Computing, Information Systems & Development Informatics Journal
Vol. 11 No. 2, 2020 - www.isteams.net/cisdijournal

An Email Classification Model for Detecting Advance Fee Fraud: A Conceptual Approach

Oyegoke, T.O., Amoo, A.O., Aderounmu, G.A. & Adagunodo, E.R.
Department of Computer Science & Engineering
Obafemi Awolowo University
Ile-Ife, Osun State. Nigeria
Email: temitayooyegoke@yahoo.com; olawunmikemi@yahoo.com, gaderoun@oauife.edu.ng; eadagun@oauife.edu.ng
Phone Nos: +2348033923880; +2348139638708; +2348035177940; +2348037250909

ABSTRACT

Advance Fee fraud commonly called 419 is a fraudulent process involving the process of enticing a potential victim with bogus propositions with a promise to transfer cash. The Electronic mail is by far the most common means via which this fraudulent act is committed which has led to Billions Dollar losses annually. Existing techniques for fraud detection using machine learning algorithms have also been observed to slowly adapt to the dynamics of fraudulent activities. As a result of this, nature-inspired algorithms have been adopted to address the limitations of adapting to the dynamics of fraudulent activities. The ant colony (ACO) and particle swarm optimization (PSO) algorithms have been proposed in this study for the identification of relevant fraudulent features within e-mail content and for the development of a hybrid model with the back-propagation algorithm for fraud detection. This study also presents a framework for the development of a fraud detection and incident reporting system which consists of three main modules. Data collection and feature extraction module for extracting contents of incoming e-mails into structured data; a fraud detection module for detecting fraud based on information about extracted contents and an incident reporting module which allows a user to report false alarms made by the system so as to ensure adaptation to changing features in undetected fraudulent e-mails. Future works is focused on the development of a fraud detection and incidence reporting system based on the proposed framework.

Keywords: Fraud, E-mail, Classification model, Advance fee.

CISDI Journal Reference Format
Oyegoke, T.O., Amoo, A.O., Aderounmu, G.A. & Adagunodo, E.R. (2020): An Email Classification Model for Detecting Advance Fee Fraud: A Conceptual Approach. Computing, Information Systems, Development Informatics & Allied Research Journal. Vol 11 No 2, Pp 91-104.
DOI - https://doi.org/ 10.22624/AIMS/CISDI/V11N4P6. Available online at www.computing-infosystemsjournal.info

1. INTRODUCTION

Fraudulent activities usually involve the perpetrator using false and misleading representations to disguise activities with the victim in order to avoid detection as long as possible with an effort to maximize the effects of their fraudulent behaviour (Alexopoulos *et al.*, 2007). Advance Fee Fraud commonly known as 419, is a fraudulent process of enticing a victim with a bogus business proposal which is done with a promise to transfer large sums of money (Reich, 2004). This money which is usually in foreign exchange is purported to be part of the proceeds of certain contracts to the addressee's bank account to be shared in some proportion between the parties involved (Tive, 2006). The stages of advance fee fraud are shown in Figure 1.

