An Opportunistic Forwarding Protocol with Relay Acknowledgement for Vehicular Ad-hoc Networks

Inhyeok Jang, Wooyeol Choi, and Hyuk Lim

Abstract

Robust and efficient data delivery in vehicular ad hoc networks (VANETs) with a high mobility is a challenging issue due to unstable wireless links and dynamic topology changes. In this vehicular environment, the exchanges of topology information and discoveries of routing paths on an end-to-end basis are not desired, because the network topology changes rapidly and the information is out-of-date in short. In this paper, we propose a contention based forwarding mechanism, which locally exploits *geographic* location information to achieve robustness of data delivery for vehicular communications. A node that is closer to the destination node has a higher priority in competing for the wireless channel, and forwards packets as a relay node earlier than other adjacent nodes. When a packet is forwarded by a relay node, its preceding node informs adjacent nodes that it has been already relayed to prevent the duplications of packet delivery. Through ns-2 simulations, we show that the proposed scheme can significantly improve the network performance in terms of the packet delivery ratio and the end-to-end delay in various synthetic and trace-based realistic VANET scenarios with a wide range of vehicle speed and density.

Index Terms

Routing, forwarding protocol, duplication suppression, vehicular ad-hoc networks,

I. Introduction

A mobile ad hoc network (MANET) is an infrastructure-less wireless network, which was originally developed for military purpose and has received a considerable attention as a promising next generation network for realizing ubiquitous computing because of its capability of building networks without a pre-existing infrastructure in a wide range of networking environments such as battlefield communications, vehicular communications, aeronautical communications, personal communications, environmental monitoring, emergency rescue, disaster recovery, and indoor applications. Among various applicable areas for MANETs, the demands for vehicular ad-hoc networks (VANETs) for vehicle-to-vehicle and vehicle-to-infrastructure communications have dramatically increased for road safety, transport management, ubiquitous connectivity provisioning for mobile vehicles. Many governments, companies, and research communities have acknowledged the potential of wireless vehicular communications (e.g., European eSafety initiative [2] and CVIS [3], Wheels project sponsored by the German Ministry of Education and Research [4], the US programs derived by the Intelligent Vehicle Initiative [5], and the Japanese Internet ITS [6] and AHS programs [7]). In addition, US have already allocated the dedicated frequency spectrums, and European countries are working on the spectrum regulation for vehicular communications.

In VANETs, a reliable routing for vehicles with high mobility is a challenging issue because mobility causes dynamic changes of network topology, which may increase the routing overhead for maintaining topology information and degrade the routing performance in terms of the packet delivery ratio and latency [8]–[10]. As the speed of vehicles ranges from a few to over 100 Km/h, the neighbor information gathered by a routing protocol may easily become invalid in a short time in VANETs.

A considerable amount of research has been conducted on routing protocols for MANETs. Reactive routing protocols such as the dynamic source routing (DSR) and the ad-hoc on demand distance vector (AODV) routing are proposed to cope with mobility of ad-hoc networks. They require topology information to be gathered and exploit it to find paths to a mobile destination. However, if the degree of mobility becomes higher than a certain value, the performance significantly degrades because the information gathered becomes out-of-date rapidly [11]. If a node tries to send packets using invalid out-of-date routing information, route rediscovery and retransmission happen because of a routing failure while the packets are relayed towards

the destination. Frequent route rediscovery and packet retransmission can cause a significant performance degradation, resulting in the increase of the latency and network congestion.

To mitigate the overhead incurred by exchanging and managing topology information, stateless routing approaches have been extensively studied. These approaches usually do not gather an end-to-end full path information. Instead, they rely on a greedy forwarding using a local information about neighboring nodes. Geographic routing is one of the most promising stateless approaches. As a geographic location based approach, the greedy perimeter stateless routing (GPSR) [12] and the contention based forwarding scheme [13] enable a mobile node to explicitly determine its next-hop node before it transmits a packet. However, when the sender attempts to transmit a packet, the next-hop node selected in advance may move away. Note that the next-hop node is usually located near the boundary of transmission range for maximum stretch toward the destination. The beacon-less routing algorithm [14] does not require a mobile node to determine the next-hop node in advance, but restricts an area where the competition among potential relay nodes occurs for suppressing the duplication of packets. While these stateless routing approaches are suitable for VANETs with high mobility because they do not rely on static topology information, their routing performances in VANETs can be significantly affected by how to reduce duplicated forwardings and how to mitigate contention among relay candidates. (We will provide a more detailed summary of existing work in Section II.)

In this paper, we propose an opportunistic contention based forwarding that can achieve robust packet delivery performance in VANETs with high mobility. We incorporate the contention mechanism for selecting a next-hop node among competing neighbor nodes into the contention algorithm in the media access control (MAC) layer (e.g., binary exponential backoff (BEB) in the IEEE 802.11 DCF) for mitigating the collision due to concurrent transmissions. This contention mechanism enables mobile nodes to exploit wireless channel adaptively to the contention and congestion level of networks, while differentiating the priority among neighboring nodes competing for the channel with each other. In order to keep the rate of duplicated packet deliveries low, we devise a novel suppression mechanism with an explicit acknowledgement. Unlike the existing forwarding protocols, this suppression mechanism neither restricts any forwarding area nor incurs an exchange of complicated control messages. Instead, if a packet is forwarded by a relay node, the preceding node sends a relay acknowledgement to explicitly inform the other neighboring nodes that the packet has been successfully forwarded by the relay node.

The performance of the proposed forwarding has been evaluated by extensive ns-2 simulations

in a variety of random and vehicular environments. The simulation results show that the proposed forwarding achieves both high delivery ratio and low end-to-end latency in a wide range of vehicle speed and density under the synthetic and trace-based VANET scenarios.

The contributions of this paper are as follows:

- We propose a duplication suppression scheme with relay acknowledgement for reliable stateless routing in highly dynamic mobile environments.
- We propose a contention based next-hop relay selection scheme that adaptively adjusts the contention window size in order to mitigate the contention among neighboring nodes competing for packet forwarding.
- We present an analytic result for choosing a proper maximum contention window size that minimizes forwarding delays with a low collision probability when the number of contending nodes is given.

The rest of this paper is organized as follows. In Section II, we give a summary of related work in the literature. In Section III, we propose the forwarding protocol with relay acknowledgement for VANETs with a high mobility, and in Section IV, we show the analytical results for collision probability and forwarding failure probability of the proposed forwarding protocol. The simulation results of the proposed forwarding protocol follows in Section V. Finally, we conclude the paper in Section VI.

II. RELATED WORK

In recent years, stateless routing approaches for data delivery in mobile/vehicular ad hoc networks have been extensively studied, because they provide a robust data delivery performance with low control overhead in highly dynamic wireless network. Here, we classify these stateless routing approaches in two categories; *sender driven* and *receiver driven* approaches, depending on who is responsible for determining a next-hop relay node among multiple candidate neighbor nodes.

A. Sender driven approaches

The Geographic Perimeter Stateless Routing (GPSR) [12] is a greedy forwarding protocol using location information. Every node periodically exchanges geographical position information with its neighbor nodes, and maintains a list for positions of neighbor nodes. Based on the positions information, senders make a packet forwarding decision before they transmit a packet.

In GPSR, by greedy forwarding, packets are forwarded to nodes that are always progressively closer to the destination. In mobile networks with high mobility, the neighbor information should be updated more frequently because the network topology rapidly changes. This may cause the increase of routing control overhead and contention level. Moreover, if the information is out-of-date, the packet delivery performance degrades significantly.

B. Receiver driven approaches

The receiver driven approaches such as [13]–[15] do not impose the exchange of topology information before a sender node transmits data packets. Rather, a packet is broadcast without determining which node will forward the packet, and on receiving the packet, every node becomes a potential relay node. When a node receives the packet, it evaluates its priority depending on its own geographical information. The node with the highest priority is responsible for forwarding the packet as a relay node. These approaches are opportunistic because a next-hop node is not pre-determined by a sender and is selected opportunistically by contention among multiple receivers. Depending on how to prevent duplicated packet forwarding during the contention, the existing receiver driven approaches are classified into two groups: area based suppression and reservation based suppression.

