Towards Reliable Multi-Hop Broadcast in VANETs: An Analytical Approach

Mozhdeh Gholibeigi‡, Mitra Baratchi‡, Hans van den Berg*‡, Geert Heijenk‡
‡ University of Twente, The Netherlands
*Netherlands Organization for Applied Scientific Research (TNO), The Netherlands
Email: [m.gholibeigi,m.baratchi,j.l.vandenberg,geert.heijenk]@utwente.nl

***Abstract*—Intelligent Transportation Systems in the domain of vehicular networking, have recently been subject to rapid development. In vehicular ad hoc networks, data broadcast is one of the main communication types and its reliability is crucial for high performance applications. However, due to the lack of acknowledgment techniques in IEEE 802.11p standard, it is challenging to ensure communication reliability. In this work, we analytically model a receiver-oriented reliability mechanism, with the objective of on-demand error recovery for multi-hop broadcast vehicular communication. In particular, using absorbing Markov modeling and probabilistic graphical modeling, we analyze its performance in terms of relevant indicators, such as overhead and delivery ratio. Further, the model is validated using simulations. The results are useful in tuning the influential parameters and accordingly adjusting trade-offs and meet performance requirements in various circumstances.**

***Index Terms*—Multi-hop broadcast, reliability, analytical modeling, vehicular ad hoc networks.**

I. Introduction

Intelligent Transportation Systems (ITS) have much potential in improving traffic efficiency and safety in the domain of vehicular networking. Data broadcast, as one of the main communication types in vehicular networks, receives significant attention in the research community. At the same time 802.11p/1609, which is the standard communication protocol suite for vehicular networks, does not have a MAC-Layer acknowledgement scheme for broadcast [3] [1]. This leads to challenges in providing reliable communication. Accordingly, delivery of data and evaluating its reliability are among the main performance concerns.

In our work in [9] [10], we analytically modeled and evaluated a receiver-based end-to-end reliability assurance mechanism for vehicular networks in the context of single-hop (geo-)broadcast. This mechanism, which is described in Section III, enables loss detection and correction by means of sequence numbering and checking at communicating parties, as shown at high-level in Figure 1 and is suitable for various ITS applications, demanding strict requirements on delivery assurance. In vehicular networks, it is quite common that nodes act as relays in order to disseminate information over multi-hops beyond the transmission range of individual nodes. Hence, in this paper we aim to model the aforementioned reliability mechanism in the context of multi-hop broadcast. That is, those nodes failing to receive a packet after several hops of rebroadcasts and having detected such a loss through this mechanism, will broadcast retransmission requests to their one-hop neighborhood. Accordingly, the neighbors possessing that missing packet will reply by retransmitting it.

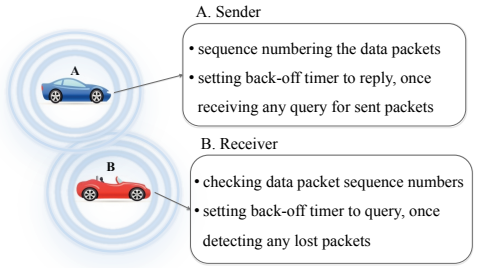


Figure 1: The operation principle of the E2E reliablility mechanism.

Our analytical approach is not based on a specific network scenario and can provide results for various settings and accordingly basic insights regarding the system behavior. As a result of this, our model is extensible by introducing more detail into the parameterization of the model. The main contributions of this paper are as follow:

(i) we analytically model the error recovery process in the context of multi-hop broadcasting.
(ii) based on the analytical model, we analyze its functionality, in terms of relevant performance indicators.
(iii) we validate observations from the analytical analysis, using simulations.

The rest of the paper is organized as follows. At first, the related work is discussed in Section II. Section III introduces the proposed reliability mechanism and the multi-hop broadcast model. The analytical modeling and analysis of the integrated multi-hop broadcast and reliability mechanism is discussed in Section IV. Numerical results are presented in Section V. Conclusions and the future work are given in the final section.

II. Related work

In this section, we discuss some of the mechanisms for multi-hop (geo-)broadcasting in vehicular networks and approaches proposed for reliability improvement in the literature.

Klein Wolterink et al. proposed Constrained Geocast in [14] which targets the vehicles based on their future and not the current position. Dynamic Time-Stable Geocast (DTSG) protocol [17] improves reliability by keeping the geocast message alive

for some adjustable time in the geocast region. Geocache [15] is a peer-to-peer pull-based geocast protocol which improves reliability by allowing vehicles to cooperatively collect and disseminate data.

The approach in [19] is a cooperative repetition by piggybacking old messages to the new ones as a means of improving reliability. Hassanabadi et al. proposed Synchronized Persistent Coded Repetition (SPCR) algorithm [11] as the main functionality of the reliability sub-layer in the application layer of the WAVE stack to increase the reliability of safety applications.

In Grid-based Predictive Geographical Routing (GPGR) [4] node positions are predicted during the relay node selection as a means of increasing reliability. The authors in [7] proposed an analytical model to predict the number of nodes in the neighborhood and adjust the probability of rebroadcast for a better coverage. In [18] an end-to-end geocast acknowledgement scheme is proposed, in which individual ACKs are accumulated into larger messages in an aggregator and further forwarded back over multiple hops to the original source.

In summary, most of the related work consider modeling or simulating a specific scenario and rely on some default settings (e.g., rebroadcast decision making mechanisms) as a means of reliable delivery. Works with definitive and event-driven reliability schemes, either mostly do not rely on accurate information, regarding the necessity of error recovery, from the receiving side or rely on, probably unnecessary, overloading acknowledgments. Given this, in this work we develop a novel approach to model and analyze the receiver-based reliability assurance mechanism, as described in Section III. This mechanism is based on on-demand error recovery from the receiving side, avoiding redundant traffic. Our approach is not dependent to a particular setting and hence is applicable to various scenarios.

