Chapter 12
An Adaptive WiFi/WiMAX Networking Platform for Cognitive Vehicular Networks

Dusit Niyato, Ekram Hossain, and Teerawat Issariyakul

Abstract This chapter presents an adaptive networking platform using WiFi/WiMAX technologies for cognitive vehicle-to-roadside communications, which can be used to transfer safety messages and provide Internet access for mobile users inside vehicles. The proposed platform is based on a heterogeneous multi-hop cluster-based vehicular network, where a vehicular node can choose to play the role of a gateway or a client. The gateway nodes communicate directly with a roadside base station through a WiMAX link. The client nodes connect to the gateways through WiFi links. Traffic from client nodes are relayed by the gateways to a roadside base station. The vehicular nodes are the self-interest (i.e., rational) and have capability to learn and adapt decision to achieve their objectives independently. A decision-making framework is proposed for this WiFi/WiMAX platform. This distributed decision-making framework, which enables the vehicular nodes with cognitive capability, is modeled and analyzed using game theory. Also, a Q-learning algorithm is used in vehicular nodes to provide the cognitive capability to learn and adapt their decision. Dynamics of Q-learning algorithm can be modeled as an evolutionary game.

12.1 Introduction

Wireless vehicle-to-vehicle (V2V) and vehicle-to-roadside (V2R) communications and networking technologies are the keys to providing Internet connectivity to mobile users in the vehicles. Vehicular networks using wireless access technologies (e.g., WiFi and WiMAX technologies) can support data communications for safety and intelligent transportation systems (ITS) applications (e.g., reporting traffic condition to the driver) and infotainment applications (e.g., providing interactive media and advertisement to the passengers). The vehicular nodes can form a heterogeneous *cognitive* multihop wireless network, where each node is able to dynamically choose among different radio access technologies for V2V and V2R communications. Illustrated in Fig. 12.1, two major components in a cognitive vehicular network are the network model and the decision-making framework. A network model incorporates all the basic functionalities necessary for data communications. A decision-making

D. Niyato (✉)
School of Computer Engineering, Nanyang Technological University, Singapore
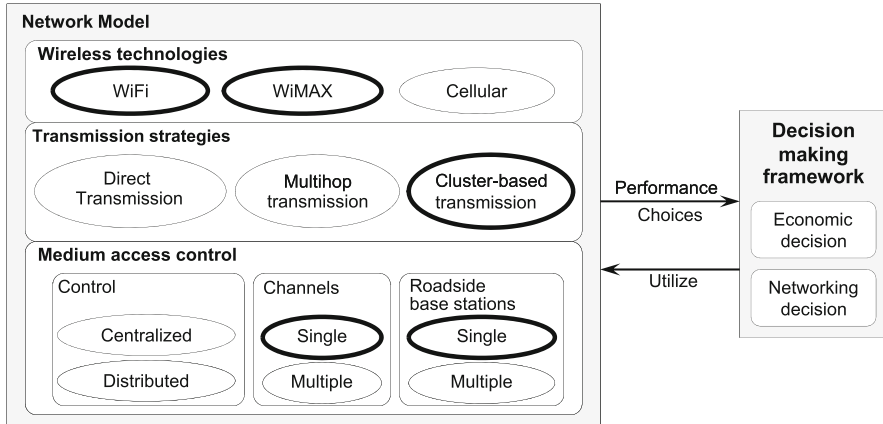e-mail: dniyato@ntu.edu.sg

Fig. 12.1 Cognitive vehicular network model: The *rectangles* represent the key components of the model, while the *ellipses* stand for choices for a component. The *ellipses* with *thick edges* correspond to the components considered in this chapter

framework, which is required for packet routing and distributed resource management in a vehicular network, is composed of the economic and the networking decisions to optimize the utility of the vehicular nodes in terms of both performance and cost. In a vehicular network, which can be considered as a distributed dynamic system, the vehicular nodes can be considered as an independent rational agents. That is, each vehicular node is an autonomous computational entity with a flexible dynamic behavior in an unpredictable environment. In such an environment, a vehicular node must be able to learn and adapt to the ambient environment to achieve its goal.

The different components of the cognitive vehicular networking model are described next.

12.1.1 Wireless Technologies

Currently, there are several enabling technologies for V2V and V2R communications. The IEEE 802.11-based WiFi technology supports short-range high-speed data transmission. However, its short transmission range leads to frequent transmission interruption (e.g., while vehicle speed is high) and the high deployment cost (e.g., many access points have to be deployed along the road) [1, 2]. IEEE 802.11-based services would be viable in a congested area where the vehicles move slowly. In this scenario, the users would benefit from high data rate and infrequent transmission interruption, while the service provider needs to install only few roadside access points at the selected hot spots.

The IEEE 802.16-based WiMAX technology provides large coverage area and high-speed connectivity [3]. While WiMAX helps overcome the range limitation of WiFi, its achievable data rate for low mobility may not be as high as that of WiFi.

In addition, the price for WiMAX access is comparatively higher than that for WiFi access.

Besides WiMAX and WiFi, the 3G cellular wireless technology for V2R communications provides a very broad coverage and supports high-mobility vehicles [4]. Due to lower data rate, the services in a wireless cellular network are usually less expensive than that in an IEEE 802.16 network.

12.1.2 Transmission Strategies

A transmission strategy determine how data packets are delivered from a vehicular node to a roadside base station (and vice versa). The strategy can be *direct transmission* where a roadside base station can be reached directly from a vehicular node (e.g., [5, 8]). If a roadside base station is located far away from a vehicular node, a *multihop transmission strategy* can be employed. In this scenario, the data packets from a vehicular node are relayed by other vehicular nodes until these data packets reach the designated roadside base station (e.g., coordinated external peer communications (CEPEC) in [3]). In a multihop vehicular network, traffic from a vehicular client node can be relayed through a vehicular gateway node to a roadside base station. Since the traffic from multiple vehicular nodes are aggregated and transmitted through this gateway, the utilization of the vehicle-to-roadside wireless link can be improved while the cost of a network is reduced due to bandwidth sharing. This *client–gateway* model is similar to the cluster-based vehicular network which was proposed in [6].

In a *cluster-based transmission strategy*, the vehicular nodes form groups (i.e., clusters) of vehicles, delegate a representative (i.e., a cluster head or a gateway) for each group, and transmit data through this selected representative [6]. A cluster-based vehicular network can improve the communication efficiency by not only reducing the signaling overhead but also alleviating congestion of the channel access which is fully controlled by a cluster head. In the *client–gateway* model, the gateway node acts as a cluster head which controls the transmission of traffic from the cluster members to a roadside base station. In this scenario, a vehicular node can use WiFi radio for local communications with the cluster head and a cluster head can use WiMAX radio for broadband communications with a roadside base station.

12.1.3 Medium Access Control Protocols

Medium access control (MAC) protocols refer to how vehicular nodes and roadside base stations share common radio channels. The MAC protocols can be classified based on three following criteria:

- *Centralized or distributed MAC protocols*: With a centralized MAC protocol, the decision of when and how the channels are accessed is determined by a central controller (e.g., scheduling in [6, 8]). With complete node information,

a centralized MAC protocol can be optimally designed at the expense of overhead needed to acquire such information. A distributed MAC protocol, on the other hand, determines how the channels are accessed based on local information (e.g., contention in [6, 7]). Despite decreasing overhead, a distributed MAC protocol is usually not optimal due to incomplete node information.

- *Single or multiple channels*: In presence of multiple channels (possibly different technologies), a MAC protocol needs to select the channels to satisfy application requirements. For example, a high data rate high attenuation-sensitive channel (e.g., in 5.8 GHz) should be used for typical data exchange, while a low data rate (usually more robust) channel (e.g., VHF or UHF) should be used for transmitting control and safety messages [5].
- *Single or multiple roadside base stations*: When considering multiple roadside base stations, a MAC protocol needs to control the handover process when a vehicle moves from one roadside base station to another. For example, [7] designed a distributed MAC protocol which quickly associates and disassociates a vehicular node with a roadside base station. The maximum freedom last (MFL) scheme minimizes the handover rate subject to a given delay constraint [8].

