Vehicular Networking: A Survey and Tutorial on Requirements, Architectures, Standards and Solutions

Georgios Karagiannis, Onur Altintas, Eylem Ekici, Geert Heijenk, Boangoat Jarupan, Kenneth Lin, and Timothy Weil

Abstract—Vehicular networking has significant potential to enable diverse applications associated with traffic safety, traffic efficiency and infotainment. In this survey and tutorial paper we introduce the basic characteristics of vehicular networks, provide an overview of applications and associated requirements, along with challenges and their proposed solutions. In addition, we provide an overview of the current and past major ITS programs and projects in the USA, Japan and Europe. Moreover, vehicular networking architectures and protocol suites employed in such programs and projects in USA, Japan and Europe are discussed.

Index Terms—Vehicular networking, V2V, V2I, SAE, IEEE 802.11p, WAVE, IEEE 1609, ISO CALM, ARIB, Intellidrive(sm), VII, SEVECOM, VSC, SAFESPOT, CVIS, SMARTWAY, ASV, ITS-Safety 2010, eSafety, COMeSafety

I. INTRODUCTION

VEHICULAR networking serves as one of the most important enabling technologies required to implement a myriad of applications related to vehicles, vehicle traffic, drivers, passengers and pedestrians. These applications are more than novelties and far-fetched goals of a group of researchers and companies. Intelligent Transportation Systems (ITS) that aim to streamline the operation of vehicles, manage vehicle traffic, assist drivers with safety and other information, along with provisioning of convenience applications for passengers are no longer confined to laboratories and test facilities of companies. Prime examples of such services include automated toll collection systems, driver assist systems and other information provisioning systems. This grassroots movement has also been backed up by coordinated efforts for standardization and formation of consortia and other governmental and industrial bodies that aim to set the guiding principles, requirements, and first takes on solutions for communication systems that primarily involve vehicles and users within vehicles.

The excitement surrounding vehicular networking is not only due to the applications or their potential benefits but also due to the challenges and scale of the solutions. Among technical challenges to be overcome, high mobility of vehicles, wide range of relative speeds between nodes, real-time nature of applications, and a multitude of system and application related requirements can be listed. Furthermore, considering ITS applications that require information to be relayed multiple hops between cars, vehicular networks are poised to become the most widely distributed and largest scale ad hoc networks. Such challenges and opportunities serve as the background of the widespread interest in vehicular networking by governmental, industrial, and academic bodies.

Between the years 2000 and 2009 several excellent survey papers have appeared in the literature in the area of vehicular networking covering topics ranging from intelligent vehicle applications to routing protocols [1], [2], [3], [4], [5], [6], [7], [8], [9], [10], [11]. This survey paper differs than the ones listed above since it provides a comprehensive overview of the state of the art applications, architectures, protocols, challenges and their solutions applied in vehicular networks. This work aims to serve as both an introduction to vehicular networking for readers of diverse technical backgrounds, and as a detailed analysis and classification of the state-of-the-art. Moving from high-level goals and objectives towards more detailed solutions, the paper is structured to lead the reader through the evolution of vehicular networking arena without losing the sight of the big picture. More specifically, starting from motivating and driving applications leading to vehicular networks, we present both concerted efforts such as standardization efforts and large projects as well as individual works mostly available in academic publications.

First, in Section II, we introduce the basic characteristics of vehicular networks and provide an overview of applications and their associated requirements as well as the challenges and solutions proposed. In Section III, standardization efforts, ITS programs, and projects are presented in their original structure, highlighting their original scope and objectives. These projects are grouped geographically (i.e., USA, Japan and Europe), reflecting their common regulatory constraints and perceived preferential emphasis on different problems. These projects are also important, as their outcomes are relevant to standardization efforts. In Japan the outcome of such projects is used during the deployment of vehicular networking infrastructures, such as the deployment of ETC (Electronic Toll Collection) infrastructure and the ongoing rollout of the infrastructure for
vehicle safety communications. In EU and USA, the outcome of these projects is mainly used for standardization efforts carried out by industry consortia, such as C2C-CC (Car 2 Car Communication Consortium) and standardization bodies. In particular, in USA the research and development activities are mainly contributing to the standardization of the IEEE 1609 protocol suite (Wireless Access for Vehicular Environments). In EU the results of such activities are contributing to the ETSI (European Telecommunications Standards Institute) ITS and ISO (International Organization for Standardization) CALM (Continuous Air-interface Long and Medium range) standardization. Moreover, in Japan such research and development activities are contributing to the ARIB (Association of Radio Industries and Businesses) and ISO CALM standardization, via the ISO TC (Technical Committee) 204 committee of Japan. Following this, Section IV is dedicated to challenges in vehicular networking environments. This detailed view of problems help set the stage for many different aspects of vehicular networking that may or may not have been covered in concerted large-scale programs. Based on this classification, we present a detailed and comparative study of existing solutions in Section V. Each of the studied challenges and their solutions are followed by a critical evaluation of existing approaches. Finally, the paper is concluded in Section VI with open research problems.

II. VEHICULAR NETWORKING APPLICATIONS AND REQUIREMENTS

This section discusses major vehicular networking applications and use cases. A use case represents the utilization of a vehicular networking application in a particular situation with a specific purpose. Moreover, this section discusses the requirements imposed by such applications on the vehicular networking architecture.

A. Applications and use cases

Vehicular networking applications can be classified as 1) Active road safety applications, 2) Traffic efficiency and management applications and 3) Infotainment applications.

1) Active road safety applications: Active road safety applications are those that are primarily employed to decrease the probability of traffic accidents and the loss of life of the occupants of vehicles [7], [12], [13], [14], [15], [16]. A significant percentage of accidents that occur every year in all parts of the world are associated with intersection, head, rear-end and lateral vehicle collisions. Active road safety applications primarily provide information and assistance to drivers to avoid such collisions with other vehicles. This can be accomplished by sharing information between vehicles and road side units which is then used to predict collisions. Such information can represent vehicle position, intersection position, speed and distance heading. Moreover, information exchange between the vehicles and the road side units is used to locate hazardous locations on roads, such as slippery sections or potholes. Some examples of active road safety applications are given below as derived from use cases described in [12], [15], [13], [16], [17], [18].

Intersection collision warning: in this use case, the risk of lateral collisions for vehicles that are approaching road intersections is detected by vehicles or road side units. This information is signaled to the approaching vehicles in order to lessen the risk of lateral collisions.

Lane change assistance: the risk of lateral collisions for vehicles that are accomplishing a lane change with blind spot for trucks is reduced.

Overtaking vehicle warning: aims to prevent collision between vehicles in an overtake situation, where one vehicle, say vehicle1 is willing to overtake a vehicle, say vehicle2, while another vehicle, say vehicle3 is already doing an overtaking maneuver on vehicle2. Collision between vehicle1 and vehicle2 is prevented when vehicle2 informs vehicle1 to stop its overtaking procedure.

Head on collision warning: the risk of a head on collision is reduced by sending early warnings to vehicles that are traveling in opposite directions. This use case is also denoted as “Do Not Pass Warning”, see [18].

Rear end collision warning: the risk of rear-end collisions for example due to a slow down or road curvature (e.g., curves, hills) is reduced. The driver of a vehicle is informed of a possible risk of rear-end collision in front.

Co-operative forward collision warning: a risk of forward collision accident is detected through the cooperation between vehicles. Such types of accidents are then avoided by using either cooperation between vehicles or through driver assistance.

Emergency vehicle warning: an active emergency vehicle, e.g., ambulance, police car, informs other vehicles in its neighborhood to free an emergency corridor. This information can be re-broadcasted in the neighborhood by other vehicles and road side units.

Pre-crash Sensing/Warning: in this use case, it is considered that a crash is unavoidable and will take place. Vehicles and the available road side units periodically share information to predict collisions. The exchanged information includes detailed position data and vehicle size and it can be used to enable an optimized usage of vehicle equipment to decrease the effect of a crash. Such equipment can be actuators, air bags, motorized seat belt pre-tensioners and extensible bumpers.

Co-operative merging assistance: vehicles involved in a junction merging maneuver negotiate and cooperate with each other and with road side units to realize this maneuver and avoid collisions.

Emergency electronic brake lights: vehicle that has to hard brake informs other vehicles, by using the cooperation of other vehicles and/or road side units, about this situation.

Wrong way driving warning: a vehicle detecting that it is driving in wrong way, e.g., forbidden heading, signals this situation to other vehicles and road side units.

Stationary vehicle warning: in this use case, any vehicle that is disabled, due to an accident, breakdown or any other reason, informs other vehicles and road side units about this situation.

Traffic condition warning: any vehicle that detects some rapid traffic evolution, informs other vehicles and road side units about this situation.
Signal violation warning: one or more road side units detect a traffic signal violation. This violation information is broadcasted by the road side unit(s) to all vehicles in the neighborhood.

Collision risk warning: a road side unit detects a risk of collision between two or more vehicles that do not have the capability to communicate. This information is broadcasted by the road side unit towards all vehicles in the neighborhood of this event.

Hazardous location notification: any vehicle or any road side unit signals to other vehicles about hazardous locations, such as an obstacle on the road, a construction work or slippery road conditions.

Control Loss Warning: in [18] an additional use case is described that is intended to enable the driver of a vehicle to generate and broadcast a control-loss event to surrounding vehicles. Upon receiving this information the surrounding vehicles determine the relevance of the event and provide a warning to the drivers, if appropriate.

2) Traffic efficiency and management applications: Traffic efficiency and management applications focus on improving the vehicle traffic flow, traffic coordination and traffic assistance and provide updated local information, maps and in general, messages of relevance bounded in space and/or time. Speed management and Co-operative navigation are two typical groups of this type of applications [13].

a) Speed management: Speed management applications aim to assist the driver to manage the speed of his/her vehicle for smooth driving and to avoid unnecessary stopping. Regulatory/contextual speed limit notification and green light optimal speed advisory are two examples of this type.

b) Co-operative navigation: This type of applications is used to increase the traffic efficiency by managing the navigation of vehicles through cooperation among vehicles and through cooperation between vehicles and road side units. Some examples of this type are traffic information and recommended itinerary provisioning, co-operative adaptive cruise control and platooning.

3) Infotainment Applications:

a) Co-operative local services: This type of applications focus on infotainment that can be obtained from locally based services such as point of interest notification, local electronic commerce and media downloading [12], [13], [16], [19].

b) Global Internet services: Focus is on data that can be obtained from global Internet services. Typical examples are Communities services, which include insurance and financial services, fleet management and parking zone management, and ITS station life cycle, which focus on software and data updates [12]. [13], [16], [19].

B. Requirements

Vehicular networking requirements are derived by studying the needs of the vehicular networking applications and use cases [12], [13], [15], [16], [19]. In this paper we use the requirements classification given in [13]. In the following, Section II.B.2, based on [13], presents a number of system performance requirements derived from the use cases given in Section II.A.

1) Classification of requirements: Vehicular network requirements can be grouped into the following classes:

a) Strategic requirements: These requirements are related to: (1) the level of vehicular network deployment, e.g., minimum penetration threshold and (2) strategies defined by governments and commissions.

b) Economical requirements: These requirements are related to economical factors, such as business value once the minimum penetration value is reached, perceived customer value of the use case, purchase cost and ongoing cost and time needed for the global return of the invested financial resources.

c) System capabilities requirements: These requirements are related to the system capabilities, which are:

Radio communication capabilities, such as (1) single hop radio communication range, (2) used radio frequency channels, (3) available bandwidth and bit rate, (4) robustness of the radio communication channel, (5) level of compensation for radio signal propagation difficulties by e.g., using road side units.

Network communication capabilities, such as (1) mode of dissemination: unicast, broadcast, multicast, geocast (broadcast only within a specified area), (2) data aggregation, (3) congestion control, (4) message priority, (5) management means for channel and connectivity realization, (6) support of IPv6 or IPv4 addressing, (7) mobility management associated with changes of point of attachment to the Internet.

Vehicle absolute positioning capabilities, such as (1) Global Navigation Satellite System (GNSS), e.g., Global Positioning System (GPS), (2) Combined positioning capabilities, e.g., combined GNSS with information provided by a local geographical map.

Other vehicle capabilities, such as (1) vehicle interfaces for sensors and radars, (2) vehicle navigation capabilities.

Vehicle communication security capabilities, such as (1) respect of privacy and anonymity, (2) integrity and confidentiality, (3) resistance to external security attacks, (4) authenticity of received data, (5) data and system integrity.

System performance requirements: These requirements are related to the system performance, which are: (1) vehicle communication performance, such as maximum latency time, frequency of updating and resending information, (2) vehicle positioning accuracy, (3) system reliability and dependability, such as radio coverage, bit error rate, black zones (zones without coverage), (4) performance of security operations, such as performance of signing and verifying messages and certificates.

e) Organizational requirements: These requirements are related to organizational activities associated with deployment, which are: (1) common and consistent naming repository and address directory for applications and use cases, (2) IPv6 or IPv4 address allocation schemes, (3) suitable organization to ensure interoperability between different Intelligent Transport Systems, (4) suitable organization to ensure the support of security requirements, (5) suitable organization to ensure the global distribution of global names and addresses in vehicles.
f) Legal requirements: These requirements are related to legal responsibilities, which are: (1) support and respect of customer’s privacy, (2) support the liability/responsibility of actors, (3) support the lawful interception.

g) Standardization and certification requirements: These requirements are related to standardization and certification, which are: (1) support of system standardization, (2) support of Intelligent Transport System station standardization, (3) support of product and service conformance testing, (4) support of system interoperability testing, (5) support of system risk management.

