# Leveraging Neuro-Inspired Reinforcement Learning for Secure Reputation-based Communication in Connected Vehicles

Heena Rathore
*Department of Computer Science*
*Texas State University*
San Marcos, Texas
heena.rathore@ieee.org

Henry Griffith
*Department of Engineering*
*San Antonio College*
San Antonio, Texas
hgriffith5@alamo.edu

2023 IEEE Conference on Communications and Network Security (CNS) | 979-8-3503-3945-1/23/$31.00 ©2023 IEEE | DOI: 10.1109/CNS59707.2023.10289058

***Abstract*—Secure communications in connected vehicles (CV) is essential to ensure the safety of drivers, passengers and pedestrians. As vehicles are becoming more connected and autonomous, they are reliant on communication and data exchange with infrastructure and other vehicles. Although Public Key Infrastructure has the potential to offer secure communication, it does not have ground truth information of vehicle location. Reputation-based communication can provide a more reliable approach to securing communication in a dynamic and constantly changing environment. This paper proposes a neuro-inspired reinforcement learning (RL) approach for reward estimation in CV networks. Vehicles estimate the reputation of neighboring vehicles by comparing broadcasted kinematic data with onboard sensor estimates along with connectivity topology inspired from brain, thereby forming a local graphical representation with reputation distribution. This information is shared with a centralized RL agent, which provides reward signals to each vehicle from combined reputation scores to incentivize accurate reputation estimates.**

***Index Terms*—connected vehicles, security, reinforcement learning, graph neural networks.**

## I. Introduction

Connected vehicles (CVs) have the potential to revolutionize transportation by improving driver safety and reducing congestion [1]. These benefits are achieved through cooperative sensing and maneuvering across the CV network, enabled by the periodic exchange of navigation and traffic messages, specifically basic safety messages (BSMs) [2]. However, the exchange of these messages also makes CVs vulnerable to attacks that could disrupt the safety of the network [3]. Large-scale CV networks are susceptible to various types of attacks, including external attacks like denial-of-service and spoofing, as well as internal attacks like BSM falsification and controller area network bus hacking [4]. Detecting and preventing these attacks is crucial for ensuring the safety and security of CV networks [5].

Cryptography-based algorithms have traditionally been used to detect external attacks, while machine learning (ML) and consensus-based reputation systems are used to detect internal attacks [6]. Though ML have gained success in CV network, they do not take into consideration dynamic environments with constantly changing interactions and relationships between vehicles. Recently, reinforcement learning (RL) has been studied to address the challenge of a fixed number of vehicles with a fixed observation and action space size [7]. Similarly, it is becoming a popular tool for investigating navigation in cooperative decision-making [8]. RL relies on using rewards to encourage the learning of desired behaviors, through a combination of trial-and-error interactions with the environment and the internal motivation provided by the brain's reward system. We believe that the neural representation of this reward mechanism has potential benefits for developing novel RL algorithms suited that can be tailored for CV [9].

Chen et al. [7] proposed a graph based RL for cooperative control and information flow among vehicles. Here, the authors use a state space comprised of speed, position, location, intention of the vehicles, along with an adjacency matrix depicting how the vehicles are connected to each other. The methodology does not capture the connectivity, flow and clusters which can further enhance the reliability of the entire network. Similarly in [6], the authors use RL combined with Dampster Shafer theory for updating reputation policy in CV network. However, the technique can lead to overestimation of the degree of belief in a hypothesis, which can result in incorrect decisions. Additionally, the technique is computationally expensive and may not scale well to large datasets.

This paper addresses the above listed issues by proposing a framework for reward estimation in CV networks inspired from neuroscience. The graph structure of [7] is modified to capture qualitative aspects of the network as a whole [10]. In addition, estimates of reputation formed at the vehicle level are also integrated into the algorithm, which leverage the ability of vehicles to validate peer-reported BSM broadcasts using their onboard sensors. Vehicles learn to formulate accurate estimates of reputation through a feedback reward framework which fuses two information sources-1) the central entity, incorporates Laplacian-based variable reward scheme and 2)

