JES
Journal of Engineering Sciences

Vol 13, Issue 04, APRIL /2022

ISSN NO:0377-9254

An Intelligent Data-Driven Model to Secure Intravehicle Communications Based on Machine Learning

Vipin Sagar Nagelli, BTech, Department of Mechanical Engineering, vipinsagar2807@gmail.com

Kondamu Sai Teja, BTech, Department of Mechanical Engineering, kondamusaiteja@gmail.com

ABSTRACT: The high relying of electric vehicles on either in-vehicle or between-vehicle communications can cause big issues in the system. This paper is going to mainly address the cyberattack in electric vehicles and propose a secured and reliable intelligent framework to avoid hackers from penetration into the vehicles. The proposed model is constructed based on an improved support vector machine model for anomaly detection based on the controller area network bus protocol. In order to improve the capabilities of the model for fast malicious attack detection and avoidance, a new optimization algorithm based on social spider optimization algorithm is developed, which will reinforce the training process offline. Also, a two-stage modification method is proposed to increase the search ability of the algorithm and avoid premature convergence. Last but not least, the simulation results on the real datasets reveal the high performance, reliability, and security of the proposed model against denial-of-service hacking in the electric vehicles.

Keywords- *Anomaly detection, controller area networks (CAN) bus, intravehicle*

1. INTRODUCTION

Technically, vehicles are composed of many hardware modules, namely electronic control units (ECUs), being controlled by different software tools. All sensors installed in a vehicle will send their data to the ECU, where this data are processed and the requiring orders are sent to the relevant actuators [1]. Such a highly complex hardware–software data transfer process may happen through the use of different network protocols such as controller area network (CAN), LIN, FlexRay, or MOST [2]. Among these protocols, CAN bus is the most popular one not only in vehicles but also in medical apparatuses, agriculture, etc., due to its high capability and promising characteristics. Some of the main advantages of the CAN bus standard may briefly be briefly allowing up to 1 Mb/s data rate transfer, reducing the wiring in the device, saving cost and time due to the simple wiring, auto-retransmission of lost messages, and error detection capability [3]. Unfortunately, since CAN bus protocol was devised at a time where vehicles were almost isolated, this standard suffers from some security issues in the new dynamic environment of smart grids. This will motivate the hackers to attack the electric vehicles through the ECU and inject malicious messages into their systems.

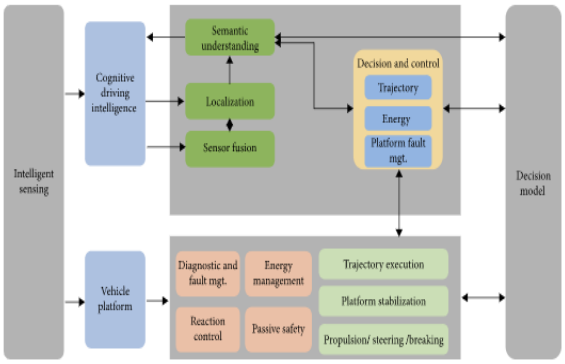

Fig.1: Intelligent driving vehicles' decision-making framework

In [4], some cyber intrusion scenarios are modeled and applied on the electric vehicles to assess their vulnerabilities and possible side effects getting finally into the power grid. In [5], a new classification method is developed for cyber intrusion detection in vehicles. In [6], a data intrusion detection system is developed, which can detect the cyberattack based on the CAN bus message frequency increase or CAN message ID misuse. This will help the driver to detect that an attack has happened so to stop the vehicle immediately. In [7], authors suggest that all CAN messages should pass a data management system to avoid any cyber intrusion. In [8], an algorithmic solution is used to stop attacks of types of denial-of-service (DoS) or error flag in the vehicle. In [9], it is suggested to assign an ECU as the master ECU in the

manufacturing stage of the vehicle so to run an attestation process in the system. In [10], a firewall is introduced for the vehicle to sit between the CAN bus and the communicating system and stop the cyberattack commands to the CAN bus. In [11], an intrusion detection system based on the traffic entropy of an in-vehicle network communication system of the CAN bus is suggested. In [12], an anomaly detection approach is developed, which is capable of detecting faults of known and unknown type without requiring the setting of expert parameters.

