WILEY | Hindawi

*Research Article*

# A Network Coding-Based Braided Multipath Routing Protocol for Wireless Sensor Networks

## Zhihua Li [ID],[1] Minda Xu,[1] Tingxu Liu,[2] and Lei Yu[3]

[1]*Department of Computer Science and Technology, School of Internet of Things Engineering, Jiangnan University, Jiangsu, Wuxi 214122, China*
[2]*Jiangsu Provincial Engineering Laboratory of Pattern Recognition and Computational Intelligence, Jiangnan University, Jiangsu, Wuxi 214122, China*
[3]*School of Computer Science, College of Computing, Georgia Institute of Technology, Atlanta, GA 30332, USA*

Correspondence should be addressed to Zhihua Li; zhli@jiangnan.edu.cn

In wireless sensor networks (WSNs), energy efficiency can simultaneously guarantee robustness to link loss and node failure and is a key design goal of routing protocols because WSNs are strongly constrained in terms of transmission reliability, transmission delay, and energy consumption. Braided multipath routing has become a powerful tool for tolerating node failures and link losses, with high reliability and efficient data transmission rates. In this paper, we propose a novel network coding-based braided multipath routing called NC-BMR protocol. It integrates a data compression-based network coding method with the construction of hierarchical multiparent nodes (HMPNs) topology for the routing with coordinated data forwarding manner and a multi-packets-based time scheduler strategy (MTSS). Its perfect transmission efficiency is achieved by only attaching a little control information with data packets. We validated NC-BMR based on the TOSSIM platform and compared it to several previous methods. Theoretical analysis and simulation results demonstrate its performance improvement in terms of the transmission reliability, delay, and overhead.

## 1. Introduction

The link losses and random node failures are common issues for reliable data transmission in wireless sensor networks (WSNs). Therefore, the routing protocol design is a challenging issue for WSNs, and many works [1–4] have been proposed for reliable data transmission from various perspectives. Among those works, multipath routing has been exploited to improve the routing reliability. It discovers and establishes multiple paths from source to destination nodes via a single routing process and leverages multilink redundancy to improve the transmission reliability.

In multipath routing, the node-disjoint multipath routing builds an optimum path as the primary path and also multiple suboptimal paths. When one path fails, another suboptimal path supplements for transmission [5]. With the node-disjoint multipath routing, the reliability, communication

efficiency, and network scalability can be significantly improved. However, the majority of node-disjoint multipath routing methods are only applicable to networks with dense node distributions. When a network becomes sparse, the performance of this method cannot be guaranteed [6]. Instead of using an alternative disjoint path, the braided multipath routing (BMR) [5] improves reliability through creating a backup node for each node on the main path. If any node in the primary path fails, the backup node is used to guarantee path connectivity. Many braided multipath routing-based protocols [7, 8] require a sink to broadcast the routing paths, which leads to excessive energy consumption to the centralized nodes. To address these deficiencies, Toledo and Wang [9] proposed dynamic next-hop node selection, by which multipath routing was formed hop-by-hop and a node can select the next-hop node from the candidate forwarding nodes set at the moment of forwarding. Both braided

multipath routing and node-disjoint multipath routing improve routing reliability through node and link redundancy. However, they also cause more energy consumption and resource usage during routing. Additionally, topology formation mechanism can enable each node to obtain the information of its neighbors and determine the forwarding candidate node set. Sun et al. [6] proposed a distributed width-controllable braided multipath routing (WC-BMR) based on local neighbor's information for data collections in WSNs. The experimental results show its higher reliability and efficiency.

To improve the previous routing scheme, Tarique et al. [5] proposed a network coding-based routing scheme in which an intermediate node can perform encoding on the received packets. It shows the network performance improvement on the reduction of transmission overhead. Network coding is usually employed to improve throughput and transmission efficiency by encoding to reduce the number of data copy delivered during data transmission. Several studies [7, 10] also demonstrate that the network coding can improve transmission reliability and reduce data redundancy. In [11], a model for multipath and redundancy-based linear network coding scheme was presented, and a hop-by-hop retransmission strategy was developed to guarantee transmission reliability. Xu et al. [12] proposed a collaborative data collection with opportunistic network erasure coding, which delivers encoded packets in batches via iterative encoding to reduce its decoding complexity. Furthermore, there are still several deficiencies in this field. First, in a network coding-based routing, an intermediate node usually needs to perform a large amount of encoding or decoding operation; meanwhile, it is hard to ensure the linear correlation of encoding coefficients, which easily cause excessive energy consumption, transmission overhead, and reduction of decoding success ratio. Second, a large number of retransmissions are required to guarantee reliability against the relatively high packet loss ratio, which results in low transmission efficiency. Finally, transmission collision often occurs and reduces the transmission efficiency.

In this paper, to improve the above deficiencies, through integrating innovations on the network coding, topology formation with braided multipath routing, we propose a network coding-based braid multipath routing (NC-BMR) for data collection in WSNs. Our design inherits the advantages of braid multipath routing. The proposed NC-BMR protocol improves the existing approaches on several aspects which targets to handle the above deficiency.

Our major contributions are detailed as follows.

(1) First, to degrade the excessive energy consumption and transmission overhead on intermediate nodes, a data compression-based random linear network coding scheme and algorithm are proposed.

(2) Then, an efficient HMPN topology formation and HMPN-based coordinated data forwarding scheme are developed, which allow parent nodes to choose the best main route dynamically as well as avoid data retransmission while guaranteeing transmission reliability.

(3) Next, a multipackets-based time scheduler scheme (MTSS) strategy is investigated to reduce the transmission collision in WSNs router protocol.

(4) At last, the proposed schemes or subalgorithms integrate and cooperate to yield the presented NC-BMR protocol. Validation and experimental comparison are performed using the TOSSIM simulator. The experimental results indicate that the proposed NC-BMR protocol has an advantage in terms of transmission reliability, delay, and overhead.

The rest of this paper is organized as follows. The related works are introduced in Section 2. In Section 3, the models and assumptions related to the proposed protocol is given first. Then, the data encoding/decoding algorithms, construction of hierarchical multiparent nodes (HMPNs) topology, strategy to the selection of path width, coordinative data forwarding, and time scheduling strategy are proposed, respectively, in Section 4. In Section 5, the performance of NC-BMR is analyzed theoretically. Promising experiment results are given in Section 6, and from the effectiveness and efficiency perspective, some validations and comparisons are performed, which are followed by the concluding remarks and future works in Section 7.

## 2. Related Works

A great deal of research studies [1, 4–7, 9, 13–19] have been carried out on designing routing in WSNs and some new emerging applications [20–25] about them, where multipath routing technique and its improvements are one of the most prominent routing. Braided multipath routing, especially network coding-associated braided multipath routing, has become a powerful tool for solving node failures and link loss because of its low redundancy and high reliability during data transmission. Under the premise of the main development in this paper, we address the instruction of multipath routing and network coding-based multipath routings.

*2.1. Multipath Routings.* By leveraging the redundancy of wireless links, multipath routing can enhance the reliability for end-to-end communications and degrade the control consumption and transmission delay. Multipath routing mainly includes cluster-based multipath routing, opportunistic routings, and braided multipath routing. Multipath routing with cluster technology can improve energy efficiency and extend the lifetime of the network effectively. Yin et al. [14] present a single path (Singlepath) routing, a more deliberated scheme combined with compressed sensing (CS) for wireless sensor networks to improve the data transmission efficiency. Here, joint multipath routing is used to transfer the data from cluster head to base station. Sharma and Jena [4] proposed a cluster-based multipath routing protocol (CBMRP), which utilizes the clustering and multipath techniques to decrease energy consumption and increase the transmission reliability. However, the CBMRP was a single layer cluster. Due to the simple structure of the single cluster, the scalability was poor, which limited its application range; the distributed piecewise search path

method for the scale of a larger network, of which inevitably large computation complexity, results in a rapid increase of transmission overhead. To monitor burst events in a kind of reactive WSNs, a multipath routing protocol (MRP) [15] based on dynamic clustering and ant colony optimization (ACO) was proposed. Due to selecting cluster heads according to the residual energy, searching the multiple paths by the ACO algorithm and dynamically choosing a route to transmit data packets, the MRP maximized the network lifetime and degraded the energy consumption. However, in the calculation of the pheromone, it was necessary to get the state information of each node on the path, which reduced transmission rate and increased the transmission delay.

Opportunistic routings is good at handling unreliable and unpredictable wireless links. ExOR [26] is a representative opportunistic routing protocol. The strategy of postponing selection provides multiple opportunities for each spread, thus making progress; as a result, ExOR was able to use long radio links with high link loss rates. Due to the lack of effective mutual recognition and coordination mechanism between the various forwarding nodes, the duplication rate of packets received by the destination node was relatively high. A biased backoff scheme [27] during the route-discovery phase was proposed to find a robustness guidance path, which was able to provide cooperative forwarding opportunities for any pairwise nodes between two adjacent hops. However, reliable reactive routing enhancement (R3E) [27] did not deal well with the problem of data collision between two adjacent hops, which degraded the transmission gain of opportunistic routing.

Braided multipath routing, as a powerful scheme to provide reliable and energy-efficient transmission ways, tolerate link losses and node failures in WSNs. Braided multipath routing, without completely disjoint paths but rather many partially disjoint alternation paths taken, allows nodes or links to interleave over the transmission paths. Yilmaz et al. [7] presented a novel shortest hop multipath (SHM) algorithm for WSNs to generate efficiently shortest hop braided multipath for data transmission or routing with the iterative method, where nodes with less hops to the sink relay the packets without the support of ACK/NACK. Each iterations from the sink at a network level only helps the addition of one hop's marginal nodes into the network. Thus, a new node intends to connect the WSNs at each time, and it inevitably requires a new network flooding, which results in poor scalability. In addition, SHM's construction may perform worse under the situation of high link losses because frequent packet losses affect the convergence of algorithms. Wang et al. [16] developed a reliable braided multipath routing protocol that allowed parent nodes dynamically select the next-hop node according to the current link-achievable quality, which is able to guarantee its reliability at the same time. This scheme effectively degraded the transmission overhead.

*2.2. Network Coding-Based Multipath Routings.* Recently, network coding has been studied well in WSNs [28]. The main contribution of network coding lies in that the intermediate nodes can aggregate or re-encode the received packets [29] such that they degrade the data redundancy and improve the transmission reliability. Random linear network coding (RLNC) [30] provides a theoretically efficient method for data encoding. In the RLNC scheme, each packet should be re-encoded at each hop until it is routed to the sink; since each intermediate node's locality encoding coefficients are independently generated, the encoding coefficients of the origin packets received by the sink may be linearly correlated. The attached deficiency is the computing complexity of decoding operation; thus, the extra accompanying transmission overhead results from the encoding vector. Guo et al. [31] proposed a segment linear network coding (SLNC) scheme. Due to adding constraints to the encoding coefficients and dividing a complex matrix inverse operation into several simple ones at receivers, the decoding complexity dramatically declines. However, since the method ignores considering the packet loss rate and other issues, the transmission reliability cannot be ensured. An energy-efficient broadcast algorithm using network coding for GBR (GBR-NC) [32] was proposed; this algorithm aims to degrade data traffic flow and energy consumption and prolongs the network lifetime. In order to guarantee the transmission reliability, GBR-NC adopts the strategy of hop-by-hop retransmission, which easily causes higher data redundancy and resource consumption between end-to-end communications. A cooperative multicast protocol named MWNCast [29], proposed based on a novel moving window network coding technique, can achieve a better balance between throughput and transmission delays, meanwhile, keeping the packet loss probability and decoding complexity at a very low level without explicit feedback. However, in this method, the transmission reliability is negatively influenced by the window size, thus how to choose an appropriate window size is still a challenge. Different from the previous methods, Jiang et al. [17] considered that network nodes were satisfied with a certain random distribution. In such a case, authors presented an energy-efficient multicast routing algorithm to construct the network topology that network coding requires, which effectively increases its throughput; however, in such algorithm, the computing complexity of the algorithm is high. Johnson et al. [18] developed a routing protocol called dynamic source routing (DSR). DST had low overhead and was able to react quickly to changes in the network, which provided a highly reactive service to help ensure successful delivery of data packets. An end-to-end automatic repeat request (EARQ) routing was proposed [19], and the EARQ routing enhanced the transmission efficiency depending on the mechanism of automatic data retransmission. This retransmission mechanism not only increased the reliability of the data but also increased the transmission overhead.

The differences between the proposed NC-BMR protocol and existing protocols include the following: data compression-based random linear network coding scheme selects the new encoding matrix to guarantee the linear correlation of the encoding coefficients, resulting in improving the decoding success rate; HMPN topology is not

just for the end-to-end pattern but for the many-to-one pattern, one of the popular communication patterns in WSNs; selection of the optimum path width, with respect to parent nodes, in guaranteeing the transmission reliability and decreasing transmission overhead; and HMPN-based coordinating data forwarding allows parent nodes to choose the best main route locally and dynamically against the data retransmission.

## 3. Models and Assumptions

In this section, the models and assumptions used in the NC-BMR protocol are introduced in detail.

*3.1. Probability Model for Receivers.* We use the following model to quantify the probability of receiving packets on a node. Since the signal propagation in WSNs is negatively impacted by factors such as noise, heterogeneous media, and signal attenuation, which decides the probability of receiving transmitted packets. Based on the previous probability model [33], the relations among the probability of received packet $P(d)$, communication distance $d$, and path loss coefficient $\alpha$ were discussed, which are shown in the following equation:

$$P(d) = \begin{cases} 1 - 0.5\left(\frac{d}{R}\right)^{2a}, & d < R, \\ \\ 0.5\left(\frac{2-d}{R}\right)^{2a}, & R \le d < 2R, \\ \\ 0, & d \ge 2R, \end{cases} \tag{1}$$

where $R$ is transmission range, the path loss coefficient $\alpha \in [2, 8]$ is determined as that in [30]. Equation (1) has been validated on its correctness and effectiveness by experiments, and it is known that having error between the received probability and the real probability less than 4% with $\alpha \in [2, 8]$ is acceptable.

*3.2. Assumptions of Energy Consumption.* Power consumption in WSNs primarily results from packets sending/reception, computation storage, and communication [32], among which communication contributes the most. In this paper, we use the effective energy consumption model [13]. The energy consumed by one node transmitting $l$ bit data to a distance of $d$ is formulated as follows:

$$E_{TX}(l, d) = \begin{cases} l \times E_{elec} + l \times \xi_{fs} \times d^2, & d < R, \\ l \times E_{elec} + l \times \xi_{mp} \times d^4, & d \ge R, \end{cases} \tag{2}$$

where $E_{elec}$ represents the energy consumption of processing per one bit data from circuits at the sender; $\xi_{fs} \times d^2$ and $\xi_{mp} \times d^4$ are the energy consumption from the transmission power amplifier based on the free space propagation model and the multipath fading propagation model, respectively; and $R = \sqrt{\xi_{fs}/\xi_{mp}}$ represents the threshold distance determined by the free space propagation model and the multipath fading propagation model. When the

transmission distance $d < R$ is satisfied, the power consumption from the transmission power amplifier is estimated by the free space propagation model; instead, the multipath fading propagation model is used to estimate the power consumption.

*3.3. Assumptions of the Experiment WSNs.* Assume a WSN consisting of $N$ sensor nodes is randomly deployed in an area, where the sink is randomly located. In addition, it still meets the following assumptions: (1) the location information of the sink is known to all; (2) each node with a unique ID has the same capability on sending and receiving data; and (3) each node periodically broadcasts the probe packet to its neighbors, which helps to get the neighbors' ID and identify their hierarchy in the WSNs. Meanwhile, the nodes automatically calculate the packet loss ratio on all link-achievable links.

## 4. NC-BMR Protocol

In this section, the proposed protocol, mainly consisting of the network coding method, topology formation, selection of the optimum path width, coordinated data forwarding, and the multipackets-based time scheduling scheme are discussed.

