Analyzing Behavioral Patterns of Authorized User's Access Interruption on CAM using Random Forest

Goodness Oluchi Anyanwu, Jae-Min Lee, and Dong-Seong Kim
Department of IT Convergence Engineering, Kumoh National Institute of Technology Gumi, South Korea
(anyanwu.goodnes, ljmpaul, dskim)@kumoh.ac.kr

***Abstract*—One of the biggest challenges facing the automotive industry in recent times is the increase in the value of data generated and the increase in connectivity. In modern vehicles, data flow on the network is facilitated by Cooperative Awareness Messages (CAM). CAMs are reliably and promptly sent from node to node as communication and network accessibility in vehicle communication must be attained at all times. To secure CAM from interruption in an authorized user's access, a supervised learning solution that adopts Random Forest is proposed in this work. The framework is validated through simulations to demonstrate the effectiveness of the solution in terms of denial-of-service attack detection. The proposed solution achieved an overall accuracy of 99.99% in all attack cases.**

***Index Terms*—Cooperative Awareness Messages, Machine Learning, Denial of Service, Random Forest.**

I. Introduction

As technology-enabled vehicles unfold with continuously evolving sensors and software, there is a need for secure and reliable communication between these systems. Furthermore, with the rise of connected vehicles, the sharing of Co-operative Awareness Messages (CAMs) will become commonplace. CAMs are exchanged in the Intelligent Transport System (ITS) network between ITS nodes to create and maintain awareness of each other [1]. However, due to lack of embedded security and message authentication mechanisms, attackers easily exploit vulnerability in sharing CAM data [2]. Therefore, securing CAM is vital for the smooth operation of ITS.

In light of the above, this work proposed an intrusion detection system (IDS) to protect CAM from interruption in an authorized user's access popularly known as Denial of Service (DoS) attacks. Using Machine Learning (ML) technique, the proposed model is examined using CAM data from an unmodified licensed vehicle-injected with a DoS attack at different intervals and frequencies. ML is a promising technology and has been noted to perform well in cellular vehicle-to-vehicle communications [3]. To use ML to identify intruders, nodes' logs and other data points are examined to identify anomalous occurrences deviating from regular behaviors.

IDS techniques must be robust enough to handle various communication paradigms. An attack that causes these nodes to behave badly could severely disrupt the network [4]. To address DoS intrusion on CAM data, existing literature proposes mechanisms that aim to detect nodes and perpetrators causing network unavailability. However, the specificity and novelty of the proposed intrusion detector in this work on attacks on CAM in VANET rather than the whole VANET communication. Scholarly works on IDS for VANET are extensive. Since we focus on the DoS attack on CAM in VANET, we review some related works on this research topic.

Given this, works related to malicious interruption in an authorized user's access, were reviewed, especially intrusions meant to shut down machines or networks. ML procedures are explored for intrusion detection to achieve highly accurate results in IDS applications. An example of the use of ML method for intrusion in the environment of connected vehicles has been presented in [5]. The researchers introduced an IDS based on an optimized Radial Basic Function of the Support Vector Machine (RBF-SVM) classifier. The RBF-SVM model was initialized and then its kernel parameters were optimized using a grid-search technique.

The design of the ML method used [6] is based on adaptive Deep Belief Networks (DBN). The DBN was developed to handle the cooperative and dynamic nature of ITS topologies. It is important to mention that some of the existing ML methods on DoS detection on CAM were tested using very old datasets. In this article, we have tested the proposed model using a more recent CAM dataset. Moreover, to our knowledge, an accurate IDS solution for securing CAM messages remains an open issue. In this study, we investigated and analyzed DoS attacks on CAM, their intensity and frequency. This work in progress is aimed at making the following contributions:

1) Develop an ML-based methodology to evaluate DoS attacks on CAM data towards repressing network unavailability.
2) Relying on the n_estimators parameter of the proposed model to get insight into the learning behavior of the proposed model.
3) Evaluating the classification performance of the ML using a side-by-side comparison with two competing techniques.

II. Methodology.

The framework proposed in this work is based on the Random Forest (RF) classifier designed to mitigate DoS attacks on CAM. Specifically, the framework was applied to CAM data from a readable Simulation of Urban MObility (SUMO) environment. The data used is obtained from the SUMO platform in Open Street Map (OSM) format. In describing the DoS intrusion mechanism, malicious interruption in an authorized user's access are launched by chosen vehicles at random intervals. The attack durations are arbitrary and range

from 0-30 seconds. In the course of the DoS attack, the attacker sends normal messages, but much more frequently.

Consecutively, there is a randomness in the frequency of the messages sent. Three different DoS attack frequencies were simulated in this dataset [7]. Intervals of 0-10, 10-20, and 20-30, respectively. The aim of the DoS ML detection model is to distinguish between accurate and overwhelming CAM data. Following the behavior of a DoS attack, the ML aims to improve the dependability of CAM data. For evaluation, the ML detection technique is compared with two competing ML techniques (Logistic Regression (LR) and k-Nearest Neighbor (kNN)). This evaluation aims to gauge how well the IDS performs when placed side by side with competing approaches.

III. Result Discussion

Based on the three selected possible ML techniques for detecting a DoS attack on VANET, the results of proposed RF, LR, and kNN are presented in this section. Fig. 1 was developed, summarizing the performance levels/frequency of threats with respect to DoS attack intervals on CAM. The analysis was made possible on the basis of intervals in which various DoS attacks are allowed to occur on CAM. From this table, it is evident that the RF model outperformed the other ML models and is therefore suitable for accurate DoS attack detection on CAM. For all attack frequencies and behavior, the proposed RF achieved an accuracy of 99.99% outperforming benchmark models.

