have acquired relevant academic qualification diploma and above in the field of specialty.

On the issue of timely update of medical records, I was noted that if Electronic Health System Integration Framework for Secure M-Health Service, 65.90% of staff would find it helpful.

On the issue of protection of medical records data, the researcher noted that a number of security measures have been put in place for this, including rights and roles in level of access, offsite backup availability, anti-virus installation and IT support team in place with at least 65% basic IT knowledge provisioned.

Patient records have not been Pseudonymized, thus making information gathered to easily allow the individual to be directly identified.

6.3 Conclusion

The study findings indicated the framework applied at University of Nairobi hospital has a gap to address pseudonymized records.

The study also revealed that the introduction of the proposed M-health framework based on service architecture model would improve security levels of the system.

The Electronic health system integration framework for secure m-health services demonstrated significance medical health records secure environment, enabling use of data and or information without infringement to patients.

6.4 Future Research Recommendations

The introduction of The Electronic health system integration framework for secure m-health services in University of Nairobi Hospital is an enhanced solution that would improve in healthcare and associated policies frameworks to be revised and improved for better health services not only in the Hospital but in the country at large.

A more focus should also address on expanding and integrating of similar systems and solutions deployed in other similar and blended environments and thus expand its usage in other hospitals in the country.

Explore and innovate possibilities to increase related services on the framework including simulations on some of the hospital activities that require similar facility in improving secure M-health.

Acknowledgments

We greatly appreciate the valuable contributions of our Mount Kenya University fraternity more so the school of computing and Informatics for offering us learning materials and a conducive environment to do the study. We also thank the University of Nairobi Hospital employees for their immense contribution of participating in this study. We extend our thanks to Lucy Namaswa for her immense contribution in logistics for the whole study.

Authors contributions

Samuel Nandasaba drafted the manuscript while Prof. Gregory Wanyembi and Dr. Geoffrey Mariga revised it. Prof. Gregoy Wanyembi was responsible for coming up with the topic of study as Dr. Geoffrey Mariga assisted in coming up with the objectives of the study. Samuel Nandasaba was responsible for Literature review as Dr. Geoffrey Mariga and Prof. Gregory Wanyembi revised the literature review. Dr. Geoffrey Mariga was responsible with designing of data collection Instruments and in analysis of data collected. Samuel Nandasaba was responsible for data collection. All authors read and approved the final manuscript.

Funding

Not Applicable

Competing interests

The authors of this paper declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Informed consent

Obtained.

Ethics approval

The Publication Ethics Committee of the Canadian Center of Science and Education.

The journal's policies adhere to the Core Practices established by the Committee on Publication Ethics (COPE).

Provenance and peer review

Not commissioned; externally double-blind peer reviewed.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Data sharing statement

No additional data are available.

References

Ahmed, I., & Mousa, A. (2016). Security and Privacy Issues in Ehealthcare Systems: Towards Trusted Services. *International Journal of Advanced Computer Science and Applications, 7*(9), 229-236. https://doi.org/10.14569/IJACSA.2016.070933

Beebe, N. L. V. S. R. (2005). Using Situational Crime Prevention Theory to Explain the Effectiveness of Information Systems Security. Proceedings of the 2005 SoftWars Conference, Las Vegas, 1-18. Las Vegas.

Bendiek, A., & Metzger, T. (2015). Deterrence theory in the cyber-century.Informatik 2015. Bonn:Gesellschaft fur Informatik e.V..(S.553-770).

Bendiek, A., & Metzger, T. (2015). *Deterrence theory in the cyber-century.* Working Paper, Research Division EU/Europe Stiftung Wissenschaft und Politic, German Institute for International and Security Affairs.

Bowman S. (2013). Impact of electronic health record systems on information integrity: quality and safety implications. Perspect. Health Inf. Manag. 10, 1c

Brands, S. (2003). Privacy and Security in Electronic Health. *Security*, 1-12.

Capra, F. (1997). *The web of life.* New York: Doubleday-Anchor Book

Charitoudi, K., & Blyth, A. (2013). A Socio-Technical Approach to Cyber Risk Management and Impact Assessment. *Journal of Information Security, 4*(January), 33-41. https://doi.org/10.4236/jis.2013.41005

David, F. (2006). *Mobile application security system: Bell labs technical journal, 11*(3). https://doi.org/10.1002/bltj.20188

Elkhodr M., Shahrestani S., & Kourouche, K. (2012). A proposal to improve the security of mobile banking applications. InProc. of ICT and Knowledge Engineering, pages 260–265. IEEE. https://doi.org/10.1109/ICTKE.2012.6408565

Gejibo S., Mancini F., & Mughal, K. (2015). Mobile data collection: A security perspective. In: Sasan A, editor. Mobile Health: A Technology Road Map. Switzerland: Springer, Cham; 2015:1015-1042. https://doi.org/10.1007/978-3-319-12817-7_42

Gundu, T., & Flowerday, S. V. (2013). Ignorance to awareness : towards an information security awareness process. 104(August 2012), 69-79. https://doi.org/10.23919/SAIEE.2013.8531867

Health, M. O. F. (2017). *Kenya Standards and Guidelines for mHealth Systems.*

Hong, K. S., Chi, Y. P., Chao, L. R., &Tang, J. H. (2003). An integrated system theory of information security management. *Information Management & Computer Security, 11*(5), 243-248. https://doi.org/10.1108/09685220310500153

Innab, N. (2018). Availability, Accessibility, Privacy and Safety Issues Facing Electronic Medical Records. *International Journal of Security, Privacy and Trust Management, 7*(1), 01-10. https://doi.org/10.5121/ijsptm.2018.7101

Iwaya, L. H., Ahmad, A., & Babar, M. A. (2020). Security and Privacy for mHealth and uHealth Systems: A Systematic Mapping Study. *IEEE Access 2020, 8,* 150081-150112. https://doi.org/10.1109/ACCESS.2020.3015962

Kellerman, A. L., & Spencer, J. S. (2013). What it will take to Achieve The As Yet -unfulfilled promises of Health information technology, Pubmed/23297272.

Leon, N, Schneider, H., & Daviaud, E. (2012). Applying a framework for assessing the health system challenges

to scaling up mHealth in South Africa. BMC Med Inform Decis Mak. 2012, 12: 123-10.1186/1472-6947-12-123. https://doi.org/10.1186/1472-6947-12-123

Liddick, D. (2013). Techniques of Neutralization and Animal Rights Activists. *Deviant Behavior, 34*(8), 618-634. https://doi.org/10.1080/01639625.2012.759048

Lizasoain, A., Tort, L. F., Garcia, M., Gomez, M. M., Leite, J. P., Miagostovich, M. P., Cristina, J., Colina, R., & Victoria, M. (2015). Environmental assessment reveals the presence of MLB-1 human astrovirus in Uruguay. *J. Appl. Microbiol., 119,* 859-867. https://doi.org/10.1111/jam.12856

Lucas J. (2013). *Oracle,an Introduction to the basics of data integrity enforcement in a variety of environments.* Amis technology blog.

Maranda, A., & Majchrzycka, A. (2016). Secure development model for mobile applications. *Bulletin of the polish academy of sciences technical sciences 64*(3). https://doi.org/10.1515/bpasts-2016-0055

Ministry of health. (2016). Kenya national ehealth policy. 2016-2030.

Nganji, J. T., & Nggada, S. H. (2011). Disability-aware software engineering for improved system accessibility and usability. *International Journal of Software Engineering and Its Applications, 5*(3), 47-62.

Nkosi, M., & Mekuria, F. (2010). Cloud computing for Enhanced mobile health applications: IEEE. https://doi.org/10.1109/CloudCom.2010.31

Pernebekova, A. P., & Ahbergenovich, B. A. (2015). Information Security and the Theory of Unfaithful Information. *Journal of Information Security, 06*(04), 265-272. https://doi.org/10.4236/jis.2015.64026

Salim, H., & Salim, H. M. (2014). A Systems Thinking and Systems Theory Approach Cyber Safety : A Systems Thinking and Systems Theory Approach to Managing Cyber Security Risks By. (May).

Schein, R., Wilson, K., & Keelan, J. (2010). Literature review on effectivess of the use of social media: A report for Peel Public Health. *Challenges, 129*(1), 63. Retrieved from http://www.peelregion.ca/health/resources/pdf/socialmedia.pdf

Serhani, M. A., Benharref, A., & Nujum, A. R. (2014). Intelligent remote health monitoring using evident-based DSS for automated assistance, 2674-7. https://doi.org/10.1109/EMBC.2014.6944173

Shifali, A., Yttri, J., &Nilsen, W. (2014). *Privacy and security in mobile health(Mhealth) research.*

Simplicio, M. A., Iwaya, L. H., Barros B. M., Carvalho, T. C. M. B., & Naslund, M. (2015). Secourhealth: A delaytolerant security framework for mobile health data collection. *Biomedical and Health Informatics, IEEE Journal of, 19*(2), 761-772. https://doi.org/10.1109/JBHI.2014.2320444

Tarus, J. K., Gichoya, D., & Muumbo, A. (2015). Challenges of implementing e-learning in Kenya: A case of Kenyan Public universities. *The international review of research in open and distributed learning, 16*(1). https://doi.org/10.19173/irrodl.v16i1.1816

University of Nairobi. (2021,January 23). Fact file. https://uonbi.ac.ke/fact-file

Vimalachandran, P., Wang, H., Zhang, Y., & Whittaker, F. (2018). Ensuring Data integrity in electronic health records: A quality Health care implication.

W. Li, L. Cheng, (2013). *Effects of neutralization techniques and rational choice theory on Internet abuse in the workplace.* Pacific Asia Conference on Information Systems, Jeju Island, Korea.

Wallis, L., Blessing, P., Dalwai, M., & Shin, S. (2017). Integrating mHealth at point of care in low- and middle-income settings : the system perspective. *Global Health Action, 10*(00). https://doi.org/10.1080/16549716.2017.1327686

Wausi, A. N., & Waema, T. M. (2009). *Organizational implementation of Information systems innovations* (Unpublished doctoral thesis). University of Nairobi, Kenya.

WHO. (2018). *Use of appropriate digital technologies for public health* (Vol. 28).

Computer and Information Science; Vol. 16, No. 4; 2023
ISSN 1913-8989 E-ISSN 1913-8997
Published by Canadian Center of Science and Education

Malware Investigation and Analysis for Cyber Threat Intelligence: *A Case Study of Flubot Malware*

Uchenna J. Nzenwata[1], Frank Uchendu[2], Haruna Ismail[3], Eluwa M. Jumoke[2], & Himikaiye O. Johnson[1]

[1] School of Computing and Engineering Sciences, Babcock University, Ogun State, Nigeria

[2] School of Computing and Information Security Studies, Salford University, Manchester, UK

[3]Silesian University of Technology, Gliwice, Poland, Europe

Correspondence: Uchenna J. Nzenwata, School of Computing and Engineering Sciences, Babcock University, Ogun State, Nigeria.

Received: September 19, 2023 Accepted: November 10, 2023 Online Published: November 29, 2023

doi:10.5539/cis.v16n4p47 URL: https://doi.org/10.5539/cis.v16n4p47

Abstract

Android operating systems have swiftly outpaced other operating systems (OS) in popularity, making them vulnerable to assaults since hackers are continuously looking for flaws to exploit. This is why several organisations have long been plagued by various types of mobile security threats. Utilizing a cyber-threat intelligence tool to evaluate, track, and prevent planned attacks is one crucial strategy to combat this effect. This paper discusses and investigates the FluBot malware, using the Dagah tool and Android Studio to phish, harvest and exploit malicious applications over SMS on Android devices. The Capability Maturity Model (CMM) was adopted and used for the investigation. The methodology adopted describes the operation of the FluBot malware through a cloned website, and demonstrates how FluBot is used to share a malicious link through the short message service (SMS), which is then used to grab a victim‟s credentials. The outcome of the study displayed the information on the FluBot malware, including its source, domain, and destination. Similar malware analysis and assessments of cyber threat intelligence may be conducted using the techniques used in this study.

Keywords: low case, comma, paper template, abstract, keywords, introduction

1. Introduction

Advanced Persistent Threats (APTs) are becoming more frequent in today's world and it is getting harder to secure wireless networks and private files as hackers always come up with new ways to steal data. Since Android is the most used smartphone operating system worldwide, it is the mobile operating system that gets targeted the most (Garg & Baliyan, 2021). Short Message Services (SMS) are delivered to iPhones and Android smartphones by a malicious program named Flubot (Salsabila, Mardhiyah & Hadiprakoso, 2022). The Flubot SMS messages come in a broad range of formats, and scammers regularly change them, (Blázquez & Tapiador, 2023). To find out what kind of malware the Flubot is, questions are being posed. Chapin, Piscitello and Strutt (2022), found that Flubot is an Android malware actively spreading over SMS, collecting passwords, online banking information, and other sensitive information from affected smartphones world-wide. The primary mode of Flubot dissemination, according to the study in (Van Haastrecht et al., 2021), is text message notifications. These notifications urge consumers to download a security update or an app. The program asks for various permissions (such as SMS, call, contact permissions, and many more) during installation that essentially gives it power over the device.

1.1 Android OS

The Android Operating System (OS) was created by Andy Rublin, Rich Miner, Nick Sears, and Chris White for Android Inc. in October 2003 (Callaham, 2018). The OS, based on Linux-Kernel and created for smartphones and tablets, is open source and source code released by Google under an Apache licence. Its system architecture is made up of four main layers which includes: Application Layer, Framework Layer, Middleware Layer and Kernel Core layer (Meng et al., 2018). Each of the layers contains the specifics as illustrated in Figure 1.

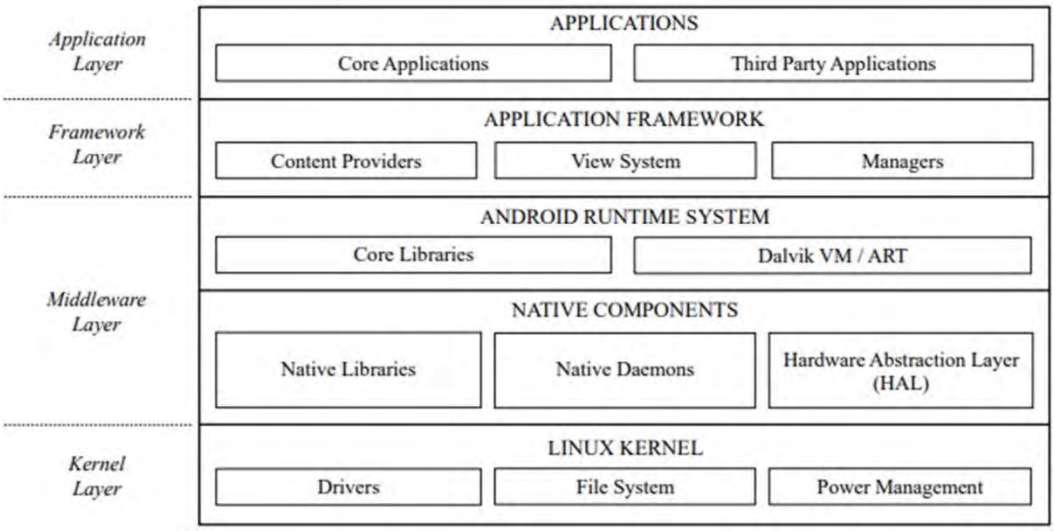


Figure 1. A Layered System Architecture of the Android OS (Meng et al., 2018)

Android quickly surpassed other operating systems in popularity, rendering it vulnerable to attacks because hackers are constantly on the lookout for weaknesses to exploit. The fact that several suppliers offer services that are marketed without well-established security measures makes defending the Android OS the most challenging issue.

1.1.1 Vulnerabilities And Security Issues Associated with The Android OS

Some of the known vulnerabilities and security issues associated with the Android OS includes the following as stated in (Özdemir and Zaim, 2021), which is summarized in Figure 2.

i. Denial of Service (DOS): This prohibits users from accessing the target system and prevents the target system from offering services.

ii. Code Execution: This happens when an attacker inserts malicious code into a string or file that is then used by the software to perform its operations.

iii. Overflow: Sequential data of the int and char types are stored in memory by buffers. When the variables of a program made up of flawed functions store more data than they can hold, it results in a buffer overflow.

iv. Gain Information: This happens when useful data about the target system is obtained during the attack phase and made easily accessible if it is in the public domain. The majority of it is completed using a tool for information gathering.

v. Gain Privilege: This is the process where the attacker searches for vulnerabilities discovered while gaining information and then exploits those vulnerabilities to get user rights.

Attack type/ using vulnerability	Android Vulnerabilities					
Attack types	DOS	Code Execution	Overflow	Bypass	Gain information	Gain privileges
Remote Attacks	✓				✓	✓
Client-Side Attacks		✓			✓	✓
Attacks Using Malicious Apps	✓	✓	✓	✓	✓	✓
Mobile Post Exploits	✓	✓	✓	✓	✓	✓

Figure 2. Android vulnerabilities/security issues with its attack type (Özdemir and Zaim, 2021)

Other related security concerns of the Android OS include: Version fragmentation, Rooting, Google Play malware, insecure apps, lack of hardware data encryption, spyware, data leaks and SMShing.

1.2 Flubot

Flubot is thought to have originated from Spain and was first discovered in December 2020 as shown in Figure 3. (Threatfabric, n.d). A report by a cybersecurity firm ThreatFabric, claims the malware is disseminated through phishing assaults, in which attackers send messages (smishing) to potential victims that contain dangerous links (Threatfabric, n.d). Clicking this link compromises the device thus, grabbing the credentials and other personal identifiable information (PII) of the victim by the attacker. Flubot's agents have a variety of motives, including monetary gains, development of botnets, undercover activities, information gathering and social engineering. There have been a number of fraudulent Short Message Service (SMS) campaigns between the end of 2020 and the beginning of 2021 that announced the arrival of a package while posing as different logistics companies, such as FedEx, DHL, or Correos. Recipients were invited to download an app on their mobile device in order to find out where the package is (Liu et al., 2021). In terms of the malicious code's functionality, once the user installs the application on their device, it begins to track the identifiers of all the applications it starts and is capable of injecting superimposed pages when it detects a session log-in in one of the target applications, so the user believes they are entering their credentials on the original website when, in reality, they are sending them to the command-and-control server (C2) controlled by the attacker. To avoid detection and analysis, the malware employs code injection, code obfuscation, and encryption. It poses a serious threat to Android users since it can spread to other devices via SMS messaging (Mayrhofer et al., 2021). Figure 4, shows the FluBot propagation pattern.