12.1.4 Distributed Decision-Making Framework

The decision-making framework, which considers both network pricing and network quality-of-service (QoS) issues, enables the vehicular nodes with cognitive radio capability. A pricing model characterizes the service fee (i.e., price) for using wireless access service. To access the radio resources in a wireless system, a mobile node has to pay to the radio resource owner (i.e., the service provider). Similarly, in most V2R communication scenarios, every vehicular node needs to pay to the service provider. For example, if a vehicular node (e.g., a gateway) uses direct WiMAX link to a roadside base station, the price has to be paid to the corresponding WiMAX service provider. However, if a vehicular node (e.g., client) uses a gateway to relay its traffic to a roadside base station, a price has to be paid to the gateway. In a vehicular network, the decision on price setting has to be optimally made by the wireless service provider.

In a V2R communications scenario, a vehicular node also has to make different networking decisions. In a cluster-based vehicular network, a vehicular node can choose to act as a *client* (i.e., a cluster member) or as a *gateway* (i.e., a cluster head). As a client, a vehicular node forwards its data traffic through the associated gateway. As a gateway, a vehicular node shares the link to the roadside base station with its client. Also, a client has to select the best gateway to relay its traffic to gain the highest benefit.

In general, a vehicular node can be considered as an independent and rational entity in a vehicular network. It will make a decision to maximize its benefits. For example, a vehicular node may decide to become a gateway and use WiMAX interface to provide the relaying functionality for other vehicular nodes if a gateway

receives high benefit from bandwidth sharing or "reselling." Alternatively, a vehicular node may decide to become client and use WiFi interface to transmit data to the gateway. Here, the decision-making framework for each vehicular node needs to be implemented in a *distributed* manner, considering both the networking aspects and the economic aspects. To this end, a supporting theory to obtain a stable solution for the above decision-making framework is required.

This chapter presents an adaptive decision-making framework for cognitive vehicle-to-roadside communications in a vehicular network. In this network, WiFi and WiMAX interfaces are used adaptively for client-to-gateway and gateway-to-roadside base station communications, respectively. A vehicular node with this platform forms a cluster-based network. With the assumption that a vehicular node is independent and rational to maximize its benefit (i.e., net utility), a vehicular node has to make a decision according to the vehicular network condition to use different wireless interface for data transmission. The first decision is whether a vehicular node should become a client or a gateway (i.e., role selection). If a vehicular node decides to become a client, it uses WiFi interface for data transmission and selects a gateway to relay its traffic. However, if a vehicular node decides to become a gateway, it uses WiMAX interface and determines the price of bandwidth sharing to be charged to its clients. The decision of a vehicular node affects not only its own benefit but also the benefits of other vehicular nodes. For example, if many clients select the same gateway, the portion of bandwidth given to each sharing node will decrease. Also, if a gateway charges high price, its clients will switch to other gateways which offer lower price of bandwidth sharing.

The rest of this chapter is organized as follows. Section 12.2 reviews the related work on cognitive vehicular networks. Section 12.3 presents an overview of distributed decision making based on game theory and reinforcement learning. The adaptive WiFi/WiMAX framework is described in Section 12.4. Section 12.5 presents the game model for the distributed decision-making framework. Performance evaluation results for the proposed framework are presented in Section 12.6. Section 12.7 summarizes the contribution of the chapter.

12.2 Cognitive Vehicular Networks: Related Work

Research on dynamic spectrum access-based cognitive vehicular networking has become popular recently. The spectrum sensing problem for cognitive-radio-enhanced vehicular ad hoc networks was addressed in [9] and [10]. In [11], communication protocols were proposed for universal wireless access in vehicular networking scenarios. In such a scenario, for V2R and V2V communications, a vehicular node is able to communicate on multiple frequency bands using different medium access control (MAC) and physical (PHY) layer interfaces. Since there are multiple interfaces, a new routing protocol is required to efficiently forward data packets. A cognitive communication for vehicular networking (CCVN) layer over multiple

MAC and PHY interfaces was introduced to support optimal connectivity, seamless mobility management, and QoS guarantee.

In [12], a cognitive MAC protocol, namely, CMV protocol, was proposed for multi-channel access in vehicular ad hoc networks. The protocol can support high mobility and spectrum handover by introducing the concepts of long-term and short-term spectrum access. For a long-term spectrum access, the channel is probed for every CCH period defined in the IEEE 1609.4 standard. Then, the spectrum status table (SST) is updated. For a short-term spectrum access, a spectrum pooling technique is used so that the best channel can be selected and the packet loss probability can be reduced. The performance evaluation results showed that the proposed CMV protocol outperforms existing multi-channel MAC protocols by up to 72%.

In [13], a dynamic spectrum access technique was adopted for inter-vehicle communications. Specifically, dynamic per-hop channel switching schemes for multi-hop VANET were proposed. These schemes were referred to as metric-based dynamic channel selection schemes with/without spatial awareness. These schemes are based on transmission rate, rate and utilization, and rate and idle probability metrics which are used together with the information about spatial movement of vehicular nodes (e.g., transmission range) to adapt the channel access. Simulation results showed the advantages of the proposed schemes in terms of communication duration and amount of transmitted data especially in the multi-hop and highly congested environments with high-speed mobility.

A framework for optimal channel access for vehicular nodes utilizing the exclusive-use and shared-use channels in cognitive radio network was proposed in [14]. The objective is to maximize the utility of data transmission by cluster members under QoS constraints (e.g., packet loss probability due to buffer overflow, average packet delay) and collision probability with primary users. Three major components in this framework are the queue-aware opportunistic access to shared-use channels, the reservation of bandwidth in the exclusive-use channel, and the cluster size control. To optimally design these components, a hierarchical optimization model was developed. With this framework, the cost of channel access to support various ITS applications can be minimized while guaranteeing the QoS requirements for the mobile nodes.

In [15], the vehicular public safety cognitive radio (VPSCR) platform was introduced. VPSCR has the ability to scan the radio spectrum over multiple public safety frequency bands. Then, commonly used public safety waveforms and networks can be identified such that VPSCR can adapt the spectrum access for network interoperation accordingly. This VPSCR platform was designed to communicate with a personal digital assistant (PDA) through existing fixed infrastructure (e.g., IEEE 802.11 or Bluetooth) to remotely control and access services.

In [16], a cognitive security protocol for sensor based VANET (S-VANET) was introduced. This protocol can distribute the security information to support the prevention of data aging, efficient QoS, and robustness against denial-of-service (DoS) attack. The reliability and optimality of the protocol were evaluated in terms of response time, ability to maintain message authentication, integrity, confidentiality, and non-repudiation.

12.3 Distributed Decision Making

Three mechanisms to achieve the distributed decision making are the agent-based computing, intelligent algorithm, and game theory. Since it is impossible for a vehicular node to anticipate and estimate the consequence of all situations to encounter in a dynamic vehicular environment, cognitive or learning capability becomes crucial. With the use of a learning algorithm (specifically reinforcement learning), evolutionary game theory can be used to study the dynamics of a multiagent system. An evolutionary game theory model can be used to obtain the solution of rational agents (i.e., agent with self-interest). The relationship among multiagent systems, evolutionary game theory, and reinforcement learning is shown in Fig. 12.2 [17].

12.3.1 Evolutionary Game Theory

Evolutionary game theory is a branch of game theory developed to provide a basis to understand rational decision making in an uncertain environment. Evolutionary game theory complements traditional noncooperative game theory in following aspects.

