

Vehicular Ad Hoc Networks and Dedicated Short-Range Communication

Jinhua Guo* and Nathan Balon†

University of Michigan - Dearborn

Telephone: +1.313.583.6439

(Dated: June 26, 2006)

*Electronic address: jinhua@umich.edu

†Electronic address: nbalon@umd.umich.edu

Contents

I. Introduction	3
II. VANET Characteristics	5
III. Dedicated Short-Range Communication	7
A. Characteristics of 5.9 GHz DSRC	8
B. Evaluated Wireless Technologies for DSRC	10
C. Complementary Technologies	12
D. Vehicle-to-Vehicle versus Vehicle-to/from-Infrastructure	13
IV. Intelligent Vehicle Applications Enabled by DSRC	14
A. Public Safety Applications	14
B. Non-Public Safety Application	18
V. Physical Layer	19
A. Channel Assignment	19
B. Control Channel Access	21
C. Dynamic Power Control	21
VI. Media Access Control (MAC) Protocols	23
A. Unicast	24
1. IEEE 802.11 MAC	24
2. Priority Access	27
3. Multi-Channel Coordination	28
B. Broadcast	29
1. Location Based Broadcast	31
2. Urban Multi-Hop Broadcast Protocol	31
3. Adaptive Adjustment of the Contention Window	33
VII. Routing	36
A. Context-Assisted Routing Protocol	38
1. Spatial Model	39
2. Context Assisted Routing	40

3. Road-Side Access Points Aware Routing	41
4. Dealing with Lossy or Intermittent GPS Reception	42
5. Direction-Aware Routing	42
6. Opportunity Forwarding	43
B. Applications of CAR	44
1. Query and Discovery	44
2. Multipath Routing	44
VIII. Security and Privacy	45
A. Potential Attacks	45
B. Security and Privacy Needs	46
C. Threat Mitigation	47
1. Anonymity - Removing Identifying Marks	48
2. Key Safety: Tamper-Resistant Devices	50
D. Group Signatures	50
IX. Conclusion	52
References	53

I. INTRODUCTION

Rapid advances in wireless technologies provide opportunities to utilize these technologies in support of advanced vehicle safety applications. In particular, the new Dedicated Short Range Communication (DSRC) offers the potential to effectively support vehicle-to-vehicle and vehicle-to-roadside safety communications, which has become known as Vehicle Safety Communication (VSC) technologies. DSRC enables a new class of communication applications that will increase the overall safety and efficiency of the transportation system.

Intelligent Transportation Systems (ITS) [23] are the future of transportation. As a result of emerging standards, such as 5.9 GHz dedicated short-range communication, vehicles will soon be able to talk to one another as well as their environment. A number of applications will be made available for vehicular networks that improve the overall safety of the transportation infrastructure. For instance, the system will be able to monitor traffic to

coordinate traffic lights so that traffic flows smoothly. Sensors will use feedback from vehicles to detect traffic jams. Public safety vehicles will broadcast, via the wireless channel, to change traffic signals in order to respond quickly to an emergency. Cars will communicate with one another to drive cooperatively, therefore avoiding collisions and improving efficiency. These are some of the possible applications, in the future, that will be possible with the advent of the DSRC standard.

Considering the tremendous benefits expected from vehicular communications and the huge number of vehicles, it is clear that vehicular ad hoc networks (VANET) are likely to become the most relevant realization of mobile ad hoc networks. The appropriate integration of on-board computers, roadmaps, and GPS positioning devices along with communication capabilities, opens tremendous opportunities, but also raises formidable research challenges.

DSRC [2], which is a candidate for use in a VANET, is a short to medium range communication service that supports both public safety and private communication. The communication environment of DSRC is both vehicle-to-vehicle and vehicle-to/from-roadside. The VANET aims to provide a high data rate and at the same time minimize latency within a relatively small communication zone.

A number of novel problems are associated with a VANET because of the unique characteristics of the network. To begin, the main differences between a VANET and a MANET is a MANET typically has no infrastructure available. In the case of a VANET, it is possible to strategically place access points along the side of the road, and in turn allow vehicles' access to the services available from the infrastructure. Also, one of the greatest challenges is the vehicles in the network move at greater speeds than most other MANETs, leading to a network that can frequently become fragmented. Furthermore, security and privacy are a crucial concern for a VANET.

In this chapter we discuss the challenges associated with a VANET, along with some possible solutions. To begin, the characteristics of a VANET are presented in Section II. Next, the DSRC standard is described in Section III. In addition, Section IV provides some of the applications that are possible in a vehicular network. Next, the issues related to the physical layer of DSRC are addressed in Section V. Furthermore, Section VI describes the issues related to the MAC layer of a VANET for both unicast and broadcast. Also, Section VII addresses the issues related to routing in a VANET, and introduces the CAR routing protocol that overcomes some of the problems related to routing in a VANET. Following

this, some of the issues related to security and privacy are explored in Section VIII, along with some possible security mechanisms. Finally, Section IX gives a conclusion along with the future direction of vehicular ad hoc networks.

II. VANET CHARACTERISTICS

The characteristics of a vehicular ad hoc network are unique compared to other mobile ad hoc networks. The distinguishing properties of a VANET offer opportunities to increase network performance, and at the same time it presents considerable challenges. A VANET is fundamentally different [5] from other MANETs. First, a VANET is characterized by a rapid but somewhat predictable changing topology. Second, fragmentation of the network frequently occurs. Third, the effective network diameter of a VANET is small. Fourth, redundancy is limited both temporally and functionally. Fifth, a VANET poses a number of unique security challenges.

The topology of the VANET changes frequently because of the high mobility of vehicles. Due to the frequent topology changes, the time that a communication link exists between two vehicles is brief. The reason why the link in a VANET is short lived is because vehicles travel at high speeds, approaching speeds of up to 200 km/h. One solution to increasing the duration a link is valid is to increase the transmission power. The problem associated with increasing a vehicle's transmission range in order to maintain a communication link is that it also decreases the throughput in the network. When vehicles travel in opposite directions, as can be expected, a link is maintained for a very small period of time. Even when vehicles travel in the same direction, with each vehicle having a transmission range of 500 ft, the wireless link between vehicles exists on the average for about a minute. Because vehicles exhibit a high degree of mobility it is difficult to maintain any form of group membership. For example, it is difficult to establish an accurate list of neighboring vehicles. Protocols that rely on group membership are difficult to implement for a VANET. Nevertheless, the topology of a VANET is also beneficial because a vehicle's movement is constrained by the road. The future movement of a vehicle is predictable.

The initial deployment of a VANET has the problem of only a small percentage of vehicles on the road being equipped with transceivers. The limited number of vehicles with transceivers will lead to frequent fragmentation of the network, causing a portion of the

network to become unreachable. Even when a VANET is fully deployed, fragmentation may exist in rural areas or during periods of light traffic, such as late at night. Since it could take years before the majority of cars are equipped with a transceiver, the VANET protocols should not assume that all vehicles can communicate.

A result of having poor connectivity between nodes is that the effective diameter of the network is small. For this reason, it is unrealistic for a node to maintain the complete global topology of the network. The limited effective diameter results in problems when trying to apply existing routing algorithms to a VANET. Traditional routing protocols are either proactive or reactive. To begin, proactive routing algorithms maintain routes by using tables. Frequent exchanges are needed between nodes to keep the routing information valid. Because the topology changes so rapidly, the routes maintained in the routing tables quickly become invalid. Traditional table-based routing approaches, such as DSDV, consume a great deal of bandwidth. Subsequently, reactive routing aims at establishing a route only when one is needed. The problem with the reactive approach is that a route must be discovered before the first packet is sent, which increases the time to send a message. Neither of these two approaches performs particularly well in a VANET. The problem with the proactive approach is that it does not scale well. The problem with the reactive approach is that even when a route to a destination is found right before transmitting a message, that route may also be very short lived because of mobility. In addition, the expected path life of a route decreases as the number of hops increases. A path may cease to exist almost as quickly as it was discovered. Sending a message a distance greater than three or four hops using traditional ad hoc routing algorithm is likely to result in a routing error. Routing is not likely to play as large a role as it does in other networks. In a VANET, it is more important to send a message towards a certain location.

Redundancy is crucial in order to provide specific services such as security. In a VANET redundancy is limited both temporally and functionally. Since links between nodes fail to exist for a significant period of time, it is extremely difficult to implement any form of redundancy.

Privacy and security are other issues that must be addressed. First, in order to gain support for the adoption of a VANET the anonymity of the driver must be preserved. For instance, the general public is unlikely to support a VANET if a driver's movement is recorded. If Anonymity features are not included it would be possible for third parties to

monitor a driver's daily activities. For this reason, mechanisms are needed to ensure the driver's privacy. Second, a VANET requires a high degree of security. It should not be possible to tamper with the messages in the VANET. To illustrate, the tampering of safety messages would result in automobile accidents occurring, which the system was designed to prevent. If strict security measures are not put in place an attacker would be able inject false data into the network resulting in the flow of traffic being altered and chaos within the transportation system.

These are some of the unique challenges related to a VANET. These are not the only unique characteristics of a VANET but they give a basic understanding of some of the issues in a VANET.

III. DEDICATED SHORT-RANGE COMMUNICATION

Dedicated Short-Range Communication (DSRC) is a standard that aims to bring vehicular networks to North America. Traffic fatalities have been a long standing problem in the United States, as in the rest of the world. As an indication of the severity of the problem, in 1999 there were 6,279,000 motor vehicle accidents that accounted for 41,611 deaths in the United States[12]. In 1991, the US Congress passed the Intermodal Surface Transportation Efficiency Act of 1991 that resulted in the creation the first generation of Intelligent Transportation System (ITS). The goal of the ITS program is to incorporate technology into the transportation infrastructure to improve safety. The first generation of the Dedicated Short-Range Communication (DSRC) system operates at 915 MHz and has a transmission rate of 0.5 Mb/s. This project had limited success and was used primarily by commercial vehicles and for toll collection. One example of a first generation DSRC application is E-ZPass that is used for electronic toll collection. The second generation of DSRC started in 1997 when ITS America requested that the Federal Communication Commission (FCC) allocate an additional 75 MHz of bandwidth. In October 1999, the FCC [12] allocated the 75 MHz of bandwidth in the 5.9 GHz band for the second generation of DSRC.

Since the allocation of the bandwidth, standardization bodies have been working on the implementation details of 5.9 GHz DSRC. The North American DSRC standards program aims at creating an interoperable standard for use in the US, Canada, and Mexico. The primary goal of the project is to enable drivers to receive up-to-date information regarding

their surrounding environment, thereby reducing traffic accidents. Furthermore, 5.9 GHz DSRC must have a low cost and be very scalable. In addition, the 5.9 GHz DSRC should require no usage fee from the users to access the network.

In this section, the characteristic of 5.9 GHz DSRC are given along with a comparison to 915 MHz DSRC. Next, a comparison of the possible wireless solutions for DSRC is given. Following this, some of the additional technologies that are used in DSRC are explained. Finally, the architecture of the VANET is described.

A. Characteristics of 5.9 GHz DSRC

DSRC is meant to be a complement to cellular communications by providing very high data transfer rates in circumstances where minimizing latency in the communication link and isolating relatively small communication zones are important. DSRC is also known as WAVE (Wireless Access in Vehicular Environments). Furthermore, an IEEE task group is currently working on the IEEE 802.11p standard for both the PHY layer and the MAC layer of DSRC. The primary reason why the MAC and PHY layers are being developed under 802.11 is to ensure that the standard remains stable over time. One of the cited problems of the original 915 MHz DSRC is that few implementations completely followed the standard. Instead, most of the original DSRC implementations were based on proprietary solutions. Realizing that proprietary implementations were one of the main causes of 915 MHz DSRC's lack of success, the new 5.9 GHz DSRC is an open standard.

The 5.9 GHz DSRC overcomes many of the weaknesses associated with 915 MHz DSRC. To begin, an increased amount of bandwidth is available for 5.9 GHz DSRC. Also, the 5.9 GHz DSRC spectrum is composed of seven channels of 10 MHz each. One channel is reserved for the control channel and six additional channels are service channels. Whereas, 915 MHz DSRC standard only supports the use of one or two channels. Next, 5.9 GHz DSRC supports high speed data transfers ranging from 6 Mb/s to 27 Mb/s. Under certain circumstances, the data rate can reach 54 Mb/s when two service channels are combined to form one 20 MHz channel. On the contrary, 915 MHz DSRC supports a data rate of only 0.5 Mb/s. Also, the transceivers used in vehicles required a reduced transmit power compared to 915MHz DSRC. In addition, the communication range is increased for 5.9 GHz DSRC. Transmission ranges of up to 1000 m are supported by 5.9 GHz DSRC, but typically the

TABLE I: Comparison of DSRC Technologies

	902 - 928 MHz Band	5850 - 5925 MHz
Spectrum	12 MHz	75 MHz
Data Rate	0.5 Mbps	6 Mbps - 27 Mbps
Interference Potential	High	Low
Coverage	One communication zone	Overlapping communication zones
Maximum Range	300 ft	1000 m
Minimum Separation	1500 ft	50 ft
Channel Capacity	1 to 2 channels	7 channels
Downlink Power	Nominally less than 40 dBm	Nominally less than 33 dBm
Uplink Power	Nominally less than 6 dBm	Nominally less than 33dBm

transmission range is shorter to promote greater frequency reuse. The transmission range that is used is based on the type of application and the channel in use. Next, the interference potential for 5.9GHz DSRC is much lower than for 915 MHz DSRC. The only interference in the 5.9 GHz band comes from sparsely located military radars and sparsely located satellite uplinks, whereas 925 MHz DSRC suffers from considerable interference. The 902-928MHz band is full of traffic. Other devices that occupy the band are 900 MHz phones, rail car AEI readers, and wind profile radars. Table I contains a comparison between 925 MHz DSRC and 5.9 GHz DSRC.

A large array of applications are being developed for DSRC. The applications of DSRC are categorized into the following four classes.

