Deep Anomaly Detector Based on Spatio-Temporal Clustering for Connected Autonomous Vehicles

Rachid Oucheikh[1,2(✉)], Mouhsene Fri[3], Fayçal Fedouaki[2], and Mustapha Hain[2]

[1] National High School for the Arts and Professions, Casablanca, Morocco
[2] School of Engineering, Jönköping University, Jönköping, Sweden
rachid.oucheikh@ju.se
[3] Euro-Med University of Fes, (UEMF), Route de Meknes, Rond-point de Bensouda, 3007 Fès, Morocco
m.fri@ueuromed.org

Abstract. Connected Autonomous Vehicles (CAV) are expected to revolutionize the transportation sector. However, given that CAV are connected to internet, they face a principal challenge to ensure security, safety and confidentiality. It is highly valuable to provide a real-time and proactive anomaly detection approach for Vehicular Ad hoc Network (VANET) exchanged data since such an approach helps to trigger prompt countermeasures to be undertaken allowing the damage avoidance. Recent machine learning methods show great efficiency, especially due to their capacity to handle nonlinear problems. However, an accurate anomaly detection in a space–time series is a challenging problem because of the heterogeneity of space–time data and the spatio-temporal correlations. An anomalous behavior can be seen as normal in different context. Thus, using one deep learning model to classify the observations into normal and abnormal or to identify the type of the anomaly is usually not efficient for large high-dimensional multi-variate time-series datasets. In this paper, we propose a stepwise method in which the time-series data are clustered on spatio-temporal clusters using Long Short Term Memory (LSTM) auto-encoder for dimension reduction and Grey Wolf Optimizer based clustering. Then, the anomaly detection is performed on each cluster apart using a hybrid method consisting of Auto-Encoder for feature extraction and Convolution Neural Network for classification. The results shows an increase in the accuracy by 2% in average and in the precision by approximately 1.5%.

Keywords: Connected autonomous vehicles · Anomaly detection · Vehicular Ad hoc network · Deep learning

1 Introduction

CAV are expected to revolutionize the transportation sector. Equipped with different sensors and Internet of Things (IoT), CAV technology demonstrates

L. Foschini and M. El Kamili (Eds.): ADHOCNETS 2020, LNICST 345, pp. 201–212, 2021.
https://doi.org/10.1007/978-3-030-67369-7_15

the potential to offer high level of safety, reliability and efficiency. In fact, CAV can alleviate road congestion [8], reduce energy consumption [17], accidents and pollution [11,14]. Considering that CAV are connected to internet, they face a principal challenge to ensure security, safety and confidentiality. In fact, communication between Vehicle-to-Infrastructure (V2I) and Vehicle-to-vehicle (V2V) are based on VANET. VANET is composed of three main components: vehicular, road side unit (RSU) and Service. The mentioned challenges motivate several authors to propose frameworks to detect anomalies in the exchanged data. The anomaly detection is defined as a process of identifying outliers or abnormalities which present significantly deviation from normal activities and may indicate suspicious activities.

Anomaly detection consists of three main components **i) Point anomalies** appear randomly as a fluctuation or irregularity at a data point of data without any particular interpretation. **ii) Collective or Group Anomaly** which describes an abnormal behavior of a group of data points. Even if the isolated and individual points can be considered normal, their co-occurrence in a particular way, defined by the problem, can be anomalous. **iii) Contextual anomaly** or conditional anomaly detection refers to the problem of identifying anomalies in specific conditions. A concrete example in CAV data is that the same numbers quantifying speed or flow of vehicles in rush hours are considered anomalous in midnight. To further formalize the notion, we assume that data points are described by contextual features and behaviour features. The contextual features are generally time and space. If the behaviour attributes of a data point are considered anomalous relatively to the behaviour attributes of the data subset having the same or similar contextual attributes, then the corresponding data point is classified as a contextual anomaly.

