RESEARCH   **Open Access**

Network coding-based multipath routing for energy efficiency in wireless sensor networks

Lei Wang, Yuwang Yang[*] and Wei Zhao

Abstract

This article studies the energy efficiency of wireless sensor networks with network coding-based multipath routing (NCMR). The employed multipath model is braided multipath model, and the network coding scheme is random linear network coding. Braided multiple paths to the sink node are established for each source node, and the packets encoded at source nodes are transmitted on the braided multipath network. Then, intermediate nodes re-encode the received packets and forward the new packets to next cluster. Finally, the sink node decodes the packets received from different paths and recovers the original data. When network coding is combined with multipath routing, the number of required routes and the total times of transmission in sensor networks are reduced, which leads energy consumption of NCMR lower than that of traditional multipath routing, and this is proved by the theoretical analysis results in this article. Meanwhile, results of the analysis show that NCMR provides more reliability. Extensive simulations are carried out, and the results are consistent with those of the theoretical analysis. Moreover, based on the study of the influence of different network parameters (e.g., number of hops, number of paths) on the performance of sensor networks, an optimal combination scheme of different parameters are proposed, which makes the network accomplish transmissions with less resource. Finally, in order to study the robustness of NCMR, some simulation experiments are carried out under special conditions such as interference on channels with memory, congested environments, and failed nodes, which show that NCMR is more effective in adapting to these scenarios.

Keywords: network coding, multipath routing, energy efficiency, wireless sensor networks

1. Introduction

Wireless sensor networks (WSNs) consist of distributed and networked sensors which jointly monitor the physical or environmental conditions such as temperature, sound, vibration, pressure, and motion at different locations. There are two distinctive characteristics in WSNs. First, nodes in WSNs communicate with others through wireless channels. Compared with wired networks, link quality of wireless channels is bad due to instability of wireless channels. Second, the nodes are always deployed in complex environments, and most often, batteries are used as their energy supply. However, a realistic problem is that energy produced from batteries is always limited. Therefore, reliable transmission in WSNs can be considered at two levels which are lower energy consumption and higher successful delivery ratio. In

order to improve the energy efficiency and successful delivery ratio of WSNs, many researchers have done much work, and several kinds of technologies are employed.

Multipath routing [1-3] is employed to increase reliability of wireless networks, and it creates several paths to the sink node for each source node. Several copies of the initial packets are transmitted through these paths. As long as one of these packets is received successfully, the whole process can be considered a success. In this sense, multipath routing improves reliability by increasing redundancy. However, high redundancy results in more transmissions, the energy consumption accordingly becomes greater.

Error correction mechanism [4,5] is another technology used to increase reliability. There are three kinds of error correction mechanisms, namely Automatic Repeat Request (ARQ), Forward Error Correction (FEC) [6], and Hybrid Error Correction. The principle of ARQ is

* Correspondence: yuwangyang@mail.njust.edu.cn
Computer Science School, Nanjing University of Science and Technology (NUST), Nanjing 210094, Jiangsu, China

that once an error occurs at sink nodes or intermediate nodes, a request will be sent to the source node for retransmission. However, the cost of retransmission is great in WSNs since it may bring some problems such as twice energy consumption and longer transmission delay. Therefore, researchers are inclined to employ the technology which can correct the inaccurate data at sink node. So, FEC is an ideal scheme in WSNs. In FEC scheme, the sender encodes the initial data into new codeword with some redundant bits. Some bits of the data may become inaccurate in the channel during the transmissions, but the receiver can recover the original data with the redundant bits as long as the number of inaccurate bits is not more than the correcting ability of FEC. Therefore, FEC improves the reliability without retransmission, which is an advantage for wireless networks. However, the disadvantage of FEC is that FEC puts all the original bits and redundant bits in the same packets, if one of the nodes in the route fails, the packets cannot successfully be delivered, so FEC scheme is not robust enough for WSNs. Therefore, the technology that satisfies the following three requirements is more ideal for WSNs. First, it reduces the data redundancy. Second, it does not require retransmission. Third, it is robust to failure of nodes. Recent researches show that network coding may meet the requirements.

Network coding was firstly developed by Ahlswede and Cai [7]. The traditional architecture of networks is store-and-forward, and a most widely recognized attitude is that coding at the intermediate nodes, such as switch and router, will not bring any benefit. However, Ahlswede and Cai reversed this attitude in their study for the first time. After that, many scholars [8-11] focused on this field. Li et al. [8] proved that linear network coding is sufficient to achieve the multicast capacity. Chou et al. [10] proposed practical network coding to be considered as a representative coding scheme which is also explored in this article. Katti et al. [11] implemented the first wireless network (COPE) based on network coding. They showed that high performance can be achieved in wireless networks with network coding, and they considered network coding is powerful in both unicast networks and multicast networks. Recently, Nguyen et al. [12,13] derived some theoretical results on bandwidth efficiency of network coding in wireless networks.

The main contributions of this article can be listed as follows. First, we propose a network coding-based energy-efficient multipath routing scheme which reduces the number of required paths and the times of network transmissions. Second, optimal combination scheme of different network parameters is proposed and proved valid, which leads to lower energy consumption of the whole network. Third, we show that NCMR is more

effective in extreme environments such as channels with memory and congested environment.

The remainder of this article is organized as follows. In Section 2, we begin with the introduction of some related studies. In Section 3, the network and energy models used in our analysis and simulations are explained. In Section 4, we present the mathematical analysis of traditional multipath routing (TMR) network and NCMR network. In Section 5, we design routing discovery algorithm, data disseminating algorithm, and routing maintenance algorithm to enable network coding to be performed on braided multipath networks so that we can obtain the simulation results. Then, in Section 6, we show the results of analysis and extensive simulations. Finally, the last section summarizes the main conclusions.

2. Related study
There are some works related to ours. Ganesan et al. [14] showed that braided multipath is energy efficient to recover from isolate and patterned failures. And the authors also proposed the initial definition of braided multipath: For each node on the primary path, find the best path from source to sink that does not contain that node. In this article, we employ the braided multipath network to serve as network model. In the network model, we find N-1 braided paths for each intermediate node.

There are some previous studies [15-20] addressing the performance of network coding-based multipath routing (NCMR). Guo et al. [15] employed multipath routing with network coding in underwater sensor networks (UWSNs), which proves that network coding can achieve high reliability in UWSNs. The authors discussed the reliability of NCMR, followed by the performance of NCMR being compared with that of TMR. Part of work in this article also addresses this comparison. Guo et al. [15] assumed that the probability that nodes in (i)th cluster receive packets sent by source node only depends on the nodes in $(i$-1)th cluster. However, actually the results based on this assumption are not accurate. In this article,, the accurate results are obtained with recursion method. It is confirmed that the comparison based on these accurate results is more reasonable. We will further detail this issue in Section 4.

Li et al. [16] insisted that the reliability can be guaranteed by introducing network coding. Random linear coding scheme is also employed in their work. They show that the energy consumption is lower since network coding reduces the number of paths required to deliver data. In addition, a method to obtain the optimal number of paths is provided. In this article, the results of analysis and simulations show that there is another parameter (namely, the number of subpackets) which

may affect the energy consumption as well. One of the aims of this article is to find the optimal scheme which is the best combination of different network parameters. It is supposed that higher energy efficiency can be achieved with the optimal scheme.

Our previous work [17] presents network coding-based reliable disjoint and braided multipath routing (NC-RMR) for sensor networks in which improving reliability of multipath routing with network coding is the primary aim. In this article, network coding is employed to achieve energy efficiency in WSNs, and the mathematical models of multipath routing with and without network coding are inherited and developed. In [17], we employed two metrics (namely, SDR and NEC) to evaluate the performance of sensor networks. SDR is the number of successful delivered packets over the total packets sent by the source node, and NEC is normalized energy consumption which is equal to total energy consumption divided by SDR. In this article, the two metrics are also employed. There are some differences between this article and the previous study [17]. First, a typical energy model is introduced to evaluate the energy consumption of sensor networks, and the influence of different network parameters on performance is also detailed. Second, we provide routing discovery algorithm, data disseminating algorithm, and routing maintenance algorithm to enable network coding to be performed on a braided multipath network model. Third, some additional data and different analysis such as energy consumption, reliability, and optimal scheme for energy efficiency are provided.

