ERP: An Efficient and Reliable Protocol for Emergency Message Dissemination in Vehicular Ad Hoc Networks

I-Hong Hou, Yan Gao, Yu-En Tsai, Jennifer Hou
University of Illinois, Urbana, IL 61801
{ihou2,yangao3,ytsai20}@.uiuc.edu, jhou@cs.uiuc.edu

Abstract—**Many safety-related applications in Vehicular Ad Hoc Networks require fast and reliable emergency message dissemination through multi-hop broadcast. However, the conventional broadcast mechanism is neither efficient nor reliable because it results in serious contention and collisions, which is usually referred to as the *broadcast storm* problem. In this paper, we propose ERP, a two-phase broadcast protocol that improves both efficiency and reliability. The first phase, a "fast-propagation phase", is designed to improve efficiency. We explicitly designate forwarders to relay the message and thus ensure both collision-free and quick propagation. The second phase, a "loss-recovery phase", enhances reliability. In this phase, nodes overhear the message and repeatedly broadcast it for the benefit of nodes which have not received the message in the first phase. We analytically show that using a density-aware power control mechanism in the second phase can efficiently improve the recovery rate. We also demonstrate how to find the optimal transmission power. Simulation results illustrate that our protocol outperforms probabilistic forwarding, which is currently the most widely studied solution, by a factor of 2 to 3.**

I. Introduction

Vehicular Ad Hoc Networks (VANETs) are of significant interest due to their capability of providing a variety of services to car drivers. Several applications have been identified and classified [1]. Communication-based automotive applications include safety warnings, traffic efficiency messages (e.g. free-flow payment), and *infotainment* (e.g. internet access). Among all these, safety-related applications attract the most attention for their potential of improving transportation safety on the road through infrastructureless, vehicle-to-vehicle (V2V) wireless communications [20][5].

In this paper, we study the safety application specified by the *backward emergency warning*: after detecting dangerous situations such as a road hazard or a sudden accident, the in-front cars notify their backward cars to assist their drivers to decelerate, brake or detour. This application requires disseminating a small amount of emergency messages to all reachable nodes within a certain geographical area and within a short time span, in order to prevent potential accidents. A broadcast mechanism that is both *efficient* and *reliable* is therefore required to ensure every driver receives the information timely.

However, in the conventional broadcast mechanism where every node forwards whatever message it receives, a problem called *broadcast storm* can arise. This arises because of the severe contention and packet collision among neighboring nodes that result from the naive transmitting mechanism in the shared wireless medium. It has been shown to result in high packet loss ratio [13]. This can be dangerous to safety critical applications in VANETs, thus motivating a better protocol for emergency message dissemination.

Previous work [17] has suggested the usage of different variations of probabilistic forwarding to mitigate the broadcast storm problem. Being probabilistic in nature, it's possible the message will die out. Moreover, the solutions do not consider node density in the neighborhood. Different node densities can result in very different performance. Hence, omitting this factor can cause the solutions to fail in specific scenarios since they do not always use the best probability to forward. A density-aware solution is thus desired.

In this paper, we propose a two-phase broadcast protocol for emergency message dissemination. In the first phase, a "*fast-propagation phase*", we use unicast with ACK in order to avoid the broadcast storm problem. The destination of the unicast is chosen from its backward neighborhood, starting from the farthest neighbor. All nodes can potentially overhear this emergency message, but since only one node transmits at any given time, the broadcast storm problem does not arise. The emergency message can thus be propagated quickly in this first phase. Further, the reliability of unicast can be guaranteed since we require an ACK from the receiver. In the second phase, a "*loss-recovery phase*", nodes that overhear the emergency message also periodically broadcast it after the fast-propagation phase. We apply power control in this loss-recovery phase to improve spatial reuse and reduce contention and collisions. The tradeoff between one-hop reliability and link redundancy is evaluated by a mathematical model. This analytical model is based on the DSRC protocol, a variation of IEEE 802.11 with extension for the outdoor high-speed vehicle environment. We show that there is an optimal number of transmitters, which is independent of the channel condition, that minimizes the total dissemination latency. We further show that a certain degree of connectivity can be ensured under our power control algorithm.

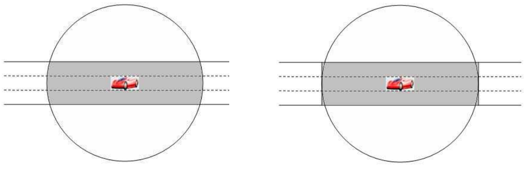
Another contribution of this work is that we propose two novel metrics, *coverage probability* and *coverage latency*. Most current metrics in the context of VANETs fail to capture the most important feature of emergency message dissemination: *every* vehicle needs to receive the message within a *limited* time. Our metrics, on the contrary, can provide a global view of both propagation speed and reliability. In terms of these two metrics, we theoretically study the performance of our protocol. We prove that the coverage latency of our protocol is at most a constant factor from a theoretical lower bound. Simulations show that our analytical results can accurately predict the performance of the proposed protocol. Moreover, compared to the probabilistic forwarding scheme, our protocol can enhance performance by 2 to 3 times in terms of the two metrics.

The rest of the paper is organized as follows. Section II introduces the traffic, channel, and network models we use throughout the paper. Section III presents the ERP protocol. Section IV describes how to optimize the loss-recovery phase by applying power control. Section V theoretically evaluates the performance of ERP and shows its coverage latency is at the same order of the theoretical lower bound. We demonstrates that simulation results concur with theoretical analysis in Section VI. Section VII provides some background information and Section VIII concludes the paper.

II. Traffic, Channel, and Network Models

We model a vehicular ad hoc network as comprised of a set of cars X, evenly distributed along multiple lanes. For each lane, the inter-vehicle distance is d. We will show how to modify our protocol under non-uniform topology in Section 4.2. Also, we assume that each node[1] is equipped with a GPS device. Each node periodically broadcasts its current GPS position, velocity and heading so that all its neighbors within the maximum transmission range receive the information. Vehicles can learn the number of their neighbors and the location of each neighbor by this information.