On the other hand, the network data anomalies can be classified according to their evolution over time into three main types: abrupt, intermittent and gradual. An abrupt anomaly is a abnormal sudden change in the features of the data. This change can be expressed using the step-function which equals $x(t)$ if $t < t_f$ and $x(t) + b$ otherwise. Gradual or incipient sensor anomalies described by the function 1 are usually due to a gradual deterioration of the sensor over a long period of time. Intermittent anomalies are anomalies that appear and disappear periodically.

$$x^f(t) = \begin{cases} x(t) & if\ t < t_f \\ x(t) + s * (t - tf)\ if\ t \geq t_f \end{cases} \tag{1}$$

This paper aims to detect in real-time these types of anomalies in the sensor data received by CAV taking into account their context. To achieve this goal, we propose two-phases approach. In the first phase, clustering of data is performed in order to unveil the latent the spatio-temporal features. Then, according to these resulting features, the second phase tries to detect the anomalies using a hybrid framework built upon two deep learning techniques, namely the auto-encoder and convolutional neural network (AE-CNN). The approach is

lightweight and can be deployed on vehicles to provide real-time and proactive way for anomaly detection.

The remaining of this paper is organized as follows: the second section provides the background related to time series clustering. The third section introduces our proposed approach for the detection of anomalies in CAV data. Then, in the third section, a performance evaluation of the model is described and its obtained results are discussed. Finally, conclusions are drawn in Sect. 5.

2 Background

Connected autonomous Vehicle is able to connect to wireless networks and can communicate with other vehicles V2V and to infrastructure V2I. Such communication enables CAV to navigate autonomously, avoid collisions and congestion and optimize many constraints such as time and energy. While CAV increase reliability, convenience and safety of both riders and pedestrians, they also present an attack target that could be exploited. Some researches already demonstrate exploitable vulnerabilities in ordinary vehicles [1]. These vulnerabilities become more and more prevalent with the increase of number of connected vehicles. To enhance safety and security of vehicular networks, anomaly detection techniques are used to identify in real-time data points or events which do not follow an expected pattern.

The authors in [10] used machine learning on VeReMi dataset to identify and analyse anomaly behaviours using different features. They demonstrate that machine learning is an efficient method to detect anomalies. The authors in [9] proposed a spatio-temporal framework to detect anomalies in VANET network. The framework is formed by convolutional neural network and trained on dataset from real traffic simulation of a RSU and 12 On-Board Units (OBUs). The authors in [13] propose framework to detect anomaly in alert message comuunicated through VANET, the authors evaluated multiple machine learning algorithms using hyperparameters tuning.

The authors in [6] provide a mechanism, called SHIELDNET, which serves to detect and defend against the vehicular botnets. To provide effective data anomaly detection, [5] combines the characteristics of the vehicular network data and the driver's emotional state extracted using sentiment analysis. The authors design a driver's emotion quantification model to reflect the driving style. Then, output of this model and vehicle driving data are fed to the anomaly detection algorithm designed based on Gaussian mixed model. F. Ghaleb et al. proposed a context-aware data-centric misbehaviour detector which uses sequential analysis of temporal and spatial features of neighbouring vehicles' mobility information [7]. The evaluation of data consistency is performed using jointly the innovation error of the Kalman filter algorithm and Hampel filter.

The main challenge for anomaly detection on VANET is the fact that the data are unlabeled. Various techniques have been developed to draw inferences from this type of data, but they are not effective for complex, high-level structures and functionality such as those dealing with labeled datasets. Standard unsupervised techniques include clustering approaches that are used to group a set of objects together so that objects in the same group, called a cluster, are more similar to each other than to objects in other groups. Such techniques differ in the method for organizing the data as well as the metrics to measure similarity. While clustering techniques have been successfully applied to static data, their extension to time series data remains an challenging problem. Formally, the clustering task is defined as follows:

