Hindawi
Security and Communication Networks
Volume 2021, Article ID 8871230, 13 pages
https://doi.org/10.1155/2021/8871230

Research Article

Analysis of Challenges in Modern Network Forensic Framework

Sirajuddin Qureshi,[1] **Jianqiang Li**,[1] **Faheem Akhtar**,[2] **Saima Tunio**,[1] **Zahid Hussain Khand**,[2] **and Ahsan Wajahat**[1]

[1]*Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China*
[2]*Department of Computer Science, Sukkur IBA University, Sukkur 65200, Pakistan*

Correspondence should be addressed to Faheem Akhtar; fahim.akhtar@iba-suk.edu.pk

Received 3 September 2020; Revised 15 July 2021; Accepted 13 August 2021; Published 29 August 2021

Academic Editor: Neetesh Saxena

Network forensics can be an expansion associated with network security design which typically emphasizes avoidance and detection of community assaults. It covers the necessity for dedicated investigative abilities. When you look at the design, this indeed currently allows investigating harmful behavior in communities. It will help organizations to examine external and community this is undoubtedly around. It is also important for police force investigations. Network forensic techniques can be used to identify the source of the intrusion and the intruder's location. Forensics can resolve many cybercrime cases using the methods of network forensics. These methods can extract intruder's information, the nature of the intrusion, and how it can be prevented in the future. These techniques can also be used to avoid attacks in near future. Modern network forensic techniques face several challenges that must be resolved to improve the forensic methods. Some of the key challenges include high storage speed, the requirement of ample storage space, data integrity, data privacy, access to IP address, and location of data extraction. The details concerning these challenges are provided with potential solutions to these challenges. In general, the network forensic tools and techniques cannot be improved without addressing these challenges of the forensic network. This paper proposed a thematic taxonomy of classifications of network forensic techniques based on extensive. The classification has been carried out based on the target datasets and implementation techniques while performing forensic investigations. For this purpose, qualitative methods have been used to develop thematic taxonomy. The distinct objectives of this study include accessibility to the network infrastructure and artifacts and collection of evidence against the intruder using network forensic techniques to communicate the information related to network attacks with minimum false-negative results. It will help organizations to investigate external and internal causes of network security attacks.

1. Introduction

Modern research on network forensics has identified several investigation techniques through which vulnerabilities and security breaches can be highlighted. Most of these investigation techniques depend on discovering, capturing, and analyzing traffic passing through the infrastructure and network devices [1]. It is necessary to determine the objective of investigating network forensics when network security suspects are present. The research indicates several ways of conducting an investigation, which may include a retort to a specific network incident [2], analysis of archives in case of internal corporate investigation [3], and performing a criminal investigation [4]. The intentions and procedures followed in these kinds of network investigations are different; however, one of the common objectives is to analyze the traffic observed during network susceptibilities. These investigations are carried out in response to the network attacks and explain such attacks' impact on the networks. The investigation also analyzes the digital events that occur after the suspected event has occurred [5]. It helps in analyzing the pattern of events that occurred during the attack on the network. The network forensics also involves capturing the network traffic to reconstruct the entire attack and then transmitting the traffic to another device to understand the attack [6, 7]. However, this process may lead to time delays in carrying out forensics because it requires transmitting a large quantity of data from one device to another [8]. Besides, this process also affects the incident response because network forensics performance is abysmal. It means that forensic experts need to identify more efficient ways of performing network forensics and improving network security. The authors in [9–11] proposed a number of different network forensic methods over the years, which can enhance network forensic techniques' efficiency. One of these techniques' key and common objectives is to extract legal evidence from network communication channels and network security devices.

Admissible evidence plays a vital role in identifying the origin of the attack. For example, Jeong and Lee[12] proposed capturing the evidence by extracting the traffic from the router. This data can help identify the origin of the attack and the potential intruder. Regardless of the number of studies that scholars performed on network forensic techniques, Pilli et al. [3] are the only one to survey network forensics. This study outlined the tools used in network forensics, the process models, and the implementation frameworks. The scholars have not yet explored the modern network forensic techniques, specifically a comprehensive cybercrime investigation with network forensic techniques. Similarly, no evidence was found in which the scholars emphasized the implementation frameworks and the target datasets of the network forensic techniques. This particular study has been conducted considering the diverse nature of digital evidence and the difficulties that arise from the digital evidence's diverse nature while analyzing different kinds of attacks in the networks. The distinct objectives of this study include accessibility to the network infrastructure and artifacts and collection of evidence against the intruder using network forensic techniques to communicate the information related to network attacks with minimum false-negative results. Setting these objectives highlights the digital evidence, which indicates that the intruder has to invest more time and effort in carrying out the attack. This study also aims to highlight the state-of-the-art challenges existing in carrying out network forensic techniques. This study is a key contribution specifically for the security agency committees and the legislators because it can help to develop standard legal frameworks.

This study's significance is that it explores the basic structure of network forensic techniques (i.e., represented in Figure 1) and how they work to assess the nature and impact of network attacks. This paper also proposed a thematic taxonomy of classifications of network forensic techniques based on an extensive literature review. The classification has been carried out based on the target datasets and implementation techniques while performing forensic investigations. For this purpose, qualitative methods have been used to develop thematic taxonomy. The similarities and differences between different network forensic techniques have been carried out based on their objective functions, execution definition, investigation time, forensic processing, target instance, target dataset, mechanism, and nature of the framework. Finally, this study has discussed the open research challenges that may occur while selecting the domain for further research within network forensics and identifying the most effective techniques.

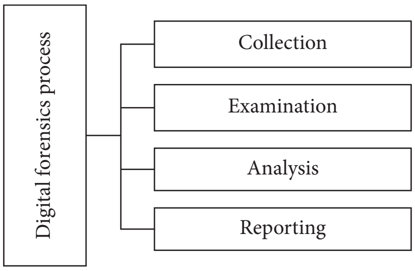

Figure 1: Digital forensics.

1.1. Related Work and Significance of Network Forensics.

One of the key motivational factors that emerged within the forensic network domain includes the emergence of the information technology (IT) industry and its apprehension on security. Most of the world's modern organizations are concerned about the safety of their data and networks because of the cybersecurity attacks observed in the last decade [13]. In the previous decade, many attempts were made on different social media websites, including Twitter, Facebook, and Google Blogger.

Disgruntled users have attacked different social media websites using DDoS, and the principal objective of these attacks was to crash the functioning of these platforms. Some of these attacks can be categorized as phishing attacks in which the intruders attack to acquire personal information, that is, bank account passwords, to steal money from the bank accounts. It is a criminal activity, and the conviction of these intruders requires digital shreds of evidence. The purpose of raising security is to stop such attacks from intruders. Perry [14] argued that the network traffic flow passes through the Internet service providers (ISPs), so the ISPs should be held responsible for attacks from outside networks. Furthermore, the ISPs should stop the malicious packets of data, which may result in network attacks. Most of the world's large companies are utilizing the online business transactions and face a higher threat of breach in the security. Most of such business operations are very large, and any breach of security may push these businesses to file for financial bankruptcy. The scholars argue that cybersecurity is the backbone of large and small companies and is the primary concern for these companies in the current and future time. Federal Information Security has defined comprehensive cybersecurity programs. Federal Information Security has defined comprehensive cybersecurity programs (Management Act (FISMA) for the federal agencies). [15]. Similarly, the healthcare data is also susceptible to malicious attacks, and the Health Insurance Portability and Accountability Act (HIPAA) has defined security plans for healthcare organizations (1996). Different companies have developed their market portfolios based on the security of their e-business, e-transactions, and other Internet-based activities, and they are using their portfolios to attract more customers. These companies claim that they can protect the interest of their customers. These companies continuously analyze the traffic to detect the potential malicious attacks as soon as possible and deal with them in time.

