International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Create an Innovative Intrusion DetectionSystem for the Internet of Things by Improving Feature Weighting Through Heuristics

[1]Amol B. Gadewar, [2]Dr. Ritesh V. Patil, [3]Dr. Surendra A. Mahajan

Submitted: 04/12/2023 **Revised**: 14/01/2024 **Accepted**: 25/01/2024

Abstract: Recently, the "Internet of Things (IoT)" industry has developed as a tool for developing intelligent models of operation. Real-world applications that rely on the IoT system view privacy and security as major issues. Security issues in IoT-enabled devices pose obstacles to progress in the smart economy. As a result, "Intrusion Detection Systems (IDSs)" tailored to the IoT industries are desperately needed to curb the escalating number of attacks based on the Internet of Things. Because of their limited processing power, memory size, and battery life, traditional IDSs cannot be used in broad IoT-aided networks. Several IDSs have been proposed in academic publications as potential solutions to these issues. However, many IDSs run into problems with false positives and false negatives when looking for anomalies. In order to detect intrusion in the IoT industry and fix the problems with traditional systems, a deep learning ensemble model is suggested. In the first stage, we obtain the raw data from established sources.Consequently, the model is verified usingcomplementary metrics. The proposed approach, on the other hand, not onlyovercomes the greater detection rate, but also aids in avoiding intrusion from thirdparties.

Keyword: *Ensemble Networks, IoT Intrusion Detection Systems, and the Internet of Things.*

1. Introduction

The term "Internet of Things" (IoT)is used to describe the global network of electronic devices that can gather and share data through the use of built-in sensors, processors, and network connections. While there are many upsides to the Internet of Things, there are also some security concerns. Protecting Internet of Things (IoT) infrastructure from threats like hacking, data breaches, and other hostile actions relies heavily on intrusion detection systems. In order to identify security threats, intrusion detection systems (IDS) analyse data from all devices and traffic on a network. Due to resource limits, heterogeneity, and changing topologies in IoT networks, traditional IDS approaches generally fall short in IoT contexts. Due to its ability to automatically learn complicated patterns and features from massive volumes of data, deep learning has emerged as a potential solution for intrusion detection in IoT. In order to detect intrusions in the Internet of Things using deep learning, artificial neural networks are trained to examine data from sensors and other sources. Deep learning models can detect anomalous behavior, identify attempted intrusions, and categories threats by learning from past data. In conclusion, deep learning-based intrusion detection in the Internet of Things has considerable potential to strengthen IoT network and device security. As research and development in this area proceeds, it willlikely play an increasingly important role in safeguarding the security, confidentiality, and privacy of Internet of Things (IoT) infrastructure.

There are a number of obstacles specific to intrusion detection in the IoT that must be overcome to assure adequate safety.The limited processing power, memory, and energy resources of many IoT devices is just one of the major obstacles. It can be difficult to implement resource-intensive intrusion detection algorithms on such devices. To get beyond these restrictions, effective methods like slimmer models or distributed computing need to be created. The devices that make up an IoT network can be made by a number of different companies and run a wide variety of software and communication protocols. Due to the wide variety of devices and communication protocols, it is challenging to create a unified intrusion detection system. For successful intrusion detection across various IoT devices, it is necessary to solve

compatibility and interoperability difficulties. Adversarial assaults, in which an attacker manipulates input data to trick the detection model, can compromise deep learning-based intrusion detection systems. False negatives and false positives caused by malicious attacks weaken an IDS's ability to protect a network. It is a huge problem to create deep learning models that are both

[1]*Research Scholar, SKNCOE*
Vadgaon (Bk.), Pune amolgadewar22@gmail.com
[2]*Principal, PDEA'S College of Engineering Manjari (Bk.), Pune*
rvpatil3475@yahoo.com
[3]*Associate Professor, PVGCOET & GKPIOM,Pune*
sa_mahajan@yahoo.com

resilient and strong enough to withstand such attacks. In order to overcome these obstacles, researchers, IoT device manufacturers, network operators, and security experts will need to cooperate together. Detecting and preventing intrusions requires taking into account the particularities of IoT deployments and coming up with novel approaches, standards,and protocols.

