ISSN: 0067-2904

An Artificial Intelligence-based Proactive Network Forensic Framework

A. Abirami*, S. Palanikumar
Information Technology Department, Noorul Islam Centre for Higher Education, Tamilnadu, India

Received: 23/3/2022 Accepted: 20/11/2022 Published: 30/11/2023

Abstract

is at an all-time high in the modern period, and the majority of the population uses the Internet for all types of communication. It is great to be able to improvise like this. As a result of this trend, hackers have become increasingly focused on attacking the system/network in numerous ways. When a hacker commits a digital crime, it is examined in a reactive manner, which aids in the identification of the perpetrators. However, in the modern period, it is not expected to wait for an attack to occur. The user anticipates being able to predict a cyberattack before it causes damage to the system. This can be accomplished with the assistance of the proactive forensic framework presented in this study. The proposed system combines a reactive and proactive framework. The proactive part will use machine learning-based classification algorithms to forecast the attack. Once the assault has been predicted, the reactive element of the proposed framework is used to investigate who is attempting to initiate the attack. The suggested system further emphasizes integrity and confidentiality by proposing an encryption method that encrypts the proactive module's report before decrypting it in the reactive module. The suggested elliptical curve cryptography-based security model was compared to several existing security methods in this paper.A comparison of multiple machine learning-based categorization algorithms is also performed in order to determine which is the most suitable for the proposed Network Forensic Framework. Accuracy, recall, precision, and F1 value are the performance metrics used to evaluate the various machine learning-based algorithms. According to the analysis, the suggested Network Forensic Framework is best implemented using the Extreme Gradient Boosting (XGB) technique.

Keywords: Artificial Intelligence, Machine Learning, Network Forensic Framework, Cyber-crime, Cyber security.

1. Introduction

The concept of network forensic operations works by capturing, recording, and analyzing a network suspected of being used for cyber vulnerabilities and investigations in a fashion that works to detect errors in the network and existing IT infrastructure and to go back to the attacker source to prosecute cybercriminals[1]. Network forensics is a small part of digital forensics. Due to the rapid increase in internet connectivity, difficulties have been achieved by increasing the level of crime committed within networks, forcing law enforcement agencies and organizations to conduct special investigations. It is a process of capturing, recording, and analyzing events; identifying access to computer programs; and searching for evidence of such a thing. A skilled attacker can detect traffic flow on the forensic network, which requires expertise and resilience.The forensic network helps the investigator track the causes and effects of the attack with many challenges, such as time, speed, accuracy, storage location, performance, etc. The biggest challenge to network security is legal reliability; networks need

*Email: abi.lecturer@gmail.com

to be configured, maintained, and updated[2].
The purpose of network forensics is to provide sufficient evidence to allow perpetrators to be successfully prosecuted (e.g., effective hacking applications, fraud, data theft, software privacy, pornographic publication, etc.), taken from the movement of objects between computer devices, and to create evidence-based authentication records related to the planned motives for disrupting services or preventing data breaches [2-3].

After many years of research, network forensics looked at young science, in which many stories are still unknown. Network security protects systems, detects potential attack patterns, and monitors the network 24 hours a day, seven days a week.The forensic network can be started in real time as long as the necessary resources and infrastructure are available to manage traffic when it is analyzed[2-3].

The Network Investigation (NFI) process has two phases: online and offline. The online category includes retrieval, recording of network packets, and subsequent tests performed in the offline category, which are important data retrieval methods. Although a criminal investigation is essential, a framework will be followed. Therefore, the basic framework has three stages, which are preparation, investigation, and presentation [4].
Cryptography comes from the Greek word for secret writing. By encrypting and securely encrypting, cryptography ensures a third-party secure account that protects data from theft and user authentication and explicitly transforms it into an encrypted form, and vice versa. Only designated users can view, access, and process it. It has two types, which are symmetric key and asymmetric key [5] and [6].

Machine learning is an important topic to talk about with more machines, such as training and implementing their programs with minimal human intervention. The automatic learning method is also updated based on machine functions during the process. Furthermore, equipment comes with reliable data, and many techniques are used to build ML models to train equipment based on data.For example, in standard applications, the input is selected from the data. In machine learning, data and output are provided as the input and output systemsare installed. In addition, machine learning systems read and monitor network data to test official and distinct ideas. However, there are still two obstacles to be identified: creating false alarm numbers and finding the source of the attack [7-11].

The main effect of this paper is to propose a network forensic model. Six machine learning-based algorithms are utilized to analyze and evaluate the network-based cyberattack.Six techniques are:decision tree (DT), K-nearest neighbors (KNN), gradient boosting machine (GBM), random forest classifier (RFC), extreme gradient boosting (XGB), and artificial neural networks (ANN).

The remainder of the paper is organized as follows: Section I discusses the literature review. Section II described the proposed framework for the forensic network as well as the flow process of active forensic network investigations, with an emphasis on the transmission of encrypted messages from one user to another. Phase III contains the role of machine learning in the forensic network, the various machine learning algorithms used to test novel lab setups, comparative analysis, and the existing forensic network database. Finally, we conclude the paper on Phase VI and provide various indications for future research.

2. Literature Survey

Due to the rapid growth of technology, the intruder enters with new and advanced

techniques to create attacks. Therefore, it is important to develop a framework between methods, recording systems, saving and translating large amounts of real-time data, and communicating with management in accordance with the organization's policy. Network forensics has two types of investigations: reactive and proactive. The reactive investigation process begins after an incident has occurred to determine the cause of the attack [1]. The biggest problem with the forensic network framework is that the process of investigation begins after the incident; it is very difficult to find the perfect source of the attack for further transmission to legal entities. A method is used to detect live site attacks by performing this practice with minimal human intervention. Some of the available reactive network frameworksare illustrated in Figure 1.