The term “malicious” may refer to the malicious packets of data or malicious traffic programs. The malicious traffic programs are irregular traffic patterns. Malicious data packets can be identified as those that violate network communication principles. These malicious packets attack the network by exploiting the vulnerabilities of the devices installed for security purposes, which may occur, including the gateways and the attempts to gain unauthorized access in the network. In the case of malicious packets, various packet fields contain forged information, that is, port numbers, TCP flags, and IP address [16]. For example, IP spoofing is a specific kind of attack in which the attacker uses a spoofed IP, and it may appear as a trusted node. The attacker can be connected with the victim node by registering as a trusted user on the network [17]. The land attack is another form of spoofing in which the intruder uses the victim’s destination and source of IP addresses. The victim of this attack enters into a loop of self-connection attempts. The intruder can alter the TCP flags to indicate several events, including pushing off the data, highest priority of data, starting of connection, and ending. The SYN TCP flag indicates the starting of the connection, and the FIN TCP flag indicates the ending of the connection. A combination of these two flags, that is, SYN-FIN, can be used by the intruder to avoid detection by the system’s security. Most of the available intrusion detection systems detect the connection’s starting (SYN TCP flag) or ending (FIN TCP flag). Therefore, any attempts to start or stop the connection are 112 considered unauthorized. The intruders can also attack by altering the packet’s port numbers for source and destination ports and can communicate the packets abnormally. The networking device discards the packet if the intruder assigns the same port numbers to both of the ports. The communication also becomes suspicious when the intruder tampers with the packet or performs fragmentation of the packets [18]. The packet fragmentation is performed when the packet’s size is too large to be transmitted. The intrusion can be conducted in the form of tiny fragment attacks; two TCP fragments are formed from each packet. Each of these fragments contains little information, and they are transmitted in the form of tiny fragments. The network devices cannot detect the tiny fragments, and bypassing the security protocols is the outcome of such transmission. On the other hand, some intruders use large packets of information to perform an attack. The packets are reassembled at the receiving device; however, when the reassembled packets become too large, the reassembling process becomes disturbed, affecting the entire network. Such an attack is known as the Ping of Death Attack. The intruder uses echo request messages, and the size of packets in such a message is larger than the regular size packets [19]. The network operators carry out active monitoring of the events to detect malicious programs and packets. Anomaly detection is one of the techniques of active monitoring. Some other techniques also include honey pots, access control lists, intrusion detection systems, and signature scan detection. In anomaly detection techniques, the system creates patterns of behavior of the user and the network’s resources. The irregular pattern of traffic is detected as malicious by the anomaly detection technique [20]. The signature scan is one of the detection techniques in which traffic signatures are stored in the network’s database. Recognition of irregular patterns is performed with the help of passive scans [21]. The malicious activity is detected when the signature matches the one stored in the database. This technique is particularly useful when the attackers in the system are known. The intrusion detection systems work based on statistical anomaly and matching patterns [22]. The 136 malicious traffic is detected using statistical anomaly when the usage patterns deviate from the normal usage patterns. The system forms the standard usage patterns, and the purpose of creating them is to identify any deviation from the standard usage patterns. Access control list is a technique in which the rules for determining malicious activities are predefined, and the intruders are detected based on matching packet headers [23]. The honeypot technique is the one that acts as the trap for the intruders. It prevents users from entering the secure areas of the network [21]. The trap that honeypot forms is a disguise whose role is to protect the server by replicating and persuading the attacker to interact with the network. The honeypot requires open ports to invite the intruder, and the attack is detected when the intruder interacts with one of these ports. Network forensics reconstructs the sequence of attacks and detects the intrusion based on historical network data [9]. It collects and captures the network packets to form emails, FTP traffic, messages, and other communication forms. Network forensics is significant to detect attacks in identifying the problems in critical business systems. Some of the other functions performed by the network forensics may include monitoring the workflow regularly for corporate defiance, enhancing the network’s performance, shielding against the viruses, and locating the device that has the potential to generate an attack.

2. Research Methodology

2.1. Classification of Modern Forensic Techniques. This section consists of the proposed taxonomy of network forensic techniques; their evaluation, implementation, and the critical review of these techniques are presented as presented in related domain [24–28]. Figure 2 shows the components of modern forensic techniques.

2.2. Traceback-Based Network Forensic Technique. Traceback is a specific term used when the origin of the packet is to be identified in a network. It is also known as the IP traceback [12]. IP tracing is a useful tool for analyzing and attribution of network assaults, Figure 3. This technique determines the origin of the attack by identifying the device from where packets are generated. The traceback technique is useful when packets’ origin is to be identified in case of spoofing attacks and DDoS attacks [29]. The DDoS and botnet attacks are mostly observed in the distribution

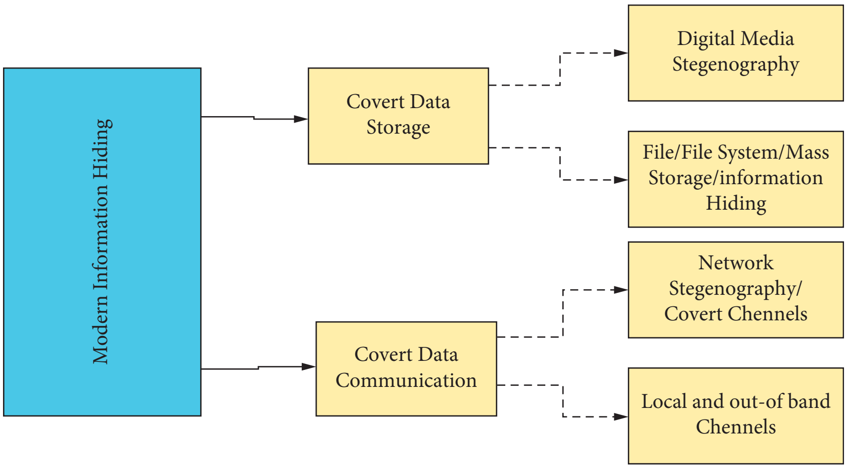
Modern Information Hiding
Covert Data Storage
Covert Data Communication
Digital Media Stegenography
File/File System/Mass Storage/information Hiding
Network Stegenography/ Covert Chennels
Local and out-of band Chennels

Figure 2: Components of modern forensic techniques.

