A Novel Overlay Token Ring Protocol for Inter-Vehicle Communication

Jingqiu Zhang, Kuang-Hao Liu, and Xuemin (Sherman) Shen
Dept. of Elec. & Comp. Eng., University of Waterloo, Waterloo, ON Canada, N2L 3G1
Email: {j47zhang, k8liu, xshen}@brcr.uwaterloo.ca

Abstract—Reliable and fast message dissemination for safety application and quality-of-service (QoS) guarantee for data service are considered to be the most demanding requirements in vehicle ad-hoc networks. Current medium access control (MAC) protocols are insufficient to meet both requirements, while the standardization process is still ongoing. In this paper, an overlay token ring protocol (OTRP) is proposed for inter-vehicle communication (IVC). In the OTRP, the vehicular network is considered as overlapped virtual rings, each of which has a token passed in the ring as the right for transmission. The ring structure is dynamically adjusted according to the movements of vehicles. A salient feature of the OTRP is that two operation modes, namely the normal and emergency modes, are devised, whereby timely emergency message dissemination is guaranteed and desired quality-of-service (QoS) for data service can be provided. Theoretical analysis and simulations under saturated traffic condition are conducted. The results show that the OTRP can meet the stringent requirements of vehicle communications with fast and reliable emergency message broadcasting.

I. INTRODUCTION

With the increasing number of vehicles and more time spent in vehicles, on-wheel mobile services have attracted great attentions both in academic and industry. Emerging vehicular applications include safety indication and warning during driving [1], coordination and information exchange among vehicles on the road, and entertainment and Internet applications such as gaming and booking hotels and restaurants. Vehicular communications can be classified to inter-vehicle communication (IVC) [2], and roadside-to-vehicle communication (RVC) [3]. Since the vehicular networks extend to thousands of miles with traffic merging and splitting, building enough infrastructures to cover all the road area is impractical. The infrastructureless ad-hoc network is thus recommended as the de facto mode for IVC, known as vehicular ad-hoc networks (VANETs) [4]. Different from traditional wireless ad-hoc networks, the frequent change of network topology in VANETs due to high mobility and fast speed of vehicles challenges the protocol design.

For a fixed physical layer, VANETs require the medium access control (MAC) layer adapting to the volatility of vehicular networks. Several MAC protocols have been proposed for VANETs. In the United States, the dedicated short range communication (DSRC) technology [5] uses IEEE 802.11a as its MAC standard and reserves the 5.9 GHz frequency band for safety applications in vehicular networks. Since 2006, the IEEE 802.11 working group has been developing a new standard, namely the IEEE 802.11p, for vehicular communication networks [6]. IEEE 802.11p employs the same core mechanism as IEEE 802.11a, i.e., carrier-sense multiple access with collision avoidance (CSMA/CA) but takes into account the safety application requirements of both IVC and RVC. It inherits the frequency band and multi-channel structure of DSRC with one control channel for safety and control messages, and six service channels for data traffic. The Enhanced Distributed Control Access (EDCA) defined in IEEE 802.11e is also applied in 802.11p to provide prioritized services. In Europe, a time division multiple access (TDMA) based protocol named reliable reserved-ALOHA (RR-ALOHA) [7] is proposed for VANETs to provide reliable broadcast to vehicles on road. Besides these standardization activities, extensive research efforts on developing MAC protocols for IVC can also be found in the literature. Directional-antenna based MAC is deemed to be an effective way in reducing collisions [8]. The idea of token-ring originated from the wired networks have been brought into wireless domain in several recent works, in virtue of its merit in providing the bounded message delay. In [9], a token-based scheme is proposed to provide guaranteed priority for different traffic classes in wireless local area network (WLAN). A token-based broadcasting protocol called reliable neighbor-cast protocol (RNP) [10] is proposed to provide reliable multi-cast within a vehicle’s neighboring area. In RNP, tokens are used as the acknowledgement control message, and the neighboring group is determined by a voting scheme. Aggressive and time-driven acknowledgement mechanisms are used to assure the reliable broadcast in the neighboring area. In [11], the wireless token ring protocol (WTRP) is devised for Intelligent Transportation System (ITS) with guaranteed QoS and can recover from multiple transmission failures. However, it is not capable of adapting fast topology change notable in vehicular communications, and no reliable transmission mechanism for warning message is provided.