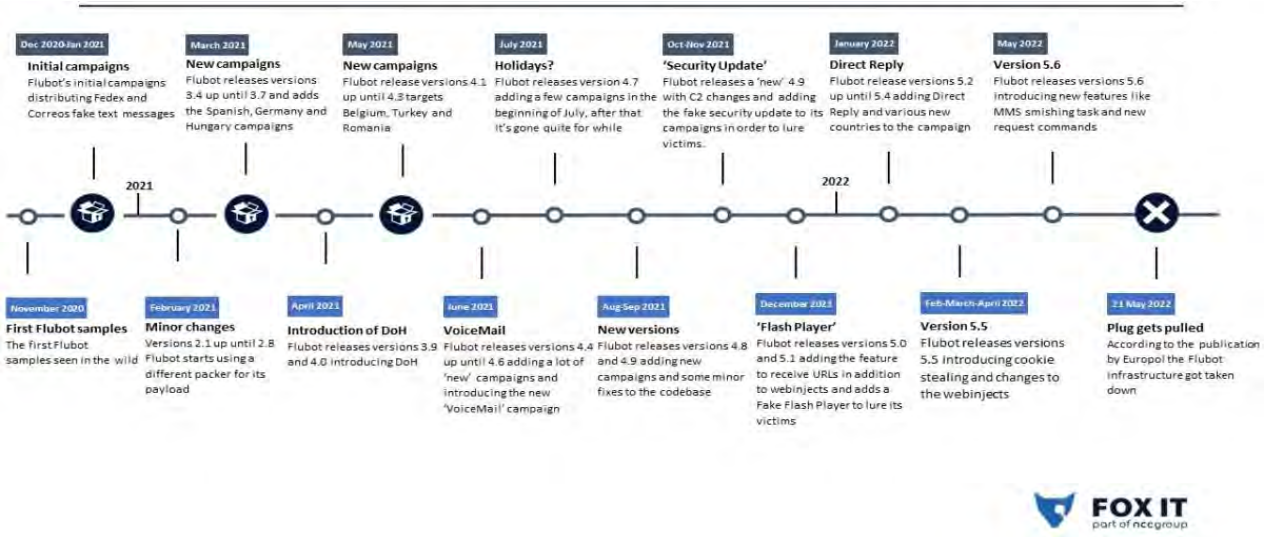


Figure 3. Evolution of the FluBot Malware (Fernick, 2022)

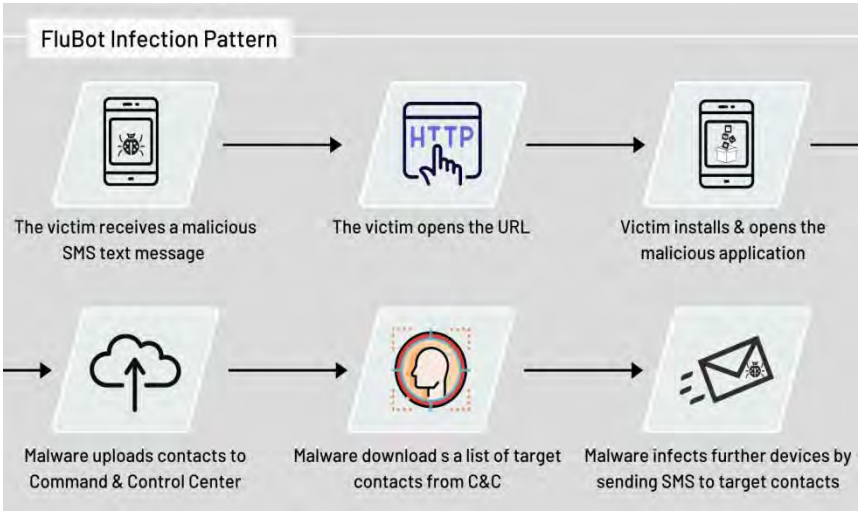


Figure 4. FluBot infection and propagation pattern (Gibbs, 2021)

2. Literature Review

FluBot has been the subject of some studies, including malware comprehension and analysis. A study was carried out in (García-Teodoro, Gómez-Hernández and Abellán-Galera, 2022), where three unknown malware samples were analysed. It was identified from the codes that these samples are FluBot malware. From the study in (García-Teodoro et al., 2022) and shown in Figure 5, FluBot was also referred as Fedex Banker or Cabassous. According to the studies done by the Swiss company PRODAFT (Mogicat & Zermin, n.d), it is possible that FluBot infected over 60,000 terminals and listed over eleven million phone numbers, which is equal to 25% of the population of Spain.

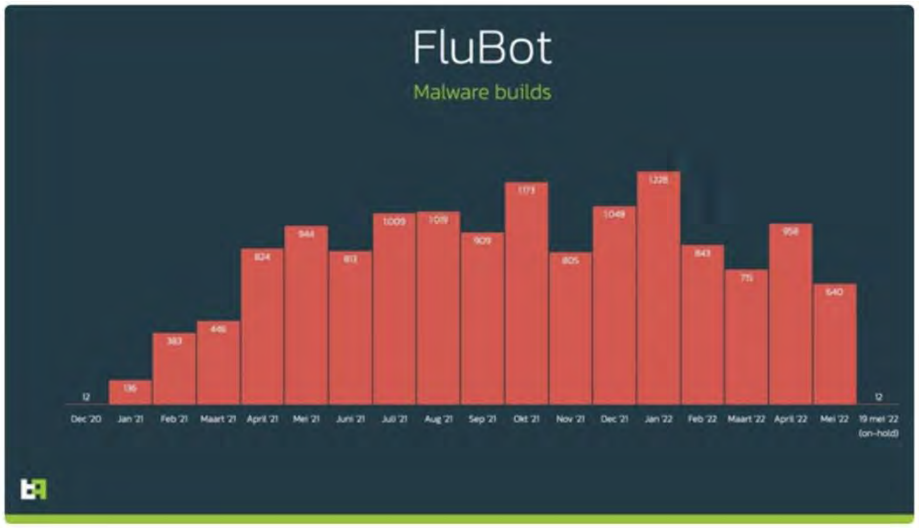


Figure 5. Graphical representation of FluBot malware spread from December 2020 to May 2022 (Threatfabric, n.d).

Android is now the most popular mobile operating system, accounting for 43.43% of the market (Riasat, Batool and Iqbal, 2022) and 70.93% of the global market share worldwide as at March 2023 by (StatCounter, 2022).

The ability to simply build and submit programmes to the official store (Google Play) not only attracts developers, but it also boosts the number of new users of this platform. Because of its popularity and market dominance, Android is frequently attacked by rogue applications. While Google claims to have eliminated up to 1.2 million dangerous apps security experts and threat intelligence firms continue to discover malicious malware disguised as legitimate programs.

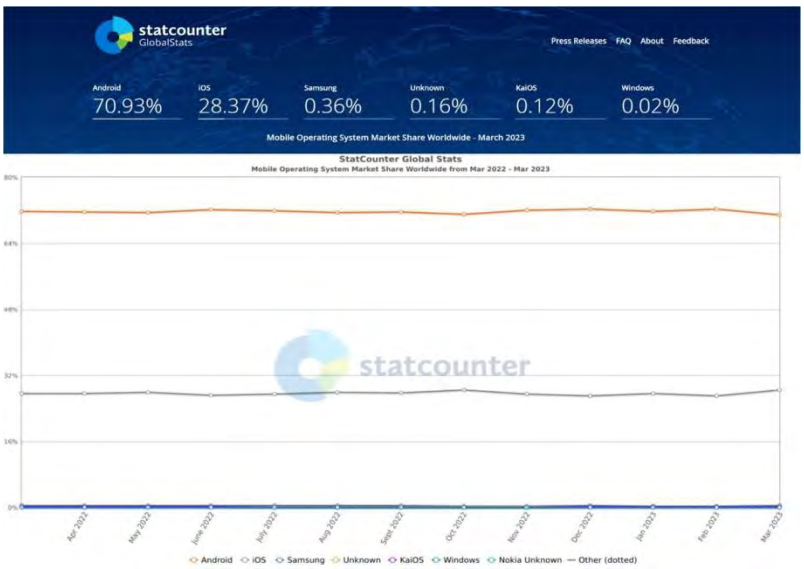


Figure 6. Graphical representation of Mobile OS Market Share Worldwide from April 2022 to March 2023 (StatCounter, 2022).

2.1 Scope and Limitation

The investigation of a mobile security threat at ABC organisation using Dagah for exploitation and Android Studio for Android device simulation, as well as carrying out a threat intelligence assessment to protect data leakage, secure wireless network communication, malware, and malicious programme propagation, is the focus of this paper. The study does not consider all types of mobile security threats; instead, it concentrates on a specific Trojan for Android devices named FluBot. The two observed limitations of this study are the usage of an android emulator in place of an actual android smartphone and Bitly's refusal to shorten URLs even after we successfully generated our access token.

2.2 Related Tools

The present ecosystem of Android tools contains various frameworks aside the Dagah tool that are intended to carry out further specialised analytic tasks. The DroidBox (Chaurasia, 2015) is used to perform dynamic analysis of Android. Another tool is the ConDroid (Schütte, Fedler and Titze, 2015), which is used execute specific code locations with no app manual interaction. For the Network analysis, the Wireshark (Ndatinya et al., 2015) is a good dynamic tool.

3. Methodology

The model considered and implemented for this investigation is the Capability Maturity Model (CMM). There are two levels of CMM and its implementation to this investigation: Threat intelligence collection capability and threat intelligence integration and dissemination. The CMM was used in this study because it has a well-defined and efficient processes, which are crucial for detecting, analyzing, and mitigating threats effectively.is a good tool for malware analysis.

3.1 Level 1: Threat Intelligence Collection Capability

This is the first phase of the model where requisite data and Indicators of Compromise (IOC) are gathered and filtered by the tactical intelligence team for threat intelligence operations. The following elements are the indicators of compromise identified as illustrated in Table 1.

Table 1. Indicators of Compromise (IOF)

Indicators of Compromise	Details
Name	FluBot
Attack family	Malware (banking Trojans)
Type of attack	Mobile Malware Attack
Target OS	Android
Country of origin	Spain
Attack Vector	SMS messaging
Resource materials	URLs, journals and books
Year of inception before propagation	2020
Risk and impact	Critical

Accordingly, in order to ascertain and understand critical information, attack and motives of the FluBot malware, Alien Vault was considered and used.

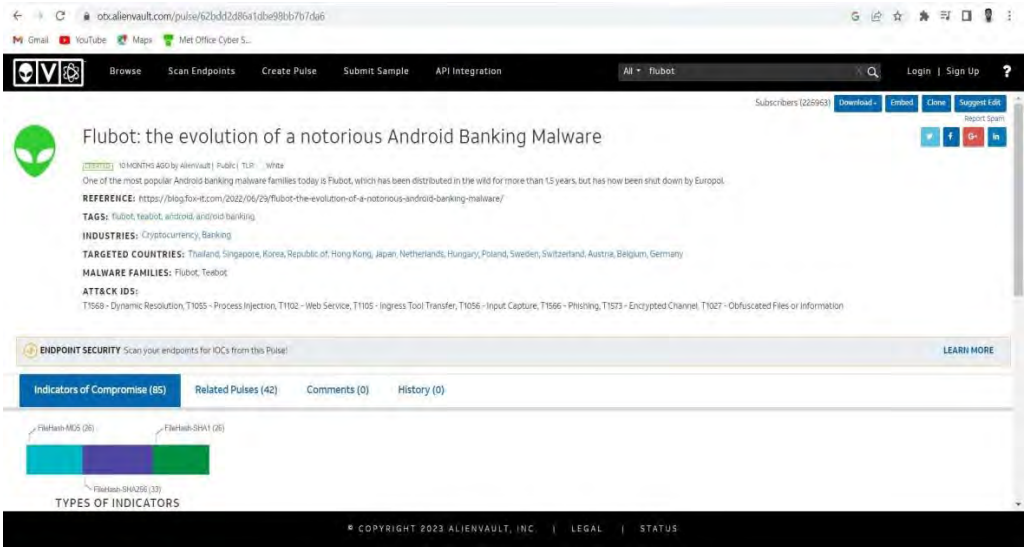


Figure 7. Research of critical information about FluBot using Alien Vault (Exchange, 2020)

3.2 Level 2: Threat Intelligence Integration and Dissemination

This is the second phase of the model where actions are taken based on the identified data or indicators of compromise collected from Level 1 to respond to the attack/threat (FluBot).

3.3 Investigation and Analysis

In investigating the FluBot Android Malware that has been a major global mobile security concern, the 5 steps of OPSEC were also considered. This includes: identification of critical information about the APT, FluBot; analysis of the APT, analysis of possible vulnerabilities; risk assessment; and use of applicable countermeasures. A static analysis was conducted on VirusTotal to generate basic metadata about FluBot.

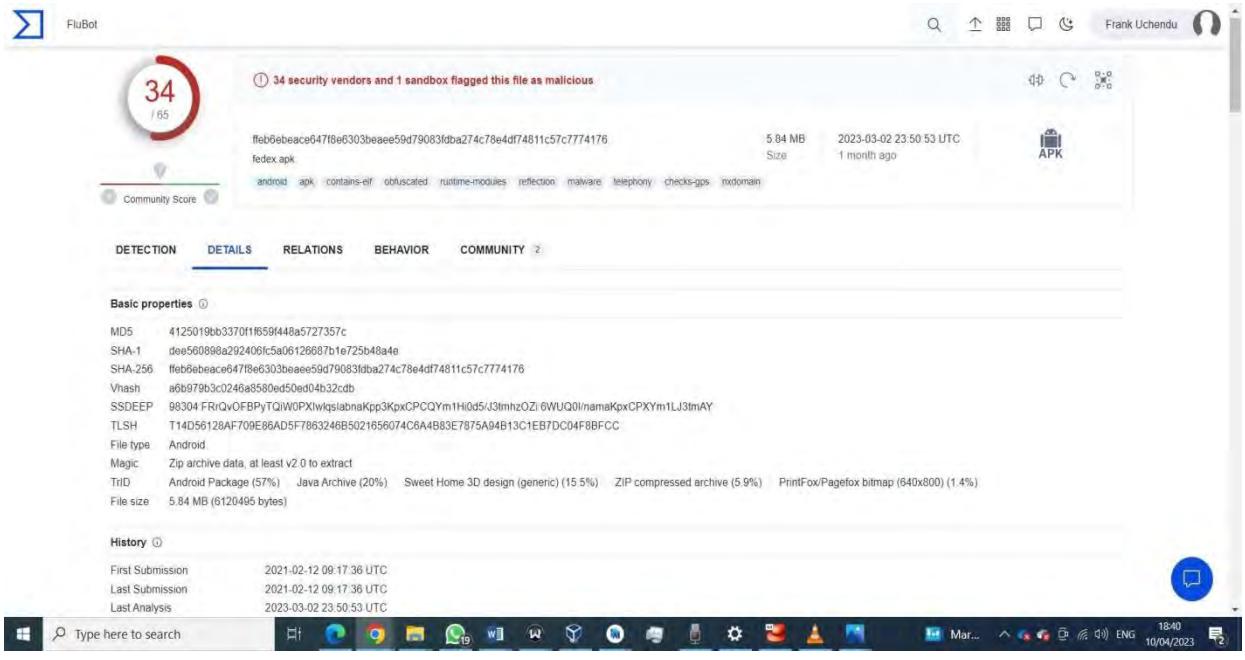


Figure 8. Basic metadata of the FluBot Android Malware (VirusTotal, n.d)

Accordingly, in order to analyze the malware dynamically, we must first understand its tactics and indicators of compromise (IP address, bad domain, OS) through the Capability Maturity Model (CMM) phases of threat intelligence analysis.

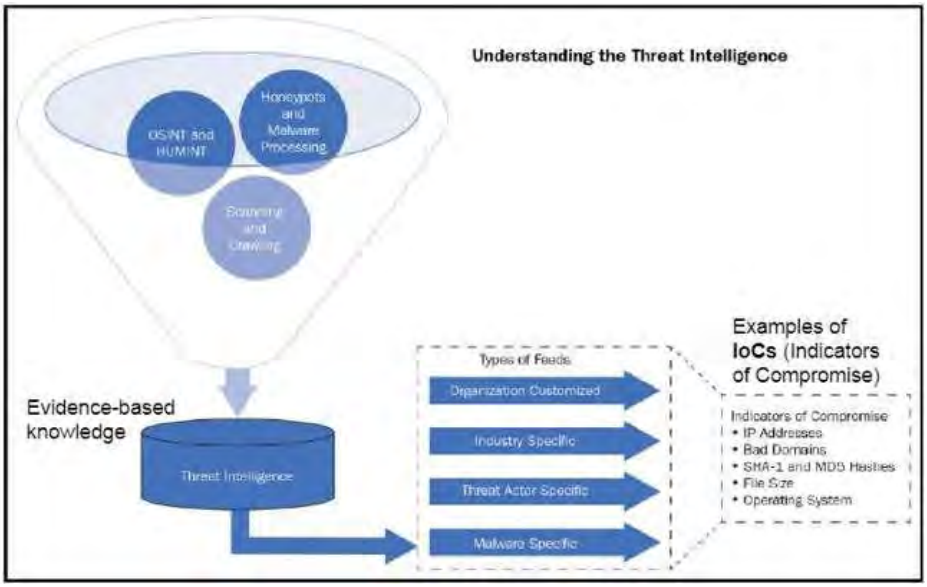


Figure 9. Depicts the Threat Intelligence Maturity Model (Abouzakhar, 2023)

Moving forward, the tools used to develop the security scenario (FluBot) in this investigation were downloaded, set up and configured respectively. They include: Dagah**,** which was installed on a virtual environment (Virtual Box) for designing and launching of attacks against Android Emulator**,** which are the simulated targets.

