

Research Article

DDGS: A Network Coding Scheme for Dynamic Adaptation to Heterogeneous Vehicular Networks

Zongzheng Wang , Ping Dong , Tao Zheng , and Hongke Zhang 

School of Electronic and Information Engineering, Beijing Jiaotong University, China

Correspondence should be addressed to Ping Dong; pdong@bjtu.edu.cn

Received 27 August 2020; Revised 19 October 2020; Accepted 15 April 2021; Published 7 May 2021

Academic Editor: Alessandro Bazzi

Copyright © 2021 Zongzheng Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The rapid development of the transportation industry has brought about the demand for massive data transmission. In order to make use of a large number of heterogeneous network resources in vehicular network, the research of applying network coding to multipath transmission has become a hot topic. Network coding can better solve the problems of packet reordering and low aggregation efficiency. The determination of coding scale is the key to network coding scheme. However, the existing research cannot adapt to the different characteristics of network resources in vehicular network, leading to larger decoding time cost and lower bandwidth aggregation efficiency. In this paper, we propose a network coding scheme called Delay Determined Group Size (DDGS), which could adaptively adjust the coding group according to the heterogeneous wireless networks state. The mathematical analysis and process design of the DDGS scheme are discussed in detail. Through a large number of simulations, we proved that the DDGS scheme is significantly superior to other coding group determination schemes in terms of decoding time cost and bandwidth aggregation efficiency.

1. Introduction

With a new wave of urbanization, the transportation industry is developing rapidly. A great quantity of vehicle services and applications emerges, creating a large number of application data to be transmitted to the ground [1]. As the development of wireless network technology, network device with multiple access interfaces has the ability to access heterogeneous wireless networks [2] (e.g., LTE, 5G, and satellite link) [3], which provide sufficient wireless channel resources for growing demands for high bandwidth vehicular applications [4] (e.g., HD video, online games, and live online). Based on the development of access technology, the multipath transmission scheme which uses multiple wireless interfaces simultaneously is widely studied in vehicle-to-ground communication [5, 6]. The aggregation of heterogeneous network resources is realized by distributing data on heterogeneous networks. In order to overcome the disorder of data packets caused by the heterogeneity of network resources in multipath transmission, scholars apply network coding [7, 8] to data scheduling process.

The encoding will encrypt [9] the data to a certain extent, and it brings a certain degree of security [10]. The core of network coding is to group packets and add redundancy. By encoding packets, the packets of the same group are equivalent to the decryption process. This feature eliminates the necessity for packets to arrive in strict order. To a certain extent, the reduction of communication efficiency caused by packet loss is also avoided. However, the addition of redundancy will cause the loss of bandwidth resources. Network coding is essentially a trade-off between link bandwidth and reliability. In order to efficiently utilize network coding schemes, the design of the coding scale is particularly important. We use group size (GS) to describe the number of data packets contained in a coding group. Designing appropriate GS according to the network status is the focus of this article.

The design of GS is related to the coding and decoding complexity, reception delay, bandwidth resource utilization, etc. The well-designed GS should be combined with the real-time state of the link to reasonably divide the data stream. Excessive GS design leads to greater computational complexity and delay, which affects the performance of

end-to-end transport protocol (TCP) [11]. Too small GS design will cause coding redundancy and reduce bandwidth utilization. The current research on the determination of group size has certain limitations in both method and theory. Most of the existing research on group size determination ignores the extremely different path characteristics in heterogeneous wireless networks. This is a prominent feature of vehicle-to-ground communication systems and has an important impact on the performance of network coding.

In order to make up for the deficiencies of current research, this paper proposes a Delay Determined Group Size (DDGS) scheme, which is aimed at achieving efficient bandwidth aggregation of heterogeneous wireless networks in vehicle-to-ground communication. This scheme can realize the dynamic adjustment of GS according to the real-time multipath characteristics. We treat the same round of data packets arriving at the receiving end as a batch [12, 13], which can be calculated as the difference in the number of packets sent between the fastest path and the slowest path in the transmission delay. In DDGS, the same batch of transmitted data packets is divided into fine-grained data packets with time scales on different links. A theoretically reasonable time segment division scheme can ensure a reasonable size of GS and achieve a reasonable decoding delay, thereby ensuring the continuity of the data stream at the receiving end.

At the sending end, before each group of data packets is encoded, the DDGS scheme dynamically senses the multilink delay difference [14] and converts the transmission data packets of each link within time Δt into a group. Then, the data packets are encoded by any k -row linearly independent encoding matrix, and k refers to the number of packets in the group. We define the coding rate k/n according to the packet loss rate of each link and send redundant packets on different links to compensate for the loss characteristics of the link. At the receiving end, we use the equivalent features of the same set of encoded data to decode, and the original data can be restored after receiving any k encoded data packets. We carry out the agreement of the coding matrix before data transmission at both ends of the communication and realize the information interaction required for decoding [15] through the header design containing the coding information of k and n . The vehicle network scenario is a typical heterogeneous wireless network scenario. In the vehicle network, the heterogeneous characteristics of various wireless networks have caused great differences in impact. DDGS scheme well supports the change in link delay difference caused by the heterogeneity and time-varying characteristics of wireless channels and compensates for the high reception and decoding delay and low bandwidth aggregation efficiency caused by the link delay difference.

Our contributions can be summarized as follows. First, we conducted a mathematical analysis on the formulation of coding groups in network coding. Second, we proposed a DDGS scheme based on mathematical analysis in heterogeneous network scenarios and designed the workflow of the scheme. DDGS scheme could achieve a smaller decoding time cost and higher bandwidth aggregation efficiency compared with the current research. Third, we did a large number of simulations to verify the superior per-

formance of DDGS in the heterogeneous network scenario of vehicular network.

The structure of the paper is organized as follows. We discuss the related work and study the limitation of the existing research status of coding group size and coding rate in Section 2. In Section 3, we introduce the system architecture of DDGS. In Section 4, we analyze and introduce the DDGS scheme in detail. In Section 5, we take lots of tests and simulations to prove that our scheme is effective and reliable. In Section 6, we conclude the paper.

2. Related Work

In order to cope with the development of mobile transportation and the consequent demand for massive data communication [16], scholars focus on the research of the multipath transmission architecture and corresponding transmission schemes of vehicular networks [17]. In this section, we summarize the relevant research on multipath transmission in vehicular networks. This section will expand from two aspects. The first part is the research on the related protocols and architecture of multipath transmission. The second part is about the research progress of improving the packet loss and disorder in multipath transmission. Part of the main research-related direction is network coding.

2.1. Research on Protocol and Architecture of Multipath Transmission

2.1.1. Multipath Transmission Protocol. Research on multipath transmission protocols first focused on the middle layer of the network layer and the transport layer, such as HIP [18] and Shim6 [19]. The idea of these protocols is to introduce a mapping between the network layer and the transport layer to achieve network layer multipath implementation for upper-layer applications. However, due to the transparency of the transport layer, out-of-sequence packets caused by path differences will seriously affect throughput. Therefore, academia turned to the study of transport layer protocols. Multipath transmission support based on SCTP and TCP is the two main research directions.

