Research Article

Data Dissemination Based on Fuzzy Logic and Network Coding in Vehicular Networks

Xiaolan Tang, 1 Zhi Geng, 1 Wenlong Chen, 1 and Mojtaba Moharrer 2

1 College of Information Engineering, Capital Normal University, Beijing 100048, China
2 Schepens Eye Research Institute, Massachusetts Eye and Ear, Harvard Medical School, Boston, MA 02114, USA

Correspondence should be addressed to Wenlong Chen; chenwenlong@cnu.edu.cn

Received 31 March 2017; Revised 20 July 2017; Accepted 3 August 2017; Published 10 September 2017

Abstract

Vehicular networks, as a significant technology in intelligent transportation systems, improve the convenience, efficiency, and safety of driving in smart cities. However, because of the high velocity, the frequent topology change, and the limited bandwidth, it is difficult to efficiently propagate data in vehicular networks. This paper proposes a data dissemination scheme based on fuzzy logic and network coding for vehicular networks, named SFN. It uses fuzzy logic to compute a transmission ability for each vehicle by comprehensively considering the effects of three factors: the velocity change rate, the velocity optimization degree, and the channel quality. Then, two nodes with high abilities are selected as primary backbone and slave backbone in every road segment, which propagate data to other vehicles in this segment and forward them to the backbones in the next segment. The backbone network helps to increase the delivery ratio and avoid invalid transmissions. Additionally, network coding is utilized to reduce transmission overhead and accelerate data retransmission in interbackbone forwarding and intrasegment broadcasting. Experiments show that, compared with existing schemes, SFN has a high delivery ratio and a short dissemination delay, while the backbone network keeps high reliability.

1. Introduction

Vehicular networks, as a promising technology for intelligent transportation systems, usually utilize vehicle-to-vehicle (V2V) communications to support data services when no fixed infrastructures are deployed. They are designed to improve driving safety and enhance the driving experience by supporting smart applications such as collision warning, traffic congestion alarm, and sharing parking information [1–3]. When a collision occurs, this warning information should be immediately disseminated to those vehicles which might be affected, in order to avoid new rear-end collisions and potential traffic jams after this accident. An example scenario of safety alert dissemination is shown in Figure 1.

Data dissemination in vehicular networks faces many challenges including variation in vehicle densities, frequent topology change, and limited wireless communication bandwidth [4, 5]. In order to design an efficient data dissemination scheme aiming for high delivery ratio, short propagation delay, and low resource consumption, two aspects need to be considered: driving environments and content broadcasting. Complicated driving environments, as the first aspect of designing efficient data dissemination schemes, result in multiple factors affecting the intervehicle communication performance. Existing research analyzes some parameters in relay node selection, but the comprehensive influence of intervehicle distance, channel quality, and other factors still require further study. For content broadcasting, the second aspect of designing efficient data dissemination schemes, some schemes select different forwarding nodes for different data flows. Since wireless signals are likely to overlap with others in a geographical area, data dissemination in vehicular networks by flooding easily results in serious redundancy, contention, and collision [6]. With transmission demand increasing, the probability of broadcast storms may sharply rise. To address this issue, some studies select several nodes with excellent communication capabilities to disseminate data, which solve the broadcast storm problem in dense scenarios. However, there is still a lot of work to be done on
how to select appropriate nodes and improve the efficiency of data propagation in vehicular networks.

In this paper, a data dissemination scheme based on fuzzy logic and network coding in vehicular networks is proposed, named SFN. A primary backbone and a slave backbone are selected in each road segment to construct a backbone network, through which all the packets are transmitted between different road segments. The backbone nodes are stable for a certain period of time and serve for different data flows. Therefore, the SFN scheme avoids too frequent relay node selection and decreases the probability of data resource contention. Specifically, the backbone nodes are selected according to the transmission abilities of all the vehicles in this segment, which are calculated by fuzzy logic and comprehensive consideration of the vehicle velocity change rate, the velocity optimization degree, and the channel quality. Additionally, based on network coding, an efficient forwarding and retransmission algorithm is designed for both intersegment and intrasegment communications. It helps in reducing and accelerating the retransmissions; therefore, the resource consumption decreases and the dissemination latency is shortened.

The main contributions of this proposal are in three aspects.

1. Construct a backbone network composed of two backbone nodes in each road segment to support data transmissions between road segments. This backbone network utilizes fuzzy logic to select backbones based on three factors: the velocity change rate, the velocity optimization degree, and the channel quality. This is done to improve the stability and reliability of the backbone network.

2. Network coding in data forwarding and retransmission algorithm improves the backbone-to-member and interbackbone transmissions. This improvement is because backbone nodes encode and decode the data packets, leading to a small transmission overhead and a quick data recovery.

3. Conduct sufficient experiments to evaluate the performances of this proposal and analyze some parameters. In particular, real taxi trajectory data in Sanya, China, are used to construct vehicular scenarios. The experimental results show that SFN has a higher delivery ratio and a shorter dissemination delay than compared schemes.

The remainder of the paper is organized as follows. In Section 2, the related work on data dissemination in vehicular networks is introduced. In Section 3, an overview of SFN scheme is provided. The details of the backbone network construction are discussed in Section 4. The coding and forwarding algorithm is presented in Section 5. Then, experimental results and analysis are shown in Section 6. Finally, Section 7 provides the conclusion of this paper.

2. Related Work

Currently, a big challenge in vehicular networks is to achieve stable and reliable data transmission in the scenarios with frequent topology change and limited bandwidth. Some early researches use flooding for data transmission, in which a node rebroadcasts packets to its neighboring vehicles once it receives them. When there is a traffic jam on the road, it is easy to cause broadcast storms and information congestion, and hence the dissemination delay sharply increases. Then, some studies try to reduce the transmissions of redundant packets. In order to improve the information propagation reliability and address the broadcast storm problem, Korkmaz et al. only allow the node farthest from the sender to broadcast the packet [7]. In [8], the receiver calculates the forwarding probability based on the distance between the sender and the receiver and sets the forwarding waiting time for retransmission checking according to the current time slot. The next-hop node is selected considering a single factor, which may cause velocity instability and decrease channel quality. Shen et al. design a data scheduling framework, which avoids the collisions and improves the dissemination efficiency by providing the transmission opportunity to nodes with maximum utilities [9].