- **Vehicle-to-Vehicle** applications transmit messages from one vehicle to another.
- **Vehicle-to/from-Infrastructure** are applications in which messages are sent either to or from vehicle to a Road Side Unit (RSU).
- **Vehicle-to-Home** is a class of application that is used when a vehicle is parked at the driver's residence, for purposes such as transferring data to the vehicle.
- **Routing Based** applications are used when the intended recipient is greater than one-hop away.

Based on these four application classes, the DSRC applications can be further categorized as safety and non-safety applications. Furthermore, the application messages of DSRC applications may be either event driven or periodic. The event driven messages are sent when a certain event occurs. For instance, when a vehicle is involved in a collision it will generate a message to warn other vehicles that an accident has occurred. On the other hand, periodic messages are repeatedly transmitted at a specific interval, such as a vehicle announcing its state or a RSU broadcasting the status of a traffic light. This forms the basis of the type of applications that are possible in a VANET.

DSRC supports a number of different network protocols for interoperability in the hope of gaining widespread adoption. To begin, DSRC supports the long-established TCP/IP protocol, which allows IP based routing in DSRC. As a result of supporting TCP/IP, most of the traditional Internet applications are available in the VANET. Next, WAVE Short Message Application is used for the majority of vehicle-to-vehicle safety communications. The reason that IPv6 is not used for many of the safety applications is because of the size associated with IPv6 headers. The IPv6 headers are a minimum of 40 bytes which is close to the size of a typical safety message. The average size of a safety message is approximately 100 bytes. To increase the overall performance of the network and allow more vehicles access to the network, the requirement of using the IPv6 protocol was removed in favor of the WAVE Short Message Application protocol. There is also the C2C-CC protocol that is being developed for VANETs in Europe. Figure 1 contains the protocol stack for DSRC.

The DSRC standard is still a work in progress. Many of the final details of DSRC are unknown at the present time. As the standardization process continues, new features are sure to be added and some of the original features of the proposal may be removed.

B. Evaluated Wireless Technologies for DSRC

A number of wireless solutions were evaluated for use as the primary communication medium for DSRC [34]. A requirement of the wireless technology is its latency must be 100 ms or less, offer high throughput, and have a communication range of 100 m to 1000 m. In addition, the wireless technology must support a number of diverse communication schemes [2]. First, the wireless technology should support both one-way communication allowing a vehicle to send a broadcast message and two-way communication allowing two

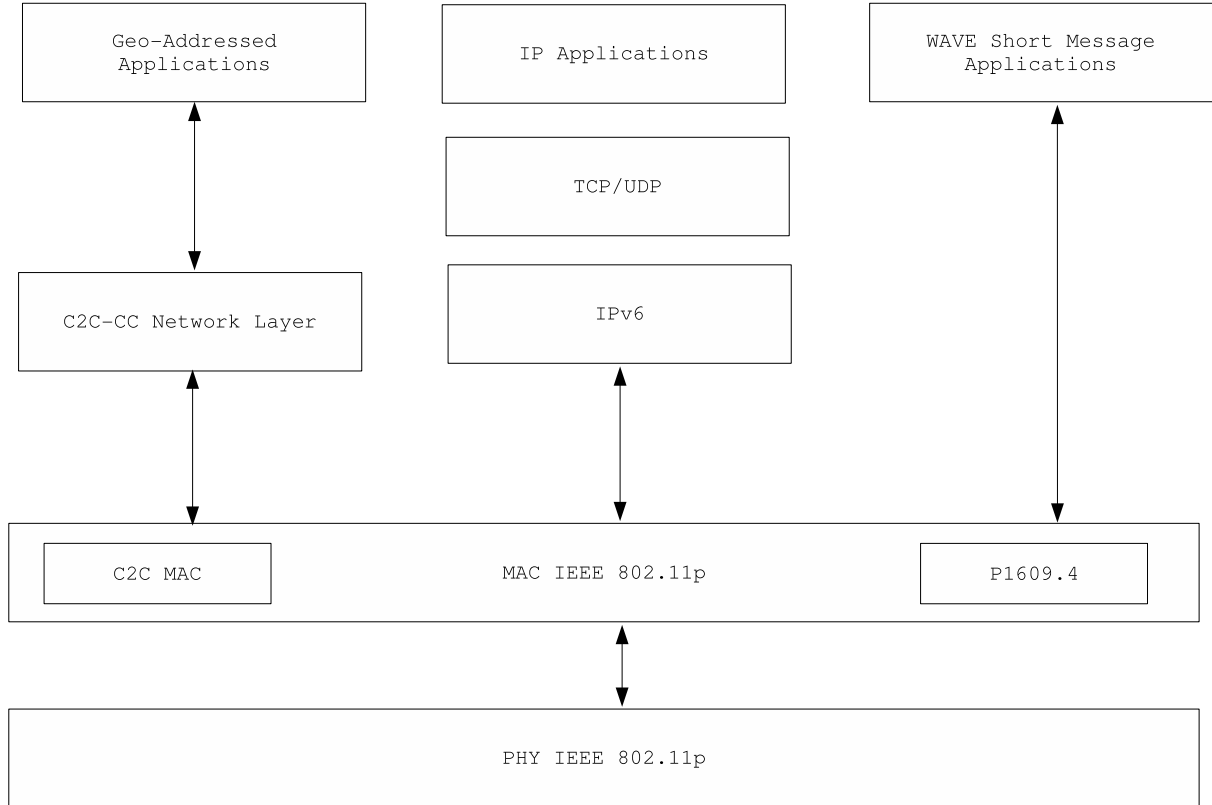


FIG. 1: DSRC Protocol Stack

vehicles to establish a dialog with each other. Second, the technology must also support both point-to-point communication where a message is intended for a specific location and point-to-multipoint communication where a message is intended for multiple receivers. Third, one-way or two-way communication may be either point-to-point or point-to-multipoint.

The wireless technologies were evaluated based upon how well they meet the requirements of DSRC. In the end, a modified version of 802.11a was chosen as the primary means of communication for DSRC. A number of the other evaluated technologies were found unacceptable for one reason or another. For instance, both cellular systems and satellite systems offer a significant amount of bandwidth but have too high of latency to be considered useful for some applications of DSRC. A further drawback of cellular technology is its lack of broadcast support. Furthermore, the cost of the wireless technology must be low. At the present time both cellular and satellite technologies are expensive. In comparison, the cost of wireless access for DSRC is free because the technology is based on ad hoc networks. Also, infrastructure costs of DSRC are much cheaper than both cellular and satellite. Table

TABLE II: A Comparison of Wireless Technologies

	DSRC	Cellular	Satellite
Range	100 1000 meters	Kilometers	Thousands of kilometers
Latency	200 μ s	1.5 to 3.5 s	10 to 60 s
Cost	None	Expensive	Very expensive

II contains a comparison of the wireless technologies.

C. Complementary Technologies

A number of additional technologies will be used in DSRC, as a complement to 802.11a. To begin, a digital map is required by each vehicle in the VANET. A digital map enables the application of an enhanced vehicle navigational system. Another use of a digital map is for location-based routing. A major challenge of DSRC is the dissemination of new maps, such as when traveling to a city never visited before or dissemination of an updated map when a road is altered. A further requirement of DSRC is that all vehicles must be able to determine their location. Global Positioning Systems (GPS) provides a great solution to the problem of determining a vehicle's location. The main drawback of GPS is the location can only be determined when there is a clear path to the satellite, which means that GPS will not work in all situations (such as when vehicle pass through a tunnel). Next, sensors are used to provide additional input to the system. Both vehicles and RSUs (i.e., RSU are stationary devices that are mounted road side that function similar to access point) are equipped with sensors that monitor the local conditions. To illustrate, sensors placed along the roadway can detect conditions such as ice on the road so that drivers are able to alter their driving. A RSU receives input from the sensor that ice is on the road and then transmits this information to the vehicles in the location. Finally, another technology that is sure to be initially included in DSRC is radar because not all vehicles in the future are likely to be DSRC enabled. A radar device is added to a RSU so that it is able to detect vehicles lacking a DSRC transceiver. The RSU can then relay location information about a non-DSRC equipped vehicle to the other vehicles in the VANET. Although the 802.11a protocol is the core component of the system, a number of complimentary technologies will also find their way into DSRC.

D. Vehicle-to-Vehicle versus Vehicle-to/from-Infrastructure

Two types of DSRC devices are used for communication in the VANET: an On-Board Unit (OBU) and a Road-Side Unit (RSU). First, each vehicle is equipped with an OBU which is a transceiver mounted within a vehicle along with a computational device. Each vehicle also has an omni-directional antenna that the OBU uses to access the wireless channel. Furthermore, each vehicle has sensors to provide input to the OBU. The sensors record the local conditions of the vehicle. Second, RSU are stationary devices that are mounted road side. The RSU is similar to an OBU in that it has a transceiver, antenna, processor, and sensors. The RSU are strategically placed along the road in order to provide services to vehicles. For instance, a RSU may be placed near an intersection to improve the flow of traffic through that intersection and reduce accidents. Also, a commercial entity can deploy a RSU to provide value-added services to their customers. As an illustration, a gas station can use a RSU to collect electronic payments from their customers. The RSU may use either a directional antenna or an omni-directional antenna depending on the type of application provided by the RSU. A directional antenna is beneficial when the signal only needs to propagate in a specific direction. For example, a distribution company could use a RSU for access control at the gate of a warehouse. In doing so, only pre-approved vehicles would be allowed through the gate. Since the transmission is to a specific location, in this case the gate, a directional antenna is used. To conclude, a VANET is composed of OBUs and RSUs.

Vehicular ad hoc networks are not pure ad hoc networks. An infrastructure of RSU will exist, which allows the VANET access an external network such as the Internet. Also, a RSU can communicate with another RSU through a wired infrastructure, making the communication between RSUs more reliable. To conclude, each RSU will require a license to operate the unit at a specific location and a specific frequency. The FCC requires a license for each RSU to maintain the integrity of the network and so that the services provided by commercial entities does not detract from the primary purpose of the network, thus ensuring safety. If the FCC did not regulate the spectrum, applications (such as multimedia services) would hinder the safety applications.

IV. INTELLIGENT VEHICLE APPLICATIONS ENABLED BY DSRC

A number of unique applications are being standardized for DSRC and similar projects worldwide. The goal of the standardization is to create a common set of application protocols. While there will be a common set of application protocol, the automobile manufactures will be able to differentiate their products based on the user interface they provide to the driver. For instance, a simple user interface may only give the driver audio feedback. On the contrary, a more advanced user interface may provide the driver with a touch screen mounted within the dashboard, allowing the driver a visual display of the road. To conclude, each vehicle has an OBU that follows the DSRC specification, but each automobile manufacturer is able interface the OBU with a proprietary user interface.

DSRC is composed of public safety and non-public safety applications. First, the objective of the public safety applications is the improvement of the overall safety of the transportation infrastructure. Second, the non-public safety applications increase the comfort of the driver by adding value-added services. Public safety applications are always given priority over the non-public safety applications.

A. Public Safety Applications

The public safety applications protect the safety of life, health, or property. The public safety services of DSRC are provided by either a governmental agency or a non-governmental organization under the authorization of a governmental agency. The Vehicle Safety Communication (VSC) project [34] determined 34 possible safety applications for DSRC. These applications were analyzed to determine the potential safety benefit provided by the application. The application analysis was based on the saving in terms of years saved from life lost, for both fatal and non-fatal accidents. Next, the applications were rated in terms of the estimated time before the application is commercially deployable. Near-term applications are deployable between 2007 and 2011. Mid-term applications are deployable between 2012 and 2016. Long-term applications are deployable beyond 2016. Furthermore, the applications were rated on their effectiveness in preventing accidents. Last, the applications were rated on their capability to operate based on the market penetration of DSRC enabled vehicles. The VSC project team used this criterion to determine the safety benefit of the applications.

TABLE III: Public Safety Applications

Safety Application	Deployment Timeframe
Traffic Signal Violation Warning	Short-term
Curve Speed Warning	Short-term
Emergency Electronic Brake Lights	Short-term
Preocrash Warning	Mid-term
Cooperative Forward Collision Warning	Mid-term
Left Turn Assistance	Mid-term
Stop Sign Movement Assistance	Mid-term
Lane Change Warning	Mid-term
Cooperative Collision Warning	Long-term
Intersection Collision Warning	Long-term

Based upon the evaluation criteria, the VSC project determined the safety applications contained in Table III are the highest priority applications to implement because they provide the greatest benefit in terms of saving lives.

Traffic signal violation warning application provides the greatest benefit in estimated functional-life years saved by the applications that could be implemented in the short-term. Passing through an intersection is one of the most dangerous activities that one encounters while driving. The goal of this application is to reduce collisions at intersections. In this scenario, a RSU is placed near an intersection that has a traffic light as depicted in Figure 2. Infrastructure-to-vehicle communication is used to warn approaching vehicles of the status of the traffic light and to alert drivers of a potential light violation. The data sent to approaching vehicles includes the status of the light, the time of light changes, the traffic light location, and the direction of the light signals. When a vehicle receives a traffic signal violation warning message, computation is performed on the received data to determine if the driver is at risk of inappropriately entering the intersection and if so a warning is issued to the driver. The traffic signal violation warning is a simple one-way application that provides the greatest safety benefits of the VANET applications. More complex variations of this scenario are used for applications such as left-turn assistance and stop sign movement assistance.