2. LITERATURE REVIEW

2.1 Multisource software on multicore automotive ECUs—Combining runnable sequencing with task scheduling:

As the demand for computing power is quickly increasing in the automotive domain, car manufacturers and tier-one suppliers are gradually introducing multicore electronic control units (ECUs) in their electronic architectures. In addition, these multicore ECUs offer new features such as higher levels of parallelism, which ease the compliance with safety requirements such as the International Organization for Standardization (ISO) 26262 and the implementation of other automotive use cases. These new features involve greater complexity in the design, development, and verification of the software applications. Hence, car manufacturers and suppliers will require new tools and methodologies for deployment and validation. In this paper, we address the problem of sequencing numerous elementary software modules, called runnables, on a limited set of identical cores. We show how this problem can be addressed as the following two subproblems, which cannot optimally be solved due to their algorithmic complexity: 1) partitioning the set of runnables and 2) building the sequencing of the runnables on each core. We then present low-complexity heuristics to partition and build sequencer tasks that execute the runnable set on each core. Finally, we globally address the scheduling problem, at the ECU level, by discussing how we can extend this approach in cases where other OS tasks are scheduled on the same cores as the sequencer tasks.

2.2 Gateway system with diagnostic function for LIN, CAN and FlexRay:

In real-time systems such as automotives, a distribution system is used to increase the reliability of the system. As the demand and complexity of the distribution system have increased, several automotive communication protocols have been introduced such as LIN, CAN, and FlexRay. Each node of the system chooses the communication protocol that is suitable for the specific purpose. Each node doesn't need to have all of communication protocols because of cost, space, efficiency, and other factors. Therefore, the gateway system was introduced in the automotive system and has became one of the most important components. The gateway makes possible node-to-node communicate over different communication protocols. However, the gateway system has high probability of error because each protocol has different features such as signaling rate, data length, and so on. Moreover, it is difficult to detect the reason and location of errors. If the gateway reports the protocol conversion result when each protocol is converted into another protocol, this report helps developers find the reason and location of errors to debug errors easily. In this paper, we implement the gateway system with a diagnostic function. LIN, CAN, and FlexRay are used as communication protocols.

2.3 Efficient protocols for secure broadcast in controller area networks:

Controller Area Network is a bus commonly used by controllers inside vehicles and in various industrial control applications. In the past controllers were assumed to operate in secure perimeters, but today these environments are well connected to the outside world and recent incidents showed them extremely vulnerable to cyber-attacks. To withstand such threats, one can implement security in the application layer of CAN. Here we design, refine and implement a broadcast authentication protocol based on the well known paradigm of using key-chains and time synchronization, a commonly used mechanism in wireless sensor networks, which allows us to take advantage from the use of symmetric primitives without the need of secret shared keys during broadcast. But, as process control is a time critical operation we make several refinements in order to improve on the authentication delay. For this we study several trade-offs to alleviate shortcomings on computational speed, memory and bandwidth up to the point of using reduced versions of hash functions that can assure ad hoc security. To prove the efficiency of the protocol we provide experimental results on two representative microcontrollers from the market: a Freescale S12X and an Infineon TriCore, both devices were specifically chosen as they are located somewhat on the extremes of computational power.

2.4 Advancing cyber–physical sustainability through integrated analysis of smart power systems: A case study on electric vehicles:

Satisfying the growing energy demand, power systems are required to increase their capacity and be able to distribute energy over wider geographical area. To maintain the reliability of such power systems, more dependence is placed on automating the process controlling the physical system. Such power systems are known as smart grids, where data is transmitted in real-time across the power grid facilitating automated actions. These smart systems have intrinsic vulnerabilities. A smarter power grid is more reliant on an ICT backbone. Such reliance renders the physical power system subjected to an ICT realm whose security depends on a set of metrics, and standards alien to those of classical power systems. The sustainability of a power system will depend on the secure and reliable operation of the new smart system. This paper proposes a comprehensive approach to solve the challenges enumerated above. The proposed approach sets the ground for new metrics of overall system sustainability, by dissecting the overall smart grid into three layers: the physical system, the SCADA system, and the ICT infrastructure. The proposed methodology is tested on the case of electric vehicles, where cyber intrusion scenarios are studied in light of their effect on the physical layer.

