

Chapter 2

Basics of Vehicular Communication

Due to the recent advances in the wireless technology and widespread use of 802.11 networks, WLAN-based vehicular communication has attracted significant research attention. 802.11-based vehicular communication is a challenging research area with several associated issues. This book is concerned with two of these issues, namely disruption and handover delay. Recall from Sect. 1.2.2 that vehicular communication is classified as V2V and R2V communications. In V2V communications, since vehicles do not communicate with the roadside APs, unplanned deployment of APs is an irrelevant issue. Similarly, since vehicles in V2V scenario do not connect to the APs at all, the delay in handing over to APs is also of little interest. The tolerance of disruption and the associated handover issues are important considerations for R2V communications only and have little relevance to V2V paradigm. Therefore, for the rest of this book, the primary focus stays on 802.11-based R2V communications only.

While the basic concepts pertinent to R2V communication have been discussed in the previous chapter, here we explore relevant previous works that have been reported in literature for addressing disruption and handover latency. First, the research issues in tolerating disruption have been discussed in the light of existing works. The adopted approaches have been divided into several categories for a classified discussion. This is followed by a review of previous works pertinent to the second issue of interest, handover latency. The literature review on handover latency has also been divided into three categories. Towards the end of this chapter, three IEEE standards have been briefly discussed that can affect the performance of 802.11-based vehicular communications.

2.1 Disruption Tolerant Networking

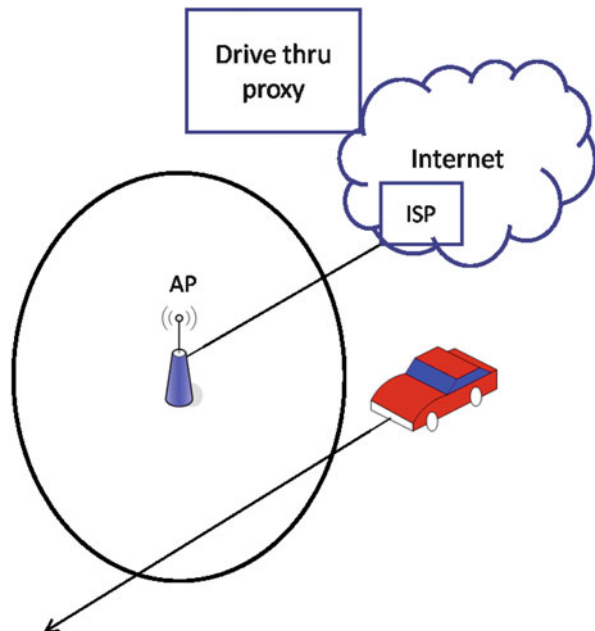
802.11-based vehicular communication suffers from the unplanned deployment of the roadside Access Points. In the case of cellular systems, an operator plans the deployment of the base stations so as to bring the maximum number of subscribers within its coverage region. In case of WLANs, on the other hand, the AP owners deploy WLANs for their own personal (or commercial) use only. The present day WLAN deployment focuses on providing the network services only for a certain business and/or home use and not for all users within an area of interest. The unplanned placement of APs leaves uncovered regions in between the footprints of the neighboring APs. The coverage gap between two consecutive APs is felt even more when the user is moving at a fast speed in the outdoor environment. Thus, unlike the cellular systems, the deployment of WLAN APs does not allow continuous network services on the move. When a mobile node attempts to use WLAN APs while moving at vehicular speeds, it encounters periods where a connection with an AP is possible and periods where no AP is available to make a connection with. This intermittency or irregularity in the network services offered by the WLAN in vehicular scenarios is termed as “disruption,” which is experienced primarily due to the unplanned placement of the WLAN APs.

It follows that for seemingly continuous user experience, the disruption in wireless services must be reduced or at least tolerated. This section focuses on exploring various techniques adopted in the previous works to address disruption in vehicular communication. A comprehensive survey article exploring several dimensions of DTN can be seen in Khabbaz et al. (2012a). The survey article notes that unicast routing has received considerable attention in the domain of DTN research. The survey reported by Pereira et al. (2012) also addresses DTN but from a vehicular perspective. This survey is more inclined towards examining the routing techniques and the applications support in typical vehicular DTN environments. It also outlines some of the existing challenges in this area. The following discussion categorizes the previous works done in the area of disruption tolerance as (1) those proposing new systems and architectures for disruption tolerance, (2) those proposing new or modified protocols for reducing disruption, and (3) those which are based on predicting the upcoming disruption period. In the following, each of these categories is discussed with regard to the previous works reported in the literature.

2.1.1 Systems and Architectures

The terms system and architecture are often used interchangeably. Both imply an interconnection of entities to perform a particular task. System may have different definitions in different contexts. In computer science, software system is a more popular term. A software system is concerned with the interconnection of software

Fig. 2.1 Drive thru internet showing the intermediary proxy in between the mobile node and the network



entities (codes and sub-routines) that often communicate with each other. An architecture is concerned with the functional organization of its constituent units. It may also include consideration of principles and procedures involved in executing a task. The internet architecture, for example, conforms to different principles of the Internet Protocol (IP) suite. Here, some of the systems and architectures are discussed that tolerate disruption in vehicular communications.

Ott and Kutscher (2004a) have proposed a network architecture for using WLAN APs from vehicles (see Fig. 2.1). The proposed architecture requires all mobile nodes to communicate via the “drive thru” proxy placed between the mobile node and the internet. The proxy serves the mobile nodes when the direct connectivity from a roadside AP is available, otherwise, it buffers the requested contents until the connectivity is reinstated.

The main idea of drive thru internet is to conceal the frequent disruption periods from the rest of the network, and to manage or “tolerate” disconnectivity rather than offering continuous service. Since this architecture is not designed for supporting continuous services, the real-time applications such as VoIP cannot be used with satisfactory performance in this setup. In the drive thru architecture, the mobile node uses the standard protocols when communicating with the local hosts but uses an optimized protocol for communicating with the drive thru proxy (Ott and Kutscher 2004b). Using the specific protocols, the drive thru proxy saves the ongoing session when a vehicle exits the footprint of one AP. The same session is restored by contacting the proxy again via the new AP as shown in Fig. 2.2. Pegasus is another system which runs on the in-situ WiFi networks for providing wireless roaming