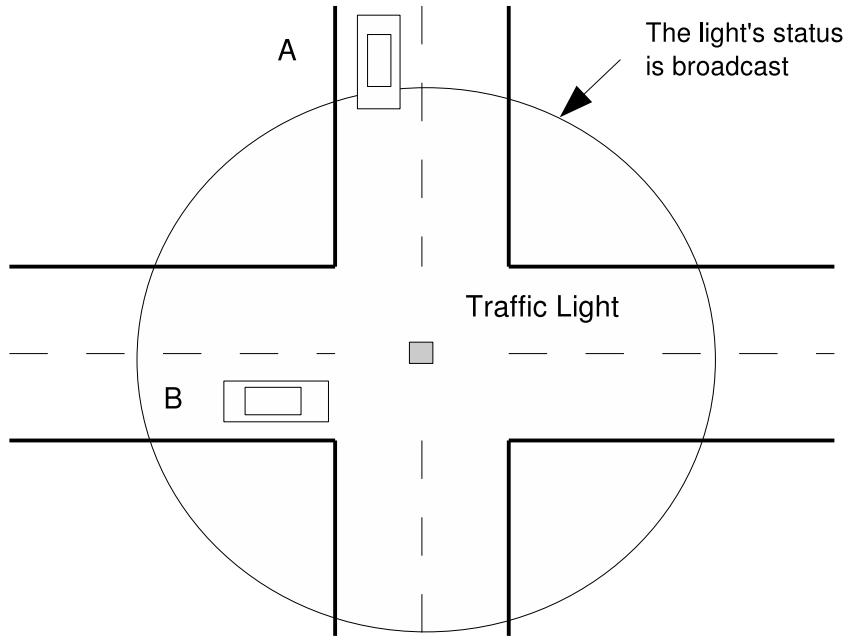


FIG. 2: Traffic Signal Violation Warning

Emergency electronic brake lights application is another short-term solution that provides a warning to a trailing vehicle when a vehicle in front of it applies its brakes. Figure 3 illustrates vehicle *A* broadcasting a warning message after applying its brakes. The emergency electronic brake light application is beneficial in situations where visibility is limited, such as poor weather conditions. The data contained in vehicle *A*'s broadcast message is the deceleration rate and braking vehicle's location. When vehicle *B* receives the warning, an algorithm is invoked to determine the relevance of the message and whether or not the vehicle is endangered. If so, a warning is sent to the driver. The emergency electronic brake light application significantly reduces accidents by giving the driver a warning before they are able to visually sense the danger.

Curve speed warning application aids a driver as he approaches a winding stretch of road. Typically, a sign is posted on the side of the road to warn drivers to reduce their speed. The success that a driver has going through the curve is based solely upon his/her judgment. A curve warning system can tremendously improve the accuracy by guiding a driver through a curve using information such as the characteristics of the vehicle, the weather conditions, and the curves geometry. A RSU is placed at a potentially dangerous curve. Furthermore, the RSU can improve safety by using sensors to estimate the condition of the road. The RSU unit then periodically broadcasts warnings of the condition of the road through the curve.

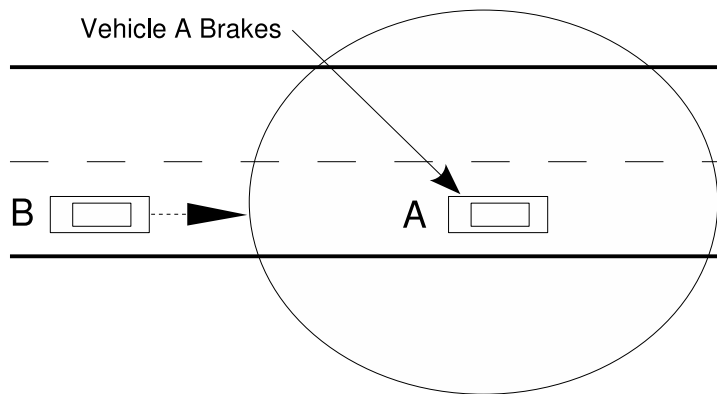


FIG. 3: Emergency Electronic Brake Lights

When a vehicle receives the broadcast from the RSU, the information is then processed by the OBU. If the vehicle's velocity exceeds a safe speed of travel, a warning is issued to the driver.

Lane change warning is an application that is expected to be implemented in the mid-term and assist a driver while changing lanes. The lane change warning application is a vehicle-to-vehicle application. Each vehicle receives periodic broadcast from the surrounding vehicles. Also, each vehicle maintains a table containing the vehicles in the immediate proximity. For this application to be successful, the vehicle locations maintained in the table must be very precise. When the driver signals his or her intent to change a lane, the OBU uses the received data to determine if the road conditions are safe to perform a lane change. One means triggering the application is when the turn signal is applied by the driver, which then invokes the lane change algorithm. If the attempted lane change puts the driver in danger, a warning is generated. The main drawback of the lane change warning application is that it requires that a high percentage of vehicles are DSRC equipped.

These are just a few examples of the safety applications that are a possible in a VANET. The actual implementations of these applications may change over time. For example, more complex and accurate implementations of the traffic signal violation warning are possible. Also, the size of the application packets are typically small (between 100 and 500 bytes). The size of the application packets presents little problem in the realization of these applications. On the contrary, one of the initial barriers of implementing many of the DSRC applications is the low initial penetration rate of vehicles that are DSRC enabled. To conclude, as time passes and more vehicles become DSRC equipped, more DSRC applications will be

TABLE IV: Non-Public Safety Applications

Access Control	Gas Payment	Point-of-Interest Notification
Drive-Thru Payment	Data Transfer	Instant Messaging
Car Rental	Fleet Management	Enhance Route Guidance
Truck Stop Data Transfer	Parking Lot Payment	Toll Collection

implemented.

B. Non-Public Safety Application

The primary focus of DSRC is for the creation of safety applications, but a number of additional non-safety applications have been proposed. The non-public safety services require licenses to provide the DSRC-based services. The FCC requires a license for service providers in an effort to eliminate services that would be detrimental to the VANET. Table IV lists the non-public safety applications that have been proposed for DSRC.

Non-public safety applications increase the overall comfort of the driver. Electronic toll collection is one possible non-safety application. Instead of a driver having to stop at a toll booth to make a payment, the payment is made electronically through the network. Also, a number of entertainment features have been proposed for vehicular networks, such as the transferring of music and video files for in-car entertainment. Applications such as these will probably not be implemented in DSRC in the foreseeable future because of the limited bandwidth and the fundamental focus on safety applications. The in-car entertainment application would consume a large amount of network resources. Although the organization have been give the approval to begin work on multi-media applications as long as they do not constrain the safety application or require any modifications to the safety protocol. Another, possible application is instant messaging which enables the driver to send a message to another vehicle. The sent message could be either predefined or custom. In addition, enhanced route guidance and navigation enables a driver to make decisions on the path of travel, based upon the received information. In this application, a RSU transmits up-to-date navigational information to the vehicles. Some of the possible information that is transmitted is construction advisories, road closings, detours, and parking restrictions. Finally, point-of-interest notifications are transmitted from the RSU containing information

regarding places of interest in the area. Some of the possible data exchanged is the location of gas stations, restaurants, and lodging. For example, a RSU might broadcast the location of the gas stations in the area along with the prices of gas. These are some of the non-safety applications that are possible in a VANET. It is expected that commercial organizations will find numerous other uses for DSRC and the greatest innovation of DSRC will come from the non-safety applications.

V. PHYSICAL LAYER

The physical layer is responsible for transmitting the raw bits on to wireless channel. First, the channel assignment of DSRC is described. Next, the control channel access is discussed along with the problem of coordinating the access of multiple channels in a vehicular network. Finally, the topic of dynamic power adjustment is explored.

A. Channel Assignment

The FCC allocated 75 MHz of the radio spectrum for DSRC. The 5.9 GHz DSRC spectrum is composed of six service channels which are each 10 MHz. Also, one control channel is provided by the DSRC standard, which is also 10 MHz. As stated earlier, the FCC recommends no unlicensed use of DSRC band. Figure 4 provides the channel layout for DSRC. The data rates possible for a 10 MHz channels are 6, 9, 12, 18, 24, and 27 Mb/s with a preamble of 3 Mb/s. The modulation scheme used by DSRC is Orthogonal Frequency Division Multiplexing (OFDM). Also, the subcarrier frequency spacing of 802.11a is double that of DSRC. In addition, DSRC doubles the guard period in comparison to 802.11a. The following list contains the channels of DSRC and the type of applications that are supported by the channel.

- Channel 172 is reserved for medium power safety applications.
- Channel 174 is reserved for medium power applications that are shared by all.
- Channel 175 is a combination of channels 174 and 176.
- Channel 176 is reserved for medium power applications that are shared by all.

5.850 GHz		CH. 175			CH. 181		5.925 GHz
Reserved	CH. 172 Service	CH. 174 Service	CH. 176 Service	CH. 178 Control	CH. 180 Service	CH. 182 Service	CH. 184 Service
5 MHz	10 MHz	10 MHz	10 MHz	10 MHz	10 MHz	10 MHz	10 MHz

FIG. 4: DSRC channels

- Channel 178 is the control channel it support all power levels, safety application broadcasts, service announcements, and vehicle-to-vehicle broadcasts messages.
- Channel 180 is reserved for low power configurations and provides little interference when units are separated by 50 ft or more.
- Channel 181 is a combination of channels 180 and 182.
- Channel 182 is reserved for low power configurations and provides little interference when units are separated by 50 ft or more.
- Channel 184 is reserved for a high power service channel that is used to coordinate intersection applications.

The current wireless technology is only able to listen to one channel at a time. In the initial deployment of DSRC, each vehicle will have a single transceiver. The drawback of having a single transceiver is that only one channel at a time is able to be monitored. To overcome this problem, it is possible to equip either an OBU or RSU with multiple transceivers allowing them access to multiple channels simultaneously. To illustrate, if an OBU is equipped with two transceivers, one transceiver can monitor the control channel while communication is underway on a service channel. The drawback to having multiple radios is it increases the complexity and the cost. For the initial roll out of DSRC, it is envisioned that vehicles will have only a single transceiver. As a result of only being able to listen to single channel at time is channel coordination is needed.

B. Control Channel Access

Channel 178 is reserved for the control channel. The control channel is the most important channel of DSRC, and the efficient use of this channel is critical. Each OBU monitors the control channel for both broadcast safety messages and brief service channel announcements. The control is monitored by each vehicle and RSU. Since there is a limited amount of bandwidth available, communication on the control channel is brief. The FCC recommends that the control channel is used for messages that take less than 200 μ s to transmit. If the communication last longer than 200 μ s another channel must be used.

Vehicles must periodically switch to the control channel to receive safety messages. A requirement of DSRC is that all vehicles must switch to the control channel every 100 ms and remain on the channel for a minimum amount of time. The purpose of vehicle switching to the control channel every 100 ms is to allow the reception of the safety broadcast from the surrounding vehicles. To guarantee that safety messages are not sent before the vehicles switch to the control channel, the time that the vehicles switch to the channel must be synchronized. One possible way to synchronize the control channel access is with the time received from a GPS unit. There are a number of proposals for the DSRC standard as to how to best implement synchronization with GPS for control channel access.

The control channel is also used for service announcements. When a service discovered is of interest to the OBU, it will switch from the control channel to the service channel to use the service. For instance, a RSU may provide the service of a map update. An updated map is then transferred to a vehicle. The OBU of a vehicle will discover the map update and switch to a service channel to begin the transfer of the new digital map. If the transfer takes too long to complete, the vehicle must switch to the control channel to receive safety messages and then switch back to the service channel to resume the file transfer. The control channel coordination allows a vehicle to correctly receive safety messages and also use the available services in the network.

C. Dynamic Power Control

In a wireless network, there are two ways to improve the overall efficiency of the network. The efficiency is improved by either reducing the transmission rate or reducing the trans-

```

a is a constant
MR is the maximum transmission range
INPUT: fraction of time stopped Ts/T
OUTPUT: transmission range TR

K = estimate_K(Ts/T)
IF Ts/T == 0 THEN
    TR = MR
ELSE
    TR = min(MR * (1-K), sqrt(MR * ln(MR)/K + a * MR)
END

```

FIG. 5: Dynamic Transmission Range Algorithm

mission range. Lowering the transmission rate is not always possible since some messages are required to be received within a specified time period. The greatest chance of increasing the throughput of the network is to reduce the transmission range.

Varying the transmission range is used to maintain connectivity. Increasing a vehicles transmission range, when the distance between vehicles is large, results the connectivity being maintained. When the distance between vehicles is small (such as in a traffic jam) a reduction in transmission range increases the network throughput.

One solution that dynamically adapts the transmission range of a vehicle is based on a local estimation of the density of vehicles[3]. The Fundamental Traffic Flow Relationship is determined by $q = u * k$, where q is the number of vehicles that pass a given point per time period, u is the speed, and k is the density. The goal of the local density estimate is to determine the Minimum Transmission Range (MTR) that still provides connectivity. One way that the MTR can be determined is from a Minimum Spanning Tree (MST), where the longest edge of MST equals the MTR. An estimate of the MTR is given by the transmission range algorithm in Figure 5. The algorithm is based on a mapping function $r = g(T_s/T)$ where T_s is the time stopped and T is the time period. The algorithm is periodically run to estimate of the local density of vehicles. As a result of running the algorithm, the transmission power is adjusted so that connectivity in the VANET is maintained.

The Fair Power Adjustment for Vehicular environments (FPAV) [33] algorithm is another algorithm that adjusts the transmit power of a vehicle. The FPAV algorithm differs from the previous discussed power adjustment algorithm in terms of the goal of the algorithm. While the previous dynamic transmission range algorithm's goal was to maintain connectivity, the

goal of the FPAV algorithm is to adjust the power so that additional bandwidth is available to non-safety applications.

VI. MEDIA ACCESS CONTROL (MAC) PROTOCOLS

Media Access Control protocols such as TDMA, FDMA, or CDMA are difficult to implement for VANET. For any of these protocols to be used either time-slots, channels, or codes need to be dynamically allocated, which requires synchronization that is difficult to achieve in a network where the nodes have a high degree of mobility [36].

The objective of the media access control layer is to arbitrate the access to the shared medium, which in this case is the wireless channel. If no method is used to coordinate the transmission of data, then a large number of collisions would occur and the data that is transmitted would be lost. The ideal scenario is a MAC that prevents nodes within transmission range of each other from transmitting at the same time, thus preventing collisions from occurring. Equally important, the media access control must be fair, efficient, and provide the ability to prioritize traffic.

Another obstacle restricting the wide-spread adoption of vehicular ad hoc networks is that is based on the wireless protocol IEEE 802.11, that was designed for networks with different characteristics than a VANET. A large focus of the 802.11 standards has been on wireless LANs. The majority of the 802.11 protocols are designed around the fact that a centralized controller is present in the network, the access point (AP). In vehicular ad hoc networks the use of an AP is limited to situation where a RSU is present. In a WLAN communication tends to be point-to-point. On the other hand, a large portion of the communication in a VANET is broadcast in nature. For these reasons, some modifications to the 802.11 protocols are necessary.

