Efficient Emergency Information Transmission in Vehicular Ad Hoc Networks: Technology Development at MAC and Network Layers

by

Gao Feng

A Thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfilment of the requirements of the degree of

Master of Science

Department of Electrical and Computer Engineering

University of Manitoba

Winnipeg, Manitoba, Canada

THE UNIVERSITY OF MANITOBA

FACULTY OF GRADUATE STUDIES

*****

COPYRIGHT PERMISSION

Efficient Emergency Information Transmission in Vehicular Ad Hoc Networks: Technology Development at MAC and Network Layers

By

Gao Feng

A Thesis/Practicum submitted to the Faculty of Graduate Studies of The University of Manitoba in partial fulfillment of the requirement of the degree

Of

Master of Science

Abstract

For the purpose of improving road safety and reducing billions of dollars property loss every year due to traffic accidents, research communities are working on the development of vehicular ad hoc networks (VANETs) as an important component to realize active auto protection. Besides the emergency warning application, VANETs have potentials to provide other kinds of applications such as road information inquiry, Internet connection, multimedia application, and so on.

In VANETs, medium access control (MAC) protocols at the MAC layer and routing protocols at the network layer are two of the most important issues. For medium access control, this thesis proposed a new protocol, called reliable busy tone multiple access protocol with neighboring information table (RBTMA-NIT), to support emergency warning transmission with the consideration of the priority of copies, distance and neighboring information table (NIT). It has been shown that the RBTMA-NIT can greatly shorten the covering time of emergency message. Meanwhile, another MAC protocol, called multiple channels busy tone multiple access with snooping (S-MBTMA), has been proposed by us to provide high throughput transmission in the multi-channel VANETs environment. By simulation, the S-MBTMA demonstrated its improvement on the system utilization. For routing, this thesis discussed the challenging factors in the routing design, reviewed and analyzed the proposed routing protocols in terms of cost and disadvantage, and proposed guidelines for routing design in VANETs.

Acknowledgement

The author would like to first thank his research adviser Dr. Jun Cai for all his patience and guidance during this two years' program, and his encouragement and help in preparing for this thesis.

The author also wants to thank all the academic professors in the Electrical and Computer Engineering department at the University of Manitoba, who have given lectures or offered help to me.

Many thanks go to all my colleagues in the Robert Alan Kandy Communication Lab and Communication and Network Engineering Research (CNER) group who helped me with the preparation of this thesis.

Finally, I want to specially thank my parents, especially my mother. All of my achievements are impossible without your boundless love and support.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

1.1 General Background

With the development of vehicles and transportation, human beings could enjoy the convenience of automobiles and highways. However, in every year, most countries in this world have to pay costly price for such convenience due to thousands of injuries and billions of dollars loss caused by traffic accidents. For instance, traffic accidents accounted for 230.6 billion U.S. dollars in damaged property (estimation), 2,575,000 nonfatal injuries and 42,642 deaths in USA in 2006 [1]. Although many passive safeguard equipments, such as anti-skid brake system (ABS) and airbags, have been widely adopted, the amount of traffic accident loss remains an astonishing number. In order to improve the road safety and reduce the billions of dollars property loss, research communities are working on the development of vehicular ad hoc networks (VANETs) as an important component to realize active auto protection. By VANETs, vehicles will be notified about the upcoming emergency events with additional common or emergency data, so that drivers or vehicles themselves can take necessary actions in time to prevent potential traffic accidents. Additionally, the designers of VANETs try to add some extra

functions, e.g., road information inquiry, Internet connection and on-board payments. VANET is expected to become a multi-function and comprehensive network in the future. Even though VANETs can be classified as a special case of mobile ad hoc networks (MANETs), their unique properties required that a number of standards adopted for the MANETs have to be modulated or even redesigned for VANETs. Different from MANETs, VANETs face the time-variant network topology under the real-time traffic constraint. On one hand, since every vehicle can carry out lane changing, acceleration, or deceleration randomly, the network topology will change at any time and such variance is much faster comparing with that happened in MANETs. On the other hand, the vehicular behaviors are regulated by traffic regulations, traffic density, weather, terrain, etc. For example, on the highway, the vehicles can only drive along the road or drive to the exit and resting area, and on some intersection, turning left or U-turn are not allowed according to the traffic rule. Beyond the fast varying topology, the complicated propagation environment, where neighboring vehicles, roadside buildings and terrain can put a direct effect on the fading and shadowing of transmitted signals and bring various noises, will become another big challenge for VANET design.

In VANETs, the medium access control (MAC) protocols and routing protocols are two of the most important issues. The medium access control protocols provide reliable connection and decrease contention loss on the common channel, while the routing protocols can improve the validity, reliability and robustness of selected routing path within complicated real road environment. In addition, the routing in VANETs

need to provide high data rate transmission in order to support future comprehensive applications. Therefore, designing the MAC and routing protocols becomes the emphasis of this thesis.

1.2 Research Motivation and Objectives

Our research motivation is to design efficient and practical MAC and routing protocols for VANETs. Considering the different requirements of MAC and routing protocols for VANETs, distinct objectives have been settled. The newly designed medium access control protocols should provide reliable and low delay transmission for emergency warning data, while support gusty and high data rate transmission for common and multi-media data. The developed routing protocols should provide high data rate, low disconnection probability path, and other specific demands according to the practical requirements, such as the high security for personal and bank account information and low delay for the Voice over IP (VoIP) packets.

1.3 Contributions of This Thesis

This thesis focuses on the technology development of VANETs both at MAC and network layers. The contribution of this thesis comes from two aspects: 1) Two new types of medium access control protocols have been proposed. One is called reliable busy tone multiple access protocol with neighboring information table (RBTMA-NIT), which sup-

ports emergency warning message transmission and greatly reduces warning messages covering time. This work has contributed to a conference paper, which was submitted to IEEE International Conference on Communications (ICC) 2010. The other one is the multiple channels busy tone multiple access with snooping (S-MBTMA), which aims to improve the throughput of VANETs for general data; 2) a comprehensive survey on routing protocols in VANETs has been finished. The design factors, challenges, proposed routing mechanisms in literature and their comparison are identified. A framework of routing design in VANETs is provided for the purpose of boosting the development of design work in this area.

1.4 Outline of This Thesis

The rest of the thesis is organized as follows. Chapter 2 introduces the fundamentals of VANETs and mobility models, which are the key differences between common MANETs and VANETs. Motivated by the limits of existing mechanisms, a new MAC protocol, called reliable busy tone multiple accesses with neighboring information table (RBTMA-NIT), has been proposed for emergency message broadcasting in VANETs in Chapter 3. However, RBTMA-NIT is not suitable for the applications other than emergency warning. Therefore, in Chapter 4, another MAC protocol, multiple channels busy tone multiple access with snooping (S-MBTMA), has been proposed by integrating the advantage of existent methods. In Chapter 5, routing protocols in VANETs has been discussed and a useful guideline for routing protocol design has been proposed. Finally,

Chapter 6 concludes the whole thesis and discusses the possible future extensions about designing MAC and routing protocols in VANETs.

2. OVERVIEW OF VEHICULAR AD HOC NETWORKS

Vehicular ad hoc networks (VANETs), as promising extensions of mobile ad hoc networks (MANETs), have been invented to provide various applications on road, such as traffic safety, road information inquiry, Internet connection, etc., and have become more popular recently [2]. In this chapter, we will introduce some general background knowledge about VANETs, including the history, current development of VANETs and some major VANETs applications. In addition, the mobility models, which are used to describe the realistic vehicle movements, have also been discussed in this chapter.

2.1 The Development of Vehicular Ad Hoc Networks

A Vehicular Ad-Hoc Network (VANET) is an extended form of a Mobile Ad-Hoc Network (MANET). Therefore, most concerns in VANETs are similar with those in MANETs. However, the details are totally different. The key difference lies in the fact that vehicles (nodes in VANETs) will move in an organized fashion rather than moving at random as the nodes in MANETs. Since VANETs can provide multiple on-board services, it has been considered as one of the most promising MANET applications. In addition, the VANET has the potential to decrease the traffic congestion, which is a

major problem in modern society, and reduce the amount of vehicle exhaust emitted to the air, which is beneficial on environment protection.

The VANET is critical to realize the Intelligent Transportation Systems (ITS), which has been proposed by the Department of Transportation (DOT) of USA [3]. The expected structure of ITS has been displayed in *Figure 2.1.* In the ITS, the travelers, vehicles, roadside units (RSUs, which refers to field part in *Figure 2.1*) and centers have been connected as a whole system. Every vehicle will be equipped with an on-board device (a special electronic device supplying ad-hoc network connectivity for vehicles), and there are RSUs settled near roads as information processing and switching centers. The major function of VANETs is to provide communication among neighboring vehicles or between vehicles and RSUs, which corresponds to the vehicle to vehicle (v-to-v) and vehicle to roadside unit (v-to-r) communication, respectively.

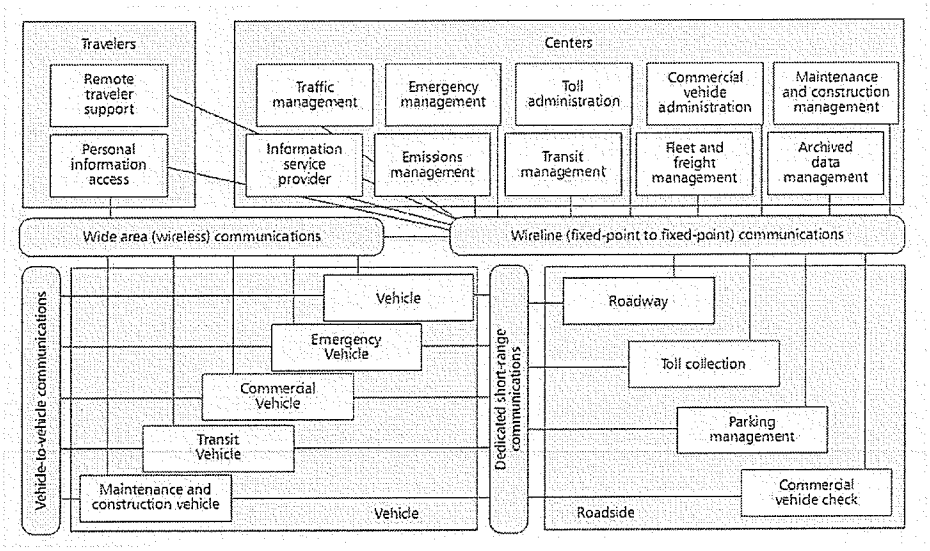
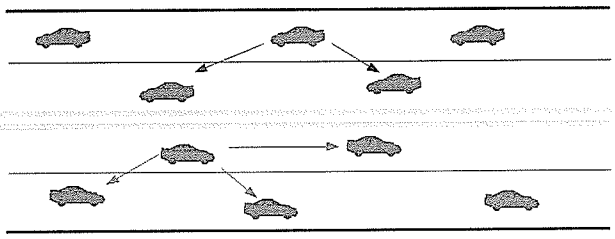


Fig. 2.1: The Structure of Intelligent Transportation Systems

By considering the fact that VANETs are entirely new networks, a huge amount of investment and construction time have to be spent before their wide deployments. Widely accepted VANETs will experience three-stage development, as shown in *Figure 2.2*, *Figure 2.3* and *Figure 2.4* [4][5]. In the first stage (*Figure 2.2*), only a part of vehicles equip on-board devices while the others do not (They will be seen as the "blind points" in VANETs.). In this scenario, the VANET will display as a purely wireless ad-hoc network (By considering the cost of installing RSUs and on-board devices, such expectation is reasonable.). The VANET can only support basic safety related applications, such as emergency warning and driving assistance. The multi-hop transmission between arbitrary nodes will be limited by the fact that the vehicle can only have the driving information around it. In the second stage as shown in *Figure 2.3*, the on-board devices have already been fully equipped. RSUs have been installed at some road sections and can provide service to the surrounding areas. Some applications, e.g., detailed electrical map downloading and electrical toll collection, will become possible under this stage. Notice that the vehicle has high speed and the valid connection time between vehicle and RSU is transitory. Therefore, short-range but high data rate networking technologies such as Orthogonal Frequency Division Multiplexing (OFDM), Worldwide Interoperability for Microwave Access (WiMAX), Ultra Wideband (UWB) and the Very High Frequency (VHF) may be adopted to improve the throughput of the communication between vehicle and RSU. In the third stage (*Figure 2.4*), RSUs have been widely adopted along the road. They have full connections between each other and such con-

nections are assumed to be lossless and highly reliable. In addition, RSUs will be linked to higher-level control centers via wireless or wire line connections. Under this stage, VANETs can maximize their effects on the traffic management, emergency warning, toll administration, and so on. Some applications like user ID authentication and user account security will be available after the integration of control centers, which have enough security level to deal with such sensitive and important data. On account of the fact that VANETs may be more vulnerable than common MANETs, the security of the link will also be strengthened in order to provide a highly reliable network. However, by the consideration of cutting hardware investment, VANETs may share the same control center with other wireless communication system, like cellular networks.



Stage I

Fig. 2.2: Vehicular Ad Hoc Network Development (Stage I)

Given the great prospect of VANETs, the Federal Communications Commission (FCC) has allocated 75MHz spectrum in the 5.9GHz band (5.85-5.925 GHz) for dedicated short-range communications (DSRC), which aims to provide high-speed communications between vehicles and RSUs or among vehicles in the ITS system within a covering range of up to 1,000 meters [6]. (As the comparison shown in Table 1, the

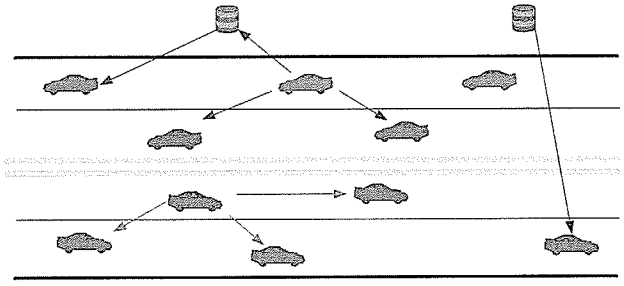


Fig. 2.3: Vehicular Ad Hoc Network Development (Stage II)

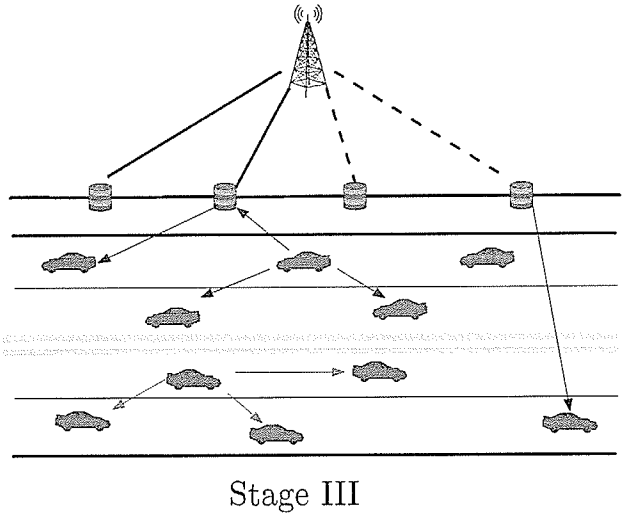


Fig. 2.4: Vehicular Ad Hoc Network Development (Stage III)

DSRC owns the properties of high data rate, low cost, but limited transmission range and directionality. That fulfills the transmission requirement of VANETs.) In Europe, the European Telecommunications Standards Institute (ETSI) has allocated 20 MHz (5.855-5.875 GHz) for immediate and reliable communications between cars, and between cars and roadside infrastructure in August 2008. The employment of 5GHz spectrum can offer low weather dependence and high data rate communications within

	DSRC	FM Radio	Cellular Phone	Satellite
Range	1000 meter	$\geq$ 100 km	several km	$\geq$ 1000 km
Data Rates	6 to 27 mbps	$\leq$10 kps	$\leq$ 10kbps to 2-3 mbps	1Mbps
Directionality	line of sight	area	area	area
Cost (per bit)	free	free	low	expensive

Tab. 2.1: Comparison of dedicated short range communications (DSRC) and other communication ways

1000 meters (maximum), which is suitable for the unique propagation characteristics of VANETs. IEEE and ASTM adopted the DSRC in the standard ASTM E2213-03 [7] in July 2003 [8]. Later, the IEEE 802.11p, also known as Wireless Access in Vehicular Environment (WAVE), has been declared as a draft amendment to the IEEE 802.11 standard for applications in rapidly changing and short-duration communications exchanges environment. According to the official IEEE 802.11 work plan predictions, the approved 802.11p amendment will be published in November 2010 [9]. The IEEE 1609 Standards defines higher layer standard based on the IEEE 802.11p [10].

As a reflection of such popularity, a series of projects have been carried out to develop comprehensive VANETs. Typical examples include Advanced Driver Assistance Systems in Europe (ADASE II) [11], CarTalk 2000 [12] and FleetNet [13].

2.2 The Applications of Vehicular Ad Hoc Networks

Vehicular ad hoc networks have wide applications including emergency warning, driving assistance, information service and on-board entrainment.

2.2.1 Emergency Warning

Emergency warning is one of the major applications of VANETs. Although many passive vehicular safeguard equipments have been widely applied, such as anti-skid brake system (ABS) and airbags, the death and property loss caused by traffic accident remains a big problem in modern societies. The VANET has been invented as an active vehicular protecting method, where any vehicle equipped with VANET devices can get warning about possible emergency event before human can properly react it to greatly shorten the reaction time of emergency situation and take necessary measures before the accidents happen. In the *Figure 2.5*, three typical emergency warning events have been displayed. In the uppermost lane, after the middle car B detects the inter vehicle distance between car A and itself is too close, it will take hard braking. However, the drivers of following vehicles C and D may not react in time partially because they may be preoccupied on the phone or chatting, rather than entirely focusing on the driving. In this case, traffic accidents may happen. But with VANET devices, cars C and D are able to receive the warning message sent by car B, which adopts hard breaking. Therefore, cars C and D will warn the drivers about this event or they may automatically decelerate the vehicles to prevent the potential hazard. The second emergency event happens due to the obstacle on the road. In this case, if VANETs are exploited, the leading car E detecting the existence of an obstacle on the road, such as rock, barrier, or forgo, can send the warning about the obstacle to the following vehicle F. Car F will get this information before the driver can see the real road situation by eyes. Therefore, the road

safety is significantly increased. In the third situation, a collision occurs on the road (as cars G and H in the *Figure 2.5*). Most of passive vehicular safeguard equipments will lose their effects after collision happens. At this time, any collision from the following vehicles will be lethal to the injuries in car G and H. Fortunately, VANET devices can still broadcast the emergency warning of this collision to the following vehicles (Cars I, J and K) in order to prevent the chained collision, and accordingly decrease the loss due to this accident.

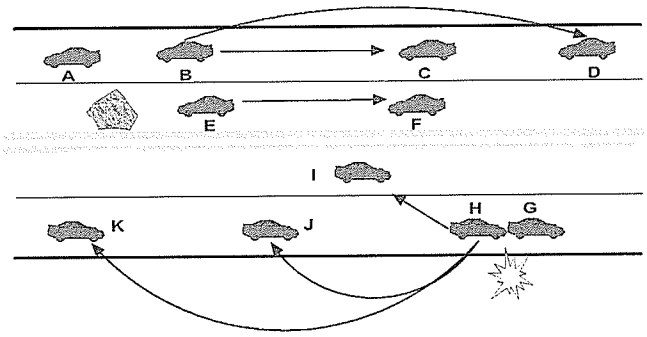


Fig. 2.5: Emergency Warning

2.2.2 *Driving Assistance*

The driving assistance is another important application in VANETs. Although, the global position system (GPS) can provide the general road map and driving direction service, it can not support the real-time driving assistance. VANET devices are capable of filling this gap and make the driving on the road more intelligent. *Figure 2.6* describes four potential application scenarios. In *Figure 2.6(a)*, car A wants to switch its driving lane to the other lane. It has to determine whether the inter vehicle spacing between

vehicles B and C is enough for its interlope and the velocity of car C is slower than car A. Traditionally, such operation is mainly based on the driver's judgment and experience. But with VANETs, the driver can get the exact inter-vehicle spacing from car C. Meanwhile, it will send its lane changing request to car C and ask car C to decelerate. The lane changing process will become more smooth and safer. In the second scenario, since vehicle A wants to go through this road quickly for some emergency reasons, it can send out the overtaking request to the vehicles driving ahead of it. After receiving the request, other vehicles will slow down and give enough space for the overtaking of vehicle A. The third scenario will take place on some small roads. Car A is following a truck, and it wants to overtake the truck. However, the sight view of the driver on car A is greatly blocked by the truck. Therefore, when it overtakes the truck, there is a probability that a vehicle comes from the other direction at the same time. In that case, a collision will happen if the driver of vehicle A can not quickly react to this dangerous situation. With VANET devices, car A can get the real-time road information from the truck and pick the optimal overtaking time. The VANET is also helpful when vehicles go across the intersection, as shown by *Figure 2.6(d)*, where two cars are crossing the intersection. The VANET devices can provide the approaching vehicle information at the orthogonal direction to avoid the collision at the intersection during the mid-night, where the traffic light did not work in a normal way.