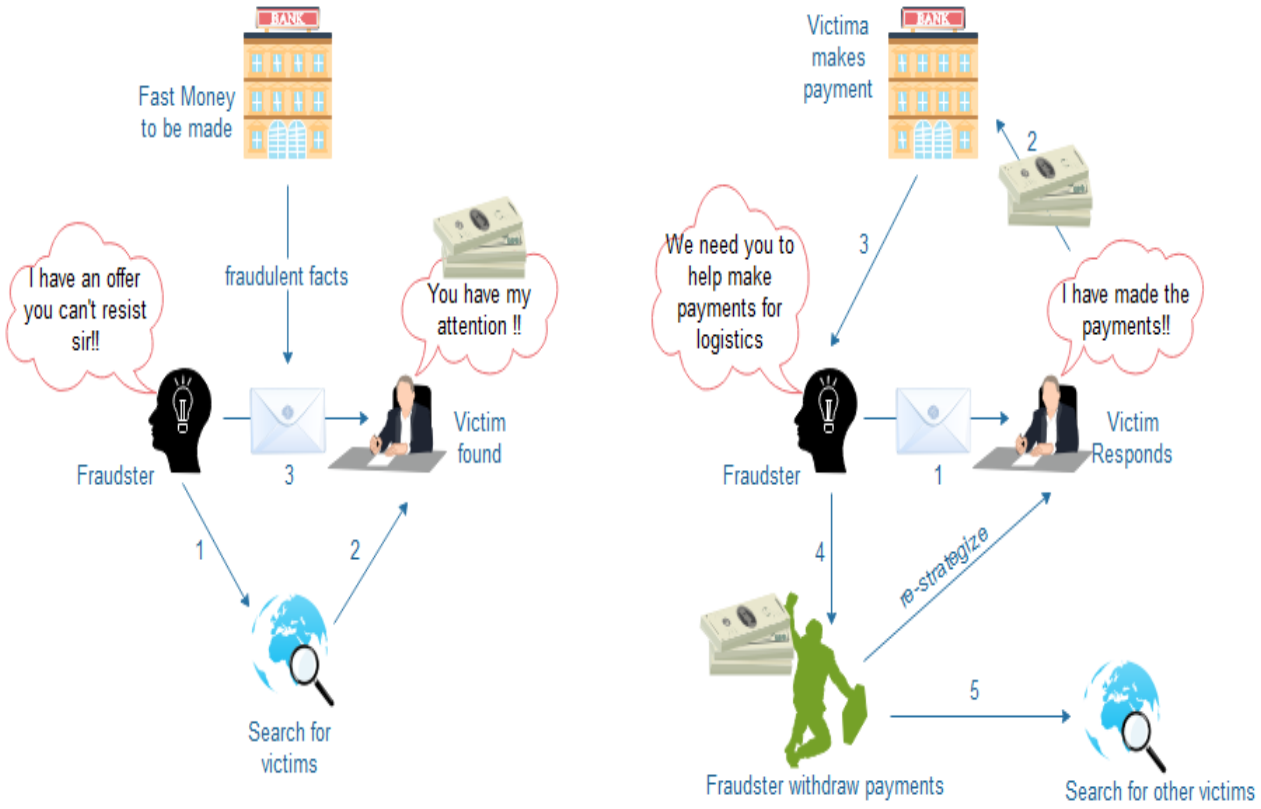


Figure 1: Stages Involved in Advance Fee Fraud Practice

The Australian Competition and Consumer Commission (ACCC) (2017), stated that advance fee fraud has led to huge loss of money among the various fraudulent practices since the past years amounting to $3.9 billion in 2015 and $6.5 billion in 2016 as shown in Table 1. Advance Fee Fraud has also proven to have the highest number of financial related reports compared to other fraudulent practices due to the high rate of financial related losses accounted in the reported crimes which has remained between 5 to 10 % in the past years.

The rapid growth of the Internet which has led to a significant increase in the number of email users at the same time has also led to an increase in spam emails rate by fraudsters. A statistical report showed that 70% of the email in circulation during the second week of 2014 was spam or illegitimate emails (Nizamani *et al.*, 2014).

Table 1: Distribution of Financial Crimes between 2015 and 2016

YEAR	Scam Type	Amount Lost ($)	Number of Reports	Percentage with Financial Loss (%)	Amount Lost per Report ($/report)
2016	**Advance Fee Fraud (AFF)**	6,499,604.00	16830	5.7	6,775.29
	Identity Theft	715,896.00	12731	1.6	3,514.53
	Phishing	373,860.00	24925	0.8	1,874.92
	Hacking	2,851,145.00	4050	5.3	13,282.76
	Ransomware/Malware	241,881.00	6210	3.7	1,052.71
2015	**Advance Fee Fraud (AFF)**	3,899.312.00	11502	7.8	4,346.30
	Identity Theft	1,255,423.00	9328	3.4	3,958.43
	Phishing	363,270.00	15430	1.2	1,961.92
	Hacking	710,797.00	3132	7.3	3,108.86
	Ransomware/Malware	388,167.00	4439	3.5	2,498.42

Source: ACCC, 2017

The Email has been considered as a convenient way of written communication in the 21st century as it is deemed to be an economical and steadfast method of communication which can be sent to a single receiver or broadcasted to groups of people (Nizamani *et al.*, 2011). For these and countless other motives, the email has also become a widely used medium for communication of the people having ill intentions such as fraudsters (Nizamani *et al.*, 2012). ACCC (2017), stated that the email has led to the loss of over $1.3 billion in 2015 and $4.5 billion in 2016 thus, resulting as the source of the greatest financial loss due to fraudulent crimes in the last 2 years as shown in Tables 2.

Table 2: Distribution of Financial loss due to AFF between 2015 and 2016.

YEAR	Amount Lost ($)			
	Phone	**Email**	**In Person**	**Internet**
2015	1,316,039.00	1,397,762.00	323,905.00	384,092.00
2016	914,039.00	4,515,758.00	230,499.00	361,357.00

Fraud detection protects customer and enterprise information, assets, accounts and transactions through the real-time, near-real-time or batch analysis of activities by users and other defined entities (Mallika, 2017). Fraud Incident Reporting refers to reported fraudulent incidents and suspicions where there has been, or could be, a breach of security; whether internal or external (e.g. perpetrated by customers, suppliers, or any other party), that impacts system users, regardless of materiality (Del Pozzolo *et al*., 2014).

Fraud detection is not intrusive to a user unless the user's activity is suspected. It tries to detect and recognize fraudulent activities as they enter systems and report them to a system manager. The reason for this failure is the fact that fraud detection must deal with some uniquely challenging properties of fraudulent activities which include: experience imbalance, online learning, adaptive adversaries, concept drift, noise, unequal misclassification costs and fast processing and large volume of data (Behdad *et al.*, 2012; Longe et al, 2019). The advantage of nature-inspired techniques over other artificial intelligence techniques lies in their ability to maintain a population of diverse individuals, similar to natural systems or ecosystems, each of which is good at something, and are flexible and capable of changing in response to changes in their environment.

The email has also proven to be the most popular means of perpetuating fraudulent activities such as advance fee fraud and is responsible for financial losses of up to $6 billion in the last 2 years (ACCC, 2017). Using the Artificial Neural Network (ANN), the Back Propagation (BP) algorithm is commonly used to perform the training task of fraud detection models, some drawbacks are often encountered using this gradient-based method which include: very slow training convergence speed and getting stuck in a local minimum easily (Nasser and Seyed, 2014). This study will attempt to develop a hybrid classification model for improving the detection of fraudulent emails by using a combination of nature-inspired algorithms with an existing artificial intelligence algorithm thereby improving the effectiveness of incident reports required for reporting suspected fraudulent activities. This classification model will be integrated into a framework used for the development of a fraud detection and incidence reporting system. A number of related works surrounding the area of electronic mail (e-mail) fraud and associated fraudulent activities such as credit card fraud, spam and phishing email detection by applying diverse number of machine learning algorithms were reviewed. Some of the papers focused on the application of heuristic machine learning algorithms while others focused on the application of meta-heuristic algorithms such as the nature inspired algorithms. The review is presented in the following paragraphs.