Definition 1 (Clustering). Consider m data $x_1, x_2, ..., x_m$ set of k clusters $C = \{C_1, C_2, ..., C_k\}$ and a distance measure D. Each cluster C_i is represented by its centroid $\bar{c_i}$. The objective of a clustering algorithm is to partition the data into similar groups such as the optimal clustering minimizing the within-cluster sum of squares, denoted as C^* is expressed in Eq. 2, where C_j is the cluster j and E_c is the space of cluster combinations:

$$C^* = \underset{E_c}{\operatorname{argmin}} \sum_{j=1}^{k} \sum_{x_i \in C_j} \|x_i - \bar{c_j}\|^2 \tag{2}$$

Time series are sequences of data points indexed in time order. Time series data from different domains exhibit considerable variations in important properties and features, temporal scales, and dimensionality. Further, time series data from real world applications often have temporal gaps as well as high frequency noise due to the data acquisition method and/or the inherent nature of the data. They are classified into three categories: hierarchical clustering, pure partitioning space–time series clustering and overlapping partitioning space–time series clustering.

Deep clustering is a family of clustering methods that make use of deep neural networks. It aims usually to optimize a loss function which is typically composed of two components: a) network loss L_n which allow the network to learn the most useful and representative features and avoid trivial solutions and b) clustering loss L_c that plays a discriminative role by driving the feature data extracted from the network to form disjoint groups. The total loss is expressed as: $L = \lambda L_n + (1 - \lambda)L_c$, where $\lambda \in [0, 1]$ is a hyperparameter to balance the two functions.

Recently, clustering algorithms are applied for anomaly detection in different applications including intelligent transportation systems. In [16], N. Peri et al. aim to detect anomalies such as stalled vehicles and collisions using vehicle re-identification and multi-camera vehicle tracking. The bottom up Agglomerative clustering method is used to obtain the multi-camera trajectories. In [18], the authors introduced a vehicle trajectory clustering method that employs a dynamic network representation learning upon a k-Nearest Neighbors (KNN) based internet of vehicles.

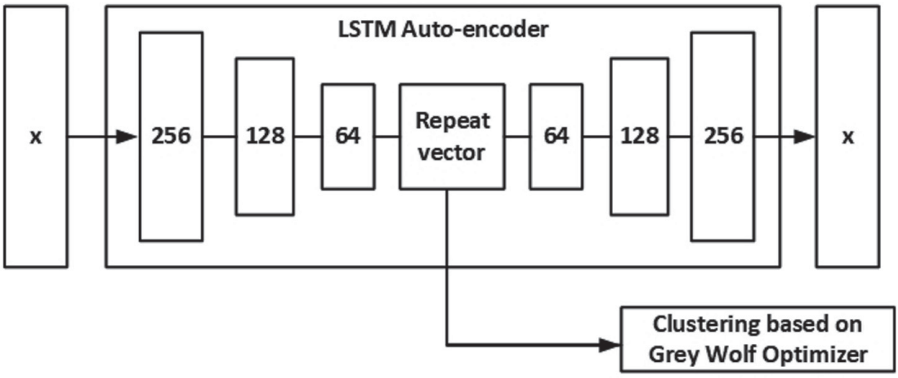

Fig. 1. Spatio-temporal clustering model's

We define the distance between two space–time series x_1 and x_2 as: $d(x_1, x_2) = d_s(x_1, x_2) + d_t(x_1, x_2)$ where d_s is the spatial distance that computes the similarity between feature points of the time-series regarding the spatial component and defines the temporal distance that determines the similarity between the time series components. The time series distance can be euclidean distance L_p or dynamic time warping (DTW).