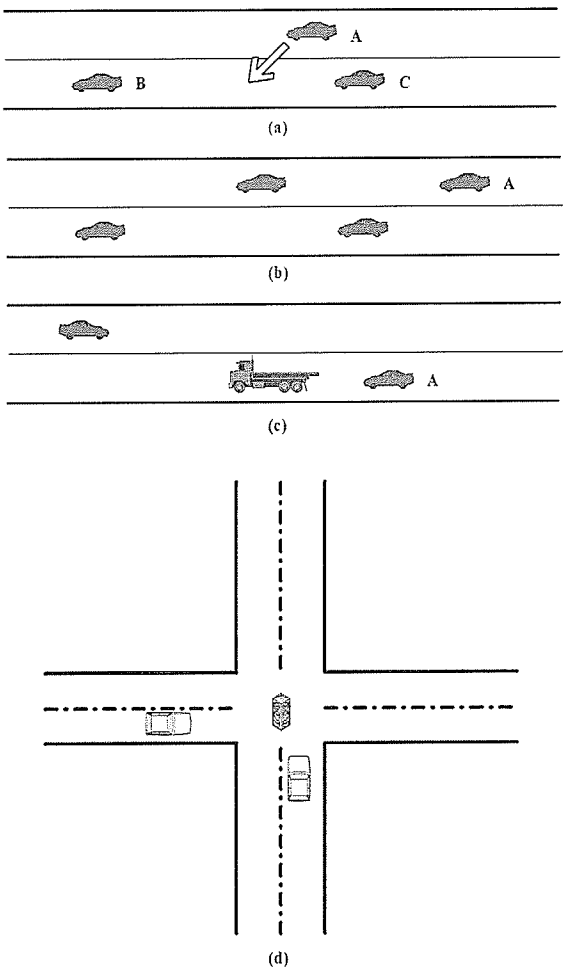


Fig. 2.6: Driving Assistance

2.2.3 Information Service

Information service refers to vehicles' inquiry of the driving related data through VANETs. GPS system is not able to describe the road information in a precise way due to the system limitation. However, the VANET can provide very detailed electrical map (including small roads, dangerous areas and speed limit on each road), which can be stored in some specific RSUs. When vehicles drive through them, the electrical map can be

downloaded according to the users' demand. Meanwhile, VANETs support real-time inquiry. For example, when a driver is looking for a parking space in the downtown area, the VANET device will send request to all neighboring public parks. The RSUs settled beside these parks will feedback the current usage status. Based on these replies, the driver can pick the most suitable parking place. Through this way, the public parks will be used in a more efficient way. Similarly, other public service branches, such as banks and hospitals, can also take advantage of VANETs. With VANETs, the searching and waiting time can be greatly reduced, so that probability of traffic jam is also decreased.

2.2.4 On-board Entertainment and Other Services

The on-board entertainment is one of advanced applications of VANETs. The vehicles with VANET devices can directly access the Internet through the gateway which is installed in the RSU, so that the passengers on the vehicle are capable of browsing webpage, watching Video-on-Demand (VOD), downloading multi-media file, playing on-line games and so on. Meanwhile, VANETs allow the passengers to order meals or do on-line shopping with their credit cards. The video talking or video conference can be realized between the passengers on different driving vehicles within the valid communication range. Beyond that, the on-board electrical ticket charging is possible through VANETs. For example, if one driver violates the traffic law, like going through the red light, the digital camera will record this illegal activity and report it to the control center. Then the control center will analyze this record and send out ticket through

VANETs if the illegal driving behavior has been confirmed. The driver can choose on-board payment when they receive the ticket. This will increase the efficiency of traffic administration. Electronic toll collection (ETC), which is adopted to eliminate the payment delay on toll roads, will become more effective by the integration of VANETs, since VANETs do not require the transponders like current ETC system. However, note that since VANETs are not as reliable as wire line networks, and are more vulnerable to the hostile attack than common wireless networks, the security of private information and account information will be vital to realize these applications.

2.3 Mobility Models

In VANETs, a mobility model refers to an abstract model, which can describe the detailed movement of vehicles in the realistic road environment. It is important to know that a good selection of mobility models will be beneficial on simplifying the modeling work of vehicle movements while increasing the reliability of research results in the real traffic environment. Simple random-pattern, graph-constrained mobility models are widely adopted [14]. However, this model does not take the details of road traffic into consideration. For example, it can not describe the acceleration and deceleration process, waiting process at the intersection, and the effect of neighboring vehicles on driving behaviors. It has been proved that a mobility model which is closer to realistic traffic differs greatly with common random movement model in terms of the system performance [15]. Meanwhile, a reliable mobility model will increase the correctness of

predicting the future vehicle movement, which is beneficial to most routing protocols in this area. Therefore, a suitable mobility model is the premise of practical routing or medium access control protocols design in VANETs.

Figure 2.7 displays a framework for the realistic vehicular mobility model description, which is based on the one proposed in [16]. It shows that a whole mobility model includes two major functional blocks: the motion constraints and traffic generator. The former comes from the topological maps (artificial map or extract from the electrical map) and describes how vehicle moves, while the later captures the vehicle interaction with its surrounding environment.

For the motion constraints, besides the resolution and accuracy of the topological map on reflecting lengths, widths and positions of the real roads, three important factors need to be considered to improve the reliability of the mobility models. The first one is propagation and movement limit. The propagation limit reflects the effect of surrounding environment on the wireless communication signal propagation. For example, in the urban area, the VANET device will suffer more interference from neighboring VANET devices, other wireless communication systems and more serious noises. The roadside building may also affect the transmitted signal. While in the mountain area, such propagation limits are not so serious. The movement limit indicates the vehicle's movement is different at some specific sections on the road, such as intersection, viaduct and roundabout. The second factor is time and geographic factor, which means the traffic will experience a variety on different time periods (rush hours, common hours, or

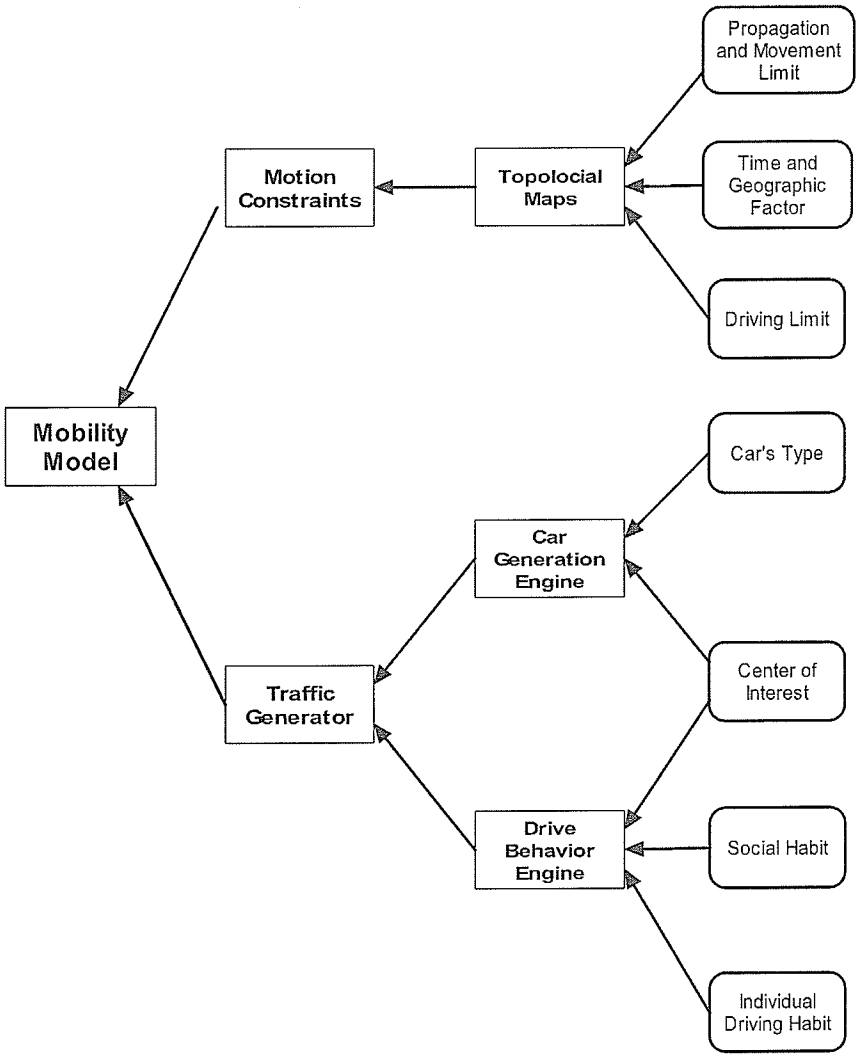


Fig. 2.7: A Framework of Describing Mobility Model

late in night) and different regions (urban, suburban, or rural area). Furthermore, traffic variety will make some regions become hot spots (heavy traffic area) in the specific time. Typical examples include the center business district (CBD) on the rush hours during weekday, and the shopping malls and resorts on the weekend and public holidays. Third, the traffic law and traffic sign will also have an effect on the vehicle movement. The speed limit is a good example. Other instances include situations that most vehicles will stop at the stop sign or when the traffic lights turn red, and the dangerous warning sign makes the driving vehicles slow down.

The traffic generator is composed of two parts, i.e., the car generation engine and driver behavior engine. The car generation engine concerns different vehicle types, (e.g., buses and ambulances will follow easily estimated routes, which differs from the private cars) and the non-uniform distribution of vehicles. The driver behavior engine needs to consider the common social habit, e.g. staying on the slow lane if the driver does not want to overtake the vehicles ahead of it on the same lane and the individual diversity on the dangerous assessments and polite driving.

A great amount of mobility models have been proposed. There are three main classes of mobility models, i.e., synthetic models, survey-based models and trace-based models [16], as shown in *Figure 2.8*. Synthetic models include all the mathematical mobility models, and are the most frequently adopted ones. From macroscopic level to microscopic level, they can be further divided into stochastic models, traffic stream models, queue models, car following models and behavioral models [17]. The stochastic

models cover all models containing purely random motions. Since the VANET is an extension of the MANET, the Random Way Point (RWP) Model [18], which has been frequently adopted in the research of MANETs, was adopted to describe the vehicle behaviors at early stage. The RWP model assumes each node randomly chooses a destination and continues to move forward to that destination at a uniform speed. Since the RWP model did not fit the realistic vehicle movement, a series of improved models have been proposed, such as Manhattan model [19], freeway model and 'virtual track' model [20]. In these models, the vehicle movement will follow the defined path as shown in *Figure 2.9*, *Figure 2.10* and *Figure 2.11*. Recently, electrical map information has been integrated in the mobility model like that shown in *Figure 2.12*. Typical examples include the STreet RAndom Waypoint (STRAW) [21] and GrooveSim [22], where the node [1] travels along the roads indicated on the map. Rama et al. [23] have even proposed a 3-D traffic model to discuss the routing in the more realistic environment, by adding the effect of real buildings on the signal propagation. However, such extension requires a high demand on the simulation hardware. Traffic stream models treat the vehicular movement as a hydrodynamic problem and try to describe the vehicular movement in a macroscopic level by three variables: velocity, density and traffic flow. It avoids the complexity of handling large amount of individual vehicle movement, but it neglects the effect of individual mobility on the connectivity and links duration which reduces the reliability of outcome. Queue model regards each road as a

[1] In this paper, we treat vehicle and node equivalently

First-In-First-Out (FIFO) queue, and each vehicle as a queue client. It can obtain good approximation on some parameters, such as individual travelling time on the given road without high computational costs. But it faces the same problem as the traffic stream models. Car following model tries to use some mathematical model to describe the maintaining of safe distance with previous car and lane changing, which are two of the most important and frequent driving behaviors on the road. With the inter-vehicle spacing and vehicle velocity information, the car following and lane changing can be abstracted by a series of equations. As the scheme proposed by [16], the driver perception and decision making process have been integrated in car following model, so that the car following model is one of the important methods to realize the human-vehicle interaction. The behavioral models make a further step, where every human movement is determined by the behavioral rules (The behavioral rules are affected by the psychological and physiological factors and the neighboring environment). It emphasizes the realistic vehicle drivers' behaviors but with an expensive cost on the computation.

Besides the behavior models, some major large scale surveys, such as the US Department of Labor Survey about US worker's behaviors (including the commuting time, lunch time, traveling distance, etc.) can provide a more realistic and general description about human behaviors [16]. By including these statistics data, the survey-based models can avoid the major limitation in the stochastic models, which assume the vehicle movement is random behavior. However, the correctness and reliability of survey-based models highly depend on the timeliness and reliability of the adopted survey. Meanwhile,

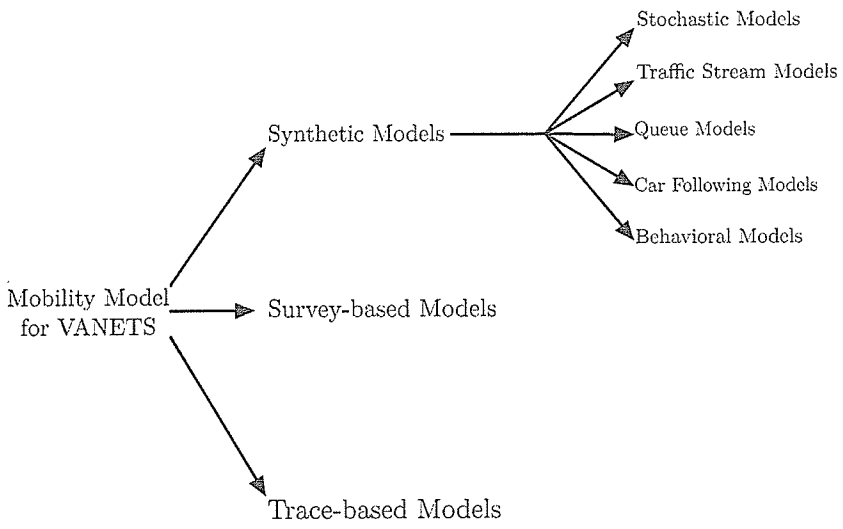


Fig. 2.8: The Classification of Mobility Model

the realistic human behavior may differ at different regions and the detailed status at each area may not be available. The trace-based model is another alternative to depict the vehicle mobility. Comparing to the other options, the trace-based model can closely depict the realistic vehicle movements without applying complex mathematic models. It utilize the vehicular trace data as the realistic vehicle traffic description, which are either abstracted from the measurement of the realistic environment (CrawDaD [24], MIT Reality Mining [25], etc.), or from traffic simulator (PARAMICS [26], CORSIM [27], etc.). They have shown that the vehicular speed and pause time distributions follow a log-normal distribution, while the inter-contact time follows a power law distribution. However, most of trace based models are from commercial groups, and the high price is a major limitation on their wide application. Fortunately, there are still some free tools for the vehicular mobility generation. [28] provides an on-line vehicular trace file

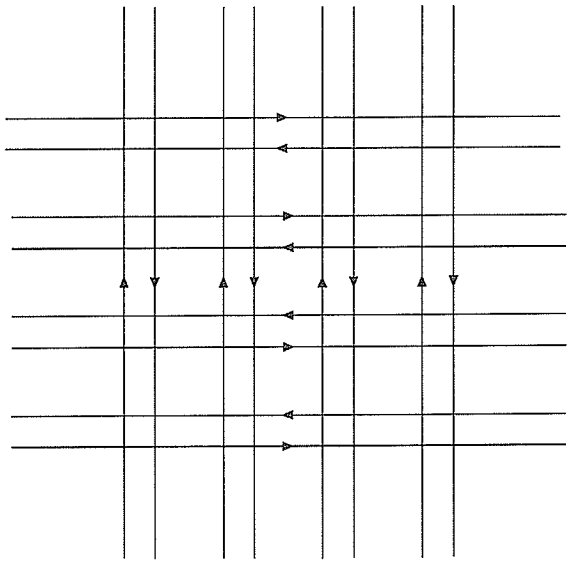


Fig. 2.9: Manhattan Model Topology

generator, which is able to generate trace data under three typical road scenarios chosen from the Swiss geographic information system (GIS). The output is supported by most popular general network simulators, like the NS-2 and Qualnet. Füßler et al. [29] used a set of realistic highway movement data derived from typical situation on German Autobahns to create the node movement trace file. Other examples include Realistic Vehicular Traces [30] and VanetMobiSim [16]. Actually, most of recently proposed network simulators for VANETs, such as NCTUns 5.0 [31], have already integrated this realistic mobility model.

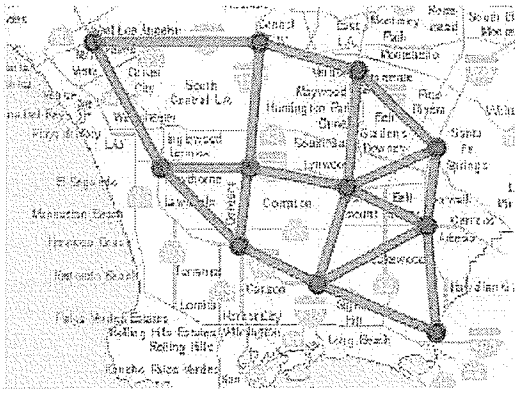


Fig. 2.10: "Virtual Track" Model Topology

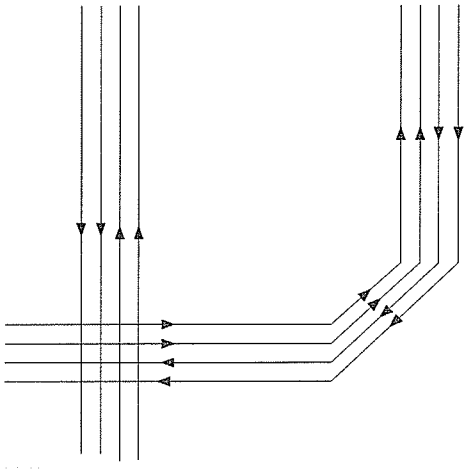


Fig. 2.11: Freeway Model Topology

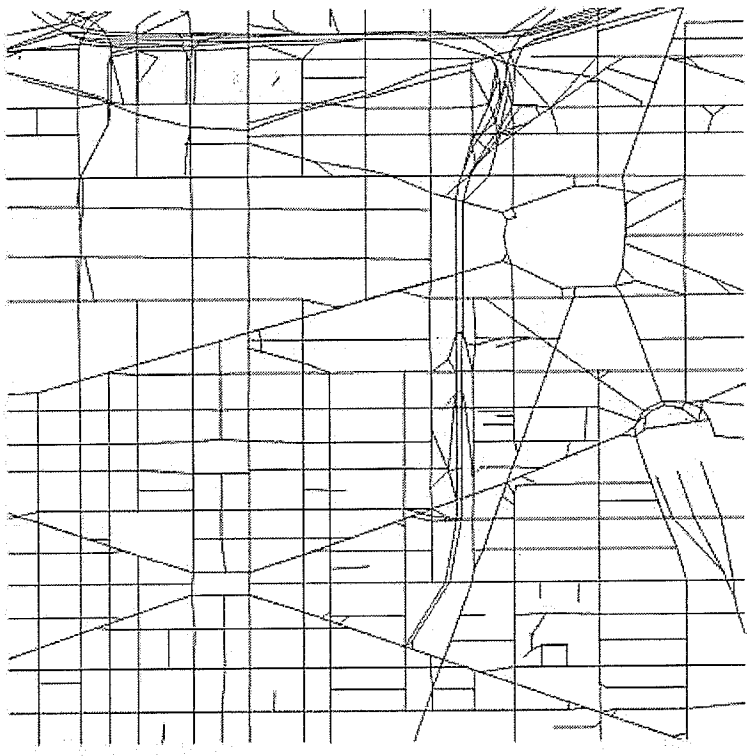


Fig. 2.12: Tiger Map Topology

3. SINGLE CHANNEL MEDIUM ACCESS CONTROL PROTOCOL IN VANETS

Medium access control (MAC) protocols have been proposed to allocate the common channel to multiple users for the successful information transmission. There are two major MAC design concerning issues. One is to avoid or decrease the possibility of collisions while the other is to improve the transmission throughput. In this section, different types of MAC protocols have been introduced and their effects on solving these issues have been verified.

Considering the properties of VANETs, the busy tone multiple access (BTMA) [32] has been selected as the basis for MAC design in VANETs. However, by realizing the fact that the traditional BTMA can not guarantee the shorter delay for high priority node, a new type of MAC protocol, reliable busy tone multiple access with neighboring information table (RBTMA-NIT), has been proposed as an improvement for BTMA application in VANETs. The simulation results demonstrate that the RBTMA-NIT can effectively reduce the covering time during which all nodes in the system have successively received the emergency message. Therefore, it is helpful to improve the safety level of vehicles that adopted VANET device on the road.

3.1 Overview

Multiple access means that several users attempt to utilize the same channel for data packets propagation during the process of communication [33]. Medium access control (MAC) allocates the common channel to multiple users for the successful information transmission with the predefined agreements (protocols) among those users. One of the main aims of MAC protocol design is to avoid or decrease the possibility of collisions when more than one independent user tries to access the common medium simultaneously. Additionally, improving the transmission throughput is another important issue to MAC.

After the first appearance of ALOHA in 1970, numerous MAC protocols have been proposed. There exist many methods to classify the MAC protocols. A typical example is shown in *Figure 3.1.* Accordingly, the MAC protocols can be classified into two major groups based on whether the contention is introduced, and each group is further divided into two subgroups. In the following section, different types of MAC protocols will be briefly introduced. As the repeated random access contention MAC is more relevant to our work, it will be described in more details.

