Intrusion Detection Using Blockchain Technology

Mrs. Kavitha B.Tech, M.E[1], Raghul Dhakshin P R[2],Vikraman R[3], Guru Bharathi M[4]
[1]Associate Professor,Dept of CSE
[2, 3, 4]Dept of CSE
[1, 2, 3, 4] Velammal College of Engineering and Technology, Madurai

Abstract- *Intrusion detection systems that have emerged in recent decades can identify a variety of malicious attacks that target networks by employing several detection approaches. However, the current approaches have challenges in detecting intrusions, which may affect the performance of the overall detection system as well as network performance. For the time being, one of the most important creative technological advancements that plays a significant role in the professional world today is blockchain technology. Blockchain technology moves in the direction of persistent revolution and change. It is a chain of blocks that covers information and maintains trust between individuals no matter how far apart they are. Recently, blockchain was integrated into intrusion detection systems to enhance their overall performance. Blockchain has also been adopted in healthcare, supply chain management, and the Internet of Things. Blockchain uses robust cryptography with private and public keys, and it has numerous properties that have leveraged security's performance over peer-to-peer networks without the need for a third party.*

To explore and highlight the importance of integrating blockchain with intrusion detection systems, this paper provides a comprehensive background of intrusion detection systems and blockchain technology. Furthermore, a comprehensive review of emerging intrusion detection systems basedon blockchain technology is presented. Finally, this paper suggests important future research directions and trending topics in intrusion detection systems based on block chain technology.

Keywords- Blockchain; intrusion detection system; network security; malicious attacks

I. INTRODUCTION

Blockchain is an emerging technology that underlies the infrastructure of Bitcoin. In 2008, Nakamoto discovered blockchain's potential to be used in other domains, thus making Bitcoin the first of blockchain's many implementations. Blockchain technology has been increasingly used in different fields, especially in the security field, which has an important presence in different network environments, such as traditional networks, the Internet of Things (IoT), and cloud computing. Blockchain technology has been applied to cryptocurrency networks, wherein the blockchain provides cryptocurrency its basic infrastructure, which allows financial operations to be performed in a secure manner and be distributed within networks.

The main contribution of this paper is to provide a comprehensive analysis of blockchain-based IDSs.

This review will accomplish the following:

- Present an overview of blockchain technology and its importance, and introduce the advantages of
- and threats to blockchain;
- Discuss and analyze existing blockchain-based IDSs to provide a clear analysis of the current works
- conducted in this field;
- Compare and analyze the proposed techniques to highlight current research gaps;
- Provide future research directions and open research issues concerning IDSs based on blockchain.

This review is scientifically significant because it allows researchers to analyze blockchain's role in IDSs by providing them with a clear view of the advantages, threats, and opportunities that result from using blockchain in IDSs. decentralization, non-repudiation, and security protection, blockchain is a compelling alternative to intrusion systems.

The basic functionality provided by a blockchain is a cryptographically secure mechanism for obtaining a publicly verifiable and immutable sequence of records (referred to as *blocks*) chronically ordered by discrete time stamps. Blockchains are typically shared and synchronized across a peer-to-peer network, and as such are typically used as a public, distributed ledger of transaction records[3]. Every participant in the blockchain network can see the record data and reject or verify it based on a consensus protocol. Once accepted, records are appended to the blockchain in chronological order of their verification.

II. BACKGROUND

Because of the well-known initiatives in Bitcoin and Ethereum, the first things that come to mind when thinking about the blockchain are cryptocurrency and smart contracts. The blockchain data structure was first employed in Bitcoin, the first crypto-currency solution. Ethereum introduced smart contracts, which take advantage of the immutability and distributed consensus of the blockchain while providing a crypto-currency solution that is equivalent to Bitcoin.

A smart contract is a piece of code that is deployed to the Ethereum network for everyone to see. The outcome of running this code is validated by a consensus method as well as by each individual member of the network. Blockchain is a peer-to-peer network that builds a chain of blocks. As seen in Figure 1, each block in the blockchain has a cryptographic hash and a timestamp added to the previous block. The Merkle tree block header and numerous transactions are contained within a block. Cryptography is a secure networking method that combines computer science and mathematics to keep data and information hidden from others. It enables data to be securely transported across an unsecured network, both encrypted and decrypted forms.