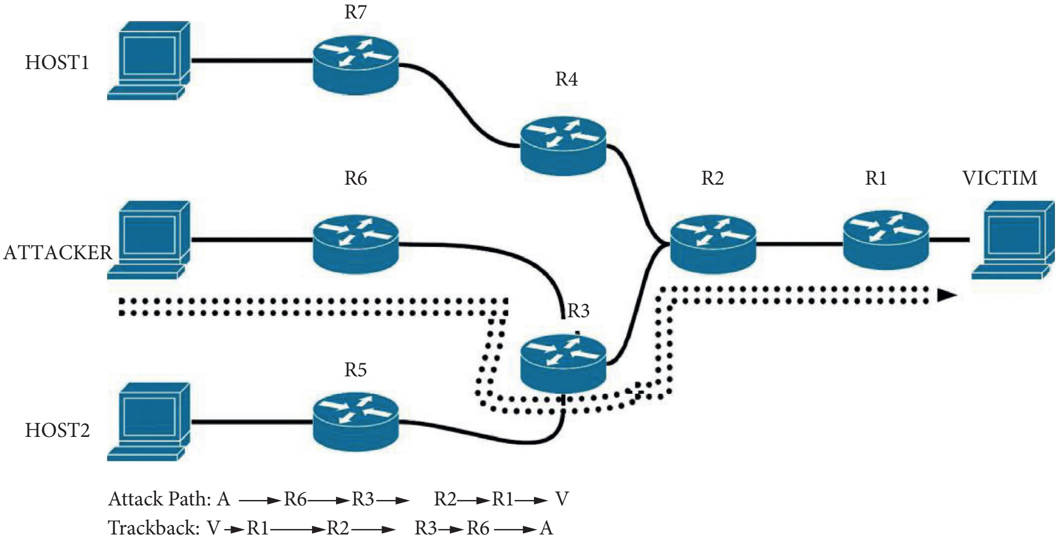
R7
HOST1
R4
R6
R2
R1
VICTIM
ATTACKER
R3
R5
HOST2
Attack Path: A → R6 → R3 → R2 → R1 → V
Trackback: V → R1 → R2 → R3 → R6 → A

Figure 3: Network forensic using IP traceback.

networks, and the traceback technique is useful in such attacks. The distributed networks are more susceptible to attack because they collaborate with the Internet, and the atmosphere is favorable for the attacks by the bot-masters [30]. The network requires various traceback systems to overcome these attacks efficiently:

Network forensics: the employment of scientifically proved processes to gather, fuse, determine, examine, correlate, evaluate, and document evidence from this is undoubtedly electronic, definitely processing and transmitting digital resources for the intended purpose of uncovering facts related to the planned intent or assessed success of unauthorized tasks supposed to interrupt, corrupt, and/or compromise system components too as providing information to help in response to or recovery from these tasks

Analysis time: forensics covers real-time and includes security for live network surveillance and its monitoring system. However, investigation of post-mortem captures packet is operated offline

Source of data: flow-based process mainly collects statistical records in the form of the flow of network traffic where packet-based tool includes thorough packet inspection

2.3. Converge Network-Based Network Forensic Technique. Converge network-based network forensic techniques are specifically useful in identifying the digital evidence found in the converged networks. The VoIP communication is a specific example of the converged network. VoIP requires a medium for data communication due to which it faces several kinds of vulnerabilities, security threats, and attacks. The communication signals in VoIP are divided into the form of frames, and these frames are embedded as voice codes in the data packets. These data packets are communicated as simple voice packets on the IP network. When the voice packets are transmitted from a sender to a receiver without any modification and interference, they are known as 181 normal voice packets. Typically, the voice packets are transmitted over the IP networks using H.3231 and SIP protocols [31]. The port and IP addresses information is enclosed in the voice packets, assisting the communication protocols. The communication protocols act as session control 184 protocols. The ports and IP addresses attached in the voice packets are not encrypted because the address translation devices have to translate the voice packets. The lack of encryption of the voice packets makes them susceptible to attacks from intruders. The intruder can exploit the voice packets during transmission, which changes the normal voice packets to the 188 malicious voice packets. The malicious voice packets may take several forms including the exploitation of VoIP devices, degrading call integrity, privacy leakage, eavesdropping, man-in-the-middle, buffer overflow, hijack calls, and flooding. Figure 4 can be referred for details.

VoIP network forensic analysis involves identifying the malicious packets from the normal packets [10]. The intruders inject malicious or abnormal packets during the transmission process. Lin et al. [10] offered a solution to such attacks by collecting digital shreds of evidence while performing a forensic investigation. The digital proof is in the form of information received from packet value, TTL, service type, protocol, and the packet's payload. A change found in the packets means that alteration has occurred, and the packet becomes malicious. The scholars have identified different fields of voice packets to differentiate between different types of VoIP-NFDE. However, a common issue is scalability, which is considered during the investigation of large integrated networks. Another problem is that the investigation process may be prolonged because of the reestablishment of a communication link between IP phone users and the SIP registrar on its disconnection. Reestablishing the connection is time-consuming, and some useful data may be lost during the process. Similarly, storage resources are also required to investigate the attack patterns collected from the voice data.

2.4. Attack Graphs-Based Network Forensic Technique.

Attack graph-based network forensic technique utilizes the attack graphs to recognize all the potential attack paths which an intruder used while performing the attack. This process requires analyzing networks, hosts, and other security devices [32]. An attack graph constitutes vertices, and each vertical is a potential attack node. The edges in Figure 5 represent the state transitions between different attack nodes. The attack graphs are very useful in network forensics because they visualize the nodes that can be attacked and highlight the worst paths with the most significant threat of attack [5]. Identifying such nodes can help the network administrator design the security before the actual attack occurs. The attack graphs are used for several other purposes, including the cost-benefit security harden [33], evidence collection [33], recognizing multistage network attacks [29], and impact analysis [34].

The interaction and visualization framework is used in the attack graphs, and the purpose of using this framework is to study the intrusion behavior of the attack. There are thousands of edges and vertices, and it is a very time-consuming process to identify those susceptible to attack. Many research studies have depicted the attack graphs that can be used for different aspects, including critical systems, data reduction, attack dependency graph, virtual exploitation information, and others. However, only a few studies have emphasized the visualization process. The attack graphs are very complex in large integrated networks, and RAVEN architecture was proposed to embed the visualization feature in the attack graphs. The principal purpose of introducing this architecture was to reduce the graphs' complexity in large networks. The RAVEN architecture has a visualized interface, and the investigators can interact with it with less complexity due to this visualized interface. RAVEN has a collaborative environment that is analytical, and it has several gesture controls as well. The RAVEN architecture has a human-computer interaction platform that allows the investigators to manage the attack graph more effectively. A multitouch technology was integrated into the human-computer interaction platform to make it easier for the investigator to view each node and how they interact with the entire network. Some of the other vital features of the RAVEN architecture include non-real-time visual support in the real-time environment. However, the RAVEN does not have a composite layout, due to which the investigators cannot observe the multiple attack paths in parallel. Another disadvantage of using RAVEN is that it cannot mine the related information from the network efficiently. Assaults tend to be foiled by the utilized services and products, and novel attacks still circumvent prevention services and products without being detected. During these circumstances, examining the assaults is just a task; this is certainly very challenging. Quite often, severe attackers tend to be skillful at concealing evidence. Consequently, firewall logs and intrusion detection notifications may totally miss these assaults or may prove to be insufficient for the examination. This is certainly extensive, especially when the target is to apprehend the perpetrator.