*4.1. RLNC-Improved Network Coding.* The random linear network coding (RLNC) [30] and the growing Frequent Pattern-Growth (FP-Growth) encoding algorithms [34] are the common network coding algorithms that are employed in the WSNs. The RLNC algorithm has been widely used. To guarantee the successful decoding of packets, RLNC delivers an additional coefficient matrix. When the original data are transmitted in large volume, the coefficient matrix grows accordingly. Therefore, the volume of delivering packets may far exceed the amount of original ones, and resulting in the increase in transmission overhead accordingly. Since a node in WSNs always collects much duplicated data during a period of time, there exists a large amount of data redundancy in WSNs. Thus, it is necessary to realize the network data compression and encode them to improve the network's throughput capacity and reliability. Apparently, data compression is useful before network coding. A data compression-based network coding scheme, consisting of encoding and decoding methodology, is proposed consequently. During transmission, the formation of an encoded packet is listed in Table 1.

In Table 1, the field of sequence denotes the coding sequence, the location information represents the site of nodes, and the other fields are shown as the meaning of their strings.

*4.1.1. Network Coding.* Here, the collected data is first compressed at a source node before transmission. Assume that the collected data at a source node during a period of time are denoted as $X = \{x_1, x_2, \ldots, x_i, \ldots, x_n\}$. The uniform random Bernoulli-like matrix $G$ is initialized to

| Coefficient length | Data length | Coefficient vector | Sequence | Data vector | Location information |
| --- | --- | --- | --- | --- | --- |
| | | | | | |

compress the data, which the matrix elements show as in equation (3). The two values each randomly account for one-half in the matrix $G$.

$$g_{ij} = \begin{cases} +\dfrac{1}{\sqrt{n}}, & \text{with probability } \dfrac{1}{2}, \\ & \qquad\qquad\qquad g_{ij} \in G, \\ -\dfrac{1}{\sqrt{n}}, & \text{with probability } \dfrac{1}{2}, \end{cases} \quad (3)$$

where $n$ is the total number of the collected data.

After the matrix $G$ is determined, the $m$ rows in $G$ are randomly selected as the compression matrix without losing the original information due to the insignificant duplication of the collected data. Next, the $m$-dimensional compressed data $Z$ is obtained via equation (4), wherein $m < n$.

$$Z = \begin{bmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_m \end{bmatrix} = GX = \begin{bmatrix} g_{11} & \cdots & g_{1n} \\ g_{21} & \cdots & g_{2n} \\ \vdots & \ddots & \vdots \\ g_{m1} & \cdots & g_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}. \quad (4)$$

Since the compression matrix is a random Bernoulli matrix, its completely linear independence cannot be guaranteed. Namely, if the source node directly delivers data $Z$, the linear independence of the compression matrix is also unable to be held, i.e., $\text{rank}(Z) = m$, which is likely to cause data restoration failure. To overcome this deficiency, the elements in $G$ are alternatively reconstructed via equation (5), i.e., the identifications (IDs) of the source nodes that sends packets in the WSNs are defined as diagonal elements of the matrix; the node's $x$ coordinate $\text{node}_x$ and $y$ co-ordinate $\text{node}_y$ are defined as matrix minor diagonal elements. This alternative scheme keeps the coding sequence of the same nodes consistent across the overall process of encoding and decoding operation. The reconstructed compression matrix $G'$ is shown in equation (5), wherein the matrix element is replaced with the meta-data including the identification and coordination of nodes. Additionally, since $G'$ is an asymmetric matrix, the remaining elements are filled with the corresponding elements at the same position in the previous matrix $G$.

$$g'_{ij} = \begin{cases} \text{ID}, & i == j, \\ \text{node}_x, & i == j+1, \\ \text{node}_y, & i == j-1, \ g'_{ij} \in G, \end{cases} \quad (5)$$

$$Z' = G'X. \quad (6)$$

Thus, the compressed matrix in equation (4) is replaced by the reconstructed matrix $G'$; as a result, a new compressed matrix is obtained, which is re-expressed as equation (6).

If one intends to reconstruct the $n$-dimensional vector from the $m$-dimensional subvector, wherein $m \ll n$, it is necessary to satisfy the condition of $m > CS * \log(n)$ [33], where $C$ is a constant and $S$ means that the coefficient of $x$ can be nonzero at most options of $S$. In contrast, the data collected by sensors are usually $S$-sparse; thus, the following theorem holds.

**Theorem 1.** *Assuming that $G'X = Z$, $G'$ is a matrix of equation (5), $X$ is the collected data, $y$ is a $m$-dimensional vector, any $2S$ column in $G'$ are linearly independent, and then the $S$-sparse vector $x$ can be only reconstructed by $Z'$ and $G'$.*

*Proof.* That can reconstruct the two components of $x$, $x'$ at the same time to meet $Gx = Gx' = Z$, where $x$ and $x'$ are $S$-sparse; then, $G'(x - x') = 0$, because the number of zero elements is less than or equal to $2S$, so $x - x'$ is $2S$-sparse, and for any given condition, $2S$ columns in $G'$ are linear independent, with $G'(x - x') = 0$ contradiction, so the assumption does not hold, and the proposition is proofed.

Here, we call an incident out-going data $Z' = \{z'_1, z'_2, \ldots, z'_i, \ldots, z'_m\}$ of the source data on source nodes. In the RLNC scheme, a set of $m$ packets with the identical length needs to be transmitted. A source node encodes all the packets by the random linear coding method, generating a new encoded packet.

$$y' = \sum_{i=1}^{m} \gamma_i \cdot z_i, \quad (7)$$

where the encoding coefficients $\gamma_i (1 \leq i \leq m)$ are chosen randomly from a Galois field. In other words, each delivered packet is a linear combination of the $m$ encoded packets, and the $m$ encoding coefficients are represented by a $1 \times m$ vector which is called the encoding vector of packets. When an encoded packet is broadcasted, its encoding vector is attached in the header and broadcasted simultaneously. All encoded packets generated by the same data source (e.g., the same node) form a new coding sequence expressed as $Y(y'_i \in Y)$. Obviously, the length of the coding sequence equals to the number of source packets.

The above process is summarized in Algorithm 1, a data compression-based encoding algorithm for source nodes under the assumption of an available optimum path width.

Usually, it is impossible to improve the performance of data transmission in the network if the random linear coding is performed in different coding sequences; the previous studies [14] also proved it. In this paper, the reconstructed compression matrix (e.g., equation (6)) records the location information of nodes, thus keeping the consistency of coding sequences from source to destination node available. Hence, the intermediate node only allows encoding the packets belonging to the same coding sequence. When the number of received packets by the intermediate node is enough, then the intermediate node re-encodes the forwarded packets according to the following equation:

$$Y' = \sum_{j=1}^{m} \beta_j \cdot z_j, \quad (8)$$

where $Y'$ is the new encoded packets yielded from the $y'$ by random linear coding algorithm on the intermediate node

---

**Input**: *Signal X*
**Output**: *Coding packets set CP*
Sink node generate compressed matrix $(G'_{ij})_{m \times n}$
(1) Use equation (5) and broadcast $(G'_{ij})_{m \times n}$
(2) Determination of the *ForwardedNum* in terms of the optimum path width
(3) Use equation (6) to calculate compressed data $Z'$
(4) for $i = 1$; $i < ForwardedNum$; $i$++ do
(5)     Selecting coefficients $\alpha_i$ in $GF(2^8)$
(6)     Use equation (7) to generate new encoding packets $y'_i$
(7)     broadcast the encoded packet $\{COLength, DataLength, \alpha_i, seq_i, y'_i, location\}$ of $CP_i // *$ *seq* means the sequence in Table 1
(8) end for

---

ALGORITHM 1: Encoding algorithm on source nodes.

and $\beta_j$ is also the encoding coefficient randomly selected from the Galois field. In combination with equation (7), it can be found that the encoded package $Y'$ is also generated by the random linear coding method, thus the encoded packet $Y'$ also belongs to the same coding sequence against being changed.

Algorithm 2 is the data compression-based encoding algorithm for intermediate nodes.

*4.1.2. Network Decoding.* Since the network decoding operation is performed on the sink or base station, we call them destination nodes.

In RLNC-associated transmission protocols, the destination node only needs to receive $m$ linearly independent encoded packets with the same coding sequence to decode the original $m$ packets from a source node. If a node in the network has $m$ linearly independent encoded packets that belong to the same coding sequence, the coding sequence order of this node is denoted as $m$. This manner is useful to distinguish the different packets belonging to the different coding sequences. Once the destination node has received encoding packets, it saves them temporarily. When the number of received encoded packets is greater or equal to the coding sequence length of $m$, it may successfully perform decoding operation. Once decoding operation starts, $m$ encoding coefficient vectors should be extracted from all of the received encoding packets; i.e., $m$ encoded packets that belong to the same coding sequence should be selected. The coefficient vectors in these encoded packets are extracted; the decoding coefficient matrix DE is restored depending on these coefficient vectors. The restored decoding coefficient matrix DE is shown in the following equation:

$$
\mathrm{DE} = (\mathrm{de}_{ij})_{m \times m} = \begin{bmatrix} \gamma_{11} & \gamma_{21} & \cdots & \gamma_{m1} \\ \gamma_{12} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \gamma_{1m} & \cdots & \cdots & \gamma_{mm} \end{bmatrix}, \tag{9}
$$

where $\gamma_i$ is the field of "coefficient vectors" in Table 1 which is extracted from the encoded packets.

The independence among encoding vectors is usually critical to the corresponding packet decoding. If $m$

independent encoding coefficient vectors $\gamma_i$ cannot be extracted from the received packets at a destination, then decoding fails. Lin et al. [35] has shown that as long as the number of encoded packets exceeds the coding length and Galois field (GF) capacity only needed to be set as $2^8$, and it can achieve a rather high success decoding rate in practice. Therefore, the GF capacity is also set as $2^8$ in this paper.

After the decoding coefficient matrix is successfully achieved, the source packets will be decoded via $Z' = Y\mathrm{DE}^{-1}$. Since packets have been compressed on the source node, once it performs success decoding operation, data should continue to undergo a restoration operation. In fact, a restoration computing solves the equation $X = G'^{-1}Z'$. According to the description in Section 4.1.1, this computation restores $n$-dimensional data from $m$-dimensional data ($m \ll n$). Here, if $X$ could be sparsely decomposed, the restoration equations could be solved via the proposed approximate solution strategy. In this paper, the orthogonal matching pursuit (OMP) algorithm [36] is employed for it. Consequently, this solution process is summarized as Algorithm 3.

Note that the compression matrix $G$ is not a part in the transmitted data packets; namely, it does not have the transmission overhead since it yields in the phase of topology formation and broadcasts via the sink to each node initially. Therefore, it is unable to increase the communication overhead.

*4.2. HMPN Topology Formation.* In HMPN topology, each node needs to initialize its information, which includes the following 5 tuples {*node ID, node hierarchy, parent node list, adjacent node list, subnode list*}. Each node will receive the information of its neighbor nodes and select the candidate forwarding nodes set based on the quality of link. HMPN topology allows a node partition of different hierarchy and lays the foundation for coordinated data forwarding.

HMPN topology is extended with hierarchy linking to form a hierarchical topology. In the topology formation, there are four types of message format defined to support the topology construction: probe message, probe acknowledges (ack), level update (lup), and level update response (lupack). The message format is shown in Table 2. The detailed topology formation process comprises three phases: first, the

**Input**: Received packets set $RP$
**Output**: Encoding packets set $EP$
(1) Determination of the *ForwardedNum* in terms of the optimum path width
(2) for $i = 1$; $i <$ ReceivedNum; $i++$ do
(3)  Selecting coefficients $\beta_i$ in GF$(2^8)$
(4)  Use equation (8) to generate new encoding packets $Y'_i$
(5)  broadcast the encoded packets $EP_i\{COLength, DataLength, \alpha_i, seq_i, Y'_i, location\}$ of $EP_i$
(6) else
(7)  Selecting *ForwardedNum* packets with the same *seq* and broadcast them
(8) end for

ALGORITHM 2: Encoding algorithm on intermediate nodes.

**Input**: *Packets set SP*
**Output**: *Signal SR*
(1) **if** $PacketsNum \geq Coding\,packets\,Num$ **then**
(2)  Selecting $m$ packets with the same *seq* in $SP$
(3)  **if** cannot select $m$ packets **then**
(4)   **Break**
(5)  **else**
(6)   Use equation (9) to calculate DE
(7)   Calculate $Z' = YDE^{-1}$
(8)   Use **OMP** to recovery the Signal $SR$
(9)  **end if**
(10) **end if**

ALGORITHM 3: Decoding algorithm.

sink node sets its hierarchy as "0" and broadcasts the hierarchy information to its neighbor nodes; second, after these neighbor nodes have received this broadcasted message, they set their hierarchy as "1" and respond to the sink node with a probe acknowledges (ack) message; and finally, after all neighbor nodes complete the hierarchy configuration, the sink broadcasts a level update (lup) message, and then a neighbor node that has received this message starts the next cycle of hierarchy updates. The process continues in such order until all nodes in WSNs successfully receive their hierarchy information, hereto, topology formation is done.

The interaction between each message shows that a probe message is a small link-layer broadcast packet. A node in the network determines its hierarchy via the probe message. A probe acknowledges (ack) message is a response to a probe message. After a node has successfully received a probe message and set the message-sending node as a parent node, it responds to the message-sending node with a probe acknowledges (ack) message. A level update (lup) message is generated by the sink to notify the corresponding hierarchy to start sending probe messages. A level update response (lupack) message is a response to a level update (lup) message. This message has two types of format including positive (lupack) and negative (lupnack) responses, which are distinguished with the lable of Lupflag. This message is also the flag on whether HMPN topology is accomplished or not.

In order to show the advantages of the above topology formation scheme, an example is given as follows: a network consisted of 100 nodes is shown in Figure 1. Figure 1 shows that HMPN topology is an infrastructure that resembles an inverted tree. Most of the edge nodes have more than two parent nodes, which ensure that each edge node has sufficient forwarding opportunity. In Figure 1, the sink node coordinate is (50, 50); a black hole emerges in the oval area between the sink and node 63; i.e., a non-intermediate node exists between node 63 and the sink, thus the sink cannot directly deliver hierarchy information to node 63. However, node 63 is still able to establish a link-achievable link with the sink via node 73 or node 10 in a multihop communication path, so it prevents link losses from occurring. Consequently, a hierarchy multiparent node topology is still available to a sparse network, that is, even if a black hole emerges, nodes in the network are capable of link-achievable to the sink via multihop paths.

All nodes including the sink simultaneously perform HMPN topology formation according to Algorithm 4.

Additionally, the HMPN accomplishes the encoding operation during the slot time before the generation of the final hierarchy formation (e.g., during lines 31–37). Thus, such an approach does not extend the time delay for end-to-end communication. Finally, since all nodes are unlikely to perform the same operation, the time complexity of the Algorithm 4 is approximately $O(myParents)$. $myParents$ also indicates the number of intermediate nodes on the main route of braided multipath routings, i.e., the path width.

*4.3. Selection of the Optimum Path Width.* Usually, the transmission path width is restricted by the number of intermediate nodes on the main route of braided multipath routings. Further, it is well known that communication dominates the energy consumption in WSNs. Reducing the total volume of data transmissions as much as possible becomes the objective of any energy-efficient methods. Although, network coding facilitates the reduction of the number of data copy needed to be transmitted, just network coding is not sufficient to guarantee the transmission reliability. There are two other factors to be concerned. One is how many packets are proper to form an encoded packet. Too small or too large data segments are unable to take advantage of the potential capability of network coding technique and even take up much storage space that is impractical for little sensors. The other is how many paths are needed to guarantee at least effectiveness successfully

TABLE 2: Four message formats.

