

AN OVERVIEW OF THE APPLICATIONS OF DATA MINING TECHNIQUES TO NATIONAL SECURITY – NIGERIA AS A CASE STUDY

Ogwueleka, Francisca N

Department of Computer
Science, Nigerian Defense
Academy, Kaduna.
Nigerian Defense Academy,
Kaduna, Kaduna State.
ogwuelekafn@gmail.com

Damwesh Tapshik T

Department of Computer
Science, Nigerian Defense
Academy, Kaduna.
Nigerian Defense Academy,
Kaduna, Kaduna State.
damweshtt@gmail.com

ABSTRACT

Data mining involves getting useful, unseen knowledge from large datasets in a form that humans can comprehend. It has gained a lot of applications in the advertising of goods and services, marketing, artificial intelligence research, biological science, and national security. This paper provides an overview of the various relevancies of this technological feat, especially how it could be credibly harnessed and utilized in the interest of national security in Nigeria. It focused on counter-terrorism and reviewed some data mining techniques that can be applied for crime analysis and control. The noted data mining techniques used in national security form research include classification, clustering, link analysis, string comparator, and deviation detection. The paper also reviewed solutions to data mining problems and data privacy legislation in Nigeria and identified the need for proper legislation. Data mining techniques and algorithms can be applied to combat, prevent and detect crime, which would improve national security, especially in terms of counter-terrorism. Data mining is however not an end in itself but a means to an end and although it has some negatives effects when not used properly, most of these effects can be reduced to a large extent by proper data protection legislation.

Key words: Data mining, national security, counter terrorism, security, data protection legislation, terrorism

1. INTRODUCTION

The capability for generating and collecting data from a variety of sources has greatly been enhanced by the computerization and digitalization of the society. Almost every aspect of our lives is characterized by a flood of a tremendous amount of data. The high availability of data has generated great and urgent need for new techniques and automated tools intelligent enough to assist in the transformation of this huge data into useful information and knowledge (Han et al., 2012). Thuraisingham (2002) defined data mining as “the process of posing queries and extracting useful patterns or trends often previously unknown from a large amount of data using various techniques such as those from pattern recognition and machine learning”. This implies that unimaginable information can be inferred or derived from this huge collection of data if the right data mining tools and techniques are applied to such a large data set.

Data mining has recently gained a lot of attention and interest from both researchers and the industrial community. This is due to the immeasurable significant benefits that data mining offers. For example, data mining could be applied to detect fraudulent behavior, unusual patterns, and terrorist activities (Thuraisingham, 2004). There is no doubt that in this 21st century of science and technology, data mining and application can contribute immensely to humans and their sustainable development in every facet of life. If used appropriately, data mining can save lives and improve national security. Although data mining has a lot of important applications, such as in biomedical and deoxyribonucleic acid (DNA) data analysis, financial data analysis, retail industry (Bagga and Singh, 2012), corporate surveillance, intrusion detection, lie detection, fraud detection, and criminal investigation, this same technology raises great privacy and national security concerns (Ahmed, 2013). This is because, as much as data can be mined for positive reasons, it can also be mined for negative reasons. For example, patterns in surveillance data can be mined to know when best to stage a terrorist attack.

This paper reviews the various possible relevancies of this technological feat, especially as to how it could be credibly harnessed and utilized in the interest of national security in Nigeria. It explores its potential dangers and examines possible ways forward towards using data mining for the positive sustainable development of Nigeria, in this era of intense insecurity. It also highlights the need for proper data privacy protection legislation.

2. RELATED LITERATURE

Research and studies have been carried out on crime data mining in the recent decade. The areas of research have been in predicting, detecting and analyzing crime data. Tolan and Soliman (2015) applied classification to predict terrorist groups responsible for attacks in Egypt from 1970 to 2013, comparing the results of five different classifiers namely; Support

Vector Machine (SVM), K-Nearest Neighbour (KNN), Tree induction (C4.5), Iterative Dichotomiser (ID3) and Naïve Bayes (NB). They obtained the highest accuracy of 72.5352% from KNN but concluded that SVM is more acceptable than other classifiers in terms of overall performance. Saltos and Cocea (2017) predicted the frequency of different types of crime by Lower Layer Super Output Areas (LSOA) code (an administrative system of areas used by the UK police) and frequency of anti-social behavior crime. They used data from the UK police that have over 600,000 records and applied three different algorithms namely; instance-based, regression and decision tree. The result obtained in terms of predictive performance and processing time shows that decision tree (MP5 algorithm) is more reliable in predicting crime frequency and frequency of anti-social behavior. An interactive query based interface has been proposed by Gupta et al. (2008) as a crime analysis tool to assist Indian police in their activities. The system is used to extract useful information from the comprehensive crime database maintained by National Crime Record Bureau (NCRB) to find crime hot spots using crime data mining techniques such as clustering. The efficiency of the system has been illustrated in Indian crime records.

