논문 20-45-03-22 The Journal of Korean Institute of Communications and Information Sciences '20-03 Vol.45 No.03
https://doi.org/10.7840/kics.2020.45.3.622

멀티홉 V2V에서 DCNN과 Q-러닝 기반 커버리지 확장 연구

마누엘 에우제니오 모로초 카얀셀라•, 임 완 수°

Expanding the Coverage of Multihop V2V with DCNNs and Q-Learning

Manuel Eugenio Morocho-Cayamcela•, Wansu Lim°

요 약

차량 대 차량(V2V) 통신은 지리적 위치, 제동 정보, 속도, 회전 신호 상태 및 주행 방향 등 다양한 안전메시지를 송수신한다. 안전메세지 전송을 위한 통신 프로토콜은 전용단거리통신(DSRC)이며, DSRC는 통신범위 제한으로 다중 홉을 이용하여 차량에 메시지를 전송한다. 본 논문은 V2V로 구성한 네트워크에서 기계학습을 이용하여 다중 홉 연결 커버리지를 확장하는 방법을 제안한다. 먼저 심층콘볼루션신경망(DCNN)을 이용하여 다양한 환경으로 구성된 V2V 네트워크의 무선채널을 학습하고 각 환경에 적합한 전파 모델을 이용한다. 무선채널을 세분화한 후 Q-러닝을 이용하여 전파 손실이 가장 적은 최적의 다중 홉 경로를 찾음으로써 V2V 안전메세지 전송 커버리지를 확장한다.

Key Words : Vehicle-to-vehicle, machine learning, Q-learning, multi-hop, wireless coverage

ABSTRACT

One of the most critical challenge in a vehicle-to-vehicle (V2V) scenario is the transmission safety messages (BSMs) e.g., geographical location, braking information, speed, the status of the turn signal, and direction of travel. The protocol adopted to transmit BSMs in V2V is refered as Dedicated Short-Range Communications (DSRC). The limited communication range of DSRC have shown that is necessary to employ a multi-hop communication strategy to reach as many target vehicles as possible. In this paper, we overcome the coverage limitation of multi-hop connectivity in V2V networks and propose a methodology consisting of two machine learning (ML) tasks. First, two deep convolutional neural networks (DCNN) are created and tuned to segment terrestrial imagery into different environments. The multi-environments are anticipated to have different propagation models. The second part uses a Q-learning algorithm to find the optimal multi-hop path with the lowest propagation loss, based on the results of the environment segmentation. The optimal multi-hop link is simulated and compared with a direct link transmission, showing that our proposal can extend the coverage of multi-hop wireless links by transmitting the BSMs via the optimum path.

※ This work was supported by Kumoh National Institute of Technology(2019-104-155)
• First Author : Kumoh National Institute of Technology, Department of Electronic Engineering, E-mail: eugeniomorocho@kumoh.ac.kr, 학생회원
° Corresponding Author : Kumoh National Institute of Technology, Department of IT Convergence Engineering, E-mail: wansu.lim@kumoh.ac.kr, 정회원
논문번호 : 201911-261-C-RE, Received October 30, 2019; Revised December 30, 2019; Accepted December 30, 2019

Ⅰ. Introduction

A real V2V communication scenario constitutes a wireless network where automobiles exchange information about their status. This standardized communication data, also known as BSM, include geographical location, speed, the direction of travel, etc. V2V technology generally adopts DSRC, a specifically designed wireless communications channel and its corresponding set of protocols and standards for automotive use[1,2]. However, since the range of DSRC is limited to a few hundred of meters, it is necessary to employ a multi-hop communication to extend the range of communication. Under this setting, the vehicles will act as message relays in order to a given destination, and eventually reach as many targets as possible[3]. In the U.S., the IEEE 1609 WAVE Wireless Access in Vehicular Environments protocol stack builds on IEEE 802.11p WLAN operating on seven reserved channels in the 5.9 GHz frequency band. Research in VANETs started as early as 2000, in universities and research labs, having evolved from researchers working on wireless ad hoc networks[4,5]. Many researchers have focused on media access protocols, routing, warning message dissemination, and VANET application scenarios, but there is a lack of literature on the application of ML algorithms to solve propagation problems on vehicular networks.

