Indonesian Journal of Electrical Engineering and Computer Science
Vol. 38, No. 3, June 2025, pp. 1642~1651
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v38.i3.pp1642-1651

SVM algorithm-based anomaly detection in network logs and firewall logs

John Benito Jesudasan Peter[1], Nitin Rakesh[2], Puttaswamy Rekha[3], Tammineni Sreelatha[4], Velusamy Sujatha[5], Surulivelu Muthumarilakshmi[6], Shanmugam Sujatha[7]
[1]Cyber, Data Science and Engineering, Capital One Services, LLC, Richmond, USA
[2]Symbiosis Institute of Technology (SIT), Symbiosis International (Deemed University), Nagpur, India
[3]Department of Electronics and Communication Engineering, BNM Institute of Technology, Bangalore, India
[4]Department of Electronics and Communication Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, India
[5]Department of Electronics and Communication Engineering, Shree Sathyam College of Engineering and Technology, Salem, India
[6]Department of Computer Science and Engineering, Chennai Institute of Technology, Kundrathur, India
[7]Department of Biomedical Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Thandalam, India

Article Info

Article history:

Received Feb 24, 2024
Revised Dec 21, 2024
Accepted Feb 28, 2025

Keywords:

Cyberattacks
Firewall logs
Genetic algorithm
Network logs
Support vector machine algorithm

ABSTRACT

The purpose of many advanced forms of cyberattack is to deceive the monitors, and as a result, these attacks often involve several kinds, levels, and stages. Existing anomaly detection systems often examine logs or traffic for indications of attacks, ignoring any additional analysis regarding attack procedures. This is done to save time. For example, traffic detection technologies can only identify the attack flows in a general sense. Still, they cannot reconstruct the attack event process or expose the present condition of the network node. In addition, the logs kept by the firewall are significant sources of evidence; nevertheless, they are still challenging to decipher. This paper introduces support vector machine algorithm-based Anomaly detection (SVMA) in network logs and firewall logs to provide robust security against cyberattacks. This mechanism consists of three modules: preprocessing, feature selection and anomaly detection. The genetic algorithm (GA) selects the better feature from the input. Finally, the support vector machine (SVM) isolates an anomaly powerfully. The investigational outcomes illustrate that the SVMA minimizes the required time to select the features and enhances the detection accuracy.

Corresponding Author:

Nitin Rakesh
Symbiosis Institute of Technology (SIT), Symbiosis International (Deemed University)
Nagpur, Maharashtra, India
Email: nitin.rakesh@gmail.com

1. INTRODUCTION

Cyberattacks almost always leave traces on the devices that make up the network. The attack vector of an attacker often involves jumping via numerous routers or servers before the attacker uploads malicious code (for example, an XSS script), implants malware (for example, a botnet), and submits an unapproved patch that contains dangerous payloads [1]. In most cases, the traces of a cyberattack may be found scattered over the activity logs of a number of different computers belonging to various victims [2]. Conversely, since the logs are scattered over a variety of unconnected sources, it is still necessary to have human intervention in order to piece together the contextual information of each bad footprint. Therefore, relying on the logs for anomaly identification will not provide the desired results [3]. This kind of information is an example of what

is known as complementary evidence. However, using merely the data from network traffic is inadequate to correctly identify attack patterns and provide a comprehensive understanding of assaults.

In addition, firewalls provide the function of control gates that network packets must go through to pass through the firewall. System administrators are responsible for configuring firewalls in a manner that is compliant with the standards of their business [4]. Firewall is an essential part of today's networks because of the critical role they act in safeguarding the network against any threats, whether they originate from inside the network or from the outside. Because they depend on the company implementing the firewall, these regulations are subject to change. In addition, because of technological improvements and the ever-shifting behavior of the surrounding environment, revising these guidelines is a labor-intensive and ongoing process [5]. Actions are done, such as Grant, Omit, Refuse, or Readjust -both. If choose the wrong step to take while handling a session, it might lead to security vulnerabilities, which would then make it possible for unwanted events to take place, such as the shutdown of devices, the loss of service, indirect loss of profit, or the leaking of sensitive materials [6].

On the other hand, the management of and the development of the rules that decide the proper actions have grown intricate and prone to mistakes [7]. Machine learning (ML) techniques have enormous promise in various disciplines, including cybersecurity. There is a growing understanding, in the realm of cybersecurity, of the positive usage of ML algorithms. These records keep a wealth of information that is hidden from view. This information, coupled with the patterns in the properties of the network traffic, may be found and worked via ML to develop models that help identify network risks [8]. Training these systems enables them to produce warnings when dangers are discovered, identify new forms of malware and secure enterprises' personal details. In addition, ML approaches vital tools for many years owing to their capacity to improve decision-making without the need for direct human involvement, which paves the way for more effective analysis on a broader scale [9]. Hierarchical anomaly-based detection of dissiminated domain name system (DNS) attacks (HADA) inside corporate networks. This system recognizes strange patterns in DNS traffic. This system makes it possible to identify malicious actions; for example, denial of service (DoS) attacks. This solution uses numerous layers of analysis to distinguish between benign as well as malicious DNS traffic, which ultimately results in an improved security posture for the network. However, this mechanism can't detect the firewall logs. Furthermore, it detects the DoS attack, and it is not able to detect the anomaly [10].