3.1.1 Contentionless MAC Protocols

The contentionless (or scheduling) MAC protocols assign all terminals to transmit in an orderly scheduled manner for the purpose of avoiding the occurrence of collision when two or more users attempt to utilize the channel simultaneously. Usually, all the users

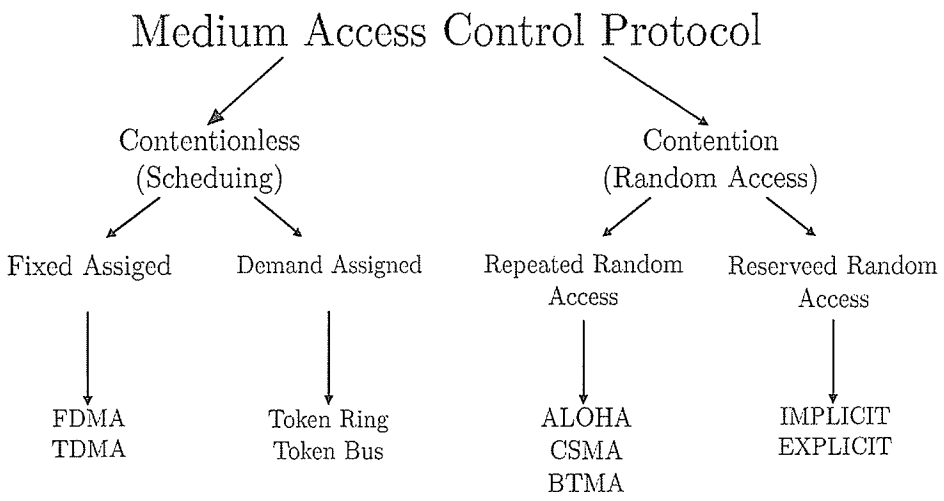


Fig. 3.1: The Classification of MAC Protocols [33]

have equal chance to be allocated exclusive channel resource without the consideration of priority. Due to scheduling, no contention will happen and every transmission will be successful under lossless propagation environment.

Based on the scheduling policy, the contentionless MAC protocols can be further divided into fixed assignment scheduling and demand assignment scheduling. In fixed assignment contentionless protocol, the available channel is divided into several sub-channels either in time domain or frequency domain, which correspond to the Time Division Multiple Access (TDMA) and Frequency Division Multiple Access (FDMA) respectively. Each user is allocated a part of whole bandwidth when it begins to transmit data. In TDMA, the transmitter will be assigned one or several time slots, while it will be allocated one or several non-overlapping frequency bands in FDMA. During one transmission, the assigned channel resource can only be used by the appointed user. In

demand assignment contentionless protocol, users are only able to transmit when they are in active state, i.e., it has data to transmit [33]. According to the decision process of active state, there are two types of demand assignment contentionless MAC protocol: one requires single central control station to manage all transmissions, while the other is distributed where all users will join the process of scheduling. In the central control based contentionless MAC, e.g., the roll-call polling, a central control station polls each user orderly and circularly. If the inquired user has data to send, it will send out data packets to the central control station. Otherwise, it will transmit a negative acknowledgement packet to the central station. The distributed control type contentionless MAC does not require any central control unit. A token, i.e., the permission of data transmission, will be passed among all users, and any data transmission has to wait for the arrival of the token. Since the token is unique, there is at most one transmission at any time. After the transmission is over, the token will be released and passed to next user in a sequential manner. Some good examples of MAC protocols in this type include token ring and token bus.

The advantage of contentionless MAC is that it eliminates the occurrence of collisions when more than one terminal tries to access the common channel simultaneously and it is possible to fully utilize the channel without idle time when demand is heavy. However, the contentionless MAC has more requirements on equipment compared to the contention based MAC. For instance, perfect timing and synchronization are the prerequisites of TDMA MAC protocol, and specific nodes for maintaining the token is

an indispensable part in token ring. Meanwhile, especially for the demand assignment contentionless MAC protocol, the waiting for the qualification of data transmission may cost much time and bring down the network utilization when the network is lightly loaded. These drawbacks have limited the deployment of contentionless MAC protocol in VANETs.

3.1.2 Contention Based MAC Protocols

In contention based MAC protocols, no central control unit exists to manage the data transmission, and there is no scheduling to promise all users transmit in an orderly manner. Therefore, every user has to attend the contention for the common channel when it has data to send. Notice that, all users which are ready to transmit new packets can only be aware of the ongoing data in the common channel, but are not able to know whether other users are trying to send data at the same time. In all likelihood, two or more users would transmit data simultaneously, thus a data collision will occur on the common channel and all of these transmissions will fail. If data collision happen frequently, the throughput, i.e. average rate of successful message delivery over a communication channel, will degrade drastically. As a result, all the contention based MAC protocols should try to suppress the occurrence of data collision.

According to the behavior of the node after successfully accessing the channel, the contention based MAC protocols can be further divided into repeated random access protocols (e.g. ALOHA, slotted(s)-ALOHA, Carrier Sense Multiple Access (CSMA))

and random access protocols with reservation (e.g. reservation ALOHA) [33]. In the former ones, every transmitter can send data immediately when it finds the common channel is idle. Otherwise, it has to wait for the future channel idle period. To eliminate the data collision, channel monitoring and random backoff policy are commonly used. The latter ones are able to reduce the frequency of channel contention. The initial transmission of a user employs the similar access policy as the random access protocols. However, after the user successfully access the channel, part of the channel capacity will be reserved to that user for its future transmission, and other users will be inhibited to use the assigned channel capacity until this user finishes its transmission and the allocated capacity is released. With the reservation, the occurrence of data collision will be greatly decreased.

The contention based MAC protocols are more flexible and suitable for full distributed network, for example VANETs, compared with the contentionless MAC protocols, especially the repeated random access protocols, which do not require any additional devices to realize the perfect timing and synchronization or maintain the token. However, note that the synchronization is necessary for random access protocols with reservation, since the channel reservation requires the channel to be divided into a series of slots as what has been done in the TDMA. But, the throughput of some contention based MAC protocols is quite low. For example, the theoretical maximum throughput for ALOHA and S-ALOHA are 18.4% and 36.8% [33], respectively. The reason is that the data collision will appear frequently when the network is under heavy load. The low

throughput will greatly affect the transmission of the delay-sensitive data, especially the multimedia data, and limit the application of contention based MAC protocol in VANETs.

3.2 Busy Tone Multiple Access Protocol

Due to the cost constraints, it is impractical for VANETs to provide an infrastructure by introducing some central control elements, e.g. access points (APs) in wireless LANs or base stations (BSs) in cellular networks. As a result, all nodes in VANETs have to communicate in a distributed way and share the same channel. To reduce the possible collisions in such environment, Carrier Sense Multiple Access with Collision Detection (CSMA/CA) was first applied [34] [35]. In CSMA/CA, every station waiting for transmission has to listen to the channel for a predetermined time period, and is permitted to transmit when the channel is idle. If the channel is sensed as "busy", then the station intended to transmit has to defer its transmission with a random interval. Such mechanism, on one hand, reduces the probability of collision occurrence compared with basic Carrier Sense Multiple Access (CSMA), while on the other hand, it compensates the defect that the collision detection, which has been shown to reduce the occurrence of collision greatly, is not applicable in the wireless channel due to the difficulty of detection on collision in wireless channel.

However, one of the major drawbacks of CSMA/CA is that it can not solve the hidden terminal problem. *Figure 3.2* explains the occurrence of hidden terminals. In this figure,

node A and node C are outside the carrier sensing range of each other, so that each of them does not know the existence of the other. When node A and node C transmit data packets to node B simultaneously, a collision will happen at receiver B and both transmissions will fail. The hidden terminal problem has been criticized as one of the major limitations on the network throughput and causes of extra delay in most wireless networks. In order to solve this problem, the Distributed Coordination Function (DCF) of IEEE 802.11 utilizes the Request To Send (RTS) /Clear To Send (CTS) handshake, as shown in *Figure 3.3*. In this mechanism, the transmitter will first send a RTS packet to the desired destination node, and the destination node will reply a CTS packet if it is available. After receiving CTS, the transmitter will begin to transmit data. Since the transmitter sends the RTS, which has shorter length comparing with common data, the probability of collision due to hidden terminal will decrease greatly. Meanwhile, as the RTS and CTS are broadcast, the neighboring nodes of both transmitter and receiver can receive them, and those nodes received RTS or CTS are not allowed to send any data or RTS packet. In this way, the hidden terminals apparently do not exist. However, the RTS/CTS handshake can only partly solve the hidden terminal problem. The reason is the occurrence of possible "chained" hidden terminal problem [36]. As shown in *Figure 3.4*, nodes A, B, C and D locate along a line. Assume that the transmitter A attempts to communicate with node B, and sends a RTS packet. Node B successfully receives it and replies with CTS. In that case, node A and node C are expected to receive the CTS. However, coincidently, node D is sending RTS request to node C. Obviously, RTS and

CTS packets will collide at node C, which will cause hidden terminal problem, i.e., node C is the hidden terminal of node B at that time because node C did not know current transmission between nodes A and B. If node C initializes a transmission to node B during the transmission process between nodes A and B, a data collision will occur.

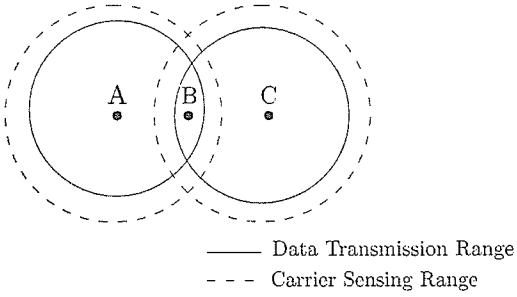


Fig. 3.2: Hidden Terminal Problem

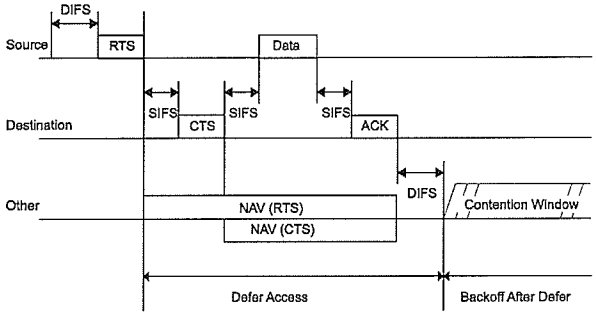


Fig. 3.3: Handshake Progress

In order to combat hidden terminal problem, another method, called busy tone multiple access (BTMA), was proposed [32]. In BTMA, as shown in *Figure 3.5*, two kinds of channels, i.e., control channel (or called busy tone (BT) channel) and data channel, are employed. Two transceivers which correspond to each individual channel are necessary to make sure that every node in the busy tone coverage area can know

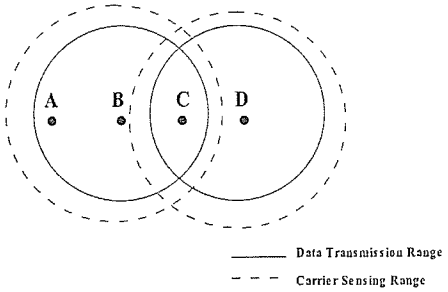


Fig. 3.4: Chained Hidden Terminal Problem

the information about current transmission in time. Note that the transceiver working on the busy tone channel usually is much simpler comparing with the one on the data channel. Therefore it will not bring too much additional cost on equipment. The busy tone (BT) channel is used to transmit single frequency pulses which indicate the ongoing transmission on the data channel, while the data channel is used to carry control packets (RTS, CTS, etc.) and common data. As the BT channel only needs to carry single frequency pulses and do not require any encode or decode operation, it can provide much larger covering range than data channel at the same power consuming level. During the transmission process, the node winning the contention for the channel will simultaneously send information and a continuous signal (busy tone pulse) on the data channel and the BT channel, respectively (The length of busy tone pulse is the same as data length or a little bit longer than the data length by the consideration of propagation delay.). Since the BT channel has larger coverage area than the data channel, the node can hear the busy tone pulse of the current transmitter in a larger distance and will keep silent to avoid the potential hidden terminal problem. However,

one critical drawback of traditional BTMA is that all data packets are treated equally. In fact, in VANETs, depending on their importance, different services ordinarily have distinct requirements on the transmission priority. For instance, the message about the traffic jam is less important than that about the traffic accident. Thus, treating them with different priorities is necessary for VANETs to facilitate the reduction of delay of more important information, such as safety-related messages.

Fig. 3.5: Busy Tone Multiple Access (BTMA)

3.3 Reliable Busy Tone Multiple Access Protocol with NIT

Although the traditional BTMA can suppress the hidden terminal problem, it is not able to solve two important problems in the deployment of realistic VANET, especially safety applications in VANETs: 1) The node in VANETs always has different priority depending on its transmission type. How can the higher priority node access the busy tone as soon as possible? 2) Transmitting multiple copies is used to guarantee data reliability instead of adopting automatic repeat request (ARQ). How can each node acquire the opportunity to occupy the common channel in the procedure of multiple data copies transmission? To deal with both problems, a novel MAC protocol, called the busy tone multiple access with neighboring information table (RBTMA-NIT), is

proposed in this section.

3.3.1 Related Works

In [38], one MAC protocol supporting multiple priorities in data transmission has been proposed, which allows higher priority nodes' interruption to low priority transmissions. In this mechanism, two channels, i.e., BT channel and data channel exist, and a separate transceiver is used for each channel. On the BT channel, a periodic on-off pulse with a single frequency is transmitted. Each period of such on-off pulse, as illustrated in *Figure 3.6*, is divided into two parts. The active part indicates there is an ongoing emergency packet transmission on the data channel. The silent part is reserved for channel contention among higher priority nodes. It is further partitioned into a contention section and a residual random section. In the contention section, except current sender, all nodes with emergency packets waiting for transmission can contend the channel with a random back-off timer. Meanwhile, the current transmitter switches its transceiver on the BT channel into monitor state in order to detect if other users generate BT pulse during that period. The duration of the residual random section is randomly selected to prevent collision among nodes that terminate their back-off timers simultaneously by the identical contention windows. The node who wins the contention will transmit corresponding signals on the BT channel and data channel simultaneously. Message priority is determined by its contents. For example, some urgent emergency message such as the detection of a barrier or an accident on the road has higher priority than

other emergency messages like the following vehicle takes a break since the inter-vehicle distance with the previous car is less than the safe distance. It is defined that the longer active section indicates a higher priority of ongoing transmission, and only the node with higher priority emergency message can interrupt the current transmission. For example, in *Figure 3.7*, node A acquires the channel first. Then node B occupies the channel, but its transmission is interrupted by node C since node C has more important information to be transmitted at that time. The active length of on-off pulse sent by node C is longer than that by node B.

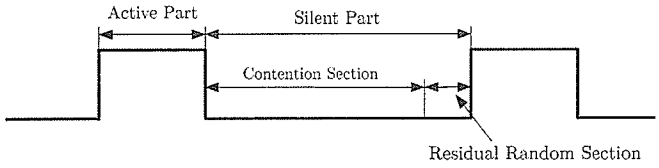


Fig. 3.6: The Structure of On-off Pulse

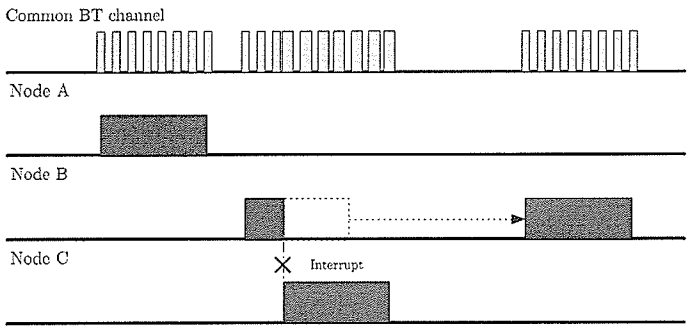


Fig. 3.7: BTMA Working Process

3.3.2 System Model

The vehicular network under consideration is shown in *Figure 3.8*, where several vehicles travel along a single lane [37]. Such scenario corresponds to the typical situation on the highway. Each vehicle is equipped with a Global Positioning System (GPS) to identify geographical position information, and a VANET device for emergency message generation and traffic data record. In the network, only one type of emergency message is taken into account for explanation purpose and multi-copy rebroadcasting is applied to guarantee the reliability of the transmission. However, note that our scheme can be easily extended to the system with multiple emergency message types. In *Figure 3.8*, the arrow denotes the vehicle moving direction, while the emergency message, which has one packet length including information about, for example, the transmitter address, event type and source node ID, is forwarded along the reverse direction with a same modulation and coding scheme. For simplicity, we assume the BT channel can cover all nodes in the network, while the data channel can only cover two neighboring nodes in both sides of the transmitter.

Our pulse structure is similar to the one adopted in the related work. Since single type of emergency message is transmitted, the interruption introduced in the related work may not happen. However, in practice, the copies of an emergency message are less important compared with the original one, since the nodes already successfully receiving the emergency message could help broadcast it. Consider the fact that emergency messages are normally broadcast, and transmitting multiple copies is used to guarantee

the reliability instead of adopting automatic repeat request (ARQ), such differentiation is meaningful to decrease the overall emergency transmission time. In addition, intuitively, the vehicle closer to the source of emergency should have less delay to receive the emergency warning message. Therefore, the vehicle with smaller distance to source should access the common channel faster to broadcast the emergency warning message to its neighbors than the ones in the far distance. The designing of effective contention policy should take all these factors into consideration.

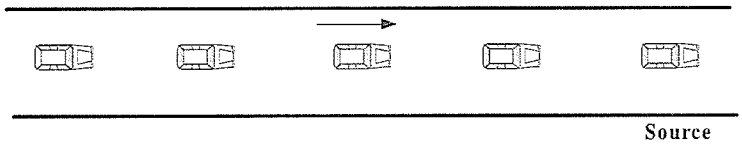


Fig. 3.8: Simple Network Structure

3.3.3 RBTMA-NIT

In this section, the proposed RBTMA-NIT will be introduced with three creative aspects, i.e., diverse priorities assignment, neighboring information table building and contention policy improvement.

Diverse Priorities Assignment

At each node, the original message is defined as the one which is first transmitted by such node, while the repeated transmissions of the same message are called its copies. In RBTMA-NIT, the original message has higher priority than its copies since after the transmission of the original one, such message may have already been successfully

received by all its neighboring nodes which can relay the same message in the future. As in [38], the higher priority of the ongoing transmission will be represented by a longer length of the active section in corresponding on-off pulse on the BT channel.

Neighboring Information Table (NIT) Building

During the transmission of emergency messages, the knowledge about receiving status of neighboring nodes can help reduce the transmission redundancy in the network since each node can terminate rebroadcasting when all its neighboring nodes have received the emergency message correctly. However, since the ARQ policy is not available for broadcasting, the information of neighboring nodes can not be obtained by the ACK packet on the data channel. Fortunately, as the emergency message includes the information of the broadcasting node, the one receiving it can deduce the transmitter who has received this message correctly. Therefore, through this way, each node can build a NIT recording the MAC addresses of its neighbors who have received the emergency message. In addition, the NIT will be integrated into emergency message, so that each node can refresh its own NIT according to the latest received NIT. With the transmitter address and up-to-date NIT, each node can efficiently get the receiving status of its neighboring nodes. Note that since the number of neighbors for each node in the network is finite (4 in this paper), integrating NIT in data packet will not introduce high overhead, i.e., the required additional data byte in the RBTMA-NIT packet will not greatly increase. Meanwhile, the NIT information needs more data processing time. However, these cost

are meaningful on account the system improvement brought by RBTMA-NIT.

Node Contention Improvement

Since the nodes in the network contend for the channel, the contention policy definitely influences the delay of emergency message transmission. To clarify our proposed protocol, two kinds of nodes are defined. They are complete nodes and incomplete nodes. Along the direction of emergency message transmission, the complete node refers to the node whose previous two nodes (on the driving direction) and at least one of its following two neighboring nodes have successfully received the emergency message. All other nodes in the network are defined as incomplete nodes. According to NIT, each node can easily judge its category. Then, we discuss the determination of feasible nodes that are able to contend for the channel. In fact, with the increment of the number of nodes, the collision will be increased accordingly. Based on [38], in the coverage area of the BT channel, the nodes successfully receiving the emergency message will contend for channel during the silent part of on-off pulse on the BT channel. However, some nodes such as the complete nodes need not contend for the channel any more. The reason is that along the direction of the emergency message transmission, one of its following nodes who received the message correctly can fully cover the other one in the range of the complete node. Therefore, to reduce collision, only nodes, whose messages waiting for transmission have higher priority than the current transmission or whose NIT indicates they are incomplete nodes, are allowed to join the competition to interrupt the current

transmission.

Another issue related to contention policy is contention window setting (It is used to define the range of possible back-off timer, and the back-off timer refers the random chosen delay before start RTS or data transmission.) for each feasible node. In general, the node with larger distance to the source of the emergency message can tolerate longer transmission delay. Therefore, in RBTMA-NIT, by taking the distance factor into account, we set the contention window of each feasible node as:

$$CW = \lceil \frac{d}{R} \cdot CW_{Total} \rceil \tag{3.1}$$

where d is the distance between the feasible node and the source of emergency information. Each feasible node can individually calculate such distance by comparing the integrated source information (MAC address, geographic position) in the emergency message with its own GPS information. R is coverage radius of busy tone, CW_{Total} is the maximum length of contention window for each node, and $\lceil x \rceil$ means the minimum integer larger than x. For the source node, its contention window has minimum length. However, its further transmission may also be blocked when it hears another busy tone pulse, which indicates that at least one of its following nodes has successfully received the emergency packet and it becomes complete node. Moreover, the vehicles with smaller distance to the source of emergency message are provided statistically shorter delay in accessing the channel. Therefore, the proposed contention window setting scheme can help the reduction of transmission delay. With such method, after the minimum

back-off timer expires, the corresponding node will begin to generate on-off pulse on the BT channel and transmit emergency messages on the data channel, while other nodes that participate in the contention and hear the on-off pulse will keep silent and wait for future idle time.