The blockchain is the name for the data structure, as previously stated. All of the written data is separated into blocks, with each block's data including a hash of all of the previous block's data. The goal of a data structure like this is to establish verifiable immutability. If a piece of data is modified, the hash of the block containing that piece of data must be computed, as must the hashes of all following blocks. To ensure that all data remains unaltered, just the hash of the most recent block must be used.

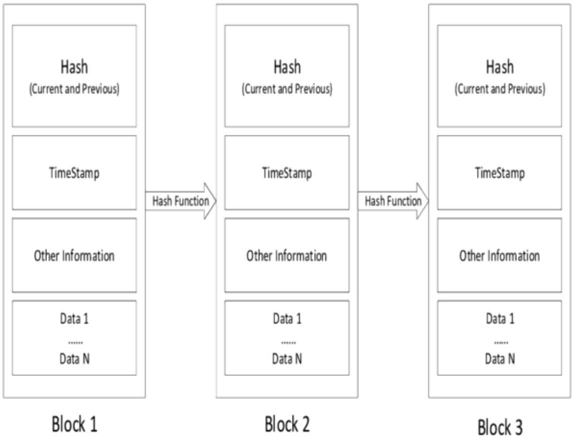


Figure 1

Data contained in blocks in blockchain solutions is produced during their construction from all validated transactions, which means no one can insert, delete, or edit transactions in an already validated block without being noticed. The "genesis block," or the first zero-block, usually contains certain network parameters, such as the original set of validators (those who issue blocks).

III. CORE COMPONENTS OF BLOCKCHAIN ARCHITECTURE

As depicted in Figure 2, the major architectural components of Blockchain are as follows:

Node:In a blockchain structure, a node is a user or a computer (each device has a different copy of the entire ledger from the blockchain).

Transaction:The smallest building block of the blockchain system (records and details), which blockchain uses.

Block:A block is a group of data structures that are used to perform transactions throughout the network and are sent to all nodes.

Chain:A chain is a set of blocks that are arranged in a specific order.

Miners: Correspondent nodes that validate transactions and add them to the blockchain system.

Consensus:A group of instructions and organizations that work together to carry out blockchain activities.

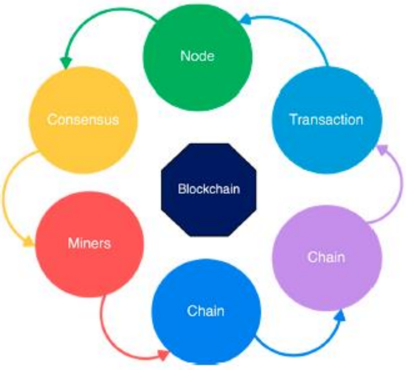
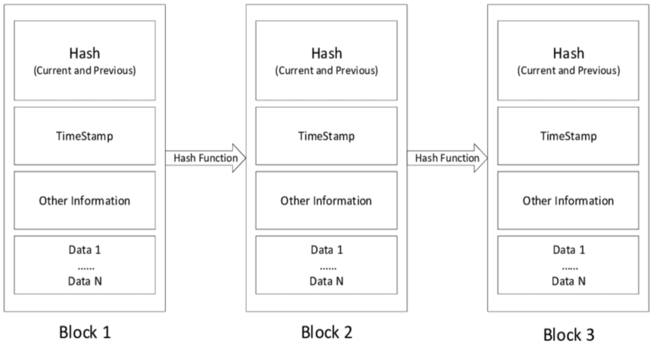


Figure 2

IV. BACKGROUND ON INTRUSION DETECTION

This section introduces the background of a single IDS, collaborative IDSs and two major challenges: data sharing and trust management. Intrusion detection describes the process of monitoring network or system events for any sign of possible incidents . An IDS is an application to realize the process of intrusion detection. Basically, an IDS can provide two main functions.