2) System performance requirements of some use cases: This section, based on [13], presents a number of system performance requirements derived from some use cases mentioned in Section II.A.

a) System performance requirements of “Active road safety applications”: System performance requirements of active road safety applications are given in Table I. The coverage distance associated with this type of application varies from 300 meters to 20000 meters depending on the use case [12], [13].

b) System performance requirements of “Traffic efficiency and management” applications: System performance requirements of Speed management applications are given in Table II. The coverage distance associated with this type of application varies from 300 meters to 5000 meters depending on the use case [12], [13]. System performance requirements of co-operative navigation application are given in Table III. The coverage distance associated with this type of application varies from 0 meters to 1000 meters, depending on the use case [12].

c) System performance requirements of “Co-operative local services”: System performance requirements of “co-operative local services” application is given in Table IV. The coverage distance associated with this type of application varies from 0 m to full communication range, depending on the use case [12], [13].

d) System performance requirements of “Global Internet services”: System performance requirements of “communities services” applications are given in Table V. The coverage distance varies from 0 m to full communication range, depending on the use case [12], [13]. System performance requirements of the ITS station life cycle application are given in Table VI. The coverage distance associated with this type of application varies from 0 meters to full communication range [12], [13].

III. VEHICULAR NETWORKING PROJECTS, ARCHITECTURES AND PROTOCOLS

This section discusses major vehicular networking projects, programs, architectures and protocols in the USA, Japan, Europe. These projects are presented with the objective of retaining their original scopes and structures so as to highlight their emphasis on different problems. These concerted efforts are grouped by regions mainly due to common constraints and regulations they are subject to. Within each group, standardization efforts, projects, and architectures are presented where applicable. This structure also helps identify different schools of approaches to solving ITS problems in different parts of the world.
A. ITS projects, architecture and standards in USA

Industrial, governmental and university research efforts have created significant opportunities in projects such as US IntelliDrive(sm)\(^1\), CAMP/VSC-2; CICAS, SafeTrip21, California PATH. The vehicular networking protocol standards used in such projects, except the SafeTrip21, are the WAVE protocol standards that are standardized by the IEEE in the IEEE 802.11p and IEEE 1609 protocol set. The SafeTrip21 project uses communication medium other wireless technologies than IEEE 802.11p, such as cellular technologies.

1) ITS Standardization: In 1991 the United States Congress via ISTEA (Intermodal Surface Transportation Efficiency Act) requested the creation of the IHVS (Intelligent Vehicle Highway Systems) program [23]. The goals of this program were to increase traffic safety and efficiency and reduce pollution and conserve fossil fuels while vehicles use the national road infrastructure. The U.S. Department of Transportation (DOT) got the responsibility of the IHVS program, which sought the cooperation of the ITSA (Intelligent Transportation Society of America). Currently, the research and innovation associated with DOT is administrated and managed by RITA (Research and Innovative Technology Administration). By 1996, a framework, denoted as National ITS Architecture (National Intelligent Transportation System Architecture), has been developed where IHVS services could be planned and integrated. The IHVS services are currently known as Intelligent Transportation System (ITS) [24]. National ITS Architecture supported the use of wireless communications for the implementation of many ITS services. The first ITS services, such as the automated toll collection, were using a frequency spectrum between 902 MHz and 928 MHz. This band was unfortunately too small, therefore, in 1997 the National ITS Architecture petitioned the FCC (Federal Communications Commission) for a frequency bandwidth of 75 MHz in the 5.9 GHz frequency range, having as goal the support of the DSRC (Dedicated Short-Range Communications). The allocation for the DSRC-based ITS radio spectrum was granted in 1999, which is a 75 MHz bandwidth in the 5.85 - 5.925 GHz. By 2002 the ITSA started lobbying in order to convince the FCC on matters such that DSRC licensing, service rules and possible technologies for the DSRC frequency band. In particular, it was recommended to adopt one single standard for the physical and medium access protocol layers and proposed to use the one that was specified by the ASTM (American Society for Testing and Materials), see Figure 1. This specification was specified in ASTM E2213-02 [25], based on the IEEE 802.11 [26]. FCC adopted this proposal during 2003 - 2004. The IEEE Task Group p, started in 2004, developing an amendment to the 802.11 standard to include vehicular environments, which is based on the ASTM E2213-02 specification. This amendment is currently known as IEEE 802.11p [27]. The IEEE working group 1609 started specifying the additional layers of the protocol suite. These standards are: IEEE 1609.1-resource manager [28], IEEE 1609.2-security [29], IEEE 1609.3-networking [30], IEEE 1609.4-multichannel operation [31]. The combination of IEEE 802.11p and the IEEE 1609 protocol suite is denoted as WAVE (Wireless Access in Vehicular Environments).

Another ITS standardization body that is active in the USA is the SAE (Society of Automotive Engineers) International [32], inaugurated in 1905. SAE is active in many areas. One of these areas is the SAE standardization, which in cooperation with IEEE 1609 group, is working on standardizing the message format that can be used by the IEEE 1609 protocols. An example of this is the SAE J2735 standard that is meant to be used by the IEEE 1609.3 WSMP (Wave Short Message Protocol).

2) US Federal and State ITS Projects: A comparative summary of major US ITS projects is given in Table VII. Main results and recommendations derived from some of the US ITS projects currently completed, are the following:

**IntelliDrive(sm):** Several recommendations are derived from the IntelliDrive(sm) tests that were performed in 2009, see [34], [21]:

**Communications:** The Vehicle Infrastructure Integration proof of concept (VII POC) communications systems met the basic requirements, however numerous shortcomings in the DSRC/WAVE standard were identified that mainly relate to

---

\(^1\) Source documents for this article, [21] and [22], were developed from the US DOT Vehicle Infrastructure Integration (VII) project. Subsequently, the US DOT has branded this research area as ‘IntelliDrive(sm)’ as cited in this article. At the time of publication, US DOT has replaced the VII/IntelliDrive(sm) program with the ‘Connected Vehicle Research’ program. In this paper ‘IntelliDrive(sm)’ will be used to identify VII-related project research cited in our work.

---

**TABLE IV**

<table>
<thead>
<tr>
<th>Use case</th>
<th>Communication mode</th>
<th>Minimum transmission frequency</th>
<th>Critical latency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Point of interest notification</td>
<td>Periodic, permanent broadcasting</td>
<td>1 Hz</td>
<td>&lt; 500 ms</td>
</tr>
<tr>
<td>ITS local electronic commerce</td>
<td>Full duplex comm. between road side units and vehicles</td>
<td>1 Hz</td>
<td>&lt; 500 ms</td>
</tr>
<tr>
<td>Media down-loading</td>
<td>User access to web</td>
<td>1 Hz</td>
<td>&lt; 500 ms</td>
</tr>
</tbody>
</table>

**TABLE V**

<table>
<thead>
<tr>
<th>Use case</th>
<th>Communication mode</th>
<th>Minimum transmission frequency</th>
<th>Critical latency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Insurance and financial services</td>
<td>Access to internet</td>
<td>1 Hz</td>
<td>&lt; 500 ms</td>
</tr>
<tr>
<td>Fleet management</td>
<td>Access to internet</td>
<td>1 Hz</td>
<td>&lt; 500 ms</td>
</tr>
</tbody>
</table>

**TABLE VI**

<table>
<thead>
<tr>
<th>Use case</th>
<th>Communication mode</th>
<th>Minimum transmission frequency</th>
<th>Critical latency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vehicle software/data provisioning and update</td>
<td>Access to internet</td>
<td>1 Hz</td>
<td>&lt; 500 ms</td>
</tr>
</tbody>
</table>
the dynamic nature of users and roadway environment. The specification of the protocols has not adequately considered that the transmitter and receiver are in motion relative to each other. In particular, the DSRC/WAVE standards and the resulting radio communication implementations need to be refined and should include measures such as signal quality, for UDP and IP-based two way transaction, an improved services design logic, improved management of applications and arbitration of competing services from nearby providers.

**Positioning:** Positioning functionality is required, but the specific provisioning means should not be prescribed since not all terminals may be able to include GPS positioning system for economic reasons. The position requirements must be refined and extended to take into account the variations under static and dynamic environments. Furthermore, significant work has to be done to improve position accuracy and position availability in all circumstances, meaning that GPS based and non-GPS based solutions should be investigated.

**Security:** The VII tests demonstrated that the basic security functions can be implemented and work in the context of the system. However, more work has to be performed in analyzing security threats and understand how to detect and solve such threats and attacks. Furthermore, it is recommended that the anonymous signing scheme be further analyzed, simulated and implemented. The message signing and verification strategy for the high rate messages, such as the Heartbeat messages should be refined and analyzed to accomplish an optimal blend for security and system throughput.

**Advisory Message Delivery Services (AMDS):** The AMDS performed well during the VII POC tests, but it could be improved to be more robust and more easy to use. It is recommended that the system should be improved such that it is clear how priority of messages should be interpreted in the context of other user activities. In particular, the activation criteria, e.g., which message is relevant, needs to be refined. Furthermore, the overall management of system in terms of properly setting configuration parameters and defining AMDV parameters should be refined.

**Probe Data Service (PDS):** This service was shown to work, but it was not clear if the huge amount of data from all vehicles was necessary, since under most conditions, messages sent from vehicles on the same roadway are strongly redundant. Furthermore, the rules used to prevent the availability to track a vehicle and to maintain privacy are quite complex. It is recommended that the probe data collected during the VII proof of concept be analyzed and that representative models of probe data user applications are developed to assess the mathematically requirements on vehicle density and the scope of the sampled vehicle parameters. The privacy rules used for PDS need also to be integrated in the data collection process, such that it could be understood and controlled when PDS should be used and when not.

**Vehicle Safety Communications (VSC):** The VSC consortium specified several performance requirements derived from the traffic safety applications, see [17]. From these requirements, the most significant ones are: (1) safety messages should have a maximum latency of 100 ms, (2) a generation frequency of 10 messages per second and (3) they should be able to travel for a minimum range of 150 meters.

3) ITS architecture and protocol standards: This section describes two ITS architectures.

The first ITS architecture introduced in this section is the one that is defined by US DOT and is denoted as National
ITS Architecture [35]. National ITS Architecture reflects the contribution of many members of the ITS community in USA, such as transportation practitioners, systems engineers, system developers, technology specialists, consultants. It provides a common framework that can be used by the ITS community for planning, defining and integrating ITS. This ITS architecture defines (1) the functions that are required for ITS, e.g., gather traffic information or request a route, (2) the physical entities or subsystems where these functions reside, e.g., the field, the road side unit or the vehicle, (3) the information flows and data flows that connect these functions and physical subsystems together into an integrated system. Figure 2 represents the highest level view of the transportation and communications layers of the physical architecture. The subsystems roughly correspond to physical elements of transportation management systems and are grouped into 4 classes (larger rectangles): Centers, Field, Vehicles and Travelers.

The second ITS architecture introduced in this section has been specified by the VII (now IntelliDrive(sm)) project (Figure 3).

This ITS architecture consists of the following network entities: 1) On Board Equipment (OBE), 2) Road-Side Equipment (RSE), 3) Service Delivery Node (SDN), 4) Enterprise Network Operation Center (ENOC), 5) Certificate Authority (CA).

WAVE is the protocol suite used by this architecture, (Figure 4 and Figure 5). The protocol layers used in this protocol suite are summarized below.

- IEEE 802.11p: specifies the physical and MAC features required such that IEEE 802.11 could work in a vehicular environment. 802.11p defines PLME (Physical Layer Management Entity) for physical layer management, and MLME (MAC Layer Management Entity) for MAC layer management.
- IEEE 802.2: specifies the Logical Link Control (LLC).
- IEEE 1609.4: provides multi-channel operation that has to be added to IEEE 802.11p.
- IEEE 1609.3: provides routing and addressing services required at the WAVE network layer. WSMP (WAVE Short Message Protocol) provides routing and group addressing (via the WAVE Basic Service Set (WBSS)) to traffic safety and efficiency applications. It is used on both control and service channels. The communication type supported by WSMP is broadcast.
- IEEE 1609.2: specifies the WAVE security concepts and it defines secure message formats and their processing in addition to the circumstances for using secure message exchanges.
- IEEE 1609.1: describes an application that allows the interaction of an OBE with limited computing resources and complex processing running outside the OBE, in order to give the impression that the application is running on the OBE.

B. ITS Projects, architecture and standards in Japan

In July 1996, five related government bodies jointly finalized a “Comprehensive Plan for ITS in Japan” [37], [38]. These government bodies are the National Police Agency (NPA), Ministry of International Trade and Industry (MITI), Ministry of Transport, Ministry of Posts and Telecommunications (MPT), and Ministry of Construction.
This ITS plan has been based on the “Basic Guidelines for the Promotion of an advanced Information and telecommunication Society”, which was determined by the Advanced Information and Telecommunication Society Promotion Headquarters in February 1996. The five government bodies listed above, recognized the need to develop a design that could respond to changes in social needs and development in technology in the future. In August 1999, these five government bodies jointly released a first draft of the “System Architecture for ITS”. The draft was released so as to collect opinions from the industrial and academic sectors and to actively address the information worldwide. In November 1999, the “System Architecture for ITS” has been finalized.

Currently, the main public and private organizations that influence the initialization, research, realization, and standardization of ITS in Japan are the following organizations:

- **ITS Info-communications Forum, Japan**
- **Public and Private sectors Joint Research**: MIC (Ministry of Internal Affairs and Communications), MLIT (Ministry of Land Infrastructure and Transport), NILIM (National Institute for Land and Infrastructure Management), Private corporations.
- **DSRC Forum Japan**: HIDO (Highway Industry Development Organization), ARIB (Association for Radio Industry and Businesses), JARI (Japan Automobile Research Institute), JSAE (Society of Automotive Engineers Japan), Private corporations and organizations.
- **Others**: ITS Japan, AHSRA (Advanced Cruise-Assist Highway System research Association), JAMA (Japan Automobile Manufacturers Association) ASV (Advanced Safety Vehicle), JEITA (Japan Electronics and Information Technology Industries Association)

1) **Japanese ITS Projects**: Major programs and projects in the ITS area in Japan are summarized in Table VIII. A couple of numbers, facts and results regarding these activities are as follows: By May of 2008, approximately 20 million vehicles were equipped with ETC OBUs. In particular, as of June 5, 2008, in the expressways nationwide, 74.1 % of all vehicles used ETC and on the metropolitan Expressways, 81.1 % of
all vehicles used ETC. In comparison, in March 2006, the annual distribution of VICS onboard units was approximately 3 million and in November 2007, the aggregate distribution of VICS onboard units surpassed 20 million.