In this paper, we propose to:

- Build and train a semantic segmentation network to create a categorical matrix based on an aerial imagery from the location of the vehicles.
- Set the transmitting vehicle as the agent and the segmented aerial image as the environment, to be able to train a model-free RL algorithm capable of finding the optimal path among the vehicles that minimizes the path loss in the wireless link.
- Simulate the scenario under a V2V channel and compare the dissemination of a wireless signal using our proposal, and a direct link transmission.

Ⅱ. Proposed Methodology

The proposed methodology consists in two ML tasks, the first is creating and tuning two deep convolutional neural networks (DCNN) to segment the terrestrial environment into sub-environments (urban, suburban, and rural), with different propagation models. The second part is the application of a reinforcement-learning algorithm to find the path with the lowest propagation loss, in order to establish the multi-hop link.

2.1 Semantic Segmentation with DCNNs

With the popularity of deep learning in recent years, many problems in the field of wireless communications are being tackled using artificial intelligence and DCNN architectures[6,7]. In our problem, even though a free-space propagation model could be used, in a real-world scenario we find obstructions like buildings or trees. We employ a semantic segmentation network to find where these obstructions are positioned and solve the scene understanding problem. Figure 1 shows at a high-level a semantic segmentation network on top, where the output matches every pixel in the image with one of our classes (i.e., open field, buildings, and street). The DCNN takes the input image through several processes (i.e. convolution, batch normalization, image down-sampling, ReLu activation, etc), and repeat them until the last layer computation. The first pre-trained network is followed by an up-sampling network, with a reverse architecture of a normal DCNN, where the series of new layers up-sample the results of the first pre-trained network back into an image[8]. As a result, the input is an image and the result is the image with every pixel labelled with a pre-set class. Pairs of images with their corresponding labels are given to the algorithm during the initial learning process, and once the segmentation network has learned the features, the network will automatically do the inference on any new input image. Our semantic segmentation network parameters initialize using the weights of the VGG16 architecture with an encoder-decoder framework, dropping the fully connected layers of the network[9]. The decoder sub-network is constitutes as mirror copy of the encoder sub-network.

2.2 The Reinforcement Learning Task

At a high level, our reinforcement learning routine comprises two main parts, an agent, and the environment (i.e., the categorical matrix generated by our DCNN in Section II.A). Figure 1 illustrates how the agent starts by taking an action A_i on the environment and receives a reward R_{i+1} for every time step S_{i+1}[10].

The value of the reward R_{i+1} can be positive or negative (interpreted as a penalty). The action of the agent is taken based on some policy γ, that is $A_i = \gamma(S_i)$. Our final goal is to find the policy

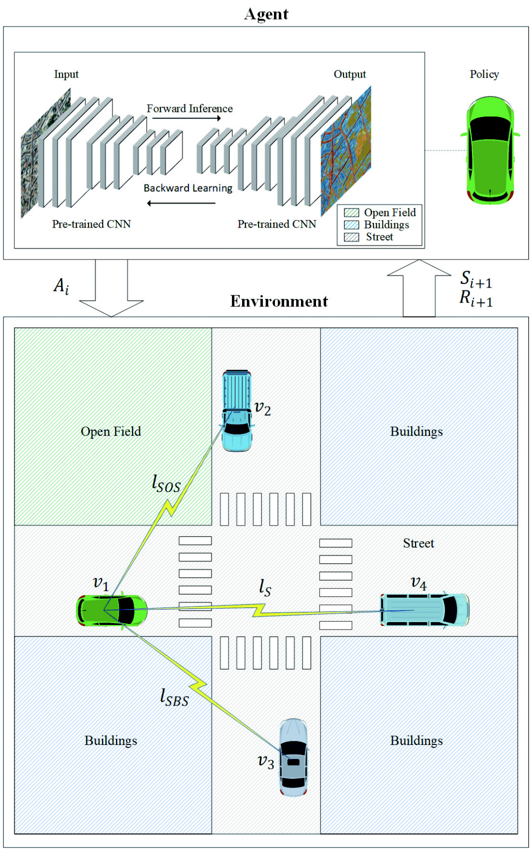


Fig. 1. An illustration of the RL algorithm, showing the agent taking an action A_i on the environment, and receiving a new state S_{i+1}, and a reward R_{i+1}.

