Fuzzy Q-learning Based Vertical Handoff Control for Vehicular Heterogeneous Wireless Network

Yubin Xu[1], Limin Li[1, 2], Boon-Hee Soong[2] and Cheng Li[3]
[1]Communication Research Center, Harbin Institute of Technology, Harbin, China
[2]INFINITUS, School of Electrical & Electronic Eng., Nanyang Technological University, Singapore
[3]Electrical & Computer Eng., Faculty of Engineering, Memorial University, St. John's, NL, Canada

Abstract—**As a novel fundamental platform for providing real time access to wireless network, vehicular communication is drawing more and more attentions in recent years. IEEE 802.11p is a main radio access technology which supports communication for high mobility terminals. Due to the limited coverage, it is usually deployed coupling with cellular network to achieve seamless mobility. In cellular/802.11p heterogeneous network, vehicular communication has the characteristics of short span of time associating with Road Side Unit (RSU). Moreover, the media access control (MAC) scheme of IEEE 802.11p decides that packet collision probability increasing followed by the increasing of user quantity, which leads to the decreasing of network throughput. In response to these compelling problems, we propose a fuzzy Q-learning based vertical handoff (FQVH) control strategy for supporting the mobility management. FQVH has online learning ability and can give optimal handoff decisions adaptively with no need for prior knowledge on handoff behavior. Simulation results verify that it can adjust handoff strategies to different traffic conditions, and keep users always connected to the best network.**

Keywords- vehicular communication; heterogeneous network; vertical handoff; reinforcement learning

I. INTRODUCTION

Vehicular communication, as a novel fundamental platform for providing real time access to infotainment and safety related services, is expected to play an important role in future Intelligent Transportation Systems (ITS). Especially, Infotainment application is regarded as the key to open up and expand the new market, for the reason that it can improve driving experience innovatively. Besides traditional cellular network with a global coverage, IEEE 802.11p, which is particularly designed for high mobility terminals, gives another competitive choice for vehicle-borne users. It can enhance the capability of data transmitting to a great extent, and further enable a variety of infotainment applications. Additionally, it will allow many derivative services of road safety and traffic management, i.e., electronic toll, traffic control, parking lot management [1]. For these reasons, the communication between RSU and vehicle equipped with On Board Unit (OBU), which is also named as Vehicle to Infrastructure (V2I), is an important supplement to cellular network. In general, cellular system with global coverage performs better in the services that require high mobility and low latency, while it usually provides lower data rate than RSU. For these reasons, it is evident that the inter-working and cooperation of these two different types of wireless network will become an important issue in the future vehicular communication field.

With the increasing in diversity of wireless networks, an interesting concept comes into view. It is the notion to establish connections with the most suitable network at any given conditions, which is called Always Best Connected (ABC) [1]. Vertical handoff, which meant transfer of active session across heterogeneous access networks, is an integral component of heterogeneous mobility management. It provides the capability of seamless communication and exploring better Quality of Service (QoS), and therefore, it is the most relevant issue of meeting the goal of ABC concept in heterogeneous network. However, it is notable that vehicular communication has the characteristic of short span of time associating with RSU. For this reason, executing a handoff behavior is likely to face the trouble that switching into RSU followed by another handoff back to the original cellular network immediately. Additionally, the MAC scheme of IEEE 802.11p leads to the issue that network throughput decreases dramatically with the increasing of user quantity. Therefore a reliable handoff control strategy should be devised to guarantee the network working at efficient status. As a counter-example, an inappropriate handoff control strategy may result in the degeneration of the overall network performance, in the situation where many vehicles switch into a same RSU. Then it will be, in effect, lower achievable data rate in WLAN than that in cellular network due to the increased congestion delay. Up to the present, only a few vertical handoff control algorithms were proposed particularly for vehicular communication particularly. A brief summary of the representative work is demonstrated as follow.

A centralized Location Service Server (LSS) is proposed in [2], for the vertical handoff decider. Vehicles periodically report their current positions and receive the information of potential networks nearby them. After that, a utility function is used to calculate the satisfaction of users for the available networks, and it is fed back to LSS. The handoff decisions are finally given by LSS. A distributed vertical handoff strategy for vehicle to vehicle and vehicle to infrastructure communication is proposed in [3]. Base on the method, the communication cost and transmitting time is discussed. This work presented a heuristic analytical handoff control strategy based on an ideal scenario, where the coverage radius of WLAN, the data rate of cellular and WLAN are considered as fixed. The work of [2] and [3] are based on the same assumption that vehicles are equipped with GPS and the accuracy positions are known by themselves at any time. This prerequisite might limit their feasibility and universality. In [4], the authors developed a vertical handoff algorithm for the aspects of battery lifetime maximizing and load balancing. Meanwhile, they devised a

