Research Article

QoSHVCP: Hybrid Vehicular Communications Protocol with QoS Prioritization for Safety Applications

Ahmad Mostafa,1 Anna Maria Vegni,2 Talmai Oliveira,1 Thomas D. C. Little,3 and Dharma P. Agrawal1

1 Center for Distributed and Mobile Computing, The School of Computing Sciences and Informatics, University of Cincinnati, Cincinnati, OH, USA
2 Department of Applied Electronics, Roma Tre University, Rome, Italy
3 Department of Electrical and Computer Engineering, Boston University, Boston, MA, USA

Correspondence should be addressed to Anna Maria Vegni, amvegni@uniroma3.it

Received 12 January 2012; Accepted 15 February 2012

Academic Editors: H. M. Sun and Y. M. Tseng

Copyright © 2012 Ahmad Mostafa et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper introduces a hybrid communication paradigm for achieving seamless connectivity in Vehicular Ad hoc Networks (VANETs), wherein the connectivity is often affected by changes in the dynamic topology, vehicles’ speed, as well as the traffic density. Our proposed technique named QoS-oriented Hybrid Vehicular Communications Protocol (QoSHVCP) exploits both existing network infrastructure through a Vehicle-to-Infrastructure (V2I), as well as a traditional Vehicle-to-Vehicle (V2V) connection that could satisfy Quality-of-Service requirements. QoSHVCP is based on a V2V-V2I protocol switching algorithm, executed in a distributed fashion by each vehicle and is based on the cost function for alternative paths each time it needs to transmit a message. We utilize time delay as a performance metric and present the delay propagation rates when vehicles are transmitting high priority messages via QoSHVCP. Simulation results indicate that simultaneous usage of preexisting network infrastructure along with intervehicular communication provide lower delays, while maintaining the level of user’s performance. Our results show a great promise for their future use in VANETs.

1. Introduction

Vehicular Adhoc Networks (VANETs) are emerging as a preferred network design for Intelligent Transportation System (ITS), particularly for relaying data in a multihop mode using Dedicated Short-Range Communication (DSRC). DSRC provides data communications among nearby vehicles, supports Internet access and safety applications [1], thereby exploits the use of flooding in a vehicular system.

Vehicle-to-Vehicle (V2V) communication is supported by smart vehicles equipped with on-board multiple network interface cards (e.g., Wi-Fi, HSDPA, and GPS), and emerging wireless technologies (e.g., IEEE 802.11p, WiMax, and LTE). V2V aims to provide low-latency short-range vehicular communications and multi-hop connectivity between vehicles. However, V2V may not always be available due to dynamic changes in the network topology, varying vehicle speeds, and traffic density [1]. In a sparsely connected or totally disconnected scenario, vehicles are not always able to communicate with each other, and V2V does not seem to be the most appropriate communication scheme, especially for non-safety-critical applications [2, 3], even though V2V forms multiple clusters of vehicles.

A long-range vehicular connectivity can exploit preexisting network infrastructure such as wireless access points, called Road-Side Units (RSUs), in order to provide communications between disconnected cluster of vehicles. This relies on Vehicle-to-Infrastructure (V2I) protocol. For instance, in either very low traffic or even totally-disconnected scenarios, intervehicle communications are difficult to maintain, and the use of network infrastructure appears to be a viable solution, at least for applications that require bridging between the networked cluster fragmentations inherent in any multi-hop network formed by moving vehicles. Drive-thru Internet
systems represent those emerging wireless technology that provides Internet access to vehicles by enabling connections V2I when a vehicle crosses an RSU [4].

In order to assure a seamless connectivity within a VANET, V2V and V2I need to be combined into a hybrid communication protocol and are assumed to complement each other [5–7]. Indeed, connectivity management is a real challenge for a VANET. Exploiting both V2V and V2I represents an effective integrated solution in avoiding disconnections and guaranteeing continued data communication independent of the traffic scenarios (i.e., dense, sparse, and totally disconnected neighborhoods). For instance, V2V is largely used in rush hours, as well as in sparsely connected rural areas with no network infrastructure; while V2I represents a viable solution for maintaining connectivity in urban areas with low vehicular density.

In this paper, we present a QoS-oriented Hybrid Vehicular Communications Protocol (QoSHVCP), for improving the vehicular connectivity with the support of network infrastructure, as well as intervehicle communications. QoSHVCP allows vehicles to decide which communication protocol (i.e., V2V and V2I) is the most appropriate for temporal and local connections. Our proposed technique takes advantage of both V2V and V2I, and based on the minimization of message delivery time propagation delay, it lets the vehicles select one of them by a handover (The handover mechanism takes origin in cellular systems, in order to maintain user services in mobility scenarios [8]. In this work we rely on handover concept to identify a protocol switching that guarantees seamless connectivity in VANETs) mechanism.

QoSHVCP works in a traditional VANET scenario with network infrastructure, assuming that both V2I and V2V are viable solutions for communications. QoSHVCP offers twofold scope for vehicles to (i) follow multi-hop communication when available via V2V, and (ii) employ communications with the network infrastructure via V2I. As a result, using QoSHVCP handover mechanism, each vehicle can switch from V2V to V2I and vice versa. The handover decision criteria depend on minimizing the message transmission delay.

The QoS requirement has also been considered in this work, since apart from achieving connectivity in a VANET, different priority levels of message are considered in each protocol switching decision. Two classes of priorities are considered: high and low priority levels (i.e., HP, LP). HP messages are preferred to be forwarded mostly via V2I connections since V2I can guarantee low delays and high performance. On the other hand, LP messages prefer V2V mechanism which, depending on the vehicles’ density, can be expected to achieve low delays as well. The QoS management is supported by a load-balancing mechanism, capable of meeting QoS requirement, thereby avoiding traffic overload on the network infrastructure.