To automatically learn and detect anomalies or malicious behaviors in IoT networks, researchers have turned to deep learning techniques for intrusion detection. The deep learning model must be regularly updated to account for the ever-changing nature of intrusion patterns and attack methods. The model's adaptability and effectiveness can be maintained through regular retraining with newly gathered data. To test how well the trained model performs, it is compared against an independent labelled data set that was not utilized during training. Measures of the model's efficacy such as accuracy, precision, recall, and F1-score are used in this analysis. The results of the analysis can then be used to refine and better implement the model. The model's capacity to track changing intrusion strategies depends on several factors, including the architecture selected, the availability of computational resources, and the regularity with which the model is monitored and updated.

2. Related Works

An effective Intrusion Detection System (IDS) has been offered as a possible solution by Bhawana Sharma et al. [1] forthe year 2023. In this research, we propose a novel anomaly-based intrusion detection system (IDS) for Internet of Things (IoT) networks that makes use of Deep Learning. In particular, we introduced a Deep Neural Network (DNN) model that filters out features that are highly associated. In addition, several parameters and hyper parameters were used to fine-tune the model. For this, we used the UNSW-NB15 dataset, which includes information on four different types of attacks. The proposed model was 84percent accurate. To fix the class imbalance problems in the dataset, we employed Generative Adversarial Networks (GANs) to generate minority attack synthetic data, and we got an accuracy of 91% with a balanced class dataset.

In 2022, Tanzila et al. [2] propose introducing a wide variety of smart, interconnected gadgets and apps across several fields in an effort to improve people's quality of life. However, the biggestissue for the devices in an IoT context was security concerns. Although numerous methods have been presented to secure IoT devices, there is room for improvement. Machine learning has proven its capacity to identify patterns even when other approaches fail. Using deep learning was an innovative approach to improving IoT security. This creates a smooth method for detecting anomalies. Taking advantage of the potential of the IoT, this study provides a CNN-based technique for anomaly-based intrusion detection systems (IDS), allowing features to quickly investigate complete traffic throughout the IoT. The proposed approach successfully identified anomalous traffic patterns and potential intrusions.

IDS utilizing a Short-Term Long Memory (LSTM) model to detect cyberattacks and justify its conclusions was published by Marwa et al. [3] in 2023. Specifically, a novel SPIP (S: Shapley Additive explanations, P: Permutation Feature Importance, I: Individual Conditional Expectation, P: Partial Dependence Plot) framework is used to extract a novel collection of input features for use in training and evaluating the LSTM model. The NSL-KDD, UNSW-NB15, and TON_IoT datasets were used to verify the framework. When compared to its contemporaries, the SPIP framework outperformed them in terms of detection accuracy, processing speed, and the interpretability of data characteristics and model outputs. Potentially helping administrators and decision-makers make sense of complex attack behavior is the proposed framework.

To create an effective intrusion detection system for IoT-cloud based systems, Mohamed et al. [4] in 2023 proposed combining the latest breakthroughs in swarm intelligence (SI) techniques with those in deep neural networks. To begin, ideal characteristics were extracted from the IoT IDS data using deep neural networks. Next, a method for effective feature selection was provided using the current SI optimizer, the Capuchin Search Algorithm (CapSA).In addition, we take into account thorough empirical comparisons to different optimization techniques utilizing various indices of classification performance. The results showed that the devised method had respectable overall dataset performance.

In 2022, Monika et al. [5] proposed the special characteristics of the limited devices; a previous approach was insufficient to safeguard the whole scope of IoT networks' safety. Attacks were detected and classified using an Intrusion Detection System (IDS) that looks for anomalies. Various security concerns can be addressed by applying machine learning (ML) and deep learning (DL) methods, which are adept at embedding intelligence in IoT devices and networks. In this work, we present a real-time intrusion detection system based on deep neural networks for the

identification of malicious packets. We have trained the model using recently developed benchmark NetFlow-based datasets. In order to detect attacks in real time, we have proposed a technique for capturing and analyzing packets. We also provided evidence for the reliability of our proposed model.