3.3.1 Dagah Environment Set-up

```
        ether 08:00:27:76:73:6d  txqueuelen 1000  (Ethernet)
        RX packets 266  bytes 86535 (84.5 KiB)
        RX errors 0  dropped 0  overruns 0  frame 0
        TX packets 240  bytes 26911 (26.2 KiB)
        TX errors 0  dropped 0 overruns 0  carrier 0  collisions 0

enp0s8: flags=4163<UP,BROADCAST,RUNNING,MULTICAST>  mtu 1500
        inet 192.168.56.105  netmask 255.255.255.0  broadcast 192.168.56.255
        inet6 fe80::a00:27ff:fed1:e8bb  prefixlen 64  scopeid 0x20<link>
        ether 08:00:27:d1:e8:bb  txqueuelen 1000  (Ethernet)
        RX packets 5093  bytes 826256 (806.8 KiB)
        RX errors 0  dropped 0  overruns 0  frame 0
        TX packets 6311  bytes 8106109 (7.7 MiB)
        TX errors 0  dropped 0 overruns 0  carrier 0  collisions 0

lo: flags=73<UP,LOOPBACK,RUNNING>  mtu 65536
        inet 127.0.0.1  netmask 255.0.0.0
        inet6 ::1  prefixlen 128  scopeid 0x10<host>
        loop  txqueuelen 0  (Local Loopback)
        RX packets 56  bytes 5684 (5.5 KiB)
        RX errors 0  dropped 0  overruns 0  frame 0
        TX packets 56  bytes 5684 (5.5 KiB)
        TX errors 0  dropped 0 overruns 0  carrier 0  collisions 0

[dagah@localhost ~]$
```

Figure 10. The IP address to use it to log on the Dagah Web interface via HTTP

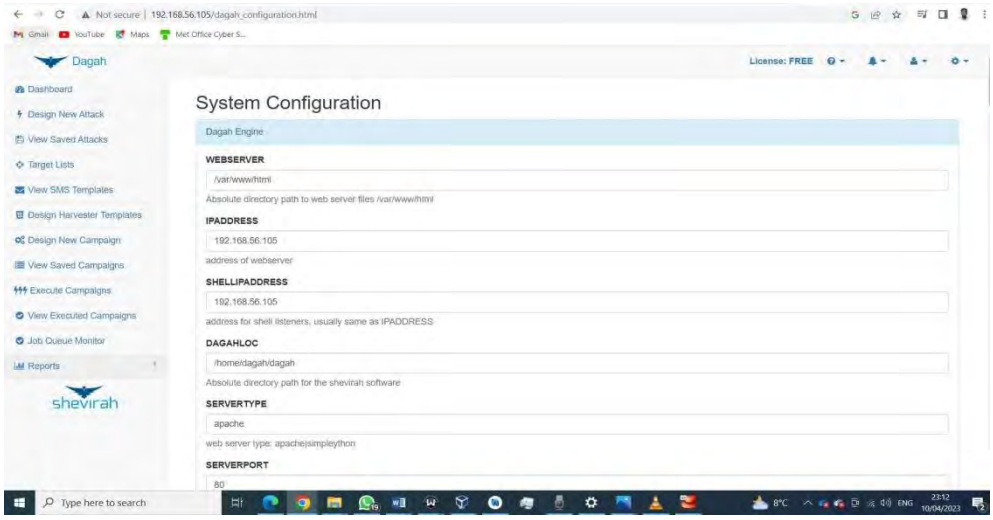


Figure 11. Configuration of the system by setting the IP address to 192.168.56.105

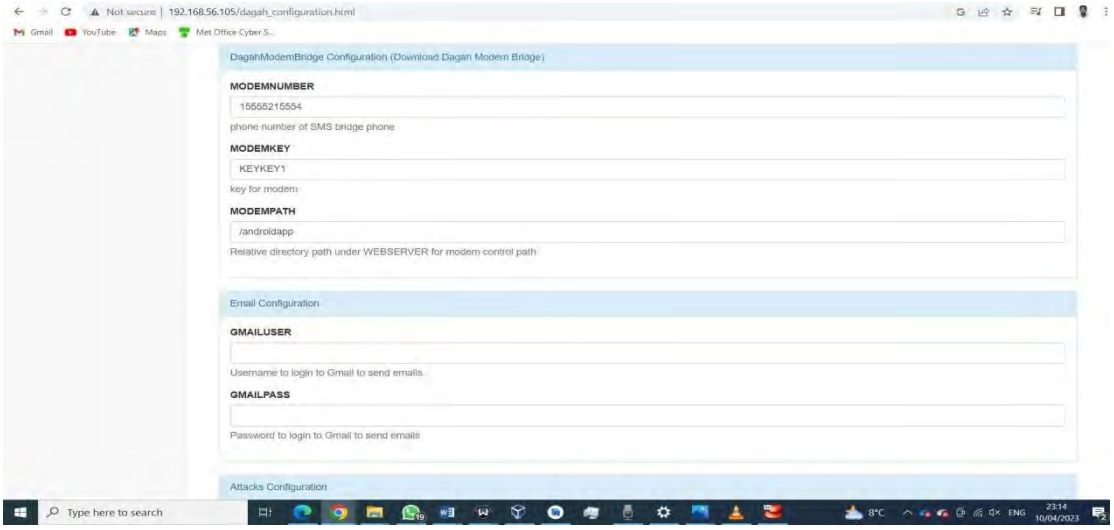


Figure 12. Configuration of the Dagah modem bridge on the GUI

we downloaded and installed the Dagah modem bridge App on the Attack Android simulator. The Dagah Modem App ensures that SMS attacks are sent from the Dagah GUI to the android device

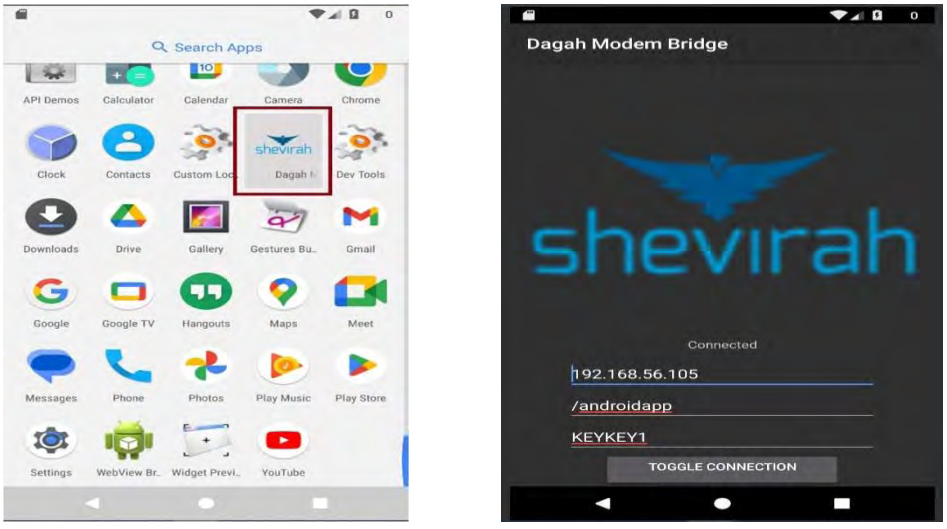


Figure 13. App on the Attackers" device, setting of the IP address, path and key on the Dagah Modem Bridge

3.3.2 Android Studio Setup

Furthermore, we installed the Android studio and created two Android Nexus 5X virtual devices for emulation of the operating system, as shown in Figure 14.

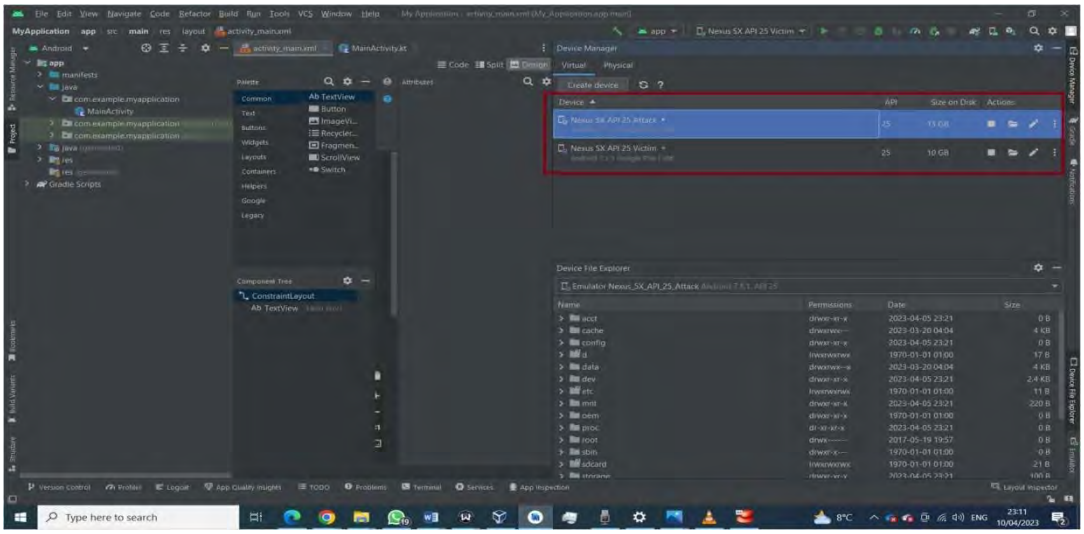


Figure 14. Installation of Android studio and creation of the devices

To ascertain that the respective devices are functional and connected, we dialed the Victim"s device using that of the Attacker"s as shown in Figure 15.

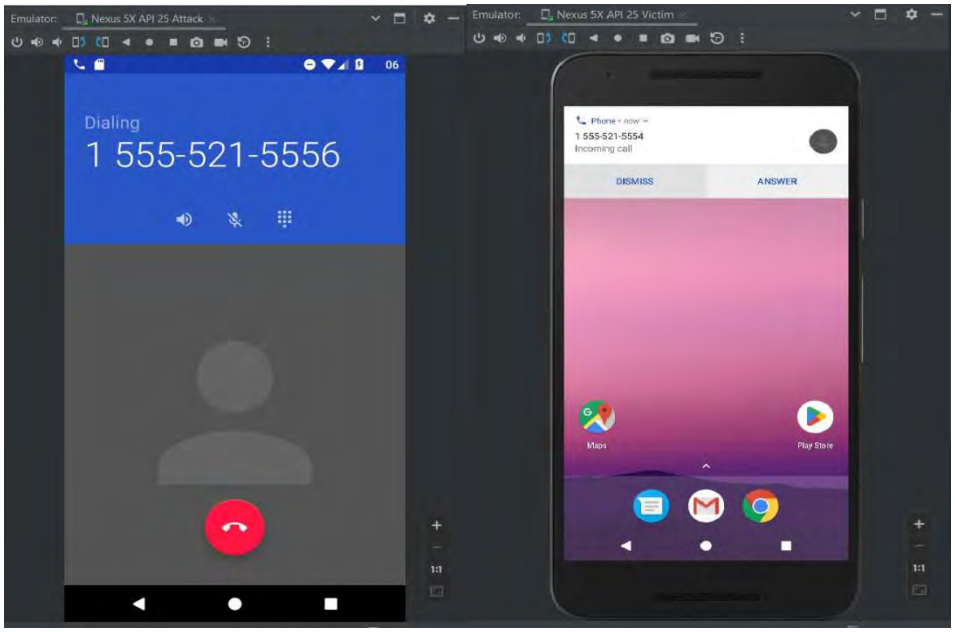


Figure 15. The Attacker"s device calling the Victim"s

3.3.3 Practical Experiment

FluBot is distributed through phishing attacks using SMS as its mode of transmission. It propagates by harvesting users" credentials through deceptive links and for this experiment, we conducted two harvester phishing types of attack:

i. An email phishing using the built-in Gmail template on Dagah GUI to harvest the users" Gmail log-in credentials; and

ii. we designed a harvester template by cloning and editing a website (https://gradintel.com), then harvested the credentials submitted to the website.

For both attacks, the victims‟ Android simulating mobile device receives an SMS from the attacker‟s device. When the victim clicks on the malicious link and inputs his credentials, the attacker grabs and stores the credentials in the campaign results of the Dagah GUI, as shown in Figure 16.

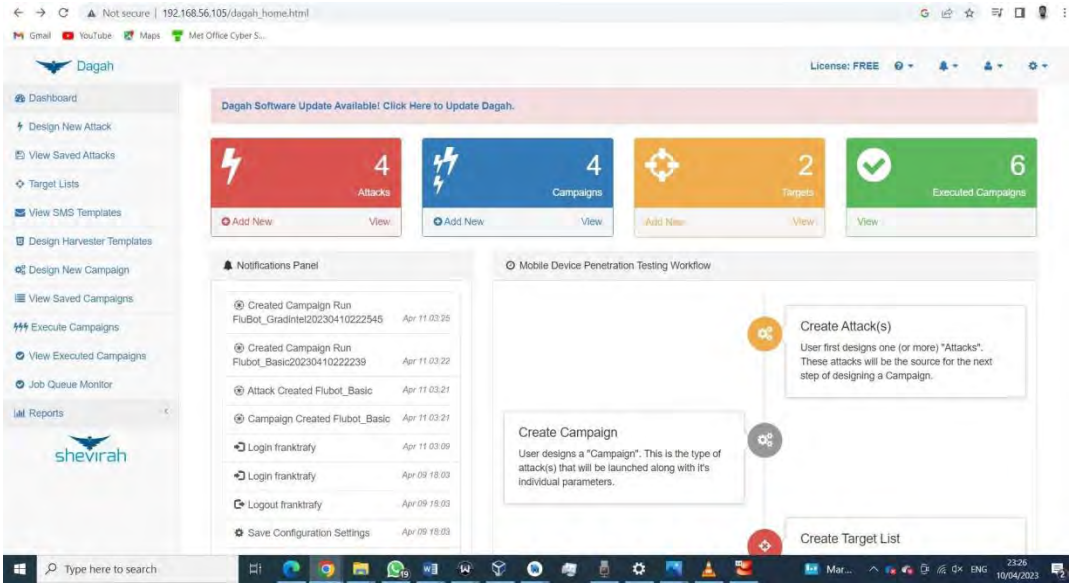


Figure 16. The Dagah GUI dashboard showing designed attacks, campaigns, target lists and the executed campaigns

For the built-in Gmail template harvester phishing attack, we designed a new attack with the harvester type of attack selected, delivery method set to SMS and harvester template (gmail.com) selected, created our target list with the Android Victim‟s phone number and designed a campaign before executing.

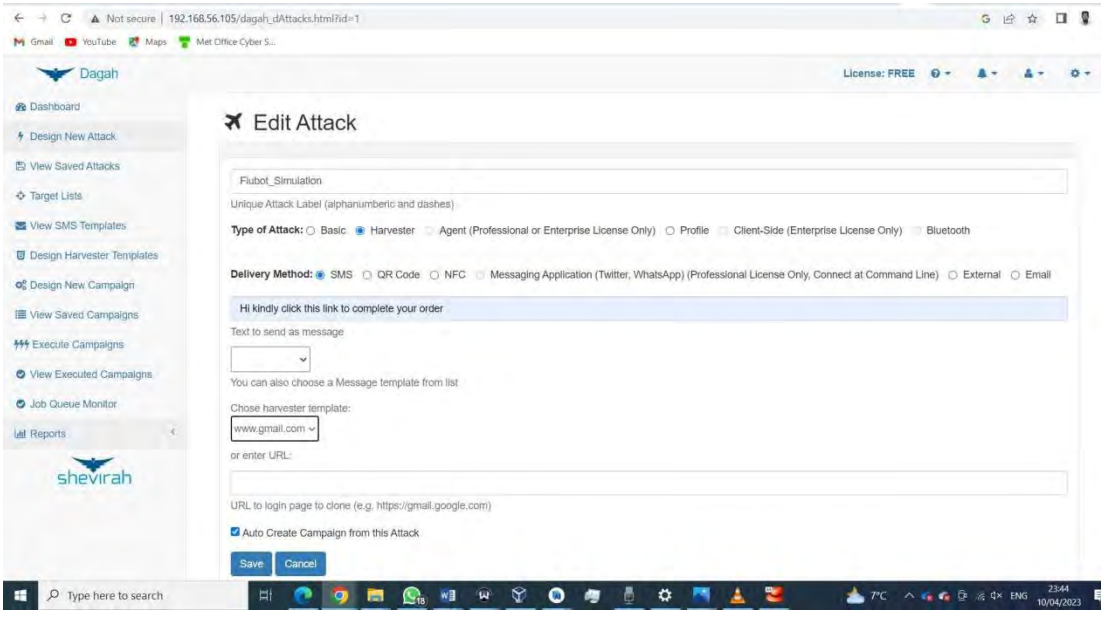


Figure 17. Designing a new Gmail harvester phishing attack on Dagah GUI

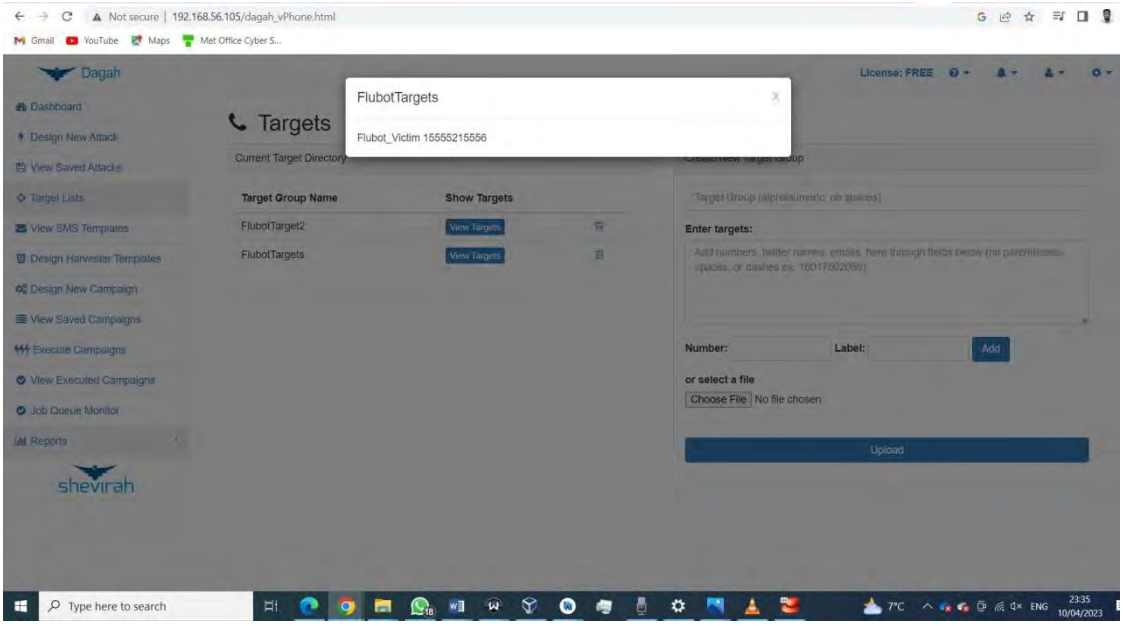


Figure 18. Target list created with the Victim"s number

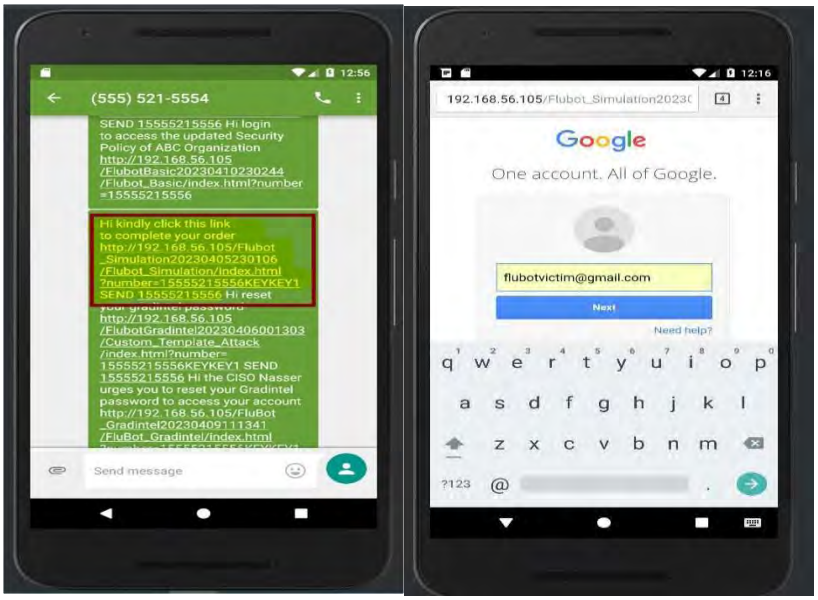


Figure 19. Depicts the victim"s device receiving the malicious link and opening the phony Gmail login page