SVM with clustering to differentiate among normal as well as abnormal traffic patterns. Clustering assists in organizing data points that are quite similar to one another. The model improves the accuracy of anomaly detection (AD) in network traffic due to the integration of different approaches, which in turn makes the model more efficient and dependable [11]. Isolation Forest and two deep autoencoders are used in the proposed unsupervised learning for log message anomaly identification that uses the model. The autoencoder networks are used first for training and feature extraction and subsequently for AD, while the isolation forest is utilized for positive sample prediction [12]. The GA is used for feature selection, while the Naive Bayes classification method is used for data analysis. To optimize the feature subset for AD, a GA is used to choose the characteristics from the dataset that are most relevant. The data that was picked for analysis is then sent into a Naive Bayes classifier, which looks for irregularities in the data that was collected via fog computing. This hybrid approach improves the accuracy of AD while simultaneously increasing its efficiency [13]. One-class SVM algorithm is developed for one-class classification, which means that it is trained only on the data belonging to the normal class and recognizes anomalies. In this particular scenario, it refers to the logs generated by the Juniper router devices. The scope of this research is to enhance the security and performance monitoring capabilities of Juniper router devices by reliably identifying and flagging aberrant patterns or occurrences inside the router logs [14].

The J48 technique for ML is based on decision trees, and the updated version contains Kendall's correlation coefficient, which evaluates the degree to that two variables are connected with one another. The purpose of this research is to investigate if adding Kendall's correlation coefficient into the J48 algorithm for classification tasks improves its accuracy and reliability [15]. In internet of things (IoT) contexts, data of poor quality may include information that is noisy, fragmentary, or erroneous. The purpose of this mechanism is to analyze how ML models can successfully deal with such difficult data situations. This mechanism represents an AD model which can effectively distinguish unexpected patterns or occurrences within IoT data despite the low quality of the data [16].

A technique that uses distance to identify unusual characteristics inside log data. Log files include records of a variety of actions and occurrences, and it is essential for the security of the system and for troubleshooting purposes to identify any irregularities within these data. The distances between characteristics in log entries are being measured as part of the suggested technique [17]. The approach may discover features that considerably differ from the predicted patterns by computing these distances and comparing them. This method allows the identification of aberrant activity inside log files, giving system administrators and security experts a helpful to discover possible security concerns or behaviors that are not

normal [18]. An AD system known as SmartRadar was developed primarily for use in scenarios that include remote work. SmartRadar uses the SVM [19] for AD [20]. The system is an intelligent piece of software developed to monitor and assess the goings-on in remote working environments. It is able to differentiate between usual patterns of behavior when working remotely and aberrant or suspicious activity thanks to the use of SVM [21]. The major objective of SmartRadar is to improve safety and productivity by recognizing and notifying users of potentially dangerous or unlawful behaviors in remote working settings [22]. An artificial neural network to rapidly sift through Spark logs data and operating system observation to precisely notice and categorize anomalous behaviors by the Spark resilient features [23]. A graph-based approach to unsupervised log anomaly identification has been developed and named Logs2Graphs. This approach first transforms event logs into graphs that are attributed, directed, and weighted and then uses graph neural networks to conduct AD at the graph level [24].Using data mining and ML, distributed firewalls may implement AD of rules, which looks for unusual behavior. The solution that has been suggested is used on large logs that come from distributed firewalls [25]. Cloud computing [26] allows the proposed seizure prediction mechanism is more accessible and scalable [27]. Cloud-based decision-support system [28] applying K-nearest neighbors' algorithm with IoT structure system integrates IoT devices [29] to collect and transmit the data [30]. Fuzzy logic with bayesian decision algorithm isolates the DoS packets efficiently. The fuzzy logic system is applied to validate the data packets [31].

2. PROPOSED METHOD

This work mostly intends to execute AD during incorporating the network with firewall logs. Particularly, the network logs and firewall logs are incorporated by association rules. This approach can efficiently enhance the AD performance and rebuild the network attack procedure that changes to comprehend an entire view of the network surroundings. The environmentalism of the network is highly complicated since there are so many different kinds of attacks. For example, botnets have to first transmit control instructions to each command and control (C&C) Server before sending them to the controlled host. In contrast, worms have to earliest upload malicious code to the target host before infecting additional computers via the target host. As a result, the network logs and traffic flows have the potential to play extremely significant roles in the identification of cyberattacks.

Initially, we retrieve the traffic flows via port mirroring and then we use TCPDUMP to extract relevant traffic parameters (for instance, the port of the sender and the receiver, the protocol number, the IP of the sender and receiver, the packet size, and the transmit time). This grants to collect data for detecting anomalies. In addition, we get log information from the inner routers, firewalls, switches, and servers of the gateway. Next, we use association rules in an effort to reverse engineer the mapping relations that exist between the logs and the traffic. In conclusion, the recovered connections have the potential to be put to use in order to produce the time stamps of log entries and rebuild the development of attacks. This mechanism contains 3 processes: preprocessing, feature extraction, and AD. Initially, the log collectors primarily gather the traffic and logs conceived as the input data. This mechanism uses a GA algorithm for the filtering process that filters the unrelated data. The feature selection contains five components: correlation of traffic that examines the malicious traffic packets, correlation of temporal to get the time features of malicious, correlation of combination to form the strong significance of malicious traffic, TCP flag to record the forwarding and replying traffic data, and attack re-enactment. Next, we apply an SVM to detect the anomaly effectively.