However, with such contention policy, the deadlock problem which brings some useless rebroadcasts may happen. For example, as shown in *Figure 3.9*, the square and the cross represent the nodes receiving the emergency message successfully and unsuccessfully, respectively. Consider that the transmission from nodes 1, 3 and 4 to node 2 are all successful, but node 2 fail to transmit data to node 1, 3 and 4 due to channel error or data collision. Then the receiving state of node 2 will not be included in the NIT of nodes 1, 3 and 4. According to RBTMA-NIT's rule, node 2 will stop further transmission, while nodes 1, 3 and 4 will keep on sending message to node 2 and expect to receive node 2's transmission. On the other hand, since node 7 did not receive the message, nodes 5 and 6 will keep broadcasting the information. However, since nodes 1 and 4 have closer distance to the source, they will have higher probability to obtain the channel. As a result, node 2 would receive more useless copies, while node 7 keeps blind to the emergency. The deadlock problem will become more serious under bad channel conditions.

To solve this problem, traditional methods, such as 1-persistence, p-persistence and consecutive transmission limit [39] can be integrated into our proposed MAC protocols. In this paper, for simplicity, we select the third one, which sets a threshold to the

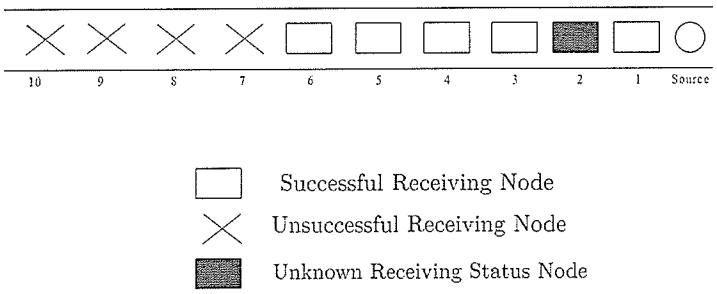


Fig. 3.9: The Occurrence of Deadlock

transmission times of message on each node. Given that the successful transmission probability is P_r, the threshold M is set so that the following condition holds:

$$(1 - Pr)^M < \epsilon \tag{3.2}$$

where ϵ is a predefined threshold. The ϵ is set to give the transmitter a reference about the unreceiving probability when the transmitter reach the transmission threshold M but neighboring node does not become a complete node. With a smaller ϵ, the transmitter is more confident to stop its transmission. But the effect on the deadlock suppression has been abated since it allows more transmission on the transmitter, even they are the redundancy transmissions. In the real world, the setting of ϵ will be based on a large amount of survey results.

3.3.4 Simulation Results

In this section, a VANET with 11 nodes in a line is simulated. We created our own Matlab discrete time simulation code. The node at one end is the source of the emer-

gency message. The maximum contention window CW_{Total} is 150 μs. The active part length of the busy pulse for priorities 1 and 2 are 400 μs and 200 μs, respectively. The busy tone channel has the capability to cover 500 meter area. The data channel bit rate is 2 Mbits/s. Each packet has identical size of 1200 bytes. ϵ is assumed as 0.01. The car following model utilized to describe the inter-vehicle space is [40]

$$S' = L + \beta' V \tag{3.3}$$

where S' is the space between two adjacent vehicles, L is the vehicle length, V is vehicle moving speed, and β' is driver reaction time (Normally, 1.0-1.5 seconds). In the simulation, we choose $L = 4m$, $V = 27.8m/s$, and $\beta' = 1.5s$, so that the inter vehicle space is around 50 meters. In addition, by taking the effect of variable channel condition into account, different successful transmission probabilities have been used to describe various propagation conditions.

As we know, the traffic accident would happen unless all nodes in the BT channel coverage area successfully receive the emergency message in time. Therefore, an important parameter T_{all}, called covering time, is used to measure the system performance. Here, T_{all} defined as the time period between when emergency message is first transmitted and all nodes in the BT channel coverage area have successfully received it.

According to IEEE 802.11p [41], T_{all} should be bounded by 0.5s for the BT channel coverage area with a radius of 500 meter. In the simulation, three different schemes, i.e.,

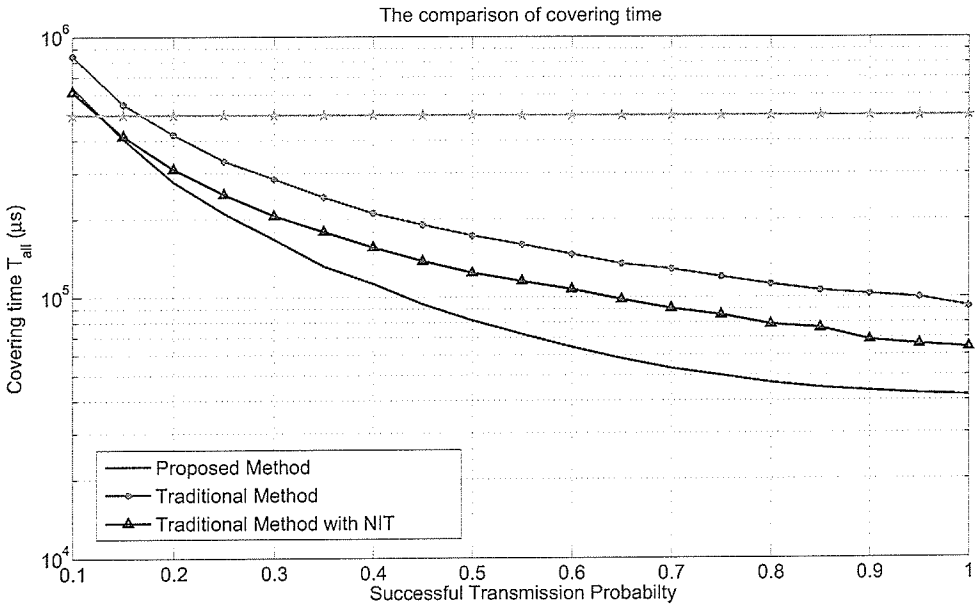


Fig. 3.10: The Comparison of T_{all}

the traditional method in [38] by considering only one kind of emergency message, the traditional method with NIT, and RBTMA-NIT, are compared in term of T_{all}, as shown in *Figure 3.10*. We add NIT in the traditional method to verify the advantage of NIT on reducing the redundant broadcasting. In *Figure 3.10*, the x-axis is the successful transmission probability, y-axis represents average T_{all} in μs, and the horizontal line represents the boundary of T_{all}. Each point on the curve is derived by taking the average value of 1000 samples. From the figure, it can be observed that the average covering time achieved by the proposed protocol is much smaller than the boundary (0.5s). With the aid of NIT information, the covering time achieved by the traditional method can be reduced by around 27% on average. For RBTMA-NIT, it is capable of

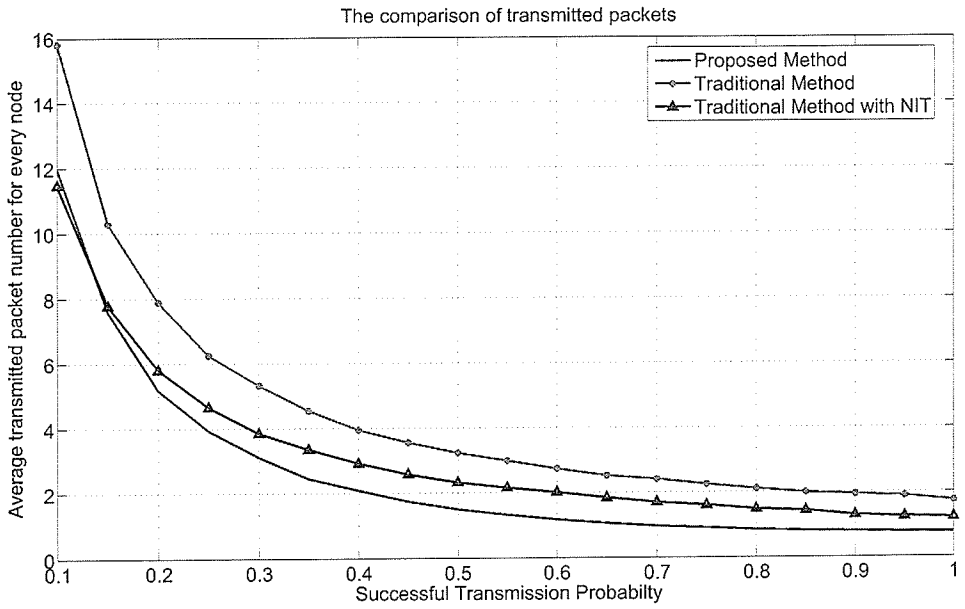


Fig. 3.11: The Comparison of Average Number of Transmitted Packets on Each Node

further reducing T_{all} with the improvement of the transmission condition. For example, if a single transmission has successful probability of 0.9, the improvement is close to 45%. Such performance improvement results from the fact that the average number of transmitted packets from each node is decreased largely by the newly introduced features, which will be clearly demonstrated in *Figure 3.11*. In addition, in *Figure 3.10*, for lower successful transmission probability (less than 0.15), the improvement on T_{all} achieved by RBTMA-NIT is smaller than the higher successful transmission probability case. The reason is that the poor transmission condition requires more transmission trials to realize one successful transmission, which hides the influence from the contention policy. Meanwhile, even if the higher priority node interrupts the transmission of lower

priority node, the transmission after interruption may fail; therefore, the diverse priority assignment almost loses its effect. Under this condition, the improvement is mainly brought by NIT information, which can also be demonstrated by *Figure 3.11*.

Figure 3.11 shows the comparison of the average number of transmitted packets on each node. In this figure, it can be seen that, integrating NIT with the traditional method can reduce the average number of packets transmission for more than 25%, which verifies the capability of NIT on reducing the repetitive emergency message. Furthermore, with the integration of multi-copy priority and improved contention policy, the RBTMA-NIT can further decrease the number of the packet transmissions. The deduction comes from the following three aspects: i) the diverse priority assignment allows the higher priority transmission (original message transmission on each node) to interrupt less important ones (transmission of the copies of the original message), so that the probability of transmitting duplicated packets can be decreased; ii) in the proposed contention policy, the feasible node selection and contention window setting reduce the collision; and iii), the limitation defined in (2) restricts the transmission times of each node. In addition, setting contention window by considering distance factor reduces the access delay of nodes with closer distance to the source of the emergency message. As shown in *Figure 3.12*, *Figure 3.13* and *Figure 3.14*, for different successful transmission probabilities, with the distance increment (i.e., the node number increment), the RBTMA-NIT can bring more improvement on the average receiving delay. The reason is that, with the consideration of distance, the nodes close to the source can receive

emergency message faster, and then become effective relaying nodes. Notice that node 1 and node 2 have similar average receiving time under three mechanisms since these two nodes mostly rely on the transmission of source node, and few relaying nodes will join this process.

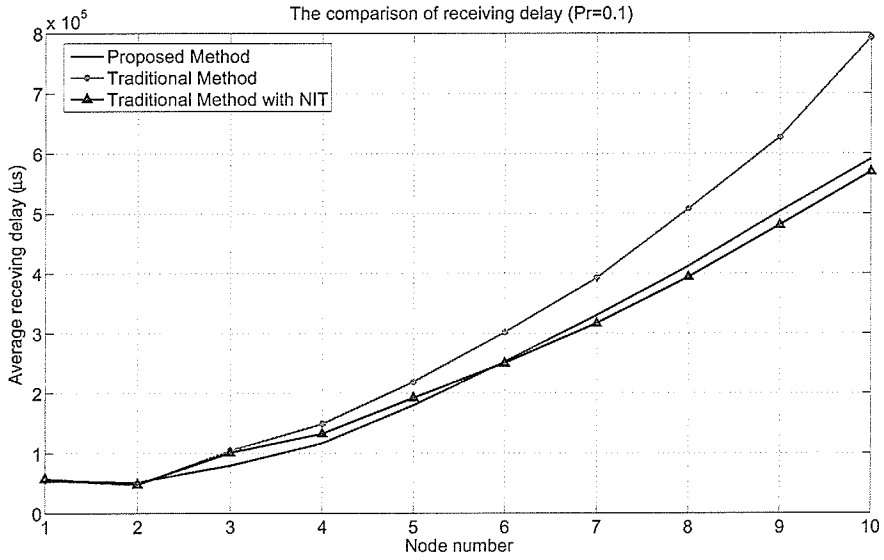


Fig. 3.12: The Comparison of Average Receiving Delay At Different Distance (Pr=0.1)

3.4 Conclusions and Comments

In this chapter, a new MAC protocol, called RBTMA-NIT, is proposed by taking into account the priorities of emergency message and its copies, receiving state of each node and contention policy. The simulation results demonstrate that the RBTMA-NIT is able to greatly shorten the covering time and reduce the average number of transmitted packets.

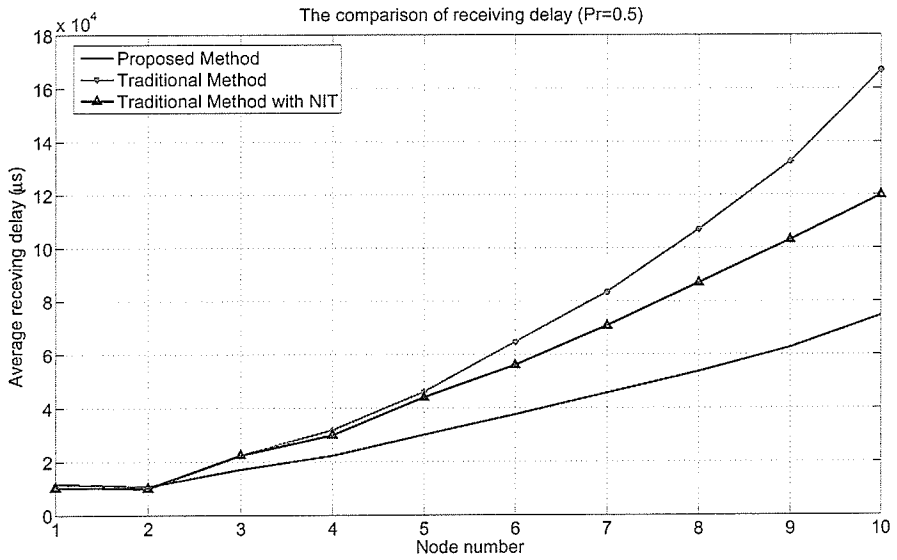


Fig. 3.13: The Comparison of Average Receiving Delay At Different Distance (Pr=0.5)

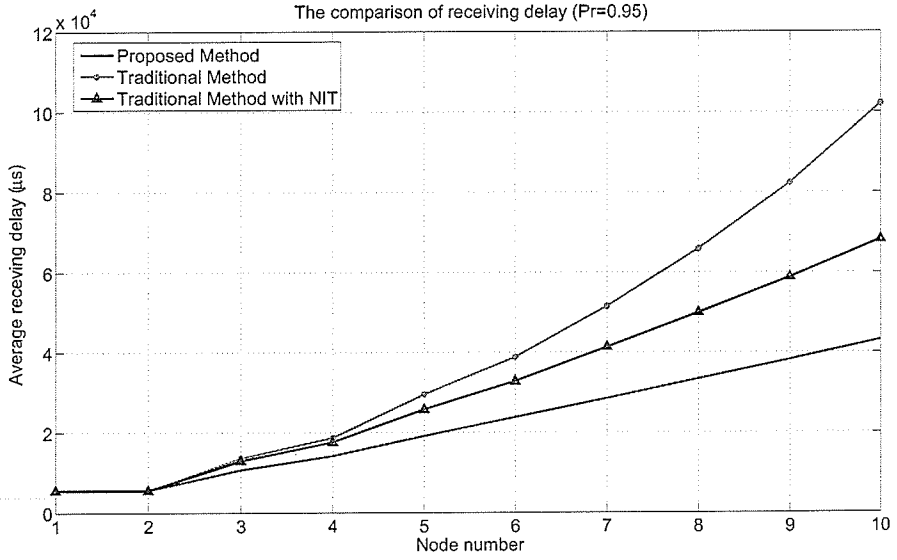


Fig. 3.14: The Comparison of Average Receiving Delay At Different Distance (Pr=0.95)

4. MULTIPLE CHANNEL MEDIUM ACCESS CONTROL PROTOCOL IN VANETS

Although the RBTMA-NIT is able to suppress the hidden terminal and decrease the broadcasting delay for the emergency message, it can not solve the low efficacy of single channel BTMA protocol, i.e., the single channel BTMA protocol only allows at most one transmission at any time even if the potential transmission will not cause hidden terminal problem. For common data, such drawback will greatly decrease the network throughput and limit the application of the BTMA protocol in VANETs. There are two ways to solve this problem. One is to introduce multiple channels in the network, and the other is improving the utilization of each single channel. In this chapter, we will apply both methods to design a new multi-channel MAC protocol. However, notice that multiple channel medium access control (MMAC) protocols have their own problems, and some improvement works are necessary.

4.1 Overview

MMAC protocols have been proved to be much more efficient than the single channel MAC in wired Local Area Network (LAN), as shown in [42]. In the wireless network, multiple channels accessing are also widely deployed and can bring three major benefits [45] [44] [42]. First, it can increase the system throughput and decrease the average accessing time. As the maximum throughput on each channel is limited by the channel bandwidth, the increment of available channels can bring the augmentation of the overall system throughput. Second, the transmitter will experience less normalized propagation delay (Normalized propagation delay is defined as the ratio of the propagation time over the packet transmission time.) per channel than the single channel counterpart [42]. It is straightforward to notice that the potential transmitter, when it detects all channels are busy, has multiple choices for the next propagation channel with the lowest delay other than the unique one in the single channel environment. The third one is that using multiple channels is more favorable to support quality of service (QoS) request. Multiple channels allows transmitter to access distinct channels with different delays [43], while these operations are hard to be realized in a single channel. In this thesis, the discussed circumstance refers to the condition that multiple channels exist, but each node is only equipped with one transceiver. How to effectively use these channels and maximize their throughput is our main concern.

4.2 The Main Challenges

Although multi-channel accessing can bring some benefits, a series of challenges exist when it is applied in the real world. In the following, we assume that the node in the network is only equipped with a single transceiver, and some potentially difficult problems for MMAC are discussed.

4.2.1 Multi-channel Hidden Terminal Problem

In the single channel environment, the hidden terminal problem occurs when two nodes outside the carrier sensing range of each other transmit data to the same node in their overlapping data transmitting area, as shown in *Figure 3.2.* When multi-channel accessing is allowed, the same problem still exists. In addition, the well-known Distributed Coordination Function (DCF) of 802.11 standard is more likely to lose its function under our discussion circumstance, and lead to more severe hidden terminal problem. For example, as shown in the *Figure 4.1*, at first, nodes 3 and 4 are transmitting data on the channel 2, while the transceivers of nodes 1 and 2 are staying on the channel 1. At next step, node 2 wants to send data to node 1 and broadcasts RTS to its neighboring nodes. Node 1 is able to receive this RTS packet, while node 3 can not because node 3 is in the data transmitting status on channel 2. Node 1 replies a CTS packet, and then node 2 begins to send data to node 1. However, since node 3 did not receive node 2's RTS packet, it will not know the exact transmitting behavior of node 2. Therefore, after the data transmission between nodes 3 and 4 finished, node 3 may switch to channel

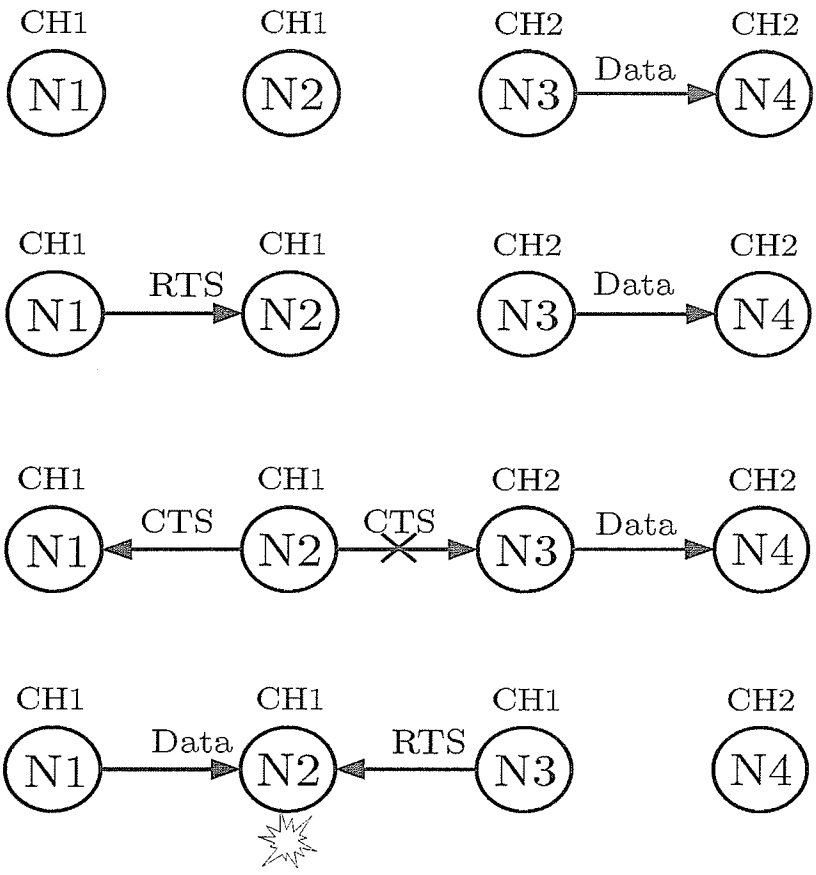


Fig. 4.1: Multiple Channel Hidden Terminal Problem

1, and try to send data to node 2 by initiating RTS packets. In that case, the RTS packets and data packet will collide on node 2. To sum up, the channel disagreement of the transmitter and possible hidden terminal cause the CTS packet not to be transmitted to the possible hidden terminal, which increases the occurrence of hidden terminal problem. It has been shown that multi-channel hidden terminal problem will affect the system throughput more severely [42].

