

Neural-network based representation framework for adversary identification in internet of things

Thanuja Narasimhamurthy, Gunavathi Hosahalli Swamy

Department of Computer Science and Engineering, Bangalore Institute of Technology, Bengaluru, affiliate to Visvesvaraya Technological University, Belagavi, India

Article Info

Article history:

Received Aug 1, 2024

Revised Mar 18, 2025

Accepted May 23, 2025

Keywords:

Dynamic threats

Indexing

Internet of things

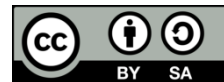
Machine learning

Neural network

ABSTRACT

Machine learning is one of the potential solutions towards optimizing the security strength towards identifying complex forms of threats in internet of things (IoT). However, a review of existing machine learning-based approaches showcases their sub-optimal performance when exposed to dynamic forms of unseen threats without any a priori information during the training stage. Hence, this manuscript presents a novel machine-learning framework towards potential threat detection capable of identifying the underlying patterns of rapidly evolving threats. The proposed system uses a neural network-based learning model emphasizing representation learning where an explicit masked indexing mechanism is presented for high-level security against unknown and dynamic adversaries. The benchmarked outcome of the study shows to accomplish 11% maximized threat detection accuracy and 33% minimized algorithm processing time.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Thanuja Narasimhamurthy

Computer Science and Engineering, Bangalore Institute of Technology affiliate to Visvesvaraya

Technological University

Vishweshwarapura, Basavanagudi, Bengaluru, Karnataka 560004, India

Email: nthanuja@bit-bangalore.edu.in

1. INTRODUCTION

Internet of things (IoT) comprises various forms of software and hardware connected in a larger context to disseminate and analyze data from various application perspectives [1]. Irrespective of various related concerns, security issues dominate the research communities, even if a more extensive set of security solutions is being introduced for IoT [2]. From the perspective of IoT security, various challenges are associated with performing frequent sets of operations, e.g., data encryption, access control, firmware updates, and device authentication [3]. Some key security issues in IoT are mainly associated with scalability, limited resources, diverse ecosystems, privacy concerns, and physical security [4]. There is an evolving landscape of cyber threats in IoT, which are consistently being researched [5]. Out of various security solutions, artificial intelligence (AI) has eventually played a robust solution to deal with this complex security problem with various evolving machine learning algorithms [6]–[9]. Existing studies have noted the adoption of AI in anomaly detection [10], adaptive security [11], automated threat response [12], and predictive capabilities [13]. However, there are rising concerns about adopting machine learning algorithms in IoT security, too. The primary limitation is associated with the dependency of massive data volumes for machine learning models which is a rising concern towards sensitivity and data privacy. This problem is also linked to the increasing requirement for training data of higher quality and model complexity. Further, there is always the risk of overfitting thereby restricting the capability of the algorithm to adapt for fluctuating data. However, one notable problem that is normally overseen by the majority of existing studies is that once

the dataset is trained, the learning model completely depends on it for its predictive analysis. Under such a scenario, if there is any form of deviation in the network scenario or if there is any form of unknown or dynamic threats, the trained model fails to support the higher and more reliable threat detection by the existing machine learning model. Hence, existing studies using machine learning may sound good, but in reality, their applicability is restricted to static environment implementation only. Hence, the prime gap of the existing study is mainly related to the non-updating process for the trained model during the actual deployment process by existing machine learning models. Further, there is a need to develop such a model that can further address the overfitting and reliability problems while using machine learning in order to balance the demands of successful, updated, and reliable threat detection with a cost-effective learning model.

The related background work of literature shows that there are various contributions of machine learning-based approaches towards boosting the security features in IoT [14]. There are various studies reviewed considering adopted methodologies, features, and datasets. Table 1 highlights the characteristic being obtained from existing studies [15]–[30]. Hence, it can be seen that XGBoost, support vector machine (SVM), random forest (RF), and decision tree (DT) are frequently adopted approaches noted in existing approaches.

Table 1. Review of existing studies

Studies	Methods	Features	Dataset
[15]	Multilayer perceptron, SVM, RF	Health-related	WUSTL-EHMC2020, ICU-IOMT
[16]	DT	18 sensor-related features	WSN-DS
[23]	Collaborative learning	Traffic data	ECU-IoHT
[24]	XGBoost, Fisher score model, genetic algorithm	115 features from traffic data	N-BaIoT
[25]	Ensemble	4 Features from network traffic	NSL-KDD
[26]	CNN, nature-inspired algorithm	34 features from network traffic	IoT-MQTT
[27]	XGBoost, RF	115 features from traffic data	N-BaIoT
[28]	SVM	Data features related to activity, 80-traffic feature	Synthetic, CSE-CIC-IDS2018
[29]			
[30]	SVM (multi-class), DT	20 features from IoT traffic	Multistep cyber-attack dataset
[17]	SVM, K-nearest neighbor, RF	A large number of features	Multiple datasets
[18]	CNN	Features for 25 classes	MaleVis
[19]	Machine learning, encryption	Features related to traffic flow	Synthetic dataset
[20]	Multiple machine learning	Traffic-related feature	IoT-23
[21]	Multiple machine learning	45-features	TON-IoT
[22]	Multiple machine learning	6-features related to traffic flow	NFUQ-NIDS-v2

The research problems associated with existing studies are manifold. With the advancement of varied networking in large scale and distributed manner, there is also an immense increase in different threats. Such threats have been reported to have high potential and can bypass any security monitoring system in IoT (e.g. firewall). The existing security firewall system can only resist the threats already considered while developing the security patch for the firewall. Hence, firewall systems cannot identify any threats that are not declared and defined well in their standard database system. Hence, there is a need for an innovative system capable of identifying a security threat of unknown form different from that used in developing a firewall system.