There are many articles [21-25] focusing on energy efficiency in wireless network. Limited energy can be considered as the main bottleneck of communication ability, which makes providing energy efficiency a challenge. To reverse this trend, many scholars have done lots of work using a number of different methods. Some authors [21-23] tried to solve this problem through topology control, Cardei et al. [24,25] addressed this issue through coverage control, and another method to solve this problem is data fusion [26]. The method described here can be grouped under data fusion as well. There are some theoretical articles [27-29] addressing energy efficiency in wireless networks with network coding, which have considered establishing minimum cost, multicast connections over networks with encoded packets. When coupled with existing decentralized schemes for constructing network codes, they yield a fully decentralized approach for achieving minimum-cost multicast. The thought of improving energy efficiency with network coding is used here, but the difference is that network coding is combined with multipath routing, and the aim is to provide an energy-efficient and robust scheme for WSNs.

3. Models of multipath in WSN
3.1. Network model
In general, multipath network can be divided into disjoint and braided models [30]. This article mainly discusses the latter. Figure 1 shows a typical braided multipath routing model which consists of a source node, a sink node, and many intermediate nodes. We assume that all the sensory data generated at different source nodes must be sent to a data processing center, so the network model we employ is a unicast model. Previous study [11] proved that network coding is useful in both multicast networks and unicast network. In Figure 1, the number of paths is equal to the number of nodes in a cluster. We employ N to denote the number of paths, so the number of nodes in a cluster is equal to N as well. For each intermediate node, we find N-1 nodes to serve as its backup nodes, and therefore these N nodes form a cluster.

The mathematical analysis presented in Section 4 is based on such network model, a routing discovery algorithm proposed in Section 5 establishes this network model for each source node, and simulations based on the established network are made in Section 6.

3.2. Energy model
Figure 2 shows the energy consumption of main components in WSNs, which is proposed by Estrin [30]. After that, the figure is widely accepted in the field of sensor networks. Figure 2 indicates that sending, receiving, and idling consume the most energy while sensing, programming, and sleeping consume very little. Therefore, the energy consumption of sensing, programming, and sleeping is neglected in the following analysis. The energy consumption of idling is always spent by the nodes to avoid collisions, which is the function of MAC layer, and the energy consumption in MAC layer does not affect the energy analysis in network layer. While the main aim of this article is to propose a network coding-based multipath protocol which is performed on network layer, and therefore we do not take into account the energy consumption of idling in this article. So, the sending energy $E_{TX}(k, d)$ and the receiving energy $E_{RX}(k)$ become the focus, thus total energy consumption can be considered as the energy consumption of sending and receiving a packet multiplied by the total transmission times. In wireless networks, as long as the nodes are deployed in the transmission range of the transmitting nodes, they receive these packets for free even if these packets do not belong to their duties. This phenomenon is considered as the broadcasting characteristic of wireless networks. So, the total energy consumption is

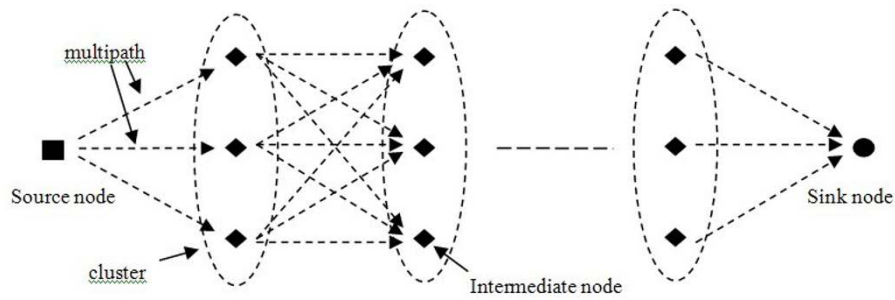$$E_{\text{MULT}} = (E_{TX}(k, d)) + N \times (E_{RX}(k)) \times NR \qquad (3.1)$$

Figure 1 Network Model

Figure 1 Network model.

where N is the number of nodes in a cluster, NR is the transmission times in the network, k is the packet size, and d is the transmission radius. To simplify the discussion, it is assumed that all the sensor nodes have the same transmission radius. The relationship among energy consumption, packet size k, and transmission radius d for one node is

$$E_{TX}(k, d) = k \times E_{txElec} + k \times \varepsilon_{amp} \times d^{\gamma} + E_{start} \qquad (3.2)$$

where E_{txElec} is the energy of sending one bit data, γ is the path loss factor whose value is constant in a typical condition, ε_{amp} is the value of signal amplifier, and E_{start} is the energy consumption of starting transmission. On the other hand, the energy consumption of receiving is

$$E_{RX}(k) = k \times E_{rxElec} + E_{start}. \qquad (3.3)$$

In Equation (3.3), E_{rxElec} is the energy consumption of receiving one bit data. For convenience, it is assumed that energy consumption of sending a bit is the same as that of receiving one, which is shown in Equation (3.4).

$$E_{rxElec} = E_{txElec} = E_{elec} \qquad (3.4)$$

From Equations (3.2), (3.3), and (3.1), the result of total energy consumption is achieved. In our analysis,

energy consumption of encoding and decoding during network coding operations is also neglected, and the reasons why we do not take into account it are as follows. First, the main work of encoding is to simply obtain some new linear combinations of original packets. Second, the operations of encoding and decoding are based on a finite field on which all the operations such as addition and multiplication are closed, and there is no floating-point operation in sensor nodes. Therefore, compared with energy consumption of sending and receiving, the energy consumption during network coding operations is much lower.

4. Performance analysis of multipath routing
4.1. TMR
In traditional multipath transmission models, when source node S intends to transmit data, it establishes N braided paths. Then, node S floods its packets to all nodes in next cluster. Intermediate nodes forward these packets in the same way. The network model introduced in the third section is employed here. It is assumed that bit error rate is BER, and the length of packet is L. So, the probability of successfully sending a packet is
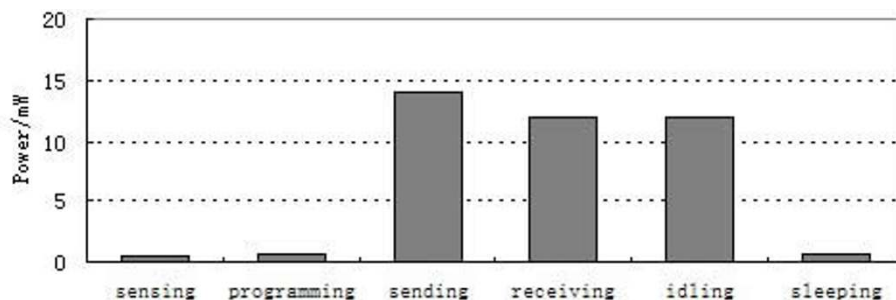
$$p = (1 - \text{BER})^{L}. \qquad (4.1)$$

Figure 2 Energy consumption of different components.

Let i be the number of clusters, H be the total hops of network, and $a_i(1 \leq i \leq H)$ be the probability that a node in (i)th cluster successfully receives the packet sent by source node. Equation (4.2) shows the probability of the nodes in the first cluster.

$$a_{i=1} = p \tag{4.2}$$

We have mentioned that the study in [15] also addresses braided multipath routing network. But, the mathematical analysis model is different from ours. Those authors employ the following method to obtain the values of a_i under the situation with $i > 1$.

$$a_i = \sum_{n=1}^{N} \alpha_{i-1,n}[1 - (1 - p)^n] \quad i \geq 2 \tag{4.3}$$

In Equation (4.3), $\alpha_{i-1,\,n}$ stands for the probability that n nodes in $(i\text{-}1)$th cluster successfully receive the packet sent by source node. Therefore, the probability can be obtained with Equation (4.4).

$$\alpha_{i-1,n} = \binom{N}{i} a_{i-1}^n (1 - a_{i-1})^{N-n} \quad i \geq 2 \tag{4.4}$$

From Equations (4.4) and (4.3), it is shown that the value of a_i only depends on the value of a_{i-1} in the method. In our study, it is proved that the results of a_i obtained with Equation (4.3) are not accurate, and the error increases as i becomes larger. In this article, we first explain why accurate results cannot be obtained with Equations (4.3) and (4.4), and then we propose a new method to obtain the accurate values of a_i.