3 Proposed Approach

The objective of the proposed framework is to provide CAV with a robust tool to detect in real-time anomalies in the networking data and signals received from on-board and external sensors. In our previous work [15], an approach is introduced to identify the type of anomaly: abrupt, intermittent or gradual. The specific aim of this paper is to increase the accuracy of anomalous event detection by considering the contextual information of the event occurrence. To achieve this goal, our proposed approach includes two main parts. First, a spatio-temporal clustering of the raw data is carried out in order to get the context of the events. The result of this phase is the clusters and the space-time subspaces in which the sensor data is similar or has the same features. Then, in the second phase, each anomaly detector is trained on a specific cluster. Regarding the inference process, the cluster of each data sequence received from the CAV is first determined, and then the anomaly detector associated to this cluster is used to classify this sequence as normal or abnormal.

The architecture of the sensor data clustering is illustrated in the Fig. 1 and consists of two main parts:

– **Dimensionality reduction:** Since the data received from the sensors are high-dimensional multivariate time-series, a LSTM auto-encoder is used to reduce their dimension and clean it from noise. The objective of the LSTM

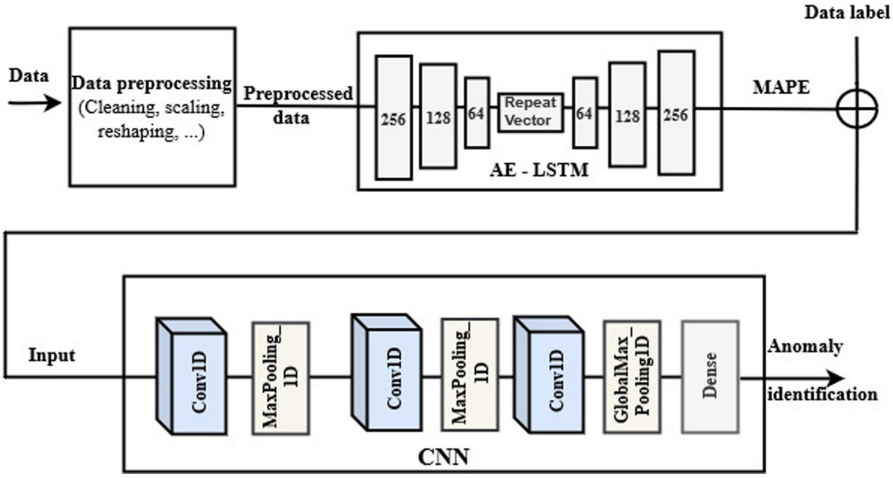

Fig. 2. The proposed model architecture to detect anomaly

auto-encoder which is an unsupervised model is to learn an approximation of the identity function since the target values are equal to those of the input. To learn this mapping, it first compresses the raw data to low dimensional space and then decompresses it to reconstruct the same data as in the input. The idea here is to make use of the compressed version of the data which is less complex, denoised and includes essential features of the original data. The asset of this step is to obtain an informative and efficient latent representation which will afford high quality clustering. As shown in Fig. 2, the network architecture consists of a consecutive LSTM layers with different sizes. This dimensionality reduction is important for further processing to avoid very long sequences which can lead to poor performance.

– **Clustering:** As shown in the Fig. 1, data with reduced dimension which is the compressed version in the decoder output is clustered using Grey Wolf Optimizer (GWO). GWO is a swarm-based metaheuristic developed by Mirjalili et al. [2]. This algorithm is inspired from hunting mechanism of grey wolves. the population of wolves is divided into four groupes, : alpha (α, γ, β and δ, the wolves omega attack the prey. GWO has widely applied in various area, such as, engineering [4], performance measurement system [12], clustering [3].