1) Area based suppression: An area based suppression scheme limits candidate nodes according to the position of each node. Heissenbüttel *et al.* proposed a beacon-less routing mechanism (BLR) [14], which does not require periodic hello-messages unlike other position based routing protocols such as GPSR. Under BLR, nodes located in a certain area, called "forwarding area", are allowed to become a next-hop relay node in order to prevent unnecessary message duplicates. The forwarding area is constructed in a way that the nodes in the area can communicate with any other nodes in the same forwarding area. Therefore, if any node in a forwarding area transmits a packet, then all the other nodes can overhear the transmission of the packet, and they do not forward the packet anymore.

Ho *et al.* proposed a virtual cell based connectionless approach for mobile ad hoc networks (CLA) [15]. A network area is divided into a number of virtual cells, and then a route path, called grid path, is formed with some of virtual cells, which are selected with source and destination node's coordinates. Only the nodes in the selected grid path are responsible for forwarding a data as a relay node. In [16], the performance of CLA has been evaluated in city street scenarios of vehicular ad hoc network.

These area based suppression schemes work well and consume a small amount of bandwidth, but their performance highly depends on the node density, and they require an additional recovery strategy for wireless networks with a low node density.

2) Reservation based suppression: A reservation based suppression scheme reserves a specific link by sending control packets. Füßler *et al.* proposed a contention-based forwarding (CBF) for mobile ad hoc networks [13], which uses a suppression scheme called "active selection". This suppression scheme transmits control packets for selecting a relay node as like in the MACA scheme [17] and the IEEE 802.11 Standard [18]. In detail, a node sends RTF (Request to Forward) first, and its neighbor nodes compete with each other to reply with CTF (Clear to Forward) packet with a certain delay time. The delay time is adjusted according to how close each node is to the destination. The node that is closest to the destination among all the neighbors is selected as a relay node. Then, the data is transmitted to the selected relay node through a unicast packet. While this suppression scheme effectively prevents packet duplications, an additional constant delay is always incurred for each hop forwarding due to the RTF/CTF exchanges and the waiting time for CTF reply. In [19], the performance of CBF has been evaluated in highway street scenarios of vehicular ad hoc network.

Unlike the above suppression schemes, our proposed scheme does not restrict any forwarding area. Rather, by sending a relay acknowledgement message to all the nodes in the transmission range, it actively informs that the packets have been successfully forwarded by a relay node. Furthermore, our scheme incurs a relatively small delay for controlling the competition among multiple neighbor nodes, because we differentiate the transmission priority among competing nodes by exploiting the contention mechanism of the MAC layer in an adaptive manner to the network contention and congestion level.

C. Broadcast storm problem

Ni *et al.* presented a broadcast storm problem in [20], and insisted that the flooding without care in ad hoc networks may cause a significant degradation of network performance due to redundant rebroadcasts, severe contention, and collisions. To resolve this problem, they introduced several mechanisms, which include an additional delay before each rebroadcast of packets or to drop packets probabilistically. These mechanisms are further improved by using a counter for rebroadcasts or location information. They also introduced a cluster based approach that exploits a graph modeling. The authors concluded that the location based scheme gives the best

performance in terms of the redundant rebroadcast elimination ability and reachability. The result of this work implies that the location based suppression is the most effective for the duplication of packet delivery.

D. Routing strategies for VANETs

There have been many studies on routing and forwarding mechanisms for VANET environments. Most of them make use of geographical information such as the shapes and locations of roads and junctions provided by a map database in vehicle navigation systems. The geographic source routing (GSR) [8] is a position based routing protocol, where the path to a destination node is composed of a series of junctions a packet has to traverse. The authors showed that GSR outperforms non position based routing strategies such as DSR and AODV in a city environment with a realistic vehicle movement pattern in terms of delivery rate and latency.

Morris *et al.* suggested a scalable vehicular ad hoc network system called CarNet [21]. In [21], they used a grid routing with the grid location service (GLS) [22] for achieving scalability in a large ad hoc mobile network. The simulation results showed that the grid routing achieves high delivery ratio despite the increase of the number of nodes.

Seet *et al.* proposed the the anchor-based street and traffic aware routing (A-STAR) for metropolitan vehicular communications with uneven distribution of vehicular nodes in [23]. The authors proposed the usage of city bus route information in order to identify anchor paths with higher connectivity. The simulation results showed that A-STAR achieves better delivery ratio and reasonable end-to-end delay in comparison with GPSR and GSR under the M-Grid mobility model.

Naumov and Gross proposed the connectivity-aware routing (CAR) [9], which enhances anchor based geographic routing by adding the connected path finding capability. The CAR locates the position of a destination node and finds a connected anchor path by using a preferred group broadcast mechanism, which enables relay nodes to forward packets without route rediscovery when the moving speed and direction of a destination node are changing. The simulation results showed that CAR achieves better delivery ratio and smaller end-to-end delay in comparison with GPSR in the city and highway vehicular scenarios.

Kihl *et al.* proposed the robust vehicular routing (ROVER) [24], which is a multicast routing protocol for supporting QoS sensitive applications. The ROVER delivers messages to a geographically specified zone of relevance (ZOR) by forming a multicast tree. The authors argue

that the multicast tree built by the ROVER could be used for QoS guaranteed transport layer protocols. The simulation results showed that ROVER achieves a high delivery ratio with a low delay under various realistic highway scenarios.

III. A FORWARDING MECHANISM WITH RELAY ACKNOWLEDGEMENT

We propose an opportunistic forwarding mechanism for achieving efficient and reliable packet delivery in VANETs with high mobility. The overall procedure of the proposed forwarding mechanism is as follows:

(S1)   A *current node* i that has a packet to forward broadcasts the packet to its neighbors. Every neighboring node receiving the packet becomes a potential relay nodes.

(S2)   On receiving a packet, each node computes its priority depending on geographic or virtual location [1] . A neighboring node j with the highest priority will relay the packet earlier than the others.

(S3)   The node i broadcasts a *relay acknowledgement* to its neighboring nodes when it overhears the packet being relayed by the node j.

(S4)   On receiving the relay acknowledgement, the other neighboring nodes that are attempting to forward the packet drop the packet in their queue except the selected relay node j.

In (S2) and (S3), if a node is the destination of the packet, the highest priority is assigned to the node, and it broadcasts a relay acknowledgement in a short time to prevent undesirable packet duplications near the destination node.

The key features of the proposed mechanism are two-fold: (i) in order to make a node with a higher priority relay a packet earlier than the other nodes, we use a contention window based next-hop selection, in which a node with a higher priority has a smaller value of contention window, and (ii) we reduce the duplication of packet delivery by using a relay acknowledgement. After a node broadcasts a packet, it overhears the wireless channel to check whether the packet is forwarded by its neighboring node. If it has been forwarded, the node informs the other neighbor nodes of the relay of the packet to prevent the duplication of the packet.

---

[1]Without GPS devices, relative proximity information among nodes can be inferred with the use of hop-distance, signal strength, time of arrival, time difference of arrival, and angle of arrival. It makes it possible to construct virtual coordinates (e.g., hop-counts from anchor nodes) in ad hoc networks [25]–[28].
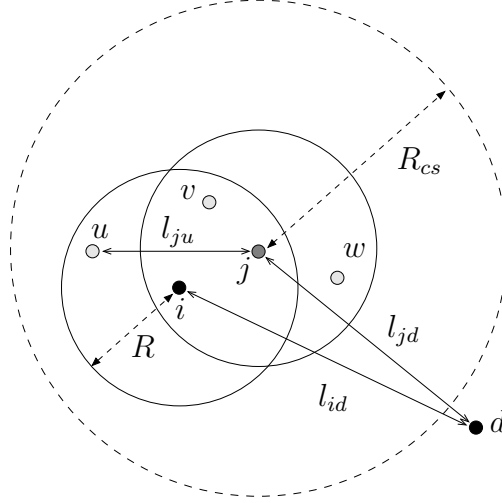
Fig. 1. A wireless network, where R and R_{cs} are the transmission range and the carrier sense range, respectively. The nodes i and d are the current node and the destination, and $l_{ju} > R$.