Preprocessing: the preprocessing procedure is employed to the data earlier than any analysis to build the file ready for utilization based on training and testing. The main objective of this process is data loading, filtering, manipulating, and exchanging the data. The log dataset consists of numerous data entries in that the deny action presents 2.68% and Allow denotes 88.23% of the data. This indicates that the data lost from inequality when equated to other class labels.

2.1. GA-based feature selection

Frequently, several cyberattacks, for example, phishing emails and botnets, necessitate repeatedly forwarding the commands through programs. In most cases, the wrapper technique is superior than a filter approach because the process of feature selection is optimized. The traditional wrapper method is accomplished in two stages: searching for a subset of characteristics, and evaluating the features that were chosen. It continues to cycle through the first stage and second stage until either the target level of learning performance is reached or specified halting conditions are met. However, due to the fact that the wrapper technique is excessively costly, it needs additional time and raises computational complexity.

GA is an optimal search algorithm that is able to be expeditiously used in feature selection. A GA contains three functions like reproduction, crossover, and mutation. Reproduction picks out the strong

strings; crossover aggregates good strings that develop better offspring; and mutation alters a string locally to provide a better string. First, the GA generates a population of possible solutions by seeding it with random information. A bit may be used to represent each chromosomal gene in a dataset, depending on the characteristics of the dataset as a whole. After that, individuals are graded according to the fitness function they possess.

2.2. SVM-based anomaly detection

Identifying an action into normal and anomaly categories based on the SVM algorithm. SVM technology is used to categorize dissimilar types of data coming from a wide variety of fields. These have been used for challenges involving the classification of two classes of data, and they are relevant to linear as well as non-linear data classification endeavors. The SVM [31] constructs a hyperplane that divides data into various classes. When forecasting the margins, a non-linear classifier will utilize a selection of kernel functions. Exploiting the margins among hyperplanes is the primary goal of the kernel functions, all of which are described above. The SVM has seen extensive use in image processing and pattern identification. Figure 1 explains SVM algorithm-based AD.

The radial basis function (RBF) kernel is the key component of the suggested system for carrying out the execution of the SVM. The kernel operation constitutes an advantage that translates input data to a high dimensional space to best segregate the provided data into the attack classes to which they belong. Thus, the kernel RBF is a useful method for distinguishing between data groups that have shared complicated borders. The SVM-based detection module is given training using the statistics of both the normal data and the suspicious data. The average of one-second intervals is used to generate the statistics pertaining to each characteristic. The data are then normalized and used as input into the system after being separated into a training segment and a test section. Because the detection module is trained to understand the behavior of the network under both usual and attack settings, the detection module labels as intrusion any large deviations from the typical behavior of the network. For this AD, the SVM classifier algorithm efficiently distinguishes the anomalies.

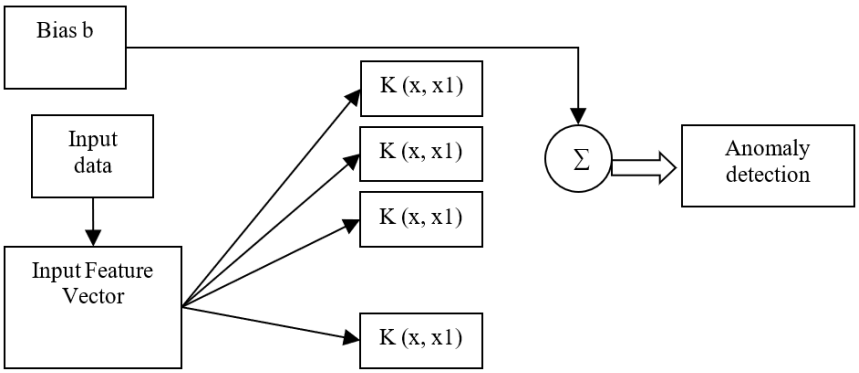

Figure 1. SVM algorithm-based AD

3. EXPERIMENTAL RESULTS

This mechanism examines four types of attacks, such as botnets of HTTP, P2P, XSS, and phishing cyberattacks, which are interposed into the normal effects. Both the network and the firewall log data are gathered from university servers. To authorize the function of incorporating traffics with logs for identifying anomaly, we accomplish comparison experimentations by leveraging the traffic data. The network traffic nor logs separately attain attractive results in identifying cyberattacks. In this mechanism, we analyze the classification performance utilized by calculating their precision, categorization accuracy, recall, required time, as well as F-measure.