2.5. Distributive-Based Network Forensic Technique.

Distributive-based network forensic techniques can distribute the data agent systems and forensic network servers to resolve scalability for network forensic techniques. The forensic network servers for analysis collect the data from different data server agents located at various locations in the network. The distributive network forensic technique performs the evidence collection process, recognizes the origin of the attack, and performs investigation [35]. The distributive network forensic technique keeps the server secured from the attackers by creating an overhead. The forensic servers are distributed throughout the network, and they are susceptible to attack. A distributed framework proposed by [36] works with different network devices and records their network logs. The network logs are stored at various locations in a network environment. The normal process is to

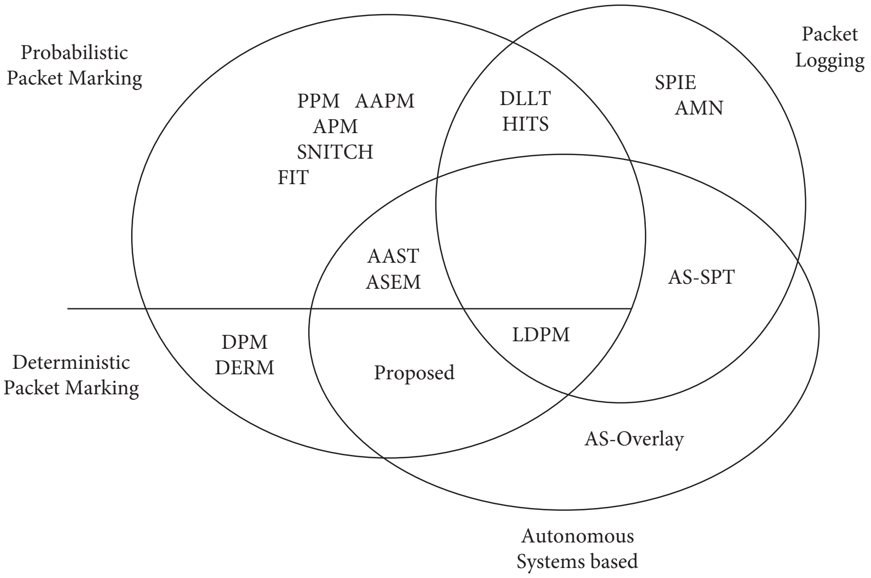

Figure 4: Relationship among traceback techniques.

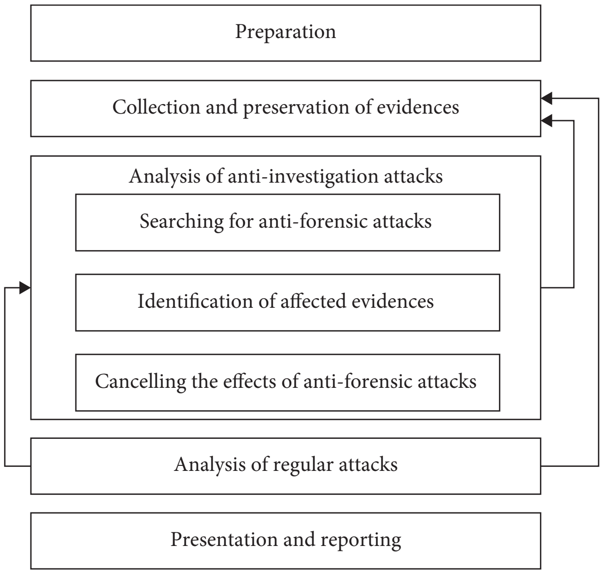
Preparation
Collection and preservation of evidences
Analysis of anti-investigation attacks
Searching for anti-forensic attacks
Identification of affected evidences
Cancelling the effects of anti-forensic attacks
Analysis of regular attacks
Presentation and reporting

Figure 5: Flow of process of attack graph-based forensic technique.

analyze the host and packet logs manually. The manual analyses create several problems because of delays in the response time, synchronization among records, improper logging mechanism, and low response time. For Net framework is the Network Forensic Framework that improves evidence collection and resolves most of the abovementioned issues. The pieces of evidence are collected from the network devices such as the routers and switches by installing an application on each network. This application is known as SynApps. This application does not only collect the data but also summarize the information collected over a long period of time. Evidence regarding the network's vulnerabilities is taken from the packet header, which is more credible than data collected from the payload data [37]. For Net uses bloom filter tracking to investigate session creation among different hosts, maintaining port connection records and IP connections. However, an essential feature of the For Net framework is the storage of raw data regarding networks, specifically in an extensive integrated network. However, the intrusion detection system on For Net is lightweight and cannot detect some attacks. Most of these attacks are DDoS, which sends the rogue queries, and the purpose of sending queries is to utilize the resources of the investigating servers. The intruder can perform modification of packets during this process. Furthermore, the intruders can also modify the logs which are transmitted through insecure 260 communication channels.

2.6. NFT Using Intrusion Detection System. The intrusion detection system is a network forensic technique that monitors and prevents malicious attacks, especially when the intruder tries to exploit the network [37]. The IDS detects the intrusion and triggers the alert system in the form of a message. The IDS informs the management system of the network to take appropriate actions. The IDS is especially essential when the intrusion threatens the confidentiality and integrity of the network [37]. IDS uses the logging approach to analyze the network intrusion, reliability of the evidence, and dynamic forensics and to describe the forensics. If an incoming packet label is in NFT, as shown in Figure 6, the packet is routed normally. Several network forensic techniques use the IDS to identify the network breaches and assist the forensic process.

Sy [38] proposed Analytical Intrusion Detection Framework (AIDF) by merging the message alert system from the IDS and forensic analysis conducted due to intrusion. The outcome of AIDF is a forensic explanation based on unreported signature rules and observed IDS alerts. AIDF uses a probabilistic approach to minimize the number of attacks that unfolds the hidden information and model the attacks. The AIDF recognizes the intrusions hidden in

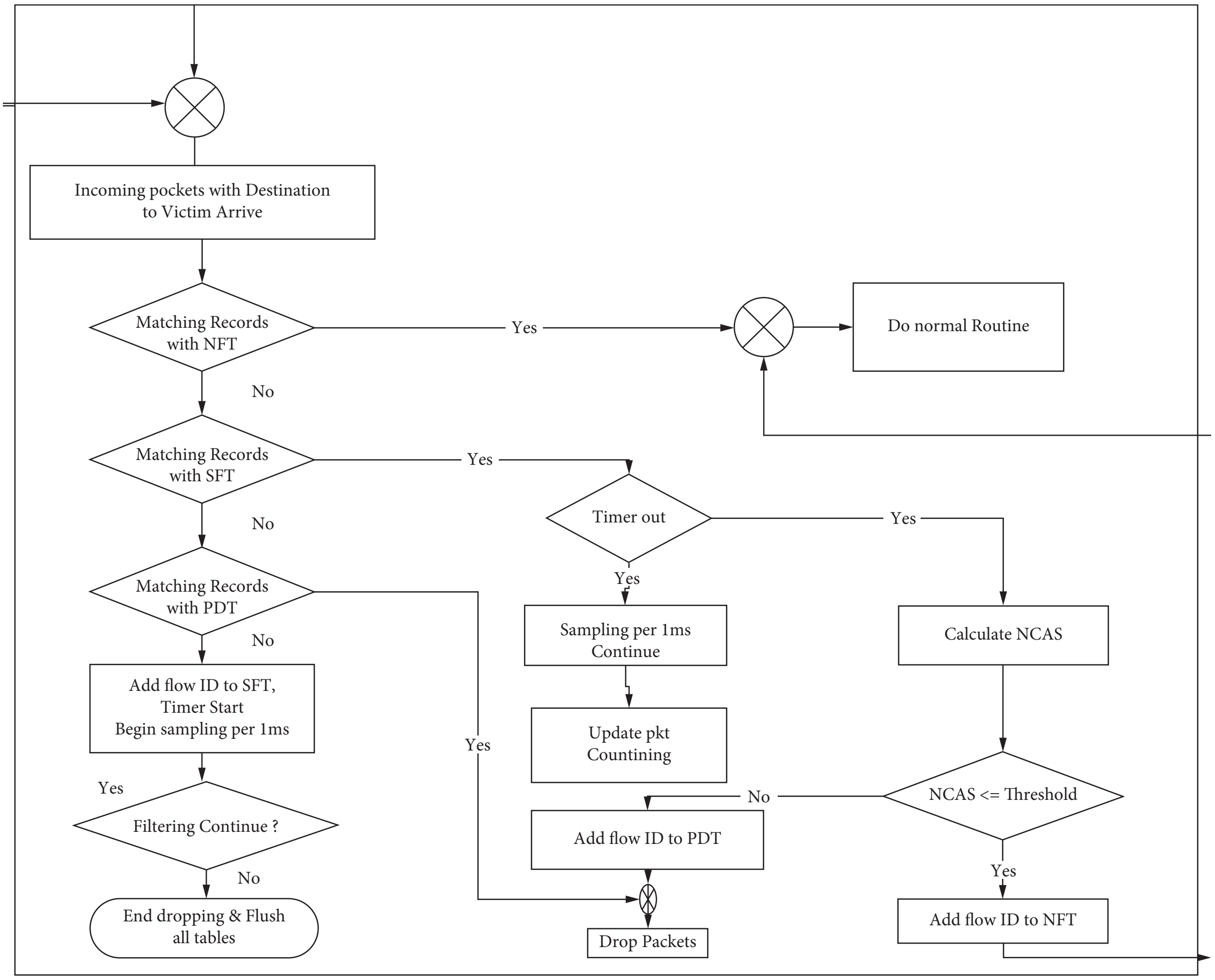
Incoming pockets with Destination to Victim Arrive
Matching Records with NFT
Yes
Do normal Routine
No
Matching Records with SFT
Yes
Timer out
Yes
No
Matching Records with PDT
Yes
Sampling per 1ms Continue
Calculate NCAS
No
Add flow ID to SFT, Timer Start Begin sampling per 1ms
Yes
Update pkt Countining
Yes
No
NCAS <= Threshold
Filtering Continue ?
Add flow ID to PDT
Yes
No
End dropping & Flush all tables
Drop Packets
Add flow ID to NFT