- *Refinement of traditional solution concept*: In a traditional noncooperative game, the Nash equilibrium is the most common solution concept. However, in any game, the Nash equilibrium cannot be guaranteed to exist if the player is restricted to use only pure strategy. Also, there could be multiple Nash equilibria in the game. In this case, the solution of evolutionary game theory (i.e., evolutionary stable strategies (ESS) or evolutionary equilibrium) can serve as a refinement to the Nash equilibrium especially when multiple Nash equilibria exist.
- *Bounded rationality*: In a traditional noncooperative game, the agent is assumed to be rational. That is, an agent will always maximize the payoff in which this assumption is derived from the utility theory. This rationality of agent requires complete information and well-defined and consistent set of choices. However, in reality, this assumption is rarely held. Evolutionary game theory has been developed to model the behavior of biological agents (e.g., insects and animals) which does not require the strong rationality assumption. Therefore, evolutionary game theory will be suitable for the problem which involves human being as the agents. These agents may not have hyperrational behavior.
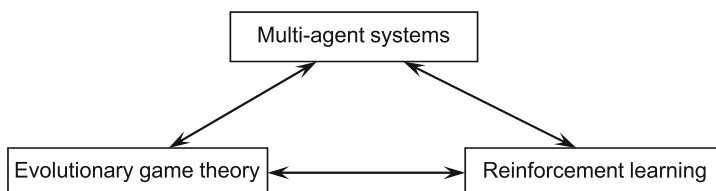

Fig. 12.2 Relationship among multiagent systems, evolutionary game theory, and reinforcement learning [17]

- *Dynamics of game*: Traditional noncooperative game has been developed mostly for the static analysis. It cannot model the adaptation of agents to change their strategies and to reach the equilibrium solution. Evolutionary game theory is based on the evolutionary process which is dynamic in nature. Evolutionary game establishes the dynamics model of interactions among agents in the population (i.e., strategy adaptation over time).

In evolutionary game, a game is played repeatedly by the agents. These agents are randomly selected from a large population. Two major mechanisms of an evolutionary process are mutation and selection. While the mutation mechanism is used to provide diversity in the population, the selection mechanism is used to promote the agents with higher fitness over other agents. In an evolutionary game, the mutation mechanism is described by the evolutionary stable strategies (ESS). The selection mechanism is described by the replicator dynamics. In other word, ESS is used to study a static evolutionary game while replicator dynamics is used for a dynamic evolutionary game.

12.3.1.1 Evolutionary Stable Strategies (ESS)

With a large population, let most of the players adopt the same strategy (e.g., mixed strategy s), and there is a small fraction $\varepsilon \in (0, 1)$ of a population adopting a different strategy (e.g., mixed strategy s'). Then, if the reproductive success of the new strategy s' is smaller than the original strategy s, the entire population will not be overruled by the new strategy s' and this new strategy s' will disappear eventually. In this case, the original strategy s is said to be an ESS which is robust against the evolutionary pressure from any appearing mutant strategy s'. Specifically, the payoff of the player adopting original strategy s is denoted as $\mathscr{U}(s, (1-\varepsilon)s + \varepsilon s')$, where $\mathscr{U}(\cdot, \cdot)$ is the utility function whose first parameter is the current strategy and the second parameter is the strategy of an opponent. Then, the payoff of a player adopting new strategy s' is denoted as $\mathscr{U}(s', (1-\varepsilon)s + \varepsilon s')$. Strategy s is an ESS if $\forall s' \neq s$, there exists $\delta \in (0, 1)$ such that the following condition holds:

$$\mathscr{U}(s, (1-\varepsilon)s + \varepsilon s') > \mathscr{U}(s', (1-\varepsilon)s + \varepsilon s'), \quad \forall \varepsilon : 0 < \varepsilon < \delta. \qquad (12.1)$$

In general, ESS is a subset of the Nash equilibrium, since the conditions for an ESS are stricter than those of the Nash equilibrium. That is, the Nash equilibrium of a player is required to be the best response to the strategy of the opponent. To be ESS, this strategy s has to be also optimal against itself. Otherwise, there would be other strategy s' which yields higher payoff and this new strategy s' will successfully invade strategy s.

12.3.1.2 Replicator Dynamics

In an evolutionary game, the dynamic process is related to the evolution of population adopting different strategies. As has been mentioned before, the evolution is based on selection and mutation. While selection is used to select the fraction of a population with the higher payoff, a mutation provides a variety of strategies in

the population. The replicator dynamics is a system of differential equations used to describe the selection process of evolution, i.e., how the population choosing different strategies changes over time. Each replicator represents a pure strategy s. The new offspring is reproduced for different strategy which can be modeled by

$$\frac{\mathrm{d}x_s}{\mathrm{d}t} = \dot{x}_s = x_s \left(\mathscr{U}(s) - \overline{\mathscr{U}}(\mathbf{x}) \right) \tag{12.2}$$

where x_s represents the fraction of population adopting strategy s, $\mathbf{x}$ is a vector x_s, and it is referred to as the state of population. $\mathscr{U}(s)$ is a payoff of player adopting strategy s, and $\overline{\mathscr{U}}(\mathbf{x})$ is the average payoff of the population. At the steady state (i.e., for time $t \to \infty$), the replicator dynamics will be $\dot{x}_s = 0$ if the strategy is stable. The fraction of the population adopting different strategy at the stable steady state is referred to as the evolutionary equilibrium. It is known that every Nash equilibrium is an evolutionary equilibrium of replicator dynamics. However, an evolutionary equilibrium may not be a Nash equilibrium.

The theory of evolutionary game has been adopted to solve various problems in wireless networks (e.g., [18–22]). In [18], an evolutionary game theory was used to model the network selection behavior of the mobile users in a heterogeneous wireless network, which is composed of multiple wireless access technologies (e.g., cellular, broadband wireless access, and WLAN). The mobile users can adapt their network selection strategy based on the perceived performance and the cost of a wireless connectivity. In [19], the traffic routing problem was modeled by an evolutionary game. The players can choose the routing path to avoid any congestion so that the performance is maximized. In [20], a similar game model was developed for the IEEE 802.16 multihop wireless backhaul. In this case, the traffic routing has to also take the wireless channel quality into account. In [21], an evolutionary game theory was applied to study the problem of power allocation in the cooperative relay networks. In such a network, a relay node can select the different power levels for relaying traffic from the source nodes. The power level to be used can evolve based on the benefit of relaying (e.g., higher transmission rate). In [22], the cooperative spectrum sensing problem for cognitive radios was modeled by an evolutionary game. The cooperation behavior of the secondary users can evolve due to benefit of performing cooperative spectrum sensing with other secondary users to detect the primary user.

12.3.2 Reinforcement Learning

Distributed decision making using reinforcement learning algorithm is based on the optimization model of Markov decision process (MDP). An MDP is defined by a set of states, a set of actions, and a set of rewards. At each time t, the agent observes the state x_t. Then, the agent chooses an action s_t given state x_t. Then, the system transits to the new state and the agent receives the reward u_t (or experiences the cost). The agent with a learning algorithm (e.g., reinforcement learning) will develop a policy, which is a mapping from state to action, to maximize the long-term reward. This long-term reward can be the sum of a immediate reward in the finite time horizon

case with limit T (i.e., $U = \sum_{t=0}^{T} u_t$) or the sum of a discounted immediate reward (i.e., $U = \sum \lim_{t=0}^{\infty} \gamma^t u_t$, where γ for $0 < \gamma < 1$ is a discounting factor). The most popular reinforcement learning algorithm for MDP is the Q-learning. A simple example of this algorithm is shown in Algorithm 1 where $rand()$ is a random number generator, α is the learning rate, and γ is the discounting factor. This algorithm is divided into two steps, i.e., exploration and exploitation. The algorithm performs exploration step randomly with a certain probability (line 4 of Algorithm 1). In this exploration step, the algorithm tries different action randomly so that the knowledge (i.e., Q-value) of the action can be obtained. In the exploitation step, this knowledge is used to make the optimal decision (line 6 of Algorithm 1). This Q-value is updated according to the equation in line 8 of Algorithm 1.