Let $r_T(x)$ denote the transmission radius of a car $x \in X$. We consider the wireless channel as a lossy channel in which the nodes within $r_T(x)$ have the probability p of correctly receiving a message in absence of interference. This channel model is widely used to characterize the unreliable nature of wireless communication. To capture the effects of interference, we use a variation of the protocol model introduced by [7]. For each transmitting node x, its interference radius is $r_I(x) = cr_T(x)$, where c is a constant greater than 1. We define the transmission range and the interference range of a node as the circles centered at that node with radii r_T and r_I, respectively. A node can correctly receive a message with probability p if and only if it lies in the transmission range of a sender and outside the interference range of all other senders.



(a) Strip of a circle

(b) Rectangular approximation

Fig. 1: Effective transmission range

Since vehicles are confined within the roads, we define the *effective transmission range* of a node as the intersection of its transmission range and the road. Typically, the effective transmission range is a strip of a circle, as shown in Fig. 1a. However, the transmission radius is usually much larger than the width of the road. We can therefore approximate the effective transmission range as rectangular with length $2r_T$, as depicted by the grey area in Fig. 1b. The *effective interference range* is defined similarly.

Let $N(x)$ denote the set of one-hop neighbors of x. The one-hop neighbor of x in this paper refers to the node that lies in the effective transmission range of x. We define the heading of car x as the forward direction and the opposite direction as the backward direction. Then let $N^+(x)$ denote the set of forward neighbors which belong to $N(x)$ and lie in front of x. Similarly, let $N^-(x)$ denote the backward neighbor set consisting of the neighbors lying behind x.

Finally, we assume each node uses the IEEE 802.11-based protocol, DSRC, as its MAC protocol. IEEE 802.11e supports Quality of Service by tuning values of Arbitration Inter Frame Space ($AIFS$) and Contention Window (CW) [12]. Before transmitting a packet, a node needs to sense the channel. After detecting the channel being idle for an $AIFS$, the node chooses a backoff counter randomly from the interval $[0, CW - 1]$. The backoff time counter is decremented by 1 whenever the channel is sensed idle for a slot_time, which is $9\mu s$. The node transmits when the backoff time reaches 0. By choosing smaller $AIFS$ and CW, nodes with higher priority can be granted greater chance to access the channel. When the transmitted packet is an unicast

[1] Since we are considering VANETs. Throughout this paper, the terms "car" and "node" are used interchangeably

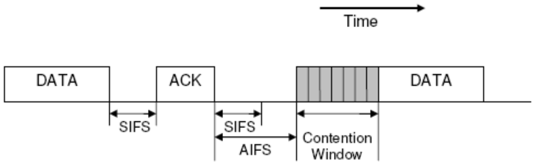


Fig. 2: An example of IEEE 802.11e backoff mechanism

packet, the receiver will reply with an ACK after a Short Inter Frame Space($SIFS$), which is $16\mu s$. The value of $AIFS$ in time is defined as $SIFS$ plus $AIFS$(in number)$*$slot_time. For example, by setting $AIFS = 7$, the duration of an $AIFS$ is $16\mu s + 7 * 9\mu s = 79\mu s$. An example of the backoff mechanism is demonstrated in Fig. 2. We assume $AIFS$ and CW are also tunable in DSRC.

The 2004 FCC ruling [3] specifies DSRC will have six service channels and one control channel. The control channel is reserved for the use of safety-related applications to grant higher priority for safety messages. While there may be other safety applications that access the control channel, such as location exchange applications that periodically broadcast the position of a vehicle, these applications access the control channel infrequently. Normally, vehicles exchange their GPS information every several seconds, whereas an emergency message is to be propagated from the source to far-way vehicles with latency within tens of milliseconds. Hence, interference caused by these applications is negligible and we assume there is no other applications going on during the period of emergency message propagation.

III. Protocol Overview

In this section we describe ERP, a two-phase *emergency message dissemination protocol*. As stated in Section 1, emergency message dissemination is both time-critical and reliability-critical. ERP satisfies both criteria by its two-phase mechanism: in the *fast-propagation phase*, ERP disseminates an emergency message to the farthest car as fast as possible. In the *loss-recovery* phase, ERP efficiently recoveries the nodes who failed to receive the message in the first phase.

A. Phase 1: Fast-Propagation Phase

As mentioned before, the naive flooding strategy results in a broadcast storm problem [13]. The first phase of ERP is designed to solve this problem and guarantee fast packet penetration. Its design is described below. In the description, we define the node that initiates the process as the *initiator*, and nodes that help relay the message as the *forwarder*.

1) A forwarder $f^{(n)}$ first chooses the farthest neighbor in $N^{-}(f^{(n)})$ as the potential next forwarder $\hat{f}^{(n+1)}$. Then it uses the maximum transmit power to unicast the emergency message to $\hat{f}^{(n+1)}$. If $n = 0$, *i.e.*, the node is the initiator, we set $AIFS = 2$ and $CW = 16$. Otherwise, we set $AIFS = 7$ and $CW = 4$. Further, if the forwarder overhears the same emergency message during its backoff process, it aborts its own transmission. The larger value of CW for the initiator is to avoid the risk where multiple vehicles initiate the same emergency message dissemination simultaneously. Under this setting, two initiators will choose the same backoff counter, which will lead to a collision, with a small probability $1/16$. The larger value of $AIFS$ for other forwarders is to make sure there is only one forwarder in each step, which will be explained in following paragraph.
2) Upon the receipt of the message, only the designated next forwarder $\hat{f}^{(n+1)}$ replies an ACK to $f^{(n)}$, all other nodes in the effective transmission range overhear the message and buffer it for the loss-recovery phase. If no ACK comes back in $SIFS$, $f^{(n)}$ will attempt to reach another neighbor in $N^{-}(f^{(n)})$ closer than the last attempted node. If all nodes in $N^{-}(f^{(n)})$ have been attempted, the selection will roll back to the farthest neighbor. $f^{(n)}$ does not stop the retransmission process until it receives an ACK. In this step, we set $AIFS = 2$ and $CW = 4$.
3) After sending the ACK, the potential forwarder $\hat{f}^{(n+1)}$ becomes the next forwarder $f^{(n+1)}$ and repeats all the above from step 1.

