Scientific and Practical Cyber Security Journal (SPCSJ) 5(1): 9-15 ISSN 2587-4667 Scientific Cyber Security Association (SCSA)

A PROPOSED NOVEL LOW COST GENETIC-FUZZY BLOCKCHAIN-ENABLED INTERNET OF THINGS (IoT) FORENSICS FRAMEWORK

Faisal A. Garba, Department of Computer Science Education, Sa'adatu Rimi College of Education, Kano
Kabiru I. Kunya, Department of Computer Science Education, Sa'adatu Rimi College of Education, Kano
Zahrau Ahmad Zakari, Kunya, Department of Computer Science Education, Sa'adatu Rimi College of Education, Kano
Shazali A. Ibrahim, Kunya, Department of Computer Science Education, Sa'adatu Rimi College of Education, Kano
Abubakar Abba, Department of Computer Science, Federal College of Education, Zaria
Jameel Shehu Yalli, Federal University Gusau
Zaharaddeen Karami Lawal, Department of Computer Science, Federal University Dutse
Aliyu Lawan Musa. Department of Computer Engineering Technology, School of Technology, Kano State Polytechnic

ABSTRACT

Practitioners of network forensics often employ automated software and hardware tools for the collection and preservation of data, however, the process of performing a forensic examination is not well defined. This has resulted in the emergence of various digital forensic frameworks, which determine the correct course of action during an investigation, separating the process into autonomous stages and suggesting appropriate tools and techniques for each task. Even though many forensic frameworks have been proposed, existing solutions give emphasis on acquisition and neglect examination and analysis. Privacy is also a key element in maintaining the confidentiality of data in forensics as it may lead to exposure of personal identifiable information. Furthermore, accountability is one of the IoT forensics challenges. The widespread adoption of an estimated 30.9 billion IoT devices by 2025 (Statista, 2021), as well as the increasing interconnectivity of those devices to traditional networks, not to mention to one another with the advent of fifth generation (5G) networks, underscore the need for IoT forensics. This work proposed a novel low cost IoT forensic framework to tackle: (a.) the examination and analysis phase of IoT forensics using genetic-fuzzy expert system (b.) the issue of guarding the privacy and chain of custody of IoT forensics data using hyperledger fabric, private-permissioned blockchain that is both free and open source. The framework will be implemented and evaluated with related works using BoT-IoT dataset. The BoT-IoT dataset includes Distributed Denial of Service (DDoS), Denial of Service (DoS), Operating System (OS) and Service Scan, Keylogging and Data exfiltration attacks, with the DDoS and DoS attacks further organized, based on the protocol used. The genetic-fuzzy IoT forensics framework will be compared against related work and Network Forensics Analysis Tool (NFAT) to evaluate the performance and accuracy of the proposed framework. The private permissioned blockchain IoT forensics framework will be compared against a related work to evaluate the security and cost of the proposed private permissioned blockchain framework. The genetic-fuzzy blockchain-enabled IoT forensic framework will be compared with, related works and NFATs to evaluate the speed and accuracy performance of the proposed framework. The result of this study is a low cost genetic-fuzzy blockchain-enabled IoT forensics framework.

KEYWORDS*: IoT, forensics, blockchain, genetic-fuzzy*

INTRODUCTION

Internet of Things (IoT) will soon be present in all areas of our life. While it is true that this development makes the lives of humans easier, said development also gives rise to numerous