The paper is structured as follows. In Section 2, we investigate main issues of seamless connectivity in VANETs and highlight the related work on hybrid vehicular communication protocols. Section 3 gives the details of our proposed QoSHVCP technique; in Section 3.1, we describe the protocol switching mechanism, while in Section 3.2 we introduce the QoS prioritization adopted in QoSHVCP. Section 4 deals with the message delivery time delay propagation rates in our QoSHVCP. The proposed technique is then validated through simulation results, and compared to traditional V2V and V2I protocols in Section 5. Simulations results are obtained in terms of average message propagation delay and the network overload, for different QoS requirements, vehicle densities, and speeds. Finally, conclusions are drawn in Section 6.

2. Related Work

Many factors can affect a VANET topology and its dynamic behavior. Traffic density (i.e., well-connected, sparsely connected, and totally disconnected neighborhood), vehicles’ speed (i.e., low, medium, and high speed), and the heterogeneous network environment (i.e., technologies of wireless networks around the VANET and their deployment methods) are the main aspects depicting a VANET. It may be noted that fast mobility of vehicles make most traditional MANETs routing protocols inefficient for VANETs, mainly due to lack of maintaining the same topology for a reasonable amount of time.

As a consequence, communications among vehicles are an open issue since it cannot always be supported, and messages can be either lost or never received. Opportunistic forwarding is a traditional technique adopted in a Delay Tolerant Network (DTN) [9] and has been extended in VANETs to achieve connectivity between vehicles via V2V and to disseminate information [2, 10, 11]. It provides message propagation through dynamically changing links as a bridging technique, where any vehicle can be used as the next hop and vehicles forward the message via RSU to the final destination. The authors in [12] define an opportunistic forwarding technique in VANET as an advanced information dissemination communication pattern with an objective for disseminating the information among vehicles during a certain period of time. Traditionally, schemes for advanced information dissemination use single-hop broadcasts or store-and-forward technique, and forward messages multiple times to all those vehicles unreachable due to the network partitioning.

Message and time delay propagation in a VANET via opportunistic networking have been largely investigated in the literature, and different broadcasting techniques have been proposed which can be classified by distance, location, probability, and topology based [13]. The distance and location-based approaches simply exploit the intervehicular distance and the vehicles’ positions through GPS devices in order to select the next hop to forward a message. Beacon messages are implemented in many location-based approaches, where vehicles’ position information is embedded. In [14], each vehicle has the knowledge of its neighbors in terms of both numbers of neighbors and their respective positions. The next hop selection occurs for the furthest vehicle from the source vehicle. In [15], a fast multi-hop broadcast technique is proposed. It relays estimates of vehicles distance and reduces the number of hops and
associated delay required to forward a broadcast message. It is well known that a nonoptimal number of hops, used by a message in forwarding to a destination vehicle, cause higher delays, and the network performance could be affected drastically. However, one drawback of position-based broadcasting approach is the need for global information about the network topology, as well as the geographical distribution of the vehicles. A large quantity of data information has to be typically sent by a dedicated logical channel.

In the probability-based broadcasting techniques, it is assumed that the probability of collision is reduced and transmitted messages are decreased. Upon message reception, each vehicle retransmits with a probability depending on the distance from the source vehicle [16]. It follows that greater the vehicle’s distance is, higher will be the retransmission probability. In [11], Resta et al. consider multi-hop emergency message dissemination through a probabilistic approach and derive a lower bound on the probability that a vehicle correctly receives a message within a fixed time interval. Similarly, Jiang et al. [17] introduce an efficient alarm message broadcast routing protocol and estimate the receipt probability of the alarm messages sent to vehicles.

All the previous methods are effective only for V2V with dense traffic scenario and are quite limited when vehicles are in a low density neighborhood. A RSU could represent a viable solution to enhance the vehicular connectivity. Many authors investigated techniques to allow vehicles to be seamlessly connected. Such approaches rely on using both V2V as well as V2I techniques. This combination is commonly referred to as V2X.

V2V and V2I communication technologies have been developed as a part of the Vehicle Infrastructure Integration (VII) initiative [18]. As described in [19, 20], the use of a vehicular grid, along with an opportunistic infrastructure placed on the roads, can guarantee seamless connectivity in a dynamic vehicular scenario. In [5], the authors propose a Cooperative Infrastructure Discovery Protocol, called CIDP, which allows vehicles to gather information about encountered RSUs through direct communication with the network infrastructure, and subsequent exchange messages with neighboring vehicles via V2V. The authors show the effectiveness of their approach. But, it is limited to the message exchange about the infrastructure discovery. In [21], Wedel et al. use V2X communications for an enhanced navigation system which intelligently help drivers to circumnavigate congested roads and avoid traffic roadblocks. Their contribution highlights advantages of V2X communication protocols for numerous safety applications.

Finally, in [6] Seo et al. analyze the performance of a general hybrid communication protocol, based on the IEEE 802.11p wireless access in vehicular environments (WAVEs) system. The authors focus on packet error rates, while connectivity and reliability of vehicles have not been considered.

In order to provide a more efficient resource management and in an attempt to satisfy soft real-time requirements in a distributed system, there has been a significant number of works that looked at how to implement load-balancing mechanism in a distributed system [22]. Most of the work focus on the distribution and/or migration of the workload among the many different servers [23, 24]. Some even go as far as adapting the functionality of the clients in the system [25]. When there is an option of redistribution, load-balancing could also be obtained by distributing the traffic generated among multiple paths and servers [26, 27]. However, such load-balancing mechanisms might not provide optimal resource management in a VANET scenario due to the fact that there is often a lack of multiple resources, or routing options are limited. In many cases, vehicles would be limited to either go through the route using the infrastructure, or to use the formed adhoc network.