Some studies analyze the impacts of relevant factors on the dissemination performance in vehicular applications. Zhu et al. propose a data forwarding strategy based on relative velocity and distance between vehicles, which improves forwarding efficiency [10]. In [11], the vehicles parked on the pavement or in parking lots are selected as backbones, which extend the coverage of the network. A scheme in [12] introduces a delay model and an improved greedy broadcast algorithm as well as a coverage elimination rule, taking into account road topology and traffic signals. Specifically, vehicle density is considered in some routing schemes. In [13], routing protocols are adapted for vehicular applications in a real-time way, according to the current level of vehicle density. For highway or urban scenarios, a protocol supports multidirectional data dissemination by combining a generalized time slot scheme based on directional sectors and a store-carry-forward algorithm [14].

In recent years, as a classical mathematical method, fuzzy logic has been utilized to improve data propagation
in vehicular networks. In [15], a protocol incorporates fuzzy logic with geographical routing when making forwarding decisions. It takes the moving direction and the distance as the inputs of fuzzy logic and improves the delivery ratio. A seamless streaming dissemination system for vehicular networks is designed in [16]. It uses fuzzy logic to check if a roadside unit or a vehicular node can be a candidate to transfer stream data for users or not. In [17], Wu et al. propose a fuzzy-logic-based algorithm considering link quality, intervehicle distance, and vehicle mobility and design a redundancy transmission approach to enhance reliability.

Besides, since network coding can enhance data delivery in wireless communications, some existing researches focus on its benefits in data dissemination [18]. The paper [19] uses network coding to rebroadcast the messages, which improves the overall reliability and delivery ratio. In [20], cache solutions utilize network coding to reduce bandwidth cost and shorten latency. An abstract model of a general network coding process is developed to support distribution of content in vehicular networks. Furthermore, some studies combine fuzzy logic with network coding in vehicular networks. In SBN scheme, the factors including vehicle velocity, vehicle density, and channel quality are taken into account in fuzzy logic, while it also uses network coding to improve the transmission efficiency [21]. However, since this scheme prefers to select the vehicles with slow velocities as backbone nodes, its performance is greatly affected when most of the vehicles have much higher velocities than the backbones. FUZZBR scheme in [22] models the forwarding ability based on distance, velocity, and communication quality and selects two relays within a particular range by using fuzzy logic. In [23], fuzzy logic is utilized to select next-hop nodes, considering factors such as the distance between vehicles and vehicle velocity and density.

In order to improve the data dissemination in vehicular networks, how to select appropriate backbone nodes and how to fully explore network coding both require further study. Therefore, the scheme applied in this paper uses fuzzy logic to comprehensively consider the state of traveling vehicles and the channel quality of V2V communications. In addition, two backbone nodes in each road segment are selected to construct a reliable backbone network and use network coding in data forwarding to shorten propagation delay and reduce bandwidth consumption.

3. SFN Scheme Overview

In SFN, the focus is on data dissemination scenario, where the data (such as collision warning messages) have one source (such as the vehicle which has collided) and several destinations (such as the vehicles behind which might be affected). This proposed scheme aims to deliver the data from its source to all the destinations quickly with a small overhead. It is noteworthy that, with the provision of backbone network, this scheme also applies in unicast transmission applications each having one source and one destination, with some simple adjustments.

To collect real-time information of traveling vehicles, each vehicular node gets its location and velocity from onboard equipment like GPS and speedometer. It is assumed that all the vehicular nodes have the same communication radius, denoted by $R$. For efficient multihop data transmissions, a long road is divided into multiple segments, and the length of each segment is $R/2$, which ensures the connectivity of the backbone network [20]. It should be noted that the road width is negligible when compared with the transmission radius.

In the scheme, there are two kinds of vehicular nodes: backbone nodes and member nodes. In every segment, two backbone nodes are selected, named primary backbone and slave backbone, and other nodes are member nodes. The backbone nodes transmit packets between different road segments and propagate packets to member nodes in the same segment, while member nodes only generate and receive packets. In each segment, primary backbone has a higher transmission ability than slave backbone, and it takes more data transmission tasks. Having two backbone nodes in each segment, as opposed to one backbone node, has two main benefits. First, if one backbone loses data or gets incorrect data, the other can transmit correct data to the next segment immediately. In addition, the missing or incorrect data can be recovered through exchanging packets between backbones. In this way, the scheme reduces the transmission overhead and shortens the dissemination delay.

When a data packet such as a collision warning message is generated and is ready to be disseminated in a specific area, the sender transmits it to the backbone nodes in the same segment first. In the next step, it is transmitted through the backbone network to all the target segments. In an example scenario shown in Figure 2, a highway has four segments: $S_1$, $S_2$, $S_3$, and $S_4$. In $S_1$, $V_2$ is the primary backbone and $V_3$ is the slave backbone. Similarly, $V_5$, $V_{10}$, and $V_{11}$ are primary backbones in their segments, while $V_8$ and $V_{12}$ are slave backbones. The source $V_1$ generates two data packets $p$ and $q$, which need to be broadcasted in all four segments. First, $V_1$ transmits the original data to the backbones $V_2$ and $V_3$ in its segment. Then, $V_2$ encodes them to $p + q$ and $2p + q$ and forwards the encoded packets to the backbones $V_4$ and $V_5$ in the next segment. If $V_2$ loses $p + q$ and $V_3$ loses $2p + q$, they firstly send the received packets to the next backbone node $V_{10}$ and then recover their lost packets through packet exchange between themselves. Backbones $V_{11}$ and $V_{12}$ obtain the data through backbone network in a similar way. Additionally, the primary backbones $V_2$, $V_5$, $V_{10}$, and $V_{11}$ decode and broadcast the data to the member nodes in their segments.

If there is only one vehicle in a segment, such as $V_{10}$ in $S_3$, it is selected as a unique backbone, the primary backbone. Data dissemination in sparse scenarios will be discussed later. This paper uses linear network coding [24] in the examples and experiments; however, this suggested scheme also supports other network codes.

It is obvious that how to select backbone nodes greatly affects the data dissemination performance. In SFN, backbone nodes are selected periodically according to the transmission abilities of the vehicles. Specifically, each vehicle calculates its transmission ability and shares it with others.
within the same segment. The two nodes with the highest abilities work as backbone nodes, among which primary backbone has a higher ability than slave backbone.