III. THE END-TO-END RELIABLE GEOCAST

The mechanism discussed in this paper, is a receiver-based end-to-end reliability assurance for (geo-)broadcast in vehicular environments. It enables error recovery by implementing a sequence numbering and checking functionality at communicating parties, as shown in the flowcharts in Figure 2 and Figure 3. Through this mechanism, on failure of receiving a packet, a retransmission request is broadcast to all one-hop neighbors. Accordingly, the neighbors possessing that missing packet will reply by retransmitting it. As a means of collision avoidance, due to concurrent requests and replies, back-off timers are considered prior to requests and replies. Note that request and reply packets are broadcast to the single-hop neighborhood and they would not be sent, if overheard during the back-off (i.e. other nodes send the same request or reply that is overheard by the node backing off to send it). For further detail regarding its functionality, we refer to our work in [9]. It is worth pointing out that here we assume that the underlying ETSI/ITS standard geo-networking infrastructure [5] is functionally available, as the basis to implement such a reliability mechanism.

IV. MODELLING AND ANALYSIS

The focus of this section is on modeling and analyzing the performance of the proposed approach, in a multi-hop broadcast scenario. For this purpose, at first we need modeling of multi-hop broadcasting, which is described in Section IV-A. We call multi-hop broadcasting as the first phase of operation. Next, in Section IV-B the error recovery functionality is modeled, using absorbing Markov chains. We call this the second phase of operation and its performance is evaluated in terms of relevant performance indicators, introduced later in Section IV-C. The notation used in this paper, is listed in Table I.

Table I: Notation.

Parameter	Definition
N	The number of network nodes
b	The number of the rebroadcasts (first phase)
p_{rb}	The probability of data packet rebroadcast (first phase)
k	The number of the retransmission requests (second phase)
p_T	The probability of successful packet transmission
R	Transmission range (m)

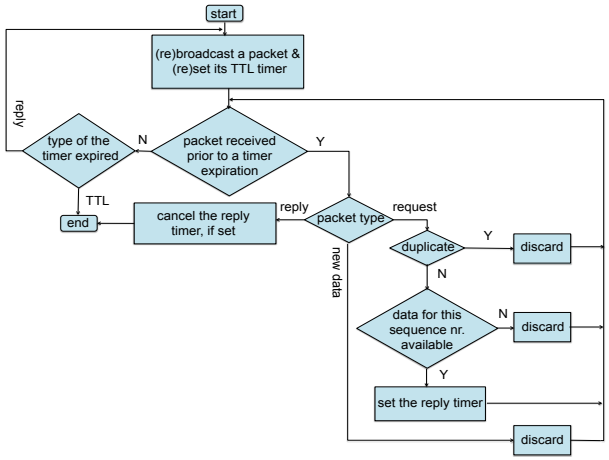


Figure 2: The sender side of the E2E reliablility mechanism.

A. PRELIMINARY WORK ON MULTI-HOP BROADCAST MODEL

In this section, we briefly describe our work on modeling multi-hop broadcast in [8], as the preliminary steps of the work in this paper. That is, we analytically modeled a single cycle of multi-hop broadcast in two-dimensional vehicular networks. The model incorporates network topology in its basis. As a result, the effect of major network characteristics such as network density and the number of 1-hop neighbors are considered in the model as described in the following. These characteristics play a crucial role in the overall dissemination penetration rate.

Let tuple $T = (N, R, p_{rb}, f(\theta), b, U((x_{min}, x_{max}), (y_{min}, y_{max})))$ denote a network topology of N nodes with transmission

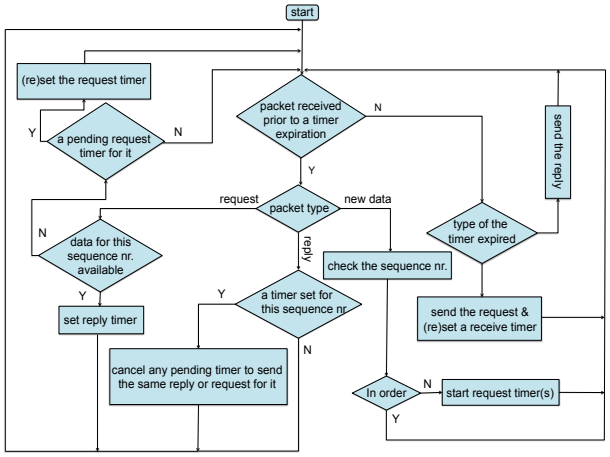


Figure 3: The receiver side of the E2E reliablility mechanism.

range R, uniformly distributed in a two dimensional area $(x_{min}, x_{max}), (y_{min}, y_{max})$. In such a network arbitrary nodes i and j, within the threshold distance R are reachable. Accordingly, we constructed the $N \times N$ reachability matrix D, with d_{ij} elements either 0 or 1, for i and j as non-reachable (i.e. their euclidean distance is more than R) and reachable (i.e. their euclidean distance is less than or equal to R) pair of nodes, respectively. Figure 4 shows an example reachability graph of $N = 7$ nodes in a $2km \times 2km$ area, with the transmission range $R = 300m$ and its corresponding matrix D of reachability. All the diagonal elements are set to 1. We

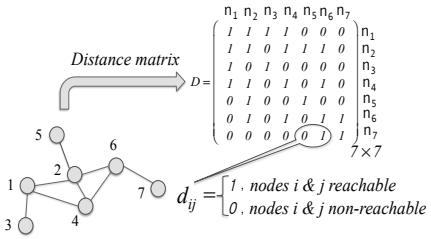


Figure 4: The reachability graph of a topology of $N = 7$ nodes.