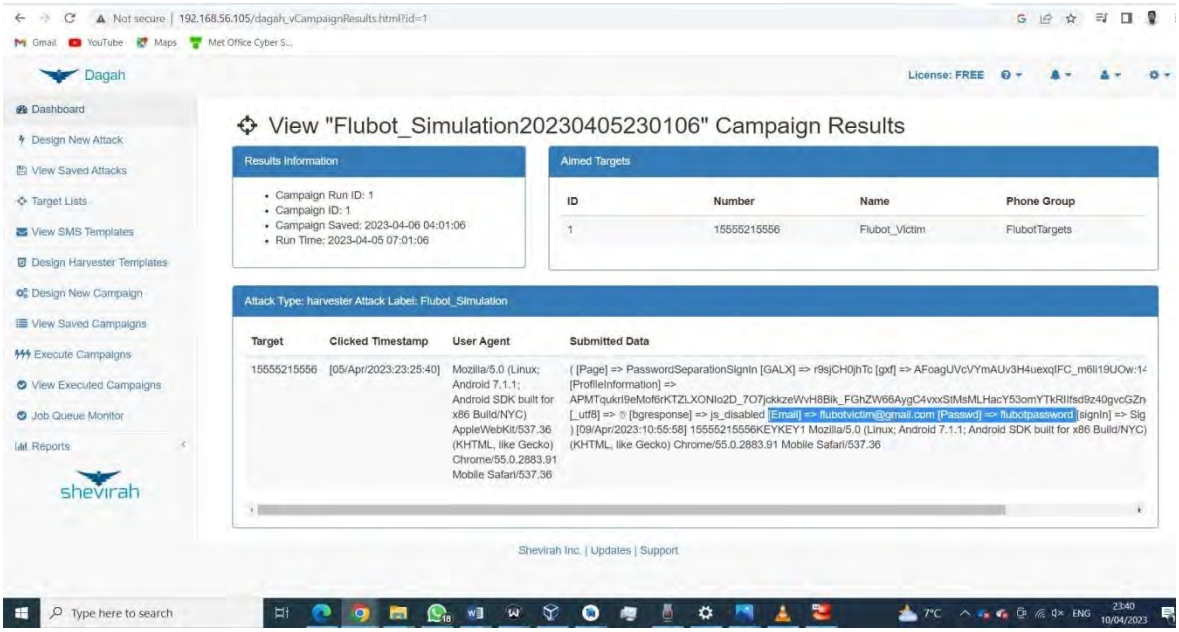


Figure 20. Results of the executed campaign showing the grabbed credentials of the Victim

For the cloned and edited Website template harvester phishing attack, we clicked on the Design Harvester Template on the Dagah GUI, clicked on Add new Harvester Template, entered the new template name (gradintel.com) and saved it. Then we removed the client-side validation script and updated the form method"s POST action in line 252 of HTML code to „./post.php", as shown in Figure 21.

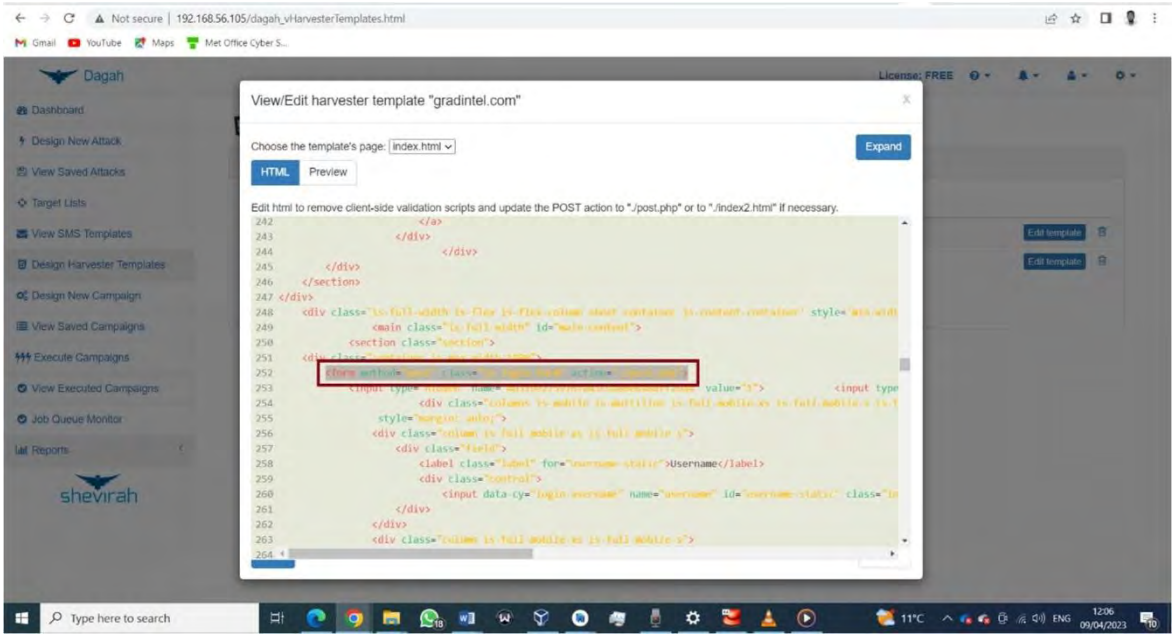


Figure 21. Editing the HTML code to clone the website

we then created a new attack, choosing the *harvester* attack type, setting the delivery mechanism to *SMS*, choosing the harvester template (*gradintel.com*), choosing our target list containing the Android Victim's phone number, and creating a campaign before launching it as depicted in figures 22, 23, 24 and 25.

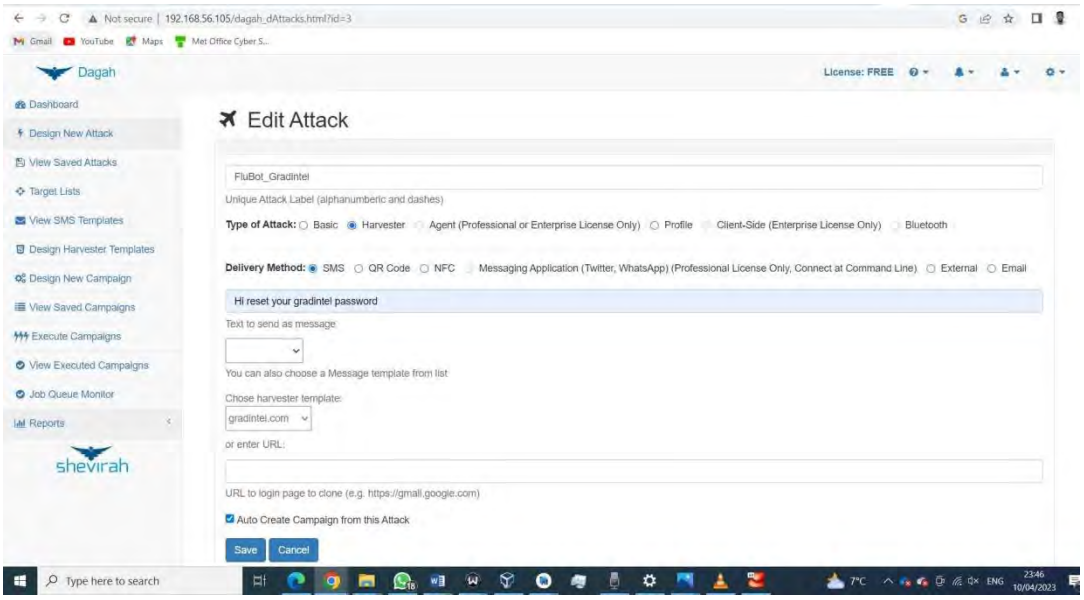


Figure 22. Designing a new website (gradintel.com) harvester phishing attack on Dagah GUI

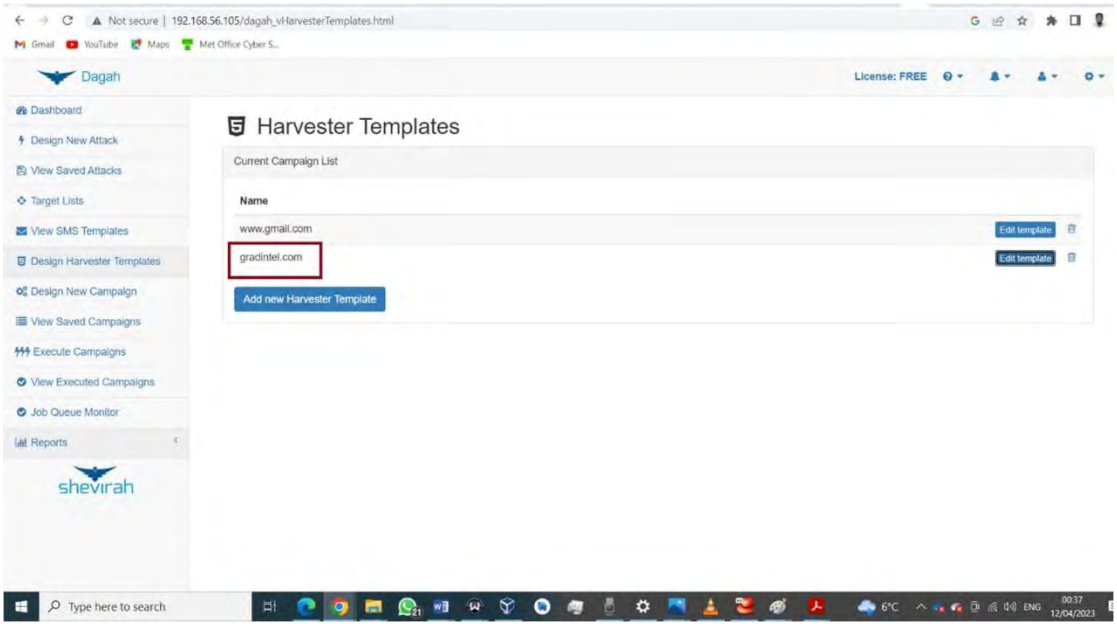


Figure 23. Choosing the newly created „gradintel.com“ harvester template

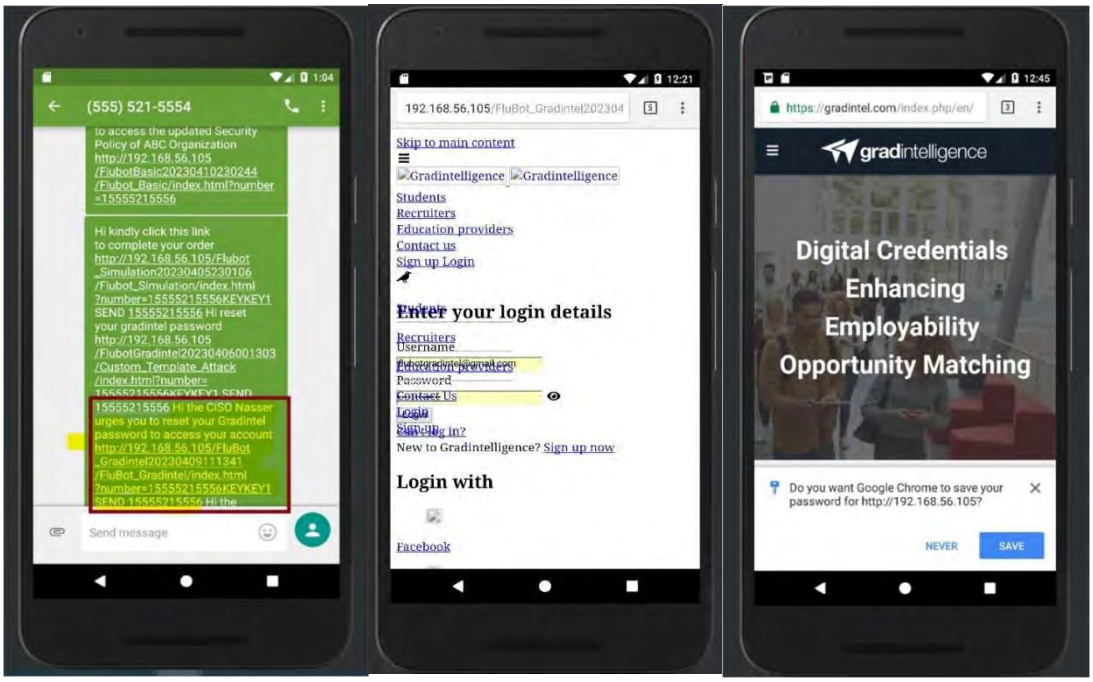


Figure 24. Victim"s device receiving the malicious link, opening the phony Gradintel login page and landing page of the main Website

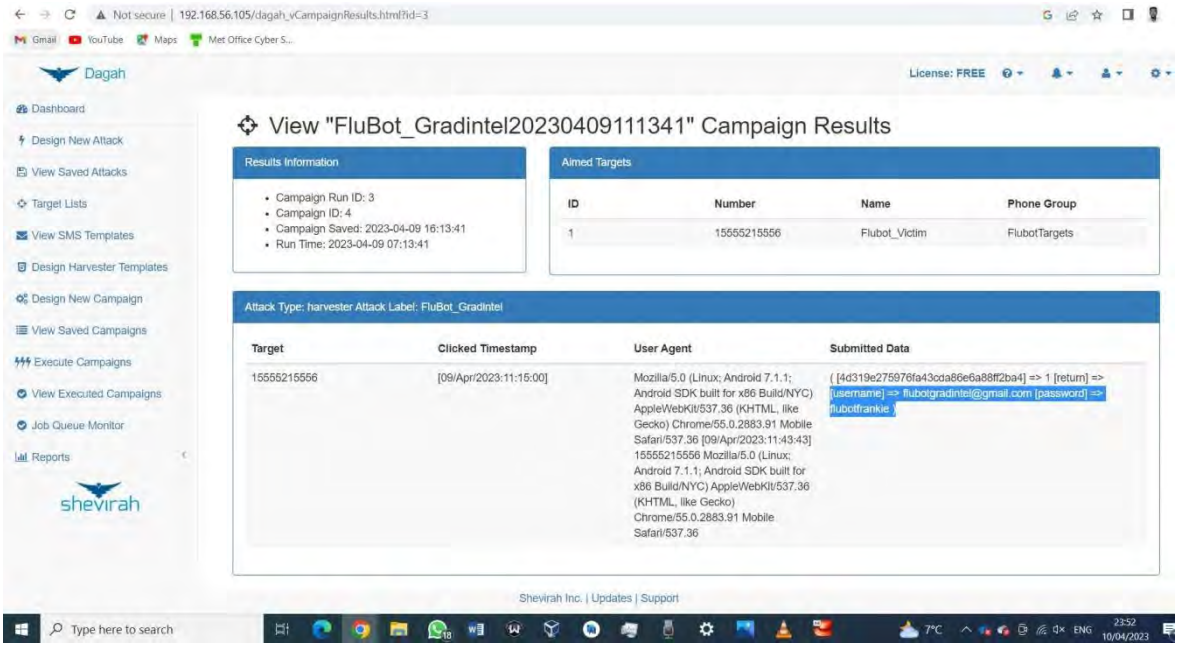


Figure 25. Results of the executed campaign showing the grabbed credentials of the Victim

3.4 Dissemination

3.4.1 Techniques and Procedures

Based on our research and the analysis conducted, we observed the following to be the techniques and procedures instigated by FluBot to obstruct mobile security:

i. String Encryption: FluBot uses a unique encryption method to encrypt all relevant strings. Each class has a function in charge of encrypting any dubious strings it comes across.

ii. MultiDex (Multi Davik Executable): APK files contain DEX (Dalvik Executable) which are executable codes that ensures the running of your Android app. When more than one DEX is generated

to run your app, it is called MultiDex. FluBot conceals its harmful code from reversers and static analyzers by using MultiDex.

iii. DEX Decryption: FluBot employs a decrypted and loaded encrypted dex file from the assets to carry out its malicious behaviour.

iv. Domain Generation Algorithm (DGA): This is an algorithm used in generating domain names. FluBot uses this algorithm to locate and communicate with the C2 server in order to bypass security safeguards.

v. DNS Tunneling over HTTPS: Flubot resolves the IP addresses of DGAs after generation, then communicates with the C2 server through DNS Tunneling over HTTPS port 443.

vi. Error Logging: The C2 records any application errors that are not noticed. This enables the attackers to update and fix the FluBot's code.

