

Attack Classification and Detection for Misbehaving Vehicles using ML/DL

Saleha Saudagar¹, Dr. Rekha Ranawat²

¹Computer Science and Engineering
SAGE University, Indore
India

salehasaudagar@gmail.com

²Computer Science and Engineering
SAGE University, Indore
India

rekharathod23@gmail.com

Abstract— Vehicle ad hoc networks are a crucial component of the next Intelligent Transportation System created to build a reliable and secure connection between various network components to establish a safe and effective transportation network. Because of open nature of VANETs become vulnerable to numerous assaults such as forgery, Denial-of-Service (DoS), and false reports, which can ultimately cause traffic jams or accidents. The earlier study concentrated on misbehaving vehicles rather than RSUs. Proposed method integrates data from two subsequent BSMs for testing and training by employing machine learning (ML) methods. The framework merges the data from two BSMs in the right manner and utilizes machine learning/Deep learning methodology which identify the running vehicle as a legal or hostile one.

Keywords- Vehicular ad hoc network(VANET), Smart transportation System(STS), Bi-GRU, Intrusion detection system(IDS), Misbehaviour Detection System(MVDS), Machine Learning, Deep Learning.

I. INTRODUCTION

In today's scenario road accidents are major reason of rising mortality rate for the people of age from 5 to 30 years. As per the survey of World Health Organization's (WHO), the report generated in 2018 Global Status Report on Road Safety [1]. Vehicular communication [2] are a crucial component of the futuristic Smart Transportation System [3] created to build a reliable and secure connection between various network components to establish a safe and effective transportation network. Vehicle to vehicle and vehicle to infrastructure communication are made possible via the Vehicular infrastructure (also known as VANET), a form of Mobile ad hoc network (MANET) [4]. Vehicle communication has become more prevalent thanks to telecommunications advancements such as the use of high-definition mapping, intelligent transportation systems, autonomous driving, and coordinated driving [5]. While doing so, it also draws attention to the key traits of VANETs, such as self-operated, decentralized infrastructure where each vehicle coming and leaving the range are responsible for communication, and one of the most dynamic topologies [6]. When compared to traditional networks, these communication features provide significant problems and distinctions in terms of security and safety requirements [7, 8].

Exchange of messages are mainly navigation messages, and event-oriented communications, are hardly ever encrypted in automotive communication networks [9]. Because of this, the open nature of VANETs makes them vulnerable to numerous assaults such as forgery, Denial-of-Service (DoS), and false reports, which can ultimately cause traffic jams or accidents [10-12]. Drivers are really put at risk since malicious groups may also monitor participants' messages and identity. Therefore, malevolent vehicles for VANETs should be tracked down and punished in the case of any misbehavior [13, 14] from the perspective of maintaining security. Traditional prediction-based protocols have been developed for a variety of specialized applications, including routing, traffic control, safety, and others. To improve performance, Machine Learning (ML) technique for further data analytics, has been encouraged. By examining the data flow in the VANET system, various machine learning algorithms reveal these issues. At various VANET components, numerous types of data are produced. Road Side installed Devices, Various Vehicles, and the Central communication authority, all have access to data such as the specifics of nearby vehicles in the range, routing information, congestion-related information, weather-related information, and any information related with communication among vehicular infrastructure. Due to the dynamic nature of VANET, this tremendous volume of data is constantly generated. Various

Machine learning methodology can be chained with Vehicular technology which yields efficient results and help in improved road trip and can avoid accidents [15]. Machine learning is a clever technology that can handle these tasks well, as shown in a very broad range of applications [16]. The real-world characteristics of VANET's technological and societal features make it susceptible to hackers. Intra-vehicular and Inter-vehicular attacks on vehicular systems can both occur. Intra-vehicular attacks concentrate on connecting equipment inside a vehicle, as opposed to inter-vehicle assaults, which try to disrupt communication between vehicles and infrastructure [17].