further defined the function $f(\theta_{ij})$ that parameterizes the probability of successful transmission between an arbitrary pair of nodes i and j that are reachable (i.e. $d_{ij} = 1$). $f(\theta_{ij})$ is a function of effective factors on successful packet transmission between an arbitrary pair of nodes i and j, such as transmission rate, probability of propagation loss, hidden nodes and probability of collision (represented in general by θ_{ij}). Now, by multiplying each $(ij)^{th}$ element of matrix D to the corresponding $f(\theta_{ij})$, we get the successful transmission probability matrix, denoted by Q_T. The elements q_{ij} of this matrix are either 0 or $f(\theta_{ij})$. Without loss of generality, for any given i and j, we made the assumption of $f(\theta_{ij}) = p_T$ as the average probability of successful transmission between any given pair of nodes.

Given this, we consider a scenario where a source node s in a particular topology T broadcasts a packet and nodes that receive the packet rebroadcast it with probability p_{rb}. Accordingly, using the concept of recursive function definition, we define $P(x, b, Y)$ as the probability for an arbitrary node x of the network to have this packet, after b number of rebroadcasts have been done, as follows.

$$P(x,b,Y) = \begin{cases} P(x,b-1,Y) + \big(1 - P(x,b-1,Y)\big)\times \\ \Bigg(1 - \prod\limits_{\forall n \in Y \setminus \{x\}} \Big(1 - \big(P(n,b-1,Y\setminus\{x\}) - \\ P(n,b-2,Y\setminus\{x\})\big) \times p_{rb} \times q_{nx}\Big)\Bigg), \quad \text{if } b \geq 2 \\ \\ q_{sx} + \big(1 - q_{sx}\big) \times \Bigg(1 - \prod\limits_{\forall n \in Y \setminus \{x\}} \\ \Big(1 - P(n,0,Y\setminus\{x\}) \times p_{rb} \times q_{nx}\Big)\Bigg), \quad \text{if } b = 1 \\ \\ q_{sx}, \quad \text{if } b = 0. \end{cases} \tag{1}$$

Y in $P(x, b, Y)$ represents the set of all network nodes, as the potential rebroadcasters that node x may receive the packet from. Note that $b = 0$ is considered for the original broadcast by the source node s. See [8] for more detail regarding formulation, though this is not requisite to follow the work here.

Considering such a scenario, at the end of a given number of rebroadcasts b, there would be nodes failed to successfully receive the broadcast packet. Assuming such nodes detected the loss, using the reliability mechanism described earlier in section III, they are supposed to broadcast a retransmission request for each lost packet. As mentioned earlier, we call this phase, after the end of the multi-hop rebroadcast (first phase), the error recovery (second phase). In the next section, we continue by modeling this phase of operation and later analyzing its performance on providing reliable data delivery.

B. ERROR RECOVERY MODEL

At the end of the multi-hop rebroadcast phase, we end up with a situation where some nodes may have and some other may not have a broadcast packet with some probability. Such probabilities are computed according to Eq. 1. Let n_i denote the status of an arbitrary node i regarding having ($n_i = 1$) or not having ($n_i = 0$) a packet. We could reasonably model the process of error recovery after the end of the multi-hop rebroadcast as a discrete-time absorbing Markov chain, where each state is a combination of the status of all N nodes. Hence, the system state space could be characterized as the binary representation of the states as $B = \sum_{i=1}^{i=N} n_i \times 2^{i-1}$. Since the source node has status 1, in principle the number of states is $2^{(N-1)}$. However, due to topology constraints regarding node connectivities, in practice the state space would have fewer number of states and accordingly would be different for each particular topology.

Figure 5 shows an example of the state space for $N = 4$ nodes, with principally $2^{(4-1)} = 8$ states. In the binary representation of the states, the leftmost position represents the source node. Since the nodes with status "0" may receive the packet, as a result of request / reply rounds of error recovery

mechanism, the system jumps between the transient states. Clearly, the transition between states, in the direction where the status of any given node changes as $1 \rightarrow 0$, has logically zero probability to take place. At the end, the desired state, where all nodes have the packet, is the absorbing state and once entered cannot be left.

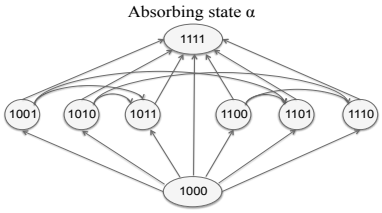


Figure 5: The Markov chain of 4 nodes, with binary state numbering.

An absorbing Markov chain with M transient states and one absorbing state is canonically characterized by [13]

$$T = \begin{bmatrix} Q & \overrightarrow{q} \\ \overrightarrow{0} & I_1 \end{bmatrix}, \quad (2)$$

where Q is an $M \times M$ transition probability matrix, with elements $\{q_{ij} \mid i,j = 1,2,\ldots,M\}$, representing the probability of transition from the transient state i to the transient state j. $\overrightarrow{q} = (q_{1\alpha}, q_{2\alpha}, ..., q_{M\alpha})$ is a $1 \times M$ absorbing probability vector, containing probabilities of direct transition from either of the transient states to the absorbing state α. Note that matrix Q and its elements are not the same as matrix Q_T, described earlier in Section IV-A. I_1 is a 1×1 (because of one absorbing state) identity matrix and $\overrightarrow{0^T}$ is a $1 \times M$ vector of zeros. The condition $\overrightarrow{q} + Q\overrightarrow{e} = 1$ must hold for an absorbing Markov chain, given by Eq. 2. That is, the probabilities of leaving any arbitrary system state to other states or staying at the same state, all should sum up to one, where $\overrightarrow{e}$ is a vector of ones of size $1 \times M$. Besides Q and $\overrightarrow{q}$, a $1 \times (M+1)$ vector $\overrightarrow{P} = (P_\alpha, P_1, ..., P_M)$ of the initial probability distribution of system states, is needed to fully parameterize an absorbing Markov chain. In our case, it includes the probabilities of ending up in either of the system states (including the absorbing state with probability P_α), after the end of the multi-hop rebroadcast (first phase). These probabilities could be derived from Eq. 1.