that maximizes the cumulative reward $\sum_i^N R_i$ over a finite number of iterations N. The agent will take a deterministic policy, in other words, the agent will observe the state of the system and then it will choose an action. The RL problem can be mathematically represented as a Markov Decision Process (MDP), consisting on a finite set of states S, a set of allowable actions A_i for each $i \in S$.

A transition function $T: S \times A \rightarrow S$, and a reward function $R: S \times A \rightarrow \mathcal{R}$. We define the policy as $\pi: S \times A \rightarrow S$, and denote $R(t)$ as the reward in time t. Therefore, the objective will be to learn the policy that maximizes the cumulative reward $r(0) + \gamma r(1) + \gamma^2 r(2) + \ldots$, where $\gamma \in (0,1)$ represents a discount factor, in order to make the sum converge. To evaluate the quality of our policy π, a value function $V: S \rightarrow \mathcal{R}$ assigns a real value to each of the states, as follows:

$$V^{\pi}(s) = \sum_{t=0}^{\infty} \gamma^t r(t) \tag{1}$$

where the state at time t, is generated from the state at time $t-1$, by applying the action according to the policy $\pi(s_t - 1)$. Our problem will be solved when the algorithm finds an optimal policy that satisfies the condition in (2).

$$V^{\pi^*}(s) = \max_{a \in A_s}[R(s,a) + \gamma V^{\pi^*}(T(s,a))] \tag{2}$$

Our agent can take four different transmitting actions, we write it as $a \in A\{\mathrm{up}, \mathrm{down}, \mathrm{left}, \mathrm{right}\}$. In Fig. 1, the agent vehicle v_1 will receive the highest reward if it transmits through a street, since the propagation loss experienced in l_s is the lowest. If v_1 transmits through an open field (e.g., large parks, or areas that might include vegetation), the agent will also receive a reward, but the value will be lower than the previous case. On the contrary, if v_1 decides to transmit across buildings, where the propagation loss is the highest, the agent will receive a negative reward. The propagation loss can be a sum of sub-propagation estimations, for instance Fig. 1 illustrates the transmission between v_1 and v_2, where the propagation will be estimated as the sum of the propagation in the classes street-buildings-street l_{SBS}. In the same manner, l_{SOS} in Fig. 1 indicates a propagation across the evironments street-openfield-street.

The proposed algorithm converges to the optimal value V^*, and policy π_t provided that $\gamma \in (0,1)$. Expressing the values in terms of Q-functions, the value of the policy π that starts in the state s and takes an action a, is defined as $Q^{\pi}: S \times A \rightarrow \mathcal{R}$. For an optimal policy π^*, the definition represented in (3)-(4) needs to be satisfied.

$$Q^{\pi}(s,a) = R(s,a) + \gamma V^{\pi^*}(T(s,a)) \tag{3}$$

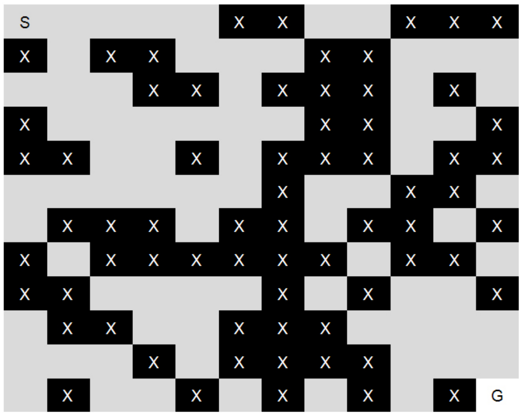
$$V^{\pi^*} = \max_a Q^{\pi^*}(s,a) \tag{4}$$

Since the transition and reward functions are not assumed to be known in advance, the agent will be the learner, whose task is to maximize its rewards.