Zavvar *et al*., (2016), worked on the classification of spam mail using a combination of particle swarm optimization and artificial neural network and support vector machines. The features of 4601 e-mails were collected from UCI database were extracted using PSO algorithm and used to formulate the spam detection model using ANN and SVM. The features extracted using PSO outperformed the original features extracted from the e-mails. The study was limited to the feature extraction of email features using particle swarm optimization (PSO) for improving spam detection performance. Adwan and Abdelmunem (2016), worked on the classification of phishing emails using relevant features. Weights were assigned to phishing terms extracted from emails using information gain after text stemming following model building. The results of the study showed there was an improvement in the performance of the classification model using phishing terms extracted with high information gain values compared to using all terms extracted. Feature extraction of entire feature space of phishing email terms were limited to the use of information gain.

Kathiravan and Vasumanthi (2015) worked on the classification of phishing e-mails using Artificial Immune Systems (AIS). Information within the header and HTML body of sample e-mails were analyzed and contents extracted (using features parsing, stemming and tokenization) after which 15 features were extracted for 2981 spam and 4129 ham e-mails. The AIS was afterwards used to formulate the classification model compared with support vector machines (SVM). The AIS showed the capacity of detecting phishing e-mails based on the characteristic features extracted compared to general purpose spam filters. The model was limited to the use of SVM for the detection of phishing mails. Dilek et al., (2015), worked on the application of nature inspired algorithms for cyber-crime detection. A review of cyber-crime detection systems using artificial neural networks, genetic algorithms, artificial immune systems and intelligent agents. Nature-inspired algorithms have been used to develop systems that are flexible, adaptable, robust, and able to detect a wide variety of threats and make intelligent real-time decisions. The study was limited to the review of the application of nature-inspired algorithms to cyber-crime detection.

Choudhary (2015), worked on the classification of spam mails using a spam mail filtering algorithm. 2448 E-mails were collected with their features collected and extracted based on weights assigned using the probability of word outcome and compared with features extracted using genetic algorithms. The features extracted using genetic algorithms outperformed the features extracted using assigned weights. The study was limited to the comparison of feature extraction of email features using genetic algorithm and assigned weights. Nizamani (2014), worked on the classification of different categories of extracted features in fraudulent e-mails. Various kinds of features were extracted from 8000 e-mails and classified as family-related and financial-related e-mails following which the performance of each category of features were compared in terms of the fraudulent email detection rates. The results showed that equal detection rates as high as 96% were achieved using all features and using financial-related features which were better than the results of using family-related features alone. The study is limited to feature extraction of family-related and financial-related features using human judgement which can be biased.

2. Materials and Methods

This section presents the materials and method that were used for the design and development of the framework required for developing a fraud detection and incidence reporting system. The section presents the process of data identification and collection required for the identification of features needed for fraud classification. Also, the method of feature selection using the proposed nature-inspired algorithms were presented alongside the fraud detection model required for the classification of e-mails. The design and components of the proposed framework for guiding the development of a fraud detection and incidence reporting system was also presented.

2.1 Data identification and collection

The datasets required consists of fraudulent and non-fraudulent e-mails datasets which will composed of the header, email contents and the classification type (fraudulent and non-fraudulent emails). This study adopted the Spambase datasets and the CLAIR dataset both consisting of emails that have been already classified according to their respective class. The spambase dataset contains 4601 pre-processed datasets consisting of 1813 spam mails (39.4%) and 2788 ham mails (60.0%). Each dataset record is composed of 57 continuous features (frequency of words and characters in each email records) and a target class named 1 for spam and 0 otherwise. The CLAIR datasets consists of 2500 collection of raw fraudulent e-mails containing e-mail header information, the body of the e-mail text and the target class defining whether an e-mail has been opened or read. The CLAIR dataset requires the need of text preprocessing for the purpose of converting the unstructured format of the emails into a structured form for each email contents which were mapped to their respective class (Fraudulent or Not). Therefore, there are three (3) classes for identifying the emails collected from the two (2) data sources, namely: spam mail (suspicious but not necessarily fraudulent emails), ham mails (non-fraudulent emails) and fraudulent emails as shown in Table 1.

Table 1: Dataset Source and Dataset Size

S/N	Data Source	E-Mail Type	Frequency	Percentage (%)
1.	**Spambase**	Ham mail	2788	39.3
		Spam mail	1813	25.5
2.	**CLAIR**	Fraudulent mail	2500	35.2
			7101	100.0

2.2 E-mail text preprocessing

For the purpose of the development of the fraud detection model for emails, there is the need of converting the unstructured contents within the CLAIR e-mails collected into a structured format similar to the Spambase features (but not necessarily the same features). The features extracted are the words that are found within the body of the CLAIR email contents. The text preprocessing of emails require the use of the Python® Natural Language Tool-Kit (NLTK) for the purpose of performing the different text preprocessing stages required for converting the unstructured data into a structured format. Using the Python® NLTK, the unstructured CLAIR emails will be tokenized in order to convert every content of the emails into sets of words found within each email which will be followed by the removal of stop words from the extracted email contents. The stemming process is applied to the extracted contents in order to convert all words into their root word (for example, *families, familiar* becomes *family* and so on). The application of the stemming algorithms will reduce the feature space of the words in each document to their root words following which frequency of occurrence of the words found in each document if taken into account.

