Reinforcement Learning in Ad Hoc Vehicular Networks

Subjects: Computer Science, Artificial Intelligence

Contributor: Lincoln Herbert Teixeira , Árpád Huszák

Ad hoc vehicular networks have been identified as a suitable technology for intelligent communication amongst smart city stakeholders as the intelligent transportation system has progressed. In a highly mobile area, the growing usage of wireless technologies creates a challenging context. To increase communication reliability in this environment, it is necessary to use intelligent tools to solve the routing problem to create a more stable communication system. Reinforcement Learning (RL) is an excellent tool to solve this problem.

advanced vehicular ad-hoc network   routing network   reinforcement learning

1. Introduction

A vehicular ad-hoc network (VANET) allows cars and roadside devices to connect with one another [1]. Even if they have no prior knowledge of other vehicles in the region, automobiles are viewed as nodes in a self-organizing network for communication reasons [2]. This means that there is no preexisting infrastructure required for this decentralized wireless ad-hoc network. Each node acts as a host as well as a router, delivering and receiving data between nodes. To communicate between devices, dedicated short-range communication radios (DSRC) are employed [3]. DSRC is a wireless communication system based on the IEEE 802.11p standard that allows automobiles to communicate directly at high speeds and with high security without the need for cellular infrastructure. Also, the 5.9 GHz frequency is used by DSRC to allow low-latency information sharing between automobiles.

This architecture is in charge of delivering and receiving safety alerts, as well as maintaining passenger and pedestrian safety [4]. It also improves the flow of traffic and the effectiveness of the traffic management system. An onboard unit (OBU) has several sensors such as a GPS, accelerometer, resource command processor, user interface, and read/write storage for storing data. Over an IEEE 802.11p wireless connection, OBUs are in charge of communicating between surrounding devices [5]. The most difficult component of VANET is managing and routing the data required for optimum connectivity.

The data broadcasted and received by vehicular units in ad-hoc vehicular networks comprise information on the telemetry of the related cars. When information is transmitted over the air, it is very sensitive to interference, which can result in a network outage, putting the lives of drivers and anyone near to them in danger. Because there is no large infrastructure involved in a shared wireless medium, there is a high risk of car-to-car communication failure.

In multi-hop communication, nodes or vehicular units function as hosts and routers, forwarding and receiving data from other nodes [6].

As a result, nodes are chosen based on their connectivity, so the routing algorithm selected is a sensitive issue that must be handled carefully in this type of system. Routing protocols developed for legacy networks may not adequately serve vehicular networks in the near future. Therefore, new alternatives must be developed to solve the routing problem in these complex networks.

2. Reinforcement Learning in Vehicular Networks

Related work address the two traditional routing protocols widely used in ad-hoc networks, DSR and DSDV. These protocols were used in the comparative performance of the planned method. It will also address solutions related to routing path lifetime where current machine learning research is applied in VANET. Lastly, the two Reinforcement Learning algorithms used in the proposed approach were PPO and A2C.

VANETs (vehicular ad hoc networks) have become a prominent research topic in recent years. VANETs are confronted with new development patterns as new technologies arise. Advanced VANETs, which integrate regular VANETs with these upcoming technologies, can increase transportation safety and efficiency while also improving automobile owner experiences. Advanced VANETs, on the other hand, have additional obstacles. To overcome these, new architectures, procedures, and protocols must be devised [7].

Owing to stochastic node movements, interference, multipath propagation, and path loss, wireless ad-hoc mobile networks lack a consistent topology due to the absence of physical links between nodes. Many routing protocols have been suggested and are continuously being researched in order to reduce the possibility of communication failures with this technology.

VANETs are currently confronting new development trends as a result of the advent of new technologies such as 5G, cloud/fog computing, blockchain, and machine learning [8][9]. Advanced VANETs, which combine standard VANETs with these future technologies, have the potential to boost efficiency dramatically. Advanced VANETs, on the other hand, have new obstacles. To overcome them, new architectures, procedures, and protocols must be developed, as proposed by the journal [10].

Reinforcement Learning

Reinforcement learning (RL) is the process of determining what actions to take to enhance a quantitative reward value. The learner is not taught which actions to perform; rather, he is encouraged to try them all and see which ones provide the greatest outcomes. Actions can have an influence not only on the immediate, but also on a problem, a class of problem-solving techniques that successfully act on the problem, and the field that analyzes the problem and its answers, under the most intriguing and difficult situations [11].

Reinforcement learning is distinct from supervised learning, which is the focus of the majority of contemporary machine learning research. Learning from a set of labeled examples given by an expert external supervisor is known as supervised learning. Each example is a description of a scenario followed by a specification—the label—of the proper reaction the system should take in response to that condition, which is frequently to identify a category to which the situation belongs. The goal of this sort of learning is for the system to be able to generalize its responses to situations outside of the training set. While this is an important sort of learning, it falls short of what is required for learning through interaction. It is difficult to conceive of examples of desired conduct that are both correct and indicative of all the scenarios in which the agent must respond to interactive concerns. An agent must be able to learn from its own experience in an uncharted territory where learning would be most helpful [12].

RL is also separate from unsupervised learning, which is focused on detecting structure in vast volumes of unlabeled data, as defined by machine learning specialists. The terms supervised and unsupervised learning appear to categorize machine learning paradigms entirely, but they do not. While it is easy to mistake reinforcement learning for unsupervised learning, rather than discovering a hidden structure, the goal of reinforcement learning is to maximize a cumulative reward. While studying the structure of an agent's experience can help with reinforcement learning, it doesn't address the problem of maximizing a reward value.