979-8-3503-3945-1/23/$31.00 ©2023 IEEE

TABLE I: Qualitative comparative analysis of trust and reputation models in CV

| Type | Description | Advantages | Disadvantages |
|---|---|---|---|
| Bayesian networks [17] | A probabilistic framework for representing the interactions and relationships between vehicles and the infrastructure | Able to handle uncertainty and can be used to infer the trustworthiness of a vehicle when there is incomplete or uncertain data | Computationally expensive and may not scale well to large systems with many vehicles, they require a significant amount of data to train and test |
| Machine learning [18] | Learn from data and can be used to infer the trustworthiness of a vehicle based on its past behavior and the behavior of other vehicles in the system | Can be easily scaled to handle a large number of vehicles and interactions in the system, can handle non-linear relationships and can automatically extract features from the data | Require a significant amount of data to train and test, may also be vulnerable to adversarial attacks if the training data is not representative of the real-world scenario |
| Social network analysis [19] | Uses techniques from graph theory to represent the interactions and relationships between vehicles and the infrastructure | They can provide insights into the structure and dynamics of the interactions between vehicles | Computationally expensive and may not scale well to large systems with many vehicles. |
| Game theoretic approaches [20], [23] | Use game theory to model the interactions between vehicles in the system | They can provide insights into the incentives and motivations of vehicles | Complex to model and analyze, and may not accurately reflect the real-world interactions between vehicles |
| Cryptographic methods [11], [12], [24] | Use cryptographic techniques such as digital signature and encryption to establish trust between vehicles and the infrastructure | A secure and tamper-proof way to establish trust and can be used to authenticate vehicles and protect sensitive information | Computationally expensive, particularly for large systems with many vehicles |
| Distributed ledger techniques [21] | Use blockchain technology to create a tamper-proof and decentralized ledger of trust and reputation scores. | Secure and decentralized | Require large amount of memory to store the transactions |

the individual vehicle, where estimates of trustworthiness (denoted as reputation for distinction) are computed by validating reported BSM data of nearby peers from onboard sensor readings.

## II. Related Work

Public Key Infrastructure (PKI) is a security framework used in CV that employs cryptographic algorithms like encryption and digital signatures to ensure the confidentiality and integrity of data transmitted between vehicles [11], [12]. However, PKI may not be well-suited for the highly dynamic and decentralized environment of CV and often lacks ground truth information about the vehicles [13], [14]. To address these issues, trust and reputation models can be used to analyze data transmitted from vehicles through BSM and identify and isolate security breaches [15]. By establishing trust between vehicles, efficient and safe interactions can be enabled. Trust and reputation mechanisms can also be used to detect and prevent spoofing attempts by malicious actors who may try to impersonate other vehicles or infrastructure to launch attacks [16]. Several approaches, including game theory, social networks, machine learning, and Bayesian networks, can be employed in trust and reputation models for CV. This section provides the state of the art of trust and reputation algorithms that are utilized in CV [22].

Bayesian networks provide a probabilistic framework for representing interactions between vehicles and can handle incomplete or uncertain data. The authors in [17] proposed a Gaussian-distribution-based trust management model which utilizes a Bayesian network. Here, the direct and indirect trust values were combined into a final trust value, which was further enhanced by incorporating third-party recommendations. To evaluate the model's effectiveness, the authors simulated an on-off attack in Matlab and used trust value and detection time as the performance metrics. The results showed that the model had a shorter detection time and higher accuracy in detecting the on-off attack.

Machine learning (ML) can handle large amounts of data, non-linear relationships, and extract features automatically. In [18], the authors present artificial intelligence and statistical data classification framework to analyze messages in CV. The model is trained on the US Department of Transportation Safety Pilot Deployment Model, which integrates a ML algorithm and a local trust manager. Experimental results show that the trained model can accurately predict false alerts, achieving a 98% accuracy rate and a 0.55% standard deviation on 25% malicious data.

Social network analysis uses graph theory to identify potential security threats. In [19], the authors utilize social network-based bootstrapping techniques for trust management in CV networks. The model incorporates initial trust values, node similarity, and a trust and reputation management system that considers the behavior and history of nodes for long-term trust establishment. Simulations conducted with Colt libraries in Java show that the proposed approach is resilient to up to 60% of malicious nodes and achieves a worst-case accuracy of 74%.