Therefore, the proposed study aims to develop a computational model capable of harnessing the potential of a machine-learning approach towards dynamic, unknown, and uncertain threat detection in IoT. The value-added contributions in current work are as follows: i) a novel threat detection scheme has been presented capable of learning the dynamic patterns and correlation of potential vulnerabilities, ii) a distinct updating scheme has been introduced that can optimize the data quality of training model to realize the dynamic events in incoming traffic, iii) a simplistic and yet innovative neural network based machine learning model has been presented that can perform representation learning to extract more significant underlying patterns of threats, iv) a unique masked indexing mechanism has been presented towards safeguarding the labels of trained data from adversaries. The paper's organization is: Section 2 presents a discussion of methods, while the result is discussed in Section 3, and the conclusion is presented in Section-4.

2. METHOD

The proposed system aims to develop a novel detection scheme of potential threats by harnessing machine learning towards securing future communication systems *i.e.* IoT. An IoT infrastructure is formed

by considering a specific number of IoT devices that capture the data and forward it to a processing unit hosted within an edge device. The processing unit, considering an additional 50+ features using a standard dataset related to traffic flow, further executes the algorithms towards scrutinizing threats that are hosted in a connected cloud environment. Analytical modelling is considered for developing the research methodology, where the complete operation is classified into dual implementation stages. The first implementation is towards identifying the potential threat, while the second implementation is about applying a neural network based machine learning model to perform updating of incoming threats. Figure 1 highlights the adopted architecture of the proposed methodology.

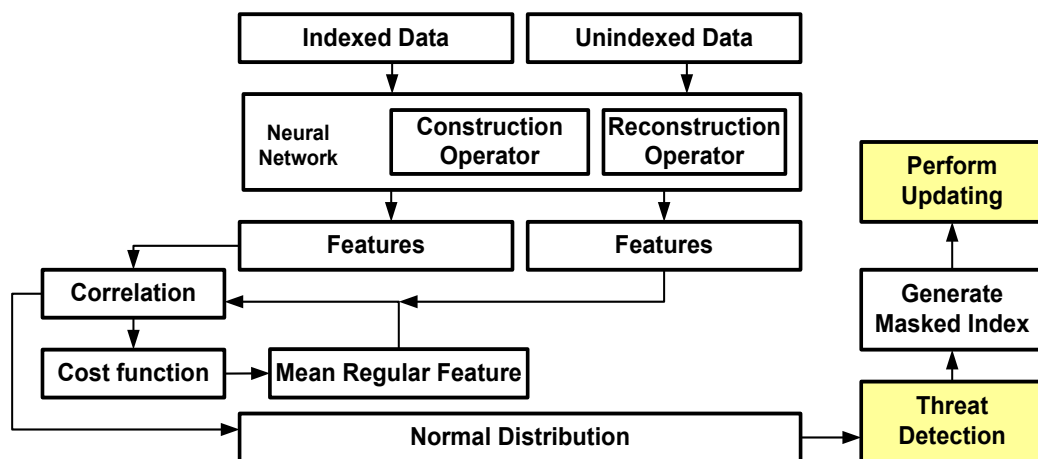


Figure 1. Architecture of proposed research methodology

According to Figure 1, the traffic information is split into indexed and unindexed data which is subjected to a neural network model with an inclusion of construction and reconstruction operator. The neural network model adopted towards performing learning of traffic data representation where the input traffic data is initially subjected to forward pass where construction operator is used to generate a latent representation while reconstruction operator is used to generate reconstruction of original input using latent representation generated in prior stage. The operation yields feature that are subjected to correlation analysis followed by using cost function with an objective to make the reconstruction matching with the original input. The cost function adopted in the proposed system is basically designed for learning representation that is capable of extracting both similarities and differences among data points. This is done with the objective to incentives regular traffic and penalize malicious traffic. The novelty of this neural network model is that it uses its operator and cost function to reduce the distance between the positive traffic data (representing regular traffic) and increasing the distance between dissimilar data (representing malicious traffic).

The operations of the cost function result in mean regular features, which, along with extracted values of normal distribution, are subjected towards assessment of regular and malicious traffic, thereby accomplishing the threat detection operation. After the threat detection is performed, the next task is to increase the adaptability of the model to perform detection over a stream of dynamic traffic information. The same neural network model is trained with a random weight, resulting in the masked index which represents a second layer of encoded value of labelled data of traffic. The beneficial aspect of this operation is that the adoption of masked label offers further data integrity as well as data privacy as no attacker will ever be able to compute the actual label information of the traffic data. Another beneficial aspect of this module is that – the model can now be able to easily identify any form of spurious traffic if an attacker attempts to tamper with either the data or the label of it. The tampered label of the data by the attacker is less likely to match with the computed value of the masked index presented in the proposed system and hence strongly acts similarly to a trapdoor function used in a conventional cryptography-based approach.