The trade-off between exploration and exploitation is one of the issues that occur in reinforcement learning but not in other types of learning. To obtain a large number of rewards, a reinforcement learning agent must select activities that it has previously done and found to be useful in terms of reward provision. He must, however, do actions he has never performed before in order to uncover such acts. In order to receive a reward, the agent must not only examine what it has already experienced but also discover how to make better future action options. The issue is that neither exploration nor exploitation can be carried out merely for the purpose of achieving success. Before favoring the actions that appear to be the most successful, the agent must test a number of them. To get a reliable estimate of an action's anticipated reward, it must be tried multiple times on a stochastic task. Reinforcement learning also has the advantage of being able to handle any issue involving a goal-directed agent interacting with an unknown environment.

Reinforcement learning begins with a fully functional, interactive, goal-seeking agent. All reinforcement learning agents have specified goals, are able to monitor parts of their surroundings, and may influence their environments by taking actions [13]. Furthermore, it is customary to anticipate that the agent will have to function in an environment with a great deal of uncertainty from the start.

One of the most exciting aspects of today's reinforcement learning is how successfully integrates with different technological and scientific domains. Reinforcement learning is part of a long-standing trend in AI and machine learning to combine statistics, optimization, and other mathematical subjects. The ability of some reinforcement learning algorithms to train using parameterized approximates, for example, tackles a long-standing problem in operations research and control theory: dimensionality. Reinforcement learning, in particular, has had a fruitful collaboration with psychology and neuroscience, with considerable benefits for both sides. Many of the best

reinforcement learning algorithms were influenced by biological learning systems. Reinforcement learning is the closest sort of machine learning to the kind of learning that humans and other animals do.

Therefore RL has been recognized as one of the most effective optimization tools in solving complex problems. Existing RL-based systems, on the other hand, suffer from sluggish convergence for optimum communication due to the improper design of RL elements (i.e., reward and state) for complicated traffic dynamics.

Meanwhile, most optimization approaches assume that the network communication environment's phase length is constant to simplify RL modeling, which severely limits the RL's ability to seek up network route management policies with a reduced average number of hops and better communication time stability [14].

References

1. Malik, S.; Sahu, P.K. A comparative study on routing protocols for VANETs. Heliyon 2019, 5, e02340.

2. Liu, T.; Li, Y.; Wang, M.; Shi, S.; Gu, X. An optimized mobile similarity and link transmission quality routing protocol for urban VANETs. Wirel. Netw. 2021, 1–11.

3. Abboud, K.; Omar, H.A.; Zhuang, W. Interworking of DSRC and Cellular Network Technologies for V2X Communications: A Survey. IEEE Trans. Veh. Technol. 2016, 65, 9457–9470.

4. Li, J.T.; Li, X.J.; Li, Z.F.; Qiu, R. Design of Cooperative Driving System Based on VANET. In Smart Communications, Intelligent Algorithms and Interactive Methods; Jain, L.C., Kountchev, R., Hu, B., Kountcheva, R., Eds.; Springer Nature Singapore: Singapore, 2022; pp. 7–12.

5. Chatterjee, T.; Karmakar, R.; Kaddoum, G.; Chattopadhyay, S.; Chakraborty, S. A Survey of VANET/V2X Routing From the Perspective of Non-Learning- and Learning-Based Approaches. IEEE Access 2022, 10, 23022–23050.

6. Latif, S.; Mahfooz, S.; Jan, B.; Ahmad, N.; Cao, Y.; Asif, M. A comparative study of scenario-driven multi-hop broadcast protocols for VANETs. Veh. Commun. 2018, 12, 88–109.

7. Anh, D.V.; Tan, N.D.; Truong, C.D.; Hau, N.V.; Quy, V.K. Performance Analysis of Routing Protocols for Mobile Ad Hoc Networks in Urban Scenarios. J. Commun. 2021, 16, 545–552.

8. Darabkh, K.A.; Alkhader, B.Z.; Khalifeh, A.F.; Jubair, F.; Abdel-Majeed, M. ICDRP-F-SDVN: An innovative cluster-based dual-phase routing protocol using fog computing and software-defined vehicular network. Veh. Commun. 2022, 34, 100453.

9. Skondras, E.; Michalas, A.; Vergados, D.D. Mobility management on 5G Vehicular Cloud Computing systems. Veh. Commun. 2019, 16, 15–44.

10. Zhang, l.; Meng, M.; Liang, K. Sensors Special Issue: Advanced Vehicular Ad Hoc Networks. Available online: https://www.mdpi.com/journal/sensors/special_issues/Vehicular_Ad_Hoc_Networks (accessed on 13 April 2022).

11. Sutton, R.S.; Barto, A.G. Reinforcement Learning: An Introduction; MIT Press: Cambridge, MA, USA, 2018.

12. Yan, H.; Cui, Z.; Chen, X.; Ma, X. Distributed Multi-agent Deep Reinforcement Learning for Multi-line Dynamic Bus Timetable Optimization. IEEE Trans. Ind. Inform. 2022, 1.

13. Hernandez-Leal, P.; Kartal, B.; Taylor, M.E. A survey and critique of multiagent deep reinforcement learning. Auton. Agents-Multi-Agent Syst. 2019, 33, 750–797.

14. Chen, L.; He, Y.; Wang, Q.; Pan, W.; Ming, Z. Joint Optimization of Sensing, Decision-making and Motion-controlling for Autonomous Vehicles: A Deep Reinforcement Learning Approach. IEEE Trans. Veh. Technol. 2022, 71, 4642–4654.