• **Information Recording**: An IDS can monitor the target objects and record information locally. Then, the collected data can be sent to other facilities for analysis, like a central event management system. • Alert Generation: The main task of an IDS is to generate alerts (alarms) to inform security administrators of important identified anomalies. False alarm rates are an important measurement to decide whether an IDS is effective or not. FIGURE 1. The deployment of HIDS and NIDS in a network environment. As mentioned, an IDS can be generally classified into HIDS and NIDS, whereas such classification can be more specific according to the deployed locations like wireless based IDS, which identifies malicious activities through monitoring wireless network packets and protocols. In practice, an IDS product often combines these two types of detection, as they can complement each other and provide a more thorough protection. The typical architecture of a collaborative intrusion detection network. Based on the detection approaches, an IDS can be either a signature-based or an anomaly-based system. The signature-based detection method, also called misuse-based detection, is usually effective in detecting known exploits but would be ineffective for unseen threats and the variants of known threats. For instance, given a signature that searches a filename of ‘malware.exe’, an attacker can write a malicious application named as ‘malware1.exe’ to easily bypass it. By contrast, the anomaly-based detection has the capability of detecting unknown threats (or zero-day threats). Such detection firstly establishes a normal profile by monitoring the system or network events for a period of time, and then identifies any behavior that would be significantly different from the established profiles. In literature, various machine learning classifiers have been researched in building a normal profile. In particular, profiles can be either static or dynamic in practical usage [2]. A static profile would not be updated while a dynamic profile would be updated periodically based on the security policies. High false rates are a big limitation for the anomaly-based detection. In addition to the above two basic detection approaches, there exists another detection method, called specification-based detection, which identifies deviations between predetermined benign profile and observed events. The benign profile is different from the normal profile in that the former defines generally accepted events in advance. For example, a benign profile can specify how particular protocols should and should not be used.



V. PROBLEMS AND SOLUTIONS OF DEVELOPING INTRUSION SYSTEMS

Although intrusion detection has been studied for nearly 40 years, data sharing and trust computation in a collaborative environment are still two major challenges. • Data Sharing: Data sharing is a major issue for a collaborative detection system, as it is not a trivial task to let all participating parties trust each other. For example, PKI technology can help build some kind of trust, but it does not always work for intrusion detection. Moreover, due to privacy concerns, some parties are not willing to share the data. Without enough data, it is unable to optimize detection algorithms and to build a robust model for identifying suspicious events.

• **Trust Management**: It is known that CIDNs/CIDSs are vulnerable to insider attacks, where the intruders have authorized access to the network. Typically, computational trust is often used to quantify the trust levels among various nodes. In practice, a central server is deployed to collect nodes’ traffic and behavioral data and to compute the trust value of each node. However, the trust management would become an issue when the organization becomes large, as it is hard to find a trusted third party, i.e., central server can be compromised. By design, blockchain technology is a decentralized and distributed ledger that enables recording transactions across a set of nodes. It can be implemented in a peer-to-peer network without the need of a trusted third party. The blockchain integrity can be enforced by strong cryptography, making it nearly impossible to compromise by any individual. Due to the nature of blockchains, there is a chance of applying such emerging technology for solving the above challenges in intrusion detection. Data Sharing: The data sharing problem is mainly caused by two requirements: mutual trust and data privacy. Mutual trust means that when sharing the data, collaborating parties have to trust others who would not disclose the data. For instance, two IT organizations would like to make an agreement that they will not share the data with others. Data privacy indicates that the shared data may contain some information linked to an actual organization, i.e., the shared traffic including IP addresses and packet payloads that can be utilized to refer the privacy of an organization. Blockchains are one of the solutions that can be used to mitigate this challenge. More specifically, data sharing can be considered as a series of transactions. Firstly, collaborating parties should make a data-sharing agreement, which digitally signed by each party. Then, the agreement can be kept in a blockchain box, which is public and unalterable. In this case, other parties can access the blockchain box, read the agreement, and confirm the ownership of the data. Such permanent visibility of the agreement ensures that one party cannot unilaterally repudiate it. Similar to the application of blockchains in the healthcare domain [6], building an open accounting system is able to offer trust among various collaborating parties. For data privacy, one solution is to share transformed data instead of raw data. For example, suppose that a collaborating party (say Party A) wants to verify the performance of their designed classifier using the data from another party (say Party B). As part of the data-sharing agreement, Party A can deposit the classifier into the blockchain box, and then Party B can retrieve the classifier, run it locally with the data and send back the result to Party A. In this case, Party B actually maintains the privacy of the raw data. On the whole, for data sharing issue, blockchains can help build mutual trust among collaborating parties and preserve data privacy by working as a permanent public ledger of contracts between data owners and other parties. Trust Computation: Generally, a collaborative network architecture can be classified as centralized, hierarchical and distributed. In literature, distributed architecture has been widely studied, while the other two are believed to suffer from scalability and an issue of single point of failure. For a CIDN, alert exchange is extremely important among various IDS nodes, which can be used to help decide whether there is an anomaly. In addition, alert exchange can be used to compute the trustworthiness of a node within the network. For example, Fung et al. [4] designed a type of challenge-based CIDNs, in which the trustworthiness of a node could be computed based on the satisfaction of received alert-related information. Their proposed architecture can be robust against some insider attacks like newcomer attack and Betrayal attack, but is still vulnerable to advanced collusion attack where a group of malicious peers cooperate together by providing false alarm rankings in order to compromise the network, e.g., passive message fingerprint attacks [7]. Therefore, how to perform trust computation in a robust way remains a challenge. Blockchain technology provides a potential way to mitigate this issue. For instance, Alexopoulos et al. [8] introduced a blockchain-based CIDS, which applied blockchains for enhancing trust among IDS nodes. In particular, they considered the raw alerts generated by each IDS node as transactions in a blockchain, which could be replicated among the collaborating nodes of a CIDN. Then, all collaborating nodes adopted a consensus protocol to guarantee the validity of the transactions before putting them in a block.