The main challenges in vehicular MAC protocol design include efficient coordination among nearby vehicles, data transmission with quality-of-service (QoS), and reliable safety message exchange. In this paper, we propose an overlay token ring protocol (OTRP) to address the aforementioned issues. Vehicles are logically organized into multiple overlapped virtual rings. A unique token in each ring can be used as the sole right of transmission, and it performs the predefined control functions. The proposed OTRP can jointly work with many existing MAC mechanisms (e.g., CSMA based or multi-
channel based) to handle multiple overlapped rings in the vicinity. The ring can be dynamically adjusted according to the present traffic condition. Vehicles may join/leave the ring when they move in/out the ring coverage. Different from the existing token-ring based protocols, the OTRP contains two modes, namely the normal and emergency modes, whereby timely emergency message dissemination is guaranteed and desired QoS for data service can be provided.

The remainder of this paper is organized as follows. The proposed OTRP is described in Section II. The performance of the OTRP is theoretically analyzed in Section III. Analytical results are compared with simulation ones in Section IV. The impact of the key MAC parameters on the protocol performance is also discussed. Concluding remarks are given in Section V.

II. PROTOCOL DESCRIPTION

In OTRP, the vehicular network is considered as overlapped virtual rings, and each of which is composed of multiple vehicles within the same communication range, as shown in Fig. 1. We use the term “vehicle” and “node” interchangeably. Vehicles are classified into three states: FREE (nodes that are not belonged to any ring), INRING_IDLE (nodes in a ring without token on hand), and INRING_BUSY (nodes currently hold the token). Moreover, the protocol operates in two modes: the normal mode and the emergency mode. The former is mainly used for data communication among vehicles, while the latter is devised for exchanging safety messages. The protocol principle is described as follows.

A. Ring Formation

A new ring is formed when a FREE node fails to join an existing ring for a predefined duration, denoted as \( T_r \). It may also happen if the existing ring has no room to accept any new node. When a new ring is formed, the node will generate a token and set its status to INRING_BUSY. The capacity of each ring, namely the ring size, is the maximum number of nodes that a ring can contain. \( T_r \) is set a little longer than the token rotation time (as will be defined in Sec. III) to prevent the situation that multiple unsaturated rings exist without FREE nodes to join thus increase the unnecessary contentions.

B. Token Type

Different token types are distinguished by frame control field in the token frame. Five types of control messages are defined.

1) DATA token: the DATA token serves as the sole right for a node to access the channel. When the token resides at a node, it means that a maximum token holding time (MTH) is allocated to the node for data transmission.

2) OPEN token: this token is used to indicate that the node number has not reached the ring size. Hence nodes in FREE status and hear the OPEN token can join the ring by responding to the current token holder.

3) Responding message (RS): the responding message is sent from FREE nodes when hearing the OPEN token; it contains the node address (id).

4) Acknowledgement Token (ACK): control message is modified to ACK when the token site hears an emergency message. It is used to acknowledge the emergency message at nodes in the ring. If the node receives the message, ACK is passed to the next node. Otherwise, it is sent to its previous node as a request for retransmission of the emergency message.

5) Request for message (RQ): it is used when an ACK confirms that a node is lack of the emergency message. The token is used to request the destination node to resend the message.

C. Ring Operation

The ring operation in OTRP consists of two aspects. For each node, after joining the ring, it needs to be aware of its transmission order among other members. This information is stored in a field of the token frame, called Order List (OL). OL is updated by the token whenever join or leave events happen. This allows the node to maintain the most up-to-date OL. On the other hand, for each ring, its structure dynamically changes due to fast mobility and speed of vehicles. The ring thus operates differently in response to the join, leave, and reformation processes as described in the following.

Join process: After data transmission, the current token holder will broadcast OPEN token if the ring size does not reach the maximum. By overhearing the OPEN token, a FREE node can join the ring through contending with other FREE nodes. A timer \( T_j \) is triggered at the instance as the OPEN token is sent. When it expires, the token site will update its OL according to the received information and release the token to the succeeding node. \( T_j \) is an adjustable parameter, depending on the number of nodes that are allowed to join the ring in each process.