Further, the proposed study introduces updating operations in order to increase the capability of coverage of trained data by the proposed neural network model to be able to resist potentially dynamic forms of threats. Interestingly, the internal operation of the scheme is quite progressive and less iterative. Eventually, it uses machine learning methodology which contributes towards an effective balance between security features and computational features. An illustration of design implementation follows next:

2.1. Potential threat identification

This is the first module of implementation that is responsible for threat identification. The prime task of this algorithm is basically to perform the identification of potential threats, followed by the generation of a masked index. Basically, the term masked index is meant for safeguarding the labels assigned for the incoming traffic data, which otherwise could be tampered with or eavesdropped by an attacker. The steps of the proposed algorithm towards threat identification are as Algorithm 1:

Algorithm 1. Algorithm for threat identification

```

Input: T
Output: TI
Start
1. init  $T \rightarrow (T_r, T_m)$ 
2.  $(T_r, T_m) \rightarrow [(\alpha_r, \beta_r) (\alpha_m, \beta_m)]$ 
3. compute  $\rho = f_1(tr_i, tr_j)$ 
4.  $\lambda = f_2(\rho, \delta)$ 
5.  $\lambda_1 = \arg_{\min}(\Sigma \lambda)$ 
6.  $TI = f_3(\rho(r, m))$ 
End

```

The illustration of the algorithm is as follows: The algorithm takes the input of T (traffic data), which, upon processing, will lead to the generation of TI (threat identification). The algorithm initializes the traffic data T concerning regular traffic T_r and malicious traffic T_m (Line-1). Further, the traffic data is represented in the form of input traffic vector (α_r, α_m) of regular and malicious forms as well as associated indexes (β_r, β_m) of regular and malicious forms, respectively (Line-2). It will eventually mean that the traffic data T will consist of input traffic vector α of s size while each index β is initialized with a binary range of $[0, 1]$ where 0 and 1 depict regular and malicious traffic, respectively. The proposed algorithm considers the trained representation associated with each input traffic data as $tr(\in TR)$, while the proposed neural network model used for threat detection is represented by π_w , where w represents the weight associated with the neural model.

The proposed system uses a unique cost function that targets to increase the correlation between regular traffic $(tr_{r,i}, tr_{r,j})$ and decrease the same for reference sample $tr_{r,i}$ and malicious traffic $tr_{m,k}$. The variable $tr_{r,i}$ represents the reference sample associated with regular traffic information while the value of the subscripts (i, j) is $(1, 2, 3, \dots, I_r)$, where variable I_r represents the index for regular traffic. The value of subscript k is $(1, 2, 3, \dots, I_m)$, where variable I_m represents the index for malicious traffic. In the consecutive part of algorithm implementation, the scheme computes correlation attribute ρ by constructing an explicit function $f_1(x)$ considering input arguments of regular traffic information tr_i and tr_j (Line-3). The function $f_1(x)$, which is intended to evaluate the similarity between pairs of traffic data points, is used by the system to calculate the correlation attribute ρ . In particular, $f_1(x)$ computes the dot product of two traffic samples' feature vectors and normalizes the outcome using the Euclidean distance. This process makes sure that the correlation accurately represents how similar the samples are, which is essential for differentiating malicious traffic from legitimate traffic. The neural network learns to distinguish between malicious and legitimate traffic by using the correlation score as input to the cost function λ . The computation of this explicit function $f_1(x)$ is empirically expressed as (1):

$$f_1(x) = (tr_i \cdot tr_j) \cdot [abs(tr_i, tr_j)]^{-1} \quad (1)$$

After the explicit function $f_1(x)$ computation, the system computes the value of the correlation attribute ρ , which is further utilized towards its consecutive steps for computing cost function λ . A closer look into Line-4 showcases that the computation of cost function λ is carried out by developing another explicit function $f_2(x)$ with an input argument of correlation attribute ρ and adjustment parameter δ . The step-wise computation of cost function λ is shown as (2) and (3):

$$\lambda_{i,j} = c \cdot \frac{A_1}{A_2} \quad (2)$$

$$\lambda = [I]^{-1} \sum \lambda_{i,j} \quad (3)$$

A closer look into the expressions (2) and (3) show that the proposed scheme initially computes expression (2), which is then used in expression (3). In the initial computational step of expression (2), the variable A_1 represents an exponential form of correlation attribute ρ for (i, j) divided by adjustment parameter δ . At the same time, the variable A_2 represents the summation of the newly accomplished value of A_1 and all

additive values of A_1 with the lower range of (i, k) and index range of (I_r, I_m) . In the next computational step of expression (3), the variable I represents $I_r(I_r-1)$, where I_r is the index for regular traffic and the second component represents all additive values of expression (2). The algorithm minimizes the cost function associated with both construction co and reconstruction re to efficiently learn the traffic representation to obtain the new value of cost function λ_1 (Line-5). The variable $\Sigma\lambda$ represents the summation of λ_{co} and λ_{re} . In the final step of the algorithm (Line-6), another explicit method $f_3(x)$ is used for higher probability assessment (HPA) in order to derive an accurate correlation score ρ with the normal distribution of traffic with respect to both regular and malicious. This information of the identified threat is stored in matrix TI.

