Annals of Emerging Technologies in Computing (AETiC)
Vol. 7, No. 1, 2023

Research Article

A Predictive Cyber Threat Model for Mobile Money Services

Mistura Laide Sanni[1], Bodunde Odunola Akinyemi[1,*], Dauda Akinwuyi Olalere[2], Emmanuel Ajayi Olajubu[1] and Ganiyu Adesola Aderounmu[1]

[1]Obafemi Awolowo University, Ile-Ife, Nigeria
msanni@oauife.edu.ng; bakinyemi@oauife.edu.ng; emmolajubu@oauife.edu.ng; gaderoun@oauife.edu.ng
[2]MTN, Nigeria
DaudaO@mtnnigeria.net
*Correspondence: bakinyemi@oauife.edu.ng

Received: 21st March 2022; Accepted: 22nd December 2022; Published: 1st January 2023

Abstract: Mobile Money Services (MMS), enabled by the wide adoption of mobile phones, offered an opportunity for financial inclusion for the unbanked in developing nations. Meanwhile, the risks of cybercrime are increasing, becoming more widespread, and worsening. This is being aggravated by the inadequate security practises of both service providers and the potential customers' underlying criminal intent to undermine the system for financial gain. Predicting potential mobile money cyber threats will afford the opportunity to implement countermeasures before cybercriminals explore this opportunity to impact mobile money assets or perpetrate financial cybercrime. However, traditional security techniques are too broad to address these emerging threats to Mobile Financial Services (MFS). Furthermore, the existing body of knowledge is not adequate for predicting threats associated with the mobile money ecosystem. Thus, there is a need for an effective analytical model based on intelligent software defence mechanisms to detect and prevent these cyber threats. In this study, a dataset was collected via interview with the mobile money practitioners, and a Synthetic Minority Oversampling Technique (SMOTE) was applied to handle the class imbalance problem. A predictive model to detect and prevent suspicious customers with cyber threat potential during the onboarding process for MMS in developing nations using a Machine Learning (ML) technique was developed and evaluated. To test the proposed model's effectiveness in detecting and classifying fraudulent MMS applicant intent, it was trained with various configurations, such as binary or multiclass, with or without the inclusion of SMOTE. Python programming language was employed for the simulation and evaluation of the proposed model. The results showed that ML algorithms are effective for modelling and automating the prediction of cyber threats on MMS. In addition, it proved that the logistic regression classifier with the SMOTE application provided the best classification performance among the various configurations of logistic regression experiments performed. This classification model will be suitable for secure MMS, which serves as a key deciding factor in the adoption and acceptance of mobile money as a cash substitute, especially among the unbanked population.

Keywords: ***Cyberspace; Machine Learning; Mobile Money; Predictive Model; Threats***

1. Introduction

Innovations such as mobile money, enabled by the proliferation of mobile phones, have facilitated an exceptional opportunity for the financial inclusion of a large number of unbanked populations in developing nations [1]. Due to a lack of financial incentives, traditional nationalised banks do not have branches in villages. Furthermore, with the ubiquitous presence of mobile devices, the number of connections to cyberspace has astronomically increased. The number of World Unbanked Adults (WUA) is 1.7 billion, with nearly half (46%) living in less developed countries, 80% of whom are Sub-Saharan Africans

Mistura Laide Sanni, Bodunde Odunola Akinyemi, Dauda Akinwuyi Olalere, Emmanuel Ajayi Olajubu and Ganiyu Adesola Aderounmu, "A Predictive Cyber Threat Model for Mobile Money Services", *Annals of Emerging Technologies in Computing (AETiC)*, Print ISSN: 2516-0281, Online ISSN: 2516-029X, pp. 40-60, Vol. 7, No. 1, 1st January 2023, Published by International Association for Educators and Researchers (IAER), DOI: 10.33166/AETiC.2023.01.004, Available: http://aetic.theiaer.org/archive/v7/v7n1/p4.html.

[2]. However, mobile phone penetration is rapidly increasing in these countries[1]. Hence, Mobile Financial Services (MFS) applications are among the most promising mobile applications in the developing world [3-4]. The advent of and increased access to mobile devices have created opportunities for various self-service innovations such as MFS solutions, mobile money, and mobile commerce in cyberspace. Such innovations have helped to provide financial instruments to many unbanked populations in the financial systems of third world countries and, as such, have been a major contributor to the financial inclusion of the unbanked in these emerging markets [5-6]. Consequently, Sub-Saharan Africans are responsible for 75% of global Mobile Money Service (MMS) transactions[2]. In Figure 1, the adoption of MMS by the unbanked is exemplified.

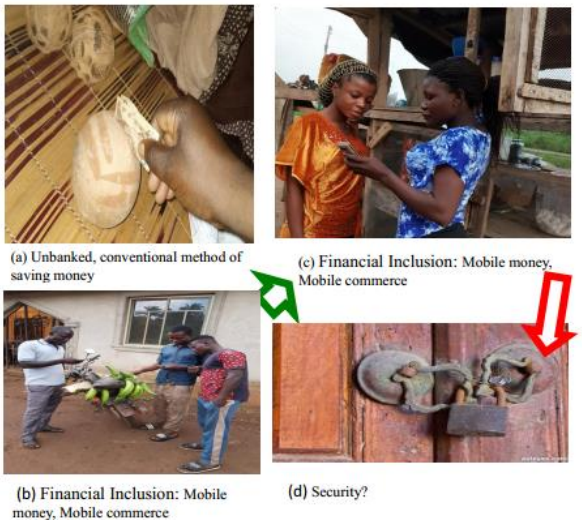


Figure 1. Mobile Money Services (MMS) and financial inclusion (a) the unbanked conventional way of saving, (b) &(c) Financial inclusion, unbanked mobile commerce with Mobile Money at Mokore Farm Settlement, Orile-Owu, Osun state, Nigeria, (d) Security concerns.

Mobile money is an excellent alternative for bridging the Financial Inclusion (FI) gap in mobile commerce [7]. According to [7], MMS must be used for money transfers, making and receiving payments via mobile phone, unbanked accessibility, providing a network of physical transactional points (e.g., agents) outside of bank branches and ATMs, and while MMS exclude mobile banking or payment services (such as Apple Pay and Google Wallet). Despite the potential opportunities of MMS in terms of adoption, the fear of losing money to cybercriminals remains a major concern among customers. This innovation is supposed to be widely adopted, but the perceived trust and security awareness of the service have remained the principal adoption determinants for this new innovation [8-9]. Knowledge related to the security of the environment and the framework to uncover and detect mobile money cyber threats in developing nations is underrepresented in the literature [10]. Furthermore, fraudsters are getting more innovative and finding loopholes in new security controls very quickly. The risks of cybercrime are increasing, widespread, and exacerbating. This is being aggravated by the poor security practises of both service providers and the attendant criminal mind-set of many of the customers or potential customers whose goal is to compromise the system for financial gain.

Traditional threat modelling and standard security requirements for mobile payment solutions such as mobile money for the unbanked are no longer effective and comprehensive enough to curb cybercrime because they are based on standard checklists (e.g., PCI-DSS, ITSEC) and implement standard protocols (e.g., SSL, DNSSEC). Thus, security measures are limited to the implementer's responses to each checklist item and the standard security requirements, which have been flawed by the increasing trend of fraud-related cases on MFS even after meeting the standard security requirements. As a result, the ability of tools, methods, or models to automate the prediction of these cyber threats would be useful in addressing MMS's

[1] Brahima Sanou, ICT Facts and Figures 2017, International Telecommunication Union (ITU), pp. 1-8, 2017, Available: https://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2017.pdf.

[2] Aramé Awanis, Christopher Lowe, Simon K. Andersson-Manjang and Dominica Lindse, State of the industry report on mobile money, Global System for Mobile Communications Association (GSMA), London, Available: https://www.gsma.com/sotir/wp-content/uploads/2022/03/GSMA_State_of_the_Industry_2022_English.pdf.

cyber threat challenges, because anything that can be predicted is owned, and anything owned can be decided on and the desired action can eventually be taken as desired. Having models that can predict potential mobile money cyber threats will provide an opportunity to take action on such threats before cybercriminals exploit them to impact mobile money assets or perpetrate financial cybercrime. An ability to predict a cyber-threat event in mobile money cyberspace helps to take ownership of the decision in order to take the necessary actions. Although joint efforts from industry and government stakeholders have culminated in the publication of standards, frameworks, and guidelines, for example, by the National Institute of Standards and Technology, to mitigate the risks of cybercrime, the wave of increased MFS security is still on the rise. For example, in Nigeria, MFS fraud cases increased by 3,015% between 2015 and 2016; Nigeria lost N12.3 billion between 2014 and 2017. Also in Nigeria, MFS fraud in 2018 was the highest in the last four years [11].

In many of the successful cyberattacks against MFS, the role of humans cannot be overemphasized. They can be the originator, the medium, or the actual executor of the attack. Hence, MMS providers' methods or processes of onboarding, modifying, and terminating have important security implications. If a customer's intention could be predicted from the information supplied at any of the customer management process stages, such as customer onboarding or Subscriber Identity Module (SIM) registration, modification, or termination, it would help beef up MMS security. Meanwhile, some organizations have implemented a second stage of validation, such as manually going through each customer record (eyeballing) after registration to determine whether or not the customer has fraudulent intent before activating the MMS. For instance, most developing nations' MMS providers use the mandatory SIM registration as KYC (Know Your Customer). The second-stage validation process is tedious, ineffective, and inefficient. Therefore, for human-vectored cyber threat prevention to be effective, countermeasures must be robust and intelligent enough to predict and prevent it [12]. There is a need to focus on the on-boarding stage in MMS activation to build predictive models that would detect and prevent cyber threats vectored via the on-boarding process without the need for manual human checks to prevent fraudulent customer on-boarding. Hence this study.

The remainder of the paper is divided into the following sections: Section 2 discusses related works, while Section 3 presents the ML technique used in identifying and forecasting cyber threats associated with MMS. Section 4 covered the findings, and Section 5 gave the conclusion.

2. Related Works

Secure MMS is a major determinant of mobile money adoption [9]. Because of the rapidly increasing use of the World Wide Web and the Internet nowadays, there is an increase in the volume and complexity of MMS insecurities. The upsurge of security flaws has significantly deteriorated the Quality of Services (QoS) of the MFS. Critical cyber security issues relating to mobile payment systems include identity theft, agent-driven fraud, sharing personal identification numbers (PINs), phishing, vishing, and authentication attacks [13-14]. Several fraud cases that threaten the security of MMS include false transactions and the misuse of PINs [15]. The use of better access controls, customer awareness campaigns, agent training on acceptable practises, strict measures against fraudsters, service providers' monitoring of high-value transactions, and the creation of an extensive legal document to operate MMS, among others, were some of the proposed mitigation measures in the literature. Studies found that MMS operators are aware of the need to improve mobile money security and such improvements will enable operators to protect themselves, their customers, and agents and assist in the successful provision of MMS. It was stressed that a mobile money-enabling environment should be properly regulated to avoid any potential risks. Thus, the user management or activation process of MMS for customer onboarding, for example, via SIM registration, requires thorough regulatory monitoring and considerable research attention to uncover the risks inherent in the process.