issues related to digital forensics and security (Atlam *et al.,* 2020). Computer or digital forensics is the practice of investigating computers, digital media, and digital communications for potential artifacts. In this context, the word artifact indicates any object of interest (Messier, 2017). Network forensics is one of the sub-branches of digital forensics where the data being analyzed is the network traffic going to and fro the system under observation. The purposes of this type of observation are collecting information, obtaining legal evidence, establishing a root-cause analysis of an event, analyzing malware behavior, and so on (Jaswal, 2019). Unlike other areas of digital forensics, network forensic investigations deal with volatile and dynamic information (Datt, 2016). IoT forensics comprises three digital forensics schemes in total: network forensics, device-level forensics, and cloud forensics (Atlam *et al*., 2020; Zawoad and Hasan, 2015). As the majority of the IoT devices are characterized by low storage and computational capability, any data which is produced by the IoT network and IoT device is kept and sorted in the cloud. IoT infrastructures are made up of different kinds of networks, such as Wide Area Networks (WAN), Body Area Network (BAN), Home/Hospital Area Networks (HAN), Personal Area Network (PAN), and Local Area Networks (LAN). Crucial pieces of evidence can be gathered from any one of the above-mentioned networks. If a vital piece of evidence must be gathered from the IoT devices, device-level forensics comes into play. The device level forensics scheme is employed when there is the need to collect, from the IoT devices, a vital piece of evidence (Zawoad and Hasan, 2015). IoT forensics remains in the process of maturing, particularly since there are numerous challenges in existence and fewer studies in the field. Accountability is a major requirement in IoT forensics (Lutta *et al.,* 2020; Singh *et al.,* 2018). IoT forensics framework at network level has been proposed to handle the accountability issues with the use of public-permissionless blockchain. Public-permissionless blockchain however, comes at a cost which is usually paid in the form of cryptocurrency (for instance Bitcoin or Ether depending on the platform used) to the miners as an incentive for validating a transaction. Aside from being not free (since gas fee is paid for transaction validation), with public-permissionless blockchain there is no control and restriction to who should join the blockchain. Anyone can join the blockchain platform. This research work therefore proposed a private permission blockchain framework which preserves provenance of IoT forensic data.

Even though many forensic frameworks have been proposed existing solutions neglect examination and analysis and instead give more emphasis on acquisition (Koroniotis and Moustafa, 2020). In the examination phase, evidence collected is searched methodically to extract specific indicators of the crime. These indicators of crime are then classified and correlated to deduce important observations using the existing attack patterns during the analysis phase. Statistical, soft computing and data mining approaches are used to search the data and match attack patterns. The attack patterns are put together, reconstructed and replayed to understand the intention and methodology of the attacker (Pilli *et al*., 2010). Soft computing is viewed as a foundation component for the emerging field of computational intelligence (Cabrera *et al*., 2009). According to Mankad (2013) soft computing is a good option for complex systems where: the required information is not available; the behavior is not completely known; and the existence of measure of variables is noisy. Soft computing is a consortium of computing methodologies that provides a foundation for the conception, design, and deployment of intelligent systems to provide economical and feasible solutions with reduced complexity (Mankad, 2013). Members of this consortium include: Fuzzy Logic (FL), Neural Network (NN), Evolutionary Computations (EC) and Probabilistic Reasoning (PR). Each of these techniques has their own strengths and limitations. Integration of two or more techniques can provide significant advantages for intelligent system design. The hybridization of major constituents of Soft Computing can be represented as EC-FL, EC-NN, PR-FL and PR-NN. Fuzzy logic is used to process human-like classification of things into groups with the representation of fuzzy linguistic variable. Hybridization of genetic algorithm with other soft computing components, results in natural evolution of a solution. It has been observed that genetic algorithm provides the following major advantages: genetic algorithm can be easily interfaced to obtainable simulations and models; genetic algorithm is easy to

hybridize and easy to understand; genetic algorithm uses little problem specific code; genetic algorithm is modular, separate from application; genetic algorithm is capable to obtain answers always and gets better with time; and genetic algorithm is inherently parallel and easily distributed (Williams, 2020). The major limitations of fuzzy systems are: inability of self-learning, adaption or parallel computation; cannot support optimization; answers obtained once cannot get better with time. In order to solve the stated problems, the use of genetic algorithm to find optimized values for the membership function parameters, particularly when manual selection of their values becomes difficult or takes too much time to attain has been proposed (Mankad, 2013). Liao *et al.,* (2009) and Kim *et al.,* (2004) have both proposed a fuzzy expert system network forensics investigation. Liao *et al.,* (2009) evaluated the fuzzy expert network forensic system performance with DARPA 2000 dataset and compared the proposed fuzzy expert network forensic system with other proposed studies that utilizes Support Vector Machine (SVM), Naïve Bayes Algorithm (NB), C4.5 and SMO algorithm (another training algorithm for SVM) to which the proposed fuzzy expert network forensic system outperform them all in attack detection accuracy. Kim *et al.,* (2004) on the other hand uses DARPA 1998 dataset and the proposed fuzzy expert network forensic system has a detection accuracy of 92% but no performance comparison with other related study was done. Mankad (2013) successfully applied genetic-fuzzy to measure multiple intelligence. In this work we seek to employ a similar approach to use genetic algorithm to improve fuzzy expert system performance in examination and analysis of IoT network traffic data. We intend to use Bot-IoT dataset (Koroniotis *et al*., 2019) to evaluate the performance of the proposed Genetic- Fuzzy IoT Network Forensic Framework.