Using a single semi-supervised autoencoder and a threshold-setting technique, Marta et al. [6] in 2023 proposeda revolutionary intrusion detection system. The method was "outlier-aware," meaning it takes advantage of outlier detection to smooth out defects in the training data. Direct studies with normal and intrusion data points from separate sensing devices, an HTTP server, and four fully functional systems, including CPSs, were used to evaluate CPS-GUARD. Six state-of-the-art datasets were used as the basis for the experiments' simulated attacks. Depending on the system, CPS-GUARD's detection of intrusions yielded a recall of 0.949 to 1.000, precision of 0.961-0.999, and a false positive rate of 0.006-0.027. The outcomes held their own in comparison to alternative intrusion detection strategies. A comparison study of several threshold selection and outlier detection methods was also conducted to supplement the evaluation.

3. Problem Statement

Intrusion detection, the process of keeping an eye out for and reporting on suspicious or harmful actions taking place within a computer network or system, is an essential part of cyber security. The fundamental objective of intrusion detection is to identify security problems as soon as possible and take appropriate action tomitigate losses. Several existing technologies, such as DNN [1], canautomatically learn hierarchical representations of data through multiple layers of convolutional and pooling operations; this allows for high accuracy andgeneralization on large-scale datasets; however, training and deploying such a system requires substantial computational resources, including high-performance GPUs. However, ML [2] requires a lot of computational resources, such as powerful processors and high-performance GPUs, to process and analyses large volumes of data much faster than humans, allowing for faster decision-making and improved efficiency. While LSTM [3]can learn and remember relevant patterns over long periods of time,it requires a large amount of training data to generalize well and avoid overfitting. This is because it is designed to overcome the vanishing gradient problem in traditional RNNs. Inadequate or low-quality training data can hinder the LSTM model's performance and generalization capabilities. CNN [4] can learn spatial hierarchies, where lower layers capture low-level features, and it can recognize objects and patterns in images regardless of their location or position, but it is prone to overfitting, which is when the model does well on the training data but fails to generalize to new, unseen data. Furthermore, CNN [5] can effectively learn from massive datasets, capitalizing on the abundance of information to make accurate predictions or decisions, and it can learn end-to-end mappings from Cyber-Physical Systems [6] have improved healthcare, transportation, entertainment,and many other aspects of human life, bringing us greater ease, comfort, and well-being. However, their use can become habit forming and even addictive. When dealing with long sequences or huge networks with many layers and parameters, LSTM [7] can manage them by utilizing its memory cell and gating processes, but it can be computationally expensive. such as imageclassification, object detection, and image generation, but it needs a lot of labelled training data to [18] learn meaningful representations and achieve high accuracy. Particularly for specialized fields or niche applications, acquiring and annotating huge datasets can be a time-consuming and money-consuming process.

Table I: Features and challenges of Intrusion detection-IoT based Deep learning

Author [citation]	**Methodology**	**Features**	**Challenges**
Bhawana Sharma *et al.*[1]	DNN	• High accuracy and generalisation may be achieved on large-scale datasets, and the system can learn hierarchical representations of data automatically using many layers of convolutional and pooling algorithms.	• Training and deploying this model take a lot of computing power, including high-performance graphics processing units.

Tanzila *et al*.[2]	ML	• It is able to process and analyse vast amounts of data far more quickly than humans, which translates to quicker decision-making and higher productivity. • It can find connections and patterns in data that humans would miss.	• It calls for a lot of processing power, in the form of fast CPUs and graphics processing units.
Marwa *etal*[3]	LSTM	• It was created to fix the issue of vanishing gradients that plagues conventional RNNs. • To learn and recall useful patterns over time, the network is given the ability to selectively retain information across long sequences.	• To generalise effectively and prevent overfitting, a substantial amount of training data is necessary. Inadequate or low-quality training data can hinder theLSTM model's performance and generalisation capabilities.
Mohame d *et al*.[4]	CNN	• It is capable of learning spatial hierarchies, in which lower layers record low-level data, and it can recognise objects and patterns in images.	• When working with huge input photos or deep network structures, it mightbe computationally intensive.