SCTP was first specified in RFC2960 and redefined in RFC4960. Supporting multihosts is one of the most important features of the standard SCTP protocol, which can be used to establish multiple paths between two multihosts. However, only one path is allowed for data transmission, and the other paths can only be used as backup paths. There are many studies on the realization of multipath parallel transmission through modification. For example, Iyengar et al. [20] proposed the CMT-SCTP protocol, whose main idea is SFR (Split Fast Rete) scheme to manage each path separately, thereby realizing multipath transmission. Some other researches like WiMPSCPT [21] and cmpSCTP [22] also use similar ideas. However, the asymmetry of the path can easily cause disorder, and the resulting unnecessary retransmission and reduction of the congestion window are difficult to solve.

There is also a large amount of research foundation about the expansion of tcp multipath support. A series of TCPbased multipath transmission protocols are proposed, pTCP [23]

divides the transport layer function into central engine and TCP-v pipe, realizing multipath management of data packets through double sequence. mTCP [24] introduces virtual queues, and each end-to-end path uses subflows with independent congestion control mechanisms. Early previous research provided the basis for MPTCP, and MPTCP divides the transmission layer into MPTCP and TCP layers. The MPTCP layer is responsible for path management, packet scheduling, and congestion control. The TCP layer transmits data in the form of subflows and uses the sliding window mechanism for congestion control [25, 26]. To solve the problem that MPTCP throughput will be significantly reduced in the case of path heterogeneity. Schemes like ECF [27], BLEST [28], and BBP [29] are proposed. The main idea of these studies is to study the scheduling strategy of the sender and congestion control. Regardless of the middle layer protocol, SCTP, or TCP-related research, they all need terminal equipment to support the protocol stack, so it is difficult to deploy under existing conditions. Moreover, most of their application environments are static environments, and their performance is poor in the mobile heterogeneous environment of the vehicular networks. In contrast, our method is independent of the protocol stack and transparent to the transport-layer protocols.

2.1.2. Multipath Transmission Architecture. Research on multipath transmission architecture mainly focused on the network layer. There are two main ways to implement the architecture, NAT, and IP-in-IP encapsulation. Rodriguez et al. [30] introduced the MAR system, which acts as a NAT box. MAR can work in two modes: with or without the proxy server at the receiving end. When there is a proxy server, the mar strategy uses a packet-based scheduling scheme. In the opposite case, the scheduling scheme is flow-based.

Dong et al. [31] proposed a multipath network architecture for vehicular networks. It uses tunneling mechanisms through IP-in-IP encapsulation between two proxies. This method performs multipath scheduling of data packets based on the historical status data of wireless heterogeneous network links, thereby achieving bandwidth aggregation. This method has a poor response to packet loss on unreliable wireless links [32]. In this paper, we improved this architecture by introducing network coding to enhance the robustness of the architecture, and at the same time, designing dynamic coding schemes to improve communication performance.

2.2. Research on Network Coding in Multipath Transmission. Due to the mobile nature of vehicular networks, the status of wireless heterogeneous link resources is dynamically changing. The difference of paths will cause the inevitable out-of-sequence problem of multipath transmission. In order to solve this problem, scholars have introduced network coding strategies.

The general process of network coding is that at the sending end, the group size (GS) is appropriately formulated, and the data packets are grouped accordingly. The grouped data packets are redundantly coded through the coding matrix according to the coding strategy requirements. Then, the

encoded packets are sent through multiple links to the receiving end. After the receiver receives the same number of coded packets as the original packets, the receiver recovers the original packets by solving the linear equation through Gaussian elimination according to the extracted coding information. Through the analysis of the above process, we can conclude that GS parameters have an important impact on the efficiency of network coding to a certain extent. At present, the research on GS determination can be divided into three types: (1) fixed group size and coding rate, (2) feedback-based group size, (3) GS determination based on network state. [33, 34] adopt fixed group size and coding rate. These strategies cannot adapt to the dynamic changes of wireless heterogeneous network resources in vehicular networks, and it is difficult to achieve the best communication performance. [35–37] determines the group size based on the end-to-end feedback results. The disadvantage of feedback based is that it will introduce additional delay, which does not meet the low delay requirements of vehicular networks. The decision of the GS based on the real-time network parameters can make the network code better meet the needs of multipath transmission. MPTCP-PNC [13] and CMT-NC [12] fully consider the real-time status of the network, especially the difference in link delay. By using the same batch of packets as a group, packets of a group can arrive at the receiver within the same time round and avoid out-of-order packets between the two groups. These policies treat the same batch of packets as a group, and the packets of a group can arrive at the receiver within the same time round, thus, avoiding out-of-order packets between two groups. However, if only the same batch of data packets is used as a group, under the circumstances caused by the huge performance difference between heterogeneous networks, it will cause an excessively large group size. It greatly increases the computational complexity and decoding time of the receiver. We will perform mathematical analysis in Section 3.2 for further explanation. In our DDGS strategy, we divide the same batch of data packets into fine-grained groups for large delays. In our DDGS strategy, we divide the same batch of data packets into fine-grained groupings for large differences in link delays, thus, avoiding the above problems.

3. The System Architecture of DDGS

In this section, we will introduce the system architecture of DDGS from two aspects: network topology and system module. In order to facilitate the understanding of the formulas and diagrams in the article, we have summarized the symbols appearing in this article in Table 1.

3.1. The Network Topology of Multipath Transmission. The Delay Determined Group Size (DDGS) scheme we proposed is designed for the high bandwidth requirements in the vehicle-to-ground transmission scenario. In order to support multiple communication terminals on the vehicle and reduce the improvement of the protocol stack, we adopt the multipath transmission architecture of the network layer, as shown in Figure 1.

TABLE 1: Summary of symbols.

Symbols	Description
S	Link set between the sender and the receiver
i	Sequence number of a link
l	The link with the largest delay
m	The link with the minimum delay
b	Sequence number of a batch
N_b	Number of packets arriving in the same batch b
BW_i	Bandwidth of link i
RTT_i	Round-trip delay of link i
pe_i	Packet loss rate of link i
T_i	Delay of single packet on link i
q	Sequence number of a group
GSq	Group size of the group q
Mq	Number of redundant packets appended to group q
CMq	Coding matrix for group q
DMq	Decoding matrix for group q

At the sending end, the mobile router completes the aggregation of packets generated by device onboard and realizes grouping, coding, and distributing of packets. At the receiver side, the aggregation router decodes packets from heterogeneous networks and forward them to the server.

We use multiaccess router (MAR) and aggregation router (AR) as the two ends of multipath communication proxy on the vehicle and on the ground, respectively. As the gateway, the multiaccess router is responsible for aggregating the application traffic of users on board. The multiaccess router has multiple network interfaces and can access multiple heterogeneous channels. Traffic is transmitted in parallel to heterogeneous channels and converges at the aggregation router. Our DDGS scheme deploys in MAR and AR. At the MAR, DDGS dynamically groups data packets based on channel state parameters. Each group of data packets is encoded according to the packet loss rate of each channel and sent to the heterogeneous network. At the receiver side, the aggregation router decodes packets from heterogeneous networks and forward them to the server.