In their method, it is considered that the number of received packets in $(i\text{-}1)$th cluster subjects to binomial distribution. According to probability theory, if the number of received packets subjects to binomial distribution, nodes in the same cluster must be independent. However, the nodes in the same cluster of braided multipath network must be dependent. Take the network model shown in Figure 1 for example, all the nodes in the second cluster depend on the first node in the first cluster, in other word, the first node in the first cluster affects all the nodes in second cluster, so the nodes in second cluster are not independent. Therefore, the results of a_i in the previous study [15] are not accurate, and the errors become more and more large when i becomes larger. In this article, the accurate results of a_i are achieved with recursion method below.

$$a_i = \begin{cases} p & i = 1 \\ G(1,1) \times a_{(i-1)} + \sum_{z=2}^{N} G(1,z)F(i,z) & i > 1 \end{cases} \tag{4.5}$$

Here, $F(i, z)$ is achieved with the following recursion.

$$F(i,z) = \begin{cases} 1 - (1-p)^z & i = 2 \\ G(z,1) \times a_{(i-2)} + \sum_{n=2}^{N} G(z,n)F(i-1,n) & i > 2 \end{cases} \tag{4.6}$$

where $G(x, y)$ is the probability that y nodes successfully receive the packets which are sent by x upstream nodes, and it is obtained with Equation (4.7).

$$G(x,y) = \binom{N}{y}[1 - (1-p)^x]^y[(1-p)^x]^{N-y} \tag{4.7}$$

In Equation (4.7), N is the number of nodes in a cluster which is equal to the number of paths in a braided multipath network model. When such a network model is established, N is a constant. This recursion method to obtain the value of a_i will be validated through extensive simulations in Section 6.

To achieve the energy consumption, the number of sending times in (i)th cluster σ_i should be obtained first. Since any node sends the same packet no more than once, and the probability of receiving the packet successfully by each node is a_i, so σ_i is calculated in Equation (4.8).

$$\sigma_i = N \times a_i \qquad (i \geq 2) \tag{4.8}$$

Then, normalized redundancy (NR) of the network is obtained by total sending times divided by SDR which is equal to a_H.

$$NR = \frac{1 + \sum_{i=1}^{n-1} \sigma_i}{SDR} \tag{4.9}$$

From Equations (4.9) and (3.1), the total energy of TMR is achieved.

4.2. Multipath routing with network coding

In general, there are two kinds of linear network coding, namely, random linear network coding (RLNC) and deterministic linear network coding. Since RLNC [10] has widely been studied and applied, it is also employed in this article. First, the RLNC scheme should be introduced briefly. When there are some sensory data to be transmitted at the source node, these packets are divided into K packets. The source node randomly chooses $(K+n) \times K$ elements from a Galois field to be a coefficient matrix, then encode these K original packets into $(K+n)$ new packets. Multiply each row vector of the coefficient matrix by the K original packets is a new packet which is a linear combination of these original packets. Intermediate nodes need to recode the received packets, and then send the new packets to next cluster. As long as the sink node successfully receives K packets of these $(K+n)$ packets, it may decode and recover the

original data. In our scheme, we choose 256 to be the size of Galois field since previous researchers [10] proved that when the size of the Galois field is 256, probability of decoding these packets successfully at the sink node is 0.996. And another reason for choosing 256 to be the field size is that, in current computer system, the data type of BYTE (namely, unsigned char) also has 8 bits. It is considered that the utilization rate of system memory is high in this manner. If we choose 2^9 as the field size, we need to employ the data type of unsigned integer to store the elements which only requires 9 bits, while the data type of unsigned integer always have 16 or 32 bits, and therefore many bits are wasted. Moreover, the cost of multiplication operation between two variables with the data type of unsigned integer is much more than that between two variables with the BYTE data type. Therefore, both the coefficients and data of packets are stored, encoded, and computed with BYTE data type in our scheme. Figure 3 explains the encoding scheme.

When the sink node receives K or more than K packets, it decodes the original packets by the following calculation.

$$\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_K \end{bmatrix} = \begin{pmatrix} g_{11} & \cdots & g_{1K} \\ \vdots & \ddots & \vdots \\ g_{K1} & \cdots & g_{KK} \end{pmatrix}^{-1} \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_K \end{bmatrix} \qquad (4.10)$$

After the data are encoded at the source node, the number of packets changes from K to K' ($K' = K+n$, $K' > K$). Here, we denote $a_{i,k}$ as the probability that any node in (i)th cluster receives k packets successfully. To formulate the value of $a_{i,k}$, linear correlation of all the packets should be first considered. Since the probability of decoding these packets successfully at each node is 0.996 when the field size is 256, it is assumed that the node can decode these packets and recover the original data as long as it receives K or more than K packets successfully. For the nodes in the first cluster, $a_{i,k}$ subjects to binomial distribution, and therefore the probability of receiving k packets successfully for each node in the first cluster is obtained with Equation (4.11).

$$a_{1,k} = \begin{pmatrix} K' \\ k \end{pmatrix} \times p^k \times (1-p)^{(K'-k)} \qquad (4.11)$$

The nodes are supposed to be the same in the network, therefore, for all nodes in (i)th cluster, the probability that receiving k packets at each node from (i-1)th cluster is the same. If i is greater than 1, $a_{i,k}$ is equal to the probability that any node in cluster i successfully receives k packets which come from (i-1)th cluster in discretional combinations. For $2 \leq i \leq H$, $1 \leq j \leq N$, $0 \leq k \leq K$, $a_{i,k}$ is calculated as follows.

$$a_{i,k} = \sum_{k_j \in \{0, \cdots k\}, s.t. \left(\sum_{j=1}^{N} k_j \right) = k} \prod_{j=1}^{N} a_{i,j,k_j} \qquad (4.12)$$
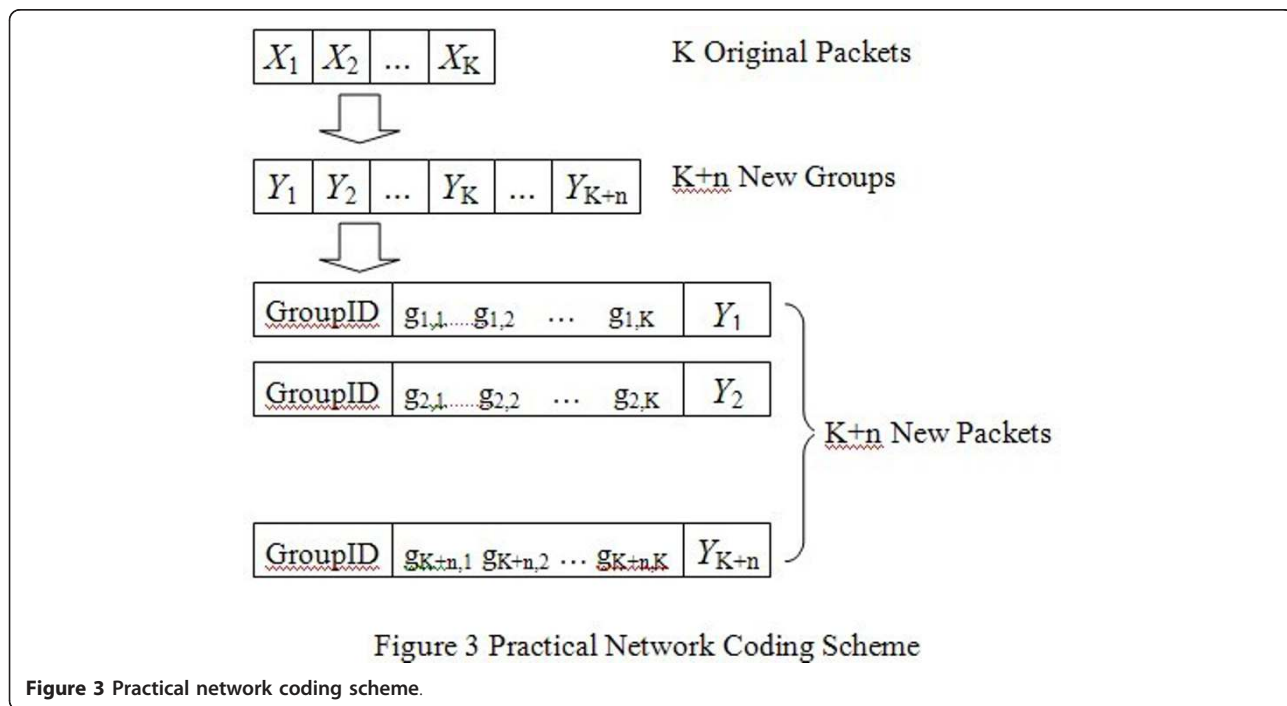

Figure 3 Practical network coding scheme.

