Reliable Periodic Safety Message Broadcasting in VANETs Using Network Coding

Behnam Hassanabadi and Shahrokh Valaee, *Senior Member, IEEE*

***Abstract*—Reliable local information dissemination is the primary concern for periodic safety broadcasting in VANETs. We propose a sublayer in the application layer of the WAVE stack to increase the reliability of safety applications. Our design uses rebroadcasting of network coded safety messages, which considerably improves the overall reliability. It also tackles the synchronized collision problem stated in the IEEE 1609.4 standard as well as congestion problem and vehicle-to-vehicle channel loss. We propose a discrete phase type distribution to model the time transitions of a node state. Based on this model, a tight loss probability upper bound for the network coding algorithm is derived. Numerical results based on our analysis as well as ns-2 simulations show that our method outperforms the previous repetition-based algorithms.**

***Index Terms*—Ad-hoc networks, network architectures and protocols, quality of service assurance, multiple access techniques, systems and services.**

I. INTRODUCTION

VEHICULAR Ad-hoc NETworks (VANETs) have emerged as a potential framework for future active safety systems. An architecture for Wireless Access in Vehicular Environment (WAVE) is presented in the IEEE 1609.0 standard. IEEE 802.11p has been approved as an amendment to IEEE 802.11 standard and specifies the MAC layer enhancements for vehicular environment. The significant changes are increased maximum transmission power and reduced channel bandwidth to 10Mhz, which provide a more reliable communication. The IEEE 1609.4 standard also specifies some enhancements to the MAC layer to support multi-channel operation. Both IEEE 802.11p and 1609.4 are part of the WAVE architecture.

Periodic broadcast and its related safety applications are one of the major driving forces for implementing VANETs [1]. In VANETs, a safety message is periodically generated (10Hz frequency) and transmitted to one-hop neighbour vehicles. These periodic "heartbeat" messages are the building blocks of many safety applications. By aggregating this local information, each vehicle can construct and maintain a local neighbourhood map that can be utilized by safety applications. The message contains the state of the vehicle, which consists of various sensor readings such as location, velocity, acceleration, etc. The goal of the communication is to provide a reliable up-to-date neighbourhood information. Having an up-to-date local map could prevent accidents and collisions. Also this extra information assists the driver to choose alternative driving strategies such as taking over or turning.

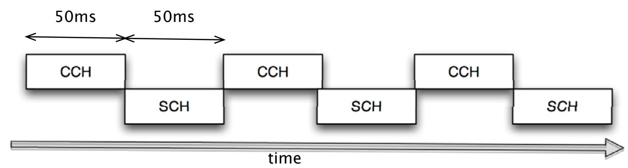

Fig. 1. IEEE 1609.4 periodic channel switching.

IEEE 1609.4 explains how channel coordination is performed for a WAVE device with a single radio [2]. The time is divided into 100ms *sync* intervals. Each *sync* interval consists of a Control CHannel (CCH) and a Service CHannel (SCH) (Fig. 1). Every 50ms the WAVE device switches to CCH (channel 178). The periodic broadcast of Wave Short Messages (WSM) takes place in the CCH intervals. SCH intervals are used for IP packets. Since WAVE devices should be in the same channel in order to be able to communicate, the channel switching for all nodes is synchronized. The synchronization mechanism is detailed in the IEEE 1609.4 standard [2].

The periodic broadcast is shown to perform well for low node densities (less than 10 nodes) [3]. In a dense network, however, congestion becomes a serious problem. It can produce excessive number of collisions and result in unacceptable reliability measures for safety applications.

In [4] and [3], the authors have proposed a congestion control mechanism based on the channel occupancy. The message rate is adjusted in conjunction with the transmission range to alleviate the congestion problem. Other than the message rate control, adaptive control of the contention window size based on the node density is another way to avoid congestion [5]. The congestion problem is also discussed in IEEE 1609.4 where many nodes have a newly generated WSM for transmission and they all switch to CCH at the same time (synchronized collision). In this scenario, collision is expected if two nodes pick the same timeslot in the contention window. Naturally, for high node densities the collision probability increases.

Aside from the congestion problem, unreliable vehicle-to-vehicle channel and channel errors are other concerns for safety applications. In the IEEE 802.11p broadcast mode there is no acknowledgement. Also, unlike the unicast transmissions, there is no retransmission for lost messages. In [6], several repetition based access schemes are proposed to

Manuscript received December 19, 2012; revised June 3 and September 17, 2013; accepted October 30, 2013. The associate editor coordinating the review of this paper and approving it for publication was J. Huang.

The authors are with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, M5S 2E4 Canada (e-mail: {behnam, valaee}@comm.utoronto.ca).

Digital Object Identifier 10.1109/TWC.2014.010214.122008

guarantee a low message loss probability. Although message repetition improves the reliability, it can potentially aggravate the congestion problem.

We previously studied the gain of network coding in safety applications [7], [8]. We showed, through simulations, that network coding can significantly improve the successful message delivery. In this paper, we extend our previous results through analysis, and propose a more comprehensive design. We propose a sub-layer in the application layer of the WAVE architecture (Fig. 2). This sub-layer handles the periodically generated messages in the application layer to improve the overall reliability of the updated local neighbourhood map and cope with the congestion problem and channel loss. Solid arrows in Fig. 2 represent the handshakings between the sublayer and the rest of the networking stack. The dashed arrow represents the timing feedback from the MAC layer over the intermediate layers. Since our sublayer is in the application layer it needs the timing information for synchronization. For example, the synchronization is necessary for the congestion control algorithm in order to transmit in predefined subframes. We study how and when the message repetition can be helpful to achieve more reliability. In our scheme, each node not only can send its own WSM, it can also transmit a random linear combination of previously received WSMs to cooperatively help delivering all WSMs.

The low-level design and the exact description of the protocol is beyond the scope and space of this paper, which aims to tackle the problem analytically. The main contribution of this paper is the analytical study of random linear network coding in the periodic broadcast of “heartbeat” messages. To the best of our knowledge, this is the first analytical study of the application of random network coding for safety messages in vehicular networks. Moreover, we propose a novel method based on pseudo random number generator to minimize the network coding overhead. This is specially critical for small safety message size. Repetitive safety message rebroadcasting is an efficient way to maximize the reliability, on the other hand congestion is a serious problem in vehicular network. We also show, through simulation, that in a congested network, network coding can substantially minimize the safety message loss probability. Finally, the network coding algorithm is implemented in ns-2 which provides a more realistic simulation framework.

In Section II, we review some of the related works. Our system model and performance metric definition are presented in Section III. In Section IV, we present the proposed reliability sublayer design. Analysis of network coding can be found in Section V. Section VI contains the numerical and simulation results. Finally, Section VII concludes the paper.

II. Related Works

The two main components of the proposed sub-layer in this paper are safety message retransmission and the use of network coding for safety message broadcasting. In the following subsections, first we review the previous works on safety message repetition, then we discuss some of the previous works on the application of network coding in VANETs. We also review the applications of network coding in gossip-based algorithms, which are closely related to the application in this paper.

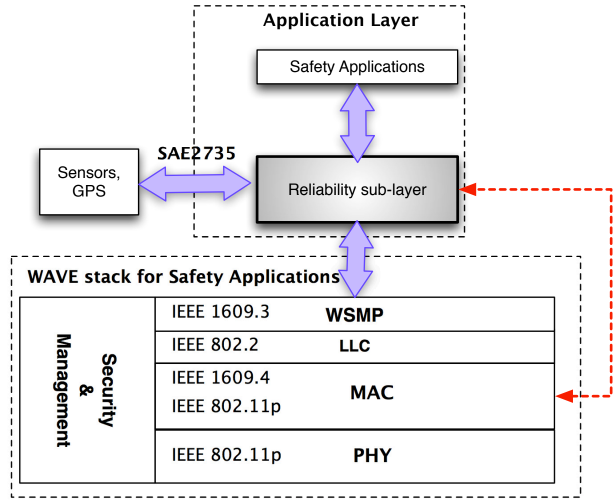

Fig. 2. Reliability sub-layer: interfaces with WAVE Short Message Protocol (WSMP), sensors, and safety applications.

A. Message Rebroadcasting

Repetitive broadcast of safety messages in VANETs was first proposed in [9] and [6]. When a safety message is generated it should be delivered to neighbours within the message lifetime. The time is slotted and the message lifetime is assumed to be L timeslots. In IEEE 802.11p, safety messages are only transmitted once during a time frame. This is due to the fact that in IEEE 802.11p broadcast mode there is no retransmission or acknowledgement. To assure the message delivery by the end of its deadline, several random rebroadcasting schemes have been proposed. The basic idea is to rebroadcast the safety message multiple times in order to ensure reliability. In Synchronized Fixed Repetition (SFR), w timeslots are randomly chosen (out of L) for retransmission. In Synchronized Persistent Repetition (SPR) at each timeslot a message is transmitted with probability p. To limit the number of collisions, Positive Orthogonal Codes (POC) is proposed in [10] and [11]. The retransmission pattern of each node is assigned based on predetermined binary codes. The 1’s represent the transmitting timeslots in the frame. Any pairwise shifted version of two POC codewords has limited correlation. This limited correlation has been shown to increase the message delivery ratio by limiting the number of collisions.

