

# A Network Coding and Genetic Algorithm Based Power Efficient Routing Algorithm for Wireless Sensor Networks\*

Wen-wei Lu, Jian Pan, and Yi-hua Zhu

College of Computer Science and Technology, Zhejiang University of Technology,  
Hangzhou, Zhejiang 310023, China

lu.wen.wei@yahoo.com.cn, pj@zjut.edu.cn, yhzhu@ieee.org

**Abstract.** In a wireless sensor network (WSN), retransmission and acknowledgement (ACK) are required to make reliable packet delivery. In this paper, a Network Coding based Power Efficient Routing (NCPER) algorithm integrated with multi-path routing algorithms is proposed to eliminate retransmission and ACK, which guarantees that the receiving node can decode the original data sent by a source node so that the source node need not care whether the transmitted packets are lost or not. In addition, numeric experiments are conducted to show impacts of packet size and finite field size on energy cost of the NCPER. The NCPER expends energy efficiently and alleviates radio interferences among nodes.

**Keywords:** Wireless sensor network; routing; network coding.

## 1 Introduction

In a Wireless Sensor Network (WSN), data packets are delivered to a sink via multi-hop manner. To overcome the problem of packet loss, retransmission and acknowledgement (ACK) are applied in some protocols to make packets delivered reliably. Clearly, the probability that a packet is successfully delivered from a node to the sink declines when the number of communication hops increase, which may cause timeouts or ACK packet loss at some nodes so that these nodes retransmit periodically, i.e., retransmission becomes more seriously.

Most nodes in a WSN are powered by battery with limited energy. Hence, retransmitting and ACK lead nodes to quickly expend their available energy so that lifetimes of the nodes and the WSN are shortened. As a result, reducing number of retransmissions and ACKs is significant for a WSN to run effectively and efficiently.

Multi-path routing[1] is one of effective approaches to enhancing reliability of packet delivery. In addition, network coding [2][3] has the ability of making a

---

\* This project is supported by National Natural Science Foundation of China under grant No. 60873228 and by Zhejiang Provincial Natural Science Foundation of China under grant No. Y1080483.

destination node recover data packets sent by several source nodes if the destination node receives a sufficient number of coded packets. Moreover, network coding can significantly reduce the number of retransmissions in lossy networks compared to an ARQ (automatic repeat request) scheme [4]. In this paper, we integrate multi-path routing with network coding and propose a routing algorithm called Network Coding based Power Efficient Routing (NCPER) to eliminate retransmission and ACK in the WSNs so that energy is expended efficiently. Under the NCPER, a source node splits data into multiple pieces and chooses a group of linear independent vectors to code these data pieces to generate some new coded packets. Then, the coded packets are transmitted via distinct paths to the sink. The main contribution of the paper are: 1) under the NCPER, retransmission and ACK are not needed; and 2) the NCPER can reduce energy cost and alleviate radio interference.

## 2 NCPER

We assume there is only one sink in the WSN and it knows the topology of the WSN. We use  $S$  to represent a source node in the WSN and assume there are  $K$  paths from node  $S$  to the sink, denoted by  $P(1), P(2), \dots, P(K)$  respectively, which are constructed by a multipath routing algorithm (e.g., AOMDV [1]). As depicted in Fig.1, each path consists of some wireless links. Besides, we assign triple  $(h_i, P_i, E_i(k))$  to  $P(i)$ , where  $h_i, P_i$  and  $E_i(k)$  stand for the number of hops, probability of successfully delivering a packet from  $S$  to the sink, and energy expended for delivering a  $k$ -bit packet to the sink through  $P(i)$ , respectively. We denote  $P(i)$  by  $(e_{i,1}, e_{i,2}, \dots, e_{i,h_i})$ , where  $e_{i,j}$  stands for the  $j$ -th wireless link in  $P(i)$ ,  $j = 1, 2, \dots, h_i$ .