**Smartway**, in contrast, supports vehicle to infrastructure communication at 5.8 GHz, combining ETC, e-payment services and VICS traffic information and warning in a single OBU. The Smartway driver warning system was successfully demonstrated in field trials on public roads in 2004 and 2005. The Smartway OBU was publicly presented in February 2006, while the Smartway driver information and warning service became operational in Summer of 2006.

**ASV (Advanced Safety Vehicle)** program is divided into four phases: ASV-1, which was conducted during 1991 to 1995, ASV-2 between 1996 to 2000, ASV-3 between 2001-2005 and ASV-4 between 2006 to 2010. ASV-1 and ASV-2 mainly focused on traffic safety and efficiency applications supported by vehicle to infrastructure communications, while ASV-3 and ASV-4 focused on the direct communication between vehicles and the infrastructure-based communication is only used for augmentation. The main purpose of ASV-3 and ASV-4 is to develop a vehicle to vehicle based driver information and warning system. The demonstration project results took place on a test track in Hokkaido in October 2005. Partial market introduction is envisaged soon.

**ITS-Safety 2010** defines the frequency bands that will be used for vehicle to vehicle, vehicle to road and for radar communication (Figure 6). In particular, one interesting point to observe in Japan is that the frequency band of 700 MHz is expected to be introduced for V2V safety applications. The frequency spectrum reallocation in Japan for UHF (Ultra High Frequencies) and VHF (Very High Frequencies) are given in Figure 7. In 2008 and 2009 verification testing on public roads has been accomplished. The start for a nation-wide deployment is planned to take place soon.

2) ITS architecture and protocol standards: In Figure 8, the ITS architecture used in the Smartway project [40], is used as an example. An On-Board Unit (OBU) provides similar functionalities as the OBE used in the USA ITS architecture. In particular, it is the processing and communication feature that is located in each vehicle and it provides the application run time environment, positioning, security, and communications functions and interfaces to other vehicles and other entities. Such entities can be central servers used by service providers that are communicating with OBUs using cellular technologies. The RSU represents the road side unit, which provides similar functionalities as the RSE used in the USA ITS architecture. The RSU is located along highways, intersections and in any location where timely communications is needed. Its main functionality is to provide communication support to OBUs via the 5.8 GHz DSRC radio communication link and to communicate with network entities, e.g., servers and car navigation systems used by the service provider and by road administrators, located far away and that are using the Internet infrastructure. Note that the DSRC communication link is synchronous and it uses as medium access, the TDMA/FDD (Time Division Multiple Access - Frequency Division Duplex), which is different then the medium access used by the IEEE 802.11p.

The protocol suite used in Japan is depicted in Figure 9. Similar to the WAVE protocol suite two types of protocol suites can be distinguished. In the left part of the protocol suite the applications are supported directly by the DSRC protocol, which is specified in the ARIB standard [41]. On the right side of the protocol suite applications are supported...
via the ASL (Application Sub-Layer), which is specified in the ARIB standard [42]. In Figure 10, an overview of the service interfaces and the protocols of the DSRC-ASL protocol suite are given. The ARIB STD-T75 is composed of three protocol layers: OSI Layer 1 provides the physical layer functionalities, OSI Layer 2 provides the data link layer functionalities and OSI Layer 7 provides the application layer functionalities. Note that if needed, layer 7 could also provide the functionalities of the OSI Layers 2, 4, 5 and 6. The ARIB STD-T88 layer provides some extension to the link layer protocol, and the network control protocol.

C. ITS Projects, architecture and standards in Europe

The scope of many European programs and projects is to provide the ability to its citizens that use European roads to benefit from improved traffic safety, reduced traffic congestion, and more environmentally friendly driving. This can be realized by providing standardized and common communication means between vehicles driving on these roads as well as between vehicles and road infrastructure.

1) ITS standardization: Three bodies are responsible for planning, development and adoption of the European standards [43]. These are: (1) the European Committee for Standardization (CEN), which is a general standardization body and is responsible for all sectors excluding the electro-technical sector, (2) the European Committee for Electro-technical Standardization (CENELEC), which is responsible for the electro-technical part of the standardization, (3) ETSI (European Telecommunications Standards Institute), which is responsible for the standardization in the telecommunications sector.

CEN is currently standardizing the European ITS DSRC 5.9 GHz radio communication technology. ETSI ITS Technical Committee (TC) has several working groups: (1) WG1, which describes the basic set of application requirements, (2) WG2, which provides the architecture specification, (3) WG3, which provides the 5.9 GHz network and transport protocols, (4) WG4, which provides the European profile investigation of 802.11p, (5) WG5, which provides the security architecture. The European standardization bodies are heavily cooperating with international standardizations, such as the ISO (International Organization for Standardization), the IEC (International Electro-technical Commission) and the ITU (International Telecommunication Union) as depicted in Figure 11.

ISO, in 1993, created the ISO/TC 204 that covers ITS activities, excluding the in-vehicle transport information and control systems, which are covered in ISO/TC 22. The ISO/TC 204 activities are performed in 16 working groups. In particular, the general communication system for all types of ITS communications is the focus of ISO/ TC 204 WG16. The protocol suite that is standardized by this working group is denoted as Continuous Air-interface Long and Medium range (CALM). CALM considers infrared communications, as well as radio systems that are following different standards and communication technologies, such as GSM, UMTS, DAB, CEN DSRC, etc. ISO/TC 204 WG 16 is closely cooperating with ETSI TC ITS.

ERTICO ITS Europe [44], is an organization that was founded at the initiative of leading members of the European Commission, Ministries of Transport and the European Industry. It represents a network of Intelligent Transport Systems and Services stakeholders in Europe. The main goal of ERTICO is to accelerate the development and deployment of ITS across Europe and beyond.
to progressively achieve the creation of technologies for the optimization of the motor vehicle of the future.

**eSafety**: The European Commission organized together with the automotive industry and other stakeholders a meeting over Safety in April 2002 and as a result of this meeting eSafety Working Group was established. Currently, eSafety [47], can be considered to be a joint initiative of the European Commission, industry and other stakeholders and it aims to accelerate the development, deployment and use of Intelligent Vehicle Safety Systems that use ICT such that the road safety is increased and the number of accidents on Europe’s roads is reduced. eSafety plays an important roll on the realization of the i2010 (Intelligent Car Initiative).

2) **ITS projects**: The European Commission research and development programs are structured in “framework programs” covering several years of broad activity with topics ranging from biology to environment. The current program is FP7 [49]. Most of the R&D activities associated with ITS are covered by the Information and Communication Technology (ICT) work in FP7. Some of the ITS projects within FP6 and FP7 are introduced in Table IX, Table X and Table XI.

Main results and recommendations derived from some of the EU ITS projects currently completed, are the following: Currently, technologies developed in SAFESPOT [50] are being verified in test beds located in six European countries, i.e., France, Germany, Italy, Netherlands, Spain and Sweden.

CVIS has developed several vehicular applications such as guidance of the fastest possible path towards the destination and emergency vehicle warning. Currently CVIS technologies and applications are being tested in test beds in seven European countries, i.e., France, Germany, Italy, Netherlands, Belgium, Sweden and the UK.

NoW [51] provided solutions for (1) position based routing and forwarding protocols, (2) adaptation of wireless LAN under realistic radio conditions, (3) fundamental questions on vehicular antennas, (4) data security in vehicular ad hoc networks, (5) secure and fast communication between vehicles.

SEVECOM provided a security architecture that is used as input for security related ETSI ITS WG5 and ISO CALM standards.
3) ITS architecture and protocol standards in Europe: The ITS ISO CALM architecture [52], [53] is shown in Figure 12. CALM is being used and is enhanced by ITS European projects, such as COMeSafety and CVIS. Figure 13 shows the European system architecture used by the COMeSafety project. Major difference with the USA and Japanese ITS architectures is that European architecture includes the ISO CALM protocol suite which provides interfaces that specify how several existing wireless technologies can be used by the upper layers. These different interfaces are:

- CALM 2G/2.5G/GPRS Cellular [54].
- CALM 3G [55].
- CALM Infra Red (IR) [56].
- CALM M5, includes IEEE 802.11p and WiFi (5 GHz) [57], [58]. Supported logical channels are control channel, service channel and auxiliary channel.
- CALM Millimetre (MM), in frequency band 62-63 GHz [59].
- CALM Mobile Wireless Broadband IEEE 802.16 / WiMax.
- CALM Mobile Wireless Broadband IEEE 802.20.
- CALM Mobile Wireless Broadband - Existing Systems.
- CALM Satellite.

The ISO CALM protocol suite architecture is shown in Figure 14. The ISO CALM first layer represents the physical and link layers, which corresponds to OSI layers 1 and 2, respectively. The second ISO CALM layer represents the network and transport layers, which corresponds to the OSI layers 3 and 4, respectively. The third ISO CALM layer...
represents the CALM services and applications layer, which corresponds to the session, presentation and application OSI layers 5 through 7.

The left part of Figure 14 shows the ISO CALM management functions [63], which reside outside the communication protocol suite. The purpose of these functionalities is to set-up and release connections between media and services. The top layer is not part of the ISO CALM protocol suite, but is shown here to emphasize that user services and applications can use the ISO CALM protocol suite via the Application Programming Interfaces (APIs).

In Figure 15, a more detailed representation of the CALM CI (Communication Interface) [61], and CALM networking layers are given. The CALM CI layer (equivalent to physical and link layers) supports different types of interfaces as described previously. The CALM networking layer can be divided in two main parts:

- **CALM IP networking and transport ([62]):** uses IPv6 mobility support protocols for Internet reachability, session continuity and seamless communications. The protocols defined in the IETF working groups NEMO and MEXT will probably be applied. UDP and optionally TCP are used on top of IPv6.

- **CALM non-IP networking and transport ([64], [65]):** Does not use the IP layer, but a new network layer is defined for the support of user applications with strict latency requirements. Instead it uses the CALM FAST network protocol for unicasting and broadcasting on a single hop basis. This protocol is currently specified by the C2C-CC. The CALM FAST protocol also provides transport layer functionalities. It uses the CALM geo-networking for unicast, broadcast, geo-unicast, geo-unicast, geo-broadcast, topo-broadcast and store and forward functionalities.

### D. Conclusions

The ITS vehicular networking standardization and research activities in USA, Europe and Japan are rapidly progressing, but they cannot be considered as completed. In Japan however, the ETC infrastructure is deployed and the rollout of the infrastructure for vehicle safety communications is ongoing. These standardization and research activities are strongly supported by the US states and European and Japanese national governments, as well as the US federal administration and the European Commission.

In USA the research and development activities are mainly contributing to the standardization of the IEEE 1609 protocol suite. In EU the results of such activities are contributed to the ETSI ITS and ISO CALM standardization, while in Japan such research and development activities are contributed to the ARIB and ISO CALM standardization, via the ISO TC 204 committee of Japan.

One of the common factors associated with the standardization activities in these parts of the world is that the IEEE 802.11p technology is targeted to be the common V2V data link technology used for traffic safety applications.

### IV. Vehicular Networking Challenges

Section II discussed several applications and use cases that make use of vehicle-to-vehicle, vehicle-to-roadside units and vehicle-to-infrastructure communication technologies. Variety of applications, ranging from infotainment applications, such as media downloading, to traffic safety applications, such as driving assistance co-operative awareness, impose diverse requirements on the supporting vehicular networking technologies. These diverse requirements lead us to a number of research challenges. This section describes these research challenges.
TABLE VII  
MAIN US ITS PROJECTS

<table>
<thead>
<tr>
<th>US projects</th>
<th>ITS projects</th>
<th>Start / End years</th>
<th>Goals</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Enabling secure wireless communication among vehicles and between vehicles and roadway infrastructure.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Design of new ITS services, where 110 use cases are identified, but only 20 were available at initial deployment of IntelliDrive(sm) system [22].</td>
<td></td>
</tr>
<tr>
<td>Vehicle Safety Communications (VSC) [17]</td>
<td>2002 / 2004</td>
<td>Development of traffic safety applications. In particular: (1) cooperative forward collision warning, (2) curve speed warning, (3) pre-crash sensing, (4) traffic signal violation warning, (5) lane-change warning, (6) emergency electronic brake light, (7) left turn assistant, (8) stop sign movement assistant.</td>
<td></td>
</tr>
<tr>
<td>Vehicle Safety Communications (VSC-A) [18]</td>
<td>2006 / 2009</td>
<td>Develop and test communication-based vehicle safety systems to determine whether vehicle positioning in combination with DSRC at 5.9 GHz can improve the autonomous vehicle-based safety systems and/or enable new communication-based safety applications.</td>
<td></td>
</tr>
<tr>
<td>CICAS (Cooperative Intersection Collision Avoidance System) [67]</td>
<td>2004 / 2009</td>
<td>Develop vehicle infrastructure cooperative systems used to address intersection crash problems, traffic sign violations, stop sign movements and unprotected signalized left turn maneuvers.</td>
<td></td>
</tr>
<tr>
<td>SafeTrip21 (Safe and Efficient Travel through Innovation and Partnership for the 21st century) [68]</td>
<td>2008 - ongoing</td>
<td>Accomplish operational tests and demonstration in order to accelerate the deployment of near-market-ready ITS technologies that have the ability and the potential to deliver safety and mobility benefits.</td>
<td></td>
</tr>
<tr>
<td>PATH (California Partners for Advanced Transit and Highways) [69]</td>
<td>1986 - ongoing</td>
<td>Collection of research projects funded by the Caltrans Division of Research and Innovation (DRI) [70] Policy and behavior research Transportation Safety Research Traffic Operation Research (1): traffic management and traveler information systems.</td>
<td></td>
</tr>
<tr>
<td>V2V communication for safety [71]</td>
<td>2009 - ongoing</td>
<td>Facilitate and help the deployment of the V2V communication based safety systems that should enhance safety across the vehicle fleet within the USA.</td>
<td></td>
</tr>
</tbody>
</table>

