Hindawi
Security and Communication Networks
Volume 2022, Article ID 6546004, 13 pages
https://doi.org/10.1155/2022/6546004

Research Article

IoV-SDCM: An IoV Secure Data Communication Model Based on Network Encoding and Relay Collaboration

Yan Sun,[1] **Lihua Yin**,[1] **Ying Ma**,[2] **and Chonghua Wang**[3]

[1]*Cyberspace Institute of Advanced Technology, Guangzhou University, Guangzhou, China*
[2]*State Information Center, Beijing, China*
[3]*China Industrial Control Systems Cyber Emergency Response Team, Beijing, China*

Correspondence should be addressed to Lihua Yin; yinlh@gzhu.edu.cn and Chonghua Wang; chonghuaw@live.com

Received 25 July 2022; Revised 26 September 2022; Accepted 11 October 2022; Published 26 November 2022

Academic Editor: Xiaofan Liu

Internet of Vehicles (IoV) is a significant 5G application scenario. As it developed rapidly, more and more vehicles are connected to Internet of Vehicles. The data security and privacy are the premises to ensure its service quality in an open communication network. This paper proposes IoV-SDCM, a secure data communication model in IoV. It includes a self-organizing relay forwarding network and an assured delivery mechanism. The relay forwarding network is used for constructing a dynamic collaboration network with the vehicle as the node. The security delivery mechanism is that network coding is used for data fragmentation and re-encoding to improve network communication reliability. Homomorphic encryption is used to encrypt and protect the encoding vector, improving information leakage and anticollusion attacks. The theoretical proof proves that the model has the ability of data transmission confidentiality and better antiattack capabilities, while it has privacy protection capabilities. Furthermore, the experiment verifies that the model also has the advantage of high and stable data delivery efficiency.

1. Introduction

Internet of Vehicles (IoV) is a new form of mobile Ad hoc networks (MANETs) in the field of road traffic, which is an intelligent information network service [1]. Vehicles are connected, and real-time information exchange is performed on roadside facilities, the Internet, and transportation systems in an open and joint environment [2]. However, the rapid movement of IoV nodes, dynamic topology changes, and on-demand connectivity challenge security and privacy [3]. Its security can be roughly divided into security of terminals, communication networks, and cloud platforms. Especially in the IoV communication process, the high openness is well used to build a transmission network. However, with the high-speed mobility of vehicles, the network topology changes dynamically, and the quality of data transmission service is difficult to guarantee. At the same time, fake nodes or malicious attack entities will also be introduced to launch Sybil attacks, DDoS, and APT attacks. Or infer some privacy information for identity, location, preferences, motion trajectory by traffic analysis, packet analysis and tracing, and collusion attacks. These issues cause that the vehicle control signals and alarm signals are unable to be transmitted to the vehicle terminal in a timely and fast manner or are intercepted and tampered by attackers. They may even greatly impact the safety of users' personal and property [4, 5]. There is a further problem with the effective combination and balance of communication efficiency and security.

To solve the communication efficiency problem, some researchers study V2V relay communication strategies based on node self-organizing cooperation to improve the throughput of in-vehicle networks. For collaborative data distribution and transmission, most researchers use content relay and perfecting schemes [6], and some researchers use game theory-based incentive mechanisms [7, 8] or some learning algorithm [9] to promote cooperation between nodes. These methods have made contributions to the transmission delay and transmission flow and have a significant role in promoting vehicle data transmission in high-

speed road scenarios or urban road scenarios. At the same time, many researchers have paid great attention to the above security issues in the new scenarios of 5G IoV communication and some researchers have also focused on the effective combination and balance of communication efficiency and security. The concept of network coding was proposed in 2000 [10]. On the one hand, network coding technology can improve the robustness of the network, resist the impact on network links, reduce retransmissions, and reduce network management overhead. On the other hand, it can improve the security of information. Through the XOR operation, it is equivalent to encrypting the information, making the information more difficult to be eavesdropped. Even if the information is eavesdropped, it is difficult for the eavesdropper to decode the information correctly because he does not know the processing method of the information and cannot obtain valid information. Therefore, some researchers have used it for vehicle network data communication [11–13] and secure communication [14, 15]. Meanwhile, some researchers use anonymity, encryption, and one or more technologies to design IoV privacy protection mechanisms for solving node identity privacy, location privacy, and data privacy leakage during communication [16, 17].

After research and analysis, it is necessary that a secure data communication scheme of IoV is constructed, configuring a relay collaboration strategy to improve transmission performance and a security policy to guarantee security and privacy. This paper proposes IoV-SDCM, a secure data communication model in IoV. Our main contributions include the following:

(1) We design a self-organizing data communication network composed of relay cooperative vehicle nodes. Through a pseudonymous strategy and a broadcast policy, it achieves source node anonymity, target node anonymity, and communication relationship anonymity in the self-organizing relay forwarding network.

(2) A data delivery strategy is proposed utilizing random network coding aided by homomorphic encryption in the data communication. It increases network throughput while protecting the confidentiality of information. And it can defend against collusion attacks during network coding transmission as the global network coding vectors are encrypted by homomorphic encryption.

(3) After the theoretical proof and performance analysis, the proposed model reduces the overhead of encryption and decryption and the computing cost of nodes. Furthermore, it ensures the confidentiality of data transmission and privacy protection capabilities. We conclude that the model has high reliability and good performance by simulation experiments.

The rest of the paper is organized as follows. Section 2 provides the related work. The proposed IoV-SDCM model is described in Section 3 and the theoretical proof is analyzed in Section 4. Section 5 reports the experiments in detail and discusses the experimental results. Finally, a brief conclusion is drawn in Section 6.

2. Related Work

2.1. Network Coding. Initially, some researchers have made some progress in using network coding to improve the communication performance of the Internet of Vehicles. Ahlswede et al. [10] proposed the concept of network coding in 2000. Ho et al. [18, 19] proposed an algorithm of random network coding (RNC). It is simple in construction and easy to implement in relay collaborative IoV communication. Some scholars have used network encoding to provide IoV data transmission and improve the stability and security of data transmission between dynamic nodes. Kai et al. [20] proposed an auxiliary scheduling algorithm based on network coding to achieve data sharing and collaboration between V2X, which improved data services' performance and bandwidth efficiency and reduced the risk of direct data exposure. Kwon and Park [21] proposed a V2I real-time data distribution system for system network encoding, aiming at the validity and reliability of V2I data transmission. It could effectively reduce the network delay in V2I communication caused by packet loss in the channel. Gao et al. [22] proposed a network coding system that assists D2D transmission, which improved the total network capacity using a payoff function balancing relay selection and resource allocation under complex interference conditions. The above research could effectively improve the throughput of the network, but no further research has been conducted on possible security risks.

Next, some researchers attempt to solve the transmission performance and some security problems by network coding. Khan and Chatzigeorgiou [23] proposed an opportunistic relay framework based on random network coding. It simulated the probability that it could partially or wholly recover confidential data if an eavesdropper intercepted a certain number of packets. And it also validated the trade-offs between security and reliability. Xu et al. [24] proposed a transmission scheme using adaptive relay selection, in which users promote secure communication through collaboration. It had a stable performance gain and effectively suppressed eavesdropping channels. However, due to its security problems by itself, it is still unable to solve the security problems such as conspiracy attacks in data transmission.