3.2. System Module of DDGS. In order to clearly describe the workflow of DDGS, we abstract the communication process as Figure 2. Sending end mainly consists of four modules, network measurement module, group size determination module, network coder, and scheduling module. The network measurement module is responsible for monitoring the status parameters of each link, including RTT, bandwidth, and packet loss rate. The RTT and packet loss rate are obtained by a modified “ping” program. We get these state data by sending probe packets regularly. The bandwidth is passively measured by calculating the payload combined with the probe packet interval. The group size determination module determines the size of the encoding group according to the real-time network status parameters, and the specific

calculation process will be introduced in Section 4. The general grouping process is shown in Figure 2. The DDGS algorithm divides the application layer packets into different batches according to the real-time status of the network, which is represented by packets of different colors in Figure 2. On this basis, the packets are divided into fixed time slices according to the link delay difference. In this way, the data packets in transmission can be divided into boxes as shown in the figure. The network coder dynamically selects the encoding matrix according to the group size and coding rate, and the encoded data packets are forwarded to each link through the scheduling module. At the receiving end, the encoded packets enter the receiving buffer. Decoder designs a decoding matrix to decode by extracting the encoding information of the packet header.

The encoding and decoding information interaction will be introduced in Section 5.

4. The Analysis and Principle of DDGS Scheme

When packets are transmitted through multiple paths, the problem of out-of-sequence of data packets caused by link delay is particularly prominent. Network coding makes the same group of data packets equivalent to decoding, so that the same group of data packets does not have to be in strict order when arriving at the receiving end. However, if the determination of the encoding group is unreasonable, when data packets belonging to different groups arrive, it will cause the buffer of the receiving end to be blocked. In the heterogeneous multilink scenario, there is a delay difference between links. During the link delay difference, the low-latency link will continue to receive data packets before the data packet reaches the receiving end through the high-latency link. We regard the packets sent in the transfer delay difference between the fastest path and the slowest path as the same batch b . [12, 13] use the number of packets in the same group as the code group size which can be expressed as N_b . In our paper, we added consideration of the influence of link bandwidth on N_b , and N_b can be mathematically formulated as.

$$N_b = 1 + \sum_{i \in S} \left\{ \frac{BW_i * (1 - pe_i)}{MTU} * \left[\left(\frac{RTT_l}{2} + \frac{MTU}{BW_l} \right) - \left(\frac{RTT_i}{2} + \frac{MTU}{BW_i} \right) \right] + 1 \right\}. \quad (1)$$

BW_i , RTT_i , and pe_i are measured in real-time by the network measurement module. $RTT_l/2 + MTU/BW_l$ represents the transmission delay of the packet on the path with the maximum delay. $RTT_i/2 + MTU/BW_i$ represents the transmission delay of the packet on the path i . $BW_i * (1 - pe_i)/MTU$ represents the number of transmitted packets per unit time of path i . Through the correction of the calculation, N_b is obtained. Taking N_b as the group size can prevent the data packets of different groups from arriving out of order to a certain extent. However, there is a big difference in the delay of vehicle heterogeneous links. This situation will lead to a large N_b value. We can calculate $N_b = 56$ through the test data in Table 2.

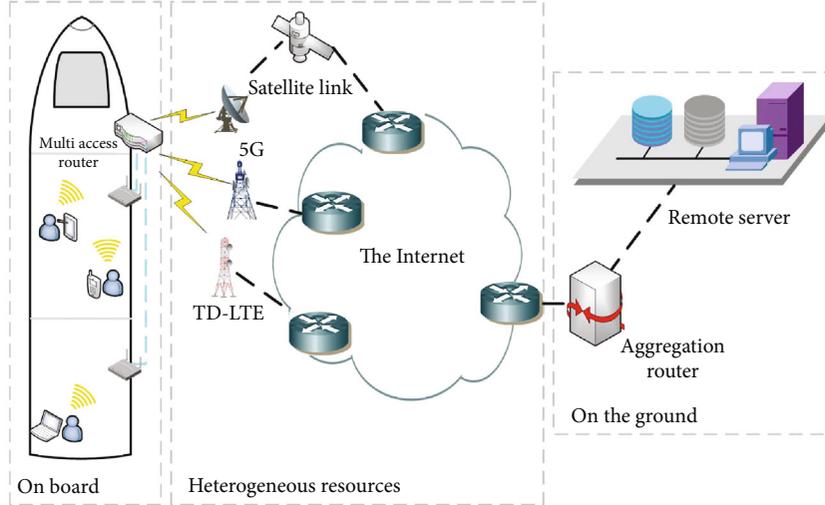


FIGURE 1: Multipath transmission architecture.

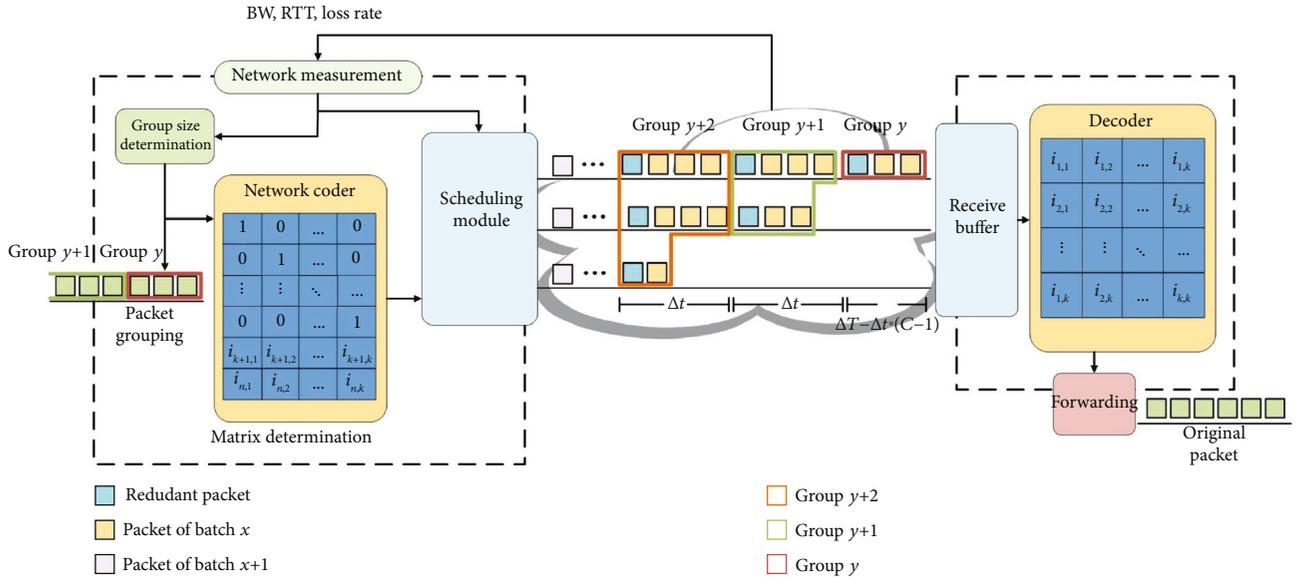


FIGURE 2: Workflow of DDGS scheme.

TABLE 2: Link configurations in simulations.