2.5 A taxonomy and survey of cyber-physical intrusion detection approaches for vehicles:

With the growing threat of cyber and cyber-physical attacks against automobiles, drones, ships, driverless pods and other vehicles, there is also a growing need for intrusion detection approaches that can facilitate defence against such threats. Vehicles tend to have limited processing resources and are energy-constrained. So, any security provision needs to abide by these limitations. At the same time, attacks against vehicles are very rare, often making knowledge-based intrusion detection systems less practical than behaviour-based ones, which is the reverse of what is seen in conventional computing systems. Furthermore, vehicle design and implementation can differ wildly between different types or different manufacturers, which can lead to intrusion detection designs that are vehicle-specific. Equally importantly, vehicles are practically defined by their ability to move, autonomously or not. Movement, as well as other physical manifestations of their operation may allow cyber security breaches to lead to physical damage, but can also be an opportunity for detection.

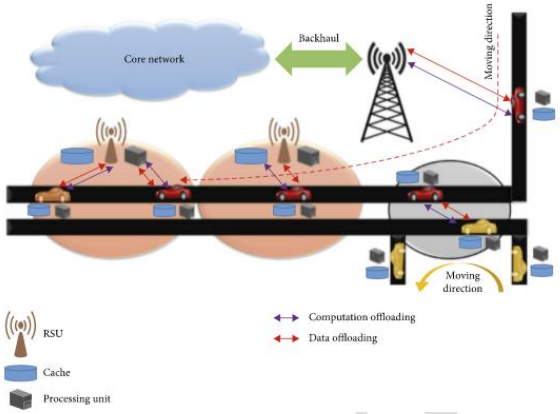

Fig.2: IoV mobility-aware caching and computational scenario.

2.6 Security threats to automotive can networks practical examples and selected short-term countermeasures:

The IT security of automotive systems is an evolving area of research. To analyse the current situation and the potentially growing tendency of arising threats we performed several practical tests on recent automotive technology. With a focus on automotive systems based on CAN bus technology, this article summarises the results of four selected tests performed on the control systems for the window lift, warning light and airbag control system as well as the central gateway. These results are supplemented in this article by a classification of these four attack scenarios using the established CERT taxonomy and an analysis of underlying security vulnerabilities, and especially, potential safety implications. With respect to the results of these tests, in this article we further discuss two selected countermeasures to address basic weaknesses exploited in our tests. These are adaptations of intrusion detection (discussing three exemplary detection patterns) and IT-forensic measures (proposing proactive measures based on a forensic model). This article discusses both looking at the four attack scenarios introduced before, covering their capabilities and restrictions. While these reactive approaches are short-term measures, which could already be added to today's automotive IT architecture, long-term concepts also are shortly introduced, which are mainly preventive but will require a major redesign. Beneath a short overview on respective research approaches, we discuss their individual requirements, potential and restrictions.

2.7: On the need of data management in automotive systems:

In the last decade, automotive systems changed from traditional mechanical or mechatronical systems towards software intensive systems, because more and more functionality has been implemented by software. Currently, this trend is still ongo- ing. Due to this increased use of software, more and more data accumulates and thus, has to be handled. Since it was no subject up to now to manage this data with soft- ware separatly, we think that it is indispensable to establish a data managent system in automotive systems. In this paper we point out the necessity of data management, supported by exemplary scenarios, in order to overcome disadvantages of current solu- tions. Further, we discuss main aspects of data management in automotive systems and how it could be realized with respect to the very special restrictions and requirements within such a system.