Studies on mobile money financial crimes have also been conducted[3], with the goal of providing guidelines on regulatory policy and frameworks [16-17]. It was established that the regulations to fight

[3] John Villasenor, Smartphones for the Unbanked: How Mobile Money Will Drive Digital Inclusion in Developing Countries, Issues in Technology Innovation, No. 25, September 2013, Published by The Brookings Institution, Available: https://innovation.luskin.ucla.edu/sites/default/files/VillasenorMobileMoney.pdf.

crime should not impede MMS adoption, but instead adapt the traditional financial systems of combating crime to the mobile money industry in appropriate ways. The use of "frameworks, standards, and countermeasures" for MFS to provide mechanisms for mitigating cybercrime threats has a lot of challenges, as no workable solution has specifically been provided, especially in the context of developing nations [11].

There have been few studies that have focused on formal methodologies for threat modelling of network systems, such as mobile money solutions, as available for specific software systems [18]. For threat modelling purposes, a network system is viewed via a network model for threat analysis, which allows analysts to determine communications between computers with different roles [18-20]. While threat modelling of network-based solutions and its methodology are relatively scarce in the literature, it is more common to find works that provide threat models for specific software applications. A software application threat model can be modelled using a Data Flow Diagram (DFD) to describe the system. Examples of such targeted, specific application threat modelling include threat analysis of on-line banking systems by combining the STRIDE threat model and the threat tree analysis [21], the threat analysis of Web services and grids [22-23], and the threat modelling of identity federation protocols [24-25]. A goal-oriented approach to security threat modelling and analysis has been applied to model different systems, for instance, using visual model elements to explicitly capture threat-related concepts [26].

In mobile money solution services, research has been conducted on strengthening MFS technical security countermeasures [27-29] and improving MFS security [30]. Some of these techniques are structural equation modelling [31], biometric techniques [32], two-factor authentication [8], quantitative analysis of subject matter experts (SME) [33], and a host of others. Biometric techniques were proposed for providing the highest security to mobile payments in e-banking, particularly at the wireless transmission level. In the model, the image of a fingerprint is captured in real time and sent to the server for authorization. A fuzzy logic-based fingerprint matching algorithm was used on the server side for authorization [32]. The detection of fraud in mobile banking was also investigated with user input patterns when mobile banking services are being used, as well as the transaction pattern. The study's findings revealed that user input and transaction pattern data contain information that can be used to identify a specific user, allowing abnormal transactions to be detected [34].

A probabilistic-based model that was leveraged for the formulation of a mathematical derivation for an adaptive threshold algorithm for detecting anomalous transactions was reported in [35]. The model was optimized with Baum-Welsh and hybrid Posterior-Viterbi algorithms. A credit card transaction dataset was simulated, trained, and predicted for fraud. And finally, the proposed model was evaluated using different metrics. The results showed that the detection model performed well for credit card anomalous transaction detection; however, this has not been established for MMS. A framework was used to study the banking environment's information systems as they relate to information security initiatives; the case study used the Kenya banking sector [36]. The research objectives were to identify common banking information system vulnerabilities; analyse and define gaps in existing frameworks in order to evaluate banking programme initiatives and security; develop a framework for use in evaluating security programmes for the banking industry; and validate the developed security investment framework. The findings revealed that people pose the greatest threat to information systems, and customer security awareness was identified as a major barrier to security effectiveness. The increased risk exposure in banks was also traced to fraud, careless or unaware employees, and internal attacks, which were cited as the causes. The study concluded that people, process, and technology alignment are very important in transforming an organisation's information security.

Meanwhile, the use of all these traditional mechanisms for preventing fraud in MMS is not effective. They do not provide effective security as cybercrime issues persist in MFS, and when such security serves as the last resort, user experiences are often impacted [37]. The frameworks, standards, and countermeasures for MFS to provide mechanisms for mitigating cybercrime threats have a lot of challenges as no workable solution has been specifically provided, especially in the context of developing nations, which showed in the survey conducted in [30] that MFS was the least preferred method of payment compared to instruments like payment cards and cheques [12]. Meanwhile, there is a dearth of information in the literature about the AI-based detection and prevention models for cyber threats associated with MMS. The developing world security issues for MFS are peculiar, and the security issues are not well focused in the literature [38].

Since the use of Artificial Intelligence (AI) in cybersecurity has become ubiquitous, it may be trained to create threat warnings, recognise novel malware strains, and safeguard sensitive data for organisations in different domains. Therefore, defending mobile money solution services against cyber threats in real-time using AI-based cyber threat detection and prediction models is imperative.

Thus, an attempt is made in this study to develop a predictive cyber threat model for MMS by employing ML techniques.

3. Methodology

This work focused on predictive models for MMS cyber threats (detection and prevention) vectored via the customer life cycle management process in developing nations, using Nigeria as a case study. The study focused only on mobile phone subscriber biodata registration details for MMS, with respect to the following research questions: What security threats are involved in customer management across different mobile money solution ecosystems? And how can predictive models be built to detect and prevent these cyber threats based on intelligent software defence mechanisms?

Therefore, a ML algorithm was employed for the formulation of the predictive model for cyber threat detection and prevention.

3.1. Architectural Modelling

Modelling the conceptual view of the mobile money customer life-cycle management process from information gathered from technical interviews revealed that the mobile money customer onboarding process was an integral part of the customer life cycle and security management. In this study, the customer management life cycle was defined as comprising three basic activities. Customer creation (C), Modification (M), and Blocking or Termination (B), or CMB. These were defined as follows with respect to the totality of MMS subscribed to by the customer at a given instance of time.

Customer Creation (C): encompasses the entire onboarding or customer setup process, from SIM registration to MMS activation on the mobile money system. This study was focused on customer creation or customer registration, otherwise called Subscriber Identity Module (SIM) registration processes, which are vectored cyber threats.

Modification (M): is any change to a customer's profile in the system, such as a Subscriber Identification Module (SIM) exchange for a customer, also known as a SIM swap.

Blocking or Termination (B): also known as customer service termination or removal. The customer termination function disconnects the customer's MMS from the system. For financial inclusion for unbanked customers, mobile money heavily relies on customer SIM registration details and is used for "Know Your Customer" (KYC).

To ease cyber threat challenges in the existing model, i.e., the eyeballing model shown in Figure 2, the proposed predictive model architecture shown in Figure 3 was developed to utilize ML techniques as well as intelligent software agents to help in notifying system administrators of detections of cyber threats. A ML algorithm was used to detect anomalies in customer biodata registration records for both new and existing customers during SIM registration and MMS activations. The algorithm (supervised ML) was used for model prediction to flag cyber threats or anomalous customer data records. From Figure 3, the customer creation process flow can be summarised as follows.

Step 1: Customer approaches a Subscriber Identity Module (SIM) registration agent to purchase a SIM and requests SIM registration from a GSM service provider's designated agents. This registration is activated on the GSM network and stored in the customer database or Customer Relationship Management (CRM) database. A customer may be an adversary or a genuine customer.

Step 2: The customer dials the USSD code for mobile money registration after SIM activation on the GSM network.

Step 3: The mobile money system then pulls the KYC information from the customer's database to fulfil the requirements for registering the customer for MMS.

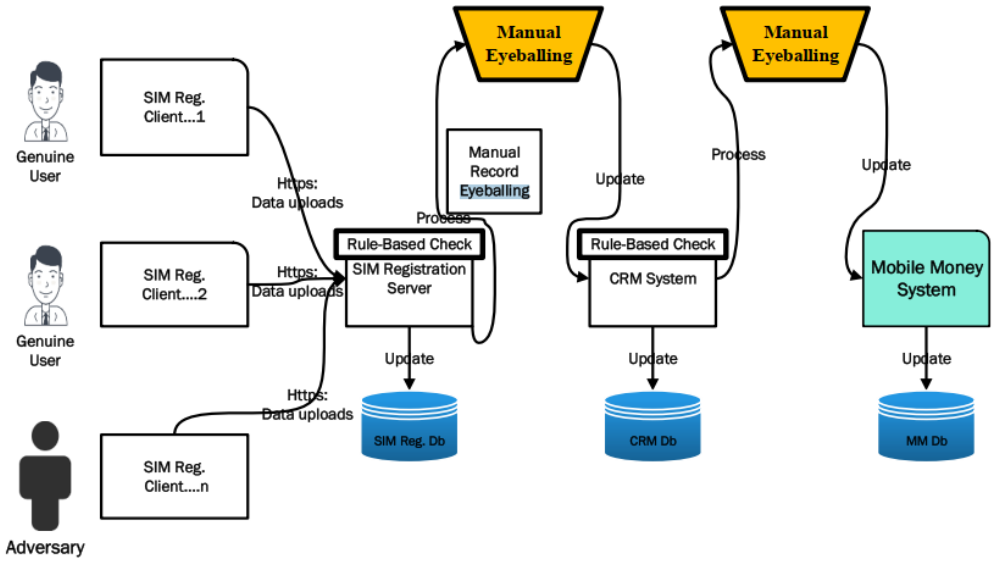


Figure 2. The existing model-eyeballing method

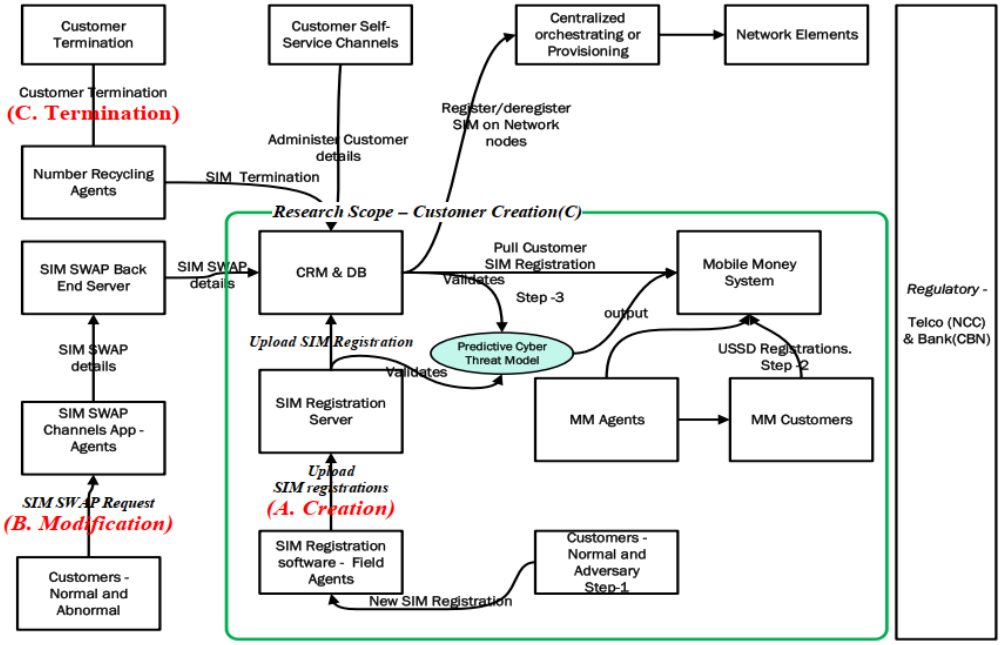


Figure 3. Block diagram of proposed architecture for mobile money customer life-cycle management

3.1.1. Mathematical Model Interpretation of the Conceptual View

Translating the mobile money customer life-cycle management view in Figure 3 into a mathematical model, the following was deduced:

Given Customer Creation = C, Customer Modification = M, Customer Blocking/Termination = B.