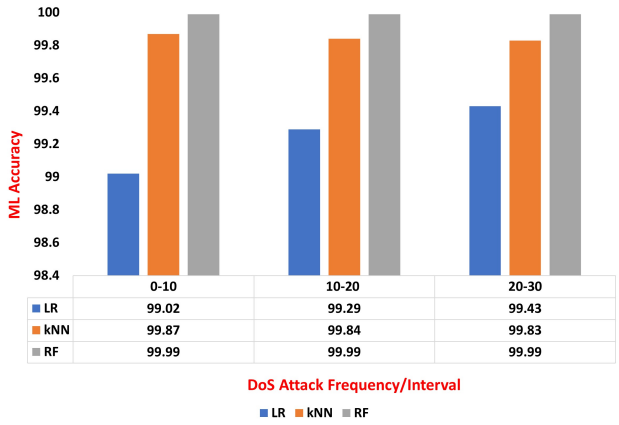

Fig. 1. **Accuracy Comparison of ML Models with respect to DoS Attack Intervals/Frequencies**

Fig. 2 shows the plots of the accuracy performance over a number of RF estimators. For testing purposes, an early stop with only a range of 10–90 estimators was considered. The number of trees in the proposed model's forest is specified by the n_estimators parameter. This parameter is a base parameter that determines how well the RF model performs. As shown in the figure below, evidently, the performance of the RF model increases as the number of estimators increases. This signifies one will need to model with a higher number of estimators beyond 90 to get good performance for the RF configuration. Plotting the result of each estimator epoch is a quick way to get insight into the learning behavior of the RF model.

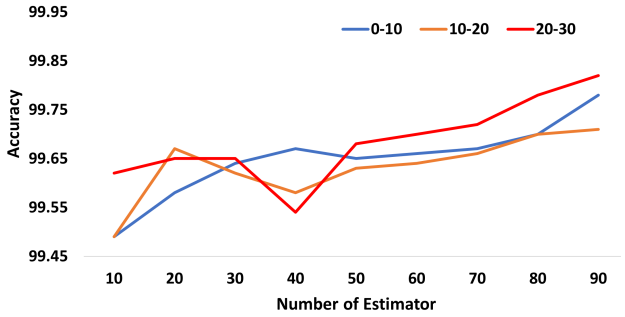

Fig. 2. **RF Cross-validation Score over a Range of 10-90 Estimators**

IV. Conclusion

Safety is the primary concern to many road users. Safety requirements can be powerfully supported by many AI applications. ML application has the opportunity to provide such safety requirements. In this paper, through simulation, we have analyzed the DoS attack frequencies that may be applicable to CAM data. We have developed an accurate RF model to provide a solution to DoS attacks whose intention is to ensure CAM message unavailability between nodes. The proposed solution achieved an overall accuracy of 99.99% across all attack intervals. We intend to expand the proposed approach in the near future to make use of diverse attack scenarios gathered through vehicular communication.

Acknowledgment

This research work was supported by Priority Research Centers Program through NRF funded by MEST (2018R1A6A1A03024003) and the Grand Information Technology Research Center support program (IITP-2022-2020-0-01612) supervised by the IITP by MSIT, Korea.

References

[1] C. Zoghlami, R. Kacimi, and R. Dhaou, "Dynamics of Cooperative and Vulnerable Awareness Messages in V2X Safety Applications," in *2022 International Wireless Communications and Mobile Computing (IWCMC)*, 2022, pp. 853–858.

[2] M. Safwat, A. Elgammal, E. G. AbdAllah, and M. A. Azer, "Survey and Taxonomy of Information-Centric Vehicular Networking Security Attacks," *Ad Hoc Networks*, vol. 124, p. 102696, 2022.

[3] L. Lusvarghi and M. L. Merani, "Machine Learning for Disseminating Cooperative Awareness Messages in Cellular V2V Communications," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 7, pp. 7890–7903, 2022.

[4] K. Stepień and A. Poniszewska-Marańda, "Security Measures in the Vehicular Ad-Hoc Networks in the Aspect of DoS Attack," in *Complex, Intelligent and Software Intensive Systems*, L. Barolli, A. Poniszewska-Maranda, and T. Enokido, Eds. Cham: Springer International Publishing, 2021, pp. 222–232.

[5] G. O. Anyanwu, C. I. Nwakanma, J.-M. Lee, and D.-S. Kim, "Optimization of RBF-SVM Kernel using Grid Search Algorithm for DDoS Attack Detection in SDN-based VANET," *IEEE Internet of Things Journal*, pp. 1–1, 2022.

[6] S. A. Almalki, A. Abdel-Rahim, and F. T. Sheldon, "Adaptive IDS for Cooperative Intelligent Transportation Systems Using Deep Belief Networks," *Algorithms*, vol. 15, no. 7, 2022.

[7] F. Gonçalves, B. Ribeiro, O. Gama, J. Santos, A. Costa, B. Dias, M. J. Nicolau, J. Macedo, and A. Santos, "Synthesizing Datasets with Security Threats for Vehicular Ad-Hoc Networks," in *GLOBECOM 2020 - 2020 IEEE Global Communications Conference*, 2020, pp. 1–6.