2.8: An algorithm for detection of malicious messages on can buses:

Control systems are encountering increasing security threats; as one of the common systems, CAN control system is very facile to be attacked. Aiming at the improvement of CAN control system security, an algorithm for detection of malicious CAN messages is given, and it has been implemented in the CANoe simulation environment. The result shows that this algorithm has powerful detection function and valuable practice signification.

3. IMPLEMENTATION

Technically, vehicles are composed of many hardware modules namely called electronic control units (ECUs) being controlled by different software tools. All sensors installed in a vehicle will send their data to the ECU, where this data are processed and the requiring orders are sent to the relevant actuators [1]. Such a highly complex hardwaresoftware data transfer process may happen through the use of different network protocols such as CAN, LIN, FlexRay or MOST [2]. Among these protocols, CAN bus is the most popular one not only in vehicles, but also in medical apparatuses, agriculture, etc due to its high capability and promising characteristics. Some of the main advantages of the CAN bus standard may be briefly named as allowing up to 1Mbps data rate transfer, reducing the wiring in the device.

Propose an intelligent and highly secure method to equip the electric vehicles with a powerful anomaly detection and avoidance mechanism. The proposed method is constructed based on support vector machine and the concept of one-class detection system to avoid any malicious behavior in the vehicle [13]. Here the experimental CAN bus data are used to let the support vector machine learn the normal frequency of the different message frames at different commands. In order to get into the maximum capability of the model, a new optimization algorithm based on social spider optimization (SSO) algorithm is proposed to adjust the SVR setting parameters, properly [14].

In this paper author is describing concept to secure communication between intravehicle to avoid accidents which can trigger due to malicious intruder attack. In electric vehicles different sensors attached to different parts (steering, brakes, engine etc.) of vehicle will sense data and send to ECU (Electronic Control Units) and ECU will process data and send result back to sensors and sensors will act based on response received from ECU. Sometime some malicious users may hack sensors and ECU communication and send false information to ECU and this false information cause sensors to work improperly and can cause accidents. All existing techniques were using AES encryption to have secure packet transmission between sensors and ECU but hackers are still able to decode encrypted packets and hack communication between ECU and sensors. All ECU and Sensor communicate with each other using CONTROLLER AREA NETWORK (CAN) protocol.

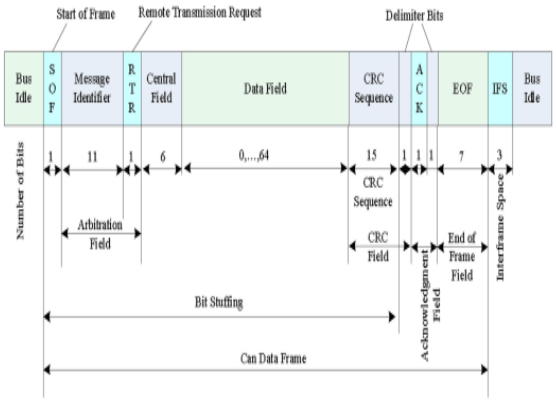

Fig.3: Structure of a message frame in CAN bus

To overcome from this problem author using machine learning algorithms to detect intrusion or anomaly packet received by ECU or sensors. Machine learning algorithms will be trained and model to predict attack based on request frequency received by ECU. All hackers will send packet with

high frequency and priority to make ECU busy and to process high priority packets and other genuine sensors request will keep on waiting. To avoid this problem machine learning will build train model with attack class label as 1 when high frequency of packets received with same id or device. If packets receiving in normal mode then class label 0 will be assigned which indicates received packet is normal.

Author using various machine learning algorithms such as Conventional SVM algorithm, Decision Tree, KNN Algorithm and propose Social Spider Algorithm with SVM by selecting optimal features and evaluating their performance with indices such as HR (Hit Rate), MR (Miss Rate), CR (Correct Rejection Rate) and FR (False Alarm Rate).

Here HR refers to machine learning metric called TRUE POSITIVE (TP) which means classifier able to predict given record correctly as positive.