Link type	Delay	Bandwidth	Loss rate
TD-LTE	60 ms	5 Mbps	2%
Satellite link	180 ms	2 Mbps	5%

Using N_b as the number of coding groups makes the data transmission delay of this group equal to the delay of the longest path. What is more, when N_b is too large, it will increase the computational complexity of the encoding and decoding. In order to remedy the above problems, we adopt the DDGS scheme which could dynamically divide batches based on the performance of heterogeneous network resources. At the sender side, the parameter N_b indicates the size of the same Figure 2. Workflow of DDGS scheme batch of packets is sent through the parallel link at the same time. Through our pro-

posed DDGS strategy, we divide the batch of packets into fine-grained ones to avoid the problem caused by too large N_b value. In the next part, we will introduce the specific algorithm of DDGS.

4.1. The Grouping Principle of DDGS. Combined with the real-time network state parameters measured by the network measurement module, we can calculate that the time consumption of one packet from sender to receiver on link i can be depicted as

$$T_i = \left(\frac{RTT_i}{2} + \frac{MTU}{BW_i} \right). \quad (2)$$

According to formula (2), we can get that the difference between the transmission delay of the data packet on link i

and the transmission delay on the path with the maximum delay is

$$\Delta T_i = \left(\frac{RTT_i}{2} + \frac{MTU}{BW_i} \right) - \left(\frac{RTT_i}{2} + \frac{MTU}{BW_i} \right). \quad (3)$$

According to Section 4 part A, the packets transmitted within ΔT_i on the link i can arrive at the receiving end in the same batch as the packets transmitted simultaneously on the path with the highest delay. In DDGS, we choose Δt to further divide ΔT_i . In the actual deployment, we make Δt equal to the interval time of delay ACK feature of TCP (usually 40 ms). The advantage of this choice is that it can make full use of the end-to-end TCP characteristics, so that the decoded packets of the same group can be acknowledged in one ack. By calculating the transmission delay difference between the maximum delay path and the minimum delay path, we can get the number of groups C_b that this batch of packets can be divided into.

$$C_i = \lceil \frac{((RTT_i/2) + (MTU/BW_i)) - ((RTT_i/2) + (MTU/BW_i))}{\Delta t} \rceil. \quad (4)$$

In formula (5), m refers to the link with minimum delay, l refers to the link with largest delay, and we use q to represent the group sequence, $q \in 1, 2, \dots, C_b$. For any link i , the distribution of data packets in the same batch can be expressed as a sequence number starting from C_i to ending at C_b .

$$C_i = \lceil \frac{((RTT_i/2) + (MTU/BW_i)) - ((RTT_i/2) + (MTU/BW_i))}{\Delta t} \rceil. \quad (5)$$

According to the group distribution analysis for link i , we can express the number of packets transmitted on link i with group sequence q as n_i^q .

$$n_i^q = \begin{cases} \frac{[\Delta T_i - \Delta t * (C_i - 1)] * BW_i}{MTU} & q = C_i, \\ \Delta T * \frac{BW_i}{MTU} & q < C_i. \end{cases} \quad (6)$$

Then the group size of q th group transmitted in all paths can be expressed as

$$GS_q = \sum_{i \in S} n_i^q. \quad (7)$$

Next, we introduce the packet encoding process based on group.

4.2. Dynamic Coding and Encoding Based on GS. Due to the mobile characteristics and heterogeneity of vehicular networks, the link will cause serious packet loss and out-of-order data packet transmission. We alleviate the two major problems by encoding ordinary data packets to generate redundant data packets. According to the packet loss rate of each link, we can conclude that the number of additional

redundant data packets required to ensure the completion of the decoding on each link is

$$r_i^q = \lceil \frac{pe_i}{1 - pe_i} * n_i^q \rceil. \quad (8)$$

Adding the redundant data packets on each link to obtain the number of redundant data packets required for the code group q is

$$M_q = \sum_{i \in S} r_i^q. \quad (9)$$

Based on the above analysis, we can calculate the coding ratio $r = GS_q / (GS_q + M_q)$ of the coding group q . We design a coding matrix that can be dynamically expanded so that the network coder module dynamically determines the coding matrix according to different values of r . The scalable coding matrix is designed as follows.

$$CM_q = \begin{bmatrix} I \\ RM \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \\ 1 & 2 & \dots & GS_q \\ GS_q & 1 & \dots & GS_{q-1} \\ \vdots & \vdots & \ddots & \vdots \\ 2 & 3 & \dots & 1 \end{bmatrix}. \quad (10)$$

The coding matrix CM_q for group q has $GS_q + M_q$ rows and GS_q columns. The first GS_q rows of CM_q is a GS_q order identity matrix I , and the last M_q rows is a special form of the Toeplitz matrix RM . [34] has proved that any GS_q rows of the constructed matrix are linearly independent. Encoding data packets by matrix CM_q ensures that the receiver can decode the original data packets after receiving any GS_q encoded packets. The construction prototype of coding matrix with different coding ratio r is negotiated by the sender and the receiver before communication. At the receiving end, after receiving any GS_q encoded packets, decoder module can trigger the decoding process. The decoding matrix DM_q for group q can be constructed by inverting the linear combination matrix of the received encoded packets.

Through the analysis of the above process, for realizing the interaction of decoding information, the encoded data packet only needs to carry two pieces of information, encoding ratio r and line number N which the encoding packet is linear combined through, then, the receiving end can realize the construction of the decoding matrix. We add these fields in the encoded packet header, as shown in Figure 3. The decoder module extracts the header information at the

receiving end and strips the header to decode. This design can effectively increase the payload of the data packet.

4.3. Packet Scheduling. According to formula (7), we can get the number of packets of group q transmitted on link i . Therefore, in the process of dispatching packets at the sending end, we can calculate the coding group of the batch of packets sent on each link in advance, so as to obtain the sequence of packets on the transmission path. The scheduling process is shown in Figure 2, in order to achieve the effect of the same batch of packets arriving at the receiving end at the same time, the data packets with larger coding group number are sent on the high delay link, and the coding groups are sent sequentially on the low delay link. The scheduling process is carried out according to the group number formulated by the DDGS scheme, which improves the receiving efficiency of the receiver. After completing a batch of data packet scheduling, in order to ensure real-time feedback of the link status, the scheduling module will trigger the update of GS_q and M_q , as well as coding matrix.

5. The Analysis and Principle of DDGS Scheme

In this section, the performance of DDGS is illustrated by simulation. We choose two main research methods to determine the coding group as a comparison. One is based on the single transmission process to determine the coding group size, abbreviated as BOST, and the other is the fixed coding rate, abbreviated as FIXED. For the research point of DDGS, we evaluate the impact of different network parameters on encoding and decoding delay and average throughput. We deploy the above three schemes on Network Simulation 3 (NS3) version 3.29 and dynamically adjust the link parameters for simulation.

5.1. Network Topology for Simulation. The simulation topology is shown in Figure 4. We use Node2 as MAR and Node3 as AR. We create three links between Node2 and Node3 to simulate a heterogeneous network scenario. The three links are used to simulate heterogeneous links with different delay, bandwidth, and packet loss rate in subsequent simulation. For different simulation objectives, we choose the link parameters which have the greatest impact on the proposed scheme as variables to verify the performance of the strategy. We create a virtual device tun for DDGS deployment. We design from a client-server program from Node1 to Node2 to perform one-way communication simulation. At Node2, data packets are grouped, encrypted, and forwarded to three links. At Node3, the received data packets are cached, decrypted, and forwarded to node4. The pseudocode on the sender and receiver is shown below.