Taking the components as threat vectors, the cyber threats induced by the adopted customer management approach for Mobile Money System can be expressed as an overall threat profile, denoted as P_t, and summarised mathematically as Equation 1 as follows:

$$P_t = f(C_t M_t B_t) \quad (1)$$

Where:

P_t = the threat profile for the mobile money solution.

C_t = the threat profile elicited by the customer creation and SIM registration processes.

M_t = the threat profile induced by the customer modification process, e.g., SIM swaps.

B_t = threat profile resulting from a customer service blockage or termination process.

If the function in Equation 1 is subjected to a continuous probability density distribution, then the threat profile probability can assume a non-negative value of a to b, where a=0 to b=1.

3.1.2. Description of the Proposed Model

In the proposed model shown in Figure 4, the analytics module (i.e., the ML module) examined the incoming SIM registration data records for mobile money applicants, existing or new, and classified them based on cyber threat infections. In this operation, clean records were flagged as compliant, while unclean records were flagged as non-compliant, and both were eventually stored in a permanent repository. In real-life operations, this is usually the organization's Customer Relationship Management (CRM) database repository.

The second ML analytics module is another layer for deeper analysis, and this module scans and classifies both new and existing customers for mobile money eligibility. The system administrator reviews the classified records and/or updates the rules database as required.

The predictive model was formulated to check incoming online applicants' registration data in real-time for cyber threats. A supervised ML algorithm was employed to determine whether a mobile money applicant's incoming registration or activation data record detail is a legitimate transaction or not. The classification algorithm works by classifying applicants' registrations into compliant (non-fraudulent) and non-compliant (fraudulent) records. The non-compliant records are the suspicious records for cyber threats based on predictive ML algorithms. If an applicant's records are compliant, the MMS is activated; if not, it is flagged as an anomalous registration, and the customer is rejected for MMS activation. The processed applicants' records were subsequently added to the historical data records database for future algorithm learning, and the intelligent agents logged the anomalous registrations. The flow diagram is shown in Figure 5.

3.1.3. Framework for Implementation of the Predictive Analytical Model

An implementation design of the analytical model is presented in Figure 6. The model design framework would detect mobile money application fraud in real time using ML algorithm models. This would detect anomalous transactions from incoming mobile money applicants' biodata registration details. This framework uses the Apache Spark stack, that is, the spark streaming module for collecting online registration data and the spark ML module for building, training, and retraining of the predictive model. ML packages (*scikit* learn) in the Python programming language contain packages to train and re-train ML algorithms to build the predictive model. This ensures that the model is updated in real time, so that real-time registration data analysis can be performed and fraudulent customer registrations can be flagged and rejected Data from different sources is meant to be pre-processed into a format that ML algorithms can work on. The historical data for SIM registrations with the right predictive features selected and stored in a database is used by ML algorithms to process and generate the predictive models.

3.2. Dataset Collection, Pre-processing and Analysis

Five (5) million dataset records of applicants' registration for MMS were gathered from the mobile money practitioners' interviews with the Nigerian Telecoms on the issues with the registration of customers' data for the purpose of mobile money registration. Because providing customer transaction details is considered a breach of confidentiality, the practitioners masked and transformed the majority of the features in the dataset. The sample masked registrations for valid and invalid registrations were obtained, and the generated dataset was highly imbalanced as there were more valid registrations than invalid registrations.

3.2.1. Data Preprocessing

The following pre-processing steps were carried out:

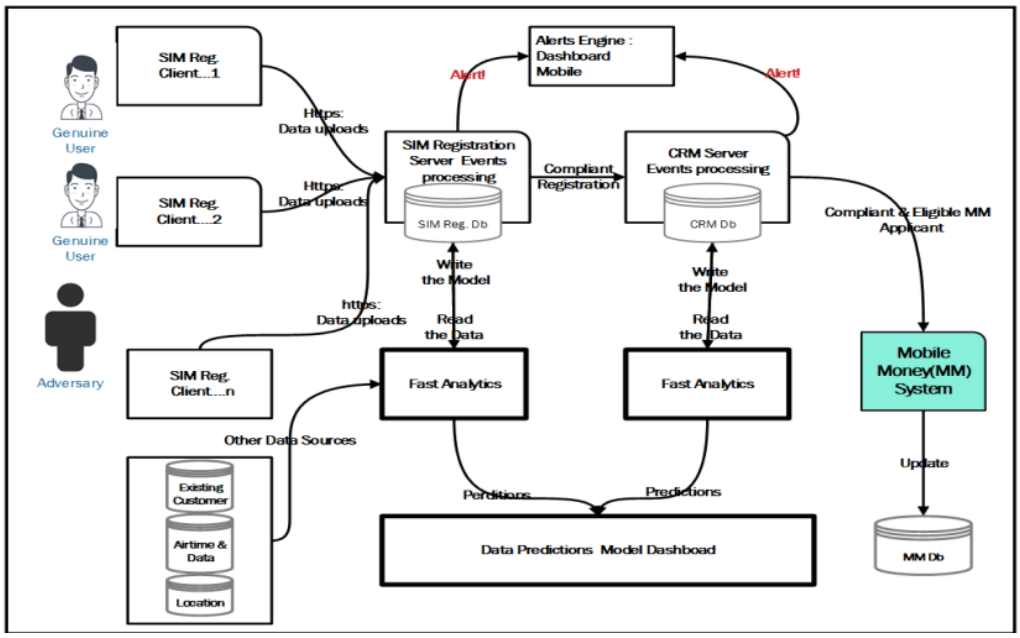


Figure 4. Proposed cyber threat predictive model

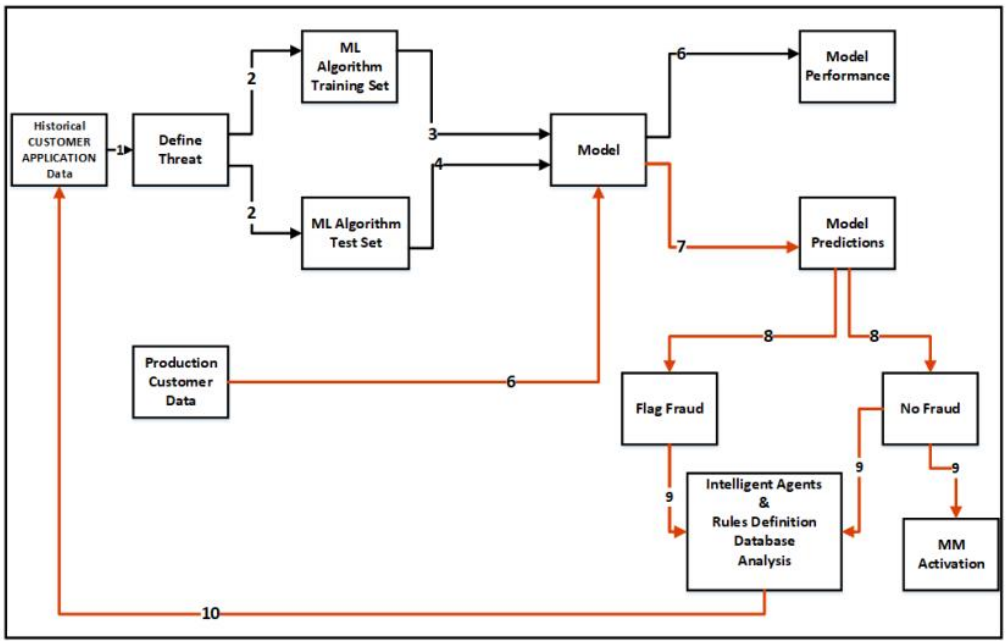


Figure 5. Design flow of the proposed model

Data Masking and Vectorization: To maintain the confidentiality of user information, the string data type in the data records was masked and vectorized into numeric values. Masking entailed concealing the customer's real phone number as well as personal information such as their address and zip code.

Dataset creation and cleaning: The datasets were generated based on the features in the sample data using the python faker library, *Faker().* Faker is a Python library that generates fake data to anonymize data taken from a production service for confidentiality reasons or to generate large quantities of data. Irrelevant features and those with null values were removed.

Features Selection: Regardless of the classification algorithm used, a feature selection procedure was performed on all factors suggested by mobile money technical and business experts as the most likely to affect the fraudulent behaviours of mobile money applicants. These resulted in 13 factors for each applicant.

Surname, first name, gender, mother's maiden name, region, customer reputation, agent reputation, and agent identity are all examples.

Bag-of-Words (BoW): The dataset features comprise many strings or character data types, such as applicants' names, regions, or addresses, that are not usable by ML algorithms, which only work on numeric data types. Hence, the BoW approach was used to convert string features to numeric representation.

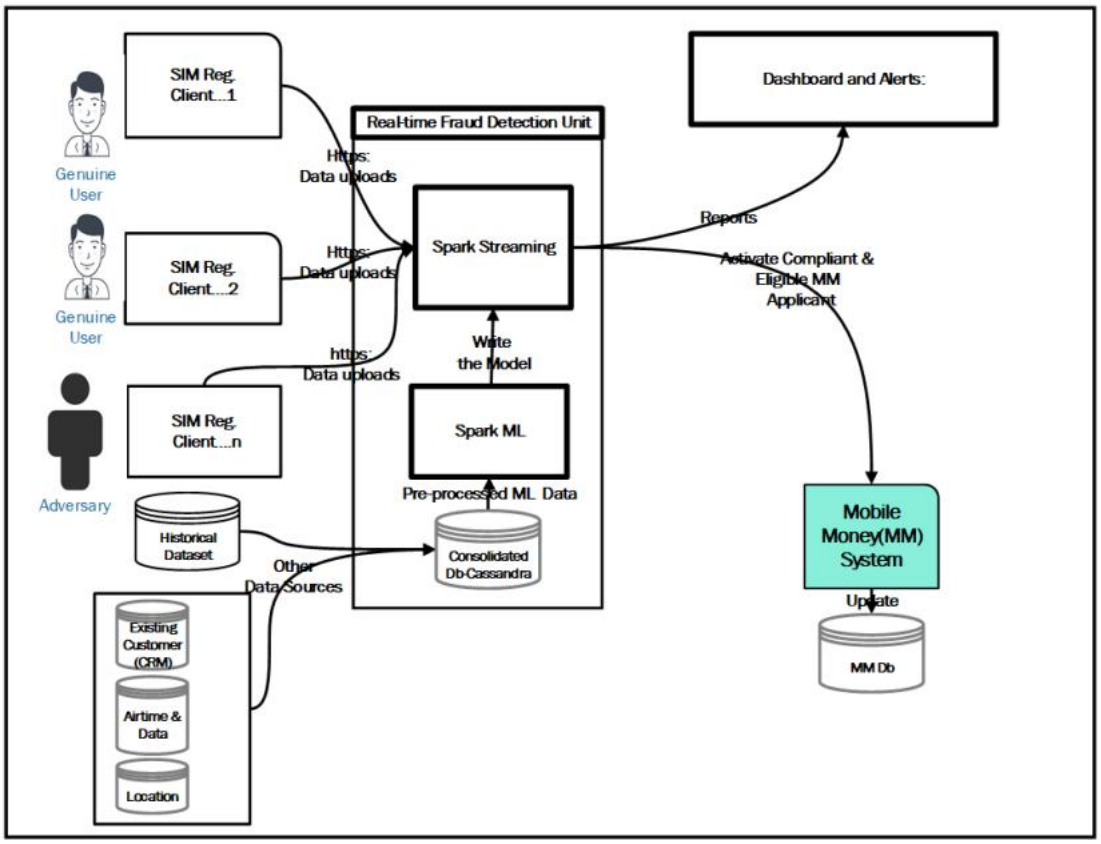


Figure 6. A framework of the proposed model

Classification Rules: Rules were developed for different predictive indicators in the dataset features for proper labelling of the dataset to label the applicant's historical records or the dataset as fraudulent or non-fraudulent. The term "non-fraudulent" refers to mobile money applicant records that are valid and compliant for MMS activation. The fraudulent records were grouped into two categories in relation to the cyberthreat risk potential of the customer, i.e., high and low.