Note that step 1 explicitly designating the next forwarder is designed to avoid collisions because at each run time of the fast-propagation phase, only the forwarder can access the channel. The "access" is passed to the next forwarder when the DATA-ACK handshake succeeds. Further, the $AIFS$ used by the forwarder is larger than the sum of $AIFS$ and CW used in the retransmission step. In case the ACK of $f^{(n+1)}$ is dropped, $f^{(n)}$ will always retransmit before $f^{(n+1)}$ attempts to access the channel, making $f^{(n+1)}$ gives up its role as a forwarder. Finally, when there are multiple vehicles to initiate the process, only one of them will become the initiator since all other vehicles will overhear its message and abort their own transmissions. In sum, the fast-propagation phase can largely eliminate the possibility of packet collision by ensuring that only one packet is propagated in the network. We can hence achieve fast packet penetration since we won't waste time for packet collisions.

In the fast-propagation phase, the emergency message is rapidly disseminated backwards since each node favors its farthest one-hop neighbor to forward the message. However, note that only the selected forwarders are guaranteed to receive the message. It is possible that a small number of nodes may fail to receive the message due to their channel loss. Hence the second phase of ERP is designed to recovery this loss.

B. Phase 2: Loss-Recovery Phase

One distinct feature of emergency message dissemination is that it requires high reliability for all nodes of interest. In the fast-propagation phase, we trade reliability for rapidness because in each one-hop broadcast, only the forwarder is guaranteed to receive the message. Instead of waiting for success of all one-hop neighbors, the current forwarder allows its designated forwarder to immediately broadcast the message to the next hop. Because of channel loss or interference, a small number of nodes may not receive the message after the fast-propagation phase. So we need a loss-recovery phase to cover these nodes. To improve reliability, a conventional solution, such as epidemic algorithms [4], has the nodes that hold the message periodically broadcast at random. This conventional way is simple and easy to implement. However, it may not efficiently recover the lost nodes in VANETs. This is because, on one hand, if the network is of high density, the frequent occurrence of collision degrades the performance; if the network is sparse, on the other hand, the low connectivity also reduces the recovery likelihood.

To enhance the performance of the loss-recovery phase, we propose a density-aware protocol. We find that Phase 2 is most efficient when each node has a specific number of neighbors, T_{opt}, and a specific value of CW, CW_{opt}. The values of T_{opt} and CW_{opt} are derived in Section 4. Since nodes periodically exchange their locations and velocities before ERP is initiated, they can tune their power levels to ensure they have the desired number of neighbors. They then use this power level and CW_{opt} to broadcast the emergency message repeatedly.

One major concern of the loss-recovery phase is that it should not interfere with the fast-propagation phase. To this end, we set $AIFS = 12$ when a node first attempts the loss-recovery phase. It will use $AIFS = 2$ in latter transmissions or if it overhears a loss-recovery phase packet from its neighbors. Note that the sum of $AIFS$ and CW in the fast-propagation phase is at most 11. By setting $AIFS = 12$, nodes can conclude that there are no nodes operating in the fast-propagation phase within their effective interference ranges after they sense the channel being idle for an $AIFS$. The probability of interference between the two phases is hence eliminated. A complete overview of ERP is shown in Algo 1.

Algorithm 1 ERP

```
while TRUE do
   if detects an emergency event then
      $AIFS \leftarrow 2$;$CW \leftarrow 16$
      {acts as an initiator}
      $f^{(1)} \leftarrow$ farthest node
      repeat
         send an emergency message to $f^{(1)}$
         $AIFS \leftarrow 2$;$CW \leftarrow 4$
         $f^{(1)} \leftarrow$ next farthest node
      until receive an ACK or overhear another emergency message
   if receives an emergency message then
      if $ID =$ destination of the message then
         reply with an ACK
         {acts as a forwarder}
         $AIFS \leftarrow 7$;$CW \leftarrow 4$
         $f^{(n+1)} \leftarrow$ farthest node
         repeat
            send an emergency message to $f^{(n+1)}$
            $AIFS \leftarrow 2$;$CW \leftarrow 4$
            $f^{(n+1)} \leftarrow$ next farthest node
         until receive an ACK or overhear another emergency message
      else
         $AIFS \leftarrow 12$;$CW \leftarrow CW_{opt}$
         {operates in the loss-recovery phase}
         tune power level according to Section 4
         repeat
            broadcast the emergency message
            if broadcast succeeds or overhear another broadcast message then
               $AIFS \leftarrow 2$
         until the node is outside the emergency area
```

IV. Optimization of Loss-Recovery Phase

Power control is known to be able to reduce contention and collisions [10][11]. Using a small power to transmit packets can reduce the interference imposed on other on-going transmissions. Thus, the reliability of one-hop transmission can be improved. However, the benefits of power control come with a price. Smaller power will result in smaller transmission range and hence lesser link redundancy. The resulting topology might not be robust against a small number of failed transmissions. To offer a reliable loss-recovery mechanism, we need to evaluate the tradeoff between one-hop reliability and link redundancy. In this section, we first derive the optimal transmission power under models described in Section 2. The resulting optimal transmission power is related to the contention window size and hence we derive the optimal contention window.

The optimal transmission power, together with the optimal contention window, can achieve the highest probability for lost nodes to receive the emergency message. Finally, we show how to choose the optimal power under non-uniform traffic.

A. Finding the Optimal Power

For the ease of analysis, we assume every node uses the same power and has the same transmission radius, r_T. Also, we assume most of the nodes have received the emergency message in the fast-propagation phase. This assumption is validated in Section 6.1, where we find that even when the channel reliability, p, is as low as 60%, more than 90% of the nodes can receive the emergency message in Phase 1 due to its retransmission mechanism.

Suppose there are a total of L lanes. Every node is in the effective transmission range of $T = 2r_T \times L/d$ nodes, which are referred to as *transmitters*, and in the effective interference range of $I = 2cr_T \times L/d = cT$ nodes, which are referred to as *interferers*. By tuning transmit power, we can set T to a desired value. Hence, we focus on choosing the optimal T.