MR refers to machine learning metric called FALSE NEGATIVE (FN) which means classifier unable to predicted given record correctly

CR refers to machine learning metric called TRUE NEGATIVE (TN) which means classifier able to predict given record correctly as negative

FR refers to machine learning metric called FALSE POSITIVE (FP) which means classifier predicting negative records as positive.

For any classifier whose HR and CR is high then its performance will be consider as better and efficient. In propose work to secure CAN bus (electric vehicle communication) author is doing enhancement to SVM algorithm by analysing frequency of received packets and if received packets from same device ID has high frequency then SVM mark that records as anomaly and propose SVM performance will be evaluated using above four indices such as HR, FR, MR and CR. In propose SVM to select optimal features from dataset author is using SOCIAL SPIDER OPTIMIZATION (SSO) algorithm. In this algorithm dataset features vector will be consider as SPIDERS and fitness will be calculated between all features and features which has high similarity will be consider as related and will have high fitness score and all those high fitness score features will be selected and low fitness features will be removed out. Comparison between one features to other features will be consider as MALE and FEMALE spiders. After applying SSO algorithm we will have optimal features using which classifier can efficiently predict anomaly from new and old records.

4. ALGORITHMS

Author using various machine learning algorithms such as Conventional SVM algorithm, Decision Tree, KNN Algorithm and propose Social Spider Algorithm with SVM.

SVM:

SVM is a supervised machine learning algorithm which can be used for classification or regression problems. It uses a technique called the kernel trick to transform your data and then based on these transformations it finds an optimal boundary between the possible outputs.

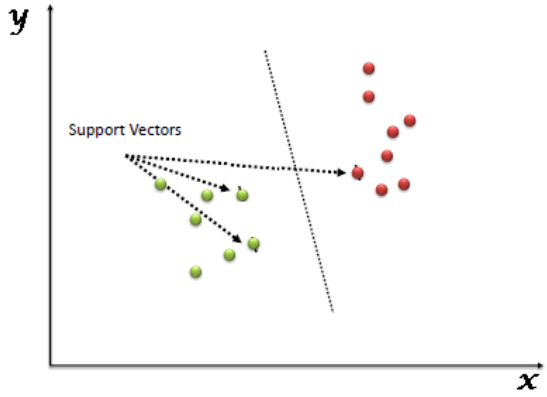

Fig.4: SVM architecture

The kernel trick takes the data you give it and transforms it. In goes some great features which you think are going to make a great classifier, and out comes some data that you don't recognize anymore. It is sort of like unraveling a strand of DNA. You start with this harmelss looking vector of data and after putting it through the kernel trick, it's unraveled and compounded itself until it's now a much larger set of data that can't be understood by looking at a spreadsheet. But here lies the magic, in expanding the dataset there are now more obvious boundaries between your classes and the SVM algorithm is able to compute a much more optimal hyperplane.

DECISION TREE:

Decision Tree algorithm belongs to the family of supervised learning algorithms. Unlike other supervised learning algorithms, the decision tree algorithm can be used for solving regression and classification problems too. The goal of using a Decision Tree is to create a training model that can use to predict the class or value of the target variable

by learning simple decision rules inferred from prior data(training data). In Decision Trees, for predicting a class label for a record we start from the root of the tree. We compare the values of the root attribute with the record's attribute.

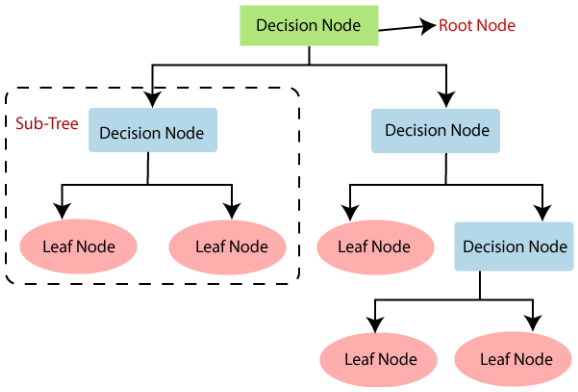

Fig.5: Decision tree architecture

KNN:

The k-nearest neighbors (KNN) algorithm is a simple, easy-to-implement supervised machine learning algorithm that can be used to solve both classification and regression problems. Machine learning models use a set of input values to predict output values. KNN is one of the simplest forms of machine learning algorithms mostly used for classification. It classifies the data point on how its neighbor is classified.