VI. PROPOSED MODELS

Pattern Matching: Pattern matching compares new strings that enter the system with strings in the system’s database to verify that there is no malicious attack occurring. If there is any matching pattern, then the system detects an attack and will generate an alarm; if there is no matching pattern, then no attack is detected. There are two kinds of pattern matching algorithms: single and multiple. Single pattern matching algorithms are simple because they search for one pattern at a time. Multiple pattern matching algorithms search for all patterns at the same time, require more time and resources. A popular pattern matching algorithm applied to IDSs is the Boyer–Moore single pattern algorithm compares strings from the rightmost character. Although it has achieved the best performance in searching operations, the Boyer–Moore algorithm does not have feature scalability. Meanwhile, the Aho–Corasick and Wu–Manber algorithms are multiple pattern matching algorithms that search for more than one pattern simultaneously; however, the Aho–Corasick algorithm requires more memory than the Wu–Manber algorithm. Pattern matching algorithms have a trade-off between their search speed and consumed memory. Some researchers have proposed ways to optimize these algorithms, while others have proposed new algorithms to enhance the performance of detection techniques in IDSs .

Rule-based: This technique is used in both signature and anomaly approaches. Signature detection diagnoses packets and detects malicious attacks through rules that are predefined in the system, where as anomaly detection diagnoses the behavior of the system and detects differences between normal and abnormal behavior depending on predefined rules in the system, such as programmers’ sequence of system calls. Both detection methods must update a network’s rules to acquire more security. Updating the rules using the signature approach is simple, easy, and automatic; updating the rules using anomaly detection, however, is more complex because it needs time to record new training rules.

State-based: Signature detection uses the state transition analysis technique to describe attack scenarios. This technique contains two main elements, namely, state, and arc. The state represents the user or process, and the arc represents an action; if the user or process reaches the final state, then an attack occurs and the system detects it. The first tool to implement the state transition analysis technique was the Unix State Transition Analysis Tool, which executes host-based intrusion detection. The Unix State Transition Analysis Tool is a rule-based expert system that looks for known attacks in the audit traces of multi-user computer systems. However, it suffers from some limitations, such as its features being difficult to extend or adapt to different operating systems.

Data Mining: The signature detection approach can use data mining techniques to discover new patterns for IDSs and to overcome its main disadvantage. Although data mining is used mainly in the signature approach, much research has also applied data mining to anomaly detection. However, data mining requires data from various machine learning techniques, such as rule-based, classification, and clustering, to gather knowledge for network intrusion detection.

Statistical-based Intrusion Detection: This technique deals with two profiles in anomaly detection: one for observing current network traffic, and the other for statistical training. When an event occurs, the anomaly detection system evaluates it by comparing two behaviors. If the anomaly score exceeds the threshold, then the intrusion detection system generates an alarm [9]. Most model-based statistics assume multivariate statistical techniques, such as the chi-square statistic, Canberra technique, and Hotelling's T-squared distribution. Numerous anomaly detection mechanisms find outliers in the dataset by analyzing behavior, as each element in the dataset has specific features and a local outlier factor that could be used to detect the abnormal behavior [10,11].