This section first discusses media access control for unicast frames. To begin, a brief explanation of the IEEE 802.11 MAC, upon which DSRC is based, is given. The next topic that is discussed is the differentiation of MAC frames, followed by multi-channel coordination at the MAC layer. Next, the problems related to sending broadcast messages are discussed, along with some possible solutions to overcome the problems related to broadcasting.

A. Unicast

The MAC layer of DSRC is based on 802.11; in turn it provides the standard IEEE 802.11 MAC layer mechanisms to reliably transmit a unicast frame. The following sections describe the MAC layer for a VANET. First, a quick overview of the 802.11 MAC is given. Second, the issue of priority access in a vehicular environment is explored. Third, multi-channel coordination is examined.

1. IEEE 802.11 MAC

The 802.11 standard defines two MAC protocols, the Distributed Coordination Function (DCF) and the Point Coordination Function (PCF). The DCF is an asynchronous contention based access protocol. In a contention based protocol, all nodes that have data to send contend for access to the channel. On the other hand, PCF is a contention free protocol that provides access to the medium by scheduling when a node can transmit. Contention free protocols, such as PCF, enable the use of real-time services. Although there are benefits to using the PCF, it is not applicable for a VANET in most cases because it relies on central node to support the real-time delivery of packets. For this reason, the majority of communication that takes place in DSRC uses the DCF.

The 802.11 family of protocols uses Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) with an acknowledgment to restrict the number of collisions and to reliably transmit packets. The DCF achieves collision avoidance with a random back-off procedure. The IEEE 802.11 standard uses the concept of slot time. Each time-slot for 802.11a is $9\mu\text{s}$, but in general the slot time will vary based on the physical layer characteristics of the IEEE 802.11 protocol. When a node begins a transmission, it randomly selects the number of time slots it must wait before transmitting, which is known as the back-off process. One clock tick of the back-off timer expires, when the medium remains free from transmission for one time-slot. Collisions are avoided by nodes randomly selecting different values for their back-off timers. The value of the back-off timer is chosen randomly in the range of $[0, CW)$, where CW is the size of the contention window. If two nodes wish to transmit a frame at the same time and they select different values for their back-off timers, then a collision is avoided because the nodes will transmit at different times. On the other

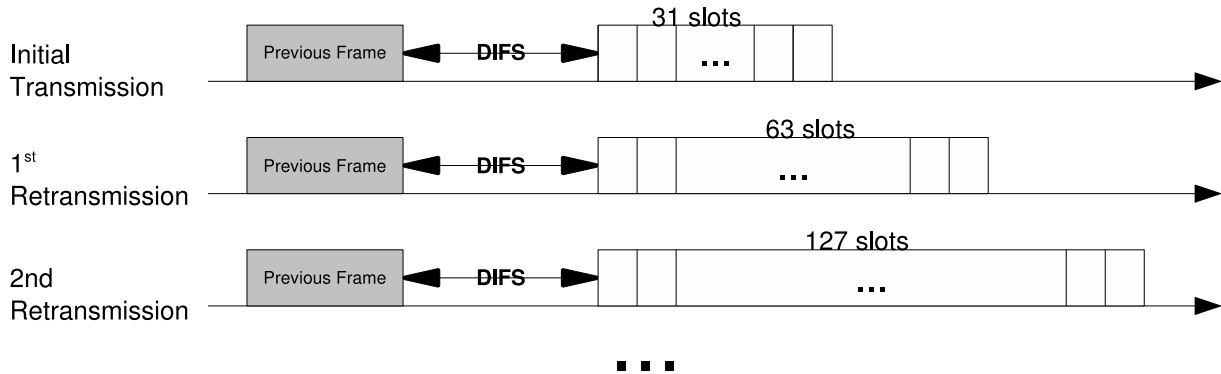


FIG. 6: Back-off Procedure

hand, if two nodes both decrement their back-off timers to zero at the same time, a collision occurs. The back-off time ($T_{backoff}$) is calculated with Equation 1 where j is the number of retransmissions:

$$T_{backoff} = Rand(0, 2^j * CW_{min}) * T_{slot} \quad (1)$$

The contention window continues to increase, as illustrated in Figure 6, after each failed transmission until CW_{max} is reached. If the transmission of a frame does not succeed after a predefined numbers of attempts the frame is discarded. After a frame is successfully received the CW is reset to CW_{min} .

In addition, the DCF use a number of different inter-frame spaces. When a node wishes to transmit a frame, it must wait for the Distributed Inter-Frame Spacing (DIFS) of $34 \mu s$ to expire, for 802.11a. During this time the wireless medium must remain free. If a transmission is overhead while a node is waiting for the DIFS expire, the node then defers its attempted access to the medium until the medium becomes free. When the overheard transmission is complete the node will then begin to listen to the medium until the DIFS has expired. Once the DIFS is complete, the node will begin to count down its back-off timer. If a transmission occurs before the back-off timer reaches zero, the node will then pause its back-off timer. The node must then wait for the medium to remain free for the DIFS to resume the back-off timer. Finally when a node back-off timer reaches zero it will begin its transmission.

The wireless transmission is made reliable with the introduction of an explicit acknowledgment mechanism. The intended receiver of a frame transmits an acknowledgment (ACK) which alerts the sender that the frame has been successfully received. If an ACK is not re-

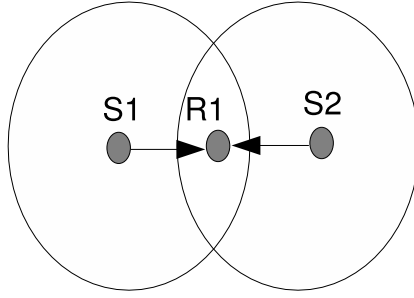


FIG. 7: Hidden Terminal Problem

ceived by the sender of a frame, it is assumed that the frame was not successfully received, and another attempt is made to transmit the frame.

One of the main problems affecting the reliability of the DCF is known as the hidden terminal problem. The hidden terminal problem is the main cause of collisions in a wireless network. The hidden terminal problem occurs when there are two nodes that are outside the transmission range of each other but each transmits to a node that is shared between them. In Figure 7 below, nodes $S1$ and $S2$ cannot sense each other's transmissions. Therefore, the medium appears free to both $S1$ and $S2$. If both $S1$ and $S2$ were to transmit to $R1$ at the same time, a collision would occur at $R1$ and neither of the frames would be successfully received.

The hidden terminal problem is addressed by 802.11 with an optional RTS/CTS exchange before any data is transferred[21]. Figure 8 illustrates the RTS/CTS sequence. When node $S1$ has data to send, once it is able to gain access to the medium (e.g., after the nodes back-off timer expires), node $S1$ first transmits a RTS to the intended receiver $R1$. When $R1$ receives the RTS after a Short Inter-Frame Spacing (SIFS) has expired, node $R1$ will respond with a CTS. When the CTS is received at $R1$ it will signal to node $R1$ that $S1$ is ready to receive a data frame. The hidden node problem is mostly eliminated, when $S2$ overhears the CTS transmitted from $R1$ it then sets the network allocation vector (NAV) for the amount of time it takes to complete the communication. Node $S2$ will then defer from accessing the wireless medium until the NAV expires and the transmission between $S1$ and $R1$ is complete. When $S2$ overhears the ACK sent from $R1$ it knows the transmission is complete. After the DCF Inter-Frame Spacing (DIFS) has elapsed nodes in the network can then begin to contend for access to the channel.

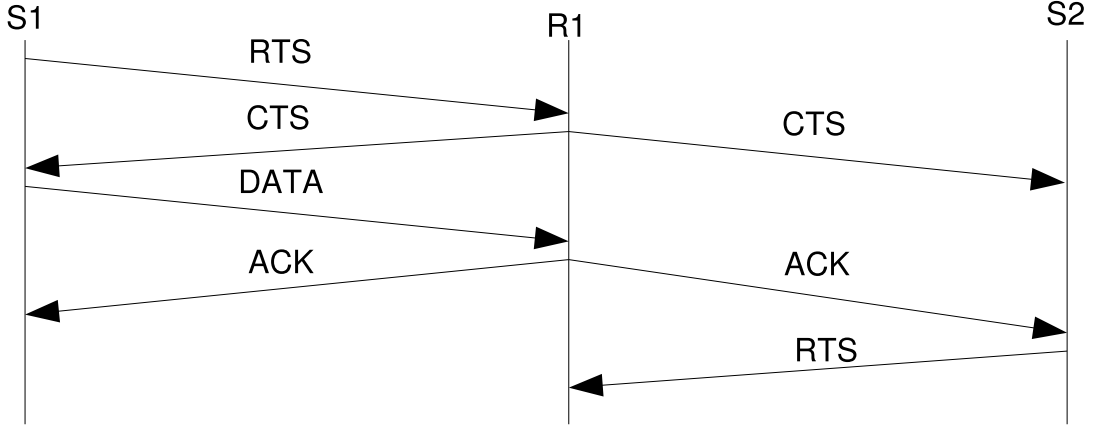


FIG. 8: RTS/CTS Handshake

2. Priority Access

A requirement of DSRC is that safety messages must have priority access over non-safety messages. In order to timely deliver high priority messages, such as those used by collision warning applications, DSRC adopts the Enhanced Distributed Coordination Function [14, 20] (EDCF) of 802.11e.

EDCF provides differential access to the wireless medium by assigning eight priority classes which are referred to as Access Categories (AC). The ACs are labeled 0 to 7, with TC 0 having the highest priority. EDCF functions similar to DCF. The primary difference is that EDCF uses a different set of access parameters for each AC. The AC parameters are used to set $CW_{min}[AC]$, $CW_{max}[AC]$, and $AIFS[AC]$. The parameters $CW_{min}[AC]$ and $CW_{max}[AC]$ control the minimum and maximum size of the contention window. Assigning larger values to the $CW_{min}[AC]$ and $CW_{max}[AC]$ for a low priority class increases the average time that a low priority class has to back-off before transmitting. On the other hand, the inter-frame spacing is used for the duration of time that a station must wait before it can begin the back-off process. EDCF makes use of the Arbitration Inter-Frame Space (AIFS) to vary the amount of time a station must remain idle before it can decrement its back-off timer. Equation 2 is used to calculate the AIFS value.

$$AIFS[i] = SIFS + AIFSN[i] * slottime \quad (2)$$

Choosing a smaller value for $AIFSN[i]$ means that the station will be able to back-off sooner and it will be able to access the channel faster. As a result of using priority access,

contention is mainly between the same AC.

There is also the question of how to implement the priority queues. The DSRC standard recommends four priority queues for the access categories. The difficulty in this is that there are a number of different channels for DSRC. One possible implementation is to have a separate set of priority queues for each channel. Another possible implementation is to have one set of queues for the services channels and another set of queues for the control channel, simplifying the implementation. In either case, the control channel is always given priority over a service channel. For example, if a frame in a service queue and a frame in control queue back-off timers expire at the same time, the frame from the control queue is transmitted. The control channel frame is always given priority over a service channel frame. Using EDCA ensures that priority traffic is given preferential treatment.

3. Multi-Channel Coordination

Coordination is necessary between the control channel and service channels. The current wireless technology allows a receiver to only listen to one channel at a time. Since both safety applications and non-safety applications will coexist in the VANET, coordination is necessary so that safety messages are not missed while a vehicle is using the service channel. For example, if a vehicle is using a service channel and an emergency warning message is sent, the vehicle will miss the warning. In this case, the vehicle may end up in an accident that would have otherwise been avoided if the warning was received. When a vehicle is using a non-safety service, a method is needed to guarantee that safety messages are still correctly received.

One solution is the use of an AP coordinated mode [19] (similar to the PCF) to coordinate channel access. Within the coordinated mode architecture, two types of access points exist: service access points and coordinating access points. First, the service access points provide non-safety services at a RSU. Second, the coordinating access points coordinates the transmissions of vehicles within its range, and the coordinating access points radio is set to the control channel. A number of different configurations are possible at a RSU. One possible implementation is to have the coordinating access point and service access point co-located at a RSU.

The coordinated mode consists of three states for a vehicle. First, the ad hoc mode is

the state a vehicle is in when not within the range of a coordinated access point, and the DSRC ad hoc protocols are used. Second, AP coordinated mode is the state a vehicle is in when under control of the coordinated access point. Third, service mode is the state that a vehicle is in when accessing a service access point. These are the three states required for multi-channel coordination.

Each coordinating access point will periodically broadcast beacons to alert vehicles that they are entering a region that contains a coordinating access point. When a vehicle approaches a region providing a non-safety service, the vehicle switches from ad hoc mode to AP coordinated mode. A repetition period is created T , which is the lower bound on the latency of a safety message. The period T is used so that the safety messages that are transmitted meet the latency requirements of the application. The repetition period T is then subdivided into a contention free period (CFP) and a contention period (CP). The start of the CFP is signaled with a CF_{start} frame and terminated with a CF_{end} frame. During the CFP, each vehicle is polled by the coordinating access point to exchange its safety message. Vehicles will not transmit during the CFP unless they are polled. After the CFP expires the vehicles switch from the AP coordinated state to the service state, and the CP begins. During the CP a vehicle may switch to a service channel and use a service provided by a service access point. The vehicles then switch back to the AP coordinated state at the expiration of the CP. When a vehicle leaves the region where a service is provided, it switches its state back to ad hoc mode. This is one possible solution to ensure that safety messages are not lost.

B. Broadcast

Broadcast messages will play a larger role than unicast messages in a vehicular environment. Some of the uses for broadcasting are to send emergency warning messages and periodically broadcast a vehicle's state. A large percentage of the messages sent in a VANET will be periodic messages that announce the state of a vehicle to its neighbors (e.g., speed, location, direction of travel, etc.). However, the 802.11 technology, on which DSRC is based, is known for not being able to manage the medium resources very efficiently, especially in case of broadcast messages. Providing reliable delivery of broadcast messages in a VANET introduces several key technical challenges.