This work is supported by National Nature Science Foundation of China (Grant No. 61071105)

route selection algorithm for forwarding data packets to the most appropriate attachment point. Besides the aforementioned work, an adaptive MAC scheme based on the consideration of traffic density is proposed for WLAN in [5]. It can adjust the back off time to suit for different traffic and network conditions, and then achieve a lower collision and drop rate than traditional method. [6] proposed an optimization framework for solving the interface switching in vehicular heterogeneous network. It is designed for accommodating different performance metrics such as data transfer efficiency, monetary cost, and interface switching overhead. It can adapt the decision based on user-specified relative importance of the various metrics. In [7], the authors proposed a vehicular-IP MAC scheme to handle the limitation in the 802.11p standard for the operation and differentiation of IP applications.

There still lie ahead many challenges in the vertical handoff control for vehicular heterogeneous network. Conventional vertical handoff strategies are mostly designer intervened. The handoff decisions are derived by integrating the metrics about network load, terminal's status, user's preference, etc. However, it is usually difficult to establish an appropriate model for exploring optimal handoff solution with respect to so many related input parameters. Additionally, many demanded metrics, such as channel parameter with respect to attenuation factor and standard deviation of shadow fading, are quite different in diverse environment [8]. They need to be measured practically, which may cost additional human and financial resources. Moreover, less effort has been dedicated to explore the handoff strategy based on a relatively realistic scenario, including the conditions that, i.e., multiple vehicles simultaneously access a RSU, and MAC layer adopts adaptive Modulation and Coding Scheme (MCS). The aforementioned challenges severely restrict the feasibility of traditional vertical handoff algorithms in practice, and motivate the work in this paper.

The key contribution of this paper is stated as follow. We propose a fuzzy Q-learning based vertical handoff control strategy for vehicular heterogeneous network. Comparing to traditional designer intervened handoff control algorithms, it is an objective-oriented system that continuously exploring for the optimal solutions with no need for prior knowledge on handoff behavior. It can adjust handoff strategies to different conditions with respect to vehicle velocity and arrival rate of sessions adaptively. In addition, the learning capability helps for saving efforts on the measurement of channel parameters, which are generally recognized as requirements for channel modeling in literatures. We focus on the service type of infotainment application, which benefits from high throughput, and not sensitive to delay and jitter. The objective is ensuring WLAN supply relatively long service time on the premise that providing higher data rate than cellular network.

II. General description

In this work, we focus on a heterogeneous network consists of a global covered cellular network complemented by V2I communication mode, as depicts in Figure 1. Since WLAN provides potential better channel condition with low (or free) monetary cost as common trends, V2I communication is expected to improve the QoS with high throughput requirement and balance the network load for cellular network.

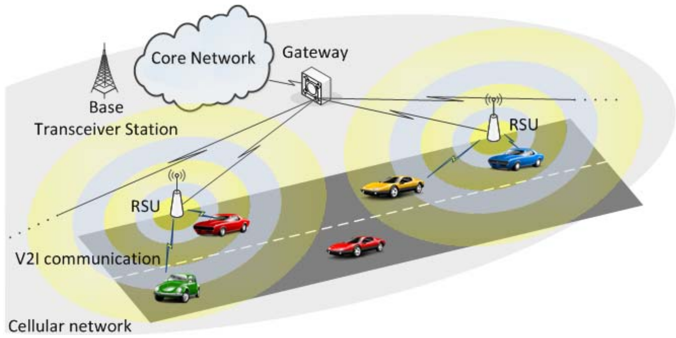

Figure 1. The reference scenario of vehicular heterogeneous network

We mainly focus on the service type of infotainment application, i.e., file transmitting, non-real time streaming. The services benefit from high data rate, while not sensitive to delay and jitter. It is notable that, IEEE 802.11p defines a scheme for OBUs to exchange data with RSU without association and authentication procedures. It benefits some particular emergent services, which need low access latency, while the data confidentiality cannot be guaranteed. Considering the service type we focus on is not safety related, so that without loss of generality, we apply a conventional IEEE 802.11 link initiation procedure, whereas vertical handoff behaviors followed by relatively considerable latency. Moreover, in order to achieve relatively realistic simulation results, adaptive MCS is adopted in the MAC layer of WLAN through the work.

Q-learning, as the core component of FQVH, achieves the function of sensing and learning. Since proposed by Watkins and Dayan [9], it has been widely explored in the field of automatic control. For overcoming the limitations of classical Q-learning algorithm, i.e., discrete state perception and discrete actions, Neural Fuzzy Inference System (NFIS) is adopted to retain continuous perception of the state space. It infers the global policy, which is relative to state, from local policies associated with each rule of the learner. FQVH is a centralized controller, and it can be integrated as a module in RSU. Vertical handoff can be categorized into two classes, which are Downward Vertical Handoff (DVH), for the case that OBU switches from cellular to WLAN and Upward Vertical Handoff (UVH) for the opposite process correspondingly.

