Research on Transmission Based on Collaboration Coding in WSNs

LV Xiao-xing, ZHANG Bai-hai
School of Automation
Beijing Institute of Technology
Beijing 100081, China
lvxx@mail.btvu.org

Journal of Digital Information Management

ABSTRACT: *In order to reduce the increase of network delay and excessive energy consumption, which are caused by network code in WSNs, a fast random linear mechanism of network coding based on nodes collaboration is introduced. Energy consumption of nodes is reduced by transmitting the initial code of the node, and transmitting linear independent packets received merely. After gathering node coding, ACK return along the best path to reduce the network delay. Simulation results show that delay of data transmission is reduced greatly by introducing random linear network coding for sensor network, and the improved coded mode balances the network load.*

Categories and Subject Descriptors:
C.2.1 [Network Architecture and Design]: *Wireless Communication*; E.4 [**Coding and Information Theory**]

General Terms:
Wireless Networks, Collaboration Coding, Sensor Networks

Keywords: Wireless Sensor Network, Nodes Collaboration, Random Linear Network Coding

Received: 30 June 2013, Revised 3 August 2013, Accepted 10 August 2013

1. Introduction

Among existing wireless communication networks, the information of multicast is transmitted from source node to target node with the method of storing and forwarding in intermediate nodes. Because of the influence of channel quality, the traditional mode of storing and forwarding is difficult to reach the maximum capacity of the network. To improve the performance of multi-hop and self-organized wireless network, A. Ahlswede and others proposed the concept of network coding [1] in 2000. They theoretically proved that if network node is allowed to encode the information with appropriate method, such as mode 2 plus and operations on finite field, rather than restricted to storing and forwarding, then multicast based on this mode can achieve the theoretical maximum transmitting capacity. The realization of network coding depends on intermediate nodes, which encode and transmit received information. Nodes with capabilities of network coding get information from there input link, then information is transmitted to every output links after corresponding processes to improve the use of network bandwidth and increase throughput of network. The complexity of coding algorithms is an important basis to measure whether the network coding can be implemented effectively. From the practical view, polynomial-time algorithm does not apply to the network, which has dynamic changed topology and large scale.

Random network coding proposed by Ho. Medard and others in 2003 is no longer limited to determined topology, but involves information cache and matrix operations, and additional overhead is larger than traditional route transmission in terms of wireless sensor networks. Even the simplest random network coding, the coding and decoding time can not be ignored. This paper linearly pre-judges forwarding nodes in the basis of random linear network coding, to reduce the complexity of network coding with minimum cost. [4] study the use of intra-session random linear network coding (RLNC) in wireless sensor networks. In RLNC, intermediate nodes buffer the packets received from upstream nodes. Using intra-session RLNC, intermediate nodes transmit coded packets by performing coding on the packets of various flows. In [5], a distributed algorithm is developed to effectively detect, locate, and isolate the Byzantine attackers in a wireless ad hoc network with random linear network coding (RLNC).

2. Network coding ideas

According to maximal flow and minimal cut of network

information flow, the maximal value W of flow that from source S to information sink T equals to the capacity of its minimal cut. Namely, max flow $(S, T) = \min C(S, T)$. Consider a network shown in Figure 1, W is bottleneck of network transmission because multicast is a simple mode of storing and forwarding. It can not transmit b_1 and b_2 once a time. We add coding mode to nodes of W and S and take a coding plan shown as figure 2. As shown in figure, b_1 and b_2 are transmitted to information sinks Y and Z after XOR processing, $\oplus$ represents mod 2 plus. In this case, receiver Y receives b_1 and $b_1 \oplus b_2$, then b_2 can be recovered by plusing b_1 and $b_1 \oplus b_2$. For the same season, receiver Z can recover b_2. One more bit is sent if there is no coding, so network coding reduces the number of forwarding and the energy consumption of network. Network coding is an effective way that approximates the limited theorem of network transmitting capacity.