2.2. Updating operation for incoming threat

This is the second part of implementation which is mainly concentrated towards performing updated operations for all incoming threat-prone traffic. A closer look into the adoption of machine learning approaches towards threat detection in existing systems is always found to depend on its predictive operation on its trained data. The very assumption that all trained data will consist of complete possibilities of traffic events upon exposing them in a deployment scenario is impractical, especially regarding dynamic threats. Hence, the prime objective of this algorithm is to increase the adaptability of the threat detection model to understand and realize the dynamicity involved in the traffic environment by generating updated information on incoming threats. The core contribution of this model is its freedom from any form of demands of manual indexing tasks while performing training operations over streams of traffic information. The algorithmic steps are as shown in Algorithm 2:

Algorithm 2. Algorithm for updating incoming threat

```

Input:  $T_o, w$ 
Output:  $up$ 
Start
1. init  $w$ 
2. train param
3.  $\alpha, \beta \leftarrow \alpha_o, \beta_o$ 
4. For  $i=\alpha_i$ 
5.    $N[(co, re)(x, m)] = f_4(param_1)$ 
6.   For  $j=0: (h-1)$  do
7.      $res(co, re)_{i+j} = \pi_w(\alpha_{i+j})$ 
8.      $\beta_1 = f_5(param_2)$ 
9.      $up(\alpha, \beta) = f_6[(\alpha, \beta), (\alpha_{i+j}, \beta_{i+j})]$ 
10.  End
11. End
12. initiate training( $param_2$ )
End

```

The above-shown algorithm takes the input of T_o (dataset for training) and w (weight), which, after processing, yields an outcome of up (updated information of threat). Neural network framework π_w is initialized considering weights w of arbitrary form (Line-1) is used for training the system considering a pilot sample of indexed traffic information that consists of input traffic vector α_o and associated indexes β_o considered for pilot epoch round e_o (Line-1 and Line-2). The variable $param$ represents input traffic vector α_o , associated indexes β_o , and epoch e_o (Line-2). The input traffic vector α_o and associated indexes β_o are further assigned to the new mapping attribute of input traffic vector α and associated indexes β (Line-3). The following line of operation considers all the input traffic vector α_i (Line-4), followed by an evaluation of the normal distribution N (Line-5). For this purpose, an explicit function $f_4(x)$ is constructed to assess the suitability of normal distribution considering $param_1$ representing pilot input traffic vector α_o , input traffic vector α , and neural network framework π_w (Line-5). The evaluated normal distribution suitability score is then assigned to matrix N considering construction operator co , reconstruction operator re associated with regular traffic r and malicious traffic m (Line-5). In simple words, this operation (Line-5) extracts the normal distribution of construction operator co and reconstructor operator re using pilot traffic information T_o .

Further, the algorithm generates the masked index (Line-5 to Line-8) followed by the generation of updated information (Line-9 to Line-12). Considering the phases involved in Training ranging from 0 to $(h-1)$ as shown in Line-6, the algorithm obtains results res concerning construction operator co and reconstruction operator re (Line-7). It eventually infers towards accomplishing results concerning input traffic vector α (Line-7). In the consecutive process, the algorithm allocates the masked index to all the recently generated input traffic aj that has already been detected in the first algorithm using the Neural network framework π_w . The function $f_5(x)$ is meant for undertaking a decision considering newly generated result res in the previous step concerning $param_2$ (Line-8) that can be further empirically expressed as (4):

$$\beta(co)_{i+j} \rightarrow res(co), N(co, r), N(co, m) \quad (4)$$

$$\beta(re)_{i+j} \rightarrow res(re), N(re, r), N(re, m) \quad (5)$$

In the expression (2) and (3), the $\{res(co), N(co, r), N(co, m)\}$ and $\{res(re), N(re, r), N(re, m)\}$ together constitute of $param_2$. The function $f_5(x)$ generates a value of masking index β as 1 if it finds error-prone input followed by a flagging alert to the system. It will also mean that the previous algorithm plays the role of inference to yield outcomes in this model while performing generation for masked indexes. Only the outcome of masked index β_{i+j} that is found to have a higher score is considered further, followed by randomly permuting masked index β . This algorithm's final step is updating both input traffic vector α and associated masked indexes β (Line-9). For this purpose, function $f_6(x)$ has been constructed, which is mainly performing union operations between α , α_{i+j} , and β_{i+j} while the Training is carried out on $param_2$, which consists of α , β , π_w , and new epoch e_1 (Line-12). It can be seen that training dataset T is now expanded with the inclusion of a new input traffic vector α_j and its associated masked index β . Considering h number of incoming traffic being newly identified, new epoch e_1 is used for adjusting the neural network model π_w , which eventually means that the final index consists of both ground-truth information index and masked index. The following section discusses about accomplished results.

By doing away with manual labeling and ongoing retraining, the masked indexing mechanism presents a revolutionary method for dynamic threat detection. By employing "masked indexes" to conceal traffic labels and thwart label manipulation by adversaries, it enables the system to automatically identify and adjust to changing threats in real-time. Even in dynamic, label-free situations, which are typical of IoT networks, this approach guarantees reliable threat detection. Because iterative re-training of the system is not necessary, it also reduces computing overhead, increasing efficiency. Because of the increased accuracy, shorter processing times, and more trustworthy threat detection that results, it is especially useful in settings where data patterns are ever-changing.