In order to be able to convert the preprocessed words in the emails into a structured form required for fraud detection, the words identified from both email samples (CLAIR and Spambase) will be used to form a term-document matrix, $\mathcal{D}_{ij}$ which represents the occurrence of each term w_i within each document d_j. In the term-document matrix $\mathcal{D}_{ij}$ the rows represent the occurrence (or absence) of a word w_i in a document. The value is 0 when the word does not occur in a document and the value if greater than 0 based on the technique of representation to be adopted. For this study, the Boolean model and the term frequency-inverse document frequency (TF-IDF) will be adopted. Using the Boolean model, weight $w_{ij} > 0$ is assigned to each term $w_i \epsilon d_j$ while for any term that does not appear in d_j then, $w_{ij} = 0$.

Using the TF-IDF, let *q* be this term weighting scheme, then the weight of each word w ∈ d is computed as shown in equation (1). In TF-IDF the term frequency is normalized by inverse document frequency, IDF. This normalization decreases the weight of the terms occurring more frequently in the document collection, making sure that the matching of documents be more effected by distinctive words which have relatively low frequencies in the collection. Following the process of creating the term-document matric for all 7101 email samples collected, the class of each email will be used to map each row of the term-document matrix. Therefore, the term-document matrix consisting of 7101 rows and *n* columns (features/words/terms from emails) is mapped to a fraudulent classification column vector consisting of 7101 values of either of {Ham, Spam, Fraudulent} by a function *f*. This relationship is presented by the function shown in equation (2).

$$q(w) = f_d(w) \cdot \log \frac{|\mathcal{D}|}{f_{\mathcal{D}}(w)} \quad (1)$$

$$f: \mathcal{D} \rightarrow \mathcal{F} \quad (2)$$

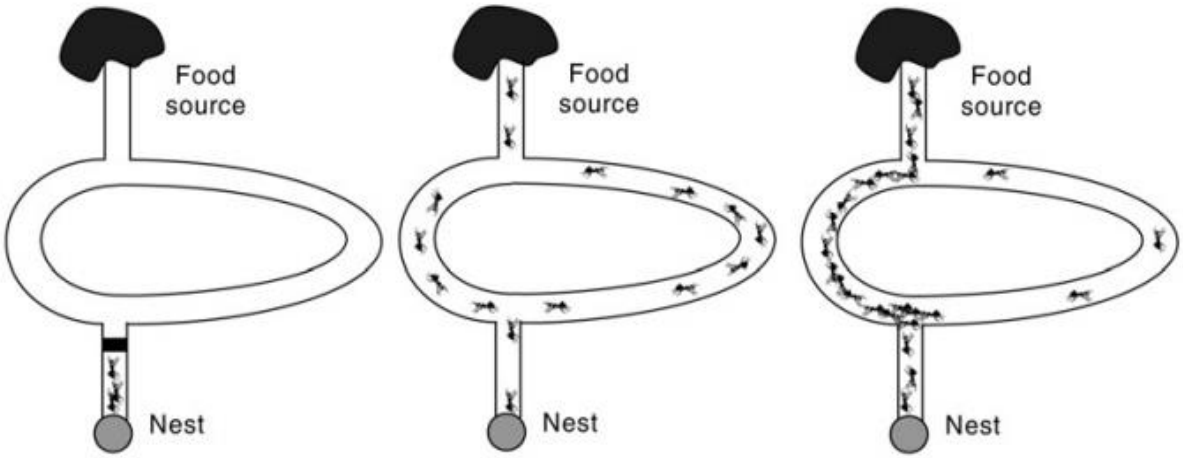
Defined as: $f(w_1, w_2, \ldots\ldots\ldots, w_n) = \begin{cases} Ham \\ Spam \\ Fraudulent \end{cases}$

2.3 Nature Inspired Swarm Intelligent Techniques

For this study among the identified nature-inspired algorithms available, the set that fall under the swarm intelligence (SI) were selected due to their appropriateness to this problem. Therefore, the ant colony optimization (ACO) and the particle swarm optimization algorithms (PSO) were selected. The algorithms were used to implement the feature selection algorithms and to implement the hybrid algorithm proposed. SI is a property of systems of unintelligent agents of limited individual capabilities exhibiting collectively intelligent behavior. SI includes any attempt to design algorithms or distributed problem-solving devices inspired by the collective behavior of social insects and other animal societies.

2.3.1 Ant colony optimization (ACO)

Ant Colony Optimization (ACO) techniques are based on the behaviour of real ant colonies used to solve discrete optimization problems (Babatunde *et al*., 2017). Therefore, ants can discover the best feature combinations as they proceed throughout the search space as shown in Figure 3 (left). Figure 3 (right) shows the pseudo-code for the ACO algorithm. In the ACO algorithm, features extracted from the e-mails are the ants within a colony. Initial pheromone values are the heuristic merit of the features extracted from e-mail contents while initial solution set is an empty set of optimal features. The pheromone value reflects the amount of information needed for the identification of the most relevant features needed for the classification of fraudulent e-mails. When an ant (feature) contributes to the high deposits of pheromone then the best solution contains the set of relevant features with the greatest deposit (optimal solution) of pheromone needed for the identification of fraudulent e-mails.