STATEMENT OF THE PROBLEM

Privacy is a key element in maintaining the confidentiality of forensics data as it may lead to exposure of personal identifiable information (Lutta *et al.,* 2020). Singh *et al.,* (2018) mentioned accountability as one of the IoT forensics challenges. Singh *et al.,* (2018) stress that this is because different entities manage the composition and the interactions between the IoT components. The distributed and immutable characteristics of blockchains suit the demands of IoT Forensics. An ideal solution for IoT Forensics is a private-permissioned blockchain where the number of nodes is restricted and access is only provided to selected users as suggested by Sadineni *et al*., (2019). Even though many forensic frameworks have been proposed existing solutions give emphasis on acquisition and neglect examination and analysis (Koroniotis and Moustafa, 2020) In the examination phase, evidence collected is searched methodically to extract specific indicators of the crime. These indicators of crime are then classified and correlated to deduce important observations using the existing attack patterns during the analysis phase. Statistical, soft computing and data mining approaches are used to search the data and match attack patterns. The attack patterns are put together, reconstructed and replayed to understand the intention and methodology of the attacker (Pilli *et al*., 2010). Soft computing is viewed as a foundation component for the emerging field of computational intelligence (Cabrera *et al*., 2009). According to Mankad (2013) soft computing is a good option for complex systems where: the required information is not available; the behavior is not completely known; and the existence of measure of variable is noisy. Soft computing is a consortium of computing methodologies that provides a foundation for the conception, design, and deployment of intelligent systems to provide economical and feasible solutions with reduced complexity (Mankad, 2013). Members of this consortium include: Fuzzy Logic (FL), Neural Network (NN), Evolutionary Computations (EC) and Probabilistic Reasoning (PR). Each of these techniques has their own strengths and limitations. Integration of two or more techniques can provide significant advantages for intelligent system design. The hybridization of major constituents of Soft Computing can be represented as EC-FL, EC-NN, PR-FL and PR-NN. Fuzzy logic is used to process human like classification of things into group with the representation of fuzzy linguistic variable. Hybridization of genetic algorithm with other soft computing components, results in natural

evolution of a solution. It has been observed that genetic algorithm provides the following major advantages: genetic algorithm can be easily interfaced to obtainable simulations and models; genetic algorithm is easy to hybridize and easy to understand; genetic algorithm uses little problem specific code; genetic algorithm is modular, separate from application; genetic algorithm is capable to obtain answers always and gets better with time; and genetic algorithm is inherently parallel and easily distributed (Williams, 2020). The major limitations of fuzzy systems are: inability of self-learning, adaption or parallel computation; cannot support optimization; answer obtained once cannot get better with time. In order to solve the stated problems, the use of genetic algorithm to find optimized values for the membership function parameters, particularly when manual selection of their values becomes difficult or takes too much time to attain has been proposed (Mankad, 2013).

AIM AND OBJECTIVES

The aim of this research is to develop a novel low cost fuzzy-genetic blockchain enabled network forensics framework to address the preservation of digital provenance, examination and analysis challenges of IoT forensics.

The specific objectives of this work are to:

a. design a blockchain fuzzy-genetic IoT forensics framework
b. implement a blockchain fuzzy-genetic IoT forensics framework
c. evaluate the blockchain fuzzy-genetic IoT forensics framework with related works.

SIGNIFICANCE OF THE STUDY

It has been proven that IoT devices are vulnerable to both well established and new IoT-specific attack vectors. In a 2018 report by Symantec regarding the security threats found in the Internet, it was reported that the total number of attacks targeting IoT devices for 2018 exceeded 57,000, with more than 5,000 attacks being recorded each month. Hackers have compromised vulnerable, unpatched or unencrypted IoT devices in order to steal sensitive data, corrupt the device's normal operation, spread malware infections or even compromise the security of a smart home by disabling smart locks and garage doors (Koroniotis and Moustafa, 2020). The widespread adoption of an estimated 30.9 billion IoT devices by 2025 (Statista, 2021), as well as the increasing interconnectivity of those devices to traditional networks, not to mention to one another with the advent of fifth generation (5G) networks, underscore the need for IoT forensics (Zhang *et al*., 2020).