In a WSN, Energy consumption is mainly caused by radio communications. Thus, we only consider energy consumption of the radio and ignore other energy consumption in the sensor node. In addition, we apply the same model as proposed in [5], i.e., the energy costs of transmitting and receiving a  $k$ -bit packet between two nodes being  $d$  meters apart are, respectively, as follows [5]:

$$E_{Tx}(k, d) = k(E_{elec} + \epsilon_{amp}d^\gamma) \tag{1}$$

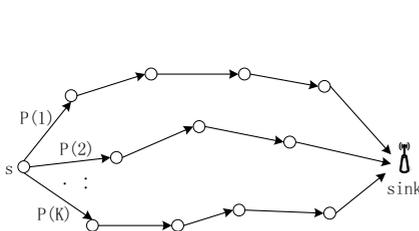


Fig. 1.  $K$  paths from  $S$  to the sink

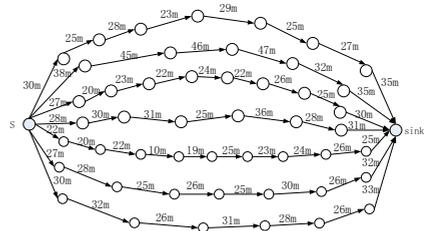


Fig. 2. Multiple paths from source node  $S$  to the sink

$$E_{Rx}(k) = kE_{elec} \tag{2}$$

where  $\gamma \in [2, 4]$  is the path loss exponent;  $E_{elec}$  denotes the energy consumption due to digital coding, modulation, filtering, and spreading of the signal, etc., and  $\epsilon_{amp}$  is the energy consumed by the transmitter power amplifier.

Assume failure probability of transmitting a packet on wireless link  $e_{i,j}$  is  $p$ , i.e., successful probability is  $1 - p$ , where  $i = 1, 2, \dots, K, j = 1, 2, \dots, h_i$ . Let  $H(i)$  be the maximum number of hops of a packet being transmitted on  $P(i)$  before it is lost. Then, we have the following

$$p_i = (1 - p)^{h_i}, i = 1, 2, \dots, K$$

$$Pr\{H(i) = n\} = \begin{cases} p(1 - p)^n, 0 \leq n < h_i \\ (1 - p)^{h_i}, n = h_i \end{cases}$$

$$\bar{H}(i) = \sum_{n=0}^{h_i} nPr\{H(i) = n\} = \frac{1 - (1 - p)^{h_i+1}}{p} - 1 \tag{3}$$

where  $\bar{H}(i)$  is the average number of hops. Thus, the energy cost for a  $k$ -bit packet to reach the sink from node  $S$  through  $P(i)$  is as follows:

$$E_i(k) = \sum_{j=1}^{\bar{H}(i)} [E_{Tx}(k, d(e_{i,j})) + E_{Rx}(k)]. \tag{4}$$

Assume the message to be delivered to the sink from the source node  $S$  is composed of a string of characters from finite field  $\mathbb{F}_{2^n}$  [6]. Although data are generated continuously by sensor nodes, they are collected periodically in some applications. Thus, we consider the situation that there is a fixed number of packets at a sensor node to be delivered to the sink. Especially a node only sends  $N$  blocks of data, say  $\alpha_1, \alpha_2, \dots, \alpha_N$ , once a time. Assume each block consists of  $m$  characters from  $\mathbb{F}_{2^n}$ . Additionally, node  $S$  produces the following coded data block (CDB) based on network coding:

$$\beta_i = r_{i1}\alpha_1 + r_{i2}\alpha_2 + \dots + r_{iN}\alpha_N \tag{5}$$

where  $r_i \equiv (r_{i1}, r_{i2}, \dots, r_{iN})$  is termed as  $i$ -th coding vector (CV) and  $r_{ij} \in \mathbb{F}_{2^n}$ .