Leave process: Leave process occurs when a ring member moves out of the coverage of the ring. A node will transmit data immediately after obtaining the token. When there is no
frame heard from the next node after the token is released from the current token site, it can be a sign that the node has left the ring. Thus the leave process is detected by not hearing transmission from “the next node” by a period of time. Under this situation, the node address will be removed from OL and the token will be passed to its next node in updated OL. The leaving node will set its status as FREE if not obtaining the token for a time duration that allows itself being able to join a nearby ring.

**Reformulation**: If the token site leaves the ring, the token can not be successfully passed to the succeeding node with the allowable number of trials. In this case, the token site will set itself status as FREE.

### D. Emergency Mode

In case an accident happens on the road, it is of vital importance to notify other members in the network as soon as possible. The emergency mode is designed to provide reliable message dissemination in a timely manner. When a node detects something dangerous, it becomes the emergency site and generates an emergency message immediately. The warning message has the highest priority over the normal token, which can be implemented along with the underlying channel access mechanism. For instance, the token site needs to sense the channel at the beginning of each slot before transmission. If the channel is sensed to be busy, indicating that the emergency site is transmitting, the token site will defer its transmission to the next slot. The warning message detection may also be implemented using the busy tone scheme [12].

When the token site receives the emergency message, it changes the token type to ACK, which is then passed in sequence according to the OL. Otherwise, the token type is set to RQ and sent to its predecessor for retransmission until the message is acknowledged. The acknowledgement process stops when the token reaches a node whose message is already acknowledged. During the process, nodes can still join and leave the ring. On the other hand, data transmissions are suspended for a period of time, during which no ACK is received. Fig. 2 shows the OTRP emergency mode.

### III. PERFORMANCE ANALYSIS

In this section, we analyze the performance of the proposed OTRP. To get insight into the interaction between various parameters and the protocol behavior, we focus on a single ring case as the first step. For a large-range vehicular network in which vehicles are virtually grouped to multiple rings, further research is underway.

Suppose there is a ring already formulated, which consists of \( N \) nodes and a number of FREE nodes in the vicinity. The ring size is \( N_m \). Saturated network traffic condition is assumed, and all the control messages are of the same size. We define the “transmission process” to describe the course that the token travels from the current token site to the tagged node, denoted by \( m \), as depicted in Fig. 3. To model the dynamic change of ring structure, FREE nodes can join an unsaturated ring with probability \( p_1 \) when it hears the OPEN token. During each SLOT, a node in the ring has the probability \( p_2 \) of leaving the ring. We also define the number of FREE nodes that can join the ring within each join process as \( n_f \).

We aim to derive the following important performance metrics. The average access delay, denoted as \( \overline{W} \), is the time interval between a frame arrives the buffer and the time it is about to be sent. The average throughput, denoted as \( \overline{S} \), is the average amount of data payload transmitted per unit time. The average token rotation time, denoted as \( \overline{T} \), is defined as the average duration that the token consecutively arrives at the same node. For emergency mode, the average emergency delay denoted as \( \overline{E} \) is the interval that the emergency message is generated to the end of acknowledgement process. They are obtained by calculating the corresponding time cost during the transmission process in Fig. 3, including the data and token transmission of the passed nodes, and the time cost for accepting new nodes and detecting leave nodes.

We use the following auxiliary variables in the analysis. The length of the transmission process measured in SLOT is represented by \( d_m \), which is uniformly distributed between \( [0, N - 1] \). During the transmission process, the number of
leave nodes, \(N_t\), is binomial distributed with mean
\[
\mathcal{N}_t = \mathcal{A}_m \cdot p_2,
\]
where \(\mathcal{A}_m = (N - 1)/2\). The variable \(N_a\) denotes the number of nodes added during the transmission process, which consists of two components. One counts for the number of the current open positions in the ring, and the other counts for the potential open positions due to the leave nodes. The mean number of \(N_a\) can be obtained as
\[
\mathcal{N}_a = \left\lceil \frac{N_m - N}{n_f} \right\rceil \cdot n_fp_1 + \mathcal{N}_t \cdot p_1
\]
The variable \(N_i\) denotes the number of initiated join processes, which has the mean value given by
\[
\mathcal{N}_i = \left\lfloor \frac{N_m - N}{n_f} \right\rfloor + \left\lceil \frac{\mathcal{N}_t}{n_f p_1} \right\rceil
\]
Finally, a timeout value, \(\tau\), is defined as the time needed to confirm that the token has been correctly passed to the next node. Assuming a node transmits data immediately after it gets the token, and ignoring the propagation and processing delay, \(\tau\) can be approximated by the token transmission time, denoted by \(T_t\).