The first part of our approach provides us with spatio-temporal clusters. The second phase builds the deep models which serve to data training and inference for each cluster. Each model network consists of two components: LSTM auto-encoder which builds the data features and CNN classifier which uses the built features to detect anomalous in the corresponding cluster. The overview of the proposed framework is shown in Fig. 2. Each model is trained on specific cluster

and will serve for inference of that cluster. The two parts of the network model can be described as follows:

– **Extraction of signal features:** Once we identify the cluster of a data sequence representing the signal data received from the sensors, it will be fed in its raw form to the LSTM auto-encoder shown in Fig. 1. The role of this component is to learn the latent features of the normal data which is free of anomalies. For this reason, only the normal data are used to train the auto-encoder. After a full training of the auto-encoder, it will be able to accurately reconstruct the normal data and get normal outputs approximately similar to that of the original data i.e. input of auto-encoder. This means that the error of the reconstruction will be very small. In contrast, if the data fed to the trained model includes the abnormal data, the predictions will have big margin error. This way, we will be able to build a new relevant feature space. In fact, the problem of anomaly detection is originally a one class classification problem in which only the normal data is known with details. The LSTM will then differentiate the normal data from any other data whatever the anomalies it includes. In real-world the patterns of the anomaly data could be of wide spectrum. In this paper, three anomaly types are used and their error reconstruction features are obtained using the LSTM auto-encoder.
We build the feature space using distance measure called Mean Absolute Percentage Error (MAPE) described by the Eq. 3 such that $y_{p,x}$ is the x^{th} predicted result, $y_{a,x}$ is actual value, N is the number of predictions. The correspondent error tensor of MAPE is fed to the input of the CNN classifier.

$$MAPE = \frac{100}{N} \sum_{i=0}^{N} \frac{(y_{p,i} - y_{a,i})}{y_{a,i}} \tag{3}$$

– **Anomaly classification:** This component aims to exploit the spatial features of the vector resulting from the previous phase. The CNN model network consists of three One-dimensional convolutional layers with 32 filters and different sized kernels to allow multi-resolution data processing and thus the model can learn from these different levels and granularities. The dropout is used to reduce the overfitting by training the neural network with different architectures in parallel and dropping combination of layers in each iteration. The maxpooling operation is performed after each convolution layer to downsample the feature map representation reducing its dimensionality and allowing for detection of features contained in the sub-regions. A global pooling is applied on the 32 feature maps obtained in the output of the last convolution layer. The reason behind this layer is to downsample and summarize each feature map to a single value ready to feed to the dense layer.

4 Results and Discussion

4.1 Data Collection

To evaluate the proposed approach, we used the dataset collected by Basic Safety Messages (BSM) inside the Wyoming Connected Vehicle (CV) Pilot project [19]. The project aims to reduce incident-related delays and to improve safety in the corridor Wyoming between Canada, United States, and Mexico.

These data are received by other vehicles via Vehicle-to-Vehicle (V2V) communications to help determine immediate threats and alert drivers as necessary. The dataset includes more than 2.5 lines and consists of 69 features such as longitude, latitude and altitude (position), speed, timestamp, acceleration.

4.2 Evaluation and Results

The collected data is first cleaned and then clustering in multi-dimensional space is performed using the deep clustering model based on Grey Wolf. The results of the clustering are depicted in the Figs. 3 and 4, which show the projection in 2-dimensions and each color represents one cluster. The abscissa presents the time converted to real number under format "HHMM" such that "HH" denotes the hour and "MM" denotes the minutes. The ordinates present the scaled latitude in Fig. 3 and the scaled longitude in Fig. 4. The clustering results show that the data having the same features lie in the time intervals $[00:00, 07:00]$, $[07:00, 10:00]$, $[10:00, 15:30]$, $[15:30, 20:00]$, $[20:00, 00:00]$. Regarding the space, the clustering shows that the pieces of road having the same features are delimited by longitude and latitude values that are dynamically changing depending on the time as depicted in the figures. For example, longitude can be divided in two intervals in the early morning and three intervals in between 10 h and 15 h 30.

The data is reshaped into 3D array (samples, timesteps, features) to fit in the LSTM auto-encoder input. Afterwards, four operations are required: normal distribution scaling, window generation, shuffling, and train/test splitting. The obtained preprocessed data is used to train the auto-encoder consisting of 6 layers as described in the Fig. 1. The trained model is then used to build features of normal and anormal data. Before feeding these features to the CNN, two operations have to be performed: dimension shuffling and increase of the sample width. The CNN input is a multi-channel with a single timestep.