3. NATIONAL SECURITY THREATS

National security is the security of a nation-state, including its people, economy and institutions, which is considered a responsibility of government (Onyemachi and Stephen, 2014). There are different dimensions of national security, which includes physical security, economic security, political security, ecological security, computer security, security of energy and national resources, and infrastructure security, all of which could be threatened. Threats to security could be through terrorism, extremism, sabotage, information warfare, natural disaster, cybercrime and ransomware, climate change, espionage, proliferation of weapons of mass destruction, and perception management, which entails manipulating data, disseminating false information and correspondence aimed at distorted public perception. These threats have the potential to hamper the peaceful coexistence of any nation.

3.1 Security Concerns with Data Mining

Data mining techniques while providing the ability to extract hidden knowledge, on one hand, introduce several security and privacy threats on the other hand (Niranjan and Nitish, 2017). Security from a computer standpoint is the ability a system must pose to protect data or information and its resources in terms of confidentiality, integrity, and authenticity (Niranjan and Nitish, 2017). Problems in database security can be handled by applying data mining techniques (Lin et al., 1996). On the other hand, these same techniques can be employed to cause security problems (Lin et al., 1996). Thuraisingham (2002) reviewed security both from positive and negative aspects. According to her data mining can be a threat to privacy because data mining tools can easily be accessed and are available on the web and other sources. This implies that even ordinary users can apply these tools to extract information from different sources, thereby increasing the tendency of violating the privacy of individuals, where privacy often refers to the right of individuals to have power over information about themselves. The emergence of technology for Knowledge Discovery and Data Mining (KDDM) has re-energized concern about general privacy issues such as the secondary use of personal information, handling misinformation, and granulated access to personal information (Brankovic and Estivill-Castro, 1999).

Since the announcement of information sharing and data mining initiatives, more attention has been focused on the implications for privacy (Serfert, 2007). These privacy concerns are both focused on actual projects proposed and the possibility for data mining applications to be used for purposes beyond the original, a concept referred to as mission creep (DeRosa, 2004). Observers have different opinions about privacy issues that relate to data mining. Some opinions are that “tradeoffs may need to be made regarding privacy to ensure security. Serfert (2007) suggested that the degree to which government agencies should use and mix commercial data with government data should be considered.

Thuraisingham (2002) identified “inference problem” as a security concern for data mining which has been discussed by several other researchers. Thuraisingham (2002) described inference as the process where users pose queries and deduce unauthorized information from legitimate responses that they receive. She maintained that data mining tools make inference problem quite dangerous. Without the availability of these tools, it will be quite difficult for ordinary users to deduce information because they have to be complicated in their reasoning. Ramasastry (2005) also identified adverse inference as a problem that can be caused by governmental data mining especially in a case where a criminal suspect is innocent. Often, certain rights such as access to governmental services can be wrongly denied to citizens as a result of an adverse inference. Ramasastry (2005) also opined that the more data government mines, the larger the number of adverse inferences that may be drawn.

Mission creep has been identified by Seifert (2008) as one of the leading risks of data mining, which is always referenced by libertarians. Seifert (2008) described mission creep as the use of data for a purpose that was not originally intended or collected, which usually can occur regardless of the method of collection. Some reasons for mission creep as identified include the sense of urgency usually posed by efforts to fight terrorism and pressure on government officials concerned with the safety of others. These problems may lead to combining databases to find potential threats. The danger with such is that it can lead to unintended outcomes and can produce misleading results.