Several studies have introduced analytical models for the data delivery rates and delay time within vehicular networks. Some have addressed propagation delay for safety critical warning messages in a vehicular environment [28–30]. In [28], the authors develop an analytical model that evaluates the message delivery delay in critical safety applications and its relation to the buffering and switching mechanism within the WAVE protocol. The same problem has been considered by Abboud and Zhuang in [29]. However, they observe the tradeoff between the message delivery delay versus the cluster size used by the vehicles travelling on the highway. Finally, in [30], the authors present an analytical model to show dependence on the vehicular density in the highway.

Yousefi et al. [31] have also developed an analytical model for message delivery delay in a VANET by exploring queuing theory in studying the vehicular connectivity when the traffic follows a unidirectional model. We derive an analytical model for message delivery in a typical dynamic network for a bidirectional traffic. In general, our work concentrates on a different aspect of the VANET that represents a more realistic view of such networks. We present an analytical model when such a network appears as a partitioned network that incorporates different connectivity phases a vehicle encounters during its trip on a highway scenario. We also consider [32], where the authors present an analytical model that characterizes the connectivity of the VANET on a unidirectional road. Rather than only considering the network connectivity aspect in a unidirectional traffic scenario, we compute an expected delay for the message delivery.

In this paper, we investigate a hybrid approach for enhancing the connectivity among vehicles. Our approach, that is, QoSHVCP, is a hybrid vehicular protocol, providing appropriate switching from V2V to V2I and relying on a vehicular grid with neighboring wireless network infrastructure.

QoSHVCP is a broadcast protocol by means of intervehicle communications (V2V), which can be conditionally relayed by one or more RSUs (V2I). This approach considers a protocol switching (from V2V to V2I, and vice versa), aiming at seamless connectivity and is expected to improve communication performance independent of any specific traffic scenarios or vehicle speeds. It consists of a handover procedure from V2V to V2I (and vice versa), resulting in improving opportunistic connectivity with respect to traditional intervehicles communications. Our QoSHVCP also has a load-balancing component that considers two different classes of message priorities. It allows the network to
gracefully degrade, while still maintaining good performance for high priority messages.

3. QoSHVCP Technique

QoSHVCP technique is a hybrid approach that links both vehicles (i.e., V2V) and from vehicles to the infrastructure (i.e., V2I) communications. The cooperation and coexistence of these two different methods can assure a good connectivity in a VANET scenario, especially in sparsely connected neighborhoods where V2V communication is not always feasible.

QoSHVCP is a broadcast protocol that reduces the time required by a message to propagate from a source vehicle to the farthest vehicle inside a certain strip-shaped area of interest. QoSHVCP represents a realistic communication protocol, since vehicles can establish opportunistically both V2V and V2I communications and reduce the message delivery time, as well as avoid disconnections due to changed traffic density and dynamic topological changes.

Based on the estimation of the link utilization time (i.e., the message delivery time for one hop) of vehicles, QoSHVCP is then used to reduce the amount of hops needed to deliver the message. In a previous work [33], we presented a limited version of protocol switching algorithm, which assumed a known and constant transmission range of vehicles. This represents a strong limitation of the protocol resulting in an unrealistic implementation of the algorithm. In this paper, we adapt QoSHVCP to be a more pragmatic broadcast protocol where vehicles’ actual transmission data rates are subjected to continuous changes due to physical obstacles, vehicle density, speed, network overload, and so forth.

Apart from achieving seamless connectivity in a VANET [33] through dynamic protocol switching, our proposed technique is QoS oriented and guarantees message delivery with low delay, specially for HP messages. In particular, QoSHVCP treats HP messages (e.g., warning, safety, and soft-real time messages) to be forwarded via V2I; while LP messages (e.g., delay-tolerant) via V2V. The main characteristic of QoSHVCP is to exploit the connectivity in the network infrastructure for HP messages whenever available, as RSU can directly forward a message to the next RSU, resulting in an increase in the message propagation gap inside the vehicular grid.

In the following Sections 3.1 and 3.2, we respectively describe the protocol switching mechanism, and the QoS prioritization adopted in QoSHVCP.

3.1. Delay-Based Protocol Switching Mechanism in QoSHVCP

Let us consider the vehicular scenario depicted in Figure 1. Several RSUs of different wireless technologies are deployed, partially covering a given area. The local information—assumed as global—comprises the key data defining the network scenario, since the traffic density is directly detected by the vehicles. Each vehicle continuously monitors its local connectivity by storing HELLO broadcast messages and is then able to determine if it is within a cluster or is travelling alone on the road. A vehicle will be aware of Internet access on the basis of broadcast signals sent by the RSUs.

The knowledge of RSUs’ presence in the range is indicated by a routing parameter, defined as Infrastructure Connectivity (IC). This parameter indicates the ability of a vehicle to be connected directly with one or more RSUs. The IC assumes two values, that is, IC = {0, 1}, corresponding to no RSU, and one or more available RSUs respectively. For instance, when a vehicle has IC = 1, it means that it is driving inside the radio coverage of an RSU wireless cell and is potentially able to directly connect to the RSU. Otherwise, the value of IC is 0 when no accessible wireless cell is available.

Let us consider a cluster C comprised of a set S of n vehicles (i.e., S = {1, 2, ..., n}). We assume m RSUs (i.e., m < n) displaced in the network scenario as depicted in Figure 1. Each vehicle is able to communicate with all the other vehicles around it via V2V. At the same time, we assume that only a limited subset of vehicles in the cluster C, (i.e., S’ = {1, 2, ..., I} ⊂ S, with I < n), is able to connect to an RSU via V2I. For example, not all the vehicles might have an appropriate network interface card and/or are not in the range of connectivity of an RSU. Analogously, we assume that only k RSUs (i.e., k = {1, 2, ..., h} with h < m) are available to V2I communications (These are only assumptions, based on monetary cost of RSU displacement and the availability of V2I communications).