The Hamming weight (i.e. the number of nodes with status "1") of an arbitrary state i, denoted by $H(i)$, is the number of potential replying nodes in that state and accordingly, $N - H(i)$, denoted by $F(i)$, is the number of potential requesting nodes (with status "0") in the same state. The Hamming distance between two states i and j (with logically non-zero probability of transition), denoted by $HD(i,j)$, shows the number of "$0 \rightarrow 1$" changes. In other words, it shows the number of nodes that received a packet, as a result of transition from state i to j. Given this, we aim to find the probability of transition from arbitrary state i to j, denoted by q_{ij}.

Since in a "multi-hop" setting not all nodes are in one-hop neighborhood of each other, for modeling we need to differentiate among nodes, based on their positions in the network. Hence, we explicitly consider:

(i) the set of node(s) sending request(s), among the potential requesters in a particular state (denoted by the random variable A);
(ii) the set of node(s) receiving the request(s), among the potential repliers (denoted by the random variable B);
(iii) the set of node(s) replying back to the received request(s) (denoted by the random variable C);
(iv) the particular set of node(s) receiving the reply(ies), among all the existing nodes with status "0" (denoted by the random variable D, with the only instance d). This results in the corresponding $0 \rightarrow 1$ turns and consequently, the changed system state.

One may note that these four random variables correspond to four steps of a transition from an arbitrary state i to an arbitrary state j. Such randomness implies the fact that a transition might be achieved through various ways, depending on the instances of these random variables, as shown in the example in Figure 6. Each of these ways may result in a different transition probability which contributes to the overall q_{ij} probability of transition from state i to state j.

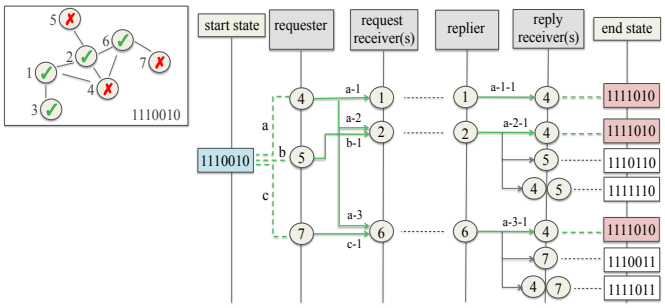


Figure 6: Transition from $i = 1110010$ to $j = 1111010$.

Recalling from the system state space and the Hamming weight property of states, we attribute an arbitrary state i to two vectors $\overrightarrow{rep_i}$ of size $1 \times H(i)$ and $\overrightarrow{req_i}$ of size $1 \times F(i)$, representing the ID's of the potential repliers and requesters in state i, respectively. For instance, considering $i = 1110010$ as a transient state of a network of $N = 7$ nodes and labeling nodes from left to right, we have $\overrightarrow{rep_i} = (1,2,3,6)$, $\overrightarrow{req_i} = (4,5,7)$.

Due to the topology-dependent reachability limitations between nodes, the random variables A, B, C and D are conditionally dependent. That is, the identity of nodes that act at each step of a transition, depends on the identity of nodes that acted in the previous step. Bayesian Networks can capture such conditional dependencies. Given this, we define a Bayesian network, represented by the following Directed Acyclic Graph (DAG) [16], as shown in Figure 7.

The rationale behind considering "maximum independent" subsets in the first and third step, lies in the fact that according to the reliable (geo-)broadcast protocol, neighboring nodes hear each other and do not send identical requests and replies. This implies that in these steps, we just need to consider

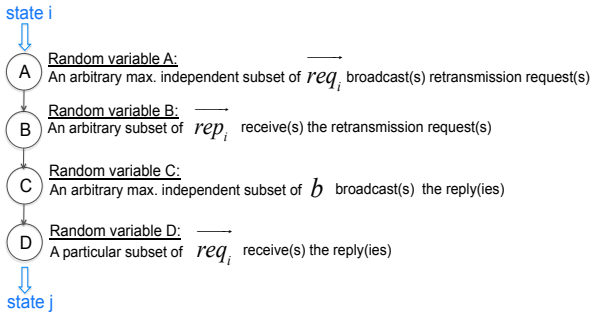


Figure 7: The DAG of the transition from state i to j.

the maximum independent subsets (i.e. pairwise non-adjacent nodes) of potential requesters / repliers, as the ones that may really send requests / replies for the same packet.

Following this, we have the joint probability of A, B, C and D, by the chain rule of Bayesian networks [16] as

$$P(A,B,C,D) = P(A)P(B \mid A)P(C \mid B)P(D \mid C). \tag{3}$$

Recalling the last step of the transition, we are interested in extracting $P(D=d)$ (denoted by $P(d)$) in Eq. 3, which is indeed q_{ij}. We can compute $P(d)$ from the joint probability by the method of marginalization [12], which is summing out variables A, B and C. Accordingly, $P(d)$ is given by

$$P(d) = \sum_{\forall a \in A} \sum_{\forall b \in B} \sum_{\forall c \in C} P(a)P(b \mid a)P(c \mid b)P(d \mid c). \tag{4}$$

By pushing the sums we have $P(d)$ as

$$P(d) = \underbrace{\sum_{\forall c \in C} P(d \mid c) \underbrace{\sum_{\forall b \in B} P(c \mid b) \underbrace{\sum_{\forall a \in A} P(b \mid a) \underbrace{P(a)}_{\text{P(a)}}}_{\text{P(b)}}}_{\text{P(c)}}}_{\text{P(d)}}, \tag{5}$$

This is computationally less complex and could be solved by an inside-out dynamic programming method. In what follows, we calculate $P(a)$, $P(b \mid a)$, $P(c \mid b)$ and $P(d \mid c)$ in four steps.