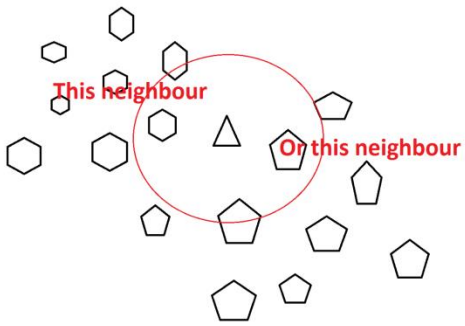

Fig.6: K-NN architecture

5. EXPERIMENTAL RESULTS

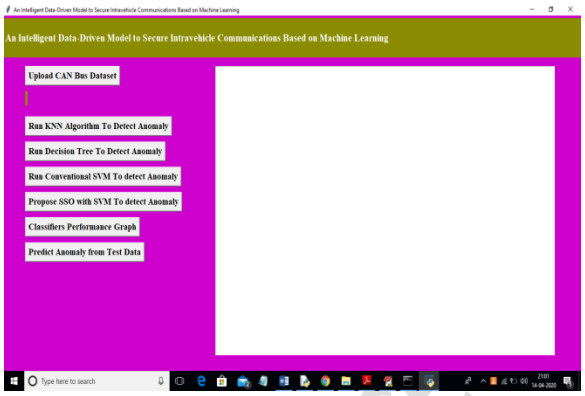

Fig.7: Home screen.