For the connectivity link from the i-th to the j-th vehicle, we define link utilization time \( q_{(i,j)} \) as the time needed to transmit a message of length \( L \) [bit] from the i-th to the j-th vehicle, at an actual data rate \( f_{(i,j)} \) [bit/s], which can be given by

\[
q_{(i,j)} = \frac{L}{f_{(i,j)}}.
\] (1)

For a direct link between i-th vehicle and k-th RSU, the data rate is computed by the nominal data rate \( f_{(i,k)} \) [bit/s] by applying a Data Rate Reduction (DRR) factor (i.e., \( \rho_{(i,k)} \)) that depends on the distance from the vehicle to the RSU, namely, \( f_{(i,k)} = \rho_{(i,k)} f_{(i,k)} \). The DRR factor decreases when a vehicle is far from the center and located within the bounds of an RSU wireless cell.

Let us define a path from i-th vehicle to k-th RSU, comprising of a sequence of M hops, where a single hop represents both a link between two neighboring vehicles, and from a vehicle to the RSU. The path length represents the number of hops M for a single path. (The path length is assumed to be known in advance for each available path in a cluster, before transmitting messages. Each path is built on the basis of local connectivity and IC parameter information.) It follows that the maximum number of directed links from a vehicle to an RSU is \( a = l \cdot h \), while the maximum number of different paths that can connect i-th vehicle to k-th RSU is \( n \cdot a \).

From the definition of a path, we define the path utilization time (i.e., \( Q_{(i,k)} \) [s]) from the i-th vehicle to the k-th RSU as the sum of single link utilization time parameters (i.e., \( q_{(x,y)} \) [s]), for each available hop (i.e., \( (x,y) \)) that constitutes the path, such as

\[
Q_{(i,k)} = \sum_{\{x,y\} \in S} q_{(x,y)} = L \sum_{\{x,y\} \in S} \left[ f_{(x,y)} \right]^{-1}.
\] (2)
where $x$ and $y$ are the indexes for available links, in the range $[i, k - 1]$ and $[i + 1, k]$, respectively.

Among all the $na$ paths, the *optimal path* will be the one with minimized path utilization time, given by

$$\min_{s=1,2,...,na} Q_{(i,k)}^{(s)} = L \cdot \min_{s=1,2,...,na} \sum_{\{x,y\} \in S} \left[ f_{(x,y)}^{(s)} \right]^{-1}.$$  \hspace{1cm} (3)

Equation (3) provides a fast message transmission from a source vehicle to an RSU. Notice that the *optimal path* can be comprised of both V2V and V2I multi-hops. The switching mechanism from V2V to V2I, and vice versa, occurs on the basis of minimization of the overall message propagation delays.

### 3.2. Load-Balancing Mechanism in QoSHVCP

QoSHVCP aims to guarantee connections either through V2V or through V2I on the basis of minimizing path utilization time. However, in a VANET, various applications require different communication modes and QoS levels. For instance, two most important safety applications are the Extended Emergency Brake Light (EEBL), and the Cooperative Intersection Collision Avoidance System (CICAS). EEBL is based on V2V communications, while CICAS exploits V2I mode [1]. Leveraging on such considerations, we assume that two sets of vehicles are, respectively, transmitting EEBL and CICAS safety messages. EEBL and CICAS messages are classified to have low and high priority, respectively. QoSHVCP will force HP messages to be transmitted via V2I, while LP messages using V2V.

As will be presented in the section on simulation results, when the traffic density increases, the message propagation delay decreases due to enhanced connectivity in the network. However, this is somewhat unrealistic due to the fact that when the traffic density increases, the overload on the network infrastructure also increases. This results in a decrease in the bandwidth available for each vehicle, which leads to increased message propagation delays.

In order to avoid this traffic overload in the network infrastructure, a *load-balancing mechanism* is used. We define the *channel utilization*, that is, $\rho(\nu)$, as the percentage of the traffic load in a wireless network, where $\nu$ is the number of vehicles connected to Internet inside an RSU wireless cell. The *channel utilization* is expressed as

$$\rho(\nu) = \begin{cases} \exp\left[\frac{\nu - \nu_{\max}}{\nu_{\max}}\right], & \text{for } \nu \leq \nu_{\max}, \\ 1, & \text{otherwise}, \end{cases}$$ \hspace{1cm} (4)

where $\nu_{\max}$ is the maximum number of vehicles that can be served by the RSU. When $\nu > \nu_{\max}$, the traffic load in the RSU wireless cell will be maximum, and new HP messages cannot be served successfully.

The analytical trend of *channel utilization* is depicted in Figure 2. In the region with a low number of connected vehicles (i.e., $\leq 15$), the analytical trend converges to a low channel utilization percentage (i.e., $\leq 40\%$). In this case, there is no dependence on the channel utilization threshold. For increasing number of connected vehicles, the analytical trend varies for different channel utilization thresholds, which shows an overall availability with different slopes.

Notice that the threshold $\nu_{\max}$ strictly depends on the particular wireless network technology. It needs to be updated constantly and is expressed as $\nu_{\max} \propto (r, B)$, where...
realistic analytical and simulation models; (iii) the channel assumptions and will allow us to incorporate more the objective is to present a framework that will eliminate the infrastructure network overload on the performance; (ii)

Before any overload, any High Priority packets are routed via V2I only depends on (i) the effective transmission data rate within the network infrastructure as the ratio between the message length \( L \) RSU [s] as the transmission time delay for a message propagating within a cluster \( C \) is \( d \) [s] which is defined as the difference between the timestamps of message reception (i.e., \( t_{Rx} \) [s]) and transmission (i.e., \( t_{Tx} \) [s]), respectively

\[
d = t_{Rx} - t_{Tx}.
\]

For a successful transmission of a message of length \( L \) [bit] between two vehicles \( (i,j) \), (5) can be also expressed as the link utilization time from (1), that is, \( d_{(i,j)} \) [s], where \( r \)-th vehicle transmits a message to \( j \)-th vehicle at a transmission data rate \( f_{(i,j)} \) [bit/s], such that

\[
d_{(i,j)} = q_{(i,j)} = \frac{L}{f_{(i,j)}}.
\]