Biological Models: Prior works have proven that the human immune system and computer network security are similar in nature. Both systems have a complex network and aim to protect its nodes from any malicious attack. In addition, both systems have security policies and security levels. The human immune system sets its policies to depend on natural selection phenomena, and its security levels should meet disposability, correction, integrity, and accountability requirements. Meanwhile, computer network systems establish a set of rules to defend against attacks and detect illegal actions that may occur in the network that break specific security levels [15–17]. In recent years, several algorithms inspired by biological processes, such as genetic algorithms and artificial neural network algorithms [18,19], have been widely applied to the anomaly detection approach to enhance the performance of intrusion detection.

Learning Models: Artificial learning techniques have increased the effectiveness of the anomaly detection approach. Anomaly detection can be supervised or unsupervised. Supervised anomaly detection is taught by a labelled dataset that distinguishes between normal and abnormal behavior. Supervised learning algorithms include support vector machines and the k-nearest neighbor. Unsupervised anomaly detection is taught by unlabeled training data; therefore, it uses several techniques to distinguish between normal and abnormal behavior in the system. One of these techniques is clustering, which has been used in anomaly intrusion detection to find outliers exhibiting anomalous behavior. The k-mean clustering algorithm is the most popular such algorithm, and has been applied to intrusion detection [20–22].

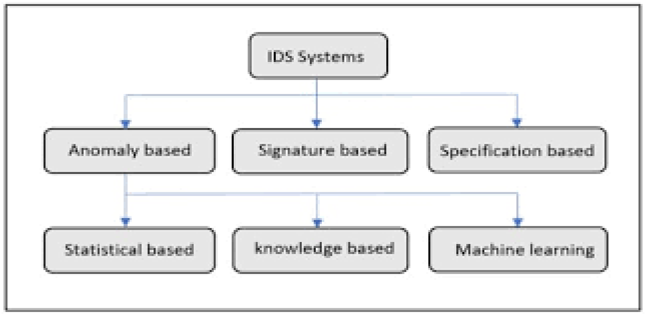


Figure 4

VI. CONCLUSION

Blockchain technology is an emerging solution for decentralized transactions and data management without the need of a trusted third party. It is an open and distributed ledger, enabling the recording of transactions among various parties in a verifiable way. To date, blockchains have been studied in several domains like healthcare and supply chain management, but there has been little work investigating its potential application in the field of intrusion detection. Motivated by this observation, our work mainly discusses the applicability of blockchain technology to mitigate the challenges of data sharing and trust computation in a collaborative detection environment. We identify that blockchains have a potential impact on the improvement of an IDS, whereas not all IDS issues can be solved with this technology.

Recently, blockchain technology has emerged within several fields to ensure high level of security. This paper discussed the structure of blockchain, presented an overview of IDSs, and compared between existing blockchain-based IDS models. However, few research has been conducted on this topic, and no standard approaches or real applications have been demonstrated. In addition, this paper identified future directions that need to be addressed and investigated by researchers to improve the performance of IDSs based on blockchain technology. From the authors' perspectives, the CIDS architecture is the most proper architecture for building general architecture for IDSs based on blockchain technology because CIDSs can share data between nodes over a P2P network, which is considered an important feature in a blockchain structure.

REFERENCES

[1] Bitcoin.org (2009) Bitcoin Developer Guide. Available at: https://bitcoin.org/en/developer-guide#block-chain-overview (Accessed: 27 September 2016)

[2] Genesis block (2015) Available at: https://en.bitcoin.it/wiki/Genesis_block (Accessed 27 September 2016)

[3] Meng, W., Tischhauser, E. W., Wang, Q., Wang, Y., & Han, J. (2018). When Intrusion Detection Meets Blockchain Technology: A Review. IEEE Access, 6, 10179 - 10188.

[4] C. J. Fung, O. Baysal, J. Zhang, I. Aib, and R. Boutaba, ‘‘Trust management for host-based collaborative intrusion detection,’’ in Managing LargeScale Service Deployment (Lecture Notes in Computer Science), vol. 5273, F. De Turck, W. Kellerer, and G. Kormentzas, Eds. Heidelberg, Germany: Springer, 2008, pp. 109–122.

[5] F. Gong, ‘‘Next generation intrusion detection systems (IDS),’’ McAfee Netw. Secur. Technol. Group, Santa Clara, CA, USA, White Paper, 2003.