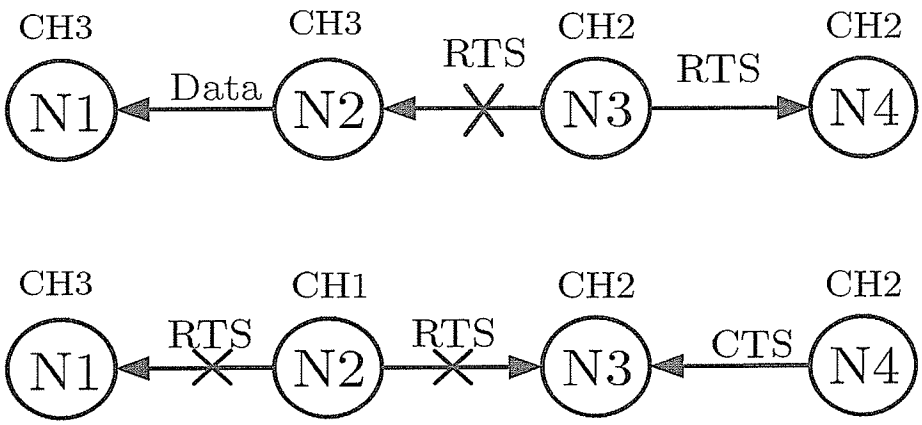


Fig. 4.2: Control Packets Missing Problem

4.2.2 CTS Missing Problem

The CTS missing problem is mainly caused by the unsuccessfully receiving RTS packets. As shown in *Figure 4.2*, node 1 and node 2 are communicating with each other on channel 3. At that time, node 3 wants to initiate transmission with node 2 on channel 2 by broadcasting a RTS packet. However, node 2 can not receive this RTS packet since its transceiver is working on another channel. Therefore, no CTS will be sent back by node 2, and node 3 will continuously broadcast RTS packets until the maximum RTS transmitting limit. Although the CTS missing problem will not result in packet loss, it prolongs the access waiting time of node 3 and decreases the whole network efficiency.

4.2.3 Broadcasting Problem

By broadcasting, data can be delivered to all neighboring nodes (inside the broadcasting covering area) around the transmitter. In the multi-channel environments, the

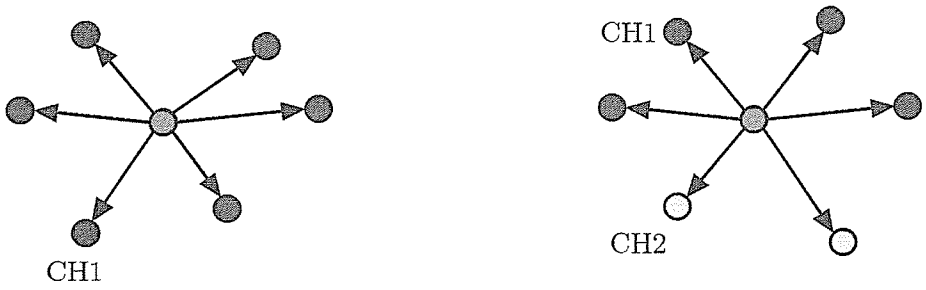


Fig. 4.3: Broadcasting Problem [42]

broadcasting mechanism may lose its function. For example, as shown in *Figure 4.3*, the middlemost node is the initiator of broadcasting, while the other nodes are the receivers. The left sub-graph corresponds to the single channel case, where only one data channel is available and the transceivers of all nodes are always monitoring the channel 1 (CH1). Therefore, when the middlemost node broadcasts data, all of its neighboring nodes can successfully receive this packet if there is no error during the broadcasting process. The right sub-graph demonstrates the multiple channel case, where any node can switch among multi-channels, but, each one can only work on one channel at anytime. For example, the transceivers of bottom two nodes are staying on the channel 2 (CH2), while the other four nodes remain on channel 1 (CH1). Then if the middlemost node broadcasts data on channel 1, only the upper four nodes are able to receive the delivered data. Some applications of VANETs, like the emergency warning message, mostly rely on broadcasting for data transmission. The occurrence of broadcasting problem will greatly reduce the effectiveness of broadcasting, and put a negative influence on those applications.

4.2.4 Channel Switching Delay Problem

In the multi-channel communication environment, switching among channels is necessary and may take considerable time. As indicated by IEEE 802.11b, switching interval from one operating channel frequency to another is 224 μs [42]. No sending or receiving packet is allowed during this process. If the data transmission time is comparable with the channel switching time, the delay brought by the channel switching can not be ignored (short packet length). Thus, the frequent channel switching will bring a great amount of delay and degrade the overall network throughput.

4.3 Related Work

The proposed multi-channel MAC protocols in the literature can be roughly divided into four groups based on their specific characteristics [42][44].

4.3.1 Dedicated Control Channel Approach

Dedicated control channel approach requires every node installing two transceivers. One transceiver is specifically used on the control channel, where the control packets are transmitted and the negotiation about next employed channel is carried out. The other transceiver is utilized to transmit data, and it will switch among multiple data channels as shown by *Figure 4.4*. During the negotiation process about the future channel usage, the transmitter firstly sends a RTS packet on the control channel and one of the available data channels will be selected as the next employed channel. Such channel selection

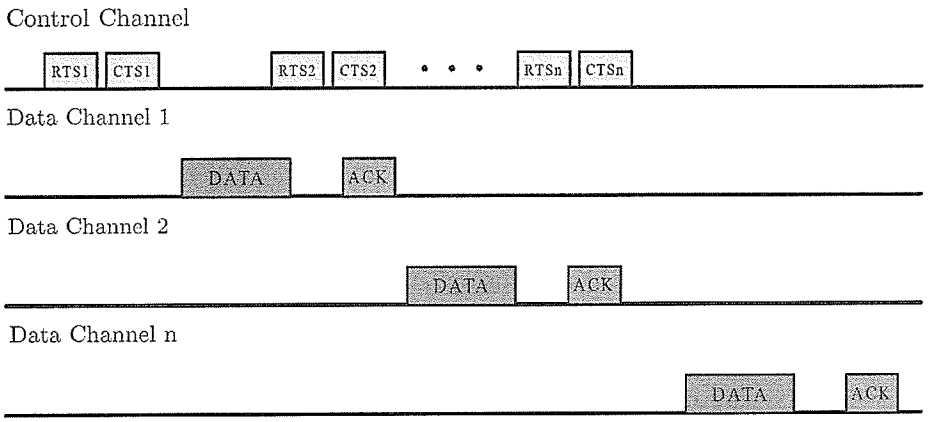


Fig. 4.4: Dedicated Control Channel Approach

can either be based on randomly idle channel selection or maximizing the signal-to-interference ration at the receiver [45]. The desired receiver will reply a CTS packet on the control channel if it is in the idle state. Then the data packet is transmitted on the chosen data channel. (There are some solutions only need one transceiver [46], however, the frequent channel switching delay and possible control packets missing during the data transmission prevent their applications in practice.)

4.3.2 Channel Hopping Approach

By adopting channel hopping approach, all users only need one half-duplex transceiver. All idle transceivers will continuously and simultaneously hop among different data channels. The transmitter and the desired receiver will stop hopping in order to exchange data when they make an agreement for transmission. After the transmission finishes, this pair of nodes in the communication will return back to the common hopping pattern. As shown in *Figure 4.5*, the hopping order is CH1 (data channel 1)-CH2-CH3-CH4,

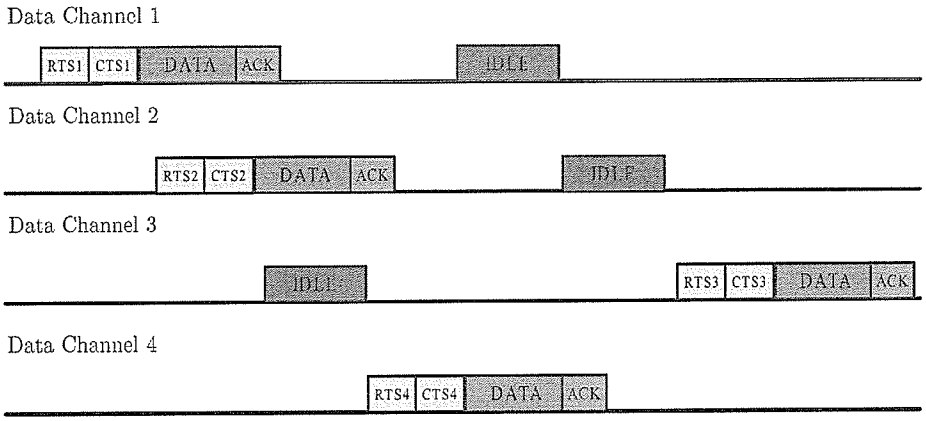


Fig. 4.5: Channel Hopping Approach

i.e., all idle transceivers will firstly stay on the channel 1 for a short period, which is longer than the maximum RTS-CTS handshake time. If no data transmission has been initialized in this period (idle), all nodes will hop to channel 2, channel 3, channel 4 and then hop back to channel 1. When node A wants to communicate with node B on the channel 1, it firstly sends a RTS packet on its current channel, and node B will reply a CTS packet on the same channel if the RTS is correctly received. Then nodes A and B will stay on channel 1 until the data transmission terminates (Channel 1 in *Figure 4.5* indicates this process). When the data transmission is completed, the sender and receiver will re-synchronize to the current common hopping sequence and rejoin it [45]. The other devices will avoid transmitting on occupied channel for a certain period after they detect ongoing transmission.

4.3.3 Time Split Approach

Time division approach divides the whole transmission time period into alternating sequence of control phase and data exchange phase. Each node only needs one transceiver. During the control phase, all transceiver will switch to the control channel in order to negotiate the channel usage of next data phase. In following data exchange phase, the data will be transmitted on the assigned channel. For example, as shown in *Figure 4.6*, channel 1 takes the role of control channel when it is in control phase. Therefore, all nodes will tune to the channel 1 to receive the control packets. One node wants to use channel 3 in the next data exchange phase. Through contention, it transmits a RTS3, which asks for reserving channel 3, on the control channel (channel 1). The desired destination will reply a CTS3 on the same channel when it properly receives RTS3, and then the channel 3 will be reserved. At next data exchange period, the transmitter and the receiver will switch to channel 3, and begin the data transmission process. After the data transmission finishes, the transceiver of transmitter and the receiver will tune to the control channel again.

4.3.4 Multiple Data Transceivers Approach

In the multiple transceivers approach, every node is equipped with more than one transceiver. As what discussed in [48], each node equips two data transceivers (the primary transceiver and secondary transceiver) and one tertiary transceiver. The primary transceiver is used to send data on the predefined primary channel, while the

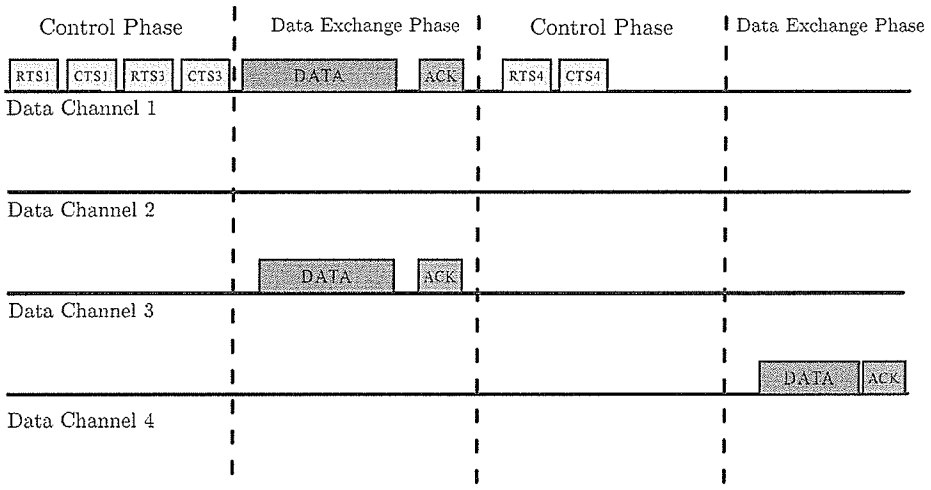


Fig. 4.6: Time Split Approach

secondary transceiver does not work on a fixed channel. The tertiary transceiver is mainly used to transmit or receive broadcast messages. It can also serve as data transmitter. Another example is that shown in [47], where each node is divided into several subnodes, and each subnode has one half duplex transceiver. Every subnode maintains one common control channel and several data channels. During the data transmission, the transmitter will first send a RTS packet on the control channel, and then send out data on the desired data channel with the corresponding subnode when it hears the CTS packet from the destination node on the control channel. With multiple transceivers, the maximum utilization of channel bandwidth has been achieved.

4.3.5 *Comparison Among MMAC Protocols*

By going through these mechanisms, we will find out each approach has its advantages and drawbacks. The dedicated control channel approach does not need time synchro-

nization, therefore it is quite flexible. The disadvantage is the dedicated control channel decreases the spectral efficiency especially when the number of available channels is small. In addition, as all channel usage negotiations are carried out on the control channel, the control channel may become a bottleneck when the network is highly loaded [44]. The channel hopping approach decreases required transceiver number, and allows multiple and parallel rendezvous on different channels. However, it needs perfect time synchronization, which is hard to be realized in realistic traffic environment. Therefore, the channel hopping approach is not feasible in VANETs. The time split approach also needs only one transceiver and requires looser synchronization. However, the bottleneck problem of control channel under high data load may still exist if this protocol is applied. Meanwhile, since every channel reservation in the control phase must follow a data exchange phase, the transmitter may wait a long time before it is allowed to start channel request, which greatly reduces channel utilization efficiency. The multiple transceivers approach requires more transceivers and increases the hardware complexity. In addition, the multi-transceiver multiple access (MTMA) [47] faces the control channel bottleneck problem, while primary channel assignment based MAC (PCAM) [48] needs a complex method to assign the primary channel. Both limit their applications in VANETs. After comparison, the dedicated control channel approach seems the best candidate as the MAC protocol for VANETs. However, in order to compensate its drawbacks, a series of modifications are necessary.

	Dedicated Control Channel	Busy Tone Channel	Data Channel
Covering Range(meter)	10s to 100s	100s-1000s	10s-100s
Bandwidth Consumption	medium	small	large
Information Type	Control Packets	Pulses	Data Packets
Information carrying methods	Encode, decode	Pulses Detection	Encode, decode
Robustness	bad	good	bad

Tab. 4.1: Comparison of dedicated control channel and busy tone channel

4.4 Proposed MMAC Mechanism

Our idea is to use modified BTMA to replace the dedicated control channel approach, i.e., using busy tone channel to realize the function of control channel. By our modification, the major problem in the dedicated control channel approach, i.e. the occupied bandwidth problem and control channel bottleneck problem, can be solved.

As our protocol will be based on the BTMA MAC and partial functions of dedicated control channel have to be realized by busy tone channel, a brief comparison between busy tone channel and dedicated control channel is necessary to evaluate the possibility for the replacement of a dedicated control channel by a simple busy tone channel. Dedicated control channel refers to a separate channel which only transmits control packets, such as RTS packets and CTS packets. Busy tone channel is an individual channel which sends various pulses to help nodes know the current transmission status inside busy tone coverage area, but the pulses themselves do not contain any information. The detailed comparison is summarized in *Table 4.1.*

From the table, it can be seen that the busy tone channel is superior on the large coverage area, small bandwidth cost (the occupied bandwidth is negligible compared

with the data channel) and robustness to complicated propagation environment. But the busy tone channel can only include limited information (pulses), so that the channel assignment may be not so efficient compared with the dedicated control channel. Therefore, some modification on the busy tone channel should be introduced to overcome its common drawbacks.

4.4.1 System Description

The vehicular network under consideration is shown in *Figure 4.7*, where several vehicles are travelling along on a road with three lanes. The arrow represents the driving direction. The uppermost lane is the lane allows the fastest vehicle speed. Correspondingly, the inter-vehicle space is the largest. While the lowest lane only permits the minimum vehicular speed limit but can tolerate smallest inter-vehicle space. Such scenario corresponds to the typical situation on the highway. In this network, each vehicle is equipped with a Global Positioning System (GPS) to identify geographical position information, and a VANET device for transmitting multiple types of data including vehicular driving status, common data, emergency warning data and multimedia data. For explanation purpose, all of these data will be treated equally with single packet data. Note that our work is easily extended to the heterogeneous case.

The proposed MAC protocol requires two types of independent channels, i.e., BT channel and data channel, and a separated transceiver is used for each channel. On the BT channel, a busy tone pulse with a single frequency is transmitted to indicate the

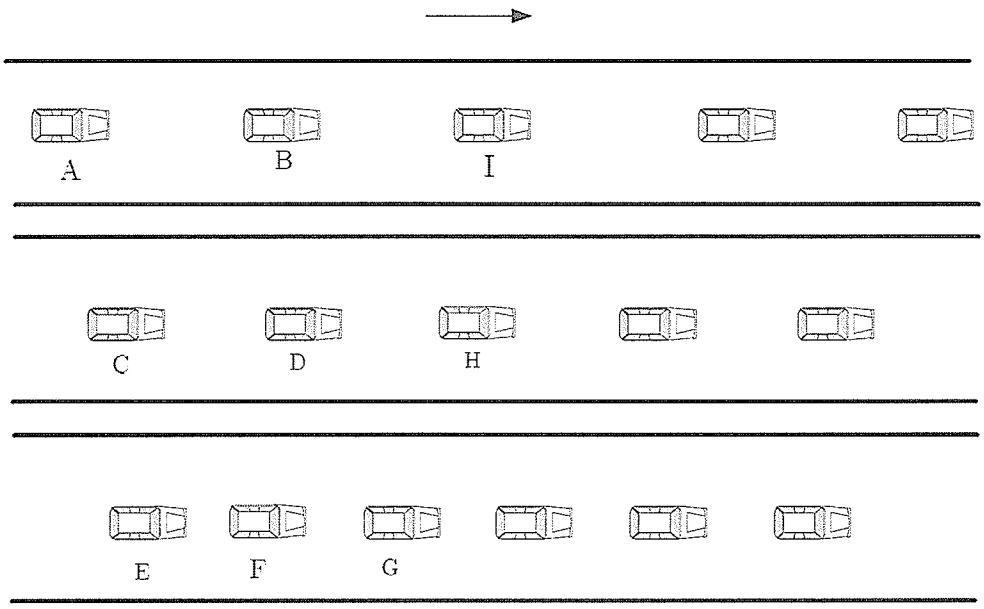


Fig. 4.7: Network Structure

ongoing transmission on the data channel. On the contrast, the data packets will be communicated between the sender and receiver on the data channel. For simplicity, we assume the BT channel can cover all nodes in the network, while the data channel can only cover the neighboring nodes in the valid communication range. Our main focus is how to improve the network efficiency in this scenario.

4.4.2 S-MBTMA

In this section, a new multi-channel MAC protocol, called multiple channels busy tone multiple access with snooping (S-MBTMA), is proposed. By this mechanism, we use busy tone channel to solve the multi-hidden terminal problem and CTS missing problem. In addition, the snooping policy can further boost system performance in term of

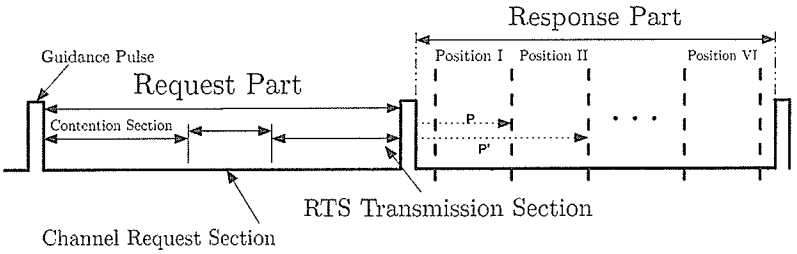


Fig. 4.8: The Structure of On-off Pulse

throughput.

The structure of adopted busy tone pulse in the S-MBTMA is defined in *Figure 4.8*. Each period of on-off pulse is divided into two parts: the request part and the response part. At the beginning of each part, a guidance pulse (the short pulse) will be generated in order to mark different parts and work as a reference to locate the position of following pulses. The generation of guidance pulse is discussed in details as follows. Given the fact that the perfect synchronization is difficult to be realized among vehicles on the road, the generation of guidance pulse individually will not be a good choice. Therefore, an additional source of guidance pulse generation is expected. There are three possible candidates: the first one is the vehicle at the geographical center of concerned motorcade; the second one is the roadside facilities, note that since the synchronization among roadsides facilities are much easier to be achieved, the guidance pulse generated at different roadside facilities can be assumed to be perfectly synchronized; the third one is GPS or cellar communication system. After comparison these three solutions, the second and the third one may be better choices, because the geographical center of con-

cerned motorcade may vary frequently (The guidance pulse needs a stable and reliable generator). The fully covered roadside facility may require large amount of investing and construction time. For the third method, by extracting the built-in guidance pulse from received GPS signal, every vehicle in the discussed motorcade can get reliable and synchronous guidance pulse. However, how to add and abstract the guidance pulse is beyond the discussion of this work, and some proposed solutions to this problem can be found in [49][50]. The request part can be further divided into 3 sections. The first section is the contention section, where the potential transmitters will contend the usage of common data channels based on their observation about channel status during the response part of last on-off pulse period. Each qualified one will initiate a back-off timer, and the first one terminate its back-off process will begin to generate the on-off pulse and get the chance to reserve the channel. The second section is channel requesting period, where short on-off pulse will be generated to indicate the channel number which the sender wants to use. The third section is the RTS transmission period and a corresponding on-pulse period will be transmitted on the busy tone channel to indicate the current RTS transmission. At the response part, a fixed number of positions, which are identical with the number of data channels, will be assigned based on the relative position between the given area on the pulse and the priority pulse (P and P' in the *Figure 4.8*). And the on pulse on one position represents the transmission of CTS packet or data on the corresponding channel.