By assuming that the cluster \( C \) comprises of a set of vehicles connected with each other through \( U \) hops (i.e., \( u = \{1, 2, \ldots, U\} \) ), the average propagation time delay within a cluster (i.e., \( d \) [s]) adds delays due to each single link \( (i,j) \) such as

\[
d = \frac{1}{U} \sum_{i,j} d_{(i,j)} = \frac{L}{U} \sum_{i,j} \frac{1}{f_{(i,j)}}.
\]

Analogous to (6), let us consider \( d_{RSU} \) [s] as the propagation time delay within the network infrastructure as

\[
d_{RSU} = \frac{L}{f_{RSU}},
\]

which defines the link between the \( m \)-th and \( (m+1) \)-th RSU as the ratio between the message length \( L \) [bit], and the effective data rate \( f_{RSU} \) [bit/s]. It represents the time necessary to forward a message of length \( L \) between two consecutive RSUs at rate \( f_{RSU} \) [bit/s]. Equation (8) represents the time delay propagation rate within the preexisting Internet network infrastructure.

Each RSU works as a relay node and forwards the message to vehicles crossing its wireless cell. According to Figure 1, we shall also consider the propagation time delay in uplink (downlink), when a vehicle sends a message to an RSU (and vice versa) such as

\[
d_{UP} = \frac{L}{g_{(i,m)}}, \quad d_{DOWN} = \frac{L}{g_{(m,i)}},
\]

where \( g_{(i,m)} \) and \( g_{(m,i)} \) are the effective transmission data rate for the link \( (i,m) \) (uplink), and \( (m,i) \) (downlink), respectively.

From (8) and (9), it follows that the propagation time delay \( d_{V2I} \) [s] for the communication between vehicles and RSUs via V2I only depends on (i) the effective transmission data rates in uplink and downlink (i.e., \( d_{UP} \) [s] and \( d_{DOWN} \))

\[
\text{Figure 2: Percentage of channel utilization in an RSU wireless network for HP messages delivery, with } \gamma_{\text{max}} = [40, 60, 80].
\]
accessible network infrastructure. No connectivity via V2V is assumed to be available during the wireless cell and can connect via V2I and forming a cluster with other vehicles. It enters an RSU (Long-range connectivity). A vehicle is traveling available during this phase.

In a similar way, we define message delivery time for V2V communications (i.e., $d_{V2V} [s]$) as:

$$d_{V2V} = d + \Delta T,$$

where $d [s]$ is the propagation time delay within a cluster, as defined by (7), and $\Delta T [s]$ is the minimum time interval necessary for intervehicle connection at distance $\Delta x [m]$. $\Delta T$ is defined as

$$\Delta T = \frac{\Delta x}{c},$$

where $c [m/s]$ is the radio propagation rate at the speed of light.

Notice that when no connectivity is present (i.e., a vehicle is traveling alone), the propagation time delay is equal to $\Delta T [s]$. In V2V communications, the message delivery time delay drastically increases for low traffic density scenario. In our vision, we can model the overall system as an alternating renewal process where vehicular connectivity structure alternates between three phases as follows.

**Phase 1** (No connectivity). A vehicle is traveling alone in the vehicular grid. It represents a typical totally-disconnected traffic scenario where no connectivity via V2V is available. Moreover, we assume that no connectivity via V2I is assumed to be present during this phase (no network infrastructure).

**Phase 2** (Short-range connectivity). A vehicle is traveling and forming a cluster with other vehicles. V2V connectivity is available within the transmission range of the sender/forwarder. No connectivity via V2I is assumed to be available during this phase.

**Phase 3** (Long-range connectivity). A vehicle is traveling and forming a cluster with other vehicles. It enters an RSU wireless cell and can connect with the associated RSU via V2I. No connectivity via V2V is assumed to be available during this phase. Vehicles are forced to connect to the Internet with accessible network infrastructure.

Each phase is described as follows. During Phase 1, the vehicles are completely disconnected to very low vehicle density and no available network infrastructure. Data packets are cached within a vehicle and traverse the network once a connectivity link becomes available. The minimum time necessary for a vehicle to be connected with a neighboring vehicle is $\Delta T [s]$, as expressed by (12).

When a vehicle is in Phase 2, the messages propagate in a multihop fashion via V2V within the cluster. The transmission time delay to forward a message within a cluster is $d [s]$, which depends on the effective transmission data rates for each hop within the cluster as defined by (7).

We assume that a traditional opportunistic networking in a VANET depends on exploiting connectivity in both Phase 1 and Phase 2. In order to avoid disconnections, the bridging technique connects separated vehicles in Phase 1 with those in Phase 2. It follows that the propagation time delay via V2V (i.e., $d_{V2V} [s]$), respectively, comprises of two components from Phase 1 (i.e., $\Delta T [s]$), and Phase 2 (i.e., $d [s]$).

Finally, in Phase 3, time period necessary for a vehicle to transmit a message via V2I to an RSU is $d_{UP} [s]$, which depends on the RSU’s wireless technology. End-to-end time delay between two separated vehicles for communications via V2I comprises of the uplink (i.e., $d_{UP} [s]$), the inter-RSU link (i.e., $d_{RSU} [s]$), and the downlink (i.e., $d_{DOWN} [s]$) time delays.