3. RESULT

The scripting of the proposed study is carried out on a python environment on a normal 64-bit Windows environment considering two standard benchmarked datasets implemented on a standard 64-bit Windows system with NVIDIA GeForce GTX with 16 GB RAM and Intel Core i5 processor. The environment was retained the same for assessing the proposed existing system. The first dataset is the NSL-KDD dataset [31], where 25 classes of malicious forms are considered for training operations encapsulating attackers that attempt to explore various types of attacks as follows: i) Attacker Type-1: This type of attackers exploit the weakness of network structure, e.g., Port scanning, ping sweeping, network mapping, service enumeration, vulnerability scanning. ii) Attacker Type-2: This type of cyber-attack to gain illegitimate access to control the network or machine of the victim node remotely, e.g., SQL injection, buffer overflow, remote code execution. iii) Attacker Type-3: This type of cyber-attack is used for gaining access to root administrative accounts illegally, e.g., Kernel exploits, misconfiguration of set group ID (SGID) or set user ID (SUID). iv) Attacker Type-4: This attacker is meant to flood illegitimate requests in order to disrupt the normal services from a server, e.g., Volume-based attacks, protocol attacks, application layer attacks, denial-of-service (DoS), and distributed DoS (DDoS). The second dataset is the UNSW-NB15 dataset [32], with more than 100 GB of raw traffic information extracted from artificially constructed adversaries with 9 classes of attack with 49 features. This dataset provides vulnerable environmental assessment for DoS, worms, and malware. There is a total of 175, 341, and 82, 332 records used for training and testing. The proposed scheme uses 5 layers with following sizes i) for first dataset (size of input layer=121, size of constructor layer=64, size of hidden layer=32, size of reconstructor layer=64, size of output layer=121), and ii) for second dataset (size of input layer=196, size of constructor layer=128, size of hidden layer=64, size of reconstructor layer=128, size of output layer=196). The adjustment parameter δ is set to 0.02 with 128 batch size, while 0.001 is set for the rate of learning considering rectilinear unit (ReLU) as the activation function. The scheme uses gradient descent of stochastic form to play the role of optimization in the proposed machine learning model. The NSL-KDD dataset consists of network-based attacks, while the UNSW-NB-15 dataset consists of more diverse attack coverage on more realistic scenarios of IoT. Adoption of these datasets facilitates to simulation of general attacks based on network vulnerabilities in IoT. The outcome is assessed for varied accuracy parameters (precision, recall, and F1-score), overall accuracy, and algorithm processing time. The benchmarking is carried out by comparing the proposed system with existing machine learning algorithms that are frequently reported in the literature, e.g., XGBoost, SVM, RF, and DT as shown in Figures 2 to 6. The outcome of the analysis eventually shows the proposed scheme of a neural network to perform better performance in contrast to existing machine learning approaches with respect to accuracy and processing time-based evaluation metric.