In Equation (4.12), $a_{i,j,k}$ is the probability that the node in (i)th cluster receives k_j packets from the (j)th node in $(i$-$1)$th cluster. To obtain the value of $a_{i,j,k}$, we need to first obtain the value of a_k which is the probability that the node transmits l packets while k packets are received successfully by a node in the next cluster. When nodes in any cluster transmit their packets to next cluster, some packets would become inaccurate. Whether the packets are transmitted successfully or not is independent. Therefore, this kind of events subjects to Bernoulli distribution. So, a_k is obtained with Equation (4.13).

$$a_k = \binom{l}{k} \times p^k \times (1-p)^{(l-k)} \qquad (4.13)$$

As mentioned above, the sink node can decode the packets as long as the K packets are received successfully, and therefore intermediate nodes would just send K packets when they receive K or more than K packets. So, $a_{i,j,k}$ is obtained with Equation (4.14).

$$a_{i,j,k} = \sum_{l=k}^{K} a_{i-1,l} a_k \qquad (4.14)$$

Denote $a_{i,K}$ as the probability that any node in cluster i successfully receives K or more than K packets, and therefore $a_{i,K}$ is obtained with Equation (4.15).

$$a_{i,K} = 1 - \sum_{n=0}^{K-1} a_{i,n} \qquad (4.15)$$

SDR of the network system is the probability that nodes in cluster H successfully receive K or more than K packets which is equal to $a_{H,K}$, and $a_{H,K}$ is calculated as follows.

$$\text{SDR} = a_{H,K} = 1 - \sum_{n=0}^{K-1} a_{H,n} \qquad (4.16)$$

Then, the number of sending times is obtained with Equation (4.17).

$$\sigma_i = N \times \sum_{k=1}^{K} (a_{i,k} \times k) \quad i \geq 2 \qquad (4.17)$$

Consequently, NR of NCMR is shown in Equation (4.18).

$$NR = \frac{N \times \sum_{k=1}^{K} (a_{1,k} \times k) + \sum_{i=2}^{N} \sigma_i}{\text{SDR} \times K} \qquad (4.18)$$

The new packets transmitted in the network are encoded so that the original packet size is modified. The new packet size is equal to the sum of the length of the new group, the length of the coefficient vector, and the length of group ID. At last, from Equations (4.18), (3.1), and the new packet size, the energy consumption of NCMR is achieved.

5. Protocol design

We have analyzed energy consumption of TMR and NCMR, and we will show the results of analysis in Section 6. To validate the correctness of the analysis results, we need to make extensive simulations and compare the results of analysis with those of simulations. So, first we need to design a protocol to enable network coding to be performed on a braided multipath routing network. Like other routing protocols, the designed protocol also mainly consists of routing discovery algorithm, data disseminating algorithm, and routing maintenance algorithm.

5.1. Routing discovery

The procedure of routing discovery is the first step of a routing protocol. The work of our routing discovery algorithm is to establish $N(N \geq 2)$ paths from the source node to sink node. Deb et al. [31] provided a typical computation method to obtain the number of N in multipath network which is also employed in our previous work [17]. And the method is based on the assumption that every node is assumed to have the knowledge of link quality, so the source node could know the value of BER. The number of required transmission paths could be determined by BER and the hops to sink (H). A rule is that as BER or H becomes greater, N must be greater since the source node must select more paths to guarantee reliability. In this article, we also employ the computation method but we do not detail it any longer due to the limited space.

When a great number of nodes are deployed into a specific location, the source node needs to establish paths to transmit sensory data to the sink node. Algorithm 1 shows our routing discovery algorithm.

Algorithm 1: Routing discovery algorithm

Requirement: The source node is assumed to have the knowledge of link quality.

1: //Step 1. To get the minimum hops to sink node
2: The sink node floods a route discovery packet (RDP).
3: For each non-sink node receiving RDP
4:     Forwards the packet.
5:     Updates the shortest route to sink and records all the nodes in the shortest route.
6:     Records and updates the number of hops of the shortest route (H).

7: End

8: //**Step2. Routing establishment**

9: The source node S calculates the number of required paths (N).

10: Then node S sends the value of N to its downstream node A which is the first hop in the shortest route to the sink.

11: Node A finds other N - 1 nodes to serve as its backup nodes. These N nodes form a cluster.

12: Node A sends the value of N to its downstream node B, and node B performs the same function.

13: If sink node is reached

14:　Finish routing establishment.

15: Else

16:　Continue routing establishment.

17: End

In routing discovery phase, if any intermediate node is selected to relay the packets generated at source nodes, it must record the address of source nodes in its cache. Once an intermediate node receives a packet with sensory data, it must check the source address. If the source address does not exist in the cache, the intermediate node must drop the packet, and the node only receives the packets whose source address is recorded in the cache.

5.2. Data disseminating

Now, braided multipath routes have been established, the following step is to transmit sensory data to sink node. In traditional multipath network, the main work of intermediate nodes is to forward the received packets to downstream nodes. While in NCMR network, intermediate nodes need to recode the received packets and then transmit the new packets to the nodes in next cluster. Algorithm 2 shows the procedure of data disseminating.

Algorithm 2: Data Disseminating

Requirement: Operations on Galois field are already implemented in sensor nodes.

1: //**Step 1. Coding at source node**

2: The source node S encodes the original K packets into K' new packets.

3: Send the new packets to next hop.

4: //**Step2. Coding at intermediate nodes**

5: For each intermediate node

6:　If R > = K % R is the number of received packets

7:　　Randomly generates a $K*K$ matrix M to be local coding kernel.

8:　　Encodes these R packets into K new packets with matrix M.

9:　Else

10:　　Randomly generates a $R*R$ matrix M to be local coding kernel.

11:　　Encodes these R packets into R new packets with matrix M.

12: End

13: Forwards the new packets to next hop.

14: End

15: //**Step3. Decoding at the sink node**

16: If R > = K

17:　The sink node recovers the original packets with Gaussian elimination.

18: End

In fact, the encoding operations of RLNC at source node are very similar to that of traditional source coding (e.g., erasure coding). But, the main difference between the two is that network coding requires encoding operations at intermediate nodes while traditional source coding only encodes packets at the source node. In Algorithm 2, local coding kernels M of intermediate nodes are randomly obtained from GF(256), and then each new packet is a new linear combination of the received packets.

5.3. Routing maintenance

Since the proposed scheme is a decentralized one, the intermediate nodes do not require the knowledge of overall network topology to establish end-to-end paths. Therefore, the main routing maintenance work of intermediate nodes is to maintain the communications with its upstream nodes, downstream nodes and its backup nodes in the same cluster. In our algorithm, the cluster-head nodes in different clusters are responsible for maintaining the communications with cluster-head nodes of upstream and downstream clusters. And they also need to keep in touch with its backup nodes. For this reason, a cluster-head node must have a counter to record the number of its backup nodes. In a routing updating period, if the counter is less than the required N, the cluster-head node need to find some more free nodes to be its backups so that the number of braided paths can be guaranteed.

Algorithm 3: Routing maintenance

Requirement: The source node periodically broadcasts a routing maintenance packet (RMP).

1: For each cluster-head node receiving RMP

2:　If RMP is from upstream cluster

3:　　Forwards it to downstream cluster

4:　Else if RMP is from its backup nodes

5:　　Records the node ID and changes the counter with the increment of 1

6:　　Replies a message to the backup node.

7:　End

8:　Forwards RMP to the cluster-head node of next hop.