By employing such assumptions, we shall define the average transmission propagation time delay (i.e., $d_{avg} [s]$) as the average time delay necessary to propagate a message in a vehicular network, where vehicles are able to opportunistically communicate either via V2V and/or V2I. Basically, the average time delay alternates between (i) the time delay occurring in Phase 1 (i.e., $\Delta T [s]$), (ii) the multihop time delay in Phase 2 (i.e., $d [s]$), and (iii) the time delay in Phase 3 via V2I (i.e., $d_{V2I} [s]$), respectively. Let us denote $T_v^{(n)}$ with $\tau = \{1, 2, 3\}$ the random amounts of time a vehicle spends in one of the three phases during the $n$-th cycle. $T_v^{(n)}$ are independent and identically distributed variable, due to the memory-less assumption on the intervehicular distances, and the expected time spent in the $\tau$-th phase is $E[T_v^{(n)}]$. It follows that the long-run fraction of time spent in each of these phases is

$$p_{\tau} = \frac{E[T_v^{(n)}]}{\sum_{\tau} E[T_v^{(n)}]},$$

where $E[T_v^{(n)}]$ has been assumed to approximate $E[T_v^{(n)}]$.

We are now able to compute the average propagation time delay (i.e., $d_{avg} [s]$), which occurs in a vehicular scenario, where connectivity is alternating between three main phases, such as

$$d_{avg} = p_{\tau=1} \Delta T + p_{\tau=2} d + p_{\tau=3} d_{V2I}.$$ (14)

Each term in (14) represents effective propagation time delay which occurs each time a vehicle is in a given connectivity phase, that is, for $\tau = \{1, 2, 3\}$. The probability that a vehicle lies in one of the three phases can be expressed as the probability that a vehicle is not connected, connected with neighbors and RSUs, respectively.

In order to determine the probability that a vehicle is connected with other vehicles traveling in the same or opposing direction, it is useful to assume that a vehicular grid is discretized in terms of a number of cells; that is, the gap between two vehicles is equivalent to $N$ cells. Basically, we considered two bounds for the cell size, that is, $R$ an upper bound, and $R/2$ a lower bound.
Figure 3: Vehicular grid comprised of $\tilde{l}$-size RSU virtual and wireless V2V cells. The probability that a vehicle is connected via V2V and V2I depends on the cells occupancy.

Figure 3 depicts how the vehicular grid is assumed to be composed of virtual RSU and wireless V2V cells, each of them with a variable size (i.e., $\tilde{l}$ [m]). We consider a cell to be occupied if one or more vehicles are positioned within that cell.

For a vehicle traveling alone on the southbound (northbound), the probability that it will be connected in the Phase 1 via multi-hop with a next vehicle on the southbound (northbound) depends on whether each of the $N$ southbound (northbound) cells within the gap is occupied by at least one vehicle, given by

\[
(p_{s,n})^{N} = (1 - \exp(-\lambda_{s,n}R))^{N},
\]

where $\lambda_{s,n}$ [veh/km] is the traffic density distribution on southbound and northbound, respectively. In this case, the number of cell is $N = 1$ since the gap equals the minimum intervehicle distance, that is, $G = R$ [m]. Equation (15) becomes

\[
p_{s,n} = (1 - \exp(-\lambda_{s,n}R)).
\]

Again, in Phase 2, the vehicles along southbound (northbound) are connected via V2V if each of the $N$ northbound (southbound) cells in the gap is occupied by at least one vehicle. This is an event which occurs with the probability expressed by (15), but the number of cell $N$ is equal to

\[
N = \left\lfloor \frac{G}{R} \right\rfloor,
\]

where $G$ [m] is the gap between two separated vehicles. However, in the event that not all of the $N$ cells in the northbound direction are occupied, the vehicles along southbound are deemed to be disconnected. A message is then buffered in the vehicle’s cache until connectivity is again achieved.

Finally, in Phase 3, the probability that a vehicle traveling in the northbound (southbound) will be connected via V2I with a northbound (southbound) next vehicle depends on if each of the $N$ northbound (southbound) cells in the gap is occupied by at least one RSU, such as

\[
(p_{n,s})^{N} = (1 - \exp(-\lambda_{n,s}R))^{N},
\]

where the number of cell $N$ is

\[
N = \left\lfloor \frac{G}{KR} \right\rfloor,
\]

since we assumed that the RSU wireless networks have a larger cell size than that in the vehicular V2V grid, that is, $\tilde{l} = K \cdot R$ [m], with $K > 0$.

Figure 4 shows the analytical trend of the probability of connected vehicles for three different connectivity phases. The probability has been evaluated for $R = 125$ m; in Phase 2 vehicles are assumed to be separated for $G = 150$ m, while in Phase 3 the RSU wireless cells are greater than V2V cells for $K = 3/2$, and vehicles are separated for $G = 1000$ m.

We can now introduce the following theorem.

**Theorem 1** (average propagation time delay). The average time delay necessary for a vehicle, being driven in a vehicular grid partially covered by a wireless network, to forward a message of length $L$ is

\[
d_{\text{avg}} = p_{s,n}(\lfloor N = 1 \rfloor) \cdot \Delta T + p_{s,n}(N = \left\lfloor \frac{G}{R} \right\rfloor) \cdot d_{\text{V2V}} + p_{n,s}(N = \left\lfloor \frac{G}{KR} \right\rfloor) \cdot d_{\text{V2I}}.
\]
As we assumed two bounds for the cell size (i.e., the upper and lower one, for $\tilde{t} = R$ and $\tilde{t} = R/2$, resp.), the average propagation time delay in (20) will be comprised of a lower bound and an upper bound.

5. Simulation Results

In order to properly authenticate our theoretical model, we performed an extensive simulation. In this section, we compare the delay propagation rates in a VANET scenario using different communication method as defined by the three phases of connectivity previously described in Section 4. We also evaluate the performance of the load-balancing mechanism and observe the improvements in the message delivery delay.

Sections 5.1 and 5.2, respectively, introduce the simulation setup and the obtained results.

5.1. Simulation Setup. We have developed our own simulator using Java, which includes the highway model scenario with 4 different car speeds. The simulator measures the propagation delay as the main performance metric. We considered both asymmetric and symmetric bidirectional traffic flows, where the traffic density on respective southbound and northbound traffic is different and assumed equal. However, in this paper, we assume a symmetric traffic flow, that is, a typical configuration, illustrating the propagation behavior and message transmission performance when cars are faced with during all the three connectivity phases.