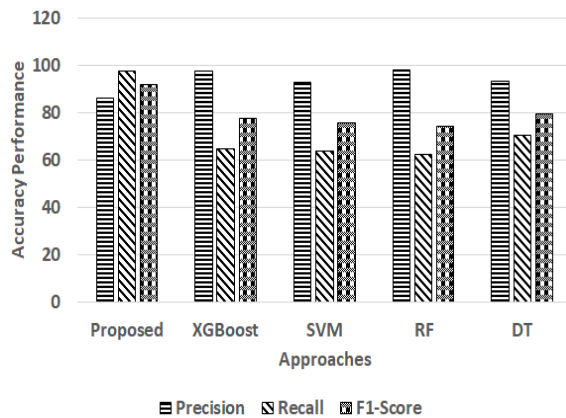


Figure 2. Accuracy for NSL-KDD dataset

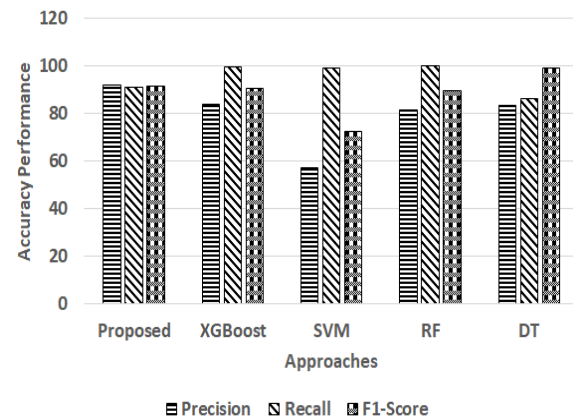


Figure 3. Accuracy for UNSW-NB15 dataset

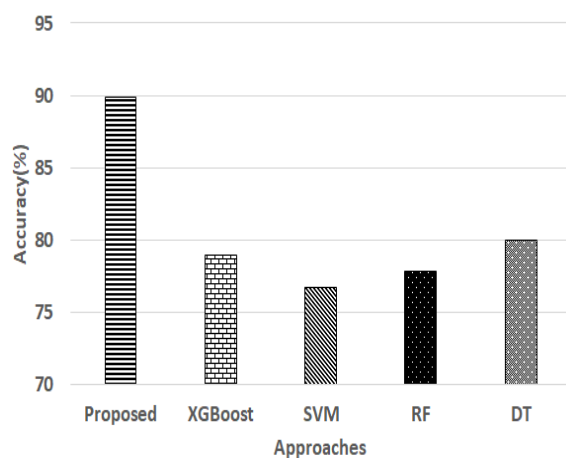


Figure 4. Accuracy score for NSL-KDD dataset

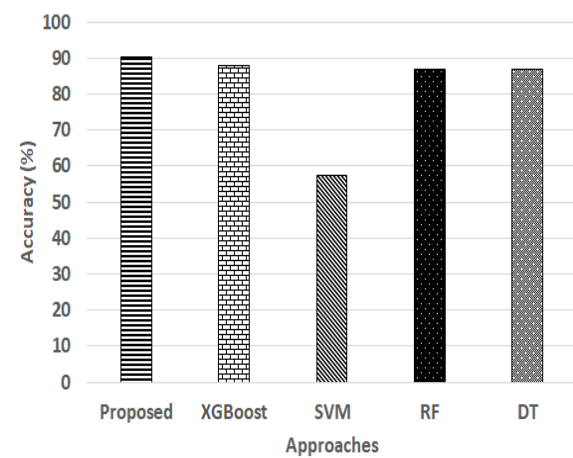


Figure 5. Accuracy score for UNSW-NB15 dataset

Discussion of accuracy outcomes is illustrated in Figures 2 to 5. The quantification of accuracy outcomes shows that there is no significant difference in accuracy outcomes for the first and second datasets. The proposed system is found to offer approximately 12% and 10% better accuracy scores recorded for the first and second datasets, respectively, in contrast to existing machine learning approaches. The justification is as follows: Although the XGBoost approach is known for its higher accuracy performance with better handling capacity of imbalanced data, it is the computationally intensive model when exposed to a larger dataset. It works well for small datasets, but its accuracy starts declining when exposed to streams of data. A similar problem is also encountered for the SVM approach, and its performance is potentially dependent on an appropriate kernel function selection. Although this selection is somewhat easier for smaller and static sets of data its accuracy starts degrading when assessed with increasing random sizes of incoming traffic of data. The performance of RF is found to be quite good in accuracy for the second dataset, while the DT approach is found to be better for both datasets. However, the DT approach cannot be considered suitable in this implementation scenario as it can lead to a higher degree of instability in case of smaller changes in data. However, the proposed scheme offers better accuracy mainly because of the second algorithm towards the updating process, minimizing all the extra time required to find optimal results. Owing to fitment testing with processed features with normal distribution, the detection outcome is quite reliable with respect to its overall accuracy score. The adoption of correlational score between trained representation offers a better form of data representation considering optimal cost function associated with the constructor and reconstructor operator. This is further justified by evaluating its computational efficiency with respect to algorithm processing time results exhibited in Figure 6.

Discussion of algorithm processing time is illustrated in Figure 6. The quantified outcome shows that the proposed system offers approximately 28% and 38% of reduced algorithm processing time in comparison to the mean of existing machine learning algorithms for the first and second datasets,

respectively. The justification behind this outcome is as follows: Both XGBoost and SVM are noted with considerably higher algorithm processing time, which is mainly due to their extensive hyperparameter tuning and slower computational process respectively. For both datasets, their performance is nearly the same. RF algorithm generates a large number of trees that eventually leads to increased complexity especially during the training operation. DT has better performance in contrast to XGBoost, SVM, and RF, which is mainly due to its independence from normalization or feature scaling step; however, when exposed to unindexed data, it is noted to be prone to overfitting, leading to inferior generalization of unseen data. However, the methodology involved in the proposed scheme calls for unsupervised learning, while the identification of outliers is easily done by estimating reconstruction error without much dependency on labelled data. Further, the model is highly adaptive to complex patterns of new types of intrusion by finetuning the data, as noted in the second algorithm of the proposed system. Hence, without the inclusion of any sophisticated iterative steps, the algorithm can carry out progressive operations to identify even smaller changes in traffic with higher accuracy. This ground of fact is attributed towards reduced algorithm processing time for the proposed machine learning-based security scheme.

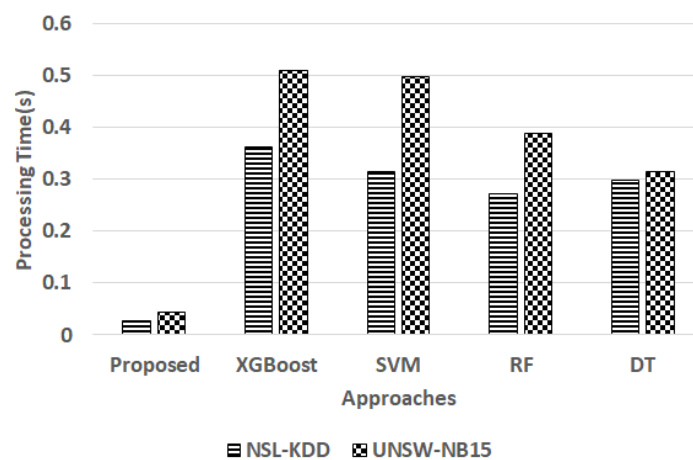


Figure 6. Processing time analysis

4. CONCLUSION

The adoption of machine learning algorithms has been frequently witnessed in existing literature towards threat detection on various forms of network systems. Keeping aside the potential capability of machine learning algorithms towards addressing security threats, one of the challenging problems identified is that – a trained machine learning model is not sufficient to mitigate dynamic forms of threats, which is a major gap in the literature. This gap is addressed in proposed system by following notable contribution: i) proposed scheme implements a highly adaptable and flexible architecture which is capable of identifying dynamic threats with faster response and higher accuracy, ii) neural network-based learning model has been designed with an inclusion of two additional layers (construction operator and reconstructor operator) which assists towards better form of learning representation of variable and uncertain traffic pattern, iii) proposed scheme is completely free from any form of human intervention which can undertake its own decision threat detection and updating process, iv) an innovative concept of masked index has been introduced to address the task of labelling so that an involuntary threat detection scheme can be defined on dynamic environment, and v) proposed system offers approximately mean of 11% increased accuracy and 33% of reduced processing time when compared with frequently used machine learning towards cyber security. For real-world applicability, the model needs to be synchronized with existing edge devices for particular service providers of IoT-cloud systems. Various challenges that could surface are handling and processing massive incoming traffic from IoT devices with respect to designated edge devices. The complete system operation is carried out considering fault-tolerant processing carried out within edge devices and servers that execute algorithms in a cloud environment. This dependency could be one of the prime limitations although it will not affect algorithmic performance. Hence, Future work will be continued towards improving the same model with more parametric inclusion of cloud attributes and edge device attributes apart from IoT devices, which was the prime focus of this study. The future work will have an inclusion towards novel formulation of threat prevention strategy considering large IoT environment.