A. DVH process

Figure 2 depicts the flow of FQVH. OBUs initiate vertical handoff process when they approach RSUs and receive beacon signal. Similar to [3], we assume OBU can periodically send handoff request with beacon message including the measures of $\overline{\mathrm{RSS}}(t)$, $V(t)$, and $D(t)$, where t is time instant. We adopt the mean value of Received Signal Strength (RSS) during the interval of two handoff requests, for reducing the deviation caused by shadow fading. Symbol V and D are the velocity of vehicle and data quantity to be transmitted, respectively. The handoff controller sends these measures together with S, which is the quantity of users associated with the target RSU, into fuzzy Q-learning module. RSS and V relates to the information of position and motion. D reflects the necessity of vertical handoff and the quantity of real time users that impact on the current utility of network. Handoff decisions are given by the output of fuzzy Q-learning. NFIS follow Takagi-Sugeno structure, whereby each fuzzy rule is in the form of Eq. (1).

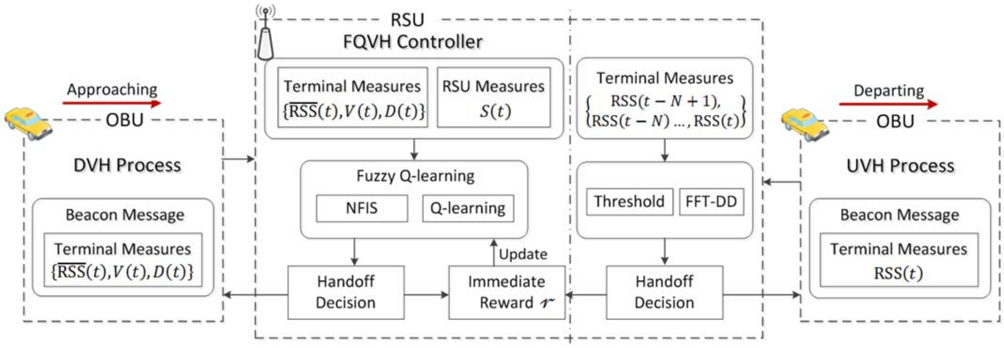

Figure 2. The architecture of FQVH

IF $\overline{\text{RSS}}(t)$ is $\mathcal{M}_{1,j_1}$ **AND** $V(t)$ is $\mathcal{M}_{1,j_2}$ **AND** $D(t)$ is $\mathcal{M}_{1,j_3}$ **AND** $S(t)$ is $\mathcal{M}_{1,j_4}$ (1)

THEN handoff action is $\mathcal{A}$ with Q.

where $\mathcal{M}_{i,j_i}$ is the j_ith linguistic variable related to the ith input metric, and Q is the reward value which guides the adaptive learning of FQVH.

B. UVH process

Since vehicle departing RSU, UVH procedure is going to be triggered. It can be relatively simpler than the reverse process, for the reason that achievable individual throughput is known information. For reducing computational complexity, we apply the handoff control method integrated with threshold and Fast Fourier Transformation based Decay Detection (FFT-DD) [10], as depicts in the right part of Figure 2. FFT-DD is employed to estimate the trend of WLAN signal, and then identify the vehicle is approaching or departing RSU. For the past Ω WLAN signal samples, as shown in Eq. (2), a positive value for $X(t)$ indicates rising signal trend, or vice versa. We assume the data rate of cellular network is steady during the duration of associating with RSU. UVH process is initiated when achievable data rate is lower in WLAN than that in cellular, plus FFT-DD indicates RSS signal is decaying. Then related measurements are sent to layer 1 as shown in Figure 3.

$$X(t) = \sum_{\omega=0}^{\Omega-1} RSS(t-\omega) sin\left(\frac{-2\pi\omega}{\Omega}\right) \quad (2)$$

When OBUs switch back to cellular network, the backward propagation of FQVH is triggered. It achieves the functions of evaluating the complete handoff actions involving DVH and UVH, and then tunes the parameters of FQVH. The detailed description is presented in section 3.

III. MATHEMATICAL MODEL OF FUZZY Q-LEARNING

The topology of fuzzy Q-learning is shown in Figure 3. Essentially, it is a feed forward network integrated with Q-learning and NFIS. The neurons in the different layers achieve different functions, and the only available information for learning is the system feedback. We use u_i^j and O_i^j to represent the input and output of the ith node in jth layer, respectively. Note that, in order to demonstrate the mathematical model in a general form, we present the equations with unfixed dimension of input state vector (N) and membership functions (T).