Input: Initialize a set of solutions; Initialize pheromone values (τ)
1 **while** *termination condition is not met* **do**
2 **for** *each ant* $i = 1$ *to* N **do**
3 Construct a solution for each ant;
4 Apply a local search for each solution (optimal);
5 Update the best-so-far solution s_{bs};
6 Utilized a pheromone updating strategy
Output: The best-so-far solution s_{bs}.

Figure 3: Foraging behaviour of Ants in Search for Food (left) and Algorithm for the Ant Colony Optimization (ACO) (right)
Source: de Castro, 2007; Robinson *et al*., 2005

2.3.2 Particle Swarm Optimization

Particle Swarm Optimization (PSO) mimics the flocking behavior of birds whenever they are in search for food (Gamal, 2016). Hence, they learn from the experience of their local best solutions and global best solutions as shown in Figure 4 (left) while Figure 4 (right) shows the pseudo-code for the PSO algorithm. In the PSO algorithm, features extracted from the e-mails are the swarm particles. The Stopping criteria is the amount of information about a swarm (fraudulent e-mail or not) that is possessed by a particle. The fitness value of each particle is the heuristic merit (initial velocity) possessed by the features extracted from e-mail contents and is used to evaluate the best position (local solution) of the particle and that of the swarm (global solution). If stopping criteria is not met then the amount of information missing in a feature (or particle) is determined by the amount of change in velocity needed to move a particle from its best position to that of the swarms'.

```
1 Initialize velocity and position randomly for each particle;
2 while the stopping criteria is not satisfied do
3     Calculate each particle's fitness value;
4     Determine each particle's best position, and the best position of entire swarm;
5     for each particle do
6         Update particle's velocity;
7         Update particle's position;
```

Figure 4: Swarms of Birds in Flight (left) and Algorithm for the Particle Swarm Optimization (PSO) (right)
Source: Gamal, 2016; Robinson *et al*., 2005

Let $\mathrm{E} = \{E_1, E_2, E_2 \ldots\ldots E_i\}$ be the set of E-mails gathered for fraud detection. Let Y= $\{Ham, Spam, Fraudulent\}$ represent the set of outcome for each e-mail collected. Let F_{ri} represent the set of features *r* extracted from e-mails *i* following tokenization. By applying ACO and PSO, the feature set of e-mail is reduced from *r* to m and n where m<r and n<r as shown in equation (3).

$$PSO(E_{F_{ri}}) = F_{mi} \qquad \mathrm{ACO}(E_{F_{ri}}) = F_{ni} \qquad (3)$$

It is the attempt of this study to show that using PSO or ACO to update the weights of back-propagation algorithm, the performance will improve. Therefore, the fraud detection model developed using the features extracted using ACO or PSO and used to develop the fraud detection model will perform better than the fraud detection model developed using the hybrid ACO/PSO back-propagation model which in turn will perform better than the fraud detection model developed using the features extracted by PSO/ACO and the ordinary back-propagation model according to equations (4a) and (4b).

$$BP_{ACO}(F_{ni}) > BP_{ACO}(F_{ri}) > BP(F_{ni}) > BP(F_{ri}) \qquad (4a)$$
$$BP_{PSO}(F_{mi}) > BP_{PSO}(F_{ri}) > BP(F_{mi}) > BP(F_{ri}) \qquad (4b)$$

The procedure provided above will be used for the development of the classification model for the detection of fraud in e-mails as shown in figure 5. The PSO and ACO algorithm are used to extract relevant features after which the features identified are used as a basis of developing the classification model using the hybrid back-propagation algorithm. The developed hybrid model will be compared in order to identify the most relevant and effective among the classification models.

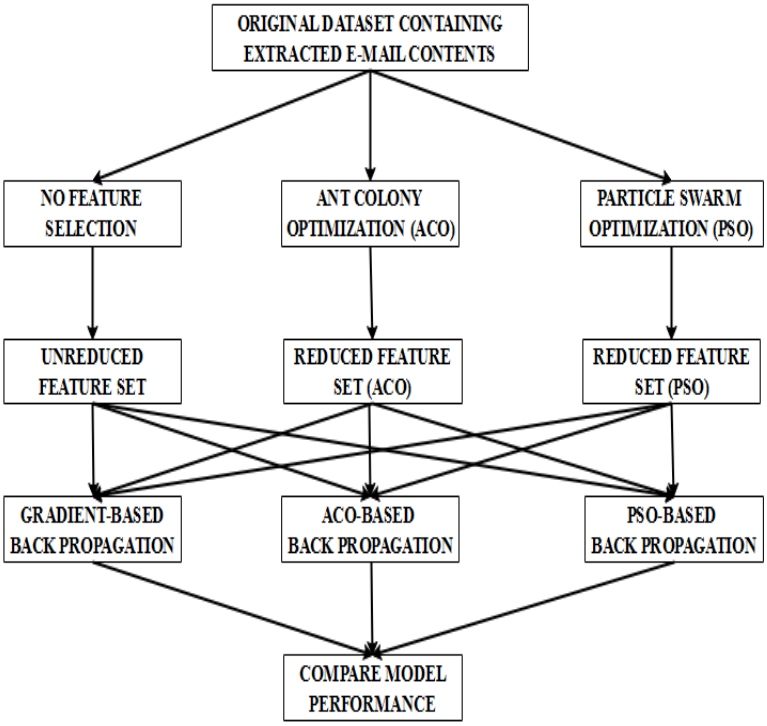


Figure 5: Conceptual Diagram of the Model Development Process

3. THEORY

During the feature selection process, the swarm intelligence algorithms selected for this study (PSO and ACO) will be handling the problem as an optimization problem with their own objective function and constraints. The objective function requires the selection of the optimal number of features w_r from the initial identified number of features w_n that will maximize the classification accuracy of the fraud detection problem according to equation (5).