Game-theoretic approaches model the interactions between vehicles like in [20], which explores the practicality of evaluating reputation management schemes for CV under dynamic and diverse attack scenarios using evolutionary game theory-based solutions. The authors simulate a CV network to introduce malicious actors with randomly initialized attack plans and compare the results of trust factor gained using evolutionary decision making versus static decision making.

Blockchain-based algorithms create a decentralized ledger of trust and reputation scores. In [21], the authors propose a multi-tier authentication and trust-building framework that

leverages blockchain technology to enhance the safety and validity of exchanged information in CV. The experiments were conducted by mimicking platoon structure and the simulation results showed the tradeoff between blockchain mining time and the number of generated blocks, as well as the impact of vehicle speed on block generation. The comparative analysis of trust and reputation models in CV is shown in Table I. In this paper, we use neuro-inspired RL for trust and reputation systems in CV for several reasons:

- CV are dynamic environments with constantly changing interactions and relationships between vehicles. Though RL is better suited to handle dynamic environments they are not able to generalize well for dynamic complex environment. Proposed work can adjust the behavior to changing conditions in the environment, similar to how the brain adjusts its activity in response to changing stimuli.
- In CV, the trust and reputation of vehicles can be uncertain, which can make it difficult to train ML algorithms. RL algorithms can handle uncertainty more effectively, as they can learn from their interactions with the environment and adapt to changing conditions. In the case of reputation building in CV which rely on the collective feedback of multiple vehicles to evaluate the behavior of individual vehicles in the network, fixed RL reward function may fail. Reward function inspired from neuroscience can accelerate learning convergence rates and improve the prediction accuracy.

## III. Methods

*a) Formation of Graph:* In order to make proactive and safe decisions, the CV needs not only information pertaining to vehicles in proximity but also information pertaining to vehicles which are far away. This makes the CV network as a complex network like brain. Thus, at any time-step $t$, each vehicle would have a set of $N$ vehicles sharing local information and a centralized authority (RSU) having global information of $G$ vehicles where $G > N$. The dynamic topology of vehicles can be represented using a graph $(N, R \subseteq N \cdot N)$ consisting of a set of vehicles $N$ and a relation $R$ that specifies a directed edge from a vehicle $n$ to another one $m$ whenever they are in sensing range with each other and $R$ is equivalent to the reputation score. This information is shared to RSU where it computes the laplacian of the combined computed graph to know the continuous measure of how well the graph is connected [10].

*b) Neuro-inspired Reward Estimation:* In addition to developing the foundational theoretical algorithm for graph based learning, we also leverage variable reward structure in RL to mimic brain signals to speed up learning [25]. This is based on the hypothesis that each vehicle will share the same dynamics of the environment, however, will have different rewards based on how they have made the reputation decisions [32], [33]. In traditional RL algorithms, an agent transitions between different states $s_t$ by taking actions $a_t$ that leads to maximum rewards $r(t)$.

The state space contains reputation scores reported by individual vehicles through BSM. Each vehicle calculates the reputation score based on the reported value by the sensed vehicle vs the actual value estimated by peer vehicle (See Figure 1 (a)). RSU stores the reputation matrix $(R_{i,j})$ reported by

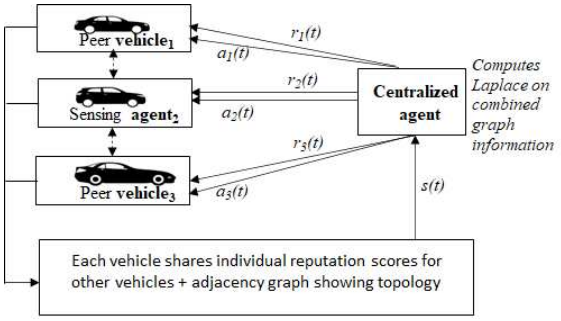



Fig. 1: Proposed framework

vehicles and uses the beta distribution to model the accuracy of each vehicle in providing information about other vehicles. The beta distribution is a continuous probability distribution with two positive shape parameters, denoted by $\alpha$ and $\beta$ where $\alpha, \beta > 0$. given by [27]:

$$f(x; \alpha, \beta) = (1/B(\alpha, \beta)) * x^{\alpha-1} * (1-x)^{\beta-1} \quad (1)$$

where $x$ is the random variable $0 < x < 1$, and $B(\alpha, \beta)$ is the Beta function, defined as [27]:

$$B(\alpha, \beta) = \Gamma(\alpha) * \Gamma(\beta)/\Gamma(\alpha + \beta) \quad (2)$$

where $\Gamma$ is the Gamma function [28]. We model the prior and posterior belief about the accuracy of each vehicle. The two shape parameters, $\alpha$ and $\beta$, are initialized to 1 for each vehicle, which represents our initial belief about the accuracy of each vehicle. The values of $\alpha$ and $\beta$ can be thought of as representing the number of "high" reputation and "low" reputation for each vehicle. Once we have chosen the prior values of $\alpha$ and $\beta$, we can update the distribution as we receive new evidence. The updated values of $\alpha$ and $\beta$ then represent our posterior belief about the accuracy of the vehicle, based on both the prior belief and the new evidence. During the simulation, the beta distribution is updated based on the prior and likelihood as below [29]:

$$posterior^i_\alpha = prior^i_\alpha + likelihood^j \quad (3)$$

$$posterior^i_\beta = prior^i_\beta + (1 - likelihood^j) \quad (4)$$

The likelihood is calculated based on the reputation score of vehicle *'i'* from the perspective of neighboring vehicle *'j'*. For each of the vehicle, the reputation is updated using:

$$R_j = w_j * prior^j_\alpha + (1 - w_j) * R_{i,j} \quad (5)$$

Here '$w$' is the weight maintained by the RSU for each of the vehicle and $prior$ is the current prior. Vehicles with higher

confidence (lower variance in their reputation scores) are given more weight. The RSU agent adjusts its weights over time based on the discrepancy between the predicted value and the actual reputation score estimated by RSU. It penalizes the vehicle which has variance in the reputation values as compared to RSU. The rewards to individual vehicles are given as:

$$r_i(t) = L(t) * w_i(t) \tag{6}$$

where $L(t)$ is the Laplacian matrix as defined in [10] and $w_i(t)$ is the weight for vehicle $i$.

## IV. Results

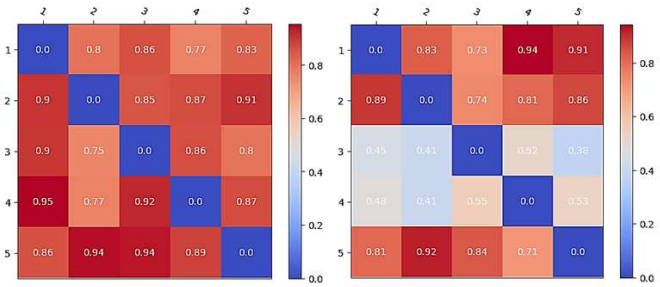
*a) Vehicle Simulation:* Experiments were done to simulate the motion of multiple vehicles in a highway setting, using the Python libraries Matplotlib and Numpy. The simulation involves a set of vehicles traveling on three lanes, with the simulation time set to 200 seconds and a time step of 0.1 seconds. The safe distance between vehicles is set to 5 meters, and the width of each lane is 3 meters. The motion model for each vehicle is defined using the position $x(t)$ and $y(t)$, velocity $v(t)$, and lane of each vehicle, and the simulation updates the positions of each vehicle based on the motion model which is described as below:

$$\begin{aligned} x(t+dt) &= x(t) + v(t) * dt \\ y(t+dt) &= lane\_width * (lane + 0.5) - y(t) \end{aligned} \tag{7}$$

where $lane$ is the current lane of the vehicle, $lane\_width$ is the width of each lane, and $dt$ is the time step size. The motion model assumes that the velocity of the vehicle is constant over the time step $dt$, and that the vehicle will maintain its desired lane position. The simulation also checks for collisions between vehicles and updates the lane positions if necessary. If the vehicles are in closer proximity with each other, adjacency matrix is updated and the reputation estimates are made.