A. Forward propagation

L_1 input layer: This layer consists of N neurons, and all of them transmit the input value directly by:

$$O_i^1 = u_i^1, \forall i \in \{1, 2, \ldots, N\} \quad (3)$$

We use $\overline{\text{RSS}}(t)$, $V(t)$, $S(t)$ and $D(t)$, which have been discussed in the previous section, as the input state vector. Therefore the input can be represented by Eq. (4).

$$\mathbf{U}^1(t) = [\overline{\text{RSS}}(t), V(t), S(t), D(t)] \quad (4)$$

L_2 input linguistic layer: The functions of the neurons in this layer are fuzzification. As shown in Eq. (5) and (6), $\mathcal{M}(\cdot)$ is membership function, which is also recognized as linguistic variable. O_i^2 is the membership value related to input metric. The outputs of this layer reflect the degree that input values correspond with membership functions, which are defined by Gaussian Function in Eq. (6), where m and σ are mean and variance, respectively. Each row in matrix $\mathbf{M}^{N\times T}$ is a linguistic variable set related to one dimension of the input state vector.

$$O_i^2 = \mathbf{M}^{N\times T} = \begin{bmatrix} \mathcal{M}_{1,1}(u_1^2) & \cdots & \mathcal{M}_{1,T}(u_1^2) \\ \vdots & \ddots & \vdots \\ \mathcal{M}_{N,1}(u_N^2) & \cdots & \mathcal{M}_{N,T}(u_N^2) \end{bmatrix} \quad (5)$$

$$\mathcal{M}_{i,j}(u_i^2) = exp\left(-\frac{1}{2}\left(\frac{u_i^2 - m_{i,j}}{\sigma_{i,j}}\right)^2\right) \quad (6)$$
$$\forall i \in \{1, 2, \ldots, N\}; \ \forall j \in \{1,2 \ldots, T\}$$

L_3 rule layer: This layer achieves the fusion of fuzzy rules. We use the rules, which are the numerical multiplication of the membership values related to each of the input metric, to implement the precondition of the Takagi-Sugeno logic. These fuzzy rules divide the continuous state space of input metrics into discrete sub-state space. Denote s_k and $\mathbb{R}^{T^N}$ as the kth sub-state and the rule space, respectively. Each element in $\mathbb{R}^{T^N}$, which is $\delta_k(s_k)$, is regarded as the weighting of s_k. The output of layer 3 can be written as Eq. (7), where $u_{i,j}^3$ is the jth membership of ith input metric.

$$O_k^3 = \delta_k(s_k) = \prod_i u_{i,j}^3 = \prod_i (\{\mathcal{M}_{i,j_i}(u_i^2) | \forall \mathcal{M}_{i,j_i}(u_i^2) \in \mathbf{M}_i\}) \quad (7)$$
$$\forall i \in \{1, 2, \ldots, N\}, \forall j \in \{1,2, \ldots, P\}, \forall k \in \{1,2, \ldots, T^N\}$$

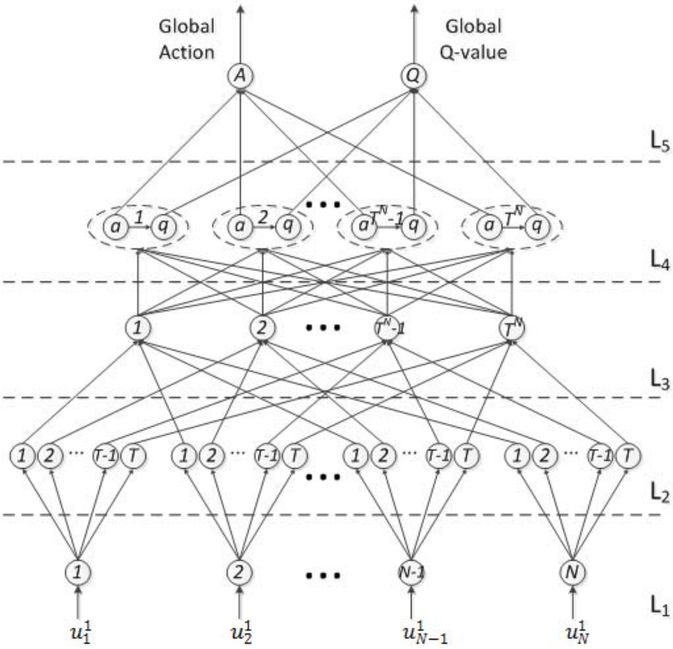

Figure 3. The topology of fuzzy Q-learning

L_4 output linguistic layer: Every neuron in this layer includes a local action-reward pair, which is represented by (a_i, q_i). Action is a predefined finite set including the probable solutions in output space. In this work, the local action set is defined as $\mathbf{A} = \{a_1(\text{strongly reject}), a_2(\text{reject}), a_3(\text{access}), a_4$ (strongly access)}. Note that, a more detailed local action set is used in pursuit of high resolution of output linguistic space, and enhancing the fitting capability. For each sub-state s_k, the local action a_i is guided by the reward $q(s_k, a_i)$. Assuming the optimal local action is a_k^*, which satisfies Eq. (8).