II. LITERATURE SURVEY

Several researches had carried out for misbehavior detection in VANET using machine learning.

Some of the researches have been analyzed in the following literature survey. Hind Bangui et al. [25] suggested a new machine learning model that uses Random Forest and posterior detection based on coresets to boost detection efficiency and improve detection accuracy in order to enhance IDS performance. The model uses an unsupervised clustering approach based on coresets to filter out unknown attacks and has incorporated the random forest as a classifier to identify well-known attacks.

Aekta Sharma and Arunita Jaekel [18] proposed a brand-new machine learning-based method for categorizing position falsification attacks in VANET. By implementing the misbehavior detection mechanism in the RSUs, which can communicate this information broadly with other RSUs and cars, the suggested technique shifts the computational burden from vehicles (OBUs). The VeReMi dataset, which only includes five particular attack types and does not include all position falsification assaults that could be used in VANETs, was used to train the proposed models in this study. To operate a VANET securely, it is necessary to create robust models that can identify unexpected threats as well as inaccurate data in other BSM characteristics (such as speed, acceleration, heading, etc.).

Agria Rhamdhan and Fadhil Hidayat [19] suggested a defense mechanism to address the concerns with Sybil attack detection in VANET related to accuracy, privacy, safety, and real-world application. A hybrid system and a trust-based approach are the foundation of the defense mechanism. A reputation system based on machine learning is used by RSU to offer central trust, and message exchange is used to manage neighbor trust in a data-centric manner. They intended to put the suggested protection mechanism into practice and assess its effectiveness and accuracy in identifying the Sybil assault as future work.

Kumar Sharshembiev et al., [20] proposed the detection of protocol misbehavior is discussed utilizing cutting-edge machine learning frameworks and entropy. In order to detect misbehavior in VANETs, they investigated the use of opportunistic selective sampling and entropy. A realistic WAVE environment simulation was used to show the effects of selective sampling versus packet sampling. We anticipate the model and metrics will advance, allowing a more precise classification of broadcast misbehavior flows when the data collected grows larger and approaches some of the other image and text machine learning datasets in scale. To accomplish unsupervised learning or possibly an active learning framework by having the driver provide feedback when the broadcast misbehavior is noticed, would like to employ the same Tensor Flow framework in future work and are interested in learning more about transmission and computational latency, a crucial factor in VANET safety applications.

Pranav Kumar Singh et al., [21] utilized various machine learning techniques in the research to identify position falsification attacks in VANETs. SVM with normalization outperformed logistic regression with or without normalization in terms of performance. The choice of features has a significant impact on model correctness. Multiple misbehavior modelling in VANETs and detection using an ML-based technique can both be the subject of future research. Artificial neural networks can also be used to evaluate the performance on the dataset, even though traditional approaches perform well.

Ayoub Alsarhan et al. [22] presented a unique method for reducing the percentage of VANET packet transmissions that are invalid. The suggested system utilizes information from both past and present behavior to assess the reliability of both data and nodes in order to detect unexpected traffic. Four stages—a rule-based security filter, a Dempster-Shafer adder, a node's history database, and a Bayesian learner—are used to implement a new intrusion detection technique. They intended to expand the suggested model in the near future to make use of massive data gathered from actual systems.

Abhilash Sonker and R. K. Gupta [23] proposed that VANETs have drawn a lot of attention since they have significantly improved driving conditions and road safety. The misbehavior in VANETs can be found to determine whether or not a node is malevolent. In the study, five distinct algorithms are used to identify the five attacks, and the accuracy of each approach is assessed independently. The best method that may be used on the combined dataset is used in a novel procedure for the multiple detection of the attacks. This new method can also be used as a generic idea or methodology for the detection of malicious nodes. By selecting the optimum method, this strategy is ideal for the detection of improper behavior in VANETs. The use of hybrid machine learning approaches can

be used to advance the work in the study. The deployment of various situations and assaults may also be taken into consideration in the future for the purpose of detecting inappropriate behavior.

