BILINGUAL PUBLISHING CO.
Pioneer of Global Academics Since 1984

Semiconductor Science and Information Devices
https://ojs.bilpublishing.com/index.php/ssid

REVIEW

Cybersecurity and Cyber Forensics: Machine Learning Approach Systematic Review

Ibrahim Goni[1*] Jerome M. Gumpy[2] Timothy U. Maigari[3] Murtala Mohammad[1]

1. Department of Computer Science, Adamawa State University, Mubi, Nigeria
2. Department of Computer Science, Federal University, Gashua, Nigeria
3. Department of Computer Science, Federal College of Education Gombe, Nigeria

ARTICLE INFO

Article history
Received: 26 October 2020
Accepted: 16 November 2020
Published Online: 30 November 2020

Keywords:
Cybersecurity
Cyber forensics
Cyber space
Cyber threat
Machine learning and deep learning

ABSTRACT

The proliferation of cloud computing and internet of things has led to the connectivity of states and nations (developed and developing countries) worldwide in which global network provide platform for the connection. Digital forensics is a field of computer security that uses software applications and standard guidelines which support the extraction of evidences from any computer appliances which is perfectly enough for the court of law to use and make a judgment based on the comprehensiveness, authenticity and objectivity of the information obtained. Cybersecurity is of major concerned to the internet users worldwide due to the recent form of attacks, threat, viruses, intrusion among others going on every day among internet of things. However, it is noted that cybersecurity is based on confidentiality, integrity and validity of data. The aim of this work is make a systematic review on the application of machine learning algorithms to cybersecurity and cyber forensics and pave away for further research directions on the application of deep learning, computational intelligence, soft computing to cybersecurity and cyber forensics.

1. Introduction

Cyber space is a platform that support internet of things, networks, telecommunication systems and all other recent information and communication technologies raining today [1,3]. Global cyber security index 2017 revealed that 3.5 billion users are connected to the internet and also predicted that there will be 12 billion devices connected on the cyber space by 2020. It was also predicted that by the year 2020 80% of the youth population in this world would have a smart phone [48], moreover, almost 49.7% out of the 80% will be connected to the internet with an exponential growth of 936% globally between 2000-2017 [49]. Although threats and attack to these devices are becoming order of the day. It is at the heart of this research work to explore the critical research contributions of researches that used machine learning algorithms in cybersecurity and digital forensics.

2. Digital Forensics

Digital forensics is a field in forensic science that presented the methodologies of investigating crimes that take place on a digital devices of individual, private organizations or government institutions be it national or international [2]. Moreover, Nickson et al. [3] explored that Digital

**Corresponding Author:*
Ibrahim Goni,
Department of Computer Science, Adamawa State University, Mubi, Nigeria;
Email: algonis1414@gmail.com

 DOI: https://doi.org/10.30564/ssid.v2i2.2495

forensics is a field of computer security that uses software applications and standard guidelines which support the extraction of evidences from any computer appliances which is perfectly enough for the court of law to use and make a judgment based on the comprehensiveness, authenticity and objectivity of the information obtained. The information obtained should be able to present facts about the evidences; like the profile of who obtained the information? The address where the information obtained and where it has been stored and what happened to the information after collection. Moreover, Rukayat et al., [4] presented the major goals of forensics evidences are finding out the evidence, proper documentation and storage of the evidences, maintaining the evidence and moving it to the appropriate location without any alteration. In Anwar & Riadi [38] argued that digital forensics is perfectly relied on the information obtained with a degree of clear understanding and show clearly evidence of security breaches. In [46] "cyber forensics are scientific methods and methodologies in recent technologies to investigate, trace, and obtain and information from digital device which is going to be used in the court of law as evidence to make a judgment. Cyber forensics science is presented graphically by [50] as:

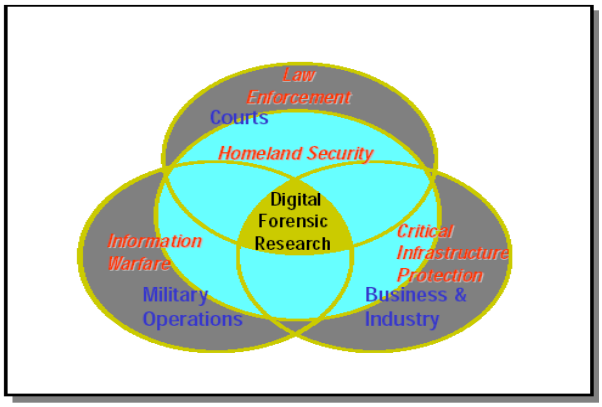

Figure 1. Digital forensics science [50]

2.1 Digital Crime

Intrusions, cyber-attacks, system, information and network breaches are becoming alarming day by day among internet of things [6]. David et al. [10] believed that cyber terrorism is the use of digital device to perpetrate crime that will affect either socio-economic, religion or political agenda of an individual, private organization, government institution or a nation. Capgemini research institute revealed that there was an incidence in which a hacker was able to breaches in to 27000 vehicles which led to the total shutdown the engine. "There is a compelling necessity among researchers to come out with methods and methodologies, tactics and techniques to help forensics scientist in finding out evidences that will be used in the court of law as evidence or make a case [5].

2.2 Cybersecurity

Cybersecurity is a field in computer science that includes; information security, cloud security, and system security. It is also agreed by the literatures that cybersecurity is mainly depend on three major things; confidentiality, integrity and availability, of information [28] highlighted that the major principles of cybersecurity are prevention detection and reaction. In addition to Central Intelligence Agency (CIA) show that the central goals of cybersecurity are confidentiality of information, integrity of information and availability of information. Furthermore, National Cyber Security Center (NCSC) UK highlighted ten (10) tips of cybersecurity these includes; Internet security, Public awareness, threat prevention, device access control, prevent access to configurations, monitoring users, monitoring and management mobiles communications [35]. Gyun [33] also revealed that artificial intelligence/computational intelligence techniques and deep learning and machine learning techniques are among the cyber tools for modeling behaviors of attacks and building systems for defense.

2.3 Cyber Attacks

According to United State intelligent unit revealed that in 2016 and 2017 there has been sponsored cyber-attack against Ukraine and Saudi Arabia which ended in targeting both government and non-governmental organizations. And also classified cyber-attacks based on the following; identity theft Unauthorized access and Deniel of service (DoS, DDoS). Cybersecurity experts revealed that in 2019 ransomware will damages almost $11.5 billion [44]. There has been a ransomware attack targeting England's National Health Service which affected 60 health trusts, 150 countries, and more than 200,000 computer systems [45].

3. Machine Learning Algorithm

Machine learning algorithms are technique in artificial intelligence and computational intelligence that uses algorithm to parse data, learn from the data and make a decision or classification, These algorithms are technically depend on the statistical and mathematical models. In recent time machine learning algorithm are applied in clustering, regression, anomaly detection, intrusion detection systems artificial immune systems, network security, pattern recognition and even forensics investigations [34]. Basically there are three types of machine learning algorithms these

DOI: https://doi.org/10.30564/ssid.v2i2.2495

are; supervised learning, unsupervised learning and reinforcement. Supervised learning algorithms involve the use of datasets for training and testing the performance of the system build. Some supervised learning algorithm includes; decision tree, logistic regression, support vector machine, relevance vector machine, random forest, K-NN, bagging neural networks, linear regression and naïve Bayes which has been applied to cybersecurity, intrusion detection systems, network security and digital forensics [33]. Unsupervised learning algorithm required datasets for training and testing the system performances but require no labeled on the datasets. The two most common unsupervised learning algorithms are Principal Component Analysis (PCA) and clustering. Some of the unsupervised learning algorithms that are applied in cybersecurity are hierarchical, k-means, mixed model, DBSCAN, OPTIC, self-organizing mapping, Bolzan machine, auto encoder, adversarial network which has yield results [34].