Step 1. $\mathbf{P(a)}$ This is the probability of a given maximum independent subset a of nodes in $\overrightarrow{req_i}$, to be the requesters for a missing packet. We reasonably assume that in practice all the nodes that miss a packet and pairwise are far enough not to hear each other, will send a request for that packet. This implies that only the nodes in a Maximum independent subset of nodes in $\overrightarrow{req_i}$ will send such a request. Considering all the maximum independent subsets are equally likely to be chosen, the probability $P(a)$ is given by $1/nr(MIS(\overrightarrow{req_i}))$, where $nr(MIS(\overrightarrow{req_i}))$ represents the number of all maximum independent subsets of $\overrightarrow{req_i}$.

Step 2. $\mathbf{P(b \mid a)}$ We define the following two probabilities for a given node $n \in \overrightarrow{rep_i}$, to either belong to the subset b or its complement $b^{'}$ related to $\overrightarrow{rep_i}$.

$$P(n \in b) = 1 - (1 - p_T)^{|a_n|}, \tag{6}$$

where $|a_n|$ represents the cardinality of the set of nodes in a that can reach node n. Such probability implies that node n should receive a request, at least from one of the nodes in a_n. Accordingly, the probability $P(n \in b^{'})$ is given by the following equation.

$$P(n \in b^{'}) = (1 - p_T)^{|a_n|}. \tag{7}$$

Consequently, $P(b \mid a)$ is given by

$$P(b \mid a) = \prod_{\forall x \in b} P(x \in b) \prod_{\forall y \in b'} P(y \in b^{'}), \tag{8}$$

implying that the probability that only and only the nodes in the subset b and no other node in $\overrightarrow{rep_i}$ receive the request(s), sent by nodes in the subset a.

Step 3. $\mathbf{P(c \mid b)}$ Inline with the assumption in Step 1, $P(c \mid b)$ is given by

$$P(c \mid b) = 1/nr(MIS(b)), \tag{9}$$

where $nr(MIS(b))$ represents the number of all maximum independent subsets of the subset b.

Step 4. $\mathbf{P(d \mid c)}$ Similar to Step 2, in order to compute $P(d \mid c)$, we define the following two probabilities for a given node $n \in \overrightarrow{req_i}$, to either belong to the subset d or its complement $d^{'}$ related to $\overrightarrow{req_i}$.

$$P(n \in d) = 1 - (1 - p_T)^{|c_n|}, \tag{10}$$

where $|c_n|$ represents the cardinality of the set of nodes in c that can reach node n. Such probability implies that node n should receive a reply, at least from one of the nodes in $|c_n|$. Accordingly, the probability $P(n \in d^{'})$ is given by the following equation.

$$P(n \in d^{'}) = (1 - p_T)^{|c_n|}. \tag{11}$$

Consequently, $P(d \mid c)$ is given by

$$P(d \mid c) = \prod_{\forall x \in d} P(x \in d) \prod_{\forall y \in d'} P(y \in d^{'}), \tag{12}$$

where it implies the probability that only the nodes in d and no other node in $\overrightarrow{req_i}$ receive the reply(s), sent by nodes in the subset c. Now, by substituting all the relevant terms in Eq. 5, we arrive at $P(d) = q_{ij}$ as follows.

$$\begin{aligned} P(d) = q_{ij} = \sum_{\forall c \in C} \prod_{x=1}^{|d|} P(x \in d) \prod_{y=1}^{|d^{'}|} P(y \in d^{'}) \times \sum_{\forall b \in B} 1/nr(MIS(b)) \\ \times \sum_{\forall a \in A} \prod_{x=1}^{|b|} P(x \in b) \prod_{y=1}^{|b^{'}|} P(y \in b^{'}) \times 1/nr(MIS(\overrightarrow{req_i})). \end{aligned} \tag{13}$$

$P(d) = q_{ij}$ in Eq. 13 is capable of capturing all paths maybe traversed through, during the transition. Note that in a given scenario the transition from a particular state i to a particular state j could be logically feasible (i.e. there is no $1 \rightarrow 0$

change in the status of any node). However, due to the reachability constraints q_{ij} maybe just 0, implying it is not practicable to transit directly from state i to j.

The elements of the absorption vector $\overrightarrow{q}$ are derived similarly by considering $d = \overrightarrow{req_i}$ in Eq. 13. Given this, we have the main attributes of the AMC and this completes the parameterization of the system.

C. PERFORMANCE METRICS

The Key Performance Indicators used to assess the performance of the mechanism, are introduced in this section.

1) Complete delivery probability: Considering the retransmission request-based property of the introduced error recovery technique, it is relevant the probability of successful packet delivery to all receivers within k retransmission steps. Hence, for the modeled AMC, we focus on the Cumulative Distribution Function (CDF) of the number of retransmission steps k, until absorption. Note that for an AMC, such a probability is given by a First Passage Time (FPT) distribution, interpreted as the number of steps, required to end up in the absorbing state [13]. Given this and denoting the random variable, describing the number of retransmission steps until absorption by S, the Cumulative Distribution Function (CDF) of S is given by

$$F_S(k) = Pr(S \leq k) = 1 - \vec{P}^{'}(Q^k)\vec{e}, \tag{14}$$

where $\vec{P}^{'}$ is the initial state probability vector $\overrightarrow{P}$, without P_α.