False Positive is also a concern with data mining. Ramasastry (2006) stated that when data mining generates an incorrect inference, the outcome is referred to as “False Positive”. According to DeRosa (2004), one significant concern with data

mining is the tendency that law enforcement agencies might get it wrong and innocent people will be stigmatized as terrorist or criminals simply because they exhibit unusual patterns of behavior or have some link to suspected terrorist or criminals as the case may be. It is termed false because the person in question is not truly exhibiting the behavior that the data mining system is supposed to uncover while it is positive because the system produced a result by flagging that person as a possible risk. Since False Positive incorrectly identifies people as matches or links, the danger is that government actors might be tempted to take action based on the result of data analysis queries alone. These actions could include arrest, detention, and denial of benefits. Even if mistakes are later corrected by the government, damage to reputation could already be done which will have long-term negative consequences to the individual (DeRosa, 2004).

3.2 National Security and Data Mining Solutions

Terrorism is both a national and international concern. Most national security issues are those that relate to terrorism. Thuraisingham (2004) broadly categories terrorist activities under non-information related, information related, bioterrorism, and chemical attacks. Most data mining solutions to terrorism are those that carefully examine a dataset and make positive predictions. These techniques are efficient for handling particular problems as circumstance warrants (Okonkwo and Enem, 2011). Data mining solutions to terrorism are identified by Okonkwo and Enem (2011), Thuraisingham (2003) and Tamilselvi et al. (2013) to include data mining techniques such as making associations, link analysis, forming clusters, classification and anomaly detection. For example, classification, when applied to a dataset, can identify and group suspects of a crime based on predefined criteria. Likewise, a link analysis algorithm can build links between suspected individuals. It can analyze the suspect's activities, associations, and relations by forming a link of their activities. Another application of data mining is in intrusion detection. According to Niranjana and Nitish (2017) intrusion is any action that compromises the availability, integrity, and confidentiality of one or more computer resources. An intrusion detection system (IDS) by employing data mining techniques can discover consistent patterns of features of a system that are useful and can detect anomalies and known intrusions using a relevant set of classifiers. Using some basic data mining techniques such as classification and clustering, intrusion can be detected easily (Khan et al., 2012). Other National Security concerns include privacy, inference, mission creep, and false positive. These issues can amount to serious security in the long run. For this category of security concerns, proper data protection legislation will go a long way in reducing such threats (Akinsuyi, 2015).

In addition to proper data protection legislation, security challenges posed by data mining can also be minimized by employing privacy preserving data mining (PPDM). PPDM aims at safeguarding sensitive information from unwelcome or unsanctioned disclosure (Zang et al., 2011). PPDM generally composed of different approaches that can be applied to enforce privacy. These approaches include suppression, randomization, data aggregation, data swapping, and noise addition or perturbation.

It is worth noting that although data mining could contribute towards national security, including counterterrorism, it is not a claim that data mining will solve all national security problems. However, the ability to extract hidden patterns and trends from large quantities of data is very important for detecting and preventing terrorist attacks (Thuraisingham, 2004).

3.3 Data Privacy Legislation

Privacy is one serious concern of data mining. Efforts have been in place as far back as 1974 to protect data by some government, such as the Federal Privacy Act of 1974 of the U.S government, which applies to commercial data brokers who supply information to the government (Ramasastry, 2006). The United Kingdom (UK) has an effective data protection and privacy legal environment. The UK also has a strong regulator which recently allowed its courts to rule that Google had used cookies without the consent of users to track people surfing the web using Apple's popular Safari browser, contrary to the promise made by Google in its privacy policy. The English court ruled that Google's act was unlawful under English privacy law and found that Google has violated both English data protection act and committed a tort of misuse of private information (Akinsuyi, 2015).

In developing countries such as Nigeria on the other, there is no adequate data protection legislation, as stated by Jemiluhon and Akomolede (2015). Although it could be argued that the 1999 constitution of the Federal Republic of Nigeria guarantees the protection of privacy of its citizens, it is worthy to note that the second schedule to the Constitution, which deals with legislative powers mentions no such thing as information technology directly (Jemiluhon and Akomolede, 2015). Akinsuyi (2015) affirmed that Nigeria has two data protection related bills (one dated 2008 and the other 2010), yet none has been passed into law. According to Jemiluhon and Akomolede (2015), several bills have been drafted that should have addressed areas bothering information and communication technology in Nigeria, but none have been passed into law. These bills include the Computer Security and Critical Infrastructure Protection Bill 2005, the Cyber Security and Data Protection Bill 2008, the Electronic Fraud Prohibition Bill 2008, the Nigerian Computer Security and Protection Agency Bill 2009, the Computer Misuse Bill 2009, and the Economic and Financial Act (Amendment) 2010.