2.2. Privacy Protection. At the same time, some researchers have designed the IoV privacy protection using one or more technologies such as anonymity and encryption. Kang et al. [25] proposed that IoV edge resources and fog computing technology can effectively manage and distribute pseudonyms for identity authentication. It improves the ability of identity privacy protection. Wang et al. [26] designed a binary privacy-preserving scheme. The scheme used decentralized CA and biometric password-based authentication to reduce authentication costs and achieve conditional privacy protection. Rajput et al. [27] designed a hierarchical pseudonym authentication protocol that relied solely on CA no longer and reduced the burden on IoV systems. Rabieh et al. [28] gave a route privacy protection

method using homomorphic encryption and error-checking technology, which protects the driver's trajectory data privacy and prevents collusive attacks between malicious vehicles. The above research has only improved in security and privacy, but limited improvement in network performance.

2.3. Our Motivation. Our paper focuses on a model for secure data communication based on network coding and relay collaboration. The data communication network is constructed based on relay collaboration vehicle nodes. In the relay collaborative communication network, the relay node could expand communication coverage, effectively improving communication quality and increasing the eavesdropped information risk. We deeply study the stability and security of data transmission and give the corresponding mechanism for security and privacy protection.

3. IoV-SDCM

Section 3.1 gives IoV-SDCM's model definition and components. For the detailed components, the relay forwarding network and security data delivery mechanism are described in Section 3.2 and Section 3.3.

3.1. Model Definition

Definition 1. IoV secure data communication model (IoV-SDCM): IoV-SDCM could be defined as quadruples (S, D, N, T), where S is the source vehicle node, D is the target vehicle node, N is the forwarding network including relay cooperative vehicle nodes, and T is the secure transmission policy, as shown in Figure 1.

(i) Source vehicle node S: The source vehicle node S divides the information into m slices and uses the pseudonymous strategy to generate m virtual source vehicle nodes, each of which has a data fragment.

(ii) Target vehicle node D: target vehicle node D receives the encoding information, global encoding vectors, and local encoding matrix from the last relay vehicle nodes and decodes to restore the original data.

(iii) Forwarding network N: The forwarding network is defined as a wireless multihop network composed of $m \times n$ relay nodes $R_{(i,j)}$ ($i = 1,2, \ldots, m; j = 1,2, \ldots, n$; $R_{(i,j)}$), where $R_{(i,j)}$ means the j node in the ith hop group. The nodes and their links in the forwarding network meet the following properties at the same time: (1) there are m neighbour nodes within a hop of source node S as the entry nodes $R_{(i,j)}$ ($i = 1,2, \ldots, m; j = 1$); (2) there are m neighbour nodes within a hop of target vehicle node D as the exit nodes $R_{(i,j)}$ ($i = 1,2, \ldots, m; j = n$); (3) the relay nodes within the adjacent one-hop range are all within the communication range of each other; (4) the length of each path is n; (5) there are m disjointed data forwarding paths from the entry node to the exit node.

(iv) Secure transmission policy T: The source node uses the homomorphic encryption function to encrypt the initial global encoding vector and uses a random coefficient to encode the information slices for network encoding, which are transmitted to the entrance nodes, respectively. The relay node encodes the data slices by random coefficient selection. It uses the splitting forwarding strategy in the anonymous forwarding network to transmit the encoding information, global encoding vectors, and local encoding matrix. Finally, the exit node broadcasts encoding information, a global encoding vector, and a local encoding matrix to the target node.

3.2. Relay Forwarding Network. It is necessary to meet the requirements of source node anonymity, target node anonymity, and communication relationship anonymity in the self-organizing relay forwarding network. We use the pseudonymous strategy for generating multiple virtual source nodes to achieve the anonymity of the source node. We build an anonymous relay forwarding network to achieve anonymous communication relationships. In the anonymous relay forwarding network, each hop contains a group of nodes, the groups can communicate with each other, and each group of nodes only knows its previous hop group and the next hop group. The exit node broadcasts information to the target node to achieve receiver anonymity. The specific forwarding network construction process and related anonymity strategies are shown in Figure 2.

3.2.1. Initialization. First, a forwarding link from a source vehicle node to the target vehicle node is generated.

The source vehicle node S routs a request message *RREQ* to target vehicle node D. The structure of *RREQ* is shown in Table 1. If the adjacent node is not the target vehicle node, it is logged to the *RREQ* message, and the number of paths is increased by 1. Then, it continues to be forwarded. Otherwise, it stops forwarding and gets a forwarding path from the source vehicle node S to target vehicle node D if an adjacent node is the target vehicle node. Finally, the target vehicle node D sends an answer message containing the path information *path* to route request node S.

3.2.2. Source Vehicle Node Anonymity Policy. A virtual source node strategy is proposed to achieve the source vehicle node S anonymity. It generates m virtual source nodes as forwarding nodes using the pseudonymous mechanism and generates m forwarding paths.

Supposing the source vehicle node S is identified as ID_s, S presets a hash function H and a random number generator that results in a random number α_i, where ID_s and H are l bits. Finally, S uses the hash function to generate the pseudonym set $S' = \{S_1, S_2, \ldots, S_m\}$. The pseudonym of the source node is

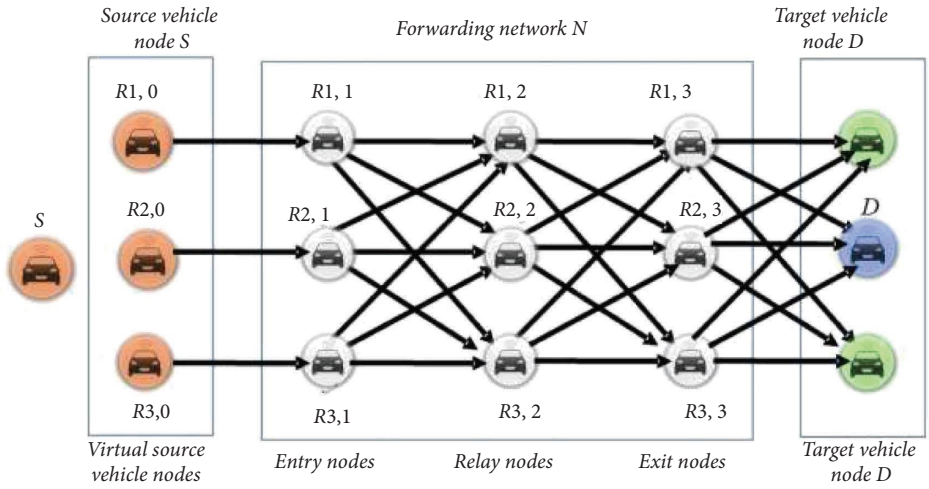

Figure 1: IoV data security delivery model.

$$\begin{aligned} S_1 &= H\left(ID_S \oplus \alpha_1\right), \\ S_2 &= H\left(ID_S \oplus \alpha_2\right), \\ S_m &= H\left(ID_S \oplus \alpha_m\right). \end{aligned} \tag{1}$$

3.2.3. Relay Node Anonymity Policy. In a routing path generated by initialization from source vehicle node S to target node D, node i as the group header ($Header_i$) is selected by neighbouring nodes, constructing the ith hop m-anonymous group (Hop_Group_i). And its member node g_j meets the following conditions:

(i) The ith hop m-anonymous group Hop_Group_i is within the communication scope of its previous anonymous group ($Hop\_Group_{i\text{-}1}$) and the next hop anonymous group (Hop_Group_{i+1})

(ii) The group head node $Header_i$ could set the forwarding path sequence g_j for the member nodes in the group, and the forwarding path sequence of the member nodes is shown in Table 2.