ACKNOWLEDGEMENTS

We wish to confirm that no known conflicts of interest are associated with this publication and all the authors have contributed equally.

FUNDING INFORMATION

Authors state no funding involved.

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
Thanuja Narasimhamurthy	✓	✓	✓	✓	✓	✓		✓	✓	✓			✓	
Gunavathi Hosahalli		✓		✓		✓		✓	✓	✓	✓	✓		
Swamy														

C : Conceptualization

M : Methodology

So : Software

Va : Validation

Fo : Formal analysis

I : Investigation

R : Resources

D : Data Curation

O : Writing - Original Draft

E : Writing - Review & Editing

Vi : Visualization

Su : Supervision

P : Project administration

Fu : Funding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

The data that support the findings of this study are available on request from the corresponding author.

REFERENCES




- [1] A. S. Mashaleh *et al.*, "Evaluation of machine learning and deep learning methods for early detection of internet of things botnets," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 14, no. 4, p. 4732, Aug. 2024, doi: 10.11591/ijece.v14i4.pp4732-4744.
- [2] T. Nagaraj and R. K. Channarayappa, "An efficient security framework for intrusion detection and prevention in internet-of-things using machine learning technique," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 14, no. 2, p. 2313, Apr. 2024, doi: 10.11591/ijece.v14i2.pp2313-2321.
- [3] L. Zhang and L. Wang, "A hybrid encryption approach for efficient and secure data transmission in IoT devices," *Journal of Engineering and Applied Science*, vol. 71, no. 1, p. 138, Dec. 2024, doi: 10.1186/s44147-024-00459-x.
- [4] T. Rajmohan, P. H. Nguyen, and N. Ferry, "A decade of research on patterns and architectures for IoT security," *Cybersecurity*, vol. 5, no. 1, p. 2, Dec. 2022, doi: 10.1186/s42400-021-00104-7.
- [5] K. Kandasamy, S. Srinivas, K. Achuthan, and V. P. Rangan, "IoT cyber risk: a holistic analysis of cyber risk assessment frameworks, risk vectors, and risk ranking process," *EURASIP Journal on Information Security*, vol. 2020, no. 1, p. 8, Dec. 2020, doi: 10.1186/s13635-020-00111-0.
- [6] T. Mazhar *et al.*, "Analysis of IoT security challenges and its solutions using artificial intelligence," *Brain Sciences*, vol. 13, no. 4, p. 683, Apr. 2023, doi: 10.3390/brainsci13040683.
- [7] S. Selvarajan *et al.*, "An artificial intelligence lightweight blockchain security model for security and privacy in IIoT systems," *Journal of Cloud Computing*, vol. 12, no. 1, p. 38, Mar. 2023, doi: 10.1186/s13677-023-00412-y.
- [8] M. A. Ferrag, L. Maglaras, and M. Benbouzid, "Blockchain and artificial intelligence as enablers of cyber security in the era of IoT and IIoT applications," *Journal of Sensor and Actuator Networks*, vol. 12, no. 3, p. 40, May 2023, doi: 10.3390/jsan12030040.
- [9] K. Chakraborty, D. Kapila, S. Kumar, Bhupati, N. Shaik, and A. Singh, "Intelligent machine learning based internet of things (IoT) resource allocation," in *RaiSE-2023*, Dec. 2023, p. 73. doi: 10.3390/engproc2023059073.
- [10] T. Lai, F. Farid, A. Bello, and F. Sabrina, "Ensemble learning based anomaly detection for IoT cybersecurity via Bayesian hyperparameters sensitivity analysis," *Cybersecurity*, vol. 7, no. 1, p. 44, Jun. 2024, doi: 10.1186/s42400-024-00238-4.
- [11] C. Liu *et al.*, "Dissecting zero trust: research landscape and its implementation in IoT," *Cybersecurity*, vol. 7, no. 1, p. 20, May 2024, doi: 10.1186/s42400-024-00212-0.
- [12] S. Mishra, A. Albarakati, and S. K. Sharma, "Cyber threat intelligence for IoT using machine learning," *Processes*, vol. 10, no. 12, p. 2673, 2022, doi: 10.3390/pr10122673.
- [13] M. E. E. Alahi *et al.*, "Integration of IoT-enabled technologies and artificial intelligence (AI) for smart city scenario: recent advancements and future trends," *Sensors*, vol. 23, no. 11, p. 5206, May 2023, doi: 10.3390/s23115206.
- [14] T. Narasimhamurthy and G. Hosahalli Swamy, "Insights of machine learning-based threat identification schemes in advanced network system," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 14, no. 4, p. 4664, Aug. 2024, doi: 10.11591/ijece.v14i4.pp4664-4674.