2) PMF of the number of retransmission steps: Besides absorption within a limited number of retransmission steps, it may be of interest as well, the probability of absorption in exactly a predefined number of retransmission steps. That is the Probability Mass Function (PMF) of the number of retransmission steps and obtained from CDF, as follows.

$$\begin{aligned} &f_S(0) = F_S(0), \\ &f_S(k) = F_S(k) - F_S(k-1), k = 1, 2, \ldots, K. \end{aligned} \tag{15}$$

3) Overhead of error recovery: Based on the fact that this receiver-based error recovery mechanism is built up on retransmission requests, it is also relevant to know the imposed overhead of such a scheme and investigate how many retransmission steps k are to be done in a given scenario, for all the nodes to successfully receive a packet. This is given by the following equation.

$$H = \sum_{k=1}^{\infty} f_S(k) \cdot k. \tag{16}$$

It is interpreted as the expected number of retransmission steps, obtained by taking average over probabilities of absorption, for all the given number of retransmission steps k.

4) Residual loss probability: In order to keep a reasonable trade-off between the overhead and improved reliability, it is also equally important to keep track of the system failure, in the case of the limited number of allowed retransmission steps. For this, we define the residual loss probability metric. That is, the average probability for an arbitrary station to be failed, after $k{=}\ K$ retransmission steps, given by

$$p_{res}^K = \frac{\sum_{i=1}^{M} |S_i(n=0)| \cdot P_i^K}{M}, \tag{17}$$

where $|S_i(n=0)|$ is the cardinality of the set of nodes with status "0" in state i and P_i^K is i^{th} element of vector $\overrightarrow{P}.T^K$, representing the probability of being in state i after $k{=}\ K$ retransmission steps.

V. NUMERICAL RESULTS

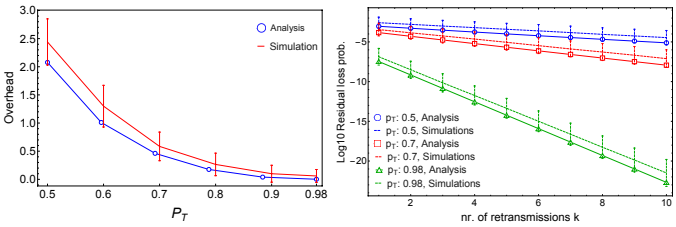
The results of evaluating the performance of the system, in terms of the introduced metrics, are presented in this section. First, we implement the analytical model, using a Wolfram Mathematica implementation and validate it using simulations. Then we continue analyzing the system in detail, relying on the validated analytical model. Key system parameters such as the number of nodes, successful packet transmission probability, the transmission range and the number of retransmission requests are considered as the inputs to the model. The results for each numerical setting are obtained by averaging over twenty random topologies of a given number of nodes and transmission range, uniformly distributed in a $2km \times 2km$ two-dimensional area. The rebroadcasting probability p_{rb}, in the multi-hop rebroadcast phase (first phase), is considered 1. Accordingly, the number of rebroadcasts b in the first phase is considered equal to the number of nodes N, assuming each node rebroadcasts only once, unless it is explicitly mentioned. Also note that the number of retransmission steps k refers to retransmission requesting steps in the error recovery phase (second phase), not to be confused with the number of rebroadcasts b in the first phase. Logarithmic scales are used in some graphs, as a means of reasonable presentation.

A. Validation by simulation

We use NS-3 to validate our model. For this, we set up a multi-hop wireless network of nodes, considering the above-mentioned specifications, with constant positions. This is a reasonable setting, considering the dissemination time of a packet, which is in the order of milliseconds, with regard to velocity of moving vehicles. The propagation loss is assumed to be negligible. Probabilistic packet transmission loss and the reliability assurance mechanism are implemented at the application level on all nodes. 0.95% confidence limits are used for the purpose of validation.

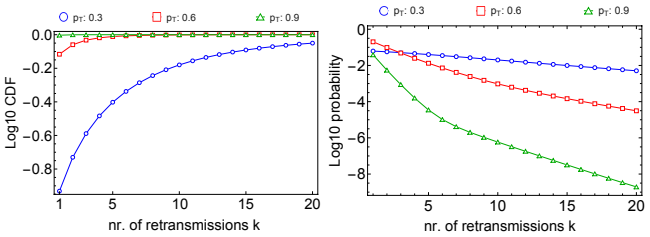
Figure 8a shows the results, obtained via simulations and the analytical analysis, as the overhead of error recovery, in terms of the number of retransmission steps, against increasing probability of successful packet transmission p_T, in a network of $N = 10$ nodes with the transmission range $R = 0.5km$. As expected, overhead follows a descending flow by an increase in p_T and the two curves follow each other closely. Figure 8b shows the results, obtained via simulations and the analytical analysis, for the residual probability of loss in a network of $N = 10$ nodes and $R = 0.5km$. This probability decreases at the expense of more retransmissions. In particular, we

observe a sharper decrease supported by a higher probability of successful packet transmission $p_T = 0.98$. The results obtained from simulations are inline with the analysis, being within the confidence limits. Though they slightly fall apart, due to practical implementation differences between platforms and simplifications of the mathematical modeling. Relying on these observations, confirming the validity of the analytical modeling, in the rest of the paper, we merely consider the analytical model.