3.2.4. Exit Node Broadcast Policy. For the exit node ($hop = n$), a set of forwarding nodes containing m nodes is generated. In the communication range of the previous hop forwarding node, there are a developed set of forwarding nodes. The target node D, i.e., ID_R, is within the broadcast range of the exit node set.

3.2.5. Update Policy. We set the communication cycle T, in which the source vehicle node carries data transmission along the constructed anonymous forwarding network. When the next communication cycle arrives, the original anonymous forwarding network is abandoned, and an anonymous forwarding network is re-established for data transmission. It could prevent the failure of the routing node and balance the energy consumption.

The specific process of constructing a self-organizing anonymous forwarding network is shown in Algorithm 1.

3.3. Secure Data Delivery Mechanisms. As shown in Figure 3, we suppose a trusted authority distributes a key pair (k_e, k_d) for each node, where k_e is an encryption key, and k_d is a decryption key, and the encryption key k_e is issued to all other nodes.

Phase 1. The source vehicle node S divides the information to be sent M into m slices of information ($M_1, M_2, \ldots, M_m$). Taking $m = 3$ as an example, it generates m virtual nodes, assigning the encoded fragments to the virtual nodes.

If the source vehicle node does not know the target vehicle node key, we preprocess the original data using the information-slicing strategy. This mechanism aims to ensure the confidentiality of the data during transmission. A specific method of slicing information is given as follows.

An original message M of the source vehicle node is sliced into m data fragment. The length of a data fragment is d, and then, the original message M can be represented as an m-dimensional vector:

$$M = \left(M_1, M_2, \ldots M_m\right). \tag{2}$$

Since plaintext transmission of shared information leaks content to relay nodes in the forwarding network, plaintext transmission is not ideal. By introducing a random but reversible transformation matrix A to construct a perturbation source information slice, the original information after the disturbance M' is as follows.

$$\begin{aligned} M' &= AM \\ &= \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{pmatrix} \left(M_1, M_2 \ldots, M_n\right) \\ &= \left(A_1 M_1, A_2 M_2, \ldots, A_n M_n\right) \\ &= \left(M_2', M_2', \ldots, M_n'\right). \end{aligned} \tag{3}$$

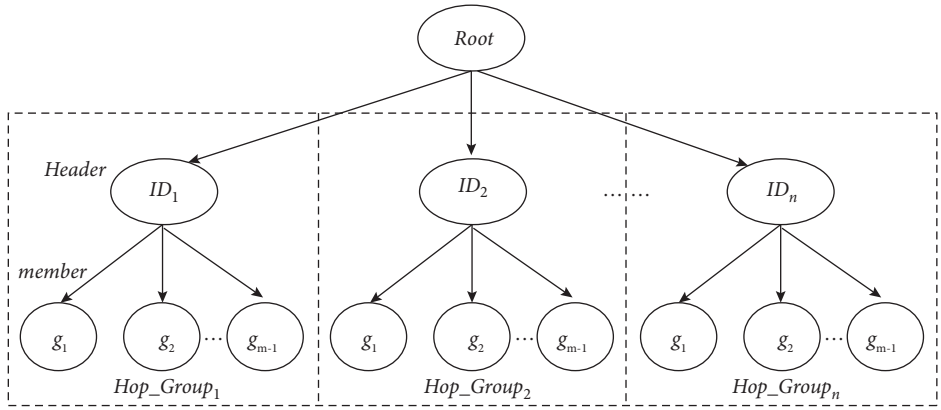

Figure 2: A self-organizing anonymous forwarding network based on groups.

Table 1: Routing request message RREQ.

Source	Target	*Num*	path [1]	path [2]	. . . . .	path [n]
ID_S	ID_R	n	ID_1	ID_2	. . . . .	ID_n

Table 2: The m-anonymous group of the ith hop (Hop_Group_i).

hop	hop_1	hop_2	. . . . .	hop_j	. . . . .	hop_m
i	g_1	g_2		g_j		g_m

Thus, the source information transmitted in the anonymous forwarding network is the information M transformed information.

$$M' = (M_1', M_2', \ldots M_m'). \tag{4}$$

The information-slicing policy avoids the direct transmission from leaking content to relay nodes in the forwarding network.

For the target vehicle node, as long as all slicing information of M' and transformation matrix A are received, the original information M can be restored, i.e.,

$$\begin{aligned} M &= A^{-1}M' \\ &= A^{-1}(M_1', M_2', \ldots M_m'). \end{aligned} \tag{5}$$

Phase 2. The virtual source node encodes each slicing information separately and then sends the encoded data and the encrypted global encoding vector to the entry node, respectively.

The source node S builds m different forwarding paths for the data slices. For the original data M, the source node divides it into m information slices (M_1, M_2, . . ., M_m) in the source node data slicing strategy. Its encoding forwarding strategy is that the source node selects a random coefficient for each slicing data and computes the network encoding, and then, it sends the encoded data and the encrypted global encoding vector to the next hop node along m different paths.

Assuming the length of each slice data d, the specific steps for the source node data encoding are as follows:

Step 1. Random coefficient selection for network encoding.

Randomly select m coding coefficient vectors of length d on a finite field F, $C_{(i,0)}{}^j$ ($i = 1,2, \ldots, m$; $j = 1,2, \ldots, m$), which forms a local encoding matrix, denoted $C_{(i,0)}$ ($i = 1,2, \ldots, m$), i.e.,

$$C(i,0) = \begin{bmatrix} C_{(i,0)}^1 \\ C_{(i,0)}^2 \\ \vdots \\ C_{(i,0)}^m \end{bmatrix} (i = 1, 2, \ldots, m). \tag{6}$$

Step 2. Network encoding for each slice M_i.

The local encoding matrix $C_{(i,0)}$ and each slice M_i operate a binary bit addition. That is, each encoding coefficient vector of $C_{(i,0)}$, $C_{(i,0)}{}^j$ performs an XOR operation on M_i, separately, denoted as follows:

$$\begin{aligned} M_i^{(0)} &= C_{(i,0)} \oplus M_i \\ &= \left(C_{(i,0)}^1 \oplus M_i, C_{(i,0)}^2 \oplus M_i, \ldots, C_{(i,0)}^m \oplus M_i\right) (i = 1, 2, \ldots, m). \end{aligned} \tag{7}$$

Step 3. Calculating the global encoding vector.

The global coding vector $V_i^{(0)}$ consists of the ith coding coefficient component of $C_{(i,0)}$, and then, the global coding vector is as follows:

$$V_i^{(0)} = \left(C_{(1,0)}^i, C_{(2,0)}^i, \ldots, C_{(m,0)}^i\right). \tag{8}$$

Step 4. Encrypting the global encoding vector.

The global encoding vector $V_i^{(0)}$ is encrypted using a homomorphic cryptographic function, denoted as

$$\begin{aligned} EV_i^{(0)} &= E_H\left(V_i^{(0)}, k_e\right) \\ &= E_H\left(\left(C_{(1,0)}^i, C_{(2,0)}^i, \ldots, C_{(m,0)}^i\right), k_e\right) \\ &= \left(EC_{(1,0)}^i, EC_{(2,0)}^i, \ldots, EC_{(m,0)}^i\right). \end{aligned} \tag{9}$$

Step 5. Forwarding the encoded message.