4. Reviewed of Related Literature

Bandir [7] revealed that machine learning algorithms such self-organizing mapping, clustering will be very effective for digital forensics especially in a situation where large amount of data is going to be used. In [8] applied memetic algorithm in forensics analysis. In addition to [9] showed how machine learning algorithms are applied to security breaches. Malware classification system was also implemented using machine learning algorithms [16]. Hybrid system that is the combination of deep learning and machine learning algorithms was used to implement cybersecurity system in [15]. Intrusion detection system was also implemented in [14,51]. In [13] a systematic review was made on the combination of machine learning algorithms and data mining approach to cybersecurity. In [12] described how machine learning algorithms are good in the feature of cybersecurity. In the literatures researches and white papers are presenteds and published regarding the application of computational intelligence/ artificial intelligence techniques, machine learning, deep learning are applied to system security [16] and [18-22]. Cybersecurity system was modeled in [29]. Cybersecurity Framework was implemented in [30] using fuzzy logic algorithm.

Furthermore, [23] combined machine learning algorithm and deep learning algorithm for intrusion detection system. [24] Conducted a systematic survey on the anomaly based intrusion detection system. [25] Implemented intrusion detection system using machine learning algorithms for cloud mobile system in a heterogeneous network. Hybrid system for intrusion detection system was implemented in [26]. In addition to [27] has pave away for further implementation of industrial anomaly detection using machine learning algorithms. Anomaly detection system for mobile networks and automobile network was presented by [31]. Hybrid system for traffic control and monitoring was implemented in [32]. A review was made by [40] on the methods that are used for malware detection, and [41] applied machine learning algorithm to detect malware in android mobile devices. In [42] they conducted a review on malware detection using parallel computing. [43] made a comparative analysis on malware detection between static, dynamic and hybrid system. Digital Forensics analysis was also made on WhatsApp and Facebook to identify those that are using the application to commit a crime or illegal businesses [36-39]. In addition to Parag [47] Digital forensics framework was proposed and made a comparative analysis with other framework made with many artificial intelligence techniques and machine learning algorithms.

5. Conclusion

In recent time machine learning algorithms computational intelligence techniques, artificial intelligence techniques deep learning among other intelligent techniques are used to modeled or build a cybersecurity system such as internet security, information security, identity access security, cloud computing security, Internet of Things security, intrusion detection system, artificial immune systems, although majority of the security systems depend on the detection, prediction and response. Moreover, the main goals of cyber security are confidentiality, integrity and availability. In this research work it is also noted that there are ten steps to cybersecurity; network security, user education and awareness, malware prevention, removable media control, secure configuration, managing user privileges, incident management, monitoring and home and mobile working. In addition to AI and machine learning are among the good cyber tools for modeling the investigation system in digital forensics.

References

[1] Shahzad S. protecting the integrity of digital evidence and basic human rights during the process of digital forensics. Ph.D. thesis Stockholm University, 2015.

[2] Abdalzim A. M. A., Amin B. A. M. A survey on mobile forensics for android smart phones IOSR. Journal of computer engineering, 2015, 17(2): 15-19.

[3] Nickson M. K., Victor R. K.,Venter H. Divergency deep learning cognitive computing techniques into cyber forensics. Elservier Forensics Science international synergy, 2019, 1: 61-67.

[4] Rukayat A. A., Charles O. U., Florence A. O. Com-

 DOI: https://doi.org/10.30564/ssid.v2i2.2495

puter forensics guidelines: a requirement for testing cyber crime in Nigeria now? 2017.

[5] Casey E. Editorial- A sea change in digital forensics and incident response. Digital investigation evidence Elservier Ltd., 2016, 17: A1-A2.

[6] Ehsan S., Giti J. Seminars in proactive artificial intelligence for cyber security consulting and research. Systematic cybernetics and informatics, 2019, 17(1): 297-305

[7] Bandir A. Forensics analysis using text clustering in the age of large volume data: a review. International journal of advanced computer and application, 2019, 10(6): 72-76.

[8] Al-Jadir I., Wong K. W., Fing C. C., Xie H. Enhensing digital forensics analysis using memetic algorithm feature selection method for document clustering. IEEE international conference on systems, Man and cybernetics, 2018: 3673-3678.