9: End

10: If counter <N

11: Selects (N-counter) free nodes to be its backup nodes.

12: End

13: For each backup node receiving RMP

14: Broadcasts RMP to cluster-head node and waits for a reply.

15: If no reply is received % Cluster-head node fails.

16: Elects a new cluster-head node and finds a new backup node.

17: End

18: End

19: At last, the sink node discards RMP.

So far, we have accomplished the design of algorithms. And the following work is to make extensive simulations based on the braided multipath network.

6. Analysis and simulation results
6.1. Establishment of braided multipath network
According to the routing discovery algorithm presented in Section 5, we need to establish a hop-by-hop braided multipath network which consists of a number of randomly deployed nodes. We employ Omnet++ [32] and MiXiM (Mixed Simulator) [33] to implement the algorithms. OMNeT++ is an object-oriented modular discrete event network simulation framework which is a powerful tool for WSNs. However, a shortcoming is that nodes in Omnet++ are lack of design of layer. To make up the shortcoming, we combine MiXiM with Omnet++ since MiXiM provides detailed node and channel models. When MiXiM is employed in our simulation, network coding is performed on network layer, CSMA is selected for MAC layer, and free space model is chosen for physical layer. The work of application layer is to randomly generate some sensory data. Before giving the results of simulations, we list the main parameters used in the simulations in Table 1.

Figure 4a, b shows the results of routing discovery algorithm. Figure 4a, b is based on the same network topology. The number of nodes is 80, and the area is 400×200 m^2. In the two figures, green node refers to sink node, yellow node refers to source node, red nodes

Table 1 Main parameters used in the simulations

Parameters	Values	Parameters	Values
Network area	400×200 m^2	E_{elec}	50 nj/bit
Carrier frequency	2.412e+9Hz	E_{start}	2500 nj/packet
Signal sensitivity	-84 dBm	Transmission radius	10 m
Max sending power	100 mW	Signal amplifier	2
Mac layer head length	24 bit	Path loss coefficient	2.0
Net layer head length	32 bit	Mac layer protocol	CSMA

form the main route, and blue nodes are backup nodes. In Figure 4a, there are three nodes in a cluster and four hops to the sink. And in Figure 4b, there are two nodes in a cluster and three hops to the sink. The differences of the above two topologies are the number of hops to a sink and the number of the cluster size, and these different values are determined by the method mentioned in Section 5.1. Analysis and simulations in this article are based on this kind of topology.

6.2. Result analysis
Since different network parameters may result in different results, the default network parameters should be first configured. After obtaining the results of performance with default network parameters, we will address the influence of different network parameters on performance in Section 6.3. The total hop of the network model H is set to be 5, the number of nodes in a cluster N is set to be 3, and the length of packets L is set to be 1024 bits. Each packet is split into four groups, and then, these groups are encoded into six subpackets. For consistence, the performance curves of TMR in different figures are denoted as TMR, and the performance curves of NCMR in Figure 5.

From the comparison of TMR-Ana and NCMR-Ana, NCMR achieves higher successful delivery ratio. For normalized energy consumption of the network model, if there are few or no errors which occur in the channels, it is found that energy consumption in TMR is lower. While NCMR consumes more energy since more redundant bits such as encoding coefficients and group ID are transmitted. If BER becomes larger, the normalized energy consumption of NCMR becomes lower than that of TMR. The results show that NCMR provides higher reliability and lower energy consumption in the network with low link quality which is a character of WSNs.

To validate the results of analysis, simulations are carried out 10,000 times with OMNET++. The simulation results are consistent with those of the analysis. Meanwhile, it is further proven that the results of Equation (4.5) are accurate. Consequently, the comparison based on the accurate results is more reasonable.
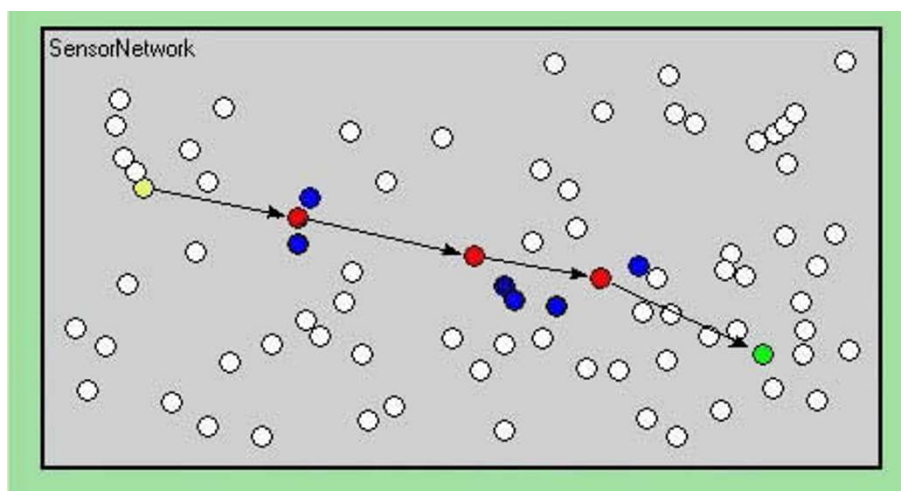
There is a difference between the results in this article and those in previous theoretical studies [27-29]. Previous studies consider that higher energy efficiency can be achieved with network coding. However, the results in this article indicate that when the link quality is fine, the network coding scheme would exert more energy due to the coding schemes being different. The scheme in this article requires some redundant bits which would cause more energy consumption when the link quality is fine.
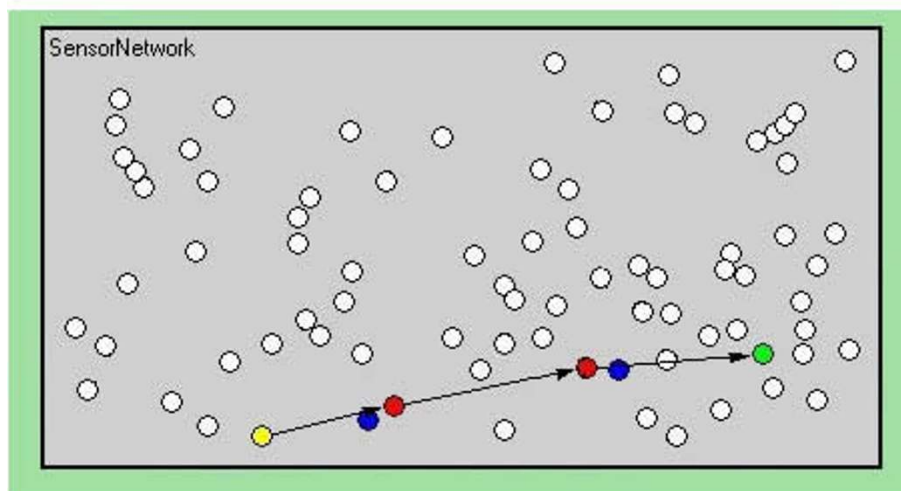
6.3. Influence of different network parameters
Successful delivery ratio and energy consumption are related to the number of hops H in the network model,

(a) 4 hops to the sink and 3 nodes in a cluster

(b) 3 hops to the sink and 2 nodes in a cluster

Figure 4 Simulation results of routing discovery algorithm.

the number of nodes N in a cluster, and the number of subpackets K. Next, the influence of different parameters on performance will be detailed. This article is aimed at finding an optimal combination of these parameters. It is supposed that higher energy efficiency can be achieved through the optimal scheme.

6.3.1. Influence of H on performance
As shown in Figure 6, when TMR is employed, there is an intersection of the two curves. Another similar intersection is found in the network coding scheme. To simplify the statement, denote J as x coordinate of this kind of intersection. Figure 6 shows that when BER is less than J, the smaller H is, the lower SDR is. In contrast, if BER exceeds J, the smaller H is, the higher SDR is. In particular, if BER is equal to J, no matter how much H is, the value of SDR is constant. On the other hand, the

influence of different hops on NEC is considered. Figure 6b suggests that the larger H is, the more energy should be consumed. The reason is that more packets are needed to be transmitted. Therefore, the relation between NEC and H is linear.