Algorithm 1 Q-learning algorithm

1: Initialize q-value $Q(x_t, u_t)$ where x_t is state and s_t is action at time $t \leftarrow 0$
2: **loop**
3:     **if** $rand() <$ Exploration probability **then**
4:         Select action s_t randomly given state x_t at time t
5:     **else**
6:         Select the best action $s_t = \arg\max_s Q(x_t, s)$
7:     **end if**
8:     $Q(x_t, s_t) \leftarrow Q(x_t, s_t)(1 - \alpha) + \alpha (u_t + \gamma \max_s Q(x_{t+1}, s))$
9: **end loop**

The Q-learning algorithm has been applied to solve distributed decision-making problems in wireless networks (e.g., [23–28]). In [23], Q-learning was used to obtain the distributed handoff decisions for the mobiles in a heterogeneous wireless network. The objective is to maximize the expected total utility of a connection subject to the constraint on the total access cost. The utility is defined as the quality of wireless connection, with a penalty on the signal and call dropping. In [24], Q-learning was used to obtain a distributed buffer management policy for mobiles transmitting biosignal data from patients to different wireless access networks in a heterogeneous wireless telemedicine system. The objective is to minimize the cost while the QoS requirements (i.e., delay and loss) are met. In [25], a Q-learning algorithm was adopted in a cognitive radio network where the secondary base station chooses a wireless channel to access given the states of the channels. The reward was defined in terms of the signal-to-interference and noise ratio (SINR). In [27], a Q-learning algorithm was used for solving a routing problem in the multihop cognitive radio networks. The number of available channels is estimated and the optimal route is selected based on this information. In [28], a Q-learning algorithm was used to optimize the spectrum sensing in the cognitive radio networks. The reward is defined in terms of the accuracy of channel sensing result.

12.3.3 Reinforcement Learning and Evolutionary Game Theory

Reinforcement learning (i.e., Q-learing) of agents can be modeled as an evolutionary game [17]. For a system with two players, let $\mathbf{U}_1$ and $\mathbf{U}_2$ denote payoff matrices

of players 1 and 2, respectively. The dynamics of player 1 can be expressed as
follows:

$$\frac{dx_{s,1}}{dt} = \dot{x}_{s,1} = x_{s,1}\alpha((\mathbf{U}_1\mathbf{x}_2)_s - \mathbf{x}_1\mathbf{U}_1\mathbf{x}_2) + x_{s,1}\alpha \sum_{s'} x_{s',1} \ln\left(\frac{x_{s',1}}{x_{s,1}}\right) \quad (12.3)$$

and dynamics of player 2 can be expressed as follows:

$$\frac{dx_{s,2}}{dt} = \dot{x}_{s,2} = x_{s,2}\alpha((\mathbf{U}_2\mathbf{x}_1)_s - \mathbf{x}_2\mathbf{U}_2\mathbf{x}_1) + x_{s,2}\alpha \sum_{s'} x_{s',2} \ln\left(\frac{x_{s',2}}{x_{s,2}}\right). \quad (12.4)$$

These dynamics represent the evolution of both players using a Q-learning algorithm
in terms of a probability for selecting strategy (i.e., $x_{s,j}$ is a probability of selecting
strategy s of player j). It can be observed that the first terms of (12.3) and (12.4),
which account for the strategy selection process of players, are the same as those
of replicator dynamics. The second terms account for the mutation process. Specif-
ically, the mutation and selection processes can be considered as the exploration
and exploitation steps in the reinforcement learning. Alternatively, the evolutionary
game formulation for a Q-learning algorithm can also be modeled as a Markov chain
since the population can make decision randomly due to bounded rationality [29].
In this case, the fractions of population selecting different strategies are modeled
as the states of the Markov chain. The transition rates or transition probabilities are
determined by the payoffs corresponding to different strategies.

12.4 Adaptive WiFi/WiMAX Networking Platform

In this section, we present an adaptive multihop and clustered WiFi/WiMAX-based
cognitive vehicular networking platform.

12.4.1 Network Model

Consider a cluster-based vehicular network with N vehicular nodes moving in
the same direction (Fig. 12.3). Each of these N nodes is equipped with a dual-
mode WiFi/WiMAX transceiver. A WiFi/WiMAX transceiver conforms to the IEEE
802.11 and the IEEE 802.16 MAC protocols.

   All vehicular nodes need to communicate with a roadside base station. These
vehicular nodes may establish a direct wireless link to the roadside base station
using a WiMAX transceiver unit. These nodes are referred to as *gateways*. Others,
referred to as *clients*, communicate with the roadside base station through one of
the gateways. These clients connect to a gateway using a WiFi transceiver and share
the WiMAX link with the gateway. n_g and n_c denote the numbers of gateways and
clients, respectively, where $n_g + n_c = N$. Also, the number of clients associated
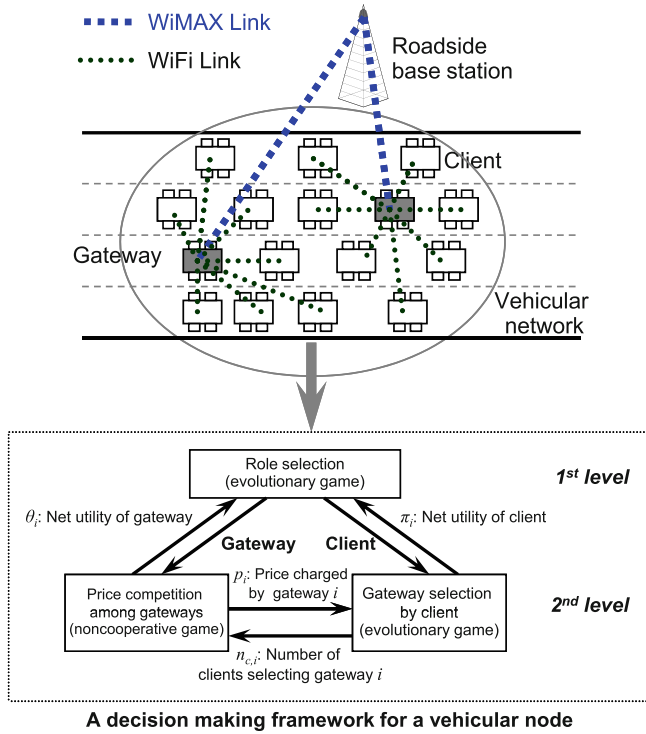with gateway i is denoted by $n_{c,i}$. It is assumed that the bandwidth on a WiMAX

A decision making framework for a vehicular node

Fig. 12.3 A cluster-based vehicular network: A gateway is represented in *gray* and has a direct WiMAX link to the roadside base station. A client is represented in *white* and connects to the roadside base station via a gateway. Both the gateway and the client use a two-level decision making framework

link is B_b bps. This link is shared among $1 + n_{c,i}$ nodes (a gateway and its clients). Correspondingly, each of the $1 + n_{c,i}$ vehicular nodes is allocated with a logical WiMAX link to a roadside base station with bandwidth $B_b, (1 + n_{c,i})$ bps. Let B_c denote the aggregated bandwidth on all WiFi links associated with a gateway. The bandwidth between gateway i and each of its clients would be $B_c, n_{c,i}$, bps in which the IEEE 802.11 MAC protocol is based on point coordination function (PCF). Under this model, the bandwidth of a link between a vehicular node and a roadside base station is $b = \min(B_b/(1 + n_{c,i}), B_c, n_{c,i})$ bps.

12.4.2 Decision-Making Framework

Communication services are offered to the vehicular nodes by the service provider and the gateways. The service provider offers WiMAX services with bandwidth B_b. The price for bandwidth B_b bps is P_b monetary units (MUs). A gateway which purchases a WiMAX link may share the link with its clients. Gateway i offers the traffic relaying service to its clients and charges price $p_i < P_b$ MUs to each client.

Under the above network model, each vehicular node needs to make a two-level decision (Fig. 12.3). In the first level, a vehicular node decides whether to be a client or to be a gateway. An evolutionary game model based on a Markov chain is used to analyze this decision of vehicular nodes implementing the Q-learning algorithm. After deciding its role, a vehicular node defines its parameters based on the selected role. As a client, a vehicular node selects a gateway for relaying its traffic. The decision process for gateway selection is modeled by an evolutionary game. As a gateway, a vehicular node determines the price to charge its clients for bandwidth sharing. The decision on price setting can be modeled as a noncooperative game. Given the decisions of other gateway nodes, a vehicular gateway node makes its pricing decision to maximize its net utility. At an equilibrium point, none of the vehicular nodes would be willing to change its decision. That is, in the second level, the clients and gateways determine their gateway and competitive price, respectively.