3.2.2. Handling the Data Imbalance Problem

It was noted that the number of observations for the majority and minority classes in the acquired dataset was not equally distributed. Random oversampling could lead to an overfitting problem and, consequently, biased classification. In order to avoid a class imbalanced classification, the widely used Synthetic Minority Over-sampling TEchnique (SMOTE) [39], which is an oversampling approach that creates a synthetic minority class, i.e., synthesizes new minority class samples, was employed. This is accomplished by concentrating on the feature space and interpolating the positive instances that span together.

3.2.3. Historical Records Labelling, Fraud Score Calculation, and Risk Categorisation Design

For predictive classification, the dataset records were labelled as mostly "non-fraudulent intent" and "fraudulent intent" applicants. The fraudulent category was further broken down into two categories based on risk profile: high and low, as defined in Table 1. The table shows the fraud score and category based on the weight of fraud intent, which impacts the final determination of the status of the customer record as follows:

Fraud Rating: A fraud rating was assigned per issue rule per feature in the applicant's data record for likely potential cyber threat issues.

Fraud Score*:* Each record's fraud score was calculated from the fraud rating for each feature per the defined rule fraud rating as in Table 2, and the correct label for each dataset record was determined based on the number of invalid or valid rating rule summaries are expressed mathematically in Equation 2. Hence, for a record with features i to n, the fraud score was calculated thus:

$$Fraud\ Score = \sum_{i=1}^{n} Rating \quad (2)$$

That is, the sum of the fraud ratings per issue was compared with the risk range in Table 1 to arrive at the risk category. This was used to determine the critical level of the issues per feature per applicant registration record, hence the label and the risk category for the potential cyber threat per record.

Table 1. Dataset fraud score

Risk Category	Fraud Score	Flag
Compliant (C)	0	0
Low Risk (M)	< 0.60	1
High Risk (H)	>= 0.60	2

Table 2. Classification rules for invalid or valid mobile money applicant data record based on fraud rating and score

S/N	Invalid/Valid based on Labelling Rule	Label Rule Code	Fraud Rating	Fraud Score	Risk Category
1	First Name against Region	First Name/Region	0.5	0.5	Low
2	Surname against Region	Surname/Region	0.9	0.9	High
3	MMName against Region	MMName/Region	0.1	0.1	Low
4	First Name against Gender	FName/Gender	0.3	0.3	Low
5	Surname agminst Gender	SName/Gender	0.1	0.1	Low
6	Organisation /institution	Organisation/Institution	0.9	0.9	High
7	Not Name Letters	Not Name Letters	0.9	0.9	High
8	User Reputation	User Reputation	0.9	0.9	High
9	Agent Reputation	Agent Reputation	0.9	0.9	High
10	Agent Identity	Organisation/Institutio	0.9	0.9	High
11	Compliant Records	Compliant Records	0	0	Compliant

3.3. Formulation of the Predictive Model

For modelling cyber threat prediction during MMS activation via customer onboarding or SIM registration, a Logistic Regression ML algorithm model was used. This approach entails predicting the continuous value of one field (the target) from a set of values of the other fields (attributes or features).

A Regression model usually produces a continuous prediction value, which is usually in the form of a probability and is described as follows: ML classifiers require a training corpus of M input/output pairs (*x (i), y (i)).* Logistic regression uses the logistic curve for fraud detection, and it is a probabilistic statistical supervised learning model. There is data on a dummy dependent variable y_i (with values of 1 and 0) and a column vector of explanatory variables x_i (including a 1 for the intercept term) for a sample of n cases ($i = 1... n$). The logistic regression model is shown in Equation 3, as follows:

$$Pr(y_i = 1|x_i) = \frac{1}{1+exp(-\beta x_i)} \quad (3)$$

where β is a row vector of coefficients. In logit form, i.e. by taking, natural logarithms, the model may be written in "logit" form in Equation 4 as follows:

$$\ln\left[\frac{Pr(y_i = 1|x_i)}{Pr(y_i = 0|x_i)}\right] = \beta x_i \quad (4)$$

The goal of maximum likelihood estimation is to find a set of values for β that maximizes this function in the model. Hence, this model used a dependent variable, y_i, for each mobile money applicant, i, for each feature x of the applicant's SIM or mobile money registration biodata record details, x, representing the occurrence of fraudulent intent (1 = fraud; 0 = non-fraudulent or compliant registrations). The fraudulent category was divided into two based on risk quantification levels: low risk (class 1) and high risk (class 2).

The logistic curve ranges from a value between 0 and 1, so it can be interpreted as the probability of class membership to predict the occurrence of cyber threats or intent in a customer MMS registration or activation request, as shown in Equation 5.

$$P(y_i = 1) = exp(\beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k + \varepsilon)_1 + (\beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k + \varepsilon)$$

Defining:

$$\pi_i = P(y_i = 1) \text{ and } 1 - \pi_i = P\ (y_i = 0) \quad (5)$$

The Logistic Regression has been known to be sensitive to the class-imbalance problem of the dataset, which may impede the classification capabilities in terms of its predictive accuracy, precision, and sensitivity and thus may be biased. In this study, to investigate the effect of the class imbalance problem, the Logistic Regression algorithm for predicting mobile money threats was designed using various configurations based on two variants:

Classification Configurations: This involves the Binary Classification Configuration, which classified the mobile money applicants' dataset as compliant (class 0) and non-compliant (class 1), and also the Multiclass Classification Configurations, which classified the mobile money applicants' dataset into three distinct categories of cyber threat risks: low risk (class 1), high risk (class 2), The biodata of the applicant was classified into several categories based on cyber threat non-existence as compliant, class 0, existence as low-risk, class 1, high-risk, and class 2.

Dataset Distribution: This involves imbalanced data (NO-SMOTE), which are the datasets acquired with unequal class distribution, and balanced data (SMOTE), which are the processed datasets using the Synthetic Minority Over-Sampling Technique (SMOTE).

This brought four different algorithms for Logistics Regression variants based on configurations, as shown in Table 3, finally resulting in a total of four (4) predictive models. The models were trained on the historical data set of SIM registrations for new applications and existing customers. This was done in many iterations to learn the model.

Table 3. Description of variants of logistic regression algorithms used

	Algorithms	Description
A	Logistics Regression	
1	LR Binary-No SMOTE	Logistic Regression with Binary feature configuration and without SMOTE application to Dataset
2	LR Binary-SMOTE	Logistic Regression with Binary feature configuration and with SMOTE application to Dataset
3	LR Multiclass-No SMOTE	Logistic Regression with Multiclass feature configuration and without SMOTE application to Dataset
4	LR Multiclass-SMOTE	Logistic Regression with Multiclass feature configuration and with SMOTE application to Dataset

4. Results and Discussions

The data for the entire sample was subjected to a logistic regression analysis model, with the applicant's mobile money registration biodata record details serving as the predictor variables and the dependent variable being the applicant's mobile money registration status. The simulation of the Logistic Regression-based predictive models for the detection and prevention of mobile money cyber threats during the customer registration process was carried out using the Python programming language with its ML library (i.e., *Pandas, NumPy, SciPy, Scikit-Learn, Matplotlib*). A Jupyter notebook was used for the coding, while Pandas, a data analysis library, was used for the pre-processing of the dataset. The dataset was split into 70% training and 30% testing according to accepted heuristics (other split values yielded similar results). The detailed results are presented as follows:

4.1. Data Pre-processing Output

During the data collection stage, the Python faker library was used to generate 1000 names, which were then combined with 8000 harvested Nigerian names. Pseudonymized data records gathered from the field with valid and invalid registrations from the practitioners were used as the basis for generating a larger number of valid and invalid records with data record-based simulations. Bag-of-Words (BoW) in Python was used to transform, store, and convert string values into names and other string features in the dataset. The Label encoder library in Python was used to encode the categorical variables in the dataset, e.g., male = 1, female = 0, while some were encoded with codes. The dataset was then simulated to five million records, of which 25 thousand were sampled for subsequent simulations.

The dataset for the analytical model building was drawn from the 5 million original records after each record was labelled according to the labelling rules to conserve processing resources. The distribution histograms of the imbalanced sampled dataset and the balanced dataset by Synthetic Minority Over-Sampling Technique (SMOTE) application are presented in Figures 7 and 8, respectively. After pre-processing, the ML algorithm classifiers were trained on the dataset with and without the SMOTE operation to observe the differences in performance and select the best algorithm. The SMOTE was performed on the dataset to observe the algorithm's performance under the two scenarios.

Other pre-processing activities include generating applicants' names with the faker () library and converting string features in the dataset to numeric values in a format that can be processed by a ML algorithm.

Figure 7. Distribution of the imbalanced sampled dataset of mobile money applicants

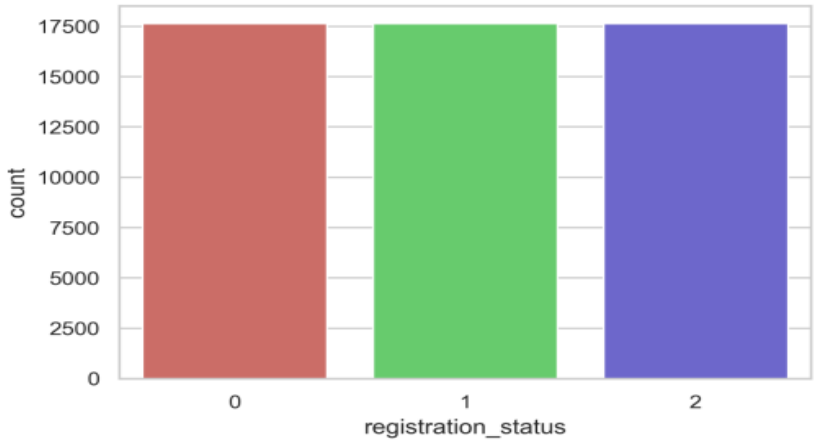


Figure 8. Distribution of the sampled dataset of (balanced with SMOTE) of mobile money applicants

4.2. Simulation Results

The proposed predictive algorithm was simulated on the defined dataset to determine the effectiveness of the logistic regression classifier for cyber threat prediction during the mobile money applicant on-boarding process. The total dataset was divided into eighty percent for training and twenty percent for testing. For validation of the dataset, a cross-validation technique was applied using *the train_test_split* of the *scikit*-learn library.

Simulations were used to systematically explore the behaviour of the logistic regression-based classifier with both the binary and multiclass classification capabilities for classifying the applicants' records into compliant (0) and bi-level fraudulent registration statuses—Low risk (1) and High risk (2). The simulation experiments were grouped as follows:

Group I: Binary classification algorithm configurations with and without SMOTE for dataset rebalancing. In Experiment I, an investigation was carried out on the classifier's binary ability to predict multi-level cyber threat categories in relation to SMOTE's pre-application to the dataset. While in Experiment II, investigation was carried out on the binary ability of the classifier with respect to the non-application of SMOTE to the dataset.