Heena Khanna and Manmohan Sharma [24] proposed in this paper, a better security algorithm for VANET. This algorithm can handle attacks like DoS, Sybil, and Replay. The clusters for the different attacks are created using the Enhanced K-Mean method in the proposed work, and the classifier's accuracy is tested using a hybrid strategy combining Support Vector Machine (SVM) and Feed-forward back propagation. A method that addresses more than only DoS, Sybil, and Replay assaults is one of the future works that is planned to be concentrated.

Based on the aforementioned studies, some of the problems are studied and that are stated below. The VeReMi dataset, only includes five particular attack types and in order to operate a VANET securely, it is necessary to create robust models that can identify unexpected threats as well as inaccurate data in other BSM characteristics (such as speed, acceleration, heading, etc.). The detection coverage of the detection system should be further expanded by including more attacks in the dataset. Additionally, the dataset should be parsed to improve the label accuracy. Multiple misbehavior modelling in VANETs and detection using an ML-based technique can both be the subject of future research. Artificial neural networks can also be used to evaluate the performance on the dataset, even though traditional approaches perform well. The use of hybrid machine learning approaches can also be used to advance the work. The deployment of various situations and assaults may also be taken into consideration in the future for the purpose of detecting inappropriate behavior. Hence a more efficient method should also be suggested, in order to address the drawbacks.

A hybrid machine learning technique was put up by Bangui et al. [26] to easily carry out thorough intrusion detection in VANET. The suggested approach combines coresets-based clustering and data categorization. It makes use of coresets to reduce overhead in computational time consumption and improve IDSs' inference capabilities in VANET. It is still having trouble getting excellent detection accuracy, though.

In order to build a shared trust value for each vehicle on the network, Shams et al. [27] presented a full IDS in VANET using the combination of modified promiscuous mode for data collecting and Support Vector Machine (SVM) for data analysis (TSIDS). This procedure makes sure that the source vehicle or node, as well as any intermediary network nodes, are aware of the activities of their next hop and will react appropriately to maintain the highest possible network performance in the event of malicious conduct or breakdown. However, it is necessary to

enhance reliability indicators like Packet delivery ratio and End to End Delay.

III. MOTIVATION

VANET communication is susceptible to a number of threats, therefore message integrity and vehicle authentication frequently require cryptographic approaches. However, using only cryptographic methods might not be enough to find off insider attacks. The earlier work relay on the routine transmission of Basic Safety Messages (BSMs) from nearby vehicles, which carry crucial status details about a vehicle. The method integrates data from two subsequent BSMs for testing and training by employing machine learning (ML) methods. The framework merges the data from two BSMs in the right manner and utilizes machine learning/Deep learning methodologies to classify the running vehicular nodes as a legal or hostile one.

The earlier study concentrated on misbehaving vehicles rather than RSUs (which serves as the network center in a VANET, and monitors it for possible threats). Attacks in RSU is not concentrated. One of VANET's most critical and difficult goals is network availability which had not been concentrated in the work i.e., threatening the RSU availability. Hence, an improved model should be proposed to overcome these limitations.

IV. RESULTS AND DISCUSSION

TABLE I. CLASSIFICATION RESULTS OF PROPOSED MODEL FOR LOW DENSITY

Algorithm Used	Accuracy	Precision	Recall	F1 Score
SVM	100	99.88	100	100
KNN	99.99	99.89	100	99.80

TABLE II. CLASSIFICATION RESULTS OF PROPOSED MODEL FOR LOW DENSITY

Algorithm Used	Accuracy	Precision	Recall	F1 Score
SVM	99.99	99.98	100	99.97
KNN	100	100	99.70	99.98

V. PROPOSED METHODOLOGY

Some of the challenges experienced by the previous work has been given above. In order to overcome the above mentioned challenges an improved Intrusion Detection System (IDS) that runs at RSU, has to be proposed. For this persistence we have

proposed, Model for Attack Detection and Classification. We specifically focus on detecting and classifying position falsification attack, low rate as well as high rate DoS/DDos and Sybil attacks on RSUs. Our proposed model does not specify any data type and thus it can be used for any type of data.