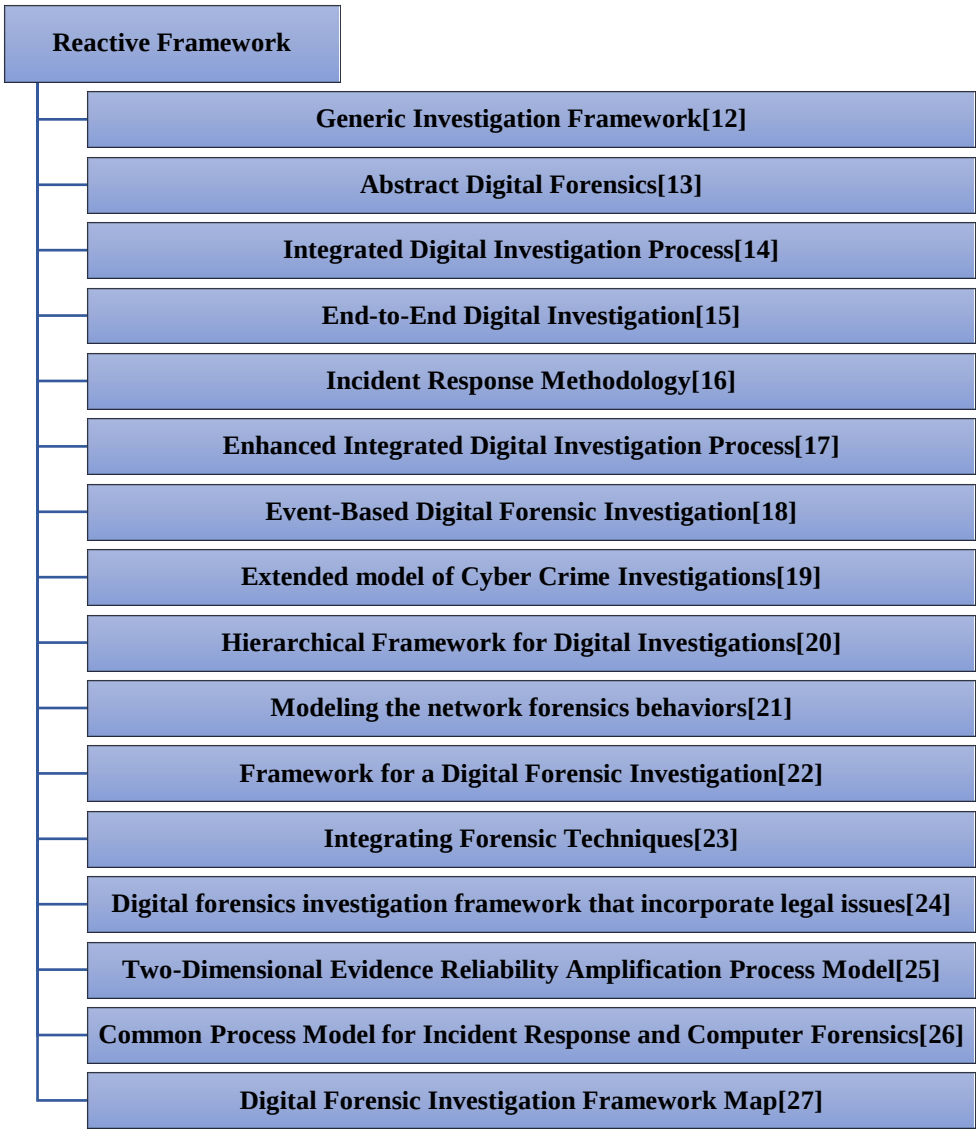

Figure 1: Reactive Framework[12-27]

Proactive investigation provides more reliability and accuracy in real time when an outbreak occurs. Early detection reduces the possibility of evidence distortion while increasing the processing of final heads, identifying patterns of attack proclivity, and keeping the evidence in real time [28].

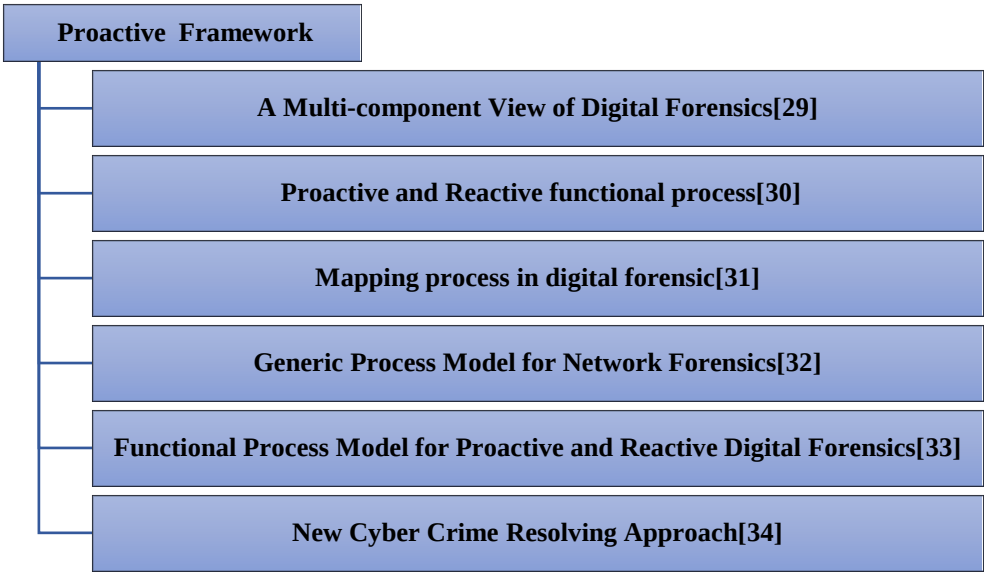

Figure 2: Proactive Framework[29-34]

Many proactive frameworks are being proposed by different authors, but their implementation is still pending. The proposed proactive frameworks are depicted in Figure 2. Grobleret al. [29]proposed a proactive network forensic framework, which is the multi-component view of the Digital Forensics Framework. The components presented in the multi-component view of the digital forensics framework are extremely challenging to apply in the various phases and obtain an efficient outcome.

Alharbi et al. [30] defined proactive and reactive functional processes.In comparison to the multi-component view of the digital forensic framework, the proactive and reactive functional processes are well defined, and components that can automate the output are designed, but the framework's problem has yet to be solved.

Rahayuet al. [31] represented a mapping process in digital forensics. The redundancy in each component of the other proposed systems is reduced throughout the mapping process in digital forensics.

Kaur et al. [32] projected the Network Forensic Process Model and Frameworkto get rid of unneeded phases.The processes are outlined in a more precise and detailed manner. The tools and techniques utilized in all steps of the generic process model for network forensics framework are not fully mentioned, making implementation extremely difficult.
Barik et al. [33] proposedrestricting functionality in the Functional Process Model for Proactive and Reactive Digital Forensics Framework because it does not mention all anti-forensic tactics; this is the proposed framework's flaw.

Mohammad Rasmiet al. [34] proposed a New Cyber Crime Resolving Approach,which is also a proactive framework. The problem with this system is that the phases are not well articulated, making implementation impossible.

Reactive frameworks are ineffective and inefficient because they only react after the damage has occurred. The proactive frameworks defined are not implemented, and as mentioned, there are many flaws in the proactive framework. Artificial intelligence plays a vital role in making the proactive framework successful [35]. Artificial intelligence can predict cyber-attacks by launching a cyber-attack model based on network packets collected [36].

An artificial intelligence-based framework is proposed that requires little user intervention and solves most problems by providing good training based on the model and the dataset [37].Various machine learning algorithms are incorporated into the framework to classify the network packets that are accumulated and captured during the live data transmission [38–39].

Kumar et al. [40] proposed anintrusion detection system using a decision tree algorithm. This intrusion detection system is trained to classify anomalies and misuse attacks. The intrusion detection systems available on the market are signature-based, which means they are not capable of finding unknown attacks. The decision tree-based intrusion detection system provides a better result compared to the signature-based detection system.