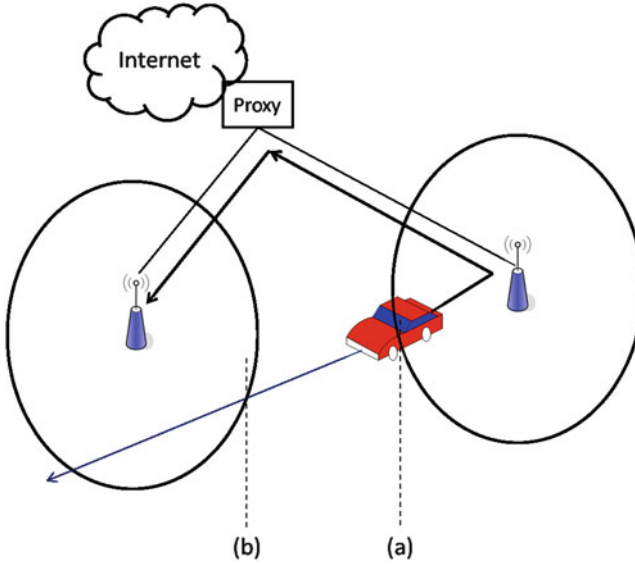
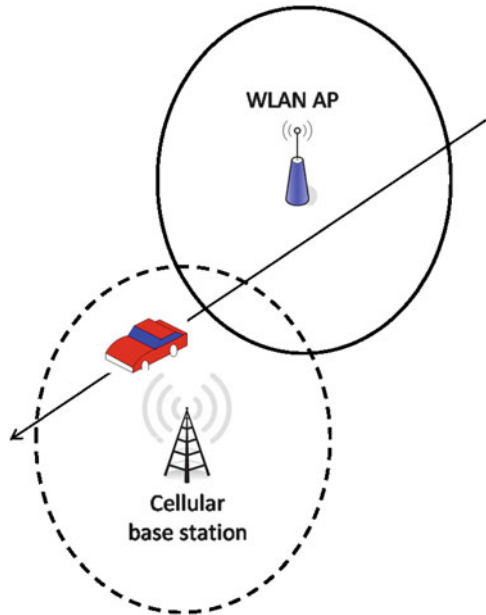


Fig. 2.2 The vehicle saves its session in the proxy at point (a) and retrieves the same from the next AP at point (b)

(Frangiadakis et al. 2007). Pegasus houses a special cache to keep records of all assigned IP addresses which are continuously reused. The primary idea, however, is similar to that proposed in Ott and Kutscher (2004a), in which the overall connection is split into two: one between the mobile node and the manager proxy, and the other between the manager proxy and internet.

The same connection splitting mechanism is also used by Mancuso et al. (2004) to support streaming services in the disrupted networks. The proxy in this case downloads the required stream and delivers the contents to the mobile node whenever the connectivity is available. On top of this proxy system, A2M (All to Minimum) bandwidth sharing algorithm is applied for faster transfers between the proxy and the mobile node. Thedu is also a new system that pre-fetches the web search responses during the disruption periods and transmits the results to the mobile node during the connectivity periods (Balasubramanian et al. 2007). Using Thedu, the mobile nodes can download relevant web responses within 2.7 min using the unplanned network infrastructure. A system that exploits wireless diversity for improved end user experience has been developed as MAR (Mobile Access Router) by Rodriguez et al. (2004). In essence, the idea is to switch to other wireless networks if the prevailing connection is facing disruptions. Figure 2.3 shows that a vehicle switches to the cellular network as soon as the WLAN connectivity disappears. MAR supports several wide area interfaces such as GPRS, UMTS, and CDMA to which the mobile node can connect for getting seemingly continuous network services. An innovative application of the roadside infrastructure for supporting a mobile sensor network has been given in Hull et al. (2006). CarTel is a distributed system for sensing and communicating the sensed stimuli to a central portal via the wireless hot spots (WiFi, Bluetooth, etc.).

Fig. 2.3 MAR allows the mobile nodes to connect to another wireless technology in the absence of WLAN APs



Khabbaz et al. (2012b) have focused on reducing the time delay involved in delivering data in typical vehicular DTNs. According to the proposed method, a source carries bundles of data that are only released to selected passing by vehicles for relaying. The vehicular relays are selected on the basis of their velocities. It has been shown that the proposed Probabilistic Bundle Release Scheme (PBRs) outperforms the legacy Greedy-Based Release Scheme (GBRS). Unlike PBRs, GBRS does not consider the speed of passing by vehicle before releasing a bundle to it for relaying. It has been pointed out in Baccelli et al. (2012) that vehicle density significantly affects information dissemination in vehicular networks. The authors have developed exact expressions for thresholds after which information delivery increases quasi-exponentially with the density of the nearby vehicles.

2.1.2 New and Modified Protocols

In the context of communication networks, protocols define the rules and regulations of data transfer between two hosts. Transmission Control Protocol (TCP) is the most commonly used protocol for exchanging information over the internet. Since the conventional protocols (such as TCP) are not meant to support vehicular communications, several previous works have either modified the existing protocols or have proposed completely new ones.

The Cabernet Transport Protocol (CTP) (Eriksson et al. 2008), for example, has been introduced as a replacement for TCP for 802.11-based vehicular communications. While TCP reduces data rate as soon as it detects packet loss, CTP distinguishes between congestive and non-congestive packet losses and imposes rate adaptation only when packet loss occurs due to network congestion. Consequently, CTP achieves twice as high throughput as the TCP. Along with CTP, ViFi is another protocol that minimizes disruption by using the base station diversity (Balasubramanian et al. 2008). The basic idea behind this protocol is that the frequent disruptions can be masked if a mobile node is associated with more than one base station at the same time. The main challenge in this regard is to develop coordination between the associated base stations. According to ViFi, when a source transmits a packet, it reaches other base stations (auxiliaries) along with the intended destination, as shown in Fig. 2.4. If the destination receives the packet, it broadcasts an ACK message. If an ACK is not heard from the destination, an auxiliary relays the same packet with a certain probability and broadcasts an ACK. Table 2.1 gives the brief flow of procedures in the ViFi protocol. The performance of ViFi is a close approximate of the scenario where the mobile node opportunistically communicates with all nearby base stations.

Fig. 2.4 Exploiting base station diversity as envisaged by ViFi

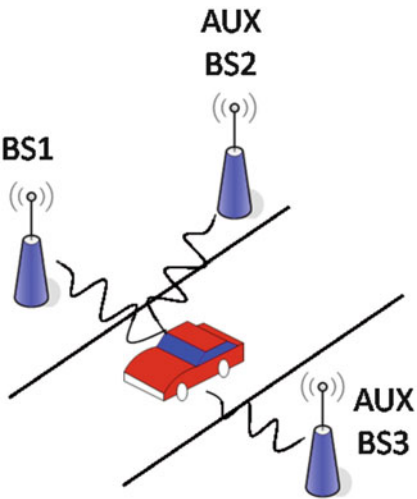


Table 2.1 Flow table for ViFi

SRC transmits message M
If DST receives M, it broadcasts ACK
If AUX hears M but no ACK for a certain time t , it relays M with probability P
If DST receives relayed M, it broadcasts M
If SRC does not receive ACK within a certain time t , it retransmits M

Subramanian et al. (2007) propose several changes in the link-level protocols to “survive” the disruption periods. The link-level enhancements are in conjunction with the loss-tolerant TCP (LT-TCP) protocol proposed for congestion detection. Predicting the vehicle’s route to destination for enhanced routing between the vehicle and the roadside infrastructure has been addressed in Leontiadis et al. (2010). The idea is to determine the vehicle’s route in advance using the navigation system on board the vehicle, and to communicate with the AP closest to the vehicle’s current location. The proposed two-way protocol facilitates both vehicle-to-AP and AP-to-vehicle communications.