Figure 6: NFT network forensic.

the network traffic using the signature rule, which compares the packets with the already encoded packets. The probabilistic inference process is carried out through using Snort, which triggers the rule when a specific packet matches the pattern prerecorded in the signature rule. These rules assist the investigators in recognizing the potential attacks on the network, and they also help the investigators to define new rules to prevent future attacks. AIDF also has some disadvantages; for example, the AIDF does not know based modules because it cannot store hidden and untreated data. AIDF cannot be used to prevent future attacks because of this disadvantage. The storage of unprocessed information can save a lot of time during the investigation process, and it can help to generate precise results. This mechanism can be used in creating intrusion detection alerts that can be used in real-time situations.

2.7. Review Analysis of Modern Network Forensics. This section of the paper discusses the open challenges faced by modern network forensic techniques (Figure 7). These challenges are significant to be studied while investigating large integrated networks such as cloud computing and software-defined networks. Following are the open challenges extracted from review analysis and are in line with the network forensic techniques.

2.8. High-Speed Data Transmission. High-speed data transmission is one of the biggest challenges for network forensics because it cannot capture and record all the packets on the network because of high speed. Millions of data packets are transmitted on networks within a short period, and these packets pass through a vast number of interconnected devices. The network devices can play a significant role as evidence as the network data transmit through them. To identify the network data's susceptibilities, it is necessary to record the data packets at high speed; however, it is a very time-consuming process. Most companies enhance and expand their network structures. For this purpose, the companies connect distributive infrastructures to their high-speed networks. However, in most cases, network traffic is not entirely captured by the distributive infrastructures, and incomplete logs of network information are obtained. The reconstruction of suspicious attacks becomes more difficult because of these incomplete logs, and it becomes tough to recognize the origin of the attack. This problem can be resolved only when capturing, indexing, preserving, and analyzing the data packets on a network are carried out in a real-time situation. A review of the literature suggests three distinct solutions for the aforementioned problems. These solutions include distributive-based solutions, software-based solutions, and hardware-based solutions [19]. The hardware-based solution requires installing a separate high-speed network traffic capturing device. This device can obtain specific data, and it can carry out a real-time analysis. The response time of this device is quick. The software-based solution requires installing software on the network. The nCap library is the software that is specifically designed to capture high-speed traffic on a network. This software is used to program the customized hardware for this purpose. This software is installed using the space of the user rather than utilizing kernel space. The programmers can quickly deploy this software, and it can perform the functions of capturing and carrying out analysis quickly. The distributed solution involves using a distributed packet capturing technique on high-speed networks. This technique provides the load balance within different nodes and minimizes the cost of CPU cycles and memory.

2.9. Data Storage on the Network Devices. The amount of data captured and stored on the network for carrying out the investigation is tremendous. Such a large amount of data creates problems for forensic experts while retrieving relevant information from these networks [39]. The interconnectivity devices' storage capacity is low, and huge storage space is required to store the captured data packets. Apart from this, the problem is resolved by designing a framework for capturing data by a machine based on time [9]. This framework minimizes the need for a massive storage speed and can also enhance the investigation process's pace.

Moreover, the General Processing Unit (GPU) offloads the indexing packets when it carries out a compressed bitmap index in real-time. The GPU performs deterministically, which results in parallel operations at the same speed because of the advance and faster memory interfaces. The storage speed reaches approximately a million records per second. In addition to this infrastructure, n2disk architecture can also be used for this purpose. It can be used for single and multiple threat packet consumers, and it can search the packets from the dump files efficiently. This infrastructure can resolve the issue of storage space in high-speed networks.

2.10. Data Integrity. Data integrity is one of the major concerns for the investigators while performing the network forensics. Data integrity means that the network must have the most consistent, complete, and accurate data. Analyzing data integrity on the networks is one of the most challenging and critical tasks for the investigators. Maintaining data integrity is difficult, considering several factors, including velocity, size, and scope of data. Network complications become higher when the trust and integrity of the data and data system become low. Several causes of little integrity may include frequent mobility of data, system malfunctioning, malicious attacks, software errors, and hardware errors. The process of network forensic is adversely affected when the data loses its integrity because of deliberate and intentional efforts [5]. Data integrity is an essential factor while prosecuting the intruder in the court of law.

The data integrity should be maintained using the end-to-end approach. It means that the use of software, as well as the hardware, should be seamless. Modern networks are growing at a breakneck pace, and it is necessary to get early updates about the problems and resolving the issues of the network as soon as possible. Standards and appropriate methodologies are required to efficiently achieve the objective of cost-effectiveness to maintain the integrity of data, specifically in large and distributed networks. Reference [5] proposed a GUI-based monitoring system in which the server carries out the analysis of the network packets and then transmits them to client nodes for storage. It is a reliable system because it improves the data packets' analysis on the network with real-time characteristics 348 and then stores them in the storage spaces owned by the clients where they are safe from different kinds of vulnerabilities.

2.11. Data Privacy. Data privacy holds one of the critical positions within the realm of network forensics. The problem mentioned above can be resolved using a forensic attribution solution. The forensic investigator can have a look at the particular data but only by verifying the signature. This process is known as forensic attribution in any network. Cryptographical tools can be used for this purpose, which may include BBS short group signatures and group signatures. It means that anyone of the group members can create a signature, which is verifiable by the rest of the group; however, the identification of the creator cannot be performed without authentication from the rest of the members of the group. BBS small signature group has short signatures as compared to the group-signature scheme. It is in the form of a clue that is associated with the optimization of the evidence. The cryptographical tools used for this purpose ensure that only known parties have awareness about the hardware's physical identity, which is transmitting the IP packets on a network. In contrast, others cannot identify their physical status. As a result, the issue of data privacy is resolved, which arises while analyzing the integrated networks.