In this work we investigate the benefits of utilizing random linear network coding in conjunction with a repetitive broadcasting scheme. More specifically we propose the Synchronized Persistent Coded Repetition (SPCR) algorithm. To see the potential benefit of network coding in our application, let us consider the example in Fig. 3. In their first transmission opportunity, U_1 and U_2 broadcast their safety messages (m_1 and m_2 respectively). Suppose m_1 is received but m_2 is not received by U_3. If U_1 has a second transmission opportunity during the CCH interval, it rebroadcasts m_1 which is not helpful for U_3. If instead U_1 broadcasts a random linear

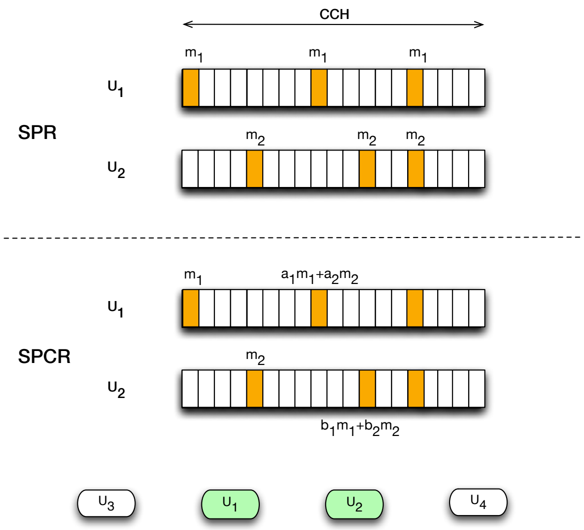

Fig. 3. Potential benefits of network coding.

combination of m_1 and m_2 and U_3 receives the transmission, it can decode m_2. U_2 can act in the same way to deliver m_1 to U_4. However, if U_3 does not receive the first U_1's transmission but receives the second transmission, if the second transmission is coded, U_3 can not decode any messages. This shows that network coding can be potentially beneficial for message delivery but a careful study is needed to provide insights into its performance gain. In this paper, we analytically quantify the gain of network coding.

B. Network Coding

In [12], authors have proved that random linear network coding could asymptotically achieve the multi-cast capacity in a lossy wireless network. The work in [12] has not been proposed for vehicular networks, but is the closest, in terms of the techniques, to our work. This well-known result cannot be extended to the application of network coding in safety message broadcasting in VANETs. In a typical vehicular network, the number of vehicles in one-hop neighbourhood does not increase asymptotically. In addition, the analysis in [12] only considers long-term throughput, but for delay sensitive safety messages short-term throughput or the successful message reception in a small time window is of interest. This is the main drawback of the analysis scheme in [12] for our application which we address in this work. More specifically, we derive an upper bound for the message loss probability in the CCH interval (50ms according to IEEE 1609.4 standard) which can not be inferred from the analysis in [12].

1) Vehicular Networks: Most of the previous works on the application of network coding in vehicular network deal with the content distribution from a Road Side Unit (RSU) to multiple On Board Units (OBUs) or consider only throughput performance ([13],[14],[15]). To the best of our knowledge there are only few works on network coding application in safety message broadcasting.

In [16], [17], Symbol-Level Network Coding (SLNC) is utilized for multimedia streaming from RSUs to vehicles. In [18], SLNC is used for content distribution, in order to maximize the download rate from the access points. SLNC is shown to achieve better performance compared to packet level network coding in unicast transmissions. This is effective for large packet sizes. For small safety messages, if each message is decomposed into smaller symbols, the network coding overhead can be comparable to the symbol size leading to network inefficiency. To address this drawback, our proposed network coding algorithm does not need any message decomposition and the network coding overhead is minimized through a novel algorithm proposed in Section V.A to further maximize the reliability.

In [19], [20], authors propose a XOR-based network coding in which each user blindly picks a number of original messages for encoding. It is assumed that Conditional Reception Probability (CRP), defined as the probability that a packet is received by a neighbour vehicle given it is successfully received by the transmitter, is known. The optimal number of packets to encode is then determined based on the CRP. The optimized metric is the average number of pairwise recoveries. This means, if vehicle A is transmitting, the number of XORed packets is optimized in order to maximize the expected number of recovered packets by vehicle B. However, in practice, after vehicle A's transmission, multiple destinations can potentially decode and a more reasonable metric should be the expected number of decoded packets by all neighbours. The main drawback of the XOR-based coding schemes is the need for feedback information which can consume a fair amount of network resources. Unlike the coding scheme in [19], [20], the proposed random linear network coding algorithm in this paper does not need any side information. Even though the network coding does not benefit from the feedback information, it can still achieve the same reliability as XOR-based network coding algorithms as has been shown in [8].

In [21], [22], a 1-D vehicular network is considered. A source node sends a file, consisting of M packets, to all other vehicles. An ideal scheduler is assumed. The source keeps transmitting random linear combinations of packets till another node receives M linearly independent packets and acts as the source node for the second round. The probability density function of completion delay for a 3-node case scenario is then computed. For the 3-node scenario, node 0,1 and 2 are placed in sequence and node 0 has all the packets. In the first round, node 0 starts broadcasting random linear combinations of the packets. Depending on the channel, it takes a random time T_1 for node 1 to collect all M linearly independent coded packets. Meanwhile node 2 has overheard $C < M$ coded packets. In the second round node 1 starts broadcasting till node 2 collects M linearly independent coded packets. The second round takes T_2 seconds. The probability density function of $T_1 + T_2$ is derived and has been compared to the uncoded scenario. The proposed analysis scheme in this paper is more general and is not limited to constrained topologies. Moreover, the analysis in [21], [22] does not consider the repetitive broadcast of the safety messages. So in this paper we not only provide a general analysis framework but also consider the more challenging repetitive broadcasting traffic model.

2) Gossip Algorithms: In the context of gossip algorithms, in [23] the problem of disseminating k messages in a large network of n nodes is considered. At each timeslot, each node selects a communication partner in a random uniform fashion and only one message is transmitted, while in the broadcast scenario potentially more than one node can receive the message. In [23], the network is large and asymptotic bounds are considered. In contrast, in a typical VANET topology n does not increase asymptotically.

For periodic safety message broadcasting, a fast local information broadcast is of interest. In the gossip algorithms studied in [23], at each timeslot, each node can contact at most one neighbour while a node can be contacted by multiple nodes. Our communication mechanism follows the opposite scenario: each node can potentially communicate with multiple nodes, but if a node is contacted by multiple nodes simultaneously, there will be collision. In the PULL mechanism [23], a node i contacting node j leads to a transmission from node j to node i. If all the nodes contact node j, it transmits a message to all nodes, which can be cast as a broadcast transmission. However, in a single cell scenario if more than two nodes are contacted, transmissions collide in the wireless channel while in the communication model in [23] transmissions are successful.

The information dissemination considered in [23] intends to deliver global information through local communication. However, in our problem, fast local information dissemination is done through local communication and vehicles are not specifically interested in global information.

Gossip algorithms for complete graphs have linear stopping time as compared to $O(nlog(n))$, the complexity of the sequential algorithms. Most recently it has been shown that the stopping time of algebric gossip (for EXCHANGE gossip algorithm) is $O(\Delta n)$ in which Δ is the maximum degree of the network graph [24].

More recently in [25], broadcast gossip algorithms are proposed for sensor networks to compute the average initial nodes measurements over the network. It is shown that the broadcast gossip almost surely converges to a consensus. In safety broadcasting however, nodes are interested to receive the state information of all neighbour nodes not to achieve consensus over the network.

III. System Model and Performance Metrics

A. System Model

We consider a cluster of N nodes in the radio communication range of each other. At the start of each CCH interval (Fig. 1) each node has a safety message consisting of its state information such as the GPS location, velocity and braking status. Each user needs to receive the messages of the active nodes by the end of each CCH interval in order to update its neighbourhood map. Due to the congestion problem not all the nodes should transmit. Active nodes are the ones that transmit their generated messages and are selected according to a congestion control algorithm. This will be discussed further, in the next section.

The IEEE 802.11p broadcast mode does not have a retransmission mechanism. The presented rebroadcasting algorithms in [9] do not take into account the recently emerged congestion problem in VANETs. In this section, we study how a rebroadcasting scheme performs when there is both congestion and channel loss.

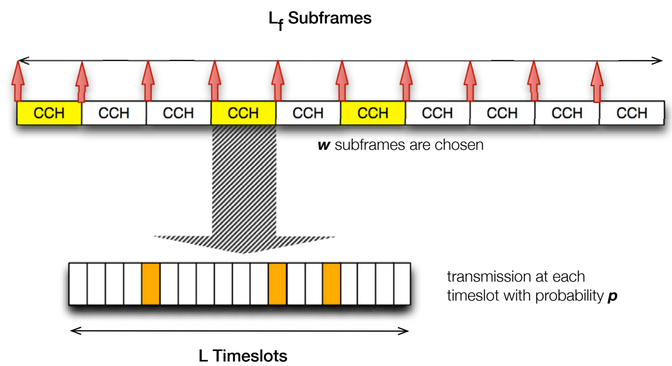

Fig. 4. Message retransmission in active CCHs.

At first glance, the use of message retransmission, in the presence of congestion, seems counter intuitive. On the other hand, since we intentionally drop some messages to avoid congestion, further message losses might not be tolerated. Based on the presented congestion control algorithm in the previous section, the number of active nodes in each subframe can be controlled and therefore retransmission can still be a viable solution to alleviate message loss.

We assume each subframe is slotted and the size of each timeslot is equal to the message transmission time $\frac{M}{R}$. We consider the Synchronized Persistent Repetition (SPR) scheme [9]. At each timeslot, each active node retransmits the generated message with probability p. We define a frame to consist of L_f CCH intervals. We call a CCH interval a subframe. We assume nodes are synchronized with synchronization information inferred from the MAC layer (Fig. 2). Based on the congestion control algorithm, each node is active in w out of L_f subframes (Fig. 4). If there are n active nodes in a subframe it is easy to show that the optimal p is $\frac{1}{n}$. We assume each node attaches L_f bits to its message to represent its activity pattern in each subframe inside the frame. The 1 bits represent the subframes in which the node is active. Also we assume nodes do not change their transmission pattern. Since each node knows its neighbours and theirs transmission pattern based on the received messages, it knows the number of active nodes in each subframe and can adjust its retransmission probability accordingly.