Figure 4.9 displays the MBTMA (the S-MBTMA without snooping) working pro-

cess. Node A, who wants to transmit data packet to the expected destination B, will first contend with other candidate transmitters for the usage of common channel. If node A wins, it will randomly select an idle channel (Node A selects channel 6 in the *Figure 4.9.* Note that it can get the idle channel information through the observation of last response period) and generate a channel request pulse, which indicates the request channel number, in the busy tone channel. Any node which is not in the transmission state, will switch to the chosen data channel after it hears the channel request pulse. And then the sender begins to transmit its RTS packet. If the desired destination node B is available, it will generate a CTS packet on the same data channel (channel 6), and transmits a busy tone pulse on the busy tone channel on position VI. When node A detects the generated busy tone pulse at position VI before the maximum transmission limit expires, it will know the channel 6 has been reserved, and begin to transmit data packet on that data channel. In the contrast, if node A does not observe the busy tone pulse in this limit, it will recognize that node B is not available right now and wait for future transmission chances. The pulse on the position VI in the response part will last during the whole data transmission time in order to indicate the occupation of channel 6 during this period. After one whole period (including the request period and response period), the transmitters not winning the contention will attend next competition if there exist idle channels.

However, a major problem of the aforementioned method is that only a single transmission is allowed on each data channel at one time, even if the hidden terminal problem

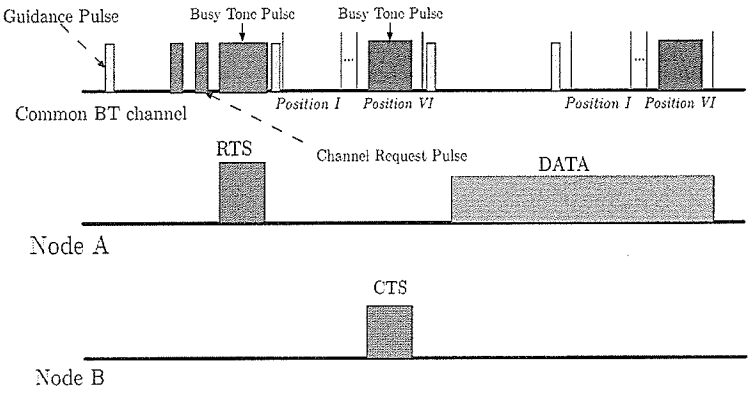


Fig. 4.9: S-MBTMA Process

will not occur. As shown in *Figure 4.10*, all vehicles on the road are inside the same busy tone coverage area and each vehicle owns identical data transmission range. Node S is transmitting data to node D on the assigned channel. At that time, busy tone channel will indicate this channel has been occupied, and any other nodes are not allowed to use this channel in this period. However, in fact, node A or node B should be allowed to initiate data transmission to its neighboring nodes on the same channel since they locate far from the existing transmission and will not bring any data collision.

This defect will greatly reduce the network efficiency. Therefore, in the S-MBTMA, the snooping policy is adopted. The snooping refers to a procedure, which monitors the data channels' usage of neighboring nodes within a fixed period. If data channel power has been detected at a level lower than a threshold (In the real world, the setting of this threshold will be based on a large mount of testing results.), it means that the data channel can be reused to send data even though the busy tone channel tells there is an ongoing transmission in this channel. The reason is that, in this case, busy tone

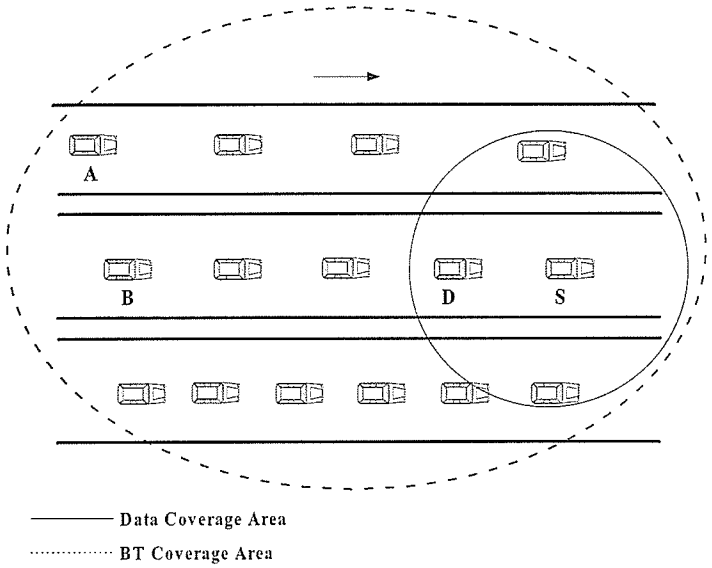


Fig. 4.10: The Problem of Channel Re-utilization in MBTMA

pulse only indicates the transmission between the nodes outside the influence area, and will not affect attempting data transmission. If the data channel power is higher than a threshold, the transmitter will switch to monitor the next data channel. Assuming the transmitter concluded that there is no idle channel by monitoring the busy tone channel and one channel is available through snooping, it will still generate channel request pulse on the busy tone channel, and transmit RTS at the available data channel. Note that, as the synchronization is realized through the BT channel, and the transceiver on the BT channel will not involve in the snooping process, the synchronization among nodes still can be maintained. If the desired destination node is available on that data channel, it will feedback a CTS packet, and generates busy tone pulse on the corresponding position. Notice that since a busy tone pulse exists in the corresponding position, two

busy tone pulses will overlay, and the amplitude for the busy tone pulse at that position will be increased. If the transmitter recognizes the amplitude variance on that position, it means the channel reservation is successful on that channel. And then, the transmitter will begin to transmit data on that channel. Through this way, the data channel can be re-utilized.

Besides the hidden terminal suppressing, the S-MBTMA is capable of warning all vehicles about the occurrence of dangerous situation by a specific pulse on the BT channel, and then all vehicles can switch to the emergency communication state to receive or forward the emergency message. This method has shorter reaction time compared with the control channel protocol like dynamic channel assignment [53].

4.4.3 Simulation Results

In this section, a VANET with 39 nodes in a road with 3 lanes is simulated. The maximum contention window CW_{Total} has been set as 100 μs. The busy tone channel has the capability to cover 500 meter area. There are 6 data channels available in the simulation. The transmission range covers 100 meters and the bit rate of each data channel is 1 Mbits/s. Each packet has identical size equal to 1200 bytes. As Chapter 3, we imply car following model to describe inter-vehicle spacing, and the vehicle length $L = 4m$ and driver reaction time $\beta' = 1.5s$. Additionally, we assume the vehicle speeds on three lanes are 100 km/h, 80 km/h, 60 km/h, respectively. Therefore, the inter vehicle space at each lane is about 50 $meters$, 45 $meters$ and 35 $meters$ accordingly.

The inter space at neighboring lane is 5 *meters*. The arrival of data packets is assumed to follow Poisson distribution, and the total packet arrival rate varies from small value to large value to verify its effect on the system performance.

The average network utilization has been used to describe the system throughput, which is defined by following equation:

$$S = \frac{L_p * N}{T_{sim}} \tag{4.1}$$

where S is average network utilization, L_p is the data packet length, N is the transmitted data packet number, and T_{sim} is the maximum simulation time. In the simulation, the MBTMA (without snooping) and S-MBTMA are compared in terms of average network utilization. In *Figure 4.11*, the x-axis represents the total packet arrival rate and y-axis denotes the average network utilization. From the figure, it can be observed that the average network utilization increases with the total packet arrival rate, but it will tend to saturation after the arrival rate is larger than 900 (for the MBTMA). The observation results from three aspects. First, the bandwidth of each channel is limited, so a large data arrival rate may exceed the total channel capacity. The second is that each transmission cycle involves a request part and a response part. The potential users have to wait the next request part to begin channel reservation, which will increase the waiting time. Third, with the increment of arrival rate, the probability that the desired destination node is busy in other transmission, which is defined by "blind transmission",

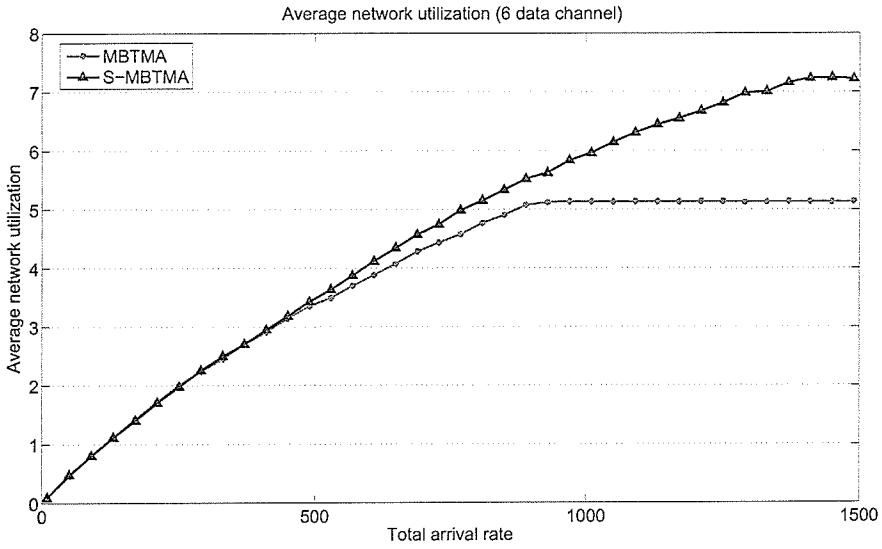


Fig. 4.11: Average Network Utilization (6 data channels)

also increases. Since the busy tone channel can not carry the detailed transmission information in the whole network like the dedicated control channel, the frequent "blind transmission" will become a bottleneck for system throughput. By adding snooping policy, S-MBTMA can improve the average network utilization around 20% when the total data arrival rate is 1010, and such increment reaches over 40% when the total data arrival rate is larger than 1400. This can be explained by the fact that the S-MBTMA successfully reuses the data channels without bringing interference to other transmissions. In addition, for the S-BTMA, there also exists a maximum utilization threshold. However, such threshold is much higher than that in the MBTMA.

Compared with the single data channel case as shown in *Figure 4.12*, the implementation of multiple data channels does enhance the system performance in terms of

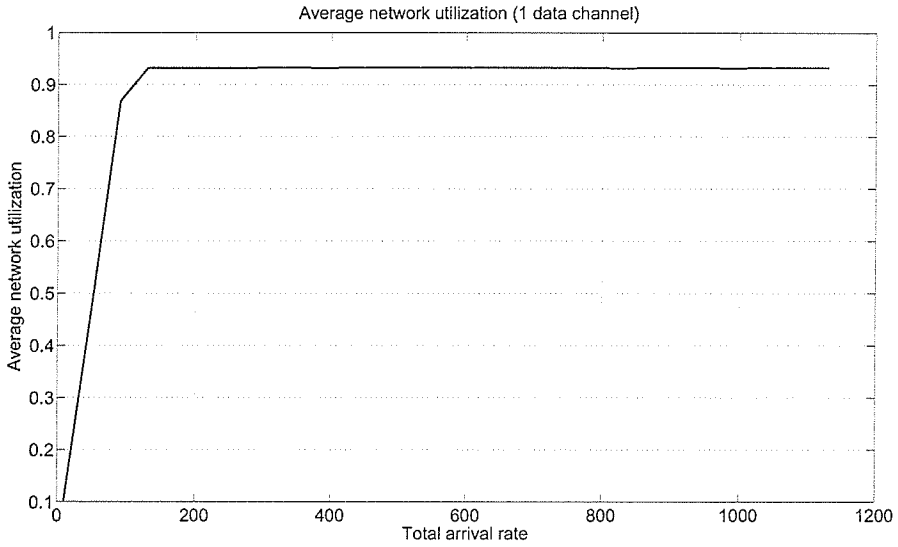


Fig. 4.12: Average Network Utilization (1 data channel)

average network utilization. However, notice that the maximum average network utilization for MBTMA (for 1 data channel case) is 0.93, which is larger than the maximum average network utilization for MBTMA (6 data channels) equal to 0.85. This is because the increase of data channels extends the total busy tone pulse period and increases the waiting time of potential transmitters.

4.4.4 Conclusions and Comments

In this chapter, a new multi-channel MAC protocol, called multiple channels busy tone multiple access with snooping (S-MBTMA), is proposed. By simulation, we demonstrate that the S-MBTMA is able to greatly improve the system throughput than the traditional busy tone mechanism.

5. ROUTING IN VANETS

Besides MAC protocols, designing routing mechanisms is another important topic in VANETs. Similarly, although VANETs can be classified as MANETs, their properties predestine that a number of specifics need to be modified according to the deployment environment. One of main challenges of routing protocol design in VANETs is how to handle the limited time-variant topology. Technically, every vehicle could carry out lane changing, acceleration, deceleration or U-turning randomly. However, such vehicular behaviors are regulated by traffic regulations, traffic density, weather, terrain, etc. The routing protocol in VANETs will experience more frequent and fast topology change comparing with the counterpart in the MANETs. Moreover, in VANETs, more strict demand should be put on the propagation delay, especially for the emergency information. It has been shown that traditional ad hoc routing protocols, such as dynamic source routing (DSR) [51] and ad hoc on-demand distance vector routing (AODV) [52], could hardly achieve this aim. Meanwhile, synchronization is hard to realize in VANETs due to the shortage of roadside units (RSUs) and the intolerance to high synchronization cost. Therefore, new routing protocols for VANETs, which are entirely distributed with low delay, are waiting to be proposed.

5.1 Challenge and Design Factors for Routing in VANETs

The realistic traffic is quite complex and a series of elements have to be integrated in designing an efficient and practical routing protocol. In this section, some critical design factors are summized and evaluated about their effects on the routing mechanism design.

5.1.1 Market Penetration

As we introduced in Chapter 2, the development of VANETs will experience a long process. Therefore, it is straightforward to expect that most vehicles do not equip VANET devices at the early stage. Market penetration has been used to describe the proportion of VANET and GPS devices adopted on the vehicles and the percentage of road coverage by the RSU. Evidently, larger market penetration provides more chance for valid connection among vehicles and has better system performance. Unfortunately, the designers of routing protocol in VANETs have to face the low level market penetration at current stage. As a result, a practical routing protocol should be applied even at low market penetration environment.

5.1.2 Traffic Density

From network point of view, traffic density will affect the valid connection number and the quantity of data which need to be processed by VANETs. The traffic density on the road varies during different time periods. Berkeley Highway Laboratory (BHL)'s empirical data [54] demonstrated that the vehicle number reaches its peak during rush

hour (7 am-9 am and 3 pm - 5 pm) while drops to the floor during midnight (1 am -3 am) on the observed highway. Considering a multi-hop transmission case, if the valid transmission range is fixed, the information source node (node S in *Figure 5.1*) can easily find a neighboring vehicle as the transmission relay node, and therefore the probability of maintaining a full connection between source node and destination node (node D in *Figure 5.1*) is much higher under the dense traffic case. Notice that the data from the vehicle to the west direction may not be useful to the vehicle to the east direction but only bring interference under this situation. Thus, it is necessary to suppress such redundant packets. When VANETs face sparse traffic, which corresponds to the case shown in *Figure 5.2*, the information source node can not find a relay vehicle at the same driving direction within its valid communication range immediately. As a consequence, the transmission from source to destination will experience longer delay, and the current connection may suffer from frequent link disconnection. Unfortunately, till now, little work has been done on the routing in this scenario. Besides the aforementioned time factor, the traffic density will also be affected by the terrain, weather, etc. Meanwhile, the traffic volume is not evenly distributed on the same road. Therefore, an efficient and practical VANETs routing protocol should be self-adaptive according to the neighboring traffic density.

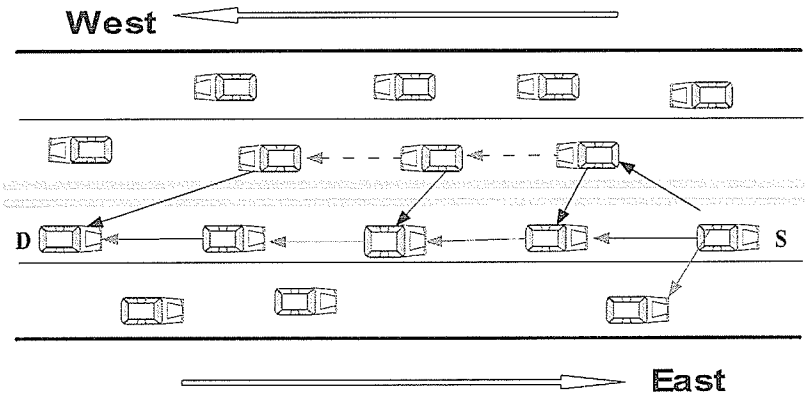


Fig. 5.1: The Effect of Traffic Density (Dense Traffic)

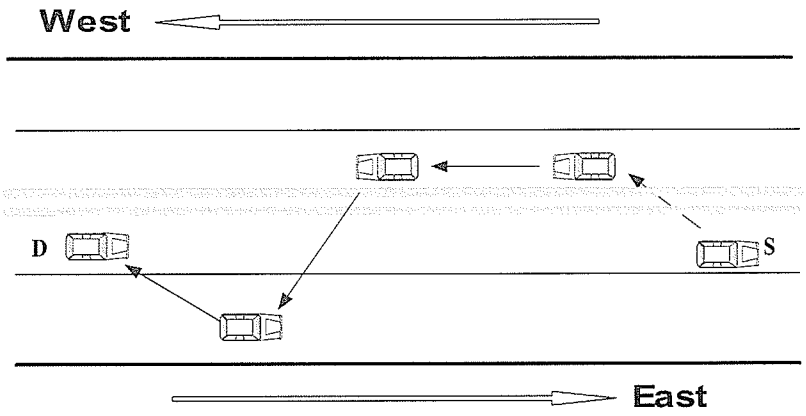


Fig. 5.2: The Effect of Traffic Density (Sparse Traffic)

5.1.3 Road Constraint and Traffic Rules

The driving vehicles have to follow the road direction (road constraint) and obey the traffic rules in most cases, which result in some unique properties of VANETs. On the one hand, differing from the random moving nodes in the MANETs, nodes in VANETs will move along the road. Turing left, turning right and U-turn are three major vehicle movements. On the other hand, traffic rules will put a limitation on vehicle movement.

The traffic rules have two levels. One kind of traffic rule is general, e.g., the leftmost lane owns the highest speed on the multiple lane roads, and the overtaking always happens at the leftmost lane. The other type of traffic rule is restricted to some specific areas. For example, there are some traffic signs indicating the speed limitation, the driving direction limitation, stop sign and so on. These signs are effective only at the specified area. The road and traffic information are valuable because it is possible to predict the future vehicle movement at a certain degree when the starting point and the destination point are known. In the real world, the detailed road constraint and traffic rules information will be marked on detailed electrical map, which can be downloaded from specific RSUs.

5.1.4 Driver Behaviors

The effects of driver behaviors will be reflected in two aspects. One aspect is the driver will choose a destination before or during his/her driving procedure. Therefore, the general driving path is known. On the other hand, every driver will react variously to the real time road situation and unexpected events (traffic accident, road obstacle, neighboring vehicles' dangerous behaviors). This variance depends on the age, experience, fatigue, etc. Therefore, it is hardly to precisely describe the behaviors of every driver. Fortunately, some reliable surveys like US Department of Labor Survey can provide some statistical data about driver behaviors including commuting time, lunch time, traveling distance and so on. As a result, it is possible to predict behavior of the large

group. Meanwhile, the well-known car following model [40] and lane changing models [55] [56] are beneficial to describe the car following and lane changing (two of the most frequent behaviors on the road) in a mathematical way, which can be considered as a good approximation of drivers' reaction to real-time road situation. The integration of driver behaviors can further boost the maneuverability of VANETs routing protocols on the road.

5.1.5 Propagation Environment

VANETs have to work in a complicated propagation environment, where neighboring vehicles, roadside buildings and terrain can play a direct effect on the fading and shadowing of transmitted signal and bring kinds of noise. In addition, considering the fast moving speed, the Doppler Effect for the signal propagation between vehicles will inevitably become another important issue. Although the two ways ground model and Nakagami model have been chosen to describe fading and shadowing of the signal propagation in VANETs, there is little evidence that those two models can match the real situation on the road. Thus, some improvements have to be proposed. As in [23], a 3-D model, with the integration of measured roadside buildings' parameter from the electrical map, has been created to evaluate their effect on the signal propagation in the realistic environment. But, till now, there is still no widely accepted model which can perfectly describe the signal propagation between vehicles on the road. Such model remains an open topic in VANETs.

5.1.6 Service Types

The comprehensive VANETs have to deal with multiform data with distinct QoS requirements. For instance, the emergency warning data (the packet size may be smaller comparing with the common data), which notifies the potential dangerous situation on the road, demands high successful receiving ratio and low transmission delay. For common data, some of them, e.g., the neighboring driving data, can tolerate longer delay (The neighboring vehicles will not change its position in short time during its travelling. Therefore, this type of data can accept longer delay), while others, like the on-board video or Voice over Internet Protocol (VoIP), needs to be transmitted in high data rate and low delay. Some data including private information and account information deserve to be treated with high level security. A reliable and practical VANETs routing protocol should handle all these data to provide individual requirements. Additionally, the importance and requirement of some data may be variable according to different receivers. A good example in this case is the emergency warning data. For the vehicle around the initial source of emergency warning data, such emergency warning data should be broadcasted with reliability and low-delay, and all other types of data should be interrupted to promise the priority of emergency warning data. On the contrary, for the vehicle outside the affected area (several kilometers away), the received emergency warning data should be recognized as low priority driving assistance data, since it loses the function for collision avoidance. Notice that the emergency warning data, when they are used as driving assistance data, is useful to predict the traffic density upswing

around that area. And then the driver may change the driving direction on the basis of this information. This variety of information importance would bring another challenge for routing design in VANETs.