[9] Sunil B., Preeti B. Application of artificial intelligence in cyber security. International journal of engineering research in computer science and engineering, 2018, 5(4): 214-219.

[10] David O. A., Goodness O., Etecte M. A. Unbated cyber terrorism and huma security in Nigeria. Asian social science, 2019, 15(11): 105-115.

[11] April. Threat start-SMS spam volume by month of each region SC magazine. 2014. Available online at: http//www.scmagazine.com/april-2014-threat-stats/slideshowz

[12] Apruzze G., Colajanni M. F., Ferreti L., Marchett M. On the effectiveness of machine learning for cyber security in 2018. IEEE international conference on cyber conflict, 2018: 371-390.

[13] Buckza A. L., Guven E. A survey of data mining and machine learning methods for cyber security intrusion detection. IEEE communication survey and totorials, 2016, 18(2): 1153-1176

[14] Biswas S. K. Intrusion detection using machine learning: A comparison study. International Journal of pure and applied mathematics, 2018, 118(19): 101-114

[15] Y. Xin, Kong L., Liu Z., Chen Y., Zhu H., Gao M., Hou H., Wang C. Machine learning and deep learning methods for cyber security. IEEE Access, 2018, 6: 35365-35381.

[16] N. Miloseivic, Denghantanh A., Choo K. K. R. Machine learning aided android malware classification. Computer and electrical engineering, 2017, 61: 266-274.

[17] B. Geluvaraj, Stawik P. M., Kumar T. A. The future of cyber security: the major role of Artificial intelligence, Machine learning and deep learning in cyber space. International conference on computer network and communication technologies Springer Singapore, 2019: 739-747.

[18] H. Mohammed B., Vinaykumar R., Soman K. P. A short review on applications of deep learning for cyber security, 2018.

[19] M. Rege, Mbah R. B. K. Machine learning for cyber defense and attack. in the 7th International conference on data analysis. 2018: 73-78.

[20] D. Ding, Hang Q. L., Xing Y., Ge X., Zhang X. M. A survey on security control and attack detection for industrial cyber physical system. Neuro-computing, 2018, 275: 1674-1683.

[21] D. Berman S., Buczak A. L., Chavis J. S., Corbelt C. L. A survey of deep learning methods for cyber security information. 2018, 10(4).

[22] Y. Wang, Ye Z., Wan P., Zhao J. A survey of dynamic spectrum allocation based on reinforcement learning algorithms in cognitive radio network. Artificial intelligence review, 2019, 51(3): 413-506.

[23] A. Abubakar, Paranggono B. Machine learning based intrusion detection system for software defined networks. 7th International conference on Emerging security techniques IEEE, 2017: 138-143.

[24] S. Jose, Malathi D., Reddy B., Jayaseeli D. A survey on anomaly based host intrusion detection system. Journal of physics. Conference series, 2018, 1000(1).

[25] S. Dey, Ye Q., Sampalli S. A Machine learning based intrusion detection scheme for data fusion in mobile cloud involving heterogeneous clients network. Information fussion, 2019, 49: 205-215.

[26] P. Deshpande, Sharma S. C., Peddoju S. K., Junaid S. HIDS: a host based intrusion detection system for cloud computing environment. International journal of system assuarance engineering and management, 2018, 9(3): 567-576.

[27] M. Nobakht, Sivaraman V., Boreli R. A host-Based Intrusion detection and mitigation framework for smart IoT using open flow in 11th International conference on availability reliability and security IEEE. 2016: 147-156.

[28] A. Meshram, Christian H. Anomaly detection in industrial networks using machine learning: A road map. Machine learning for cyber physical system. Springer Berlin Heldelberg, 2017: 65-72.

[29] R. Devakunchari, Souraba, Prakhar M. A study of cyber security using machine learning techniques. International journal of innovative technology and exploring engineering, 2019, 8(7): 183-186.

[30] E. Alison N. FLUF: fuzzy logic utility framework to support computer network defense decision making.