The condition that a node can receive a message with probability p in a given time slot is that exactly one of its transmitters is transmitting and every interferer other than the transmitter is silenced. Given the contention window W, the probability that a node transmits in a randomly time slot is $\tau = \frac{2}{W+1}$[2]. Therefore, the probability of successfully receiving a message in a slot time is

$$Succ(T,\tau) = pT\tau(1-\tau)^{I-1} = pT\tau(1-\tau)^{cT-1}.$$

Suppose T_{opt} yields the highest $Succ(T,\tau)$ for a given τ. $Succ(T_{opt},\tau)$ should be at least as large as both $Succ(T_{opt}+1,\tau)$ and $Succ(T_{opt}-1,\tau)$. We derive the following inequalities:

$$\begin{aligned}
&Succ(T_{opt},\tau) \geq Succ(T_{opt}+1,\tau)\\
\Rightarrow &pT_{opt}\tau(1-\tau)^{cT_{opt}-1} \geq p(T_{opt}+1)\tau(1-\tau)^{c(T_{opt}+1)-1}\\
\Rightarrow &T_{opt} \geq \frac{(1-\tau)^c}{1-(1-\tau)^c} = T^-_{opt}
\end{aligned}$$

and

$$\begin{aligned}
&Succ(T_{opt},\tau) \geq Succ(T_{opt}-1,\tau)\\
\Rightarrow &pT_{opt}\tau(1-\tau)^{cT_{opt}-1} \geq p(T_{opt}-1)\tau(1-\tau)^{c(T_{opt}-1)-1}\\
\Rightarrow &T_{opt} \leq \frac{1}{1-(1-\tau)^c} = T^+_{opt}
\end{aligned}$$

From the two inequalities, we show that the optimal T must be an integer between T^-_{opt} and T^+_{opt}. Note that we have $T^+_{opt} - T^-_{opt} = \frac{1-(1-\tau)^c}{1-(1-\tau)^c} = 1$ and there is exactly one integer between T^-_{opt} and T^+_{opt}. Hence, the optimal T is well-defined. Thus the above condition is necessary as well as sufficient. Further, both T^-_{opt} and T^+_{opt} are independent of p, the reliability of the channel. This implies that choosing the optimal T is not related to the channel condition.

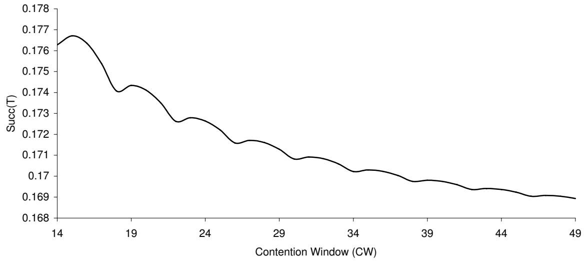


Fig. 3: $Succ(T_{opt},\tau)$ for different values of CW

B. Optimal Contention Window

As shown in the last section, both the values of T^+_{opt} and T^-_{opt} are determined by $\tau = \frac{2}{W+1}$. By choosing $T = T_{opt}$, $Succ(T,\tau)$ becomes a function depending solely on the size of the contention window. In this section, we show how to choose the optimal contention window, CW_{opt}.

Table I shows the value of T_{opt} for different values of CW. As in most current models, we set the ratio of interference radius and transmission radius, *i.e.* c, to be 2. When CW is smaller than 14, the corresponding T_{opt} is at most 3. This means each vehicle will have at most one transmitter in either the forward direction or the backward direction. The resulting topology can be disconnected due to even one failed node. To maintain a less vulnerable topology, T_{opt} should be at least 4, in which case each vehicle will have 2 transmitters both forwardly and backwardly. Thus, we only focus on the case when $CW \geq 14$.

We plot the values of $Succ(T_{opt},\tau)$ for all $14 \leq CW \leq 49$ in Fig. 3. $Succ(T_{opt},\tau)$ is maximized when $CW = 15$ and decreases as CW becomes larger. In our protocol, we choose $CW_{opt} = 15$ and $T_{opt} = 4$.

C. Enhancing Robustness Under Non-Uniform Traffic

In Section 4.1, we show that we should choose $T = T_{opt} = \lfloor\frac{1}{1-(1-\tau)^c}\rfloor$, or equivalently, $\lceil\frac{(1-\tau)^c}{1-(1-\tau)^c}\rceil$, to yield the highest probability for lost nodes to receive the emergency message. In our uniform traffic model, this implies each node has $\frac{T_{opt}}{2}$ neighbors on both forward and backward directions. However, in a real world scenario, the densities of vehicles in both directions might not be the same. Hence, we set the optimal power of a node as the minimal power such that it has at least $\frac{T_{opt}}{2}$ neighbors in both directions. In addition to enhancing performance, this setting also ensures a certain level of robustness according to the following theorem:

TABLE I: T_{opt} for different values of CW

CW	1–5	6–9	10–13	14–17	18–21	22–25	26–29	30–33	34–37	38 – 41	42–45	46–49
T_{opt}	1	2	3	4	5	6	7	8	9	10	11	12

Theorem 1: Under the power control algorithm described above, the resulting topology is $\frac{T_{opt}}{2}$-connected.

Proof: We number the vehicles in increasing order beginning from the initiator, which is numbered 1. To show the resulting topology is $\frac{T_{opt}}{2}$-connected, it suffices to show that there exist $\frac{T_{opt}}{2}$ node-disjoint paths between any two vehicles, i and j. We divide the nodes into $\frac{T_{opt}}{2}$ disjoint sets such that the k_{th} set $S_k = \{u|u = w\frac{T_{opt}}{2} + k, w \in N\}$. Each of the sets is connected since neighboring nodes in each set are $\frac{T_{opt}}{2}$ apart and each node has at least as many neighbors in each direction. Moreover, both i and j are connected to each of the sets since there exist some u and v in each S_k such that $|u-i| \leq \frac{T_{opt}}{2}$ and $|v-j| \leq \frac{T_{opt}}{2}$. Hence, we can find a path from i to j in S_k for every $1 \leq k \leq \frac{T_{opt}}{2}$, and there are at least $\frac{T_{opt}}{2}$ disjoint paths from i to j. ■

This result shows that our protocol can still work in the presence of a small number of faulty nodes. It also suggests that a designer can choose a larger T_{opt} when node failure becomes a frequent event.