A. Implementation steps

A new model framework is proposed for preprocessing, feature selection, detection and classification of the attack.

- Firstly, the dataset is cleaned or pre-processed by the following steps – Filtering out inconsistent values (Outliers), One hot encoding etc. The inconsistent values (outliers) can affect the learning, leading to missed intrusion detection so that the inconsistent values can be filtered out by using the Median Absolute Deviation Estimator.
- Next the one-hot encoding technique was used to convert the different categorical features such as protocol type, service, and flag into numerical values. Through the above steps the data has been balanced.
- Next the key features are chosen by Bi directional GRU.
- An Intellectual-Focused mechanism is consequently included and different weights are assigned to the features through max-pooling and average-pooling. It is used to preserve key characteristics and increase the model's robustness. The results of two pools are combined into one.
- The next process is testing, which involves feeding the test set into the trained model for detection and categorization of the attacks by using ResCaps.
- The great dependence on hyper-parameters, however, is a significant problem for deep learning models. Under various hyper-parameter configurations, it may change drastically. The Orthogonal Array Tuning Method (OATM) which is based on a design matrix and allowing the user to take into consideration a specified subset of combinations of various parameters at several levels, is presented in this research as a solution to the aforementioned problem.
- In order to ensure that all potential values for each hyper-parameter are taken into account equally, the OATM is balanced as well.
- The OATM can thus achieve a balance between minimal tuning time and competitive performance. Hence the proposed Framework achieve the high accuracy to detect various classes.

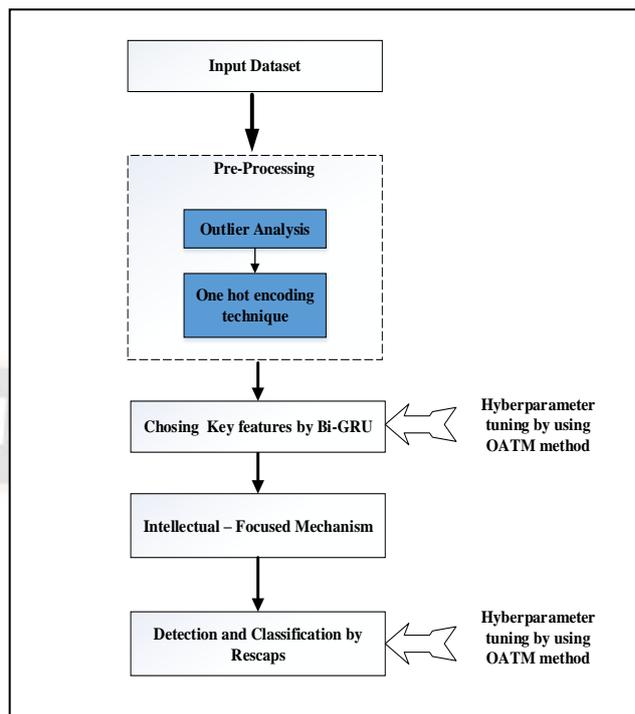


Figure 1. Flow Diagram of our proposed work, Hybrid BI-GRU and ResCaps: An Amalgamated Novel IDS Model for Attack Detection and Classification.

B. Dataset Used(VeReMi)

In order to the proposed approach and to do unbiased comparison, we decided to use well known VEREMI dataset which is labeled dataset and contains 225 simulations and 5 attacker types, this dataset has wide range of attacker densities among that we are considering high range and low range where high range density scenario simulates from 108 and 519 vehicles and generates 21 thousand individual messages. For low vehicle density, the simulation consists of 35-39 vehicles, which generate more than 1100 individual BSMs.