We illustrate the procedure of the proposed forwarding mechanism in a simple wireless network. In Fig. 1, the node i is a forwarding node that has a packet to deliver. The nodes j, v and u are neighboring nodes of the node i. The node d is a destination node. The forwarder i broadcasts the data packet destined to the destination d, and then nodes j, v and u hear its transmission and compute their initial contention window size depending on the relative distance to the destination d. Suppose that the back-off counter of each node is selected as $BC_j = 3$, $BC_v = 5$, and $BC_u = 15$. Then, the node j will forward the data packet earlier than the others. As the node v can overhear the packet relay done by the node j, it immediately cancels the pending transmission for the same packet. However, the node u cannot overhear the packet relay because it is not in the transmission range of the node j. This may cause the duplication of the packet delivery if it is not properly taken care of. In our mechanism, as soon as the forwarder node i overhears the packet relay, it sends a relay acknowledgement message to its neighboring nodes $\mathcal{N}(i) = \{j, u, v\}$. On receiving the relay acknowledgement, the neighboring nodes except the node j immediately drop the packet pending in their queue.

If a relay acknowledgement is lost (e.g., due to packet collisions), the nodes that do not hear the relay acknowledgement may relay the packet that has been already relayed, resulting in the duplication of packet delivery. However, even in this case, if multiple nodes that have the same packet in their queue hear a relay acknowledgement for the packet transmitted by one of

them, all the other nodes drop the packet from their queue and do not further relay it. This mechanism effectively reduces the duplication of packet delivery despite a certain level of relay acknowledgement losses.

A. Contention window based next-hop selection

Contention based forwarding approaches [13]–[15] broadcast packets without determining a corresponding relay node in advance. Usually, they use an additional time delay in the networking layer for electing a next-hop node among competing neighbor nodes. The delay is inversely proportional to the priority of each node.

In this paper, we exploit the contention mechanism in the MAC layer for differentiating the priority among neighboring nodes instead of inserting a time delay in the networking layer. The contention window algorithm (e.g., binary exponential backoff (BEB) in the IEEE 802.11 DCF) in the MAC layer for mitigating the collision due to concurrent transmissions is modified to be incorporated with the contention mechanism for selecting a next-hop node among competing neighbor nodes. Note that in our proposed approach, the routing layer functionality is minimized; a routing layer only manages its own location and those of its destination nodes In other words, the MAC layer of the proposed protocol takes charge of both the next-hop relay node selection and contention control among competing neighboring nodes. This cross-layer approach efficiently reduces the transmission delay and routing control overhead in an adaptive manner.

Consider the wireless networks depicted in Fig. 1. When a current node i broadcasts a packet in (S1), the next-hop relay node that is responsible to forward the packet is not selected yet. After the neighbor nodes $\mathcal{N}(i)$ receive the data packet, they compute the contention window size and compete with each other to be selected as a relay node. In this competition, the node that is closest to the destination is given the highest priority to among the neighbor nodes. For a node k, the priority is computed by its location information as follows:

$$p_k = \frac{1}{2} - \frac{\|x_i - x_d\| - \|x_k - x_d\|}{2R},$$ (1)

where x_i, x_k, and x_d are the locations of the current node i, the contending node k, and the destination node d, respectively. Note that $0 \leq p_k \leq 1$. Depending on the distance to the destination node, each node determines the priority p_k. Note that a smaller value of p_k implies a higher priority. One with the smallest value of priority is chosen as the next-hop relay node, i.e., the next-hop relay node $j = \arg\min_{k \in \mathcal{N}(i)} p_k$.

Once the next-hop relay node j with the highest priority is identified, it should be made forward the packet earlier than the other nodes in (S2). Therefore, the backoff counter BC_k is set to a value proportional to p_k as follows:

$$BC_k = \max\left(0,\ p_k \cdot CW_{max} - \text{rand}(0,\ CW_{min})\right), \tag{2}$$

where CW_{max} and CW_{min} are the maximum and minimum contention window size, respectively, and the second term is a random function that picks an integer value between 0 and $(CW_{min}-1)$. This random function is inserted to avoid the transmission collision even in a case where neighbor nodes unexpectedly have the same p_k and CW_k.

The maximum contention window is adaptively adjusted depending on the result of transmission attempt as follows:

$$CW_{max}[t] = \begin{cases} \max(CW_{max}[t-1]\ /\ 2,\ CW_{low}) & \text{for tx success} \\ \min(CW_{max}[t-1] \cdot 2,\ CW_{high}) & \text{otherwise,} \end{cases} \tag{3}$$

where CW_{high} and CW_{low} are predetermined high and low bound for CW, respectively. In (S3), as the current node i is supposed to overhear the packet being relayed by the next-hop relay node, it knows whether the packet is successfully relayed or not. If the transmission fails, the node believes that the failure is due to the collision among multiple relay candidates and doubles the contention window CW. As like the BEB in the IEEE 802.11 DCF, this dynamic adaptation of CW enables the nodes to efficiently forward packets regardless of the node density and traffic load in the wireless network.

B. Explicit acknowledgement for suppression of duplicated forwarding

In a contention based forwarding mechanism, how to prevent the neighboring nodes that are not selected as a relay node from forwarding the packets that have been forwarded by the relay node is one of the most important issue. Whenever packets are forwarded by a relay node, it should be notified to all the other nodes to suppress unnecessary duplication of the packets. If this suppression is not performed effectively, the packets may flood to the whole network, resulting in a significant performance degradation due to redundant rebroadcasts, heavy contentions, and collisions. For example, once a node with the highest priority forwards the packet in Fig. 1, all the other nodes should not re-broadcast the packet in order to prevent the network from being congested with duplicated packet forwarding. However, the node u may re-broadcast the packet because it cannot see whether the packet has been already forwarded or not. Note that l_{ju}, the
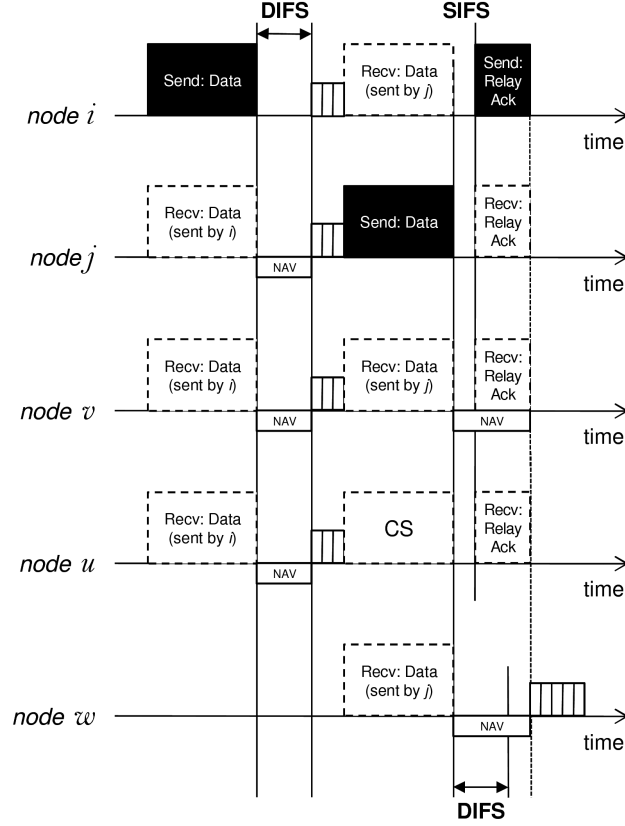
Fig. 2. Timing diagram of the proposed suppression mechanism for the wireless network depicted in Fig. 1. The propagation delays are ignored.

distance between the node j and u, is larger than R, i.e., the node u is out of the transmission range of the node j.

Here, we propose an active suppression mechanism that explicitly notifies the success of packet relay to all the neighboring nodes to prevent duplicated forwarding. This suppression mechanism lets the preceding node (the node i in Fig. 1) broadcast a control packet (called relay acknowledgement) immediately without back-off as soon as the packet is forwarded by the node j. Note that the node i can overhear the packet being relayed by the node j, and all the neighboring nodes of the node i can receive the relay acknowledgement sent by the node i. While the node j forwards the packet and i broadcasts a relay acknowledgement, the other neighboring nodes $k \in \mathcal{N}(i) \backslash \{j\}$ should be forbidden not to interrupt this transaction. Otherwise, the packet may be duplicated by some of the other neighbor nodes.