Some experiments have been led to evaluate the performance of our proposed model. First, we study the impact of using the contextual information on the classification performance metrics. Table 1 shows the impact of detecting anomalies according to different context. It is clear that the overall performance is increased slightly when using either temporal or spatial context. Spatial context is more informative for the considered data. Using spatio-temporal context gives more improved results with an increase of approximately 1% in the accuracy.

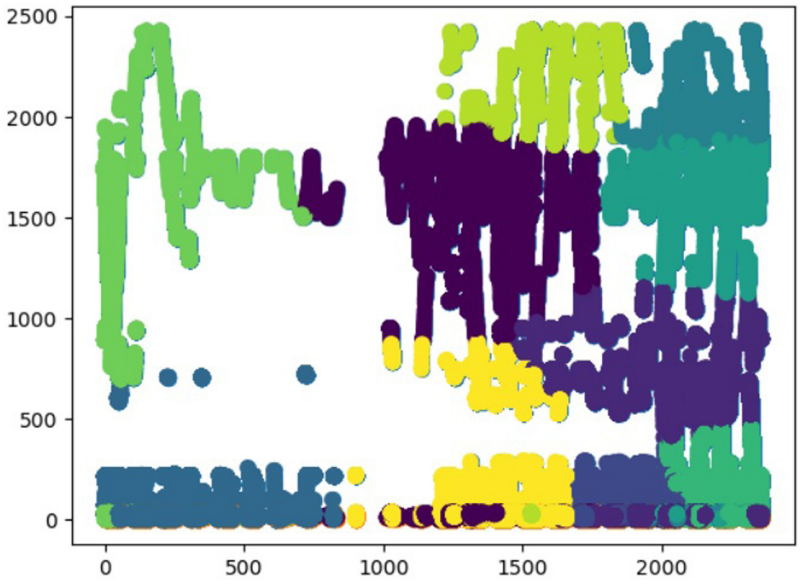

Fig. 3. Clustering results projected on time-latitude plan

The Table 2 shows the classification report of our models and includes more insights of each class. The values shown for the contextual classification represent the average of all the sixteen models for different clusters. In a binary classification scheme, a positive prediction indicates that the designated class is detected, while a negative prediction indicates that the predicted class is different. The accuracy metric allows to measure the total number of predictions a model correctly predict. The accuracies of the models lie in the interval $[92.1, 97.5]$ depending on the anomaly type and cluster.

The efficiency of the model is generally very high, but it varies depending on the anomaly type. The detection of intermittent anomalies is the easiest and the most efficient, then comes the gradual anomalies with difference of 2% in each metric, and finally the abrupt with a drop of 3%. It is worth mentioning that these results are obtained using a window size of $w = 30$.

Table 1. Impact of various contextual information on the overall detection performance

Context	Accuracy	Precision	Recall	Specificity
Temporal	93.4%	92.7%	91.5%	94.6%
Spatial	95.6%	94.5%	93.8%	95.8%
Spatio-temporal	95.7%	94.8%	93.9%	95.4%