3.4.2 Appropriate Intelligence and Cost-Effective Solutions/Countermeasures

This study proposed some appropriate intelligence and cost-effective solutions/countermeasures to protect and improve systems from the FluBot mobile security threat. They include but are not limited to:

i. The internal team (e.g., IT, legal, communications, internal audit, risk management, etc.) should be trained by the organisation using tabletop exercises or other briefings intended to test and enhance incident response function.

ii. A complete system reset or safe boot of the android devices will get rid of the malware and all current settings, including stored data.

iii. For organisations and individuals, ensure you stay informed on phishing tactics and social engineering techniques through system awareness campaigns, workshops and education;

iv. Enable two-factor authentication (2FA) or multi-factor authentication (MFA) on your accounts to provide an extra layer of protection and prevent unauthorised access to your device;

v. Obtain APKs from legitimate vendors rather than unauthorised ones, and avoid installing add-on programmes as they might include the malware Flubot;

vi. Avoid opening attachments from unreliable sources or clicking on suspicious links as they can include Flubot or other malicious programs that has tendencies of compromising your device.

vii. Formulate, review and implement when needed, an intrusion detection system (IDS) to monitor network traffics for suspicious activities and signal alerts when noticed, a Business Impact Analysis and Business Continuity and Disaster Recovery Plan for contingencies.

3.4.3 Unsolved Problem

It is known that Flubot spreads using SMS messages that entice recipients to click on harmful links. Users continue to fall prey to these phishing assaults, spreading infection, despite awareness campaigns and security precautions. It remains a challenge to stop users from clicking on these links and falling for social engineering tricks.

3.5 Utilization

3.5.1 Legal and Ethical Issues

In the event of a breach in the confidentiality, integrity, or availability of an organization's system, it is necessary to have more robust governance structures, as well as legal and ethical obligations to protect and prioritise organisational assets. When there is a legal crisis or APT of any type, managing legal privileges can become a severe problem. Security concerns must be considered when drafting legal agreements with partners, suppliers, and customers. This will enable better containment, communication, and analysis of the technical and legal dangers posed by the attack.

4. Conclusion

This survey report provided a thorough analysis and explanation of the FluBot Android malware, to evaluate it as a danger to mobile security. we were able to identify critical information of the FluBot APT using VirusTotal and Alien Vault, and we were also able to develop and execute a FluBot-like attack against simulated targets using Dagah and Android Simulator. As a result, we were able to:

i. Identify FluBot-instigated tactics and procedures to undermine mobile security;

ii. Make some pertinent intelligence, recommendations and cost-effective fixes and countermeasures to ensure mobile organisational security; and

iii. Talk about moral and legal concerns to safeguard assets.

5. Recommendation

At the end of the investigation, we came up with the following recommendations:

i. Block unknown senders or enable SMS filtering in the device's settings. By doing this, you might be able to prevent harmful SMS messages from reaching your smartphone and potentially propagating the FluBot malware.

ii. Update your operating system, programs, and security updates on a regular basis to keep your devices safe from known vulnerabilities;

iii. Ensure you periodically back up your data to a secure location in order not to lose vital information;

iv. Ensure to employ a reliable antivirus program on your Android device.

v. Avoid jailbreaking your device. This could severely reduce its security and expose gaps in protection.

Acknowledgments

We are grateful to Dr. Ernest E. Onuiri for his useful advice and for giving us some possibilities that made the work better than it was before, and we are grateful to every team member who took the time to participate in this study.

Authors contributions

Dr. Uchenna J. Nzenwata was responsible for android environment implementation, the flubot malware analysis and its documentation. Frank Uchendu was responsible for the implementation using the Dagah tool and the documentation of the Dagah tool analysis. Haruna Ismail and Eluwa M. Jumoke contributed in gathering the literature sources and the documentation. Himikaiye O. Johnson was responsible for the final proof reading and the collation of the article sections. All the authors ensured that the documentation was thoroughly proof read and ascertain equal right to the contribution of the final state of the work.

Funding

Not Applicable

Competing interests

Not Applicable

Informed consent

Obtained.

Ethics approval

The Publication Ethics Committee of the Canadian Center of Science and Education.

The journal's policies adhere to the Core Practices established by the Committee on Publication Ethics (COPE).

Provenance and peer review

Not commissioned; externally double-blind peer reviewed.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Data sharing statement

No additional data are available.

References

"FluBot Android Spyware Taken Down in Global Law Enforcement Operation," The HackerNews. Retrieved Apr. 09, 2023, from https://thehackernews.com/2022/06/flubot-android-spyware-taken-down-by.html

"VirusTotal," www.virustotal.com, Apr. 10, 2023.

Abouzakhar, N. (2023, March 06). Lecture on Advanced Persistent Attacks (APTs), Active Cyber Defence (ACD) and Threat Intelligence and Operations [PowerPoint slides]. Available: https://blackboard.salford.ac.uk/

Blázquez, E., & Tapiador, J. (2023). Kunai: A static analysis framework for Android apps. *SoftwareX, 22,* 101370. https://doi.org/10.1016/j.softx.2023.101370

Callaham, J. (2019). The history of Android OS: its name, origin and more. *Android Authority, 18.*

Chapin, L., Piscitello, D., & Strutt, C. (2022). *Malware Landscape 2022.*

Chaurasia, P. (2015). *Dynamic analysis of Android malware using DroidBox* (Doctoral dissertation, Tennessee State University).

Exchange, A. O. T. (2020). AlienVault Open Threat Exchange.

Fernick, J. (2022). *Flubot: the evolution of a notorious Android Banking Malware*. NCC Group Research, Retrieved July 5, 2022, from https://research.nccgroup.com/2022/07/05/flubot-the-evolution-of-a-notorious-android-banking-malware/

FluBot, "Partners-in-crime: Medusa and Cabassous attack banks side-by-side — ThreatFabric," Retrieved Apr. 09, 2023, from https://www.threatfabric.com/blogs/partners-in-crime-medusa-cabassous.html

García-Teodoro, P., Gómez-Hernández, J. A., & Abellán-Galera, A. (2022). Multi-labeling of complex, multi-behavioral malware samples. *Computers & Security, 121*, 102845. https://doi.org/10.1016/j.cose.2022.102845

Garg, S., & Baliyan, N. (2021). Comparative analysis of Android and iOS from security viewpoint. *Computer Science Review, 40,* 100372. https://doi.org/10.1016/j.cosrev.2021.100372

Gibbs, S. "About the flubot virus," Queensland Tech, Aug. 28, 2021. Retrieved April 10, 2023, from https://queenslandtech.com.au/flubot/

Liu, M., Zhang, Y., Liu, B., Li, Z., Duan, H., & Sun, D. (2021, December). Detecting and characterizing SMS spearphishing attacks. In Annual Computer Security Applications Conference (pp. 930-943). https://doi.org/10.1145/3485832.3488012

Mayrhofer, R., Stoep, J. V., Brubaker, C., & Kralevich, N. (2021). The android platform security model. *ACM Transactions on Privacy and Security (TOPS), 24*(3), 1-35. https://doi.org/10.1145/3448609

Meng, H., Thing, V. L., Cheng, Y., Dai, Z., & Zhang, L. (2018). A survey of Android exploits in the wild. *Computers & Security, 76,* 71-91. https://doi.org/10.1016/j.cose.2018.02.019

Mogicato, R., & Zermin, A. Design and Implementation of a Collaborative, Lightweight Malware Analysis Sandbox using Container Virtualization.

Ndatinya, V., Xiao, Z., Manepalli, V. R., Meng, K., & Xiao, Y. (2015). Network forensics analysis using Wireshark. *International Journal of Security and Networks, 10*(2), 91-106. https://doi.org/10.1504/IJSN.2015.070421

Özdemir, D., & Zaim, H. Ç. (2021). Investigation Of Attack Types in Android Operating System. *Journal of Scientific Reports-A, 46,* 34-58.

"Top 10 Android Security Risks | eSecurity Planet," eSecurityPlanet, Retrieved March 18, 2011, from https://www.esecurityplanet.com/trends/android-security-risks/

Riasat, H., Batool, T., & Iqbal, S. (2022). *Review and Comparative Studies on Mobile Operating System* (No. 8848). EasyChair.

Salsabila, H., Mardhiyah, S., & Hadiprakoso, R. B. (2022, November). *Flubot Malware Hybrid Analysis on Android Operating System.* In 2022 International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS) (pp. 202-206). IEEE. https://doi.org/10.1109/ICIMCIS56303.2022.10017486

Schütte, J., Fedler, R., & Titze, D. (2015, March). *Condroid: Targeted dynamic analysis of android applications.*

In 2015 IEEE 29th International Conference on Advanced Information Networking and Applications (pp. 571-578). IEEE. https://doi.org/10.1109/AINA.2015.238

StatCounter, “Mobile Operating System Market Share Worldwide,” StatCounter Global Stats, 2022. Retrieved from https://gs.statcounter.com/os-market-share/mobile/worldwide

Van Haastrecht, M., Golpur, G., Tzismadia, G., Kab, R., Priboi, C., David, D., ... Spruit, M. (2021). A shared cyber threat intelligence solution for smes. *Electronics, 10*(23), 2913. https://doi.org/10.3390/electronics10232913

Computer and Information Science; Vol. 16, No. 4; 2023
ISSN 1913-8989 E-ISSN 1913-8997
Published by Canadian Center of Science and Education

Proposal of a Visualization System for a Hierarchical Clustering Algorithm: The Visualize Proximity Matrix

Sulaiman Abdullah Alateyah[1]

[1] Department Of Computer Science, College Of Science and Arts, Qassim University, Unaizah, Saudi Arabia;

Correspondence: Sulaiman Abdullah Alateyah Department Of Computer Science, College Of Science and Arts, Qassim University, Unaizah, Saudi Arabia. E-mail: salateyah@qu.edu.sa

Received: October 29, 2023 Accepted: November 28, 2023 Online Published: November 29, 2023

doi:10.5539/cis.v16n4p65 URL: https://doi.org/10.5539/cis.v16n4p65

Abstract

Visual data mining is a new and effective strategy for dealing with the growing phenomenon of information overload. There is an urgent need for effective visual data mining because the growth of data sets from different domains and sources has made exploring, managing and analyzing vast volumes of data increasingly difficult. Identifying location-based trends and anomalies in data from a numeric output is challenging. The outputs of data mining procedures are often quite difficult to interpret. There is no human input to data exploration or direct involvement in the data mining process. Thus, it is difficult to analyze, explore, understand and view the data. In data mining, massive data sets are clustered using an iterative process that includes human input but data mining algorithms can miss some knowledge and observations. In addition, many users cannot take the appropriate action to address problems or opportunities because the huge amounts of data prevent them from staying aware of what is happening in and around their environments. Visual data mining can help in dealing with information overload and it is effective in analyzing large complex data sets. Visual data mining assists users in gaining a deep visual understanding of data. In the information age, users need to study and observe vast amounts of data to acquire important knowledge, and thus the need for visual and interactive analytical tools is particularly pressing. Visual cluster analysis has long utilized shaded similarity matrices, and this study investigated how they can be used in clustering visualization. We focused on proposing an agglomerative hierarchical clustering method. Ensemble cluster visualization is also presented for handling large data sets. This study proposes the adoption of a shaded similarity matrix to visually cluster knowledge discovered using data mining. Using the technology acceptance model as the measuring tool, we questioned respondents to evaluate the visualization prototype. The findings demonstrated that the visualization was effective and easy to use, and satisfied users.

Keywords: Data mining, Visual data mining, Pattern, Algorithm, Exploring

1. Introduction

In this information era, data are collected from many sites and sources such as loop detectors, marketing, technology, medical and banking. Data mining algorithms are applied to detect hidden knowledge or patterns. There are many data mining algorithms that can be adopted for analysis; for example, clustering, decision tree, genetic algorithm and neural networks. Data mining is the result of a long process of research, development and evaluation (Card, 1999; Guo et al., 2020; Spence, 2001; Ware, 2012).

In their work, (Pampalk et al., 2003; Schneiderman & Plaisant, 2005; Wickens & Hollands, 2000) stated that complicated data mining techniques were necessary for the study of complex and heterogeneous data. With new visualization paradigms, many analysis approaches can be applied which benefit from visual data processing. Visualization is a natural way to combine several data sources, has been applied in many different fields and has been confirmed to be reliable and effective. Although visual approaches cannot replace data mining algorithms, it is useful to combine visualization techniques and data mining algorithms in data exploration processes. The main purpose of visualization is to convert data into an appropriate representation or visual form. Then, users can use their recognition skills to interpret, understand, observe, analyze and query the data efficiently. Users should be directly involved in the analysis of the data in order to maximize the usefulness of the visualization tool. They should also be able to dynamically explore the visual representation of the data in order to comprehend them more quickly and easily.

As a result, visual presentation can be extremely effective to highlight patterns, outliers, clusters and data gaps.

This research also aimed to visualize the result extracted by a data mining algorithm called knowledge which enables users to interact with every step of the data mining process, making it easier to interpret and view the data. According to (Keim, 2002), Visual Data Mining (VDM) is a new and effective strategy for dealing with increasing information overload. Never in history have there been as many data points produced as there are now. Effective visual data mining, data mining and visualization are now necessary because of the difficulty of exploring, managing and interpreting the huge volumes of data in the form of ever-expanding data sets from many fields and sources (Meyer & Cook, 2000). Visual data mining is an efficient way to process huge, complicated data sets and can help with information overload. To fully grasp data visually, visual data mining is necessary (Feng et al., 2021; Mendoza-Silva et al., 2021). This study proposes a visual cluster approach to visualize the knowledge extracted by a data mining algorithm based on a tree strategy for monitoring the data, involving the user in the data discovery process and allowing the user to analyze and observe large amounts of data in order to extract valuable knowledge.

2. Literature Review

In order to extract hidden rules, expressions and meaningful patterns from these big data, scalable and reliable analytical algorithms must be developed. Huge amounts of data are amassing in numerous fields, including in scientific and engineering databases.

Visual data mining is an emerging interdisciplinary science aiming to develop automatic or semiautomatic techniques which can discover the knowledge hidden in these databases, to make decision-making processes faster and more efficient. Hence, utilization of data mining in medical, education, finance, engineering, marketing and telecommunication industries has dramatically increased in recent years. Incorporating data mining algorithm and visualization methods is potentially effective, as revealed by successful visual data mining tools such as generative topographic mapping and the self-organizing map. However, a significant amount of integration work remains to be done in order to benefit from advanced results from both domains.

There are various methods for visualizing data; for example, x–y plots, line plots and histograms. These procedures are valuable for data exploration but are generally restricted to small and low-dimensional data sets. In recent decades, a large number of novel data representation methods have been developed, enabling visualization of multidimensional data sets (Card, 1999; Guo et al., 2020; Ran et al., 2023; Spence, 2001; Ware, 2012).

The strength of visualization represents the capacity for discovery (Schneiderman & Plaisant, 2005; Wickens & Hollands, 2000) Implementing visualization tools to examine and comprehend high-dimensional information is currently proving to be an effective method of combining intelligence with the enormous capabilities of the processing power currently available (Pampalk et al., 2003; Yang & Hussain, 2023).

The key benefit is making use of the human visual system to assist the data mining process. This is achieved through the creation of visualizations of the data which allow users to identify features within the data which would not otherwise be apparent (Feng et al., 2021; Keim, 2002; Mendoza-Silva et al., 2021; Meyer & Cook, 2000). According to (Zhang, 2008), the human visual system comprises the brain and the eyes. The eyes can be regarded as a strong and highly parallel processing and reasoning engine.

(Bhadran et al., 2008), explained that visualization techniques are widely used in exploring, understanding, summarizing, interpreting, observing and analyzing large amounts of data (Ankerst et al., 2000; Grinstein & Wierse, 2002; Morrison et al., 2002; Shneiderman, 2001). Many different visualization approaches, including geometric, icon-based, pixel-oriented, hierarchical and graph-based methods, have been created to map multidimensional data sets to two- or three-dimensional space.

3. Proposed Visualizing Proximity Matrix

The shaded similarity matrix is described in this section. For the past 40 years, visual cluster analysis has primarily used shaded similarity matrices. In (Gale et al., 1984; Ling, 1973; Ran et al., 2023), the authors provide a full summary of the early work, whereas (Biedl et al., 2001; Wishart, 1999) include some recent work that is pertinent. Greater likeness is shown by darker shading, whereas lower resemblance is represented by lighter shading. Dark and light cells may be spread throughout the matrix at first, so the rows and columns are restructured so that similar things are placed adjacent to one another to show possible groupings. If there are "real" clusters in the data, they should show up as symmetrical black squares on the diagonal (Cao et al., 2023; Gale et al., 1984). This tutorial will explain how a shaded similarity matrix is constructed and how it looks using an example. The data for this example are taken from the literature data set (Merz & Murphy, 1996; Zhu et al., 2018; Zidan et al., 2020).

The dendrogram tree method is used in hierarchical cluster analysis to visualize how the cluster is merged. The visualizing proximity matrix proposed in this study shows the cluster in a contrasting color and also shows the distance between the merged clusters. The following snapshots explain the agglomerative hierarchical clustering and visualizing proximity matrix. Fig 1 shows the main page of the proposed agglomerative hierarchical clustering.

Figure 1. Main Page

The similarity measure dialog box is shown in Fig 2. This specifies the distance measure and the clustering method. The first step is to select the similarity distance measure. For interval data, the most common measure is Euclidian distance.

Figure 2. Similarity measure.

Complete, single and average hierarchical clustering methods are the linkage methods used to calculate the distance between data points. These are all based on Euclidean distance but the main difference between them is the selection of the data points that are considered as the final criterion on which the similarity or distance depends, as shown in Fig 3.

Figure 3. Clustering method.

Fig 4 shows the visualization dialog box. The Visualize Proximity Matrix and the dendrogram will graphically show how the clusters are merged and enable identification of the appropriate number of clusters. Then, the user chooses a specified range of clusters or all clusters, as shown in Fig 5.