Here, the method to compute the transmission ability is the key for backbone selection. Since the relation between the driving status and the communication quality is not clear, it is challenging to design a simple and accurate model. Considering that fuzzy logic is an efficient method to model complicated relations [25], it is used in this study to compute the transmission ability. By doing so, three factors are taken into account: the velocity change rate, the velocity optimization degree, and the channel quality. With respect to the cycle of backbone selection, from experiment results in Section 6, the backbone network in SFN is more stable than the compared schemes. Therefore, the backbone selection cycle can be longer and hence it is less costly to maintain backbone network.

In vehicular data services, a vehicle may receive several messages that need to be forwarded to others from different neighbors. In traditional schemes, a sender sends out the original messages in different directions, regardless of the existing packets at receivers. Additionally, once several packets are lost, the sender resends these packets, not considering the different lost packets at different receivers. In this way, the network has a large retransmission overhead especially in case of lots of data. However, network coding works well in these cases, which reduces the transmission consumption. In SFN, backbone nodes encode the original packets to coded packets and hence improve the communication performance. This is explained in detail in Section 5.

In dense vehicular networks, the backbone network helps in transmitting data effectively and reducing the probability of broadcast storm. However, in sparse vehicular scenarios, it is difficult to select two backbones in every segment, and hence the data dissemination may be interrupted. In order to support data propagation in sparse areas, in SFN, if there is a unique vehicular node in a segment, it is regarded as the primary backbone to maintain connectivity. If there is no vehicle in a segment $S_i$, a sender in the previous segment $S_{i-1}$ tries to find another forwarder in the next segment $S_{i+1}$. Since the length of each segment is $R/2$, two nodes in $S_{i-1}$ and $S_{i+1}$ might be in the transmission range of each other. Furthermore, if several nodes in $S_{i+1}$ could communicate with the sender in $S_{i+1}$, two nodes are selected that are moving in the same direction with the sender and having the largest transmission abilities as the next backbones. Although the delivery probability may be reduced due to a long distance between nodes in $S_{i+1}$ and $S_{i+1}$, the connectivity of the backbone network is enhanced. In case the vehicle density is so low that the sender cannot communicate with any vehicles in the next two segments, it carries packets and continues traveling until a next-hop vehicle appears.

4. Backbone Network Construction

In this section, the selection of backbone nodes and construction of a backbone network are discussed. The backbone node selection is decided by the transmission abilities of the vehicular nodes. These abilities are computed according to three transmission factors: the velocity change rate, the velocity optimization degree, and the channel quality. These factors are not independent; for example, a large velocity change rate leads to a low channel quality. Considering the complicated relations among these factors and the transmission ability, fuzzy logic is utilized to model the transmission ability. In detail, there are four steps as listed below.

(1) Calculation of the Transmission Factors. Each vehicle gets its driving and communication information from its sensors and hello messages and then calculates its velocity change rate, velocity optimization degree, and channel quality.

(2) Fuzzification. For each factor, a membership function is used to convert an original value to several fuzzy values.

(3) Calculation of the Transmission Ability Rank. All the combinations of nonzero fuzzy values of the three factors are obtained. For each combination, it is mapped to transmission ability ranks according to preset rules.

(4) Defuzzification. A function graph and a defuzzification method are designed to convert the transmission ability ranks and values to an overall transmission ability value.
4.1. Transmission Factors. In vehicular networks, the vehicle velocity and the channel quality greatly affect the transmission performance. Here, three variables are introduced to represent relevant factors.

(1) Velocity Change Rate. The relative velocity is an important factor in vehicular networks [26]. High velocity jitter usually leads to intermittent communication and transmission failures accordingly. Consequently, the vehicle with a high velocity jitter is not suitable to be selected as a backbone node. In order to indicate the velocity jitter of a vehicle, the velocity change rate is introduced, denoted by \( v_c \) and calculated by

\[
v_c = \left| \frac{v \cdot v_l}{v_s^2} - 1 \right|.
\]

Here, \( v \) is the current velocity of the vehicle, \( v_s \) is the average velocity within a recent short duration of \( t_c \), and \( v_l \) is the average velocity in a recent long period of time \( t_l \) \((t_l > t_c)\).

Since \( v_s \) and \( v_l \) imply the recent mobility features of the vehicle and \( v \) shows the current state, \( v_c \) indicates the recent velocity change rate. Because of the frequent velocity changes, here two average velocities are used rather than velocities at two specific times in the past. In (1), a big difference between \( v_s \) and \( v_l \) or between \( v \) and \( v_s \), meaning a high velocity jitter, results in a large \( v_c \). It is apparent that the larger the value of \( v_c \) is, the less chance the vehicle has to be selected as backbone. Therefore, in order to easily use fuzzy logic, only those vehicles with \( v_c \) in \([0, 1]\) are considered as backbone candidates. This study’s experiments show that SFN performs well when \( t_l = 10 \times t_{hello} \) and \( t_c = 4 \times t_{hello} \), where \( t_{hello} \) is the hello message cycle.

(2) Velocity Optimization Degree. The velocity change rate factor is not enough for backbone selection. This is because a vehicle with a stable velocity may not be a good backbone if its speed is very different from other vehicles in the same segment. The velocity optimization degree shows the relation between a vehicle’s average velocity and the optimal velocity of vehicles in the same segment, denoted by \( v_d \) and calculated by

\[
v_d = \left| 1 - \frac{v_l}{V} - 1 \right|, \quad v_l \leq 2V.
\]

Here, \( V \) is the optimal velocity of vehicles in this segment at this time.

Only the vehicles whose velocity \( v_l \) is less than or equal to \( 2V \) are considered to be backbone candidates. In this way, the value domain of \( v_d \) is \([0, 1]\), which can be easily analyzed in fuzzy logic, and a vehicle with a similar velocity to \( V \) has a high \( v_d \). In other words, a high \( v_d \) means the vehicle should have a high probability to be backbone. This is because it keeps the same pace with other vehicles and has a long encountering time to transmit data. Meanwhile, the above equation can be represented by

\[
v_d = \begin{cases} 
   \frac{v_l}{V}, & v_l \leq V; \\
   2 - \frac{v_l}{V}, & V < v_l \leq 2V.
\end{cases}
\]

There are many factors affecting the vehicle velocity, and [27] proposes an optimal velocity model by investigating the properties of congestion and the delay time of car motion. The optimal velocity \( V \) is determined by the vehicle density, the number of lanes, the traffic accidents, and some other factors. Although it is a complicated issue to select an appropriate value for \( V \), the comprehensive analysis of present and predicted traffic information may give some clues. For further information, please refer to [28, 29]. In this study’s simulation, the same optimal velocity is set for all segments except the two with smaller values, due to assumed traffic collisions.