5.2. Time Consumption of Decoding. In this section, we analyze the decoding time cost of the DDGS scheme. The design purpose of DDGS is to select a reasonable coding group size for link delay the difference in heterogeneous networks. Therefore, for the analysis of decoding time cost, we take link delay as an independent variable, and the delay settings of each link are shown in Table 3. By dynamically adjusting

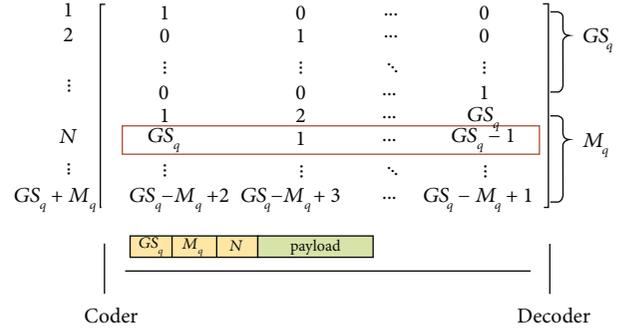


FIGURE 3: Decoding information interaction process.

the delay of different links, the performance of DDGS under different link delay combinations is displayed in Figure 5. We fix the delay of link A to 20 ms and adjust the delay of link B and C to achieve different link delay combinations. The abscissa represents the link delay of link C. The ordinates are the time required for the receiver to complete the decoding of the whole group of packets from the first packet of this group arriving in.

Through the analysis of Figure 5, we can see that when the three links have no delay difference or the delay difference is small, and the decoding time cost of the three schemes is similar. With the increase of delay difference, the time cost of the BOST scheme increases linearly due to the increase of the coding group, while the time cost of FIXED and DDGS schemes increases less. With the increase of the delay difference, the BOST scheme will lead to the excessive growth of the coding group, thus, increasing the waiting delay and decoding complexity of the receiver, resulting in the linear growth of the overhead. Although the DDGS scheme will adjust the coding group due to the increase of delay difference, the scheme ensures that the coding group will not grow excessively. The time cost is stable to a certain extent. Keeping the decoding cost time small is beneficial to reduce the end-to-end communication delay, which makes TCP communication avoid performance degradation caused by delay fluctuation. Although the FIXED scheme is slightly better than DDGS in decoding time cost, through subsequent analysis, we can see that DDGS can achieve higher throughput. According to the above analysis of simulation results, we can conclude that DDGS is suitable for delay fluctuation scenarios in heterogeneous networks.

5.3. Bandwidth Aggregation Efficiency. In this section, we will analyze the bandwidth aggregation performance of the DDGS scheme. First, we analyze the bandwidth aggregation performance of DDGS from the perspective of heterogeneous network delay. We set up two simulation environments, as shown in Table 4. In the first simulation, we set the three link delays to 10-30 ms fluctuations. This parameter setting represents the scenario with little difference in link delay for comparison. In the second simulation, the delay difference of the three links increases. This parameter setting indicates the heterogeneous network scenario with gradient delay difference, which is also the main scenario for DDGS. The simulation results are shown in Figure 6. The horizontal axis

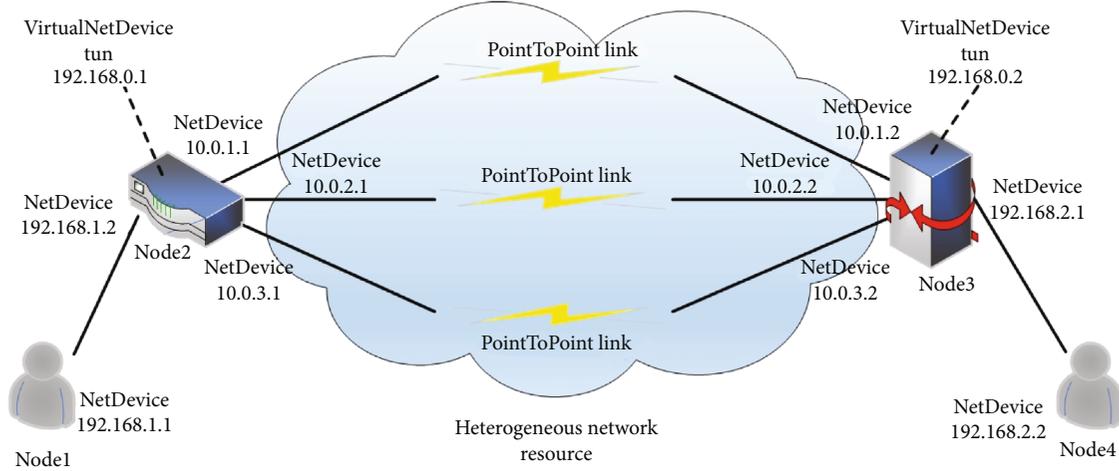


FIGURE 4: Simulation topology.

```

MAR sender.
1: for RecvPacket() do
2:   if batch.pendingqueue=NULL then
3:     batch.pendingqueue=DetermineGS()
4:   end if
5:   Codingbuffer.q.queue=Grouping()
6:   if sizeof(Codingbuffer.q.queue)=Codingbuffer.q.k then
7:     ConstructCodingMartix()
8:     Encoding()
9:     Scheduling()
10:  end if
11: end for

```

ALGORITHM 1

```

AR receiver.
1: for RecvPacket() do
2:   HeaderAnalysis()
3:   recvbuffer.q.queue=Add2RecvBuffer()
4:   if sizeof(recvbuffer.q.queue)=recvbuffer.q.k then
5:     ConstructDecodingMartix()
6:     Decoding()
7:     Forwarding()
8:   end if
9: end for

```

ALGORITHM 2

TABLE 3: Link configurations for decoding time cost.

Parameters	Link A	Link B	Link C
Delay	20 ms	20, 50, 80 ms	20, 50, 80, 110, 140 ms
Loss rate	0.05	0.05	0.05
Bandwidth	1 Mbps	1 Mbps	1 Mbps

represents simulation time, and the vertical axis represents real-time bandwidth statistics. It can be concluded that when the delay difference is small, the bandwidth aggregation performance of the three schemes is similar. When the link delay difference expands, FIXED and BOST have experienced significant performance degradation. In comparison, DDGS maintains better bandwidth aggregation performance.

Next, we will study the impact of link bandwidth on DDGS schemes. We set the delays of the three links to 20 ms, 80 ms, and 140 ms, respectively, to simulate the heterogeneous network scenario with large difference in link delays. We fixed the bandwidth of two links at 1 Mbps and adjusted the bandwidth of the other link from 0.1 Mbps to 2 Mbps to observe the impact of bandwidth changes of links with different delays on the performance of DDGS.

The simulation results are shown in Figure 7. From Figures 7(a) and 7(b), we can see that the aggregate bandwidth increases linearly with the increase of the bandwidth of the two links with a small delay. DDGS and BOST have a better convergence effect than the FIXED scheme (the slope of simulation results is larger than FIXED scheme), which is consistent with our assumption, because the FIXED scheme does not have the ability to adjust the size of the coding group according to the link delay difference, which will lead to the code group not suitable for the link state, resulting in excessive redundancy or coding groups out of order. The growth rate of Figure 7(a) is slightly larger than that of Figure 7(b). It can be seen that low delay link and high bandwidth link are more conducive to the overall performance improvement.