$$a_k^* = \underset{a_i}{\operatorname{argmax}} \{q(s_k, a_i)\}, \forall a_i \in \mathbf{A}, \forall k \in \{1,2,\dots,T^N\} \tag{8}$$

The output of layer 4 are normalized local actions and reward values, as shown in Eq. (9) and (10), respectively.

$$O_k^4 = \delta_k(s_k) \times a_k^* \times \left(\sum_{i=1}^{T^N} \delta_k(s_k)\right)^{-1} \tag{9}$$

$$\tilde{O}_k^4 = \delta_k(s_k) \times q(s_k, a_k^*) \times \left(\sum_{i=1}^{T^N} \delta_k(s_k)\right)^{-1} \tag{10}$$

L_5 defuzzy & ouput layer: The neurons in the last layer achieve the function of defuzzification by linear summation. Global action $\mathcal{A}$ and global reward $\mathcal{Q}$ are obtained by the fusion of local action a_k^* and local reward $q(s_k, a_k^*)$, which are demonstrated in Eq. (11) and (12), respectively.

$$O_i^5 = \mathcal{A}^*(\mathbf{U}^1(t)) = \sum_{k=1}^{T^N} u_i^5 \tag{11}$$

$$\tilde{O}_i^5 = \mathcal{Q}\left(\mathbf{U}^1(t), \mathcal{A}^*(\mathbf{U}^1(t))\right) = \sum_{k=1}^{T^N} \tilde{u}_i^5 \tag{12}$$

B. Backward propagation

Referring to Figure 2, the backward propagation of FQVH is triggered after OBUs finish the period associated with RSU and switch back to cellular network. Comparing to the forward propagation phase, which gives handoff decision, the reverse phase achieves the function of parametric learning. It evaluates the complete handoff actions involving DVH and UVH process, and then calculates updates the reward values according to the policies presented as below. The parametric conclusions in layer 4 of fuzzy Q-learning can be tuned. Consider the state transition of an OBU is:

$$\left(\mathbf{U}^1(t), \mathcal{A}^*(\mathbf{U}^1(t))\right) \xrightarrow{\text{handoff}} \mathbf{U}^1(t+\tau_l) \tag{13}$$

According to the update principle from classical Q-learning [10], the global reward values are updated by Eq. (14), where α is learning rate, β is discount factor for future reward, $\mathcal{r}$ is immediate reward, and τ_l is handoff latency.

$$\begin{aligned} \mathcal{Q}'\left(\mathbf{U}^1(t), \mathcal{A}^*(\mathbf{U}^1(t))\right) &= (1-\alpha)\mathcal{Q}\left(\mathbf{U}^1(t), \mathcal{A}^*(\mathbf{U}^1(t))\right) \\ &+\alpha\left(\mathcal{r} + \beta\mathcal{Q}\left(\mathbf{U}^1(t+\tau_l), \mathcal{A}^*(\mathbf{U}^1(t+\tau_l))\right)\right) \end{aligned} \tag{14}$$

Then the difference of Q value can be presented by:

$$\begin{aligned} \Delta\mathcal{Q} = \alpha\Big(\mathcal{r} - \mathcal{Q}&\left(\mathbf{U}^1(t), \mathcal{A}^*(\mathbf{U}^1(t))\right) \\ &+ \beta\mathcal{Q}\left(\mathbf{U}^1(t+\tau_l), \mathcal{A}^*(\mathbf{U}^1(t+\tau_l))\right)\Big) \end{aligned} \tag{15}$$

According to the ordinary gradient decent principle, local reward $q(s_k, a_k^*)$ can be updated by Eq. (16):

$$q'(s_k, a_k^*) = q(s_k, a_k^*) + \Delta\mathcal{Q} \times \delta_k(s_k) \times \left(\sum_{k=1}^{T^N} \delta_k(s_k)\right)^{-1} \tag{16}$$

Immediate reward $\mathcal{r}$ is evaluated according to the validity of the handoff decision after OBU logging out of RSU. As discussed in section 2, handoff behavior with payback of high individual throughput should be encouraged, and a positive immediate reward should be given. On the opposite, a negative immediate reward value should be given if a terminal obtain lower average individual throughput in WLAN than that in cellular network. The absolute value of $\mathcal{r}$ reflects the degree of that the handoff decision satisfies or dissatisfies the predefined objective. Define r_C as the average individual throughput of cellular network, and define $r_W(t)$ as the individual throughput of WLAN at time instant t. Firstly, considering the case that OBU switches into RSU, and the instantaneous achievable individual throughput is lower than that in cellular network. The immediate reward is defined as follow, where $t_{W,in}$ is the time instant that the duration of handoff latency is just passed.