$$\max_{x_r \subset x_i} \left(f\left(\sum_{j=1}^{r} \alpha_j x_j \right) \right) \; subject\ to: 0 < \alpha_j < 1 \qquad (5)$$

$$\alpha_j \epsilon \mathbb{R};\ x_j\ is\ the\ set\ of\ selected\ relevant\ features$$

During the back-propagation, the swarm intelligence algorithms (using PSO and ACO) will handle the problem of selecting the optimal value of weights of the artificial neural network as an optimization problem with their own objective function and constraints. The objective function requires the selection of the optimal value of weights w_{jk} attaching an input *i* to a node *k* that will minimize the mean square error of the back-propagation algorithm according to equation (6). By comparing the models developed according to figure 5, then one can propose the fraud detection model which can be used to classify e-mails according to equation (7).

$$\min_{w_{rs} \subset w_{ij}} \left(f\left(\sum_{k=1}^{r} \sum_{l=1}^{s} w_{kl} O_{kl} \right) - A_{kl} \right)^2 \; subject\ to: -1 < w_{kj} < 1 \qquad (6)$$

$$w_{kl} \epsilon \mathbb{R}\ such\ that\ k \epsilon (1, r)\ and\ l \epsilon (1, s);\ \ O_{kl}\ is\ set\ of\ node\ output$$

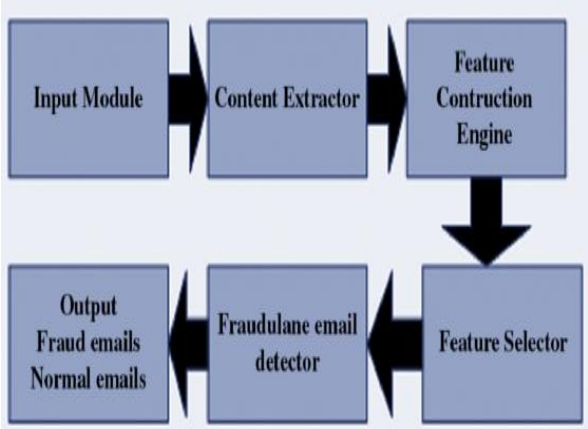
$$BP_{SI}\left(SI\left(E_{F_{ri}} \right) \right) = \begin{cases} Ham \\ Spam \\ Fraudulent \end{cases} \qquad (7)$$

3.1 Proposed Framework

The existing framework for implementing a fraud detection system that was applied by Nizamani *et al.* (2014) was reviewed. The existing system composed of six main modules, namely: input module, content extractor, feature construction engine, feature selector module, e-mail detector module and out module. Figure 6 shows the diagram of the existing framework on the left as provided by Nizamani *et al.* (2014) alongside the algorithm explaining the process on the right. The existing framework required the use of a content extractor which was used to extract the most relevant features which were then fed to the fraud detection module which was used to output the result of the fraud detection.

Following a thorough investigation of the behaviour of the existing system, a number of limitations were observed:

a. The existing fraud detection system could not adapt to changing features in e-mail writing styles;
b. The existing fraud detection system had no capacity to respond to false alarms;
c. The existing fraud detection system had no feedback mechanisms in place; and
d. The techniques used for feature selection and fraud classification were limited to the use of heuristic algorithms.



Algorithm 1	
Step1:	Input Email dataset E
Step2:	Extract email content
Step3:	Construct feature set from email content
Step4:	Make different feature selection
Step5:	For each set of features F_i
Step6:	For each set of classification algorithm C_j
Step7:	$Output_{i,j}$ = Fraud_detection (E, F_i, C_j)
Step8:	Display $Output_{i,j}$
Step9:	End loop at step5
Step10:	End loop at step6

Figure 6: Exisitng Framework for the Detection of Fraudulent E-mails
Source: Nizamani *et al.*, 2014

An incident reporting system has been proposed to complement the work of the existing framework for fraud detection system proposed by Nizamani *et al*. (2014). The proposed framework composes of three main components, namely: content and feature extraction module, fraud detection module and the incident reporting module. Figure 7 shows a description of the proposed framework for the development of the fraud detection and incident reporting system. In the proposed framework which composes of three main modules, each module will be responsible for handling the three main activities of the fraud detection system.