A socially aware protocol for unicast and multicast data delivery has been reported in Gao et al. (2012). The main idea is to select only those nodes as data relays that are socially well connected. The centrality metric has been proposed, which determines the extent of social interaction a node has with its neighbors. By definition, the centrality metric reflects on the rate at which nodes come into contact with each other. Selecting a central node for data relaying improves performance in terms of data dissemination. A Delegation Geographic Routing (DSR) scheme has been proposed by Cao et al. (2013), which uses Time To Intercept (TTI) as a metric for decision making. TTI is defined as the time interval between the current encountered node and the previous encountered node.

While the protocols covered in this section may be useful in maintaining continuous connectivity, they may require a great deal of hardware and software changes. The idea of switching between wireless technologies, for example, may require even more changes. Secondly, there is no known method to evaluate the performance improvements brought by the implementation of these architectures and protocols. The mathematical techniques given in Chap. 5 of this book can be used to assess the performance of different techniques that set out to reduce disruption. As an example, the improvements brought by WISPr have been analyzed in Chap. 6 using the developed stochastic models.

2.1.3 *Prediction-Based Techniques*

Reserving the network resources at the new AP helps in executing smooth handovers. Resource reservations require advance knowledge of the candidate APs that may be handed over to in near future. In other words, predicting the path of the mobile node can improve the overall network performance significantly. As the name suggests, prediction-based disruption tolerance is concerned with predicting the connectivity conditions t seconds ahead of time. These forecasts can help in scheduling the network usage in a more efficient manner. Some of the works addressing prediction-based techniques are highlighted in the following.

Based on a discrete-time mobility model that is personal to every device, combined with the past network conditions, Breadcrumbs predicts the near-term connectivity of the device (Nicholson and Noble 2008). It uses the state transition

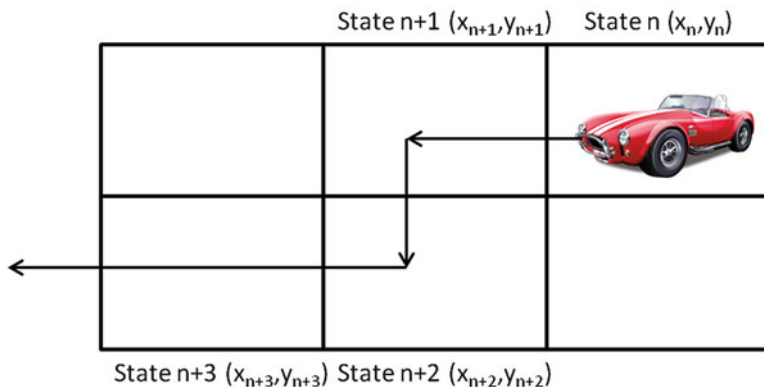


Fig. 2.5 Vehicular motion represented by state transitions in accordance with the changes in geographical coordinates

property of the Markov models to represent the motion of a vehicle as transitions from one state to another. Figure 2.5 shows that the state transition is recorded for different geographical coordinates of the moving vehicle. The location information of the vehicle is obtained by the GPS. This kind of prediction requires not only the knowledge of user mobility but also the performance of the AP encountered by the mobile node. Probably the biggest issue with Breadcrumbs is estimating the performance of the next AP. For this purpose it runs a series of tests on the encountered AP after a usable connection is set up. Not only establishing a connection with an AP is time consuming but the tests with the AP further reduce the amount of usable time.

The concept of making the network resources available at locations which are expected to be visited by the user has been introduced by A-Ghazaleh and Alfa (2010). The Markov Renewal Process (MRP) has been used to make future predictions about the user's location by dividing the geographical area of interest into cells (A-Ghazaleh and Alfa 2008). The user's motion from one cell to another represents its motion along a route. The MRP predicts which cell the mobile node is likely to dwell at t seconds ahead of time. Mahajan et al. (2007) have used the location-specific performance to predict the regions where connectivity is poor. It has been shown that the long term history of a particular location is stable enough to support predictions about the so-called gray periods over a trace. These kinds of history driven predictions have also been discussed in Deshpande et al. (2009). It has been shown that the RF fingerprint (in the form of SNR) at a certain location from a particular AP can be a handy way of predicting the location of the mobile user. The SNR and location information allows offline calculation of when (and where) handoffs should be initiated to provide smoother services and reduced disruption. The historic tracks can also facilitate pre-fetching, which improves system performance by a factor of 2.5 (Deshpande et al. 2009). Lee and Hou (2006) have proposed a model that takes into account both steady-state and

transient mobility behaviors of the WLAN users. Based on the association history of the users, the model not only forecasts about the user mobility but also predicts the prospective APs that the user may associate with in future. A 3-D Markov model has been presented in Hassan and Hassan (2009) to evaluate the performance of a proxy housed inside a vehicle. This onboard proxy attempts to mask service interruptions for scenarios such as when a vehicle (in this case a train) passes through a tunnel.

Cao et al. (2012) estimate the movement of the destination node using the history of its previous positions. The proposed method works best when the prevailing geographical information of the destination is unknown. The issue of positioning a node in DTN has also been addressed in Li et al. (2015). It has been pointed out that, in contrast to the regular networks that have fixed infrastructure, positioning a device in an intermittent network is a considerable challenge. This challenge has been addressed by:

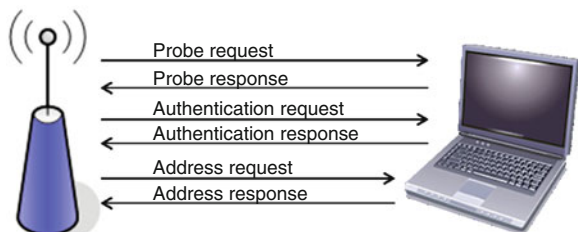
- pulse counting, which counts the steps taken by a walking user and keeps track of its orientation, and
- probabilistic tracing of user trajectory using Markov Chains.

In contrast with the aforementioned works, in this book, 802.11-based R2V communication has been modelled to measure the available disruption. It has been discussed at length in Chap. 4 that R2V communication exhibits the feature of state transition. Therefore, Markov and hidden Markov model techniques have been adopted to measure disruption in 802.11-based vehicular communication.