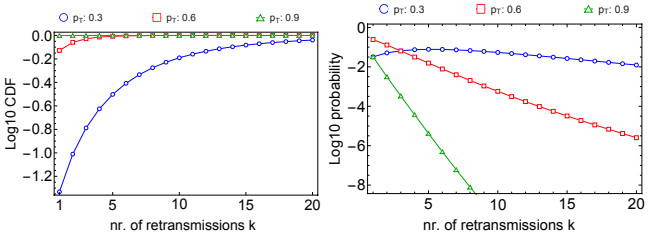


(a) Overhead of error recovery. (b) Residual loss probability.

Figure 8: Validation test cases.

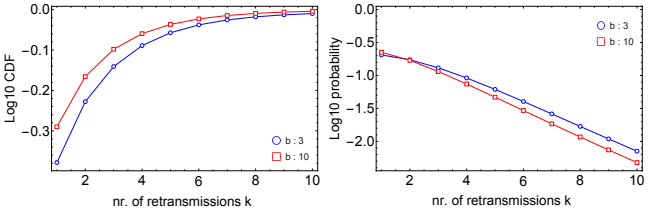


(a) $CDF, N = 5$ (b) $PMF, N = 5$



(c) $CDF, N = 10$ (d) $PMF, N = 10$

Figure 9: CDF and PMF of the number of retransmission steps for different probabilities, $N = 5, 10$.



(a) $CDF, N = 10$ (b) $PMF, N = 10$

Figure 10: CDF and PMF of the number of retransmission steps for $p_T = 0.5$ and different rebroadcast steps, $b = 3, 10$.

B. Analytical System Performance

In this section we present the results of analysing the analytical model, considering various numerical settings.

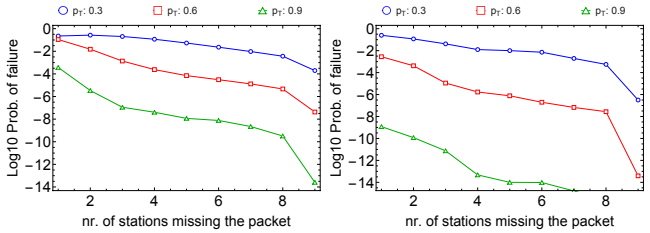
Figure 9 demonstrates CDF and PMF of the number of retransmission steps, until all stations receive the packet in a network of $N = 5$ nodes, during the error recovery phase, for different packet success probabilities, p_T. Figure 9c and 9d demonstrate the same metrics for $N = 10$ nodes. For both N, we observe that PMFs are more evenly distributed for lower success probabilities, confirming necessity of more retransmission steps to be carried out to end up in the absorbing state, where all nodes have successfully received the broadcast packet. Comparing the peer graphs for $N = 5$ and $N = 10$, we follow a faster decrease in PMF values, against increasing number of retransmission steps, for $N = 10$. This indicates the positive effect of increased network density on system absorption. The common remark regarding Figure 9b and 9d is the lightweight tail distribution of probability of absorption over the number of retransmission steps. This could be a determinative knowledge for the trade-off between improved reliability and imposed overhead, in terms of the number of retransmission steps.

It is likely that all nodes successfully receive a packet at the initial rebroadcast phase. Such a probability (P_α in $\overrightarrow{P}$), which is not included in Figure 9b and Figure 9d, contributes to the overall probability of successful packet delivery to all nodes. Such that, adding this probability to the corresponding sum of PMFs of all the number of retransmission steps (i.e. CDF), sums up to 1, as it is supposed to be and shown in Figure 9a and Figure 9c. The same is valid for all the other experiments.

In Figure 10, we see the effect of the number of rebroadcasts b in the first phase (i.e. the multi-hop rebroadcast), on the error recovery (second phase) performance, for the transmission range $R = 0.5km$ and the packet success probability $p_T = 0.5$. These figures demonstrate CDF and PMF of the number of retransmission steps k, until all stations successfully receive the packet, during the error recovery phase, given different number of rebroadcasts b, in the first phase.

These observations are mainly useful in the case of applying probabilistic rebroadcasts in the first phase. For instance, we see no big difference between results, belonging to $b = 3$ and $b = 10$ number of rebroadcasts. Hence, the rebroadcasting probability p_{rb} could be adjusted, not to overload the network with excessive rebroadcasts.

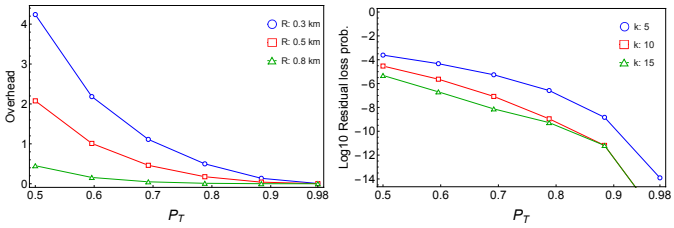
Figure 11 demonstrates the probability that a given number of stations fail to receive the packet, after $k = 2, 8$ retransmission requests, for three different probabilities $p_T = 0.3, 0.6, 0.9$. What is straight in these figures, is sharper decrease of the probability of failing having higher number of unsuccessful stations, for higher probability of success p_T. The same behavior, in a more intense scale, is observable while comparing Figures 11a and 11b. In particular, when supported by better channel conditions (i.e. for the higher $p_T = 0.9$), increasing the number of retransmission steps k, results in way more significant improvement. Figure 12a demonstrates the overhead of error recovery, in terms of the number of



(a) $k = 2$ (b) $k = 8$

Figure 11: Failure probability after k retransmission steps, $N = 10$, $R = 0.5km$.

retransmission steps k against increasing probability of successful packet transmission p_T, for three transmission ranges $R = 0.3, 0.5, 0.8km$, in a network of $N = 10$ nodes. We note a smooth decay in the imposed overhead, as either p_T or R increases. The dominant p_T factor, suppressing the gap between curves corresponding to various R, is traceable as the curves merges towards higher p_Ts.