No retransmission is possible for failed broadcast transmissions since they are undetectable. A failed unicast transmission is detected through the lack of acknowledgment (ACK) from the receiver. However, it is not practical to receive an ACK from each node for a broadcast message. If acknowledgments were used, a problem known as the ACK explosion problem [24] would exist. Each receiving node would at almost the same instance send an ACK back to the transmitting node, causing a large number of collisions.

The contention window size, CW, fails to change because there is no MAC-level recovery on broadcast frames. In order to control congestion, the contention window (CW) is exponentially increased each time a failed transmission is detected. Since there is no detection of failed broadcast transmissions, the size of the CW fails to change for broadcast traffic as it does for unicast traffic. This results in excessive collisions, if a large number of nodes are contending for access.

The hidden terminal problem exists because the RTS/CTS exchange cannot be used. The hidden terminal problem [2] is the main cause of collisions in a wireless network. The IEEE 802.11 protocols use an optional RTS/CTS handshake followed by an acknowledgment to guarantee the delivery of a unicast packet. Broadcast messages, on the other hand, cannot use the RTS/CTS because it would flood the network with traffic.

The vehicular network should support the ability to prioritize messages. When emergency warning messages are broadcast, they should be given a higher access priority than common data messages.

The collision rate of broadcast frames increases as the distance from the sender increases. Under saturated conditions, the probability of the reception of a broadcast frame sharply decreases at distance greater than 66% of the transmission range[32] . The primary reason for the decreased reception rate is the hidden terminal problem. One solution to increase the probability of reception is to implement a repetition strategy where a message is broadcast multiple times. The main drawback of repetition is it generates excessive traffic in the network.

Multi-hop broadcasts are another challenge. A naïve approach, such a flooding a broadcast frame, results in a broadcast storm [24] leading to a significant number of frames colliding and poor use of the network resources. A flooding algorithm works by each node receiving a broadcast message for the first time, then rebroadcasting the message. A message sent to n nodes results in the message being rebroadcast n times. The problem is characterized by

redundant rebroadcasts, contention, and collisions. Creating an efficient multi-hop broadcast is an open problem.

These are some of the problems associated with MAC layer broadcast for VANETs. The following protocols address some of the issues related to broadcasting MAC frames.

1. Location Based Broadcast

A repetition strategy is employed for the Location Based Broadcast [37] (LBB) that transmits a frame to all vehicles within communication range of the sender. When packets arrive, it is the receiver's responsibility to determine what action to take in terms of processing the packet based on the location from the sender and the type of message. In order to reliably deliver broadcast frames, a repetition strategy is used. Each frame has a time for which the message is useful denoted by τ . The time it takes to transmit a packet is denoted by t_{trans} . The lifetime of the message is divided into $m = \lfloor \tau/t_{trans} \rfloor$ slots. The concept of flipping an unfair coin is used to determine if a node should transmit during a time slot with $p(H) = n/m$ and $p(T) = 1 - n/m$. The packet is transmitted if a head, $p(H)$, is obtained for the time slot. If one or more packets are transmitted without a collision, then the packet is successfully received. The value n is a parameter of the protocol, and it is selected so that $n < m$. The selection of the value of n is key in the implementation of the protocol. If n is selected so that a node transmits too often, then a significant amount of bandwidth is wasted. The LBB increases the probability that a frame is successfully received but, at the same time, consumes additional resources.

2. Urban Multi-Hop Broadcast Protocol

The urban multi-hop broadcast (UMB) protocol [17] addresses the problem of transmitting multi-hop broadcast messages in areas where there is shadowing caused by large buildings. UMB protocol selects the furthest node from the transmitter to rebroadcast a message and places repeaters at intersections that rebroadcast the message in order to overcome the problem of large buildings obstructing a message's path. The goal of the protocol is to avoid collisions caused by hidden nodes, use the channel efficiently, make broadcast communication reliable, and disseminate messages in all directions at an intersection. The

protocol assumes that all vehicles are equipped with a GPS device and an electronic map. The UMB protocol is a variant of IEEE 802.11.

The first part of the protocol is the *directional broadcast* that is used to select the node farthest from the transmitter that rebroadcasts the frame. The RTS/CTS sequence of 802.11 helps alleviate the hidden terminal problem. In the case of broadcast messages, if the RTS/CTS sequence is used, as previously mentioned, a storm around the transmitter would ensue. UMB introduces an alternative to the RTS/CTS, the Request to Broadcast (RTB) and Clear to Broadcast (CTB). Only the transmitter and farthest node from the transmitter exchange the RTB/CTB messages. When a node has a broadcast message to send, it transmits a RTB.

The network is iteratively divided into segments to determine the farthest node from the broadcaster. The farthest node is determined by each node transmitting a black-burst. When a node receives a RTB, each node computes the length of the black-burst based on their distance from the sender. The length of the black-burst is computed as follows:

$$L_1 = \left\lfloor \frac{\hat{d}}{Range} * N_{max} \right\rfloor * SlotTime \quad (3)$$

where L_1 is the black-burst length of the first iteration, \hat{d} is the distance between the source and receiver, N_{max} is the number of segments, and $SlotTime$ is the length of a time-slot. Each node will then simultaneously start transmitting a black-burst. If a node finishes transmitting the black-burst and hears no others sending the black-burst on the medium, it knows that it is the farthest node, so it sends a CTB to the sender. On the other hand, if two or more nodes determine that they are the farthest away, a collision will occur when the CTB is sent. In this case, the RTB is retransmitted to the furthest non-empty segment and that segment is divided into N_{max} sub-segments. This process continues until one of the nodes CTB succeeds. The iterative black-burst is calculated as follows:

$$\begin{aligned} L_i &= \left\lfloor \frac{\hat{d} - lLongest_{i-1} * W_{i-1}}{W_{i-1}} * N_{max} \right\rfloor * SlotTime \\ i &= 2, 3, \dots, D_{max} \\ W_i &= \frac{Range}{N_{max}^i} \end{aligned} \quad (4)$$

where $lLongest$ is the longest black-burst and W_i is the segment width for the i^{th} iteration. Once a node is selected to forward the broadcast, the sender then transmits the frame to the receiver. Collisions are avoided because the surrounding nodes overhear the RTB/CTB

exchange and defer from accessing the channel. The receiver of the broadcast then sends back an ACK to indicate that the frame was successfully received. The receiver then continues the process of relaying the broadcast message.

The second part of the protocol is the *intersection broadcast* that involves relaying frames by placing repeaters at intersections. If a vehicle is in range of an intersection, as determined by the vehicle's electronic map, it sends an 802.11 unicast packet to the repeater, the RSU. The RSU will in turn relay the message in all directions except the direction from which the message was received. The sender includes directional information in the packet to prevent a RSU from rebroadcasting a message in the same direction from which it was received. The protocol also addresses the problem of loops by using a cache to determine if a packet has already been seen.

3. Adaptive Adjustment of the Contention Window

The 802.11 technology is known for not being able to manage the medium resources very efficiently, especially in case of broadcast messages. Since there is no MAC-layer recovery on broadcast frames within an 802.11-based VANET, the reception rates of broadcast messages can be very low, especially under saturation conditions.

The contention window size, CW for 802.11, has a minimum value CW_{min} that is exponentially increased by a factor of 2 each time a packet collision occurs. The size of the CW continues to increase until it reaches the maximum value, denoted as CW_{max} . Unicast transmissions in a VANET are able to adjust the contention window size to adapt to the changing conditions of the network, but this is not the case for broadcast transmissions. Because broadcast transmissions suffer from the ACK explosion problem, it is not possible to determine if a frame is successfully received or not.

On a crowded highway the number of vehicles contending for access the wireless medium is large. For instance, in a gridlocked four lane highway with vehicles placed 15 m apart, approximately 300 or more vehicles contend for channel access (e.g., 600 m diameter / (15 m between vehicles * 4 lanes * 2 directions) \approx 320 vehicles). Because a large number of vehicles are contending for access to the medium, it is necessary to vary the size of the contention window to reduce the likelihood of a collision. Vehicles can also benefit from the opposing situation where the contention window is decreased to account for light traffic.

frames received from Node B	32		34	35	36	37	38		40	41
frames received from Node C	7	8	9				13	14	15	16
frames received from Node D	15	16		18	19		21	22	23	24
frames received from Node E	62	63	64	65		67	68	69	70	71

FIG. 9: Received Frames at Node A

A node in a VANET is able to detect network congestion by simply analyzing the sequence numbers of packets it has recently received. In a VANET, each node will broadcast its status to its neighbors at least 10 times every second. While a node does not know if the packets it sent are correctly delivered or not, it knows the exact percentage of packets sent to him from neighboring nodes that are successfully received. Based on the percentage of packets that are successfully received in the last few seconds, a node is able to determine the current local conditions of the network and roughly estimate the number of neighbors in its communication range. Therefore, a node is able to dynamically adjust the parameters it uses, such as contention window size, transmission rate, and transmission power, to improve the delivery rate of broadcast messages.

In this modified implementation of 802.11, when a node sends a MAC frame, a sequence number is assigned. Each node then records the overheard sequence numbers coming from a specific node. As shown in Figure 9, node A records that it has overheard the frames coming from node B with the sequence numbers 32, 34, 35, 36, 37, 38, 40, 41. Based on the observed sequence numbers node A could conclude that frame 33 and frame 39 were corrupted or lost. Similarly, node A could conclude that three frames from node C, two frames from node D, and one frame from node E were corrupted or lost. Therefore, the percentage of frames sent to him from neighboring nodes that were corrupted in the last second is 20% (8 out of 40), and four nodes are currently in its communication range.

Prioritized access is achieved by a priority scheme similar to the one proposed in IEEE 802.11e [38]. Different levels of channel access priorities are provided through different choices of IFS and contention window size, as explained in Section VI A 2. A scheme similar

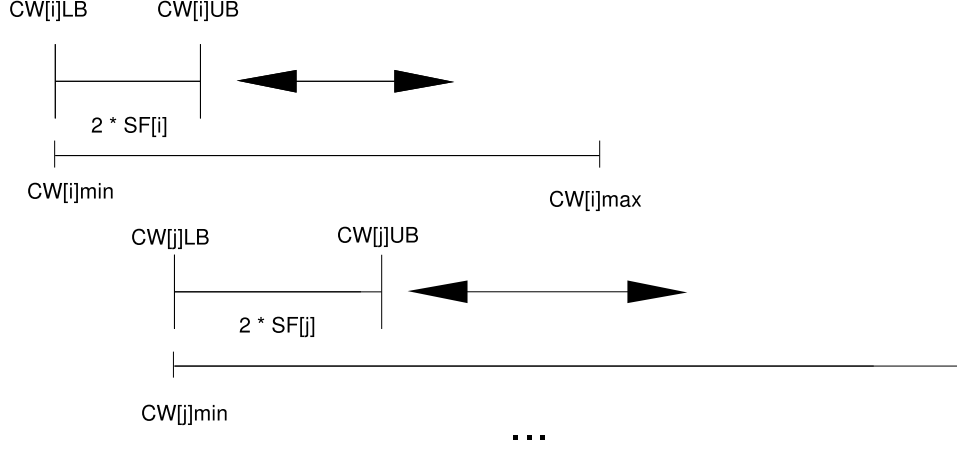


FIG. 10: Sliding Contention Window

to the Sliding Contention Window (SCW) [22] is used to dynamically adjust the CW. Each traffic class $TC[i]$ has a $CW[i]_{min}$ and $CW[i]_{max}$ which are the minimum and maximum possible values of the contention window for a traffic class. For example, $TC[0]$ could have the parameters $CW[0]_{min} = 8$ and $CW[0]_{max} = 127$, while $TC[4]$ could have the parameters $CW[4]_{min} = 64$ and $CW[4]_{max} = 1023$. The value $SF[i]$ is the scaling factor for the traffic class, which determines how much the window is slid up or slid down. $SCW[i]$ is the size of the contention window for a specific traffic class and it is set to $SCW[i] = 2 * SF[i]$. The $CW[i]$ will also contains a $CW[i]_{LB}$ and a $CW[i]_{UB}$, which are the lower bound and upper bounds of the window at any instance. The back-off that a node uses is randomly selected between $CW[i]_{LB}$ and $CW[i]_{UB}$.

$$Backoff = random() \% (CW[i]_{UB} - CW[i]_{LB} + 1) + CW[i]_{LB} \quad (5)$$

A weighted moving average is used to calculate the average reception rate. In a highly dynamic network such as a VANET, the emphasis should be placed on the most recent conditions of the network. To calculate the weighted average, an approach similar to the TCP round trip time estimation is used.

$$EstReceptionRate = \alpha * EstReceptionRate + (1 - \alpha) * SampledReceptionRate \quad (6)$$

Once the $EstReceptionRate$ is determined for each node, an average of the reception rates is used to determine the local reception rate. Once a node has determines the local reception

rate it compares the value against the previous stored local reception rate to adjust the CW that the node uses.

```
IF (average - previous average >= sliding threshold)
    Slide the window down
ELSE IF (-(average - previous average) >= sliding threshold)
    Slide the window up
ELSE
    Maintain the current window
```

Periodically each vehicle uses this algorithm to adjust the CW.

The number of collisions experienced and the number of nodes contending to access the medium determines if the current value of the contention window needs to be maintained. If a large number of collisions have occurred, $SF[i]$ is used to slide $SCW[i]$ towards $CW[i]_{max}$, as shown in Figure 10. On the other hand, if the number of collisions detected is below a threshold then $SCW[i]$ is slid toward $CW[i]_{min}$.

VII. ROUTING

Routing is the process of finding a path from a source node to a destination node. Since each node has a limited transmission range, messages often have to be forwarded by other nodes in a VANET. There are two general classes of routing protocols in ad-hoc networks: topology-based routing and location-based routing. Topology-based routing protocols use the information about the links that exists in the network to perform packet forwarding. On the other hand, location-based routing the forwarding decisions are based on a nodes location. They can be sub-divided into proactive and reactive approaches.

Proactive algorithms employ classical routing strategies such as distance-vector routing (e.g. DSDV [29]) or link-state routing (e.g. OLSR[13] and TBRPF[4]). Proactive algorithms maintain routing information about the available paths in the network even if these paths are not currently used. The main drawback of this approach is that the maintenance of unused paths may occupy a significant part of the available bandwidth if the network topology changes frequently [10].