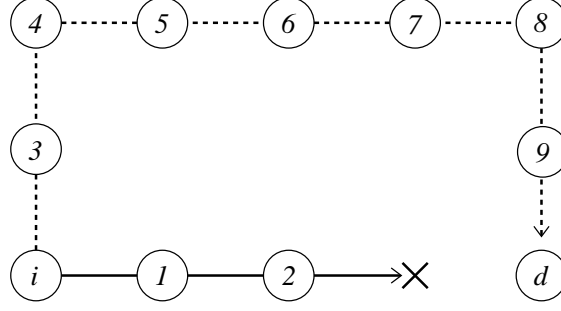To protect this transaction between the current node and the next-hop relay node, we exploit

Fig. 3. An example of the dead end problem.

the physical carrier sensing and inter-frame spacing (IFS). In our protocol, each node is allowed to start the contention to be selected as a relay node only when the wireless channel has been idle for longer than a distributed inter-frame spacing (DIFS) time interval, whereas relay acknowledgements have to be broadcast in a short inter-frame spacing (SIFS) time interval. Note that SIFS is shorter than DIFS.

Fig. 2 illustrates this suppression operation. While the current node i is transmitting a data packet, all the other nodes do not transmit any packets because they are within the carrier sense range of the node i. If the channel is idle for DIFS, the nodes j, v, u, and w start to compete with each other, and eventually the node j start to forward the packet as a relay node. Suppose that at this instance the node u wants to transmit a packet. However, the node u cannot start its transmission until the node i broadcasts the relay acknowledgement, because the node u should wait for at least DIFS after the end of the node j's transmission. Note that the node i has to wait only for SIFS before it start to broadcast the relay acknowledgement.

C. Local recovery for the dead end problem

The *dead end* problem [29] is an important issue in greedy forwarding protocols because they do not gather an end-to-end full path information. For example, if there is no other node that is closer to the destination than the current node, the packets on the current node cannot be forwarded further and are dropped at the node. This dead end problem may happen more frequently in a case where the forwarding area is restricted to a certain area or direction (e.g., beaconless routing (BLR) [14]).

Fig. 3 illustrates the dead end problem. After the node i broadcasts a packet, the node 1 becomes a relay node for the packet because the node 1 is the closest to the destination node

d. However, the packets cannot reach the destination node d along the path because there is no way to forward the packets towards the destination at the node 2 by using a greedy forwarding. In this case, the packet should have been forwarded along an alternative path of the nodes $3 \rightarrow 9$.

To mitigate the dead end problem, we let a node perform a local recovery if it recognizes that one of its neighbors has attempted to forward a packet more than a retransmission threshold. In Fig. 3, when the node 1 observes that the retransmission number of its neighbor node 2 reaches the retransmission threshold, it re-broadcast the packet with a *recovery flag* in the packet increased by 1. Then, the neighbor nodes that have canceled the forwarding before begin to contend for forwarding again. As a result, the node i is selected as the next-hop node in Fig. 3, and the packet is eventually forwarded along the path of the nodes $3 \rightarrow 9$. The nodes along the failed path for the packet with a lower recovery flag are not allowed to be selected as a relay node.

This local recovery algorithm does not resolve the dead end problem perfectly. However, it is quite useful in wireless networks with a low node density, and gives a better performance in terms of the packet delivery ratio. On the other hand, it may cause a long end-to-end delay in a specific topology. Therefore, we limit the maximum value of the recovery flag by 1. This is why it is called the recovery flag rather than the recovery number.

IV. ANALYTICAL RESULTS

In this section, we provide analytical results for the collision probability of the first replying candidate node and the forwarding failure probability of messages. Because forwardings in the proposed contention based forwarding scheme are taken place by a node with the highest priority, the analysis for the collision probability of the first replying node is more important than that for the average collision probability of all contending nodes. Watteyne *et al.* studied how to reduce the collision probability in backoff-based election mechanisms in [30]. Here, we take a similar approach to derive the collision probability of the first replying candidate node for our proposed scheme. Then, we further develop the forwarding failure probability that a packet eventually fails to be relayed due to successive collisions among all the candidate nodes in order to find an optimal value of the maximum contention window size when the number of neighboring nodes is given.
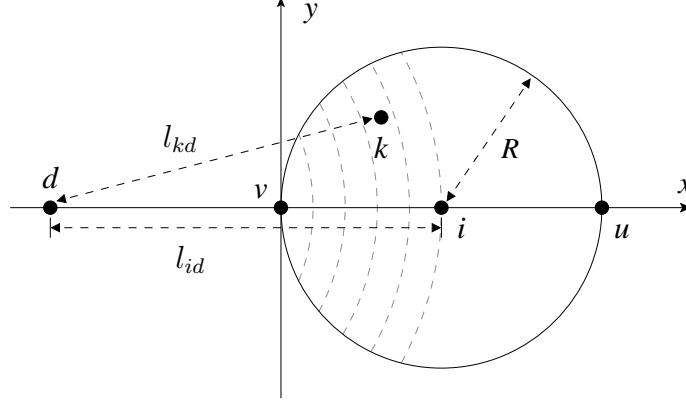
Fig. 4. A selection of relaying node for the current node i and the destination node d.

A. Collision probability for the node with the smallest backoff counter

First, we derive the probability density function (pdf) for the backoff counter of relay nodes. Consider a relaying scenario in which the current node is i, the destination node d, and the nodes k, u, and v are the candidate nodes as shown in Fig. 4. We assume that candidate nodes are uniformly distributed in a circle centered at the node i. Then, the transmission area is expressed by

$$(x - R)^2 + y^2 \leq R^2. \tag{4}$$

Note that in this case, the backoff counters for the nodes v and u become 0 and CW_{max} by (2), respectively, if we ignore the random variation in (2).

In order to make the derivation of pdf for a backoff counter simple, we make an assumption that the nodes on a same vertical line have the same backoff counter. Note that this assumption is reasonable if the destination is sufficiently far from the current node. Under the assumption, a node at $x = x_i$ has the backoff counter of $BC_i = x_i \; CW_{max}/(2R)$. The pdf of backoff counter can be easily obtained as follows:

$$f_{BC_i}(bc) = \frac{8}{\pi CW_{max}^2}\sqrt{-bc^2 + bc \; CW_{max}} \quad \text{for} \quad 0 \leq bc \leq CW_{max}. \tag{5}$$

By using (5), the pdf of the smallest backoff counter (i.e., $BC_{first} = \min(BC_1, \cdots, BC_N)$) is given by

$$f_{BC_{first}}(bc) = N f_{BC_i}(bc) \left(\int_{bc}^{CW_{max}} f_{BC_i}(x)dx \right)^{(N-1)}, \tag{6}$$

(a) Probability density function of the backoff counter
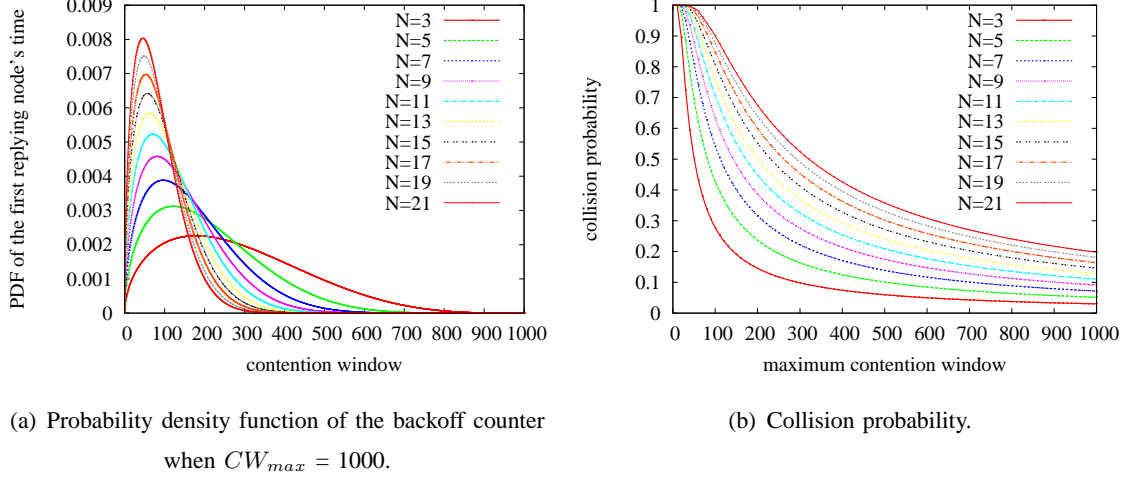when $CW_{max} = 1000$.

(b) Collision probability.