(3) Channel Quality. The channel quality reflects the reliability of intervehicle channel, denoted by \( lr \). It is related to many factors, including the network technology, the local environment, the signals traveling through channel, and the fundamental physics behind wireless transmission. A greater channel quality provides a higher delivery probability. It is difficult to estimate channel conditions in vehicular networks accurately due to the frequent changes of network topology and the complex environmental factors, such as weather and nearby buildings [30]. Thus, the delivery ratio of hello messages, which are periodically exchanged among all the vehicles in the same segment, is used in experiments to represent channel quality.

4.2. Fuzzification. A membership function presents whether a value of an element falls within a specific range and indicates the membership degree in a fuzzy set [31]. The fuzzy set is obtained by assigning a value to each level to represent its grade of membership function. In SFN, fuzzy logic uses a membership function to convert the value of every factor to a fuzzy set. Suitable membership functions are acquired for transmission factors through data analysis and simulation experiments, shown in Figure 3. For instance, if the usual velocity range is \([60, 100]\) km/h, an extreme value of the velocity optimization degree is got when \( v_l = 100 \) and \( V = 60 \) as \( v_d = 0.3 \). Therefore, in the membership function of \( v_d \), \((0, 0.3)\) is ”bad” with probability of 1. Some additional experiments with different membership functions are illustrated in Section 6.3.

As Figure 3 shows, the velocity change rate has three levels \{Low, Medium, High\}, while the velocity optimization degree and the channel quality both have three levels \{Bad, Medium, Good\}. In Figure 3(a), when the velocity change rate \( v_c \) is 0.1, its fuzzy set via mapping is \{Low: 0.5, Medium: 0.5, High: 0\}. Similarly, when the velocity optimization degree \( v_d \) is 0.74 and the channel quality \( lr \) is 1, their fuzzy sets are \{Bad: 0, Medium: 0.75, Good: 0.25\} and \{Bad: 0, Medium: 0, Good: 1\}, respectively.

Based on the three fuzzy sets, several combinations are obtained, each of which consists of three nonzero fuzzy values relevant with three factors. In the above example shown in Figure 3, four combinations are listed in Table 1.

4.3. Transmission Ability Rank. First, the influences of transmission factors on the transmission ability in theory as well
as the feedback from experiments are analyzed. In the next step, 27 rules are set to determine the transmission ability rank, denoted by $tar$, as listed in Table 2. According to the rules, map the fuzzy sets of three transmission factors into a transmission ability rank. Overall, there are six ranks of the transmission ability as \{Perfect, Good, Acceptable, Not Acceptable, Bad, Very Bad\}.

Each fuzzy value combination is mapped to a transmission ability rank according to the rules. Meanwhile, the rank value is the minimum fuzzy value in the combination. When multiple combinations are mapped to the same rank with different rank values, the maximum rank value is selected.

To sum up, max-min method [32] is used to calculate

<table>
<thead>
<tr>
<th>Combination</th>
<th>$vc$</th>
<th>$vd$</th>
<th>$lr$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c_1$</td>
<td>Low: 0.5</td>
<td>Medium: 0.75</td>
<td>Good: 1</td>
</tr>
<tr>
<td>$c_2$</td>
<td>Low: 0.5</td>
<td>Good: 0.25</td>
<td>Good: 1</td>
</tr>
<tr>
<td>$c_3$</td>
<td>Medium: 0.5</td>
<td>Medium: 0.75</td>
<td>Good: 1</td>
</tr>
<tr>
<td>$c_4$</td>
<td>Medium: 0.5</td>
<td>Good: 0.25</td>
<td>Good: 1</td>
</tr>
</tbody>
</table>
Table 2: The rules of transmission ability ranks.

<table>
<thead>
<tr>
<th>Rule</th>
<th>Vc</th>
<th>Vd</th>
<th>Ir</th>
<th>Tar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rule 1</td>
<td>Low</td>
<td>Good</td>
<td>Good</td>
<td>Perfect</td>
</tr>
<tr>
<td>Rule 2</td>
<td>Low</td>
<td>Good</td>
<td>Medium</td>
<td>Good</td>
</tr>
<tr>
<td>Rule 3</td>
<td>Low</td>
<td>Good</td>
<td>Bad</td>
<td>Not Acceptable</td>
</tr>
<tr>
<td>Rule 4</td>
<td>Low</td>
<td>Medium</td>
<td>Good</td>
<td>Good</td>
</tr>
<tr>
<td>Rule 5</td>
<td>Low</td>
<td>Medium</td>
<td>Medium</td>
<td>Acceptable</td>
</tr>
<tr>
<td>Rule 6</td>
<td>Low</td>
<td>Medium</td>
<td>Bad</td>
<td>Bad</td>
</tr>
<tr>
<td>Rule 7</td>
<td>Low</td>
<td>Bad</td>
<td>Good</td>
<td>Not Acceptable</td>
</tr>
<tr>
<td>Rule 8</td>
<td>Low</td>
<td>Bad</td>
<td>Medium</td>
<td>Bad</td>
</tr>
<tr>
<td>Rule 9</td>
<td>Low</td>
<td>Bad</td>
<td>Bad</td>
<td>Very Bad</td>
</tr>
<tr>
<td>Rule 10</td>
<td>Medium</td>
<td>Good</td>
<td>Good</td>
<td>Good</td>
</tr>
<tr>
<td>Rule 11</td>
<td>Medium</td>
<td>Good</td>
<td>Medium</td>
<td>Acceptable</td>
</tr>
<tr>
<td>Rule 12</td>
<td>Medium</td>
<td>Good</td>
<td>Bad</td>
<td>Bad</td>
</tr>
<tr>
<td>Rule 13</td>
<td>Medium</td>
<td>Medium</td>
<td>Good</td>
<td>Acceptable</td>
</tr>
<tr>
<td>Rule 14</td>
<td>Medium</td>
<td>Medium</td>
<td>Medium</td>
<td>Not Acceptable</td>
</tr>
<tr>
<td>Rule 15</td>
<td>Medium</td>
<td>Medium</td>
<td>Bad</td>
<td>Bad</td>
</tr>
<tr>
<td>Rule 16</td>
<td>Medium</td>
<td>Bad</td>
<td>Good</td>
<td>Bad</td>
</tr>
<tr>
<td>Rule 17</td>
<td>Medium</td>
<td>Bad</td>
<td>Medium</td>
<td>Bad</td>
</tr>
<tr>
<td>Rule 18</td>
<td>Medium</td>
<td>Bad</td>
<td>Bad</td>
<td>Very Bad</td>
</tr>
<tr>
<td>Rule 19</td>
<td>High</td>
<td>Good</td>
<td>Good</td>
<td>Not Acceptable</td>
</tr>
<tr>
<td>Rule 20</td>
<td>High</td>
<td>Good</td>
<td>Medium</td>
<td>Bad</td>
</tr>
<tr>
<td>Rule 21</td>
<td>High</td>
<td>Good</td>
<td>Bad</td>
<td>Very Bad</td>
</tr>
<tr>
<td>Rule 22</td>
<td>High</td>
<td>Medium</td>
<td>Good</td>
<td>Bad</td>
</tr>
<tr>
<td>Rule 23</td>
<td>High</td>
<td>Medium</td>
<td>Medium</td>
<td>Bad</td>
</tr>
<tr>
<td>Rule 24</td>
<td>High</td>
<td>Medium</td>
<td>Bad</td>
<td>Very Bad</td>
</tr>
<tr>
<td>Rule 25</td>
<td>High</td>
<td>Bad</td>
<td>Good</td>
<td>Bad</td>
</tr>
<tr>
<td>Rule 26</td>
<td>High</td>
<td>Bad</td>
<td>Medium</td>
<td>Very Bad</td>
</tr>
<tr>
<td>Rule 27</td>
<td>High</td>
<td>Bad</td>
<td>Bad</td>
<td>Very Bad</td>
</tr>
</tbody>
</table>