Group II: Multiclass Classification Algorithms configurations with and without SMOTE for dataset re-balancing. In Experiment I an investigation was carried out on the multiclass ability of the classifier with respect to SMOTE application to the dataset. While in Experiment II, investigation was carried out on the multiclass ability of the classifier with respect to the non-application of SMOTE to the dataset.

Based on the binary and multiclass classifications in Tables 5 and 6, respectively, the simulation results are presented as confusion matrices in Figures 9 and 10. It was shown that in the binary classification, the number of True Positives (TP) in the NO-SMOTE configuration is higher than that of the SMOTE configuration, whereas the SMOTE configuration has a higher number of TP in the multiclass classification. Also, the predicted probabilities are shown in Figures 11 and 12. These results showed that the Logistic Regression algorithm predicts better with two-class classifications (i.e., binary) and not with multiple classes of compliant (0), low-risk (1), and high-risk (2). This implies that the balancing nature of the dataset affected the predictive capability of the classifier with respect to its four (4) variants.

Table 5. Confusion matrix (binary classification)

		Predicted Class	
Actual Class		Positive (P)	Negative (N)
	Positive (P)	TP: Correctly Classified	FN
	Negative (N)	FP	TN: Correctly Classified

Table 6. Confusion matrix (multiclass classification)

		Predicted Class		
Actual Class		0	1	2
	0	Accurate	FN	FN
	1	FP	Accurate	TN
	2	FP	TN	Accurate

0= Compliant
1= Low Risk
2 = High Risk

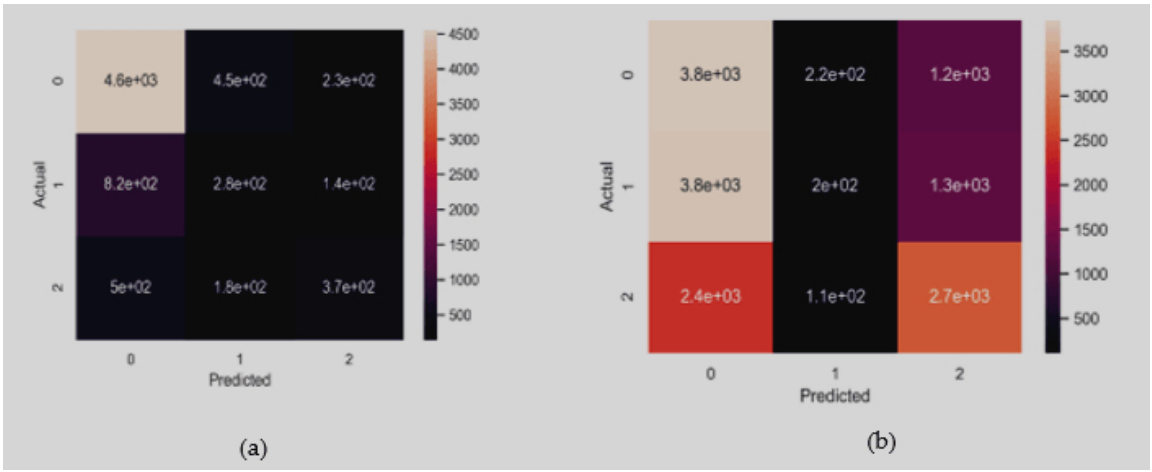


Figure 9. (a) No SMOTE binary classification; (b) SMOTE binary classification

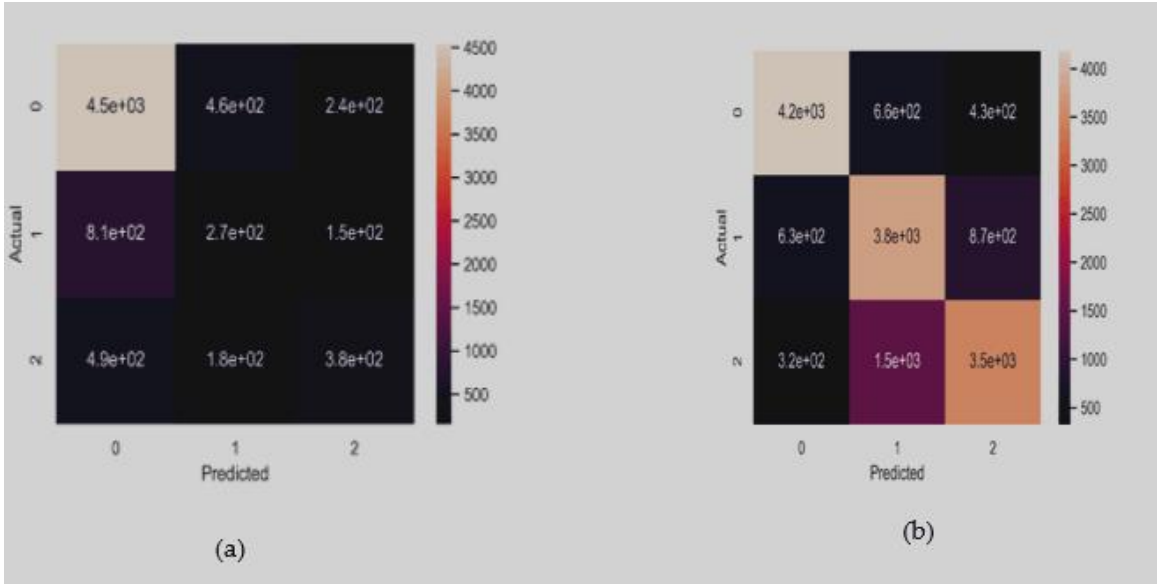


Figure 10. (a) No SMOTE, multiclass classification; (b) With SMOTE, multiclass classification

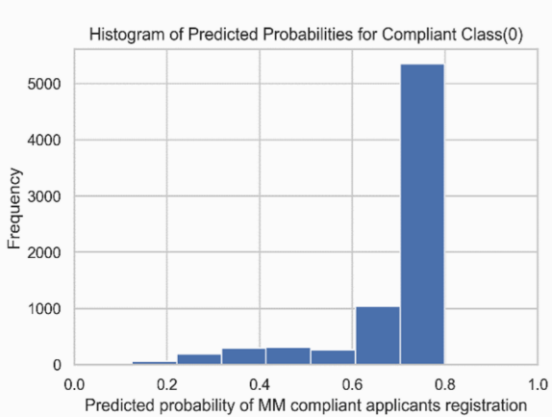
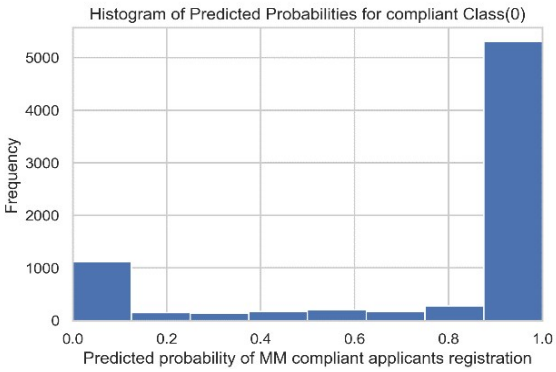
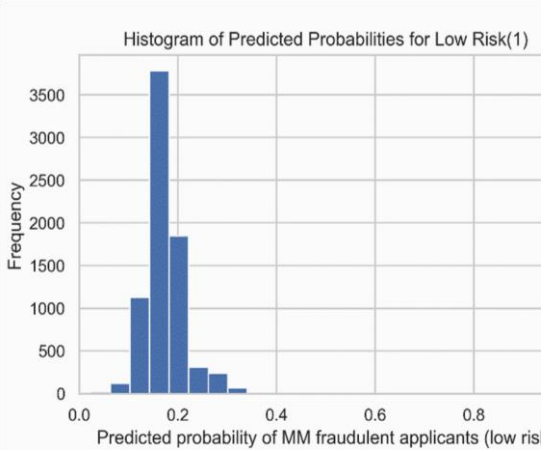
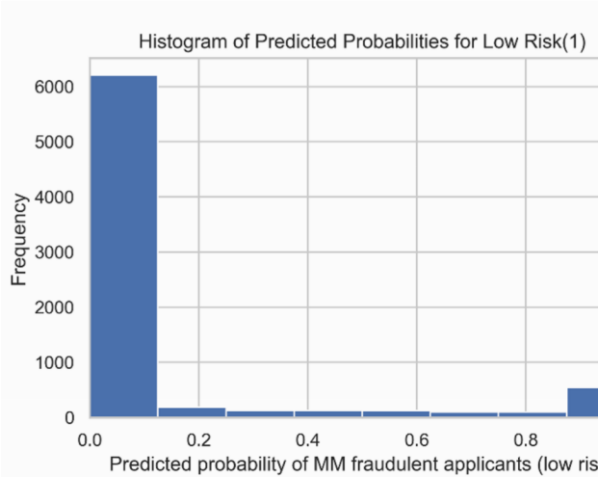
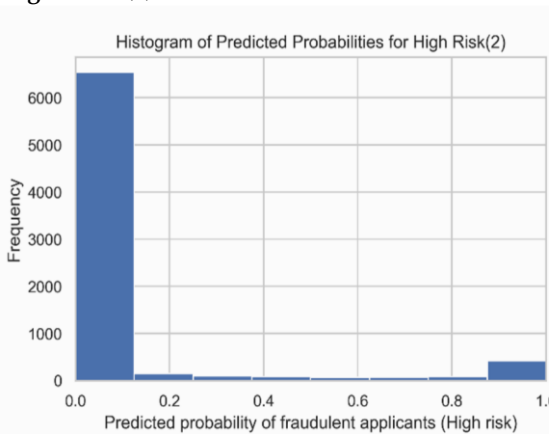
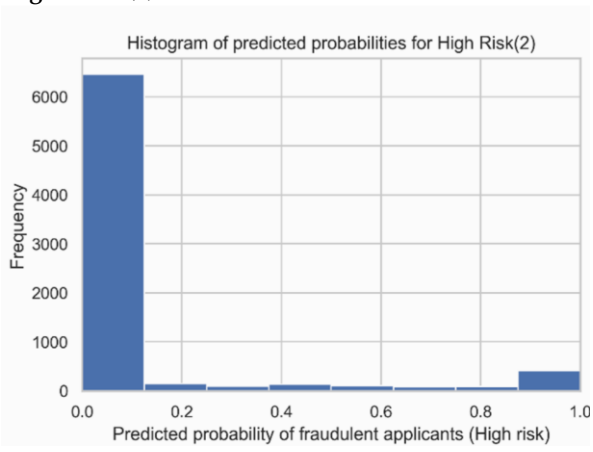
Predicted Probabilities, No SMOTE (Binary)	Predicted Probabilities, No SMOTE (Multiclass)
Low Class (0) 	**Low Class (0)** 
Medium Class (1) 	**Medium Class (1)** 
High Class (2) 	**High Class (2)** 

Figure 11. Logistic regression classifier No SMOTE- binary and multiclass LR algorithm configurations