Fig. 5. The pdf of the backoff counter and the collision probability for the first replying node when $N = 3, \cdots, 21$

where N is the number of candidate nodes within the transmission range in (4). Because the node with the smallest backoff counter replies first among N candidate nodes, the above pdf corresponds to that of the first replying node's response delay. Finally, the collision probability for the node with the smallest backoff counter is obtained by

$$
P_{BC_{first}}(CW_{max}, N) = \int_0^{CW_{max}-d} (1 - \left(\frac{CW_{max} - x - d}{CW_{max} - x}\right)^{(N-1)}) \times f_{BC_{first}}(x)dx
$$
$$
+ \int_{CW_{max}-d}^{CW_{max}} f_{BC_{first}}(x)dx,
$$

(7)

where d is the transmission time of a packet in slots. (For more details about its derivation, refer to [30].)

We compute the pdf of the backoff counter for the first replying node by (6) when $N = 3, \cdots, 21$ and $CW_{max} = 1000$, and plot it in Fig. 5(a). We observe that the first replying node has the smaller backoff counter as the number of candidate nodes increases. In (7), the collision probability for the first replying node is given by a function of CW_{max} and N. Fig. 5(b) shows the collision probability decreases as CW_{max} increases and N decreases.

B. Selection of CW_{max}

Based on the previous observation, one may choose a large CW_{max} to reduce the collision of the first replying node. However, it incurs a long delay for each packet forwarding and results in inefficient wireless channel usage. Therefore, we attempt to find a value of CW_{max} that gives

(a) Forwarding failure probability
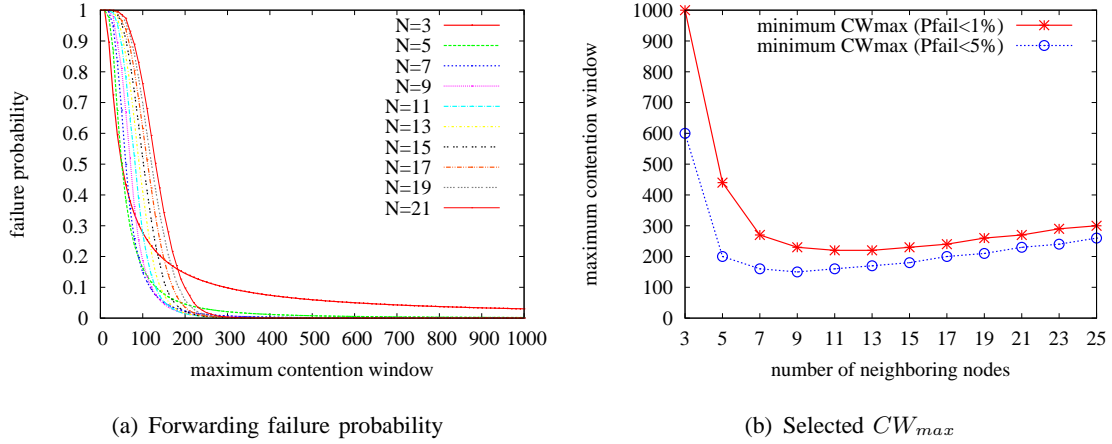
(b) Selected CW_{max}

Fig. 6. Forwarding failure probability and the optimal CW_{max} satisfying 1% and 5% requirement for forwarding failure probability.

both reasonably low collision probability and small transmission delay when the number of candidate nodes is given.

Under the proposed protocol, an opportunistic forwarding may happen when a node that has a higher forwarding priority fails to relay a packet. For example, if the first replying message collides with the second replying message, then the third node may opportunistically take the chance to relay. Despite such an opportunistic forwarding, a message eventually may fail to be relayed due to the successive collisions among all the candidate nodes. For this case, we define a "forwarding failure probability" as the probability that a message is not forwarded further by any relay nodes. Note that the forwarding failure probability can be obtained by applying (7) to the successive collisions. Then, the CW_{max} is determined by the smallest value of CW_{max} satisfying a certain failure probability.

Fig. 6(a) shows the forwarding failure probability with respect to CW_{max}. When $N \geq 5$, the forwarding failure probability rapidly decreases and becomes sufficiently small as CW_{max} increases. Fig. 6(b) shows the smallest CW_{max} value satisfying 1% and 5% requirement for forwarding failure probability. For the 5% requirement, when the number of neighboring nodes is less than 9, the selected CW_{max} value decreases as the number of candidate nodes increases. This is due to that the node density is too low to take advantage of the opportunistic forwarding. Note that in this case the forwarding failure probability is relatively high. On the other hand, the selected CW_{max} increases as N increases when N is larger than 9 because the collision

probability increases as N increases.

V. SIMULATION RESULTS

To evaluate the proposed forwarding protocol, we have performed extensive simulations using the network simulator ns-2 [31] in two scenarios: a synthetic random scenario and a vehicular scenario.

- **Random scenario**: The random network is constructed in a 800 m x 300 m area with 70 mobile nodes, which are initially located at a uniformly distributed random location and then follow the random waypoint model [32] with zero pause time. The default value for vehicle speed is 20 m/s. In this network, the average number of neighbor nodes within the transmission range of a node is approximately 7. For this scenario, 25 source–destination pairs are established and the simulation runtime is 100 s. The transmission range is 100 m, and the carrier sense range is 220 m.

- **Vehicular scenario**: For vehicular environments, we use the Generic Mobility Simulation Framework (GMSF) [33], [34] and generate the geographic information system (GIS) based mobility model traces for rural, urban, and city environments, where the simulation areas are 3000 m $\times$ 3000 m for each environment, and the numbers of vehicles are 100, 300, and 500, respectively. The speed of vehicles varies from 30 to 120 km/h and the simulation runtime is 1000 s. These mobility models are based on highly detailed road maps from a GIS and realistic microscopic behaviors in consideration of the car-following and traffic lights management. A detailed explanation of GMSF and GIS-based mobility model is found in [34].

The default values of the parameters used in the simulations are listed in Table I. All the simulations have been performed 10 times for each set of parameters and topologies.

We compared the performance of the proposed forwarding protocol with that of AODV [35] and GPSR[2] [12] in terms of the following performance metrics:

- **Packet delivery ratio**: The ratio of successfully delivered data packets to the total number of data packets sent by sources. Duplicated packets are not counted in the number of packet deliveries.

- **Average path length**: Average hop-count of successfully delivered data packets.

---

[2]The GPSR code is obtained from [36], and the beacon interval for the GPSR protocol is set to 0.5 s.

TABLE I

DEFAULT PARAMETERS USED IN THE NS-2 SIMULATIONS

Paramenter	Random scenario	Vehicular scenario
Slot time	20 μs	
SIFS	10 μs	
DIFS	50 μs	
TTL	31	
Data rate	11 Mb/s	
Basic rate	2 Mb/s	
Packet size	1000 bytes	
Traffic type	CBR / UDP	
Simulation time	100 s	1000 s
Tx range	100 m	250 m
CS range	220 m	550 m
Region	800 m $\times$ 300 m	3000 m $\times$ 3000 m

- **Average hop latency**: The per-hop delay of successfully delivered data packets. It is obtained by dividing the time required to send a packet from a source to its destination divided by the average path length.

- **Average control overhead**: The average number of control packets transmitted for one successful data packet delivery.

- **Packet duplication ratio**: The ratio of totally received data packets (including duplicated packets) at the destination to the successfully delivered data packets (excluding duplicated packets). If it is zero, there is no packet duplication; if it is one, one more same packet is delivered to the destination per one successfully delivered data packet.

- **Number of per-hop data transmissions**: The average number of data transmissions divided by the path length for successfully delivered data packets. If the value is one, there are no unnecessary data transmissions. If it is larger than one, nodes perform more than one transmission per each hop.