2.12. Access to IP Addresses. Identification of the IP address of the attacker is an essential step in carrying out network forensics. The IP address of the source provides information about the origin of the attack [12]. Identifying the IP address can lead the investigators to the intruder and prevent future attacks from the same intruder. The intruders use several techniques to hide their IP addresses from the various devices installed on the network. Spoofing the IP address is one such technique in which the intruder can show a fake IP address to the devices registered on the network. Spoofing is the technique that is mostly used in DDoS attacks. The purpose of conducting DDoS attacks is to bombard the network with enormous traffic from different suspect systems. Identifying the original IP address in case of spoofed IP address becomes very difficult for forensic investigators, specifically in large integrated networks. A Source Address Validation Improvement (SAVI) solves the problem mentioned above, as it binds the source host Mac address, IP address, and uplink port properties. It prevents intruders from spoofing the IP address by restricting the attached nodes to stay connected with the same uplink. Additionally, the SAVI uses traceback and antispoofing for IPv4/IPv6 transition by 377 extracting common and crucial properties. In short, this system works very effectively to prevent spoofing.

2.13. Location of Data Extraction. The process of network forensics becomes challenging due to the virtualized characteristics and distributive nature of networks. It becomes difficult for forensic experts to identify the device and appropriate location for extracting data. It is almost impossible to handle all links and the connected devices on the networks where thousands of devices are connected and millions of packets of data pass through each device every second. Extracting data from any location in such networks becomes a radical challenge for network forensics because they may breach privacy or affect data integrity at any point. Besides, many devices on a network are designed to extract and analyze the data for network forensics, including packet sniffers, protocol analyzers, network forensic analysis tools, firewalls, IDS, and routers. The appropriate placement of these devices to extract and analyze data from the data network is a key challenge for network forensics.

2.14. Forensic Networks in Mobile Cloud Computing (MMC). MCC network cloud services are obtained by smart device users associated with long-term evolution networks via Wi-Fi, WLAN, and 3G/4G/5G. These networks must be very quick and protected enough to send user requests to parallel computing clouds and return results to connected devices' users. However, attackers target these networks to connect and get access in the form of a network attack. The current security plan covers the use of firewalls and IDS to detect and identify attack patterns. However, smart attacks use these security approaches to spread malicious network activities. Protection systems should be smart enough to detect smart attacks that threaten the system. The network position shows a network address, which connects two entities in MCC. MCC generally includes these networks: data center, cloud access, and intercloud networks [40], as shown in Table 1. The value of the forensics process lies in every aspect of MCC's network channels of communication. NFF must monitor malicious activity through network packets if a smart device user is connected to mobile clouds or data centers are connected or linked to other cloud data centers. NFIs have limited or no access to examine various network susceptibilities [41]; therefore, the forensic investigation should become a permanent service for MCC users via channels and secure cloud resources (Table 2). The linking network is called the intercloud network between two or more clouds. This network deploys a high capacity and high-speed line rate fiber optics network. The intercloud network is used when one domain migrates or transfers an application for execution or storage to another domain. The intercloud network provides a dedicated network, and through the protocol, optimization increases the transfer speed. Table 3 provides a brief overview of each network position within MCC. All network positions in MCC are considered vulnerable to NFF attacks. Due to vulnerabilities, no network is safe from a 411 attack, which requires further investigation to find the attack's origin.

(i) Scalability
(ii) Overhead computational
(iii) Data storage
(iv) Data accuracy
(v) Complexity
(vi) Privacy/security
(vii) Adaptability

The data is analyzed using SPSS (version 16), and all the graphical illustrations of study variables are interpreted accordingly. Moreover, the description of codes is as follows: scalability: horizontal (HT), vertical (VT), both (BT), and not applicable (N/A); overhead computational: high (H), moderate (M), low (L), and not applicable (N/A); data storage: high (H), moderate (M), low (L), and not applicable (N/A); data accuracy: high (H), moderate (M), low (L), and not applicable (N/A); complexity: implementation (IM), analysis (AL), collection (CL), and investigation (IV); privacy/security: high (H), moderate (M), low (L), and not applicable (N/A); adaptability: difficult (D), high (H), moderate (M), low (L), and not applicable (N/A).

2.15. Data Analysis Results. Frequency analysis is suitable for categorical data of the current study. The data size is 22 covering each theme of Network Forensic Framework (NFF) for each variable of the study. The variables are analyzed as shown in the outputs as follows. Table 4 shows the experimentation results. It can be perceived that the scalability is the highest for the horizontal category having 50%, whereas both-sided scalability is very less and is 1%. Besides, overhead shows the maximum percentage for moderate (45.5%) while the minimum percentage exists for low (22.7%). Furthermore, the output (data storage) represents 435 that out of the total themes of NFF; moderate and low data storage is at the same percentage of 40.9% and is very less for high and not appropriate, showing 9.1%. Data accuracy shows the minimum percentage for the category high and is 9.1%, whereas not applicable is 50%, and 27.3% of the data is accurate for NFF, which is low. Besides, the highest complexity proportion exists for implementation and analysis both together and is 40.9%. In contrast, data complexity collection and analysis are less. This paper proposed a thematic taxonomy of classifications of network forensic techniques based on extensive. The category was performed

Table 1: Various positions using mobile cloud computing (MCC).

Positioning	Entities link	Example	Objective	Network accessibility
Cloud access network	User cloud services	Internet, NGN, and 4G	Dynamic routing and accessibility to cloud	Possible
Data center network	Data center	Cluster computing	Load balancing, virtualization, and intensive computing	CSP
Intercloud network	Cloud system	Cloud resource migration	Cloud collaboration	CSP

Table 2: Various problems in current network forensic and mobile cloud computing (MCC).

Issues	Current network forensic	MCC network forensics
Data acquisition	No	Yes
Access of artifacts	No	Yes
Bandwidth utilization	No	Yes
Chain of custody	No	Yes
Data integrity	No	Yes
Privacy	No	Yes
Real-time analysis	No	Yes
Volatile data	No	Yes
Forensics tools	No	Yes

Table 3: Network position within the MCC.

NFF techniques	Scalability	Overhead computational	Data storage	Data accuracy	Complexity	Privacy	Adaptability
Traceback	N/A	H	H	L	IM	L	N/A
	HT	H	M	M	AL	M	D
	VT	L	L	H	IV	L	M
	HT	M	N/A	N/A	IM	N/A	D
	N/A	H	H	M	IM	M	D
Converge network	N/A	H	L	L	IM, AL	L	D
	N/A	M	M	N/A	IM, CL, AL	L	D
	HT	M	M	N/A	IM, AL	L	N/A
Intrusion	HT	M	M	N/A	IM, AL	N/A	D
	HT	H	L	N/A	IM, AL	N/A	D
	N/A	H	L	N/A	AL	L	D
	N/A	L	L	L	IM, AL	L	D
Attack-based graph	HT	M	L	N/A	AL	L	L
	N/A	M	L	N/A	IM, CL, AL	M	L
	N/A	H	M	L	IM, AL	M	D
	HT	M	M	H	AL	M	M
	HT	L	L	M	AL	M	H
	HT	L	N/A	N/A	AL	H	M
Distributive	VT	M	M	N/A	CL, AL	L	M
	BT	M	M	N/A	AL	L	L
	HT	M	M	L	CL, AL	L	L
	HT	L	L	L	IM, CL, AL	L	M

considering the target information units and execution strategies while doing investigations is forensic. The qualitative practices were made use of to develop thematic taxonomy for this function. The objectives of this study include availability to the system infrastructure and artifacts and collection of research against the intruder system that utilizes practices to communicate the information regarding community attacks with minimal false-negative issues.