V. Performance Analysis

In this section, we theoretically evaluate the performance of the two-phase emergency message dissemination protocol. We propose novel metrics to evaluate the performance of road hazard condition notification protocols. Such an application is both time-critical and reliability-critical because *every* vehicle in the area of interest needs to receive the emergency message in a timely manner. Most current metrics fail to reflect both criteria simultaneously. For example, *packet penetration rate* defined as the rate at which the packet propagates across the network cannot offer any information on reliability. Resta, Santi, and Simon [15] use another metric, the probability that a specific vehicle can receive the message within a given time, to capture both criteria. However, their approach only focuses on per-vehicle performance and lacks a global perspective. To obtain a global view on both time and reliability criteria, we propose using *coverage probability* and *coverage latency*. The *coverage probability* is defined as the probability that *every* vehicle in a certain area receives the message within a given time. The *coverage latency*, on the other hand, is defined as the minimum time needed for the coverage probability to exceed a certain threshold. In the following, we first derive the coverage probability of our protocol. We then show the coverage latency of our protocol is at most within a constant factor of the coverage latency of an optimized, but unrealistic, protocol.

In the analysis, we divide time into *rounds*. A round is defined as the time interval from the end of a data packet to the end of the next data packet. To be more specific, a round will include a period of $AIFS$, several periods of slot_time due to contention window, and the time needed for transmitting the emergency message. The duration of a round is not constant due to the variation of time slots randomly chosen from the contention window and different $AIFS$ in different steps of the protocol. However, we show this variation is negligible. To achieve better reliability, nodes should transmit the emergency message using the base rate, *i.e.* $1\ Mb/sec$. The size of a packet is usually around $1Kb$ and we need approximately $1ms$ to transmit the emergency message. The variations of both the contention window and $AIFS$ are at the scale of tens of microseconds. Therefore, we can neglect these differences and assume the duration of each round is the same henceforth.

A. Coverage Probability

For simplicity, we set the inter-vehicle distance to be 1 and normalize transmission radius and interference radius accordingly. We follow the same traffic, channel, and network models as described in Section 2. Under these models and the optimal power derived in Section 4, each node has T_{opt} transmitters in the loss-recovery phase. We also assume each relay node in the fast-propagation phase always chooses the farthest backward neighbor, which is r_{max} away, to forward the message. We assume it takes exactly u rounds for the node at ur_{max} to receive the message in Phase 1. In other words, there is no retransmission in Phase 1. The assumption, though somewhat optimistic, is not unrealistic since the reliability in unicast is usually high. We will later offset this optimistic assumption by adding the expected number of failed transmissions to the resulting formula.

To compute coverage probability, we need to know when the loss-recovery phase can start in each part of the region. Consider a node at distance d, $ur_{max} \leq d < (u+1)r_{max}$, it can start Phase 2 only when it is outside the effective interference range of the relay node in Phase 1. The interference radius of the relay node is cr_{max}. Supposing the position of the relay node is $d'r_{max}$, we have $d'r_{max} - d > cr_{max}$ and $d' \geq u + 1 + c$. Since it takes $u + 1 + c$ rounds for

Phase 1 to reach $(u+1+c)r_{max}$, the node at distance d can start Phase 2 at round $u+1+c$.

Next we derive the probability that a lost node can receive the emergency message in a round. From the perspective of a lost node, there are four different statuses for each time slot: *Success*, meaning that exactly one of its transmitters is transmitting and every interferer is silenced; *Busy*, all of its transmitters are silenced and some of its interferers are transmitting; *Collision*, meaning that some of its transmitters are transmitting but they get into collisions with other interferers; and *Idle*, all of its transmitters and interferers are silenced. We use P_{succ}, and P_{idle} to denote the probability of *Success* and *Idle*, respectively.

The condition that a node can receive a message with probability p, which is the channel reliability, in a given round, is that there is one *Success* time slot in the round. There is exactly one non-idle time slot in each round. The probability that a lost node can receive the emergency message in a round is $q = pP_{succ}/(1-P_{idle})$. By setting $T = T_{opt}$ and $\tau = \frac{2}{W_{opt}+1}$ as in Section 4, we have P_{succ} and P_{idle} as below and the value of q can be easily obtained. The value of q is 0.27 when $p = 0.9, c = 2, T_{opt} = 4$, and $W_{opt} = 15$.

$$P_{succ} = T\tau(1-\tau)^{I-1} = T\tau(1-\tau)^{cT-1}$$
$$P_{idle} = (1-\tau)^{I} = (1-\tau)^{cT}$$

Now we derive the coverage probability, $CProb$, of the region within nr_{max} from the initiator at a given round t. Considering a node at distance d, $ur_{max} < d < (u+1)r_{max}$, it can receive the message under the following two situations: it can receive the message from the node at ur_{max} or from the node at $(u+1)r_{max}$ in the fast-propagation phase with probability p; it can also receive the message from a neighboring node in the loss-recovery phase between round $(u+1+c)r_{max}$ and round t. In each round in Phase 2, the node can receive the message with probability q. Phase 2 starts at time $u+1+c$ and there are $t-u-c$ rounds for Phase 2 by round t. The probability that the node fails to receive a message under both conditions is $(1-p)^2(1-q)^{t-u-c}$. Therefore, the probability that the node can receive the message within round t is $1-(1-p)^2(1-q)^{t-u-c}$. Under the assumption that packet loss is independent at each node, the probability that all nodes between ur_{max} and $(u+1)r_{max}$ receive the message within round t is $[1-(1-p)^2(1-q)^{t-u-c}]^{r_{max}-1}$. Multiplying through for all $0 \leq u \leq n-1$ yields:

$$CProb(t) = \prod_{u=0}^{n-1}[1-(1-p)^2(1-q)^{t-u-c}]^{r_{max}-1} \quad (1)$$

The above formula is derived under the optimistic assumption that there is no packet loss in the fast-propagation phase. To offset the assumption, we notice that the expected number of transmissions to reach nr_{max} is n/p since the channel reliability is p. The expected number of failed transmissions is $n/p - n = \frac{1-p}{p}n$ and we need as many rounds to retransmit. The starting time of the Phase 2 at each node will be delayed by at most as many rounds. Hence, we can refine the formula as:

$$CProb(t) = \prod_{u=0}^{n-1}[1-(1-p)^2(1-q)^{t-\frac{1-p}{p}n-u-c}]^{r_{max}-1}$$

Note this result becomes too pessimistic since it assumes the starting time of Phase 2 at every node is delayed by every retransmission in Phase 1. However, when the failed transmission happens at distance $(u+1+c)r_{max}$, the loss-recovery phase within distance ur_{max} is not influenced. These two formulas can serve as upper-bound and lower-bound on coverage probability, respectively.