From Figure 7, we can see that the overall performance of DDGS is better than the other two schemes. The bandwidth aggregation performance of DDGS and BOST is relatively close. The difference between DDGS and BOST lies in the segmentation of the same batch of packets. The difference in bandwidth aggregation between DDGS and BOST is due to the aggregate bandwidth reduction caused by the receiving delay experienced by the receiver. Therefore, the bandwidth aggregation performance of the DDGS scheme is slightly better than that of the BOST scheme, but according to the

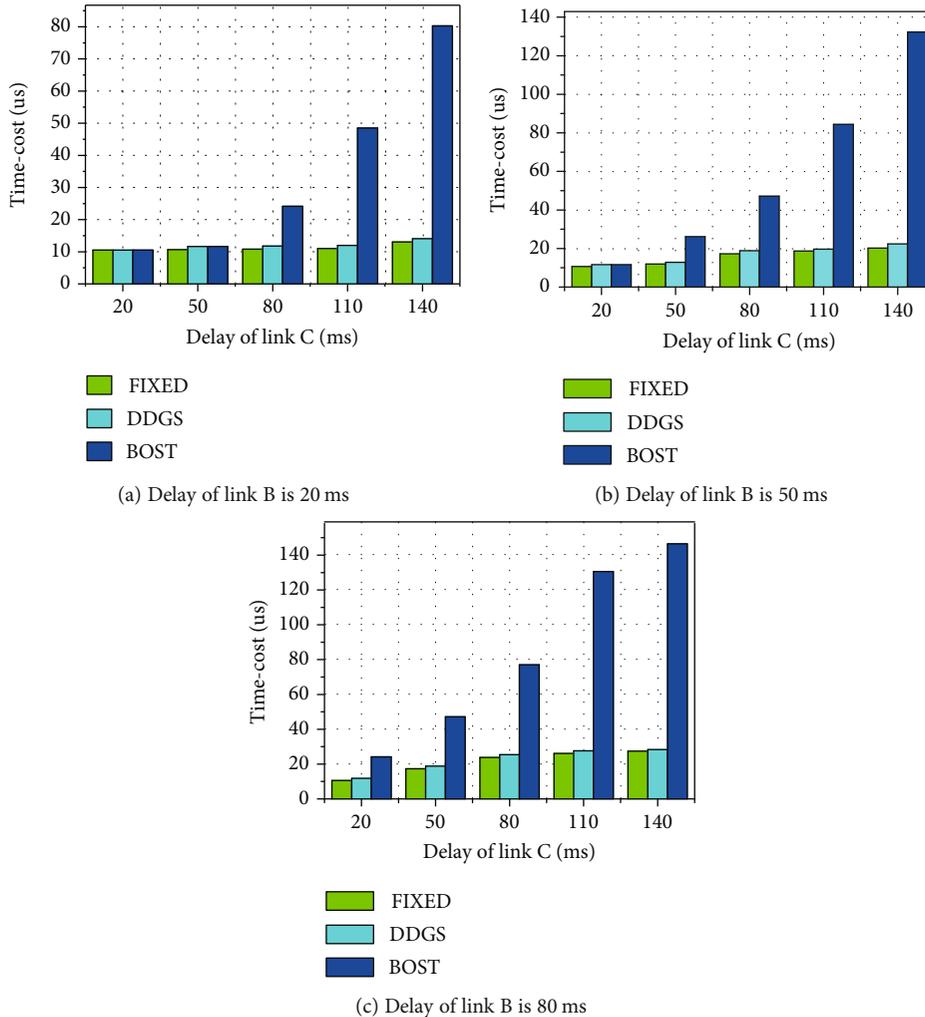


FIGURE 5: The time-cost of decoding processes at receiver for different schemes.

TABLE 4: Link configurations for bandwidth aggregation.

Parameters	Link A	Link B	Link C
Loss rate	0~0.03	0~0.03	0~0.03
Bandwidth	0.5~1.5 Mbps	0.5~1.5 Mbps	0.5~1.5 Mbps
Delay of simulation 1	10~30 ms	10~30 ms	10~30 ms
Delay of simulation 2	10~30 ms	70~90 ms	130~150 ms

analysis in Figure 5, DDGS is much better than BOST in receiving delay.

It is worth noting that in Figure 7(c), DDGS and BOST have a high-efficiency bandwidth aggregation effect when the bandwidth of the high-latency link *c* is small. When the bandwidth of the link *c* increases to a certain value, the growth rate of the aggregate bandwidth decreases significantly. The reason for this is that the bandwidth of the low-latency link is insufficient to support the timely transmission of data packets on the low-latency link, resulting in queuing. The results show that the DDGS scheme has a better aggregation effect on low-latency, high-bandwidth links and high-latency, low-bandwidth links.

Next, we will study the impact of link loss rate on DDGS schemes. We maintain the link delay in the above simulation and set the bandwidth of the three links to 1 Mbps. We adjusted the packet loss rate of the three links from 0% to 20%.

From Figure 8, we can see that when the packet loss rate is less than 4%, the three schemes can achieve effective bandwidth aggregation. When the packet loss rate is greater than 4%, the bandwidth aggregation efficiency will decrease significantly. But DDGS can still guarantee a better aggregation effect than the other two schemes. When the link packet loss rate is less than 4%, the FIXED scheme uses a smaller group size, and redundant packets encoded to compensate for packet loss will greatly occupy effective bandwidth, resulting in a waste of bandwidth resources. When the packet loss rate increases, the performance of the BOST decreases more significantly than the other two schemes. This is because when the link delay difference is large, BOST will cause a large group size. In order to compensate for the link loss, each group of data needs to wait for more encoding redundant packets to arrive, resulting in a drop in throughput.

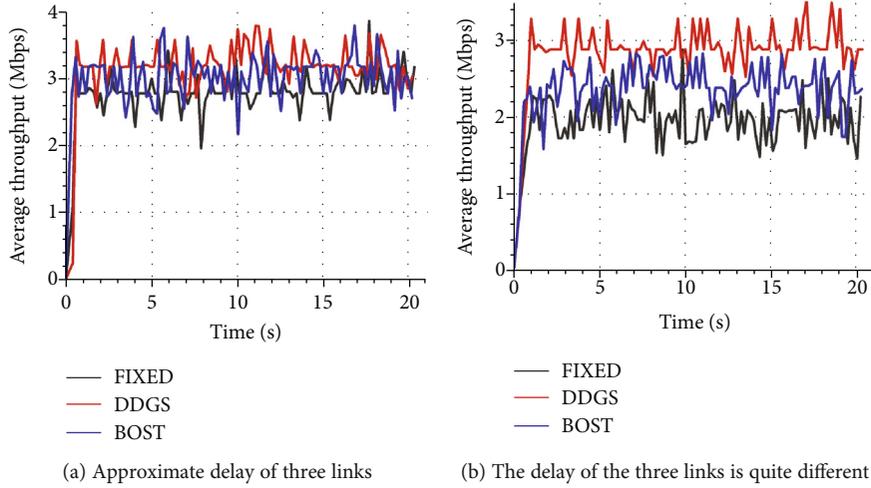


FIGURE 6: Bandwidth aggregation under different delay conditions.