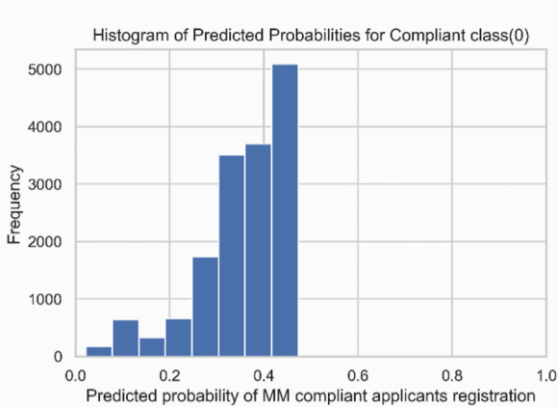
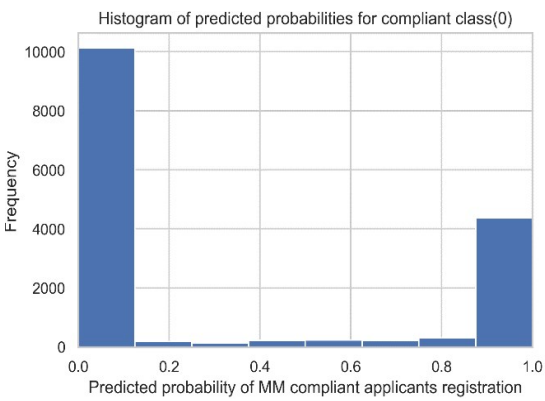
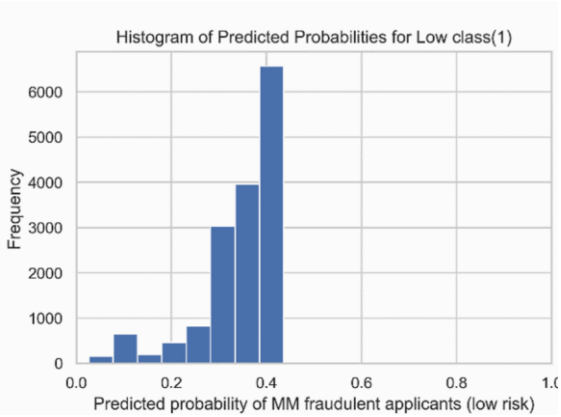
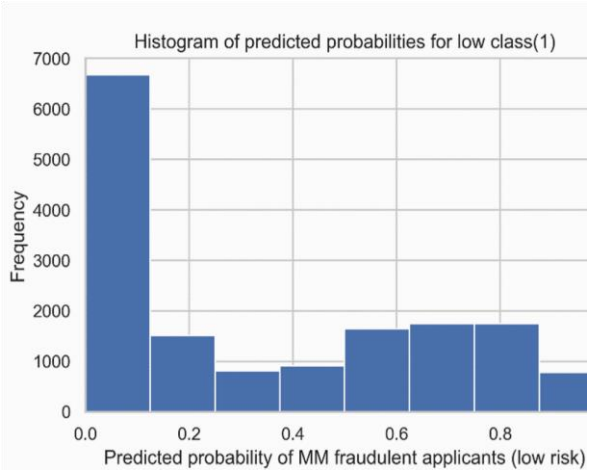
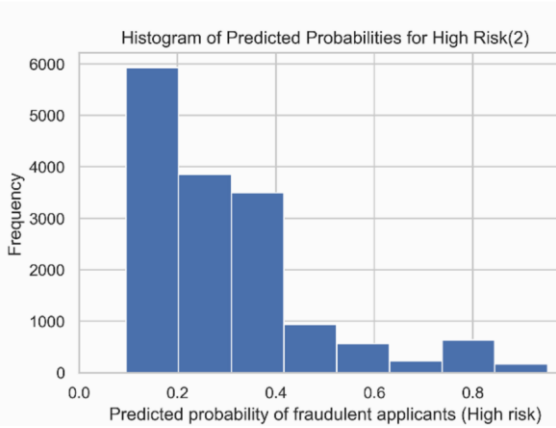
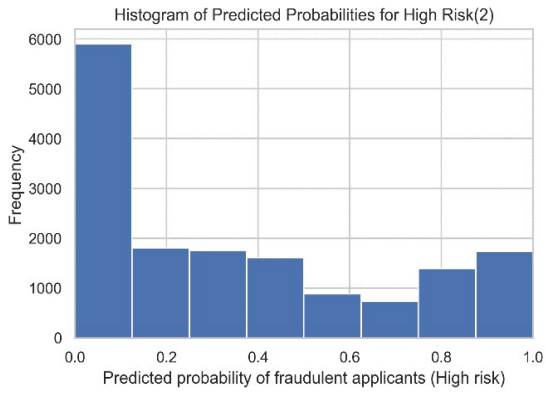
Predicted Probabilities, With SMOTE (Binary)	Predicted Probabilities, With SMOTE (Multiclass)
Low Class (0)	**Low Class (0)**
	
Medium Class (1)	**Medium Class (1)**
	
High Class (2)	**High Class (2)**
	

Figure 12. Logistic regression classifier with SMOTE- binary and multiclass LR algorithm configurations

4.3. Performance Evaluation Results

The classification performance of logistic regression in relation to imbalanced (No-SMOTE) and balanced (SMOTE) datasets was evaluated using accuracy, Matthews Correlation Coefficient (MCC), precision, recall, F1-Score, and Receiver Operating Characteristics (ROC) as performance metrics. The evaluation results based on the metrics are presented in Table 5, and in Figures 13 and 14. They are as follows:

Accuracy: With imbalanced datasets (NO SMOTE), a prediction accuracy of 72% was observed with the default binary classification feature of the algorithm, while multiclass classification gave an accuracy of 69%. With a balanced dataset (SMOTE), binary classification's accuracy dropped to 42% while multiclass classification's accuracy increased to 72%. This implies that the binary classification feature of the logistic regression classifier has the ability to predict effectively with the imbalanced dataset, while its multiclass classification feature predicts well with the balanced dataset.

MCC: The coefficient of the prediction was evaluated. It was revealed that the best classifier configuration was multiclass with SMOTE among the four logistic regression experiments performed, which had the highest MCC of 58% when compared with other experiments' MCC's of 27%, 15%, and 16%.

Precision, Recall, and F1-Score: As presented in Table 5, the precision, or specificity, for the LR binary with No-SMOTE classification experiment was 0.76, and the recall, or sensitivity, was 0.48 when compared with the multiclass classification logistic regression model specificity of 0.86 and sensitivity of 0.72.

ROC and Predicted Probabilities: The Receiver Operating Characteristics (ROC) Area Under Curve (AUC) for multiclass logistic regression of 0.84 were also higher than for all other logistic regression experiments.

The evaluation results showed that SMOTE improves the performance of the logistic regression classifiers, although the multiclass with logistic regression gave the best performance.

4.4. Validation Results

The model validation was done to benchmark the cyber threat compliance and correctness of the applicant's record in relation to existing methodologies. The predictive cyber threat model was benchmarked with the manual eyeballing process used by human agents to verify the sanctity of on-boarding mobile money applicants' registration data. Figure 15 depicts the schematic view of the human agent and ML algorithm predictive models. The result showed that the ML algorithm performed better and faster than the existing manual eyeballing method in use. The two methods are expressed in details as follows:

Manual Eyeballing method: Eyeballing each record after customer registration before final acceptance for mobile money registration was the main method in use to validate the textual correctness and completeness of a customer record. According to expert interviews, sample manual validation of customer records takes 8–10 minutes on average. This was further expressed as Average Eyeballing Duration and Eyeballing Accuracy. In an extensive eyeballing operation, say, agents are given a total of 100 customer records per day to manually eyeball. If validating a record takes t seconds, then validating n records will take Q seconds per agent: $Q = n*t$. From the table of observations of five agents, as in Table 6, eyeballing 20 records each for a total of 100 observations, the average eyeballing duration per record is 9.98 minutes, which is far more than the duration of validating 25 thousand records, which is 0.4 minutes when using Random Forest algorithms. Also, in the simulation with the proposed predictive algorithm, the accuracy was 91%. However, compared with the manual eyeballing method, the human agent becomes fatigued and tired, which frustrates the accuracy of this method. From expert interviews and observations, the accuracy is always below 71%, and the number of records validated is typically 100 per day per agent. This is far less than the ML output per second.

Machine Learning Methods: This provides improved efficiency, a lower error rate, and better data quality. From the results obtained above, it is evident that the average processing time for a record is a fraction of a second.

Table 5: Evaluation results for Logistics Regression with (SMOTE)/ without SMOTE (NO SMOTE)

	Logistics Regression	MCC (-1+1)	Accuracy	Fl-Score	Precision	MR	ROC AUC	Specificity (TNR)	FPR	Sensitivity (TPR)	FNR	RT (Min)
Group I Experiment I	LR Binary-No SMOTE	0.16	0.72	0.79	0.92	0.29	0.62	0.76	0.24	0.48	0.52	0.0115
Group I Experiment II	LR Binary-SMOTE	0.15	0.42	0.47	0.59	0.57	0.64	0.71	0.29	0.42	0.58	0.0256

Group II Experiment I	LR Multi-class-No SMOTE	0.27	0.69	0.71	0.74	0.31	0.71	0.75	0.25	0.48	0.52	0.0156
Group II Experiment II	LR Multi-class-SMOTE	0.58	0.72	0.72	0.72	0.28	0.84	0.86	0.14	0.72	0.28	0.0857

Legend
LR Binary-No SMOTE = Binary Logistic regression without SMOTE
LR Binary-SMOTE =Binary Logistic regression with SMOTE
LR Multiclass-No SMOTE = Multi-class Logistic regression without SMOTE
LR Multiclass-SMOTE = Multi-class logistic regression with SMOTE
ROC AUC = Receiver Operating Characteristics Area Under Curve
FNR = False Negative Rate; FPR = False Positive Rate; TNR = True Negative Rate
MR = Mis-classification Rate
RT= Runtime

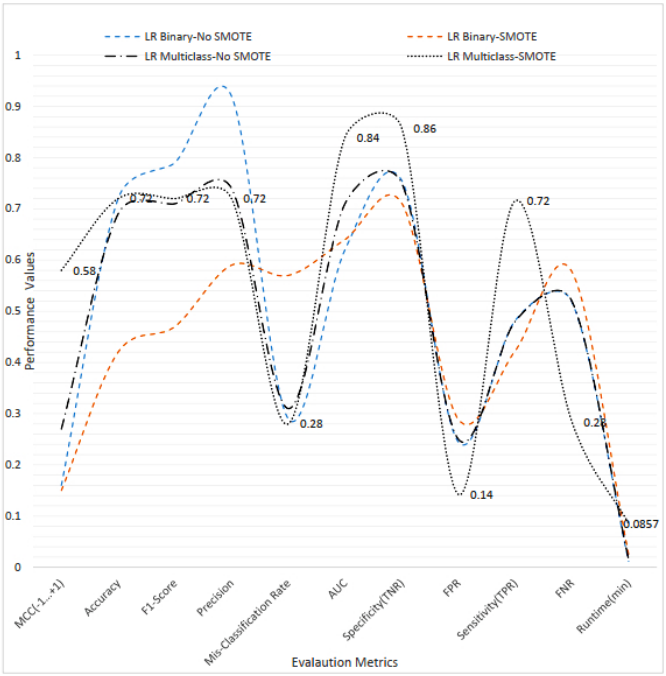


Figure 13. Results of different Logistic Regression configuration with different performance metrics