REVIEW OF PROPOSED IoT NETWORK FORENSICS FRAMEWORKS

Mercan *et al*., (2020) proposed "A Cost-efficient IoT Forensics Framework with Blockchain". The study claimed to be cost effective and reliable digital forensics framework that achieves this by exploiting multiple inexpensive blockchain networks as a temporary storage before the data is committed to Ethereum. To reduce Ethereum costs, they utilize Merkle trees which hierarchically store hashes of the collected event data from IoT devices. They evaluated the approach on popular blockchains such as EOS, Stellar and Ethereum by conducting a cost analysis. The results indicates cost savings resulting from using the proposed 'Cost-efficient IoT Forensics Framework with Blockchain'. The proposed work of Mercan *et al*., (2020) has some limitations. First, the use of public-permissionless blockchain platform in this case Etherium to preserve forensic evidence is not recommended. This is because in a public blockchain everyone can join the network and have access to all the blocks in the network. An ideal solution for IoT Forensics is a private-permissioned blockchain where the number of nodes is restricted and access is only provided to selected users (Sadineni *et al*., 2019). Secondly, the use of Ethereum blockchain comes at a cost in the form of 'gas fees' that is paid to the miners as an incentive for validating a block. There are alternative private-permissioned

blockchain platforms that are open source and free to use. One such instance is the Hyperledger Fabric which this research work intends to use to preserve forensic evidence.

Li *et al.,* (2019) proposed "Blockchain-Based Digital Forensics Investigation Framework in the Internet of Things and Social Systems". A blockchain-based digital forensic investigation framework in the Internet of Things (IoT) and social systems environment is proposed, which can provide proof of existence and privacy preservation for evidence items examination. The work has some limitations. The use of blockchain would of course guards the provenance of the forensic digital evidence, however the proposal did not go into detail to specify which type of blockchain it intends to use to implement the proposal –whether private-permission blockchain or public-permissionless blockchain. Secondly, the proposal is just a theoretical presentation, there was no implementation and evaluation to show how it advanced the state of the art.

Brotsis *et al*., (2019) proposed "Blockchain Solutions for Forensic Evidence Preservation in IoT Environments". This study presented a blockchain-based solution, which is designed for the smart home domain, dealing with the collection and preservation of digital forensic evidence. The system utilizes a private forensic evidence database, where the captured evidence is stored, along with a permissioned blockchain that allows providing security services like integrity, authentication, and non-repudiation, so that the evidence can be used in a court of law. The blockchain stores evidences' metadata, which are critical for providing the aforementioned services, and interacts via smart contracts with the different entities involved in an investigation process, including Internet service providers, law enforcement agencies and prosecutors. The proposed work of Brotsis *et al*., (2019) however has not been implemented and there was no evaluation to show how it has advanced the state of the art.

Hossain *et al.,* (2018) propose FIF-IoT – a forensic investigation framework using a public digital ledger to find facts in criminal incidents in IoT-based systems. FIF-IoT collects interactions that take place among various IoT entities (clouds, users, and IoT devices) as evidence and store them securely as transactions in a public, distributed and decentralized blockchain network which is similar to the Bitcoin network. A limitation to the work of Hossain *et al.,* (2018) is that the use of public blockchain to preserve forensic evidence is not recommended. This is because in a public blockchain everyone can join the network and have access to all the blocks in the network. An ideal solution for IoT Forensics is a private-permissioned blockchain where the number of nodes is restricted and access is only provided to selected users as suggested by Sadineni *et al*., (2019).