Table 3: Combination mapping results.

<table>
<thead>
<tr>
<th>Combination</th>
<th>Rule</th>
<th>Tar</th>
<th>Rank value</th>
</tr>
</thead>
<tbody>
<tr>
<td>c₁</td>
<td>Rule 4</td>
<td>Good</td>
<td>0.5 [= min(0.5, 0.75, 1)]</td>
</tr>
<tr>
<td>c₂</td>
<td>Rule 1</td>
<td>Perfect</td>
<td>0.25 [= min(0.5, 0.25, 1)]</td>
</tr>
<tr>
<td>c₃</td>
<td>Rule 13</td>
<td>Acceptable</td>
<td>0.5 [= min(0.5, 0.75, 1)]</td>
</tr>
<tr>
<td>c₄</td>
<td>Rule 10</td>
<td>Good</td>
<td>0.25 [= min(0.5, 0.25, 1)]</td>
</tr>
</tbody>
</table>

Table 4: Final transmission ability ranks.

<table>
<thead>
<tr>
<th>Tar</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perfect</td>
<td>0.25</td>
</tr>
<tr>
<td>Good</td>
<td>0.5 [= max(0.5, 0.25)]</td>
</tr>
<tr>
<td>Acceptable</td>
<td>0.5</td>
</tr>
</tbody>
</table>

the transmission ability rank. For the above example, the combination mapping results are shown in Table 3, and the final transmission ability ranks are listed in Table 4.

4.4. Defuzzification. Next, several transmission ability ranks are converted into a digital number, which is the transmission ability value, denoted by $taV$, according to a function graph. The function graph of SFN scheme is shown in Figure 4. $taV$ is the coordinate of the center of gravity of a shadow $sp$, which is determined by the ranks and their values. $taV$ is calculated by

$$taV = \frac{\int_{x \in sp} \mu(x) x \, dx}{\int_{x \in sp} \mu(x) \, dx}.$$  

Still in the above example, the shadow in Figure 4 is in line with the transmission ability ranks listed in Table 4. Accordingly, the transmission ability value $taV = 0.71$ is calculated.

Each node broadcasts its own transmission ability value to others in its road segment. The higher the transmission ability value is, the better the node works as a backbone. Therefore, every node sorts the received values in ascending order and selects the nodes with maximum values as backbones. For instance, in Figure 2, there are six vehicles in the segment $S_4$, which are $v_{11}$, $v_{12}$, $v_{13}$, $v_{14}$, $v_{15}$, and $v_{16}$, and their transmission ability values are 0.91, 0.88, 0.75, 0.85, 0.66, and 0.76, respectively. Each vehicle receives and sorts the transmission ability values and selects $v_{11}$ as primary backbone and $v_{12}$ as slave backbone.

5. Data Coding and Forwarding Algorithm

In data transmission process, SFN uses network coding to reduce bandwidth consumption and support fast recovery when packet loss occurs. There are two main cases as follows.

1. In wireless communications, when there are several packets to be forwarded in different directions, the relay node can send a small number of coded packets to complete the delivery. As shown in Figure 5, $v_1$ and $v_3$, respectively, send packets $q$ and $p$ to each other through a relay node $v_2$. Without network coding, $v_2$ needs to send $q$ to $v_3$ and $p$ to $v_1$,
and thus the total number of transmissions is 4. By contrast, with network coding, \(v_5\) only needs to broadcast \(p + q\) to \(v_1\) and \(v_3\). Hence, the total number of transmissions is 3. This example shows that network coding reduces the transmission overhead which results in improvement of the bandwidth utilization.

(2) Network coding also works for efficient data recovery. In Figure 5, a primary backbone \(v_3\) wants to broadcast two original packets \(p\) and \(q\) to its member nodes. First, \(v_3\) encodes \(p\) and \(q\) to coded packets \(2p + q\) and \(2p + 3q\), which are linearly independent. If \(v_6\) only receives \(2p + 3q\) and \(v_7\) only receives \(2p + q\) correctly, \(v_3\) needs to resend data to \(v_6\) and \(v_7\). Using networking coding, \(v_3\) does not need to resend \(2p + q\) and \(2p + 3q\) again. Instead, it creates and sends a new coded packet \(2p + 5q\) with a new coefficient matrix. Then, \(v_6\) and \(v_7\) individually decode two coded packets and get original packets. This example demonstrates the improvement in retransmission efficiency and saving network resources through networking coding.