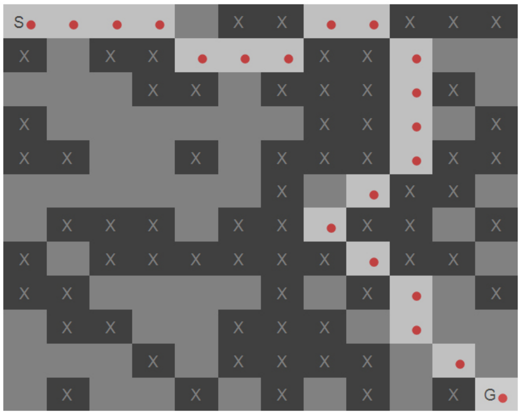
Ⅲ. Simulation Results and Discussion

The semantic segmentation network was fed using a dataset that accounts for aerial images from different locations on different cities[11]. To simulate the received power, we have gathered a dataset containing the received signal strength (RSSI) from a V2V scenario[12]. From each location, we have segmented an aerial image and set the value of the reinforcement-learning algorithm to 1 to streets, 0 for the open areas, and -1 for buildings. The problem now can be solved as a maze problem, and the optimal path can be found with a relative low computational complexity.

As an illustration, Fig. 2-(a) shows the results of the categorical matrix reduced to a 12×12 pixel area. The black squares with an “x” represents the buildings our algorithm should avoid. The grey squares represents the rewards ≥ 0. Figure 2-(b) shows the results of the optimal path for the transmission. We assume the transmitting vehicle is located in the top-left of the maze, and the location of the goal to be the square in the bottom right corner.



(a)



(b)

Fig. 2. The multi-hop problem solved as a maze game with Q-learning. (a) A map showing the start point and the goal , where represents a location area to be evaded due to high propagation losses. (b) The solved maze, with the optimal multi-hop path solution.

Figure 3 shows the results of out MATLAB simulation. We can notice that while the wireless link is lost at 300m when the multi-hop path is not optimal, the optimal path connectivity is not compromised on distances above 500m. Furthermore, the received power for the vehicles on the optimal path is higher than for the vehicles in a

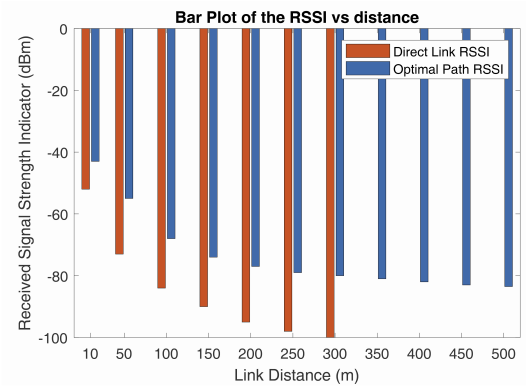


Fig. 3. RSSI comparison between direct link and proposed method.

sub-optimal route. Since the multi-propagation model is based on the propagation scenarios from the Okumura-Hata model[13], this research work can be extended to map-based and stochastic channel and propagation models, multi-hop routing, path planning, etc.

Ⅳ. Conclusions and Future Work

A combination of two ML algorithm has been developed to enhance transportation researchers and analysts to study BSM data transmission by creating multi-hop links within the path that presents the lowest propagation loss and lowest energy consumption. We have used two DCNNs to create a semantic segmentation network that will feed the Q-learning algorithm with the specific value function according to the environment of the specific terrestrial location. The methodology presented is discussed for the V2V scenario, but it can be used in additional applications such as device-to-device (D2D), relay nodes communication, nomadic node links, etc. The proposed solution has the potential to enhance the reliability of the link in vehicle-to-everything (V2X) communications, and the RL technology can be further studied to improve multi-hop network routing, or path planning.