Wazirali et al. [41] developed anintrusion detection system based on a semi-supervised learning method using a k-nearest neighbor machine learning algorithm. This method optimized the outcome by using cross-validation and hyperparameter logic to yield a high accuracy rate with a minimum false-positive rate. The result of the proposed method provides a good precision rate of 0.95 and a recall rate of 0.92.

Verma et al. [42] proposed a network-based intrusion detection system with the help of the NSL KDD dataset. The classification algorithms utilized for his implementation are XGBoost and AdaBoost. Both the machine-learning-based algorithms yield better results as compared to the existing systems.

Farnaazet al. [43] deal with intrusion detection systems using the Random Forest (RF) classification algorithm.The system categorizes attacks into four types: DOS, U2R, probe, and R2L. The author followed 10 cross-validations in the Random Forest algorithm. The feature selection methods are applied to remove the duplicated data in the dataset and the irrelevant attributes in the dataset. The dataset utilized in this approach is NSL KDD, like many of the authors'. As per the results achieved by the proposed system, this classifier generated better accuracy, detection rate, false alarm rate, and Mathew's correlation coefficient.

Shenfield et al. [44]presented an intrusion detection system with an artificial neural network to detect malicious network packets. As per the results generated by the implemented system, the accuracy rate obtained by the system is quite good as compared to the other methods, and the false alarm rate is very low. This system has the capability to significantlyimprove the effectiveness of intrusion detection systems.

The proactive network forensic model needs a classifier to classify malicious and non-malicious network packets. As per the survey, classification using Decision Tree (DT), K Nearest Neighbors (KNN), Gradient Boosting Machine (GBM), Random Forest Classifier (RFC), Extreme Gradient Boosting (XGB), and Artificial Neural Network (ANN) is providing a better result.

3. Proposed Network Forensic Framework

According to the study, the reactive form of network forensic inquiry will begin only after a cyberattack has been launched and the system has been damaged. The proactive network framework is more effective since it anticipates a cyber assault by gathering live packets and denying harmful packets access to the network.The suggested system is a network forensic architecture that combines proactive and reactive capabilities. Machine learning-based categorization methods will be used to predict harmful packets in the proactive part. The proactive component is responsible for detecting cyberattacks using live network traffic and conducting basic investigations. The proactive forensic report is forwarded to the reactive forensic section for further examination into the cyber assault. To ensure the report's integrity and confidentiality before it is delivered to the reactive component, it must first be encrypted. The proposed algorithm provides confidentiality and integrity. Confidentiality is provided by elliptical curve cryptography, and integrity is provided by the hashing methods and the digital signature. In terms of confidentiality, it has been demonstrated in [45-46] that ECC with Koblitz encoding improves security. The MD5 hashing function can be used in the proposed algorithm since it is one of the fastest hashing methods. The security gaps in this MD5 will be covered by the other modules of the proposed algorithm.The encoded message from Koblitz's encoder module is encrypted using the ECC algorithm, then a hash is generated using MD5, and the message is digitally signed to make it more secure. The reverse operation is done on the receiver side. The comparison of the existing security model with the proposed model is given in Table1.

Table 1: Comparison of proposed encryption model with the existing system

Number of Users	Methodologies (Time in Seconds)				
	CL-PRE[47]	Certificateless encryption[48]	PRE[49]	AES[50]	Proposed ECC
10	1.494	1.594	1.534	0.004	0.00212
20	1.598	1.741	1.606	0.00425	0.00235
30	1.673	2.321	1.684	0.00476	0.00286
40	1.791	1.888	1.799	0.005	0.00302
50	1.907	1.952	1.866	0.00512	0.00328
60	1.954	2.193	1.923	0.0055	0.0035
70	1.994	2.286	2.034	0.00598	0.00398
80	2.092	2.694	2.129	0.00632	0.00427
90	2.401	2.827	2.388	0.00664	0.00463
100	2.495	2.887	2.545	0.00697	0.00499

While it is in the basic course of action, incomplete information will be available for investigation, challenges with data integrity will exist, and it will be difficult toprove complete evidence to law enforcement authorities. To overcome those challenges, we propose a new framework shown in Figure 3.

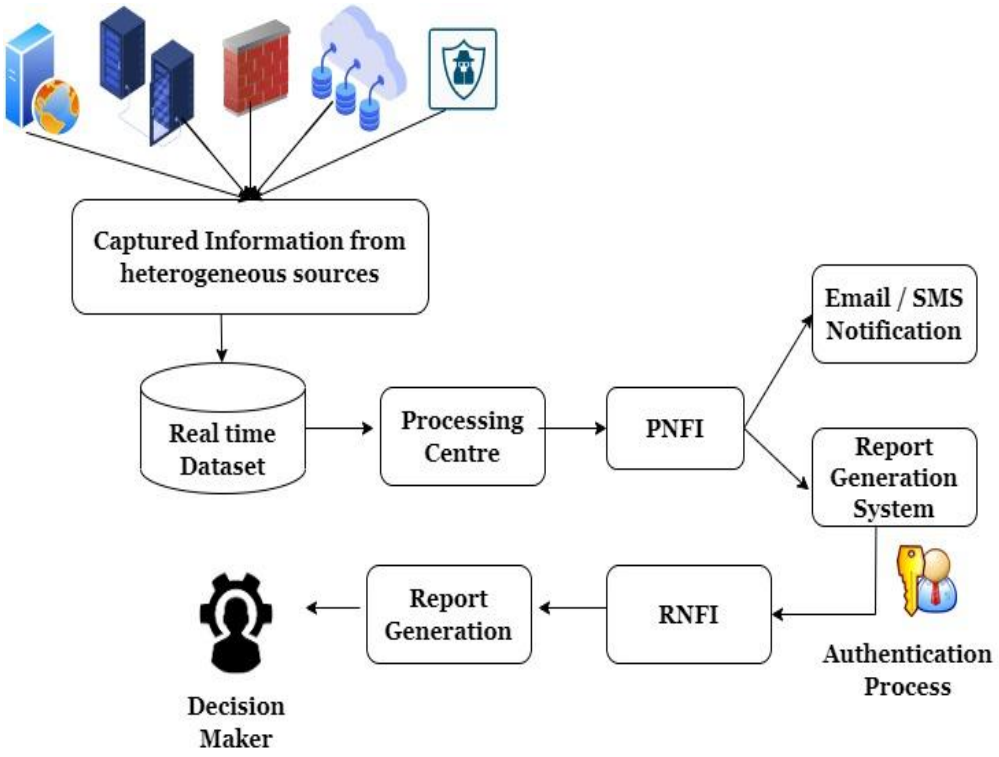

Figure 3: Proposed Network Forensic Model

Our proposed framework implements an effective and efficient research process. In our proposed framework, network traffic is collected from a variety of sources, reduced to a minimum by removing unwanted data, and useful features are extracted from the processing unit.

The feature selection is done as per the requirements of the attributes considered in the dataset. The newly released pattern is consistent with existing matching patterns and behavioral differences compared to an existing knowledge base. If any match is found, the immediate response is due to the intruder informing us of the activity. The selection of input is done by processing input data collected online, which is collected from various sources. Finally, standard practice aims to combine alerts into a single format.