A. Simulation results in the random scenario

We compare the performance of the proposed forwarding protocol, AODV, and GPSR with respect to the vehicle speed and the node density. The speed of the vehicles varies from from 0

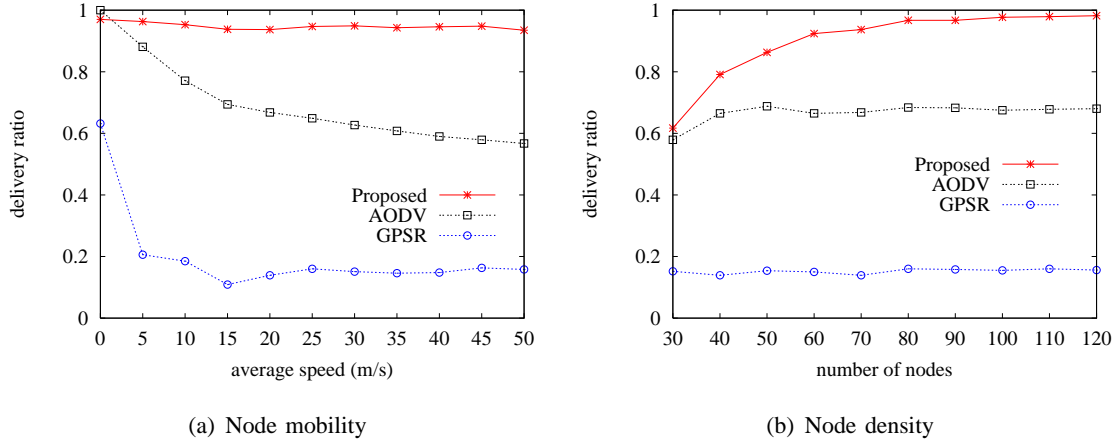(a) Node mobility   (b) Node density

Fig. 7.   Delivery ratio with respect to the node speed and the node density in the random topology.

to 50 m/s, and the node density varies from 30 to 120. Unless specified differently, the default values aforementioned are used in the simulations (i.e. 70 nodes with a speed of 20 m/s).

1) Packet delivery ratio: Fig. 7(a) shows the average packet delivery ratio with respect to the node speed. The proposed forwarding protocol maintains a high delivery ratio regardless of node mobility. Under the AODV protocol, the delivery ratio is 1 in the static case where the speed of vehicles is zero, but drops gradually as the speed of vehicles increases. The reason is that the established paths by the AODV protocols are easily broken due to the mobility of vehicles. The delivery ratio of GPSR drops rapidly as the mobility of nodes increases because the neighbor information is not properly updated on time. The beacon interval should be set small for highly dynamics networks, but this causes significant increases of routing control overhead. As a result, packets cannot reach their destinations, and trigger a number of routing rediscoveries and packet retransmissions. The average packet delivery ratio with respect to the node density is depicted in Fig. 7(b). We observe that the delivery ratio of the AODV protocol is not significantly influenced by the node density. The AODV protocol gives almost a constant delivery ratio at around 60 %, when the number of nodes is larger than 40. The GPSR protocol gives a quite low delivery ratio of about 20 % in this mobile case where the nodes move at a speed of 20 m/s. To the contrary, the proposed protocol achieves the highest delivery ratio, and its value is above 90 % when the number of nodes is 60 or more.

2) Average path length: The average path length reflects how many hops a successfully delivered packet has been forwarded through. A small value of this metric implies that routing
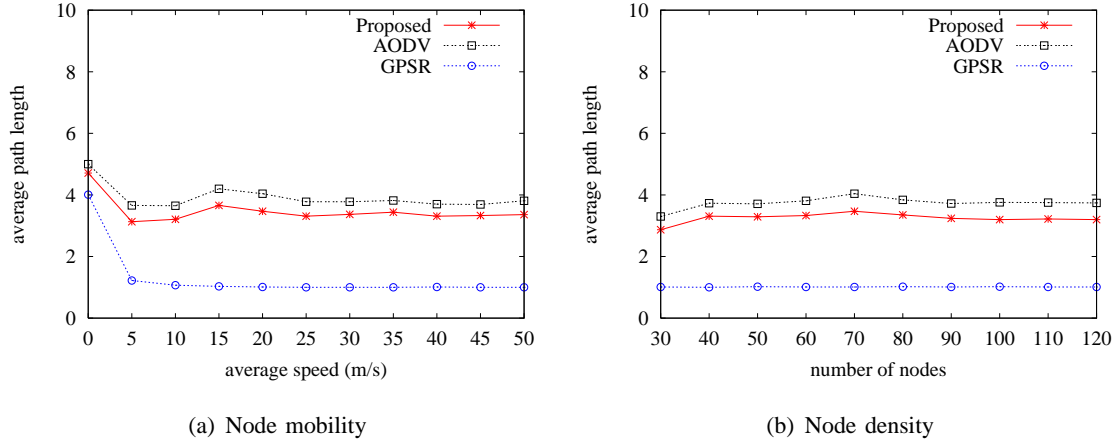
(a) Node mobility

(b) Node density

Fig. 8. Average path length with respect to the node speed and the node density in the random topology.
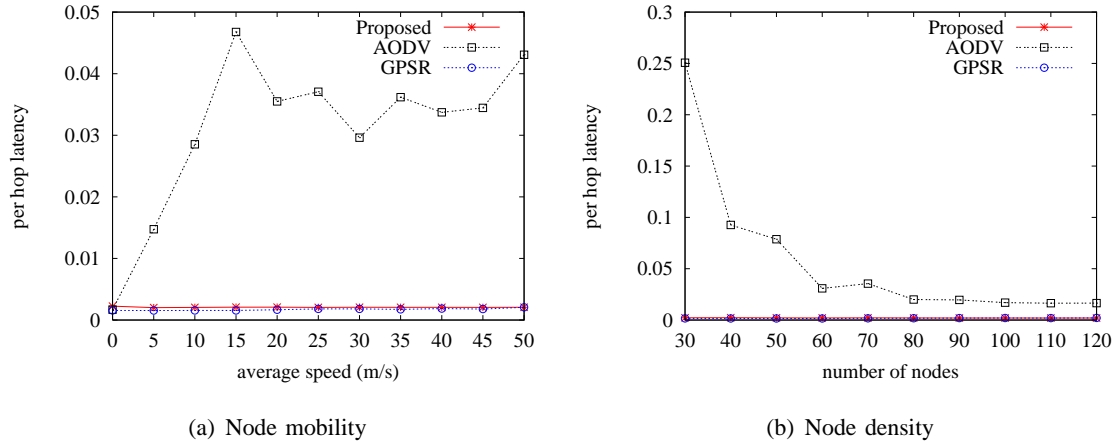

(a) Node mobility

(b) Node density

Fig. 9. Per hop latency with respect to the node speed and the node density in the random topology.

paths are efficiently set up toward their destinations. We observe that the proposed protocol finds the shorter paths in average than the AODV in all the cases of Fig. 8. This results implies that the proposed protocol adapts well to dynamically changing topology environments and maintains an efficient routing capability in a wide range of the node speed and density. It seems that the GPSR gives quite a small value of average path length. However, it is why only a small number of packets traveling a short multihop path are successfully delivered to their destinations.

3) Average hop latency: The average hop latency represents the time required to send a packet from a source to its destination divided by the average path length. Fig. 9 shows the average per hop latency for three protocols. The proposed forwarding protocol maintains quite a low hop delay, of which the value is close to that of the GPSR protocol. The reason of the
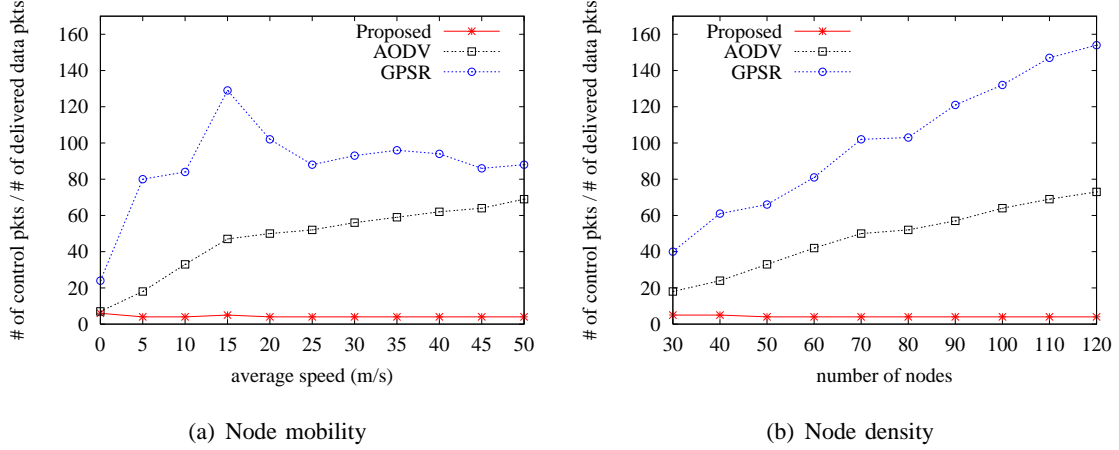
(a) Node mobility   (b) Node density

Fig. 10. Control overhead with respect to the node speed and the node density in the random topology.

small latency for the GPSR protocol is that the hop latency is averaged over the number of successfully delivered packets, which is quite small for the GPSR as shown in Fig. 7. For the AODV, the average hop latency rapidly increases as the movement of nodes increases, because of the congestion and contention caused by the retransmissions and route rediscoveries.