Neural-network based representation framework for adversary ... (Thanuja Narasimhamurthy)




- [15] D. Alsallman, "A comparative study of anomaly detection techniques for IoT security using adaptive machine learning for IoT threats," *IEEE Access*, vol. 12, pp. 14719–14730, 2024, doi: 10.1109/ACCESS.2024.3359033.
- [16] M. A. Elsadig, "Detection of denial-of-service attack in wireless sensor networks: a lightweight machine learning approach," *IEEE Access*, vol. 11, pp. 83537–83552, 2023, doi: 10.1109/ACCESS.2023.3303113.
- [17] U. Akram *et al.*, "IoTTPS: ensemble RKSVM model-based internet of things threat protection system," *Sensors*, vol. 23, no. 14, p. 6379, Jul. 2023, doi: 10.3390/s23146379.
- [18] S. Ben Atitallah, M. Driss, and I. Almomani, "A novel detection and multi-classification approach for IoT-malware using random forest voting of fine-tuning convolutional neural networks," *Sensors*, vol. 22, no. 11, p. 4302, Jun. 2022, doi: 10.3390/s22114302.
- [19] A. Hamarsheh, "An adaptive security framework for internet of things networks leveraging SDN and machine learning," *Applied Sciences*, vol. 14, no. 11, p. 4530, May 2024, doi: 10.3390/app14114530.
- [20] A. Alrefaei and M. Ilyas, "Using machine learning multiclass classification technique to detect IoT attacks in real time," *Sensors*, vol. 24, no. 14, p. 4516, Jul. 2024, doi: 10.3390/s24144516.
- [21] J. Li, M. S. Othman, H. Chen, and L. M. Yusuf, "Optimizing IoT intrusion detection system: feature selection versus feature extraction in machine learning," *Journal of Big Data*, vol. 11, no. 1, p. 36, Feb. 2024, doi: 10.1186/s40537-024-00892-y.
- [22] H. A. Gouda, M. A. Ahmed, and M. I. Roushdy, "Optimizing anomaly-based attack detection using classification machine learning," *Neural Computing and Applications*, vol. 36, no. 6, pp. 3239–3257, Feb. 2024, doi: 10.1007/s00521-023-09309-y.
- [23] Z. E. Ekolle, H. Ochiai, and R. Kohno, "Collabo: a collaborative machine learning model and its application to the security of heterogeneous medical data in an IoT network," *IEEE Access*, vol. 11, pp. 142663–142675, 2023, doi: 10.1109/ACCESS.2023.3341837.
- [24] M. Alqahtani, H. Mathkour, and M. M. Ben Ismail, "IoT Botnet attack detection based on optimized extreme gradient boosting and feature selection," *Sensors*, vol. 20, no. 21, p. 6336, Nov. 2020, doi: 10.3390/s20216336.
- [25] H. El-Sofany, S. A. El-Seoud, O. H. Karam, and B. Bouallegue, "Using machine learning algorithms to enhance IoT system security," *Scientific Reports*, vol. 14, no. 1, p. 12077, May 2024, doi: 10.1038/s41598-024-62861-y.
- [26] M. Salb *et al.*, "Enhancing internet of things network security using hybrid CNN and XGBoost model tuned via modified reptile search algorithm," *Applied Sciences*, vol. 13, no. 23, p. 12687, Nov. 2023, doi: 10.3390/app132312687.
- [27] J. Al Faysal *et al.*, "XGB-RF: a hybrid machine learning approach for IoT intrusion detection," *Telecom*, vol. 3, no. 1, pp. 52–69, Jan. 2022, doi: 10.3390/telecom3010003.
- [28] C. Ioannou and V. Vassiliou, "Network attack classification in IoT using support vector machines," *Journal of Sensor and Actuator Networks*, vol. 10, no. 3, p. 58, Aug. 2021, doi: 10.3390/jsan10030058.
- [29] K. M. Abuali, L. Nissirat, and A. Al-Samawi, "Advancing network security with AI: SVM-based deep learning for intrusion detection," *Sensors*, vol. 23, no. 21, p. 8959, Nov. 2023, doi: 10.3390/s23218959.
- [30] S. Dalal *et al.*, "Next-generation cyber attack prediction for IoT systems: leveraging multi-class SVM and optimized CHAID decision tree," *Journal of Cloud Computing*, vol. 12, no. 1, p. 137, Sep. 2023, doi: 10.1186/s13677-023-00517-4.
- [31] J. Jang, Y. An, D. Kim, and D. Choi, "Feature importance-based backdoor attack in NSL-KDD," *Electronics*, vol. 12, no. 24, p. 4953, Dec. 2023, doi: 10.3390/electronics12244953.
- [32] M. More, M. Idrissi, H. Mahmoud, and A. T. Asyhari, "Enhanced intrusion detection systems performance with UNSW-NB15 data analysis," *Algorithms*, vol. 17, no. 2, p. 64, Feb. 2024, doi: 10.3390/a17020064.

BIOGRAPHIES OF AUTHORS



Thanuja Narasimhamurthy    is Assistant Professor at Department of Computer Science and Engineering, Bangalore Institute of Technology, Bengaluru. Completed Graduation and Post-Graduation from Visvesvaraya Technological University in Information/Computer Science and Engineering Stream. Published papers in International Journals and Conferences in the domain machine learning and internet of things. Research focus is on networking systems, internet of things, and machine learning. She can be contacted at email: nthanuja@bit-bangalore.edu.in.



Gunavathi Hosahalli Swamy    is Associate Dean, Skill Development and Assistant Professor at Department of Computer Science and Engineering, Bangalore Institute of Technology, Bengaluru. She has published many research papers in international journals and international conferences. Actively involved in organizing events/conference, also organized many workshops/FDPs. Major areas of research interest are image processing, pattern recognition, artificial intelligence, machine learning, data science, and IoT. She can be contacted by gunavathihs@bit-bangalore.edu.in.