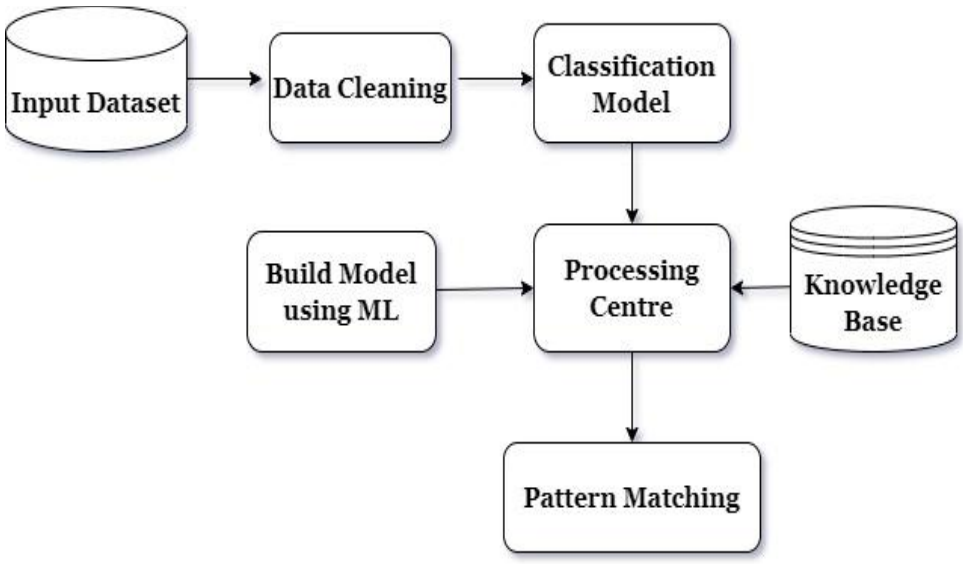

Figure 4: The Processing Unit

The processing unit employs different machine learning concepts, pattern classifications, and knowledge bases. The processing unit is shown in Figure 4.

For inclusion in the research unit of the forensic research network for analysis, process data is provided, and a warning-based system is proposed. If any suspicious activity is no longer active, the user is notified via the default email program, and an initial report is generated. The initial report from the operating procedure is considered a contributor to the process of investigating practical research. When looking at the investment process, we propose a framework based on organizational approval. After obtaining approval from the relevant authorities, the investigation process begins. We also promoted secure communications using encryption mechanisms with an additional layer of security based on two-factor authentication while transmitting confidential information to the proactive network forensic analysis unit as input. Information contained in a confidential report should not be available to all employees of the organization.It should not be disturbed, or else credibility may be lost and it may be difficult to create evidence. Figure 5 displays the proposed network forensic process model.

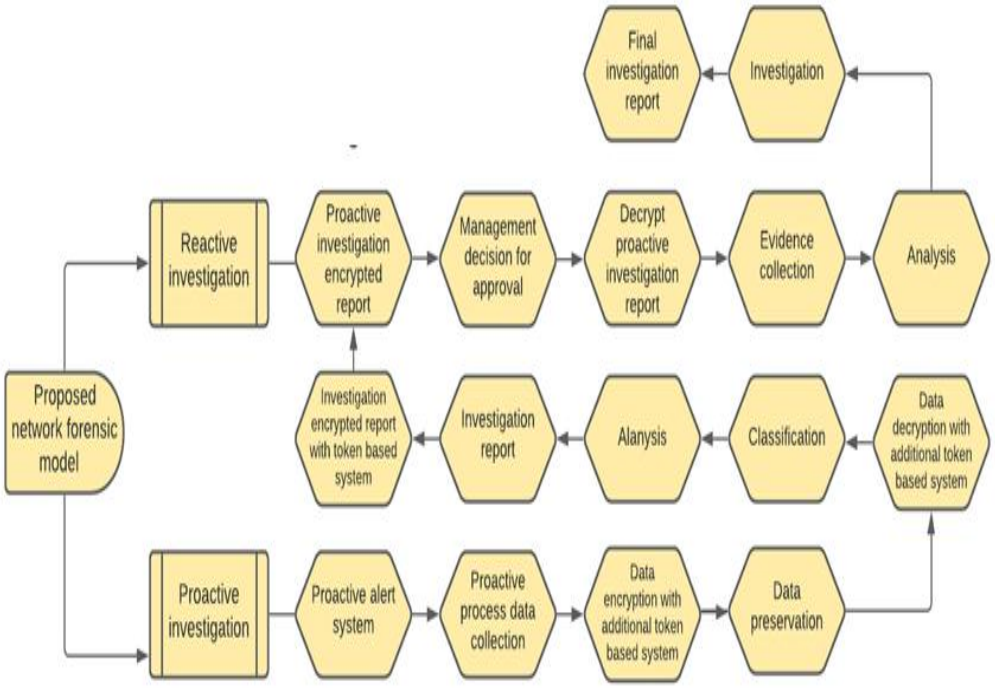

Figure 5: Proposed Network Forensic Investigation Modelwith Proactive and reactive phases.

According to the report, further investigation was conducted, and a final confidential report was made. Accordingly, a decision is made. If there is a discrepancy, an option is available to re-investigate as per requirements. When transferring a report from one user to another, it must be in a secure, encrypted format so that unauthorized users cannot access it.

4. Artificial Intelligence in Network Forensic

Machine learning is considered to be the backbone of ethical intelligence, which comes from the field of artificial intelligence. Therefore, the adoption of machine learning in digital forensics was given a prominent place. There are various methods and

algorithms used in machine learning for forensics analysis. There are seven steps to machine learning-based prediction, which are represented in Figure 6 [51].

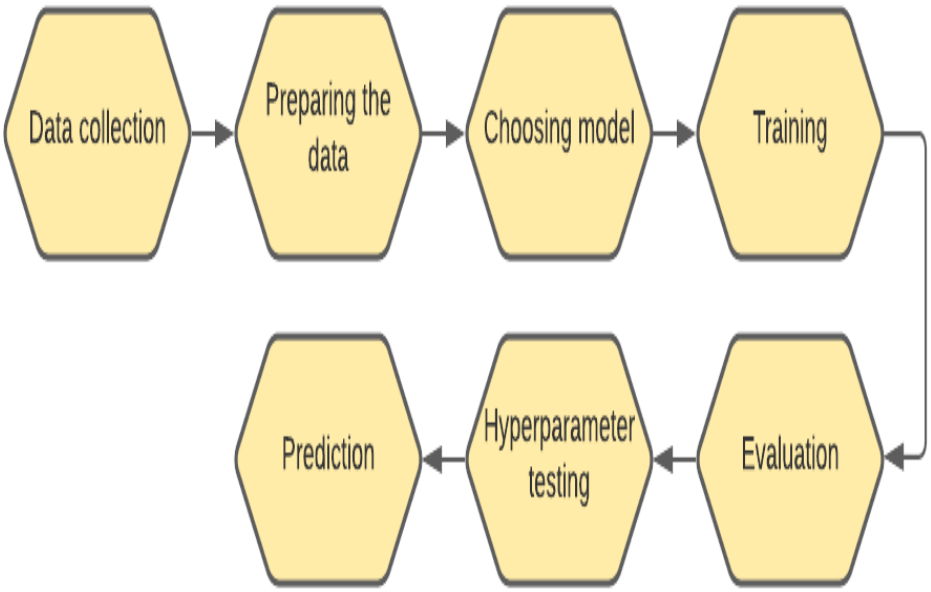

Figure 6: Machine Learning based Prediction System

4.1 Network ForensicInfrastructure

A lab setup for the purpose of the investigation is illustrated in Figure 7.