4) Average control overhead: Fig. 10 shows the average control overhead, which is the average number of control packets for one successfully delivered data packet. We observe that the control overhead for the proposed protocol is kept consistently small and those for AODV and GPSR increase with respect to the node density and mobility. Under the GPSR protocol, each node exchanges periodic beacon messages with its neighboring nodes in order to maintain its neighboring node list, with which a node can determine a next hop node by selecting the closest node to a destination. If a set of neighboring nodes of a node does not change, the message exchanges are just wasteful. On the other hand, if a neighboring node list is not frequently updated in VANETs with high mobility, the list becomes out-of-date, and the node may attempt to relay packets via unreachable nodes in its list. Under the proposed protocol, however, nodes do not maintain a neighboring node list and exploit the broadcast property of wireless links. Note that the nodes that actually are within the transmission range of a node are given an opportunity to become a relay node. Furthermore, the proposed protocol needs only one control packet (i.e., relay acknowledgment) for each packet relay without beacon message exchanges including geographical information.
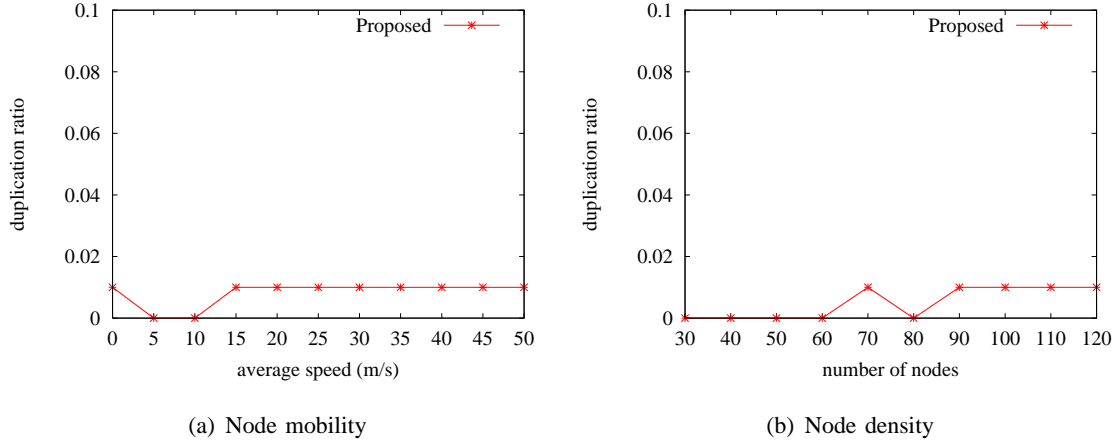
(a) Node mobility

(b) Node density

Fig. 11. Duplication ratio with respect to the node speed and the node density in the random topology.
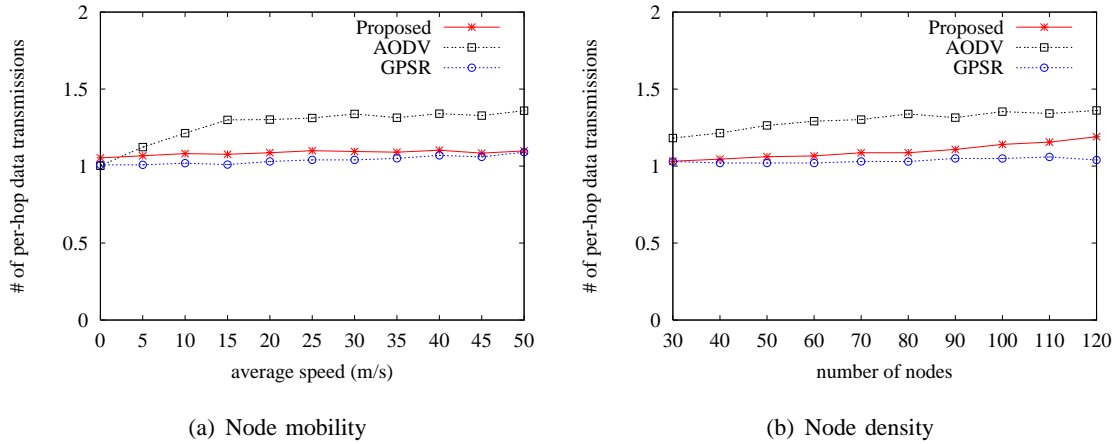

(a) Node mobility

(b) Node density

Fig. 12. Number of per-hop data transmissions with respect to the node speed and the node density in the random topology.

5) Packet duplication ratio: The packet duplication ratio reflects how many duplicated packets are delivered to the destination. Note that for the the AODV and GPSR protocols the duplication of packet forwarding does not happen because they forward the packet through unicast rather than broadcasting. Under the proposed protocol, incomplete suppressions may happen as the vehicles move at a high speed. Some nodes that have received a broadcast data packet may not hear the corresponding relay acknowledgement, if the node moves out from the transmission range of the preceding node. In this case, the packet is duplicated and starts to travel towards the destination along multiple paths. Nevertheless, the results in Fig. 11 show that the duplication rate for the proposed forwarding is kept small under 1 % in all the cases.
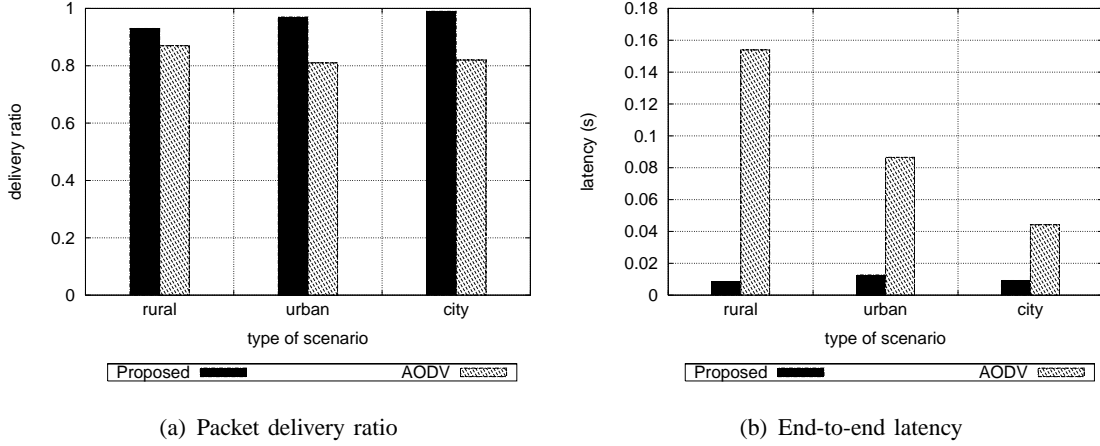
(a) Packet delivery ratio   (b) End-to-end latency

Fig. 13.   Simulation results in the rural, urban, and city vehicular scenarios.

6) Number of per-hop data transmissions: The number of per-hop data transmissions represents how many undesired data packet transmission occurs during the packet forwarding to the destination. If this value is larger than one, the nodes transmit more than one copy of the packet per each hop. This is due to the retransmission of data packets or the duplication of packet delivery. The desirable number of the data transmission is the same as the average path length. The AODV protocol shows larger number of per-hop data transmissions as the node speed and density increase, because it has to perform more route discoveries as the network topology changes faster and the number of nodes increases. For the proposed protocol, the number of per-hop data transmissions is quite close to 1, which implies that it successfully finds the good path to the destination with the packets flooding properly controlled. With this set of simulations, we confirm that the proposed contention and suppression scheme can efficiently reduce unnecessary re-broadcasts.

B. Simulation results in the vehicular scenario

Fig. 13 shows the end-to-end packet delivery ratio and the end-to-end latency for the proposed mechanism and the AODV protocol in three VANETs scenarios. As shown in Fig. 13, the proposed mechanism achieves the larger packet delivery ratio and the smaller end-to-end latency than AODV in all the scenarios. Under the AODV protocol, the source nodes have to frequently re-discover the paths that have been broken due to the mobility of vehicles, and meanwhile a large amount of packets are dropped, resulting in the performance degradation. To the contrary, the

proposed scheme does not require paths to be established in advance before the source nodes begin to transmit packets. In addition, we observe that the the proposed suppression scheme successfully maintains a low rate of packet duplication around 1–2 % in most cases. As a result, it achieves the better performance in terms of the delivery ratio and the end-to-end latency.