6.3.2. Influence of N on performance
Figure 7a shows that SDR increases as the number of N increases. N is the number of nodes in a cluster. The larger N is, the more redundant packets would be duplicated resulting in a higher SDR. On the other hand, this may cause higher energy consumption. Naturally, in Figure 7b, more energy is consumed when the channel is fine, and only when the link quality becomes low, does the normalized energy consumption become lower.

Compare the curve of TMR(N = 4) with that of NCMR (N = 3), we find that the performance of NCMR(N = 3) is
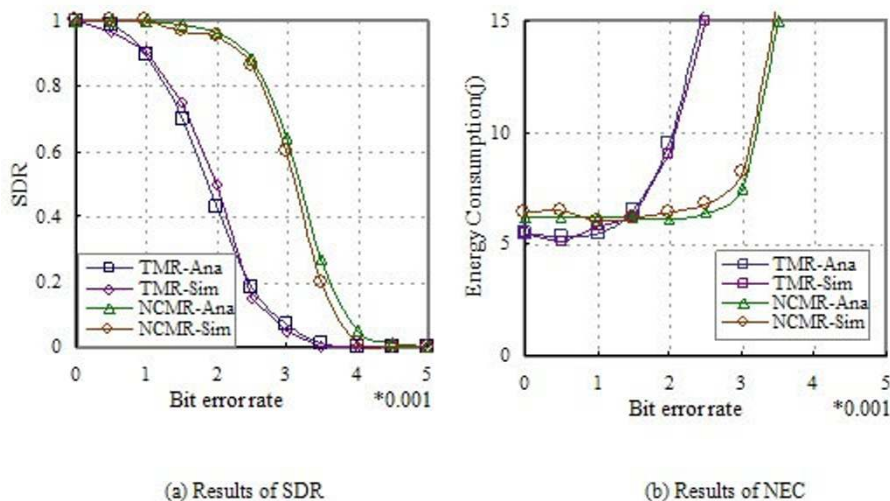
Figure 5 SDR and NEC in different routing model with the default parameters.

better than that of TMR(N = 4) although the number of required paths of the later is less than that of the former. Therefore, we consider that NCMR reduces the number of required paths, meanwhile the reliability and energy efficiency is higher. It is necessary to note that, in the experiments, N is set to be 3 or 4, but it does not mean that 3 or 4 is optimal. Here, the two values are only employed to illustrate that the number of required paths is reduced in the network with network coding.

6.3.3. Influence of K on performance

As mentioned above, NCMR divides the data into many subpackets. For fair comparison, the data in TMR should be divided as well. The reason is that splitting packets into many subpackets may improve SDR in a

multipath network. Based on this method, the comparison is reasonable.

From Figure 8, the increase of K enhances SDR in the model, but this causes more times of transmission. Therefore, it would cause more energy consumption.

6.4. Optimal combination scheme of different parameters

Through the above analysis, it is concluded that SDR and NEC of NCMR network are related to H, N, and K. And it is considered that if there is an optimal combination scheme for SDR to satisfy a specific reliability, it simultaneously consumes lower energy.

We know that there are several parameters which may affect the results of SDR and NEC. However, H is
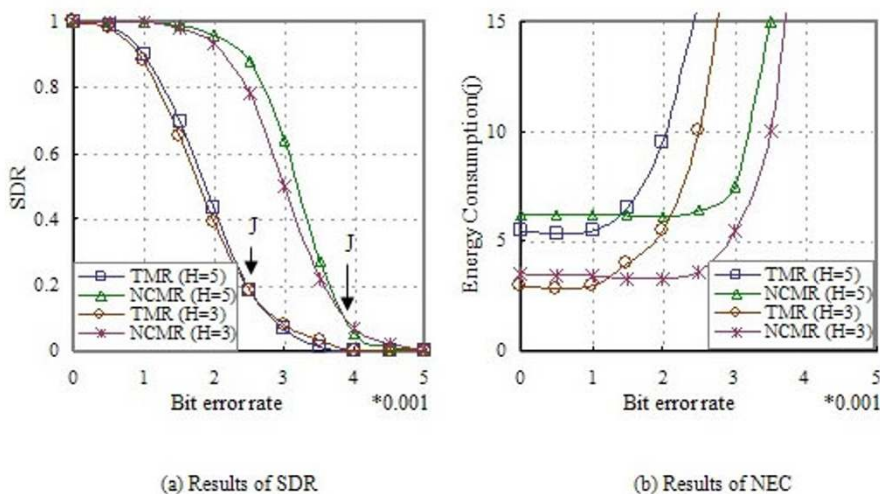
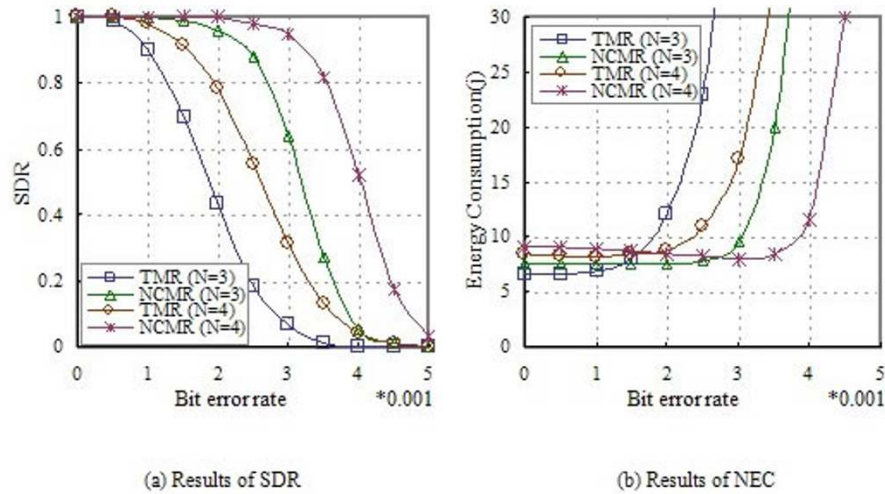**Figure 6 Influence of *H*, and other parameters are default**.

Figure 7 Influence of *N*, and other parameters are default.

determined by the distance to the sink node and the transmission radius of nodes. Therefore, the parameters which can be determined by the source node are N and K. In the previous study [16], the authors insist that the number of N influences the results of the successful delivery ratio and energy consumption and also provide a method to obtain the optimized number of N. Through the above discussion, the number of subpackets K has a relation to performance as well. Therefore, it is necessary for the source node to make an optimal combination of the two parameters. It is expected that the energy consumption would be lower due to the optimal combination. In the analysis of Section 6.3, we only change one network parameter every time. We have

known that the number of subpacket K and the number of required path N could be determined by the source node. And now we simultaneously change the two values, the two values are from 1 to 6. According to Equation (4.16), there are 36 different results, and then we obtain Figure 9 with linear interpolation method.

In Figure 9, the four curves represent different required SDR of the model, respectively. The combinations on the right of the curves can satisfy the required SDR, respectively. For example, if the required SDR does not exceed 0.975, the number of routes N and the number of subpackets K are the following combinations: (4,3), (2,4), (3,3), and so on. Then, the energy consumption can be obtained by substituting these variables in the formula.
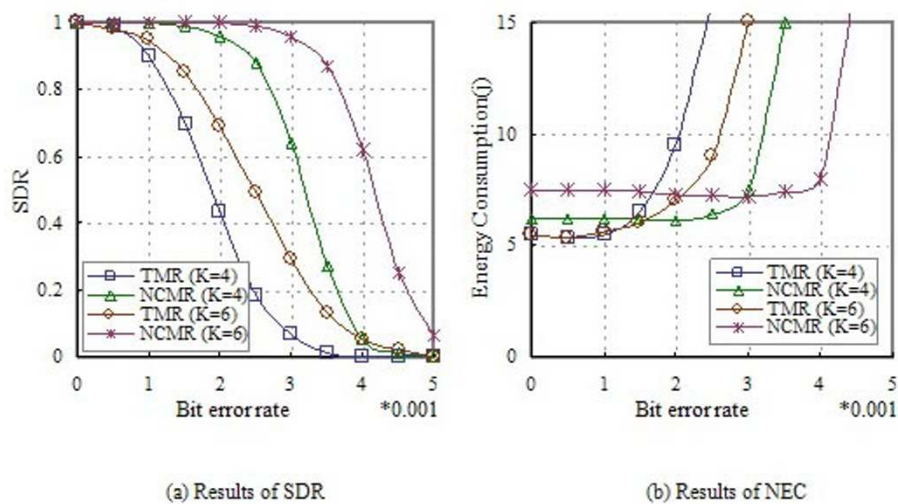