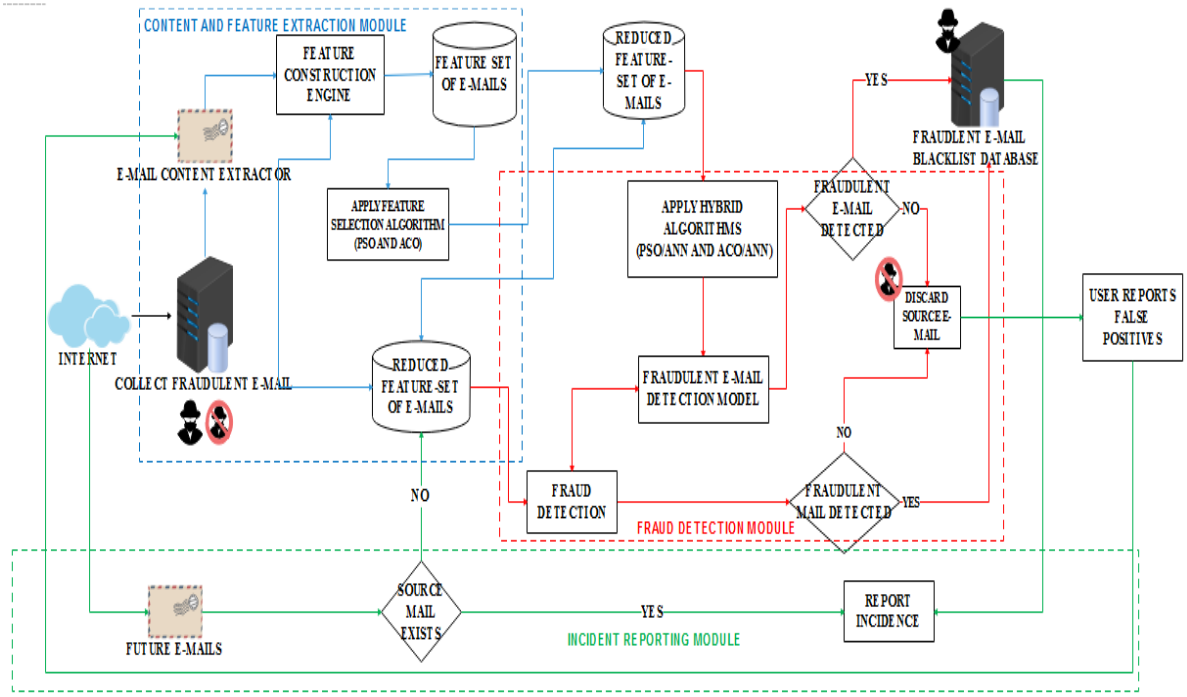


Figure 7: Proposed Framework for Fraud Detection and Incidence Reporting

First, the content and feature extraction modules is responsible for collecting incoming e-mail messages and are parsed in for tokenization, stop word removal for filtering tokenized terms, lemmatization and stemming processes for the pre-processing of incoming e-mails. This process converts unstructured e-mail messages into structures forms. Following this process, the PSO and ACO algorithms will be used for the extraction of the most relevant features out of the initially identified features and stored into a repository which will be required for identifying the presence or absence of identified terms in incoming e-mails messages.

Second, the fraud detection module is required for collecting information about the absence or presence of features identified in incoming e-mails and are parsed down to the hybrid swarm intelligent-based back-propagation algorithm for the detection of fraudulent e-mails. The source-emails of the fraudulent e-mails will be stored in a black-list database for the identification of fraudulent e-mails at the point of entry before applying the fraud detection algorithms. The implication of this is that the source IP of an incoming an e-mail can be initially used to identify an e-mail if it is found on the black-list database. If the source IP of the e-mail is not found on the black-list database and the e-mail was detected has been fraudulent then the source IP of that e-mail is added to the black-list database consisting of the list of the source IP of suspected fraudsters else the incoming e-mail is parsed to the inbox.

Third, the incidence reporting module will be required to correct the issues relating to false positives in the system. In case an e-mail whose source IP was not found in a black-list database and still identified by the fraud detection system as non-fraudulent can be reported by a client after discovery and the information about the features and the e-mail source IP can be used to review the fraud detection algorithm using the Swarm intelligent-based back-propagation algorithm proposed.

4. CONCLUSION

This paper has identified fraud as a serious barrier to the success of many individuals due to the increasing number of losses incurred by victims of financial fraud. The e-mail was also discovered as the most popular means of performing fraudulent acts among fraudsters. This led to the need of the development of an improved model for the detection of features in e-mail required for the effective detection of fraud alongside for the development of a framework that informs the development process of an incidence reporting system for detected fraudulent activities. The study will assist in the early identification of fraudulent activities thereby averting the onset of loss of valuables and money due to advance fee fraud. The study will provide a means of the early reporting of fraudulent activities through the use of emails. The study will also identify the relevant features present in fraudulent email activities thereby improving the classification accuracy of advance fee fraud detection.

This paper presented a proposed framework that can be used for the development of a fraud detection and incidence reporting system for the early detection and reporting of detected fraudulent e-mails. The framework depends on three components namely: the data collection and feature extraction module, fraud detection module and incidence reporting modules. The data collection and feature extraction module which collects incoming e-mails and extracts the necessary features via the process of natural language processing. The fraud detection module composes of two parts, a feature extraction part which handles the extraction of relevant features from incoming e-mails while the second part consists of the hybrid-model composing of the combination of the SI and back-propagation algorithm based on the identified features extracted from the e-mails. The incidence reporting modules allows a user to report the incidence of a fraudulent e-mail which was not detected which is used to re-evaluate the fraud detection model.