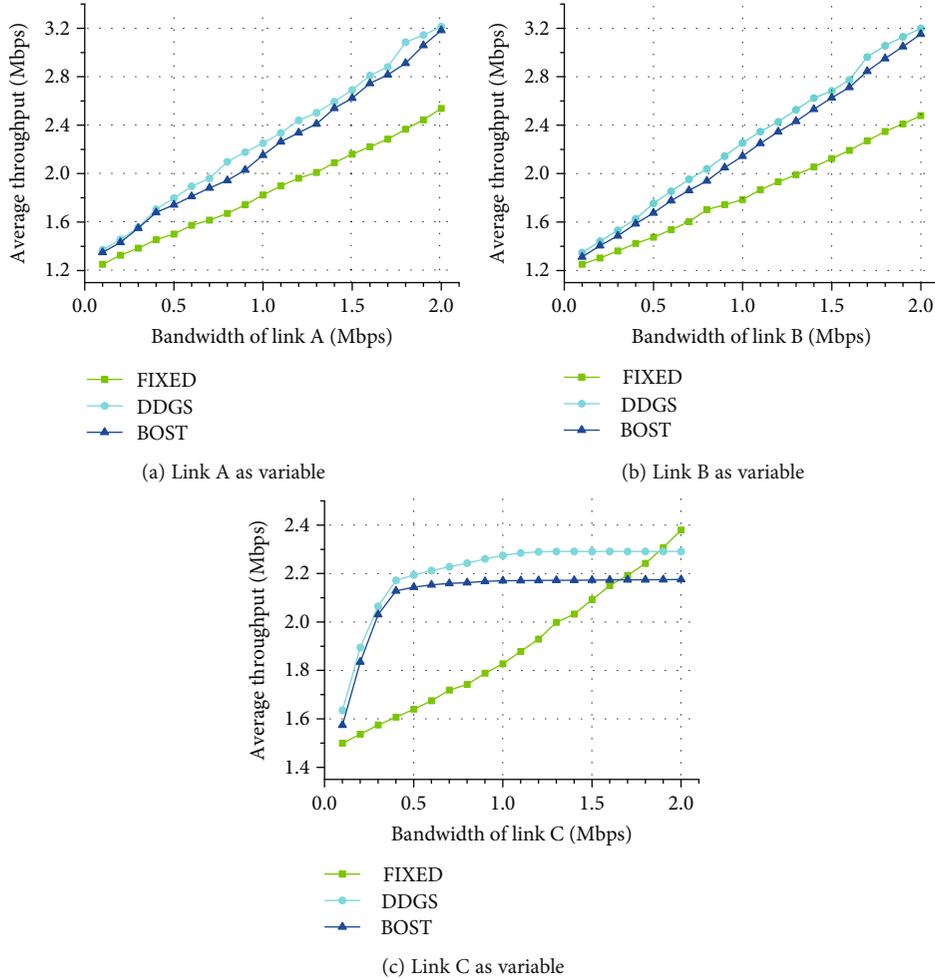


FIGURE 7: Bandwidth aggregation efficiency with link bandwidth fluctuation.

The following conclusions can be drawn from the above simulation results. (1) The DDGS scheme can guarantee a shorter packet decoding time than the BOST scheme, thereby ensuring data continuity and TCP friendliness. (2) The DDGS solution also has good bandwidth

aggregation performance when the link delay difference is large. What is more, DDGS has a better aggregation effect for heterogeneous networks with low latency, high bandwidth links, high latency, and low bandwidth links in the vehicular network.

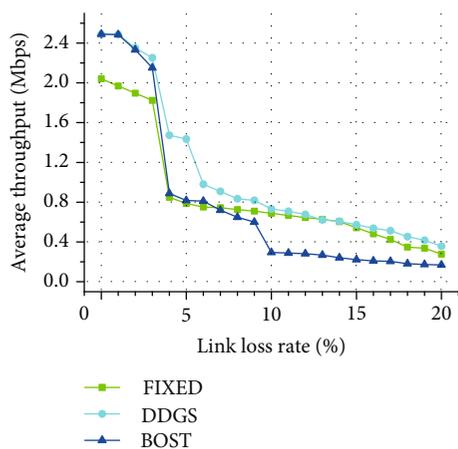


FIGURE 8: Bandwidth aggregation efficiency with loss rate fluctuation.

6. Conclusion

In this paper, aiming at solving the low efficiency of multipath bandwidth aggregation caused by the heterogeneous characteristics of network resources in the vehicular network, we propose a DDGS scheme based on network coding. The scheme we proposed could adaptively adjust the coding group according to the heterogeneous wireless networks state, so as to make full use of the advantages of network coding. We discussed in detail the mathematical analysis and process design of the DDGS scheme. Finally, we did a lot of simulations, comprehensively considering a variety of network parameters to analyze DDGS. The results show that the DDGS scheme is significantly superior to other coding group determination schemes. DDGS is very suitable for heterogeneous networks with low latency, high bandwidth links, high latency, and low bandwidth links in vehicle networks.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grant 61872029 and in part by the Beijing Municipal Natural Science Foundation under Grant 4182048.