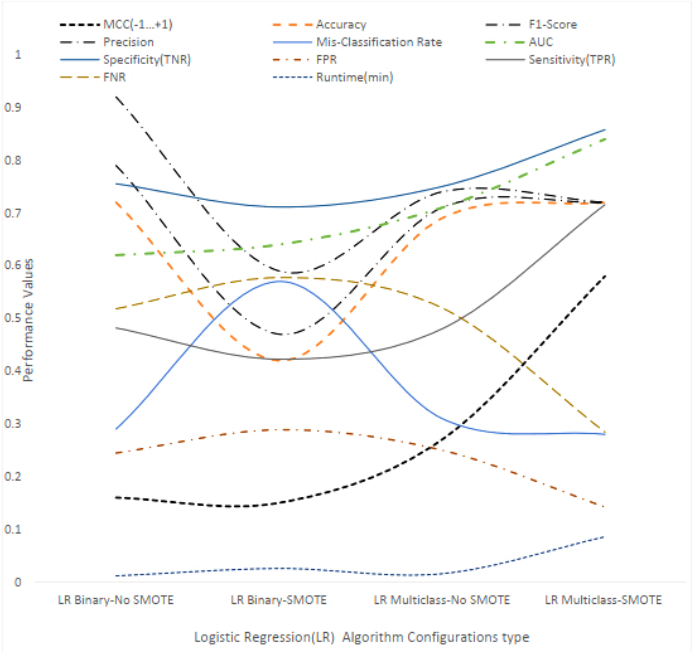


Figure 14. Logistic Regression (LR) performance results for different configurations

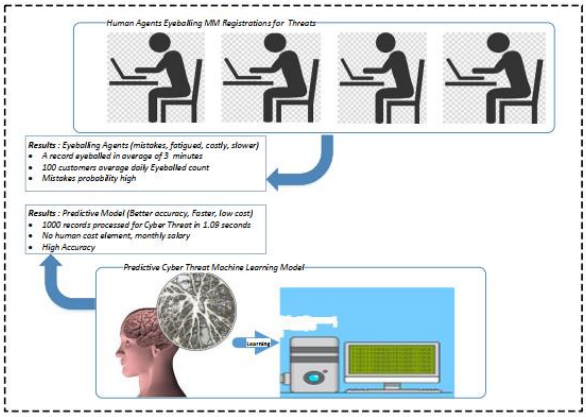


Figure 15. Predictive machine learning model Vs eyeballing method

5. Conclusion

This research aimed to develop a predictive cyber-threat model that detects and prevents any suspicious customer with cyber threat potential during the on-boarding process of mobile money customer lifecycle management in developing nations, using Nigeria as a case study. Logistic regression, a ML algorithm, was used to inhibit or prevent suspicious customers from joining the MMS by predicting the customers' intent and thus eliminating the financially criminal mindset customer. A dataset was collected, and the Python programming language was used for data analysis, data conversion, and transformation functions. SMOTE was applied to handle the imbalance of the dataset class. The effectiveness of the logistic regression-based predictive algorithm for cyber threat predictions from customer data details during the mobile money on-boarding process was determined by evaluating the predictive model based on the dataset balance status and classification configurations.

Thus, the proposed model for the mobile money initiative will provide a sustainable drive for financial inclusion and cashless policies, as well as accelerate mobile service adoption in developing nations. As shown and concluded in this study, the adoption of this predictive model would go a long way toward reducing financial fraud at MFS. Based on this conclusion, MMS providers should consider using this predictive cyber threat model to prevent suspicious customers from being onboarded onto mobile money platforms.

Table 6: Duration of eyeballing per records for 20 observations per human agent

Record	Agent-1(min)	Agent-2(min)	Agent-3(min)	Agent-4(min)	Agent-5(min)
1	10	6	7	10	9
2	8	7	11	11	7
3	10	8	12	12	8
4	10	13	10	13	9
5	10	11	14	15	10
6	8	12	12	16	11
7	10	10	15	12	7
8	10	8	13	15	10
9	8	11	12	11	7
10	10	12	15	14	8
11	10	10	10	13	9
12	7	6	8	14	10
13	10	5	9	10	9
14	10	7	13	9	9
15	10	6	7	16	11
16	6	8	8	14	8
17	10	9	9	12	9
18	10	5	7	14	8
19	10	8	8	10	9
20	10	11	9	10	10

Meanwhile, this study only focused on mobile phone subscriber biodata registration details for MMS. Other components of the customer lifecycle management process such as modification (SIM SWAP), customer biometrics, profile modification, and customer termination processes as cyber threat vectors for

MMS were not explored. Further research will be conducted to investigate the performance of other classifiers capable of predicting the likelihood of a cyber threat or fraudulent intent applicant during the MMS onboarding or service activation process, with the goal of determining the best ML model for the predictive model solution. Furthermore, investigations will be carried out using Ordinary Differential Equation (ODE) models or epidemiological models to predict cyber threat spread rates in MFS or other domains.

Acknowledgement

This Research was funded by the TETFund Research Fund and Africa Centre of Excellence OAK-Park, Obafemi Awolowo University, Ile-Ife, Nigeria.

References

[1] Babatope E. Akinyemi and Abbyssinia Mushunje, "Determinants of Mobile Money Technology Adoption in Rural Areas of Africa", *Cogent Social Sciences*, ISSN: 23311886, Vol. 6, No. 1, 2020, DOI: 10.1080/23311886.2020.1815963, Available: https://www.tandfonline.com/doi/full/10.1080/23311886.2020.1815963.

[2] World Bank, *World Development Report 2016: Digital Dividends*, ISSN: 0163-5085, ISBN: 978-1-4648-0671-1, E-ISBN: 978-1-4648-0672-8, World Bank Publications, Washington D.C., United States, 2016, DOI: 10.1596/978-1-4648-0671-1, Available: https://www.worldbank.org/en/publication/wdr2016.

[3] Kevin Donovan, "Mobile Money for Financial Inclusion", *Information and Communications for Development*, ISBN: 9780821389911, e-ISBN: 9780821395875, Vol. 61, pp.61–73, 2012, Published by The World Bank Group, DOI: 10.1596/9780821389911_ch04, Available: https://elibrary.worldbank.org/doi/abs/10.1596/9780821389911_ch04.

[4] Kevin P. Donovan, "Mobile Money, More Freedom? The Impact of M-Pesa's Network Power on Development as Freedom", *International Journal of Communication*, Vol. 6, No. 23, pp. 2647–2669, 2012, Published by University of Southern California, Available: https://ijoc.org/index.php/ijoc/article/view/1575/815.

[5] Janine Aron, "Mobile Money and the Economy: A Review of Evidence", *World Bank Research Observer*, Vol. 23, No. 2, pp. 135-188, August 2018, Published by Oxford University Press, DOI: 10.1093/wbro/lky001, Available: https://elibrary.worldbank.org/doi/abs/10.1093/wbro/lky001.

[6] Sam Castle, Pervaiz Fahad, Cassebeer Weld Galen, Roesner Franziska and Richard J. Anderson, "Let's Talk Money: Evaluating the Security Challenges of Mobile Money in the Developing World", in *Proceedings of the 7th Annual Symposium on Computing for Development (ACM DEV '16)*, 18 – 20 November 2016, Nairobi, Kenya, ISBN: 978-1-4503-4649-8, pp. 1-10, Published by the Association for Computing Machinery, DOI: 10.1145/3001913.3001919, Available: https://dl.acm.org/doi/10.1145/3001913.3001919.

[7] Sionfou Seydou Coulibaly, "A Study of the Factors Affecting Mobile Money Penetration Rates in the West African Economic and Monetary Union (WAEMU) Compared with East Africa", *Financial Innovation*, Vol. 7, No. 25, 2021, Published by Springer Nature, DOI: 10.1186/s40854-021-00238-0, Available: https://jfin-swufe.springeropen.com/articles/10.1186/s40854-021-00238-0.

[8] Adam B. Mtaho, "Improving Mobile Money Security with Two-Factor Authentication", *International Journal of Computer Applications*, ISBN: 973-93-80884-75-6, Vol. 109, No. 7, pp.9-15, 2015, DOI: 10.5120/19198-0826, Available: https://www.ijcaonline.org/archives/volume109/number7/19198-0826.

[9] Ibn Kailan Abdul-Hamid, Aijaz A. Shaikh, Henry Boateng and Robert E. Hinson, "Customers' Perceived Risk and Trust in Using Mobile Money Services—an Empirical Study of Ghana", *International Journal of E-Business Research (IJEBR)*, ISSN: 1548-1131, EISSN: 1548-114X, EISBN13: 9781522564287, Vol. 15, No. 1, pp. 1-19, 2019, Published by IGI Global, DOI: 10.4018/IJEBR.2019010101, Available: https://www.igi-global.com/article/customers-perceived-risk-and-trust-in-using-mobile-money-servicesan-empirical-study-of-ghana/219224.

[10] Andrew Harris, Seymour Goodman and Patrick Traynor, "Privacy and Security Concerns Associated with Mobile Money Applications in Africa", *Washington Journal of Law, Technology & Arts*, Vol. 8, No. 3, pp. 245, 2013, Published by University of Washington, Available: https://digitalcommons.law.uw.edu/wjlta/vol8/iss3/5.

[11] Babatunde Moses Ololade, Mary Kehinde Salawu and Aderemi Daniel Adekanmi, "E-Fraud in Nigerian Banks: Why and How?", *Journal of Financial Risk Management*, ISSN Online: 2167-9541, ISSN Print: 2167-9533, Vol. 9, pp. 211-228, 2020, Published by Scientific Research Publishing Inc., DOI: 10.4236/jfrm.2020.93012, Available: https://www.scirp.org/pdf/jfrm_2020090915162210.pdf.

[12] Stephen Ambore, Christopher Richardson, Huseyin Dogan, Edward Apeh and David Osselton, "A Resilient Cybersecurity Framework for Mobile Financial Services (MFS)", *Journal of Cyber Security Technology*, Vol. 1, No. 3-4, pp. 202-224, 2017, Published by Taylor and Francis Online, DOI: 10.1080/23742917.2017.1386483, Available: https://www.tandfonline.com/doi/full/10.1080/23742917.2017.1386483.

[13] Ali Guma, Mussa Ally Dida and Anael Elikana Sam, "Evaluation of Key Security Issues Associated with Mobile Money Systems in Uganda", *Information*, EISSN: 2078-2489, Vol.11, No. 6, pp. 309, 2020, Published by MDPI, DOI:10.3390/info11060309, Available: https://www.mdpi.com/2078-2489/11/6/309.

[14] Hakeem J. Pallangyo, "Cyber Security Challenges, its Emerging Trends on Latest Information and Communication Technology and Cyber Crime in Mobile Money Transaction Services", *Tanzania Journal of Engineering and Technology,* Print ISSN: 1821-536X, E-ISSN: 2619-8789, Vol. 41, No. 2, pp. 189-204, 2022, Published by College of Engineering and Technology, University of Dar es Salaam, DOI: 10.52339/tjet.v41i2.79, Available: https://tjet.udsm.ac.tz/index.php/tjet/article/view/792/648.

[15] Lema Aulelius, "Factors Influencing the Adoption of Mobile Financial Services in the Unbanked Population", *Inkanyiso: Journal of Humanities and Social Sciences,* E-ISSN: 2077-8317, Print ISSN: 2077-2815, Vol. 9, No. 1, pp. 37-51, 2017, Published by African Journals Online, DOI: 10.4314/IJHSS.V9I1, Available: https://www.ajol.info/index.php/ijhss/article/view/165506.