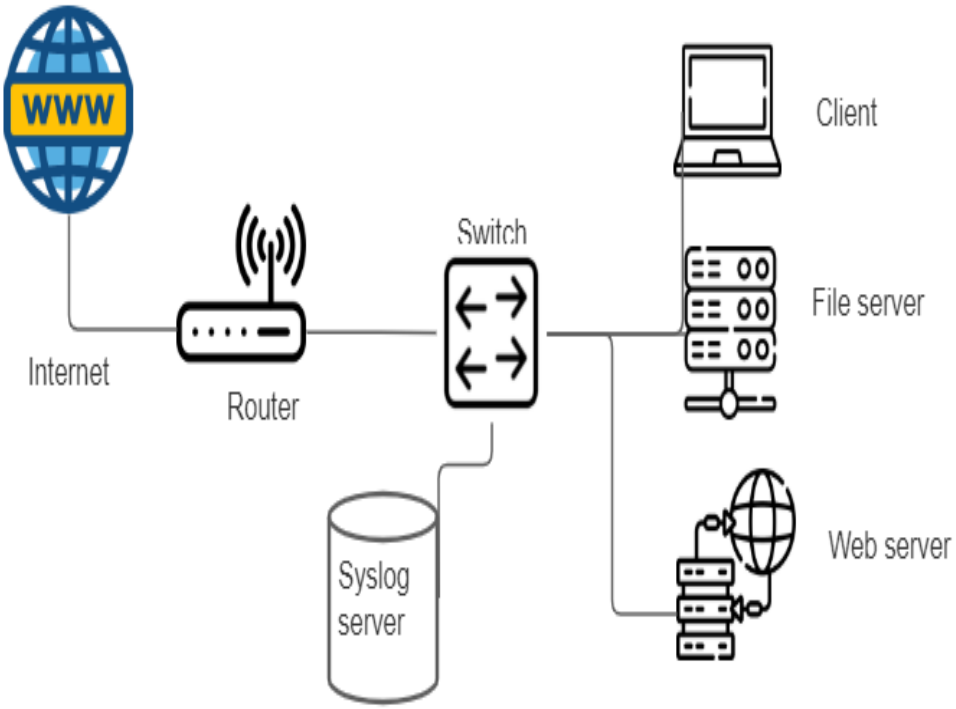

Figure 7: Lab setup for the proposed model

The proposed infrastructure comprises the following four elements: traffic log collection;network packet feature selection;machine learning-based algorithms for prediction;evaluation parameters; and detailed analysis.

4.2 Traffic Log Collection

The Graylog server is used for capturing organization logs. Graylog is open source, which means freely available software. The logs are collected based on package, flow-based, and session-based detection. The proposed log collection method excludes key parameters from the log parameters and associates them with the pre-defined sixteen categories stored in the MySQL database. In the preprocessing stage, alerts captured by the Syslog server are considered input. Next, the Data Processing section aims to integrate alerts into a single format as organized and labeled alerts.

4.3 Feature Selection

The information gain mechanism is implemented to classify the dataset in use. The logs are captured based on 6 features. The traffic is classified into five different sections, as shown in Table 2. The 16 features of the KDD dataset on which the analysis is basedare illustrated in Figure 8 in the correlation matrix. From the 41 attributes of the KDD dataset, the selected 16 attributes contribute more to the classification of the network forensic model, which is proposed.

Table 2: Dataset packet classification

Sr. No	Category	Description
1	normal	Normal traffic
2	dos	An attack to make network rescores unavailable to intended users.
3	probe	Action taken or an object used for the purpose of learning something about the state of the network.
4	r2l	To gain unauthorized access to a victim machine.
5	u2r	For illegally obtaining the root's privileges.

	protocol_type	service	flag	src_bytes	dst_bytes	count	srv_count	serror_rate	same_srv_rate	diff_srv_rate	srv_diff_host_rate	dst_host_count	dst_host_srv_count	dst_host_same_srv_rate	dst_host_diff_srv_rate	dst_host_same_src_port_rate	Attack Type
protocol_type	1																
service	-0.00579	1															
flag	-0.505	0.86603	1														
src_bytes	0.97213	-0.23101	-0.68546	1													
dst_bytes	-0.19775	-0.34365	-0.19775	-0.10807	1												
count	0.97484	0.21312	-0.30347	0.89846	-0.26812	1											
srv_count	0.99972	-0.00643	-0.50541	0.97197	-0.19841	0.9749	1										
serror_rate	-0.505	0.86603	1	-0.68546	-0.19775	-0.30347	-0.50541	1									
same_srv_rate	0.5042	-0.86466	-0.99841	0.68437	0.19743	0.30359	0.50536	-0.99841	1								
diff_srv_rate	-0.4998	0.85711	0.9897	-0.6784	-0.19571	-0.3034	-0.50116	0.9897	-0.99144	1							
srv_diff_host_rate	-0.16744	-0.29098	-0.16744	-0.10036	0.23379	-0.22899	-0.16538	-0.16744	0.16717	-0.16571	1						
dst_host_count	0.42678	0.74168	0.42678	0.24832	-0.4147	0.58042	0.4285	0.42678	-0.4261	0.42238	-0.26231	1					
dst_host_srv_count	0.57072	-0.75644	-0.93827	0.72999	0.18743	0.39248	0.57261	-0.93827	0.93833	-0.93054	0.11302	-0.2691	1				
dst_host_same_srv_rate	0.50479	-0.86567	-0.99959	0.68518	0.19766	0.30387	0.50559	-0.99959	0.99955	-0.99121	0.16737	-0.4266	0.93869	1			
dst_host_diff_srv_rate	-0.50267	0.86204	0.99539	-0.6823	-0.19684	-0.30322	-0.50334	0.99539	-0.99451	0.99022	-0.16667	0.42481	-0.93445	-0.99546	1		
dst_host_same_src_port_rate	0.96954	-0.08305	-0.55646	0.95826	-0.16498	0.92704	0.96837	-0.55646	0.55557	-0.55072	-0.06597	0.31205	0.60588	0.55623	-0.55389	1	
Attack Type	-0.00579	1	0.86603	-0.23101	-0.34365	0.21312	-0.00643	0.86603	-0.86466	0.85711	-0.29098	0.74168	-0.75644	-0.86567	0.86204	-0.08305	1

Figure 8: Correlation Matrix

4.4 Machine Learning Algorithms

A study is made to choose the best classification algorithm for the proposed network forensic model.The machine learning algorithms for the study are described in Table 2.