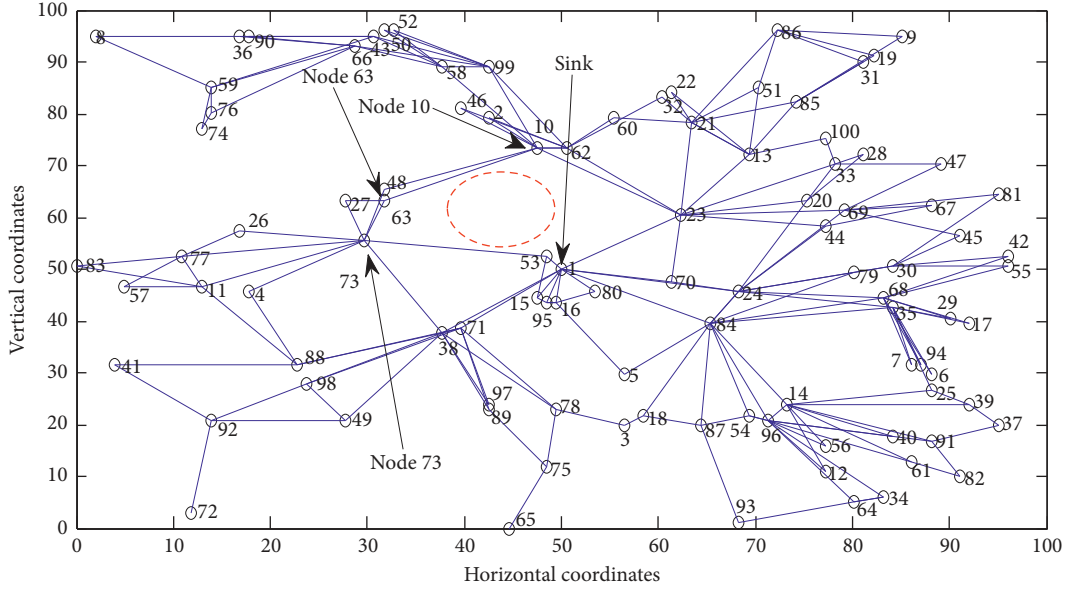| Type | SenderID | Sequence number | Level information | Ack flag |
| --- | --- | --- | --- | --- |
| Probe | SendID | SeqNum | Level | NULL |
| Ack | SendID | SeqNum | Level | NULL |
| Lup | SendID | SeqNum | Level | NULL |
| Lupack | SendID | SeqNum | Level | Lupflag |



FIGURE 1: Hierarchical multiparent node topology.

delivering paths for desire reliability. Therefore, it is feasible to determine the path width through the selected parent nodes indirectly. Hereto, we find an alternative way to compute the optimum number of parent nodes and then recessively control the path width. This way integrates two factors on the number of delivered packets and the packet loss rate between the intermediate node and its neighbors comprehensively. Especially, the covered packet loss ratio targets to calculate the proper required number of parent nodes; the number of packets required at each hop is limited with the size of its parent nodes. Thus, such way enables a parent node to adaptively select the proper number of delivered packets, substantially resulting in controlling the "path width" of data transmission. The theoretical supplement analysis is as follows.

Assume that the packet loss rate for each link is $P_{\text{link}}$; $P_u$ is node $u$'s possibility of success transmission between a hop and its next hop, which is shown in the following equation:

$$P_u = 1 - \prod_{j=1}^{|N_u|} P_{\text{link}}, \tag{10}$$

where $N_u$ is the parent node set of node $u$ and $|\bullet|$ represents its size. Thus, $P_s$ is the possibility of success delivery rate of single packet to the sink that is defined in the following equation:

$$P_s = \prod_{i=1}^{h_u} \left( 1 - \prod_{j=1}^{|N_i|} P_{\text{link}}^{ij} \right). \tag{11}$$

Further, to ensure the success delivery rate of data packets forwarding to the sink is not less than $\sigma_t$ and $\sigma_t$ is a user-predefined threshold, the desired success delivery rate on each node $u$ is $(\sigma_t)^{2^{-i}}$. Consequently, according to equation (11), $E_u$ as the possibility of the success delivery rate of a single packet to the sink is shown in the following equation:

$$E_u = \prod_{i=1}^{h_u} (\sigma_t)^{2^{-i}} = (\sigma_t)^{\sum_{i=1}^{h_u} 2^{-i}} = (\sigma_t)^{1-2^{-h_u}}, \tag{12}$$

where $h_u$ represents the total hops from source node to node $u$. Since $(\sigma_t)^{1-2^{-h_u}}$ is greater than $\sigma_t$, in order to achieve the probability that the packets are successfully received by the destination node is not less than $\sigma_t$, the success delivery rate on each node $u$ in the main path is constrained with $P_u \geq (\sigma_t)^{2^{-i}}$.

To satisfy the constraint condition $P_u \geq (\sigma_t)^{2^{-i}}$, the intermediate node selects the former $\rho$ nodes with the lower packet loss rate as the real forwarding nodes from the candidate forwarding nodes set (e.g., parent nodes) and the node with the lowest packet loss rate in the candidate forwarding nodes set are taken as the members on the main path. Meanwhile, the selected forwarding nodes are assigned a priority, the ones with the smallest packet loss rate their priority is "0," and the others with the highest packet loss rate their priority is "−$\rho$." The ones with a higher priority preferentially perform packet delivery.

The scheme for determining the number of delivered packets on an intermediate node is also investigated.

```
(1) Sink broadcast probe message
(2) Upon sink receives message from node u
(3) if messagetype == ack then
(4)     if sum of received ack messages == neighbor number then
(5)        broadcast level update message in order to give the turn to first level
(6)     end if
(7) else if messagetype == lupack
(8)     if sum of received lupack and lupnack messages == neighbor number then
(9)        broadcast level update message in order to give the turn to next level
(10)    end if
(11) else if messagetype == lupnack then
(12)    if sum of received lupack and lupnack messages == neighbor number then
(13)    if all received messages are lupnack messages then
(14)       Terminate
(15)    else
(16)       broadcast level update message in order to give the turn to next level
(17)    end if
(18)    end if
(19) end if
(20) end Upon
(21) Upon node u receives message from node v
(22) if messagetype == probe then
(23)    if message.level < v.level then
(24)       Insert node v into parentlist and send ack message to node v
(25)    end if
(26) else if messagetype == lupack then
(27)    if sum of received lupack and lupnack messages == neighbor number then
(28)    broadcast lupack to parents
(29)    end if
(30) else if messagetype == lupnack then
(31)    if sum of received lupack and lupnack messages == neighbors' number then
(32)       if node u has not received any lupack message then
(33)          call Algorithm 2 and broadcast lupnack to parents//∗Network coding
(34)       else
(35)          broadcast lupack to parents
(36)       end if
(37)    end if
(38) else if messagetype == lup then
(39)    if node v is one of the parents of node u then
(40)       if node u has no neighbor then broadcast lupnack message to parents
(41)       else if level of node u == received level value in lup message then
(42)          broadcast probe by the level value in message = level of node + 1
(43)       else if level of node u < received value in lup message then
(44)          broadcast received lup message for forwarding to children nodes
(45)       end if
(46)    end if
(47) else if messagetype == ack then
(48)    if ackNum == Neighbor number then
(49)       broadcast lupack message to parents
(50)    end if
(51) end if
(52) end Upon
```

ALGORITHM 4: HMPN topology formation.

For example, assuming that a packet from a source node reaches the destination node after two hops, the number of selected parent nodes is $n_{parent}$, and each parent node successfully sends an encoded packet; the success decoding rate on the destination node is calculated according to the following equation:

$$P_{decode} = \sum_{j=m}^{n_{parent}} C_{n_{parent}}^{j} \left( 1 - P_{link} \right)^{\left( n_{parent} - j \right)} \left( P_{link} \right)^{j}, \qquad (13)$$

where $m$ is the number of delivered packets by source node and $n_{parent} > m$.

An example is further given to show its correctness. In terms of equation (13), the relation between the number of parents and the decoding success rate is shown in Figure 2.

In Figure 2, $N_p$ indicates the number of encoding packets and $p$ indicates $P_{\text{link}}$. Figure 2 shows that the decoding success rate initially increases slowly as the number of parent nodes grows, and then the success rate rises rapidly. While $P_{\text{decode}}$ exceeds 0.95, the uptrend of decoding success rate gradually stabilizes, i.e., delivering more amount of packets impossibly improve the decoding success rate but inevitably increase extra transmission overhead. This phenomenon inspires us that an appropriate number of parent nodes is enough to guarantee the decoding success rate from another aspect, although the number of parent nodes in practical applications may be limited and deficient.

The previous works [4, 9] have also proved that delivering more amounts of packets impossibly improve the decoding success rate but inevitably increase extra transmission overhead. An appropriate number of parent nodes is enough to guarantee the decoding success rate, although the number of parent nodes in practical applications may be limited and not enough. NC-BMR leverages the different packets with the same coding sequence to proactively supplement this deficiency. Hence, the criterion to select the optimum number of forwarded packets on an intermediate node is presented, i.e., from equation (14) we get the proper number of $m'$, where $m' > m$. In combination with the equation (13), we can successfully get the optimum number of parent nodes.

$$P_{\text{decode}} = \prod_{j=1}^{n_{\text{parent}}} \sum_{k=n_{\text{parent}}}^{m'} C_{m'}^k \left(1 - P_{\text{link}}\right)^{(m'-k)} \left(P_{\text{link}}\right)^k. \quad (14)$$

Therefore, the procedure of determining the optimum number of forwarded packets on intermediate nodes is ongoing increased until $P_{\text{decode}}$ satisfies the selection criterion $P_{\text{decode}} \geq \sigma^t$. As a result, selection of parent nodes is also treated simultaneously. The descriptions above integrate on Algorithm 5.

### 4.4. Coordinated Data Forwarding.
An improved coordinated data forwarding scheme is also proposed inspired by the mechanism [8]. In which the initial phase selects its parent nodes for a source node by the proposed Algorithm 5, and then the source node simultaneously delivers the packets to multiple parent nodes. The coordinated data forwarding mainly comprises two steps: (1) the source node encodes the collected data and broadcasts the new encoded packet and (2) the packets received by the intermediate node are delivered according to the cooperative forwarding mechanism. The corresponding packet format is defined in Table 3. Specially, an encoded packet (e.g., the fields in Table 1) is carried in the field "payload."

The coordinated data forwarding mechanism can take full advantage of the transmission opportunity on each hop and limit the path width. Therefore, under maintaining the transmission reliability, the transmission is controllable and the transmission overhead is limited.
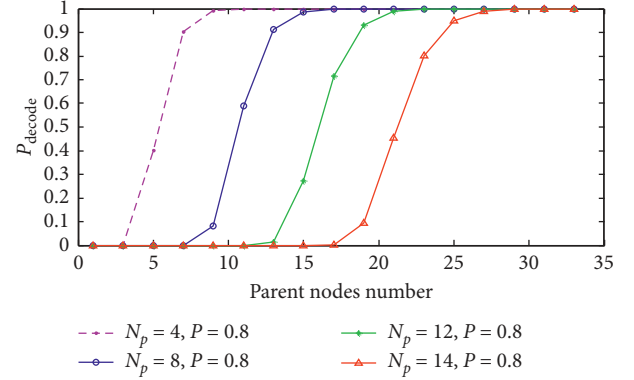


FIGURE 2: Success decoding rate vs. number of parent nodes.

```
Input: σ_t, Neighbor set Nu, Parents set myParents
Output: myParents, Forwarded packets number m'
(1)  σ̃_t ⟵ σ_t;
(2)  Sort packet loss rates in Nu in the increasing order as
     {p_ui} ∈ Nu;
(3)  P_e ⟵ 1;
(4)  For ρ ⟵ 1 to |Nu| do
(5)      P_e ⟵ P_e × p_uρ;
(6)      Insert the node in Nu corresponding to p_uρ into
         the set myParents;
(7)      If P_e ≤ 1 − σ̃_t then
(8)          |myParents| ⟵ ρ; and exit;//*|myParents|
             represents the number of the selected parent
             nodes, namely, the size of path width.
(9)      End if;
(10) End for
(11) m' ⟵ m //*m is the number of the data packets
     delivered by source nodes
(12) While p_decode < σ̃_t
(13)     calculate p_decode by equation (14);
(14)     increase m';
(15) End while
```

ALGORITHM 5: Parents and Sendlist selection for each node $u$.

To demonstrate its works clearly, an evolution example from [8] is given out. In Figure 3, the source node 1 intends to deliver packets to the destination node 9. Visually, nodes 3, 6, and 9 are the optimal parent nodes of nodes 1, 3, and 6, and then nodes 1, 3, 6, and 9 build the main path. In this situation, assuming node 1 broadcasts a packet {1, 1, 1, 3, 9, 1, PayLoad}, this packet will be received by nodes 2, 3, and 4. Once node 3 receives the packet, since the field of mainrouterID (e.g., the main path) is 3 and node 3 is the first one of the parent node set for node 1, then node 3 updates and broadcasts the packet of {1, 1, 3, 6, 9, 1, PayLoad}. Unlike node 3, node 2 receives the same packet, since the field of mainrouterID (e.g., the main path) is 3 and node 2 is not the first one of the parent node set for node 1, and node 2 updates the fields of "Flag, SenderID, and mainrouterID" as "0, 2, 6," respectively and broadcasts the updated packet of {1, 0, 2, 6, 9, 1, PayLoad}. Node 4 performs the similar operation like that of node 2, and then, node 5, 6, 7, and 8

TABLE 3: Transmission data formation.

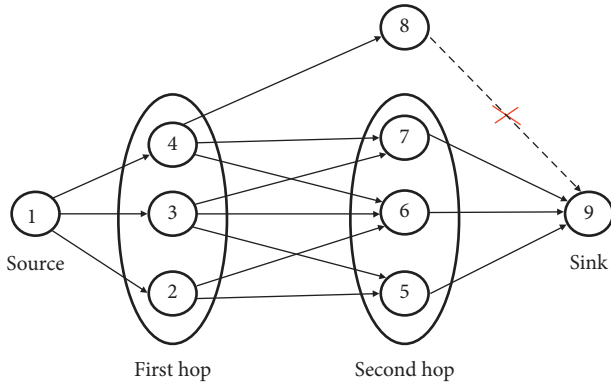| OriginID | Flag | SenderID | MainrouterID | SinkID | SeqNum | PayLoad |
|----------|------|----------|--------------|--------|--------|---------|
|          |      |          |              |        |        |         |



FIGURE 3: Cooperative forwarding mechanism.

will receive the above three or some of the updated packets, similar to the operation examined on nodes 2, 3, and 4. At this time, since node 6 is the optimal parent node of node 3, node 6 will update and deliver the packet of {1, 1, 6, 9, 9, 1, PayLoad}. Node 5 and 7 will forward packets {1, 0, 5, 9, 9, 1, PayLoad} and {1, 0, 7, 9, 9, 1, PayLoad}, respectively. In contrast, after node 8 receives the updated packet, since neither the field of mainrouterID is 1 nor node 3 is its neighbors, the updated packet will be discarded. Consequently, the destination node 9 may only receive packets

from node 5, 6, and 7. As can be seen from the above example, the cooperative data forwarding takes full advantage of the transmission opportunity on each hop and limits the path width. Therefore, while maintaining the transmission reliability, the transmission overhead is reduced.

Summarizing the above process as the coordinated data forwarding algorithm, it is shown in Algorithm 6.