References

[1] M. E. Morocho-Cayamcela and W. Lim, "Finding the optimal path for V2V multi-hop connectivity with Q-learning and Convolutional Neural Networks," in *Proc. KICS Int. Conf. Commun.*, Jun. 2019.

[2] G. Karagiannis, et al., "Vehicular networking: A survey and tutorial on requirements, architectures, challenges, standards and solutions," *IEEE Commun. Surv. Tuts.*, vol. 13, no. 4, pp. 584-616, 2011.

[3] R. He, et al., "Vehicle-to-vehicle radio channel characterization in crossroad scenarios," *IEEE Trans. Veh. Technol.*, vol. 65, no. 8, pp. 5850-5861, Aug. 2016.

[4] W. Song, F. Zeng, J. Hu, Z. Wang, and X. Mao, "An unsupervised-learning-based method for multi-hop wireless broadcast relay selection in urban vehicular networks," *IEEE Veh. Technol. Conf.*, Sydney, NSW, Australia, 2017.

[5] N. Taherkhani and S. Pierre, "Centralized and localized data congestion control strategy for vehicular ad hoc networks using a machine learning clustering algorithm," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 11, pp. 3275-3285, Nov. 2016.

[6] M. E. Morocho-Cayamcela and W. Lim, "Artificial intelligence in 5G technology: A survey," in *2018 ICTC*, no. 1, pp. 860-865, 2018.

[7] M. E. Morocho-Cayamcela, H. Lee, and W. Lim, "Machine learning for 5G/B5G mobile and wireless communications: Potential, limitations, and future directions," in *IEEE Access*, vol. 7, pp. 137184-137206, 2019. doi: 10.1109/ACCESS.2019.2942390

[8] V. Badrinarayanan, A. Kendall, and R. Cipolla, "Segnet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 12, pp. 2481-2495, 2015.

[9] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *Int. Conf. Learn. Represent.*, pp. 1-

14, 2015.
[10] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, 2nd Ed., The MIT Press, London, England, 2018.
[11] E. Maggiori, Y. Tarabalka, G. Charpiat, and P. Alliez, “Can semantic labeling methods generalize to any city? The inria aerial image labeling benchmark,” *2017 IEEE Int. Geosci. Remote Sens. Symp.*, pp. 3226-3229, 2017.
[12] Z. H. Mir and F. Filali, “Simulation and performance evaluation of vehicle-to-vehicle (V2V) propagation model in urban environment,” in *2016 7th ISMS*, pp. 394-399, 2016.
[13] A. F. Molish, “*The Okumura – Hata model, in Wireless Communications*,” Second Ed., West Sussex: John Wiley & Sons, Inc., 2011.

마누엘 에우제니오 모로초 카얌셀라
(Manuel Eugenio Morocho Cayamcela)

2012 : B.Eng. in electronic engineering, Universidad Politécnica Salesiana, Cuenca, Ecuador.
2016 : M.Sc. in communications engineering and networks, The University of Birmingham, England, United Kingdom.
2017~present : Research assistant and Ph.D. candidate, Kumoh National Institute of Technology, Gumi, South Korea.
<Research interests> communications engineering and networks, artificial intelligence, signal processing, statistical analysis, and optimization.
[ORCID:0000-0002-4705-7923]

임 완 수 (Wansu Lim)

2007 : M.Sc in electronic engineering, GIST, South Korea.
2010 : Ph.D. degree, GIST, South Korea.
2010~2013 : Research Fellow at the University of Hertfordshire, U.K.
2013~2014 : Postdoctoral Researcher at the Institut national de la recherche scientifique (INRS), Canada.
2014~present : Associate Professor at Kumoh National Institute of Technology, South Korea.
<Research interests> integrated optical/wireless access networks, device-to-device (D2D) communications, IoT, sensor networks, and artificial intelligence.
[ORCID:0000-0003-2533-3496]