An erasure channel is assumed such that an uncollided transmission is received from node i with probability $1 - p_e$. The erasure channel has also been previously assumed in the context of safety communications in vehicular networks [26], [27]. It has been shown that Nakagami distribution with properly estimated parameters would be a more realistic channel model for vehicle-to-vehicle communiations ([28], [29], [30]). To make our analysis tractable, we assume a symmetric erasure channel. This assumption does not consider higher reception probabilities for closer neighbours (unlike the Nakagami channel model) and it essentially underestimates the successful message reception probability [27]. In our simulations, we will study the Nakagami channel and will define an appropriate erasure probability by averaging the probability of reception in the Nakagami channel over a certain interval that corresponds to the coverage range in the WAVE standard.

B. Performance Metrics

1) Probability of Success: We define $P_s(n)$ as the probability of successful reception of all n messages by an inactive node if there are n active nodes. If a node receives all n messages, it can construct a full neighbourhood map based on the location and other sensory information included in the message. Since an active node does not need to receive its own message, the successful reception probability by an active node is lower bounded by $P_s(n)$. If there are no inactive nodes, we can consider a virtual inactive node and the resulted $P_s(n)$ will be a lower bound.

2) Map Throughput Metric: Safety applications in VANETs require distinctively defined network performance metrics compared to MANETs. For safety critical applications, a vehicle needs to receive the messages of all its neighbours by the end of a CCH interval in order to make sure of safety. A message delivery is successful if it provides meaningful safety information about the neighbourhood. Even if $n-1$ messages are received successfully, the missing vehicle state can be hazardous and the received messages do not guarantee local safety. So a delivery is defined to be successful only if the messages of *all* other neighbour nodes are received successfully. Hence, for this particular application, throughput is naturally defined as $\frac{nP_s(n)}{100ms}$.

We define the Map Throughput Metric (MTM) as the normalized system throughput:

$$MTM = nP_s(n),$$

which accounts for the average number of successful messages received in a *sync* interval.

IV. Safety Message Reliability Control

This section demonstrates different components of the proposed sublayer in Fig. 2. The message is generated based on the sensors output and according to the dictated format by SAE J2735 (Society of Automotive Engineers DSRC Message Set Dictionary standard). The header for the new sublayer is essentially the network coding overhead which we will explain in Section V.A. Instead of passing the message directly to the WAVE stack, our sublayer combines the received messages, attaches the network coding overhead (sublayer header) and sends it to the lower layer.

The proposed reliability sublayer consists of three parts. The congestion control algorithm determines if a node should be active in a given CCH interval. If the node is not active in a CCH interval, the generated message will be dropped. If the node is active, it probabilistically transmits at each time slot during a CCH interval. We assume that the back off parameters of the IEEE 802.11p have been set to minimum and the physical carrier sensing is disabled. This can be done through the wireless card driver interface and ensures an instant transmission when a message is sent to the lower layer. The congestion control algorithm filters the generated messages in order to reduce the number of active nodes in each CCH interval. The Synchronized Persistent Coded Repetition (SPCR) algorithm describes the main functionality of the sublayer for transmission. At each time slot a random linear combination of all the queued messages can be transmitted by an active node.

The following subsections detail the functionality and analysis of these components. First, we introduce the congestion control algorithm that limits the number of active nodes in a CCH. Then, the analysis for message retransmission is presented. Finally, message coding based on random linear network coding is used to further maximize the success probability.

A. Congestion Control

When the number of vehicles in the cluster, N, is large the message reception probability drops considerably. We first derive the $P_s(n)$ for IEEE 802.11p to see how it behaves by increasing the number of vehicles in a cluster.

Theorem 1. *The success probability for IEEE 802.11p broadcast mode in a symmetric erasure channel with n active nodes at the beginning of the CCH, and the contention window size, CW, can be written as:*

$$\begin{aligned} P_s(n) &= \prod_{i=1}^{n-1}\left(1-\frac{i}{\min\{X, CW-1\}+1}\right) \\ &\quad \cdot \left(\frac{\min\{X, CW-1\}+1}{CW}\right)^n \cdot (1-p_e)^n, \end{aligned}$$

where X is the number of idle timeslots in a subframe in which n packets have been successfully transmitted, and is given by:

$$X = \lfloor \frac{T_{CCH} - T_g - n(T_h + \frac{M}{R} + AIFS)}{\sigma} \rfloor$$

where T_{CCH} is the CCH duration, T_h is the PLCP duration, M is the Message size, R is the Channel rate, σ is the timeslot duration, T_g is the Guard time, AIFS is the Arbitration interframe spacing, and p_e is the Erasure probability.

Proof: To find the success probability, we should find the probability that all nodes pick distinct backoff counters (C_i) and all transmissions are completed within the CCH period T_{CCH}. Each successful transmission takes $(T_h + \frac{M}{R} + AIFS)$ seconds. So the maximum backoff counter must be less than $X = \lfloor \frac{T_{CCH} - T_g - n(T_h + \frac{M}{R} + AIFS)}{\sigma} \rfloor$. The success probability can be written as follows: $\forall\ i, j$: $1 \leq i, j \leq n$ *and* $i \neq j$,

$$\begin{aligned} P_s(n) &= Pr(C_i \neq C_j \wedge \max\{C_i\} \leq X) \\ &\quad \cdot (1-p_e)^n \\ &= Pr(C_i \neq C_j | \max(C_i) \leq X) \\ &\quad \cdot Pr(\max(C_i) \leq X) \cdot (1-p_e)^n. \end{aligned}$$

The first term is the complement of the Birthday Paradox Problem [31] when we have n people and $\min(X+1, CW)$ days in a year and is equal to:

TABLE I
PHYSICAL AND MAC LAYER PARAMETERS: IEEE 802.11P

R	3Mbps
σ	16μs
T_h	40μs
$AIFS$	32μs
T_g	0s
CW	1023

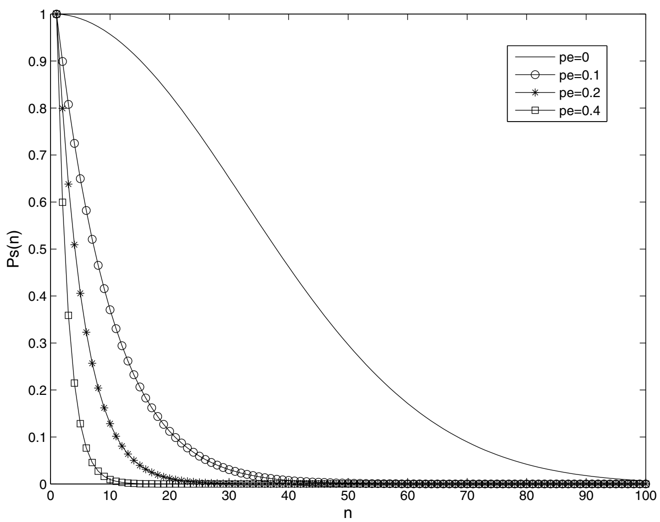

Fig. 5. Success probability vs. number of nodes: IEEE 802.11p broadcast mode.

$$\prod_{i=1}^{n-1}\left(1-\frac{i}{\min(X,CW-1)+1}\right)$$

The second probability is the probability that all C_i's are less than or equal to X. ■

To see how the IEEE 802.11p broadcast mode performs when there is traffic congestion, we evaluated $P_s(N)$ with the parameters in Table I, and for various erasure probabilities. The message size is assumed to be 200 Bytes. Fig. 5 shows how the probability of success declines by increasing the number of nodes N. Even when there is no channel loss ($p_e = 0$) the success probability degrades to less than 0.3 for 50 nodes. For higher erasure probabilities the congestion problem is more pronounced.

Ideally, we are interested in delivering all N messages with high probability, but for higher traffic densities, it is unattainable. The main solution to the congestion problem is message rate control. Rate control algorithms drop some of the messages to reduce congestion. Nonetheless, this is the only way to manage excessive collisions and provide an overall acceptable reliability.

Depending on the system requirements and constraints, congestion control can be implemented both in the MAC layer or in the application layer. Since the IEEE 802.11p is published and the network cards based on this standard is already in the market, solutions that are not based on MAC modifications are more attractive. Several proposed effective congestion control algorithms for VANETs are implemented in the application layer [3-4].

Here, we consider a simple rate control algorithm, which randomly filters the load. At the beginning of every frame each node randomly picks w subframes and only transmits in those subframes. Hence, the number of active nodes in a subframe is a random number $n \leq N$. Unlike the congestion control algorithm in [4], in which the rate is controlled based on the channel occupancy feedback, the feedback parameter in our method is the number of neighbours N, which is derived based on the most up-to-date neighbourhood map.

Since MTM is a random variable, we use $E(MTM)$ for performance comparisons. The w that maximizes the $E(MTM)$ is set based on N. There are $\binom{L_f}{w}$ ways to choose w subframes for each of N users. In order to have n active users in a subframe, $N-n$ users must choose w and the n active nodes must choose $w-1$ subframes out of the remaining $L_f - 1$ subframes. Therefore, the expected value of MTM is:

$$E(MTM) = \sum_{n=1}^{N} \frac{\binom{N}{n}\binom{L_f-1}{w}^{(N-n)}\binom{L_f-1}{w-1}^{n}}{\binom{L_f}{w}^{N}} nP_s(n). \qquad (1)$$

By maximizing $E(MTM)$, the optimal value for w can be found.