If $r_W(t_{W,in}) < r_C$, the immediate reward is:

$$\mathcal{r} = \frac{r_W(t_{W,in}) - r_C}{r_C} \tag{17}$$

Denote $t_{W,sens}$ and $t_{W,out}$ as the timing that OBU initiate the first DVH and UVH request, respectively. For exploring the feasibility of handoff action and the optimal handoff instant, we define the immediate reward as follow, where $r_{W,avr}$ is the potential average individual throughput during the period of OBU reaching the edge of RSU's coverage to the time instant that it switches back to cellular network.

$$r_{W,avr} = \frac{r_c(t_{W,in} - \tau_l - t_{W,sens}) + \int_{t_{W,in}}^{t_{W,out}} r_W(t)dt}{t_{W,out} - t_{W,sens} + 2\tau_l} \tag{18}$$

If $r_W(t_{W,in}) \geq r_C$, immediate reward is:

$$r = \frac{r_{W,avr} - r_C}{r_{W,avr}}, \quad r_{W,avr} \geq r_C \tag{19}$$

$$r = \frac{r_{W,avr} - r_C}{r_C}, \quad r_{W,avr} < r_C \tag{20}$$

IV. Performance Evaluation

Referring to the scenario shown in Figure 1, we assume RSUs are deployed along the road side with 400 meters inter-distance, and 10 meters away from the roadway. They adopt 10MHz channel spacing in PHY according to IEEE 802.11p. Depending on the MCS level, RSUs provide data rate ranging from 3Mbps to 27Mbps including signaling cost. Distributed coordination function with binary exponential backoff is adopted in MAC layer. We use CDMA2000 1x-EV with average data rate of 0.6Mbps [11] for supplying global radio access. The initial data bits follows negative exponential distribution with expectation (λ) of 200Mb. Providing vehicles arrive at the coordinates of [0, 0]m and keep uniform linear motion states with velocity range from 20km/h to 70km/h. A typical log-normal function is used as the propagation model for WLAN signal, which is presented in Eq. (21). The related parameters for the simulation are shown in Table I.

Firstly, we investigate the performance of FQVH, for the impacts of vehicle velocity and arrival rate of session, which result in different traffic density correspondingly. For giving a clear view, we mainly discuss the simulation results under three selected conditions in Figure 4. The simulation time in each stage is 1000 seconds. The handoff thresholds appear stochastic trend at the beginning of each stage, which are illustrated as "learning period". Due to none priori knowledge is added, the convergence process will spend some time when a type of state is inputted firstly. With the growing of simulation loops, FQVH is tuned, and handoff thresholds appear stable trend. It indicates that the system already obtain memories on this type of input state. The controller will execute similar decisions when similar states input again. As observed, FQVH tends to compress the accessing area in avoid of permitting too much OBUs under a relatively dense traffic condition. This policy only enables the potential handoff request initialed by OBUs with relatively higher RSS, to enhance the channel utility under heavy traffic load. The threshold in distance is about 50 meters in the first stage. Along with the traffic load becoming sparse, the controller enlarges the valid coverage, where the threshold of accessing radius reaches to about 140 meters at the third stage.