2.2 Handover Latency in Wireless Networks

Handover occurs when a mobile node leaves the footprint of the previous AP and associates with a new one. Handover procedure itself involves several sub-processes, broadly classified as probing, authentication/association, and allocation of unique identification. These individual handover phases have been shown in Fig. 2.6. The handover starts with the probing phase where the mobile node searches for a candidate AP. After finding an appropriate AP, the mobile node authenticates and associates with it by providing valid user credentials. Finally, the mobile node is assigned a unique identifier (commonly an IP address) so that it becomes reachable over the network. For a fast handover, the delays associated with all these processes should be reduced.

Fig. 2.6 The handover process can be divided into three phases, namely probing, authentication/association, and address allocation



In the vehicular context, the handovers occur very frequently. This is mainly because the WLAN APs cover a small outdoor region. When a vehicle passes by an indoor AP, it can get network services for a few seconds before it gets out of range. To continue using the network services, it has to handover to another AP. One way of reducing the number of handovers is to increase the signal strength of the APs, thus allowing the vehicle to stay within the AP footprint for a longer duration. This of course requires changes in the roadside infrastructure which is not the focus of this book. Instead of reducing the number of handovers, reducing the delay involved in performing the handover is more often addressed. Reducing handover delay is important in the vehicular context because it saves the vehicles from spending a considerable portion of their already small connection time in executing the handovers.

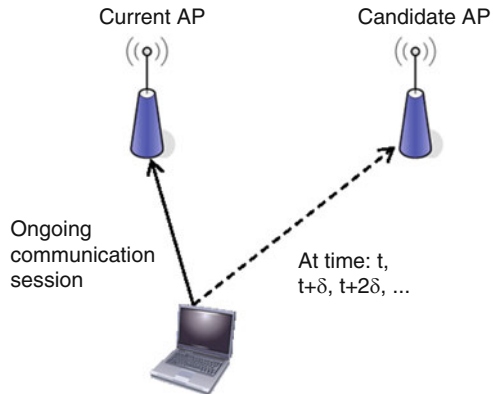
In the following, the previous research efforts in reducing the delays incurred during handover have been reviewed.

2.2.1 Detection, Search, and Probing Delay

Velayos and Karlsson (2004) have divided the handover procedure into three phases, namely detection, search, and execution. In the detection phase, the mobile node realizes the need for handover and the search phase is concerned with acquiring essential information for the handover. Their work has primarily focused on reducing the delay associated with these phases. In another approach, the detection and search phases have been omitted by introducing the pre-active scan phase which scans the neighboring APs without terminating the prevailing connection (Manodham et al. 2005). The pre-active scan phase becomes active once every 2 s and looks for the APs offering a better signal strength than the current one. Replacing the search and detection phases with the pre-active scan phase results in handovers fast enough to support VoIP services. A similar approach has been adopted in Mhatre and Papagiannaki (2006), in which long term and short term signal strength characteristics decide the handover. The mobile node continuously monitors the signal strength levels of the neighboring APs on same as well as overlapping channels. Instead of handing over to APs showing instantaneous increase in the signal strength, the proposed scheme initiates the handover when a decrease (or increase) in the trend is observed for the associated (or the next) AP.

SyncScan in Ramani and Savage (2005) synchronizes the mobile node's listening periods with the beacon transmission periods of the APs. If this synchronization is achieved, the mobile node can scan channels without spending excessive amounts of time in the scanning phase. Figure 2.7 shows that the candidate AP transmits a beacon signal after every σ seconds starting from time t . The mobile node switches to the candidate AP only during these times while still maintaining the usual communication session. In another attempt to reduce the probing delay, HaND (Chen and Qiao 2010) requires a mobile node to jump back to its present channel immediately after sending a probe request on another channel. The scheme can be

Fig. 2.7 SyncScan synchronizes the mobile node's listening periods with the beacon transmission times of the candidate APs



effective if the mobile node knows in advance that it is about to disassociate from the present AP. The concept of interleaving the scanning phase with the ongoing session of the mobile node has been presented in Wu et al. (2007). In this scheme, the mobile node keeps probing the nearby APs during its usual communication session with unnoticeable impact on the overall performance. It also addresses link asymmetry and explores the triggering mechanisms to initiate proactive scan and handover. In another approach, the handover is initiated when there are few (preferably no) packets to be delivered (Choi et al. 2010). The objective in this approach is to reduce packet loss during the handover.

Teng et al. (2009) have established thresholds for starting the pre-scan phase, and for initiating the handover. If the signal strength from an AP falls below the pre-scan threshold, the mobile node starts a background scan and selects 3 nearby APs as the prospective APs. As the AP signal strength becomes smaller, the mobile node hands over to one of these APs.

The use of multiple WNIC (Wireless Network Interface Card) for mobile nodes has been proposed in Jin et al. (2009). The approach adopted here is that one of the channels is dedicated just for scanning purposes while the rest are used for normal communication purposes. Chen et al. (2009) propose the transmission of an authentication request to the APs listed in a smart list as soon as the handover is triggered. The smart list contains the candidate APs selected on the basis of their signal quality. This algorithm is referred to as the Mesh Scan algorithm.

Instead of searching for APs on all 802.11 channels, selective channel scanning calls for examining only those channels which meet a certain criteria. The Neighborhood Graph in Kim et al. (2004), for example, selects the channels to be scanned based on the graph topology. In Tommasi et al. (2006), a scan agent assigns weights to the channels which correspond to the probability of finding a suitable AP on a particular channel. These weights are made available to the mobile nodes which use them to select the next candidate AP. Selective channel scanning has also been used in Eriksson et al. (2008) and Shin et al. (2004), in which mobile nodes prioritize a certain group of channels to scan. The main idea is to restrict the scan cycle to a smaller number of channels to avoid scanning the entire 802.11 spectrum.

In Sect. 7.4 of this book, a channel scanning scheme has been presented which reduces the scanning phase delay in the vehicular context.

2.2.2 *Authentication and Address Allocation Delay*

Security has always been an issue with the wireless networks. The vulnerabilities in the Wired Equivalent Protocol (WEP) require the use of more robust security solutions and protocols (Borisov et al. 2001). However, the robustness in the authentication mechanism should not incur extra delay in the handover procedure. Fathi et al. (2005) have evaluated the latencies in some basic authentication schemes such as Open Systems Authentication (OSA) and Shared Key Authentication (SKA). The fast handover method in Ok et al. (2008) also operates over the OSA mechanism. However, more advanced and robust security solutions are now available for the wireless networks. For instance, the Extensible Authentication Protocol (EAP) and the 802.1x security framework (Chen and Wang 2005) are commonly employed in the recent 802.11 deployments. While EAP ensures network security, its tedious handshake mechanism increases the delay in the authentication phase. Xu et al. (2008) propose to reduce the number of handshakes in the EAP protocol to minimize the authentication delay. A similar reduction in the number of handshakes has also been proposed for the 802.11i networks in Altunbasak and Owen (2004).