B. *Approximation Bound of Coverage Latency*

Coverage latency, $CTime$, is defined as the minimum time needed for the coverage probability to exceed a certain threshold, $P_{threshold}$. We can obtain the value by choosing the smallest t such that $CProb(t) \geq P_{threshold}$. Formally, it is given by

$$CTime := \min \arg_t\{CProb(t) \geq P_{threshold}\}.$$

While there is no close formula for the coverage latency, we can instead derive the approximation bound of the coverage latency compared to an optimized protocol.

We notice that there are two restrictions introduced directly by our channel and network models. First, for each round, the message can propagate no longer than r_{max}. Second, for each node within the area that the message has propagated, it has at most probability p to receive the message in every round. Let OPT be a protocol such that equality holds for both restrictions. Obviously, OPT is usually not achievable since it doesn't consider interference and packet collisions. Yet, it can serve as a performance bound to be compared to.

Before deriving the approximation bound, we need to obtain the coverage probability for OPT at round t within the range nr_{max}. Consider a node at distance d, $ur_{max} < d < (u+1)r_{max}$. It takes u rounds for the message to be propagated to ur_{max} and the node can receive the message with probability p from then on. The probability that the node at d can receive the message before round t is $1-(1-p)^{t-u+1}$. There are r_{max} nodes between ur_{max} and $(u+1)r_{max}$. Also, u can range from 0 to $n-1$. Hence, we have:

$$CProb_{opt}(t) = \prod_{u=o}^{n-1}[1-(1-p)^{t_{opt}-u+1}]^{r_{max}-1} \quad (2)$$

Suppose the coverage latency of OPT is t_{opt}. That is, $CProb_{opt}(t_{opt}) = P_{threshold}$. We want to find the minimum time t such that $CProb(t) \geq CProb_{opt}(t_{opt})$. We have:

$$CProb(t) = \prod_{u=0}^{n-1} [1 - (1-p)^2 (1-q)^{t-u-c}]^{r_{max}-1}$$
$$\geq CProb_{opt}(t_{opt}) = \prod_{u=o}^{n-1} [1 - (1-p)^{t_{opt}-u+1}]^{r_{max}-1}$$

For the inequality to hold, it suffices to choose t such that $1-(1-p)^2(1-q)^{t-u-c} \geq 1-(1-p)^{t_{opt}-u+1}$ for every u. We can further derive:

$$1 - (1-p)^2(1-q)^{t-u-c} \geq 1 - (1-p)^{t_{opt}-u+1}$$
$$\Rightarrow (1-q)^{t-u-c} \leq (1-p)^{t_{opt}-u-1}$$
$$\Rightarrow t \geq \frac{\ln(1-p)}{\ln(1-q)} t_{opt} + u + c - \frac{\ln(1-p)}{\ln(1-q)}(u+1)$$

Since q is the probability a node can receive the message when considering both interference and channel loss, we have $1 > p \geq q > 0$. This implies $\frac{\ln(1-p)}{\ln(1-q)} \geq 1$. By setting:

$$CTime = \min\{t\} = \frac{\ln(1-p)}{\ln(1-q)} t_{opt} + c - \frac{\ln(1-p)}{\ln(1-q)}$$

we can guarantee: $CProb(CTime) \geq CProb_{opt}(t_{opt})$. The constant term $c - \frac{\ln(1-p)}{\ln(1-q)}$ becomes negligible when t_{opt} is large. Hence, our protocol is a $\frac{\ln(1-p)}{\ln(1-q)}$-approximation protocol in terms of coverage latency. Both p and q are independent of the size of the network and the approximation ratio is a constant.

VI. Model Validation

To validate the theoretical analysis reported in the previous sections, we have implemented our protocol on top of ns-2. Ns-2 is a widely used network simulator that can simulate both IEEE 802.11 protocol and mobile nodes. In this section, we will present the simulation results as well as the corresponding analytical results. Meanwhile, the performance improvement over the probabilistic forwarding scheme will also be reported.

In this simulation, we consider the highway scenario in which cars are uniformly distributed along 4 lanes and all of them move towards the same direction. The velocity of each vehicle is randomly chosen between 60 km/hr and 120 km/hr. We assume nodes use omnidirectional antennas and the transmission range is a disk with radius of 200m. This value is chosen based on the result reported by [1]. It shows that the packet loss rate is below 20% when the v2v communication range is less than 200m. In addition, we assume the area of interest is a segment of 2000m behind the initiator. Other parameter values are in Table II. Unless otherwise specified, the following results have been obtained by considering the traffic with the density of 25 cars/km/lane. This implies there are 200 cars in the area of interest. We repeated each simulation 100 times.