In response to the maintenance problem, reactive routing protocols were developed. A few examples of reactive routing protocols are DSR[15], TORA [27], and AODV [28]. Reactive

routing protocols maintain only the routes that are currently in use, thereby reducing the burden on the network when only a small subset of the available routes are in use.

In location-based routing, forwarding decisions are based on the location of the forwarding node in relation to the location of the source and destination nodes. In contrast to purely topological ad-hoc routing approaches, no route set-up or route maintenance is needed with a location-based routing approach since packets are forwarded “on the fly”. Location-based routing protocols consist of location services and geographic forwarding.

Geographic forwarding takes advantage of a topological assumption that works well for wireless ad hoc networks: nodes that are physically close are likely to be close in the network topology also. Each node learns its own geographic position using a mechanism such as GPS, and periodically announces its presence, position, and velocity to its neighbors. Thus each node maintains a table of its current neighbors identities and geographic positions. When a node needs to forward a packet, it includes the destination nodes identity as well as its geographic position in the header of the packet. Each node along the forwarding path consults its neighbor table and forwards the packet toward the neighbor closest to the destination in terms of physical location, until the final destination is reached.

Although geographic forwarding works well for networks where nodes are uniformly distributed, it may not find a route to a packet’s destination when the packet has to travel around a topology “hole” – that is, when an intermediate forwarding node has no neighbors that are closer than itself to the packets destination.

For example, Figure 11 shows a mobile ad hoc network consisting of vehicles driving on the road. The source vehicle S wants to find the services available (e.g., gas stations, restaurants, etc.) within the proximity of destination D . Using geographic forwarding, packets are forwarded along the road segment $S1$. Without taking into account the road constraint, this appears to be the geographically shortest path from the source S to the destination D and seems to be the best local decision. Consequently, the packet is greedily and wrongly forwarded for potentially many hops along the road segment $S1$, before a greedy failure is recognized by Node A . This kind of dead-end (topology hole) situation is well predictable with knowledge of the road infrastructure [31].

A topology hole can also be introduced due to low vehicle density. For example, in Figure 11, there are two road paths from S to D , $S2-S3$ and $S4-S5-S6$. A naïve approach would be taking the path $S2-S3$, which has the shorter distance. However, packet forwarding

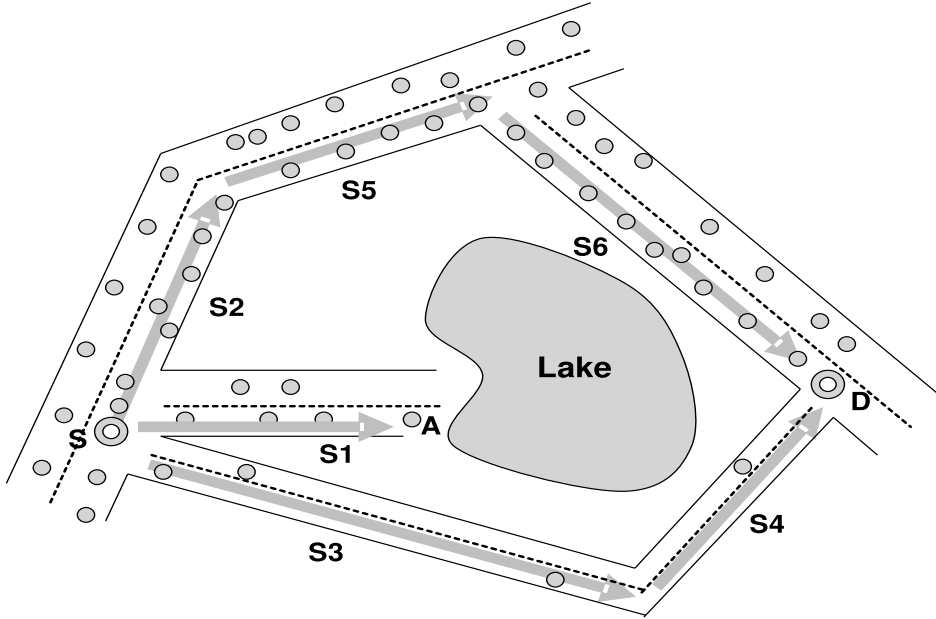


FIG. 11: Topology holes in geographic forwarding

along a road depends on the vehicle density on it. There are only a few vehicles on the road segments $S2$ and $S3$. Along the path $S2$ – $S3$, the packet will reach a vehicle B that does not know about any vehicles closer than itself to the destination. Although the packet may eventually be delivered to D by means of the vehicle movement, it will result in excessive delay. Even though the path $S4$ – $S5$ – $S6$ is longer, it has a much higher vehicle density that results in the packet usually being delivered to D in a much shorter time.

Although the topology hole problems can be solved using the planar graph face traversal method, e.g., GPSR [16], to recover from routing failures, one problem remains unsolved: since geographic forwarding is stateless, as long as a topology hole exists, each packet reaching it will initiate a routing recovery process and result in seriously degradation of the packet routing performance.

A. Context-Assisted Routing Protocol

The Context Assisted Routing (CAR) [11] protocol is considered a combination of source routing, trajectory based forwarding [25], geographic forwarding, and opportunistic forwarding [9] in the hybrid network. The key idea is to utilize the following domain specific context to assist the routing decision:

- Global context information:
 - **Road infrastructure:** CAR forwards packets along the road infrastructure. This is natural since the topology of the network matches the topology of the road infrastructure.
 - **Traffic information:** CAR favors a route in which the vehicle density is above the certain threshold to avoid partitioned networks.
 - **Roadside access points (gateways):** If available, the packet can be routed through the high speed and reliable wired network via roadside access points.
- Local context information:
 - **Vehicle location:** the forwarding decision is local and greedy, based on the vehicle location.
 - **Vehicle velocity:** the vehicle's velocity and heading is used to improve the quality of vehicle positioning, particularly, when a GPS signal is not available.
 - **Vehicle driving direction:** the packet forwarding prefers neighbors moving in the same direction. This will reduce the rate of topology change, reduce the frequency of route changes, and increase the efficiency of routing tasks.

The source node uses the global context information to compute the forwarding trajectory, which efficiently bypasses the topology holes. The intermediate nodes utilize the local context information to greedily forward the packets along the trajectory.

Along with supporting unicast packet forwarding, context assisted routing and forwarding can also be extended to support many other important network functions such as broadcasting, multicasting, multipath, discovery, and path resilience, as shown in Figure 12.

1. Spatial Model

A spatial model is constructed based on a Geographical Data File (GDF) [1] that is extracted from the topology information. Internally the spatial model is represented as a graph $G(E, V)$ consisting of a set V of *vertices* referring to the *significant places* (junction, exits) together with a set E of *edges* denoting the *road segments*, a stretch of a road between

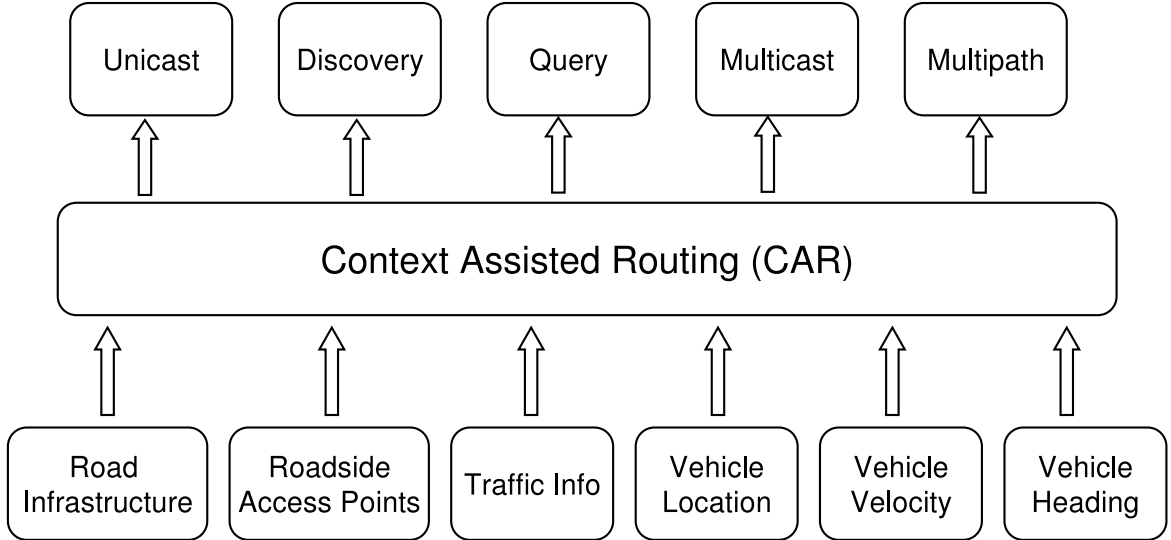


FIG. 12: Context Assisted Routing Protocols

two successive exit points (junction, exits). Hence, a vehicle moves from one location to the next, and within the graph based spatial model a vehicle moves from one vertex to the next vertex along an edge.

A tuple representing the dynamic semantic road properties is assigned to each edge, allowing the computation of dynamic shortest routes, based on different road context criteria (e.g. traveling time, distance, communication latency, etc.). The tuple consists of the *road length*, the *average speed of vehicles*, the *average number of vehicles*, and the *logical number of the route* that the road element belongs to, that can all be derived from the GDF model.

This spatial scheme allows the computation of three types of routes: *the shortest communication delay route*, *the shortest distance route*, and *the shortest traveling time route*. As a result of having three types of routes, it offers the possibility of adapting the path computation based on the application’s needs. Besides the weight of an edge, its direction represents the legal driving direction.

2. Context Assisted Routing

The Context Assisted Routing (CAR) protocol consists of Context Source Routing (CSR) and CSR-based packet forwarding. A source node computes a route path that matches the topography of the road infrastructure, and then embeds the path in the header of a packet. The intermediate nodes forward packets to nodes that lie on the path. In general, CAR is a

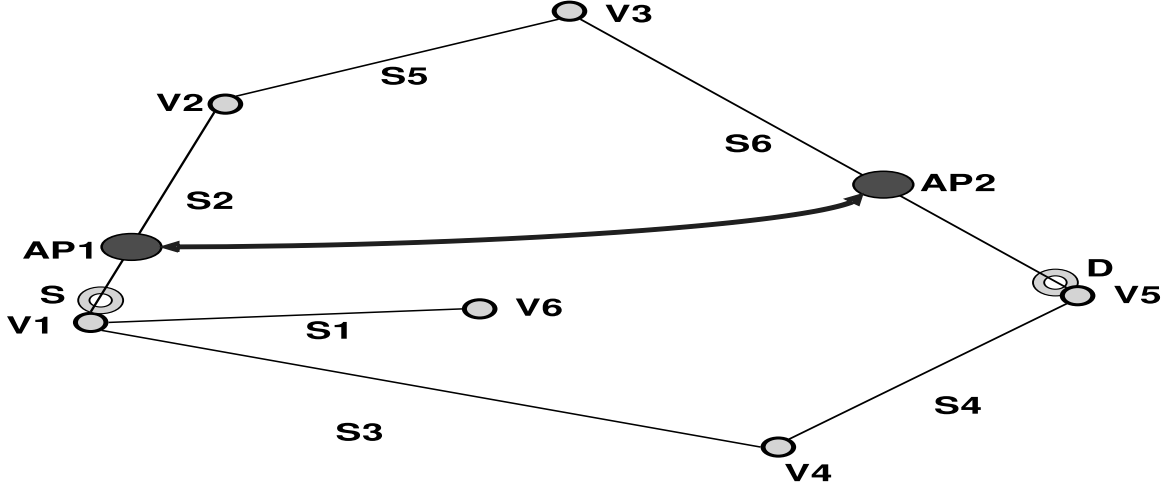


FIG. 13: Spatial Model

special case of Trajectory Based Forwarding [25], where the trajectory is specified by a list of vertices on the spatial graph.

Using the above example, an outline the major steps of CAR algorithm is given. After the deduction of the Spatial Model Graph $G(E, V)$ as shown in Figure 13, the following steps are followed:

1. source S maps itself and destination D on the graph based spatial model $G(E, V)$;
2. S calculates the shortest communication delay path to D ;
3. S sets the *Context Source Route* $CSR = \{S, V2, V3, V5\}$ to destination D ; CSR consists of a list of intermediate vertices;
4. embed the CSR in the header of all data packets from S to D ;
5. forwards the packets along the CSR path.

3. Road-Side Access Points Aware Routing

CAR does not require availability of any roadside network infrastructure. However, if available, the packet could be routed through the high speed and reliable wired network via roadside access points.

An enhanced $G(E, V)$ is maintained that includes roadside access points. The vertices added to the enhanced graph represent the roadside access points and the edges between

RSUs have a relatively lower weight compared to an edge representing a wireless link (for example, equal to one or two wireless hops) between those vertices. Therefore, when calculating the CSR, the roadside access points have a better chance of being included in the CSR. Again, using the above example, the enhanced graph is shown in Figure 13. Since the weight of the edge between $AP1$ and $AP2$ is relatively very small, the Context Source Route from S to destination D is $\{S, AP1, AP2, V5\}$. The packets from S to D go through a VPN tunnel over the wired infrastructure network. The RSUs are used as entry and exit points into and out of the more reliable wired network. In addition, the number of hops is reduced and therefore latency is also reduced.

4. *Dealing with Lossy or Intermittent GPS Reception*

In general, the location-based routing techniques assume an accurate means of localizing individual vehicles, presumably using GPS. However, GPS reception is not always available, in particular, in urban canyons and other areas (e.g., tunnels). The lack of GPS reception could be highly problematic. Fortunately, an accurate location estimate can be made based on the vehicle's velocity, heading, and previously known location.

5. *Direction-Aware Routing*

There are two special properties of highway traffic used to predict a vehicle's location. First, the traffic flow can be generalized to a line with a bidirectional traffic flow at the microscopic scale. Second, the velocity of a car is not random because each car travels along fixed set of roads.