4. Research Methodology

Recent advances in the Internet of Things (IoT) concept have made it possible to use it to create intelligent surroundings. For any IoT-based smart environment in the real world, security and privacy are critical concerns. Applications designed for use in smart environments are vulnerable to security risks caused by flaws in IoT-based systems. Because of this, intrusion detection systems (IDSs) tailored to IoT contexts are essential for protecting against IoT-related security assaults. The battery life, memory capacity [19]. While IDSs are essential, the majority of them struggle with a high false alarm rate and a lack of accuracy when it comes to detecting anomalies [20]. An ensemble deep learning model is presented for intrusion detection in the IoT sector to address the problems seen in current approaches. To begin, we'll assemble the foundational data from industry-standard sources. T-distributed stochastic neighbor embedding (TSNE), principal component analysis (PCA), and statistical features will then be used for subsequent feature extraction [24]. Next, the output features will be fused using weights that are optimized with an enhanced version of the Pelican Optimization Algorithm (POA). Ensemble networks, comprised of a Long Short-Term Memory (LSTM), a Recurrent Neural Network (RNN), a Deep Markov Random Field (DMRF), and a Ridge classifier, will next be applied to the combined features. Then, we'll use a fusion method based on the average of the scores to determine the final detected outcome [21]. As a result, we employ mutually exclusive metrics in our model's verification. In contrast, the proposed approach not only has a greater detection rate but also aids in preventing malicious outsiders from breaking into the network.

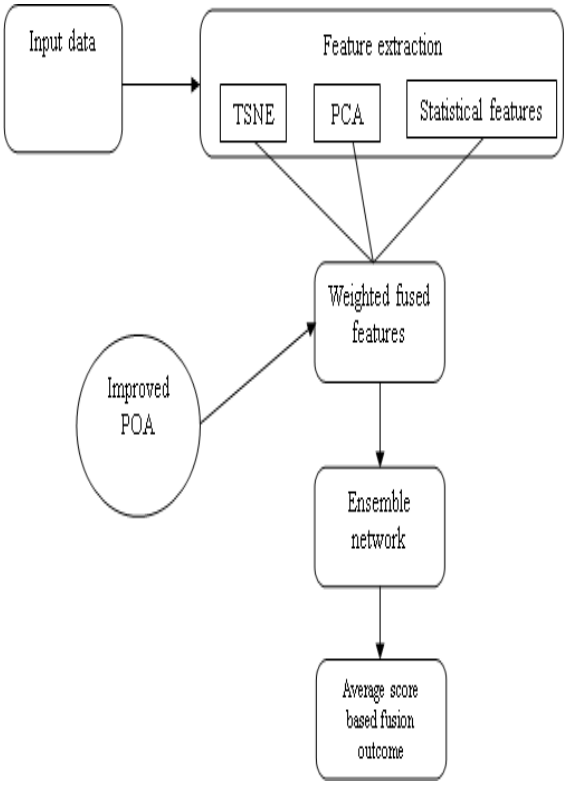

Fig 1: Illustration of the proposed Internet ofThings-based ensemble model for intrusion detection.

4.1 System Model of IoT

The advancements and benefits to different sectors, including as smart environment, automation, health care, the industrial process, and so on, have helped the IoT garner significant attention in recent years. With the internet and the smart object's convergence influencing the composition between the cyber and physical sectors, the IoT links together previously separate developments. While each IoT advancement serves a unique purpose [22], they all have commonalities. There are a few distinct phases to the IoT, including data collecting, data transfer, and data processing. The primary goal of the gathering phase is to amass information from the real world. In order to accomplish this, we are encouraging the use of sensing devices and short-range communication technology. In this phase, data gathered during the gathering phase is relayed to potential employees.