(a) Overhead of error recovery. (b) Residual loss probability.

Figure 12: Recovery overhead and residual probability of loss, $N = 10, R = 0.5km$.

The residual loss probability is shown in Figure 12b for a network of $N = 10$ nodes and $R = 0.5km$. Expected behavior of the system is confirmed by the model, as we see smaller probabilities of residual loss, against increase in p_T. Non-significant gap between curves, belonging to three different number of retransmission steps k, implies the fact that most of nodes are reachable by fewer number of retransmissions.

VI. Conclusion and Future Work

In this paper, we analytically modeled and evaluated a reliability assurance mechanism, for multi-hop broadcast setting, as a means of providing detailed insight into the system functionality. Using simulations, we proved the validity of the model. These observations are useful for designing relevant optimization schemes, for reliable data dissemination. For instance, parameters such as nodes' transmission range or, particularly in this case, the number of retransmissions could be tuned, given various network densities and channel conditions.

Our model is general, in a sense that it is not based on a specific network topology, but topologies are input to the model, in terms of connectivity graphs. Also, the model is extensible such that elaborate $f(\theta)$ functions, considering detailed parameterization of data dissemination, can be augmented into the transition probability matrix that we developed.

Ensuring ultra reliable broadcast, using existing technologies [6] is not yet achieved. Motivated by the concepts of Internet of Vehicles (IoV) and heterogeneous networking, as the main design criteria for the next generation of the mobile networking system, 5G, in our next work we will investigate the gain of using cellular communication technologies, to contribute to vehicular communications with a higher level of reliability.

References

[1] IEEE standard for information technology– local and metropolitan area networks– part 11: WLAN MAC and PHY specifications amendment 6: Wave. *IEEE Std 802.11p-2010*, pages 1–51, July 2010.

[2] T. Atechian and L. Brunie. Dg-castor for query packets dissemination in vanet. In *Mobile Ad Hoc and Sensor Systems, 2008. MASS 2008. 5th IEEE International Conference on*, pages 547–552, Sept 2008.

[3] K. Bilstrup, E. Uhlemann, E. G. Strom, and U. Bilstrup. Evaluation of the ieee 802.11p mac method for v2v communication. In *Vehicular Technology Conference (VTC), 2008. IEEE 68th*, pages 1–5.

[4] Lee K. Cha, S. and H. Cho. Grid-based predictive geographical routing for inter-vehicle communication in urban areas. *Distributed Sensor Networks*, 2012, Jan.

[5] European Telecommunications Standards Institute Technical Specification. "102 636-4-1 v1.1.1: Intelligent Transport Systems (ITS); Vehicular Communications; GeoNetworking". Technical report, 2011.

[6] A. Festag. Standards for vehicular communication–from IEEE 802.11p to 5G. *Elektrotechnik und Informationstechnik*, 132(7):409–416, 2015.

[7] R. Fracchia and M. Meo. Analysis and design of warning delivery service in intervehicular networks. *Mobile Computing, IEEE Transactions on*, 7(7):832–845, July 2008.

[8] M. Gholibeigi and G. Heijenk. Analysis of multi-hop broadcast in vehicular ad hoc networks: A reliability perspective. In *2016 Wireless Days (WD)*, pages 1–8, March 2016.

[9] M. Gholibeigi, G. Heijenk, D. Moltchanov, and Y. Koucheryavy. Analysis of a receiver-based reliable broadcast approach for vehicular networks. In *IEEE Vehicular Networking Conference (VNC), 2014*.

[10] M. Gholibeigi, G. Heijenk, D. Moltchanov, and Y. Koucheryavy. Analysis of a receiver-based reliable broadcast approach for vehicular networks. *Ad Hoc Networks*, 37, Part 1:63 – 75, 2016. Special Issue on Advances in Vehicular Networks.

[11] B. Hassanabadi and S. Valaee. Reliable periodic safety message broadcasting in vanets using network coding. *Wireless Communications, IEEE Transactions on*, 13(3):1284–1297, March 2014.

[12] E.T. Jaynes and G.L. Bretthorst. *Probability Theory: The Logic of Science*. Cambridge University Press, 2003.

[13] J. Kemeny and J. Snell. *Finite Markov Chains*, volume 40 of *Undergraduate Texts in Mathematics*. Springer-Verlag, 1960.

[14] W. Klein Wolterink, G. Heijenk, and G. Karagiannis. Constrained geocast to support cooperative adaptive cruise control merging. In *Vehicular Networking Conference (VNC), IEEE*, pages 41–48, Dec 2010.

[15] A. Lakas and M. Shaqfa. Geocache: Sharing and exchanging road traffic information using peer-to-peer vehicular communication. In *Vehicular Technology Conference (VTC Spring), IEEE 73rd*, pages 1–7, May 2011.

[16] Richard E Neapolitan et al. *Learning bayesian networks*, volume 38. Prentice Hall Upper Saddle River, 2004.

[17] H. Rahbar, K. Naik, and A. Nayak. Dtsg: Dynamic time-stable geocast routing in vehicular ad hoc networks. In *Ad Hoc Networking Workshop (Med-Hoc-Net), The 9th IFIP Mediterranean*, pages 1–7, June 2010.

[18] M. Slot and V. Cahill. End-to-end acknowledgement of geocast in vehicular networks. In *Vehicular Networking Conference (VNC), 2011 IEEE*, pages 131–138, Nov 2011.

[19] L. Yang, J. Guo, and Y. Wu. Piggyback cooperative repetition for reliable broadcasting of safety messages in vanets. In *6th IEEE Consumer Communications and Networking Conference (CCNC) 2009*, pages 1–5, Jan.