Address allocation refers to identifying the network nodes via unique tags or addresses. In a typical IP network, the task of address assignment is performed by the Dynamic Host Configuration Protocol (DHCP).¹ Like all other phases, the delay incurred in this phase should also be kept to a minimum. Cardenas et al. (2008) have proposed to reduce the number of handshakes between a DHCP server and the mobile node in an attempt to reduce the DHCP delay. While the legacy DHCP requires 4-way handshake between mobile node and server (see Fig. 2.8), their proposal calls for retaining only the first and last message (namely DHCP discover and DHCP Acknowledge). While this proposal can speed up address allocation, at the same time, a loss of discover or acknowledge packet can incur intolerable delays. For instance, if the mobile node fails to receive the acknowledge message from the DHCP server, it may keep on waiting for an acknowledgement in vain. Some extensions in the current DHCP protocol, such as reducing the number of handshakes in the protocol and detecting the subnet changes, have been discussed in Floris et al. (2003), while a seamless handoff process has been presented by Chen et al. (2007). It has been proposed in Chen et al. (2007) that the mobile node uses previous AP until a new address from the new AP has been issued to it.

¹See Sects. 7.2.2 and 7.3.1 for detailed discussion on DHCP and its associated delay.

Fig. 2.8 Handshakes involved in the DHCP procedure

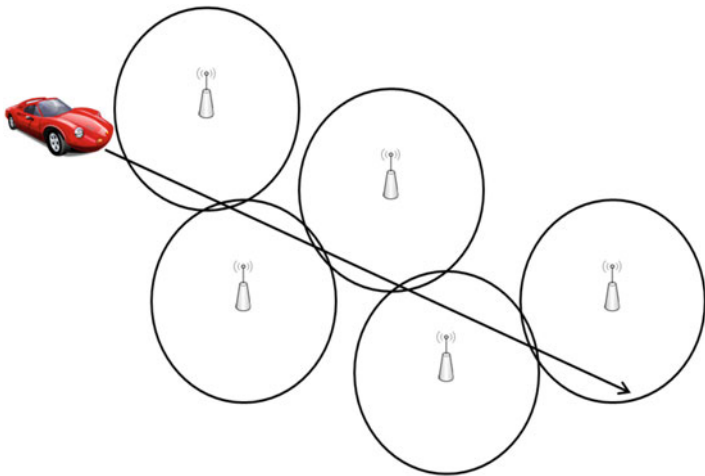
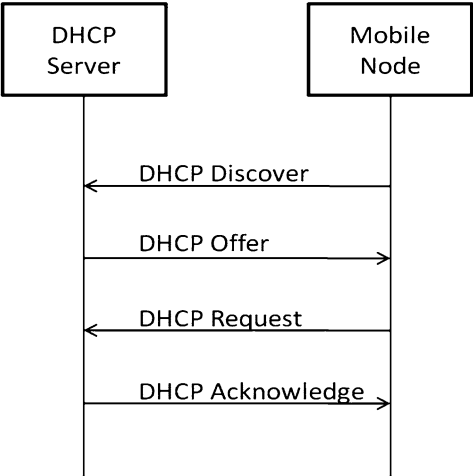


Fig. 2.9 Frequent handovers in vehicular environments

Section 7.3 of this book gives detailed measurement of authentication and address allocation delays. It also highlights the key delay contributors in the EAP-based authentication and DHCP-based address allocation methods.

2.3 Handovers in Vehicular Communication

Smooth handovers are difficult to realize in the vehicular context because of their inherent latency and high frequency. Figure 2.9 shows that the small AP coverage leads to frequent handovers in the vehicular environments. Note that the

figure shows overlapping AP footprints which is not the case with 802.11-based vehicular communications. This section discusses the research efforts in allowing fast handovers in vehicular communications.

The advance information on neighboring APs and the path followed by a vehicle may speed up the handovers in the vehicular context (Kwak et al. 2009). Paik and Choi (2003) have studied the predictability of handovers in WLANs. The proposed method relies on the user's mobility pattern to predict the handover. The impact of velocity of the vehicle has been studied in Emmelmann (2005), which leads to the conclusion that signal strength-based triggering mechanisms cannot be effective in high velocity scenarios. Giannoulis et al. (2008) have explored the disparity in the performance of APs and have used the per-AP scores in making the handover decisions. They report that different APs perform in a different way but their characteristics, although different from others, remain steady over a period. The per-AP scores may therefore be recorded and used repeatedly over a time period.

Apart from detecting and initiating the handovers, some works have focused on analyzing the authentication mechanism required for vehicular communications. Some guidelines about the authentication mechanisms have been covered in Ott et al. (2005). While the private APs have a framework to authenticate the vehicles, the main challenge is with accessing and authenticating with the unsubscribed public APs. To tackle this problem, Luo and Henry (2003) argue that the customers that are not subscribed to the APs should be given an option of one-time payment for immediate access to the network. Once the payment is made, the AP will issue the user credentials and open its resources to the user in a secure way. Rapid symmetric decryption can be employed in an efficient manner for authenticating the vehicles using the hash message authentication code (HMAC) (Zhang et al. 2008a). The simulation results show that this scheme outperforms the conventional Public Key Infrastructure (PKI) and group signature schemes. Recall from Sect. 1.5.2 that the vehicles spend very little time within the AP footprint. Therefore, any authentication framework that causes intolerable delays will not be suitable in the vehicular context.

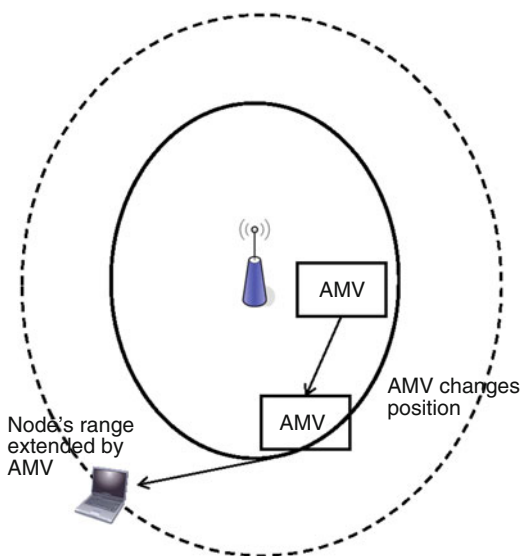
The address allocation issues in vehicular environments have also been discussed in a few works. Vehicular Address Configuration (VAC) has been introduced for VANETs in Fazio et al. (2007), which uses a distributed DHCP mechanism for address allocation. According to this scheme, a vehicular node among a group of vehicles is selected as the leader. The leader serves the IP address requests for the nearby vehicles and acts as a DHCP server for them. The leaders can communicate with each other to maintain a list of already configured addresses. Note that this scheme calls for allowing mobility for the DHCP server and proposes a new protocol (VAC) instead of using the legacy DHCP. In another approach (Mohandas and Liscano 2008), a centralized addressing scheme has been presented for the Vehicular Ad hoc Networks (VANETs). Instead of having multiple DHCP servers scattered across a certain geographical area, this scheme proposes to use one central DHCP server to assign the addresses. The roadside APs act as relays between mobile nodes and the central server to assist in obtaining an IP address. The simulation results show that this scheme leads to a smaller DHCP delay. Although the performance