DOI: https://doi.org/10.30564/ssid.v2i2.2495

IEEE, 2016.

[31] A. Taylor, Leblanc S., Japkowicz N. Anomaly detection in auto-mobile control network data with long short term memory network in data science and advance analytics. IEEE international conference. 2016: 130-139.

[32] O. Amosov S., Ivan Y. S., Amosovo S. G. Recognition of abnormal traffic using deep neural networks and fuzzy logic. International Multi-conference on industrial engineering and modern technologies IEEE, 2019.

[33] M. Gyun L. Artificial Intelligence for development series: Report on AI and IoT in Security Aspect. 2018.

[34] L. Matt. Rise of machine: machine learning & its cybersecurity applications. NCC group white paper, 2017.

[35] National cyber security center UK. www.ncsc.gov.uk

[36] A. Nuril, Supriyanto. Forensic Authentication of WhatsApp Messenger Using the Information Retrieval Approach. International Journal of Cyber Security and Digital Forensics (IJCSDF), 2019, 8(3): 206-212.

[37] A Marfianto, I Riadi. WhatsApp Messenger Forensic Analysis Based on Android Using Text Mining Method. International Journal of Cyber Security and Digital Forensics (IJCSDF), 2018 7(3): 319-327.

[38] N Anwar, I. Riadi. Forensic Investigative Analysis of WhatsApp Messenger Smartphone Against WhatsApp Web-Based. Journal Information Technology Electromagnetic Computing and Information, 2017, 3(1): 1-10.

[39] S. Ikhsani, C. Hidayanto, Whatsapp and LINE Messenger Forensic Analysis with Strong and Valid Evidence in Indonesia. Tek. ITS, 2016, 5(2): 728-736.

[40] M. Ashawa, S. Morris. Analysis of Android Malware Detection Techniques: A Systematic Review. International Journal of Cyber Security and Digital Forensics (IJCSDF), 2019, 8(3): 177-187.

[41] W. Songyang, Wang, P., Zhang, Y. Effective detection of android malware based on the usage of data flow APIs and machine learning: Information and Software Technology, 2016, 75: 17-25.

[42] Anastasia, S., Gamayunov, D. Review of the mobile malware detection approaches: Parallel, Distributed and Network-Based Processing (PDP). In: Proc. 2015. IEEE 23rd Euro micro International Conference, 2015: 600--603.

[43] D. Anusha, Troia, F. D., Visaggio, C. A., Austin, T. H., Stamp, M. A comparison of static, dynamic, and hybrid analysis for malware detection. Journal of Computer Virology and Hacking Techniques, 2017, 13(1): 1-12.

[44] S. Morgan. Cyber security Business Report. 2017. Retrieved from CSO: https://www.csoonline.com/article/3237674/ransomware/ransomware-damage-costs-predicted-to-hit-115b-by-2019

[45] R. Collier. NHS ransomware attack spreads worldwide. CMAJ. 2017, 189(22): 786-787. https://doi.org/10.1503/cmaj.1095434

[46] H. Trisnasenjaya, I. Riadi Forensic Analysis of Android-based WhatsApp Messenger Against Fraud Crime Using The National Institute of Standard and Technology Framework. International Journal of Cyber Security and Digital Forensics (IJCSDF), 2019, 8(1): 89-97.

[47] H. Parag Rughani. Artificial Intelligence Based Digital Forensics Framework. International Journal of Advanced Research in Computer Science, 2017, 8(8): 10-14.

[48] 2016: Current State of Cybercrime, RSA Whitepaper. 2016.

[49] World Internet Users and 2017 Population Stats. Accessed from http://http://www.internetworldstats.com/stats

[50] R. Mark. Computer forensics: Basics. Lecture note Purdue University, 2004.

[51] Ibrahim Goni & Ahmed L. Propose Neuro-Fuzzy-Genetic Intrusion Detection System. International Journal of Computer Applications, 2015, 115(8). Available online at: http://www.ijcaonline.com/archives/volume115/number8/20169-2320

 DOI: https://doi.org/10.30564/ssid.v2i2.2495