Table 4: Analyzed results of the selected variables using IBM SPSS (version 16).

		Frequency	Percent	Valid percent	Cumulative percent
Scalability					
Valid	HT (horizontal)	11	50.0	50.0	50.0
	VT (vertical)	2	9.1	9.1	59.1
	BT (both)	1	4.5	4.5	63.6
	N/A (not applicable)	8	36.4	36.4	100.0
	Total	22	100.0	100.0	
Overhead					
Valid	High (H)	7	31.8	31.8	31.8
	Moderate (M)	10	45.5	45.5	77.3
	Low (L)	5	22.7	22.7	100.0
	Total	22	100.0	100.0	
Data storage					
Valid	High (H)	2	9.1	9.1	9.1
	Moderate (M)	9	40.9	40.9	50.0
	Low (L)	9	40.9	40.9	90.9
	Not applicable (N/A)	2	9.1	9.1	100.0
	Total	22	100.0	100.0	
Data accuracy					
Valid	High (H)	2	9.1	9.1	9.1
	Moderate (M)	3	13.6	13.6	22.7
	Low (L)	6	27.3	27.3	50.0
	Not applicable (N/A)	11	50.0	50.0	100.0
	Total	22	100.0	100.0	
Complexity					
Valid	Implementation (IM) and analysis (AL)	9	40.9	40.9	40.9
	Implementation (IM), collection (CL), and analysis (AL)	4	18.2	18.2	59.1
	Analysis (AL)	6	27.3	27.3	86.4
	Collection (CL) and analysis (AL)	3	13.6	13.6	100.0
	Total	22	100.0	100.0	
Privacy/security					
Valid	High (H)	1	4.5	4.5	4.5
	Moderate (M)	6	27.3	27.3	31.8
	Low (L)	12	54.5	54.5	86.4
	Not applicable (N/A)	3	13.6	13.6	100.0
	Total	22	100.0	100.0	

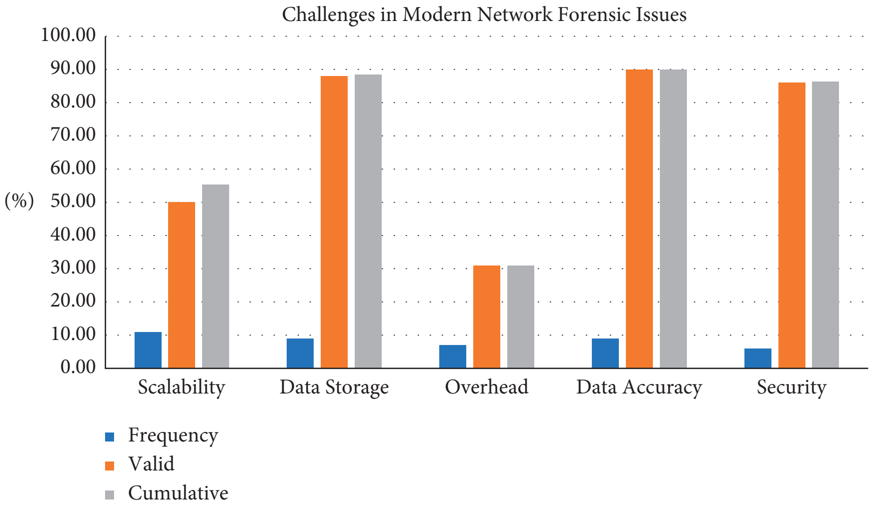

Figure 7: Challenges in modern network forensic issues.

Among all categories showing a portion of the data storage 88.6%, the percentage of security/privacy is highest, which is 88.5%, and is very less in the category of accuracy, which is high showing 90%, which means that data is not secured and has weak or no privacy system NFF. Lastly, the adaptability to MCC is very difficult in cloud computing. Data also shows the percentage of 100%, which is the highest, is least for high adaptability, and is 90.5%, respectively.

3. Conclusions and Future Research

This research reviewed the subject matter of network forensic techniques used to gather and investigate the legal information regarding the intruders. The investigators have to consider many factors, including the integrity and reliability of attack, the origin of the attack, the objectives behind the attack, determining the worst path susceptible to attacks, and highlighting the actual attack paths. The network forensic goals can be achieved when forensic experts are well aware of the nature of the attack. Also, they are aware of the challenges of network forensics associated and the tools they select to perform the network forensics. The network forensic techniques play a significant role in capturing, identifying, recording, and analyzing the legal information, specifically in integrated networks. Forensic experts face several challenges while performing forensics, and details about each of the 455 problems are provided in the paper's previous sections. The network forensic experts need to emphasize developing more intelligent network forensic tools instantly. This is the only way through which they can minimize the abovementioned challenges in network forensics. Besides, they can also reduce the storage requirements and delays in network forensics, can work in high-speed networks, and can also maintain the privacy and integrity of data. The forensics should also explore cloud computing networks, especially mobile cloud computing because mobile devices will also be the most important and widely used devices sooner. The classification has been carried out based on the target datasets and implementation techniques while performing forensic investigations. For this purpose, the qualitative methods have been used to develop thematic taxonomy. The objectives of this study include accessibility to the network infrastructure and artifacts and collection of evidence against the intruder using network forensic techniques to communicate the information related to network attacks with minimum false-negative results.

Data Availability

Experimental data available within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] K. Jiang and R. Xuan, "Book review: guide to computer forensics and investigations," *Journal of Digital Forensics, Security and Law*, vol. 3, no. 5, p. 467, 2008.

[2] R. Bejtlich, *The Practice of Network Security Monitoring: Understanding Incident Detection and Response*, p. 469, No Starch Press, San Francisco, California, 2013.

[3] E. S. Pilli, R. C. Joshi, and R. Niyogi, "Network forensic frameworks: survey and research challenges," *Digital Investigation*, vol. 7, no. 1-2, pp. 14–27, 2010.

[4] M. Rasmi and A. Jantan, "A new algorithm to estimate the similarity between the intentions of the cyber crimes for network forensics," *Procedia Technology*, in *Proceedings of the 4th International Conference on Electrical Engineering and Informatics (Iceei 2013)*, vol. 11, pp. 540–547, Malaysia, Malaysia, June 2013.

[5] B. Cusack and M. Alqahtani, "Acquisition of evidence from network intrusion detection systems," in *Proceedings of the 11th Australian Digital Forensics Conference*, Perth, Western Australia, December 2013.

[6] B.-C. Cheng, G.-T. Liao, H.-C. Huang, and P.-H. Hsu, "Cheetah: a space-efficient HNB-based NFAT approach to supporting network forensics," *annals of telecommunications - annales des télécommunications*, vol. 69, no. 7-8, pp. 379–389, 2014.

[7] D. Wang, T. Li, S. Liu, J. Zhang, and C. Liu, "Dynamical network forensics based on immune agent," in *Proceedings of the Third International Conference on Natural Computation (ICNC 2007)*, vol. 3, pp. 651–656, IEEE, Haikou, China, August 2007.

[8] M. Ibrahim, M. T. Abdullah, and A. Dehghantanha, "VoIP evidence model: a new forensic method for investigating VoIP malicious attacks," in *Proceedings of the 2012 International Conference on Cyber Security, Cyber Warfare and Digital Forensic (CyberSec)*, pp. 201–206, IEEE, Kuala Lumpur, Malaysia, June 2012.