Table 2: Machine Learning Algorithms

S. No.	Algorithm	Description
1.	**Decision Tree[52]**	The decision solution (DT) belongs to the supervised learning algorithm. Analyzes data in stages and creates a flowchart. The root illustrates an attribute that meets the primary role in the category, and the leaf classifies the class.
2.	**The k-Nearest Neighbours[53]**	The closest k algorithm (KNN) is a supervised machine learning algorithm that can be used to work on both classification and regression problems. It is an algorithm for data classification that attempts to determine which data point group it belongs to by studying the surrounding data points.
3.	**Gradient Boosting Machine[54]**	Boosting is the process of conversion into a strong signal. Gradient boosting leads many models in a slower, additive and sequential way. GBM links prediction from various decision trees to make final predictions.
4.	**Random Forest Classifier[55]**	Random Forest is not only flexible but also an easy-to-use machine learning algorithm. It can be used both as a classification and a regression algorithm. Random forest (RF) is a composite separator used to improve accuracy. The random forest contains many decision trees, has a low classification error rate, and is linked to different classification algorithms.
5.	**Extreme Gradient Boosting[56]**	It employs a gradient descent algorithm to lower the loss when inserting new models. It is an application of gradient-boosted decision trees, which are created for velocity and performance.
6.	**Artificial Neural Network[57]**	The artificial neural network (neural network) is a computational paradigm that stimulates the activity of nerve cells in the human brain. ANNs play an important role in machine learning (ML) and support the broad field of artificial intelligence (AI). Multilayer Perceptron (MLP) is a feed-forward artificial neural network model that bases the input data sets on a set of relevant results.

4.5 Performance Evaluation using Machine Learning Algorithms

Recall, precision, accuracy, and the F1-value are calculated to measure the performance of the different machine learning algorithms [58].When both the actual and predicted classes of a data point are 1, it is said to be True_Positive.True_Negatives occur when a data point's actual and predicted classes are both 0.False_Positive occurs when the actual class of a data item is 0 and the predicted class is 1. False_Negative occurs when the actual class of a data point is 1 and the predicted class is 0. The formula to calculate the parameter metrics and its descriptions are given in Table 3.

Table 3: Performance Metrics and its Description

Performance Parameter	Description	Formula
Recall	It shows how many positives the Machine Learning model has returned.	$\frac{True\_Positive}{(True\_Positive + False\_Negative)}$
Accuracy	It represents the number of right guesses as a percentage of all predictions.	$\frac{(True\_Positive+True\_Negative)}{(True\_Positive+True\_Negative+False\_Positive+False\_Positive)}$
Precision	It shows how many correct documents the Machine Learning model has returned.	$\frac{True\_Positive}{(True\_Positive + False\_Positive)}$
F1 Value	It calculates the precision and recall harmonic mean.	$\frac{2*Precision*Recall}{(Precision+Recall)}$

75% of the dataset has been used to train the data, and the remaining 25%is used to test the data. The percentage of 75:25 is taken to get better accuracy as per the dataset considered. The dataset has 7992 records that are taken for analysis, out of which 4302 (53.83%) are categorized as normal and 3690 (46.17%) are categorized as attacks.

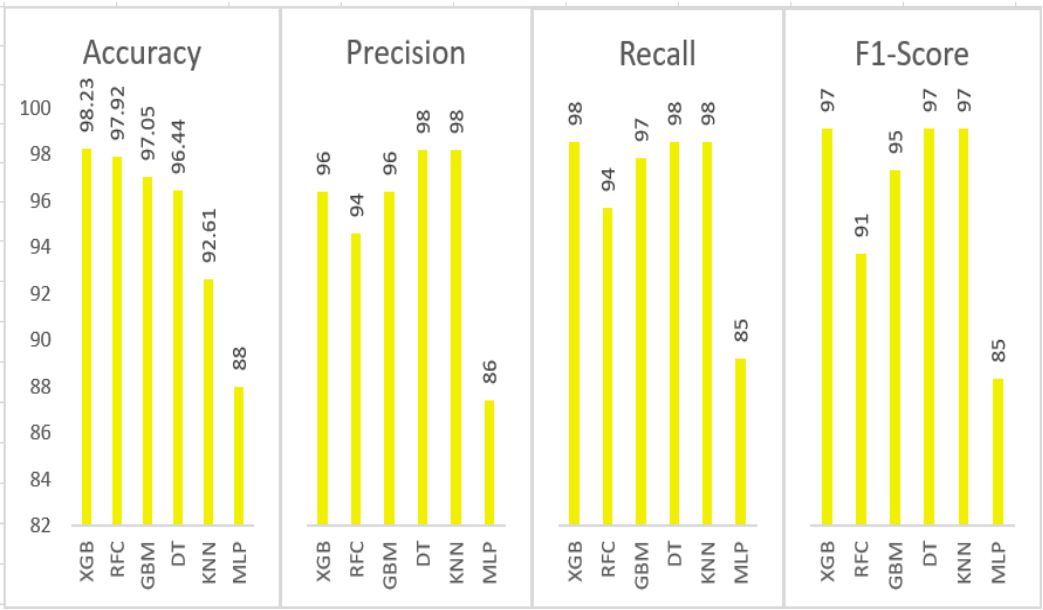

Figure 9: Comparative Analysis of Machine Learning Algorithms

Experiments show that the Extreme Gradient Boosting (XGB) Classifier is the most successful at distinguishing between suspicious and normal network traffic on a given network. The analysis concluded that XGB is the most accurate algorithm of all the tested algorithms, with an accuracy of 98.23%. The Random Forest classifier (RFC) ranks second with an accuracy of 97.92%. The Gradient Boosting Machine classifier (GBM) is the third-best classifier with an accuracy of 97.05%. The decision tree (DT) algorithm produced an

accuracy of 96.44% and ranked fourth. The K-Nearest Neighbors (KNN) achieved an accuracy of 92.61%. Lastly, the multi-layer perception classifier (MLP) showed results that were comparatively less accurate than the other six algorithms, with an accuracy of 88%. The detailed evaluation of the algorithms is shown in Table 3, and the comparison graph is represented in Figure 9

5. Conclusion

Because of current technology, cybercrime is on the rise, and all types of commerce, including education, are conducted over the Internet. As a result of this significant shift in the current period, hackers can now carry out a variety of attacks. Finding proof and predicting the hacker is pointless once the crime has been carried out and the system has been harmed. The anticipated system will be able to predict an assault before it occurs. Existing systems are forensic models that are reactive. A proactive forensic framework with a security layer is proposed in this study. A suggested ECC-based algorithm is used to make the security layer more secure, and a comparative analysis with the present system is used to show that the new security layer is stronger.The suggested security layer is made more secure by employing an ECC-based algorithm, and a comparative analysis with the present system is used to demonstrate that the proposed security layer is stronger. The system is made up of sections that are both reactive and proactive. With the use of machine learning-based classification algorithms, the initial assault packet can be predicted. A survey was conducted and a comparison analysis was performed between the Decision Tree (DT), K Nearest Neighbors (KNN), Gradient Boosting Machine (GBM), Random Forest Classifier (RFC), and Extreme Gradient Boosting Machine (EGBM) to determine the best machine learning-based algorithm (XGB). According to a comparison of performance parameter measures such as accuracy, precision, F1 Score, and recall, Extreme Gradient Boosting (XGB) produces a better outcome with an accuracy of 98.23%.