[6] J. Sotos and D. Houlding, ‘‘Blockchains for data sharing in clinical research: Trust in a trustless world,’’ Intel, Santa Clara, CA, USA, Blockchain Appl. Note 1, 2017.

[7] W. Li, Y. Meng, L.-F. Kwok, and H. H. S. Ip, ‘‘PMFA: Toward passive message fingerprint attacks on challenge-based collaborative intrusion detection networks,’’ in Proc. 10th Int. Conf. Netw. Syst. Secur. (NSS), 2016

[8] N. Alexopoulos, E. Vasilomanolakis, N. R. Ivanko, and M. Muhlhauser, ‘‘Towards blockchain-based collaborative intrusion detection systems,’’ in Proc. Int. Conf. Critical Inf. Infrastruct. Secur., 2017, pp. 1–12.

[9] W. Gao, W. G. Hatcher and W. Yu, “A survey of Blockchain: techniques, applications, and challenges,” in 2018 27th Int. Conf. on Computer Communication and Networks (ICCCN), pp. 1–11, 2018.

[10] X. Liang, S. Shetty, D. Tosh, C. Kamhoua, K. Kwiat et al., “Provchain: A Blockchain-based data provenance architecture in cloud environment with enhanced privacy and availability,” in Proc. of the 17th IEEE/ACM Int.sym. on cluster, cloud and grid computing, pp. 468–477, 2017.

[11] M. Muzammal, Q. Qu and B. Nasrulin, “Renovating Blockchain with distributed databases: An open-source system,” Future Generation Computer Systems, vol. 90, no. Supplement C, pp. 105–117, 2019.

[12] P. Garcia-Teodoro, J. Diaz-Verdejo, G. Maciá-Fernández and E. Vázquez, “Anomaly-based network intrusion detection: Techniques, systems and challenges,” Computers & Security, vol. 28, no. 1–2, pp. 18–28, 2009.

[13] M.-L. Shyu, S. C. Chen, K. Sarinnapakorn and L. Chang, A novel anomaly detection scheme based on principal component classifier. Miami Univ Coral Gables Fl Dept of Electrical and Computer Engineering, 2003.

[14] N. Ye and Q. Chen, “An anomaly detection technique based on a chi-square statistic for detecting intrusions into information systems,” Quality and Reliability Engineering International, vol. 17, no. 2, pp. 105–112, 2001.

[15] A. Boukerche, R. B. Machado, K. R. Jucá, J. B. M. Sobral and M. S. Notare, “An agent based and biological inspired real-time intrusion detection and security model for computer network operations,” Computer Communications, vol. 30, no. 13, pp. 2649–2660, 2007.

[16] E. A. E. R. Abas, H. Abdelkader and A. Keshk, “Artificial immune system-based intrusion detection,” in 2015 IEEE Seventh Int. Conf. on Intelligent Computing and Information Systems (ICICIS), pp. 542–546, 2015.

[17] P. Saurabh and B. Verma, “Immunity inspired cooperative agent-based security system,” International Arab Journal of Information Technology, vol. 15, no. 2, pp. 289–295, 2018.

[18] M. Jha and R. Acharya, “An immune inspired unsupervised intrusion detection system for detection of novel attacks,” in 2016 IEEE Conf. on Intelligence and Security Informatics (ISI), pp. 292–297, 2016.

[19] M. H. Chen, P. C. Chang and J. L.Wu, “A population-based incremental learning approach with artificial immunesystem for network intrusion detection,” Engineering Applications of Artificial Intelligence, vol. 51, no. 1, pp.171–181, 2016.

[20] M. Zamani and M. Movahedi, “Machine learning techniques for intrusion detection,” 2013. [Online]. Available at: https://arxiv.org/abs/1312.2177.

[21] F. Hosseinpour, P. V. Amoli, F. Farahnakian, J. Plosila and T. Hämäläinen, “Artificial immune system based intrusion detection: Innate immunity using a unsupervised learning approach,” International Journal of Digital Content Technology and its Applications, vol. 8, no. 5, pp. 1, 2014.

[22] H. H. Pajouh, G. Dastghaibyfard and S. Hashemi, “Two-tier network anomaly detection model: A machine learning approach,” Journal of Intelligent Information Systems, vol. 48, no. 1, pp. 61–74, 2017