of this scheme has been simulated for VANETs, its use for R2V communications needs to be evaluated. Chao et al. (2010) proposes to merge the AP and the DHCP server as a single entity. Accordingly, the roadside APs will assign addresses to the mobile nodes instead of contacting the DHCP server every time an address request is received.

Note that the aforementioned works focus on reducing the latency incurred by certain phases of the handover. On the contrary, QuickWiFi (Eriksson et al. 2008) addresses almost all phases of the handover procedure and makes modifications in almost all phases. Apart from proposing changes in the timeout and authentication/association mechanisms, it gives optimal channel scanning and connection loss detection techniques. It also analyzes the performance of individual DHCP phases in the vehicular context.

Reducing the frequency of handovers by controlling the coverage range of the AP may be another approach to reduce irregularity in the 802.11-based vehicular communication. Two range extension schemes have been given in Amdouni and Filali (2009). The Scan-based AP Selection (SAPS) and History-based AP Selection (HAPS) work particularly well for an average density of APs across the road. Both SAPS and HAPS maintain a list of the scanned candidate APs based on a certain criteria. Once the connection with one AP is lost, the next AP is chosen from the candidate list without having to scan the channels again. The motivation is to connect to the AP closest to the mobile node so that its connectivity lasts for larger time duration. Another method of AP range extension is by using the Automatic Mobile Vehicle (AMV) as discussed in Hu et al. (2009). As the mobile node begins to get out of range of the associated AP, the AMV positions itself such that it acts as a relay between the AP and the mobile node, as shown in Fig. 2.10.

Fig. 2.10 AMV changes its position to act as a relay between the vehicle and the AP



2.3.1 *Mobility Management and Heterogeneity*

Recent work in vehicular handovers has been dominated by two main research directions: (1) proposing new and modified protocols that support mobility for vehicles, and (2) exploiting heterogeneous networks for seamless communication experience. These directions are discussed in brief in the following.

In IP networks, mobility is typically managed at layer-3 of the OSI model. A number of mobility protocols exist in literature that offer a diverse range of advantages and apply to different use case scenarios. For example, Hierarchical Mobile IPv6 (HMIPv6) and Proxy Mobile IPv6 (PMIPv6) have been very popular in providing services in high and low speed conditions, respectively (Hasan 2015). Banda et al. (2013b) have developed a Very Fast Mobile IPv6 (VFMIPv6) protocol that minimizes the handover latency in addition to reducing the overhead cost and packet loss. The authors argue that since the address space of IPv6 is very large, it is possible to assign unique IP addresses to all users and access routers. Having globally unique addresses remove the need for the Care of Addresses (COA), which are assigned to a user that is visiting a foreign network. Duplicate Address Detection and other similar processes can be consequently removed, thus reducing the overall time delay. The work reported in Banda et al. (2013a) analyzes the handover of vehicles between consecutive IP subnets. The main concept is to initiate the handover process by accurately predicting the direction and speed of motion of the vehicle. Layer-2 of the OSI model has also been used to enhance the wireless channel access methods for vehicles. Contrary to the existing Point Coordination Function (PCF) for channel access, Chung et al. (2011) have introduced the WAVE Point Coordination Function (WPCF). Wireless Access for Vehicular Environments (WAVE) is the name given to a new mode of vehicular communication which is discussed in the next section. WPCF is meant for V2I communication paradigms that eliminates unnecessary overheads (TXOP, RTS, CTS, etc.). According to the proposed WPCF, the roadside units broadcast a sequence of vehicles that transmit one by one in a pre-defined order. The vehicles are ordered in a sequence based on their urgencies. The message transmitted from a vehicle that is likely to exit the coverage area is treated as urgent. WPCF is shown to outperform several existing channel access solutions.

Most of the mobility protocols discussed in this chapter apply to any one kind of networking technology. For example, WPCF and VFMIPv6 are both meant for 802.11 networks. However, more recent research efforts are geared towards exploiting a range of simultaneously available wireless networks. The so-called heterogeneous network allows a user to connect to several networks at the same time so as to optimize different performance indicators. In such an environment, handovers become more challenging because users are occasionally connecting to a completely new wireless network (Qureshi and Dadej 2012). Dias et al. (2012) examine a mobile setup in which a user hands over between 802.11 networks (802.11p and 802.11g) and the 3G cellular network, as shown in Fig. 2.11. A mobility manager has been designed that sits between L2 and L3 to take care of

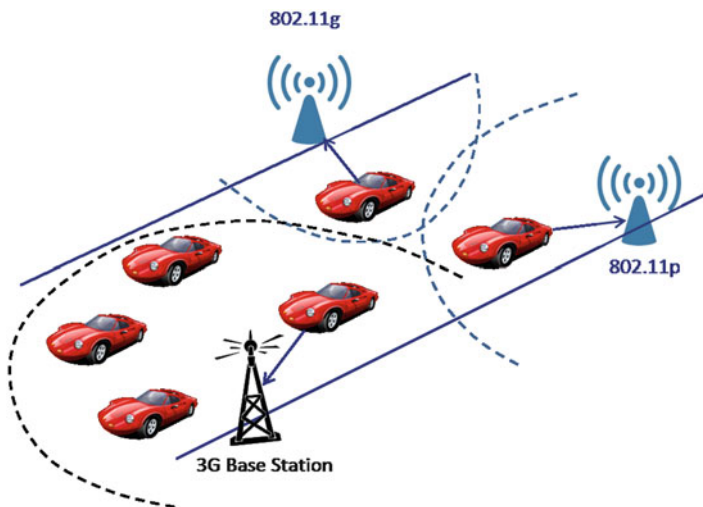


Fig. 2.11 A heterogeneous environment that shows 802.11g, 802.11p, and 3G networks

local and global handovers. The manager independently caters for all L2 handovers, and collaborates with L3 for IP-based handovers. Datta et al. (2012) have used Multi Criteria Decision Analysis to choose the best network for a given user. The parameters of interest include network traffic load, velocity of the user, throughput, and time delay. A mobility management architecture for seamless handovers has been proposed in Meneguette et al. (2013). It has been pointed out that using multiple networks can help in effective load balancing of the messages that are exchanged between the users. This results in lower packet losses and smaller time delays for a larger user population.