Thus, cars traveling in the opposite direction are only briefly connected, while those moving in the same direction are connected for extended periods of time. As shown in Figure 14, the possibility of a path break is much smaller in the route $A - B - C - E$ than in the route $A - D - E$. By choosing peers in the same direction instead of the opposite direction, the changes to the topology are greatly reduced. This peer selection policy has the greatest potential of reducing the rate of topology changes, reducing the frequency of route changes, and increasing efficiency of routing.

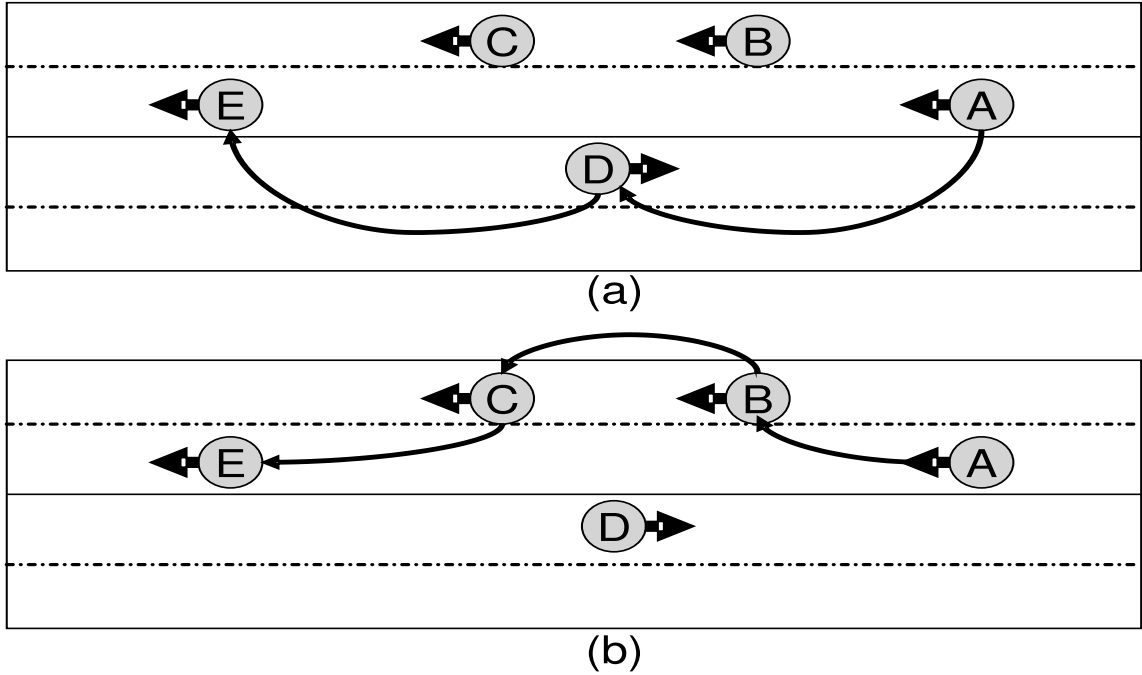


FIG. 14: Direction Aware Routing

6. Opportunity Forwarding

Due to the unavoidable slow deployment, the initial vehicular ad-hoc networks will be sparse. A VANET may experience frequent partitioning and may never have an end-to-end contemporaneous path. CAR exploits a node's movement to deliver packets opportunistically; mobile hosts exchange packets when they meet. Optimistic forwarding [9] is an opportunistic scheme for vehicular networks. It dictates that a message can have one owner at one time instant and the ownership has to be transferred from one node to another. Mobility-Centric Data Dissemination (MDDV) [35] is similar, but MDDV employs the concept of "group ownership," e.g., a group of message holders can actively propagate the message and the group membership varies with time. Optimistic forwarding is inherently more efficient and less robust than MDDV. Based on the facts that a vehicle can accurately calculate/predict the positions of neighboring vehicles and predict mobility induced errors, CAR is able to achieve efficiency similar to optimistic forwarding and at the same time exhibit robustness similar to MDDV.

B. Applications of CAR

Besides unicast routing, the CAR protocols can also be easily extended to support query and resource discovery, multipath forwarding, broadcasting, and multicasting. A brief discussion of query and resource discovery, and multipath routing follows.

1. Query and Discovery

Vehicles need traffic information several miles in advance to be able to take an alternative faster route. For this application, vehicles need to be able to send their queries to remote regions and receive timely replies. This suggests a need for the system to support geographic routing, for routing the query to the remote geographic regions.

For the query to find the best path from the source to the destination, multiple Context Source Routes (CSRs) are constructed at the source, one for each alternative path. Multiple copies of the query packet are forwarded along those paths at the same time. As the packet is forwarded towards the destination D , it will travel through all regions from the source to the destination. Each region leader then sends the region specific traffic information back to the source. The source then computes the shortest routes, the fastest routes, etc. Similarly, the Context Assisted Routing protocols are used for service discovery along the path (e.g. the closest gas station, the available parking lots, etc.). Furthermore, the discovery process can be used as a location service to find the position of an intended destination node.

2. Multipath Routing

Multipath routing is employed to increase resilience or increase bandwidth or reduce delay. Using CAR, the source may generate multiple disjoint paths. To increase bandwidth or reduce delay, each packet takes a different path, being forwarded at the same time. To improve the delivery rate and resilience, multiple copies of a packet are forwarded along the different path simultaneously.

VIII. SECURITY AND PRIVACY

Securing vehicle-to-vehicle and vehicle-to-roadside communication is an indispensable prerequisite for DSRC deployment and real world use. The system must ensure that the transmission comes from a trusted source, and that the transmission has not been tampered with. For example, with the Traffic Signal Violation Warning application, the in-vehicle system will use information communicated from the infrastructure located at traffic signals to determine if a warning should be given to the driver. An incorrect transmission from a malfunctioning, invalid or compromised unit might jeopardize the safety of the vehicle and endanger others in the vicinity. Similarly, future implementation of safety applications (such as the Approaching Emergency Vehicle Warning application) would be greatly compromised without assurance that transmissions are from an actual emergency vehicle.

Privacy and anonymity are major issues that also need to be addressed. Vehicle safety communication applications broadcast messages about a vehicles current location, speed and heading. It is desirable that users have their privacy in order to prevent their full identities from being disclosed. This is essential since *“ally consumer fears that the system might be used to build tracking mechanisms that would allow harassment, automatically issue speeding tickets, or otherwise behave in an undesirable way”* [26].

Unfortunately, unconditional anonymity may be abused; undesirable uses of anonymity include sending bogus information, denial of service attacks, and spam. Accountability hinges on the ability to attribute actions to the entity that caused those actions. A system of accountability serves a deterrent for misbehavior. In addition, individuals are often well-served by a system of accountability. The users know there will be consequences for others if their data is abused. The challenge is how to ensure anonymity and accountability at the same time. They appear conflict with each other.

In this section, we first describe potential attacks and security needs. We then discuss approaches on threat mitigation in the vehicular networks.

A. Potential Attacks

Due to the large number of independent network members and the presence of the human factor, it is highly probable that misbehavior will arise in the future vehicular networks.

Potential attacks of vehicular networks include:

- **Bogus information:** attackers send inaccurate information into the network to affect the behavior of other drivers. For example, an adversary may report false information about other parts of the vehicular networks (e.g., nonexistent traffic jams or accidents) to divert traffic from a given road and thus free the road for itself.
- **Imposture:** attackers pretend to be other vehicles by using false identities. To illustrate, a vehicle may pretend to be a police car or fire truck to issue Emergency Vehicle Approaching Warning to free the traffic.
- **Denial of service:** attackers may want to bring down the vehicular network or even cause an accident. Example attacks include channel jamming and aggressive injection of dummy messages.
- **Surveillance:** Vehicle Safety Communication technology might lead to increased surveillance of drivers engaging in everyday activities on the public roads. Potential abuses of vehicle tracking systems are rampant, including stalkers, terrorists, law enforcement tracking, automatically issued speeding tickets, or rental car agencies issuing fines for going out of state.
- **Replay legitimate messages:** a legitimate message may be intercepted and replayed at a different time and/or at different locations. For example, a vehicle may save a received Emergency Vehicle Approaching Warning and replay it later.

B. Security and Privacy Needs

In general, the vehicular network traffic must be viewed as adversarial rather than cooperative. In addition, it is essential to ensure network robustness through security protocols that work despite misbehaving participants. The future vehicular networks must assure the following properties:

- **Anonymity:** the full identity of a vehicle sending each packet/data should be kept private. The privacy principles of ITS America include an Anonymity Principle that states: “Where practicable, individuals should have the ability to utilize Intelligent

Transportation Systems on an anonymous basis.” This anonymity requirement is very important in principle. People who are concerned about tracking might disable their radio, impacting the safety and other benefits. The system also needs to reassure people that Big Brother is not in the passenger seat.

- **Authenticity:** the system must assure that the packet/data are generated by a trusted source. Privacy and anonymity might be important for our social and business well-being, but authenticity is essential for survival.
- **Integrity:** the system must assure that the packet/data has not been tampered with or altered after it was generated. Integrity is not concerned with the origin of the data who created it, when, or how - but whether it has been modified since its creation.
- **Accountability:** the system must have the ability to attribute actions to the entity that caused those actions, in case of conflict.
- **Revocation capability:** the system must have the capability to reject messages from known compromised units.
- **Real-time constraints:** the security solution must still allow low latency communication. Safety messages are very much time-sensitive. Most safety applications require latency less than 100 milliseconds.

Security is a major challenge. It is essential to make sure that life-critical information cannot be inserted or modified by an attacker. However, most security mechanism will result in significant overhead. This might seriously degrade the system capabilities in terms of latency and/or channel capacity.

Privacy is another major challenge. To ensure accountability, messages need to be uniquely signed. However, the unique signatures will allow the signer to be tracked and eventually reveal its true identity.

C. Threat Mitigation

To achieve the authenticity and integrity, messages sent should be signed and messages received should be verified. In the VANET, public-key signatures are generally more desir-

able because broadcast applications dominate and the targets of the messages are all vehicles within the vicinity.

Digital signatures provide a level of authentication for messages, ensure the integrity of messages, and maintain the original sender's accountability for their messages. In addition, all application messages should include both sequence numbers as well as timestamps. This will effectively prevent the replay of legitimate messages at different times and different places.

However, the unique signatures will disclose the identities of the senders. When vehicle safety applications broadcast messages about a vehicles current location, speed and heading with a unique signature 10 times per second, every second that a vehicle is driven, the vehicle can be easily tracked. Basically, every move a vehicle makes could be monitored.

1. Anonymity - Removing Identifying Marks

In the vehicular network, it is highly desirable that users have their privacy in order to prevent their full identities from being disclosed.

In the case of broadcast applications, the messages must not contain data that identifies the vehicle or that would allow a recipient to link messages – that is, to determine if multiple messages from dispersed locations and times have come from the same vehicle. More precisely, the chance that an attacker can link messages must drop off rapidly with the distance and time between the transmissions of the two messages. This requirement must be satisfied consistently with also requiring that messages are authenticated, in other words preventing an attacker with a radio unit from inserting messages into the system that did not actually originate from a particular vehicle.

In the case of transactional applications, the vehicle may choose to reveal its identity, or at least reveal linkable data, to a trusted respondent. However, it will not wish to reveal this data to any other entity. Therefore, all data exchanged by transactional applications must be encrypted. The encryption applied must be semantically secure, meaning that even if the same data is encrypted twice it produces two apparently unrelated ciphertexts. Secure encryption mechanisms for use with transactional applications should allow such applications to be used in an anonymous fashion.

Anonymity is difficult in vehicular networks, because so much of the information in the

messages is identifying. A vehicle could be tracked by its unique IP address, MAC address, digital signature and certificate, and account or billing information for transactional applications.

Long-lived IP addresses in theory can be used as a tracking token. However, the system is not designed for the handoff of IP sessions from one RSU to another. So, long-lived IP sessions only happen when a vehicle is stationary. All devices in the vehicular network will change their IP address when an OBU moves from one RSU communication zone to another. Therefore, it will be similar to the wired Internet. As a consequence, a vehicle is at less risk of being tracking.

Random MAC addresses are needed to ensure privacy. In computer networking a Media Access Control address (MAC address) is a unique identifier attached to most forms of networking equipment. On broadcast networks such as Ethernet and Wireless LAN, the MAC address allows each host to be uniquely identified and allows frames to be marked for specific hosts. MAC addresses are designed to be globally unique. However, it is sufficient to be unique among immediate communication neighboring hosts. To help facilitate anonymity, random MAC addresses can be used to avoid associating a particular vehicle with a particular MAC address. In vehicular networks, the immediate communication range is often within 300 meters, with a few hundreds vehicles. With a very large address space (2^{46}) and small groups, two vehicles in the same group are very unlikely to randomly pick the same MAC address at the same time. To avoid being tracked, a vehicle will need to change its MAC address very frequently, (e.g., in an order to a few minutes or less).

Digital signatures and public key certificates are attached to safety messages for the purpose of authentication and integrity, which also be used as tracking tokens. To help facilitate anonymity, one proposal suggests vehicle based units should be issued multiple digital certificates, making identification or tracking of individual vehicles more difficult. This scheme is mainly based on digital signatures under the PKI (Public Key Infrastructure). Under the PKI solution, each vehicle will be assigned a set of public/private key pair. Each message sent will contain a digital signature and a corresponding certificate. Thus, the resulting message might be three times the original message. To ensure privacy, a vehicle will have to store a large key/certificate set and frequently change keys. Each certificate contains a unique identifier, but no distinguishing information. According to the analysis by Raya and Hubaux in [18], a vehicle should change its anonymous key within an interval of

around one minute to avoid being tracked. Thus, if we assume that an average driver uses his car 2 hours per day, the number of required keys per year is approximately 43800, which amount to around 21 Mbytes. How to securely issue and store such a large number of keys will be a formidable challenge of this scheme.

Digital Cash acts much like real cash, except that it's not on paper. Money in your bank account is converted to a digital code, stored on a microchip, a pocket card, or on the hard drive of your computer, and can be used for anonymous transactions by any vendor who accepts it. Digital cash allows payment transactions (e.g., toll or parking lot fee payment) to be performed anonymously.