A glaring gap exists in Nigeria's data protection and privacy law landscape which according to Akinsuyi (2015) is not just frustrating and embarrassing but also raises economic and national security concerns when it comes to processing of

Nigerian citizen's data. Although some legislative efforts in the Nigerian landscape resemble data protection legislation, Jemiluhon and Akomolede (2015) pointed out limitations in many ways. The resemblance of these data protection legislation includes the Official Secrets Act, the Freedom of Information Act, the National Information Technology Development Agency (NITDA) Draft Guidelines for data protection, and most recently, the Nigeria Data Protection Regulation 2019 by NITDA. Although NITDA has drafted several guidelines, drafted guidelines are only a draft and have no operative effects. It is not like an enforcement social contract similar to a legislation that creates rights and liberties (Jemiluhon and Akomolede, 2015). Jemiluhon and Akomolede (2015) hold the opinion that drafts do not replace proper legislation, and therefore, there is a need for adequate legislation for data protection in Nigeria.

4. DATA MINING TECHNIQUES FOR COUNTER-TERRORISM

One of the critical applications of data mining technologies is in counter-terrorism (Thuraisingham, 2003). The criticality of national security cannot be overemphasized as security is bedrock to any meaningful development. Data mining provides different techniques and algorithms for data analysis and scrutiny. The circumstance at hand is a determinant of the technique to use (Okonkwo and Enem, 2011). This section discusses some data mining techniques for crime data analysis and counterterrorism, which include, link analysis, forming clusters, classification, and anomaly detection (Thuraisingham, 2004).

4.1 Classification

Classification is the most commonly applied data mining technique as stated in Tamilselvi and Kalaiselvi (2013). It is a data mining technique used to predict group membership for data instances. The technique makes use of a set of predefined criteria to come up with a model capable of classifying the population of records (Tamilselvi and Kalaiselvi, 2013). Given a database labeled with predefined classes or categories, the task is to learn from these objects a model that separates the objects into the predefined categories. Then for a new object, the learned model is applied to assign this new object to one category (Zaki and Wong, 2004). According to Okonkwo and Enem (2011) when classification is used, an existing dataset can easily be understood, and it will help to predict how new individuals and events will behave based on the classification criteria.

There has been a rise in the cases of violent clashes between farmers and herders in Nigeria between 2014 and 2018, as stated by Igboke (2019). Assuming a scenario where a report has been received by a law enforcement agency of an attack on a farming community, the agency may try to form the idea of the attackers. For instance; 20 males, with the age range between 21 and 34, between 4inch and 6inch tall, and speaks Hausa fluently. Based on these criteria, it is imperative to place all males that meet these criteria under proper observation. The classification algorithm to be adopted will be such that it divides the population into two distinct parts; male and female. The algorithm will also segment males based on their suspected criteria (Okonkwo and Enem, 2011). The male dataset could be further segmented. Next, historical dataset is then scanned for related matches. In a situation where the search yields no positive result, the search could be narrowed by reclassifying the already classified dataset. In a case where no class is found, the data warehouse will then be updated with the new criteria. If another related crime occurs, it becomes easier to form a match (Okonkwo and Enem, 2011). This process continues until we have a dataset that can be matched against the criteria. If there is a related match, it is placed under surveillance else the process continues until there are no more matches.

4.2 Link Analysis

Link analysis (LA) is a data mining technique that uses various graph theoretic techniques. LA is essentially about analyzing graphs as stated by Thuraisingham (2004). It is very similar to social network analysis (SNA) in the sense that both handle relationships in terms of nodes and edges, and try to establish who is more important by analyzing the whole network, not individual events (Paul, 2013). The theory behind link analysis is that events are linked to one another and therefore mutually exclusive (Okonkwo and Enem, 2011). LA often involves seeing associations between suspects and locations, be it physical, on a network or the Internet through chart and map. LA is a process of "building up a network of interconnected objects through relationships to explore patterns and trends". It is essential in identifying patterns, merging groups and the quick generation of connections (Memon and Qureshi, 2005). LA can be applied by investigators and intelligence analysts to graphically examine anomalies and inconsistencies; and establish a network of relationships and hidden contacts. Since LA theory suggests that events are linked to one another, how to relate these events is now the concern. For example, how do we relate a bank transfer to another one or a telephone call to another?