The accuracy (ACU) displays what percentage of the total number of right forecastings have been made, which is the value of cases that have been correctly categorized. The accuracy of the classification may be determined by using (1) and dividing the total number of forecastings through the total number of forecasts.

$$ACU = \frac{T_{Pos}+T_{Neg}}{T_{Pos}+F_{Neg}+F_{Pos}+T_{Neg}} \quad (1)$$

Concerning the confusion matrix contains four variables: true positive (T_{Pos}), true negative (T_{Neg}), false positive (F_{Pos}), as well as false negative (F_{Neg}). For instance, T_{Pos} proposes that the existence of dangers was exactly distinguished. Rather than, T_{Neg} represents that the deficiency of threats was accurately forecasted, that is a positive exploitation. The information that the method wrongly forecasted that there was no danger F_{Neg} and failed to distinguish the existence of the attack F_{Pos} proposes that the mechanism did not sufficiently notice the existance of the attack.

$$Prc = \frac{T_{Pos}}{T_{Pos}+F_{Pos}} \tag{2}$$

Additionally, the accuracy, Recall (Rca), and F-measure (F-M) are all derived using these four variables. When (2) is used to determine precision (Prec), the number of 'Anomalous' occurrences that are really positive is quantified as a percentage of the total number of instances that are projected to be 'Anomalous'.

$$Rca = \frac{T_{Pos}}{T_{Pos}+F_{Neg}} \tag{3}$$

Regarding the recall, it likewise forecasts the positive class forecasting, but, as shown in (3), it is computed over the whole anomalous count occurrences included inside the dataset itself, regardless of whether or not such instances were successfully forecasted. To conclude, the F-measure is a single count that aggregates the values of the precision as well as recall into one number. Figures 2 to 4 displays the precision, recall and F-measure (F-M) of SVMA and HADA mechanisms. It is computed by applying formula (4).

$$F - M = \frac{Prc \times Rca}{Prc + Rca} \tag{4}$$

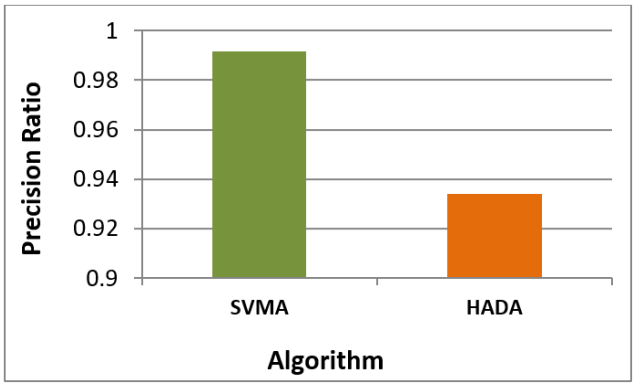

Figure 2. Precision ratio of SVMA and HADA

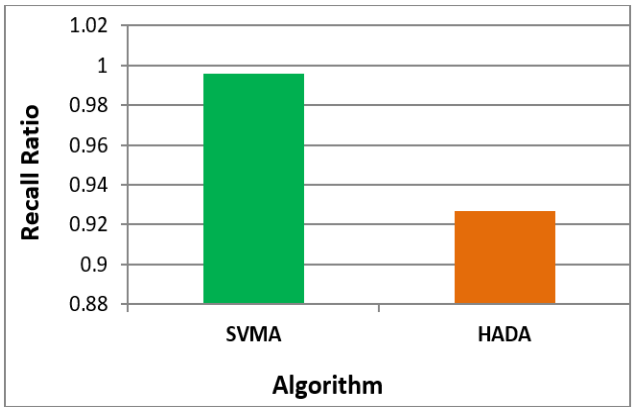

Figure 3. Recall ratio of SVMA and HADA

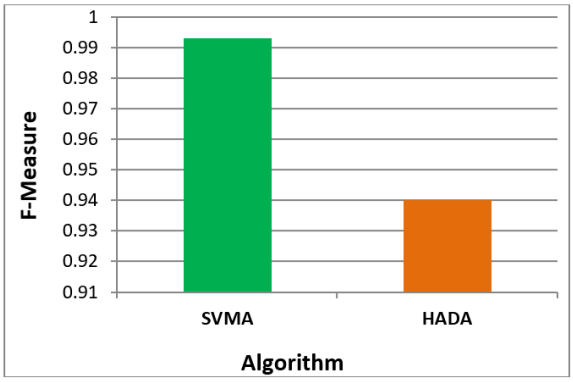

Figure 4. F-Measure of SVMA and HADA

From these figures, the proposed SVMA mechanism has a 0.9915 and HADA mechanism has a 0.9342 precision ratio. The ratio of recall value for SVMA mechanism presents 0.996, and HADA mechanism has a 0.927. Furthermore, the the proposed SVMA mechanism presents the ratio value of F-measure is 0.933 and the HADA mechanism presents the F-measure ratio value is 0.94. The SVMA mechanism compared to the HADA mechanisms, the SVMA mechanism provide better results in the network. As the figures expose, the SVMA reached the greatest operation above 0.99 than the existing HADA mechanism. Figure 5 explains the proposed SVMA mechanism precision percentage of selected features with whole features.