5. FUTURE WORKS

Future works in line with this study are focused at the extraction of the most relevant features required for the detection of fraud using the ant colony (ACO) and particle swarm optimization algorithms (PSO) alongside the development of the classification model which uses a hybrid ACO/PSO based back-propagation algorithm for the detection of fraudulent e-mails based on knowledge extracted from the dataset collected about fraudulent features and fraudulent e-mails. Finally, the fraud detection and incidence reporting system will be implemented based on the framework proposed for this study using Python-based technologies such as the Natural Language Toolkit (NLTK®) for e-mail content extraction and pre-processing alongside with the combination of Python® program, Python Flask® and Jinja® Templating framework for the implementation of the front-end of the fraud detection and reporting system. The study will contribute to knowledge by identifying the relevant features for fraud detection using nature-inspired algorithms. The study will formulate a hybrid model using swarm intelligence algorithms and back propagation algorithm for the classification of fraudulent e-mails. The study has implications in improving fraud detection in emails.

REFERENCES

1. Adwan, Y. and Abdelmunem, A. (2016). An Intelligent Model for Phishing Email Detection. *International Journal of Network Security and Its Applications 8*(4): 55 – 72.
2. Alexopoulos, P., Kafentzis, K., Athanassiadis, N., Benetou, X., Tagaris, T., Jollie, C. and Georgolios, P. (2007). An Adaptive Knowledge-Based Approach for Detecting Fraud across Different E-Government Domains. In Proceeding of the *International Conference on E-Business and Telecommunications 23*: 110 – 121.
3. Australian Competition and Consumer Commission (ACCC) (2017). Upfront Payment and Advance Fee Frauds. Accessed from https://www.scamwatch.gov.au/types-of-scams/unexpected-money/up-front-payment-advanced-fee-frauds on November 30, 2017.
4. Behdad, M., Barone, L., Bennamoun, M. and French, T. (2012). Nature-Inspired Techniques in the Context of Fraud Detection. *IEEE Transactions on Systems, Man and Cybernetics – Part C: Applications and Reviews 42*(6): 1273 – 1290
5. Choudhary, M. (2015). Automatic E-mail Classification Using Genetic Algorithm. International Journal of Computer Science and Information Technologies 6(6): 5097 – 5103.
6. de Castro, L.N. (2007). Fundamentals of Natural Computing: An Overview. *Journal of Physics of Life Reviews 4*: 1 – 36.
7. Del Pozzolo, A., Caelen, O., Le Borgne, Y.-A., Waterschoot, S. and Bontempi, G. (2014). Learned Lessons on Credit Card Fraud Detection from a Practitioner Perspective. *Journal of Expert Systems with Applications* 41: 4915 – 4928.
8. Dilek, S., Cakir, H. and Aydin, M. (2015). Applications of Artificial Intelligence Techniques to Combating Cyber Crimes: A Review. *International Journal of Artificial Intelligence and Applications (IJAIA) 6*(1): 21 – 39.
9. Gamal A. E. (2016). Nature Inspired Algorithms in Cloud Computing: A Survey. *International Journal of Intelligent Information Systems* 5(5): 60.
10. Kathiravan, A.V. and Vasumathi, B. (2015). Artificial Immune Based Classification Approach for Detecting Phishing Mails. *International Journal of Innovative Research in Computer and Communication Engineering 3*(5): 4308 – 4315.
11. Longe, O.B. Okorejior, M.S. & Etebong, B.I. (2019): Design Architecture for Training Random Forest Classifiers to Detect Phishing Attacks. Advances in Mathematical & Computational Sciences. https://www.isteams.net/mathsjournalMallika, R. (2017). Fraud Detection Using Supervised Machine Learning Algorithms. *International Journal of Advanced Research in Computer and Communication Engineering 6*(6): 6 – 10.

12. Nasser, M. and Seyed, J.M. (2014). Comparison of Particle Swarm Optimization and Back propagation Algorithms for Training Feed-forward Neural Network. *Journal of Mathematics and Computer Science* 12: 113 – 123.
13. Nizamani S, Memon N and Wiil U.K. (2011) Detection of Illegitimate Emails Using Boosting Algorithm. *Counterterrorism and Open Source Intelligence*. Vienna: Springer: 249 – 264.
14. Nizamani S, Memon N, Wiil UK and Karampelas P. (2012). Modeling Suspicious Email Detection using Enhanced Feature Selection. *International Journal of Model Optimization 2*(4): 371 – 377
15. Nizamani, S., Memon, N., Glasdam, M. and Nguyen, D.D. (2014). Detection of Fraudulent Emails by Employing Advance Feature Abundance. *Egyptian Informatics Journal 15*: 169 – 174.
16. Reich, P. (2004). Advance Fee Schemes in Country and Across Borders. In Proceeding of Crime in Australia: International Connections. Conference organized by Australian Institute of Criminology, Melbourne, Australia.
17. Robinson D. G. (2005). Reliability Analysis of Bulk Power System using Swarm Intelligence. In *IEEE.* 96-102
18. Tive, C. (2006). *419 Scam: Exploits of the Nigerian Con Man*. Bloomington: iUniverse.
19. Zavvar, M., Razaei, M. and Garavand, S. (2016). E-Mail Spam Detection Using a Combination of Particle Swarm Optimization and Artificial Neural Network and Support Vector Machine. *International Journal of Modern Education and Computer Science 7*: 68 – 74.