At this point, the candidates and objects are linked together across great distances via a network built from a combination of technologies and protocols, including WiFi and Ethernet. The final step is processing, when the data is put to use by the applications [20]. The data-driven decisions must be made within the apps. Several safety measures for the Internet of Things are described below.

- Data confidentiality: Data changes cause major issues, and protecting people's privacy in their personal and professional lives is crucial. Therefore, data privacy is crucial in the Internet of Things industry.
- Data integrity: It's an essential part. In the Internet of Things (IoT) industry, verifying data resources and identifying fraudulent assaults are crucial.
- Authentication: In order to provide data access and authorization, IoT requires a potential confirmation operation. The harmony between the data and the IoT infrastructure will be improved as a result.

It is critical to improve the IoT's security system in light of these security considerations. Network problems can be reduced if threats are identified in a timely fashion.

5. Results and Discussions

Simulation setup

The recommended IDS work in the IoT industry has been successfully implemented on the Python platform. Overall chromosomal length was 48, with 50 iterations being the maximum advised for the IDS challenge. For the planned IDS experiment, there were a total of 10 participants. For the proposed IDS task validation [21], classifiers such LSTM, RNN, Deep MRF, and ridge classifier were also used.

The overall comparative validation of the designed

IDS over conventional classifiers

Tables II illustrates the comparative examination of the implemented IDS over diverse algorithms and classifiers. This extensive solution showed that the suggested IDS has betterfunctionality.

Table II: The suggested IDS was validated against a number of different traditional classifiers for test data.

Terms	LSTM	RNN	Deep MRF	Ridge classifier	EnsNet
			"Dataset-1"		
"Accuracy"	88.8	91.9	90	93.63	94.73
"Recall"	88.91	91.98	89.96	93.65	94.77
"Specificity"	88.69	91.82	90.03	93.61	94.69
"Precision"	87.81	91.16	89.21	93.07	94.24
"F1-Score"	11.08	8.01	10.03	6.34	5.22

We confirmed that our IDS outperformedconventional classifiers on a variety ofdatasets.

The higher performance of the produced IDS on all three datasets was verified using industry-standard classifiers (Fig.02). Therefore, the created IDS'simproved functionality has been confirmed.

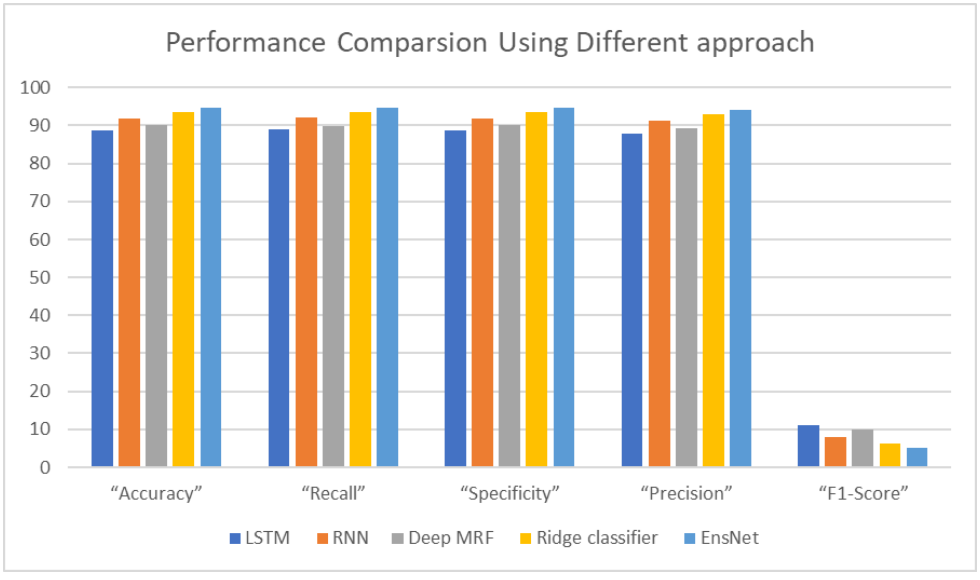

Fig 02: The performance examination of the suggested IDS for the first dataset overvarious classifiers