TABLE VIII  
MAIN JAPANESE ITS PROJECTS

<table>
<thead>
<tr>
<th>Japanese ITS projects</th>
<th>Start / End years</th>
<th>Goals</th>
</tr>
</thead>
<tbody>
<tr>
<td>ETC (Electronic Toll Collection) [72], [73], [74], [75]</td>
<td>1993 - ongoing</td>
<td>Development of a common Electronic Toll Collection system capable of both prepay and postpay systems, confirmable of usage records, which are written into IC (Integrated Circuits) cards.</td>
</tr>
<tr>
<td>VICS (Vehicle Information and Communication System) [76], [77]</td>
<td>1995 – 2003</td>
<td>Support vehicle to infrastructure communications using the communication radio at 2.5 GHz frequency range.</td>
</tr>
<tr>
<td>AHSRA (Advanced Cruise Assist Highway Systems Research Association) [78], [79]</td>
<td>1997 – 2003</td>
<td>Development of vehicle to infrastructure communication based driver information and warning system with information collection by infrastructure sensors.</td>
</tr>
<tr>
<td>Smartway [79], [80]</td>
<td>2004 / 2006</td>
<td>Reversing the negative legacy of motorization.</td>
</tr>
<tr>
<td>ITS-Safety 2010: Public-Private Co-operation program [39]</td>
<td>2006 - ongoing</td>
<td>Use millimeter wave radar system to sense the distance between vehicles or vehicle and obstacles.</td>
</tr>
</tbody>
</table>

A. Addressing and Geographical addressing

Some vehicular networking applications require that the addresses are linked to the physical position of a vehicle or to a geographic region. Mobility makes tracking and managing of “geo-addresses” extremely challenging.
**B. Risk analysis and management**

Risk analysis and management is used to identify and manage the assets, threats and potential attacks in vehicular network systems. Risk analysis and management is used to identify and manage the assets, threats and potential attacks in vehicular network systems.

**TABLE IX**

<table>
<thead>
<tr>
<th>European ITS projects</th>
<th>Start / End years</th>
<th>Goals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Communications for eSafety (COME-Safety) [83]</td>
<td>2006 – 2010</td>
<td>Co-ordination and consolidation of the research results obtained in a number of European projects and organizations and their implementation. Support of the eSafety Forum. Worldwide harmonization with activities and initiatives elsewhere. Frequency allocation, mainly for the spectrum allocation for ITS applications. Dissemination of the system properties towards all stakeholders.</td>
</tr>
<tr>
<td>SAFESPOT [50]</td>
<td>2006 – 2010</td>
<td>An FP6 IP that should develop a Safety Margin Assistant to increase the road safety, which detects in advance dangerous situations on the road and is able to extend the diver awareness of the surrounding environment in time and space. The SAFESPOT solutions should be based on vehicle to vehicle and vehicle to infrastructure communication. SAFESPOT should use safety related information provided by the communication network and the in-vehicle sensors and should be able to provide the proper warning and driving advices information to the driver.</td>
</tr>
<tr>
<td>AIDE (Adaptive Integrated DrivervEhicle interface)</td>
<td>2004 – 2008</td>
<td>FP6 IP project that had as main goal the development of an adaptive and integrated driver-vehicle interface that should be able to (1) allow a large number of individual functions, (2) maximize benefits of individual functions, (3) be safe and easy of use.</td>
</tr>
<tr>
<td>APROSYS (Advanced protection systems)</td>
<td>2004 – 2009</td>
<td>FP6 IP project that developed and introduced critical technologies that could improve passive safety for all European road users in all-relevant accident types and severities.</td>
</tr>
<tr>
<td>CVIS (Cooperative Vehicle-Infrastructure Systems) [84]</td>
<td>2006 – 2010</td>
<td>FP6 IP project that designed, developed and tested technologies needed to support vehicles to communicate with each other and with the nearby road infrastructure.</td>
</tr>
<tr>
<td>HIDENETS (Highly dependable ip-based Networks and services) [85]</td>
<td>2006 – 2008</td>
<td>FP6 STREP project that developed and analyzed end-to-end resilience solutions for distributed applications and mobility-aware services in ubiquitous communication scenarios.</td>
</tr>
</tbody>
</table>

**TABLE X**

<table>
<thead>
<tr>
<th>European ITS projects</th>
<th>Start / End years</th>
<th>Goals</th>
</tr>
</thead>
<tbody>
<tr>
<td>NoW (Network on Wheels) [51]</td>
<td>2004 – 2008</td>
<td>German project that developed communication protocols and data security algorithms for inter-vehicle ad hoc communication systems.</td>
</tr>
<tr>
<td>COOPERS (COOPerative SystEMS for Intelligent Road Safety) [87]</td>
<td>2006 – 2010</td>
<td>FP6 IP that has as main goal the enhancement the road safety by using a cooperative traffic management and direct and up to date information obtained via communication between infrastructure and motorized vehicles on a motorway section.</td>
</tr>
<tr>
<td>GeoNET [64]</td>
<td>2008 – 2012</td>
<td>FP7 IP project that develops geographic addressing and routing (geonetworking) solutions using reliable and scalable communication capabilities, which enable the exchange of information in a particular geographic area, usually located far away from the source of information. Support the deployment of IPv6 for in-vehicle onboard access and internet access to other vehicular services and applications, by combining geonetworking and IPv6.</td>
</tr>
</tbody>
</table>
| FRAME [88] | 2001 – 2004 | Enhanced the European ITS Framework architecture that was originally produced by an earlier European project, i.e., KAREN.
C. Data-centric Trust and Verification

For many vehicular applications the trustworthiness of the data is more useful than the trustworthiness of the nodes that are communicating this data. Data-centric trust and verification provides the security means to vehicular applications to ensure that the communicated information can be trusted and that the receiver can verify the integrity of the received information in order to protect the vehicular network from the in-transit traffic tampering and impersonation security threats and attacks [93]. Public key cryptosystems can be used here but the main challenge is associated with the overhead that is introduced by the use of the public key cryptosystem, see e.g., [94].

D. Anonymity, Privacy and Liability

Vehicles receiving information from other vehicles or other network entities need to be able to somehow trust the entity that generated this information. At the same time, privacy of drivers is a basic right that is protected, in many countries, by laws. Privacy can be provided using anonymous vehicle identities. One of the main challenges here is the development of a solution that is able to support the tradeoff between the authentication, privacy and liability, when the network has to (partially) disclose the communicated information and its origin to certain governmental authorities.

E. Secure Localization

Secure Localization is a Denial of Service (DoS) resilience mechanism related to the means of protecting the vehicular network against attackers that are deliberately willing to retrieve the location of vehicles.

F. Forwarding algorithms

Forwarding of packets is different than routing, where the goal of routing is to choose the best possible route to reach destination(s), whereas forwarding is concerned about how data packets are transferred from one node to another after a route is chosen.

G. Delay constraints

Data packets sent by vehicular networking applications usually have time and location significance. Primary challenge in designing vehicular communication protocols is to provide good delay performance under the constraints of vehicular speeds, unreliable connectivity, and fast topological changes.

H. Prioritization of data packets and congestion control

Data packets carrying traffic safety and traffic efficiency information usually have higher significance and therefore should be forwarded “faster” than other packets. Majority of the research activities have focused on how to provide the highest priority to the emergency type of data packets. When an emergency occurs, the channel utilization is likely to degrade due to massive broadcast of emergency messages.

I. Reliability and cross-layering between transport and network layers

Due to the wireless nature of the vehicle to vehicle communication network, a route may suddenly break. It is therefore important to provide as much reliable as possible transport service on top of the inherently unreliable network. Designing cross-layer protocols, which span between transport and routing layers, can be beneficial in vehicular networks that support real-time and multimedia applications.

<table>
<thead>
<tr>
<th>European ITS projects</th>
<th>Start / End years</th>
<th>Goals</th>
</tr>
</thead>
<tbody>
<tr>
<td>E-FRAME [89]</td>
<td>2008 - 2011</td>
<td>Further expand the European ITS Framework Architecture in order to include the support of cooperative systems and at the same time provide advice for the developmental and operational issues for a given ITS architecture.</td>
</tr>
<tr>
<td>PRE-DRIVE C2X (Preparation for driving implementation and evaluation of C2X communication technology) [90]</td>
<td>2008 - 2010</td>
<td>FP7 IP project that is establishing a pan European architecture framework for cooperative systems and is setting the road for future field operational tests on cooperative systems by answering the following questions: (1) How should a common European system look like?, (2) Which are the most promising applications?, (3) How will the system have to be implemented and deployed?</td>
</tr>
<tr>
<td>ROSATTE (Road Safety attributes exchange infrastructure in Europe) [91]</td>
<td>2008 - 2011</td>
<td>FP7 IP project that establishes an efficient and quality ensured data supply chain from public authorities to commercial map providers with regard to safety related road content.</td>
</tr>
<tr>
<td>PRECIOUSA (Privacy enabled capability in co-operative systems and safety applications) [92]</td>
<td>2008 - 2010</td>
<td>FP7 STREP that verifies whether co-operative systems can comply with future privacy regulations by demonstrating that an example vehicular based application can be endowed with technologies for suitable privacy protection of location related data.</td>
</tr>
</tbody>
</table>

Defines an approach for evaluation of co-operative systems, in terms of communication privacy and data storage.

Defines a privacy aware architecture for co-operative systems, involving suitable trust models and ontologies, a V2V privacy verifiable architecture and a V2I privacy verifiable architecture.

Defines and validates guidelines for privacy aware co-operative systems.

Investigates specific challenges for privacy.

**TABLE XI**

**MAIN EUROPEAN ITS PROJECTS (PART 3)**
V. VEHICULAR NETWORKING SOLUTIONS

This section describes solutions to the challenges described in Section IV.

A. Addressing and Geographical addressing

Packets transported within a vehicular network require particular addressing and routing features. In fixed infrastructure routing, packets are usually routed following topological prefixes and therefore cannot be adapted to follow geographical routing [95]. In [96] and [97], three families of solutions are described to integrate the concept of physical location into the current design of Internet that relies on logical addressing. These families of solutions are: (1) Application layer solutions, (2) GPS-Multicast solution. (3) Unicast IP routing extended to deal with GPS addresses. [96] specifies how GPS positioning is used for destination addresses. A GPS address could be represented by using: (1) closed polygons, such as circle (center point, radius), where, any node that lies within the defined geographic area could receive a message, (2) site-name as a geographic access path, where a message can be sent to a specific site by specifying its location in terms of real-word names such as names of site, city, township, county, state, etc.

1) Application layer solutions to addressing: The application layer solution uses an extended DNS scheme to find the geographical position. DNS (Domain Name System) is extended by including a "geographic" data base, which contains the full directory information down to the level of IP addresses of each base station and its coverage area represented as a polygon of coordinates. Four level domains are included. The first level represents the "geographic" information, the second one represents the states, the third one represents the counties and the fourth one represents polygons of geographical coordinates, or the so called points of interest. The geographic address is resolved in a similar way as the typical domain address, by using IP addresses of base stations that cover the geographic area. Two possibilities are distinguished. In the first one, a set of unicast messages is sent to the IP addresses returned by the DNS. These IP addresses correspond to the base stations located in the given geographic area. Each base station then forwards the messages to the nodes that are communicating with it, either using application layer filtering or network level filtering. In the second option, all the base stations located in the given geographic area have to join the temporary multicast group for the geographic area specified in the message. All messages that have to be sent to that given geographic area will be sent on a multicast manner using that multicast address.

2) GPS-Multicast Solution to Addressing: The GPS-Multicast solution uses the GPS Multicast Routing Scheme (GPSM). Here each partition and atom is mapped to a multicast address. An atom represents the smallest geographic area that can have a geographic address. A partition is a larger geographic area that contains a number of atoms that can also have a geographic address. A state, county, town could be represented by a partition. The main idea used by this protocol is to approximate the addressing polygon of the smallest partition, which is contained in this polygon and by using the multicast address corresponding to that partition as the IP address of that message. GPSM provides a flexible mix between application level filtering for the geographic address and multicast.

3) Unicast IP routing solution extended to deal with GPS addresses: The solutions associated with this geographic addressing type are the following:

- Geometric Routing Scheme (GEO) [96]: This routing scheme uses the polygonal geographic destination information in the GPS-cast header directly for routing. GEO routing uses a virtual network, comprised of GPS-address routers, which applies GPS addresses for routing overlayed onto the current IP internetwork.
- Geographical Positioning Extension for IPv6 (GPIPv6) [98]: This protocol is defined for distribution of geographical positioning data within IPv6. GPIPv6 requires the specification of two new option types for IPv6. These options are GPIPv6 source and GPIPv6 destination, which consist of signaling the geographical positions of the source and destination, respectively.
- Using unicast prefixes to target multicast group members [99]: In [99] an extension to IPv6 multicast architecture is described that allows for unicast-prefix-based allocation of multicast addresses. Using this specification unicast prefixes could be used to target multicast group members located within a geographic area.
- Conclusions: Three geographical addressing families can be identified: Application layer, GPS-multicast and Unicast IP routing extended to deal with GPS addresses. The most promising, but also the most complex one is the family that extends IP routing and IP addressing in order to cope with GPS addresses. While several solutions associated with this family have been proposed, more research and standardization activities are needed for a successful realization.