5.1.7 Power Limit

Usually, the power limit of the wireless devices is not regarded as a bottleneck in VANETs. Consider the fact that VANETs devices will be powered by the vehicle engine, and the power supply for them can always be considered as sufficient. Therefore, traditional power-saving mechanisms like the sleep-awake intervals and battery life report will not reflect their advantages in VANETs [57].

5.2 An Overview of Existent Routing Protocols

Based on the hardware complexity and requirements, and the difficulty of realization, existing routing protocols can be roughly divided into four categories: vehicle oriented routing protocols, RSU oriented routing protocols, cluster routing protocols and security oriented routing protocols. In the following part, we will compare their characteris and drawbacks by examples.

5.2.1 Vehicle Oriented Routing Protocols

Vehicle oriented routing protocols refer to those routing protocols which only require vehicular on-board device, and will utilize the characteristics of vehicle behaviors to

predict the future movement in order to provide a more stable routing path.

Differing from the nodes in the MANETs, the movement of a node (vehicle) in VANETs will be affected by its neighboring nodes. Therefore, a well known position-based routing protocol, greedy perimeter stateless routing (GPSR), has been chosen as the premise of many vehicle oriented routing mechanisms [5]. Under GPSR, every node is assumed to know the exact physical locations of its neighbors and the destination. During the routing path setting procedure, the packets are forwarded to the nodes that are closest to the destination by greedy forwarding method. However, the major problem of GPSR is the selected node closest to the destination may not be the best data relay node.

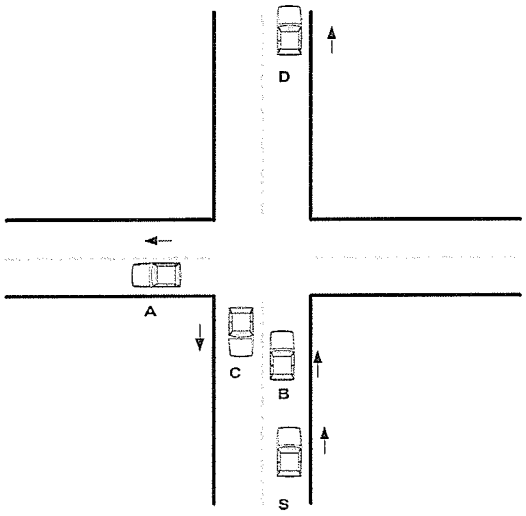


Fig. 5.3: The Node Selection Problem of GPSR

As shown by *Figure 5.3*, let the node S be the source node of the data and the node D be the desired destination. We assumed that the node D is out of the direct

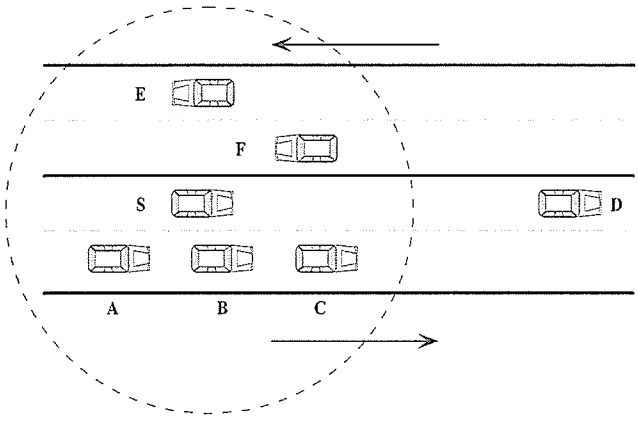


Fig. 5.4: The Node Selection Problem

transmission range of node S, therefore node S has to find an intermediate node to relay the data. According to the rule of GPSR, node S will choose node A or node C as the relay node since they have closer distance to node D than node B. If node A is selected, it may move out the transmission range of S soon, so that node S has to start another searching for relay node. If node C is selected, it will move to the node S, and may not help relay the data. In both situations, the transmission between node S and node D will be prolonged due to unsuitable relay node selection. To solve this problem, in [58], the author proposed a new amendment, called movement-based routing algorithm (MORA), which takes the vehicle driving direction into consideration. By integrating the information of the present location, neighboring nodes and the intended destination, the MORA chooses the intermediate node with largest projection distance on routing path, and the information packets will be transmitted by flood-based propagation strategy in a decentralized manner. It avoids the aforementioned unsuitable relay node selection problem, however, the MORA also may not choose the best candidate relay node under

some circumstances. For example, as shown in *Figure 5.4*, node S wants to communicate with node D. Inside its valid transmission range (the dash circle), there are five nodes can be selected as the relay node. With MORP, node E and node F will not be selected since they are on the opposite driving direction, while node C will be selected since it is on the same driving direction with the source node according to the MORA's rule. Notice that node C is close to the transmission limit of source node, thereby the link through node C may break off after a short period, and then source node will have to initiate another relay node searching and routing path setting procedure. Usually, it will cost too much time. As a result, choosing nodes A or B would be a better choice than node C. Hamid et al. [59] have solved this problem by their proposed movement prediction-based routing (MOPR) mechanism, which integrates the physical location, vehicle driving direction, vehicle speed in the selection of an effective routing track. With the assumption that two nodes within the maximum transmission range could keep valid communication, the link stability is estimated according to the time interval between current time and the time that the possible relay node reaches the maximum transmission range. Such estimation can be easily obtained with the relay node location and speed information. (Here, it assumes the vehicle speed will not vary significantly, which is close to the real situation.) Then, the node owns highest link stability will be chosen as the next relay node. Tarik et al. further improve the aforementioned works in their receive on most stable group path (ROMSGP) scheme [60]. First, the ROMSGP groups the nodes according to their speed vector, which includes both the absolute speed and driving direction. If two vehicles

belong to two different groups, the connection between them are considered as unstable and a penalties metric is introduced to decrease the probability of selection the unstable routing path, which is helpful to decrease the link breakage probability. Meanwhile, the link expiration time will be calculated at each node and transmitted in the routing reply packet. A new path selection would be initiated before the link breakage, so that the current transmission node can keep transmission without waiting for another routing path setting process. Those two methods greatly enhance the reliability of routing path.

All of these proposed routing protocols focus on handling the VANETs communication in high dense traffic environment. However, most of them neglect the fact that at some time, the number of vehicles on the road may be quite small, i.e. sparse traffic case, and the valid connection between nodes may not exist, as we have discussed in the Chapter 5.1.2. Under this situation, the using of 'store and forward' policy is necessary. As shown in the *Figure 5.5*, the source node S wants to transmit some data packets to node D, which is outside of its transmission range (The dash circle), and, there is no nearby vehicle which can be used as a relay node on the same driving direction. However, vehicle A is on the opposite lanes. Therefore, node S can first send data to node A, and node A will store this data in its buffer. Then, when node A approaches destination node D, node A will deliver the data to the desired destination. Actually, this 'store and forward' mechanism has been demonstrated to be able to improve the system performance for emergency warning in the sparse traffic situation [4]. However, there are two limitations for the 'store and forward' policy. On one hand, the valid connection

time between vehicles on the opposite direction is transitory, and can only support limit data transmission. On the other hand, the source node S needs to make sure the pass-by vehicle will approach the destination node in the future. However, without RSU and other neighboring vehicles, the location and driving data of the destination node can hardly be obtained. In that case, the forwarded data from the source node will become redundant information when it can not be transmitted to the desired destination.

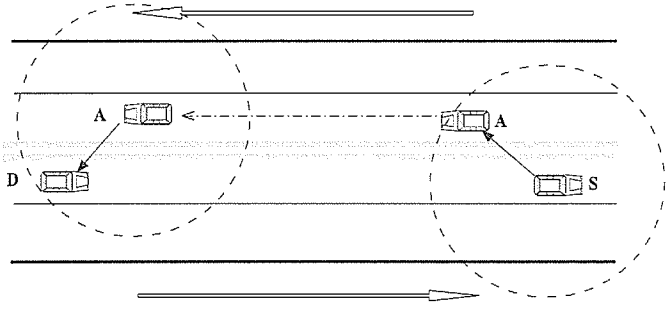


Fig. 5.5: 'Store and forward' Policy

5.2.2 RSU Oriented Routing

This kind of routing protocols makes use of the geography property of traffic system, and usually requires additional RSU as external coordinate or information relay unit [64], [65]. The integration of RSU has two benefits. On one hand, since inter-connections between RSUs are much reliable than those between vehicles to RSUs, RSU has the ability to obtain the information in far distance with low delay by RSUs' communication. The RSU oriented routing protocols own the potential to support a variety of data packets in VANETs, including basic traffic data, common data (document or E-mail)

and multimedia data. Moreover, the fixed RSU has larger data processing capability and will not experience location variance. Therefore, the RSU can be considered as a reliable local area communication center, which is helpful to improve the limitation of vehicle oriented routing. In summary, by additional RSUs, the RSU oriented routing could overcome some challenges brought by the realistic road environment.

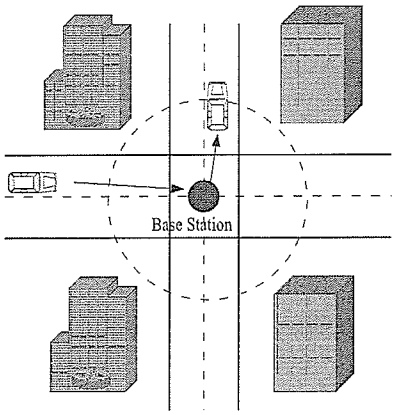


Fig. 5.6: The Data Relay at Intersection

A typical but complex traffic environment in the urban area happens at the intersection, where the information packets may not be able to deliver to the vertical direction through direct communication due to the obstruction of skyscraper or other city facilities. The vehicle oriented routing can be used to handle this problem [61] [62]. The vehicle closest to the intersection center will be chosen as the data relay center. When it received the information packets, it will rebroadcast those messages to the other three directions during its driving procedure. And the vehicles on the vertical direction will get the data through the rebroadcasting. However, at some major intersections in the metropolis, the vehicle oriented routing may lose effect. There are two reasons. First,

at these major intersections, the vehicle closest to the center of intersection may not be unique, and the selection of relay node at the intersection will cost lots of time. Moreover, the selected relay node can only deal with limited rebroadcasting request before it leave the intersection, and another searching for the relay node has to be initialized. The second is the on-board VANET device has limited channel bandwidth. If all vehicles around the center of intersection rely on one vehicle to rebroadcast their message, a possible congestion may happen at the chosen relay vehicle. Therefore, a better solution for intersection relay is to set up a fixed relay node (RSU) [63] at or near the center of the major intersection as shown in *Figure 5.6*. The RSU can broadcast received data repeatedly and cover all vehicles near the intersection. Then, the vehicles on the vertical direction could acquire the packets from the center facility and further forward data on its own direction. Since the RSU is fixed, the time for searching the relay node will be greatly reduced. Meanwhile, the RSU ordinarily has much larger system capacity, so that the congestion at relay node can be decreased. Moreover, by applying directional antenna, the received data can only be broadcasted to the desired direction, which can further reduce the redundant rebroadcasting.

Another example for the RSU oriented routing is utilizing the diversity between vehicular dense roads and other roads with smaller traffic load. Generally, vehicular dense roads have a large amount of driving vehicles and the valid connections between vehicles always exist. Therefore, the high vehicular density area can support large data stream in VANETs. While in small dense roads, the driving vehicles are much lesser, and

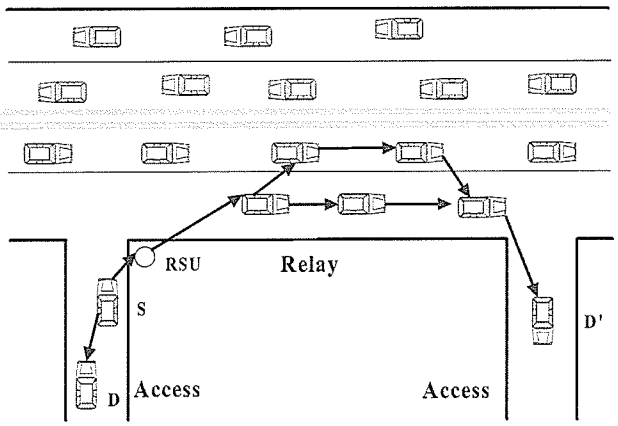


Fig. 5.7: Two Phase Routing Protocol

frequent connection disrupt may happen. The two phase routing protocol (TOPO) make use of this road property [66]. By integrating the electrical map information, the city roads have been divided into high vehicular density road and low vehicular density road, which are named by the *overlay* phase and *access* phase, respectively. If the source and the destination both located at the same *access* area (node S and node D in the *Figure 5.7*), single stage small scale routing protocols, such as DSR, AODV, will be adopted after the evaluation of relevant nodes position. If the source and destination are driving at different *access* areas, then the packet need to be forwarded via an *overlay* area, and a three stage method would be carried out. At first, the source or the relay node in the initial access area would find the nearest intersection of *access* and *overlay* by using the greedy perimeter stateless routing (GPSR) protocol. The data waiting to be transmitted is forwarded to the vehicle in the *overlay* phase directly or the roadside RSU, which will further broadcast the data to the vehicle on the *overlay* area. After the data has been

loaded on the overlay area, a greedy path selection will be executed, so that the data can be transmitted with a high speed. When the data arrives at the intersection closest to the destination, a route discovery will be carried out to establish a valid path between *overlay* and that intersection. In the third step, data will be transmitted from that intersection to the desired destination. If the intersection of *access* and *overlay* are not immediately available, some other strategies like 'carry and forward' strategies will be adopted to keep the data until the relay node finds the valid intersection. With similar policy, another RSU oriented routing can be utilized, which requires that RSUs can cover the whole urban area. If the source node and the destination node are in the same RSU coverage area, they can initiate direct communication. When the source node and the destination node are driving at different RSU coverage area, the source node will firstly transmit data to a nearby RSU. After that, data will be forwarded to the nearest RSU to the destination node through the wire line connection between RSUs. Finally, the destination node obtains the data from its nearest RSU.

5.2.3 Cluster Based Routing

Vehicles on the road naturally exhibit a series of groups, i.e., several vehicles driving with similar speed and identical driving direction, as shown in *Figure 5.8*. Therefore, it is an intuitive idea to assemble close vehicles in an area as a virtual cluster, and the inter-vehicle communication is carried out based on this cluster structure [67]. The cluster based routing follows this idea. Each cluster elects a header as the representative of

whole cluster members, and it will coordinate the intra-cluster communication, i.e., the communication inside specific cluster. If the source and destination locate in the same cluster, the source can directly transmit data to the destination by greedy broadcasting or header control limiting broadcasting. In case the source and destination are in different clusters, the source only needs to find the header of cluster including the desired node which will further forward those messages to the destination. A simple example is shown in *Figure 5.8.*

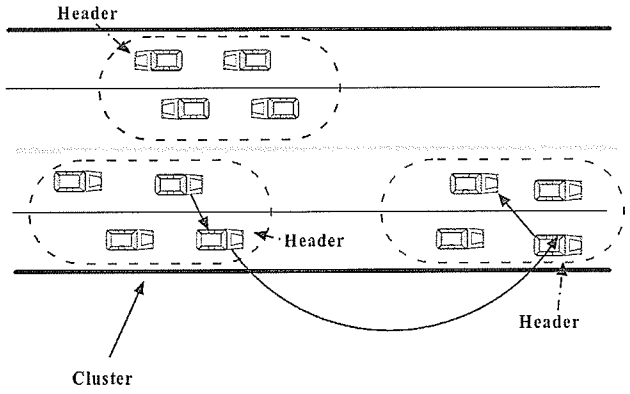


Fig. 5.8: Cluster-Based Routing Protocol

The cluster based routing protocols can either fully on-board device based, or on-board device and RSU based. The on-board device based cluster routing usually chooses the vehicle at the center of the cluster as the cluster header, while, for the on-board device and RSU based cluster routing, the RSU takes the role of cluster header or it assigns a neighboring vehicle as the cluster header.

The advantage of cluster based routing is that it can reduce the redundant message transmission. There are two typical examples. The first one is the intra-cluster com-

munication, i.e., the source and the destination nodes locate at different clusters. With cluster-based routing, the routing path will be settled as 'Source-Source Neighboring Cluster header-Destination Neighboring Cluster header-Destination', which has much less transmission comparing with traditional routing protocols. Therefore, the possible redundant messages are reduced. The other situation is that the generated data can be limited into the concerned area (The generated data will be useful in concerned area, but, out of the concern area, it can be regarded as redundant data.). For instance, when the vehicle encounters emergency situation, the generated warning message will be broadcasted only inside the concerned area (one or few cluster range) based on cluster serial number information. Meanwhile, by establishing clusters, the intra-cluster communication can be carried on the ordinance of cluster header, and the unwanted data from other clusters can be blocked by the header. For some special users, such as motorcade, the cluster based routing will be beneficial. However, the biggest challenge for the cluster-based routing is to find an efficient and practical method to select a cluster header and maintain cluster structure. If the cluster header is a vehicle, it may change its location very soon. It will bring another challenging topic: how to keep the cluster stable and find another cluster header rapidly after current one is not suitable to act as the header. If the cluster header is RSU, the expected huge hardware investment is the top issue. Another choice is the RSU assigns a cluster header. Under this condition, the optimal RSU setting position and effective assigning policy will become major problems.

Besides the maintaining cluster structure, synchronization is required in some pro-

posed protocols and will be another big challenge. The reason is that the synchronization will lead to additional delay, which may be unacceptable to VANET communication. The cluster-based routing protocol on the basis of purely on-board device will suffer this problem more severely.

5.2.4 Security and Reliability Oriented Routing

Except the basic safe-related applications, all other applications in VANETs should consider the security of data, especially for those data involving personal information, bank account information or other sensitive information. Unfortunately, VANETs are more vulnerable than other wireless networks, such as cellar networks. On one hand, the malfunction of any node could possibly be expanded to neighboring nodes, and leads a chain serious error. On the other hand, the malicious nodes can attack VANETs and jeopardize the network security. There are three main types of attacking threats, which have been described in *Figure 5.9*, *Figure 5.10* and *Figure 5.11*. The first is data intercepting as shown in *Figure 5.9*. According to [68] and [69], the malicious node A can transmit fake location formation. Because the source node relies on the received location information to locate its neighboring vehicles, it may consider that node A is driving at place A'. Since with the GPSR, the data will be forwarded to the node which is closest to the destination, the source node will always transmit data to node A' (In fact, it transmits data to node A) instead of node B. However, if node A does not relay any received information to other nodes, node A successfully intercepts the

transmitted data. The second case is data modification. For example, in the *Figure 5.10*, the malicious node A is the only node successfully obtains data from source node (To achieve this aim, node A can transmit a cheating data to mislead the source node into believing no other node exist in its transmission range). Node A can modify the original data or inject malicious packets, and then relay the infected data to other nodes. In this case, all other nodes can only receive the modified data. The third case is called data jamming. As shown by the *Figure 5.11*, source node S is transmitting data to node B. The malicious node A sends jamming signal to node B at the same time in order to suspend current transmission and exhaust the limit bandwidth of VANETs.

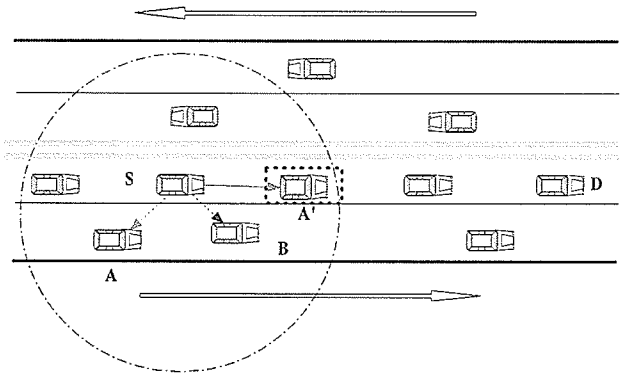


Fig. 5.9: Attack in VANETs (Scenario I)

To protect the data in VANETs, the security and reliability oriented routing protocols have been proposed. A good example is that proposed in [70]. Based on existing routing protocols, the author introduces the cryptograph and plausibility check in his work. Three safe guarding policies have applied in the proposed mechanism, including plausibility checks, digital signature and public/private key pairs. The plausibility

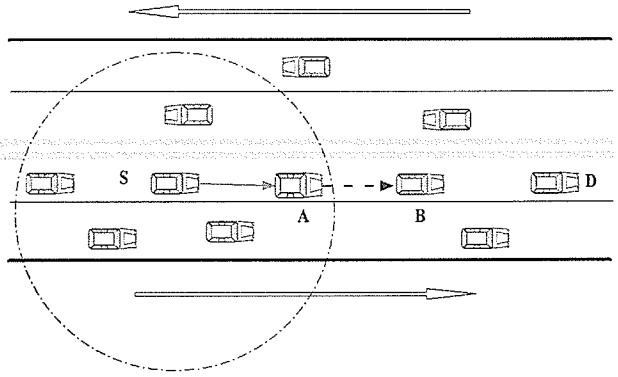


Fig. 5.10: Attack in VANETs (Scenario II)

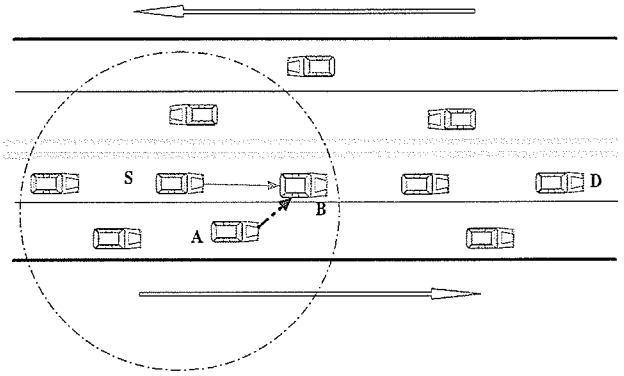


Fig. 5.11: Attack in VANETs (Scenario III)

checks make sure the timestamp and the acceptance range of coming packets are reasonable, so that false data generated by the malicious nodes can be recognized. Meanwhile, the plausibility check compares the received location data and on-board sensor detection result, and the aforementioned data intercepting by fake node position information can be prevented. The digital signature and public/private key can protect the authenticity and integrity of original data on both hop-to-hop level and end-to-end level. After reading the digital signature, the modified section in the original data node can be

distinguished and corrected. With public/private key policy, the added data by the malicious node will lose their effect, since they will be dropped on the process of decoding. In addition, the node type based rate-limiting policy has been adopted to suspend the over redundancy, occasionally or maliciously, which is useful to stop the data jamming attack.