In SFN, every primary backbone uses networking coding to transmit data packets to backbones in the next segment and member nodes in the same segment, while each slave backbone only transmits coded packets to backbone nodes. After receiving data packets, the primary backbone decodes them to recover the original data. Then, it encodes these packets with its own coefficient matrix and finally broadcasts them to member nodes in the same segment and backbones in the next segment. If some packets are lost, the receiver sends a request message to the backbones in the same segment or in the upstream segment. In the next stage, when another node receives this request, it encodes its packets and transmits them to the request sender. An example of data coding and forwarding of backbone \(B_i^p\) in a segment \(S_i\), when it receives a packet \(p\), is illustrated in Algorithm 1.

In the algorithm, \(p\) is the packet \(B_i^p\) just received, and \(OP_i^p\) and \(CP_i^p\) are the original packets and the coded packets carried by \(B_i^p\), respectively. NOP is the new original packets \(B_i^p\) gets by decoding \(p\) and other existing packets; DP is the new coded packets \(B_i^p\) creates; \(B_{i+1}^p\) is the backbones in the segment \(S_i\), consisting of primary backbone \(B_i^p\) and slave backbone \(B_{i}^S\); and \(M_i\) is the member nodes in the same segment.

Also, SOP(\(x\)) is the original packets used to generate a coded packet \(x\); YP(\(x\)) is the existing packets at \(B_i^p\), which are generated by all or part of original packets in \(x\); and DIR(\(x\)) is the forwarding direction of packet \(x\). There are two forwarding directions along the road; DIR(\(x\)) = 1 means data transfer from \(S_i\) to \(S_{i+1}\) and DIR(\(x\)) = −1 means data transfer from \(S_i\) to \(S_{i−1}\). In addition, REQ(\(x\)) is a request for data packet \(x\).

According to the data coding and forwarding algorithm, when \(B_i^p\) receives a data packet \(p\) correctly, it tries to get original packets NOP. Then, \(B_i^p\) generates and forwards coded packets of NOP to the next-hop backbones and delivers them to the member nodes. If the data packet is not received correctly, \(B_i^p\) sends a request of this data to its slave backbone \(B_i^S\) for data sharing. When \(B_i^S\) receives a request of some data, if it has that data, it encodes and replies to the request sender; otherwise, it forwards the request to the last-hop backbones.

In SFN, fuzzy logic is used to select two backbone nodes in each road segment, construct a backbone network, and use networking coding to propagate data packets to target road segments. The backbone nodes with suitable velocity and good channel quality help to improve the transmission efficiency, while the network coding in intrasegment broadcast and intersegment forwarding may enhance the utilization of limited communication resources.

6. Performance Evaluation

6.1. Network Configurations. In order to evaluate the performance of SFN scheme, the opportunistic network environment simulator (ONE [33]) is used to conduct simulation experiments. The scenario configurations are listed in Table 5. Regarding the mobility model, a vehicle-following model [34] is utilized, which results in different vehicle velocities and different densities in the road segments. For \(vd\) calculation, to simulate different optimal velocities in different segments, a small optimal velocity of 60 km/h for two segments is set, 300–450 m and 1350–1500 m, while other segments have a higher optimal velocity of 90 km/h. Moreover, the influences of different backbone selection cycles, velocity ranges, and membership functions of transmission factors are discussed in Section 6.3.

SBN [21] and FUZZBR [22] are chosen as compared data dissemination schemes in vehicular networks. Compared with SBN, the proposed scheme SFN has several advantages. Although both of them utilize backbone nodes to forward data, SFN improves the backbone selection and the backbone network construction. Firstly, SBN prefers a vehicle with slow speed to be backbone in a road segment, no matter...
Input: \( p \), OP\( _p \), CP\( _p \);
Output: NOP, DP;
(1) if \( p \) is a correct data packet then
(2) \( u_{op} = SOP(p) - OP_p \);
(3) if \( \exists sc \subseteq YP(u_{op}) \) satisfying decoding conditions then
(4) obtain original packets SOP(sc) by decoding sc;
(5) NOP = SOP(sc) - OP_p;
(6) while \( \exists [r, s] \subseteq NOP \) having DIR(r) \( \neq \) DIR(s) do
(7) create a coded packet DP([r, s]);
(8) send DP([r, s]) to \( B_{i+1} \) and \( B_{i-1} \);
(9) send two coded packets using [r, s] to \( M_i \);
(10) NOP = NOP - [r, s];
(11) end
(12) create coded packets DP(NOP);
(13) if DIR(NOP) = 1 then
(14) send DP(NOP) to \( B_{i+1} \) and \( M_i \);
(15) else
(16) send DP(NOP) to \( B_{i-1} \) and \( M_i \);
(17) end
(18) end
(19) end
(20) if \( p \) is an incorrect data packet then
(21) send REQ(p) to \( B_i \);
(22) end
(23) if \( p \) is REQ(q) from node \( z \) then
(24) if \( q \) is carried by \( B_i \) then
(25) if \( \exists \) other original packet \( t \) to be sent to \( z \) then
(26) create one or several coded packets DP([q, t]);
(27) send DP([q, t]) to \( z \);
(28) else
(29) send \( q \) to \( z \);
(30) end
(31) else
(32) if DIR(q) = 1 then
(33) send REQ(q) to \( B_{i-1} \);
(34) else
(35) send REQ(q) to \( B_{i+1} \);
(36) end
(37) end
(38) end

Algorithm 1: Data coding and forwarding algorithm for primary backbone \( B_i^p \).