B. Message Retransmission

Here, we find the $P_s(n)$ for the retransmission scheme presented in Section III. All users want to collect the messages of all n active nodes by the end of the subframe. At each timeslot, and independent of other nodes, the transmitted packet of an active node does not collide with other transmissions with probability $p_s = p(1-p)^{(n-1)}$. If we consider the messages as coupons, our problem can be seen as a modified version of the coupon collection problem [31]. The completion time D (reception delay of all n messages) can be divided into n rounds such that by the end of the j^{th} round the j^{th} distinct message is received. The duration of the j^{th} round, D_j, has a Geometric distribution with the success probability $p_j = (n-(j-1))p_s(1-p_e)$. The completion time D is the sum of n Geometric random variables with distinct success probabilities and can be modelled as a discrete phase type distribution $PH_d(\tau, T)$ with the following transition matrix:

$$T = \begin{bmatrix} 1-p_1 & p_1 & & & & \\ & 1-p_2 & p_2 & & 0 & \\ & & \ddots & \ddots & & \\ & & & \ddots & \ddots & \\ & 0 & & & 1-p_{n-1} & p_{n-1} \\ & & & & & 1-p_n \end{bmatrix}.$$

Using the CDF of D, the probability of failure (loss), $P_l(n)$, can be computed as:

$$\begin{aligned} P_l(n) &= Pr(D > L) = 1 - F_D(L) \\ &= \tau T^L \mathbf{1}. \end{aligned} \qquad (2)$$

where $\tau = [1, 0, \cdots, 0]$ and $L = \lfloor \frac{T_{CCH} * R}{M} \rfloor$ is the number of timeslots in a subframe.

C. Message Coding

In this section, we propose a novel algorithm that uses random linear network coding in combination with message rebroadcasting. Based on the introduced repetition-based scheme in the previous section, each node potentially has multiple transmission opportunities in a subframe. The same copy of the message is retransmitted to account for the channel loss. However, a node can linearly combine the already heard messages and transmit the coded message. In this section, we consider the application of random linear coding together with SPR. We call this algorithm Synchronized Persistent Coded Repetition (SPCR).

The random linear coding algorithm is simple: each node enqueues all the received messages and when it has a transmission opportunity based on its retransmission pattern it broadcasts a random linear combination of all the already received messages in its queue with the coefficients in GF(q) (Galois field with order q). At the end of the subframe, if the node has n linearly independent coded vectors it can decode all the original packets. Next, all nodes empty their queues and start a new transmission for the next subframe.

To explain the overall structure of the reliability sub-layer, here we give a summary of how the different components work together. Each frame consists of L_f subframes (CCH intervals). The congestion control determines in which subframes a node is active (w subframes chosen at random). In the active subframes, nodes start broadcasting a random linear combination of all the messages in their queues. Initially, each node only has its own message in the queue. As times goes by and nodes overhear other transmissions, their subsequent broadcasts is not only their messages but a random linear combinations of a subset of all messages (depending on their queue content). Algorithm 1 outlines the transmission algorithm of the sublayer. All nodes actively listen to the channel, when they are not transmitting, and queue all the received packets. The queue resets at the end of each CCH interval.

Algorithm 1 Transmission Algorithm

```
for every subframe do
    message generated
    (congestion control:)
    if node is not active then
        drop the message
        continue
    end if
    (SPCR:)
    for every time slot in a subframe do
        if transmitting (with probability p) then
            create a random linear combination of the messages in the queue
            broadcast the coded message
        end if
    end for
    reset the queue
end for
```

V. Network Coding Analysis

Let us consider a single cell of n active nodes (u_1, u_2, $\cdots$, u_n). The time is slotted and the nodes are synchronized. Each subframe consists of L timeslots. At the beginning of each subframe each active node generates a message that should be received by all other nodes within the subframe. The generated message is retransmitted several times during a subframe. When a node has a transmission opportunity, it transmits a random linear combination of all already received packets in its queue. If ${m_i}^k$ represents the k^{th} retransmission of u_i's message, then the ${m_i}^k$ can be expressed as a linear combination of all the original messages:

$$ {m_i}^k = \sum_{j=1}^{n} c_j m_j \tag{3} $$

in which c_j is a random coefficient in $GF(q)$ and m_j is the original generated message of node u_j. The coefficient vector $\mathbf{c_i}^{\mathbf{k}} = (c_1, c_2, \cdots, c_n)$ is a vector in q^n. Note that some of c_i's can be zero. If a node receives n linearly independent equations, it can decode all the original messages. The coefficient vectors for the original messages at the beginning of a subframe are the standard basis for q^n.

A. Coding Overhead

In random linear network coding, the random coefficients should be attached to the coded message. In a congested network with large n the coding overhead can be comparable to the size of the message. For example, for $GF(2^8)$, if there are 200 vehicles in a cluster, the maximum coding overhead will be 200 Bytes, which is comparable to the message size of 200-500 Bytes.

To reduce the overhead, in [32], the authors have suggested to attach the seed of the random number generator. The seed specifies the sequence of random coefficients in a coded message. New coded messages can be generated from received coded messages. In this case, the seeds of all the included coded messages should be attached to the new coded message. This can potentially result in excessive overhead. As a result, the introduced algorithm in [32] can only encode the uncoded messages. However, in our algorithm, we need to encode all the received coded messages. In the following, we show how we can indeed find the corresponding seed to the new coded message which considerably reduces the overhead.

Linear Feedback Shift Registers (LFSRs) are an efficient way for implementing Pseudo Random Number Generators (PRNGs) in a Gallois field. A LFSR implementation based on a primitive polynomial of $GF(2^m)$ has m states and has a period of $2^m - 1$. The state of the shift register represents the binary coefficients of the corresponding polynomial to a member of the Gallois field. Based on an initial seed (LFSR state), the sequence of the states of the LFSR is equivalent to the sequence of random numbers from a Galois field. The following lemma shows how the seed of the random coefficients of the random linear combination of coded messages relates to the seeds of the coded message coefficients.

Lemma 1. *If $s_1, s_2, \cdots, s_k \in GF(2^m)$ are the seeds corresponding to k received coded messages $m_1, m_2, \cdots, m_k$,* *the seed for the corresponding random number sequence to* $\sum_{i=1}^{k} c_i m_i$ *is* $s_c = \sum_{i=1}^{k} c_i s_i$*, where* c_i*s are randomly chosen from* $GF(2^m)$*.*

It is easy to prove this lemma. Assume $C_i(\mathrm{x})$ and $S_i^0(\mathrm{x})$ are the polynomial representation of c_i and s_i. Subsequent states are related through $S_i^{n+1}(\mathrm{x}) = XS_i^n(\mathrm{x})$ for a Gallois LFSR. If the first state of the Gallois LFSR is a linear combination of the received seeds with coefficients c_i's, we have $S_c^0(\mathrm{x}) = \sum_{i=1}^{k} C_i(\mathrm{x})S_i^0(\mathrm{x})$. Multiplying both sides by X^n, we get $S_c^n(\mathrm{x}) = \sum_{i=1}^{k} C_i(\mathrm{x})S_i^n(\mathrm{x})$, which demonstrates that every subsequent shift is also a linear combination of the shifted version of the received seeds with the same coefficients.

Using this lemma, the coding overhead is limited to m bits. If we use a large enough field, the random sequences are distinct with high probability. For example, if we use $GF(2^{32})$ and a primitive polynomial, there are $2^{32} - 1$ distinct random sequences and the coding overhead is only 4 Bytes.

Since each coded message does not necessarily contain all the original messages, we should attach another n bits together with the PRNG seed to specify the original messages included in each coded message. So the total overhead of SPCR is $M_h = \frac{m+n}{8}$ Bytes. To fairly compare SPCR with other algorithms and to account for the coding overhead, we assume the message size for SPCR is $M^{'} = M + M_h$, in which M is the original message size. For a fixed channel rate, this increased message size results in smaller number of timeslots in a subframe.

B. Reliability Bound

Assume ${S_l}^t$ is the spanned subspace by all the received equations of u_l up to time t. We can partition the set of all other nodes into two disjoint subsets: the innovative nodes, ${\mathcal{I}_l}^t$, and non-innovative nodes, $\overline{\mathcal{I}_l}^t$, defined as:

$$\forall i \in {\mathcal{I}_l}^t : \quad {S_i}^t \nsubseteq {S_l}^t$$
$$\forall i \in \overline{\mathcal{I}_l}^t : \quad {S_i}^t \subseteq {S_l}^t$$

At the beginning of the subframe (t=0), for an inactive node all the active nodes are innovative and $\left|{\mathcal{I}_i}^0\right| = n$. For an active node, all *other* active nodes are innovative and $\left|{\mathcal{I}_i}^0\right| = n - 1$. The success probability for node u_l is defined as the probability that it decodes all the original messages by the end of the subframe. The success probability for node u_l can be written as follows:

$$P_s(n) = Pr(dim(S_l^L) = n).$$

The set of original packets creates an n-dimensional subspace. Network coding creates a linear combination of the data vectors, which belongs to the same subspace since the subspace is closed under linear operation. If $dim({S_l}^t) = i$, then we should have at least $n - i$ innovative nodes since the packets of the innovative nodes and ${S_l}^t$ should span the whole space. It is impossible to span a space with dimension $n-i$ with less than $n - i$ vectors. We summarize this fact in the following proposition:

Proposition 1. *If* $dim({S_l}^t) = i$*, then the minimum number of innovative nodes for* u_l *at time* t *is* $n - i$*, or the maximum number of non-innovative nodes is* i*.*

This is an important result that shows why SPCR should perform better than SPR. In SPR, the number of innovative nodes is equal to $n - i$, which is the minimum value of that achieved by SPCR.