As soon as the sink receives  $N$  linearly independent CDBs  $\beta_1, \beta_2, \dots, \beta_N$ , we, from (5), have  $(\beta_1, \beta_2, \dots, \beta_N)^T = R(\alpha_1, \alpha_2, \dots, \alpha_N)^T$ , where  $R = (r_1, r_2, \dots, r_N)^T$ . Thus, the sink can recover the original data as follows.

$$(\alpha_1, \alpha_2, \dots, \alpha_N)^T = R^{-1}(\beta_1, \beta_2, \dots, \beta_N)^T \tag{6}$$

Obviously, solution of (6) exists if matrix  $R$  is full rank, i.e., the CVs  $r_1, r_2, \dots, r_N$  are linearly independent.

### 3 Combinational Optimization with the NCPER

The key of the NCPER is an optimization problem to find an optimal pair  $(L, N)$  such that all of the following are satisfied:

- (i)  $N \leq L \leq K$ ;
- (ii)  $L$  CDBs are generated by node  $S$  such that arbitrary  $N$  of them are linearly independent;
- (iii) The sink is guaranteed to receive at least  $N$  CDBs;
- (iv) The energy cost per bit (ECPB) for delivering  $N$  CDBs through  $L$  distinct paths is minimized.

Condition (ii) can be easily met by taking the following four steps: 1) limit  $L \leq 2^n$ ; 2) take  $L$  different elements  $\theta_1, \theta_2, \dots, \theta_L$  in  $\mathbb{F}_{2^n}$ ; 3) let  $r_i \equiv (1, \theta_i, \theta_i^2, \dots, \theta_i^{N-1})$ ,  $i = 1, 2, \dots, N$ ; and 4) generate CDBs as follows:

$$\beta_i = 1\alpha_1 + \theta_i\alpha_2 + \theta_i^2\alpha_3 + \dots + \theta_i^{N-1}\alpha_N \quad (i = 1, 2, \dots, L). \tag{7}$$

The reason that arbitrary  $N$  of CDBs  $\beta_1, \beta_2, \dots, \beta_L$  are linearly independent is that the CVs of these  $N$  CDBs form a full rank Vandermonde determinant valued  $\prod_{1 \leq j < i \leq N} (\theta_i - \theta_j) \neq 0$ .

Now, we consider the other three conditions. Let  $\Phi(x)$  be number of bits  $x$  contains. Thus,  $\Phi(r_{ij}) = n$  since  $r_{ij}$  is in  $\mathbb{F}_{2^n}$ , and besides,  $\Phi(\alpha_i) = nm$  as  $\alpha_i$  contains  $m$  characters in  $\mathbb{F}_{2^n}$  ( $i = 1, 2, \dots, N$ ). Moreover, from (6), we have  $\Phi(\beta_i) = \Phi(r_{i1}\alpha_1 + r_{i2}\alpha_2 + \dots + r_{iN}\alpha_N)$ . Noting that  $\alpha_i$  contains  $m$  characters in  $\mathbb{F}_{2^n}$  and  $r_{ij} \in \mathbb{F}_{2^n}$ , it can be proved that  $r_{i1}\alpha_1 + r_{i2}\alpha_2 + \dots + r_{iN}\alpha_N$  is also a string of  $m$  characters in  $\mathbb{F}_{2^n}$ . This is because finite field  $\mathbb{F}_{2^n}$  has the property that any linear combination of its elements is also in it. As a result,  $\Phi(\beta_i) = \Phi(\alpha_i) = nm, i = 1, 2, \dots, L$ .