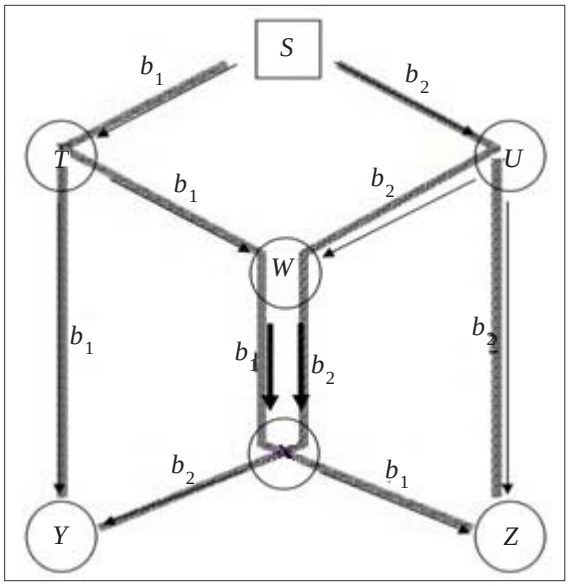

Figure 1. Multicast of WSN without coding

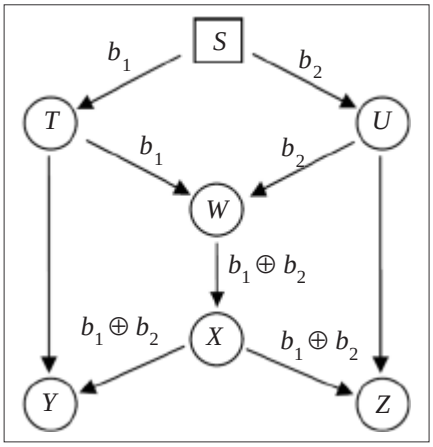

Figure 2. Multicast of WSN with network coding

3. Random linear network coding of WSNs

Linear network coding maps transinformation of nodes into a finite field, and implements encoding and decoding process [6] using linear relationship. Consider the size of each data packet is L bits. If the size is different with the data packet, with which it will compose, some 0 is added to the shorter information. S continuous bits in the data packet are composed of a sign of the field, and then there are L/S signs in one packet. Multiplication and addition are used in the linear coding that the data sent from nodes is linear combination of information received by the node.

Deterministic coding strategy and random coding strategy can be adopted for choosing coding vector. Deterministic network coding requires us to understand the whole network and its complexity is high, so it is hard to implement distributed. Because wireless sensor network has features of large scale and dynamic routing, random linear network coding is chosen in this paper. Besides, coding vector is optimized in forwarding node to reduce the transmission cost of data packet after coding.

3.1 Random Coding Strategies

We defined a coding vector in the head of data packet. A new set of linear independence coefficient is generated to modify the coding vector when data packet reached the intermediate node. As a result, the sink node does not need to know the topology of entire network and the situation of network coding when it received the information. As long as the number of received data packets reaches the default size of a batch, which is defined as 8 data packets in the test, the information of this batch can be got. Random coding also provides that number is chosen randomly in the finite field, when intermediate node distributes local coding vector for edge. As a result, receiver does not need to know the topology of the network.

When intermediate node receives a new data packet, it fetches the factor of coding to measure the linear interdependence with previous data packets. The new data packet is pushed into sending stack of the MAC if linearly independence, otherwise, the packet is discarded. So the number of forwarding and the commandeer conflicts of information channel are reduced.

3.2 Specific Algorithms

3.2.1 Data Coding

Data set to be sent is divided into $Data = \{B_1, B_2, B_3\}$. When node is sending data in the way of commandeering information channel, K chosen random number is composed of K-dimensional vector $\{r_1, r_2, \ldots, r_k\}$ according to the generating algorithm of Madhi Mersenne twister pseudo-random number firstly. The K-dimensional vector is put into CODE VECTOR field of the packet head to generate a new data Data, which is multicast.