TABLE II: Simulation Setup

area of interest	2000m	CS range	400m
max tx range	200m	data rate	1Mbps
message size	1000bit	slot_time	$9\mu s$

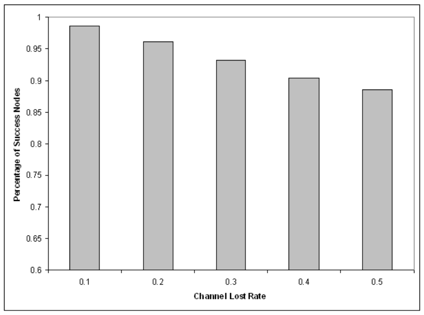


Fig. 4: Probability of Reception on the Fast-Propagation Phase v.s. Channel Loss Rate

A. Robustness of Fast-Propagation Phase

Recall that in the analysis of Section 4.1, we assume that after Phase 1, most of the nodes have received the emergency message except for a small number of nodes. We validate this assumption by counting the percentage of nodes that receive the emergency message in Phase 1. Fig. 4 plots the mean percentage of received nodes among the 100 simulation runs under different channel loss rate. The result confirms our assumption by showing that 89% of the nodes have successfully received the message after Phase 1 even the channel loss rate is as high as 50%. This result, though somewhat surprising, is due to two factors. First, every node can receive the emergency message from both the forward relay node and the backward relay node. Nodes can still receive the message even if it fails to overhear the transmission by one of the two relay nodes. Second, when the channel loss rate is high, there will be more retransmissions in Phase 1. Nodes that fail to overhear the first transmission may still receive the message in the following retransmissions.

B. ERP and Power Control

In this simulation, we demonstrate how different transmit powers can influence the behavior of Phase 2 in ERP. We assume 10% randomly chosen vehicles fail to receive the emergency message in Phase 1. We then evaluate the time required for Phase 2 to deliver the emergency packets to all these lost nodes.

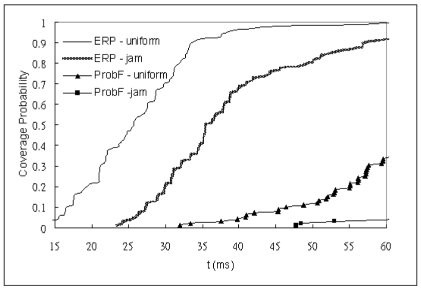


Fig. 6: $CProb$ comparison

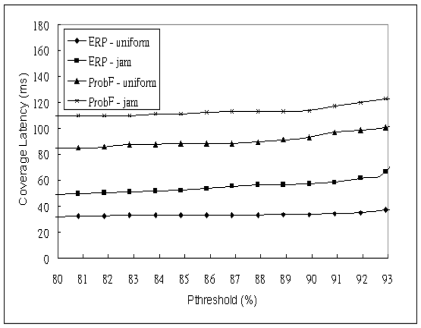


Fig. 7: $CTime$ comparison

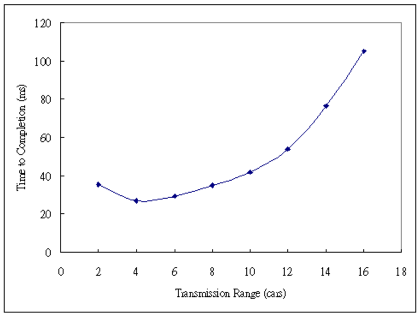


Fig. 5: The performance of ERP when the loss-recovery phase works at different powers

According to the analysis in Section IV, the loss-recovery phase will be more efficient if the power is optimized. Using a small power to transmit can reduce the interference and hence increase transmission reliability. On the other hand, it also results in lesser transmitters around the lost nodes and hence reducing the likelihood nodes overhearing each other. Fig.5 shows the time to completion in Phase 2 for different transmission powers, which are represented as the numbers of vehicles within the effective transmission range. We can observe that the time to completion is minimized when the number of neighbors is 4. This result confirms the theoretical analysis derived in Section 4.1 and Table I as the contention window size is chosen to be 16. Further, by using a large transmission power that can reach 16 vehicles, the full coverage latency can be more than 3 times worse than that by applying power control. This significant difference validates the need for power control.

C. Coverage Probability and Coverage Latency

We consider two metrics to evaluate ERP: coverage probability $CProb$ and coverage latency $CTime$. As suggested in Section IV, we assume ERP uses $CW = 15$ and $T_{opt} = 4$ in its loss-recovery phase. Two different scenarios are examined. In the first scenario, we assume vehicles are uniformly distributed with density 25 cars/km/lane on a four-lane highway. In the second scenario, we assume the density in the first half of the area of interest is 50 cars/km/lane and the density in the second half is 16 cars/km/lane. This is to simulate the scenario where a traffic jam occurs. Again, the velocities of vehicles are randomly chosen between 60 km/hr and 120 km/hr in both scenarios. Most current solutions for emergency message dissemination uses some variations of probabilistic forwarding (ProbF), in which every node that receives the message broadcasts it with probability p_f in each time slot. We compare the performance of ERP against that of ProbF. While different values are used in different work, we exhaustively evaluate ProbF for $0.01 \leq p_f \leq 0.1$ and find that setting $p_f = 0.03$ can yield better performance for the examined car densities. We hence assume $p_f = 0.03$ in the simulation.

Fig. 6 presents the coverage probability for both ERP and ProbF. The figure shows ERP always outperforms ProbF under both scenarios. This is because the two-phase mechanism in ERP can alleviate the broadcast storm problem and achieve better rapidness and reliability. Similar results are observed in Fig. 7, where we compare the two protocols in terms of $CTime$. The coverage latency of ProbF is always more than twice as large as that of ERP.

VII. Related Work

Most VANETs applications are based on the DSRC (Dedicated Short Range Communications)[8] standard. DSRC is an extension of the IEEE 802.11 technology that supports both Public Safety and Private operations in roadside-to-vehicle and vehicle-to-vehicle communication environments. The MAC layer of DSRC, based on the CSMA/CA protocol, is very similar to IEEE 802.11 with some modifications. Yin et al.[20] have shown that the DSRC standard is defective in terms of one-hop reliability. Xu et al.[18], therefore, propose protocols compatible with DSRC, setting the number of MAC layer repetitions to improve one-hop reliability. We show that low one-hop reliability does not hurt the performance of our protocol, by simulation, as shown in Section 6.1.