2.4 IEEE Standards for Vehicular Communication

Various technologies have emerged over the years which meet the increasing demands of vehicular communications. IEEE has developed various standards that are applicable to vehicular communications. A brief discussion on some of the IEEE standards that may affect the network performance in the vehicular context is presented in this section. The standards IEEE 802.11r and 802.11n are not specifically designed to facilitate vehicular communications, however, their use in this context can be helpful. On the other hand, IEEE 802.11p has been standardized specifically for use in vehicular communications.

2.4.1 Wireless Access in Vehicular Environments: 802.11p

Direct Short Range Communication (DSRC) aims at improving passenger safety by allowing communications between vehicles. The efforts on standardizing and implementing DSRC started in the 1990s (Morgan 2010). As a result of these efforts, IEEE 802.11p has been standardized to facilitate R2V and V2V communication scenarios. The 802.11p standard (often referred to as WAVE) is aimed at enabling vehicles to exchange safety messages such as lane-changing-warning messages and images, for keeping the vehicles aware of the real-time traffic situation (Bilstrup et al. 2008).

IEEE 802.11p is basically a modified version of IEEE 802.11a. The main difference between 802.11p and 802.11a standards lies in the PHY and MAC layers. On the PHY layer, 802.11p uses a reduced channel bandwidth, i.e., 10 MHz in comparison with 20 MHz. Using smaller bandwidth allows a larger guard band between adjacent channels to avoid Inter-Symbol Interference (Stancil et al. 2007). On the MAC layer, 802.11p supports a new mode of operation in addition to ad hoc and infrastructure modes of 802.11a/b/g networks. This new mode of operation is termed as WAVE (Wireless Access in Vehicular Environments) mode. In emergency situations, vehicular nodes operating in the WAVE mode can send and receive messages without associating with a BSS in the conventional manner (Jiang and Delgrossi 2008). The so-called WAVE BSS (WBSS) allows quick commencement of information exchange with an added advantage of very low overhead. The information exchange using WAVE may be between the vehicles or between the vehicles and the roadside infrastructure. Figure 2.12 shows the communication between OBUs and RSUs in a typical vehicular setup, while Table 2.2 tabulates the main differences between 802.11a and 802.11p standards.

Additionally, 802.11e-based priority schemes are also considered for DSRC communication. The objective is to allow service differentiation between safety and non-safety messages (Etemadi and Ashtiani 2011). 802.11e-based VANET and its broadcast services for safety related packets has been discussed in Ma et al. (2009).

Fig. 2.12 OBU-OBU and OBU-RSU communication scenario

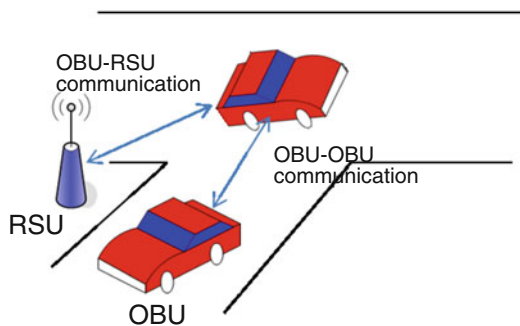
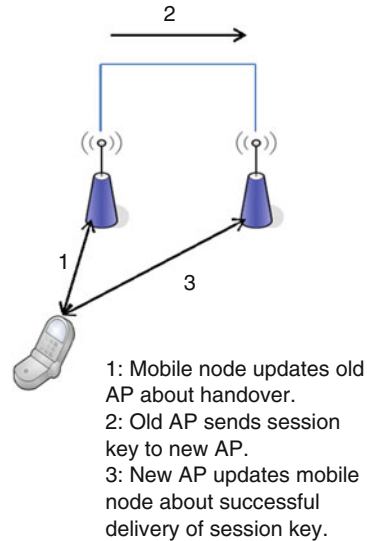


Table 2.2 PHY and MAC layer differences between 802.11a and 802.11p

Layer	802.11a	802.11p
PHY	20 MHz bandwidth, smaller guard band	10 MHz bandwidth, larger guard band
MAC	Ad hoc and Infrastructure modes	WAVE mode in addition to legacy modes

Fig. 2.13 Handover in the IEEE 802.11r FT context



2.4.2 Fast Transition: 802.11r

IEEE 802.11r, also called Fast BSS Transition (FT) standard, allows faster transition of mobile nodes from one AP to another within a single mobility domain. The main purpose of introducing 802.11r is to reduce handover latency and facilitate delay intolerant applications such as VoIP. IEEE 802.11r attains low reconnection times by skipping the authentication process at every reconnection to the AP belonging to the same subnet. Consequently, 802.11r can avoid considerably large authentication delays.

The 802.11r handover procedure is shown in Fig. 2.13 (Hasan et al. 2012), and briefly explained as follows. The first AP to which a mobile node authenticates, referred to as R0 Key Holder (R0KH), saves the Pairwise Master Key (PMK) generated from the authentication procedure. The rest of the APs of the subnet derive their unique session keys using the same PMK stored at the first AP. The session keys derived are used to encrypt messages exchanged between mobile node and AP. As the mobile node transits from R0KH to new AP, R0KH derives PMK-R1 and forwards it to the new AP (R1KH) (Clancy 2008a). In other words, the master key is issued only once; the rest of the session keys are derived from the same master key by the encountered APs. Thus, a lengthy handshake authentication process reduces to two round trip exchanges: first, in which mobile node expresses its desire of handing over to the new AP and second, in which new AP confirms delivery of the new authentication key.

Handover Keying (HOKEY) is another new protocol that incorporates the so-called re-authentication mechanism in EAP authentication protocol (Zheng and Sarikaya 2009). HOKEY adds two new messages to the conventional EAP protocol (see Sects. 7.1, 7.3.2 and Appendix B) that allows commencement and completion of the re-authentication phase. HOKEY also derives subsequent session keys from the key derived in the initial authentication session (Clancy 2008b). However, unlike 802.11r, HOKEY is specific to EAP authentication mechanism.