*b) Adversary Model:* The attacker's primary objective here is to manipulate the reputation scores of vehicles in the network in order to deceive other vehicles or the overall system. Here, the malicious vehicle tamper with the GPS or sensor readings to provide inaccurate position information of peer vehicles. If the estimated position is within a threshold (0.1) from the reported position, the reputation score of both the vehicles is assigned random floating-point numbers, where each number is drawn uniformly from the range [0.7, 1). However, if the estimated position is not within the threshold, the reputation score of the reporting vehicle is assigned random floating-point numbers, where each number is drawn uniformly from the range [0.1, 0.6). Figure 2 shows the reputation matrix for a reputable and non reputable CV network.

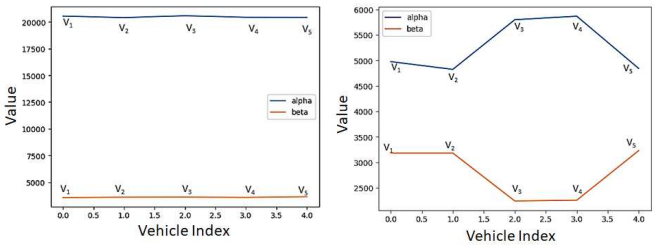
*c) RL Environment:* We then simulate a custom environment called VehicleEnv that inherits from the gym.Env class. The state of the environment includes the reputation scores of the vehicles. We also set the prior beliefs about the accuracy of each vehicle using the $\alpha$ and $\beta$ parameters of the beta distribution. The goal of the agent is to make decisions about which vehicles to trust based on their reputation scores. The environment is trained for a episode length of 100. After every episode the environment is reset to its initial state, and initializes the reputation scores of each vehicle. We implement a DQN algorithm in Python using the TensorFlow and Keras libraries for implementing RL algorithm. The DQN algorithm is implemented using a neural network with three hidden layers, each with 24 nodes and ReLU activation function. The network is trained using the Adam optimizer and Mean Absolute Error (MAE) as the metric. The training is performed for a 100 episodes, and the scores obtained in each episode are plotted using matplotlib. The DQN algorithm uses a BoltzmannQPolicy for exploration and SequentialMemory for storing the past experiences of the agent. The trained DQN agent is then tested on the environment, and the average reward obtained over 100 episodes is printed to evaluate the performance of the agent.




(a) Reputable network (b) Non reputable network

Fig. 2: Reputation score matrix (a) Five vehicles are connected to each other (b) $V_3$ and $V_4$ are malicious

*d) Results:* Figure 3 shows the number of positive and negative reputation values received for the vehicles in the reputable network by evaluating $\alpha$ and $\beta$ values. As we can see, all the vehicles have high $\alpha$ values for reputable network. Next, we simulate an environment where two vehicles are not trustworthy and start giving incorrect reputation scores to all other vehicles to increase their $\alpha$ values (See Figure 3(b)). Similarly, we also plot the cumulative reward per episode and



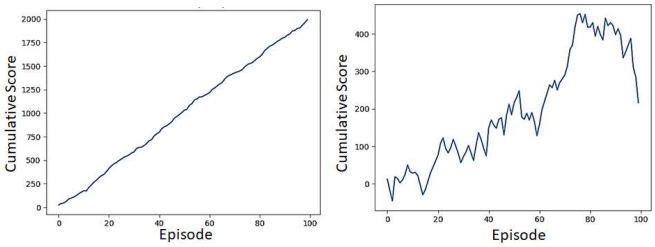

(a) Reputable network (b) Non reputable network

Fig. 3: $\alpha$ and $\beta$ values for the five vehicles

the results show that the rewards are continuously increasing (Figure 4(a)). As we can see, the cumulative reward per

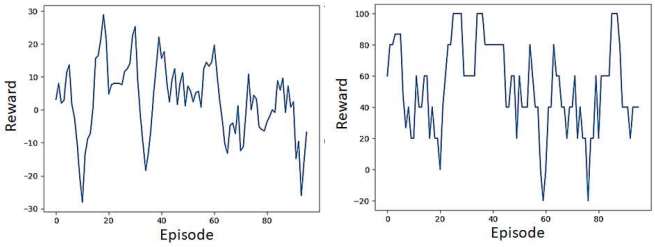
episode is less as compared to the case when all the vehicles are reputable in the network (See Figure 4(b)). We further use following metrics for performance analysis:

- *Learning curve*: Figure 5 shows the plot of the agent's performance over time, typically as a function of the number of episodes. The model improved the rewards per episode by $107.62\%$ with DQN training.
- *Mean reward*: The mean reward is the average reward obtained by the agent across all episodes (see Table II). The mean reward improved by $177.9\%$ when the agent was trained with DQN network.
- *Sample efficiency*: This metric measures how many training episodes are needed for the agent to achieve a certain level of performance. It can help identify whether the agent is able to learn efficiently with limited data. The sample efficiency of 1.01 with DQN training suggests that the algorithm requires a relatively small amount of data to achieve a certain level of performance. This was $82.51\%$ better than without training the network (see Table II).




(a) Reputable network (b) Non reputable network

Fig. 4: Cumulative score per episode of RL agent



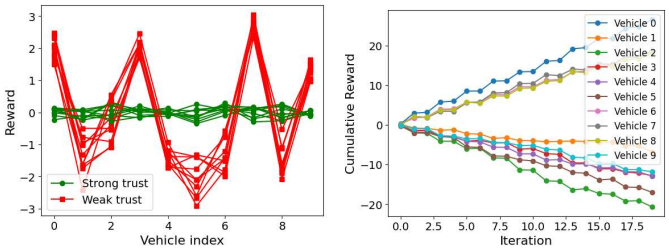

(a) Without DQN training (b) With DQN training

Fig. 5: Learning curve to evaluate agent's performance on 100 episodes

TABLE II: Performance Analysis

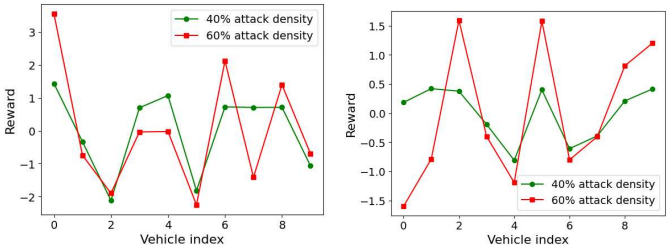
| | **Mean reward** | **Sample efficiency** |
|---|---|---|
| Without DQN training | 3.06 | 0.42 |
| DQN training | 52.34 | 1.01 |

For more comprehensive evaluation we extended the work on network of 10 vehicles where $60\%$ of the vehicles are not trustworthy. Figure 6 (a) shows the reward for different vehicles when the network is reputable (denoted by strong trust) vs non reputable (denoted by weak trust). In a strong trust scenario vehicles reward are close to each other, in contrast to weak trust scenario. Figure 6 (b) shows the cumulative reward for different vehicles when the network is non reputable. As we can clearly see the reward of trustworthy vehicle monotonically increases vs for non trustworthy vehicles. Figure 7 (a) and (b) illustrate a comparison of rewards obtained with average reputation scores of malicious vehicles set at 0.4 and 0.6, respectively. These comparisons were conducted under two different average attacker density scenarios: 40% and 60%. The proposed technique demonstrates its capability to offer adaptable rewards, even when facing an average reputation score of malicious vehicles as high as 0.6, as clearly demonstrated in the figure.




(a) Reward vs vehicle index (b) Cumulative reward vs iteration

Fig. 6: Reward performance for 60% attacker density




(a) Reputation score of 0.4 for malicious vehicles (b) Reputation score of 0.6 for malicious vehicles

Fig. 7: Reward performance for 60% attacker density

*e) Comparative Analysis:* We compare the complexity of our model with existing state of the art [6], [7]. The complexity of the Dempster-Shafer theory proposed in [6], which is used to reason with uncertain and incomplete information, can be expressed in terms of the number of focal elements in the belief function being used. The number of focal elements can grow exponentially with the number of vehicles, so the complexity of the Dempster-Shafer technique can be expressed as $O(2^n)$, where $n$ is the number of vehicles. On the other hand, the computation of the beta function proposed in our work has a complexity that can be expressed as $O(\alpha + \beta)$, which means that it grows linearly with the sum of the two shape parameters. The overall complexity of computing the probability density function of the beta distribution for $n$ vehicles is $O(n(\alpha + \beta))$ which is less in comparison to

dempster shafer technique. In this paper, we also compute the laplace of the adjacency matrix proposed in [7] which gives better representation of the network including the number of connected components, the size of the largest connected component, and the algebraic connectivity of the network.