In the multihop broadcast scenario, previous research shows that naive broadcast suffers from slow delivery and low reliability due to an excessive number of concurrent packet transmissions [13][17]. Some work has been proposed to mitigate this problem. Vehicular Collision Warning Communication (VCWC)[19] uses a backoff mechanism to perform congestion control by reducing the message retransmissions. In Smart Broadcast[6], nodes set their contention windows inversely proportional to the distance from the sender. Palazzi et al.[14] propose a broadcast scheme to adjust contention window based on the transmission range estimation. Interferences and mobility are taken into account by dynamically computing the nodes' transmission ranges estimation. Kutylowski and Zagorski [9] propose a dynamic opportunistic protocol to forward emergency messages along a motorway. Moreover, Wisitponphan et al.[17] propose techniques using probabilistic forwarding to mitigate the broadcast storm problem. All these approaches basically use a probabilistic mechanism to broadcast the packets. F. Stann et al.[16] uses a density-aware approach to recover missing nodes. However, their work focuses on enhancing reliability and lacks discussion on efficiency. Our protocol, instead, propagates the packets as fast as possible in the first phase, and deterministically adjusts the transmit power in the second phase. Hence, our protocol can guarantee both high reliability and efficiency.

Resta, Santi, and Simon[15] derive lower bounds on the probability that a car at distance d from the source correctly receives the message within time t. Their work aims at multihop broadcast and also considers interference. However, their metrics are not representative in the sense that they only analyze the probability for one node. During the process of propagating emergency messages, we definitely need to consider the whole coverage within the area of interest. In addition, they don't have a good strategy to broadcast packets. Their strategies are unrealistic because they are centralized. Therefore, we provide a two-phase broadcast protocol and propose new metrics to analyze our protocol.

VIII. Conclusions

We have presented ERP, an efficient and reliable protocol for emergency message dissemination in VANETs. ERP conducts a two-phase mechanism to address the broadcast storm problem. In the first phase, ERP explicitly designate forwarders to relay the message. No collision can occur since there is only one forwarder at a time. Nodes overhear the message will periodically broadcast it in the second phase. ERP adopts a density-aware power control mechanism in the second phase to achieve efficient message dissemination. We theoretically prove that the power used by ERP is optimal.

We evaluate ERP through both theoretical study and simulation. We show that the coverage latency of ERP is at most a constant factor from a theoretical lower bound. We also show that ERP outperforms probabilistic forwarding, which is currently the most widely studied solution, via simulations.

References

[1] F. Bai, T. Elbatt, G. Hollan, H. Krishnan, and V. Sadekar. Towards Characterizing and Classifying Communication-based Automotive Applications from a Wireless Networking Perspective. In *Proc. of AutoNet 2006*, December 2006.

[2] G. Bianchi. Performance analysis of the ieee 802.11 distributed coordination function. *IEEE Journal on Selected Areas in Communications*, 18(3), March 2000.

[3] Federal Communications Commission. Fcc 03-024. fcc report and order, Feburary 2004.

[4] A. Demers, D. Greene, C. Hauser, W. Irish, and J. Larson. Epidemic algorithms for replicated database maintenance. *ACM Principles of Distributed Computing*, Aug 1987.

[5] T. Elbatt, S. K Goel, G. Holland, H. Krishnan, and J. Parikh. Cooperative collision warning using dedicated short range wireless communications. In *Proc. of VANET 2006*, September 2006.

[6] E. Fasolo, R. Furiato, and A. Zanella. Smart broadcast algorithm for inter-vehicular communication. In *Proc. of WPMC'05*, September 2005.

[7] P. Gupta and P.R. Kumar. The capacity of wireless networks. *IEEE Trans. on Information Theory*, 46(2):388 – 404, March 2000.

[8] Dedicated Short Range Communications (DSRC) home. http://www.leearmstrong.com/dsrc/dsrchomeset.htm.

[9] J. Kutylowski and F. Zagorski. Reliable broadcasting without collision detection. *SOFSEM 2006: Theory and Practice of Computer Science*, 2006.

[10] L. Li, J. Halpern, V. Bahl, Y. Wang, and R. Wattenhofer. Analysis of a Cone-Based Distributed Topology Control Algorithm for Wireless Multihop Networks. In *Proc. of PODC*, August 2001.

[11] N. Li, J.C. Hou, and L. Sha. Design and analysis of a mst-based distributed topology control algorithm for wireless ad-hoc networks. *IEEE Trans. on Wireless Communications*, 4(3):1195 – 1207, May 2005.

[12] S. MANGOLD, S. CHOI, P. MAY, O. KLEIN, G. HIERTZ, and L. STIBOR. Ieee 802.11e wireless lan for quality of service. In *Proc. European Wireless*, Florence, Italy.

[13] S. Ni, Y. Tseng, Y. Chen, and J. Sheu. The Broadcast Storm Problem in a Mobile Ad Hoc Network. In *Proc. of ACM MobiCom*, 1999.

[14] C. E. Palazzi, S. Ferretti, M. Roccetti, G. Pau, and M. Gerla. How do you quickly choreograph inter-vehicular communications? a fast vehicle-to-vehicle multi-hop broadcast algorithm, explained. In *Proc. of CCNC/NIME 2007*, January 2007.

[15] G. Resta, P. Santi, and J. Simon. Analysis of Multi-Hop Emergency Message Propagation in Vehicular Ad Hoc Networks. In *Proc. of ACM MobiHoc'07*, September 2007.

[16] F. Stann, J. Heidemann, R. Shroff, and M. Murtaza. Rbp: robust broadcast propagation in wireless networks. In *Proc. of Sensys'06*, 2007.

[17] N. Wisitpongphan, O. K. Tonguz, J. S. Parikh, P. Mudalige, F. Bai, and V. Sadekar. Broadcast storm mitigation techniques in vehicular ad hoc networks. *IEEE Wireless Communication Magazine*, November 2006.

[18] Q. Xu, T. Mak, J. Ko, and R. Sengupta. Vehicle-to-vehicle safety messaging in dsrc. In *Proc. of VANET '04*, 2004.

[19] X. Yang, J. Liu, F. Zhao, and N. H. Vaidya. A vehicle-to-vehicle communication protocol for cooperative collision warning. In *Proc. of MobiQuitous'04*, Augest 2004.

[20] J. Yin, T. ElBatt, G. Yeung, B. Ryuand S. Habermas, H. Krishnan, and T. Talty. Performance evaluation of safety appliations over dsrc vehicular ad-hoc networks. In *Proc. of VANET '04*, 2004.