Table 5: Simulation environment configurations.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Road</td>
<td>2000 m with 4 lanes</td>
</tr>
<tr>
<td>Number of vehicular nodes</td>
<td>20, 40, 60, 80, 100</td>
</tr>
<tr>
<td>Data packet size</td>
<td>512 bytes</td>
</tr>
<tr>
<td>Vehicle velocity</td>
<td>Random in [60, 100] km/h</td>
</tr>
<tr>
<td>Communication radius</td>
<td>300 m</td>
</tr>
<tr>
<td>Backbone selection cycle</td>
<td>4 s</td>
</tr>
<tr>
<td>Simulation time</td>
<td>150 s</td>
</tr>
<tr>
<td>Hello packet exchange cycle</td>
<td>1 s</td>
</tr>
</tbody>
</table>

how fast other vehicles drive; SFN chooses a vehicle with a speed close to the optimal speed in each segment. Therefore, the backbones in SFN are probably in a similar pace to others, which results in more intervehicle communication chances. Secondly, considering that the frequent change of velocities affects the quality of wireless transmission, SFN chooses vehicles with relatively stable speeds as backbones, which enhance the efficiency of data delivery. Thirdly, SFN uses fuzzy logic to calculate the transmission abilities of vehicles, as well as a large quantity of experiments to figure out the appropriate membership functions. Fourthly, for sparse vehicular scenarios, SFN maintains the connectivity of backbone network by picking up backbones from a farther segment rather than the next one. Last but not least, while SBN has only one backbone in each segment, there are two backbones, primary backbone and slave backbone, in each segment in SFN, which not only improve the reliability of backbone network but also accelerate data delivery as a result.
of network coding. Overall, SFN is an innovative scheme with several advantages over SBN.

Comparing SFN and FUZZBR, both of them utilize fuzzy logic to select nodes for forwarding. However, the factors and their membership functions in fuzzy logic are different, and the tasks of those selected nodes also differ. On the one hand, as it was discussed above, SFN uses the speed factor in depth, which helps to select appropriate backbone nodes. On the other hand, FUZZBR selects the relays for each data delivery requirement, while SFN constructs a backbone network for all the data services during a period of time, which reduces the overhead of forwarder selection and improves the efficiency of data dissemination.

The experiments evaluate four criteria: the delivery ratio, the number of transmissions, the dissemination delay, and the backbone stability. The delivery ratio is the ratio of the number of delivered packets to the total number of transmissions during the data dissemination from the source to the destination segments. The higher the delivery ratio is, the better the performance the scheme has. Besides, the number of transmissions counts the successful and failed transmissions as well as the retransmissions, which indicates the communication overhead of the scheme. Additionally, the dissemination delay is the delay from the time the data generated to the time data dissemination finishes. A scheme with a short delay works well in delay-sensitive applications. The backbone stability is the ratio of the number of nodes working as backbones during a specific period of time to the total number of backbones. Higher backbone stability implies a less frequent backbone reselection and hence a smaller update cost of backbone network.

6.2. Simulation Results. This study’s SFN scheme is compared with SBN and FUZZBR, and the results are shown in Figure 6.

(1) Delivery Ratio. In Figure 6(a), the data delivery ratios of three schemes rise when the vehicle density increases, and SFN keeps a high and relatively stable ratio compared with SBN and FUZZBR. In a sparse network with 20 vehicles, SFN has a higher ratio than SBN and FUZZBR by about 25%, due to the data forwarding between two segments \( S_i \) and \( S_{i+2} \) when no vehicles exist in \( S_{i+1} \). In the scenario with a high density (100 vehicles), the delivery ratio of SFN is also higher than others by approximately 10%, because of its stable and efficient backbone network.

(2) Number of Transmissions. As Figure 6(b) shows, when the node density is low, SFN and FUZZBR transmit less data than SBN, and vice versa. As discussed above, SFN works well when there are few nodes, while in other schemes the sender searches for the next-hop relay all the time, resulting in a large transmission overhead. In Figure 6(b), SBN has the smallest transmission overhead among the compared schemes, when there are more than 40 vehicles. The main reasons lie in two aspects. First, because of the relatively even vehicle distributions in all segments, the intervehicle communication is relatively reliable. A small number of retransmissions make the advantages of two backbones and network coding in SFN not apparent. Second, in dense scenarios, the backbone network in SBN, consisting of one backbone in each segment, is well connected. Since it avoids the exchange among backbones in the same segment, the number of transmissions is smaller than SFN. However, if the number of vehicles is a bit less, the transmission overhead of SBN rises sharply. However, SFN keeps a relatively small communication cost, despite data sharing between the primary and slave backbones, which may be acceptable in most vehicular applications.

(3) Dissemination Delay. According to Figure 6(c), in sparse networks, the propagation latencies of all the schemes are longer than those in dense networks, due to the fewer encounter chances between vehicles. SBN has the longest delay, because the unique backbone in every segment has to wait for retransmissions in case of packet loss, before forwarding to the next segment. In general, SFN has a short delay, similar to FUZZBR.

(4) Backbone Stability. SFN and SBN use backbone nodes to forward the data packets, which help to avoid broadcast storms. Therefore, the maintenance of backbone network affects the performances of these schemes. The results of backbone stability in SFN and SBN are illustrated in Figure 6(d). This figure shows the proportions of nodes selected as backbones in consecutive 1–5 rounds (a round is 1 s) in all the backbones. It is obvious that, in SBN scheme, half of the backbone nodes only work for one round, which indicates that a majority of the backbone nodes are changed frequently. In comparison, the backbones in SFN are more stable than SBN. In detail, over 70% of backbones work for 2 to 4 rounds. Therefore, the cost of backbone update in SFN is small.

To sum up, in vehicular scenarios with different vehicle densities, SFN keeps a higher delivery ratio and a shorter propagation delay than SBN and FUZZBR and maintains an acceptable transmission overhead and a good stability of the backbone network.

6.3. Parameter Analysis. Considering that the backbone selection cycle and the velocity distribution affect the performances of data dissemination schemes, in this section, two main criteria are analyzed: the delivery ratio and the dissemination delay. Additionally, in order to select appropriate membership functions of transmission factors (as shown in Figure 3), a large number of experiments are conducted. Due to space limit, here only the coordinates of the peaks in medium levels of the membership functions are presented. It should be mentioned that, in these experiments, the number of vehicles is always 80.

In the first group of experiments, the backbone selection cycle ranges from 1 s to 6 s, and the results are illustrated in Figure 7. As shown in Figure 7(a), when the backbone selection cycle is shorter than 4 s, the delivery ratio of SFN keeps stable at around 87%. However, longer cycles such as 5 s and 6 s lead to lower delivery ratios by nearly 4% and 8%, because the network topology changes a lot during a long time and the backbones are not always suitable. In SBN, when the backbone selection cycle increases, the delivery ratio decreases gradually. In Figure 7(b), the dissemination
delay of SFN drops slightly when the cycle changes from 1 s to 4 s and then rises with increasing cycle. The reason is that too frequent backbone selection leads to a large control cost, and too long cycle has detrimental impacts on the performance of backbone network as a result of unsuitable backbones. Overall, in order to achieve a high delivery ratio and a short dissemination delay, 4 s is a good choice for backbone selection cycle in SFN.