B. Risk analysis and management

Risk analysis in vehicular networks has not yet been studied extensively. One frequently cited paper on attacker capabilities in vehicular networks is [100], which describes the work accomplished in the German project Network on Wheels (NoW) [51]. The security model used in NoW is flexible, allowing to integrate previously found attacks into the studied attack model. This model studies four major attack aspects:

- Attacks on the wireless interface;
- Attacks on the hardware and software running on OBUs and RSUs;
- Attacks on the sensor inputs to different processing units in vehicles;
- Attacks on security infrastructure behind wireless access networks, such as vehicle manufacturers, certification authorities, traffic authorities, etc.

In [101] two procedures are identified to enhance the overall security: 1) perform local plausibility checks, such as comparing the received information to internal sensor
data and evaluating the received information from different sources about a single event; 2) do regular checks on the nodes, most notably RSUs.

1) Conclusions: Risk analysis and management have been researched on a small scale. From the performed studies in this area it can be concluded that position forging attacks constitute a major vulnerability of the system. More work is needed in the area of risk management in order to cope with this vulnerability.

C. Data-centric Trust and Verification

In [102], security concepts that can be used to support the data trust and verification are categorized into proactive security and reactive security concepts.

1) Proactive security concepts: The proactive security concepts can be currently, considered as the most promising candidates for traffic safety applications in vehicular networks. This type of security solutions can be further divided into three classes, i.e., Digitally Signed Messages, Proprietary System design, and Tamper Resistant Hardware.

- Digitally Signed Messages come in two flavors: Digitally Signed Messages Without Certificates and With Certificates. The first solution is much simpler to deploy and use, while the latter provides a more secure communication, but is much more complex. Similar solutions can be found in [103], [104], [105], [106], [17], [107], [108], [109].

- Proprietary System design comprises Non-public Protocols, and Customized Hardware. The former uses non-public protocols to realize access restrictions to nodes that are not using these protocols. The latter uses customized hardware in order to achieve the same goal.

Note however, that these solutions do not prevent an attacker from doing any harm, but they aim on raising the needed effort an attacker has to spend in order to enter into the system.

- Tamper Resistant Hardware is the third proactive security context class. Even when securing the external communication part of an application, it is not possible to guarantee that the in-vehicle system is free from the generation of e.g., unnecessary accident warnings. A solution to this problem is to use tamper-resistant hardware for the in-vehicle devices. Some examples can be found in [110], [111].

2) Reactive security concepts: The reactive security concepts consist of Signature-based, Anomaly-based and Context-based approaches. The main characteristic of such systems is that they correlate the received information with information that is either already available into the system from observations on normal system operation or which is introduced additionally, see [112], [113], [114], [115].

- Signature-based: intrusion detection is comparing network traffic to known signatures of attacks to detect an attack on the system.
- Anomaly-based: intrusion detection compares the received information with the one derived from the normal operation behavior. This solution requires that the definition of the normal communication system behavior is available.

- Context verification: is an approach used by each vehicle to collect information from any information source available in its neighborhood in order to create an independent view of its current status and the current surroundings environment. When the vehicle receives data, it compares the parameters related to status and environment, e.g., position, with its own estimated information regarding status and environments to detect an intrusion. Three Context verification types are identified:
  - The Position Information verification aims to prevent an attacker from pretending to be at arbitrary positions.
  - The Time verification solution correlates the vehicle’s internal clock, which is synchronized and updated using information provided by GPS, with the time data fields of the received messages, and in particular of the beacon messages.
  - Application Context Dependent verification solution correlates the application context with a similar application context that is known to a vehicle. This solution can be realized if it is assumed that for every application, there is a set of constraints in a realistic scenario where the application can generate and deliver e.g., accident warning messages. The solutions presented in [114], [116] and [117] can be considered as being Application Context Dependent verification solutions.

3) Conclusions: The data centric trust and verification solutions can be categorized in proactive and reactive. The proactive security concept has been researched extensively. However, the tamper-resistance-hardware used in a vehicle to e.g., detect unnecessary accident warnings, needs to be further researched. The reactive security concepts have been studied in a smaller scale. More work is needed in the area of context verification, where a vehicle is able to realize an intrusion detection system by comparing received information on parameters associated with status and environment with its own available information.

D. Anonymity, privacy and liability

1) Anonymity and privacy: Pseudonyms should ensure that cryptographically protected messages should not allow for their sender to be identified. Furthermore, it should be difficult that two or more messages generated by the same node should be linked together.

a) Linkability between pseudonyms: Vulnerability to a movement tracking attack is a possible issue associated with pseudonyms. However, if the lifetime of a public key is several minutes and different vehicles update their public keys at different times, then situations can be observed where consecutive messages can be connected and thus the whole movement of a vehicle can be traced. Two solutions can be identified that reduce the movement tracking attack:

- Silent period: of [118], is proposed to reduce the linkability between the pseudonyms, or to create groups and
guarantee that the vehicles in a group cannot listen to messages sent by vehicles from another group.

- Mix Zone: of [119], is a concept, where all vehicles within a Mix-Zone share the same secret key, which is provided by an RSU located in the same Mix-Zone. Public keys are changed when the vehicles go out of the Mix-Zone. This way, the location privacy is protected.

b) Anonymity and adaptive privacy: The adaptive privacy and anonymity concept is introduced in [120] and [121]. In particular, it is argued that privacy is a user-specific concept, and a good mechanism should allow users to select the privacy that they wish to have. A higher level of privacy requirement usually results in an increased communication and computational overhead. Users may want to use different levels of privacy depending whether they are communicating with a public or a private server. The trust policies include, full trust in which the user trust both types of servers, the partial-trust in which the users trust only one type of servers and zero-trust in which users trust neither type of servers. This algorithm assumes that the zero-trust model is used. It uses a group-based anonymous authentication protocol that can trade off the computational and communication overhead with the privacy degree. By using this group-based protocol, the authentication requester only needs to be verified by only verifying that he is a member of a group. All the users are treated the same.

Another concept used to provide pseudonymity is described in [122]. A multi-layer addressing is provided, which is able to support user privacy at different levels by providing pseudonymity at the different levels. Furthermore, it provides packet forwarding schemes that use pseudonym caching. Moreover, a location service is introduced that is able to periodically change pseudonyms thus enabling unicast communications.

c) Liability: The Liability challenge is mentioned in many papers [93], but no solutions are provided. One of the anonymity solutions that has been mentioned in [120] could satisfy the liability by adapting the privacy degree of the user.

2) Conclusions: Anonymity and privacy are being extensively investigated. However, an open area is anonymity and adaptive privacy, where users are allowed to select the privacy that they wish to have. Effective liability solutions are not yet provided. A significant work in this area is necessary.

E. Secure Localization

Several solutions have been proposed for secure localization in the literature.

Tamper-proof GPS [123] proposes a system, where each vehicle has a tamper-proof GPS receiver, which can register its location at all times and provides this information to other nodes in the network in an authentic manner. The main problem with this solution is that its availability is limited in urban environments, e.g., GPS reception problems on bridges, or in tunnels. Furthermore, GPS-based systems are vulnerable to several types of attacks, such as blocking, spoofing and physical attacks.

With verifiable multilateration [123], the verification of the vehicle location is accomplished using the roadside infrastructure and by using multilateration and distance bounding. Distance bounding is used to ensure that the distance between some nodes is not higher than some value. Multilateration means that the same operation is used in several dimensions. One such operation involves the use of verifiers to establish positions. In [124], a challenge-and-response-based solution is proposed that involves verifiers. Verifier nodes are placed at special locations defining an acceptable distance for each verifier. Given a set of overlapping circles of radius R, verifiers are distributed over each such circle. The verifier requests from a node to send its position. Afterwards, the verifier sends a challenge to the node via the communication radio link. The node that receives the challenge has to reply via ultrasound. If the answer arrives in a certain time then the verifier can deduce that the node is within the region R.

Another challenge-response system involves the use of logic reception of beacons [125], which involves synchronized acceptor and rejector nodes. Acceptor nodes are distributed over region R, while rejectors form a closed annulus around the receptors. If a node sends a beacon, then the first verifier that receives the message decides whether the transmitted position by the transmitting node is acceptable. If the transmitted information first reaches a rejector then the transmitting node cannot be located with the region R. If it first reaches an acceptor, the transmitting vehicle is approved to be located within region R.

In [126], [127] the concept of “Position Cheating Detection System” is introduced. In this scheme suitable sensors are used to detect cheated position information. Two classes of position verification sensors are used. With Autonomous Sensors, sensor results contribute to the overall trust ratings of neighboring nodes independently. With Collaborative Sensors, sensors collaborate with other nodes surrounding the monitored neighbor node. In both cases, sensors only use the information provided by the routing layer. Furthermore, no additional dedicated infrastructure is needed, since only VANET nodes are included.

Another solution is based on plausibility checks. Two examples of plausibility checks are SLV and the solution developed for PBR (Position Based Routing). In [128], Secure Location Verification (SLV) is proposed to detect and prevent position-spoofing attacks. This is accomplished by using distance bounding, plausibility checks and ellipse-based location estimation to verify the position claimed by a vehicle. On the other hand, a secure localization solution is developed for PBR [129]. PBR is considered and evaluated by the CRC-CC and it provides a scalable and efficient unicast forwarding in large-scale vehicular networks. PBR is based on three features: beaconing, location service and forwarding. The location service is used when an originating node needs to know the position of another node that is not included in its location table. When this happens the originator node sends a location query message that includes the node ID, sequence number and hop limit. Nodes that receive the message and are not the ones that are searched, are rebroadcasting the location query message. When the searched node receives the location query message it replies with a location reply message carrying its current position and a time stamp. When the originating node receives the location reply message it updates its location table. In order to secure
the PBR messages, i.e., location query and location reply, each received message has to pass certain plausibility checks by using the packet’s time and location fields as inputs. If the plausibility checks fail then the message is discarded. Otherwise, the verification of the message continues. First the certificate is validated unless it was previously validated. If this verification passes then the digital signature of the message is verified. If all these verification steps pass, then the message is further processed, otherwise the message is discarded.

1) Conclusions: Secure localization can be considered as an efficient solution for the DOS attacks associated with localization. A number of solutions have been found and briefly described in this section. However, more work in the area of tamper-proof GPS and on the use of plausibility checks to prevent position spoofing attacks is needed.

F. Forwarding algorithms

The multi-hop communication between source and destination can be performed in either V2V, V2I, or hybrid fashion. Messages are forwarded to a destination by making use of multiple intermediate vehicles as relay nodes. We now describe the forwarding solutions by organizing them into two main categories. The first category focuses on unicast routing, and the second addresses broadcast routing.

1) Forwarding for Unicast Routing: There exist a number of routing algorithms proposed for VANETs. The protocols developed for unicast communication can be divided into three sub-categories: geographic, link stability-based, and trajectory-based. A brief overview of these protocols can be found in Table XII.

Geographic: Most algorithms in this category are inspired from the popular Greedy Perimeter Stateless Routing (GPSR) [130]. All forwarding decisions are taken mainly using the location information of vehicle’s immediate neighbors. Such algorithms are especially required for VANETs due to high vehicular mobility. Since each vehicle maintains only local information, these methods can scale to networks with large number of vehicles. The vehicles are assumed to be equipped with GPS or other location services so that they can determine their own location without incurring any overhead. The set of all neighbors and their respective locations are discovered using periodic beacon messages that are exchanged among nearby nodes. Subsequently, all forwarding decisions follow a greedy method where the neighbor that is geographically closer to the destination is selected as the next forwarding node. Since the vehicles do not have global knowledge of the network topology, the forwarding decisions are often locally optimal and may not be globally optimal. As a result, these protocols often encounter cases where a vehicle does not find the next forwarding node (e.g., due to dead-ends or network disconnection). To solve the local optimum problem, GPSR proposed the perimeter forwarding algorithm or the right-hand rule. This solution is not suitable for road networks, especially in urban environments where there are multiple intersections and paths. Many existing geographic routing protocols solve this problem by employing a new path recovery mechanism. Examples of such protocols are GSR [131], GPCR [132], GPSRJ+ [133].

GPSR [131] addresses various problems involved in applying position-based GPSR to city environments. It specifically targets the following issues: network disconnection due to radio obstacles; too many hops; and loopy paths. GSR discovers the current position of a desired communication partner by using a reactive location service, where the query node floods the network with a “position request” packet for some specific node identifier. Whenever the node with requested identifier receives a request, it sends a “position reply”. With this information, the sending node can compute a Dijkstra shortest path to the destination by using the underlying map of the streets. The reverse path contains sequence of junctions that packets can reach the destination.

To compute shortest paths in GSR protocol, vehicles must have access to a global street map. When such a map is not available, a protocol called GPCR [132] can be employed. Unlike GSR that relies on source routes information, GPCR makes a decision at each junction about the street the packet should follow next; and in between junctions, GPCR uses greedy forwarding.

The key challenges, since a street map is not available, are to identify nodes that are at the junctions and to avoid missing intersections while greedy forwarding is being used. One way to deal with this problem is by making the nodes at intersections (junction nodes) send special type of notification messages so that surrounding nodes can make their forwarding decisions.

The limitations of GPCR are highlighted by GPSRJ+ [133], which argues that GPCR incurs additional delay and overhead since forwarding decisions and recovery process depend primarily on the nodes at the intersections. GPSRJ+ makes an observation that all packets need not be forwarded via junction nodes. The sender vehicles forward information to neighbors of junction nodes instead of junction nodes themselves. This greatly reduces the packet load around intersections. However, this solution requires additional information exchange between the junction nodes and their neighbors.