A. Average Access Delay (\(\bar{W}\))

The access delay consists of two parts: the time for the token to travel from the token site to the tagged node \(m\), denoted by \(W_T(m)\), and the time that the frame needs to wait in the buffer before transmission, denoted by \(W_Q\). Given \(d_m\), \(W_T(m)\) can be derived as
\[
W_T(m) = T_t + (\eta + T_t)(d_m - 1 + N_a - N_t) + N_i \cdot T_j + N_t \cdot \tau,
\]
where \(T_t\) is defined in Sec. II-A, and \(\eta\) is the MTH. Since the transmission time for each frame in the buffer within MTH is the same, thus the average queueing time for a frame in the buffer can be approximated as
\[
W_Q = \eta/2.
\]
The average access delay can thus be obtained as
\[
\bar{W} = (T_t + \eta)(\mathcal{A}_m + \mathcal{N}_a + \mathcal{N}_t) + \mathcal{N}_i T_j + \mathcal{N}_t T_t.
\]

B. Average Rotation Time (\(\bar{R}\))

\(\bar{R}\) can be readily obtained if a ring operate into a stationary status under the assumption that there are always FREE nodes around. Therefore, the node number can reach ring size by accepting FREE nodes after sufficient long time. So during a rotation the actual number involved in the transmission is the ring size plus the joining nodes remedies those leave the ring. Thus the token rotation time can be calculated as
\[
\bar{R} = N_m(T_t + \eta)(1 + p_1p_2 - p_2) + \left\lceil \frac{N_m p_2}{n_f p_1} \right\rceil T_j + N_m p_2 T_t.
\]

C. Average Throughput (\(\bar{S}\))

The average throughput is calculated as the average amount of payload transmitted during the token rotation time, as given by
\[
\bar{S} = \frac{L \cdot \eta}{(L + \eta)\bar{R}} \cdot B,
\]
where \(L\) is the frame payload size, \(H\) represents the physical and MAC layer overheads, and \(B\) is the channel bandwidth.

D. Average Emergency Delay (\(\bar{E}\))

In emergency mode, the token has to pass across the entire ring to complete the acknowledge process for the emergency message. This process takes \(\bar{R}\) time. Besides, the emergency delay contains three additional parts: (1) \(T_c\) is the time that the emergency message waits in the buffer before acknowledged (see Appendix), (2) \(T_f\) is the time for the flooding period, which can be approximated as the time for an emergency message to reach the farthest node in the ring, and (3) \(T_e\) is the time cost for retransmitting the missing emergency message. Assuming the equivalent traffic moving, i.e., the number of vehicles leaving an area is identical to the number of vehicles joining the same area, the average number of nodes that have not received the emergency message and need retransmissions equals \(\mathcal{N}_t p_1\). It takes \(2T_f + T_e\), where \(T_e\) is the time for transmitting an emergency message, for each node to receive the emergency message successfully. Without loss of generality, we assume the retransmission is always successful. Therefore, the average time cost for retransmission is
\[
\bar{T}_e = \mathcal{N}_t p_1 (2T_f + T_e).
\]

Accordingly, the average emergency delay can be given by
\[
\bar{E} = \bar{R} + \mathcal{N}_a + T_f + \bar{T}_e.
\]
Note that \(\bar{R}\) is obtained from (7) by setting \(\eta\) to zero and \(N_m\) to \(N\).