$$P(d) = P_t - P_o - 10\gamma\lg(d) + \varepsilon(\mu, \sigma) \tag{21}$$

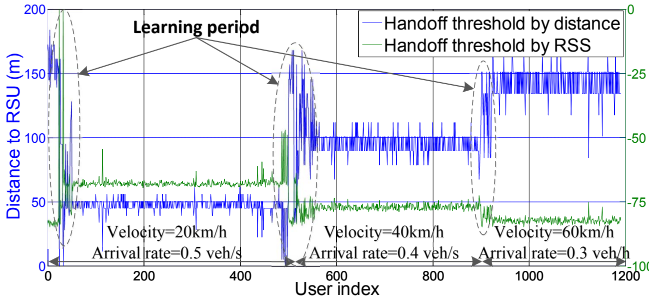

Figure 4. Vertical handoff threshold vs. different traffic conditions

TABLE I. Simulation Parameters

Parameter	Value
Tx power of RSU ($\boldsymbol{P_t}$)	100 mW
Path loss in the first meter ($\boldsymbol{P_0}$)	37.3 dB
Path loss exponent ($\boldsymbol{\gamma}$)	3
Gaussian shadow fading ($\boldsymbol{\mu, \sigma}$)	(0,5) dB
Handoff latency ($\boldsymbol{\tau_l}$)	2 s
Beacon interval	50 ms
Data rate in cellular network ($\boldsymbol{r_C}$)	0.6 Mbps
Vehicle's initial coordinate	[0,0]
RSUs' coordinates	[400,10] m, [800,10] m,…
Vehicle's velocity ($\boldsymbol{V}$)	20~70 km/h
Vehicle's arrival rate	0.3~0.5 veh/s
Average traffic density	11.1 m/veh~64.8 m/veh
Data bits required to be sent ($\boldsymbol{\lambda}$)	200 Mb
Handoff control interval	1 s
Channel spacing	10 MHz
RSS sensitivity of different MCS level	[-85,-84,-82,-80,-77, -73,-69,-68] dBm
Saturate data rate at different MCS level	[3,4.5,6,9,12,18,24,27] Mbps
MAC layer	IEEE 802.11 DCF
Minimum backoff window size ($\boldsymbol{W}$)	16
Maximum window size ($\boldsymbol{2^m W}$)	64
Duration of short inter-frame space	28 µs
Duration of distributed inter-frame space	130 µs
Slot time	51 µs
PHY header	128 bits
MAC header	272 bits
Packet payload	8184 bits
ACK	240 bits

For investigating the QoS, we define a parameter named good experience duration. It is the time duration that vehicles achieve better individual throughput in WLAN than that in cellular network. It reflects the effectiveness of handoff control algorithm. The other discussed parameter is average individual throughput of the heterogeneous network without signaling cost. For comparing, we adopt two conventional handoff control methods based on fixed thresholds plus dwell-timers. Thresholds 1 and 2 are set to the sensitivity of the lowest MCS level plus 5dB and 10dB, respectively. The dwell-timer is set to 2 seconds. Simulation results are shown in Figures 5, 6 and 7 with the arrival rate of 0.3, 0.4, and 0.5 veh/s, respectively.

The average individual throughput appears growth trend, with the increasing of vehicle's velocity. The reason is that, high velocity results in sparse traffic density by a given arrival rate, and further leads to the reduction of user quantity in RSU. Less associated OBUs result in higher saturated throughput.

Good experience duration shows a decreasing trend, for the reason that the dwelling time in RSU coverage becomes less. As observed, FQVH can guarantee the average individual throughput higher than 0.6Mbps under various conditions. The other two comparable methods cannot match this characteristic under a high traffic load condition, where vehicles arrive with a low average velocity and high arrival rate. It indicates FQVH can always execute the handoff behavior at an appropriate time instant, to guarantee a better achievable data rate in RSU than that in cellular network. Note that, the algorithm represented by the blue curve achieves a better average individual throughput than FQVH when the traffic is sparse, since a high threshold is adopted in the former scheme. As a result, the valid accessing area is small, where only a small quantity of users is permitted, and the channel utility is high for the reason of low packet collision probability. However, the potential duration that vehicle associating with RSU is very short. It results in the good experience duration demonstrated by the blue curve is quite low, which can be observed from Figures 5 to 7. The good experience duration represented by the green curve is the lowest under heavy traffic load, because the handoff threshold is too low for these given conditions. In summary, the proposed algorithm can adjust the handoff control strategy for different conditions to ensure RSUs work under an optimal status, and further guarantee high QoS for users.

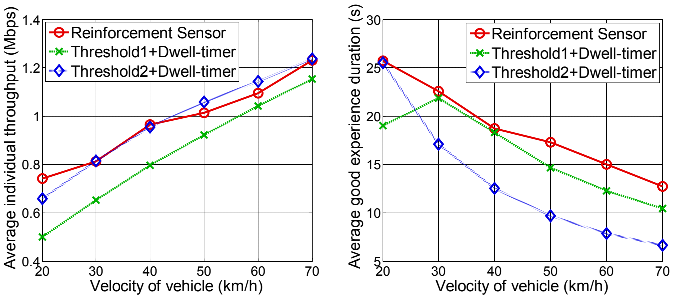

Figure 5. Performance vs. velocity of vehicle (arrival rate=0.3 veh/s)

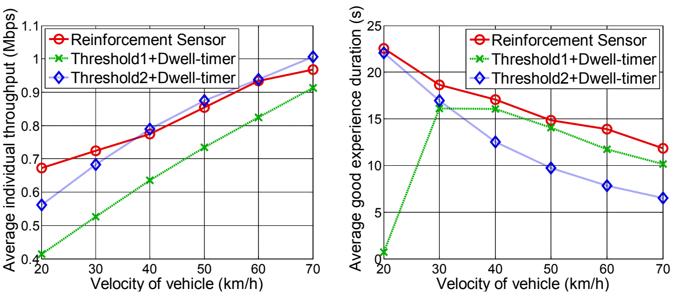

Figure 6. Performance vs. velocity of vehicle (arrival rate=0.4 veh/s)