The RSU will be useful in the security and reliability oriented routing protocols. It can check the integrity, confidentiality and correctness of the transmitting data, while detect the malicious nodes in the neighboring area. With a larger data processing capability, RSU is more suitable to take this role. Meanwhile, some high-security demand applications, such as the bank account verification and personal information confirmation, will require the attendance of RSU, because none of vehicular node can be trusted to finish those missions.

It worth mentioning that any security oriented routing protocol inevitably brings some new delays due to the additional packet header or verification operation. How to keep the source-to-destination delay within acceptable latency constraint under security and reliability oriented routing is another design challenge.

5.3 Routing Protocols Design

The routing protocols design is a complex mission. A block diagram depicting the whole process is shown in *Figure 5.12.*

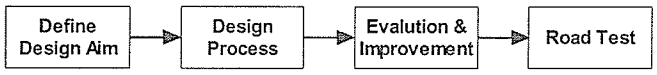


Fig. 5.12: Routing Protocol Design in VANETs

5.3.1 Define Design Aim

Different applications have their own stress on requirements. The emergency warning message may not focus on the privacy but authenticity and reliability are the main concern. Common data, especially the data with regard to personal bank account or credit information should keep authenticity, integrity, confidentiality and correctness in mind. Meanwhile, some designing works may specially focus on the typical vehicles (trucks, buses, for example.) or on the typical areas (e.g. urban area, highway). Therefore, to begin the routing design, it is necessary to define the design aim, i.e., the main applications scenario of the routing protocol. A clearly defined design aim has two benefits. 1) It is useful to estimate the hardware requirement for the given routing protocols; and 2) some properties on the discussed scenario can be utilized to boost the performance of routing protocols. For example, since the buses have fixed driving path and fixed timetable, the routing protocols specially designed for buses can make use of these features.

5.3.2 Design Process

With a clear design aim, the design process can actually be started. It usually begins with choosing an archetype. The basic archetypes are the vehicle oriented routing and

RSU oriented routing. If the vehicle oriented routing has been chosen as the archetypes, the designer should find an efficient way to predict future vehicular movement by considering the vehicular velocity, direction, destination and so on. If the RSU oriented routing has been picked, the major problem is finding the right location for the RSU. RSU has large system capacity and more reliable connection with other RSU, but the RSU is much more expensive than on board equipment. So it is necessary to find suitable places to maximize the benefits of RSUs. Typically, the intersection, urban area and some high vehicle density area need more RSUs. If the designer chooses cluster based routing as the archetype, the first problem is establishing a suitable mechanism to maintain a stable cluster structure. Otherwise the unstable cluster will lead to frequent disconnection and excessive cluster building cost. The security and reliability oriented routing can base on either one of aforementioned archetypes, but the digital signature and public/private key will be applied in the packet to boost the security level of data. Meanwhile, the security service center to provide the authentication of user ID, or some other special RSUs to monitor and record the possible malicious nodes may be necessary.

5.3.3 Evaluation and Improvement

After the draft of routing protocols has been finished, it is necessary to evaluate them under a suitable and close to realistic mobility model. By evaluation under mobility models, it is possible that some neglected details will be discovered. Meanwhile, the op-

timization and improvement work should be carried out at the same time. For example, the RSU oriented routing protocols needs the setting of RSUs. Therefore, the location for the RSU should be carefully analyzed, since it is possible to find a solution with similar capacity but less cost. Another example is that most of proposed routing protocols are more suitable for the high vehicle density case, while few focus on the sparse vehicle density situation. Therefore, the proposed mechanisms may not work properly under sparse traffic road. For the newly designed protocols, both the dense and sparse traffic cases should be considered, and the designed protocols should adapt under both two cases, which will further boost the utility value of proposed protocols.

5.3.4 Road Test

Different from other types of routing protocols, A vehicular routing protocols will be applied in the complicated traffic environment. Unfortunately, it is hard to build a model which can reconstruct the complexity of realistic propagation environment, the outburst of sudden events (including the traffic accidents, the breakdown of neighboring vehicles, the misbehavior of pedestrians, etc.), variety of vehicle density and moving direction. Therefore, if the cost is allowed, a road test for designed routing protocols should be carried out to testify if the reliability and the efficiency of routing protocols can fulfill the design aim.

6. CONCLUSION AND FUTURE WORK

6.1 Conclusion and Comments

In this thesis, the history and the development of VANETs have been introduced. Then the potential applications of VANETs have been described in details. As one of the major differences between MANETs and VANETs, different types of mobility models have been compared in terms of structure, realization complexity and calculation costs. To boost the research in this area, some popular open source mobility model resources have been introduced.

For the emergency warning messages in VANETs, it is important for MAC protocol to provide a reliable and rapid transmission. A novel MAC type protocol, called RBTMA-NIT, is proposed by comprehensively considering the priorities of emergency message and its copies, receiving state of each node and contention policy. The proposed MAC protocol solves the indiscrimination of emergency message and its copies in traditional protocols, and reduces verbose emergency transmission. Therefore, it is more suitable for VANETs. The simulation results demonstrate that the RBTMA-NIT is able to greatly shorten the covering time and reduce the average number of required

transmission packets.

In order to further improve the performance of the RBTMA-NIT for common data transmission with multiple-channel condition, a new multiple channel MAC, called multiple channels busy tone multiple access with snooping (S-MBTMA), have been proposed by taking advantage of periodic on-off pulse and priority pulse. The simulation results demonstrate that the proposed MAC protocol is able to greatly improve the system throughput.

Besides the MAC, designing an effective routing protocol is another important issue for VANETs. As a result, a comprehensive survey has been done in this thesis. By firstly defining challenges and important designing factors for routing design in VANETs, the routing protocols for VANETs in literature have been suitably classified based on the implementation complexity and hardware requirements. The advantages and disadvantages of each category have been analyzed by examples. A general rule on routing protocol designing in VANETs has been provided as the conclusion of discussion.

6.2 Future Work

Providing reliable, secure and on time information transmission in VANETs is still an open topic in communication research area. There are plenty of future works waiting for solutions. For the RBTMA-NIT, the detailed performance analysis of RBTMA-NIT will be carried out, and a comprehensive single channel MAC protocol based on RBTMA-NIT will be discussed by considering the emergency message and common data jointly.

Meanwhile, the proposed work only considers single busy tone coverage area. How to use RBTMA-NIT in the multiple busy tone coverage areas, more specifically, how to effectively broadcast the busy tone to the neighboring affected area while suppressing the influence of useless busy tone pulse are necessary to be solved.

The proposed multiple channel MAC, S-MBTMA, relays on the integration of priority signal in the busy tone, which needs additional cost of hardware and system performance. Therefore, an improvement of S-MBTMA without dependency of priority signal is one possible extension. Meanwhile, the discussion of proposed S-MBTMA focuses on the single busy tone coverage area. As a result, it will face the similar challenge as RBTMA-NIT, i.e., how to adapt the proposed protocols into multiple busy tone coverage environments. Furthermore, a more detailed system performance analysis will be beneficial to convince people about the benefits of proposed mechanisms.

Till now, there are few routing protocols which can fulfill all design factors which have been summarized in the chapter 5. A comprehensive routing protocol will be designed on the basis of these works, and can support multiple data types, including traffic data, common data, multimedia data, etc. By taking into consideration hardware cost and vehicle movement, a series of improvement and optimization work will follow the design work.

BIBLIOGRAPHY

[1] National Center for Statistics and Analysis, "Traffic Safety Facts 2006", U.S DOT., Washington DC, 2007

[2] Cuyu, C. Xiang Yong, Meilin, S. Lin Liang, "Performance Observations on MAC Protocols of VANETs in Intelligent Transportation System", CMC '09, Jan. 2009.

[3] *http : //itsarch.iteris.com/itsarch/*

[4] Nawaporn Wisitpongphan, Fan Bai, Priyantha Mudalige, Varsha Sadekar, and Ozan Tonguz, "Routing in Sparse Vehicular Ad Hoc Wireless Networks" *IEEE Journal on Selected Areas in Communications* ,Vol.25 ,No.8 ,pp. 1538-1556, October 2007

[5] Fan Li, Yu Wang, "Routing in vehicular ad hoc networks: A survey", *IEEE Vehicular Technology Magazine*, vol. 2, no.2, June 2007

[6] ITS Standards Advisory No.3, April 2003

[7] ASTM E2213 - 03 Standard Specification for Telecommunications and Information Exchange Between Roadside and Vehicle Systems 5 GHz Band Dedicated Short

Range Communications (DSRC) Medium Access Control (MAC) and Physical Layer (PHY) Specification

[8] S.Biswas, R. Tatchikou, F. Dion, "Vehicle-to-vehicle wireless communication protocols for enhancing highway traffic safety", *IEEE Communications Magazine*, vol. 44, no. 1, pp. 74-82, 2006

[9] $http://www.ieee802.org/11/Reports/802.11_Ttimelines.htm$

[10] $http://vii.path.berkeley.edu/1609_wave/$

[11] $http://cordis.europa.eu$

[12] $www.cvisproject.org/en/links/cartalk\_2000.htm$

[13] $www.et2.tu-harburg.de/fleetnet/$

[14] Brad Karp,H. T. Kung,"GPSR: Greedy Perimeter Stateless Routing for Wireless Networks", *MobiCom 2000*

[15] Yao H.Ho, Ai H. Ho, Kien A. Hua, "Routing protocols for inter-vehicular networks: A comparative study in high-mobility and large obstacles environments", *Computer Communications*, vol. 31, no. 12, pp. 2767-2780, July 2008

[16] Jérôme Härri, Fethi Filali, Christian Bonnet, "Mobility Models for Vehicular Ad Hoc Networks: A Survey and Taconomy", Reserch Report RR-06-168, EURECOM, 2007

[17] Marco Fiore, "Mobility Models in Inter-Vehicle Communications Literature", November 2006

[18] Navid W. Camp, T., "Stationary distributions for the random waypoint mobility model" ,*IEEE Transactions on Mobile Computing*, Vol.3, No.1, 2004, pp.99-108

[19] F. Bai, N. Sadagopan, and A. Helmy, "IMPORTANT: A framework to systematically analyze the impact of mobility on performance of routing protocols for adhoc networks", in Proc. IEEE Conf. on Computer Commun. (INFOCOM), San Francisco, USA, March 2003, pp. 825-835

[20] Biao Zhou, Kaixin Xu, Gerla, M. , "Group and swarm mobility models for ad hoc network scenarios using virtual tracks",*IEEE MILCOM*,pp. 289- 294, 2004

[21] David R. Choffnes, Fabián E. Bustamante,"An integrated mobility and traffic model for vehicular wireless networks", 2nd ACM international workshop on Vehicular ad hoc networks. ACM,2005.

[22] Rahul Mangharam, Daniel S. Weller, Daniel D. Stancil, Ragunathan Rajkumar, Jayendra S. Parikh, "GrooveSim: a topography-accurate simulator for geographic routing in vehicular networks", 2nd ACM international workshop on Vehicular ad hoc networks, pp.59-68, 2005.

[23] Vuyyuru, Rama Oguchi, Kentaro Collier, Clay Koch, Ed. "Automesh: Flexible Simulation Framework for Vehicular Communication", MUSC, 2006

[24] Community Resource for Archiving Wireless Data At Dartmouth (Crawdad), *http* : *//crawdad.cs.dartmouth.edu*

[25] MIT Media Lab: Reality Mining, *http* : *//reality.media.mit.edu*

[26] Paramics: Microscopic Traf?c Simulation, *http* : *//www.paramics − online.com/*

[27] CORSIM: Microscopic Traf?c Simulation Model, *http* : *//www − mctrans.ce.ufl.edu/featured/TSIS/Version6/*

[28] Generic Mobility Simulation Framework (GMSF), "http://gmsf.hypert.net/"

[29] Holger Füßler, Marc Torrent-Moreno, Matthias Transier, Roland Krüger, Hannes Hartenstein, Wolfgang Effelsberg, "Studying vehicle movements on highways and their impact on ad-hoc connectivity" *ACM SIGMOBILE Mobile Computing and Communications*, Volume 10, Issue 4, 2006

[30] Realistic Vehicular Traces,"http://lst.inf.ethz.ch/ad-hoc/car-traces/"

[31] Shie-Yuan Wang, Chih-Che Lin, "NCTUns 5.0: A Network Simulator for IEEE 802.11(p) and 1609 Wireless Vehicular Network Researches", *IEEE VTC*, 2008

[32] Leonard Kleinrock, Fouad A.Tobagi, "Packet Switching in Radio Channels: Part II–The Hidden Terminal Problem in Carrier Sense Multiple-Access and the Busy-Tone Solution", *IEEE Transactions on Communications*, vol. 23, no. 12, pp. 1417-1433, Dec. 1975

[33] Hiroshi Harada, Ramjee Prasad, "Simulation and Software Radio for Mobile Communications", Artech House, 2002

[34] IEEE 802.11, Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, Standard, IEEE, June. 2007

[35] Giuseppe Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function",*IEEE Journal on Selected Areas in Communications*, vol. 18, no. 3, pp. 535-547, Mar. 2000

[36] Jun Peng, Liang Cheng, Biplab Sikdar, "A Wireless MAC protocol with Collision Detection", *IEEE Transactions on Mobile Computing*, vol. 6, no. 12, Dec. 2007

[37] Arash Tahmasebi Toyserkani, Erik G. Ström, Arne Svensson, "An Efficient Broadcast MAC Scheme for Traffic Safety Applications in Automotive Networks", *Proceedings IEEE WCNC*, Las Vegas, Nevada, USA, 2006

[38] Jun Peng, Liang Cheng, "A Distributed MAC Scheme for Emergency Message Dissemination in Vehicular Ad Hoc Networks", *IEEE Transactions on Vehicular Technology*, vol. 56, no. 6, Part 1, Nov. 2007

[39] N.Wisitpongphan, O.K.Tonguz, J.S.Parikh, P.Mudalige, F.Bai, V. Sadekar, "Broadcast storm mitigation techniques in vehicular ad hoc networks", *IEEE Wireless Communications*, vol. 14, no. 6, pp.84-94, Dec. 2007

[40] Nawaporn Wisitpongphan, Fan Bai, Priyantha Mudalige, Varsha Sadekar, and

Ozan Tonguz, "Routing in Sparse Vehicular Ad Hoc Wireless Networks", *IEEE Journal on Selected Areas in Communications*, vol.25, no.8, pp.1538-1556, Oct. 2007

[41] Hamid Menouar, Fethi Filali, Massimiliano Lenardi, "A survey and qualitative analysis of mac protocols for vehicular ad hoc networks", *IEEE Wireless Communications*, vol.13, no.5, pp.30-35, Oct. 2006

[42] Mingfei Wang, Linlin Ci, Ping Zhan, Yongjun Xu, "Multi-channel MAC Protocols in Wireless Ad Hoc and Sensor Networks," 2008 ISECS CCCM, vol. 2, pp.562-566, 2008

[43] Christian Tchepnda, Hassnaa Moustafa, Houda Labiod, Gilles Bourdon, "Prioritizing and Enhancing Vehicular Networks Authentication Process Using DSRC Channels Diversity", *IEEE WCNC*, 2008

[44] Jeonghoon Mo, Hoi-Sheung Wilson So, Jean Walrand, "Comparison of Multichannel MAC Protocols", *IEEE Transactions on Mobile Computing*, vol.7, no. 1, pp. 50 - 65, Jan. 2008

[45] Hui Wang, Huaibei Zhou, Hang Qin, "Overview of Multi-Channel MAC Protocols in Wireless Networks", *WiCOM*, Oct. 12-14, 2008

[46] Jungmin So, Nitin H. Vaidya, "Multi-channel mac for ad hoc networks: handling multi-channel hidden terminals using a single transceiver", *5th ACM international symposium on Mobile ad hoc networking and computing*, 2004

[47] Changchun Xu, Gan Li, Wenqing Cheng, Zongkai Yang," Multi-transceiver multiple access (MTMA) for mobile wireless ad hoc networks", *IEEE ICC*, 2005

[48] J.S.Pathmasuntharam, A.Das, A.K. Gupta, "Primary channel assignment based MAC (PCAM) - a multi-channel MAC protocol for multi-hop wireless networks", *IEEE WCNC*, 2004

[49] *http : //www.gpsclock.com/gps.html*

[50] , Murtuza Chhatriwala, Christine Park, Clarence Wong, Jason B. Kenagy, Weining Ruan," TIME SYNCHRONIZATION IN A CELLULAR NETWORK DEVICE", European Patent Application EP1240769

[51] D. B. Johnson, D. Maltz, and Y. Hu,"Dynamic source routing prtocol for mobile ad hoc networks (dsr)",IETF Internet Draft, Apri 2003

[52] Charles E.Perkins, E.Royer, and S.Das,"Ad-hoc On Demand Distance Vector Routing",IETF RFC 3561,July 2003

[53] Wu ShihLin, Lin ChihYu, Tseng YuChee, Sheu Jang-Laing, " A new multi-channel MAC protocol with on-demand channel assignmentfor multi-hop mobile ad hoc networks", I-SPAN 2000.

[54] Berkeley Highway Lab (BHL), http://bhl.calccit.org/

[55] D. Chowdhury, D.E. Wolf, and M. Schreckenberg, "Particle hopping model for two-

lane traffic with two kinds of vehicles:Effects of lane-changing rules", Physica A, pp. 235-687, 1997.

[56] K, I, Ahmed, "Modeling Drivers' Acceleration and Lane Changing Behavior" ,Ph.D Thesis, MIT Press, USA, 1999.

[57] Ioannis Broustis, Michalis Faloutsos, "Routing in Vehicular Networks: Feasibility, Modeling, and Security" *International Journal of Vehicular Technology*, Volume 2008 (2008), Article ID 267513

[58] Fabrizio Granelli,Giulia Boato, Dzmitry Kliazovich, "MORA: a Movement-Based Routing Algorithm for Ad Hoc Networks", *IEEE GLOBECOM*, 2006

[59] Menouar, H. Lenardi, M.Filali, F."Movement Prediction-Based Routing (MOPR) Concept for Position-Based Routing in Vehicular Networks", *IEEE VTC* 2007.

[60] Tarik Taleb, Ehssan Sakhaee, Abbas Jamalipour, Kazuo Hashimoto,"A Stable Routing Protocol to Support ITS Services in VANET Networks", *IEEE Transactions on Vehicular Technology*, vol.56, no. 6, Part 1, pp. 3337-3347, Nov. 2007

[61] J. Zhao and G. Cao, "Vadd: Vehicle-assisted data delivery in vehicular ad hoc networks", in Proc. IEEE Conf. on Computer Commun. (INFO-COM), 2006.

[62] G. Korkmaz, E. Ekici, and F. Ozguner, "An Efficient Fully Ad-Hoc Multi-Hop Broadcast Protocol for Inter-Vehicular Communication Systems", ICC '06. IEEE International Conference on Communications,Istanbul,June 2006,pp.423-428

[63] G. Korkmaz, E. Ekici, "Urban multi-hop broadcast protocol for inter-vehicle communication systems", Proceedings of the 1st ACM international workshop on Vehicular ad hoc networks,Philadelphia,USA,2004,pp.76-85

[64] Yanlin Peng, Zakhia Abichar, J.Morris Chang, "Roadside-Aided Routing (RAR) in Vehicular Networks",Communications, *IEEE ICC*, 2006

[65] Une Thoing Rosi, Chowdhury Sayeed Hyder ,Tai-hoon Kim, "A Novel Approach for Infrastructure Deployment for VANET",*IEEE FGCNC*, 2008

[66] Wenjing Wang, Fei Xie,Chatterjee, M. "TOPO: Routing in Large Scale Vehicular Networks", *IEEE VTC*, 2007.

[67] Peng Fan, "Improving Broadcasting Performance by Clustering with Stability for Inter-Vehicle Communication", *IEEE VTC*, 2007

[68] Tim Leinmueller, Elmar Schoc, "Greedy Routing in Highway Scenarios: The Impact of Position Faking Nodes", Workshop On Intelligent Transportation (WIT),2006.

[69] Tim Leinmueller; Elmar Schoch; Frank Kargl, "Influence of Falsified Position Data on Geographic Ad-Hoc Routing" Second European workshop, ESAS, 2005

[70] Harsch, C.;Festag, A.;Papadimitratos, P. "Secure Position-Based Routing for VANETs" , *IEEE 66th VTC* 2007.