In the second series of experiments, the maximum velocity difference ranges from 0 to 40 km/h with 80 km/h as the standard velocity. In other words, the speed ranges are [80, 80], [75, 85], [70, 90], [65, 95], and [60, 100], respectively. The results are shown in Figure 8. Obviously, with a large speed range, the delivery ratio reduces and the dissemination delay grows, because the variety of mobility aggravates the unreliable wireless channels. It is noteworthy that when the speed range changes, SFN keeps a higher delivery ratio than SBN and FUZZBR, while keeping a short delay.

In the third group of experiments, the performances of SFN with different membership functions of the three factors $v_c$, $v_d$, and $l_r$ are listed in Table 6. In Figure 3, the values of $v_c$, $v_d$, and $l_r$ at the peak points in medium levels are 0.2, 0.65, and 0.5, respectively. Here, the results with other values are illustrated. Table 6 shows that, in general, the delivery ratio, the transmission overhead, and the dissemination delay are best when using this paper’s selected values. Actually, a series of experiments with typical values provide guidance to determine the assignment of complicated variables.
Figure 7: Results with different backbone selection cycles.

Figure 8: Results with different maximum velocity differences.

Table 6: Results of SFN with different membership functions of transmission factors.

<table>
<thead>
<tr>
<th>Parameters in {x, vd, Ir}</th>
<th>Delivery ratio (%)</th>
<th>Number of transmissions</th>
<th>Dissemination delay (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>{x, 0.65, 0.5}</td>
<td>x: 0.2, 0.3, 0.4</td>
<td>86.53, 83.31, 73.26</td>
<td>59.20, 72.40, 65.40</td>
</tr>
<tr>
<td>Results</td>
<td></td>
<td>86.53</td>
<td>59.20</td>
</tr>
<tr>
<td>{0.2, y, 0.5}</td>
<td>y: 0.6, 0.65, 0.7</td>
<td>80.25, 87.50, 79.67</td>
<td>64.00, 58.20, 59.00</td>
</tr>
<tr>
<td>Results</td>
<td></td>
<td>80.25</td>
<td>64.00</td>
</tr>
<tr>
<td>{0.2, 0.65, z}</td>
<td>z: 0.4, 0.5, 0.6</td>
<td>83.60, 86.26, 78.55</td>
<td>68.20, 61.00, 66.40</td>
</tr>
<tr>
<td>Results</td>
<td></td>
<td>83.60</td>
<td>68.20</td>
</tr>
</tbody>
</table>
In one word, an appropriate backbone update frequency and suitable membership functions affect the overall performance of SFN, which can be selected through sample analysis in preliminary experiments. Moreover, SFN performs better than compared schemes in the scenarios with different speed ranges.

6.4. Performance Evaluation Using Real Taxi Trajectories. Besides the simulated vehicular scenarios, a set of experiments are carried out based on real taxi trajectories. The open-access dataset is provided by the Ministry of Transport of China [35]. It includes the real-time trajectories of 4600 taxis in Sanya, Hainan Province, from 9 a.m. to 10 a.m. on November 15, 2016. The GPS data collection cycle is 10 s.

In the experiments, to improve the data reliability, five roads are selected to create data dissemination scenarios, denoted by A, B, C, D, and E. They are displayed in Figure 9, and the time period is from 9 a.m. to 9:10 a.m. In these scenarios, there are 15, 14, 20, 255, and 15 vehicles traveling on the selected roads A–E, respectively. For some missing data, estimated locations according to the existing prior and next entries as well as the velocities are inserted.

The main difference between the simulation scenarios and the real trajectory scenarios is that the vehicle distributions in simulations are relatively even in all segments, while the vehicles are unevenly distributed in real scenarios. Therefore, the performances of the three schemes in real scenarios vary from the simulation results.

The results are shown in Figure 10. Compared with the results in simulated scenarios, generally, SFN can be seen to still perform best, and SBN has a better performance than FUZZBR. This is because the nonuniform traffic distribution and the different mobility patterns of vehicles greatly degrade the performance of FUZZBR.

From Figure 10(b), compared with the simulation results in Figure 6(b), SBN does not show the smallest transmission overhead in most of the cases in real trajectory experiments. The reason is that a large variety of mobility patterns of vehicles leads to unstable communications and hence weakens the performance of SBN. Meanwhile, because of the reliable backbone network and the network coding, SFN shows obvious advantages over others.

In conclusion, SFN with a reliable backbone network composed of two backbones in each segment and network-coding-based data forwarding has a high delivery ratio and a short dissemination delay, in both simulated vehicular networks and real vehicular scenarios based on taxi trajectory data.

7. Conclusion

In order to improve data dissemination in vehicular networks, the proposed SFN scheme utilizes fuzzy logic to construct backbone network and network coding for efficient data forwarding. In each road segment, a primary backbone and a slave backbone are selected according to the transmission abilities of vehicles, which are calculated by fuzzy logic and take into account three transmission factors: the velocity change rate, the velocity optimization degree, and the channel quality. Then, these backbones construct a backbone network to support intersegment data dissemination. Moreover, when transmitting several packets in different directions or retransmitting more than one packet, SFN uses network coding to reduce the transmission overhead and hence saves wireless bandwidth and achieves quick recovery. In particular, for sparse scenarios, a specific solution is put forward to accelerate data propagation. The experimental results show that SFN has a higher delivery ratio and a shorter dissemination delay than other schemes, while keeping backbone stability.

However, the quality enhancement of wireless communications in light of MAC and physical layers could further improve the performance of data dissemination schemes [36]. Furthermore, to model the transmission ability of vehicles thoroughly in theory will be a challenging but significant attempt in the future.
Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

Acknowledgments

The authors gratefully acknowledge the support from the National Natural Science Foundation of China (61502320 and 61373161), Science & Technology Project of Beijing Municipal Commission of Education in China (KM201410028015), Youth Backbone Project of Beijing Outstanding Talent Training Project (2014000020124G133), and Cultivation Object of Young Yanjing Scholar of Capital Normal University.

References