It is assumed that every vehicular node is interested in maximizing its own satisfaction, which is modeled by the so-called *net utility*. The net utility depends on the bandwidth (b) from itself to the roadside base station, the price (p), and the revenue (r) gained from other nodes. Mathematically, net utility is defined as the rate utility $\mathscr{U}(b)$ minus cost p plus revenue r, i.e., $\mathscr{N}(b, p, r) = \mathscr{U}(b) - p + r$. The cost ($p$) of a gateway and a client denote, respectively, the price for a WiMAX link charged by the service provider, and the price charged by a gateway to a client node to relay traffic over the WiMAX link. Attributed to a gateway, the revenue r is earned by sharing the purchased WiMAX link with its clients. Finally, the rate utility is characterized by a concave logarithm utility function $\mathscr{U}(b) = u_1 \log(1 + u_2 b)$, where u_1 and u_2 are the parameters of the function.

Note that similar pricing models for traffic relaying can be found in [30, 31]. In these models, pricing was used as an incentive for one node to relay traffic for other nodes [30]. The optimal price can be determined from the bandwidth demand of neighboring nodes based on an auction mechanism [31]. However, most of the work ignored the issues of gateway selection and price competition which are common in a cluster-based network (e.g., a vehicle platoon on the highway). Also, similar networking decision model can be found in [32]. However, there are many major differences. First, in [32], the number of gateways is fixed. Second, the users are assumed to possess linear bandwidth demand function. Third, the price competition among the gateways was not considered in [32]. In short, the decision-making framework presented in this chapter is more general which can capture the independent and rational decision-making behavior of a mobile node which is common in a cognitive vehicular network.

12.5 Hierarchical Game Formulation for Distributed Decision Making Framework

The decision-making framework for independent and rational vehicular nodes is developed using a hierarchical game structure. This framework consists of two levels and three game formulations (see Fig. 12.3). In the first level, each vehicular node

applies an *evolutionary game* to determine its role (as a client or as a gateway). In the second level, a gateway applies a *noncooperative game* to obtain a competitive price, while a client uses an *evolutionary game* for the gateway selection. The following two-step backward induction procedure is applied to solve the entire hierarchical game: (1) Obtain the solution in the second level for both clients and gateways and (2) Use the solution of the second level to select the role (i.e., gateway or client) of a vehicular node in the first level.

12.5.1 Gateway Selection by Client Nodes

We use an evolutionary game model [33] for the gateway selection problem. Here, each client observes the price periodically (e.g., for every 10 s) broadcast by all the gateways, computes the expected net utility, and selects the gateway which gives rise to the highest net utility. The utility depends on both price and bandwidth. While a gateway specifies the price, its offered bandwidth depends on the number of associated clients. Therefore, the value of net utility can change after the clients make a decision (e.g., change the gateway). In this iterative algorithm, at each iteration (e.g., every 10 s) a client chooses a gateway. After reaching the equilibrium, the net utility will remain unchanged over the rest of the adaptation interval, and the clients will stick to one gateway which maximizes their net utility.

An evolutionary game for gateway selection is formulated as follows. A *player* is a client. A *population* is a group of n_c clients. The *strategy* of a client is the selection of a gateway. The set of strategies correspond to the set of gateways. The *payoff* is given by the net utility of a client. Each client decides to join one of n_g groups (i.e., gateways) which maximizes its net utility. In an evolutionary game, the proportion, x_i, of the clients selecting a gateway i can be determined, where $\sum_{i=1}^{n_g} x_i = 1$. An evolutionary equilibrium is defined as a point where no strategy can lead to a change in the proportion of clients x_i, $\forall i$. In particular, it can be expressed as

$$\dot{x}_i = \frac{dx_i}{dt} = x_i (\pi_i - \overline{\pi}) = 0 \tag{12.5}$$

where $\pi_i = \mathcal{N}(b_i, p_i, 0)$ denotes the payoff of each client selecting gateway i, and $\overline{\pi} = \sum_{i=1}^{n_g} x_i \pi_i$ denotes the average payoff of the entire population. Since the proportion x_i ceases to vary at the equilibrium, the number of clients associated with gateway i, $n_{c,i} = x_i(N - n_g)$ ceases to change. Here, the net utility of each client remains unchanged, and each client sticks to a gateway which maximizes its utility.

12.5.2 Price Competition Among Gateway Nodes

Since a client can select and switch to the gateway which provides a higher net utility, the price offered by each gateway has to be carefully chosen. For example, if the price is high, only a few clients will select this gateway to relay their traffic,

and only small revenue can be generated. However, if the price is low, a number of clients will select this gateway, and the end-to-end bandwidth can be degraded due to congestion. Since each gateway makes its decision independently and non-cooperatively, the desirable solution in terms of price has to maximize the net utility of the gateway. Therefore, a noncooperative game [34] is formulated to obtain this competitive price. This game can be described as follows. A *player* of this game is a gateway. The *strategy* of a player is the offered price. The *payoff* of a gateway is defined as

$$\theta_i = \mathcal{N}(b_i, P_b, n_{c,i} p_i) = \mathcal{U}(b_i) - P_b + n_{c,i} p_i \qquad (12.6)$$

where b_i is the end-to-end bandwidth, and the revenue from the clients is a chosen price p_i multiplied by the total number of clients $n_{c,i}$ selecting this gateway i. Since the net utility of the gateway is a function of $n_{c,i}$ which is again a function of price offered by other gateways, the net utility can be written as $\theta_i(p_i, \mathbf{p}_{-i})$, where $\mathbf{p}_{-i}$ is a vector of prices from other gateways except gateway i.

A noncooperative game is used since each gateway wants to achieve the highest payoff in terms of net utility by increasing the price. However, if one gateway increases its offered price, it is likely that other gateways will reduce the price to attract more clients, and the gateway with high price loses revenue. In this competitive situation, the Nash equilibrium is considered as a solution of this noncooperative game. The Nash equilibrium has the property that the payoff of one gateway is maximized, given the price chosen by other gateways. At the Nash equilibrium, this property applies to all gateways. Therefore, none of the gateways would unilaterally change the strategy to improve its payoff.

The Nash equilibrium can be obtained by using the best response function which is the best price from one gateway given the prices from other gateways. In particular, the best response function of a gateway is obtained by formulating a payoff maximization problem. The best response function of gateway i can be defined as follows:

$$p_i^* = \mathcal{B}_i(\mathbf{p}_{-i}) = \arg\max_{p_i} \theta_i(p_i, \mathbf{p}_{-i}). \qquad (12.7)$$

This best response can be obtained by a numerical method. Then, the Nash equilibrium can be obtained from $p_i^* = \mathcal{B}_i(\mathbf{p}_{-i}^*)$ for all i, where $\mathbf{p}_{-i}^*$ is a vector of best response of all gateways except gateway i.

The distributed algorithm that achieves the Nash equilibrium for the gateways works as follows. Gateway i observes the price broadcast by other gateways. Then gateway i chooses the price to maximize its payoff. This can be done by observing the responses of the clients to a small variation in the current price (i.e., more or fewer number of clients will select gateway i due to lower or higher prices, respectively). From these responses, the gateway can estimate marginal payoff due to variation in price. This marginal payoff is then used to obtain the best response for this gateway. This procedure is repeated for all gateways until

there is no change in price offered by all gateways. Since there is no change in the strategy adopted by all players, the solution of this algorithm is the Nash equilibrium.

12.5.3 Role Selection by Vehicular Nodes

After the solutions of the gateway selection and the price competition are obtained where the corresponding net utilities are given by π_i and θ_i, respectively, we backtrack to the first-level decision on whether the vehicular node decides to become a gateway node or a client node. Since the net utilities for being a gateway and a client are not known by the node a priori (i.e., other gateways and other clients do not reveal their net utility information), the vehicular node must learn by trials. In this case, a vehicular node can randomly become a client or a gateway and observe the net utility. For instance, if becoming a client yields a higher net utility, this node will choose to become a client in the future.