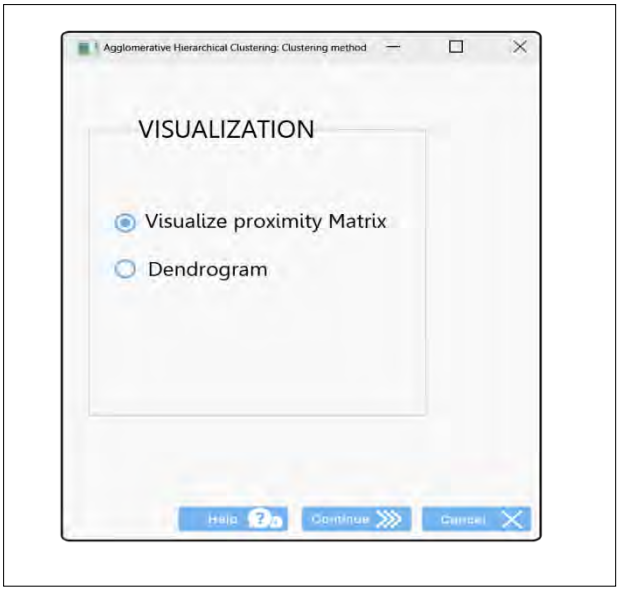


Figure 4. Visualization.

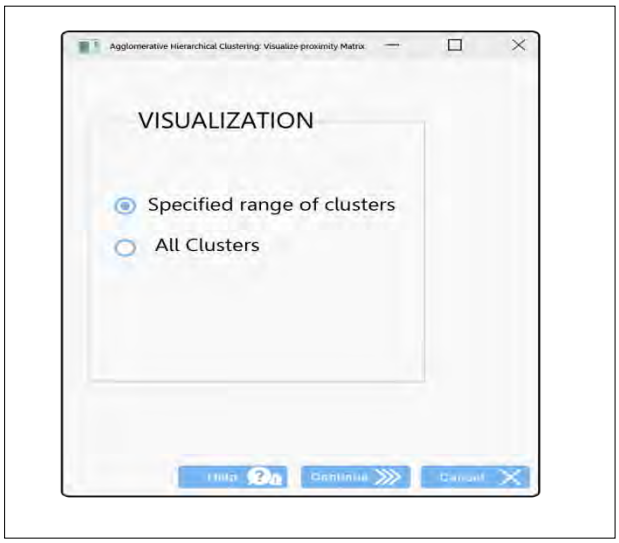


Figure 5. Specified Range Of Clusters Or All Clusters.

Using single linkage clustering, Fig 6 shows that the nearest pair is “D” and “F”, at a distance of 0.50. The D and F are merged together into one cluster called “D, F". Fig 7 shows that the nearest pair is “A” and “B”, at a distance of 0.71. The A and B are merged together into one cluster called “A, B". Fig 8 shows that the nearest pair is “E” and “D,F”, at a distance of 1.00. The D, F and E are merged together into one cluster called “D, F, E". Fig 9 shows that the nearest pair is “C” and “D, F, E”, at a distance of 1.41. The D, F, E and C are merged together into one cluster called “D, F,E,C". Fig 10 shows that the nearest pair is “A,B” and “D, F, E,C”, at a distance of 2.50. Finally, these clusters are merged together into a single cluster called “A,B, D, F,E,C".

Figure 6. Visualize Proximity Matrix for Cluster 1.

Figure 7. Visualize Proximity Matrix for Cluster 2.

Figure 8. Visualize Proximity Matrix for Cluster 3.

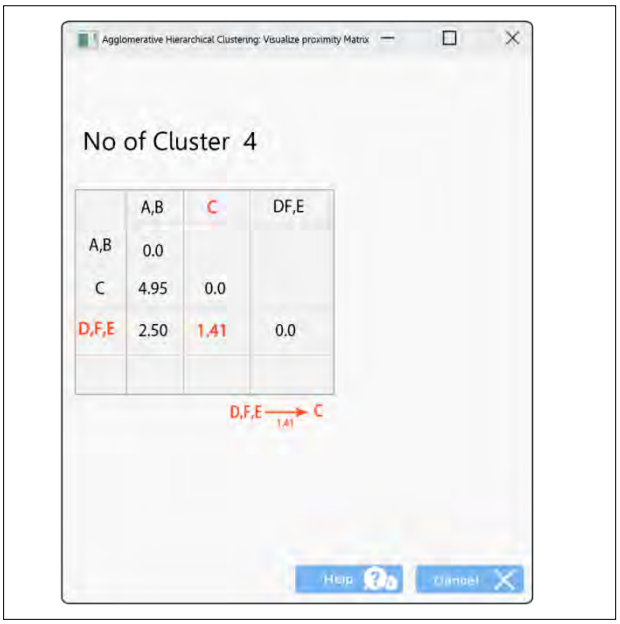


Figure 9. Visualize Proximity Matrix for Cluster 4.

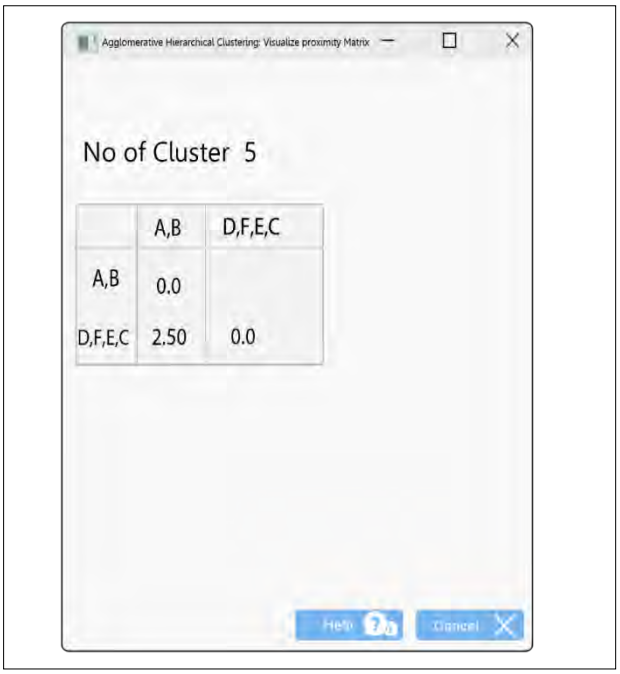


Figure 10. Visualize Proximity Matrix for Cluster 5.

4. Methodology and Hypothesis Development

Examination of the proposed proximity matrix quantifies its viability and determines whether the targeted interest group has reached the indicated locations. Evaluation should cover both the incorporation of the visualization with data mining algorithms, as well as the usability and usefulness of the visualization element for controlling the data. It should include the role of the user in the data exploration process, and whether it enables users to examine many facts to obtain useful information. According to (Kanaujiya, 2008), a visual data mining prototype must be syntactically simple to be useful. To be simple to learn, it needs to be easy to extract and interpret knowledge using intuitive and user-friendly tools. To be simple to apply, it needs to allow efficient communication between humans and data. Questionnaires have long been used to evaluate software systems and user interfaces (Root & Draper, 1983). The biggest single advantage of using questionnaires in evaluation of an interactive prototype is that they provide data on prototype acceptance from the user point of view.

(Yamazawa et al., 2008) visualized the drags based on chemical structures, after dividing the drags using the hierarchical clustering method. They evaluated user experiments, discussed the effectiveness of the presented technique using the level of detail (LOD) control technique. Eleven examinees were instructed to operate and explore the user interface in the LOD for a given time, and then rate it. All the participants in the assessment were either postgraduate or undergraduate information science students.

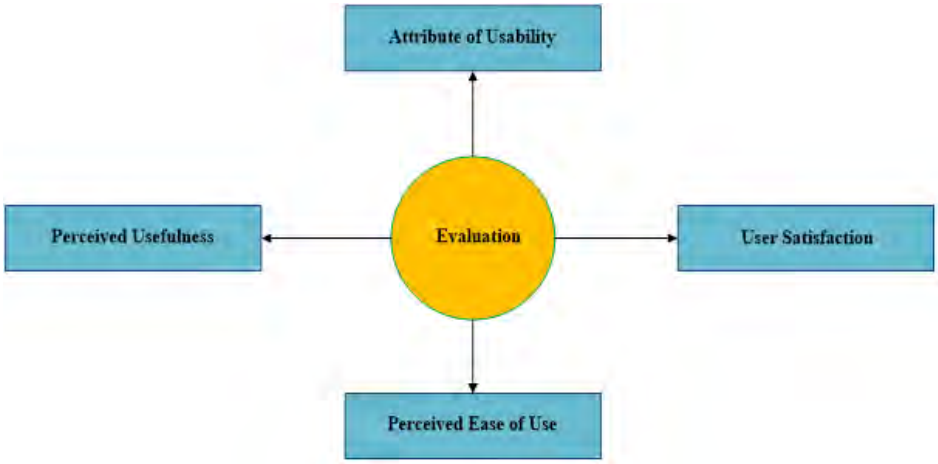


Figure 11. Evaluation Dimensions.

The sampling procedure that was adopted in this study for data collection was a target sampling method using a questionnaire survey issued to 38 respondents. The sample consisted of 23 males and 15 females and 27 were in the age group 25–35. Generally, the respondents had seen and watched the visualization prototype of the bidirectional agglomerative hierarchical clustering algorithm. The demographic profile of the respondents is presented in Table 1.

Table 1. Demographic Profile

Variable		Frequency	Percentage
Gender	***Female***	23	61%
	Male	15	39%
Age	25-35 years	27	71%
	36-45 years	6	16%
	46 >	5	13%

4.1 Dimensions

4.1.1 Perceived Usefulness

This defines the degree to which users believe that observation and exploration of knowledge will be improved by use of the visualization prototype. Table 2 shows the items used to measure the dimension of perceived usefulness. The first dimension examines HYPOTHESIS 1: Through visualization it is possible to observe and explore knowledge within the data that has been missed by data mining algorithms, and working with the visualization output is easier than working with a numeric output.

4.1.2 Perceived Ease of Use

This refers to the degree to which users believe that using the visualization system is effortless and that interaction with the visualization system is clear and understandable. Table 2 shows the items used to measure the dimension of perceived ease of use. The second dimension examines HYPOTHESIS 2: Interaction with visualization is clear, understandable and effortless. It is easy to become skilled at using visualization to explore and observe knowledge.

4.1.3 User Satisfaction

This refers to what users expect from the system, and therefore is a personal assessment about what the system should do for the end-user. Table 2 shows the items used for measurement of the user satisfaction dimension. The third dimension examines HYPOTHESIS 3: It is easy to be aware of what is happening in and around their environments from the huge amounts of data and to draw out insights by facilitating commentary or discussion regarding the experience.

4.1.4 Attribute of Usability

This is the area of human–computer interaction (HCI) with regard to the proposed visualization system. It attempts to bridge the gap between the goals of the user and the system. Table 2 shows the items used to measure the user attribute of usability. The fourth dimension examines HYPOTHESIS 4: Does the visualization introduce human issues into the design and devise practical techniques for the observation of human behavior and performance?

The hypothesis development and the four dimensions are shown in Fig 12.

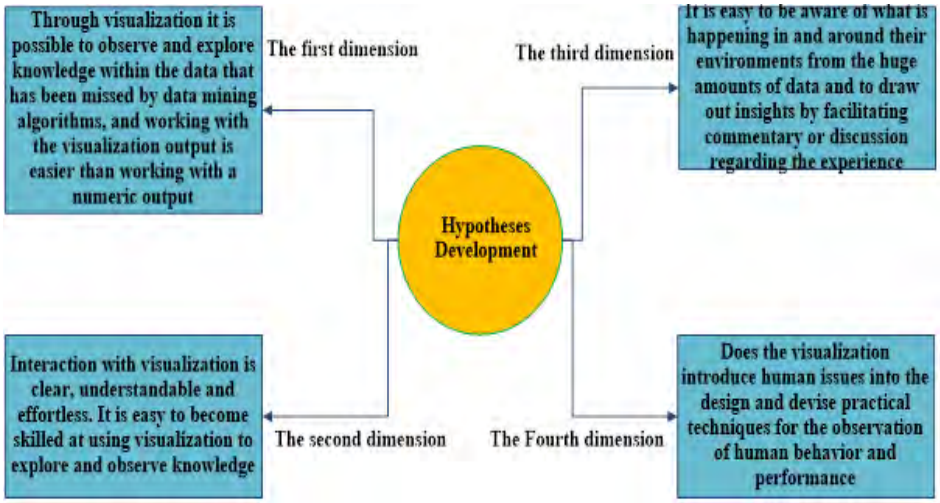


Figure 12. Hypothesis Development

5. Results and Discussion

The user testing of the visualization tool and the Visualize Proximity Matrix examined the effectiveness and usefulness of visualization tool. We asked 38 respondents to use and explore the visualization tool for several minutes and then evaluate it. All the respondents had knowledge of or connection with the fields of computer science and information technology. The descriptive statistics for the main variables in Table I revealed that all dimensions were scored higher than the midpoints of their respective scales. This shows that respondents were generally optimistic about the four dimensions of using the visualization tool to extract and view knowledge that has been missed by data mining algorithms. Additionally, users and data miners were able to be involved and interact in the data processes by exploiting the power of the human brain and visual abilities to analyze and explore data. Therefore, the visualization prototype shows promise as a useful tool to solve the problem statements for this research:

- Human beings are not included in data exploration processes.
- Some knowledge and observations are missed by data mining algorithms.

The respondents proved that working with visualization usually assists in the rapid discovery of data and understanding its structure. It was also proved that it is easier to work with a visualization output than with a numeric output, and that it is effective in enabling the users and data miners to identify and view hidden patterns and rules in data that might have been missed by the data mining algorithms.

The findings of the evaluation identify some areas of the teaching materials that require clarification. The visualization tool needs to be suitable for use by beginners and both novice and experienced users must be able to access comprehensive instructional information. Finally, based on self-reporting, some people are highly visual whereas others are not. Table 2 shows the responses to the survey concerning the four dimensions and Fig 13 shows the Hypotheses Testing for all questions.

Table 2. Dimension Items.

Dimensions Items	Responses /Mean
First dimension to examine the HYPOTHESIS 1: "Through visualization it is possible to observe and explore knowledge within the data that has been missed by data mining algorithms, and working with the visualization output is easier than working with a numeric output"	3.96%
PU1: Using visualization tool can help the user explore, observe the knowledge easier	3.8%
PU2: Proposed Visualization tool will enable the user to get information of data quickly	3.4%
PU3: Working with visualization output is easier than working with numeric output	4.2%
PU4: By using Visualization tool humans might catch and observe hidden patterns and rules in data	4.3%
PU5: By using visualize proximity Matrix is directly involve and interactive in the data processes by exploitation the power of the human sight and brain for analyzing and exploring data	4.1%
Second dimension to examine the HYPOTHESIS 2: “Interaction with visualization is clear, understandable and effortless. It is easy to become skilled at using visualization to explore and observe knowledge "	3.54%
EU1: Learning to operate Visualization tool would be easy for me	3.1%
EU2: I find it easy get Visualization tool to do what I want it to do	3.4%
EU3: My interaction with Visualization tool would be clear and understandable	4.2%
EU4: I found Visualization tool flexible to interact with	4.1%
EU5: It is easy for me to become skillful at using Visualization tool	2.9%
Third dimension to examine the HYPOTHESIS 3: " It is easy to be aware of what is happening in and around their environments from the huge amounts of data and to draw out insights by facilitating commentary or discussion regarding the experience"	3.38%
US1: I completely satisfied in using the Visualization tool	3.6%
US2: I feel very confident in using the Visualize proximity Matrix	3.1%
US3: I found it easy to explore the data by using Visualize proximity Matrix	2.8%
US4: I can accomplish the task quickly using this procedure	3.2%
US5: I believe that from using Visualize proximity Matrix it easy to stay aware of what is happening in and around their environments from the huge amounts of data	4.2%
Fourth dimension to examine the HYPOTHESIS 4: "Does the visualization introduce human issues into the design and devise practical techniques for the observation of human behavior and performance "	3.56%
AU1: It easy to interact with Visualization tool by using Visualize proximity Matrix	4.4%
AU2: The procedure through Visualize proximity Matrix is clear	4.1%
AU3: I found the use of Visualize proximity Matrix is suitable for each community groups.	2.6%
AU4: I found the various functions in this Visualization tool were well integrated	3.5%
AU5: I think that i would like to use this Visualization tool always	3.2%

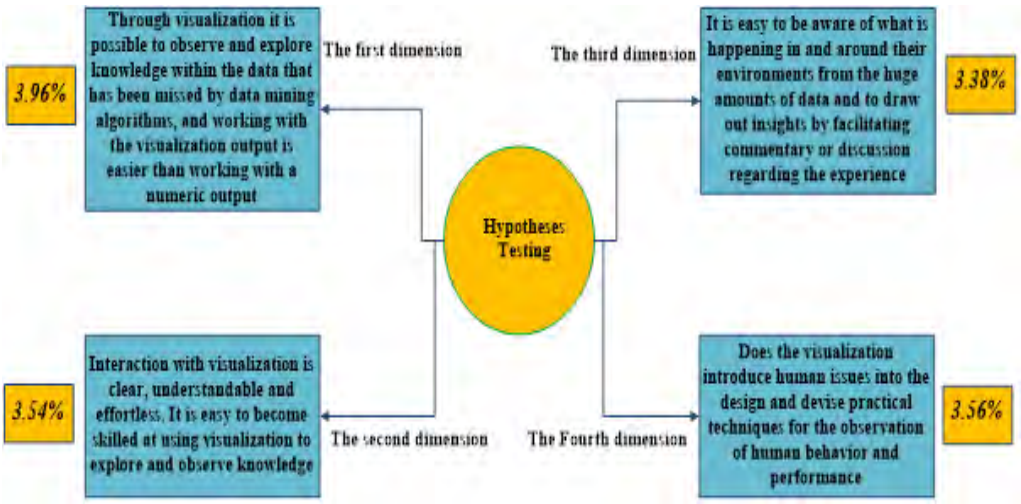


Figure 13. Hypotheses Testing.