[9] L. M. Chen, M. C. Chen, W. Liao, and Y. S. Sun, "A scalable network forensics mechanism for stealthy self-propagating attacks," *Computer Communications*, vol. 36, no. 13, pp. 1471–1484, 2013.

[10] I. L. Lin, Y. S. Yen, B. L. Wu, and H. Y. Wang, "VoIP network forensic analysis with digital evidence procedure," in *Proceedings of the 485 The 6th International Conference on Networked Computing and Advanced Information Management*, pp. 236–241, IEEE, Seoul, January 2010.

[11] W. Ren and H. Jin, "Distributed agent-based real time network intrusion forensics system architecture design," in *Proceedings of the 19th International Conference on Advanced Information Networking and Applications (AINA'05)*, vol. 1, pp. 177–182, IEEE, Taipei, Taiwan, March 2005.

[12] E. Jeong and B. Lee, "An IP traceback protocol using a compressed hash table, a sinkhole router and data mining based on network forensics against network attacks," *Future Generation Computer Systems*, vol. 33, pp. 42–52, 2014.

[13] Y. Zhu, "Attack pattern discovery in forensic investigation of network attacks," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 7, pp. 1349–1357, 2011.

[14] S. Perry, "Network forensics and the inside job," *Network Security*, vol. 2006, no. 12, p. 13, 2006.

[15] D. M. White, "The federal information security management act of 2002: a Potemkin village," *Fordham Law Review*, vol. 497, pp. 79–369, 2010.

[16] C. Wang, T. Feng, J. Kim, G. Wang, and W. Zhang, "Catching packet droppers and modifiers in wireless sensor networks," in *Proceedings of the 2009 6th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad 500 Hoc*

Communications and Networks, pp. 1–9, IEEE, Rome, Italy, June 2009.

[17] S. Zander, G. Armitage, and P. Branch, "A survey of covert channels and countermeasures in computer network protocols," *IEEE Communications Surveys & Tutorials*, vol. 9, no. 3, pp. 44–57, 2007.

[18] H. Kim, "Protection against packet fragmentation attacks at 6LoWPAN adaptation layer," in *Proceedings of the 2008 International 504 Conference on Convergence and Hybrid Information Technology*, pp. 796–801, IEEE, Wisła, Poland, October 2008.

[19] A. Sperotto, G. Schaffrath, R. Sadre, C. Morariu, A. Pras, and B. Stiller, "An overview of IP flow-based intrusion detection," *IEEE Communications Surveys & Tutorials*, vol. 12, no. 3, pp. 343–356, 2010.

[20] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection," *ACM Computing Surveys*, vol. 41, no. 3, pp. 1–58, 2009.

[21] P. Li, M. Salour, and X. Su, "A survey of internet worm detection and containment," *IEEE Communications Surveys & Tutorials*, vol. 10, no. 1, pp. 20–35, 2008.

[22] V. Igure and R. Williams, "Taxonomies of attacks and vulnerabilities in computer systems," *IEEE Communications Surveys & Tutorials*, vol. 10, no. 1, pp. 6–19, 2008.

[23] B. Yu and R. Wang, "Research of access control list in enterprise network management," Lecture Notes in Electrical Engineering, in *Informatics and Management Science VI*, pp. 121–129, Springer, Berlin, Germany, 2013.

[24] F. Akhtar, J. Li, M. Azeem et al., "Effective large for gestational age prediction using machine learning techniques with monitoring biochemical indicators," *The Journal of Supercomputing*, vol. 76, pp. 1–19, 2019.

[25] J. Li, D. Zhou, W. Qiu et al., "Application of weighted gene coexpression network analysis for data from paired design," *Scientific Reports*, vol. 8, pp. 622–628, 2018.

[26] F. Akhtar, J. Li, Y. Pei et al., "Diagnosis and prediction of large-for-gestational-age fetus using the stacked generalizationmethod," *Applied Sciences*, vol. 9, no. 20, p. 4317, 2019.

[27] A. Imran, J. Li, Y. Pei, J.-J. Yang, and Q. Wang, "Comparative analysis of vessel segmentation techniques in retinal images," *IEEE Access*, vol. 7, pp. 114862–114887, 2019.

[28] J. Li, L. Liu, J. Sun et al., "Comparison of different machine learning approaches to predict small for gestational age infants," *IEEE Transactions on Big Data*, vol. 6, no. 2, 2016.

[29] C. Liu, A. Singhal, and D. Wijesekera, "Using attack graphs in forensic examinations," in *Proceedings of the 2012 Seventh International 528 Conference on Availability, Reliability and Security*, pp. 596–603, IEEE, Prague, August 2012.

[30] A. Diamah, M. Mohammadian, and B. M. Balachandran, "Network security evaluation method via attack graphs and fuzzy cognitive maps," in *Intelligent Decision Technologies*, pp. 433–440, Springer, Berlin, Germany, 2012.

[31] A. B. Johnston, *SIP: Understanding the Session Initiation Protocol*, Artech House, Norwood, Massachusetts, 2015.

[32] D. Saha, "Extending logical attack graphs for efficient vulnerability analysis," in *Proceedings of the 15th ACM conference on Computer and communications security*, pp. 63–74, New York, NY, USA, October 2008.

[33] Y. Fen, Z. Hui, C. Shuang-shuang, and Y. Xin-chun, "A lightweight IP traceback scheme depending on TTL," *Procedia Engineering*, vol. 29, pp. 1932–1937, 2012.

[34] M. Albanese, S. Jajodia, A. Pugliese, and V. S. Subrahmanian, "Scalable analysis of attack scenarios," *Computer Security-ESORICS 2011. European Symposium on Research in Computer Security*, pp. 416–433, Springer, Berlin, Germany, 2011.

[35] S. Anwar, J. M. Zain, M. F. Zolkipli, Z. Inayat, A. N. Jabir, and J. B. Odili, "Response option for attacks detected by intrusion detection system," in *Proceedings of the 2015 4th International Conference on Software Engineering and Computer Systems (ICSECS)*, pp. 195–200, IEEE, Kuantan, Malaysia, August 2015.

[36] K. Shanmugasundaram, N. Memon, A. Savant, and H. Bronnimann, *ForNet: A Distributed Forensics Network. Computer Network Security*, V. Gorodetsky, L. Popyack, and V. Skormin, Eds., Springer Berlin Heidelberg, Berlin, Heidelberg, 2003pp. 1–16, Lecture Notes in Computer Science.

[37] M. Ponec, P. Giura, J. Wein, and H. Brönnimann, "New payload attribution methods for network forensic investigations," *ACM Transactions on Information and System Security*, vol. 13, no. 2, pp. 1–32, 2010.

[38] B. K. Sy, "Integrating intrusion alert information to aid forensic explanation: an analytical intrusion detection framework for distributive IDS," *Information Fusion*, vol. 10, no. 4, pp. 325–341, 2009.

[39] W. Wang and T. E. Daniels, "A graph based approach toward network forensics analysis," *ACM Transactions on Information and System Security*, vol. 12, no. 1, pp. 1–33, 2008.

[40] H. T. Dinh, C. Lee, D. Niyato, and P. Wang, "A survey of mobile cloud computing: architecture, applications, and approaches," *Wireless Communications and mobile Computing*, vol. 13, no. 18, pp. 1587–1611, 2013.

[41] S. Gupta, P. Kumar, and A. Abraham, "A profile based network intrusion detection and prevention system for securing cloud environment," *International Journal of Distributed Sensor Networks*, vol. 9, no. 3, Article ID 364575, 2013.