*f) Future Directions:* For the future work, we plan to use a larger topology and dataset from [30] to evaluate the effectiveness of the reputation building in practice. For the future work, we also plan to build more comprehensive model of reputation that could potentially incorporate a wider range of capabilities such as learning factor, forgetting factor, and other principles which plays an important role in estimating the reputation scores [13], [31].

## V. Conclusion

Reputation-aided peer-to-peer communication via centralized reputation learning improves reliability in CV communication. In this paper, a neuro-inspired RL algorithm for reward estimation of CVs is proposed. The environment is designed to simulate a trust-based interaction between the agents, where the goal is to learn which agents to trust and which to avoid. The reputation scores for each agent are updated using the beta distribution model to store the prior beliefs about the accuracy of each agent's reputation score. The Laplacian of the graph was subsequently used to compute the rewards provided to each vehicle based on their estimates.

## VI. Acknowledgment

This publication has been supported by NSF CISE Research Initiation Initiative (CRII) grant #2153510 and #2313351.

## References

[1] J. Zeng et al. "Congestion and energy consumption of heterogeneous traffic flow mixed with intelligent connected vehicles and platoons". *Physica A: Statistical Mechanics and its Applications*, 609, p.128331, 2023.

[2] Y. Li et al. "Using empirical trajectory data to design connected autonomous vehicle controllers for traffic stabilization", *arXiv: 2010.05440*, 2020.

[3] G. O. Anyanwu et al., "Novel hyper-tuned ensemble random forest algorithm for the detection of false basic safety messages in internet of vehicles". *ICT Express*, vol. 9, no. 1, pp.122-129, 2023.

[4] X. Ge et al., "Resilient and safe platooning control of connected automated vehicles against intermittent denial-of-service attacks". *IEEE/CAA Journal of Automatica Sinica*, 2022.

[5] K. He, D.D. Kim, and M.R. Asghar, "Adversarial Machine Learning for Network Intrusion Detection Systems: A Comprehensive Survey". *IEEE Communications Surveys & Tutorials*, 2023.

[6] S. Gyawali, Y. Qian, and R.Q. Hu, "Deep reinforcement learning based dynamic reputation policy in 5g based vehicular communication networks". *IEEE Transactions on Vehicular Technology*, vol. 70, no. 6, pp.6136-6146, 2021.

[7] S. Chen et al., "Graph neural network and reinforcement learning for multi-agent cooperative control of connected autonomous vehicles". *Computer-Aided Civil and Infrastructure Engineering*, vol. 36, no. 7, pp.838-857, 2021.

[8] Maria Hügle et al., "Dynamic Interaction-Aware Scene Understanding for Reinforcement Learning in Autonomous Driving", *CoRR*, vol. abs/1909.13582, 2019.

[9] B.W. Chen et al., "Neuro-Inspired Reinforcement Learning to Improve Trajectory Prediction in Reward-Guided Behavior". *International journal of neural systems*, pp.2250038-2250038, 2022.

[10] S.C. de Lange, M.A. de Reus, and M.P. van den Heuvel, "The Laplacian spectrum of neural networks". *Frontiers in computational neuroscience*, 7, p.189, 2014.

[11] B. Brecht, and T. Hehn, "A security credential management system for V2X communications". *In Connected Vehicles*, pp. 83-115, 2019.

[12] W. Whyte et al., "A security credential management system for V2V communications". In *2013 IEEE Vehicular Networking Conference*, pp. 1-8, 2013.

[13] H. Rathore, A. Samant, and M. Jadliwala, "TangleCV: A distributed ledger technique for secure message sharing in connected vehicles". *ACM Transactions on Cyber-Physical Systems*, vol. 5, no. 11, pp. 1-25, 2020.