The Q-learning-based algorithm of a vehicular node to decide whether to become a gateway or a client works as follows:

1: A vehicular node randomly chooses to become a gateway or a client.
2: Q-value $Q(s_t)$ for strategy $s_t \in \{\text{gateway, client}\}$ is initialized at time $t = 0$.
3: **loop**
4:     **if** Node is gateway **then**
5:         **if** $Q(\text{gateway}) < Q(\text{client})$ **then**
6:             Gateway switches back to become a client. {Becoming client yields higher net utility (exploitation)}
7:         **else**
8:             Gateway randomly becomes a client with exploration rate ρ (e.g., $\alpha = 0.1$) {Learning by trial (exploration) }
9:         **end if**
10:        A vehicular node observes its net utility (i.e., π_i for client).
11:        $Q(s_t) \leftarrow Q(s_t)(1 - \alpha) + \alpha \left(\pi_i + \gamma \max_{s_{t+1}} Q_{s_{t+1}} \right)$
12:    **else**
13:        **if** $Q(\text{client}) < Q(\text{gateway})$ **then**
14:            Client switches to gateway {Becoming a gateway yields higher net utility (exploitation)}
15:        **else**
16:            Client randomly becomes a gateway with exploration rate ρ. {Learning by trial (exploration) }
17:        **end if**
18:        A vehicular node observes its net utility (i.e., θ_i for gateway).
19:        $Q(s_t) \leftarrow Q(s_t)(1 - \alpha) + \alpha \left(\theta_i + \gamma \max_{s_{t+1}} Q_{s_{t+1}} \right)$
20:    **end if**
21: **end loop**

This algorithm is performed periodically until a vehicular node finishes the data transfer with a roadside base station or the node leaves the network. Note that the random actions in lines 8 and 16 of the above algorithm are used to try an alternative strategy periodically. This trial is required to avoid a vehicular node being locked up in the sub-optimal decision due to an obsolete information about net utility when the network condition changes.

This learning algorithm can be modeled as a stochastic evolutionary game where a vehicular node gradually learns by randomly trying the different available strategies. The game can be described as follows. A *player* is a vehicular node. A *strategy* is to become either a gateway or a client. *Payoff* is the net utility of a node. The solution of this evolutionary game is an equilibrium which can be analytically obtained by formulating a finite discrete-state and continuous-time Markov chain. The state space of this Markov chain is a random integer between 1 and N representing the current number of gateways. The transition rate between each state is a function of the net utility received by a vehicular node. In particular, if the net utility of a gateway is higher than that of a client, the transition rate from state n_g to $n_g + 1$ (i.e., the number of gateways increases by one) is $(N - n_g)(\theta_i - \pi_i)$. While $(\theta_i - \pi_i)$ indicates the "incentive" for each client to become a gateway, $(N - n_g)$ indicates that all clients can observe the higher net utility of a gateway. Therefore, every client has an equal chance to become a gateway. However, there is a small chance (i.e., ρ) of trying different strategy. As a result, a gateway can switch to become a client, although becoming a gateway can yield a higher net utility than becoming a client. The transition rate from state n_g to $n_g - 1$ is denoted by $n_g \rho$, which is non-zero. In particular, every gateway has an equal chance to become a client. The transition rates for the case that the net utility of a client is higher than that of a gateway can be obtained in a similar way. Finally, the steady-state probability can be computed from this continuous-time Markov chain which can be used to calculate the average number of active gateways in the network.

12.6 Performance Evaluation

We consider a highway with 4 lanes. The average speed of a vehicular node is 64 km/h. The roadside base station allocates 1 Mbps of bandwidth to each connection from gateway. The price of this roadside connection is fixed with $P_b = 10$ MUs. The transmission range of a WiMAX base station is 10 km, while that of WiFi is 100 m. The constants in the utility function are as follows $u_1 = 1$ and $u_2 = 1$. The simulator is developed by using MATLAB with the mobility model for the vehicular nodes similar to that in [35].

12.6.1 Gateway Selection

We first investigate an effect of the price on the number of clients selecting the gateways (Fig. 12.4). In this scenario, the number of gateways is fixed to 3. The
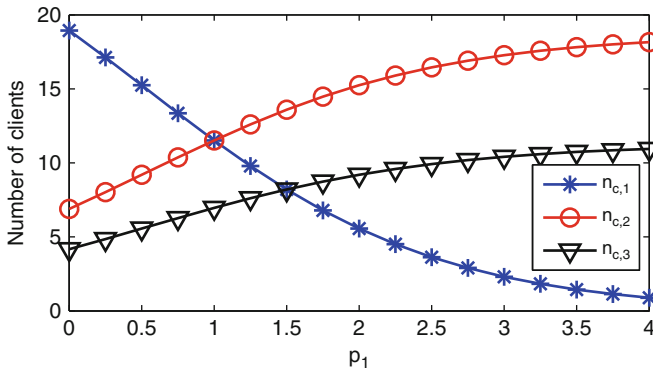
Fig. 12.4 Number of clients associated with each gateway

price of gateway 1 is varied from 0 to 4, while that of gateway 2 and gateway 3 is fixed at $p_2 = 1.0$ and $p_3 = 1.5$. When the price of the gateway 1 increases, the number of clients selecting this gateway decreases. Since the net utility decreases due to higher price, the client deviates to gateway 2 and gateway 3 with lower price to yield the higher net utility. Since the available bandwidth assigned by the roadside base station to each gateway is identical, when the prices offered by two gateways are the same (e.g., $p_1 = p_2 = 1.0$ and $p_1 = p_3 = 1.5$), the number of clients selecting all gateways is the same.

12.6.2 Gateway Selection and Price Competition

Since a client can select the gateway which yields the highest payoff, the number of clients at a particular gateway increases as the price offered by this gateway decreases. From this behavior of the clients, a gateway can optimize its price to the Nash equilibrium such that the optimal revenues can be achieved given the strategies of other gateways. In this case, the variation of the Nash equilibrium for the prices from gateway under different average speeds of the vehicular nodes is shown in Fig. 12.5. When the speed increases, the distance between each vehicle decreases [35]. Also, the density of vehicles decreases, and fewer nodes are in the same network. As the number of nodes in a network decreases, the gateway can increase its price to achieve a higher revenue and subsequently a higher net utility. The number of gateways is also varied in this scenario. The more the number of gateways, the higher the level of competition among the gateways. In such a scenario, to attract more clients, a gateway node decreases the price. Note that as the number of gateways increases, the total end-to-end bandwidth for all clients in the network increases.
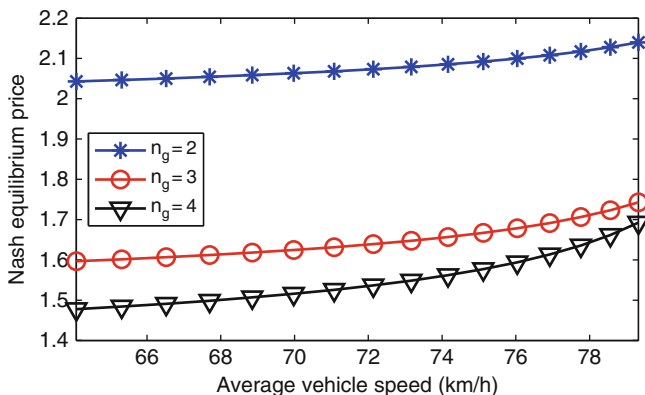
Fig. 12.5 Nash equilibrium under different speeds of vehicular nodes. Increasing mobility of the vehicular nodes increases the price charged by the gateway