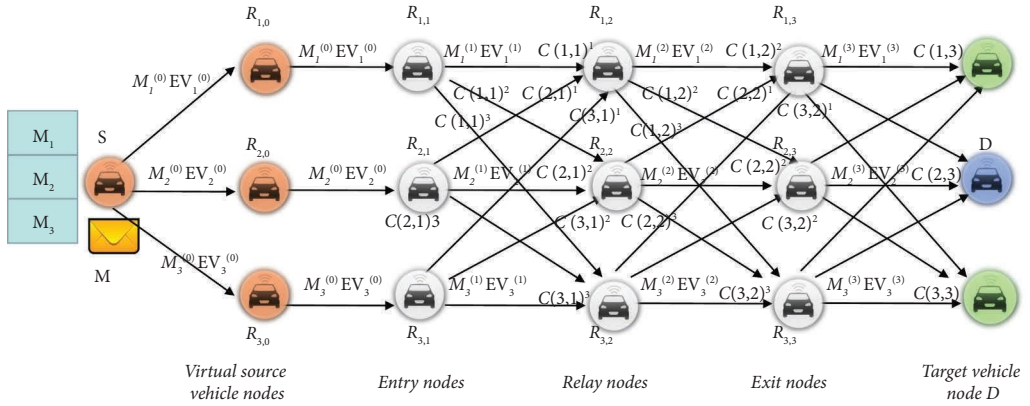

Figure 3: Secure transmission mechanism ($m = 3$, $n = 3$ as an example).

The source node sends the encoded information $M_i^{(0)}$ and the encrypted global encoded vector $EV_i^{(0)}$ along the ith path to the exit node.

The above encoding forwarding process avoids the direct decoding of a single entry node because the encoded data fragments consist of the splicing data and column vector of the local encoding matrix. The encoded data slices and encrypted global encoded vectors are forwarded along the m-path, respectively, preventing the exit nodes from colluding to recover the original data.

Phase 3. The entry nodes re-encode and use the splitting forwarding strategy for transmission after receiving the encoded information and the encrypted global coding vector. The jth relay node of the ith hop $R_{(i,j)}$ receives the m packets sent by the relay node of the previous hop m-path, selects the random coefficient, and encodes m packets. Then, according to the list of neighbour nodes $R_{(i,j)}$, the encoded information, the global encoding vector ciphertext encrypted, and the row vector of the local encoding vector are forwarded to the next hop relay node along m different paths. Repeat the above process until the exit node completes the encoding operation.

The data forwarding policy of the relay node is similar to the data forwarding policy of the source node. The relay node $R_{(i,j)}$ receives m packets sent by the relay node of the previous hop m-path. It selects the random coefficient and performs network encoding for m packets. Then, the encoded information, the ciphertext of the global encoding vector, and the local encoding row vector are forwarded along the m different paths to the next hop relay node. The specific steps for relay node data encoding are as follows:

Step 6. Random coefficient selection for the network encoding.

Randomly select m coding coefficient vectors of length d in a finite field F, ${C_{(i,j)}}^k$ ($i = 1,2, \ldots, m; j = 1,2, \ldots, m; k = 1,2, \ldots, m$), that forms a local coding matrix, denoted as $C_{(i,j)}$, i.e.,

$$C(i,j) = \begin{bmatrix} C_{(i,j)}^1 \\ C_{(i,j)}^2 \\ \vdots \\ C_{(i,j)}^m \end{bmatrix} (i = 1,2,\ldots,m;\ j = 1,2,\ldots,m). \tag{10}$$

Step 7. Relay node $R_{(i,j)}$ performs network encoding for the m packets received, $M_i^{(j)}$.

The local encoding matrix $C_{(i,j)}$ calculates a binary bit addition on $M_i^{(j)}$. That is, each encoding coefficient vector ${C_{(i,j)}}^k$ of $C_{(i,j)}$ performs an XOR operation on $M_i^{(j)}$, respectively, denoted as

$$\begin{aligned} M_i^{(j)} &= C_{(i,j)} \oplus M_i^{(j-1)} \\ &= \left(C_{(i,j)}^1 \oplus M_i^{(j-1)}, C_{(i,j)}^2 \oplus M_i^{(j-1)}, \ldots, C_{(i,j)}^m \oplus M_i^{(j-1)}\right) \\ &(i = 1,2,\ldots,m). \end{aligned} \tag{11}$$

Step 8. Calculating the global encoding vector ciphertext of the relay node $R_{(i,j)}$.

Since the global encoding vector is the ciphertext encrypted by the homomorphic encryption function in the received m-path packet, the relay node does not have the corresponding decryption key. The relay node could not decode the original packet to be directly recovered. At the same time, according to the homomorphic cryptographic function, linear changes could directly use the ciphertext of the global coding vector to generate new global coding vectors.

The global coding vector $V_i^{(j)}$ consists of the ith coding coefficient component, $C_{(i,j)}$; then, the global coding vector is as follows:

$$EV_i^{(j)} = \sum_{k=1}^{m} C_{(k,j-1)}^k EV_i^{(j-1)}. \tag{12}$$

Step 9. Forwarding the encoded message.

The relay node sends the encoded information $M_i^{(j)}$, the ciphertext of the global encoding vector $EV_i^{(j)}$, and the ith row vector of the local encoding matrix $C_{(i,j)}{}^i$ to the relay node in the ith path. In addition, the other row vectors of the local encoding matrix, $C_{(i,1)}{}^k$ $(k \neq i)$, are sent to the relay node along the j-th path, respectively.

Here, there are two points to be noted. For the entry node $R_{(i,1)}$, there is only one input link, $V_i^{(1)} = V_i^{(0)}$ and $EV_i^{(1)} = EV_i^{(0)}$. For the exit node, it broadcasts the encoded information $M_i^{(n)}$ and the ciphertext of the global encoding vector $EV_i^{(n)}$ and the local encoding matrix $C_{(i,j)}$ to the target vehicle node.

In the above encoding forwarding process, the encoded data slice $M_i^{(j)}$, the ciphertext of the global encoding vector $EV_i^{(j)}$, and the row vector of the local encoding matrix $C_{(i,j)}{}^i$ are forwarded along the unjoint m-paths, respectively. It prevents the relay node from recovering the original data.

Phase 4. The exit node broadcasts the encoding information, global encoding vectors, and local encoding matrix to the target node.

Phase 5. Without considering network errors, the target vehicle node D receives the network encoding information $M_i^{(n)}$, the global encoding vector ciphertext $EV_i^{(n)}$, and local encoding matrix $C_{(i,n)}$ from m-path. It uses the information to recover the information completing the data forwarding transmission process.

Assuming that the probability of transmission error is negligible, the target vehicle node D could receive the network encoding information $M_i^{(n)}$, the global encoding vector ciphertext $EV_i^{(n)}$, and the local encoding matrix $C_{(i,n)}$ from m-path. The decoding steps are as follows.