From Figure 5, the whole feature precision rate is low level, but the selected features precision rate is high due to the SVMA mechanism using the GA algorithm to select the feature as well. The selected features precision percentage is 98.5 and the whole features precision percentage is 93.5. The whole featues creates additional delay and overhead. Figure 6 explains the Required time for selected features with whole features.

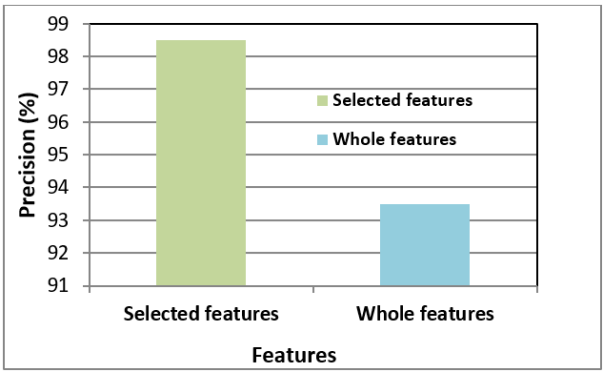

Figure 5. Precision percentage of selected features with whole features

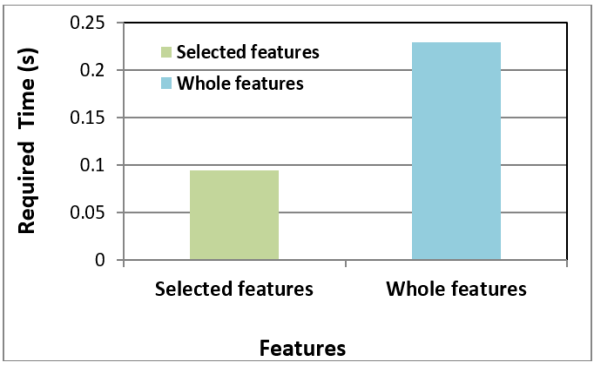

Figure 6. Required time for selected features with whole features

From Figure 6, the selected features take a lesser required time which represents the best performance accuracy. However, the whole feature takes more required time since itself have noise and unwanted and repeated data. The SVMA mechanism applies a GA to pick out the better features. Figure 7 displays an AD accuracy of SVMA and HADA mechanisms based on Experiments.

Figure 7 clearly says that the proposed SVMA mechanism raises the AD accuracy than the HADA mechanism. The SVMA mechanism utilizes the SVM algorithm to separate the anomaly in network and firewall logs against cyberattacks efficiently. Furthermore, the SVMA mechanism applies the GA to select the best features; hence, it raises the AD accuracy.

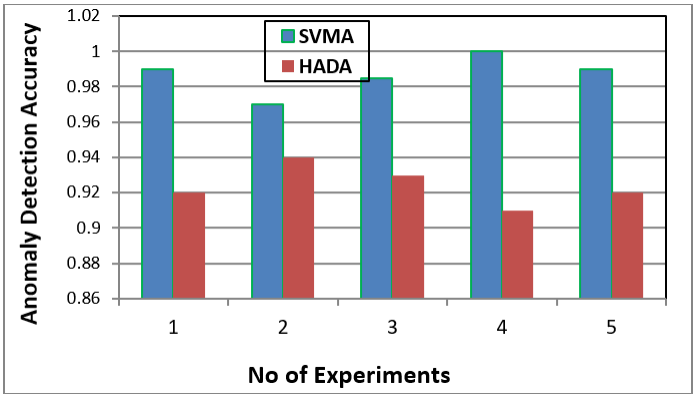

Figure 7. AD accuracy based on number of experiments

4. CONCLUSION

Specifically, a Firewall is a significant component of network security, and it can defend the data from external and internal attacks. This article presents Support Vector Machine algorithm-based AD in Network Logs and Firewall logs. It introduced to incorporate traffic with network logs for noticing cyberattacks. The preprocessing concept to filter the unwanted data from the input. This mechanism uses a GA algorithm for selecting the feature well. GA algorithm takes minimum time for feature selection, enhancing the feature selection precision. The SVM algorithm gives the input of the data; then, it categorizes the intrusion efficiently. The experimental results prove that the proposed SVMA mechanism can efficiently detect anomaly and increases the precision, accuracy, and recall ratio. The SVMA mechanism presents 99% efficiency than a conventional mechanism. This mechanism utilizes in army and fire detection applications. In future, we applying artificial intelligence and ML algorithm to improve the security and evades the network traffic in the network.

FUNDING INFORMATION

Funding information is not available.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
John Benito Jesudasan Peter	✓	✓	✓		✓	✓		✓	✓	✓			✓	
Nitin Rakesh	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓		✓
Puttaswamy Rekha	✓	✓	✓	✓	✓	✓			✓	✓	✓		✓	
Tammineni Sreelatha	✓	✓		✓	✓	✓		✓	✓	✓		✓		
Velusamy Sujatha	✓	✓	✓		✓	✓	✓		✓	✓		✓	✓	✓
Surulivelu Muthumarilakshmi	✓	✓	✓		✓	✓		✓	✓	✓	✓	✓		
Shanmugam Sujatha	✓	✓			✓	✓			✓	✓	✓	✓	✓	

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

The authors have no conflict of interest relevant to this paper.