Clearly, a CDB has to be transmitted with a CV so that the sink is able to recover the original data from  $N$  received CDBs. Consequently, a packet being transmitted in a wireless link consists of three portions: the packet header, a CDB, and a CV. We ignore the number of bits the header contains and assume the size of a packet is  $B$ . Then, the number of bits a packet contains is  $\Phi(\beta_i) + \Phi(r_i) = \Phi(\beta_i) + N\Phi(r_{ij}) = n(m + N)$ , i.e.,

$$n(m + N) = B \tag{8}$$

Let  $x_i$  be an indicator:  $P(i)$  is chosen for routing a packet if  $x_i = 1$ , otherwise if  $x_i = 0$ . Thus, the number of packets successfully reaching the sink through  $P(i)$  is  $x_i p_i$ . Besides, we have  $L = \sum_{i=1}^K x_i$ . To guarantee that there are  $N$  packets arriving at the sink, the following should hold:

$$N \leq \sum_{i=1}^K x_i p_i \tag{9}$$

From (8), the number of bits of payload in  $N$  packets is  $N\Phi(\alpha_i) = Nnm = N(B - nN)$ . Hence, from (4), we have ECPB as follows.

$$E_{bit}(X) = \frac{\sum_{i=1}^K x_i E_i(B)}{N(B - nN)} = \frac{\sum_{i=1}^K x_i \sum_{j=1}^{\overline{H}(i)} [E_{Tx}(B, d(e_{i,j})) + E_{Rx}(B)]}{N(B - nN)} \tag{10}$$

Therefore, the optimization problem earlier mentioned is equivalent to the following combination optimization problem (COP):

$$\begin{aligned} \text{Min} \{ & \frac{\sum_{i=1}^K x_i \sum_{j=1}^{\overline{H}^{(i)}} [E_{Tx}(B, d(e_{i,j})) + E_{Rx}(B)]}{N(B-nN)} \} \\ \text{s.t. } & \begin{cases} N = \lfloor \sum_{i=1}^K x_i p_i \rfloor \\ x_i \in \{0, 1\}, i = 1, 2, \dots, K \end{cases} \end{aligned} \quad (11)$$

where  $\lfloor \cdot \rfloor$  is the floor function. We adopt a genetic algorithm (GA) to solve the COP. A chromosome is defined as  $\xi \equiv x_1 x_2 \cdots x_K$ , where  $x_i \in \{0, 1\}, \xi \in \Omega \equiv \{0, 1\}^K$ . Moreover, we define fitness as

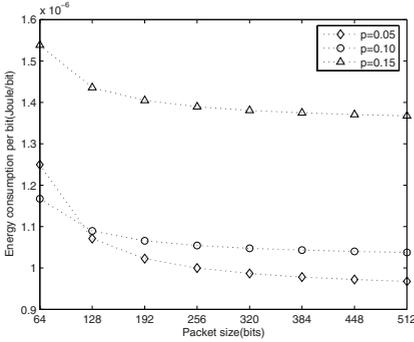
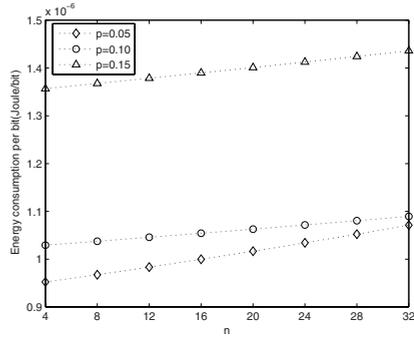
$$f(\xi) \equiv \frac{1}{E_{bit}(X)} = \frac{N(B-nN)}{\sum_{i=1}^K x_i \sum_{j=1}^{\overline{H}^{(i)}} [E_{Tx}(B, d(e_{i,j})) + E_{Rx}(B)]} \quad (12)$$

The size of population is set to  $M$ . The initial population is generated by producing  $M$  chromosomes, with each being a  $K$ -bit binary string. Selection operation is performed by calculating fitness in (12) for each chromosome and choosing a chromosome with probability  $f(\xi) / \sum_{\eta \in \Omega} f(\eta)$  according to roulette wheel. Crossover operation is based on single random cross point with probability  $p_c$ . For instance, for two chromosomes  $\xi = x_1 x_2 \cdots x_K$  and  $\eta = y_1 y_2 \cdots y_K$ , crossover operation make two new chromosomes as  $\xi' = x_1 x_2 \cdots x_i y_{i+1} y_{i+2} \cdots y_K$  and  $\eta' = y_1 y_2 \cdots y_i x_{i+1} x_{i+2} \cdots x_K$ , if the cross point is  $i (1 \leq i < K)$ . Mutation is performed by choosing a gene in a chromosome with probability  $p_m$  and changing the gene by xoring it with 1.