12.6.3 Individual Net Utility of Gateway and Client

The individual net utility is a major factor for a vehicular node to decide whether a node should become a gateway or a client. To investigate the incentive of a vehicular node to become a gateway or client, the number of gateways in a network is varied. The net utilities of gateway and client are shown in Fig. 12.6. As the number of gateways increases, the net utility of gateway decreases since the number of clients
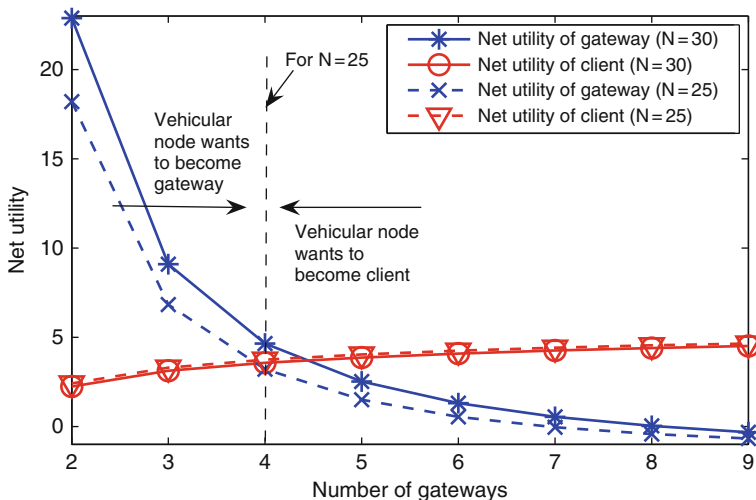

Fig. 12.6 Utility of gateway and client. As the number of gateways increases, the net utility of the clients increases, while that of the gateways decreases

per gateway decreases, and the price charged to the clients is reduced. Conversely, the net utility of a client increases as the number of gateways increases, this is due to larger end-to-end bandwidth assigned by base station and lower price charged by the gateways (i.e., due to a higher level of competition). Also, the number of nodes in a network affects the net utility. When the number of nodes is large, the net utility of gateway is high since more number of clients access the gateway. In this case, if a node can achieve a higher net utility by being a gateway rather than a client, there will be some clients who are willing to become the gateway nodes. However, if the net utility of a gateway is lower than that of a client, there is no incentive for a client to become a gateway. In addition, there is an equilibrium point (e.g., the vertical dash line in Fig. 12.6 for $N = 25$) for the left and right sides of which the vehicular node has an incentive to become gateway and client, respectively.

12.6.4 Total Net Utility and Total End-to-End Bandwidth

The total net utility (i.e., sum of net utility from all nodes in a network) versus the total end-to-end bandwidth is shown in Fig. 12.7 for $N = 30$. Evidently, when the number of gateways increases, the end-to-end bandwidth for all nodes increases linearly. On the other hand, due to decision of a node to become either a gateway or a client, and the price competition among gateways, the total net utility of the network first increases as the number of gateways increases. This increase in the total net utility is due to larger end-to-end bandwidth. However, at a certain point, this net utility decreases, since the price charged by the base station becomes higher than the utility gained from the bandwidth. Here, it is observed
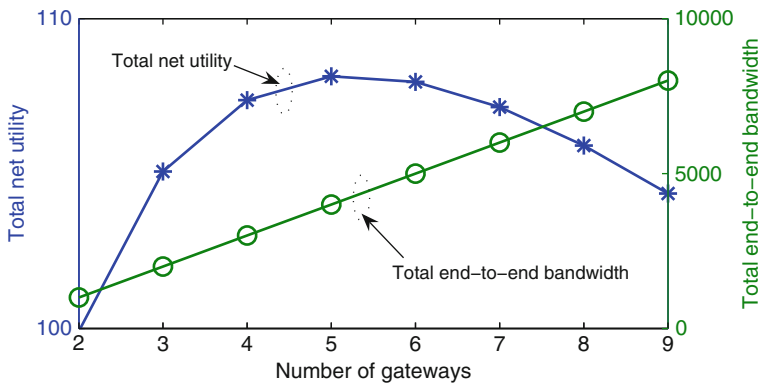

Fig. 12.7 Total utility of the network. As the number of gateways increases, the end-to-end bandwidth linearly increases, while the total utility of an entire network first increases and then decreases. There is an optimal number of gateways the total net utility is maximized
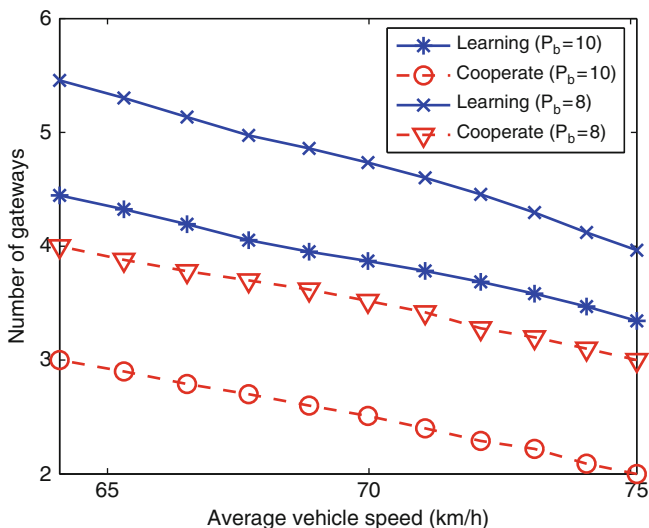
Fig. 12.8 Variation in the number of gateways when the vehicular nodes make their decision independently (i.e., by learning) and cooperatively

that there is an optimal number of gateways for which the highest total net utility is achieved.

12.6.5 Number of Gateways Under Different Vehicle Speeds

Next, the number of gateways in the case that all nodes make decision independently and in the case that all nodes make a cooperation to achieve the highest total net utility are compared. In the former case, an evolutionary game is applied to obtain the average number of gateways. In the latter case, where all nodes fully cooperate, the number of gateways is determined from the location where the total utility is maximized (e.g., 5 gateways in Fig. 12.7). The number of gateways in both cases are shown in Fig. 12.8.

As the speed of vehicles increases, the aggregated bandwidth demand (from fewer vehicular nodes) decreases, and the number of gateways decreases due to change in this demand. In particular, the vehicular node observes that becoming a client yields a higher net utility since the revenue decreases as the number of nodes in the network decreases. Also, the price charged by the base station has a significant effect to the maximum number of gateways. That is, when the base station charges a high price, the net utility of the gateway decreases. As a result, the node is reluctant to become a gateway. In the case that all nodes cooperate, the price paid to the base stations can be reduced. Therefore, the number of gateways is smaller than that in the case when all nodes make their decisions independently.

12.7 Conclusion

An adaptive networking platform for a vehicle-to-roadside communication has been introduced in this chapter. With dual WiFi and WiMAX interfaces, the vehicular nodes can form a cluster-based vehicular network. While the WiFi interface is used for intra-cluster communications, the WiMAX interface is used for cluster-to-roadside communications. A distributed decision-making framework has been developed for a vehicular node to make the decision on a wireless access intelligently and independently. The framework is adaptive to the dynamics of a cluster-based vehicular network. In particular, a vehicular node can become a client in which case its traffic is relayed by a gateway to a roadside base station. Alternatively, a vehicular node can become a gateway which connects directly to a roadside base station and also relays traffic from other clients. The framework is based on a hierarchical game formulation. The decision of a vehicular node is based on the equilibrium solution of the game which ensures that all vehicular nodes in a network are satisfied and do not want to deviate from the solution.

Based on the simulation results, the observations can be summarized as follows:

- The number of clients selecting a particular gateway depends on the price charged by that gateway.
- Vehicle mobility impacts the price at the Nash equilibrium of the gateways and the number of gateways in the network.
- As the number of gateways in the network increases, net utility of gateways decreases, while net utility of clients increases.
- There is an equilibrium number of gateways such that the net utilities of client and gateway are identical. This equilibrium number of gateways can be reached if all vehicular nodes in a network make decision independently.
- There is an optimal number of gateways in the network for which the total net utility is maximized.

Experimental evaluation of the performance and behavior of the adaptive decision framework in a practical vehicular network needs further investigation. The impact of variations in the channel state and node mobility VANET environment need to be investigated.

References

1. Susan R. Dickey, C.-L. Huang, and X. Guan, "Field Measurements of Vehicle to Roadside Communication performance," in *Proceedings of IEEE Vehicular Technology Conference (VTC) Fall*, pp. 2179–2183, September-October 2007.
2. H. Cai and Y. Lin, "Design of a Roadside Seamless Wireless Communication System for Intelligent Highway," in *Proceeding of IEEE Networking, Sensing and Control*, pp. 342–347, March 2005.