Step 10. The target node D uses the decryption key k_d to decrypt the ciphertext of the global encoding vector $EV_i^{(n)}$ to obtain the global encoding vector $V_i^{(n)}$, and the decryption operation is as follows:

$$\begin{aligned} V_i^{(n)} &= D\left(EV_i^{(n)}, k_d\right) \\ &= \left(\bigoplus_{k=0}^{n-1} C_{(1,k)}^i, \bigoplus_{k=0}^{n-1} C_{(1,k)}^i, \ldots, \bigoplus_{k=0}^{n-1} C_{(1,k)}^i\right). \end{aligned} \tag{13}$$

Step 11. The target vehicle node D restores the original slice information M_i according to the network encoding information $M_i^{(n)}$, the global encoding vector $V_i^{(n)}$, and the local encoding matrix $C_{(i,n)}$ as follows:

$$M_i = M_i^{(n)} \oplus \left(V_i^{(n)} \oplus C_{(i,n)}\right). \tag{14}$$

Step 12. The target vehicle node D recovers the original information M based on the original slice information M_i, denoted as follows:

$$M = (M_1, M_2, \ldots M_m). \tag{15}$$

Next, we will prove that the secure data delivery mechanism can be successfully decrypted after receiving encrypted data packets.

Assuming that the probability of transmission errors is negligible in a data forwarding network based on network encoding, the target vehicle node D could receive the network encoding information $M_i^{(n)}$, the ciphertext of global encoding vector $EV_i^{(n)}$, and the local encoding matrix $C_{(i,n)}$ from m-path.

Theorem 1. *Without considering network errors, the target vehicle node D receives the network encoded information $M_i^{(n)}$, the ciphertext of the global encoding vector $EV_i^{(n)}$, and the local encoding matrix $C_{(i,n)}$ from the m-path, using them to recover the information sent by the source node S correctly.*

Proof. Without considering network errors, the target vehicle node D receives the network encoding information $M_i^{(n)}$, ciphertext $EV_i^{(n)}$, and local encoding matrix $C_{(i,n)}$ from the m-path. The above information is obtained after n operations in the data forwarding network based on network encoding. First, analyze the calculation process of the above information.

The information slice M_i is encoded by n times to obtain $M_i^{(n)}$, and its calculation process is as follows:

$$\begin{aligned} M_i^{(n)} &= C_{(i,n)} \oplus M_i^{(j-1)} \\ &= C_{(i,n)} \oplus \left(C_{(i,n-1)} M_i^{(j-2)}\right) \\ &= C_{(i,n)} \oplus \left(C_{(i,n-1)} \oplus \left(C_{(i,n-2)} \oplus M_i^{(j-3)}\right)\right) \\ &= C_{(i,n)} \oplus \left(C_{(i,n-1)} \oplus \left(C_{(i,n-2)} \oplus \left(\cdots \left(C_{(i,0)} \oplus M_i\right)\right)\right)\right) \\ &= \bigoplus_{k=0}^{n} C_{(i,k)} \oplus M_i. \end{aligned} \tag{16}$$

The ciphertext of the global encoding vector $EV_i^{(n)}$ is obtained by calculating $EV_i^{(0)}$ by n times, and its calculation process is as follows:

$$
\begin{aligned}
EV_i^{(n)} &= \sum_{k=1}^{n} C_{(k,n-1)}^{k} EV_i^{(n-1)} \\
&= \sum_{k=1}^{n} C_{(k,n-1)}^{k} \left(\sum_{k=1}^{n} C_{(k,n-2)}^{k} EV_i^{(n-2)} \right) \\
&= \sum_{k=1}^{n} C_{(k,n-1)}^{k} \left(\sum_{k=1}^{n} C_{(k,n-2)}^{k} \left(\sum_{k=1}^{n} \cdots \left(\sum_{k=1}^{n} C_{(k,1)}^{k} EV_i^{(1)} \right) \right) \right) \\
&= \sum_{k=1}^{n} C_{(k,n-1)}^{k} \left(\sum_{k=1}^{n} C_{(k,n-2)}^{k} \left(\sum_{k=1}^{n} \cdots \left(\sum_{k=1}^{n} C_{(k,1)}^{k} \left(EC_{(1,0)}^{i}, EC_{(2,0)}^{i}, \ldots, EC_{(m,0)}^{i} \right) \right) \right) \right) \\
&= \left(\bigoplus_{k=0}^{n-1} EC_{(1,k)}^{i}, \bigoplus_{k=0}^{n-1} EC_{(1,k)}^{i}, \ldots, \bigoplus_{k=0}^{n-1} EC_{(1,k)}^{i} \right).
\end{aligned}
\tag{17}
$$

The target vehicle node D uses the decryption key k_d to decrypt the global encoding vector ciphertext $EV_i^{(n)}$ and obtains the global encoding vector $V_i^{(n)}$. The result is as follows:

$$
\begin{aligned}
V_i^{(n)} &= D\left(EV_i^{(n)}, k_d \right) \\
&= D\left(\left(\bigoplus_{k=0}^{n-1} EC_{(1,k)}^{i}, \bigoplus_{k=0}^{n-1} EC_{(1,k)}^{i}, \ldots, \bigoplus_{k=0}^{n-1} EC_{(1,k)}^{i} \right), k_d \right) \\
&= \left(\bigoplus_{k=0}^{n-1} C_{(1,k)}^{i}, \bigoplus_{k=0}^{n-1} C_{(1,k)}^{i}, \ldots, \bigoplus_{k=0}^{n-1} C_{(1,k)}^{i} \right).
\end{aligned}
\tag{18}
$$

According to $M_i^{(n)}$, $EV_i^{(n)}$, and the local encoding matrix $C_{(i,n)}$, the original information slice M_i can be recovered. The calculation method is as follows:

$$
\begin{aligned}
M_i^{(n)} &= \bigoplus_{k=0}^{n} C_{(i,k)} \oplus M_i \\
&= \bigoplus_{k=0}^{n-1} C_{(i,k)} \oplus C_{(i,n)} \oplus M_i \\
&= V_i^{(n)} \oplus C_{(i,n)} \oplus M_i.
\end{aligned}
\tag{19}
$$

According to the above formula, M_i could be derived and calculated as follows:

$$
M_i = M_i^{(n)} \oplus \left(V_i^{(n)} \oplus C_{(i,n)} \right). \tag{20}
$$

Finally, the original information $M = (M_1, M_2, \ldots, M_n)$ could be recovered according to M_i. End. □

4. Security Analysis

This section proves the security of the model from confidentiality and anticollusion attack. The confidentiality is shown that the relay nodes except the target vehicle node could not obtain the original information. The anticollusion attack is shown that collusion attackers could not jointly recover the original data.

4.1. Confidentiality. In IoV-SDCM, the confidentiality of the message indicates that no node in the forwarding network could obtain the content of the sender message except for the destination node.

Theorem 2. *In addition to the target vehicle node D in IoV-SDCM, the relay nodes in the network could not recover some of the slicing information of the original information M sent by the source vehicle node S.*

Proof. According to the proof conclusions of reference [29], it was proved that the homomorphic encryption function encrypts the data homomorphically to form the ciphertext, and then, the obtained ciphertext calculation result is obtained by homomorphically decrypting the plaintext, which is the same as the result of directly calculating the plaintext data, but the plaintext data cannot be obtained. Based on the demonstration, the nodes in the IoV-SDCM forwarding network could not decrypt the part of the information of the ith hop global coding vector because the homomorphic encryption function encrypts the global encoding vector. The relay nodes except the target vehicle node D could not recover the original information M sent by the source vehicle node S. Therefore, there is no early decoding phenomenon, and the relay nodes could not decrypt part of the slicing information of the original information M. End.