DATA AVAILABILITY

The data that support the findings of this study are available on request from the corresponding author, [N. R]. The data, which contain information that could compromise the privacy of research participants, are not publicly available due to certain restrictions.

REFERENCES

[1] J. C. Liu, C. T. Yang, Y. W. Chan, E. Kristiani, and W. J. Jiang, "Cyberattack detection model using deep learning in a network log system with data visualization," *Journal of Supercomputing*, vol. 77, no. 10, pp. 10984–11003, 2021, doi: 10.1007/s11227-021-03715-6.

[2] K. M. Sudar, P. Deepalakshmi, P. Nagaraj, and V. Muneeswaran, "Analysis of cyberattacks and its detection mechanisms," *Proceedings - 2020 5th International Conference on Research in Computational Intelligence and Communication Networks, ICRCICN 2020*, pp. 12–16, 2020, doi: 10.1109/ICRCICN50933.2020.9296178.

[3] M. Karanfil *et al*., "Detection of microgrid cyberattacks using network and system management," *IEEE Transactions on Smart Grid*, vol. 14, no. 3, pp. 2390–2405, 2023, doi: 10.1109/TSG.2022.3218934.

[4] L. Ayala, "Detection of Cyber-Attacks," in *Cybersecurity for Hospitals and Healthcare Facilities*, Berkeley, CA: Apress, 2016, pp. 53–60.

[5] S. Shirali-Shahreza and Y. Ganjali, "Protecting home user devices with an SDN-based firewall," *IEEE Transactions on Consumer Electronics*, vol. 64, no. 1, pp. 92–100, Feb. 2018, doi: 10.1109/TCE.2018.2811261.

[6] A. A. Ali, S. M. Darwish, and S. K. Guirguis, "An approach for improving performance of a packet filtering firewall based on fuzzy petri net," *Journal of Advances in Computer Networks*, vol. 3, no. 1, pp. 67–74, 2015, doi: 10.7763/JACN.2015.V3.144.

[7] A. I. Hajamydeen, N. I. Udzir, R. Mahmod, and A. A. A. Ghani, "An unsupervised heterogeneous log-based framework for anomaly detection," *Turkish Journal of Electrical Engineering and Computer Sciences*, vol. 24, no. 3, pp. 1117–1134, 2016, doi: 10.3906/elk-1302-19.

[8] S. Ranga, N. M. Guptha, and H. M. S, "A survey on automatic abnormalities monitoring system for log files using machine learning," *Turkish Online Journal of Qualitative Inquiry (TOJQI)*, vol. 12, no. 6, pp. 8386–8393, 2021.

[9] K. Fotiadou, T.-H. Velivassaki, A. Voulkidis, D. Skias, S. Tsekeridou, and T. Zahariadis, "Network traffic anomaly detection via deep learning," *Information*, vol. 12, no. 5, p. 215, May 2021, doi: 10.3390/info12050215.

[10] M. Lyu, H. H. Gharakheili, C. Russell, and V. Sivaraman, "Hierarchical anomaly-based detection of distributed DNS attacks on enterprise networks," *IEEE Transactions on Network and Service Management*, vol. 18, no. 1, pp. 1031–1048, Mar. 2021, doi: 10.1109/TNSM.2021.3050091.

[11] Q. Ma, C. Sun, and B. Cui, "A novel model for anomaly detection in network traffic based on support vector machine and clustering," *Security and Communication Networks*, vol. 2021, pp. 1–11, Nov. 2021, doi: 10.1155/2021/2170788.

[12] A. F. Gulliver and T. Aaron, "Unsupervised log message anomaly detection," *ICT Express*, vol. 6, no. 3, pp. 229–237, 2020.

[13] J. O. Onah, S. M. Abdulhamid, M. Abdullahi, I. H. Hassan, and A. Al-Ghusham, "Genetic Algorithm based feature selection and Naïve Bayes for anomaly detection in fog computing environment," *Machine Learning with Applications*, vol. 6, p. 100156, Dec. 2021, doi: 10.1016/j.mlwa.2021.100156.

[14] T.-B.-T. Nguyen, T.-L. Liao, and T.-A. Vu, "Anomaly detection using one-class SVM for logs of juniper router devices," in *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, LNICST*, vol. 293, 2019, pp. 302–312.

[15] A. A. Tahir and S. Anghelus, "Improving iris recognition accuracy using gabor kernels with near-horizontal orientations," *International Journal of Advances in Signal and Image Sciences*, vol. 8, no. 1, pp. 25–39, 2022, doi: 10.29284/ijasis.8.1.2022.25-39.

[16] S. Han, Q. Wu, and Y. Yang, "Machine learning for Internet of things anomaly detection under low-quality data," *International Journal of Distributed Sensor Networks*, vol. 18, no. 10, p. 155013292211337, Oct. 2022, doi: 10.1177/15501329221133765.

[17] P. A. M, C. M. Reddy, A. Anbarasi, N. Mohankumar, I. M.V, and S. Murugan, "Cloud-Based road safety for real-time vehicle rash driving alerts with random forest algorithm," in *2024 3rd International Conference for Innovation in Technology (INOCON)*, Mar. 2024, pp. 1–6, doi: 10.1109/INOCON60754.2024.10511316.