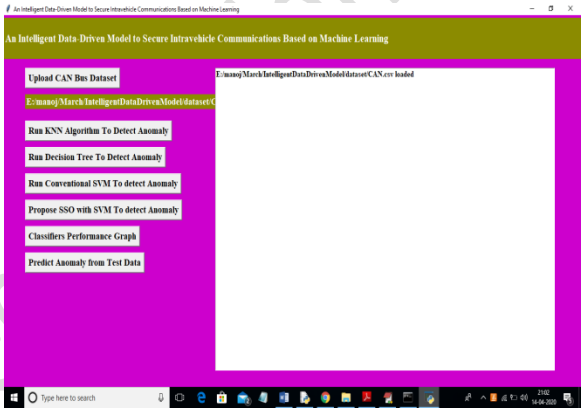

Fig.8: Data uploading

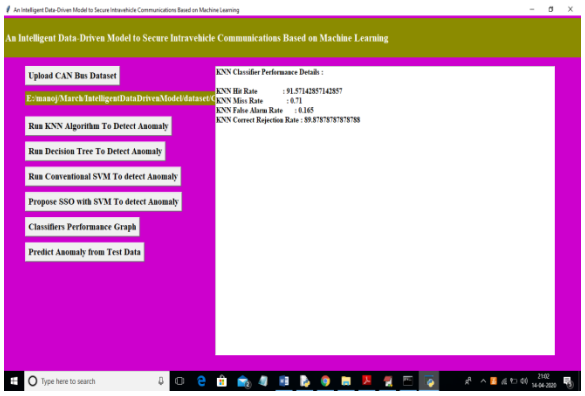

Fig.9: Run KNN algorithm to detect anomaly

Journal of Engineering Sciences

Vol 13, Issue 04, APRIL /2022

ISSN NO:0377-9254

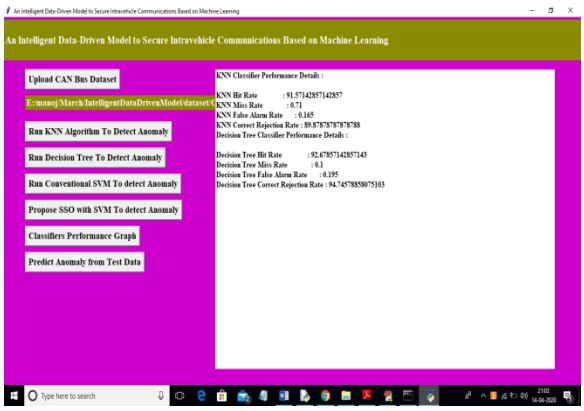

Fig.10: Run Decision tree algorithm to detect anomaly

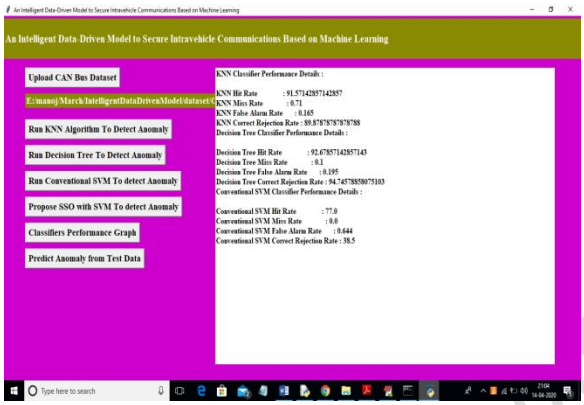

Fig.11: Run Conventional SVM To detect Anomaly

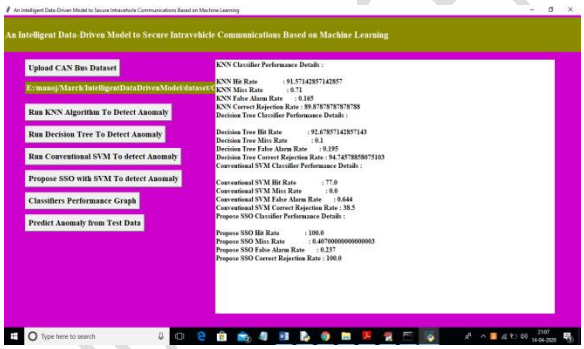

Fig.12: Propose SSO with SVM To detect Anomaly'

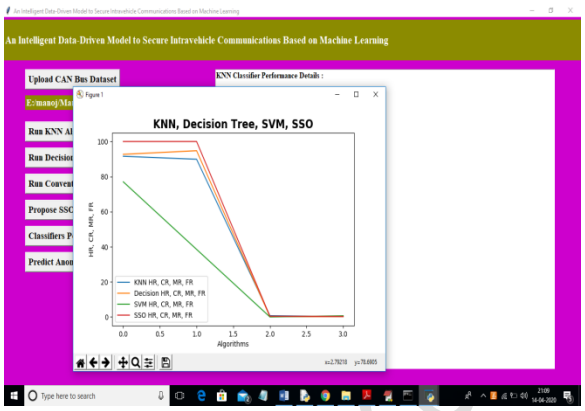

Fig.13: Classifiers performance graph

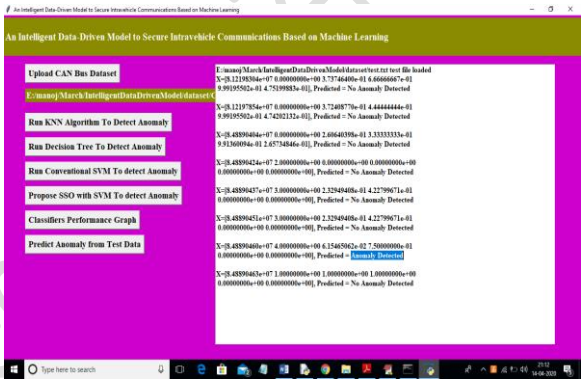

Fig.14: Prediction result

5. CONCLUSION

This paper proposed a novel intelligent and secured anomaly detection model for cyberattack detection and avoidance in the electric vehicles. The proposed model was constructed based on an improved support vector machine model reinforced by the MSSO algorithm. From the cybersecurity point of view, the proposed model could successfully detect malicious behaviors while letting the trusted message frames broadcast in the CAN protocol. The high HR% and FR% indices proved the true positive and true negative decisions made by the proposed model. Regarding the MR% and CR% indices, the very low values, which most of them were around the upper and lower bounds of the message frame frequency, showed the highly trustable performance of this model. The authors will assess the effect of other cyberattacks on the performance of different anomaly detection models in the future works.