In summary, IoV-SDCM ensures the confidentiality of the message. □

4.2. Anticollusion Attack. A collusion attack refers to the fact that multiple attackers collude with each other to decode the original information transmitted by the source vehicle node. The model proposed in this paper could prevent multiple leakers from conspiring to recover some information from the original information piece M_i.

Theorem 3. *In addition to the target vehicle node D in the model, there are multiple relay nodes in the anonymous forwarding network. The conspiratorial attackers could not obtain part of the original information M by leaking part of the information to recover the original data jointly.*

Proof. According to the evidence in the paper [30], when the global coding vector is exposed, multiple leak nodes could recover some of the information of the original information slice M_i through collusive attacks. The main reason is the leakage of the global coding vector $V_i^{(j)}$. By obtaining a partial global encoding vector $V_i^{(j)}$, an attacker could solve

for some slices of the original information. The model proposed in this paper uses a homomorphic encryption function to encrypt the global encoding vector $V_i^{(j)}$. Only the target vehicle node D could decrypt the global encoding vector $V_i^{(j)}$, while other nodes could not decrypt the global encoding vector $V_i^{(j)}$. Therefore, the leakage of the global encoding vector $V_i^{(j)}$ is prevented, so that the attacker could not solve part of the slicing information. That is, the collusion attack of multiple leaked nodes could not be successful. End □

Corollary 1. *In addition to the destination node D in IoV-SDCM, if multiple relay nodes jointly recover the original information by disclosing some data, the conspiratorial attackers could not get the original data.*

Proof. According to Theorem 3, due to homomorphic encryption functions to encrypt the global encoding vector, the nodes in the forwarding network could not decrypt part of the i-hop global encoding vector. There is no early decoding phenomenon, and the conspiratorial attackers could not obtain part of the fragmented information of the original information M. Therefore, the intermediate nodes could not obtain the information of the original information M. End □

4.3. Privacy Protection. This section proves that the mode enabled privacy protection from packet analysis, traffic analysis, and packet tracing.

Packet analysis is when the attacker analyzes the packet to obtain information such as the identity and address of the sender or receiver. In the process of anonymous data forwarding, m packets are sent by the source vehicle node containing their pseudonyms and entry node identities; the m packets are sent by the relay node $R_{(i,j)}$ containing the relay node identities and successor node identities. Those nodes do not have the identity information of the actual source vehicle node and the target vehicle node. Only one packet in the m packets sent by the exit node $R_{(i,n)}$ contains the target vehicle node. The attacker could not distinguish which is the destination node. Therefore, in the process of anonymous data forwarding, an attacker could not obtain information such as the identity and address of the sender or receiver and nor could it determine the communication relationship between the sender and the receiver.

Traffic analysis is when an attacker determines the communication relationship by observing the traffic patterns of the network. In addition to the source vehicle node in the anonymous forwarding network, the input degree of other relay nodes $R_{(i,j)}$ is m, the output degree is also m, and the network traffic pattern is balanced. Therefore, the attacker could not determine the location of the target vehicle node by observing the network traffic pattern, but only the location of the source vehicle node could be found.

Packet tracing is when an attacker listens to a wireless channel near a node and determines the source vehicle node through hop-by-hop tracing. Assuming an attacker is listening to a wireless channel at a relay node in the communication cycle T, each relay node $R_{(i,j)}$ has m front-drive nodes. The attacker could move to a front-drive node $R_{(k,j\text{-}1)}$ of the listening node each time. If the attacker could move to the source vehicle node in the same communication cycle, you can locate the sender's location. Suppose the communication cycle ends and the attacker has not moved to the source vehicle node. In that case, the original anonymous forwarding network is abandoned, and the attacker could not correctly locate the sender's location. Therefore, the success of packet tracing is affected by the length of the communication cycle T and the forwarding path.

5. Performance Analysis

The IoV simulator of OMNeT++ is used to simulate the IoV environment. We set a one-way 4-lane road shape of 30000 m × 60 m. The vehicles only conduct V2V communication, the vehicle communication radius is 100~300 m, and the vehicle speed range is 50 km/h~100 km/h. The linear mobility mobile module is used to control the movement of nodes and the vehicle. The speed of the nodes is configured to follow a random distribution (14 mps, 28 mps). For the communication between vehicle nodes, the IEEE 802.11 b PHY/MAC protocol with a data rate of 11 Mbps is configured, and IoV-SDCM is added to the application layer module.

We give two definitions to analyze the data delivery performance of the model. One is the successful decoding rate of vehicle nodes, which is the number of nodes that can decode data divided by the number of all vehicle nodes in the simulation time. The other is the data receiving rate of vehicle nodes, which are the valid data packets received by all vehicle nodes divided by the total number of data packets in the simulation time.

In the experiment, set the slice size $d = 10$, 20, and 30, the size of each data packet after the slice is 1 MB, and the communication radius of the vehicle node is 200 m. Figure 4 shows that the experiment results in the successful decoding rate of vehicles decreases as the vehicle node amount increases. Compared with the successful decoding rate, the data reception rate decreases slowly. The network topology greatly affects the successful decoding rate of vehicles. When the vehicle is in a tight state, more data packets are obtained by V2V communication. However, the successful decoding rate of the vehicle is more demanding. It not only requires the vehicle to receive the data packets but also needs to be able to obtain enough packets to be decoded.

When the amount of data to be distributed is constant, the amount of data to be distributed is set to 30 MB, the communication radius of vehicle nodes is 200m, and the slice size is set to 10, 15, 20, 25, and 30, for a total of 100 vehicle nodes. Figure 5 shows that with the increase of the slice size d, the transmission time is reduced after the data volume of each data packet is reduced. It could ensure the successful receiving rate of vehicles, so the efficiency is increasing. At the same time, we can see that the change of the decoding rate is affected by the successful reception rate. The decoding rate is consistent with the change of the successful

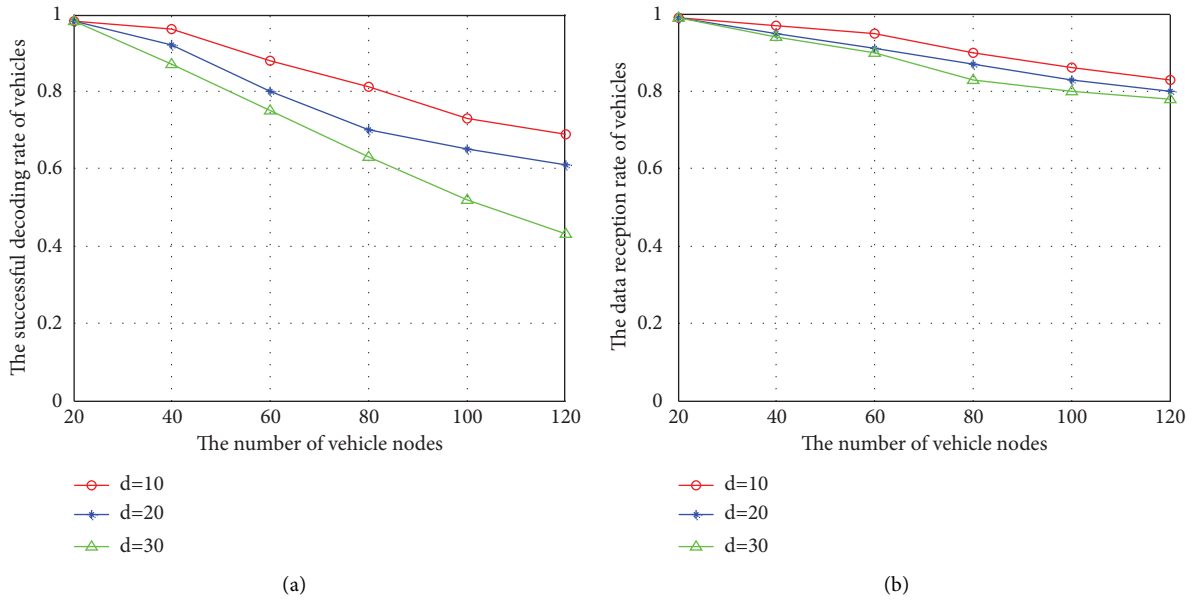

Figure 4: The influence of the vehicle node amount on the delivery efficiency.