References

[1] A. Sivaprasad, "Secured proactive network forensic framework," in *2017 International Conference on Current Trends in Computer, Electrical, Electronics and Communication (CTCEEC)*, 2017, pp. 695–699.

[2] F. M. Ghabban, I. M. Alfadli, O. Ameerbakhsh, A. N. AbuAli, A. Al-Dhaqm, and M. A. Al-Khasawneh, "Comparative analysis of network forensic tools and network forensics processes," in *2021 2nd International Conference on Smart Computing and Electronic Enterprise (ICSCEE)*, 2021, pp. 78–83.

[3] A. Abirami and S. Palanikumar, "Proactive network packet classification using artificial intelligence," in *Artificial Intelligence for Cyber Security: Methods, Issues and Possible Horizons or Opportunities*, Cham: Springer International Publishing, 2021, pp. 169–187.

[4] D. Spiekermann, J. Keller, and T. Eggendorfer, "Network forensic investigation in OpenFlow networks with ForCon," *Digit. Investig.*, vol. 20, pp. S66–S74, 2017.

[5] G. L. Matthews and A. W. Murphy, "Cryptography," in *Mathematics in Cyber Research*, 1st Edition., Boca Raton: Chapman and Hall/CRC, 2022, pp. 53–96.

[6] J. Bermejo Higuera, J. R. Bermejo Higuer, J. A. Sicilia Montalvo, and R. González Crespo, "Introduction to Cryptography in Blockchain," in *Blockchain Technologies*, Singapore: Springer Nature Singapore, 2022, pp. 1–34.

[7] N. Pise, "Application of machine learning for intrusion detection system," *INFORMATION TECHNOLOGY IN INDUSTRY*, vol. 9, no. 1, pp. 314–323, 2021.

[8] M. Takaoğlu and Ç. Özer, "The effect of machine learning on intrusion detection systems," *Uluslar. yönet. bilişim sist. ve bilgi. bilim. derg.*, vol. 3, no. 1, pp. 11–22, 2019.

[9] A. Halimaa A. and K. Sundarakantham, "Machine learning based intrusion detection system," in *2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI)*, 2019,

pp. 916–920.

[10] A. Ahmim, M. A. Ferrag, L. Maglaras, M. Derdour, H. Janicke, and G. Drivas, "Taxonomy of supervised machine learning for intrusion detection systems," in *Strategic Innovative Marketing and Tourism*, Cham: Springer International Publishing, 2020, pp. 619–628.

[11] A. Pandey and N. Badal, "Machine learning based intrusion detection system: A survey," in *Computational Methodologies for Electrical and Electronics Engineers*, Hershey, PA: IGI Global, 2021, pp. 140–149.

[12] V. R. Kebande and I. Ray, "A generic digital forensic investigation framework for internet of things (IoT)," in *2016 IEEE 4th International Conference on Future Internet of Things and Cloud (FiCloud)*, 2016, pp. 356–362.

[13] G. Shrivastava, "Forensic computing models: Technical overview," in *Computer Science & Information Technology (CS & IT)*, 2012.

[14] Y. Yusoff, R. Ismail, and Z. Hassan, "Common phases of computer forensics investigation models," *Int. J. Comput. Sci. Inf. Technol.*, vol. 3, no. 3, pp. 17–31, 2011.

[15] P. Stephenson, "End-to-end digital forensics," *Comput. Fraud Secur.*, vol. 2002, no. 9, pp. 17–19, 2002.

[16] K. Mandia and C. Prosise, *Incident response & computer forensics, 2nd ed*. McGraw Hill Professional, 2003.

[17] V. Baryamureeba and F. Tushabe, "The enhanced digital investigation process model," *Dfrws.org*. [Online]. Available: https://dfrws.org/wp-content/uploads/2019/06/2004_USA_paper-the_enhanced_digital_investigation_process_model.pdf. [Accessed: 19-Aug-2023].

[18] B. Carrier and E. Spafford, "An event-based digital forensic investigation framework," *Dfrws.org*. [Online]. Available: https://dfrws.org/wp-content/uploads/2019/06/2004_USA_pres-an_event-based_digital_forensic_investigation_framework.pdf. [Accessed: 19-Aug-2023].

[19] S. Ó. Ciardhuáin, "An extended model of cybercrime investigations," *Utica.edu*. [Online]. Available: https://www.utica.edu/academic/institutes/ecii/publications/articles/A0B70121-FD6C-3DBA-0EA5C3E93CC575FA.pdf. [Accessed: 19-Aug-2023].

[20] N. L. Beebe and J. G. Clark, "A hierarchical, objectives-based framework for the digital investigations process," *Digit. Investig.*, vol. 2, no. 2, pp. 147–167, 2005.

[21] W. Ren and H. Jin, "Modeling the network forensics behaviors," in *Workshop of the 1st International Conference on Security and Privacy for Emerging Areas in Communication Networks, 2005*, 2005.

[22] M. Köhn, M. Olivier, and J. Eloff, "Framework for a digital forensic investigation," *Information Security for South Africa*, 2006.

[23] K. Kent, S. Chevalier, T. Grance, and H. Dang, "Guide to integrating forensic techniques into incident response," National Institute of Standards and Technology, Gaithersburg, MD, 2006.

[24] R. S. C. Ieong, "FORZA – Digital forensics investigation framework that incorporate legal issues," *Digit. Investig.*, vol. 3, pp. 29–36, 2006.

[25] M. Khatir, S. M. Hejazi, and E. Sneiders, "Two-dimensional evidence reliability amplification process model for digital forensics," in *2008 Third International Annual Workshop on Digital Forensics and Incident Analysis*, 2008, pp. 21–29.

[26] F. Freiling and B. Schwittay, "A common process model for Incident Response and Computer Forensics," *International Conference on IT-Incidents Management & IT-Forensics*, 2007.

[27] Y.-D. Shin, "New digital forensics investigation procedure model," in *2008 Fourth International Conference on Networked Computing and Advanced Information Management*, 2008, vol. 1, pp. 528–531.

[28] J. MbuguaChahira, J. KinanuKiruki, and P. KipronoKemei, "A proactive approach in network forensic investigation process," *Int. J. Comput. Appl. Technol. Res.*, vol. 5, no. 5, pp. 304–311, 2016.