2. *Key Safety: Tamper-Resistant Devices*

Keys stored inside a vehicle computer can be vulnerable to use, abuse, duplication, and modification by an unauthorized attacker. To protect keys, the keys are stored in a tamper-resistant hardware device. This device offers physical protection to the keys residing inside them, thereby providing assurance that these keys have not been maliciously read or modified. In addition, the tamper-resistant device will also be responsible for verifying the access rights and signing outgoing messages.

The use of a tamper-resistant device prevents an (untrusted) member from cheating, by letting his (trusted) device both secretly store the signature keys and control their legitimate usage. The access to the contents of a tamper-resistant device requires knowledge of a PIN or password, and is restricted to only those with authorization. Furthermore, the keys should be renewed periodically.

D. **Group Signatures**

A group signature scheme allows members of a group to sign messages on behalf of the group. Signatures can be verified with respect to a single group public key, but they do not reveal the identity of the signer. Furthermore, it is not possible to decide whether two signatures have been issued by the same group member. However, there exists a designated group manager who can, in case of a later dispute, open signatures, i.e., reveal the identity of the signer.

The use of group signatures is very promising, it efficiently provides both security and privacy in the VANET. For example, if a police car issues the Approaching Emergency Vehicle Warning, all cars in its vicinity are alerted. To ensure authenticity and integrity of the messages, they are signed by a tamper-resistant chip in the police car. Using the group public key of the police cars, others can verify whether the messages are really from a police car. An additional benefit from is that the police can remain anonymous.

Groups consist of potentially thousands of members, and there is a sole manager for the group, the group manager (GM). Instead of multiple public keys for each member of the group, there exists a single group public key. Each member can produce a signature using his/her own secret signing key and the group public key [8].

The foundation of group signature schemes is in the ability of any member of the group to digitally sign a message on behalf of the group. This signature is verifiable, in that it can be verified that it came from a particular group. However, the individual within the group that signed the message is not identifiable, except by the group manager. The GM uses his/her own secret key (the group manager secret key) along with a given signature s , to determine the identity of a member of the group that generated the signature. This quality of the group manager being able to determine an individuals identity based upon the signature and the group manager's secret key is called traceability.

An entity who does not possess the group manager's secret key on the other hand, should not be able to determine the identity of a group member who signed a message s . This quality is known as anonymity. In other words, the members of the group are anonymous within the group, and are indistinguishable from other group members.

A group signature scheme consists of the following four procedures [7]:

1. **Setup**: a probabilistic interactive protocol between a designated group manager and the members of the group. Its result consists of the group's public key Y , the individual secret keys x of the group members, and a secret administration key for the group manager.
2. **Sign**: a probabilistic algorithm which, on input of a message m and a group member's secret key x , returns a signature s on m .
3. **Verify**: an algorithm which, on input of a message m , a signature s , and the group's public key Y , returns whether the signature is correct.

4. *Open*: on input of a signature s and the group manager's secret administration key this algorithm returns the identity of the group member who issued the signature s together with a proof of this fact.

It is assumed that all communications between the group members and the group manager are secure. A group signature scheme should satisfy the following properties:

- Only group members are able to correctly sign messages (unforgeability).
- It is neither possible to find out which group member signed a message (anonymity) nor to decide whether two signatures have been issued by the same group member (unlinkability).
- Group members can neither circumvent the opening of a signature nor sign on behalf of other group members; even the group manager cannot do so (security against framing attacks).

Because of the huge number of cars that generate messages, it is essential that the group signatures be short (less than 250 bytes). An appealing short group signature [6] was recently proposed. This short group signature scheme is based upon both Strong Diffie-Hellman and Linear assumptions. It utilizes signatures of length under 200 bytes, and offers about the same amount of security as an RSA signature of the same length. Group membership keys should be renewed periodically (for example, annually at the license plate renewal).

Group signatures are very promising. However, efficient group management, key certification, and key revocation issues are still challenging in the groups that consist of potentially thousands or even millions of vehicles. Some kind of multilevel hierarchical structure is essential for the efficient group management. For example, we can group vehicles based on the geographical information such as states, counties, and cities as well as vehicle types such as emergency vehicles, transit vehicles, commercial vehicles, and consumer vehicles.

IX. CONCLUSION

There are still many issues that must be addressed before vehicular networks are deployed. The standards for DSRC are still a work in progress. Currently, groups such as Vehicle Safety Communication Consortium are working on the initial prototypes of DSRC. Also,

the Crash Avoidance Metrics Partnership (CAMP) is a group of automobile manufactures, including both GM and Ford, which is currently working on the realization of the collision avoidance components of DSRC. The CAMP Intelligent Vehicle Initiative (IVI) is a research program that brings together a number of automobile manufactures and suppliers to work cooperatively with the US Department of Transportation (US DOT) [30]. The CAMP IVI includes the following four projects: Vehicle Safety Communications Project, Forward Collision Warning Requirements Project, Driver Workload Metrics Project, and Enhanced Digital Maps Project. Furthermore, IEEE 802.11p standard is not yet complete for the physical layer and MAC layer of DSRC. In the next few years, the standardization of the various aspect of DSRC will come to a close.

The technologies used for vehicular networks are still not mature and will probably not be implemented in the immediate future. The opportunities that a VANET presents are unlimited. The future introduction vehicular networks offers a tremendous opportunity to increase the safety of the transportation system and reduce traffic fatalities.

-
- [1] Geographical data file (gdf) 3.0 documentation, Oct. 1995.
 - [2] ARMSTRONG, L. Dedicated short-range communications project.
 - [3] ARTIMY, M. M., ROBERTSON, W., AND PHILLIPS, W. J. Assignment of dynamic transmission range based on estimation of vehicle density. In *Proceedings of the Second International Workshop on Vehicular Ad Hoc Networks, 2005, Cologne, Germany, September 2, 2005* (2 Sept. 2005), ACM.
 - [4] BELLUR, B., R. OGIER, AND TEMPLIN, F. Topology broadcast based on reverse-path forwarding (tbrpf), Mar. 2001.
 - [5] BLUM, J. J., ESKANDARIAN, A., AND HOFFMAN, L. J. Challenges of intervehicle ad hoc networks. *IEEE Trans. Intelligent Transportation Systems* 5, 4 (Dec. 2004), 347–351.
 - [6] BONEH, D., BOYEN, X., AND SHACHAM, H. Short group signatures. In *Advances in Cryptology - CRYPTO 2004, 24th Annual International Cryptology Conference, Santa Barbara, California, USA, August 15-19, 2004, Proceedings* (2004), M. K. Franklin, Ed., vol. 3152 of *Lecture Notes in Computer Science*, Springer, pp. 41–55.

- [7] CAMENISCH, J., AND STADLER, M. Efficient group signature schemes for large groups. In *Advances in Cryptology – Crypto ’97* (Berlin, 1997), B. S. K. Jr., Ed., Springer, pp. 410–424. Lecture Notes in Computer Science 1294.
- [8] CHAUM, D., AND VAN HEYST, E. Group signatures. In *Proceedings of Advances in Cryptology (EUROCRYPT ’91)* (Berlin, Germany, Apr. 1991), D. W. Davies, Ed., vol. 547 of *LNCS*, Springer, pp. 257–265.
- [9] CHEN, Z. D., KUNG, H. T., AND VLAH, D. Ad hoc relay wireless networks over moving vehicles on highways. In *MobiHoc* (2001), ACM, pp. 247–250.
- [10] DAS, S. R., CASTAÑEDA, R., AND YAN, J. Simulation-based performance evaluation of routing protocols for mobile ad hoc networks. *MONET* 5, 3 (2000), 179–189.
- [11] DUMITRESCU, V., AND GUO, J. Context assisted routing protocols for inter-vehicle wireless communication. In *IEEE Intelligent Vehicle Symposium (IV05)* (June 2005), pp. 594–600.
- [12] FEDERAL COMMUNICATIONS COMMISSION. Amendment of the commission’s rules regarding dedicated short-range communication service in the 5.850-5.925 ghz band, fcc 02-302. Tech. rep., FCC, November 2002.
- [13] JACQUET, P., CLAUSEN, T., LAOUITI, A., QAYYUM, A., AND VIENNOT, L. Optimized link state routing protocol for ad hoc networks, Dec. 14 2001.
- [14] JIUNN DENG, D., AND SHIUNG CHANG, R. A priority scheme for IEEE 802.11 DCF access method, Aug. 09 1999.
- [15] JOHNSON, D., AND MALTZ, D. *Mobile Computing*. Kluwer Academic Publishers, 1996, ch. Dynamic Source Routing, pp. 153–181.
- [16] KARP, B., AND KUNG, H. T. GPSR: greedy perimeter stateless routing for wireless networks. In *MOBICOM* (2000), pp. 243–254.
- [17] KORKMAZ, G., EKICI, E., ÖZGÜNER, F., AND ÖZGÜNER, Ü. Urban multi-hop broadcast protocol for inter-vehicle communication systems. In *Proceedings of the First International Workshop on Vehicular Ad Hoc Networks, 2004, Philadelphia, PA, USA, October 1, 2004* (2004), K. P. Laberteaux, R. Sengupta, C.-N. Chuah, and D. Jiang, Eds., ACM, pp. 76–85.
- [18] M. RAYA AND J. HUBAUX. The Security of Vehicular Networks. Tech. rep., EPFL Technical Report IC/2005/009, Mar. 2005.
- [19] MAK, T. K., LABERTEAUX, K. P., AND SENGUPTA, R. A multi-channel vanet providing concurrent safety and commercial services. In *Proceedings of the Second International Work-*

- shop on Vehicular Ad Hoc Networks, 2005, Cologne, Germany, September 2, 2005* (2 Sept. 2005), ACM.
- [20] MANGOLD, S., CHOI, S., KLEIN, O., AND HIERTZ, G. IEEE 802.11e wireless LAN for quality of service, Aug. 03 2002.
- [21] MATTHEW S. GAST. *802.11 Wireless Networks: The Definitive Guide*, 1st ed. O'Reily, Apr. 2002.
- [22] NAFAA, A., KSENTINI, A., MEHAOUA, A., ISHIBASHI, B., IRAQI, Y., AND BOUTABA, R. Sliding Contention Window (SCW): Towards Backoff Range-Based Service Differentiation over IEEE 802.11 Wireless LAN Networks. *IEEE Network* (Aug. 2005), 45–51.
- [23] NHTSA. Intelligent transportation systems, 2006.
- [24] NI, S.-Y., TSENG, Y.-C., CHEN, Y.-S., AND SHEU, J.-P. The broadcast storm problem in a mobile ad hoc network. In *MOBICOM* (1999), pp. 151–162.
- [25] NICULESCU, D., NATH, B., AND LAB, D. Trajectory based forwarding and its applications. Tech. rep., July 17 2002.
- [26] NTRU. Consolidated report on the requirements for public safety security in wave systems (draft 0.80), June 2004.
- [27] PARK, V. D., AND CORSON, M. S. A highly adaptive distributed routing algorithm for mobile wireless networks. In *INFOCOM* (1997), pp. 1405–1413.
- [28] PERKINS, C. E., AND BELDING-ROYER, E. M. Ad-hoc on-demand distance vector routing. In *WMCSA* (1999), IEEE Computer Society, pp. 90–100.
- [29] PERKINS, C. E., AND BHAGWAT, P. Highly dynamic destination-sequenced distance-vector routing (DSDV) for mobile computers. In *SIGCOMM* (1994), pp. 234–244.
- [30] SHULMAN, M., AND DEERING, R. K. Third Annual Report of the Crash Avoidance Metrics Partnership. Tech. rep., Crash Avoidance Metrics Partnership, Jan. 2005.
- [31] TIAN, J., HAN, L., ROTHERMEL, K., AND CSEH, C. Spatially aware packet routing for mobile ad hoc inter-vehicle radio networks. In *Proceedings of the IEEE 6th International Conference on Intelligent Transportation Systems (ITSC)* (Oct. 12, 2003), Artikel in Tagungsband, IEEE, pp. 1546–1552.
- [32] TORRENT-MORENO, M., JIANG, D., AND HARTENSTEIN, H. Broadcast reception rates and effects of priority access in 802.11-based vehicular ad-hoc networks. In *Proceedings of the First International Workshop on Vehicular Ad Hoc Networks, 2004, Philadelphia, PA, USA*,

- October 1, 2004* (2004), K. P. Laberteaux, R. Sengupta, C.-N. Chuah, and D. Jiang, Eds., ACM, pp. 10–18.
- [33] TORRENT-MORENO, M., SANTI, P., AND HARTENSTEIN, H. Fair sharing of bandwidth in vanets. In *Proceedings of the Second International Workshop on Vehicular Ad Hoc Networks, 2005, Cologne, Germany, September 2, 2005* (2 Sept. 2005), ACM.
- [34] VEHICLE SAFETY COMMUNICATIONS CONSORTIUM CONSISTING OF, BMW, DAIMLER-CHRYSLER, FORD, GM, NISSIAN, TOYOTA, AND VW. Vehicle safety communications project task 3 final report: Identify intellegent vehicle safety applications enabled by dsrc, Mar. 2005.
- [35] WU, H., FUJIMOTO, R. M., GUENSLER, R., AND HUNTER, M. MDDV: a mobility-centric data dissemination algorithm for vehicular networks. In *Vehicular Ad Hoc Networks* (2004), K. P. Laberteaux, R. Sengupta, C.-N. Chuah, and D. Jiang, Eds., ACM, pp. 47–56.
- [36] XU, Q., MAK, T., KO, J., AND SENGUPTA, R. Vehicle-to-vehicle safety messaging in DSRC. In *Proceedings of the First International Workshop on Vehicular Ad Hoc Networks, 2004, Philadelphia, PA, USA, October 1, 2004* (2004), K. P. Laberteaux, R. Sengupta, C.-N. Chuah, and D. Jiang, Eds., ACM, pp. 19–28.
- [37] XU, Q., SENGUPTA, R., AND JIANG, D. Design and analysis of highway safety communication protocol in 5.9 ghz dedicated short range communication spectrum, Spring 2003.
- [38] Y. XAIO. Enhanced DCF of 802.11e to Support QOS. *IEEE Wireless Communications and Networking* (Mar. 2003), 16–20.