2.4.3 High Throughput: 802.11n

IEEE 802.11n HT (high throughput) has been introduced to provide data rates as high as 600 Mbps to the WLAN users. 802.11n uses Multiple Input Multiple Output (MIMO) along with Orthogonal Frequency Division Multiplexing (OFDM) and a bandwidth of 40 MHz (instead of 20 MHz) in order to offer high data rates (Goth 2008). Frame aggregation is another method used in 802.11n to increase the throughput of the compliant devices. Since 802.11n networks use multiple antennae for reception and transmission, their transmission signal is not simple to comprehend for the legacy 802.11a/b/g stations. This is because the 802.11a/b/g standards use only one antenna. To prevent 802.11b/g and 802.11n stations from interfering with each other, 802.11n introduces protection schemes at PHY and MAC layers (Paul and Ogunfunmi 2009), which inform the stations about the periods when 802.11n stations are transmitting. The legacy stations stay silent during these durations and hence avoid collision and interference. 802.11n operates in 3 modes, namely legacy, mixed, and greenfield, in order to coexist with the 802.11b/g networks. Table 2.3 tabulates some characteristics of these modes (Hasan et al. 2012). With the popularity of 802.11n in domestic and commercial deployments, the vehicles are expected to attain higher throughputs, and hence an enhanced communication experience.

2.4.4 Very High Throughput: 802.11ac

The 802.11ad and 802.11ac are the two latest standards that focus on offering data rates that are comparable to the gigabit Ethernet. Both standards are improvements on the existing 802.11n standard and retain many of its capabilities. The 802.11ad standard is being considered for 60 GHz band, which results in very high data rates

Table 2.3 802.11n modes of operation

Mode	MIMO	Frame structure
Legacy	Only while receiving	Legacy structure
Mixed	Only with 802.11n stations	Legacy with 802.11b/g and HT with 802.11n
Greenfield	Yes	HT structure

(Goth 2011). The 60 GHz 802.11ad is a potential solution for backhaul communication in dense cellular networks. The high data rates promised by 802.11ad also make it suitable for uncompressed video transmissions (Verma et al. 2013). However, since 802.11ad operates on a very high frequency band (60 GHz), it is a better fit for short range communication only.

On the other hand, 802.11ac operates below 6 GHz, excluding the already populated 2.4 GHz band. The 802.11ac standard supports channel bandwidths of 40, 80, and 160 MHz, as compared to 802.11n which only has 20 and 40 MHz wide channels. One primary channel in 802.11ac is 20 MHz, which is also used for carrier sensing to detect whether another device is transmitting in close vicinity. A set of primary channels are combined in order to get channels with increased bandwidth. For example, the 40 MHz channel has a primary channel of 20 MHz. Similarly, the 80 MHz channels have primary channels of 40 MHz, which in turn have 20 MHz primary channels (Bejarano et al. 2013). The 160 MHz channel is composed of two (contiguous or non-contiguous) 80 MHz channels.

In a recent move, the Federal Communications Commission (FCC) has reserved an additional 195 MHz of bandwidth for 802.11ac devices in the 5 GHz band. The problem is that this new band overlaps with the existing DSRC band, which is being used by 802.11p vehicular communication (Park and Kim 2014). In particular, the channels 172 and 178 (used to time critical data exchange between vehicles) are set to face intense interference from the nearby 802.11ac devices. A bigger concern is that 802.11p uses 10 MHz while 802.11ac uses 20 MHz at the least. Consequently, the Clear Channel Assessment (CCA) using preambles is not possible under the present state of the 802.11p and 802.11ac (Park and Kim 2014). A detailed study has been reported by Park and Kim (2014) on how 20 MHz 802.11ac devices affect the performance of 10 MHz 802.11p vehicles.

Like 802.11n, 802.11ac also employs MIMO. In fact, 802.11ac employs Multi-User MIMO (MU-MIMO). In MU-MIMO, multiple streams of data are sent to multiple receivers at the same time and frequency. Using MU-MIMO, an AP with 4 antennae can simultaneously support 2 users each with 2 antennae. Or, a 4-antennae AP can serve 4 single antenna users at the same time (Bejarano et al. 2013). This technique is often dubbed as downlink MU-MIMO (DL-MU-MIMO) because MIMO is being used in the transmissions originating from the AP to its users.

Using spatial multiplexing, an AP can employ DL-MU-MIMO to simultaneously transmit the data streams to multiple users. DL-MU-MIMO is a concept that is realized at the MAC layer using the Transmission Opportunity (TXOP) mechanism. TXOP was first introduced in the 802.11e standard in order to provide QoS services to the users. According to the legacy TXOP, one of the Access Categories (AC) gets a contention-free period to transmit its data. As the name suggests, in this time period, no other AC contends for the transmission right. On the other hand, in DL-MU-MIMO environment, there are several ACs that wish to transmit data to several users at the same time. Therefore, instead of using dedicated TXOP,

Table 2.4 Prospective IEEE standards for vehicular communications

Standard	Modification	Application in vehicular environment
802.11p WAVE	PHY and MAC layer changes and introduction of WAVE mode	Communications between OBUs and RSUs
802.11r FT	Skips authentication at every reconnection and supports prioritizing traffic classes	Allows very fast handovers between APs
802.11n HT	Uses MIMO and larger BW (40 MHz)	Offers very high data rates
802.11ac VHT	Uses DL-MU-MIMO and larger BW (upto 160 MHz)	Data rates in Gbps

multiple ACs “share” the same TXOP (Aajami and Suk 2015). The AC that acquires a TXOP, called the primary AC, decides whether to share it with other (secondary) ACs. Different ACs that share the same TXOP may have variable-length traffic to transmit. Chung et al. (2015) have developed an analytical model to determine the saturation throughput of 802.11ac when TXOP is shared between ACs having different traffic lengths.

To this end, we have discussed some of the recent standards that relate well with vehicular communication. Table 2.4 summarizes the key characteristics of these standards.

2.4.5 IEEE 802.11ax: Work in Progress

Building on our perpetual thirst of higher throughput, work is underway to introduce 802.11ax standard for WLANs. The application area for this standard is highly dense urban environments, for example, malls, public places, and transport systems (Au 2016). The main target is to increase the spectral efficiency in the environments where a large set of users simultaneously access wireless services. The 802.11ax standard is meant for both indoor and outdoor uses. The expected completion date for this standard is in early 2019.

2.5 Summary

This chapter reviews relevant previous works that address disruption and handover latency issues. A categorized review is given by classifying different approaches adopted in these directions. This chapter highlights the research efforts in modifying and proposing different systems and protocols, in addition to the works that predict the future connectivity conditions of the mobile node. The works addressing the handover latency are categorized in terms of the individual phases of the handover, namely scanning, authentication/association, and address allocation. Some IEEE standards that can significantly affect the performance of vehicular communications are also discussed towards the end of this chapter.

Intelligent Transportation Systems

802.11-based Vehicular Communications

Hasan, S.F.; Siddique, N.; Chakraborty, S.

2018, XXIV, 183 p. 83 illus., 66 illus. in color., Hardcover

ISBN: 978-3-319-64056-3