When a crime is analyzed using classification technique, the analysis may produce several suspected individuals to put under surveillance. The issue is how to arrive at some suspected individuals. LA could be used to further analyze suspect's activities, associations, and relations by forming a link of their activities. These links might be in the form of emails exchanged, places visited, bank transactions and other tangible entities. The link analysis algorithm can build links between suspected individuals (Okonkwo and Enem, 2011). One problem with link analysis is that it can be so ambiguous and overcrowded as stated by Okonkwo and Enem (2011). The challenge is to reduce the graph into manageable chunks. The graph needs to be reduced so that the analysis is manageable and not combinatorially explosive (Thuraisingham, 2004). Using LA, a chain can be followed to find links. Take, for instance, Amos is seen with Marcus. Marcus is a friend with Abu, Abu and James are always together and James has a criminal record. What then can be inferred of Amos? It helps answer questions like who knows whom and when and where they were in contact.

4.3 Clustering

Another data mining technique is clustering, also called unsupervised learning (Zaki and Wong, 2004). As stated by Tamilselvi and Kalaiselvi (2013) clustering is descriptive tasks that identify a finite set of categories to describe data. The goal of clustering is to group similar objects into one cluster and dissimilar objects into another cluster, based on characteristics of data also known as segregation (Okonkwo and Enem, 2011). With clustering, given a database of objects without any predefined categories, we are required to partition

the objects into a group of elements that share a common set of properties. Clustering can be applied to automatically partition individuals into a defined group. The reason for clustering is that it permits the construction of simpler models of the world. Although classification and clustering share similarities, there exist a major difference between them. While classification is supervised, clustering is unsupervised. Clustering divides a population not based on predefined conditions, while classification divides a population, based on some predefined conditions.

Clustering can effectively be applied to detect crime and terrorism. Although it is similar to classification, it differs in the sense that it does not require any criteria or parameters. It is based on a real-life model that individuals with certain qualities must cluster together. Crime and acts of terrorism in Nigeria include militancy, arm robbery, kidnapping, ritual killings, cattle rustling, and most recently herdsmen attacks on farming communities. All these acts of terror can be effectively clustered. There are areas more prompt to these activities. For instance; militancy can be clustered to the Niger Delta region, kidnapping can be clustered to the southeast and some parts of northern Nigeria. Ritual killings can be clustered to the southwest as well as bank robbery, which is most common in Lagos areas. Cattle rustling and herdsmen attacks can be clustered to the north. According to Okonkwo and Enem (2011) individuals that partake in a certain crime cluster together even within crime dominated areas. For example, individuals that specialize in arm robbery will cluster together and drug traffickers usually have syndicates with interconnection amongst them. Clustering algorithms can identify clusters and their areas of operation. Law enforcement agencies can look for related clusters and examine them anytime a crime is reported. It is possible for clustering anomalies to occur, for instance, cattle rustling common in the north, suddenly happens in the south. When an anomaly occurs, it shows the starting of another cluster (Okonkwo and Enem, 2011).

4.4 String Comparator

String comparator is a data mining technique that shows the relation among textual fields in pairs of database records (Hosseinkhani et al., 2013). It computes similarities between database records by comparing their textual fields and can help detect deceptive information such as names and addresses, Bank Verification Number (BVN) or social security numbers in criminal records (Qayyum et al., 2018). String comparator, however, involves a large amount of computation and returns a numerical value by comparing two strings (Gupta, 2014). Investigators can employ the use of string comparator to analyze textual data. This technique can be effectively used in Nigeria to detect deceptive information obtained from processes such as SIM registration where criminals register multiple SIM cards with false information for the sake of perpetrating a crime.

4.5 Deviation Detection

Deviation detection also referred to as outlier detection, studies data that differs clearly from the rest of the data set (Hosseinkhani et al., 2013). The technique is particularly useful for crime detection and other crime analysis (Gupta, 2014). Areas that deviation detection can be applied by investigators include network intrusion detection and fraud detection. It is, however, difficult to identify outliers because such activities may also appear to be normal making it difficult to recognize outliers.