6. Conclusion:

The IoT industry has developed a state- of-the-art IDS task using ensemble deep learning algorithms to solve the drawbacks of traditional approaches. The original data was collected from the standard methods. The TSNE, PCA, and statistical features were all created using the feature extraction technique. In addition, an average-based fusion technique was used to produce the outcome. Therefore, the model was validated using a wide variety of criteria. It was established that the IoT industry's planned IDS obligation increased detection rates and helped block unauthorized access to the network.

References

[1] Bhawana Sharma, Lokesh Sharma, Chhagan Lal, Satyabrata Roy"Anomaly based network intrusion detection for IoT attacks using deep learning technique"Computers and Electrical Engineering, Vol. 107, pp.108626,April 2023.

[2] Tanzila Saba, Amjad Rehman, Tariq Sadad, Hoshang Kolivand, Saeed Ali Bahaj" Anomaly-based intrusion detection system for IoT networks

through deep learning model"Computers and Electrical Engineering, Vol. 99,pp.107810,April 2022.

[3] Marwa Keshk, Nickolaos Koroniotis, Nam Pham, Nour Moustafa, Benjamin Turnbull, Albert Y. Zomaya "An explainable deep learning-enabled intrusion detectionframework in IoT networks"Information Sciences Vol. 639, pp. 119000,August 2023.

[4] Mohamed Abd Elaziz, Mohammed A.A. Al-qaness , Abdelghani Dahou , Rehab Ali Ibrahim , Ahmed A. Abd El-Latif "Intrusion detection approach for cloud and IoT environments using deep learning and Capuchin Search Algorithm"Advances in Engineering Software, Vol.176, pp. 103402,February 2023.

[5] Monika Vishwakarma, Nishtha Kesswani"DIDS: A Deep Neural Network based real-time Intrusion detection system for IoT"Decision Analytics Journal, Vol. 5, pp.100142,December 2022.

[6] Marta Catillo, Antonio Pecchia, Umberto Villano "CPS-GUARD: Intrusion detection for cyber-physical systems and IoT devices using outlier-aware deep autoencoders"Computers & Security,Vol. 129, pp.103210,June 2023.

[7] Rania A. Elsayed, Reem A. Hamada, Mahmoud I. Abdalla, Shaimaa Ahmed Elsaid"Securing IoT and SDN systems using deep-learning based automatic intrusion detection"Ain Shams Engineering Journal, Vol. 14, Issue 10, pp. 102211,October 2023.

[8] Bhukya Madhu, M. Venu Gopala Char, Ramdas Vankdothu, Arun Kumar Silivery, Veerender Aerranagula"Intrusion detection models for IOT networks via deep learning approaches"Measurement: Sensors Vol. 25, pp.100641,February2023,

[9] M. Eskandari, Z. H. Janjua, M. Vecchio and F. Antonelli, "Passban IDS: An Intelligent Anomaly-Based Intrusion Detection System for IoT Edge Devices," IEEE Internet of Things Journal, vol. 7, no. 8, pp. 6882-6897, Aug. 2020.

[10] J. Gao et al., "Omni SCADA Intrusion Detection Using Deep Learning Algorithms," in IEEE Internet of Things Journal, vol. 8, no. 2, pp. 951-961, 15 Jan.15, 2021.

[11] R. Mills, A. K. Marnerides, M. Broadbent and N. Race, "Practical Intrusion Detection of Emerging Threats, "IEEE Transactions on Network and Service Management, vol. 19, no. 1, pp. 582-600, March 2022.

[12] M. Mohy-Eddine, A. Guezzaz, S. Benkirane, M. Azrour and Y. Farhaoui, "An Ensemble Learning Based Intrusion Detection Model for Industrial IoT Security," in Big Data Mining and Analytics, vol. 6, no. 3, pp. 273-287, September2023.