*4.5. Time Scheduling Strategy.* Transmission collision is one of the major factors that negatively impact the transmission reliability and cause data retransmission. Time scheduler scheme is capable enough to avoid transmission collision. In addition, the time scheduler scheme (TSS) [8] can effectively prevent transmission collision and deliver a packet within a predefined time of period rather than treating wireless channel transmission delay. Thus, it only adapts to single data flow transmission. When a parent node needs to deliver multiple packets, the TSS poorly prevent transmission collision and may cause insignificant transmission delay. To improve the TSS to treat multiple data flow, in this paper, the number of forwarded packets $n_{\text{packets}}^i$ is considered, $n_{\text{packets}}^i$ predefined time units is added, and the predefined time period is formulized as equation (15). We call it multi-packets-based time scheduler scheme (MTSS).

$$\text{Offset}(i, \theta, \varepsilon) = \begin{cases} \left(\dfrac{1}{2} + |\theta - 1| n_{\text{packets}}^i\right) T_S, & i = 1, \\ \\ \left(\left(\left(M_{i-1} - 1\right) + |(\theta - \varepsilon)|\right)\left(n_{\text{packets}}^i\right) - \dfrac{1}{2}\right) T_S, & \begin{aligned} & 1 < i < h, \\ & -M_i \le \theta \le 0, \\ & -M_{i-1} \le \varepsilon \le 0, \end{aligned} \end{cases} \tag{15}$$

where $\theta$ denotes the priority of the nodes on the $i$-th hop, $\varepsilon$ denotes the priority of the nodes on the $(i-1)$th hop, and the $i$-th hop on the main path has $M_i$ parent nodes. $T_s$ is the time slot. To give its effectiveness, we give the following theorem.

**Theorem 2.** *According to the MTSS-formalized equation (15), there does not exist interference in NC-BMR for end-to-end communication.*

*Proof.* Taking the transmission between two adjacent hops as an example, that is, as shown in Figure 4. Assuming that $N(i, -\theta)$ represents the parent node for the $i$-th hop with priority "$\theta$," the moment that node $N(1, 1)$ has received an encoded packet is defined as point "0" of horizontal ordinate on the time axis. According to equation (15), after node $N(2, 1)$ has received an encoded packet from $N(1, 1)$, it delays $3.5T_s$ prior to transmission. After node $N(2, 1)$ has received an encoded packet from node $N(1, 2)$, it delays $0.5T_s$ prior to

transmission. Node $N(2, 1)$ delivers an encoded packet within a predefined time period. This process shows that nodes $N(1, 1)$ and $N(1, 2)$ do not simultaneously deliver an encoded packet. After node $N(2, 1)$ has received all encoded packets, it forwards packets within a predefined time period. Node $N(2, 2)$ follows an identical procedure. Therefore, after the MTSS is used in NC-BMR, internode transmission overlapping is avoided and communication interference is effectively eliminated accordingly. Similarly, the situations between any adjacent hops are identical.

*4.6. Sketch of the Overall Proposed Scheme.* Integrating the overall proposed scheme on the above subalgorithms yields the final NC-BMR protocol. Namely, the aforementioned subalgorithms corporately perform the different-phased works in the NC-BMR protocol. The flow chart is shown in Figure 5.

(1) **Upon** receiving a packet{OriginID, Flag, SenderID,MainrouterID, Seqnum, PayLoad}
(2) **If** $u$ is the sink **then**
(3)   **If** has enough packets
(4)     call the decoding procedure;
(5)   **Else**
(6)     waiting for next package;
(7)   **End if**
(8) **Else if** senderID is in *myChildren* set **then**
(9)   **If** *flag* equals 1 **then**
(10)     **If** *senderID* is the first one in the parentlist of *senderID* **then**
(11)       **If** has enough packets
(12)         update *flag* and *mainrouterID* with 1 and $u$, respectively;
(13)       **Else**
(14)         reencode the packets and generate new packets, update *flag* and *mainrouterID* with 1 and $u$, respectively;
(15)       **End if**
(16)     **Else**
(17)       **If** has enough packets
(18)         update *flag* and *mainrouterID* with 0 and the first ID in the parentlist of *senderID*, respectively;
(19)       **Else**
(20)         reencode the packets and generate new packets, update *flag* and *mainrouterID* with 0 and the first ID in the parentlist of *senderID*, respectively;
(21)       **End if**
(22)     **End if**
(23) **Else if** *mainrouterID* is in myChildren set **then**
(24)   **If** *senderID* is the first one in the parentlist of *mainrouterID* **then**
(25)     **If** has enough packets
(26)       update *flag* and *mainrouterID* with 1 and $u$, respectively;
(27)     **Else**
(28)       reencode the packets and generate new packets, update *flag* and *mainrouterID* with 1 and $u$, respectively;
(29)     **End if**
(30)   **Else**
(31)     **If** has enough packets
(32)       update *flag* and *mainrouterID* with 0 and the first ID in the parentlist of *senderID*, respectively;
(33)     **Else**
(34)       reencode the packets and generate new packets, update *flag* and *mainrouterID* with 0 and the first ID in the parentlist of *senderID*, respectively;
(35)     **End if**
(36)   **End if**
(37) **End if**
(38)   generate new packets {*originID, flag, mainrouterID,senderID, seqnum, payload(i)*} and broadcast them;
(39)   **Else**
(40)     discard this packet;
(41) **End if**
(42) **End upon**

ALGORITHM 6: Transmissions procedure for each node $u$.

## 5. NC-BMR Protocol Performance Evaluation

In this section, the theoretical analysis of the proposed NC-BMR protocol is given on the transmission reliability, overhead, and delay, respectively.

*5.1. Transmission Reliability.* The scheme of best- and worst-scenario analysis benefits analyzing the stochastic phenomenon of problems. It is adopted to obtain the upper and lower boundaries of the reliability of the NC-BMR protocol. In best-case, it is always link-achievable for any pairwise nodes between adjacent hops; i.e., even if only one route exists, data delivery can be accomplished successfully. In worst-case, it is link-achievable only for the pairwise node on the main route between adjacent hops, i.e., the reliability completely depends on the nodes and links on the main route.

Assume that there are $\mathrm{hop}_{\max}$ hops transmission between a source node to sink; the $i$-th hop on the main path has $M_i$ parent nodes; node $u$ needs to forward $m'$ packets; the coding length is $m$; $P_{\mathrm{node}}$ is the node failure rate in this path; and $P_{\mathrm{link}}$ is the link loss rate. There exists the following theorem. Theorem 3 regards node failures and Theorem 4 relates link losses.

**Theorem 3.** *Under consideration of node failures, corresponding to the node on the $i$-th hop in the network, the upper and lower boundaries of the transmission reliability are shown in the following equation:*
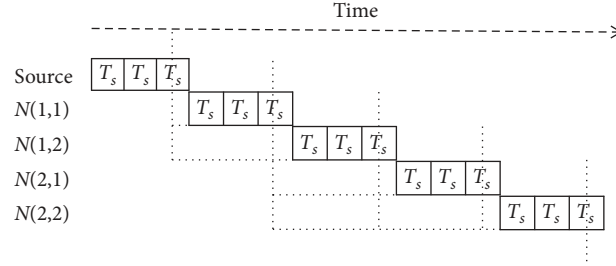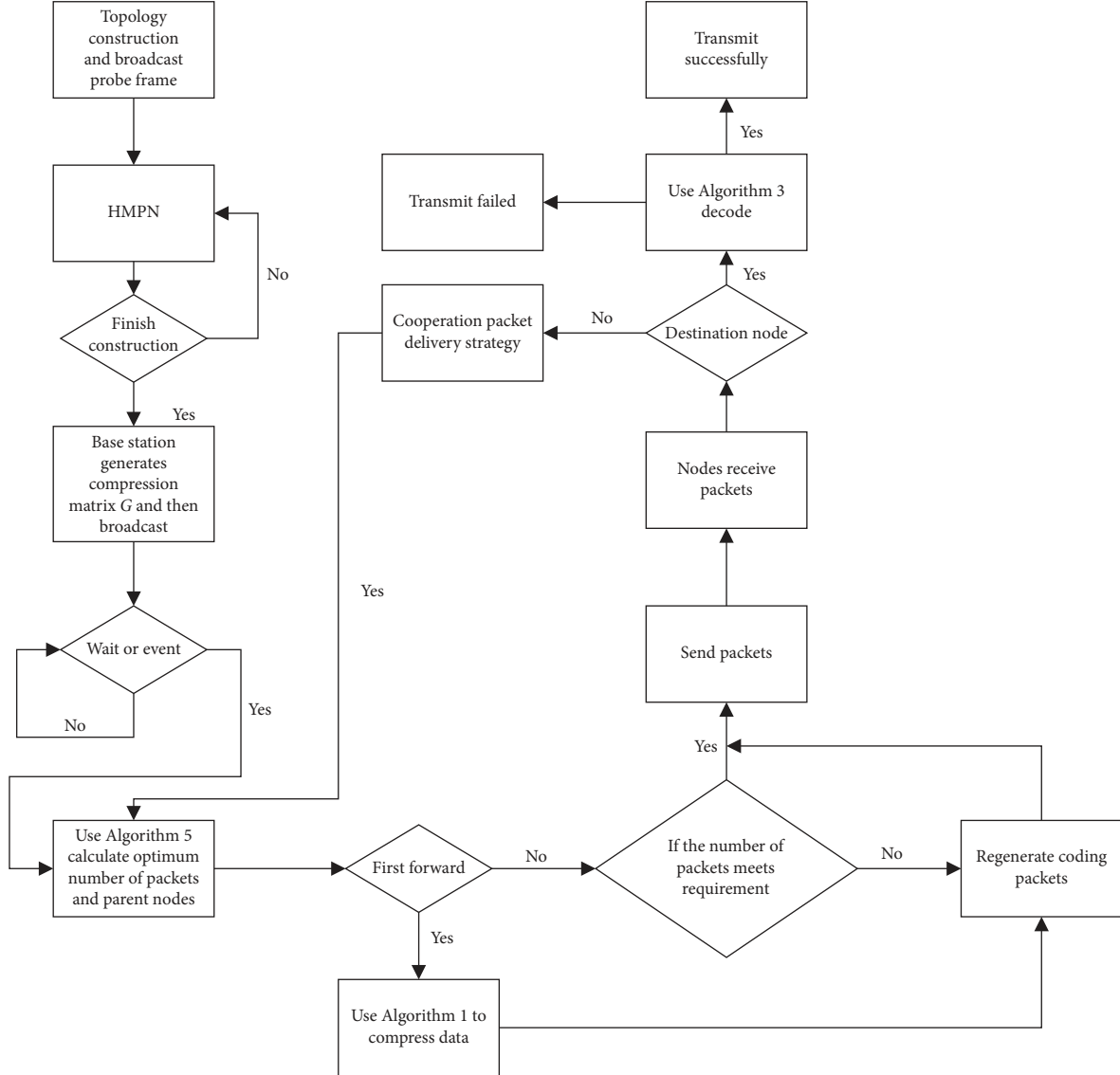
Figure 4: Example of MTSS.



Figure 5: NC-BMR protocol.

$\forall h, h \in \{1, 2, \ldots, \text{hop}_{\text{max}}\},$

$$\left( C_{m'}^{m} \left( 1 - P_{\text{node}} \right)^{m} \left( P_{\text{node}} \right)^{\left( m' - m \right)} \right)^{h-1}$$

$$\leq R_h \leq \prod_{i=1}^{h-1} C_{m'}^{m} \left( 1 - \left( P_{\text{node}} \right)^{M_i} \right)^{m} \left( P_{\text{node}} \right)^{M_i \left( m' - m \right)}. \tag{16}$$

*Proof.* Under the best scenario above, there must be at least $m$ packets received on each hop to guarantee the successful delivery. The possibility that there is at least $m$ packets received on $i$-th hop is $C_{m'}^{m} \left( 1 - \left( P_{\text{node}} \right)^{M_i} \right)^{m} \left( P_{\text{node}} \right)^{M_i \left( m' - m \right)}$. Thus, we can get the upper boundary of transmission reliability of the NC-BMR protocol as $\prod_{i=1}^{h-1} C_{m'}^{m} \left( 1 - \left( P_{\text{node}} \right)^{M_i} \right)^{m} \left( P_{\text{node}} \right)^{M_i \left( m' - m \right)}$.

Under worst-case, the link-achievable links are only for the pairwise node on the main route between adjacent hops, where the reliability completely depends on the nodes and links on the main route. It means that each node on the main route should receive $m$ packets during the transmission, so the reliability is $(C_{m'}^{m}(1 - P_{\text{node}})^m P_{\text{node}}^{(m'-m)})^{h-1}$.

**Theorem 4.** *In the case of link losses, corresponding to the link on the $i$-th hop in the network, the upper and lower boundary of the transmission reliability are shown in equations (17) and (18), respectively.*

$$R_{\text{link}}^i = C_{m'}^{m}\left(\left(1 - P_{\text{link}}^{\text{lc}}\right)^m \left(P_{\text{link}}^{\text{lc}}\right)^{(m'-m)}\right), \qquad (17)$$

*where lc is the total number of link-achievable links on the i-th hop.*

$$\forall h, h \in \{1, 2, \ldots, \text{hop}_{\max}\},$$

$$\left(C_{m'}^{m}\left(1 - P_{\text{link}}\right)^m P_{\text{link}}^{(m'-m)}\right)^{h-1} \leq R_h \leq R_{\text{link}}^1 R_{\text{link}}^{h-1} \prod_{i=2}^{h-2} R_{\text{link}}^i. \qquad (18)$$

*Proof.* First we consider the total number of link-achievable links on the $i$-th hop is as follows under the best scenario.

$$\text{lc} = \begin{cases} M_i, & i = 1, \\ M_i M_{i-1}, & 1 < i < h, \\ M_{i-1}, & i = h. \end{cases} \qquad (19)$$

Under best scenario above, to guarantee one of the nodes on the next hop is capable of receiving the packets, one link-achievable link at least is successful on the $(i-1)$th hop. At this point, the successful transmission possibility is $1 - P_{\text{link}}^{\text{lc}}$, and there must be at least $m$ packets received on each hop to guarantee the successful delivery rate. At the time, the possibility $R_{\text{link}}^i$ on the $i$-th hop is $C_{m'}^{m}((1 - P_{\text{link}}^{\text{lc}})^m (P_{\text{link}}^{\text{lc}})^{(m'-m)})$. Thus, we can get the upper boundary of transmission reliability in the proposed NC-BMR protocol $R_{\text{link}}^1 R_{\text{link}}^{h-1} \prod_{i=2}^{h-2} R_{\text{link}}^i$.

Under worst-case, the reliability completely relies on the links on the main route, the reliability is $(C_{m'}^{m}(1 - P_{\text{link}})^m P_{\text{link}}^{(m'-m)})^{h-1}$.

### 5.2. Transmission Overhead.

The transmission overhead is a typical metric to evaluate the transmission effectiveness of a protocol. Since the transmission overhead is approximately proportional to the number of packets, the transmission overhead in the NC-BMR protocol is indirectly analyzed from the number of delivering packets, as well as the best- and worst-scenario, respectively.

Before analyzing, we first consider the number of children nodes and average number of links. When NC-BMR gets the best case of topology formation where it is always link-achievable for any pairwise nodes between adjacent hops, the number of children nodes corresponding to the *i-th* hop is defined in the following equation:

$$\mu_i = \begin{cases} 1, & i = 1, \\ M_{i-1}, & 1 < i < h. \end{cases} \qquad (20)$$

Assuming that the $n_{\text{packets}}^i$ is the number of parent nodes receiving the packets on $i$-th hop corresponding to $j$ successful links from the $(i-1)$th hop, and $\overline{n_{\text{packets}}^i}$ is its average defined in the following equation:

$$\overline{n_{\text{packets}}^i} = \begin{cases} j, & i = 1, 0 \leq j < M_1, \\ \left\lceil \dfrac{j}{\mu_i} \right\rceil, & 1 < i < h, 0 < j < M_i. \end{cases} \qquad (21)$$

Since there are at most $M_i$ nodes participating in delivering packets, namely, each children node on the $i$-th hop links up to $M_i$ parent nodes and results in $M_i$ link-achievable links, thus the left inequality $n_{\text{packets}}^i \leq M_i$ always holds.