[14] H. Rathore, S. Sai, A. Gundewar, "Social Psychology Inspired Distributed Ledger Technique for Anomaly Detection in Connected Vehicles". *IEEE Transactions on Intelligent Transportation Systems*, 2023.

[15] G.D. Putra et al., "Toward Blockchain-Based Trust and Reputation Management for Trustworthy 6G Networks", *IEEE Network*, vol. 36, no. 4, pp.112-119, 2022.

[16] Z. Tu et al., "A Blockchain-based Trust and Reputation Model with Dynamic Evaluation Mechanism for IoT". *Computer Networks*, 218, p.109404, 2022.

[17] W. Fang et al.,"BTDS: Bayesian-based trust decision scheme for intelligent connected vehicles in VANETs", *Transactions on Emerging Telecommunications Technologies*, vol. 31, no. 12, p.e3879, 2020.

[18] E.O. Eze et al., "A Context-Based Decision-Making Trust Scheme for Malicious Detection in Connected and Autonomous Vehicles". in Proc. *IEEE2022 International Conference on Computing, Electronics & Communications Engineering (iCCECE)*, pp. 31-36, 2022.

[19] D. Alishev et al., "Social-aware bootstrapping and trust establishing mechanism for vehicular social networks". In *2017 IEEE 85th Vehicular Technology Conference (VTC Spring*, pp. 1-5, 2017.

[20] Z. Tian, X. Gao, S. Su, J. Qiu, X. Du and M. Guizani, "Evaluating Reputation Management Schemes of Internet of Vehicles Based on Evolutionary Game Theory," in *IEEE Transactions on Vehicular Technology*, vol. 68, no. 6, pp. 5971-5980, June 2019, doi: 10.1109/TVT.2019.2910217. 2019.

[21] F. Kandah, et al., "BLAST: Blockchain-based trust management in smart cities and connected vehicles setup". In *2019 IEEE High Performance Extreme Computing Conference (HPEC)*, pp. 1-7, IEEE, 2019.

[22] J. Zhang, "A survey on trust management for vanets". *In 2011 IEEE International Conference on Advanced Information Networking and Applications*, pp. 105-112, 2011.

[23] M. L. Y. Chiang, "Game-theoretic security and trust management in VANETs," *IEEE Communications Surveys & Tutorials*, vol. 14, no. 4, pp. 1510-1535, 2012.

[24] M.S. Rathore et al., "A novel trust-based security and privacy model for internet of vehicles using encryption and steganography". *Computers and Electrical Engineering*, 102, p.108205, 2022.

[25] N. Mehta et al., "Transfer in variable-reward hierarchical reinforcement learning". *Machine Learning*, vol. 73, no. 3, pp.289-312, 2008.

[26] J. Fan et al., "A theoretical analysis of deep Q-learning. In Learning for Dynamics and Control", textitpp. 486-489. *PMLR*, 2020.

[27] A.K. Gupta, and S. Nadarajah, "Handbook of beta distribution and its applications". *CRC press*, 2004.

[28] N.L. Johnson, S. Kotz, and N. Balakrishnan, "Beta distributions. Continuous univariate distributions". *2nd ed. New York, NY: John Wiley and Sons*, pp.221-235, 1994.

[29] J.J. Chen, and M.R. Novick, "Bayesian analysis for binomial models with generalized beta prior distributions". *Journal of Educational Statistics*, vol. 9, no. 2, pp.163-175, 1984.

[30] H. Griffith, M. Farooq, H. Rathore, "A Standardized Data Generation Workflow for Consensus-Based Connected Vehicle Security", in Proc. *IEEE 41st International Conference on Consumer Electronics*, 2023.

[31] D. Suo, and S.E. Sarma, "Real-time trust-building schemes for mitigating malicious behaviors in connected and automated vehicles". In *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, pp. 1142-1149, 2019.

[32] H. Rathore and H. Griffith, "GNN-RL: Dynamic Reward Mechanism for Connected Vehicle Security using Graph Neural Networks and Reinforcement Learning", *IEEE SmartComp*, 2023.

[33] H. Rathore and H. Griffith, "Improving Reinforcement Learning Performance through a Behavioral Psychology-Inspired Variable Reward Scheme", *IEEE SmartComp*, 2023.