[13] M. Zhou, Y. Li, L. Xie and W. Nie, "Maximum Mean Discrepancy Minimization Based Transfer Learning for Indoor WLAN Personnel Intrusion Detection," IEEE Sensors Letters, vol. 3, no. 8, pp. 1-4,Art no. 7500804,Aug. 2019.

[14] M. A. Siddiqi and W. Pak, "Tier-Based Optimization for Synthesized Network Intrusion Detection System," IEEE Access, vol. 10, pp. 108530-108544, 2022.

[15] L. Wang, J. Yang, M. Workman and P. Wan, "Effective algorithms to detect stepping-stone intrusion by removing outliers of packet RTTs," in Tsinghua Science and Technology, vol. 27, no. 2, pp. 432-442, April 2022.

[16] Z. Shi, S. He, J. Sun, T. Chen, J. Chen and H. Dong, "An Efficient Multi-Task Network for Pedestrian Intrusion Detection," IEEE Transactions on Intelligent Vehicles, vol. 8, no. 1, pp. 649-660, Jan. 2023.

[17] Kumbhare, S. , B.Kathole, A. , Shinde, S., “Federated learning aided breast cancer detection with intelligent Heuristic-based deep learning framework”, Biomedical Signal Processing and Control Volume 86, Part A, September 2023, 105080

[18] T. Yu and X. Wang, "Topology Verification Enabled Intrusion Detection for In-Vehicle CAN-FD Networks,"IEEE Communications Letters, vol. 24, no. 1, pp. 227-230, Jan. 2020.

[19] Atul Kathole , Dinesh Chaudhari “Secure Hybrid Approach for Sharing Data Securely in VANET”, Proceeding of International Conference on Computational Science and Applications pp 217–221, © 2022 The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd.

[20] Atul Kathole , Dinesh Chaudhari “Securing the Adhoc Network Data Using Hybrid Malicious Node Detection Approach”, Proceedings of the International Conference on Intelligent Vision and Computing (ICIVC 2021) pp 447–457 © 2022 The

Author(s), under exclusive license to Springer Nature Singapore Pte Ltd.

[21] S. Chundong, M. Yaqi, J. Luting, F. Ligang and K. Baohua, "Intrusion-Detection Model Integrating Anomaly with Misuse for Space Information Network," in Journal of Communications and Information Networks, vol. 1, no. 3, pp. 90-96, Oct. 2016.

[22] Atul Kathole , Dinesh Chaudhari "Securing the Adhoc Network Data Using Hybrid Malicious Node Detection Approach", Proceedings of the International Conference on Intelligent Vision and Computing (ICIVC 2021) pp 447–457 © 2022 The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd.

[23] Y. Zha and J. Li, "CMA: A Reconfigurable Complex Matching Accelerator for Wire-Speed Network Intrusion Detection," IEEE Computer Architecture Letters, vol. 17, no. 1, pp. 33-36, 1 Jan.-June 2018.

[24] Atul B Kathole, Dr.Dinesh N.Chaudhari, "Pros & Cons of Machine learning and Security Methods, "2019.http://gujaratresearchsociety.in/index.php/ JGRS, ISSN: 0374-8588, Volume 21 Issue 4.

[25] M. B. Gorzalczany and F. Rudzinski, "Intrusion Detection in Internet of Things With MQTT Protocol—An Accurate and Interpretable Genetic-Fuzzy Rule- Based Solution," in IEEE Internet of Things Journal, vol. 9, no. 24, pp. 24843-24855, 15 Dec.15, 2022.

[26] G. Abdelmoumin, D. B. Rawat and A. Rahman, "On the Performance of Machine Learning Models for Anomaly-Based Intelligent Intrusion Detection Systems for the Internet of Things," in IEEE Internet of Things Journal, vol. 9, no. 6, pp. 4280-4290, 15 March15, 2022.