[29] C. P. Grobler, C. P. Louwrens, and S. H. von Solms, "A multi-component view of digital

forensics," in *2010 International Conference on Availability, Reliability and Security*, 2010, pp. 647–652.

[30] S. Alharbi, J. Weber-Jahnke, and I. Traore, "The proactive and reactive digital forensics investigation process: A systematic literature review," in *Communications in Computer and Information Science*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 87–100.

[31] Selamat, Siti Rahayu & Robiah, Y. & Sahib, Shahrin, "Mapping Process of Digital Forensic Investigation Framework," *IJCSNS International Journal of Computer Science and Network Security*, 2008. .

[32] P. Kaur, A. Bijalwan, R. C. Joshi, and A. Awasthi, "Network forensic process model and framework: An alternative scenario," in *Advances in Intelligent Systems and Computing*, Singapore: Springer Singapore, 2018, pp. 493–502.

[33] K. Barik, A. Abirami, K. Konar, and S. Das, "Research perspective on digital forensic tools and investigation process," in *Illumination of Artificial Intelligence in Cybersecurity and Forensics*, Cham: Springer International Publishing, 2022, pp. 71–95.

[34] M. Rasmi, A. Jantan, and H. Almimi, "A new approach for resolving cyber crime in network forensics based on generic process model," 2013.

[35] A. Sivaprasad, N. Ghawalkar, S. Hodge, M. Sanghavi, and V. Shinde, "Machine Learning based Traffic Classification using Statistical Analysis," *Int. J. Recent Innov. Trends Comput. Commun.*, vol. 6, no. 3, pp. 187–191, 2018.

[36] W. Yang, M. N. Johnstone, S. Wang, N. M. Karie, N. M. B. Sahri, and J. J. Kang, "Network forensics in the era of artificial intelligence," in *Studies in Computational Intelligence*, Cham: Springer International Publishing, 2022, pp. 171–190.

[37] P. H. Rughani and Ph. D.,IFS, Gujarat Forensic Sciences University Gandhinagar, Gujarat - India, "Artificial intelligence based digital forensics framework," *Int. J. Adv. Res. Comput. Sci.*, vol. 8, no. 8, pp. 10–14, 2017.

[38] A. Abirami and S. Palanikumar, "Comparative analysis of machine learning algorithms based on the outcome of proactive intrusion detection system," in *Lecture Notes in Electrical Engineering*, Singapore: Springer Nature Singapore, 2022, pp. 177–184.

[39] M. Alrowaily, F. Alenezi, and Z. Lu, "Effectiveness of machine learning based intrusion detection systems," in *Security, Privacy, and Anonymity in Computation, Communication, and Storage*, Cham: Springer International Publishing, 2019, pp. 277–288.

[40] M. Kumar, M. Hanumanthappa, and T. V. S. Kumar, "Intrusion Detection System using decision tree algorithm," in *2012 IEEE 14th International Conference on Communication Technology*, 2012, pp. 629–634.

[41] R. Wazirali, "An improved intrusion detection system based on KNN hyperparameter tuning and cross-validation," *Arab. J. Sci. Eng.*, vol. 45, no. 12, pp. 10859–10873, 2020.

[42] P. Verma, S. Anwar, S. Khan, and S. B. Mane, "Network intrusion detection using clustering and gradient boosting," in *2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, 2018, pp. 1–7.

[43] N. Farnaaz and M. A. Jabbar, "Random forest modeling for network intrusion detection system," *Procedia Comput. Sci.*, vol. 89, pp. 213–217, 2016.

[44] A. Shenfield, D. Day, and A. Ayesh, "Intelligent intrusion detection systems using artificial neural networks," *ICT Express*, vol. 4, no. 2, pp. 95–99, 2018.

[45] K. E. Marssi, M. E. Marraki, and A. Kartit, "Koblitz's Improved Probability Mapping Method in the Elliptic Curve Cryptosystem: A comparative study and results," in *2019 Third International Conference on Intelligent Computing in Data Sciences (ICDS)*, 2019, pp. 1–10.

[46] H. Almajed, A. Almogren, and M. Alabdulkareem, "ITrust—A trustworthy and efficient mapping scheme in elliptic curve cryptography," *Sensors (Basel)*, vol. 20, no. 23, p. 6841, 2020.

[47] L. Xu, X. Wu, and X. Zhang, "CL-PRE: A certificateless proxy re-encryption scheme for secure data sharing with public cloud," in *Proceedings of the 7th ACM Symposium on Information, Computer and Communications Security*, 2012.

[48] S.-H. Seo, M. Nabeel, X. Ding, and E. Bertino, "An efficient certificateless encryption for secure

data sharing in public clouds," *IEEE Trans. Knowl. Data Eng.*, vol. 26, no. 9, pp. 2107–2119, 2014.

[49] A. N. Khan, M. L. M. Kiah, S. A. Madani, M. Ali, A. ur R. Khan, and S. Shamshirband, "Incremental proxy re-encryption scheme for mobile cloud computing environment," *J. Supercomput.*, vol. 68, no. 2, pp. 624–651, 2014.

[50] S. K. Rao, D. Mahto, and D. A. Khan, "A survey on advanced encryption standard," *Int. J. Sci. Res. (Raipur)*, vol. 6, no. 1, pp. 711–724, 2017.

[51] J. Mena, *Machine learning forensics for law enforcement, security, and intelligence*. London, England: Taylor & Francis, 2011.

[52] R. M. Panda and B. S. Daya Sagar, "Decision Tree," in *Encyclopedia of Mathematical Geosciences*, Cham: Springer International Publishing, 2022, pp. 1–7.

[53] K. Taunk, S. De, S. Verma, and A. Swetapadma, "A brief review of nearest neighbor algorithm for learning and classification," in *2019 International Conference on Intelligent Computing and Control Systems (ICCS)*, 2019, pp. 1255–1260.

[54] V. K. Ayyadevara, "Gradient Boosting Machine," in *Pro Machine Learning Algorithms*, Berkeley, CA: Apress, 2018, pp. 117–134.

[55] D. P. Mohandoss, Y. Shi, and K. Suo, "Outlier prediction using random forest classifier," in *2021 IEEE 11th Annual Computing and Communication Workshop and Conference (CCWC)*, 2021, pp. 0027–0033.

[56] G. M. Hassan, A. Gumaei, A. Alanazi, and S. M. Alzanin, "A network intrusion detection approach using extreme gradient boosting with max-depth optimization and feature selection," *Int. J. Interact. Mob. Technol.*, vol. 17, no. 15, pp. 120–134, 2023.

[57] M. S. K. Katakam, K. Devineni, P. Kanagala, and D. R. N. Raghavendra sai, "Analysis of artificial neural networks based Intrusion detection system," *Int. J. Adv. Sci. Technol.*, vol. 29, no. 5s, pp. 928–935, 2020.

[58] D. M. W. Powers, "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation," *arXiv [cs.LG]*, 2020.