The protocols that we discussed so far do not handle sparse networks. Low vehicle density of sparse networks causes intermittent network connectivity and routing failures. This problem is typically addressed by employing a technique similar to store-and-forward or data mulling. Packets are temporarily stored at moving nodes while waiting for opportunities to forward those [134]. Such techniques however are targeted only for delay tolerant applications.

Store-and-forward techniques have also been applied to urban road networks where different streets have different vehicle densities. VADD [135] attempts to address this problem by routing the packets via road segments with high vehicle density. However, such high density paths may not be the best ones in terms of delay. Since all vehicles try to route packets via high vehicle density streets, the channel utility along these streets may increase, and as a result, packets may either get dropped or incur high delays. VADD makes use of a delay estimation model to select paths with minimum delays. The model however relies on preloaded street map and traffic
statistics such as vehicle speed and traffic density at different time of the day.

D-Greedy and D-MinCost [136] are other protocols that employ the store-and-forward technique to route packets in urban networks. Unlike VADD, these protocols aim to provide bounded transmission delay while minimizing the bandwidth utilization. D-MinCost requires knowledge of global traffic conditions while D-Greedy does not require this knowledge. In D-Greedy, a source node uses geographical location of the destination to estimate the length of the shortest path. The chosen path is allocated a delay budget that is proportional to the street length of each road segment. Since each node does not have global information, D-Greedy assumes that the message delay budget can be uniformly distributed among the intersections that are part of the shortest path. Each relay node makes routing decisions based on the remaining message delay budget. The relay node is allowed to carry the packets to the next intersection as long as the time that the vehicle takes to reach the intersection is within the allocated delay budget. This mechanism reduces the number packet transmissions while guaranteeing bounded transmission delays. D-MinCost improves upon D-Greedy by incorporating additional factors such as vehicle density into the path selection process.

The problem of sending packets via non-optimal routes in VADD is addressed by SADV [137]. The main idea is to avoid non-optimal routes and reduce the packet delay by deploying static nodes at the intersections. Authors proposed two forwarding systems: in-road forwarding and intersection forwarding. In-road forwarding refers to directional packet forwarding mechanism where packets are sent greedily along the road. Around the intersections, the packets are sent to static nodes, which compute minimum delay paths based on vehicle densities on different road segments. These static nodes store the packets until they find vehicles along computed minimum delay paths.

In general, the store-and-forward technique used in VADD, D-Greedy, D-MinCost, and SADV is targeted at sparse networks, and hence it is suitable only for delay-tolerant applications. This technique typically requires vehicles to have larger buffers in order to minimize the number of packet drops.

The minimum delay estimation in these protocols is mainly based on vehicle-level information such as average speed and density. Such information alone, however, is not sufficient to find delay-optimal paths. A packet may experience small delays on some road segments but longer delays on others. Therefore, one must also consider amount of data traffic along different streets. Data-traffic characteristic can be incorporated into routing protocols by following a cross-layer design philosophy.

An example of such a protocol is PROMPT [138]. It is a position-based cross-layer data delivery protocol that uses the real-time packet traffic statistics to deliver packets along minimum delay paths. As base stations broadcast beacon messages across the network, beacons are updated with several network traffic statistics that are collected at the locations from which the beacons are reforwarded. The receivers of a beacon message can therefore construct the entire path to the base station including the data traffic statistics along the path. These statistics are later used by a sophisticated delay estimation model through which vehicles determine the minimum delay path to base stations. Furthermore, PROMPT takes advantage of digital roadmaps in mapping communication paths with positional information (obtained from beacons) to source routes along physical roads. Such a mapping is very important in the context of vehicular mobility. Given such a source route by the network layer, the MAC function determines individual relay nodes based on their locations and forward the packets towards the destination. Since beacons are sent out periodically, vehicles can always choose the delay-optimal routes as they are constantly made aware of real-time data traffic conditions along different streets. Furthermore, PROMPT can handle network sparsity issues by making the intermediate relay nodes hold the packets until suitable next forwarding nodes are found. Such a mechanism can be implemented in the MAC function.

**Link Stability-based:** Topology-based routing protocols (e.g., reactive and proactive routing), which are popular in MANETs can be applied to vehicular networks. However, the main issue for VANETs is that the overhead incurred in path discovery and path maintenance can be significantly high due to high mobility. Such protocols are mainly deployed in highway environments and small scale networks where number of hops between source and destination is small. To improve the link stability and reduce the path recovery overhead, one can also exploit mobility information to predict how long a given path will last and find a new path before the link breakage occurs.

MOvement Prediction based Routing (MOPR) [139] is a protocol that aims to improve the reactive routing process by leveraging vehicle information such as its position, speed, and direction. It estimates the lifetime of a link by predicting the future positions of vehicles involved in the link based on their current position. The source node can therefore estimate the transmission time and thereby decide upon the most stable path. During the route discovery process, MOPR specifically searches for intermediate nodes that have similar speed and direction to both source and destination. A route table that includes the position, speed, direction, street information of all neighboring vehicles is maintained by MOPR. This table is used while searching for paths with most stability. Similar techniques are implemented in proactive routing [140].

Velocity-Heading based Routing Protocol (VHRP) [141] uses vehicle headings to predict route disruption before it actually happens. Here, the vehicles are grouped according to their velocity vectors. Routes involving vehicles from same group exhibits high level of stability. Whenever a vehicle shifts to a different group, routes involving that vehicle may potentially get disrupted. To avoid such a problem, VHRP periodically sends route update message and maintains route table and vehicle groups. VHRP is particularly suitable for proactive routing protocols such as DSDV, and it can improve end to end throughput performance. Similarly, Prediction-based routing (PBR) [142] protocol makes use of a mobility model to characterize the collective motion of vehicles on a highway. PBR uses mobility model to predict route lifetimes and preemptively creates new routes before existing routes fail.

**Trajectory-based:** Trajectory-based Forwarding (TBF) [143] algorithm is a novel combination of source routing
and Cartesian (position) forwarding for ad hoc networks. The source node selects the route or trajectory to the destination. Unlike traditional source routing, the forwarding decisions in TBF are based on the relationship to the trajectory rather than ID of intermediate nodes. It essentially decouples path naming from the actual path. The framework of TBF can be used for any type of services including unicast, broadcast, multicast, multipath routing etc. The next hop node (or relay node) is chosen based on the distance between candidate relay nodes and the trajectory. However, TBF discovers trajectories using a flooding method, which causes additional overhead.

The issue of flooding overhead is addressed by TBD [143], which is a data trajectory-based forwarding scheme for low density road networks. TBD makes use of a local delay model to compute the expected data delivery delay from individual vehicles to an access point. The vehicle with the shortest expected data delivery delay is selected as the next relay node. Unlike TBF, TBD is not a source routing mechanism as it does not assume that there exists a path between source and destination. It allows intermediate nodes to make routing decisions based on its trajectory and neighbor information. The delay is estimated from the contact probability and the forwarding probability at the intersections. To obtain the path information, trajectory-based protocols are equipped with digital map and GPS system.

MDDV [144] is a mobility-centric approach for data dissemination in vehicular networks. It is designed to operate efficiently and reliably despite the highly mobile, partitioned nature of these networks. It combines the ideas from opportunistic forwarding, trajectory based forwarding and geographical forwarding. Packets are forwarded based on a predefined trajectory geometrically. The relay node must be able to store or forward the message opportunistically to the next forwarder. Each forwarder verifies the status of the message dissemination which is attached to the message header. Vehicles use the header information to gain the knowledge of the message heading location over time and apply the data propagation analysis to act accordingly. The propagation process is limited to the area of the destination to provide timely message delivery.

TBD, MDDV, and PROMPT are examples of hybrid solutions to solve the network disconnection via store-and-forward approach while minimizing the end-to-end latency using global knowledge such as trajectory or source route as well as local knowledge such as neighbor information.

2) Forwarding for Broadcast Routing: Driver safety related applications are the most important motivating applications for VANETs. In such applications, information (e.g., detour route, accident alert, construction warning) should be provided to all surrounding vehicles, thereby requiring a broadcast forwarding protocol. Traditional broadcasting techniques like flooding seriously suffer from broadcast storm problem where large amount of bandwidth is consumed by excess number of retransmissions. When node density is high, this leads to large number of collisions and high channel contention overhead. Most of research activities in broadcast forwarding algorithms propose new ideas to alleviate this problem. Solutions used to adapt the packet load by controlling the packet generation rate is discussed in several papers, see e.g., [145], [146], [147], [148], [149], [150], [151], [152], [153], [148], [154], [155]. In this paper only a subset of them will be discussed.

In [154], [148], five different techniques are proposed to address the broadcast storm problem in MANETs: probabilistic, counter-based, distance-based, location-based, and cluster-based. Their simulation results show that a simple counter-based implementation can avoid a number of redundant messages in dense networks. They showed that if location information available via services like GPS, then location-based scheme is the best choice as it can eliminate a lot of redundant rebroadcasts under all kinds of host distributions without compromising the reachability. In [147], similar techniques are performed in the context of variable thresholds where they can be adjusted on-the-fly. These studies, however, was performed in MANETs. In the following, we discuss some of the techniques developed for VANETs.

3) Probabilistic: In [145], [156], three probabilistic flooding techniques are proposed to solve the broadcast storm problem in VANETs. The solutions are denoted as weighted p-persistence, slotted 1-persistence, and slotted p-persistence schemes. The key suppression technique of these algorithms is a combination of probabilistic-based and timer-based retransmission. In weighted p-persistence methods, vehicles rebroadcast the packets according to the probability p where the higher probability is assigned to farther nodes. The slotted 1-persistence and slotted p-persistence solutions are related to the probability of re-broadcasting a packet within one time slot. The former uses a probability of 1 to re-broadcast a packet within one time slot, while the latter one uses a predefined probability p to re-broadcast the packet within one time slot. To prevent the messages dying out, vehicles buffer the message for certain time and then retransmit it if nobody in the neighborhood rebroadcasts. These techniques are designed in the network layer to reduce the number of packets sent from the network layer to the data link layer. They also quantified the impact of broadcast storms in VANETs in terms of message delay and packet loss rate in addition to conventional metrics such as message reachability and overhead.

In [153], an enhancement of the 1-persistence solution is described, denoted as microSlotted 1-Persistence Flooding, where the time slot used in the 1-persistence solution is divided into a number of micro-slots. This means that within one 1-persistence time slot, more than one node could re-broadcast. This solution, however, gives higher priority of retransmission to the furthest node within the coverage area associated with one 1-persistence slot.

4) Distance-based: In the distance-based forwarding protocol, the vehicles set the waiting time inversely proportional to the distance of the source. However, the vehicles with same distance can still contend for the channel at the same time. Time reservation-based relay node selection (TRRS) algorithm [157] aims to provide shortest end-to-end delay irrespective of the node density. According to these algorithms, all nodes in the communication range of a relay node randomly choose
their waiting time within a given time-window. The time-window range is determined by a distance from a previous relay node and a reservation ratio of the time-window. A node with the shortest waiting time is selected as a new relay node. To avoid multiple reception of broadcasting messages, TRRS further prevents the node that received many duplicate broadcast messages from previous relay nodes to be next relay node.

Similarly, the urban multi-hop broadcast protocol for inter-vehicle communication systems [154] uses directional forwarding approach that suppresses broadcast redundancy by only allowing the furthest vehicle from the transmitter to rebroadcast the packet. It determines the farthest nodes by employing a black-burst (channel jamming signal) contention approach [158]. UMB uses 802.11-based RTS/CTS handshake to avoid hidden terminal problem by divided the road into segments. To handle line-of-sight problem, UMB uses repeaters at intersections to rebroadcast the messages. This protocol is later extended to AMB [155] in order to handle intersection scenarios more efficiently. Unlike in UMB, AMB protocol selects vehicles that pass by the intersections to disseminate the packets into different directions.

In [159], it is argued that the relay nodes in UMB may potentially wait for long time periods before rebroadcasting due to its contention resolving scheme. They proposed a new technique called smart broadcast that does not spend time to resolve collisions, and hence it does not necessarily select the relay in the region that provides the largest progress. Instead of using the black burst mechanism of UMB, the potential relay nodes in this scheme selects random backoff values based on their position where the farther nodes choose a backoff value from smaller ranges. However, the delay gains are marginal.

In [160], Multi-Hop Vehicular Broadcast (MHVB) protocol is described. This protocol can be used to efficiently disseminate the information related to traffic safety applications, such as position and speed. It comprises of two main features: a traffic congestion detection algorithm that suppresses unnecessary beacons due to traffic congestion; and a backfire algorithm that efficiently forwards packets through the network. The congestion detection algorithm detects whether or not vehicles are in the middle of traffic congestion, by counting the number of nodes that are present around the concerned node. It then adjusts the transmitting interval accordingly. The backfire algorithm, on the other hand, efficiently forwards the packet through the network by selecting the next hop based on the distance from the original node. Before retransmission, the relay vehicles calculate the waiting time, which is inversely proportional to the distance from source. The backfire region is mainly contained in a circular area.

In [161], MHVB solution is enhanced in two places. First, the backfire region is changed from a circular region to a sectional region where it is implemented with its angle as an extra parameter. By adjusting the angle of the sector, the area covered by the backfire algorithm can be modified. This results in a flexible and directional backfiring region. The second enhancement is provided by using a Dynamic Scheduling algorithm that is used to differentiate between the packets that have to be transmitted. The packets are prioritized based upon "processing" of the received packets from the other vehicles. In particular, the nodes which are located at a distance farther than 200m are made to transmit the received information earlier than all the other nodes in the network. The main goal of these enhancements is to enhance the balance between the application requirements and the performance of the protocol.