4.6 Comparison of Crime Data Mining Techniques

Table 1. Comparison of data mining techniques for crime data analysis

Technique	Timeliness	Data Dependency	Accuracy
Classification	Less time consumption	Complete training set data required	Accurate with data availability
Clustering	Moderate time consumption	Less data dependency	Accurate
Link Analysis	Moderate time consumption	Moderate	Tendency of ambiguous and information overload
String Comparator	Moderate time consumption	Moderate	Accurate
Deviation Detection	Less time consumption	Moderate	Accurate

Table 1 is a comparison of the different data mining techniques for crime data analysis in terms of timeliness, data dependency, and accuracy. While string comparator and deviation detection have a moderate dependency on data with optimal accuracy and consumes a moderate and minimum time respectively, clustering has moderate timeliness with less data dependency and optimal accuracy. Classification consumes less time, requires complete training data set, and its accuracy is dependent on the availability of data.

5. CONCLUSION

National security is a concern crucial to the harmonious existence of any nation. Combating crime and terrorism to ensure national security in this information and communication age requires law enforcement agencies to be fully informed and equipped with available technologies especially the use of computers and computer analysis. Data mining techniques and algorithms can, and have successfully been applied to combat, prevent and detect crime by identifying activities of criminals and tracking them down. It is, however, important to note that data mining is not a solution to all terrorism and security problems. Data mining can only help detectives and law enforcement agencies but not replace them. To ensure effective results for counter-terrorism, data miners have to work with all parties involved in ensuring national security. Data mining certainly has the tendency to intrude privacy, a concern which has generated significant fear and

controversy. Data mining is too valuable to be rejected outrightly but embracing it with no guidelines or control for its use poses a lot of risk and the private information it analyzed, will be misused. It is therefore important for legislators to understand data mining and come up with necessary legislation for data mining.

6. REFERENCES

- [1] Ahmed, N. (2013). Security Issues in Web Data Mining, National Security: A Survey. *International Journal of Advanced Research in Engineering and Applied Sciences*. 2(11), 227-252.
- [2] Akinsuyi, F. F. (2015). Data protection and privacy laws Nigeria, a trillion dollar opportunity!!. Available at SSRN 2598603.
- [3] Berson, A., Smith, S., & Thearling, K. (1999). Building data mining applications for CRM. McGraw-Hill Professional, 11 – 20.
- [4] Brankovic, L., & Estivill-Castro, V. (1999, July). Privacy Issues in Knowledge Discovery and Data Mining. *In Australian Institute of Computer Ethics Conference*, 89-99.
- [5] DeRosa, M. (2004). Data Mining and Data Analysis for Counterterrorism. Center for Strategic and International Studies Report. CSIS press. Washington D.C, 20 – 31.
- [6] Gupta, G. K. (2014). Introduction to data mining with case studies. PHI Learning Pvt. Ltd..
- [7] Gupta, M., Chandra, B., & Gupta, M. P. (2008). Crime data mining for Indian police information system. *Computer society of India*, 40(1), 388-397.
- [8] Hosseinkhani, J., Ibrahim, S., Chuprat, S., & Naniz, J. H. (2014). Web crime mining by means of data mining techniques. *Research Journal of Applied Sciences, Engineering and Technology*, 7(10), 2027-2032.
- [9] Igbokwe, C. C., & Iwuoha, C. A. (2019). The rising Herdsmen Militancy and Security Challenges in Nigeria. *Villanova Journal of Social Sciences, Arts and Humanities*, 1(1), 35-46.
- [10] Jackson, J. (2002). Data Mining: A Conceptual Overview. *Communication of the Association for Information Systems*. 8, 267-296.
- [11] Jemilohun, B., and Akomolede, T. (2015). Regulations or legislation for Data Protection in Nigeria? A call for A Clear Legislative Framework. *Global Journal of Politics and Law Research*. 3(4), 1-16.
- [12] Khan, S., Sharma, A., Zamani, A. S., & Akhtar, A. (2012). Data Mining For Security Purpose & Its Solitude Suggestions. *International Journal of Scientific & Technology Research*, 1(7), 1-4.
- [13] Lin, T. Y., Hinke, T. H., Marks, D. G., & Thuraisingham, B. (1996). Security and data mining. In *Database Security IX* (pp. 391-399). Springer, Boston, MA.
- [14] Maxwell, T. A. (2005, January). Information policy, data mining, and national security: False positives and unidentified negatives. In *Proceedings of the 38th Annual Hawaii International Conference on System Sciences* (pp. 134c-134c). IEEE.
- [15] Memon, N., and Qureshi, A. (2005). Investigative Data Mining and its Application in Counterterrorism. *Proceedings of the 5th WSEAS International Conference on Applied Informatics and Communications*, Malta, 397- 403.
- [16] Niranjana, A., & Nitish, A. (2017). Security in Data Mining-A Comprehensive Survey. *Global Journal of Computer Science and Technology*. 16(5), 51-72.
- [17] Okonkwo, R. O., & Enem, F. O. (2011). Combating crime and terrorism using data mining techniques. In *10th International conference IT people centred development, Nigeria Computer Society, Nigeria*, 121 - 132.