REFERENCES

[1] A. Monot, N. Navet, B. Bavoux, and F. Simonot-Lion, "Multisource software on multicore automotive ECUs—Combining runnable sequencing with task scheduling," IEEE Trans. Ind. Electron., vol. 59, no. 10, pp. 3934–3942, Oct. 2012.

[2] T. Y. Moon, S. H. Seo, J. H. Kim, S. H. Hwang, and J. Wook Jeon, "Gateway system with diagnostic function for LIN, CAN and FlexRay," in Proc. Int. Conf. Control, Autom. Syst., 2007, pp. 2844–2849.

[3] B. Groza and S. Murvay, "Efficient protocols for secure broadcast in controller area networks," IEEE Trans. Ind. Informat., vol. 9, no. 4, pp. 2034– 2042, Nov. 2013.

[4] B. Mohandes, R. Al Hammadi, W. Sanusi, T. Mezher, and S. El Khatib, "Advancing cyber–physical sustainability through integrated analysis of smart power systems: A case study on electric vehicles," Int. J. Crit. Inf. Protection, vol. 23, pp. 33–48, 2018.

[5] G. Loukas, E. Karapistoli, E. Panaousis, P. Sarigiannidis, and T. Vuong, "A taxonomy and survey of cyber-physical intrusion detection approaches for vehicles," Ad Hoc Netw., vol. 84, pp. 124–147, 2019.

[6] T. Hoppe, S. Kiltz, and J. Dittmann, "Security threats to automotive can networks practical examples and selected short-term countermeasures," Rel. Eng. Syst. Saf., vol. 96, no. 1, pp. 11–25, 2011.

[7] S. Schulze, M. Pukall, G. Saake, T. Hoppe, and J. Dittmann, "On the need of data management in automotive systems," in Proc. Bus., Technol. Web, 2009, vol. 144, pp. 217–26.

[8] C. Ling and D. Feng, "An algorithm for detection of malicious messages on can buses," presented at Nat. Conf. Inf. Technol. Comput. Sci., 2012, pp. 12–18.

[9] H. Oguma, X. Yoshioka, M. Nishikawa, R. Shigetomi, A. Otsuka, and H. Imai, "New attestation based security architecture for in-vehicle communication," in Proc. IEEE Global Telecommun. Conf., 2008, pp. 1–6.

[10] L. Pan, X. Zheng, H. X. Chen, T. Luan, and L. Batten, "Cyber security attacks to modern vehicular systems," J. Inf. Secur. Appl., vol. 36, pp. 90– 100, Oct. 2017.

[11] M. J. Kang and J. W. Kang, "Intrusion detection system using deep neural network for in-vehicle network security," PloS One, vol. 11, no. 6, 2016, Art. no. e0155781.

[12] A. Theissler, "Detecting known and unknown faults in automotive systems using ensemble-based anomaly detection," Knowl.-Based Syst., vol. 123, pp. 163–173.

[13] F. Zhu, J. Yang, C. Gao, S. Xu, and T. Yin, "A weighted one-class support vector machine," Neurocomputing, vol. 189, pp. 1–10, May 2016.

[14] Y. Zhou, Y. Zhou, Q. Luo, and M. Abdel-Basset, "A simplex methodbased social spider optimization algorithm for clustering analysis," Eng. Appl. Artif. Intell., vol. 64, pp. 67–82, 2017.

[15] G. De La Torre, P. Rad, and K. K. R. Choo, "Driverless vehicle security: Challenges and future research opportunities," Future Gener. Comput. Syst., to be published, [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167739X17315066