The following theorem gives an upper bound on the loss probability of SPCR in a network of n nodes with symmetric erasure channel.

Theorem 2. *Using SPCR in a symmetric erasure channel with erasure probability* p_e*, an upper bound for the loss probability can be obtained using the following inequality:*

$$P_l \leq \tau T^L \mathbf{1}$$
$$where:$$
$$\tau = [1, 0, \cdots, 0]$$
$$t_i = \mathbf{1} - T_i\mathbf{1}$$
$$T = \begin{bmatrix} T_0 & [t_0 0] & & & & \\ & T_1 & [t_1 0] & & 0 & \\ & & \ddots & \ddots & & \\ & & & \ddots & \ddots & \\ & 0 & & & T_{n-2} & [t_{n-2} 0] \\ & & & & & T_{n-1} \end{bmatrix}$$

and T_i*s are defined as:*

$$[T_i]_{jk} = \binom{i-j}{k-j}(1-p_e)^{k-j} p_e^{i-k+1} p_j, \quad j \neq k$$
$$[T_i]_{jj} = 1 - p_j(1 - p_e^{i-j+1})$$
$$p_j = (n - (i - j))p(1 - p)^{n-1}$$
$$0 \leq j \leq k \leq i \leq n - 1$$

in which L *is the number of timeslots in a subframe,* n *is the number of active nodes in the cluster, and* p *is the transmission probability at each timeslot in the SPCR scheme.*

Proof: For an inactive user, we define the random variable D_i as:

$$D_i = \{min\ t | S^t = i + 1\} - \{min\ t | S^t = i\}.$$

The first time that the dimension of S^t becomes i, Proposition 1 guarantees that there are at most i non-innovative nodes. To find D_i, we assume the maximum number of non-innovative nodes exists. It is straightforward to show that the computed ${D_i}^*$ with this worst case assumption stochastically dominates D_i. D_i^* can be modelled as the absorption time of the Markov chain depicted in Fig. 6. This Markov chain model is inspired by [23]. The state of the Markov chain relates to the number of non-innovative nodes. At state j $(1 \leq j \leq i)$ there are $(i - j)$ non-innovative nodes. The first state represents the first time that the dimension of the node becomes i. The time transition from this state till the time that the dimension of ${S_l}^t$ becomes $i + 1$ (absorbing state) has a discrete time phase type distribution $PH_d(\tau, T_i)$. T_i is an upper triangular matrix and $\tau = [1, 0, \cdots, 0]$. The absorbing state is the first time that

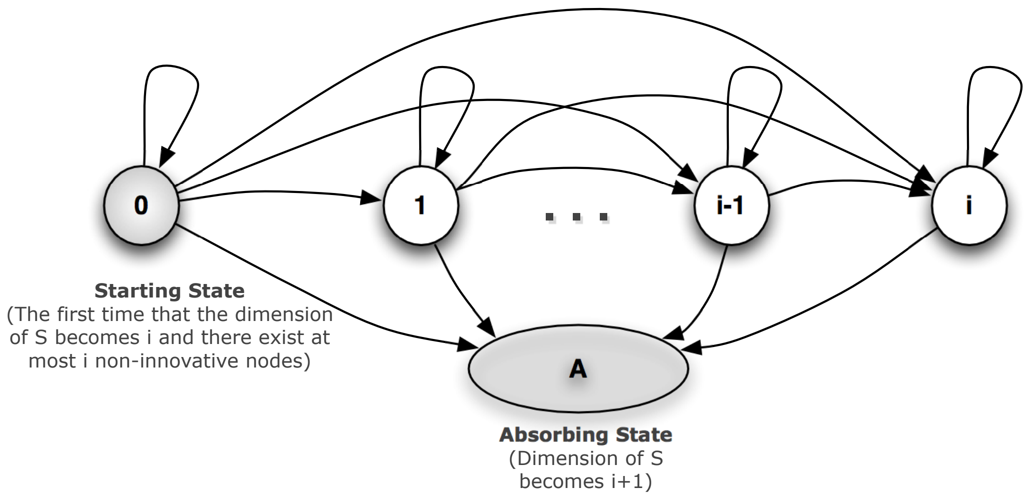

Fig. 6. Markov chain model.

the dimension of S^t becomes $i+1$. Starting from the initial state (State 0), after each transmission, if the node receives an innovative packet, it moves to the absorbing state otherwise depending on how many of non-innovative nodes receive this message, the number of non-innovative nodes may decrease.

Now, we find the transition probabilities of the Markov chain $[T_i]_{jk}$ $(0 \leq j \leq k \leq i)$. We define the following events:

$$\begin{aligned} A &= \{\text{One of } n-(i-j) \text{ innovative nodes transmits successfully}\} \\ B &= \{(k-j) \text{ non-innovative nodes become innovative}\} \\ C &= \{dim(S^t) \text{ increases to } i+1\} \end{aligned}$$

When $j \neq k$, we can compute the transition probabilities as follows:

$$\begin{aligned} [T_i]_{jk} &= P(B \wedge \overline{C}|A)P(A) \\ &= P(B|A)p_e(n-(i-j))p(1-p)^{n-1} \end{aligned}$$

The $B|A$ has a binomial distribution with parameters $(i-j)$ and p_e:

$$[T_i]_{jk} = \binom{i-j}{k-j}(1-p_e)^{(k-j)}p_e^{i-k+1}P(A) \quad (4)$$

for $j = k$, we have

$$[T_i]_{jj} = p_e^{i-j+1}P(A) + 1 - P(A) \quad (5)$$

In this Markov chain, D_i^* is the number of timeslots starting from state i to get to the absorbing state. Since D_i^* dominates D_i the loss probability P_l can be upper bounded as follows:

$$P_l = Pr(\sum_{i=0}^{n-1} D_i \geq L) \leq Pr(\sum_{i=0}^{n-1} D_i^* \geq L).$$

The random variable $D = \sum_{i=0}^{n-1} D_i^*$ is the sum of n independent phase type distributions. Theorem 2.6.1 in [33] shows that D has a discrete phase type distribution $PH_d(\tau, T)$ with the following parameters :

$$\begin{aligned} \tau &= [1, 0, \cdots, 0] \\ t_i &= \mathbf{1} - T_i\mathbf{1} \end{aligned}$$

$$T = \begin{bmatrix} T_0 & [t_0 0] & & & & \\ & T_1 & [t_1 0] & & 0 & \\ & & \ddots & \ddots & & \\ & & & \ddots & \ddots & \\ & 0 & & & T_{n-2} & [t_{n-2} 0] \\ & & & & & T_{n-1} \end{bmatrix}$$

Evaluating the CDF of D at L completes the proof. ∎

To evaluate how tight the derived bound is, simulations are performed using Matlab for a binary symmetric channel with $p_e = 0.1$. A network topology with high node density (n=50) is considered. The number of timeslots in a subframe is assumed to be a function of node density: $L = \beta\lceil nlog(n)\rceil$. As can be seen in Fig. 7, the derived loss probability bound is tight for a low erasure probability. For higher erasure probabilities ($p_e = 0.2$ and $p_e = 0.4$), the comparison is shown in Fig. 8. As can be seen for higher erasures the bound is not as tight but still good for $p_e = 0.2$. Our further evaluations shows the tightness does not change drastically by increasing the number of nodes.

If D_i is the absorption delay for the Markov chain defined in the proof of Theorem 2, then D^{SPCR} is the sum of all absorption delays: $D^{SPCR} = \sum_{i=0}^{n-1} D_i$. Since D_i has a discrete phase-type distribution $PH_d(\tau, T_i)$, $E(D_i)$ can be found using the following formula ([33]):

$$E(D_i) = \tau(I - T_i)^{-1}\mathbf{1},$$

where τ and T_i's are defined in Theorem 2. Therefore the expected delay of the SPCR can be calculated as:

$$E(D^{SPCR}) = \sum_{i=0}^{n-1} \tau(I - T_i)^{-1}\mathbf{1}. \quad (6)$$

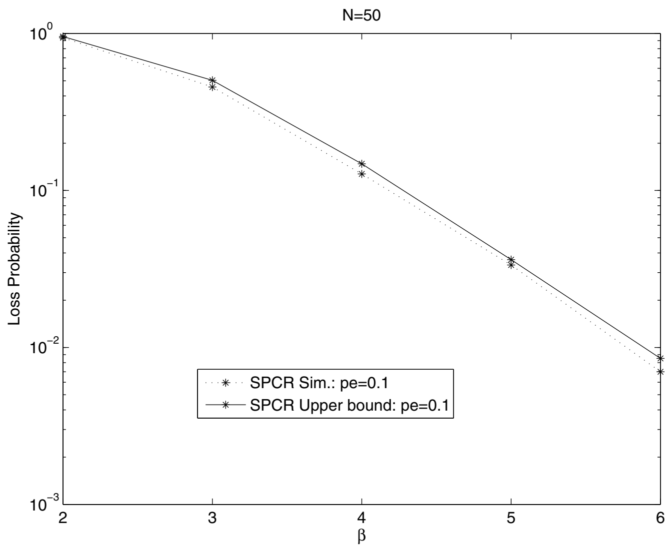

Fig. 7. Simulation results vs. theoretical upper bound: $N = 50$.