We simulated two typical safety applications, that is, the Extended Emergency Brake Light (EEBL), and the Cooperative Intersection Collision Avoidance System (CICAS), corresponding to low and high priority messages, respectively. All messages are being propagated in the vehicular grid. A large number of simulations have been performed so as to decrease any random fluctuation. We assumed an idealistic perfect conditions, that is, no dropped packets while contention or interference occurrence has been introduced. This ideal situation represents the first scenario to simulate in order to understand how delay is affected in the best case. The vehicle density on highways is varied from as low as 1 vehicle per kilometer, up to 100 vehicles per kilometer, and speed ranges from 15 to 35 m/s. These values represent a typical highway condition of a sparse, medium and heavy traffic conditions on the roadways.

The vehicular traffic has been generated using a random exponential distribution which created the intervehicle distances on the highway. The exponential distribution has been largely shown to be in a good agreement with real vehicular traces for uncongested traffic conditions, that is, up to 1000 vehicles per hour. The interarrival time of vehicles is calculated based on the vehicle density and speed of vehicle over the highway. For these reasons, the network connectivity is not always guaranteed. Consequently, at any given time, there is, a nonzero possibility that a partition may exist in the network.

For each scenario, the simulation has been run for 10000 seconds, and the average delay has been calculated from 200 different iterations to account for the randomness. Distance between RSUs is 500 m, and they are distributed uniformly (the placement of RSUs is motivated by the already existing infrastructure at known locations and has not been optimized with respect to connectivity). All RSUs are assumed to be connected over wired or other fixed communication links, for example, Internet, in order to avoid intervehicle disconnections. Hence, any two vehicles out of range from each other can still communicate as long as they both are in the range of an RSU.

Following these parameters, we have simulated our proposed QoSHVCP technique. Firstly, the effect of overload and the resulting channel utilization are considered in the traffic scenario. Packets introduced to the system are randomly assigned either high priority (HP) or low priority (LP). The fraction of HP packets of the total packets in the system is controlled. This fraction is varied from 10% to 40% of the total number of messages.

Finally, a load-balancing mechanism is added in the system. The message propagation delays for HP packets before and after the introduction of the load balancing are analyzed and compared. Complete details about the simulation setup are presented in Table 1.

5.2. Simulation Results. We compared the delay propagation rates in the three different connectivity phases for typical safety applications (e.g., EEBL and CICAS) in a VANET. However, to better understand and validate our simulator, we also included a “limited” Phase 2 that allows transmission of a single direction only. So, the message propagates strictly in an ad hoc hop-to-hop fashion from vehicle to vehicle in a single direction of the highway. This allows results to be analyzed in light of (i) no connectivity; (ii) limited one-direction communication; (iii) vehicle communication that allows transmissions to both directions through bridging; finally (iv) a hybrid QoSHVCP model in which a message

![Figure 4: Probability of connected vehicles in Phases 1–3 versus the vehicle traffic density.](image-url)
can propagate using all possible transmission means. This may be in ad hoc V2V mode or through an V2I infrastructure whenever available.

Due to space limitations, we use the following legends in the graphs: “V2V” represents the single-direction message propagation; “V2V Bridged” represents Phase 2; “QoS HVSP” our hybrid model.

Our results show that, as the vehicle density increases, the delay decreases. However, this happens only in Phases 1 and 2, while in the no-connectivity phase, the delay is constant. As shown in Figure 5(a), the graph represents a typical behavior anticipated in Phase 1; the delay in the no-connectivity phase is the ratio of physical distance covered over the vehicle speed. Thus, the delay is significantly larger as compared to Phases 2 and 3. Notice that the propagation time delay in this phase does not depend on vehicles’ density, since no connectivity is assumed.

Propagation delay also depends on the vehicles’ speed. To better understand this correlation between the delay and the vehicle speed, we have simulated vehicular movement and the delay at four different speeds, relying solely on a single direction of communication. The simulation results unambiguously show that the delay reduces with an increase in the vehicle speed as illustrated in Figure 5(b). As a consequence, Figure 6(a) depicts the average time delay propagation for a message traveling at vehicle speed, and the vehicle density is low. The time delay results in an average delay smaller than that of Phase 1 with no connectivity. Moreover, by modifying the speeds of the vehicles, the maximum time delay changes. As the vehicle density increases, the average delay decreases. This is because when the probability of connectivity among vehicles increases, the message travels faster than the vehicle speed.

Figure 6(a) shows that, in a low density situation, the average delay follows an increasing order. This is clearly expected and helps illustrate correctness of our simulator. The results show that under high-density conditions, majority of the vehicles are interconnected, and the message travels at the radio speed. Beyond this level, effect of increasing the vehicle density does not seem to be beneficial. Obviously, as vehicle density increases, more and more vehicles are connected and majority of time, the message can travel at the radio speed.

In Figure 6(b), we compare the message propagation delay for V2I communications only, since it reflects the best and worst cases of propagation time delay. As a matter of fact, V2I performance is not affected by the vehicle density since it does not rely on any multihop vehicle communications. It is only affected by variations in the uplink and downlink data rates of the network infrastructure. The transmission rates used are a combination of the maximum uplink and downlink rates for the infrastructure in order to demonstrate the impact of different thresholds. The uplink data rate ranges from 0.2 Mbps to 2.7 Mbps, and the downlink from 5 Mbps to 12.2 Mbps. Notice that V2I shows the best (i.e., 5 s) and the worst (i.e., 48 s) time delay cases, respectively, for low (i.e., uplink 0.2 Mbps and downlink 5 Mbps) and high (i.e., uplink 2.7 Mbps and downlink 12.2 Mbps) values of data rates.

As expected, the performance in Figure 6(b) leads to the following two main conclusions.