[16] Pierre-Laurent Chatain, Andrew Zerzan, Wameek Noor, Najah Dannaoui and Louis de Koker, *Protecting Mobile Money against Financial Crimes: Global Policy Challenges and Solutions,* Illustrated ed. Washington, D. C., USA: World Bank Group, 2011, ISBN:978-0-8213-8669-9, e-ISBN: 978-0-8213-8670-5, DOI: 10.1596/978-0-8213-8669-9, Available: https://elibrary.worldbank.org/doi/abs/10.1596/978-0-8213-8669-9.

[17] Whisker James and Lokanan Mark, "Anti-Money Laundering and Counter-Terrorist Financing Threats Posed by Mobile Money", *Journal of Money Laundering Control*, Vol. 22, No. 1. pp. 34-45, 2019, DOI: 10.1108/JMLC-10-2017-0061, Available: https://www.emerald.com/insight/content/doi/10.1108/JMLC-10-2017-0061/full/html.

[18] Ragib Hasan, Suvda Myagmar, Adam J. Lee and William Yurcik, "Toward a Threat Model for Storage Systems", in *Proceedings of the 2005 ACM workshop on Storage security and survivability (StorageSS '05),* Fairfax VA, USA, 11 November 2005, ISBN: 978-1-59593-233-4, pp. 94–102, Published by the Association for Computing Machinery, DOI: 10.1145/1103780.1103795, Available: https://dl.acm.org/doi/abs/10.1145/1103780.1103795.

[19] Cristina K. Dominicini, Marcos A. Simplício Jr., Rony R. M. Sakuragui, Tereza C. M. B. Carvalho, Mats Näslund *et al*., "Threat Modeling an Identity Management System for Mobile Internet", in *Proceedings of the 9th International Information and Telecommunication Technologies Symposium (I2TS'10)*, UNIRIO, Rio de Janeiro, Brazil, 2010, Available: http://www.inf.ufsc.br/~bosco.sobral/downloads/I2TS%202010%20CD%20Proceedings/www.i2ts.org/papers/full_english/78298_1.pdf.

[20] Antonietta Stango, Neeli R. Prasad and Dimitris M. Kyriazanos, "A Threat Analysis Methodology for Security Evaluation and Enhancement Planning", in *Proceedings of the 3rd International Conference on Emerging Security Information, Systems and Technologies (SECURWARE 2009),* 18-23 June 2009, Athens, Greece, Print ISBN: 978-0-7695-3668-2, pp. 262–267, Published by IEEE, DOI: 10.1109/SECURWARE.2009.47, Available: https://ieeexplore.ieee.org/document/5210987.

[21] Tong Xin and Ban Xiaofang, "Online Banking Security Analysis based on STRIDE Threat Model", *International Journal of Security and Its Applications*, ISSN: 1738-9976, Vol. 8, No. 2, pp. 271-282, 2014, Published by Science and Engineering Research Support Society (SERSC), DOI: 10.14257/ijsia.2014.8.2.28, Available: http://article.nadiapub.com/IJSIA/vol8_no2/28.pdf.

[22] Ye Xiaolie and Liao Lejian, "Verifying a Secure Session Protocol for Web Services", in *Proceedings of the International Conference on Networks Security, Wireless Communications and Trusted Computing (NSWCTC '09),* Wuhan, China, Vol. 2, pp. 301- 304, 25-26 April 2009, Published by IEEE, DOI: 10.1109/NSWCTC.2009.329, Available: https://ieeexplore.ieee.org/document/4908465.

[23] Jiancheng Ni, Zhishu Li, Zhonghe Gao and Jirong Sun, "Threats Analysis and Prevention for Grid and Web Service Security", in *Proceedings of the ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing (SNPD 2007),* 30 July 2007 - 01 August 2007, Qingdao, China, Print ISBN:0-7695-2909-7, Print ISBN:978-0-7695-2909-7, pp. 526–531, Published by IEEE, DOI: 10.1109/SNPD.2007.269, Available: https://ieeexplore.ieee.org/document/4287910.

[24] Yuri Demchenko, Leon Gommans, Cees de Laat and Bas van Oudenaarde, "Web Services and Grid Security Vulnerabilities and Threats Analysis and Model", in *Proceedings of the 6th IEEE/ACM International Workshop on Grid Computing,* November 2005, Seattle, WA, USA, Print ISBN:0-7803-9492-5, pp. 262–267, Published by IEEE, DOI: 10.1109/GRID.2005.1542751, Available: https://ieeexplore.ieee.org/document/1542751.

[25] Maurice ter Beek, Corrado Moiso and Marinella Petrocchi, "Towards Security Analyses of an Identity Federation Protocol for Web Services in Convergent Networks", in *Proceedings of the Third Advanced International Conference on Telecommunications (AICT 2007)*, 13-19 May 2007, Morne, Mauritius, pp. 31-31, Published by IEEE, DOI: 10.1109/AICT.2007.46, Available: https://ieeexplore.ieee.org/document/4215252.

[26] Ebenezer Akin Oladimeji, Sam Supakku and Lawrence Chung, "Security Threat Modeling and Analysis: A Goal-Oriented Approach", in *Proceedings of the 10th IASTED International Conference on Software Engineering and Applications (SEA 2006),* 13 – 15 November 2006, Dallas, USA, ISBN: 0-88986-642-2, pp.13-15, Published by ACTA Press, Available: https://www.actapress.com/Abstract.aspx?paperId=28899.

[27] Majda Omer Albasheer and Eihab Bashier Mohammed Bashier, "Enhanced Model for PKI Certificate Validation in the Mobile Banking", in *Proceedings of the 2013 International Conference on Computing, Electrical And Electronics Engineering (ICCEEE)*, 26-28 August 2013, Khartoum, Sudan, pp. 470–476, Published by IEEE, DOI: 10.1109/ICCEEE.2013.6633984, Available: https://ieeexplore.ieee.org/document/6633984.

[28] Shaik Shakeel Ahamad, V. N. Sastry and Madhusoodhnan Nair, "Biometric Based Secure Mobile Payment Framework", in *Proceedings 2013 4th International Conference on Computer And Communication Technology (ICCCT)*, 20-22 September 2013, Allahabad, India, pp. 239-246, Published by IEEE, DOI: 10.1109/ICCCT.2013.6749634, Available: https://ieeexplore.ieee.org/document/6749634.

[29] C. Narendiran, S. Albert Rabara and Nishanth Rajendran, "Public Key Infrastructure for Mobile Banking Security", in *Proceedings of the 2009 Global Mobile Congress*, 12-14 October 2009 , Shanghai, China, pp. 1–6, Published by IEEE, DOI: 10.1109/GMC.2009.5295898, Available: https://ieeexplore.ieee.org/document/5295898.

[30] Hossain Md. Alamgir, "Security Perception in the Adoption of Mobile Payment and the Moderating Effect of Gender", *PSU Research Review*, Vol. 3, No. 3, pp. 179-190, 2019, Published by Emerald Publishing Limited, DOI: 10.1108/PRR-03-2019-0006, Available: https://www.emerald.com/insight/content/doi/10.1108/PRR-03-2019-0006/full/html.

[31] Peter Tobbin and John K. M. Kuwornu, "Adoption of Mobile Money Transfer Technology: Structural Equation Modelling Approach", *European Journal of Business and Management*, Print ISSN: 2222-1905, Online ISSN: 2222-2839, Vol. 3, No. 7, pp.59–77, 2011, Published by International Institute for Science, Technology and Education (IISTE), Available: https://core.ac.uk/download/pdf/234624099.pdf.

[32] Belkhede Mangala, Gulhane Veena and Bajaj Preeti, "Biometric Mechanism for Enhanced Security of Online Transaction on Android System: A Design Approach", in *Proceedings of the* 2012 *14th International Conference on Advanced Communication Technology (ICACT)*, 19-22 February 2012, PyeongChang, South Korea, pp. 1193 – 1197, Published by IEEE, Available: https://ieeexplore.ieee.org/document/6174876.

[33] Hee Yeon Min, Jin-Hyung Park, Dong Hoon Lee and In-seok Kim, "Outlier Detection Method for Mobile Banking with User Input Pattern and E-finance Transaction Pattern", *Journal of Internet Computing and Services*, Print ISSN: 1598-0170, E-ISSN: 2287-1136, Vol. 15, No. 1, pp.157–170, 2014, DOI: 10.7472/JKSII.2014.15.1.157, Available: http://koreascience.or.kr/article/JAKO201409150679234.page.

[34] Samuel Oluwole Falaki, Boniface Kayode Alese, Olumide Sunday Adewale, Joshua O. Ayeni, Ganiyu Adesola Aderounmu *et al*., "Probabilistic Credit Card Fraud Detection System in Online Transactions", *International Journal of Software Engineering and Its Applications*, Print ISSN:1738-9984, Vol. 6, No. 4, pp. 69-78, 2012, Published by Science and Engineering Research Support Society (SERSC), Available: https://www.earticle.net/Article/A208418.

[35] Munirul Ula, Zuraini Ismail and Zailani Sidek, "A Framework for the Governance of Information Security in Banking System", *Journal of Information Assurance & Cybersecurity*, Vol. 2011, pp. 1-12, 2011, Published by IBIMA Publishing, DOI: 10.5171/2011.726196, Available: http://www.ibimapublishing.com/journals/JIACS/jiacs.html.

[36] Martin Graham, Robert Kukla, Oleksii Mandrychenko, Darren Hart and Jessie Kennedy, "Developing Visualisations to Enhance an Insider Threat Product: A Case Study", in *Proceedings of the 2021 IEEE Symposium on Visualization for Cyber Security (VizSec)*, 27 October 2021, New Orleans, LA, USA, pp. 47-57, Published by IEEE, DOI: 10.1109/VizSec53666.2021.00011, Available: https://ieeexplore.ieee.org/document/9629405.

[37] Stephen Ambore, Christopher Richardson, Huseyin Dogan, Edward Apeh and David Osselton, "A Soft Approach to Analysing Mobile Financial Services Socio-Technical Systems", in *Proceedings of the 30th International BCS Human Computer Interaction Conference*, 11 - 15 July 2016, Poole, United Kingdom, pp. 1-3, Published by BCS Learning and Development Ltd., DOI: 10.14236/ewic/HCI2016.103, Available: https://www.scienceopen.com/hosted-document?doi=10.14236/ewic/HCI2016.103.

[38] Precious Chikezie Ezeh and Nwukamaka Nwankwo, "Factors that Influence the Acceptance of Mobile Money in Nigeria", *Journal of Research in Marketing*, Online ISSN: 2292-9355, Vol. 8, No. 2, pp. 684-697, 2018, DOI: 10.17722/jorm.v8i2.217, Available: https://www.scilit.net/article/a37db46fd3dcccfe1952a6a1f247e042.

[39] Nitesh V. Chawla, Kevin Bowyer, Lawrence O. Hall and Philip W. Kegelmeyer, "SMOTE: Synthetic Minority Over-Sampling Technique", *Journal of Artificial Intelligence Research*, Vol.16, pp. 321–357, 2002, Published by AI Access Foundation, DOI: 10.1613/jair.953, Available: https://www.jair.org/index.php/jair/article/view/10302.