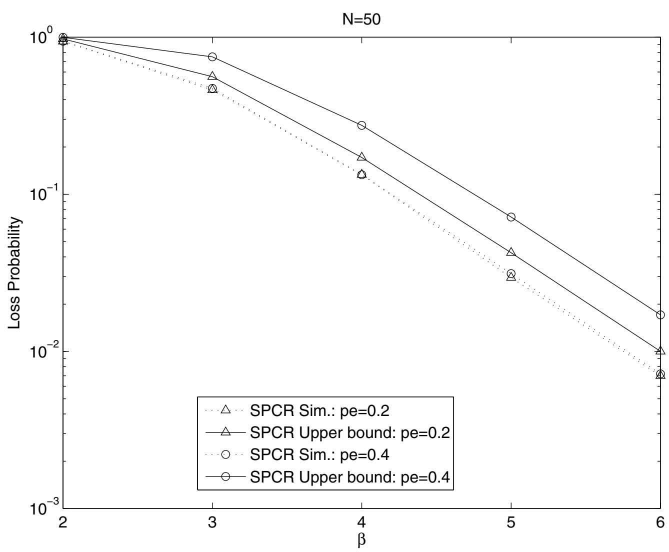

Fig. 8. Simulation results vs. theoretical upper bound: $N = 50$.

VI. Numerical and Simulation Results

In this section we first present the numerical results based on our analytical studies in the previous sections. The performance of IEEE 802.11p, SPR and SPCR are compared numerically. Next, the simulation results using a realistic vehicular channel model is presented.

A. Numerical Results

The erasure probability is a key factor for determining the reliability. To see how the erasure probability changes in a typical vehicular environment, first we consider a realistic vehicle-to-vehicle channel and find the erasure probability as a function of distance. We use the Nakagami channel model for the vehicle-to-vehicle communication link. The probability density function of the signal amplitude Y based on this

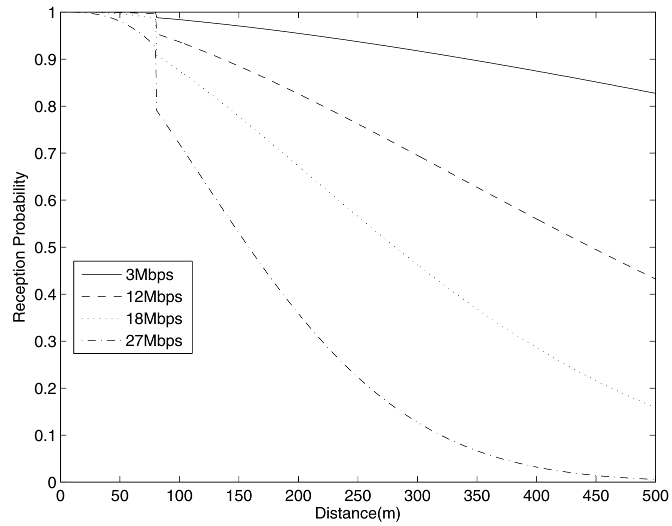

Fig. 9. Reception probability vs. distance.

channel model is:

$$f_Y(y) = \frac{2m_1^{m_1} y^{2m_1-1}}{\Gamma(m_1)\Omega_1^{m}} e^{-\frac{m_1 y^2}{\Omega}}$$
$$m_1 \geq \frac{1}{2},\ \Omega \geq 0$$

in which Ω is the average received power and m_1 is the fading figure. A path loss component of 1.8-1.9 is reported for highways [34]. Here, we assume the path loss component of 2 for our evaluations.

In [35], the parameter m_1 has been estimated based on empirical measurements for a vehicle-to-vehicle link in a highway:

$$m_1 = \begin{cases} 1.5 & d < 80 \\ 0.75 & d > 80 \end{cases} \tag{7}$$

This estimate of m_1 with the physical and MAC layer parameters listed in Table I has been used in our evaluations. We assume the transmission power is 20dBm (Class C) [36] and the reception and transmission antenna gain is 2. The message size is 200 Bytes.

The minimum reception threshold for a 10Mhz channel is specified in IEEE 802.11 standard and can be found in Table II. As we expect, for higher rates the threshold is higher. Fig. 9 shows how the successful reception probability drops as a function of distance for a few allowed rates in IEEE 802.11p.

The erasure probability is a function of distance. However, our analysis in previous section was for symmetric erasure channels. For the numerical evaluation, we assume the p_e for a communication range of R_d is the average of erasure probability over that range:

$$p_e = \frac{1}{R_d} \int_0^{R_d} p_e(x)dx. \tag{8}$$

This seems to be a valid choice as it reasonably relates the erasure probability to the rate.

Fig. 10 is the comparison of loss probability for SPCR, SPR and IEEE 802.11p as a function of the number of active nodes in a subframe and for $R_d = 500m$. We have used the

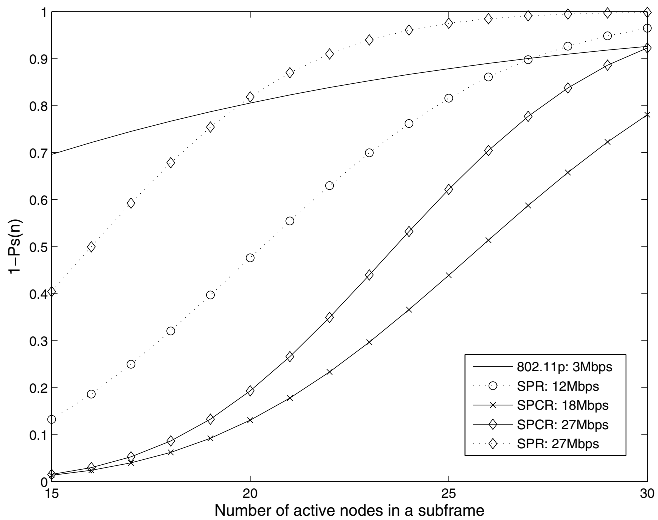

Fig. 10. Loss probability comparison for different rates.

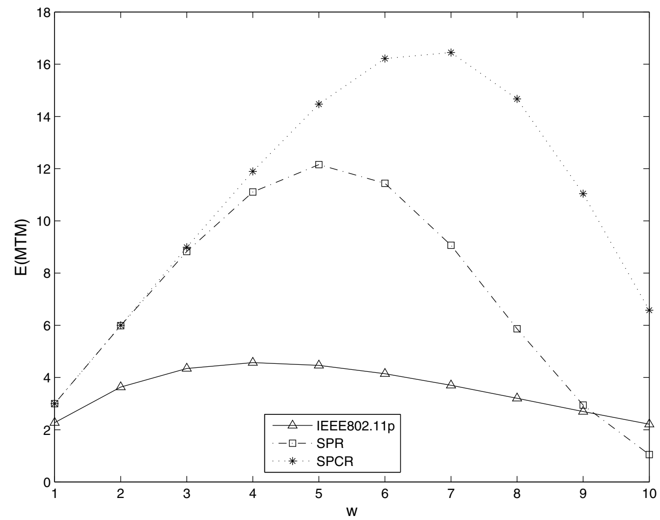

Fig. 11. E(MTM) vs. w (number of active subframes in 10 subframes): $n = 30$.

derived loss probability upper bound in the previous section for numerical evaluations of SPCR. For SPR and SPCR, increasing the rate results in more timeslots in a subframe and more retransmission opportunities. On the other hand, by increasing the rate, the erasure probability increases. For IEEE 802.11p, channel rate is set to 3Mbps. Since IEEE 802.11p cannot benefit from the increased number of transmission opportunities and suffers from poor channel quality in higher rates, the lowest rate of 3 Mbps (which corresponds to the lowest channel error) is indeed the optimal rate for IEEE 802.11p. Based on numerical results for different rates the optimal rate for SPR is 12 Mbps and for SPCR is 18Mbps. By increasing the number of nodes, the loss probability for repetition based schemes increases with a higher slope compared to IEEE 802.11p. This is due to excessive number of collisions when there is congestion. For safety applications, we are generally interested in lower loss probabilities. For lower loss probabilities, SPR performs much better than IEEE 802.11p. The SPCR performs significantly better than SPR over the entire range. It is also interesting to see that for the similar rates, increasing the rate widens the gap between the SPCR and SPR. Lower rates correspond to smaller number of timeslots in a subframe. The decoding probability for SPCR drops for smaller subframes.

For 30 nodes and $L_f = 10$, Fig. 11 shows the $E(MTM)$ versus w, the number of active subframes in L_f subframes. $E(MTM)$ is evaluated based on the analytical results in the previous sections. For SPCR, $E(MTM)$ is based on the derived loss probability upper bound in the previous section and is a lower bound. As can be seen, there is an optimal w that maximizes $E(MTM)$. The optimal w depends on the total number of nodes, which for N=30, is 4, 5 and 7, for IEEE 802.11p, SPR and SPCR, respectively. The SPCR has the highest optimal w and allows more nodes to be active in each subframe with a higher $E(MTM)$. Note that the IEEE 802.11p curve indeed represents the congestion control algorithm in [4], in which the rate is adjusted according to the channel occupancy time feedback. The number of neighbours is strongly correlated with the channel occupancy time, and different w's correspond to different rates.