From the related works we have reviewed, it can be seen that most of them (Mercan *et al*., (2020), Li *et al.,* (2019), Ryu *et al.,* (2019), Brotsis *et al*., (2019), Hossain *et al.,* (2018)) are concentrated on the preservation process of the IoT digital forensics investigation thereby neglecting the other process of the IoT digital forensics investigation such as preparation, collection, detection, incidence response, examination, analysis, investigation and presentation. There is a need for more research on the other aspects of the IoT digital forensic investigation. This work therefore proposed a genetic fuzzy network forensic framework that will cater for the examination and analysis stage of the IoT forensics investigation. Even though the reviewed works have looked into the use of both public-permissionless and private-permissioned blockchain to guard digital evidence provenance, there is still a room for improvement. Permissioned-public blockchain has been suggested as the most ideal for guarding digital evidence provenance (Sadineni *et al*., 2019). This has been proposed theoretically by Brotsis *et al*., (2019). However, the work of Brotsis *et al*., (2019) has not been implemented and evaluated to show how it has advanced the state of the art. This research proposal therefore proposed a low cost private-permissioned blockchain IoT forensics framework that will ensure the digital evidence provenance is well preserved.

A PROPOSED LOW COST NOVEL GENETIC-FUZZY BLOCKCHAIN-ENABLED INTERNET OF THINGS (IoT) FORENSICS FRAMEWORK

This section proposes a novel low cost Genetic-Fuzzy IoT Blockchained-Enabled IoT Forensics Framework as seen in Figure 1. This novel low cost IoT Forensics Framework will use genetic algorithm for fuzzy rules optimization and Fuzzy Expert System for attack identification. The fuzzy expert system maps to the examination & analysis and presentation section of the forensic investigation process. The research will make use of the Bot-IoT dataset by Koroniotis *et al.,* (2019) to evaluate the framework. A software prototype will be developed in Python 3 programming language to implement the proposed Fuzzy-Genetic IoT Forensics Framework. The prototype will be evaluated with Network Forensic Analysis Tools that carries out examination and analysis of forensics data to see how the proposed fuzzy-genetic expert system has advanced the state of the art in the IoT Forensics subdomain using results accuracy as a metric. It is a low cost private permissioned blockchain enabled IoT forensics framework in the sense that it uses a completely free and open source private-permissioned blockchain platforms unlike the work of Mercan *et al*., (2020) that uses Ethereum, a public blockchain where a cost is incurred in the form of 'gas fees'. The study will implement the blockchain component using Hyperledger Fabric and will evaluate the proposed framework in term of cost and security with the work of Mercan *et al*., (2020). The research will make use of the Bot-IoT dataset by Koroniotis *et al.,* (2019) to evaluates it's performance with Network Forensic Analysis Tools using accuracy and performance as a yardstick to see how it has advanced the state of the art. The dataset's source files are provided in different formats, including the original pcap files, the generated argus files and csv files. The files were separated, based on attack category and subcategory, to better assist in labeling process. The captured pcap files are 69.3 GB in size, with more than 72,000,000 records. The extracted flow traffic, in csv format is 16.7 GB in size. The dataset includes DDoS, DoS, OS and Service Scan, Keylogging and Data exfiltration attacks, with the DDoS and DoS attacks further organized, based on the protocol used (Koroniotis *et al.,* 2019).