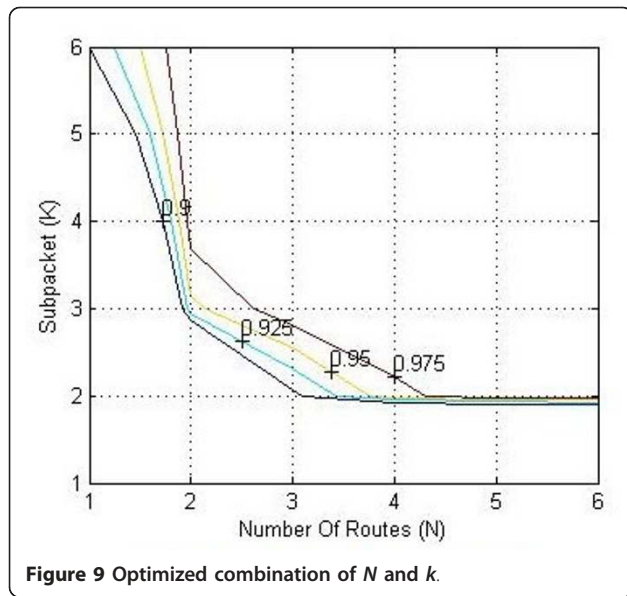**Figure 8 Influence of *k*, and other parameters are default.**

Figure 9 Optimized combination of *N* and *k*.

Figure 11 Simulation result in memory channel with default parameters.

In this way, the source node can find an optimal combination to accomplish its mission. Meanwhile, the cost would be lower. When the optimal combination scheme is obtained, the new number of required paths N can be used in our routing discovery algorithm, and the number of subpackets K can be employed by the source node to encode the original packets.

In fact, to obtain such a figure with our mathematic method may be not cheap for each source node. Fortunately, the computing only needs to be performed before routing discovery for one time. Therefore, the overhead of computing is affordable.

6.5. Interference on channels with memory

In general, channels are divided into memory channels and memoryless channels which are also called random channels in which the error occurrences of different bits are independent. While in memory channels, errors of bits are dependent. In these kinds of channels, the errors would occur in a series. The kind of interference may be lightning, dithering, or interference from the same frequency. When such interference is occurred, the channel could be considered as a memory channel. Figure 10 stands for a typical interference on channels with memory.
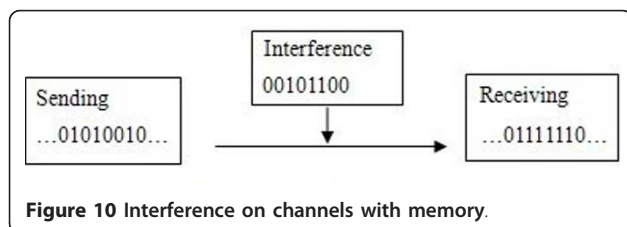

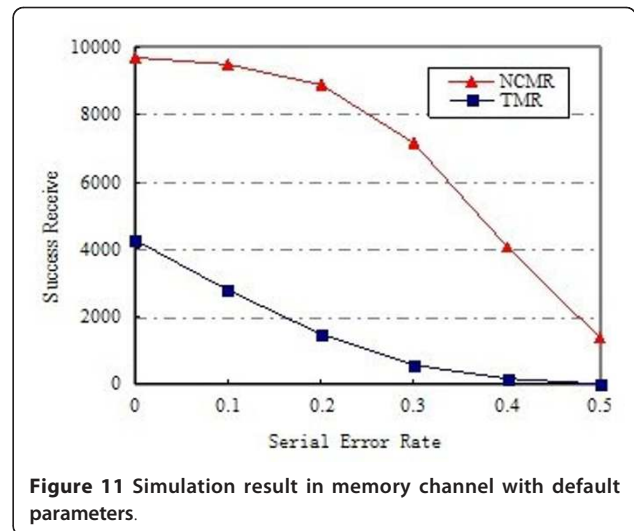**Figure 10 Interference on channels with memory**.

In the interference shown in Figure 10, the distance between the first "1" and the second "1" is called burst length l. In memory channels, l may be large. The corresponding consequence is that some groups of the packets may be seriously damaged, and as a result, the number of errors may exceed the ability of the correct mechanism. Therefore, receivers cannot decode the original data.

Under this condition, 10,000 experiments are simulated to discuss the ability to resist errors in memory channels of TMR and NCMR, respectively. In the simulations, it is assumed that both kinds of interferences exist simultaneously. Meanwhile, BER is set to be 0.002.

From the simulation results in Figure 11, NCMR would make the network more robust. Moreover, the curve shows that the number of packets received successfully decreases gradually. The TMR, however, has a lower performance. From this perspective, it is concluded that employment of network coding brings more benefits.

6.6. Performance experiment in congested environment

In practical networks, some circumstances may cause lower SDR and higher NEC as well, such as overloaded or failed nodes. In this section, this kind of event is simulated, and then the results of NEC and SDR are analyzed. For the four scenarios shown in Figure 12, 10,000 experiments are simulated, respectively. In the simulations, one node in the first cluster is set to be congested. The curves of TMRJam and NCMRJam are the simulation results with the congested node, and TMR and NCMR are the results without any node congested.

When a node blocks or fails, SDR starts to become a decreasing trend. The SDR of TMRJam is lower than
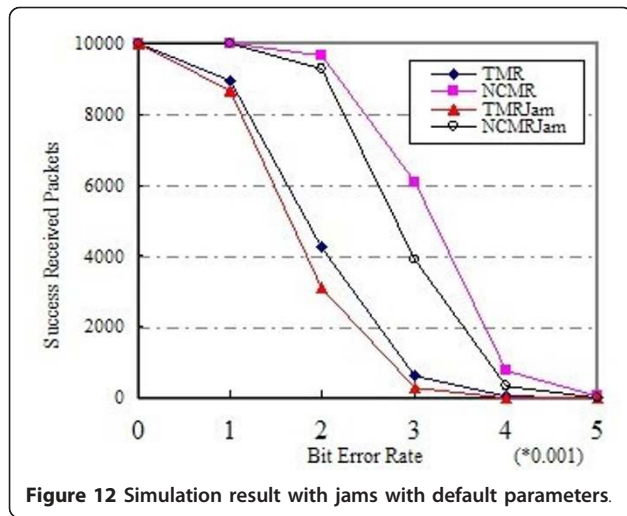
Figure 12 Simulation result with jams with default parameters.

that of TMR. Meanwhile, SDR in NCMRJam is higher than that of NCMR. And the question is about how it can be evaluated, and which scheme can effectively resist such kinds of interferences. To compare both of them fairly, rate of decline is employed and considered as a criterion to evaluate the anti-jamming abilities. The rate of decline of TRMJam is equal to (TMR-TMRJam)/ TMR, and the rate of decline of NCMRJam can be obtained with similar method.

As the SDR of TMRJam and NCMRJam are zero when BER is 0.005 in Figure 12, it is impossible to calculate the rate of decline. In Figure 13, when there are some nodes blocked, SDR of TMR decreases more quickly than that of NCMR. In other words, the anti-jamming ability of NCMR is greater than that of TMR.

7. Conclusion

This article studies the energy efficiency of WSNs. NCMR is employed to improve energy efficiency in
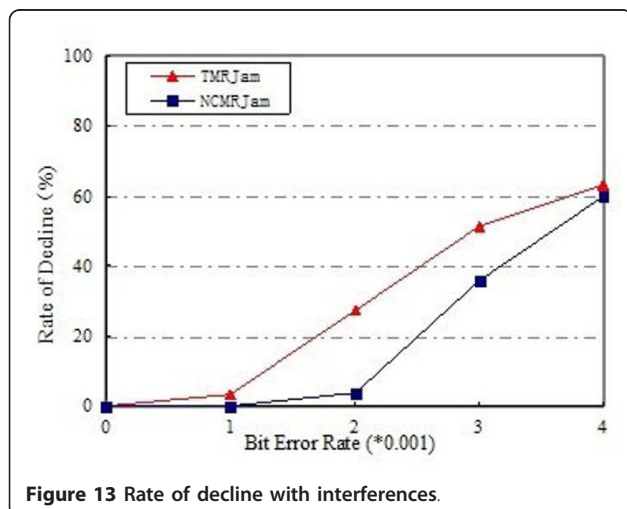

Figure 13 Rate of decline with interferences.