References

- [1] P. Dong, T. Zheng, X. Du, H. Zhang, and M. Guizani, "SVCC-HSR: providing secure vehicular cloud computing for intelligent high-speed rail," *IEEE Network*, vol. 32, no. 3, pp. 64–71, 2018.
- [2] G. Liang, H. Yu, X. Guo, and Y. Qin, "Joint access selection and bandwidth allocation algorithm supporting user requirements and preferences in heterogeneous wireless networks," *IEEE Access*, vol. 7, pp. 23914–23929, 2019.
- [3] J. Wu, B. Cheng, and M. Wang, "Improving multipath video transmission with raptor codes in heterogeneous wireless networks," *IEEE Transactions on Multimedia*, vol. 20, no. 2, pp. 457–472, 2018.
- [4] X. Du, "QoS routing based on multi-class nodes for mobile ad hoc networks," *Ad Hoc Networks*, vol. 2, no. 3, pp. 241–254, 2004.
- [5] L. Li, K. Xu, T. Li et al., "A measurement study on multi-path tcp with multiple cellular carriers on high speed rails," in *Proceedings of the 2018 Conference of the ACM Special Interest Group on Data Communication*, pp. 161–175, Budapest, Hungary, August 2018.
- [6] Y. Zhang, P. Dong, S. Yu, H. Luo, T. Zheng, and H. Zhang, "An adaptive multipath algorithm to overcome the unpredictability of heterogeneous wireless networks for high-speed railway," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 12, pp. 11332–11344, 2018.
- [7] R. Ahlswede, Ning Cai, S. Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Transactions on Information Theory*, vol. 46, no. 4, pp. 1204–1216, 2000.
- [8] Z. Ding, K. K. Leung, D. L. Goeckel, and D. Towsley, "On the study of network coding with diversity," *IEEE Transactions on Wireless Communications*, vol. 8, no. 3, pp. 1247–1259, 2009.
- [9] Z. Guan, X. Liu, L. Wu et al., "Cross-lingual multi-keyword rank search with semantic extension over encrypted data," *Information Sciences*, vol. 514, pp. 523–540, 2020.
- [10] L. Zhu, Y. Li, F. R. Yu, B. Ning, T. Tang, and X. Wang, "Cross-layer defense methods for jamming-resistant cbtc systems," *IEEE Transactions on Intelligent Transportation Systems*, vol. 99, pp. 1–13, 2020.
- [11] J. K. Sundararajan, D. Shah, M. Médard, S. Jakubczak, M. Mitzenmacher, and J. Barros, "Network coding meets tcp: theory and implementation," *Proceedings of the IEEE*, vol. 99, no. 3, pp. 490–512, 2011.
- [12] C. Xu, Z. Li, L. Zhong, H. Zhang, and G. M. Muntean, "Cmtnc: improving the concurrent multipath transfer performance using network coding in wireless networks," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 3, pp. 1735–1751, 2016.
- [13] C. Xu, P. Wang, C. Xiong, X. Wei, and G. M. Muntean, "Pipeline network coding-based multipath data transfer in heterogeneous wireless networks," *IEEE Transactions on Broadcasting*, vol. 63, no. 2, pp. 376–390, 2017.
- [14] L. Zhu, Y. He, F. R. Yu, B. Ning, T. Tang, and N. Zhao, "Communication-based train control system performance optimization using deep reinforcement learning," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 12, pp. 10705–10717, 2017.
- [15] L. Xue, Y. Yu, Y. Li, M. H. Au, X. du, and B. Yang, "Efficient attribute-based encryption with attribute revocation for assured data deletion," *Information Sciences*, vol. 479, pp. 640–650, 2019.
- [16] L. Zhu, F. R. Yu, B. Ning, and T. Tang, "Cross-layer handoff design in mimo-enabled wlans for communication-based train control (cbtc) systems," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 4, pp. 719–728, 2012.

- [17] M. Li, A. Lukyanenko, Z. Ou et al., "Multipath transmission for the internet: a survey," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 4, pp. 2887–2925, 2016.
- [18] R. Moskowitz, P. Nikander, P. Jokela et al., *Host Identity Protocol (hip) Architecture*, RFC 4423, Tech. Rep., 2006.
- [19] M. Bagnulo and E. Nordmark, *Shim6: Level 3 Multihoming Shim Protocol for ipv6*, RFC 5533, Tech. Rep., 2009.
- [20] J. R. Iyengar, P. D. Amer, and R. Stewart, "Concurrent multipath transfer using sctp multihoming over independent end-to-end paths," *IEEE/ACM Transactions on Networking*, vol. 14, no. 5, pp. 951–964, 2006.
- [21] C.-M. Huang and C.-H. Tsai, "Wimp-sctp: multi-path transmission using stream control transmission protocol (sctp) in wireless networks," in *21st International Conference on Advanced Information Networking and Applications Workshops (AINAW'07)*, vol. 1, pp. 209–214, Niagara Falls, ON, Canada, 2007.
- [22] J. Liao, J. Wang, and X. Zhu, "cmppsctp: an extension of sctp to support concurrent multi-path transfer," in *2008 IEEE International Conference on Communications*, pp. 5762–5766, Beijing, China, 2008.
- [23] H.-Y. Hsieh and R. Sivakumar, "A transport layer approach for achieving aggregate bandwidths on multi-homed mobile hosts," *Wireless Networks*, vol. 11, no. 1-2, pp. 99–114, 2005.
- [24] M. Zhang, J. Lai, A. Krishnamurthy, L. L. Peterson, and R. Y. Wang, "A transport layer approach for improving end-to-end performance and robustness using redundant paths," in *USENIX Annual Technical Conference, General Track*, pp. 99–112, Boston, MA, USA, 2004.
- [25] A. Ford, C. Raiciu, M. Handley, S. Barre, and J. Iyengar, "Architectural guidelines for multipath tcp development," *IETF, Informational RFC*, vol. 6182, article 2070–1721, 2011.
- [26] S. Barre, C. Paasch, and O. Bonaventure, "Multipath tcp: from theory to practice," in *International Conference on Research in Networking*, pp. 444–457, Springer, 2011.
- [27] Y.-s. Lim, E. M. Nahum, D. Towsley, and R. J. Gibbens, "Ecf: an mptcp path scheduler to manage heterogeneous paths," in *Proceedings of the 13th International Conference on emerging Networking EXperiments and Technologies*, pp. 147–159, Incheon, Republic of Korea, 2017.
- [28] S. Ferlin, O. Alay, O. Mehani, and R. Boreli, "Blest: blocking estimation-based mptcp scheduler for heterogeneous networks," in *2016 IFIP Networking Conference (IFIP Networking) and Workshops*, pp. 431–439, Vienna, Austria, May 2016.
- [29] B.-H. Oh and J. Lee, "Feedback-based path failure detection and buffer blocking protection for mptcp," *IEEE/ACM Transactions on Networking*, vol. 24, no. 6, pp. 3450–3461, 2016.
- [30] P. Rodriguez, R. Chakravorty, J. Chesterfield, I. Pratt, and S. Banerjee, "Mar: a commuter router infrastructure for the mobile internet," in *Proceedings of the 2nd international conference on Mobile systems, applications, and services - MobiSYS '04*, pp. 217–230, Boston, MA, USA, 2004.
- [31] P. Dong, B. Song, H. Zhang, and X. du, "Improving onboard internet services for high-speed vehicles by multipath transmission in heterogeneous wireless networks," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 12, pp. 9493–9507, 2016.
- [32] X. Du and F. Lin, "Improving sensor network performance by deploying mobile sensors," in *PCCC 2005. 24th IEEE International Performance, Computing, and Communications Conference, 2005*, pp. 67–71, Phoenix, AZ, USA, April 2005.
- [33] F. Gabriel, J. Acevedo, and F. H. P. Fitzek, "Network coding on wireless multipath for tactile internet with latency and resilience requirements," in *2018 IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6, Abu Dhabi, United Arab Emirates, December 2018.
- [34] Y. Zhang, P. Dong, X. du, H. Luo, T. Zheng, and M. Guizani, "BNNC: improving performance of multipath transmission in heterogeneous vehicular networks," *IEEE Access*, vol. 7, pp. 158113–158125, 2019.
- [35] Y. Cui, L. Wang, X. Wang, H. Wang, and Y. Wang, "Fmtcp: a fountain code-based multipath transmission control protocol," *IEEE/ACM Transactions on Networking*, vol. 23, no. 2, pp. 465–478, 2014.
- [36] N. Arianpoo, I. Aydin, and V. C. M. Leung, "Network coding: a remedy for receiver buffer blocking in the concurrent multipath transfer of data over multi-hop wireless networks," in *2014 IEEE International Conference on Communications (ICC)*, pp. 3258–3263, Sydney, NSW, Australia, June 2014.
- [37] J. Cloud and M. Medard, "Multi-path low delay network codes," in *2016 IEEE Global Communications Conference (GLOBECOM)*, pp. 1–7, Washington, DC, USA, December 2016.