TABLE II
MINIMUM RECEPTION THRESHOLD FOR IEEE 802.11P

Rate (Mbp)	Minimum RX threshold (dBm)
3	-85
4.5	-84
6	-82
9	-80
12	-77
18	-73
24	-69
27	-68

B. Simulation results

In this section, we present realistic simulation results based on the introduced Nakagami channel model in the previous section. The simulator is implemented using Matlab based on the assumed channel and system model in Section III. Unlike the previous sections, it is assumed that the erasure probability changes with distance. We assume 20 nodes are spaced horizontally with a spacing of 25m in a 500m road segment. Nodes are indexed from 1 to 20, in order, from left to right. A Nakagami channel model with the same parameter as previous section is utilized. All nodes broadcast with a channel rate of 12Mbps. The message size is 200 Bytes. The transmission power is assumed to be 20dBm for all nodes. The simulation results are averaged over 10000 runs. The loss probability $(1 - P_s(n))$ versus the node index can be seen in Fig. 12. It is observed that the SPCR loss probability of all nodes is almost the same. For SPR, the loss probability is dependent on the location and is higher for edge nodes. The location independent performance of SPCR is due to cooperative nature of network coding. All nodes act as relays for their neighbours by including the messages of farther nodes in their coded packets. In SPR, even though message rebroadcasting can mitigate the poor channel condition, since every node only repeats its own message, the edge nodes still are not able to receive all messages within a CCH interval. Although the derived loss probability upper bound in Theorem 2 is only valid for a symmetric erasure network, we have evaluated the bound for the maximum p_e in the network as

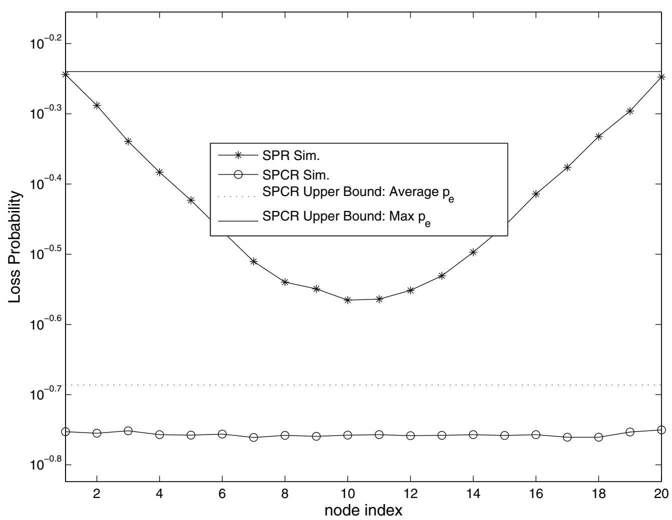

Fig. 12. Loss probability for all nodes: $n = 20$.

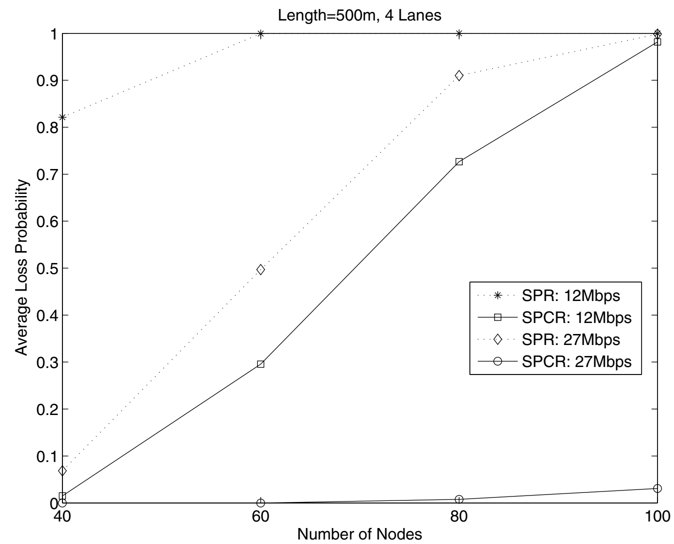

Fig. 13. Expected loss probability vs. number of nodes.

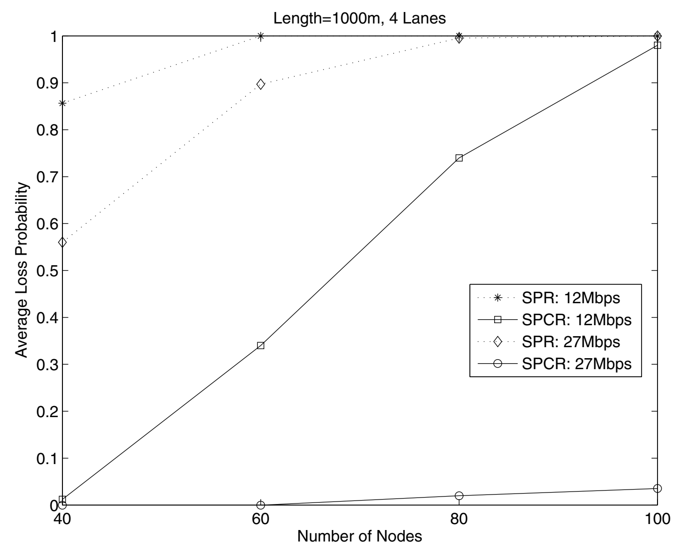

Fig. 14. Expected loss probability vs. number of nodes

well as the average p_e. Average p_e is computed over all 20×19 possible channels. As can be seen in Fig. 12, the bound based on the average p_e is a relatively tight bound for the SPCR loss probability.

To further evaluate the performance under a more realistic network model, SPR and SPCR have been implemented in the ns-2 simulator. The same Nakagami model as in Section VI.A is set in the simulator. Unlike the the assumed channel model for our analysis in Section III, in ns-2 some of the collisions can be resolved due to capture. The transmission power for every node is set to 760mw (Class D in the IEEE 802.11p standard) and the transmitter and receiver antenna gain is 2. The message size is 200 Bytes. The radio frequency, carrier and reception threshold have been set according to the IEEE 802.11p standard. A 4-lane road segment of length 500m and 1km is assumed. In each lane a vehicle is placed uniformly every d meters. Four traffic densities of 10, 15, 20 and 25 nodes per lane are assumed. All nodes are assumed to be active. The lower node densities correspond to highways in which the vehicle interspacing is larger. Higher traffic densities represent the urban areas. The simulations are performed for 100s. The average loss probability over all nodes and within the simulation time is measured. The simulation results for channel rates of 12Mbps and 27Mbps (maximum rate of IEEE 802.11p) and for the 500m road segment topology can be seen in Fig. 13. It is observed that SPCR can significantly benefit from the rate increase. For example for $n = 100$, in SPCR, by increasing the rate from 12Mbps to 27Mbps the average loss probability drops one order of magnitude to less than 0.05, while for SPR the average loss probability remains at 1. This proves the significant gain that can be achieved through network coding specially in dense topologies. As we mentioned in the previous section, increasing the rate increases the number of time slots in a sub-frame which provides more opportunities for broadcasting the coded messages. Unlike the SPR in which the additional transmission opportunities are not necessarily helpful for all the receivers, in SPCR, most of the receivers can benefit from the received coded message by expanding their subspace of received coded vectors. The simulation results for the 1km road segment topology can be observed in Fig. 14. Due to longer transmission ranges, channel quality is worse compared to the previous topology. It can be seen that SPCR performance is robust to channel error and does not change significantly compared to the previous topology. However, the SPR performance suffers from lower channel quality. For example for 40 nodes and 27Mbps rate, the average loss probability increases to more than 0.56 from 0.08. This shows that network coding not only is effective in dense topologies but also is robust to channel errors.

VII. Conclusion

We have proposed a sublayer that optimizes the reliability of periodic broadcasting in VANETs. The core of our design is the random linear network coding which is used to provide reliability for small safety messages with low overhead. We have also studied how the message rebroadcasting can be utilized when there is congestion. Numerical results based

TABLE III
TABLE OF NOTATIONS

$VANET$	:	Vehicular Ad-hoc NETwork
$WAVE$	:	Wireless Access in Vehicular Environment
CCH	:	Control CHannel
SCH	:	Service CHannel
WSM	:	Wave Short Message
SFR	:	Synchronized Fixed Repetition
SPR	:	Synchronized Persistent Repetition
$SPCR$	:	Synchronized Persistent Coded Repetition
POC	:	Positive Orthogonal Code
GPS	:	Global Positioning System
$P_s(n)$	:	The probability of receiving a full map of size n
T_{CCH}	:	The length of CCH interval
m_j	:	Original generated message of node u_j
m_j^k	:	k^{th} retransmission of node u_j
MTM	:	Map Throughput Metric
$PLCP$	:	PHY Layer Convergence Procedure
R	:	Data rate
T_h	:	PLCP duration
T_g	:	Guard time duration
p_e	:	Erasure probability
$AIFS$	:	Arbitration inter-frame spacing
M	:	Safety Message length
CW	:	Contention Window size
σ	:	Time slot duration in IEEE 802.11p
L_f	:	Number of subframes in a frame
N	:	Number of nodes in a cluster
n	:	Number of active nodes
w	:	Number of subframes in which a node is active
p	:	Transmission probability in a time slot
D	:	Completion delay
L	:	Number of time slots in a subframe
$LFSR$	:	Linear Feedback Shift Register
$PRNG$	:	Pseudo Random Number Generator
q	:	Finite field size
M_h	:	Network coding overhead
m_1	:	Fading figure
Ω	:	Average received power
R_d	:	Communication range

on our analysis confirm the superior performance for our method compared to previous schemes. Our design can be implemented in conjunction with the WAVE architecture and does not need any modification to the WAVE communication stack.

ACKNOWLEDGEMENTS

The authors would like to thank AUTO21 Network of Centers of Excellence and Kapsch Trafficcom (formerly MARK IV Industries, IVHS Division), Canada for their financial support.

REFERENCES

[1] B. Xu, A. Ouksel, and O. Wolfson, "Opportunistic resource exchange in inter-vehicle ad-hoc networks," in *Proc. 2004 IEEE International Conf. on Mobile Data Management*, pp. 4–12.

[2] "IEEE Standard for Wireless Access in Vehicular Environments (WAVE)–Multi-channel Operation," pp. 1–89, 2011, IEEE Std 1609.4-2010 (Revision of IEEE Std 1609.4-2006).