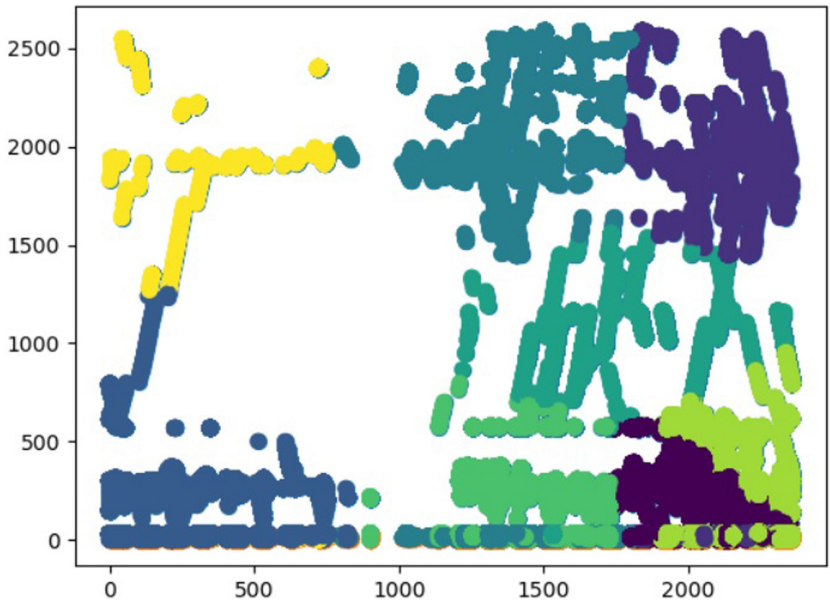

Fig. 4. Clustering results projected on time-longitude plan

The Precision score indicates the percentage of correct positive predictions among all positive predictions, this metric is useful for evaluating that the model makes correct positive predictions and not many false positive predictions. The recall score evaluates how many positive predictions are correct among the correct positive predictions, it is not affected by false-positives and is used to measure positive predictions in isolation. The Specificity score evaluates the rate of correct negative predictions, it is exactly like the Recall, but for negative predictions instead. F1-score is calculated from a balance of precision and recall. The precision increases with an average of 2.6%, the recall with 2.1% and F1-score with 2.3%.

Table 2. Impact of the contextual information on detection performance of different anomalies

Class	Precision		Recall		F1-score	
	Without context	Contextual	Without context	Contextual	Without context	Contextual
Abrupt	90.6	93.4%	89.5	91.7%	89.6	91.5%
Intermittent	94.8	96.2%	93.1	94.4%	92.3	93.8%
Gradual	91.7	94.1%	89.1	92.7%	89.7	92.9%

5 Conclusion

Anomaly detection is a crucial task required to ensure the safety and security of connected autonomous vehicles. The objective of this paper was to provide an approach based on deep learning that enhance the anomaly detector performance using the contextual information of the data received through VANET from the other vehicles. The anomalies are then detected in real-time and their type are identified accurately. The approach uses LSTM auto-encoder to reduce the data dimensions before feeding it to Grey Wolf Optimizer based clustering. To determine if a received data is anomalous, we determine first its cluster then the corresponding model network is used. In fact, each cluster has its own model which consists of LSTM auto-encoder and CNN based classifier. The error quantified by MAPE is obtained by the auto-encoder and plays the role of representative feature to the CNN based classifier which helps to distinguish normal and abnormal data. The results show good improvement comparing to the non-contextual anomaly detection. As perspective of this work, we aim to include more specific contextual events, such as festivals, sporting activities and traffic accidents. By doing this, we can handle the false alarms generated by the anomaly detection system.