6. Conclusion

Data mining processes involve multiple stages such as target data, data integration, data cleaning, data selection, data transformation and the output of the data mining algorithm. The objective of this paper is to propose a visualization tool, the Visualize Proximity Matrix, which will assist in knowledge discovery, understanding the

structure of the data and handling large amounts of data through the provision of an effective exploratory visualization tool. Interactive analysis tools reduce the gap between the human being and the flood of information that the human needs to search in order to extract valuable knowledge. Therefore, visual data mining has become a critical technological process, which uses visual data mining steps to avoid information overload. There is a need to be able to use cognitive abilities to transform the data into information that can eventually be used to make decisions, solve problems, improve products and increase understanding.

Visualization is a highly effective modality for understanding the structure of data and information. Based on the results of this evaluation, the visualization prototype shows promise as a tool for exploring, observing and increasing understanding of data. This evaluation enabled users and data miners to interact and be directly involved with the visualization prototype in order to easily explore and extract knowledge from the data that was missed by the data mining algorithms.

Acknowledgments

The researchers would like to thank the Deanship of Scientific Research, Qassim University for funding publication of this project.

Authors contributions

Dr. Sulaiman Alateyah was responsible for whole study.

Funding

Not applicable.

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Informed consent

Obtained.

Ethics approval

The Publication Ethics Committee of the Canadian Center of Science and Education.

The journal's policies adhere to the Core Practices established by the Committee on Publication Ethics (COPE).

Provenance and peer review

Not commissioned; externally double-blind peer reviewed.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Data sharing statement

No additional data are available.

References

Ankerst, M., Ester, M., & Kriegel, H. P. (2000). Towards an Effective Cooperation of the User and the Computer for Classification. *Proceedings of the Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 179-188. https://doi.org/10.1145/347090.347124

Bhadran, V., Roy, R. C., & Gopikakumari, R. (2008). Visual representation of 2-D DFT in terms of 2× 2 Data: A pattern analysis. *2008 International Conference on Computing, Communication and Networking*, 1–9. https://doi.org/10.1109/ICCCNET.2008.4787762

Biedl, T., Brejová, B., Demaine, E. D., Hamel, A. M., & Vinar, T. (2001). Optimal arrangement of leaves in the tree representing hierarchical clustering of gene expression data. *University of Waterloo, Canada*.

Cao, L., Zhao, Z., & Wang, D. (2023). Clustering algorithms. In *Target Recognition and Tracking for Millimeter Wave Radar in Intelligent Transportation* (pp. 97-122). Springer. https://doi.org/10.1007/978-981-99-1533-0_5

Card, M. (1999). *Readings in information visualization: using vision to think*. Morgan Kaufmann.

Feng, X., Nguyen, K. A., & Luo, Z. (2021). A survey of deep learning approaches for WiFi-based indoor positioning. *Journal of Information and Telecommunication*, 1-54.

Gale, N., Halperin, W. C., & Costanzo, C. M. (1984). Unclassed matrix shading and optimal ordering in hierarchical cluster analysis. *Journal of Classification*, *1*(1), 75-92. https://doi.org/10.1007/BF01890117

Grinstein, U. M. F. G. G., & Wierse, A. (2002). *Information visualization in data mining and knowledge discovery*. Morgan Kaufmann.

Guo, L., Wang, F., Sang, J., Lin, X., Gong, X., & Zhang, W. (2020). Characteristics analysis of raw multi-GNSS measurement from Xiaomi Mi 8 and positioning performance improvement with L5/E5 frequency in an urban environment. *Remote Sensing*, *12*(4), 744. https://doi.org/10.3390/rs12040744

Kanaujiya, S. (2008). Visual data mining. *Proceedings of 2nd National Conference on Challenges & Opportunities in Information Technology (COIT-2008) RIMT-IET, Mandi Gobindgarh. March 29*.

Keim, D. A. (2002). Information visualization and visual data mining. *IEEE Transactions on Visualization and Computer Graphics*, *8*(1), 1-8. https://doi.org/10.1109/2945.981847

Ling, R. L. (1973). A computer generated aid for cluster analysis. *Communications of the ACM*, *16*(6), 355-361. https://doi.org/10.1145/362248.362263

Mendoza-Silva, G. M., Costa, A. C., Torres-Sospedra, J., Painho, M., & Huerta, J. (2021). Environment-aware regression for indoor localization based on WiFi fingerprinting. *IEEE Sensors Journal*, *22*(6), 4978-4988. https://doi.org/10.1109/JSEN.2021.3073878

Merz, C. J., & Murphy, P. M. (1996). *UCI repository of machine learning databases. University of California, Department of Information and Computer Science*. Technical Report. Retrieved from http://www.ics.uci.edu/mlearn/MLRepository. html

Meyer, R. D., & Cook, D. (2000). Visualization of data. *Current Opinion in Biotechnology*, *11*(1), 89-96. https://doi.org/10.1016/S0958-1669(99)00060-9

Morrison, A., Ross, G., & Chalmers, M. (2002). Combining and comparing clustering and layout algorithms. *Tech. Rep. 148, Department of Computing Science, University of Glasgow*.

Pampalk, E., Goebl, W., & Widmer, G. (2003). Visualizing changes in the structure of data for exploratory feature selection. *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 157–166. https://doi.org/10.1145/956750.956771

Ran, X., Xi, Y., Lu, Y., Wang, X., & Lu, Z. (2023). Comprehensive survey on hierarchical clustering algorithms and the recent developments. *Artificial Intelligence Review*, *56*(8), 8219-8264. https://doi.org/10.1007/s10462-022-10366-3

Root, R. W., & Draper, S. (1983). Questionnaires as a software evaluation tool. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 83-87. https://doi.org/10.1145/800045.801586

Schneiderman, B., & Plaisant, C. (2005). *Designing the user interface: Strategies for Effective Human-Computer Interactions. The United States of America*. Pearson Education, 4th edition, Reading, MA: Addison-Wesley.

Shneiderman, B. (2001). Inventing discovery tools: Combining information visualization with data mining. *International Conference on Discovery Science*, 17-28. https://doi.org/10.1007/3-540-45650-3_4

Spence, R. (2001). *Information visualization* (Vol. 1). Springer. https://doi.org/10.1057/palgrave.ivs.9500019

Ware, C. (2012). *Information Visualization: Perception for Design, Waltham, MA*. USA: Morgan Kaufmann.

Wickens, C. D., & Hollands, J. G. (2000). *Engineering Psychology and Human Performance. (2000)*. Upper Saddle River, NJ: Prentice Hall.

Wishart, D. (1999). Clustan Graphics3 Interactive Graphics for Cluster Analysis. In *Classification in the Information Age* (pp. 268–275). Springer. https://doi.org/10.1007/978-3-642-60187-3_27

Yamazawa, M., Itoh, T., & Yamashita, F. (2008). Visualization and Level-of-Detail Control for Multi-Dimensional Bioactive Chemical Data. *2008 12th International Conference Information Visualisation*,

11-16. https://doi.org/10.1109/IV.2008.19

Yang, M. S., & Hussain, I. (2023). Unsupervised multi-view K-means clustering algorithm. *IEEE Access*, *11*, 13574-13593. https://doi.org/10.1109/ACCESS.2023.3243133

Zhang, F. (2008). The application of visualization technology on e-commerce data mining. *2008 Second International Symposium on Intelligent Information Technology Application*, *2*, 563-566. https://doi.org/10.1109/IITA.2008.18

Zhu, N., Marais, J., Bétaille, D., & Berbineau, M. (2018). GNSS position integrity in urban environments: A review of literature. *IEEE Transactions on Intelligent Transportation Systems*, *19*(9), 2762-2778. https://doi.org/10.1109/TITS.2017.2766768

Zidan, J., Adegoke, E. I., Kampert, E., Birrell, S. A., Ford, C. R., & Higgins, M. D. (2020). GNSS vulnerabilities and existing solutions: A review of the literature. *IEEE Access*, *9*, 153960-153976. https://doi.org/10.1109/ACCESS.2020.2973759

Computer and Information Science; Vol. 16, No. 4; 2023
ISSN 1913-8989 E-ISSN 1913-8997
Published by Canadian Center of Science and Education

Applying AI in the Healthcare Sector: Difficulties

Abdussalam Garba[1], Muhammad Baballe Ahmad[1], & Mukhtar Ibrahim Bello[2]

[1] Department of Mechatronics Engineering, Nigerian Defence Academy (N.D.A), Kaduna, Nigeria.

[2] Department of Computer Science, School of Technology, Kano State Polytechnic, Kano, Nigeria

Correspondence: Muhammad Baballe Ahmad, Department of Mechatronics Engineering, Nigerian Defence Academy (N.D.A), Kaduna, Nigeria. E-mail: mb.ahmad@nda.edu.ng. https://orcid.org/0000-0001-9441-7023

Received: October 9, 2023 Accepted: November 23, 2023 Online Published: November 29, 2023

doi:10.5539/cis.v16n4p78 URL: https://doi.org/10.5539/cis.v16n4p78

Abstract

Artificial intelligence (AI) broadly speaking refers to any behavior shown by a computer or system that is similar to that of a person. Computers can learn from data without explicit human programming thanks to a kind of artificial intelligence known as "machine learning". The application of artificial intelligence (AI) technologies in medicine is one of the most important current trends in global healthcare. Artificial intelligence-based technologies are radically changing the global healthcare system by allowing for a drastic rebuilding of the medical diagnostics system and a corresponding decrease in healthcare costs. Prior to beginning treatment, an illness must be classified into which class of disorders it belongs. It is possible to classify the disease kind according to the feature space of the ailment. Machine learning algorithms can help with this problem.

Keywords: Difficulties, Goals, Machine Learning, Healthcare, Artificial intelligence (AI), Chronic Diseases

1. Introduction

Because traditional programming techniques offer less flexibility, it is becoming more difficult to arrange, analyze, and respond to the daily increase in data production. Systems that can learn from data by identifying patterns and connections across data sets to improve predictions are becoming more and more in demand today. Artificial intelligence's machine learning subfield enables computers to learn from data without explicit human programming (S. Brown, 2021). In a broad sense, artificial intelligence (AI) refers to any computer or system behavior that resembles human behavior. The most basic kind of artificial intelligence is the "imitation" of human behavior by computers, which is based on considerable data on previous instances of the same behavior. The same task of utilizing computers to understand human intelligence is connected to artificial intelligence (AI), which is not always restricted to biologically logical approaches. In the twenty-first century, the field of AI has grown steadily. With dramatic revolutions influenced by both ideas and tactics, the evolution of AI has improved the development of human society in our own time (Liu et al., 2018). Deep learning, machine learning, and artificial intelligence are all areas of active research. And it appears that it may soon be able to fully replace human intelligence. The study of artificial intelligence, or AI, is a branch of computer science that focuses on creating intelligent computer systems, or systems that possess the skills that are typically associated with the human mind, such as language comprehension, learning, the capacity for reasoning, problem-solving, etc. Later, a variety of software programs and algorithms started to be referred to as AI; its defining characteristic is that they have the ability to answer some issues in the same way that a human would. For instance, AI is just starting to permeate medicine through speech processing (R. M. Fazliddinovich & B. U. Abdumurodovich, 2017), natural language text processing [M. Musaev, I. Khujayorov & M. Ochilov, 2020], object identification (M. Rakhimov, J. Elov, U. Khamdamov, S. Aminov & S. Javliev, 2021). voiceprint recognition (Khdier, Hajer & Jasim, Wesam & Aliesawi, Salah, 2021), robotics (Mihret, Estifanos., 2020), handwritten character recognition (M. Musaev & M. Rakhimov, 2020), expert systems (Brown, Carol & O'Leary, Daniel., 1995), and medical diagnostics (Kashyap, Abhishek., 2018). There have already been a lot of intriguing computer algorithms and inventions in this field, but they are still a long way from being widely used because they lack clinical evidence of their efficacy. However, it should be acknowledged that narrowly focused artificial intelligence will firmly take its position given how swiftly this subject has progressed over the past few years and the fact that computers are now outperforming people in solving specific medical problems, which will rise significantly. All areas of human activity—including medicine and healthcare—have been affected by AI technologies. The medical professional must stay current with the most recent developments in medical science. A doctor cannot treat

patients, rest, update knowledge, and maintain it in his thoughts at the same time, hence they cannot perform this task as quickly as AI (M. A. Baballe, A. M. Gambale, A. S. Bari, A. S. Lawan, & R. J. Suleiman., 2022); (A. H. Muhammad, A. Y. Abdullahi, A. Abba, A. Isah, A. A Yako, M. A. Baballe, 2022); (M. A. Baballe, et al., 2022). AI can keep all the information gathered and regularly update research data. The adoption of such technology will simplify life for medical professionals. In fact, one of the most significant aspects of healthcare that AI technologies may aid with is the treatment of chronic diseases. Broadly speaking, chronic diseases are problems that last for a year or longer and necessitate continuing medical care, restrict daily activities, or both. The main causes of death and disability worldwide are chronic illnesses such renal disease, heart disease, cancer, and diabetes. The annual investment in AI had a modest decline in 2018, however it was only momentary. The majority of total corporate investments in AI are private. The amount invested in artificial intelligence initiatives for the healthcare industry worldwide in 2021 increased to $11.2 billion from $8 billion in 2020. The Stanford Institute for Human-Centered Artificial Intelligence released such statistics in March 2022. The study found that from 2017 to 2021, the "attractive" businesses for private investment in the artificial intelligence market were those related to medicine and healthcare. During this time, core projects received a total investment of more than $28.9 billion (Artificial Intelligence Index Report, 2022). Automation and increasing the precision of diagnostics are two crucial topics. The classification of diseases is one method for increasing the precision of diagnosis. AI in the form of machine learning (ML) (Ławrynowicz, Agnieszka & Tresp, Volker., 2014). enables the classification of illness kinds that are similar to one another in terms of a parametric factor. And one of the fundamental machine learning algorithms used for classification is K-Nearest Neighbor (KNN) (Cunningham, Padraig & Delany, Sarah., 2007). A neural network can also be used to tackle the categorization problem (M. Rakhimov, T. Boburkhon & T. Khurshid, 2021). High-performance hardware is needed for deep learning algorithms that use huge datasets, such as heterogeneous computing systems (M. Rakhimov & M. Ochilov, 2021) or parallel computing techniques. At the moment, parallel and distributed computing technologies (M. Musaev & M. Rakhimov, 2019); (M. Rakhimov, D. Mamadjanov and A. Mukhiddinov, 2020) can also be used to overcome this issue. The major goal of this study is to choose significant parametric variables from the gathered disease data that produce more F1-score outcomes. For classification, two forms of coronary heart disease were chosen. It is suggested to use the KNN algorithm for categorizing coronary heart disease. It can be viewed as an algorithm that, when used with the training dataset, generates predictions based on the characteristics of other data points that are present adjacent to it (M. Rakhimov, R. Akhmadjonov, S. Javliev, 2022). In medical data mining, hidden patterns in datasets are discovered. For the early diagnosis of cardiac disease, a supervised algorithm like KNN is employed. The most well-known, successful, and efficient algorithm for pattern recognition is KNN, a frequently used lazy classification algorithm. The distance measure and K value both affect how accurate KNN is. Cosine and Euclidean distance are two other methods for calculating the separation between two instances. KNN determines its closest neighbors and determines a class by majority vote in order to evaluate a fresh unknown sample (Ma, Jabbar., 2017). When the training sample is large, lazy learning techniques like the KNN classifier can be expensive to use because they need to store the whole training sample. In order to reduce storage and processing needs, the compressed closest neighbor classifier incrementally caches a portion of the sample (Alpaydin, E., 1997). Due to its ease of use and relatively quick convergence speed, KNN is growing in popularity (Jabbar MA, Deekshatulu BL, & Priti C., 2013). Medical information technology has advanced toward intelligence as a result of the quick growth of information technology. For the intellectualization of medical information, the classification of large data in health care is extremely important. The KNN classification technique is straightforward, which has led to its widespread application in numerous disciplines (W. Xing & Y. Bei, 2020). One area of healthcare that might be categorized is coronary heart disease (CHD). The Center for Specialized Cardiology's medical personnel and the CHD statistics were both discussed. The CHD dataset was collected from the National Center for Health Statistics (NCHS) (Huang, Nur & Ibrahim, Zaidah & Diah, Norizan., 2021). The main developments in machine learning will be discussed in this paper, including automated data analysis for patient health records and data-driven prediction. The advancements in computer-aided diagnosis, medication discovery, and personalized medicine will also be contrasted (T. M. Tassew, X. Nie, 2022). It is impossible to stress the importance of using big data analytics and machine learning to improve patient outcomes and healthcare performance. With the use of these technologies, healthcare professionals are now able to gain useful insights from big datasets that were previously unexplored, opening up a whole new world of opportunities. By utilizing these information, medical professionals can decide more intelligently about tailored medicine, treatment plans, and resource allocation, ultimately improving patient outcomes and making the healthcare system more effective. Healthcare professionals may now more easily spot trends, correlations, and risk factors thanks to the ability to analyze enormous amounts of healthcare data. This information enables early disease detection, disease prevention, and patient-specific treatment approaches (F. Del

Giorgio Solfa, & F. R. Simonato, 2023). In this opinion piece, we will examine AI's enormous influence on medicine while noting both its possible advantages and impending difficulties (Mehta V., 2023). The application of AI and ML in healthcare has grown in importance, creating new opportunities for innovation, precision medicine, and better decision-making. It is essential to investigate the potential, difficulties, and ethical ramifications of integrating AI and ML into healthcare as we set out on this transformative journey. The field of diagnostics is one of the primary areas where AI and ML have demonstrated tremendous promise. These technologies can swiftly and precisely find trends, spot anomalies, and help with disease diagnosis by analyzing enormous amounts of medical data. The early diagnosis of diseases like cancer and better patient outcomes are made possible by AI-powered algorithms' outstanding accuracy in analyzing medical images like X-rays and MRIs (A. Naveed, 2023). A general taxonomy of machine learning algorithms is presented in this overview, which is followed by a more in-depth explanation of each algorithm class, its function and capabilities, and examples of applications, particularly in geriatric medicine. Additional emphasis is placed on the implications for clinical practice, the difficulties associated with depending on devices with limited interpretability, and the advancements made in overcoming the latter through the creation of explainable machine learning (R. J. Woodman, A. A. Mangoni, 2023). Examining how machine learning technologies might enhance healthcare operations management is the goal of this study. To accomplish this research goal, a machine-learning-based model to address a specific medical issue is created. This study specifically uses the CNN (convolutional neural network) technique to propose an AI solution for diagnosing malaria infection. A total of 24,958 photos were used for deep learning training using malaria microscopy image data from the NIH National Library of Medicine, and 2600 images were chosen for final testing of the suggested diagnostic architecture. The empirical findings show that, with minimal misclassification and performance metrics of precision (0.97), recall (0.99), and f1-score (0.98) for parasite cells and precision (0.99), recall (0.97), and f1-score (0.98) for uninfected cells, the CNN diagnostic model correctly identified the majority of malaria-infected and non-infected cases. The CNN diagnostic solution processed a large number of cases quickly and with a 97.81% accuracy that could be relied upon. The k-fold cross-validation test was used to further validate the performance of this CNN model. These findings imply that machine learning-based diagnostic techniques have an edge over traditional manual diagnostic techniques when it comes to enhancing operational capacities in the healthcare sector in terms of diagnostic quality, processing expenses, lead times, and productivity. In addition, by lowering the probability of unneeded medical disputes connected to diagnostic errors, a machine-learning diagnosis system is more likely to improve the financial viability of healthcare operations. Propositions with a research framework are offered to examine the effects of machine learning on healthcare operations management for safety and quality of life in international communities as an extension for future research (Y.S Cho, P. C, Hong, 2023); (Muhammad A. B., & Mukhtar I. B., 2023).