[3] Y. P. Fallah, C. Huang, R. Sengupta, and H. Krishnan, "Congestion control based on channel occupancy in vehicular broadcast networks," in *Proc. 2010 IEEE Veh. Technol. Conf. – Fall.*

[4] Y. Fallah, C.-L. Huang, R. Sengupta, and H. Krishnan, "Analysis of information dissemination in vehicular ad-hoc networks with application to cooperative vehicle safety systems," *IEEE Trans. Veh. Technol.*, vol. 60, no. 1, pp. 233–247, Jan. 2011.

[5] R. Stanica, E. Chaput, and A.-L. Beylot, "Enhancements of IEEE 802.11p protocol for access control on a VANET control channel," in *Proc. 2011 IEEE International Conf. Commun.*, pp. 1–5.

[6] Q. Xu, T. Mak, J. Ko, and R. Sengupta, "Vehicle-to-vehicle safety messaging in DSRC," in *Proc. 2004 ACM International Workshop Veh. Ad Hoc Netw.*, pp. 19–28.

[7] B. Hassanabadi, L. Zhang, and S. Valaee, "Index coded repetition-based MAC in vehicular ad-hoc networks," in *2009 IEEE Consumer Commun. Netw. Conf.*

[8] B. Hassanabadi and S. Valaee, "Reliable network coded MAC in vehicular ad-hoc networks," in *Proc. 2010 IEEE Veh. Technol. Conf. – Fall.*

[9] Q. Xu, T. Mak, J. Ko, and R. Sengupta, "Medium access control protocol design for vehicle-to-vehicle safety messages," *IEEE Trans. Veh. Technol.*, vol. 56, no. 2, pp. 499–518, Mar. 2007.

[10] F. Farnoud, B. Hassanabadi, and S. Valaee, "Message broadcast using optical orthogonal codes in vehicular communication systems," *2007 ICST QSHINE Workshop on Wireless Netw. Intelligent Transportation Syst.*

[11] F. Farnoud and S. Valaee, "Repetition-based broadcast in vehicular ad hoc networks in Rician channel with capture," *2008 Comput. Commun. Workshops, IEEE INFOCOM.*

[12] D. S. Lun, M. Medard, R. Koetter, and M. Effros, "On coding for reliable communication over packet networks," *Physical Commun.*, vol. 1, no. 1, pp. 3–20, 2008.

[13] S. Ahmed and S. S. Kanhere, "VANETCODE: network coding to enhance cooperative downloading in vehicular ad-hoc networks," in *Proc. 2006 International Conf. Wireless Commun. Mobile Comput.*,, pp. 527–532. Available: http://doi.acm.org/10.1145/1143549.1143654

[14] U. Lee, J.-S. Park, J. Yeh, G. Pau, and M. Gerla, "Code torrent: content distribution using network coding in VANET," in *Proc. 2006 International Workshop Decentralized Resource Sharing Mobile Comput. Netw.*, pp. 1–5, 116. Available: http://doi.acm.org/10.1145/1161252.1161254

[15] J.-S. Park, U. Lee, S.-Y. Oh, M. Gerla, D. S. Lun, W. W. Ro, and J. Park, "Delay analysis of car-to-car reliable data delivery strategies based on data mulling with network coding," *IEICE Trans.*, vol. 91-D, no. 10, pp. 2524–2527, 2008.

[16] Z. Yang, M. Li, and W. Lou, "CodePlay: live multimedia streaming in VANETs using symbol-level network coding," in *Proc. 2010 IEEE Int. Netw. Protocols Conf.*, pp. 223–232, 11.

[17] ——, "CodePlay: live multimedia streaming in VANETs using symbol-level network coding," *IEEE Trans. Wireless Commun.*, vol. 11, no. 8, pp. 3006–3013, 2012.

[18] M. Li, Z. Yang, and W. Lou, "CodeOn: cooperative popular content distribution for vehicular networks using symbol level network coding," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 1, pp. 223–235, 2011, 24.

[19] Z. Wang, M. Hassan, and T. Moors, "Efficient loss recovery using network coding in vehicular safety communication," in *Proc. 2010 IEEE Wireless Commun. Netw. Conf.*, pp. 1–6.

[20] Z. Wang and M. Hassan, "Blind XOR: low-overhead loss recovery for vehicular safety communications," *IEEE Trans. Veh. Technol.*, vol. 61, no. 1, pp. 35–45, 2012.

[21] F. Ye, S. Roy, and H. Wang, "Efficient inter-vehicle data dissemination," in *Proc. 2011 IEEE Veh. Technol. Conf. – Fall*, pp. 1–5.

[22] ——, "Efficient data dissemination in vehicular ad hoc networks," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 4, pp. 769–779, 2012.

[23] S. Deb, M. Medard, and C. Choute, "Algebraic gossip: a network coding approach to optimal multiple rumor mongering," *IEEE Trans. Inf. Theory*, vol. 52, no. 6, pp. 2486–2507, June 2006.

[24] M. Borokhovich, C. Avin, and Z. Lotker, "Tight bounds for algebraic gossip on graphs," in *Proc. 2010 IEEE International Symp. on Inf. Theory*, pp. 1758–1762.

[25] T. Aysal, M. Yildiz, A. Sarwate, and A. Scaglione, "Broadcast gossip algorithms for consensus," *IEEE Trans. Signal Process.*, vol. 57, no. 7, pp. 2748–2761, 2009.

[26] X. Yang, J. Liu, F. Zhao, and N. H. Vaidya, "A vehicle-to-vehicle communication protocol for cooperative collision warning," in *Proc. 2004 International Conf. Mobile Ubiquitous Syst.*, vol. 0, pp. 114–123.

[27] G. Resta and J. Simon, "Analysis of multi-hop emergency message propagation in vehicular ad hoc networks," in *Proc. 2007 ACM International Symp. Mobile Ad Hoc Netw. Comput.*, pp. 140–149.

[28] M. Torrent-Moreno, D. Jiang, and H. Hartenstein, "Broadcast reception rates and effects of priority access in 802.11-based vehicular ad-hoc networks," in *Proc. 2004 ACM International Workshop Veh. Ad Hoc Netw.*, pp. 10–18.

[29] J. Yin, G. Holland, T. ElBatt, F. Bai, and H. Krishnan, "DSRC channel fading analysis from empirical measurement," in *Proc. 2006 International Conf. Commun. Netw. China.*

[30] V. Taliwal, D. Jiang, H. Mangold, C. Chen, and R. Sengupta, "Empirical determination of channel characteristics for DSRC vehicle-to-vehicle communication," in *Proc. 2004 ACM International Workshop Veh. Ad Hoc Netw.*, pp. 88–88.

[31] R. Motwani and P. Raghavan, *Randomized Algorithms*. Cambridge University Press, 1995.

[32] Z. Liu, C. Wu, B. Li, and S. Zhao, "UUSee: large-scale operational on-demand streaming with random network coding," in *Proc. 2010 IEEE INFOCOM*.

[33] G. Latouche and V. Ramaswami, *Introduction to Matrix Analytic Methods in Stochastic Modeling*. Society for Industrial and Applied Mathematics, 1999.

[34] A. Molisch, F. Tufvesson, J. Karedal, and C. Mecklenbrauker, "A survey on vehicle-to-vehicle propagation channels," *IEEE Wireless Commun.*, vol. 16, no. 6, pp. 12–22, Dec. 2009.

[35] Q. Chen, F. Schmidt-Eisenlohr, D. Jiang, M. Torrent-Moreno, L. Delgrossi, and H. Hartenstein, "Overhaul of IEEE 802.11 modeling and simulation in ns-2," in *Proc. 2007 ACM Symp. Modeling, Analysis, Simulation Wireless Mobile Syst.*, pp. 159–168.

[36] "IEEE Standard for Information technology—telecommunications and information exchange between systems—local and metropolitan area networks—specific requirements part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 6: Wireless Access in Vehicular Environments," IEEE Std 802.11p-2010 (Amendment to IEEE Std 802.11-2007 as amended by IEEE Std 802.11k-2008, IEEE Std 802.11r-2008, IEEE Std 802.11y-2008, IEEE Std 802.11n-2009, and IEEE Std 802.11w-2009), pp. 1–51, July 2010.

Behnam Hassanabadi received the BSc and MSc degree in electrical and computer engineering from University of Tehran, Iran. He received the PhD degree in electrical and computer engineering from University of Toronto, Toronto, Canada in 2013. His research interests are in the areas of wireless networking, network coding, scheduling and medium access control design.

Shahrokh Valaee (S '88, M '00, SM '02) received the PhD degree in electrical engineering from McGill University in Canada. Currently he is a Professor and the Associate Chair for Undergraduate Studies and holds the Nortel Institute junior chair of Computer Networks in the Edward S. Rogers Sr. Department of Electrical and Computer Engineering at the University of Toronto. He is the founder and the Director of the Wireless and Internet Research Laboratory (WIRLab) at the University of Toronto.

Prof. Valaee is an Editor of IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, and an Associate Editor of IEEE SIGNAL PROCESSING LETTERS. He was the Technical Program Co-Chair and the Local Organizing Chair of the IEEE PIMRC 2011, and the Co-Chair for Wireless Communications Symposium of IEEE GLOBECOM 2006. He has served as a guest editor for several journals including *IEEE Wireless Communications Magazine*, *Wiley Journal on Wireless Communications and Mobile Computing*, *EURASIP Journal on Advances in Signal Processing*, and the *International Journal of Wireless Information Networks*. His current research interests are in wireless vehicular and sensor networks, location estimation and cellular networks.