**Theorem 5.** *The upper and lower boundary of the transmission overhead is shown in the following equation:*

$$m + \sum_{i=2}^{h-1} m' M_i \left(1 - P_{\text{link}}\right)^{M_i} \leq n_{\text{transmission}}^{\text{sum}}$$

$$\leq \left(m + m' M_1 \left(1 - P_{\text{link}}^{M_1}\right)\right) + \sum_{i=3}^{h-1} m' M_i \prod_{j=2}^{i}\left(1 - P_{\text{link}}^{M_i M_{i-1}}\right). \qquad (22)$$

*Proof.* As the same that with the transmission reliability analysis in Section 5.1, the upper and lower boundary of the transmission overhead result from the analysis of best- and worst-case in the NC-BMR protocol distinctly. Suppose that $n_{\text{transmission}}^i$ is the average number of packets delivered by parent nodes on the $i$-th hop. Obviously, $n_{\text{transmission}}^i = m$ and $n_{\text{transmission}}^2 = mM_1(1 - P_{\text{link}}^{M_1})$. Further, when $3 \leq i \leq h$, it yields the following equation:

$$n_{\text{transmission}}^i = mM_1\left(1 - P_{\text{link}}^{1}\right)\prod_{k=1}^{i}\left(1 - P_{\text{link}}^{n_{\text{packets}}^i n_{\text{packets}}^{i-1}}\right). \qquad (23)$$

By using the maximum number of $n_{\text{packets}}^i$ (e.g., $M_i$) and the number of forwarded packets (e.g., $m'$), the upper boundary of the number of delivered packets can be obtained with substituting $n_{\text{packets}}^i$ into equation (23), and then it yields the following equation:

$$n_{\text{transmission}}^{\text{sum}} = \left(m + m' M_1\left(1 - P_{\text{link}}^{M_1}\right)\right)$$

$$+ \sum_{i=3}^{h-1} m' M_i \prod_{j=2}^{i}\left(1 - P_{\text{link}}^{M_i M_{i-1}}\right). \qquad (24)$$

For the worst-scenario, it is only link-achievable for the pairwise nodes on the main route between adjacent hops; the transmission reliability completely depends on the nodes and links on the main route. As $3 \leq i \leq h$, the lower boundary is shown as the following equation:

$$n_{\text{transmission}}^{\text{sum}} = m + \sum_{i=2}^{h-1} m' M_i \left(1 - P_{\text{link}}\right)^{M_i}. \qquad (25)$$

### 5.3. Transmission Delay.

Transmission delay is one of the key metric for validating the transmission efficiency of protocols. The transmission delay of NC-BMR is as in Theorem 6.

**Theorem 6.** *The transmission delay of the NC-BMR protocol satisfies the following equation:*

$$\sum_{i=1}^{h-1} \text{Offset}(i, 0, 0) + \frac{h-1}{2} T_s \le D_{\text{delay}}$$

$$\le \sum_{i=1}^{h-1} \text{Offset}(i, -M_i, -M_{i-1}) + \frac{h-1}{2} T_s. \qquad (26)$$

*Proof.* Under best-scenario, the sink only receives packets from the nodes on the main route. At this point, the time delay is $\sum_{i=1}^{h-1} \text{Offset}(i, 0, 0) + ((h-1)/2)T_s$ in terms of MTSS. Under worst-case, the sink only receives packets from the nodes on the secondary route, and the time delay is $\sum_{i=1}^{h-1} \text{Offset}(i, -M_i, -M_{i-1}) + ((h-1)/2)T_s$. Thus, equation (26) holds.

### 5.4. Time Complexity Analysis.

In summary, the time complexity primarily comes from topology formation and sending various messages in different phases. Given $L$ layers in the WSNs, in the $i$-th layer during topology formation, which take *2i* unit time to send updating messages, whereas sending probe messages and receiving the response messages for probe message cost an additional 2 unit time. Thus, the time cost of topology formation in the whole $L$ layers is as follows:

$$t(\text{TC}) = \sum_{i=1}^{L-1} O(2i + 2). \qquad (27)$$

During data message transmission, the time consumption is $i$ unit time. So the time cost of data transmission is as follows:

$$t(\text{DT}) = O(i). \qquad (28)$$

In general, the time complexity of the above two phases is as follows:

$$t(\text{sum}) = \sum_{i=1}^{L} O(2i + 2) + \sum_{i=1}^{L} O(i)$$

$$= O(L). \qquad (29)$$

## 6. Experiment Validation and Analysis

The TOSSIM simulator [37] with TinyOS programming [38] is used to validate our proposed NC-BMR protocol. We integrate our approach into IEEE 802.15.4 protocol and disable the function of ACK/NACk, that is the default setting in most WSNs. The experiment is divided into two parts, including single data flow and multiple data flow. The proposed protocol is validated and compared against the single path (Singlepath) routing [14], single path routing with a retransmission strategy such as end-to-end automatic repeat request (EARQ) [19], and node-disjoint multipath routing such as dynamic source routing (DSR) [18] and braided multipath routing (BMR) [5] on the same indices including success delivery rate, transmission overhead, energy consumption, and transmission delay. Additionally, the packet restoration rate on the proposed NC-BMR protocol is also employed to validate its practical application.

### 6.1. Single Data Flow.

To validate the performance of NC-BMR on single data flow, 500 nodes are randomly deployed in a $100 \times 100$ area, where the maximum transmission hops is assumed to be 10. To facilitate the evaluation of the transmission performance for the minimal network coding, one packet is delivered by a single source node. Since both the path loss coefficient and number of hops affect the protocol transmission overhead and reliability, the experimental comparisons are performed according to the aforementioned indices.

### 6.1.1. Effects of Path Loss Coefficients.

As shown in equation (1), the variation of the path loss coefficient $\alpha$ will change the packet loss rate. In this experiment, a node with 6 hops transmission to the sink is randomly selected as the source node. Each group carries out 100 times of experiments on five different routing protocols, respectively, and takes their average of 100 experimental results as the final results. The results are shown in Figure 6. Figure 6(a) displays the success delivery rate, Figure 6(b) expresses the transmission overhead, and Figure 6(c) shows the energy consumption. As shown in Figure 6(a), since the EARQ protocol implements a retransmission strategy, and the success delivery rate always maintains at 100%, which is the best one. The success delivery rates of the remaining four protocols demonstrate an uptrend with an increase of the loss coefficient. In which, the success delivery rate of the NC-BMR protocol always exceeds 90%. The main reason is that NC-BMR is based on braid multipath routing, which leverages link redundancy to guarantee transmission reliability and achieve a high success delivery rate. Besides the EARQ protocol, in terms of the success delivery rate, compared with the other three compared protocols, the NC-BMR protocol has some advantages. In contrast, both the DSR and Singlepath protocols have the worst performances. The DSR and Singlepath protocols have nonimprovement on the success delivery rates and their packet loss rate declines. After the first hop, the DSR protocol only selects the parent nodes with the lowest packet loss rate from the current candidate forwarding nodes set as the next-hop, and the selected path may not be the optimum path. On the other hand, the path in Singlepath protocol is determined by the sink or source node; although the selected path in it at each hop is the optimum path under the current condition, it does not implement any strategy to improve transmission reliability, which results in low performance. In Figure 6(b), the transmission overhead of the
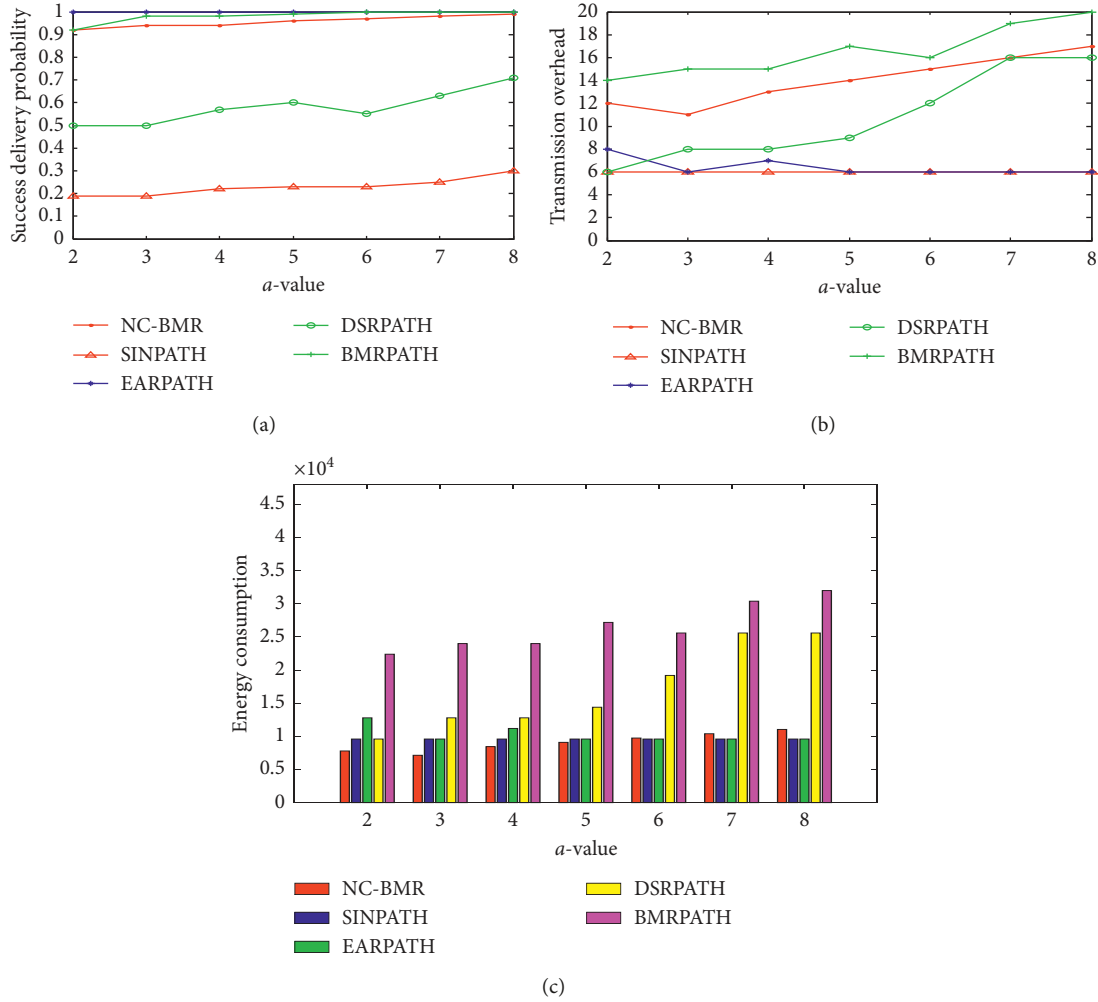
(a)



(b)



(c)

FIGURE 6: Effect of path loss coefficient on transmission performance with single data flow. (a) Success delivery rate with different path loss coefficient. (b) Transmission overhead with different path loss coefficient. (c) Energy consumption with different path loss coefficient.

multipath routing-associated protocol demonstrates an up-trend with the increase of $\alpha$ because the multipath routing protocol always leverages link redundancy to guarantee transmission reliability, which causes an ascendancy of transmission overhead. In contrast, the single path protocol has a lower transmission overhead, which always maintains at a low level. Moreover, the transmission overhead of the NC-BMR protocol is lower than that of the other braided multipath routing-associated protocols because the NC-BMR protocol additionally implements the strategy of selecting the optimum number of forwarded packets when the packet loss rate is low, which degrades the transmission overhead accordingly. Finally, a comparison of transmission energy consumption on five protocols is shown in Figure 6(c). The energy consumption of the NC-BMR protocol is lower than that of the others due to implementation of the data compression-based network coding technology in the NC-BMR protocol. The data compression and network coding technology improve the throughput and save bandwidth, which results in 1/3 to 1/2 times of the energy consumption of the BMR protocol. In the compare protocols, the BMR protocol has the highest energy consumption.

6.1.2. Effects of Hops. In this part, ten nodes within 2–11 hops transmission to the sink are randomly selected for investigation. Each group carries out 100 times of experiments on five different routing protocols, respectively, and takes their average as the final results. The results are shown in Figure 7. Figure 7(a) displays the success delivery rate, Figure 7(b) expresses the transmission overhead, and Figure 7(c) shows the energy consumption. In Figure 7(a), the Singlepath and DSR protocols have the lowest success delivery rate. When the number of hops exceeds six, the success delivery rates of the aforementioned two protocols approach "0," whereas the success delivery rate of the NC-BMR protocol always exceeds 90% because the success delivery rate at each hop is guaranteed due to the mechanism of coordination data forwarding. Therefore, even with a relatively large number of hops, a high success delivery rate always holds. Evidently, the NC-BMR has a distinct out-performance, and its success delivery rate is 20%–28% higher than that of Singlepath and 30%–41% higher than that of DSR. As shown in Figure 7(b), the transmission overhead of five protocols shows an uptrend with an increase of hops. Singlepath and EARQ have the smallest
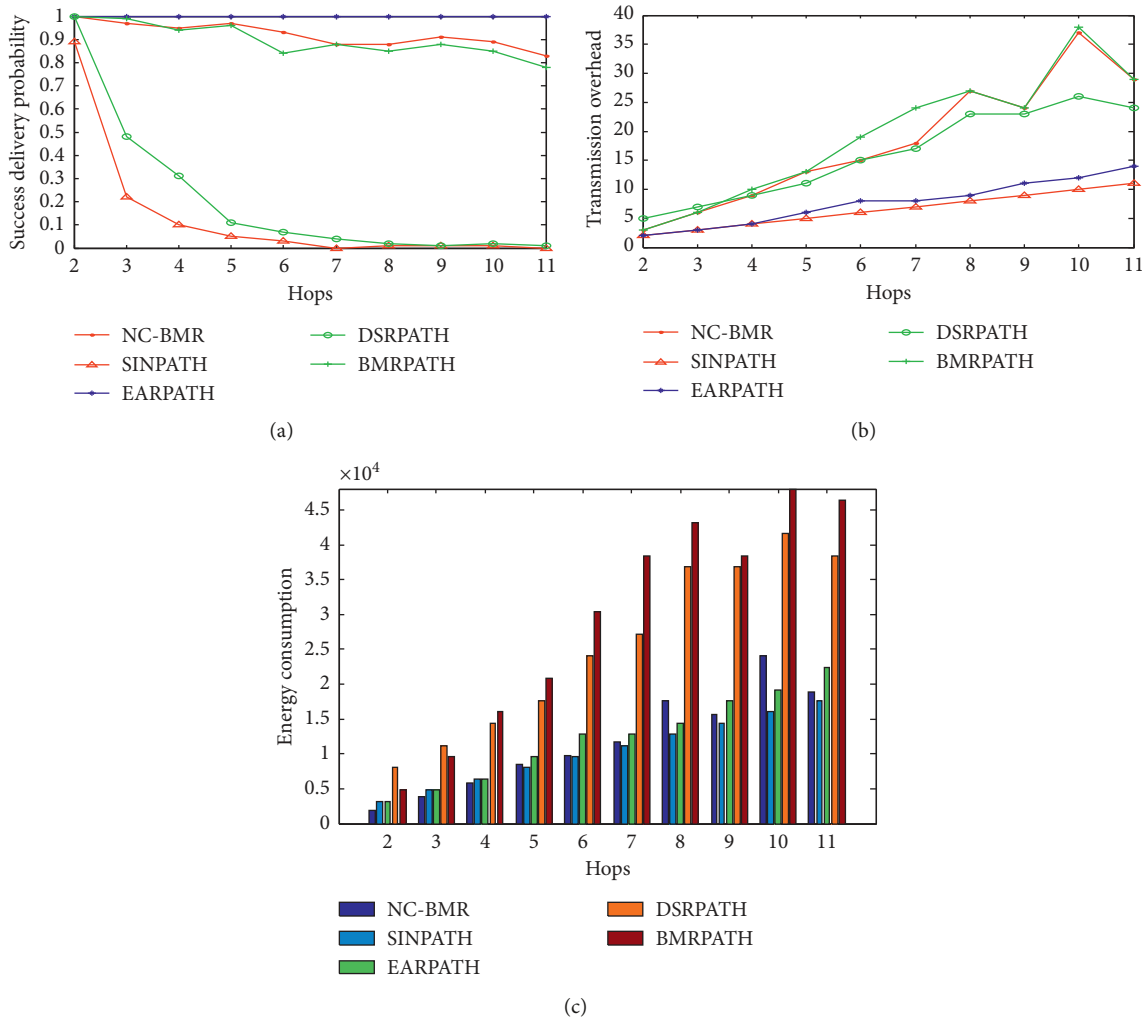
(a)



(b)



(c)

FIGURE 7: Effect of hops on transmission performance with single data flow. (a) Success delivery rate with different hops. (b) Transmission overhead with different hops. (c) Energy consumption with different hops.

transmission overhead, which are 12%–40% lower than that of NC-BMR. The BMR protocol has the highest transmission overhead due to its data retransmission and link redundancy; the NC-BMR protocol is the second highest. The reason caused the NC-BMR with a high transmission efficiency is that it does not need the excessive "re-encoding" and "re-filtering" operation at the intermediate node when single data flow is transmitted by the NC-BMR, which generates a transmission overhead that is similar with that of the BMR. Further, according to the above analysis and the discussion in Figure 7(c), the data flow degrades due to data compression-based network coding technology in the NC-BMR, finally leading to the reduction of energy consumption accordingly. In summary, despite the success delivery rate of the BMR protocol being guaranteed, the transmission overhead and energy consumption significantly increase. In contrast, the NC-BMR protocol selects the next-hop under sufficient consideration of the transmission opportunity and coordination data forwarding mechanism, and therefore, the former $\rho$ items of nodes with high propriety in the candidate forwarding nodes set are selected to take as the real parent

nodes. As a result, the NC-BMR protocol not only guarantees data transmission reliability but also maintains low energy consumption compared with the others.