[18] S. Hommes, R. State, and T. Engel, "A distance-based method to detect anomalous attributes in log files," in *2012 IEEE Network Operations and Management Symposium*, Apr. 2012, pp. 498–501, doi: 10.1109/NOMS.2012.6211940.

[19] B. J. Ganesh, P. Vijayan, V. Vaidehi, S. Murugan, R. Meenakshi, and M. Rajmohan, "SVM-based predictive modeling of drowsiness in hospital staff for occupational safety solution via IoT infrastructure," in *2024 2nd International Conference on Computer, Communication and Control (IC4)*, Feb. 2024, pp. 1–5, doi: 10.1109/IC457434.2024.10486429.

[20] B. Meenakshi, A. Vanathi, B. Gopi, S. Sangeetha, L. Ramalingam, and S. Murugan, "Wireless sensor networks for disaster management and emergency response using SVM classifier," in *2023 Second International Conference On Smart Technologies For Smart Nation (SmartTechCon)*, Aug. 2023, pp. 647–651, doi: 10.1109/SmartTechCon57526.2023.10391435.

[21] S. Srinivasan, P. Veda, P. Asha, C. Srinivasan, S. Murugan, and S. Sujatha, "SVM Classifier in IoT-Connected Doorway Thermal Scanning for Preventive Health Check Surveillance," in *2024 1st International Conference on Innovative Sustainable Technologies for Energy, Mechatronics, and Smart Systems (ISTEMS)*, Apr. 2024, pp. 1–6, doi: 10.1109/ISTEMS60181.2024.10560231.
[22] P. Maheswari, S. Gowriswari, S. Balasubramani, A. R. Babu, J. NK, and S. Murugan, "Intelligent headlights for adapting beam patterns with raspberry pi and convolutional neural networks," in *2024 2nd International Conference on Device Intelligence, Computing and Communication Technologies (DICCT)*, Mar. 2024, pp. 182–187, doi: 10.1109/DICCT61038.2024.10533159.
[23] M. Akpinar, M. F. Adak, and G. Guvenc, "SVM-based anomaly detection in remote working: Intelligent software SmartRadar," *Applied Soft Computing*, vol. 109, p. 107457, Sep. 2021, doi: 10.1016/j.asoc.2021.107457.
[24] Z. Li, J. Shi, and M. Van Leeuwen, "Graph neural networks based log anomaly detection and explanation," in *Proceedings of the 2024 IEEE/ACM 46th International Conference on Software Engineering: Companion Proceedings*, Apr. 2024, pp. 306–307, doi: 10.1145/3639478.3643084.
[25] A. Andalib and S. M. Babamir, "Anomaly detection of policies in distributed firewalls using data log analysis," *The Journal of Supercomputing*, vol. 79, no. 17, pp. 19473–19514, Nov. 2023, doi: 10.1007/s11227-023-05417-7.
[26] J. Jegan, M. R. Suguna, M. Shobana, H. Azath, S. Murugan, and M. Rajmohan, "IoT-Enabled Black Box for Driver Behavior Analysis Using Cloud Computing," in *2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS)*, Apr. 2024, pp. 1–6, doi: 10.1109/ADICS58448.2024.10533471.
[27] G. Thahniyath *et al.*, "Cloud based prediction of epileptic seizures using real-time electroencephalograms analysis," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 14, no. 5, p. 6047, Oct. 2024, doi: 10.11591/ijece.v14i5.pp6047-6056.
[28] J. Ramasamy, E. Srividhya, V. Vaidehi, S. Vimaladevi, N. Mohankumar, and S. Murugan, "Cloud-enabled isolation forest for anomaly detection in UAV-based power line inspection," in *2024 2nd International Conference on Networking and Communications (ICNWC)*, Apr. 2024, pp. 1–6, doi: 10.1109/ICNWC60771.2024.10537407.
[29] S. Srinivasan, R. Raja, C. Jehan, S. Murugan, C. Srinivasan, and M. Muthulekshmi, "IoT-enabled facial recognition for smart hospitality for contactless guest services and identity verification," in *2024 11th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, Mar. 2024, pp. 1–6, doi: 10.1109/ICRITO61523.2024.10522363.
[30] V. V. Baskar, S. Sekar, K. S. Rajesh, N. C. Sendhilkumar, T. R, and S. Murugan, "Cloud-based decision support systems for securing farm-to-table traceability using IoT and KNN Algorithm," in *2024 Second International Conference on Intelligent Cyber Physical Systems and Internet of Things (ICoICI)*, Aug. 2024, pp. 443–448, doi: 10.1109/ICoICI62503.2024.10696659.
[31] S. Sekar *et al.*, "Intrusion detection and prevention using Bayesian decision with fuzzy logic system," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 15, no. 1, p. 1200, Feb. 2025, doi: 10.11591/ijece.v15i1.pp1200-1208.