2. Benefits of AI in the Health Sector

AI is currently being extensively tested in hospitals for drug discovery as well as diagnosis and symptom prediction. Here are some of its most promising prospects:

1. Diagnostic Assessment

Electronic health records (EHRs), radiography, CT scans, and magnetic resonance images all produce large volumes of data that AI can analyze. AI systems can assist with early symptom forecasts by analyzing data from patients, identifying trends, and identifying relationships.

2. Virtual Health Assistants

Virtual health assistants are in charge of carrying out a range of duties, including returning normal patient calls and emails, monitoring medical records, safeguarding private patient information, setting up doctor appointments, and reminding patients to schedule follow-up appointments. Because it offers patients a personalized experience in managing their health and responding to their questions, it is one of the most beneficial AI applications in healthcare.

3. Treatment of Rare Diseases

In order to speed up the discovery and development of cutting-edge breakthrough medications and vaccines and revolutionize the delivery of healthcare, BERG, an AI-based clinical-stage biotech platform, aims to map diseases. It combines interrogative biology with research and development (R&D) to enable medical professionals to create durable products for people battling rare diseases.

4. Targeted Treatment

Benevolent AI, a prominent clinical-stage AI-enabled drug development business, was able to offer the proper

treatment to the required patients at the correct time with the use of technologies like deep learning and AI, leading to tailored treatment of patients with helpful insights. The company is currently focused on developing portable remedies for uncommon diseases and securing licensing for its medications.

5. Drug Discovery

Artificial intelligence uses neural networks to evaluate drug candidates' characteristics and bioactivity. With the aid of AI systems, researchers can determine the optimal therapeutic targets to investigate for specific diseases. The healthcare sector has seen an increase in speed and a decrease in investment in drug discovery as a result. It has proven important in clinical trials in the selection of the correct candidates (https://emeritus.org/blog/healthcare-challenges-of-ai-in-healthcare/).

3. Applying AI in the Healthcare Sector: Difficulties

1. Data Privacy and Security: Large volumes of patient data are needed for the application of AI in healthcare, which presents issues with data security and privacy. Ensuring that patient data is safeguarded against unauthorized access and that patients have authority over its usage is crucial.

2. Bias in the Data: If the training data for AI systems is not representative of the people they will be used to assist, then the systems may become prejudiced. This could provide unfair or erroneous results, especially for underrepresented communities.

3. Lack of Transparency: Because it might be challenging to decipher how an AI system arrived at a given choice, many of them are referred to as "black boxes". Physicians and other medical experts may find it challenging to trust the outcomes of an AI system due to this lack of transparency.

4. Regulation and Governance: At the moment, there aren't many precise laws and policies governing the application of AI in healthcare. Patients may find it challenging to know what to anticipate when interacting with an AI system, and healthcare institutions may find it challenging to employ the technology appropriately.

5. Lack of Understanding: It's possible that many patients and healthcare professionals are unaware of AI's limitations and how it functions. This may result in irrational expectations and misplaced faith in technology (https://www.forbes.com/sites/forbesbusinesscouncil/2023/02/07/top-five-opportunities-and-challenges-of-ai-in-healthcare/?sh=e1fb69d28056).

4. Conclusion

The use of machine learning techniques in the healthcare sector to manage vast amounts of patient data and reduce the time, cost, and resources required for its analysis is discussed in this article. These technologies are currently in high demand since they have been shown to improve patients' hospital experiences and be more accurate than diagnosis made by licensed doctors. It was also underlined that prior to these tools being utilized for their primary function, training is necessary, as they are meant to function in tandem with doctors. The benefits of AI in the realm of medicine are emphasized (M.A. Baballe, S. H. Ayagi, & U. F. Musa., 2023). The difficulties of applying AI are also well covered.

References

A. H. Muhammad, A. Y. Abdullahi, A. Abba, A. Isah, A. A Yako, & M. A. Baballe, (2022). The Benefits of Adopting a Wireless Nurse Call System. *Global Journal of Research in Medical Sciences, 2*(3). Retrieved from https://gjrpublication.com/gjrms/

A. Naveed (2023). Transforming Healthcare through Arti□cial Intelligence and Machine Learning. *Pakistan Journal of Health Sciences, 4*(5). https://doi.org/10.54393/pjhs.v4i05.844

Alpaydin, E. (1997). Voting over Multiple Condensed Nearest Neighbors. *Artificial Intelligence Review, 11,* 115-132. https://doi.org/10.1023/A:1006563312922

Artificial Intelligence Index Report 2022. Chapter 4: The Economy and Education. 4.2 Investment. Retrieved from https://aiindex.stanford.edu/wpcontent/uploads/2022/03/2022-AI-Index-Report_Master.pdf

Brown, C., & O'Leary, D. (1995). Introduction to artificial intelligence and expert systems. Artificial Intelligence/Expert Systems Section of the American Accounting Association.

Cunningham, P., & Delany, S. (2007). k-Nearest neighbour classifiers. *Mult Classif Syst, 54.* https://doi.org/10.1145/3459665

F. Del Giorgio Solfa, & F. R. Simonato (2023). Big Data Analytics in Healthcare: Exploring the Role of Machine Learning in Predicting Patient Outcomes and Improving Healthcare Delivery. *International Journal of*

Computations Information and Manufacturing (IJCIM). https://doi.org/10.54489/ijcim.v3i1.235

Fazliddinovich, R. M., & Abdumurodovich, B. U. (2017). *Parallel processing capabilities in the process of speech recognition.* 2017 International Conference on Information Science and Communications Technologies (ICISCT), pp. 1-3. https://doi.org/10.1109/ICISCT.2017.8188585

https://emeritus.org/blog/healthcare-challenges-of-ai-in-healthcare/.

https://www.cdc.gov/nchs/fastats/heart-disease.htm

Huang, N., Ibrahim, Z., & Diah, N. (2021). Machine Learning Techniques for Heart Failure Prediction. *Malaysian Journal of Computing, 6,* 872. https://doi.org/10.24191/mjoc.v6i2.13708

Jabbar, M. A., Deekshatulu, B. L., & Priti, C. (2013). Heart disease classification using nearest neighbor classifier with feature subset selection. Annals Computer Science 2013.

Kashyap, A. (2018). *Artificial Intelligence & Medical Diagnosis, 6,* 4982-4985.

Khdier, H., Jasim, W., & Aliesawi, S. (2021). Deep Learning Algorithms based Voiceprint Recognition System in Noisy Environment. *Journal of Physics: Conference Series, 1804,* 012042. https://doi.org/10.1088/1742-6596/1804/1/012042

Ławrynowicz, A., & Tresp, V. (2014). *Introducing Machine Learning.*

Liu, J. Y., Kong, X. J., Xia, F., Bai, X. M., Wang, L. Qing, Q., & Lee, I. (2018). *Artificial Intelligence in the 21st Century*. IEEE Access. PP. 1-1. https://doi.org/10.1109/ACCESS.2018.2819688

M. A. Baballe, A. M. Gambale, A. S. Bari, A. S. Lawan, & R. J. Suleiman. (2022). Issues with our hospitals queue management information systems. *Global Journal of Research in Medical Sciences, 2*(6), 102-106. https://doi.org/10.5281/zenodo.7330904

M. A. Baballe, A. S. Muhammad, J. Y. Abdullahi, Aminu Ya'u,, Ibrahim Idris Giwa, M. Habib Abubakar, & Z. Abdulkadir. (2022). The Impact of Hospital Queue Management Systems. *Global Journal of Research in Medical Sciences, 2*(5), 88-91. https://doi.org/10.5281/zenodo.7117651

M. Musaev & M. Rakhimov, (2019). "A Method of Mapping a Block of Main Memory to Cache in Parallel Processing of the Speech Signal" 2019 International Conference on Information Science and Communications Technologies (ICISCT), pp. 1-4. https://doi.org/10.1109/ICISCT47635.2019.9011946

M. Musaev & M. Rakhimov, (2020). Accelerated Training for Convolutional Neural Networks. 2020 International Conference on Information Science and Communications Technologies (ICISCT), pp. 1-5. https://doi.org/10.1109/ICISCT50599.2020.9351371

M. Musaev, I. Khujayorov & M. Ochilov (2020). "Development of integral model of speech recognition system for Uzbek language," 2020 IEEE 14th International Conference on Application of Information and Communication Technologies (AICT), pp. 1-6. https://doi.org/10.1109/AICT50176.2020.9368719

M. Rakhimov & M. Ochilov, (2021). Distribution of Operations in Heterogeneous Computing Systems for Processing Speech Signals. 2021 IEEE 15th International Conference on Application of Information and Communication Technologies (AICT), pp. 1-4. https://doi.org/10.1109/AICT52784.2021.9620451

M. Rakhimov, D. Mamadjanov & A. Mukhiddinov, (2020). A HighPerformance Parallel Approach to Image Processing in Distributed Computing. 2020 IEEE 14th International Conference on Application of Information and Communication Technologies (AICT), pp. 1-5. https://doi.org/10.1109/AICT50176.2020.9368840

M. Rakhimov, J. Elov, U. Khamdamov, S. Aminov & S. Javliev (2021). Parallel Implementation of Real-Time Object Detection using OpenMP. 2021 International Conference on Information Science and Communications Technologies (ICISCT), 2021, pp. 1-4. https://doi.org/10.1109/ICISCT52966.2021.9670146

M. Rakhimov, R. Akhmadjonov, & S. Javliev, (2020). Artificial Intelligence in Medicine for Chronic Disease Classification Using Machine Learning. Retrieved from https://www.researchgate.net/publication/367194708

M. Rakhimov, T. Boburkhon & T. Khurshid, (2021). Speaker Separation: Use Neural Networks. 2021 International Conference on Information Science and Communications Technologies (ICISCT), pp. 01-03. https://doi.org/10.1109/ICISCT52966.2021.9670322

M.A. Baballe, S. H. Ayagi, & U. F. Musa. (2023). Using artificial intelligence (AI) technology in the health sector has several goals. *Global Journal of Research in Engineering & Computer Sciences, 3*(5), 31-35. https://doi.org/10.5281/zenodo.10048487.

Ma, J. (2017). Prediction of heart disease using k-nearest neighbor and particle swarm optimization. *Biomedical Research-tokyo, 28,* 4154-4158.

Mehta, V. (2023). Artificial Intelligence in Medicine: Revolutionizing Healthcare for Improved Patient Outcomes. *J Med Res Innov. 2023, 7*(2). https://doi.org/10.32892/jmri.292

Mihret, E. (2020). Robotics and Artificial Intelligence. *International Journal of Artificial Intelligence and Machine Learning.* https://doi.org/10.4018/IJAIML.2020070104

Muhammad, A. B., & Mukhtar, I. B. (2023). Artificial Intelligence in the Healthcare Sector. *Global Journal of Research in Engineering & Computer Sciences, 3*(5), 10-13). https://doi.org/10.5281/zenodo.10001767.

R. J. Woodman, & A. A. Mangoni. (2023). A comprehensive review of machine learning algorithms and their application in geriatric medicine: present and future. Aging Clinical and Experimental Research, 2023, https://doi.org/10.1007/s40520-023-02552-2

S. Brown (2021). Machine learning, explained. *MIT Sloan School Of Management, 21*(04). Retrieved from https://mitsloan.mit.edu/ideas-made-to-matter/machine

T. M. Tassew, X. Nie, (2022). A Comprehensive Review of the Application of Machine Learning in Medicine and Health Care. A Comprehensive Review of the Application of Machine Learning in Medicine and Health Care. TechRxiv. Preprint. https://doi.org/10.36227/techrxiv.21204779.v1

W. Xing & Y. Bei, (2020). Medical Health Big Data Classification Based on KNN Classification *Algorithm.IEEE Access, 8,* 28808- 28819. https://doi.org/10.1109/ACCESS.2019.2955754

Y.S Cho, & P. C, Hong, (2023). Applying Machine Learning to Healthcare Operations Management: CNN-Based Model for Malaria Diagnosis. *Healthcare (MDPI), 11,* 1779. https://doi.org/10.3390/healthcare11121779

Computer and Information Science; Vol. 16, No. 4; 2023
ISSN 1913-8989 E-ISSN 1913-8997
Published by Canadian Center of Science and Education

Reviewer Acknowledgements

Computer and Information Science wishes to acknowledge the following individuals for their assistance with peer review of manuscripts for this issue. Their help and contributions in maintaining the quality of the journal is greatly appreciated.

Computer and Information Science is recruiting reviewers for the journal. If you are interested in becoming a reviewer, we welcome you to join us. Please find the application form and details at http://recruitment.ccsenet.org and e-mail the completed application form to cis@ccsenet.org.

Reviewers for Volume 16, Number 4

A. K. Sharma, National Taipei University of Business (NTUB), India

Afsana Ahamed, Arkansas Tech University, USA

Antoanela Luciana Naaji, "Vasile Goldis" Western University of Arad, Romania

Elmaallam Mina, Mohammed V University, Morocco

Franjieh El Khoury, Polytechnique Montreal, Mobile Computing and Networking Research Laboratory “Laboratoire de Recherche en Réseautique et Informatique Mobile, Canada

Frankline Makokha, University of Nairobi, Kenya

Gomathy.C.K., SCSVMV University, India

Leo John, J.J College of Arts and Science(Autonomous), India

Mohammad Taleghani, Islamic Azad University, Iran

Rushit Dave, Minnesota State University, USA

Srujan Kotikela, Texas A&M University, USA

Uchenna Nzenwata, Babcock University, Ilisan Remo, Ogun State Nigeria, Nigeria

Zhihao Xu, Qingdao University, China

ACADEMIC AND OPEN ACCESS

Canadian Center of Science and Education

CCSENET.ORG

➢ CALL FOR MANUSCRIPTS

Computer and Information Science (CIS) is a peer-reviewed, open access journal, published by Canadian Center of Science and Education. It publishes original research and applied articles in all areas of computer and information science. Authors are encouraged to submit complete unpublished and original works, which are not under review in any other journals. We are seeking submissions for forthcoming issues. All manuscripts should be written in English. Manuscripts from 3000–8000 words in length are preferred. All manuscripts should be prepared in MS-Word or Latex format, and submitted online, or sent to: cis@ccsenet.org

Paper Selection and Publishing Process

a) Submission acknowledgement. If you submit manuscript online, you will receive a submission acknowledgement letter sent by the online system automatically. For email submission, the editor or editorial assistant sends an e-mail of confirmation to the submission's author within one to three working days. If you fail to receive this confirmation, please check your bulk email box or contact the editorial assistant.

b) Basic review. The editor or editorial assistant determines whether the manuscript fits the journal's focus and scope. And then check the similarity rate (CrossCheck, powered by iThenticate). Any manuscripts out of the journal's scope or containing plagiarism, including self-plagiarism are rejected.

c) Peer Review. We use a double-blind system for peer review; both reviewers' and authors' identities remain anonymous. The submitted manuscript will be reviewed by at least two experts: one editorial staff member as well as one to three external reviewers. The review process may take four to ten weeks.

d) Make the decision. The decision to accept or reject an article is based on the suggestions of reviewers. If differences of opinion occur between reviewers, the editor-in-chief will weigh all comments and arrive at a balanced decision based on all comments, or a second round of peer review may be initiated.

e) Notification of the result of review. The result of review will be sent to the corresponding author and forwarded to other authors and reviewers.

f) Pay the article processing charge. If the submission is accepted, the authors revise paper and pay the article processing charge (formatting and hosting).

g) E-journal is available. E-journal in PDF is available on the journal's webpage, free of charge for download. If you need the printed journals by post, please order at http://www.ccsenet.org/journal/index.php/cis/store/hardCopies.

h) Publication notice. The authors and readers will be notified and invited to visit our website for the newly published articles.

More Information

E-mail: cis@ccsenet.org

Website: http://cis.ccsenet.org

Paper Submission Guide: http://cis-author.ccsenet.org

Recruitment for Reviewers: http://www.ccsenet.org/journal/index.php/cis/editor/recruitment

➢ JOURNAL STORE

To order back issues, please contact the editorial assistant and ask about the availability of journals. You may pay by credit card, PayPal, and bank transfer. If you have any questions regarding payment, please do not hesitate to contact the editorial assistant.

Price: $40.00 USD/copy Shipping fee: $20.00 USD/copy

CONTACT US

1595 Sixteenth Ave, Suite 301
Beaver Creek, Ontario, L4B 0A9
Canada
Tel: 1-416-642-2606
E-mail: info@ccsenet.org
Website: www.ccsenet.org

The journal is peer-reviewed
The journal is open-access to the full text
The journal is included in:

DBLP (2008-2019)
ERA
Google Scholar
Index of Information Systems Journals
LOCKSS
Infotrieve
SHERPA/RoMEO
Standard Periodical Directory
Ulrich's
Universe Digital Library
WJCI Report
WorldCat

Computer and Information Science

Quarterly

Publisher Canadian Center of Science and Education
Address 1595 Sixteenth Ave, Suite 301, Richmond Hill, Ontario, L4B 3N9, Canada
Telephone 1-416-642-2606
E-mail cis@ccsenet.org
Website http://cis.ccsenet.org

ISSN 1913-8989