VI. Conclusion

We proposed an opportunistic forwarding with a relay acknowledgement in order to achieve robust packet delivery performance in vehicular ad hoc networks (VANETs) with high speed mobile vehicles. The proposed forwarding is based on a contention mechanism that enables mobile nodes to exploit wireless channel adaptively to the contention and congestion level of networks, while differentiating the priority among those competing for the channel with each other. In addition, the proposed suppression mechanism effectively reduces the duplication of packets by using a relay acknowledgement, which is sent by a preceding node to inform the other neighbor nodes of the packet forwarding. This suppression does not incur an exchange of complicated control messages, resulting in the reduction of routing control overhead. The performance of the proposed forwarding has been evaluated by extensive ns-2 simulations in a variety of random and vehicular environments. The simulation results showed that the proposed forwarding achieves both high delivery ratio and low end-to-end latency in a wide range of vehicle speed and density under the synthetic and trace-based VANET scenarios.

References

[1] I. Jang, W. Choi, and H. Lim, "A forwarding protocol with relay acknowledgement for vehicular ad-hoc networks," in *IEEE LCN*, Oct. 2008, pp. 553–554.

[2] "eSafety Initiative," http://ec.europa.eu/information_society/activities/esafety/index_en.htm.

[3] "CVIS: Cooperative Vehicle-Infrastructure Systems," http://www.cvisproject.org/.

[4] "NOW: Network on Wheels ," http://www.network-on-wheels.de/.

[5] "Intelligent Vehicle Initiative," http://www.its.dot.gov/ivi/ivi.htm.

[6] "Internet ITS Consortium," http://www.internetits.org/.

[7] "Advanced Cruise-Assist Highway Systems," http://www..ahsra.or.jp/.

[8] C. Lochert, H. Hartenstein, J. Tian, H. Fuessler, D. Hermann, and M. Mauve, "A Routing Strategy for Vehicular Ad Hoc Networks in City Environments," in *IEEE Intelligent Vehicles Symposium*, 2003, pp. 156–161.

[9] V. Naumov and T. Gross, "Connectivity-Aware Routing (CAR) in Vehicular Ad-hoc Networks," in *IEEE INFOCOM*, May 2007, pp. 1919–1927.

[10] Q. Yang, A. Lim, S. Li, J. Fang, and P. Agrawal, "ACAR: Adaptive Connectivity Aware Routing Protocol for Vehicular Ad Hoc Networks," in *IEEE ICCCN*, Aug. 2008, pp. 1–6.

[11] V. Naumov, R. Baumann, and T. Gross, "An Evaluation of Inter-Vehicle Ad Hoc Networks Based on Realistic Vehicular Traces," in *ACM MobiHoc*, 2006, pp. 108–119.

[12] B. Karp and H. Kung, "GPSR: Greedy Perimeter Stateless Routing for Wireless Sensor Networks," in *ACM MobiCom*, 2000.

[13] H. Füßler, J. Widmer, M. Käsemann, M. Mauve, and H. Hartenstein, "Contention-Based Forwarding for Mobile Ad Hoc Networks," in *Ad Hoc Networks*, vol. 1, no. 4. Elsevier, 2003, pp. 351–369.

[14] M. Heissenbüttel, T. Braun, T. Bernoulli, and M. Wälchli, "BLR: Beacon-Less Routing Algorithm for Mobile Ad Hoc Networks," in *Computer Communications*, vol. 27, no. 11. Elsevier, 2004, pp. 1076–1086.

[15] Y. Ho, A. Ho, K. Hua, and G. Hamza-Lup, "A Connectionless Approach to Mobile Ad Hoc Networks," in *IEEE ISCC*, 2004.

[16] A. Ho, Y. Ho, and K. Hua, "A Connectionless Approach to Mobile Ad Hoc Networks in Street Environments," in *IEEE Intelligent Vehicles Symposium*, June 2005, pp. 575–582.

[17] P. Karn, "MACA-a New Channel Access Method for Packet Radio," in *Computer Networking Conference, ARRL/CRRL Amateur Radio*, vol. 140, 1990.

[18] ANSI/IEEE, "IEEE Standard 802.11," 1999.

[19] H. Füßler, H. Hartenstein, M. Mauve, W. Effelsberg, and J. Widmer, "Contention-Based Forwarding for Street Scenarios," in *International Workshop on Intelligent Transportation, WIT*, 2004.

[20] S. Ni, Y. Tseng, Y. Chen, and J. Sheu, "The Broadcast Storm Problem in a Mobile Ad Hoc Network," in *ACM/IEEE MobiCom*, 1999, pp. 151–162.

[21] R. Morris, J. Jannotti, F. Kaashoek, J. Li, and D. Decouto, "CarNet: A Scalable Ad Hoc Wireless Network System," in *ACM SIGOPS European Workshop*, 2000, pp. 61–65.

[22] J. Li, J. Jannotti, D. S. J. De Couto, D. R. Karger, and R. Morris, "A Scalable Location Service for Geographic Ad Hoc Routing," in *ACM MobiCom*, 2000, pp. 120–130.

[23] B.-C. Seet, G. Liu, B.-S. Lee, C.-H. Foh, K.-J. Wong, and K.-K. Lee, "A-STAR: A Mobile Ad Hoc Routing Strategy for Metropolis Vehicular Communications," in *Lecture Notes in Computer Science : NETWORKING*, 2004, pp. 989–999.

[24] M. Kihl, M. Sichitiu, T. Ekeroth, and M. Rozenberg, "Reliable Geographical Multicast Routing in Vehicular Ad-Hoc Networks," in *Wired/Wireless Internet Communications*, 2007.

[25] M. Rossi, M. Zorzi, and R. Rao, "Cost Efficient Routing Strategies over Virtual Coordinates for Wireless Sensor Networks," in *IEEE GLOBECOM*, vol. 5, Dec. 2005, pp. 6 pp.–2980.

[26] A. Caruso, S. Chessa, S. De, and A. Urpi, "GPS Free Coordinate Assignment and Routing in Wireless Sensor Networks," in *IEEE INFOCOM*, March 2005, pp. 150–160 vol. 1.

[27] B. Leong, B. Liskov, and R. Morris, "Greedy Virtual Coordinates for Geographic Routing," in *IEEE ICNP*, Oct. 2007, pp. 71–80.

[28] M.-J. Tsai, H.-Y. Yang, B.-H. Liu, and W.-Q. Huang, "Virtual-Coordinate-Based Delivery-Guaranteed Routing Protocol in Wireless Sensor Networks," in *IEEE/ACM Transactions on Networking*, vol. 17, no. 4, 2009, pp. 1228–1241.

[29] L. Zou, M. Lu, and Z. Xiong, "A Distributed Algorithm for the Dead End Problem of Location Based Routing in Sensor Networks," in *IEEE Transactions on Vehiclular Technology*, vol. 54, no. 4, 2005, pp. 1509–1522.

[30] T. Watteyne, I. Auge-Blum, M. Dohler, and D. Barthel, "Reducing Collision Probability in Wireless Sensor Network Backoff-Based Election Mechanisms," in *IEEE GLOBECOM*, Nov. 2007, pp. 673–677.

[31] "The Network Simulator - ns-2," http://www.isi.edu/nsnam/ns/.

[32] D. Johnson and D. Maltz, "Dynamic Source Routing in Ad Hoc Wireless Networks," in *Mobile Computing*, vol. 353, no. 153-181. Springer, 1996, p. 152.

[33] "Generic Mobility Simulation Framework," http://gmsf.sourceforge.net/.

[34] R. Baumann, F. Legendre, and P. Sommer, "Generic Mobility Simulation Framework (GMSF)," in *ACM MobilityModels*, 2008, pp. 49–56.

[35] C. Perkins and E. Royer, "Ad-hoc On-demand Distance Vector Routing," in *IEEE Mobile Computing Systems and Applications*, vol. 2, Feb. 1999, pp. 90–100.

[36] http://www.cs.binghamton.edu/~kliu/research/ns2code/, 2005.