BIOGRAPHIES OF AUTHORS

John Benito Jesudasan Peter is a seasoned Senior Manager in Data Science and Engineering with a track record of over 17 years in technology leadership in the Congnizant. He's a passionate and innovative leader who thrives on delivering value through people, processes, and technology. Currently at the helm of Data Science and Engineering at CapitalOne, John holds a Masters in Data Science and Engineering from BITS Pilani, underscoring his commitment to continuous learning and professional growth. John is a maestro in defining and planning ML Operations (MLOps) and the data engineering journey for organizations. He's adept at building AI-based anomaly detection models, with a focus on security, and architecting and implementing big data pipelines and data lakes in AWS. John's extensive list of certifications includes IIM Bangalore's People Management certification, AWS Certified ML Specialty, Databricks Certified Data Engineer Associate, and Google Analytics Certified. With a solid educational background, including a postgraduate degree in Data Science and Engineering and another in Artificial Intelligence, John is well-equipped for the challenges of the industry. He's also affiliated with Texas McCombs, the University of Texas at Austin. He can be cotacted at: johnbenitoauthor@gmail.com.

Dr. Nitin Rakesh is a recipient of IBM Drona Award and Top 10 State Award Winner. He is active member of professional society like Senior Member IEEE (USA), ACM, SIAM (USA), Life Member of CSI and other professional societies. He is reviewer of several prestigious Journals/Transactions like IEEE Transactions on Vehicular Technology, The Computer Journal, Oxford Press, many SPRINGER/other Scopus indexed International Journals. His research outlines emphasis on Network Coding, Interconnection Networks and Architecture and Online Phantom Transactions. He has published various SCI/Scopus Journal/Conference publication in high quality journals and conferences. Dr Nitin is having more than 90 patents/innovations/designs in his credentials. Dr. Nitin has accorded several other awards for best paper published, session chairs, highest cited author, best student's thesis guided, and many others. He can be contacted at email: nitin.rakesh@gmail.com

Puttaswamy Rekha is a Professor in the Department of Electronics and Communication Engineering at BNM Institute of Technology, Bangalore, and Karnataka. She has more than 20 years of teaching and 15 years of R&D experience. She completed her Ph.D. at Visvesvaraya Technological University, Belagavi and her postgraduate studies (2002 -2004) at University Visvesvaraya College of Engineering and undergraduate (1994 -1998) studies at R V College of Engineering, Bangalore University. Her research interests are in the area of sensors, embedded systems, IoT, network security, cyber security, robotics, data analytics, digital design, and microcontrolers. Dr. Rekha P has been part of international conferences in various capacities. She has attended, organized and conducted various workshops for studentsand faculty development programs. Till date she has published more than 20 research papers in various National, International Journals (Scopus index, Web of science, UGC impact factors) and conferences in India. She has authored book chapters with oneIndian Design patent grant. She has been working as a consultant with Agimus Technologies Pvt. Ltd. and Cogniverse Labs Pvt. Ltd. She has served as a review member in International Conferences and other Journals. She has executed funded projects received from AICTE, KSCST, and NewGen IEDC. She has received best paper award for her work presented in the ICCIIS 2024 International Conference. She can be contacted at: rekhap@bnmit.in.

Dr. Tammineni Sreelatha working as an Assistant Professor in the Electronics and Communication Engineering Department at the Koneru Lakshmaiah Education Foundation located in Vaddeswaram, Guntur District, Andhra Pradesh. She has 13 years of teaching experience and earned her Ph.D. in Electronics and Communication Engineering from Jawaharlal Nehru Technological University Anantapur. She is specialized in leveraging ML techniques for Image Processing applications. Throughout her academic career, she has contributed 20 papers to international journals in her field of expertise. She is an active life member of the Indian Society for Technical Education (ISTE) and holds memberships in both the International Association of Engineers (IAENG) and the International Association of Computer Science and Information Technology (IACSIT). Additionally, she has participated in numerous training sessions and conferences that focus on advancements in signal and image processing. She can be contacted at: sreelatha457@gmail.com.

Dr. Velusamy Sujatha received her doctoral degree and post-graduation from Anna University, Chennai, and her undergraduate degree from Madras University. Her research focuses on designing charge pump Phase-Locked Loops, with a keen interest in Analog VLSI. She began her career as a lecturer in 1993 and has served as an academician in various reputed institutions, holding several positions and guiding numerous UG and PG projects. Currently, she is a Professor in the Department of Electronics and Communication Engineering (ECE) at Shree Sathyam College of Engineering and Technology, Sankari, Tamil Nadu. She can be contacted at: info@shreesathyam.edu.in.

Dr. Surulivelu Muthumarilakshmi is an Associate Professor in Computer Science and Engineering at Chennai Institute of Technology, Kundrathur. With over 13 years of teaching experience, my primary focus revolves around Computer Networks. I am particularly interested in investigating network protocols, security measures, and ways to optimize network performance. My passion lies in researching and publishing articles that delve into these areas, aiming to enhance our understanding of robust network systems and contribute valuable insights to the academic community. She can be contacted at email: muthu3041974@gmail.com.

Shanmugam Sujatha is an adjunct professor, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, Tamilnadu, India. She published her research articles in many international and national conferences andjournals. Her research areas include network security and ML. She can be contacted at email: sujathasmvr@gmail.com.