(i) The delay performance of message propagation within an infrastructure network is highly dependent on the throughput of the infrastructure, which in turn is affected by many factors among which is the overload of the network. Hence, it is important to account for the overload of the infrastructure network, which can, in some cases (i.e., uplink 2.7 Mbps and downlink 12.2 Mbps), result in delays approaching that of basic V2V communication at low vehicle density.

(ii) From the first aspect, follows that the variation in the delay resulting from communication within the infrastructure demonstrates the necessity of incorporating the status of the infrastructure network in the decision making mechanism of the handover scheme.

Comparing Figures 6(a) and 6(b), we are able to establish thresholds for handover between a purely infrastructure-based connection to any of the other options.

Notice that the performance of V2I heavily depends on the cell size of RSUs (i.e., 500 m), as compared to the vehicular radio range (i.e., 200 m). With such a setting, the superiority of QoS HVSP, as compared with V2V, is limited, as can be seen in Figure 6(a). Better results of QoS HVSP may be achieved by using larger RSU V2I cell size; however, this depends on the technology that is being used for the network infrastructure (i.e., cellular network, Wi-Fi, or another technology).

After considering the overload factor on the network, the performance of V2V and V2I is presented in Figure 7. As mentioned earlier, when the number of vehicles communicating through V2I increases, the overload on the infrastructure network increases, and this causes the delay performance of the network to deteriorate. This is shown in Figure 7, where the V2I delay increases when the number of vehicles reflected by the vehicle density increases beyond a certain threshold. On the other hand, V2V performance is not affected by the overload of the network. Basically, two factors should affect the delay of the delay performance in V2V communication at higher vehicle density, such as (i) there is an increased connectivity coverage between the vehicles due to increase in the vehicle density, which makes the

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length of simulation</td>
<td>10000 seconds</td>
</tr>
<tr>
<td>Number of runs</td>
<td>200</td>
</tr>
<tr>
<td>Estimated vehicular Tx range</td>
<td>[0, 200] m</td>
</tr>
<tr>
<td>Infrastructure Tx range</td>
<td>500 m</td>
</tr>
<tr>
<td>Packet size</td>
<td>400 bytes</td>
</tr>
<tr>
<td>Vehicle speed</td>
<td>15 to 35 m/s</td>
</tr>
<tr>
<td>Vehicle density</td>
<td>1 to 100 veh/km</td>
</tr>
<tr>
<td>Load-balancing threshold</td>
<td>40 veh</td>
</tr>
</tbody>
</table>

**Table 1: Parameter setup used in simulations.**
message to propagate quicker across the highway, which leads to an improvement in the end-to-end message propagation delay; (ii) we ignored the collision and contention due to the increased number of packets transmitted due to increased vehicle density. In our future work, we plan to relax this assumption and simulate the effect of the contention and packet collision on the V2V delay.

As mentioned in the previous section, we considered applications with high and low priority messages in our simulations. Due to the nature of the infrastructure service, high priority messages are communicated using V2I because of the network coverage and guaranteed service, while low priority messages are communicated through V2V. We vary the percentage of high priority messages versus the percentage of

**Figure 5:** (a) Average message propagation delay for increasing vehicle density, and speed for Phase 1 (no connectivity) at different speeds (i.e., 15, 20, 25, and 35 m/s). (b) Average message propagation delay for increasing the vehicle density at different traffic speeds using V2V communication.

**Figure 6:** (a) Average message propagation delay for increasing vehicle density for V2V, V2V bridged and QoSHVCP at fixed speed. (b) Average message propagation delay for increasing vehicle density in V2I communications at different uplink and downlink rates.
low priority messages in the system. When the percentage of the HP messages increases, the overload on the infrastructure also increases which causes the delay performance of the V2I to deteriorate. In Figure 8, we compare the message propagation delays for HP and LP messages versus different vehicle densities. As expected, when the probability of HP messages increases, the message propagation delay increases for high vehicle densities. Figure 9 compares the message propagation delays for the same case after introducing the load balancing mechanism; this approach provides a delay reduction up to its minimum value also for high vehicle density.

All the previous results have shown the average delay versus the vehicle density on the highway. However, we would like to investigate the amount of message propagation delay that each packet experiences both with and without load-balancing mechanism, in a typical dense traffic scenario. Figures 10, 11, and 12 present the number of packets experiencing different message propagation delays at the vehicle density of 65 [veh/km], and for high priority packets with probability 0.2, 0.3, and 0.4, respectively. We notice that majority of the packets experience a very minimal delay with the load balancing mechanism in effect compared
to communication without our proposed mechanism. The highest peaks of packets are with the load balancing, providing very low delays; communications with no load-balancing show a low number of packets with increasing values of delay. The high the HP message probability, (i) the high the number of low delay packets with load balancing (i.e., peaks from 920 packets at 2 s for 0.2 HP message probability, up to 1800 at 2 s for 0.4 HP message probability); (ii) the high the number of packets experiencing increasing delays without load balancing (i.e., from 250 packets at 10 s for 0.2 HP message probability, up to 400 at 25 s for 0.4 HP message probability).

6. Conclusions
In this paper, we have investigated QoSHCVP, a Hybrid Vehicular Communication Protocol with QoS prioritization, relying both on V2V and V2I approaches. In order to avoid disconnections and maintain a seamless connectivity, vehicles should exploit any available connectivity link present in the vehicular grid. Based on a delay-based decision criterion, our approach represents a handover mechanism between V2V and V2I.

In this paper, we have shown effectiveness of our proposed technique in terms of message delivery time delay. Simulation results confirm our analytical work on how hybrid approach enhances the connectivity support, especially in high-mobility and low-density traffic scenarios as compared to traditional opportunistic V2V techniques. The effect of traffic overload on the message propagation delay has also been presented. The failure of achieving QoS requirements has been taken care of by our load-balancing mechanism, that decreases the message propagation delay of high priority packets regardless of the network overload.

Acknowledgment
This work is supported in part by the NSF under cooperative agreement EEC-0812056. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References