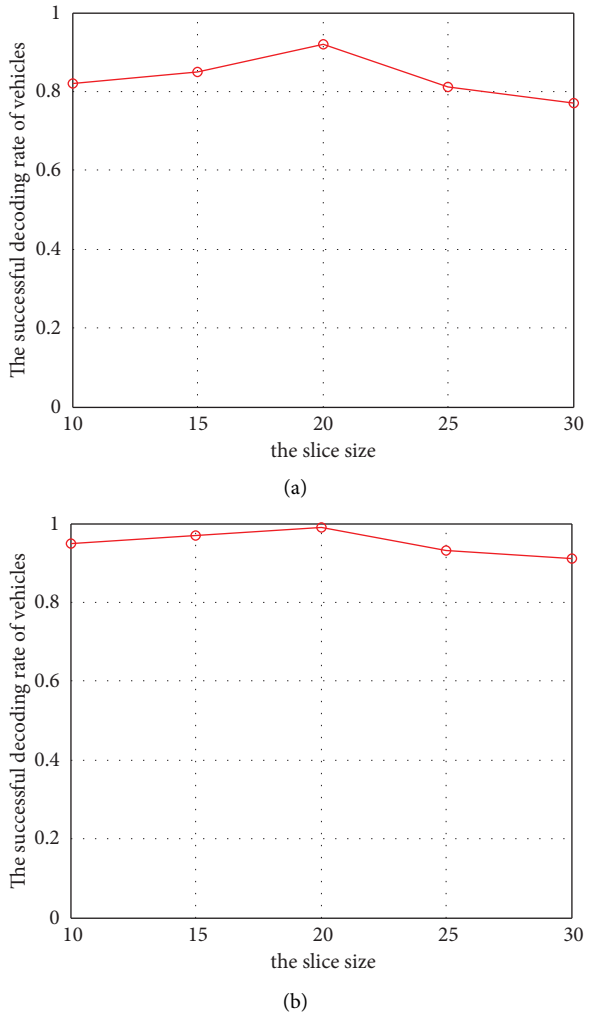

Figure 5: The influence of slice size on delivery efficiency.

Input: Source vehicle node S, target vehicle node D
Output: Anonymous forwarding network N
Process:
(1) If (Period = T)
(2) T = Null;//Using a tree structure to build an anonymous network and initialize the network.
(3) *Send* (S, D, *mesg_rreq*);//The source vehicle node S sends a routing request message to the target vehicle node *DRREQ*, and record the path information from the source node to the target node.
(4) *Path* = Receive (S, D, *mesg_path*);//The target vehicle node D sends an answer message containing the path information *path* to the source vehicle node S and complete the construction of the forwarding network. See Initialization in 3.2 Section for detail.
(5) If (*Path*)//If there is a forwarding path between the source vehicle node and the target vehicle node.
(6) For (each *Path* [i], $i < n$)//For each node in the forwarding path, it is used as a group header to build an m-anonymous group.
(7) select $m-1$ node satisfying (Hop_Group_i is in the radius of $Hop\_Group_{i\text{-}1}$ and Hop_Group_{i+1})
(8) $g_j = Hop\_Group_i[j]$;
(9) End For
(10) End If
(11) *Path* [n] = *broadcast*;//The exit node broadcasts messages to the target node
(12) End If
(13) End

Algorithm 1: An anonymous forwarding network construction.

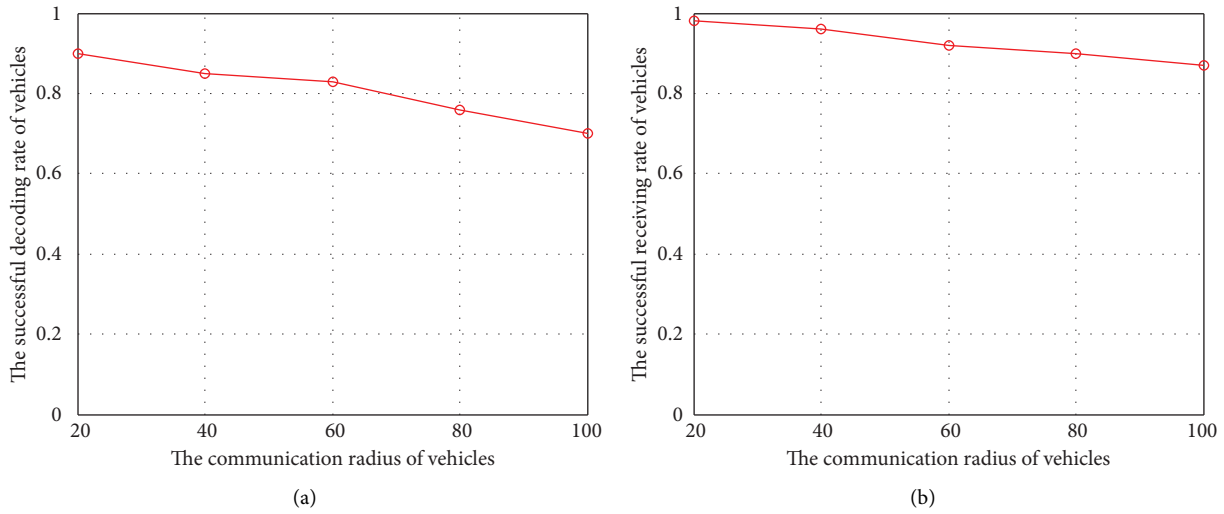

Figure 6: The influence of communication range on delivery efficiency.

receiving rate, and the decoding rate is slightly lower than the receiving rate. However, the number of slices d can neither be too large nor too small. If it is too large, the reception will be incomplete and the decoding rate is low; if it is too small, the transmission times will be increased, and the system efficiency will be reduced and the decoding rate is also low.

With the expansion of the communication radius of vehicle nodes, the vehicle nodes can communicate with more nodes. In the simulation experiment, set the slice size $d = 10$, the size of each data slice is 3 MB, and there are 100 vehicle nodes. Figure 6 shows that the successful receiving rate and decoding rate of vehicles decreases, and the decrease is relatively stable if the vehicle communication radius increases. We can see that the decoding rate of data packets is consistent with the change of the successful reception rate, and the decoding rate is slightly lower than the receiving rate. Because when the communication radius becomes larger, the vehicle node can conduct V2V communication with more vehicles, so that the average request node increases, and the time that can communicate with other nodes is wasted, thus resulting in a decrease in efficiency.