The analytical comparisons above indicate that NC-BMR inherits the advantages of braid multipath routing, such as high transmission reliability.

*6.2. Multiple Data Flow.* Multiple data flow benefits validate the effectiveness of the data compression-based network coding technology. In this section, the experiment configuration is the same as that in Section 6.1. A source node is selected to deliver multiple packets. The experimental comparisons are performed with different path loss coefficients and number of hops on the same evaluation indices including success delivery rate and transmission overhead.

*6.2.1. Effects of Path Loss Coefficient.* The results are shown in Figure 8. Figure 8(a) expresses the success delivery probability. Figure 8(b) displays the transmission overhead. As shown in Figure 8(a), in terms of the index of
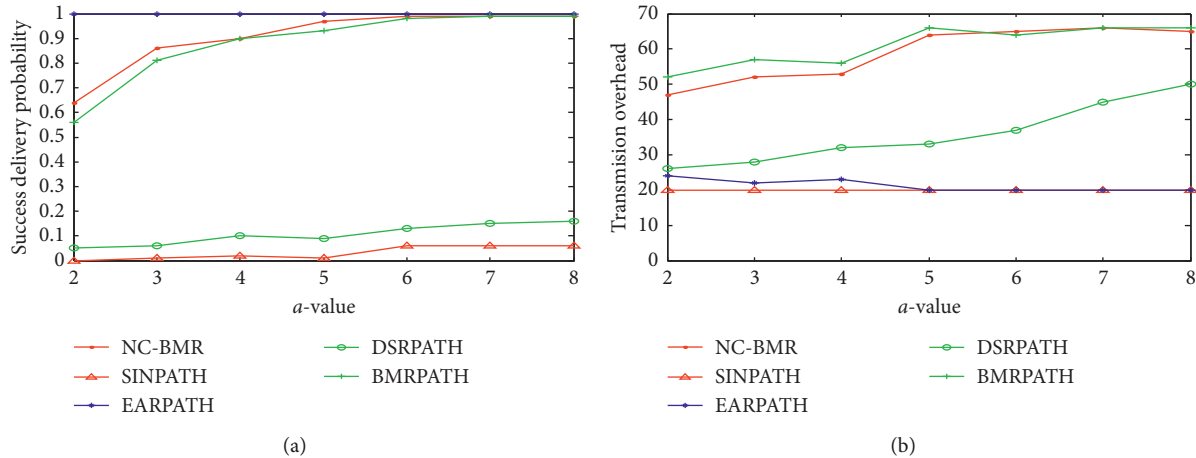
FIGURE 8: Effects of path loss coefficient on transmission performance with multidata flow. (a) Success delivery rate on different path loss coefficient. (b) Transmission overhead on different path loss coefficient.

success delivery probability, since the EARQ protocol implements a retransmission strategy, it achieves the highest success delivery rate, which is always 100%. The NC-BMR and BMR protocols have not implemented similar retransmission strategy but have the second highest success delivery rate; compared with the Singlepath and DSR protocols, they are superior to them. Further, when multiple data flows are delivered by the braided multipath routing-associated protocols, the success delivery rate far exceeds that of the DSR and Singlepath protocols. When the path loss coefficient is $\alpha < 6$, i.e., the packet loss rate is relatively high, the success delivery rate of NC-BMR even exceeds that of the BMR protocol due to the data compression-based networking code technology and coordination data forwarding mechanism. Therefore, the destination node can receive sufficient encoded packets, and the success delivery rate improves accordingly. In contrast, for the BMR protocol [5], since the link redundancy is insufficient, packet losses may occur at the source node, thereby generating the incompletely received data at the destination node and even negatively influence the final broadcast rate. On the other hand, according to the index of transmission overhead that is shown in Figure 8(b), the transmission overhead of the NC-BMR and BMR protocols are higher than the others since the braid multipath routing-associated protocols enable intermediate nodes in different paths to receive packets, which results in redundancy packets and increases the transmission overhead. Moreover, when packet loss rate is relatively high, the transmission overhead of the NC-BMR protocol is higher than that of the BMR protocol [5]. The reason is that the NC-BMR protocol forwards additional data packets to prevent the transmission rate from decreasing with the loss rate, and its transmission overhead gradually approaches that of the BMR protocol. Additionally, the experimental results in Section 6.1 and the related analysis also show that the energy consumption of the NC-BMR protocol is significantly lower than that of the BMR protocol in this situation, which indicates that the NC-BMR protocol can adapt to an environment with a

higher packet loss rate and also is not sensitive to it. Therefore, the NC-BMR protocol is relatively robust and has broader application fields.

### 6.2.2. Effects of Hops.

Figure 9 shows the experimental results affected by the number of hops. Figure 9(a) relates the success delivery probability and Figure 9(b) displays the transmission overhead. In terms of the success delivery rate, the braided multipath routing-associated NC-BMR and BMR protocols have higher ones than that of the DSR and Singlepath protocols with nonbraid multipath routing. Although, the DSR protocol achieves a high success delivery rate due to a small number of hops, as the number of hops increases, its success delivery rate rapidly declines. In contrast, the success delivery rate of the braided multipath routing-associated protocols gradually declines because it increases data redundancy, and these similar approaches improve the correction capability of network errors and the network success delivery rate. On the other hand, in terms of transmission overhead, as shown in Figure 9(b), the NC-BMR protocol has no significant improvement over the BMR protocol because the NC-BMR protocol control the "path width" and loses some potentially significant transmission opportunities. The transmission overhead of the NC-BMR protocol is approximately 10%~15% lower than that of the BMR protocol when the number of hops is relatively large. The caused reason is that the NC-BMR protocol encodes the forwarded packets which results in degrading the throughput and the whole data flow. Hence, under the case with different numbers of hops, the NC-BMR protocol demonstrates promising performance.

The analytical comparison above demonstrates that the performance of all five compared protocols decline to some extent under the multiple data flow. In which, the Singlepath and DSR protocols are suitable for single data flow against multiple data flow. Although the EARQ protocol guarantees a success delivery rate, the energy consumption of EARQ is still higher than that of the proposed NC-BMR protocol.
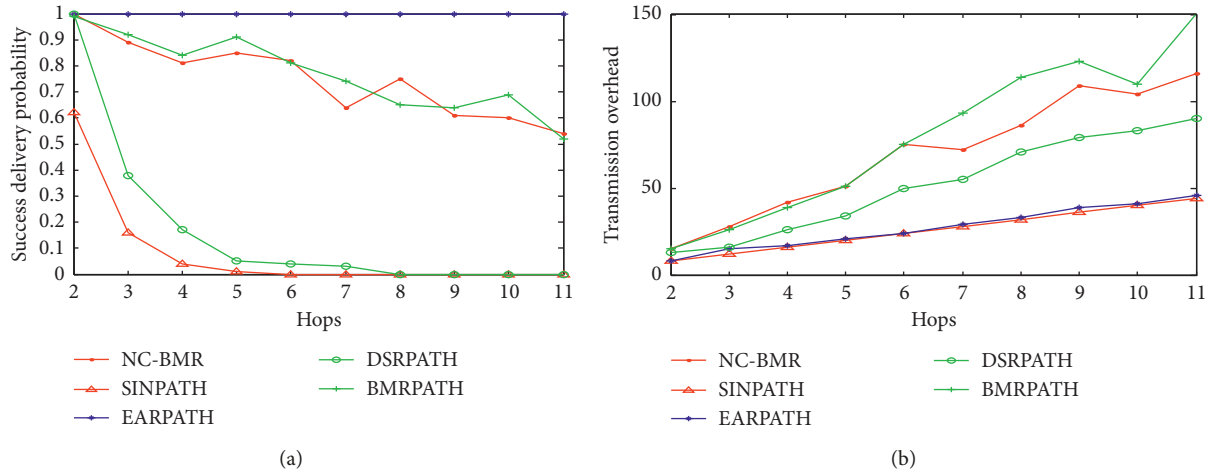
FIGURE 9: Effect of hops on transmission performance with multidata flow. (a) Success delivery rate with different hops. (b) Transmission overhead with different hops.

The NC-BMR protocol has the slowest decline tendency on the two evaluation indices, which indicates that the proposed NC-BMR protocol is more suitable for multiple data flow and even obtains a better transmission performance than that of the BMR protocol under the situation with a high packet loss rate.

*6.3. Transmission Delay.* Transmission delay is one of the primary metrics to validate the delivery efficiency for a protocol. The experiment configuration is identical to that in Sections 6.1 and 6.2. The transmission hops is randomly 6 and 10 to the sink, and the experiments are taken under different path loss coefficient and number of nodes.

*6.3.1. Comparison of Transmission Delay on Five Protocols.* Figure 10 shows the experimental results under different path loss coefficients. Arbitrarily, the number of packets delivered from the source node is four, i.e., $m = 4$. As shown in Figure 10, the transmission delay of the two braided multipath routing-based BMR and NC-BMR protocols have an obvious fluctuation, especially the BMR protocol. The main reasons are below, first, the embedded mechanism of braided multipath routing in them accompany a proper links redundancy to achieve a success delivery ratio; Next, since the complicated multipath selection mechanism, the transmission delay is negatively impacted by the path loss coefficient. It is hard to get a suitable path loss coefficient through training. Finally, the mechanism of braided multipath routing intends to leverage redundancy to improve transmission reliability without considering the transmission delay. These experimental results also validate this viewpoint from another perspective. Further, Figure 10 also shows that the transmission delay of the BMR protocol is slightly higher and more unstable than that of the NC-BMR protocol because the BMR protocol has no controlling over the "path width" and employing the coordination data forwarding mechanism, which likely spur a data storm and increase transmission delay. Additionally, the transmission
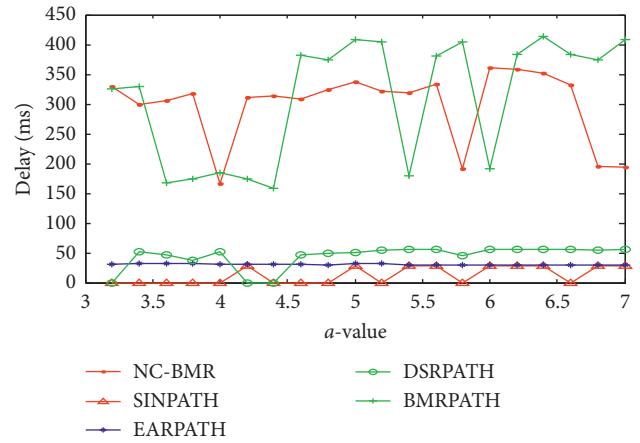


FIGURE 10: Transmission delay comparison on five protocols.

delay "0" indicates that the transmission of the corresponding protocol fails. Single path protocol has the smallest transmission delay because this protocol only implements basic mechanism of single path transmission; i.e., this protocol delivers packets on a fixed path. Therefore, this protocol has relatively stable transmission and transmission delay. The transmission delay of the EARQ and DSR protocols is lower than that of the braided multipath routing-associated BMR and NC-BMR protocols. The caused reason is that the EARQ protocol implements a retransmission strategy and selects an optimum path under the current situation, thus guarantees a smaller transmission delay. And more, the transmission delay of the DSR protocol also stabilizes after the value of $\alpha$ exceeds 4.5. The is because that the selected path in DSR determined by the source node, i.e., this protocol delivers packets via fixed paths, and each path has an identical number of hops to the sink, thus results in the stable transmission delay.

To summarize, the experimental results indicate that the braided multipath routing-based protocols generally have higher transmission delay, which is understandable due to the inherent mechanism. Therefore, the following

experiments only treat the comparisons between NC-BMR with the similar braided multipath routing-based protocols.

### 6.3.2. Effects of the Size of Nodes.

Because the amount of transmitted packets also affects the transmission delay, the experiments randomly select the value of $m$ as 10, 14, 18, and 20, namely, the number of delivered packets from source nodes. The number of nodes in the modified experiments of WSNs falls into $[200, 1200]$. The experiment results are shown in Figure 11. As shown in Figure 11(a), as the number of nodes increases, the transmission delays of the BMR and NC-BMR protocols also gradually rise. The BMR protocol demonstrates a distinct uptrend, whereas that of the NC-BMR protocol gradually stabilizes. The reason is that the NC-BMR protocol utilizes the MTSS strategy to avoid the queue congestion and data collision and also degrades the transmission delay. Further, when the number of nodes is approximately 1,100, the transmission delay of the BMR protocol has an abrupt change because the selected nodes may be located at a special area such as the center of the experiment area, which expands the delivering list and produces exponential growth of number of the links-achievable links. Meanwhile, because the intermediate node at each hop selects sufficient parent nodes due to the increase of node density, the transmission delay of the NC-BMR protocol also increases and increases the transmission overhead. Like that of Figures 11(a)–11(d)have the same variation trend under the different values of $m$. Integrating the analytical comparisons in Figures 11(a)–11(d), although the transmission delay of the NC-BMR protocol increases gradually while the source node delivers more packets, it gets relatively stabilized at the end. The primary reason is that the NC-BMR successfully obtains an optimum main path for delivering packets and utilizes the MTSS strategy to prevent data collision, thus means that has a potential limitation of the transmission delay. Hence, in the worst-case, the NC-BMR protocol is also capable of forwarding data within a predictable limitation period of time.