WSNs. The analysis shows that NCMR produces higher reliability and energy efficiency. In addition, based on the analysis of influence of different network parameters on performance in the network, a combination scheme of optimal parameters is obtained, which enables the network to deliver packets with lower energy consumption. The results of simulations show that NCMR is more effective in extreme environments such as interference of channels with memory and congested environment.

There is some further promising work in NCMR for WSN on energy efficiency.

(1) In this article, it is assumed that all nodes in the network are allowed to perform network coding, which requires high processing abilities of nodes. How to reduce the number of coding nodes requires further study.

(2) There is only one source node and sink node in current models, and previous study shows that network coding is powerful in both unicast network and multicast network. So, the network models with multiple sources and multiple sinks are worthwhile to address.

Competing interests
The authors declare that they have no competing interests.

References
1. S Mueller, RP Tsang, D Ghosal, Multipath routing in mobile ad hoc networks: issues and challenges. Lect Notes Comput Sci. **2965**, 209–235 (2004). doi:10.1007/978-3-540-24663-3_10
2. M Kim, E Jeong, Y-C Bang, S Hwang, C Shin, G Jin, B Kim, An energy-aware multipath routing algorithm in wireless sensor networks. IEICE Trans Inf Syst. **E91-D**(10), 2419–2427 (2008). doi:10.1093/ietisy/e91-d.10.2419
3. B Deb, S Bhatnagar, B Nath, ReInForm: reliable information forwarding using multiple paths in sensor networks, Proc IEEE LCN, (Bonn, Germany, 2003), pp. 406–415
4. S Choi, Y Choi, I Lee, IEEE 802.11 MAC-level FEC scheme with retransmission combining. IEEE Trans Wirel Commun. **5**(1), 203–211 (2006)
5. R Riemann, K Winstein, Improving 802.11 range with forward error correction, CSAIL, Technical Report, MIT-C, SAIL-TR-2005-011,(MIT, Cambridge, MA, USA, 2005)
6. JY Lee, WJ Kim, JY Baek, YJ Suh, A wireless network coding scheme with forward error correction code in wireless mesh networks, in *Proc of Globecom*, (Honolulu, HI, 2009), pp. 1–6
7. R Ahlswede, N Cai, Network information flow. Inf Theory. **46**(4), 1204–1216 (2000). doi:10.1109/18.850663
8. S-Y Li, R Yeung, N Cai, Linear network coding. IEEE Trans Inf Theory. **49**, 371–379 (2003)
9. T Ho, M Medard, DR Karger, M Effros, J Shi, B Leong, A random linear network coding approach to multicast. IEEE Trans Inf Theory. **52**(10), 4413–4430 (2004)
10. PA Chou, Y Wu, K Jain, Practical network coding, in *41st Annual Allerton Conference on Communication Control and Computing*, Monticello, USA, pp. 40–49 (October 20 2003)

11. S Katti, H Rahul, W Hu, D Katabi, M Madard, J Crowcroft, Xors in the air: practical wireless network coding. IEEE/ACM Trans Netw. **16**(3), 497–510 (2008)

12. D Nguyen, T Nguyen, B Bose, Wireless broadcast using network coding. IEEE Trans Veh Technol. **58**, 914–925 (2009)

13. T Tran, T Nguyen, B Bose, A joint network-channel coding technique for single-hop wireless networks, in *Proc NetCod*, Hong Kong, China, pp. 1–6 (January, 2008)

14. D Ganesan, R Govindan, S Shenker, D Estrin, Highly-resilient, energy-efficient multipath routing in wireless sensor network. Mob Comput Commun Rev. **5**(4), 10–24 (2001)

15. Z Guo, B Wang, P Xie, W Zeng, J-H Cui, Efficient error recovery using network coding in underwater sensor networks. Ad Hoc Netw. **7**, 791–802 (2009). doi:10.1016/j.adhoc.2008.07.011

16. S-S Li, P-D Zhu, X-K Liao, W-F Cheng, S-L Peng, Energy efficient multipath routing using network coding in wireless sensor networks. Lect Notes Comput Sci. **4104**, 114–127 (2006). doi:10.1007/11814764_11

17. YW Yang, CS Zhong, YM Sun, JY Yang, Network coding based reliable disjoint and braided multipath routing for sensor networks. J Netw Comput Appl. **33**(4), 422–432 (2010). doi:10.1016/j.jnca.2010.02.003

18. L Geng, F Lu, Y-C Liang, F Chin, Secure multi-path construction in wireless sensor networks using network coding, in *PIMRC, Cannes*, pp. 1–5 (2008)

19. AL Toledo, XD Wang, Efficient multipath in sensor networks using diffusion and network coding, in *CISS, Princeton*, pp. 87–92 (2006)

20. YW Yang, CS Zhong, YM Sun, JY Yang, Energy efficient reliable multi-path routing using network coding for sensor network. Int J Comput Sci Netw Secur. **8**(12), 329–338 (2008)

21. P Berman, G Calinescu, C Shah, A Zelikovsky, Efficient energy management in sensor networks, in *Proceedings of the Ad Hoe and Sensor Networks, Series on Wireless Networks and Mobile Computing*, (Nova Science Publishers, New York, 2005)

22. N Li, JC Hou, FLSS: a fault-tolerant topology control algorithm for wireless networks, in *Proceedings of MobiCom 2005, New York*, pp. 275–286 (2005)

23. XY Li, PJ Wan, Y Wang, CW Yi, Fault tolerant deployment and topology control in wireless networks, in *Proceedings of MobiHoc New York*, pp. 117–128 (2003)

24. M Cardei, DZ Du, Improving wireless sensor network lifetime through power aware organization. Proc ACM Wirel Netw. **11**(3), 333–340 (2005). doi:10.1007/s11276-005-6615-6

25. M Cardei, J Wu, Energy efficient coverage problems in wireless ad hoc sensor networks. J Comput Commun Sensor Netw. **29**, 413–420 (2006)

26. WR Heinzelman, AP Chandrakasan, H Balakrishnan, An application-specific protocol architecture for wireless microsensor networks. IEEE Trans Wirel Commun. **1**(4), 660–670 (2002). doi:10.1109/TWC.2002.804190

27. J Widmer, J-Y Le Boudec, Network coding for efficient communication in extreme networks, in *Proc ACM SIGCOMM 2005 Workshop on Delay Tolerant Networks*, Philadelphia, PA, (22–26 August 2005)

28. Y Wu, PA Chou, S-Y Kung, Minimum-energy multicast in mobile ad hoc networks using network coding. IEEE Trans Commun. **53**(11), 1906–1918 (2005). doi:10.1109/TCOMM.2005.857148

29. DS Lun, N Ratnakar, M M'edard, R Koetter, DR Karger, T Ho, E Ahmed, F Zhao, Minimum-cost multicast over coded packet networks. IEEE Trans Inf Theory. **52**(6), 2608–2623 (2006)

30. D Estrin, Wireless sensor networks tutorial part IV: sensor network protocols, in *Proc of the 8th Annual International Conference on Mobile Computing and Networking*, (Atlanta, GA, USA, ACM, 2002), pp. 23–28

31. B Deb, S Bhatnagar, B Nath, ReInForm: reliable information forwarding using multiple paths in sensor networks, in *Proceedings of the 28th IEEE Int'1 Conf. on Local Computer Networks (LCN)*, (Bonn, Germany, 2003), pp. 406–415

32. A Varga, OMNeT++ discrete event simulation system. http://www.omnetpp.org

33. MiXiM simulator for wireless and mobile networks using OMNeT++ [online]. http://mixim.sourceforge.net/