References

1. Miller, C., Valasek, C.: Adventures in automotive networks and control units. Def Con **21**, 260–264 (2013)
2. Mirjalili, S., Mirjalili, S.M., Lewis, A.: Grey wolf optimizer. Adv. Eng. Softw. **69**, 46–61 (2014)
3. Kumar, V., Chhabra, J.K., Kumar, D.: Grey wolf algorithm-based clustering technique. J. Intell. Syst. **26**(1), 153–168 (2017)
4. Alahmed, A., Taiwo, S., Abido, M.: Implementation and evaluation of grey wolf optimization algorithm on power system stability enhancement. In: 2019 IEEE 10th GCC Conference and Exhibition (GCC). IEEE, April 2019
5. Ding, N., Ma, H., Zhao, C., Ma, Y., Ge, H.: Driver's emotional state-based data anomaly detection for vehicular ad hoc networks. In: 2019 IEEE International Conference on Smart Internet of Things (SmartIoT). IEEE, August 2019
6. Garip, M.T., Lin, J., Reiher, P., Gerla, M.: SHIELDNET: n adaptive detection mechanism against vehicular botnets in VANETs. In: 2019 IEEE Vehicular Networking Conference (VNC). IEEE, December 2019
7. Ghaleb, F.A., Aizaini Maarof, M., Zainal, A., Rassam, M.A., Saeed, F., Alsaedi, M.: Context-aware data-centric misbehaviour detection scheme for vehicular ad hoc networks using sequential analysis of the temporal and spatial correlation of the consistency between the cooperative awareness messages. Veh. Commun. **20**, 100186 (2019). https://doi.org/10.1016/j.vehcom.2019.100186
8. Ma, K., Wang, H.: Influence of exclusive lanes for connected and autonomous vehicles on freeway traffic flow. IEEE Access **7**, 50168–50178 (2019)
9. Nie, L., Wang, H., Gong, S., Ning, Z., Obaidat, M.S., Hsiao, K.F.: Anomaly detection based on spatio-temporal and sparse features of network traffic in VANETs. In: 2019 IEEE Global Communications Conference (GLOBECOM). IEEE, December 2019

10. Singh, P.K., Gupta, S., Vashistha, R., Nandi, S.K., Nandi, S.: Machine Learning Based Approach to Detect Position Falsification Attack in VANETs. In: Nandi, S., Jinwala, D., Singh, V., Laxmi, V., Gaur, M.S., Faruki, P. (eds.) ISEA-ISAP 2019. CCIS, vol. 939, pp. 166–178. Springer, Singapore (2019). https://doi.org/10.1007/978-981-13-7561-3_13
11. Coelho, M.C., Guarnaccia, C.: Driving information in a transition to a connected and autonomous vehicle environment: Impacts on pollutants, noise and safety. Transp. Res. Procedia **45**, 740–746 (2020)
12. Fri, M., Douaioui, K., Tetouani, S., Mabrouki, C., Semma, E.A.: A DEA-ANN framework based in improved grey wolf algorithm to evaluate the performance of container terminal. In: IOP Conference Series: Materials Science and Engineering, vol. 827, p. 012040, June 2020
13. Khot, A., Dave, M.: Position Falsification Misbehavior Detection in VANETs. In: Marriwala, N., Tripathi, C.C., Kumar, D., Jain, S. (eds.) Mobile Radio Communications and 5G Networks. LNNS, vol. 140, pp. 487–499. Springer, Singapore (2021). https://doi.org/10.1007/978-981-15-7130-5_39
14. Kopelias, P., Demiridi, E., Vogiatzis, K., Skabardonis, A., Zafiropoulou, V.: Connected and autonomous vehicles - environmental impacts - a review. Sci. Total Environ. **712**, 135237 (2020)
15. Oucheikh, R., Fri, M., Fedouaki, F., Hain, M.: Deep real-time anomaly detection for connected autonomous vehicles. Procedia Comput. Sci. **177**, 456–461 (2020). https://doi.org/10.1016/j.procs.2020.10.062
16. Peri, N., et a l.: Towards real-time systems for vehicle re-identification, multi-camera tracking, and anomaly detection. In: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). IEEE, June 2020. https://doi.org/10.1109/cvprw50498.2020.00319
17. Qu, X., Yu, Y., Zhou, M., Lin, C.T., Wang, X.: Jointly dampening traffic oscillations and improving energy consumption with electric, connected and automated vehicles: A reinforcement learning based approach. Appl. Energy **257**, 114030 (2020)
18. Wang, W., et al.: Vehicle trajectory clustering based on dynamic representation learning of internet of vehicles. IEEE Trans. Intell. Transp. Syst. **21**, 1–10 (2020). https://doi.org/10.1109/TITS.2020.2995856
19. WY Department, of Transportation: WY DOT Connected Vehicle Pilot: Improving Safety and Travel Reliability on 1–80 in W (2020). https://wydotcvp.wyoroad.info