IV. NUMERICAL RESULTS AND DISCUSSIONS

We conduct extensive simulations to evaluate the performance of the proposed OTRP, and to verify our analytical model. We simulate the OTRP using our simulator written in C++. A network with 20 nodes is considered, where 6 of them have formed a ring. The number of nodes that join the ring during each join process is set to 3. All the results are collected from 100 runs for each figure. For emergency scenario, the emergency event is configured to happen at 7000 ms. Table I lists the main parameters used in the performance evaluation.

We first study the performance in normal mode operation. The impacts of join probability and ring size to various

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>MTH ((\eta))</td>
<td>200 ms</td>
</tr>
<tr>
<td>TOKEN size</td>
<td>800 bits</td>
</tr>
<tr>
<td>Bandwidth ((B))</td>
<td>11 Mbps</td>
</tr>
<tr>
<td>Join timer ((T_f))</td>
<td>6 ms</td>
</tr>
<tr>
<td>Frame payload ((L))</td>
<td>4000 bits</td>
</tr>
</tbody>
</table>
Fig. 4. Comparison between analytical and simulation results by varying the leaving probability and ring size.

The impact of join probability shows an increasing trend in access delay and rotation time with the increase of join probability. The change in ring size has a more significant effect than the join probability. The reason for the gently increasing rotation delay is that a higher join probability implies the ring tends to stay in the saturated condition more often. The rotation time is dominated by the number of nodes in the ring, and when the ring becomes saturated, the average rotation time only increases slightly depending on the frequency of join and leave events. For average throughput, the reduction is quite limited (6.7%) when join probability increases from 0.2 to 1, indicating that OTRP can adequately adapt to the quickly changing environment. From the figure, we can also see that the analytical results agree with the simulation ones very well.

Fig. 5. Emergency delay and complete time vs. ring size and join probability.

The impact of ring size shows that emergency delay and complete time are within 100ms. Moreover, the impact of join probability is shown in Fig. 5(b). A decreasing trend can be explained as follows. Generally the most of the time spent in emergency mode is the rotation time $R$. It is evident from Eq. (7) that the leaving probability $p_1$ weights more due to the second part (a function of $T_j$) than the first part (a function of $\eta$) does. In normal mode, $\eta$ is much larger than $T_j$, which renders the second part having minor effect. Contrarily, in emergency situation, $\eta$ becomes zero. Hence the second part, which is inversely proportional to $p_1$, has the leading effect.

Finally, we compare the average throughput of the OTRP with the existing distributed MAC protocol. The prevalent IEEE 802.11 MAC is chosen as the benchmark. For the sake of fair comparison, the distributed coordination function (DCF) mode, i.e., CSMA/CA with binary exponential backoff, is considered. The analytical model in [13] is used to obtain the average throughput of DCF. In the OTRP, the number of rings, i.e., the number of contending nodes depends on the total node number and ring size. For instance, if there are 11 nodes and the ring size is 6, then there are two rings, one with 6 nodes and the other with 5 nodes. The average throughput in the OTRP with the same collision resolution mechanism as IEEE 802.11b DCF is then obtained from Eq. (8).
During the normal mode, the occurrence of an emergency may interrupt the following four processes: data transmission, token transmission, join process, and leave process, respectively. Therefore, the probability that an emergency occurs in a particular period is the proportion of time spent in the corresponding process over the whole rotation time. For example, the probability for an emergency event happening in the data transmission period of a token holder is equal to the allowable data transmission duration divided by the rotation time. The probabilities associated with each process are given as follows.

1) Data transmission: the probability of an emergency event taking place in this period is $\frac{R(N-T_i-N_j)}{T_o}$. During this period, the average time occupied by the emergency event is $\frac{T_o}{2}$.

2) Token transmission: the probability of emergency occurrence is $\frac{T_o(N-T_i-N_j)}{T_t}$ with average duration $\frac{T_o}{2}$.

3) Join process: the emergency event may occur in this period with probability $\frac{N_jT_j}{T_o}$, and expected duration of emergency event is $\frac{T_o}{2}$.

4) Leave process: the probability of emergency occurrence is $\frac{N_jT_j}{T_o}$ with average duration $\frac{T_o}{2}$.

As a result, we can obtain the average value of $T_o$.

ACKNOWLEDGMENT

This work was supported by the Nature Sciences and Engineering Research Council (NSERC) of Canada.

REFERENCES