The NCPER is based on multiple disjoint paths, which are constructed by AOMDV [1]. Considering that a destination node in AOMDV knows all the paths from a source node to itself, we solve the COP in (11) by running GA at the sink which has ample power. Then, the optimal  $N$  together with the optimal  $L$  paths are sent to source node  $S$ , in which  $N$  is used to create CDBs that are delivered through the  $L$  received distinct paths to the sink. Eventually, the sink decodes the received packets to obtain the original data.

## 4 Numeric Experiments

Our experiments are based on Fig.2, in which there are 7 paths from source node  $S$  to the sink. An arrow represents a wireless link between two nodes, and the distance of two nodes is labeled beside the wireless link connecting the two nodes. Following [5], we set  $\gamma = 2, E_{elec} = 50nJ/bit$  and  $E_{amp} = 10pJ/bit/m^2$ . In addition, we take  $n=8$ . Moreover, for the GA, we choose  $M = 100, p_c = 0.8$  and  $p_m = 0.05$ . The GA evolves 100 generations. Fixing  $p=0.05, 0.10, 0.15$ , respectively, and letting  $B$  vary, lead to Fig. 3, which indicates that 1) for a particular  $p$ , ECPB decreases as packet size  $B$  increases, i.e., our NCPER will bring more benefits if a larger  $B$  is used; and 2) ECPB increases as  $p$  goes high when  $B$  is larger than 128bits.

Fig. 3. ECPB vs.  $B$ Fig. 4. ECPB vs.  $n$ 

Setting  $B=512$  and letting  $n$  vary, we have Fig. 4, which shows that ECPB increases with  $n$  growing. This is because the energy cost for transmitting CVs increases as  $n$  increases since more bits are transmitted. In other words, we should choose a small  $n$  so that energy cost is reduced. However, it should be stressed that network coding will be impractical if  $n$  is too small. For instance,  $n$  can not be less than 8, if node  $S$  and the sink communicate with symbols consisting of ASCII characters, with each composed of 8bits.

## 5 Conclusion

The proposed NCPER eliminates retransmission and ACK and is able to guarantee the sink can decode the original data. The numeric experimental results in the previous section show that larger size of packet and smaller size of finite field should be chosen for the NCPER. In future, we will be comparing the NCPER with some well-known multi-path routing algorithms.

## References

1. Marina, M.K., Das, S.R.: On-demand Multipath Distance Vector Routing for Ad Hoc Networks. In: Ninth International Conference on Network Protocols, November 11-14, pp. 14–23 (2001)
2. Fragouli, C., Soljanin, E.: Network coding applications. Now Publisher Inc., USA (2007)
3. Ahlswede, R., Cai, N., Li, S.-Y.R., Yeung, R.W.: Network information flow. IEEE Transactions on Information Theory 46, 1204–1216 (2000)
4. Ghaderi, M., Towsley, D., Kurose, J.: Reliability Gain of Network Coding in Lossy Wireless Networks. In: The 27th IEEE Conference on Computer Communications, INFOCOM 2008, April 13-18, pp. 2171–2179 (2008)
5. Heinzelman, W.B., Chandrakasan, A.P., Balakrishnan, H.: An application-specific protocol architecture for Wireless microsensor networks. IEEE Transactions on Wireless Communications 1, 660–670 (2002)
6. Bartee, T.C., Schneider, D.J.: Computation with finite fields. Information and Computing 6, 79–98 (1963)