Talking about the position falsification attack in VEREMI dataset, there are total five attacks are given with attack ID 1,2,4,8,16, this dataset is uneven dataset where log files are having attacks as well as normal messages. both GPS and BSM messages are included in log files which gives ground truth. GPS messages gives the position related information while BSM messages relates to the information received in V2V communication through Dedicated Short-Range Communication (DSRC).

For classification problems Precision, Recall and F1 score are considered one of the established performance calculators. Where True positive (TP), True Negative (TN), False positive (FP) and False negative (FN) will be classified for attack type 1 and 2. Highest Value 1 indicate attack done and Lowest value 0 indicate normal log.

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

$$F1\text{-score} = 2 * \frac{Precision * Recall}{Precision + Recall}$$

Figure 2. Attack detection and evaluation metrics

C. Classifier used

Supervised classification algorithm such as KNN and SVM which are having high accuracy in binary classification are used here.

D. Attack Detection Framework

The vehicular ad hoc network is decentralized architecture which is connected with central authority forming hybrid communication. Here central Authority update the shared database in periodic time. Below diagram shows V2X communication which include vehicle to Vehicle and Vehicle to RSU communication and Central authority interaction is also shown here. The diagram explained about attack detection framework.

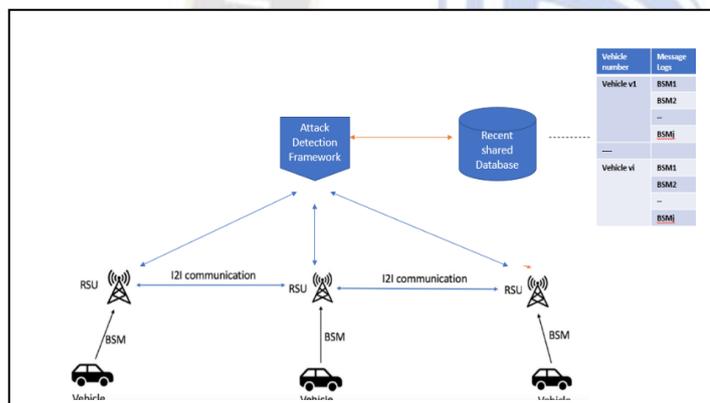


Figure 3. Shows Attack Detection framework of Vehicular Ad Hoc Network

CONCLUSION AND FUTURE PLAN

In this paper, IDS approach to detect position falsification attack in VEREMI dataset is being evaluated through KNN and SVM classifier which shows improved result. Here Proposed scheme consider only 2 attack type which can be extended to all position falsification attack of the dataset. We can further extend the work by using our own dataset and Hybrid Bi-GRU and OATM method will be impactfully removes outlier which molds the work into Deep learning approach.

REFERENCES

[1] "Global status report on road safety 2018," Accessed: Jan. 5, 2021.[Online].Available:<http://apps.who.int/iris/bitstream/handle/10665/277370/WHO-NMH-NVI-18.20-eng.pdf?ua=1>Re

[2] S. Al-Sultan, M. M. Al-Doori, A. H. Al-Bayatti, and H. Zedan, "A comprehensive survey on vehicular Ad Hoc network," J. Netw. Comput.Appl., vol. 37, pp. 380–392, 2014.

[3] S.-h. An, B.-H. Lee, and D.-R. Shin, "A survey of intelligent transportation systems," in Proc. 3rd Int. Conf. Comput. Intell., Commun. Syst.Netw., 2011, pp. 332–337.

[4] M. N. Mejri, J. Ben-Othman, and M. Hamdi, "Survey on VANET security challenges and possible cryptographic solutions," Vehicular Communications, vol. 1, no. 2, pp. 53–66, 2014.