5) Conclusions: The main challenge in designing forwarding algorithms for VANETs is to provide reliable packet transmission with minimum delay, maximum throughput, and low communication overhead. Most existing algorithms target only subset of these requirements within specific environment setups. Recently, several unicast forwarding protocols such as TBD, MDDV, and PROMPT that combine opportunistic location-based and trajectory-based solutions to provide ability to deal with the local optimum and disconnection problems are proposed. There also exists some research on addressing issues related to broadcast transmission, a primary mode of packet exchange in VANETs. Approaches such as weighted p-persistence and UMB leverage a combination of probability-based and distance-based methods to reduce broadcast storm problem. Future research must focus on protocols targeted at heterogeneous systems to handle applications with diverse QoS requirements. For instance, while location-based forwarding solutions seem to be natural for vehicular networks due to their constant topological changes, the IP-address based solutions are more desirable for internet-based applications. Respecting the requirements of applications while solving the fundamental communication problems in VANETs is a significant challenge in designing future forwarding algorithms.

G. Delay constraints

In this section, we categorize all delay-aware protocols based on the layer in which the appropriate steps are being taken.

1) Application Layer Solutions: Delay constraints at the level of application layer are necessary due to the requirements to support emergency warning messages. These messages are typically broadcasted in the affected area. To deal with broadcast storm problem, applications require a good forwarding mechanism that avoids redundant rebroadcasts which can potentially slow down message propagation speed. In [162], an overview of highway cooperative collision avoidance (CCA) is presented, which is an emerging vehicular safety application. It considers a driver model to estimate the level of emergency, and the appropriate warning signal. Emergency messages are transmitted using a direction-aware broadcast forwarding scheme with implicit acknowledgments. Authors concluded that specific context and constraint parameters should be designed in an application-specific manner. For instance, CCA messages are forwarded only in those directions in which affected vehicles are present. In [163], transmission range adaptation techniques for delay control are described. Their protocol, Fast Broadcast, allows the sender to estimate the transmission range before sending the packets. Such a method limits the number of messages that are exchanged in the network, and therefore, it reduces the total transmission time.
The QoS support for multimedia applications in VANETS is studied in [164] by considering three different types of packet flows: audio, video, and data packets. IEEE 802.11e standard is an enhancement of IEEE 802.11 that supports QoS in the MAC layer. This standard attaches a different priority value to each type of packet, and they are not readily applicable for VANETs. This is because the standard does not take link quality, vehicular mobility, and the impact of multi-hop communication into account – motivating the need for a cross layer design between MAC and routing layers. They then presented a triplet-constraint DeReHQ [164] algorithm that transmits packets via paths which have the best link reliability, the smallest number of hops, and link delay is also guaranteed to be under a desired threshold.

A mobile peer-to-peer (P2P) file sharing system that targets VANETs is introduced in CodeTorrent [165]. File swarming techniques that are based on network coding are chosen to transfer data over minimum delay paths. In wired networks, P2P systems are designed for IP address-based network, and they are not readily applicable for VANETs. The challenges here include high node mobility, error-prone wireless channel, and security-risk of information sharing. To address these problems, codeTorrent maintains communication within single-hop neighbors. The file sharing region, however, can be extended through the network of peers using network coding and mobility assisted data propagation. These techniques enable codeTorrent to maintain enough connectivity among peers with low overhead, and data is transferred with minimum download delay. Authors showed that such a strategy outperforms another file sharing protocol called CarTorrent [166].

2) Network Layer: Delay constraints can also be embedded into protocols that operate in the network layer. Designing routing protocols with delay-bound and delay-guarantee char-

<table>
<thead>
<tr>
<th>Table XII</th>
<th>Packet Forwarding Protocols Summary Table</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protocol</td>
<td>Path metric</td>
</tr>
<tr>
<td><strong>Geographic</strong></td>
<td></td>
</tr>
<tr>
<td>GSR [131]</td>
<td>shortest distance</td>
</tr>
<tr>
<td>GPCR [132]</td>
<td>greedy directional</td>
</tr>
<tr>
<td>GFSK [133]</td>
<td>greedy directional</td>
</tr>
<tr>
<td>VADD [135]</td>
<td>least delay</td>
</tr>
<tr>
<td>SADV [137]</td>
<td>least delay</td>
</tr>
<tr>
<td>D-MinCost [136]</td>
<td>least delay</td>
</tr>
<tr>
<td>D-Greedy [136]</td>
<td>shortest path</td>
</tr>
<tr>
<td>PROMPT [138]</td>
<td>least delay</td>
</tr>
<tr>
<td><strong>Link Stability-based</strong></td>
<td></td>
</tr>
<tr>
<td>MOPR [139], [140]</td>
<td>most stable path</td>
</tr>
<tr>
<td>VHRP [141]</td>
<td>most stable path</td>
</tr>
<tr>
<td>PBR [142]</td>
<td>most stable path</td>
</tr>
<tr>
<td><strong>Trajectory-based</strong></td>
<td></td>
</tr>
<tr>
<td>TBD [143]</td>
<td>least delay</td>
</tr>
<tr>
<td>MDDV [144]</td>
<td>least delay</td>
</tr>
<tr>
<td><strong>Broadcast Protocols</strong></td>
<td></td>
</tr>
<tr>
<td>Protocol</td>
<td>Forwarding decision</td>
</tr>
<tr>
<td>Probabilistic [145]</td>
<td>1-persistent, p-persistent</td>
</tr>
<tr>
<td>TRRS [157]</td>
<td>position-based</td>
</tr>
<tr>
<td>UMB [154], AMB [155]</td>
<td>directional, position-based</td>
</tr>
<tr>
<td>SmartBC [159]</td>
<td>directional, position-based</td>
</tr>
<tr>
<td>MHVB [160]</td>
<td>position-based</td>
</tr>
</tbody>
</table>
acteristics is challenging due to high vehicular mobility. We briefly describe some of the important examples here.

There exist a few location-based protocols such as VADD [135], PROMPT [138], and D-Greedy & D-MinCost [136], which obtain statistical path information to route packets over minimum end-to-end delay paths. While VADD and PROMPT perform delay estimation during the path selection phase, D-Greedy & D-MinCost considers only those paths that are within a bounded delay. A key challenge is to estimate the delay for each path before a selection among available ones is made. VADD uses preloaded statistical information such as vehicle density and speed to estimate the path delay. PROMPT, on the other hand, uses the real-time packet traffic statistics to search for low data traffic delay path. Similarly, the DeReQ [167] protocol tries to achieve its dual objectives (reliability and timeliness) by finding a route that is most reliable and also has delay within an allowed maximum bound. To estimate the reliability, DeReQ makes use of road traffic density, relative vehicle speeds, and vehicle traffic flow.

Along with such location-based strategies, there exist some topology-based source routing protocols that account for link stability by estimating the lifetime of different routes. Such an estimation mechanism is used by sender nodes to select the most reliable route to transmit the packets. The relay nodes send the route request for a new route before the existing route is broken [168]. These methods have a considerable impact on the end-to-end delay experienced by packets. In [169], two schemes viz. PGB and AGF are proposed, which aim to improve the delay performance of existing MANET protocols. Preferred Group Broadcasting (PGB) is a broadcasting mechanism that reduces the control message overhead incurred during the route discovery phase of AODV. PGB decreases the number of redundant re-transmissions of basic flooding by allowing nodes in preferred group to rebroadcast or relay the messages. All receiver nodes rebroadcast the message after waiting for a fixed time period. The receiver nodes select the waiting times based on the received signal power level. Another protocol that was proposed in [169] is Advanced Greedy Forwarding (AGF). It is an incremental improvement of GPSR that considers speed and direction along with the location information while discovering the neighbor nodes.

3) MAC Layer: The effectiveness of IEEE 802.11p amendment for traffic safety applications which require low delay, reliable, and real time communication is analyzed in [170]. It has been observed that the CSMA/CA mechanism of 802.11p does not guarantee channel access before a finite deadline and therefore it gives poor performance. The authors of [170] proposed a method known as self-organizing time division multiple access (STDMA). STDMA is a decentralized system where each vehicle determines its own slot assignment based on its positions and neighbor’s information. Such a technique helps in predicting the channel access delay, making it suitable for real-time ad-hoc vehicular networks.

Some researchers have explored the use of multiple directional antennas for fast delivery of packets. For example, RPB-MAC protocol [171] reduces the control message overhead and guarantees minimum channel access delays by making use of multiple antennas. A directional antenna with a communication channel pair is dedicated for set of neighboring vehicles depending on their positions relative to the source vehicle. Since vehicles in different directions communicate using different antennas, the number of channel collisions is reduced. Furthermore, the transmission power is adaptively adjusted to maintain the communication with its neighbors.

4) Physical Layer: The Incident Warning System (IWS) [172] utilizes direct wireless communication to transfer a variety of packets including traffic incidence reports, text messages, JPEG images etc. They divided applications into three different categories and identified specific requirements posed by applications in each category. These requirements are handled by using two different frequencies: long range frequency to reserve the channel; and short range frequency to transmit the packets.

Power adaptation is another technique which researchers have exploited to realize small end-to-end delays. Transmitting power adaptation are discussed in several papers, see e.g., [173], [174], [175], [176], [177]. Only a subset of these solutions will be discussed in this article. In [178], a vehicle can monitor the radio channel conditions by calculating the overhead sequence numbers. The receiving vehicle records the successful packet deliveries sent by neighboring nodes which uses the same radio channel and which are located within the transmission and reception range of the receiving vehicle. By identifying and counting the successfully received packets, the receiving vehicle can detect failed packets and determine the network condition, i.e., the average reception rate and the rate of packets that were not successfully delivered. From this analysis the receiver node can also calculate the minimum number of nodes that are using the same radio channel. The same vehicle can then use the calculated radio channel conditions to adapt its transmission power accordingly. The beaconing load adaptation mechanism described in [178] does not differentiate between the periodic and event driven messages transmitted by the beaconing control channel of IEEE 802.11p.

In [173], [174], other two transmission power adaptation algorithms are discussed to control the beaconing load. These algorithms are denoted as Distributed Fair Power Adjustments for Vehicular environments (D-FPAV) and Emergency message Dissemination for Vehicular environments (EMDV). The D-FPAV is a distributed transmission power control strategy that provides effective transmission for emergency event-driven messages while maintaining the fairness for periodic beacon messages. Each node evaluates the received channel utilization rate that was evaluated since the last beacon transmission. This rate can be calculated either from the link layer statistics or from the network layer statistics. Each transmitted beacon carries this value. Each node also maintains a target channel utilization rate. If the received channel utilization rate is smaller than the target value then the transmitted power is increased by a predefined amount. If on the other hand, the received channel utilization rate is higher then the transmitted power is decreased by the same predefined amount. The transmission power is not altered if both the rates are exactly the same. In addition, D-FPAV allows the prioritization of
event-based messages over periodic messages. The minimum power level assignment for a vehicle is calculated by choosing the smallest observed value of the power assignment levels among the received beacons.

The EMDV uses a contention strategy that supports the fast and effective dissemination of alerts within a target geographical area in cooperation with D-FPAV. In EMDV, a source vehicle that needs to send an alert (emergency) message chooses a relay node that is as far as possible, and has high reception probability. On successful reception, the relay node retransmits the emergency message. If the packet reception is failed at the selected relay node then other vehicles that received the message are considered as potential relay nodes. These nodes wait for a predefined time, i.e., retransmission delay timer, and rebroadcast the message only if they did not hear any rebroadcasts during the waiting period. Moreover, this algorithm is capable of differentiating between the periodic and event driven messages.

In [179], D-FPAV and EMDV are used as base protocols. D-FPAV is enhanced by modifying the algorithm in such a way that instead of processing the received utilization from one edge region, each vehicle needs to process the received utilization data from two beacon edge regions, in front of the vehicle and behind the vehicle along the road. In this case, the target utilization rate is compared with an effective edge utilization rate. The latter is calculated through linear interpolations of the received channel utilization rate and the beaconing power received in farthest beacons in front and behind the vehicle. On the other hand, EMDV is enhanced by, among others, modifying the retransmission delay timer upon each new reception of same re-broadcasted information. The delay timer adjustment is done in such a way that it results in a uniform geographic distribution of re-broadcasting relays. This can be achieved by re-evaluating the delay timer by considering the distance to the closest vehicle among the relay vehicles that can rebroadcast the same information.

In [180], [181], the transmission power is adaptively adjusted to accommodate the change in neighbors. If the number of neighbors falls below a threshold then the power is increased, and similarly when the number exceeds another threshold the power is reduced accordingly. A potential drawback here is that the thresholds are static and do not reflect different vehicle traffic conditions and quality of road segments. DB-DIPC [182] proposed power adaptation techniques for vehicular networks which rely on local information obtained via periodic exchange of beacon messages among neighbors. LOADPOW [175] uses the traffic load information in routing protocol to adjust the transmission power before sending the packet in medium access layer. Although these algorithms are adaptive and distributive, they need further analysis to understand of the effects of power adaptation on different performance measures.

5) Conclusions: The primary challenge in designing protocols is to provide good delay performance under the constraints of high vehicular speeds, unreliable connectivity, and fast topological changes. In this section, we discussed several methods that incorporate delay constraints in various layers. However, one must be aware that such individual solutions may lead to conflict between layers and among other nodes. For instance, increasing transmission range, the number of hops is reduced and this could possibly reduce the end-to-end transmission delay. However, increasing transmission range causes additional contention delay at MAC level. To provide overall system improvement, future solutions must focus on cross-layer protocols that strike a balance among conflicting issues from different layers with an objective of end-to-end delay minimization.

H. Prioritization of data packets

When an emergency event occurs, the channel utilization is likely to degrade due to massive broadcast of emergency messages. A simple approach in such situations, which many protocols adopt, is to simply drop lower priority packets. Some other protocols attempt to provide appropriate congestion control mechanisms so that the sending rate of lower priority packets is adaptively adjusted.