3. K. Yang, S. Ou, H.-H. Chen, and J. He, "A Multihop Peer-Communication Protocol with Fairness Guarantee for IEEE 802.16-Based Vehicular Networks," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 6, Part 1, pp. 3358–3370, November 2007.

4. L. Wischhof, A. Ebner, H. Rohling, M. Lott, and R. Halfmann, "Adaptive Broadcast for Travel and Traffic Information Distribution Based on Inter-Vehicle Communication," in *Proceedings of IEEE Intelligent Vehicles Symposium*, pp. 6–11, June 2003.

5. O. Maeshima, S. Cai, T. Honda, and H. Urayama, "A Roadside-to-Vehicle Communication System for Vehicle Safety using Dual Frequency Channels," in *Proceedings of IEEE Intelligent Transportation Systems Conference (ITSC) 2007*, pp. 349–354, September–October 2007.

6. H. Su and X. Zhang, "Clustering-Based Multichannel MAC Protocols for QoS Provisionings Over Vehicular Ad Hoc Networks," *IEEE Transaction Vehicular Technology*, vol. 56, no. 6, pp. 3309–3323, November 2007.

7. B. Sikdar, "Design and Analysis of a MAC Protocol for Vehicle to Roadside Networks," in *Proceedings of IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1691–1696, March–April 2008.

8. C.-J. Chang, R.-G. Cheng, H.-T. Shih, and Y.-S. Chen, "Maximum Freedom Last Scheduling Algorithm for Downlinks of DSRC Networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 8, no. 2, pp. 223–232, June 2007.

9. X. Y. Wang and P.-H. Ho, "A Novel Sensing Coordination Framework for CR-VANETs," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 4, pp. 1936–1948, May 2010.

10. H. Li and D. K. Irick, "Collaborative Spectrum Sensing in Cognitive Radio Vehicular Ad hoc Networks: Belief Propagation on Highway," in *Proceedings of IEEE Vehicular Technology Conference (VTC)-Spring*, pp. 1–5, May 2010.

11. Z. Ahmed, H. Jamal, S. Khan, R. Mehboob, and A. Ashraf, "Cognitive Communication Device for Vehicular Networking," *IEEE Transactions on Consumer Electronics*, vol. 55, no. 2, pp. 371–375, May 2009.

12. S. Chung, J. Yoo, and C. Kim, "A Cognitive MAC for VANET Based on the WAVE Systems," in *Proceedings of International Conference on Advanced Communication Technology (ICACT)*, vol. 1, pp. 41–46, February 2009.

13. K. Tsukamoto, Y. Omori, O. Altintas, M. Tsuru, and Y. Oie, "On Spatially-Aware Channel Selection in Dynamic Spectrum Access Multi-hop Inter-Vehicle Communications," in *Proceedings of IEEE Vehicular Technology Conference (VTC)-Fall*, pp. 1–7, September 2009.

14. D. Niyato, E. Hossain, and P. Wang, "Optimal Channel Access Management with QoS Support for Cognitive Vehicular Networks," *IEEE Transactions on Mobile Computing*, vol. 10, no. 4, pp. 573–591, February 2011.

15. R. Rangnekar, F. Ge, A. Young, M. D. Silvius, A. Fayez, and C. W. Bostian, "A Remote Control and Service Access Scheme for a Vehicular Public Safety Cognitive Radio," in *Proceedings of IEEE Vehicular Technology Conference Fall (VTC)-Fall*, pp. 1–5, September 2009.

16. R. Muraleedharan and L. A. Osadciw, "Cognitive Security Protocol for Sensor Based VANET Using Swarm Intelligence," in *Proceedings of Asilomar Conference on Signals, Systems and Computers*, pp. 288–290, November 2009.

17. K. Tuyls, A. Nowe, T. Lenaerts, and B. Manderick, "An Evolutionary Game Theoretic Perspective on Learning in Multi-agent Systems," in *Synthese Knowledge, Rationality & Action*, vol. 139, no. 2, pp. 297–330, March 2004.

18. D. Niyato and E. Hossain, "Dynamics of Network Selection in Heterogeneous Wireless Networks: An Evolutionary Game Approach," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 4, pp. 2008–2017, May 2009.

19. E. Altman, R. ElAzouzi, Y. Hayel, and H. Tembine, "An Evolutionary Game Approach for the Design of Congestion Control Protocols in Wireless Networks," in *Proceedings of International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks and Workshops (WiOPT)*, pp. 547–552, April 2008.

20. M. P. Anastasopoulos, P.-D. M. Arapoglou, R. Kannan, and P. G. Cottis, "Adaptive Routing Strategies in IEEE 802.16 Multi-hop Wireless Backhaul Networks Based on Evolutionary Game Theory," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 7, pp. 1218–1225, September 2008.

21. J. Hu, B. Wang, J. Wei, and S. Huang, "Evolutionary Game for Distributed Power Allocation over Cooperative Relay Networks," in *Proceedings of International Conference on Wireless Communications, Networking and Mobile Computing (WiCom)*, September 2009.

22. B. Wang, K. J. Ray Liu, and T. C. Clancy, "Evolutionary Cooperative Spectrum Sensing Game: How to Collaborate?," *IEEE Transactions on Communications*, vol. 58, no. 3, pp. 890–900, March 2010.

23. C. Sun, E. Stevens-Navarro, and V. W. S. Wong, "A Constrained MDP-based Vertical Handoff Decision Algorithm for 4G Wireless Networks," in *Proceedings of IEEE International Conference on Communications (ICC)*, pp. 2169–2174, May 2008.

24. D. Niyato, E. Hossain, and S. Camorlinga, "Remote Patient Monitoring Service Using Heterogeneous Wireless Access Networks: Architecture and Optimization," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 4, pp. 412–423, May 2009.

25. A. Galindo-Serrano and L. Giupponi, "Distributed Q-learning for Aggregated Interference Control in Cognitive Radio Networks," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 4, pp. 1823–1834, May 2010.

26. J. Gummeson, D. Ganesan, M. D. Corner, and P. Shenoy, "An Adaptive Link Layer for Range Diversity in Multi-radio Mobile Sensor Networks," in *Proceedings of IEEE INFOCOM*, pp. 154–162, April 2009.

27. B. Xia, M. H. Wahab, Y. Yang, Z. Fan, and M. Sooriyabandara, "Reinforcement Learning Based Spectrum-aware Routing in Multi-hop Cognitive Radio Networks," in *Proceedings of International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM)*, June 2009.

28. M. Li, Y. Xu, and J. Hu, "A Q-Learning Based Sensing Task Selection Scheme for Cognitive Radio Networks," in *Proceedings of International Conference on Wireless Communications & Signal Processing (WCSP)*, November 2009.

29. W. H. Sandholm, *Population Games and Evolutionary Dynamics*. The MIT Press, January 2011.

30. M. Lindstrom and P. Lungaro, "Resource Delegation and Rewards to Stimulate Forwarding in Multihop Cellular Networks," in *Proceedings of IEEE Vehicular Technology Conference (VTC) Spring*, vol. 4, pp. 2152–2156, May–June 2005.

31. K. Chen, Z. Yang, C. Wagener, and K. Nahrstedt, "Market Models and Pricing Mechanisms in a Multihop Wireless Hotspot Network," in *Proceedings of International Conference on Mobile and Ubiquitous Systems: Networking and Services (MobiQuitous)*. pp. 73–82, July 2005.

32. D. Niyato and E. Hossain, "Integration of WiMAX and WiFi: Optimal Pricing for Bandwidth Sharing," *IEEE Communications*, vol. 45, no. 5, pp. 140–146, May 2007.

33. T. L. Vincent and J. S. Brown, *Evolutionary Game theory, Natural selection, and Darwinian Dynamics*. Cambridge University Press, 2005.

34. M. J. Osborne, *An Introduction to Game Theory*. Oxford University Press, 2004.

35. M. J. Lighthill and G. B. Whitham, "On Kinematic Waves. II. A Theory of Traffic Flow on Long Crowded Roads," in *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, vol. 229, no. 1178, pp. 317–345, May 1955.