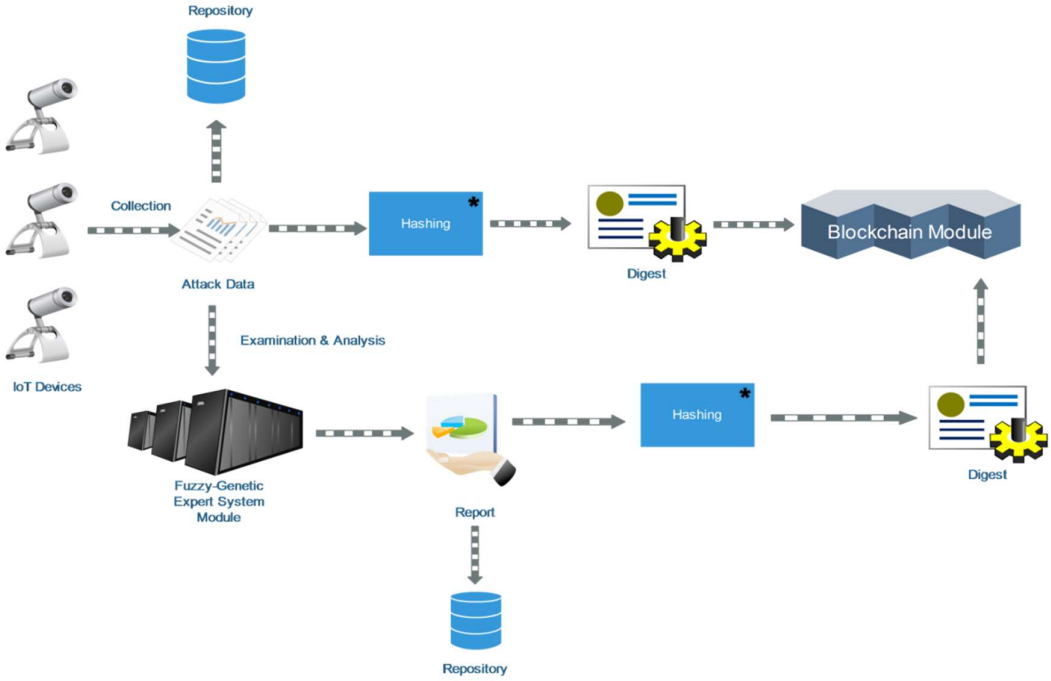

Figure 1: Proposed Fuzzy-Genetic Blockchained Enabled IoT Forensics Framework

CONCLUSION

The widespread adoption of an estimated 30.9 billion IoT devices by 2025 (Statista, 2021), as well as the increasing interconnectivity of those devices to traditional networks, not to mention to one another with the advent of fifth generation (5G) networks, underscore the need for IoT forensics (Zhang *et al*., 2020). This paper has proposed a novel low cost genetic-fuzzy blockchain-enabled Internet of Things (IoT) Forensics Framework. This novel low cost IoT Forensics Framework will use genetic algorithm for fuzzy rules optimization and Fuzzy Expert System for attack identification. It is a low cost private permissioned blockchain enabled IoT forensics framework in the sense that it uses a completely free and open source private-permissioned blockchain platforms unlike the proposed literature that uses Ethereum, a public blockchain where a cost is incurred in the form of 'gas fees'. The study will implement the blockchain component using Hyperledger Fabric and will evaluate the proposed framework in term of cost and security with related works.

REFERENCES

1. Atlam, H., Alenezi, A., Alassafi, M. O., Alshdadi, A. A., & Wills, G. B. (2020). Security, Cybercrime and Digital Forensics for IoT. In S.-L. Peng, & S. Pal, Principles of Internet of Things (IoT) Ecosystem: Insight Paradigm (pp. 551 -). Cham, Switzerland: Springer Nature Switzerland AG.
2. Brotsis, S., Kolokotronis, N., Limniotis, K., Shiaeles, S., Kavallieros, D., Bellini, E., & Pavu´e, C. (2019). Blockchain Solutions for Forensic Evidence Preservation in IoT Environments. IEEE NetSoft 2019 - 1st Workshop on Cyber-Security Threats, Trust and Privacy Management in Software-Defined (pp. 110-114). IEEE.
3. Cabrera et al. (2009) Fuzzy Logic, Soft Computing, and Applications.
4. Liao et al. (2009) Network forensics based on fuzzy logic and expert system
5. Datt (2016) et al. Learning network forensics https://www.packtpub.com/product/learning-network-forensics/9781782174905
6. Hossain, M., Karim, Y., & Hasan, R. (2018). FIF-IoT: A Forensic Investigation Framework for IoT Using a Public Digital Ledger. 2018 IEEE International Congress on Internet of Things (ICIOT). IEEE.
7. Jawal et al. (2019) Hands-On Network Forensics
8. Koroniotis, N., & Moustafa, N. (2020). Enhancing Network Forensics with Particle Swarm and Deep Learning: The Particle Deep Framework. arXiv, 1-20.
9. Koroniotis, N., Moustafa, N., Sitnikova, E., & Turnbull, B. (2019). Towards the Development of Realistic Botnet Dataset in the Internet of Things for Network Forensic Analytics: Bot-IoT Dataset. Future Generation Computer Systems, 779–796.
10. Li, S., Qin, T., & Min, G. (2019). Blockchain-Based Digital Forensics Investigation Framework in the Internet of Things and Social Systems. IEEE Transactions on Computational Social Systems, 1-9.
11. Lutta, P., Sedky, M., & Hassan, M. (2020). The Forensic Swing of Things: The Current Legal and Technical Challenges of IoT Forensics. World Academy of Science, Engineering and Technology International Journal of Computer and Information Engineering, 14(5), 159-165.