A vehicle collision warning communication (VCWC) [183] is an example of cooperative collision warning system that is enabled by vehicle-to-vehicle communication. It aims to give low latency warning message transmission at the initial state of an emergency event. The issue of packet congestion is addressed by rate adaptation scheme which assigns different priority levels to different packets based on the application requirements. Whenever a node has a backlogged emergency message, it raises an out-of-band busy tone signal, which can be sensed by vehicles located within two hops. Vehicles with lower priority messages defer their channel access whenever the busy tone signal is sensed. Furthermore, bandwidth utilization is improved by suppressing multiple warning messages regarding the same event. In [184], the authors studied the channel congestion control in 802.11p and suggested that the packets in CCH (control channel) need to be prioritized. The safety messages should have higher priority than background or control messages such as periodic beacon and hello messages. They provide various congestion control mechanisms via MAC queue manipulation. The main idea is to provide absolute priority for safety messages via manipulating (e.g., freezing) the MAC queues of lower prioritized traffic, or to dynamically reserve a fraction of bandwidth for the highest priority traffic with adaptive QoS parameters. Similarly, in [185], methods based on 802.11e protocol are proposed to provide higher priority to emergency warning messages.

In [186], a novel pulse-based control mechanism has been proposed to provide strict priority for emergency messages. According to this mechanism, as soon as an emergency event is noticed, vehicles start a random backoff timer whose value depends on the emergency of the situation. Once the timer expires, the vehicle will start to transmit pulses in the control channel. Shortly after starting to transmit pulses, the emergency packet is transmitted in the data channel. When a node detects a pulse in the control channel at any time, it aborts its transmissions to release both channels. Such a method gives strict priority for emergency messages. In [187] several random access protocols for a vehicle to send safety messages to other vehicles are proposed. These protocols fit in the DSRC multi-channel architecture, and provide high reliability and small delay for safety messages.
CVIA-QoS [188] aims to provide delay-bounded throughput guarantees for soft real-time traffic. It implements an admission control mechanism at the temporary routers and gateways to provide higher priority to real-time applications. The transmission time is divided into two periods—high priority period and low priority period. Low priority period is adopted from CVIA [189]. The high priority period provides a reliable pooling system based on channel reservation. The separation of transmission period between low and high priority packets guarantees end-to-end delay for high priority packet.

1) Conclusions: The new standards like 802.11e and IEEE 802.11p provide guidelines for packet prioritization. While there is some research in adopting these standards, more work needs to be done in effectively leveraging them. For example, cross-layer protocols that operate in multiple layers to provide priorities among different flows and different applications. Furthermore, developing efficient scheduling strategies that enable delay-aware transmission of packets with different priorities is also a matter of concern for future VANET applications.

1. Reliability and cross-layering between transport and network layers

There exist some research work activities on designing cross-layer protocols which span between transport and network layers. The motivation behind such a cross-layer design is to support real-time and multimedia applications which require a reliable end-to-end connectivity with QoS requirements. Cross-layer designs also help in congestion avoidance. Due to frequent disruptions in the routes, traditional transport layer protocols from MANETs [134], [190], [191], [192] are not directly applicable to VANETs. One must leverage the information from network layer in adjusting the packet transmission in the transport layer to adapt to the dynamic network topology in VANETs.

We first discuss several challenges in having transport layer protocols in vehicular networks. Since TCP is the most popular transport protocol, we confine our discussion to TCP in vehicular networks. TCP is originally designed for wired networks with acceptable data throughput. However, the fundamental properties of mobile networks such as dynamic topology, unreliable wireless radio transmission are highly different from wired networks. Several investigations on the impact of these properties on the performance of TCP showed that it provides poor throughput in multi-hop ad hoc networks [193]. This poor performance is mainly due to the conservative flow and congestion control mechanisms deployed in TCP. For example, TCP interprets transmission errors as a congestion situation and thus reduces the throughput. Developing effective congestion control mechanisms is also very challenging. This is because the predictions about potential congestion situations are based on local information, which may not reflect the current state of the network.

Another possibility to improve the performance is to leverage the information from other nodes in the system. For example, intermediate nodes detect congestions and signal them to other vehicles through Explicit Congestion Notification (ECN, RFC 3168) operation. The network information gathered from neighboring nodes provides a better estimation when compared to the predictions of individual node-level congestion control mechanisms. Such mechanisms require the transport layer to interact with lower layers to obtain appropriate information about current network condition, thereby motivating the need for effective cross-layer design.

In [193], the mobile control transport protocol (MCTP) that aims to provide Internet access in VANETs is proposed. The basic idea here is that MCTP observes the IP packet flow between sender and receiver in order to react appropriately in the transport layer. MCTP considers several notifications from underlying protocols as well as from other vehicles (e.g., pending congestions, number of unreachable ICMP messages etc.). Such information helps MCTP to distinguish between link errors, congestions, and disconnections from the Internet.

Reliable packet transmission of TCP is of great importance in file sharing and content distribution applications during highway driving. In [134], authors argue that robust routing protocols must be used in order to address the problem of TCP in handling route breakage in VANETs. They study the joint optimization of TCP and geographic routing parameters to handle high vehicle speeds. Under a controlled network, they show the impact of high mobility on critical system parameters of TCP and UDP such as hello message exchange rate. They then proposed an adaptive scheme where the hello interval is based (and depends) on vehicle speed: \( I = \frac{R}{k \times \text{speed}} \), where \( I \) is the interval, \( R \) is the transmission range, and \( k \) is a tunable parameter. They empirically show that the delivery ratio of both UDP and TCP is higher when used with adaptive scheme. The authors have also developed a novel scheme for out-of-order delivery.

VTP [194] relies on position-based routing such as PBR [142] to cope with temporary network partitions that interrupt end-to-end connectivity and cause packet loss. Its main approach is the utilization of statistical path characteristics for error and congestion control. VTP avoids unnecessary transmission rate reductions due to non-congestion packet loss such as routing errors by using feedback information from local neighbors. The intermediate nodes compute the minimum bandwidth that is locally available and feed the information back to sender (via piggybacking). Sender uses this information to calculate bandwidth-delay product to determine route quality. VTP provides connected/disrupted states to deal with frequent disconnection. The sender periodically sends probe message. When the relay node becomes available, the sender resumes its packet transmission.

1) Conclusions: Above mentioned work primarily focuses on applications which require unicasting. Since many safety-related and other applications require geocasting or broadcasting, there is a clear need for new approaches that are not based on traditional transport protocols [7]. It is even more challenging the case of geocasting protocols since the relay nodes in such methods do not maintain any state information. Cross-layer design holds a promising future in realizing effective protocols that address issues related to congestion and link disruption.
VI. CONCLUSIONS AND FUTURE WORK

Vehicular networking is the enabling technology that will support several applications varying from global Internet services and applications up to active road safety applications. This paper is a survey and tutorial that introduced and discussed the possible applications and use cases that could be supported by vehicular networks in the near and long term future. Furthermore, the several requirements, e.g., communication performance requirements, imposed by such applications are emphasized. Moreover, the ITS projects and programs that were and are being conducted in the USA, Japan and Europe are presented. The ITS architectures and protocol suites used in the different parts of the world are introduced and discussed. Finally the recent main research challenges associated with vehicular networking are introduced and several solutions for these research challenges are described.

The main conclusions and recommendations for future activities are listed below.

Geographical addressing: the most promising, but also the most complex one is the geographical addressing family that extends IP routing and IP addressing in order to cope with GPS addresses. While several solutions associated with this family have been proposed, more research and standardization activities are needed for a successful realization.

Data-centric Trust and Verification: the proactive data-centric trust and verification security concept has been researched extensively. However, the tamper-resistance hardware used in a vehicle to detect unnecessary accident warnings, needs to be further researched. The reactive security concept has been studied in a smaller scale. More work is needed in the area of context verification, where a vehicle is able to realize an intrusion detection system by comparing received information on parameters associated with status and environment with its own available information.

Anonymity and privacy: is being extensively investigated. However, an open area is anonymity and adaptive privacy, where users are allowed to select the privacy that they wish to have.

Forwarding algorithms: the main challenge in designing forwarding algorithms for VANETs is to provide reliable packet transmission with minimum delay, maximum throughput, and low communication overhead. Future research must focus on protocols targeted at heterogeneous systems to handle applications with diverse QoS requirements. Respecting the requirements of applications while solving the fundamental communication problems in VANETs is a significant challenge in designing future forwarding algorithms.

Delay constraints: the primary challenge in designing protocols is to provide good delay performance under the constraints of high vehicular speeds, unreliable connectivity, and fast topological changes. In this section, we discussed several methods that incorporate delay constraints in various layers. To provide overall system improvement, future solutions must focus on cross-layer protocols that strike a balance among conflicting issues from different layers with an objective of end-to-end delay minimization.

Prioritization of data packets: the new standards like 802.11e and IEEE 802.11p provide guidelines for packet prioritization. While there is some research in adopting these standards, more work needs to be done in effectively leveraging them. For example, cross-layer protocols that operate in multiple layers to provide priorities among different flows and different applications. Furthermore, developing efficient scheduling strategies that enable delay-aware transmission of packets with different priorities is also a matter of concern for future VANET applications.

Reliability and cross-layering between transport and network layers: since many safety-related and other applications require geocasting or broadcasting, there is a clear need for new approaches that are not based on traditional transport protocols. It is even more challenging the case of geocasting protocols since the relay nodes in such methods do not maintain any state information. Cross-layer design holds a promising future in realizing effective protocols that address issues related to congestion and link disruption.

ACKNOWLEDGEMENTS

G. Karagiannis and G. Heijenk wish to express their gratitude to the Dutch Connect & Drive project, which funded their part of this work.

REFERENCES

Georgios Karagiannis holds a M.Sc. degree (1993) and a Ph.D. degree (2002) in electrical engineering from the University of Twente, the Netherlands. From 1994 to 1998 he was working as a researcher and since 2003 as an assistant professor at the Design and Analysis of Communication Systems (DACS) group at the same university. In between he was working in the Wireless Multimedia Research unit of Ericsson Eurolab in Enschede, the Netherlands. His research interests are in the fields of fixed, mobile and wireless (inter)networking, end-to-end QoS signaling and provisioning, mobility and routing in vehicular communication networks, and performance evaluation. His main research interests in vehicular communication networks are related to the design and evaluation of dissemination and aggregation algorithms and protocols required for the support of vehicular applications.

Onur Altintas is a senior researcher at the R&D Group of Toyota InfoTechnology Center, Co. Ltd, in Tokyo. From 1999 to 2001 he was with Toyota Motor Corporation and from 2001 to 2004 he was with Toyota InfoTechnology Center USA, and was also a visiting researcher at Telcordia Technologies between 1999 and 2004. Before joining Toyota Motor Corporation in 1999, he was a research scientist at Ultra High Speed Network and Computer Technology Labs (UNCL), Tokyo. He received his B.S. (1987) and M.S. (1990) degrees from Orta Dogu Teknik University, Ankara, Turkey, and his Ph.D. (1995) degree from the University of Tokyo, Japan; all in electrical engineering. He served as the Co-Chair for Vehicle-to-Vehicle Communications Workshops (V2VCOM 2005 and V2VCOM 2006) co-located with ACM MobiQuitous, and V2VCOM 2007 and V2VCOM 2008 co-located with IEEE Intelligent Vehicles Symposium. He also served as the Co-Chair for the IEEE Workshop on Automotive Networking and Applications (AutoNet 2006, AutoNet 2007 and AutoNet 2008) co-located with IEEE Globecom. He is the co-founder and general co-chair of the IEEE Vehicular Networking Conference (IEEE VNC) held in Tokyo in 2009, in New Jersey in 2010, and in Amsterdam in 2011. He also served as a guest editor for a special issue on Vehicular Communications for IEEE Wireless Communications Magazine (2009) and EURASIP Journal on Wireless Communications and Networking (2009) and as Track Chair of Vehicular Electronics and Telematics for the IEEE Vehicular Technology Conference (IEEE VTC Spring 2009, 2011 and 2012). He is an IEEE VTS Distinguished Lecturer.

Boangoat Jarupan received her BS and MS degrees in Electrical and Computer Engineering from The Ohio State University in 1997 and 2000, respectively. She is pursuing her Ph.D. in the same university. Her research interests include mobile ad hoc communication systems with an emphasis on vehicular networks. She is currently working on cross-layer protocol design and multi-hop communication models for delay sensitive applications and intersection collision warning systems.

Kenneth Lin is an entrepreneur, a management consultant and a senior information security architect with extensive experience across finance, health, automotive and public sectors. He authored a few security design process published by US federal government as national guidance to secure health systems. During his tenure at Booz Allen, he played a key role in many large scale cross-organizational projects in both commercial and public sectors. He currently serves as the senior security architect for Fannie Mae. Mr. Lin holds a Master of Science degree in Electronic Commerce from Carnegie Mellon University and a Master of Computer Science degree from National Chiao-Tung University.

Tim Weil is an IT Security Architect with over twenty years of IT management, consulting and engineering experience in the U.S. Government and Communications Industry. Mr. Weil’s technical areas of expertise include enterprise security architecture, FISMA compliance, identity management, ITS standards, and network engineering. Since 2006 he worked on several Intelligent Transportation System (ITS) projects relating to the WAVE/1609 protocol suite. His degrees include an M.S. in Computer Science from Johns Hopkins University, and a B.A. in Sociology from Immaculate Heart College. He is an industry-certified Security and Privacy professional (CISSP), Project Management Professional(PMP) and IT Auditor(CISA). He currently works as the Senior Manager - Information Security for Raytheon Polar Services (Centennial, Colorado). Mr. Weil is a Senior Member of IEEE.