[5] Yasam, S. ., H. Nair, S. A. ., & Kumar, K. S. . (2023). Machine Learning based Robust Model for Seed Germination Detection and Classification . International Journal of Intelligent Systems and Applications in Engineering, 11(2s), 116–124. Retrieved from <https://ijisae.org/index.php/IJISAE/article/view/2515>

[6] S. Gyawali, S. Xu, Y. Qian, and R. Q. Hu, "Challenges and Solutions for Cellular Based V2X Communications," IEEE Communications Surveys & Tutorials, vol. 23, 2020.

[7] Mitchell, Robert, and Ing-Ray Chen. "A survey of intrusion detection techniques for cyber-physical systems." ACM Computing Surveys (CSUR) 46, no. 4 (2014): 1-29.

[8] S. S. Tangade and S. S. Manvi, "A survey on attacks, security and trust management solutions in VANETs," in Proceedings of the 2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT), pp. 1–6, Tiruchengode, India, July 2013.

[9] M. S. Sheikh, J. Liang, and W. Wang, "Security and privacy in vehicular ad hoc network and vehicle cloud computing: a survey," Wireless Communications and Mobile Computing, vol. 2020, Article ID 5129620, 2020.

[10] S. Gyawali and Y. Qian, "Misbehavior detection using machine learning in vehicular communication networks," in Proceedings of the ICC 2019-2019 IEEE International Conference on Communications (ICC), pp. 1–6, Shanghai, China, May 2019.

[11] Z. A. Biron, S. Dey, and P. Pisu, "Real-time detection and estimation of denial of service attack in connected vehicle systems," IEEE Transactions on Intelligent Transportation Systems, vol. 19, no. 12, pp. 3893–3902, 2018.

[12] Z. Yang, K. Zhang, L. Lei, and K. Zheng, "A novel classifier exploiting mobility behaviors for Sybil detection in connected vehicle systems," IEEE Internet of Things Journal, vol. 6, no. 2, pp. 2626–2636, 2018.

[13] D. He, S. Zeadally, B. Xu, and X. Huang, "An efficient identity-based conditional privacy-preserving authentication scheme for vehicular ad hoc networks," IEEE Transactions on Information Forensics and Security, vol. 10, no. 12, pp. 2681–2691, 2015.

[14] F. Qu, Z. Wu, F.-Y. Wang, and W. Cho, "A security and privacy review of VANETs," IEEE Transactions on Intelligent Transportation Systems, vol. 16, no. 6, pp. 2985–2996, 2015.

[15] A. K. Malhi, S. Batra, and H. S. Pannu, "Security of vehicular ad-hoc networks: a comprehensive survey," Computers & Security, vol. 89, Article ID 101664, 2020.

[16] Khatri, Sahil; Vachhani, Hrishikesh; Shah, Shalin; Bhatia, Jitendra; Chaturvedi, Manish; Tanwar, Sudeep; Kumar, Neeraj (2020). Machine learning models and techniques for VANET based traffic management: Implementation issues and challenges. Peer-to-Peer Networking and Applications, (), – . doi:10.1007/s12083-020-00993-4.