6. Conclusion

This paper proposes IoV-SDCM, a secure data communication model in IoV. This model is based on the relay forwarding network and the secure data transmission mechanism. It constructs a dynamic communication network with the vehicle as the node. The data communication network reaches the target vehicle through the self-

organizing relay collaboration policy with the security and privacy strategy. Theoretical proof proves that it ensures the confidentiality of data transmission with better antiattack capabilities, privacy protection capabilities, and simulation experiments verify that the model has the advantage of high and stable performance.

Data Availability

This paper adopts theoretical proof and simulation experiments to prove the correctness, security, and performance of the proposed model, so no experimental data are involved.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This study was funded by the Postdoctoral Research Fund of Guangzhou University.

References

[1] S. Chen, J. Hu, Y. Shi et al., "Vehicle-to-Everything (v2x) services supported by LTE-based systems and 5G," *IEEE Communications Standards Magazine*, vol. 1, no. 2, pp. 70–76, 2017.

[2] F. Wei, S. Zeadally, P. Vijayakumar, N. Kumar, and D. He, "An intelligent terminal based privacy-preserving multimodal implicit authentication protocol for Internet of connected vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 7, pp. 3939–3951, 2021.

[3] R. E. Navas, F. Cuppens, N. Boulahia Cuppens, L. Toutain, and G. Z. Papadopoulos, "MTD, where art thou? A systematic review of moving target defense techniques for IoT," *IEEE Internet of Things Journal*, vol. 8, no. 10, pp. 7818–7832, 2021.

[4] T. A. Butt, R. Iqbal, K. Salah, M. Aloqaily, and Y. Jararweh, "Privacy management in social Internet of vehicles: review, challenges and blockchain based solutions," *IEEE Access*, vol. 7, Article ID 79694, 2019.

[5] J. Wan, X. Cao, K. Yao, D. Yang, E. Peng, and Y. Cao, "Data mining technology application in false text information recognition," *Mobile Information Systems*, vol. 2021, pp. 1–13, Article ID 4206424, 2021.

[6] X. Chen and G. Chu, "Data cooperative distribution mechanism of Internet of vehicles using D2D technology," *Advances in Multimedia*, vol. 2022, Article ID 9722915, 10 pages, 2022.

[7] M. Hosseini, R. Ghazizadeh, and H. Farhadi, "Game theory-based radio resource allocation in NOMA vehicular communication networks supported by UAV," *Physical Communication*, vol. 52, Article ID 101681, 2022.

[8] S. M. A. Kazmi, T. N. Dang, I. Yaqoob et al., "A novel contract theory-based incentive mechanism for cooperative task-offloading in electrical vehicular networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 7, pp. 8380–8395, 2022.

[9] T. Zeng, O. Semiariy, M. Chen, W. Saad, and M. Bennis, "Federated learning on the road autonomous controller design for connected and autonomous vehicles," in *Proceedings of the IEEE Transactions on Wireless Communications*, p. 1, Austin, TX, USA, December 2022.

[10] R. Ahlswede, C. Ning, S. Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Transactions on Information Theory*, vol. 46, no. 4, pp. 1204–1216, 2000.

[11] B. Zhang, Z. Liu, S. H. G. Chan, and G. Cheung, "Collaborative wireless freeview video streaming with network coding," *IEEE Transactions on Multimedia*, vol. 18, no. 3, pp. 521–536, 2016.

[12] T. Zhu, C. Li, Y. Tang, and Z. Luo, "On latency reductions in vehicle-to-vehicle networks by random linear network coding," *China Communications*, vol. 18, no. 6, pp. 24–38, 2021.

[13] H. Song, L. Liu, B. Shang, S. Pudlewski, and E. S. Bentley, "Enhanced flooding-based routing protocol for swarm UAV networks: random network coding meets clustering," in *Proceedings of the IEEE INFOCOM 2021 - IEEE Conference on Computer Communications*, pp. 1–10, Vancouver, BC, Canada, May 2021.

[14] A. Engelmann and A. Jukan, "Balancing the demands of reliability and security with linear network coding in optical networks," in *Proceedings of the 2016 IEEE International Conference on Communications*, pp. 1–7, Kuala Lumpur, Malaysia, May 2016.

[15] C. Cheng, J. Lee, T. Jiang, and T. Takagi, "Security analysis and improvements on two homomorphic authentication schemes for network coding," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 5, pp. 993–1002, 2016.

[16] D. Jiang, Y. Wang, Z. Lv, S. Qi, and S. Singh, "Big data analysis based network behavior insight of cellular networks for industry 4.0 applications," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 2, pp. 1310–1320, 2020.

[17] R. Hussain, D. Kim, J. Son et al., "Secure and privacy-aware incentives-based witness service in social Internet of vehicles clouds," *IEEE Internet of Things Journal*, vol. 5, no. 4, pp. 2441–2448, 2018.

[18] T. Ho, M. Médard, J. Shi, M. Effros, and D. Karger, "On randomized network coding," in *Proceedings of the Annual Allerton Conference on Communication Control and Computing)*, Monticello, IL, USA, October 2003.

[19] T. Ho, M. Médard, R. Kötter, and R. K. David, "Toward a Random Operation of Networks," *IEEE Transactions on Information Theory - TIT*, vol. 50, 2004.

[20] K. Liu, J. K. Y. Ng, J. Wang, V. C. S. Lee, W. Wu, and S. H. Son, "Network-coding-assisted data dissemination via cooperative vehicle-to-vehicle/-infrastructure communications," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 6, pp. 1509–1520, 2016.

[21] J. Kwon and H. Park, "Reliable data dissemination strategy based on systematic network coding in V2I networks," in *Proceedings of the 2019 International Conference on Information and Communication Technology Convergence (ICTC)*, pp. 744–746, Jeju, Korea, October 2019.

[22] C. Gao, Y. Li, Y. Zhao, and S. Chen, "A two-level game theory approach for joint relay selection and resource allocation in network coding assisted D2D communications," *IEEE Transactions on Mobile Computing*, vol. 16, no. 10, pp. 2697–2711, 2017.

[23] A. S. Khan and I. Chatzigeorgiou, "Opportunistic relaying and random linear network coding for secure and reliable communication," *IEEE Transactions on Wireless Communications*, vol. 17, no. 1, pp. 223–234, 2018.

[24] P. Xu, Z. Ding, and X. Dai, "Achievable secrecy rates for relay-eavesdropper channel based on the application of noisy network coding," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 7, pp. 1736–1751, 2018.

[25] J. Kang, R. Yu, X. Huang, and Y. Zhang, "Privacy-preserved pseudonym scheme for fog computing supported Internet of vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 8, pp. 2627–2637, 2018.

[26] F. Wang, Y. Xu, H. Zhang, Y. Zhang, and L. Zhu, "2FLIP: a two-factor lightweight privacy-preserving authentication scheme for vanet," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 2, pp. 896–911, 2016.

[27] U. Rajput, F. Abbas, and H. Oh, "A hierarchical privacy preserving pseudonymous authentication protocol for VANET," *IEEE Access*, vol. 4, pp. 7770–7784, 2016.

[28] K. Rabieh, M. M. E. A. Mahmoud, and M. Younis, "Privacy-preserving route reporting schemes for traffic management systems," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 3, pp. 2703–2713, 2017.

[29] C. Gentry, *A Fully Homomorphic Encryption Scheme*, Stanford University, Stanford, CA, USA, 2009.

[30] C. Berge, *Graphs and Hypergraphs*, Elsevier, Amsterdam, Netherlands, 1973.