### 6.3.3. Effects of the Forwarded Packets.

As discussed in Section 6.3.2 and Figure 11, it can be found that when the number of delivered packets on source node is 10, NC-BMR has a relatively smallest fluctuation for the transmission delay. In order to fairly validate the effect of the number of forwarded packets on the transmission delay, we choose different numbers of packets to be forwarded under the same situation with the same number of delivered packets (e.g., $m = 10$) on a source node and show the corresponding performance of the NC-BMR with the same path loss co-efficients. Here, the number of forwarded packets delivered by parent nodes is randomly user-predefined as 16, 18, and 20 to carry out the comparison respectively. The experimental results are shown in Figure 12. With the increase of the number of hops, the 3 cases of different number of forwarded packets, their transmission delay shows an uptrend accordingly, in which the growth trend of the transmission delay in the raw NC-BMR protocol is relatively slow. The primary reason is that the number of forwarded

packets by parent nodes is limited by the "path width," which results in degrading the transmission delay. Figure 12 also shows that, with the number of forwarded packets increasing, the transmission delay also shows an uptrend under different hops. Such phenomenon drives us to draw two pieces of revelations: (1) parent nodes always forward packets in a potentially predefined time as shown in equation (15) and effectively prevent data collision and (2) the strategy of selection of the optimum path width in the proposed NC-BMR protocol efficiently improves the whole performance. Based on the aforementioned experiment analysis, it shows that the MTSS strategy and selection of the optimum path width is effective to degrade the transmission delay and improve the whole transmission efficiency of the NC-BMR protocol.

### 6.4. Practical Application Validation.

Since the encoding and decoding operation is embedded in the proposed NC-BMR, the practical application of the proposed protocol can be indirectly validated by measuring the accuracy of encoding and decoding computing. The positive restoration rate and negative restoration error are employed to be taken as the evaluation indexes. The positive restoration rate regards the effect of the number of encoded packets on the final restoration packets from the source node. The negative restoration error is defined as follows:

$$\mathrm{RE} = \left( \frac{\|\tilde{s} - s\|}{s} \right), \tag{30}$$

where $s$ is the original packet and $\tilde{s}$ is the restored packet.

To validate the encoding and decoding efficiency used in the NC-BMR protocol, two signals with size of 128 data points are employed as the original data segment to be delivered by the source nodes, as shown in Figures 13 and 14, respectively. In which, the horizontal coordinate in the diagram represents the sampling points, and the vertical coordinate represents signal amplitude. Original signal 1, which comprises a series of collected geological vibration data, a real data of time series, is shown in Figure 13. Single 2 in each time sequence is obtained from three easy trigonometric functions as shown in the following equation:

$$\mathrm{Sig}_2(i) = \sin\left(\frac{i}{2\pi}\right) + 0.3 \cos\left(2\pi f_1 T_{\mathrm{sample}} t_s\right)$$
$$+ 0.6 \cos\left(2\pi f_2 T_{\mathrm{sample}} t_s\right) + 0.9 \cos\left(2\pi f_3 T_{\mathrm{sample}} t_s\right), \tag{31}$$

where $f_1 = 50$, $f_2 = 100$, and $f_3 = 200$. The sampling frequency is $f_s = 800$, and the time slice length is $t_s = 1/f_s$. $T_{\mathrm{sample}} = 1 \ldots 128$ is sampling sequence.

The experiment results are shown in Figure 15. With the increase of the number of encoded packets, the restoration error tends to decrease. In Figure 15(a), the restoration error of "signal 1" has a downtrend because "signal 1" is a real signal, which contains multifrequency subsignals, and to restore it with high accuracy, more sampling data are required. Additionally, when the number of encoded packets is 15, the
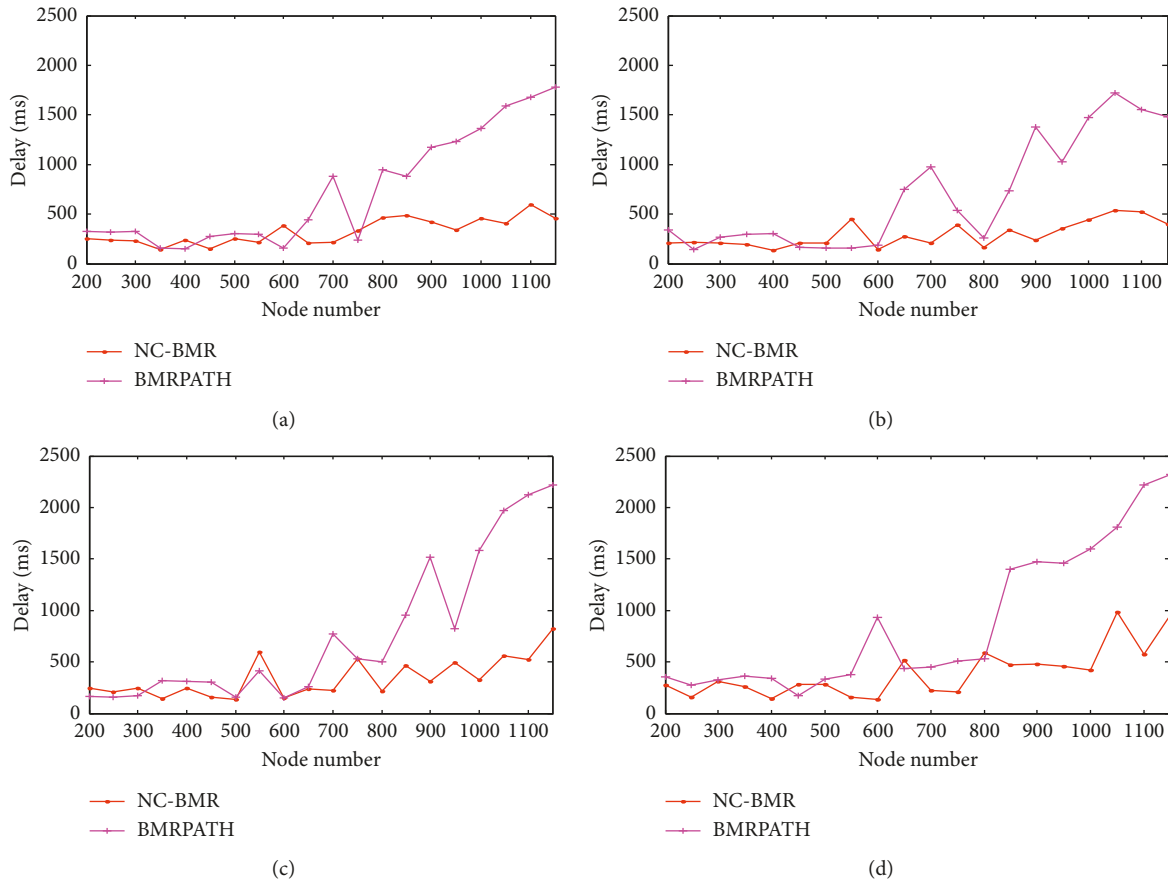
FIGURE 11: Effects of the node scale on transmission delay. (a) $m = 10$. (b) $m = 14$. (c) $m = 18$. (d) $m = 20$.
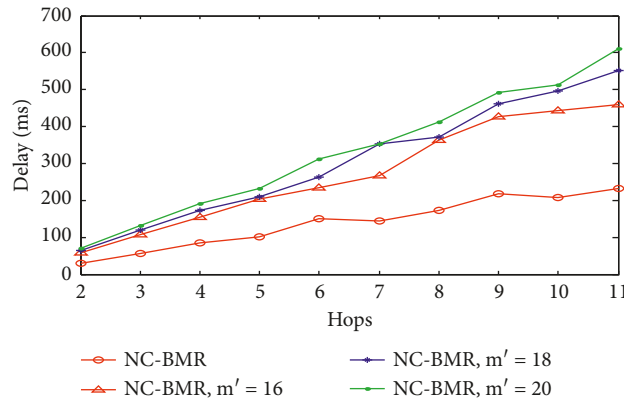


FIGURE 12: Effects of forwarded packet(s) on transmission delay.

restoration error fluctuates, and the primary reason caused this phenomenon maybe is that the feasible candidate parent nodes during transmission are limited. For example, the deployed sensors are relatively sparse in the geophysical field. Thus, each single path needs to forward additional packets; once a link-achievable link with a high packet loss rate is selected, the parent node may fail to receive the delivered packets and results in extra packet retransmission and duplication at the desti-nation node, thus increasing the restoration error. The similar reason also causes a similar situation when the encoded packet is 18 and 19 in Figure 14(a).

Figures 16 and 17 are selected to show the restoration results of "signals 1 and 2" with 11, 14, 17, and 20 items of encoded packets (e.g., the length of coding sequence), re-spectively. When the number of encoded packets exceeds 14, the restoration signal is very close to the original signal; this result leads us to perform the optimization control of "path width" meaningfully.

Based on the above experimental comparison and an-alytical studies, the NC-BMR protocol prominently guar-antees the transmission reliability and evidently demonstrates practical applications.
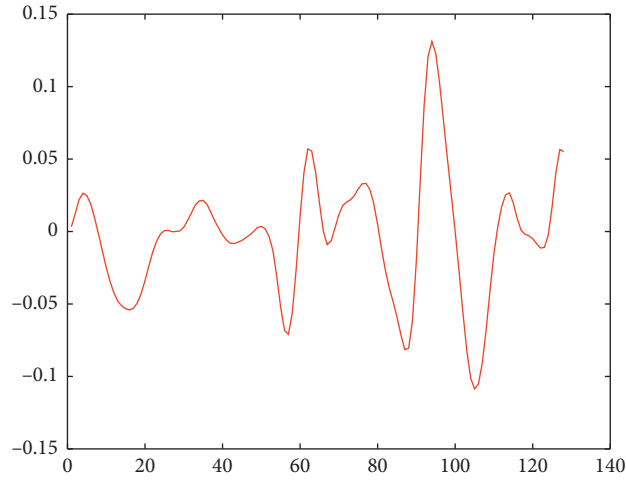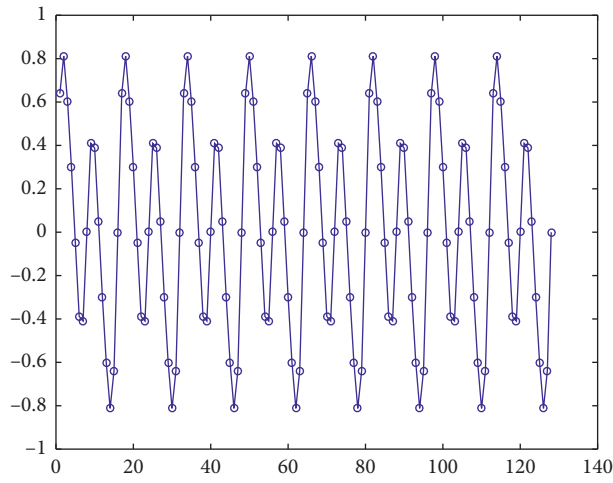
FIGURE 13: Original signal 1.



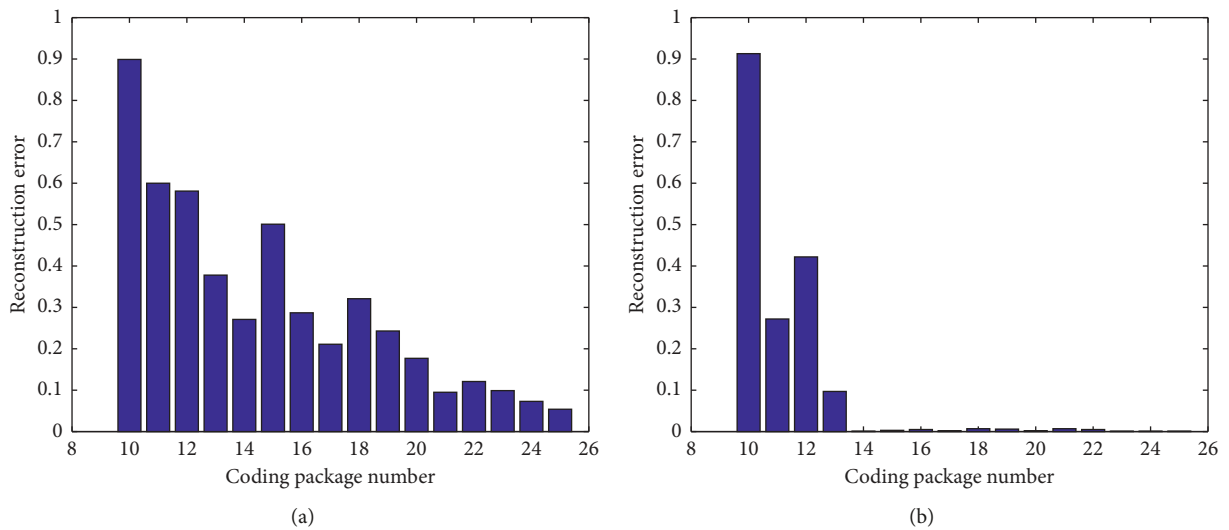FIGURE 14: Original signal 2.



(a)



(b)