- [18] Paul, S. (2013). On Some Aspects of Link Analysis and Informal Network in Social Network Platform. *International Journal of Computer Science and Mobile Computing*, 2(7), 371-377
- [19] Qayyum, S., & Dar, H. (2018). A Survey of Data Mining Techniques for Crime Detection. *University of Sindh Journal of Information and Communication Technology*, 2(1), 1-6.
- [20] Onyemachi, N.P., & Stephen, N.T. National Security Issues and Challenges in Nigeria: Which way forward, 1(1), 96-106.
- [21] Ramageri, B. (2010). Data Mining Techniques and Application. *Indian journal of Computer Science and Engineering*, 1(4), 301-305.
- [22] Ramasastry, A. (2005). Lost in Translation-Data Mining, National Security and the Adverse Inference Problem. *Santa Clara Computer & High Tech. LJ*, 22, 757.
- [23] Saltos, G., & Cocea, M. (2017). An exploration of crime prediction using data mining on open data. *International Journal of Information Technology & Decision Making*, 16(05), 1155-1181.
- [24] Seifert, J. W. (2008, April). Data mining and homeland security: An overview. LIBRARY OF CONGRESS WASHINGTON DC CONGRESSIONAL RESEARCH SERVICE.
- [25] Sharma, B., Kaur, D., and Manju. (2013). A Review of Data Mining: Its Challenges, Issues and Applications. *International Journal of Current Engineering and Technology*, 3(2), 695-700.
- [26] Tamilselvi1, R and Kalaiselvi, S (2013). An Overview of Data Mining Techniques and Applications. *International Journal of Science and Research (IJSR)*, India Online. 2(2), 506-509.
- [27] Thuraisingham, B (2004). Data Mining for Counter-terrorism. *Data Mining: Next Generation Challenges and Future Directions*. AAAI/MIT Press, 157-183.
- [28] Thuraisingham, B. (2002). Data Mining, National Security, Privacy and Civil Liberties. *ACM SIGKDD Explorations Newsletter*, 4(2), 1-5.
- [29] Thuraisingham, B. (2003). *Web Data Mining Technologies and Their Applications in Business Intelligence and Counterterrorism*, CRC Press, Florida, 191-217.
- [30] Tolan, G. M., & Soliman, O. S. (2015). An experimental study of classification algorithms for terrorism prediction. *International Journal of Knowledge Engineering-IACSIT*, 1(2), 107-112.
- [31] Zaki, M., and Wong, L. (2004). Data Mining Techniques. *Selected Topics in Post-Genome Knowledge Discovery*, 125-163.
- [32] Zang, H., & Bolot, J. (2011, September). Anonymization of location data does not work: A large-scale measurement study. In *Proceedings of the 17th annual international conference on Mobile computing and networking* (pp. 145-156). ACM.

Author's Brief Profile



Francisca Nonyelum Ogwueleka is a Professor of Computer Science and Cyber Security in Nigerian Defence Academy, Kaduna-Nigeria. Her research focuses on big data, artificial intelligence, cloud security, data mining techniques, steganography, penetration testing solutions and information security. She has nine published books, two book chapters and ninety-five original articles in international and national journals. She is an external examiner for PhD dissertation and M.Sc thesis evaluation for national and international universities. She was a resource person for the development of Nigeria Undergraduate Curriculum in Cyber Security and Information Technology programmes for National Universities Commission.



Damwesh Tapshik Thompson is currently a Masters student in the Department of Computer Science at the Nigerian Defence Academy, Kaduna, Kaduna State, Nigeria. His research areas include Data Mining, Real-time Information Systems and Computer Networks. He can be reached through E-mail damweshtt@gmail.com