- [17] Ms.Nivedita Kadam and Dr. Krovi Raja Sekhar. Machine Learning Approach of Hybrid KSVN Algorithm to Detect DDoS Attack in VANET. (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 12, No. 7, 2021.
- [18] Ammar Haydari and Yasin Yilmaz. Real-Time Detection and Mitigation of DDoS Attacks in Intelligent Transportation Systems. 2018 21st International Conference on Intelligent Transportation Systems (ITSC) Maui, Hawaii, USA, November 4-7, 2018.
- [19] A. Sharma and A. Jaekel, "Machine Learning Based Misbehaviour Detection in VANET Using Consecutive BSM Approach," in IEEE Open Journal of Vehicular Technology, vol. 3, pp. 1-14, 2022, doi: 10.1109/OJVT.2021.3138354.
- [20] Agria Rhamdhan , Fadhil Hidayat. Hybrid Trust-based Defense Mechanisms Against Sybil Attack in Vehicular Ad-hoc Networks. MSCEIS 2019, October 12, Bandung, Indonesia Copyright © 2020 EAI DOI 10.4108/eai.12-10-2019.2296524.
- [21] Sharshembiev, K., Yoo, S.-M., & Elmahdi, E. (2021). Protocol misbehavior detection framework using machine learning classification in vehicular Ad Hoc networks. *Wireless Networks*, 27(3), 2103–2118. doi:10.1007/s11276-021-02565-7.
- [22] Singh, P. K., Gupta, S., Vashistha, R., Nandi, S. K., & Nandi, S. (2019). Machine Learning Based Approach to Detect Position Falsification Attack in VANETs. *Security and Privacy*, 166–178. doi:10.1007/978-981-13-7561-3_13.
- [23] Alsarhan, Ayoub; Al-Ghuwairi, Abdel-Rahman; Almalkawi, Islam T.; Alauthman, Mohammad; Al-Dubai, Ahmed (2020). Machine Learning-Driven Optimization for Intrusion Detection in Smart Vehicular Networks. *Wireless Personal Communications*, (), -. doi:10.1007/s11277-020-07797-y.
- [24] Abhilash Sonker, R. K. Gupta. A new procedure for misbehavior detection in vehicular ad-hoc networks using machine learning. *International Journal of Electrical and Computer Engineering (IJECE)* Vol. 11, No. 3, June2021, pp. 2535~2547 ISSN: 2088-8708, DOI: 10.11591/ijece.v11i3.pp2535-2547.
- [25] Heena Khanna and Manmohan Sharma. An Improved Security Algorithm for VANET using Machine Learning. *Journal of Positive School Psychology* 2022, Vol.6, No.3, 7743 – 7756.
- [26] Bangui, H., Ge, M., & Buhnova, B. (2022). A hybrid machine learning model for intrusion detection in VANET. *Computing*, 104(3), 503-531.
- [27] Bangui, H., Ge, M., & Buhnova, B. (2021). A hybrid data-driven model for intrusion detection in VANET. *Procedia Computer Science*, 184, 516-523.
- [28] Shams, E. A., Rizaner, A., & Ulusoy, A. H. (2018). Trust aware support vector machine intrusion detection and prevention system in vehicular ad hoc networks. *Computers & Security*, 78, 245-254.
- [29] Schmidt, D. A., Khan, M. S., & Bennett, B. T. (2020). Spline-based intrusion detection for VANET utilizing knot flow classification. *Internet Technology Letters*, 3(3), e155.
- [30] Alshammari, A., Zohdy, M. A., Debnath, D., & Corser, G. (2018). Classification approach for intrusion detection in vehicle systems. *Wireless Engineering and Technology*, 9(4), 79-94.
- [31] Saleha Saudagar, Radhika P Fuke, Sonika A Chorey, Gayatri Jagnade, "Smart parking stratagem based on IOT," 2019 Cikitusi Journal For Multidisciplinary Research, Volume 6, Issue 4, April 2019, ISSN NO: 0975-6876.
- [32] Saudagar, S.I., Chorey, S.A., Jagnade, G.A. (2019). Review on Intrigue Used for Caching of Information in View of Information Density in Wireless Ad Hoc Network. In: Springer, Singapore. https://doi.org/10.1007/978-981-13-1501-5_59
- [33] Tarwani, Kanchan M., Saleha S. Saudagar, and Harshal D. Misalkar. "Machine learning in big data analytics: an overview." *International Journal of Advanced Research in Computer Science and Software Engineering* 5.4 (2015): 270-274.
- [34] Aboelfottoh, A. A., & Azer, M. A. (2022, May). Intrusion Detection in VANETs and ACVs using Deep Learning. In 2022 2nd International Mobile, Intelligent, and Ubiquitous Computing Conference (MIUCC) (pp. 241-245).IEEE