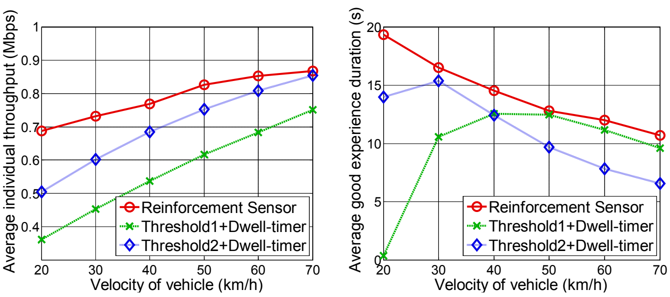

Figure 7. Performance vs. velocity of vehicle (arrival rate=0.5 veh/s)

V. CONCLUSIONS

In this paper, we propose fuzzy Q-learning vertical handoff control strategy for supporting the mobility management in vehicular heterogeneous network. The proposed algorithm has a real time learning capability and it can give optimal handoff decisions adaptively for realizing the ABC concept. Moreover, prior knowledge such as channel parameter is not a necessity. This advantage can save a lot of human and financial efforts for the measurement in practical application. We focus on a heterogeneous vehicular scenario consists of a global covered cellular network complemented by V2I communication mode. Simulation results verify that, FQVH can indicate handoff decisions to keep OBUs always connected with the better network under different conditions. Further, it can ensure RSUs work at appropriate status, and guarantee high QoS for users.

ACKNOWLEDGMENT

This paper was done when Limin Li was visiting at Nanyang Technological University under TEP program. The work was supported by the National Nature Science Foundation of China. The authors would like to thank all the members that participated in the work. Thanks also anonymous reviewers for their valuable comments.

REFERENCES

[1] M. Gramaglia, C.J. Bernardos and M. Calderon, “Virtual Induction Loops Based on Cooperative Vehicular Communications,” Sensors vol. 13, pp. 1467-1476, 2013.

[2] E. Gustafsson and A. Jonsson, “Always Best Connected,” IEEE Wirel. Commun. vol. 10, pp. 49-55, Feb. 2003.

[3] J. Zhang, H. Chan and V. Leung, “WLC14-6: A Location-based Vertical Handoff Decision Algorithm for Heterogeneous Mobile Networks,” In Proceedings of the 2006 IEEE Global Telecommunications Conference, San Francisco, CA, US, 27 Nov.- 1 Dec. 2006, pp. 1-5.

[4] K. Shafiee, A. Attar and V. Leung, “Optimal Distributed Vertical Handoff Strategies in Vehicular Heterogeneous Networks,” IEEE J. Sel. Areas in Commun. vol. 29, pp. 534-544, Mar. 2011.

[5] S. Lee, K. Sriram, K. Kim, Y.H. Kim, and N. Golmie, “Vertical Handoff Decision Algorithms for Providing Optimized Performance in Heterogeneous Wireless Networks,”. IEEE Trans Veh. Technol. vol. 58, pp. 865-881, Feb. 2009,

[6] A.B. Souza, A.L.B. Barros, A.S. Vieira, F.M. Roberto, and J.C. Junior, “An Adaptive Mechanism for Access Control in VANETs,” In Proceedings of The 10th International Conference on Networks, St. Maarten, The Netherlands Antilles, 23-28 Jan. 2011, pp. 183-188.

[7] K. Lee, X. Zhu, L. Dai, J. Zhu, S. Addepalli, R. Pan and F. Bonomi, “Intelligent Interface Switching among Heterogeneous Wireless Networks for Vehicular Communications,” In Proceedings of the 2012 Computing Networking and Communications Conference, Maui, HI, US, 30 Jan.- 2 Feb., pp. 766-770.

[8] S. Cespedes, N. Lu, and X. Shen, “VIP-WAVE: On the Feasibility of IP Communications in 802.11p Vehicular Networks,” IEEE Trans. Intell. Transport. S. vol. 14, pp. 82-97, Mar. 2013.

[9] G. Durgin, T.S. Rappaport and H. Xu, “Measurements and Models for Radio Path Loss and Penetration Loss in and Around Homes and Trees at 5.85 GHz,” IEEE Trans. Commun. vol. 46, pp. 1484-1496, Nov. 1998.

[10] C. Watkins, and P. Dayan, “Technical Note Q-learning,” J. Mach. Learn. vol. 8, pp. 279-292, May 1992.

[11] C. Guo, Z. Guo, Q. Zhang and W. Zhu, “A Seamless and Proactive End -to-End Mobility Solution for Roaming Across Heterogeneous Wireless Networks,” IEEE J. Sel. Areas in Commun., vol. 22, pp. 834-848, June 2004.

[12] CDMA2000 Wireless IP Network Standard: Packet Data Mobility and Resource Management, 3GPP2 Std. 3GPP2 X. S0011-003-C v3.0, 2006.