FIGURE 15: Effects of different number of data packets on RE. (a) RE on signal 1 under different numbers. (b) RE on signal 2 under different numbers.
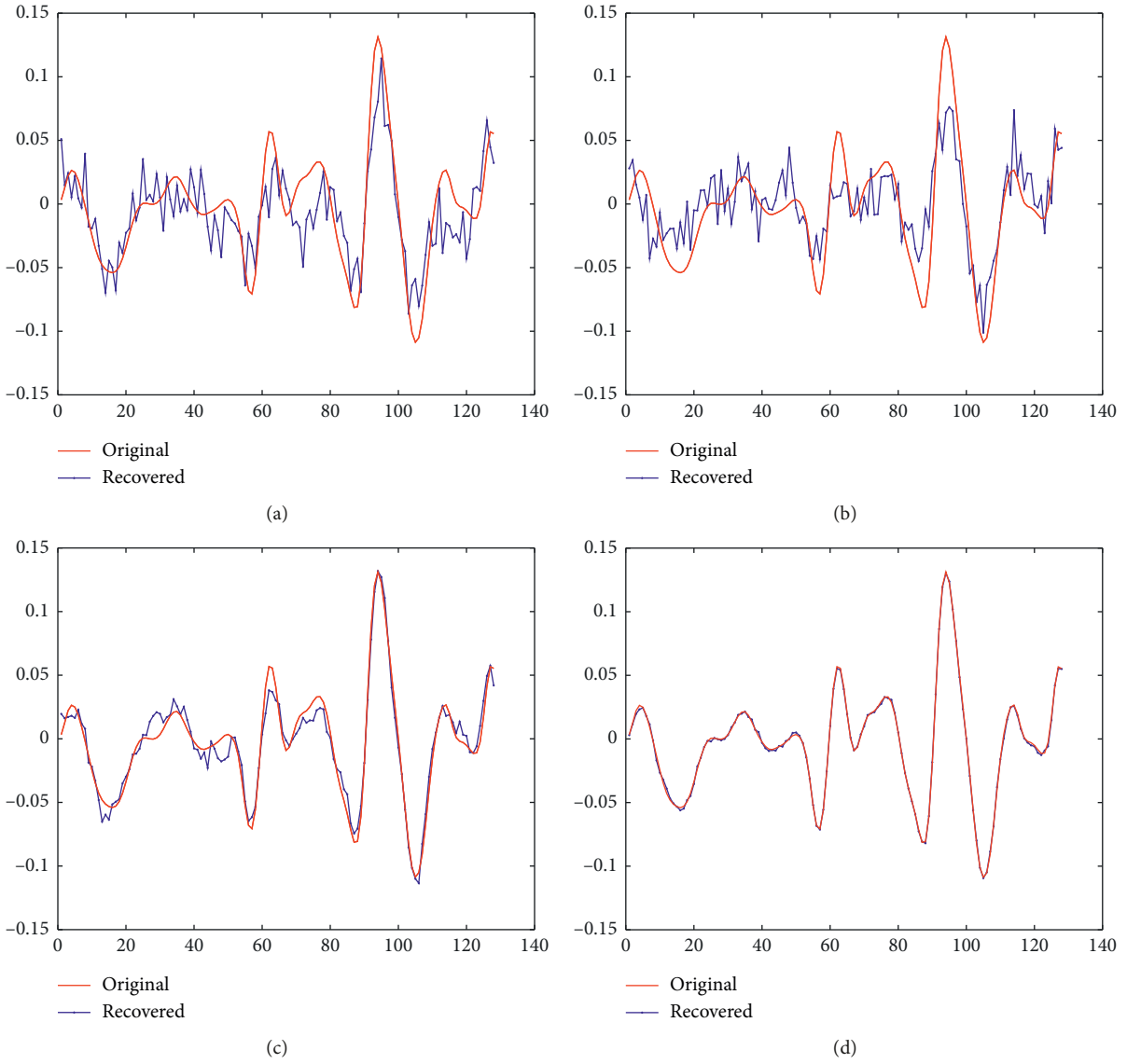
Figure 16: Restoration results on signal 1. (a) The number of encoded packets is 11. (b) The number of encoded packets is 14. (c) The number of encoded packets is 17. (d) The number of encoded packets is 20.
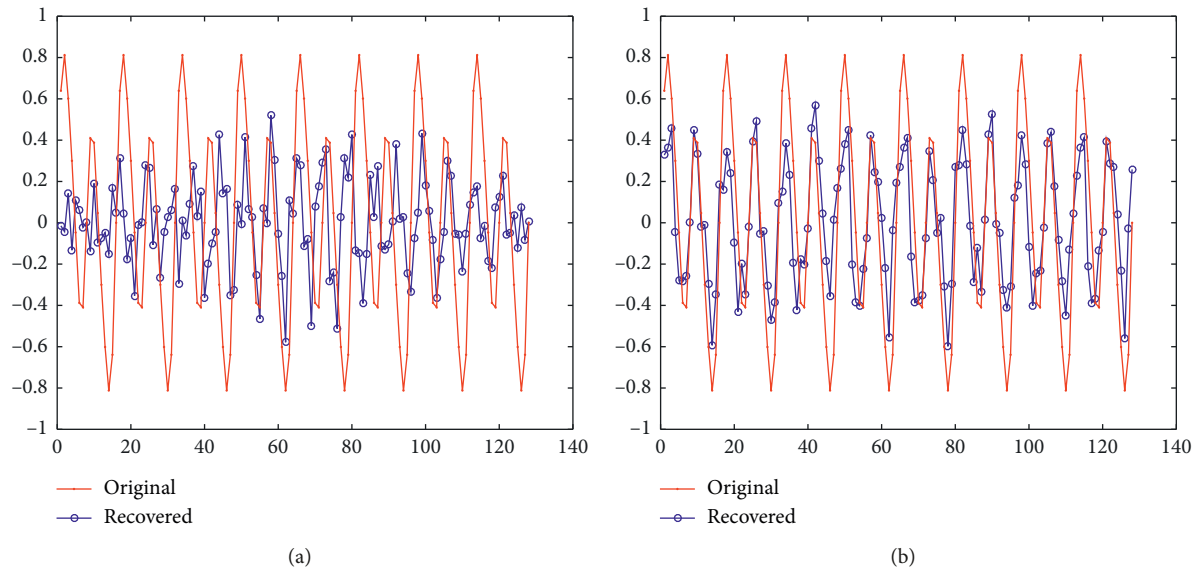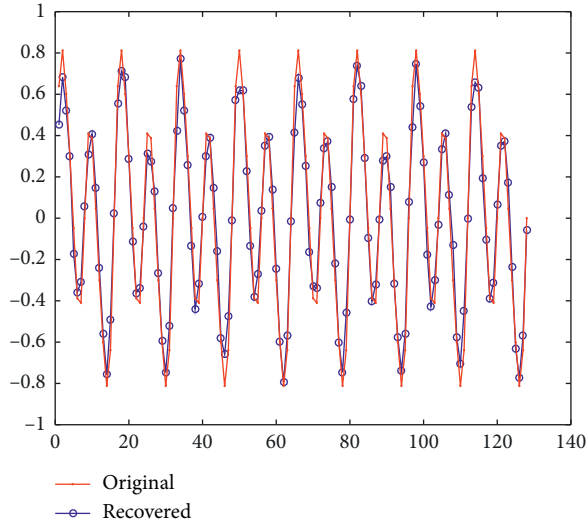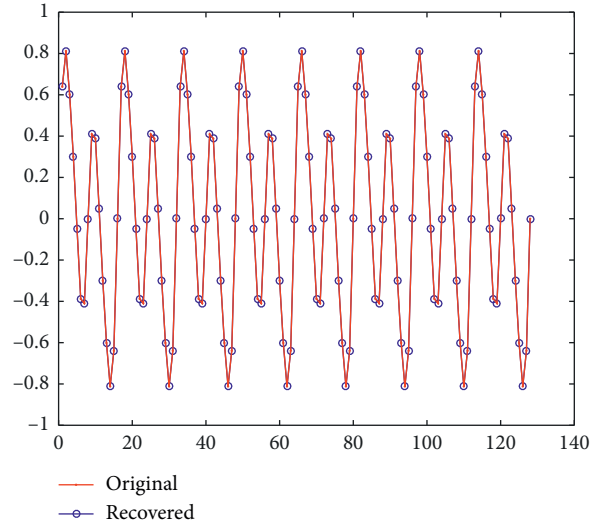


Figure 17: Continued.

(c)



(d)

FIGURE 17: Restoration results on signal 2. (a) The number of encoded packets is 11. (b) The number of encoded packets is 14. (c) The number of encoded packets is 17. (d) The number of encoded packets is 20.

## 7. Conclusions and Future Works

In this paper, we propose a network coding-based braid multipath routing (NC-BMR) for data collections in WSNs. The proposed NC-BMR integrates innovation on the following components: the encoding and decoding manner, hierarchical multiparent node topology formation, selection of the optimum path width, and coordinated data forwarding. Together with efficient network coding computing on intermediate nodes and a small amount of control information attached to data packets, NC-BMR can be implemented in practice in order to reduce the complexity. Besides, an improved time schedule strategy was presented and used to prevent transmission collisions between two adjacent hops. Moreover, by allowing nodes to choose the main route dynamically based on the real-time transmission status, the retransmission and link redundancy have been optimized, of which guarantees the reliability of hop-by-hop. As a result, the parent node set can be chosen locally by each running node to deliver packets in the WSNs. In addition, the performance of the NC-BMR transmission is analyzed in terms of the transmission delay, transmission overhead, energy consumption, and signal recovery rate. Simulation results validate that our NC-BMR protocol can achieve much better performances and obtain an ideal balance between time redundancy and link redundancy.

The network coding is helpful to improve the end-to-end transmission reliability and transmission efficiency, especially in combination with various multipath routing in WSNs. We will be studying this as our future work.

## Data Availability

If some readers want to get the data, they can contact us with their academic email.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.
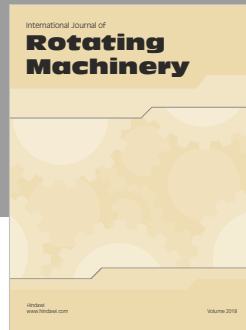
## Acknowledgments

## References

[1] Y. Yang, C. Zhong, Y. Sun, and J. Yang, "Network coding based reliable disjoint and braided multipath routing for sensor networks," *Journal of Network and Computer Applications*, vol. 33, no. 4, pp. 422–432, 2010.

[2] X. Liu, X. Gong, and Y. Zheng, "Reliable cooperative communications based on random network coding in multi-hop relay WSNs," *IEEE Sensors Journal*, vol. 14, no. 8, pp. 2514–2523, 2014.

[3] M. Z. Farooqi, S. M. Tabassum, M. H. Rehmani, and Y. Saleem, "A survey on network coding: from traditional wireless networks to emerging cognitive radio networks," *Journal of Network and Computer Applications*, vol. 46, pp. 166–181, 2014.

[4] S. Sharma and S. K. Jena, "Cluster based multipath routing protocol for wireless sensor networks," *Computer Science*, vol. 45, no. 2, pp. 14–20, 2015.

[5] M. Tarique, K. E. Tepe, and S. Adibi, "Survey of multipath routing protocols for mobile ad hoc networks," *Journal of Network and Computer Applications*, vol. 32, no. 6, pp. 1125–1135, 2009.

[6] X. Sun, H. Chen, X. Wu, X. Yin, and W. Song, "Opportunistic communications based on distributed width-controllable braided multipath routing in wireless sensor networks," *Ad Hoc Networks*, vol. 36, pp. 349–367, 2016.

[7] O. Yilmaz, S. Demirci, Y. Kaymak, S. Ergun, and A. Yildirim, "Shortest hop multipath algorithm for wireless sensor networks," *Computers & Mathematics with Applications*, vol. 63, no. 1, pp. 48–59, 2012.

[8] A. Attir, Y. Challal, A. Hadjidj et al., "Braided disjoint branch routing protocol for WSNs," in *Proceedings of the 2013 Eighth International Conference on Broadband and Wireless Computing, Communication and Applications (BWCCA)*, pp. 106–113, IEEE, Compiegne, France, October 2013.

[9] A. L. Toledo and X. Wang, "Efficient multipath in wireless networks using network coding over braided meshes," *International Journal of Sensor Networks*, vol. 7, no. 3, pp. 176–188, 2010.

[10] L. Shan-Shan, Z. Pei-Dong, L. Xiang-Ke, C. Wei-Fang, and P. Shao-Liang, "Energy efficient multipath routing using network coding in wireless sensor networks," in *Lecture Notes in Computer Science*, pp. 114–127, Springer Science+Business Media, Berlin, Germany, 2006.

[11] M. Ploumidis, N. Pappas, V. A. Siris, and A. Traganitis, "On the performance of network coding and forwarding schemes with different degrees of redundancy for wireless mesh networks," *Computer Communications*, vol. 72, pp. 49–62, 2015.

[12] M. Xu, W.-Z. Song, and Y. Zhao, "Collaborative data collection with opportunistic network erasure coding," *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 10, pp. 1941–1950, 2013.

[13] W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," *IEEE Transactions on Wireless Communications*, vol. 1, no. 4, pp. 660–670, 2002.

[14] J. Yin, Y. Yang, L. Wang, and X. Yan, "A reliable data transmission scheme based on compressed sensing and network coding for multi-hop-relay wireless sensor networks," *Computers & Electrical Engineering*, vol. 56, pp. 366–384, 2016.

[15] J. Yang, M. Xu, W. Zhao, and B. Xu, "A multipath routing protocol based on clustering and ant colony optimization for wireless sensor networks," *Sensors*, vol. 10, no. 5, pp. 4521–4540, 2010.

[16] L. Wang, Y. Yang, and W. Zhao, "Network coding-based multipath routing for energy efficiency in wireless sensor networks," *EURASIP Journal on Wireless Communications and Networking*, vol. 2012, no. 1, 2012.

[17] D. Jiang, Z. Xu, W. Li, and Z. Chen, "Network coding-based energy-efficient multicast routing algorithm for multi-hop wireless networks," *Journal of Systems and Software*, vol. 104, pp. 152–165, 2015.

[18] D. B. Johnson, D. A. Maltz, and J. Broch, "DSR: the dynamic source routing protocol for multi-hop wireless ad hoc networks," *Ad Hoc Networking*, vol. 5, pp. 139–172, 2001.

[19] L. L. Peterson and B. S. Davie, *Computer Networks: A Systems Approach*, Morgan Kaufmann Publishers, Burlington, MA, USA, 2003.

[20] X. Fu, G. Fortino, P. Pace, G. Aloi, and W. Li, "Environment-fusion multipath routing protocol for wireless sensor networks," *Information Fusion*, vol. 53, pp. 4–19, 2019, In press.

[21] G. Yang, M. Haenggi, and M. Xiao, "Traffic allocation for low-latency multi-hop networks with buffers," *IEEE Transactions on Communications*, vol. 66, no. 9, pp. 3999–4013, 2018.

[22] L. Huo and D. Jiang, "Stackelberg Game-based energy-efficient resource allocation for 5G cellular networks," *Telecommunication System*, vol. 23, no. 4, pp. 1–11, 2019.

[23] D. Jiang, L. Huo, Z. Lv, H. Song, and W. Qin, "A joint multi-criteria utility-based network selection approach for vehicle-to-infrastructure networking," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 10, pp. 3305–3319, 2018.

[24] Z. Lv, B. Hu, H. Lv, D. Jiang, H. Lv, and X. Lu, "Intelligent security planning for regional distributed energy Internet," *IEEE Transactions on Industrial Informatics*, p. 1, 2019.

[25] D. Jiang, W. Wang, L. Shi, and H. Song, "A compressive sensing-based approach to end-to-end network traffic reconstruction," *IEEE Transactions on Network Science and Engineering*, vol. 5, no. 3, p. 1, 2018.

[26] S. Biswas and R. Morris, "ExOR: opportunistic multi-hop routing for wireless networks," *CM SIGCOMM Computer Communication Review*, vol. 35, no. 4, pp. 133–144, 2005.

[27] J. Niu, L. Cheng, Y. Gu, L. Shu, and S. K. Das, "R3E: reliable reactive routing enhancement for wireless sensor networks," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 1, pp. 784–794, 2014.

[28] C. Gkantsidis and P. R. Rodriguez, "Network coding for large scale content distribution," in *Proceedings of the 24th Annual Joint Conference of the IEEE Computer and Communications Societies*, pp. 2235–2245, Miami, FL, USA, March 2005.

[29] F. Wu, C. Hua, H. Shan, and A. Huang, "Cooperative multicast with moving window network coding in wireless networks," *Ad Hoc Networks*, vol. 25, pp. 213–227, 2015.

[30] T. Ho, M. Medard, R. Koetter et al., "A random linear network coding approach to multicast," *IEEE Transactions on Information Theory*, vol. 52, no. 10, pp. 4413–4430, 2006.

[31] Y. Guo, J. Hao, and G. Yue, "Segment linear network coding in wireless sensor networks," in *Proceedings of the Wireless Communications and Networking Conference (WCNC)*, pp. 1522–1527, Las Vegas, NV, USA, April 2013.

[32] L. Miao, K. Djouani, A. Kurien, and G. Noel, "Network coding and competitive approach for gradient based routing in wireless sensor networks," *Ad Hoc Networks*, vol. 10, no. 6, pp. 990–1008, 2012.

[33] I. Stojmenovic, A. Nayak, J. Kuruvila, F. Ovalle-Martinez, and E. Villanueva-Pena, "Physical layer impact on the design and performance of routing and broadcasting protocols in ad hoc and sensor networks," *Computer Communications*, vol. 28, no. 10, pp. 1138–1151, 2005.

[34] J. Han, J. Pei, and Y. Yin, "Mining frequent patterns without candidate generation," *ACM SIGMOD Record*, vol. 29, no. 2, pp. 1–12, 2000.

[35] Y. Lin, B. Li, and B. Liang, "Stochastic analysis of network coding in epidemic routing," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 5, 2008.

[36] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.

[37] P. Levis, N. Lee, M. Welsh, and D. Culler, "TOSSIM: accurate and scalable simulation of entire TinyOS applications," in *Proceedings of the 1st International Conference on Embedded Networked Sensor Systems*, pp. 126–137, ACM, Los Angeles, CA, USA, November 2003.

[38] P. Levis and D. Gay, *Tinyos Programming*, Cambridge University Press, Cambridge, UK, 2009.

International Journal of
Rotating
Machinery

The Scientific
World Journal

Journal of
Sensors

Advances in
Multimedia

Journal of
Engineering

Advances in
Civil Engineering

Journal of
Control Science
and Engineering

Journal of
Robotics

![Hindawi logo]

Hindawi

Submit your manuscripts at
www.hindawi.com

Journal of
Electrical and Computer
Engineering

Advances in
OptoElectronics

VLSI Design

International Journal of
Navigation and
Observation

Modelling &
Simulation
in Engineering

International Journal of
Aerospace
Engineering

International Journal of
Chemical Engineering

International Journal of
Antennas and
Propagation

Active and Passive
Electronic Components

Shock and Vibration

Advances in
Acoustics and Vibration