According to theorem, the linear combination of a set of linear independent vectors is still linear independent. When other nodes in the network monitored new data package, they generate a set of linear independent random number and code continually. Consider node j received a data

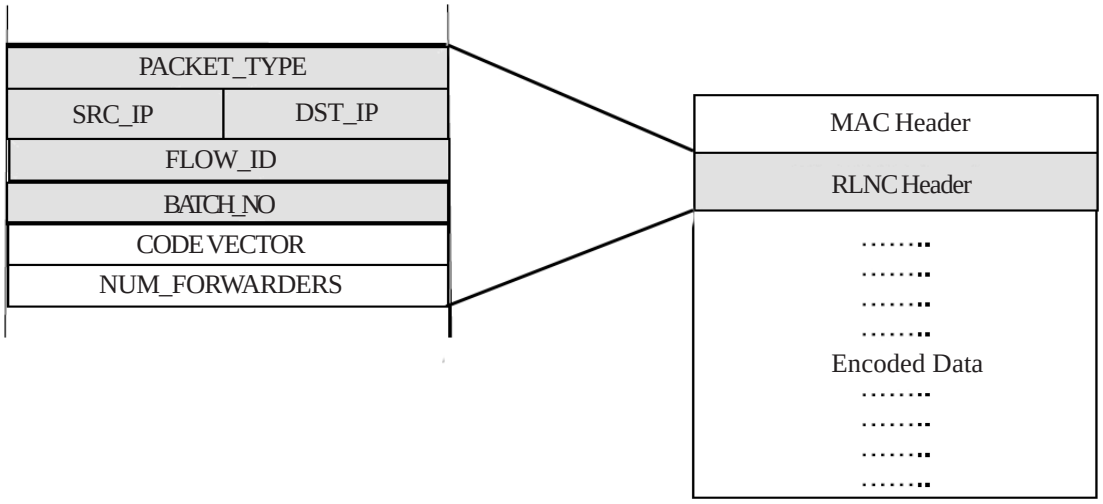

Figure 2. The mode of data packet

packet B', namely $B_j' = \sum_j C_{ji} B_j$. Then j coded is

$$B'' = \sum_j r_j B_j' \quad (1)$$

r_j is random linear factor of node j. Packet sent by node j is expressed with raw data as

$$B'' = \sum_j r_j (\sum_j C_{ji} B_i) \sum_i (\sum_j r_j C_{ji}) B_i \quad (2)$$

which still is linear combination of raw packet.

3.2.2 Sink Decoding

For each data packet that arrives at sink node, sink node will measure the interdependency of coding vectors. Dependence means the packet does not have new information, the packet is discarded. Uncorrelated packets are stored. Once K linear independent packets are received, the information of whole batch can be got through the follow matrix.

$$\begin{bmatrix} B_1 \\ \vdots \\ B_K \end{bmatrix} = \begin{bmatrix} C_{11} & \cdots & C_{1K} \\ \vdots & \ddots & \\ C_{K1} & \cdots & C_{KK} \end{bmatrix}^{-1} \begin{bmatrix} B_1' \\ \vdots \\ B_K' \end{bmatrix} \quad (3)$$

Sink node sends an ACK when it gets the raw packet $\{B_1, B_2, \ldots, B_k\}$. ACK returns along the best router to source node, to notify source node to send data packets for next batch. The priority of ACK is highest for each node. This can make some forwarding nodes with great delay stop multicasting data packets.

3.2.3 Measure of linear dependency among forwarding nodes

The coding cost is proportional to the number of data packets. When node received linear dependent packet, it means that the packet does not contain new information comparing with packets received before. Coding and broadcast does not make sense and only result in waste of energy. As a result, when one of the forwarding nodes receives data packet, data packet with new information is stored after comparing interdependency. Linear dependent packets are discarded because they do not have new information. According to linear interdependency, we can know that the number of data packets with new information has an upper limit, which is the size of a batch. View from the channel features of wireless sensor network, it is essential to consider appropriate packet discarded strategy because the number of packets broadcasted far greater than new packets received.

It is considered that the computing ability and energy of sensor node are limited. In order to avoid massive numeric operations and transmitting loss, the follow algorithm is used to code and measure linear dependency.

algorithm: linear dependent measure of coding vector
  for $i = 1 \ldots K$ **do**
    if $u[i] \neq 0$ **then**
      if $M[i]$ exists **then**
      $u \leftarrow u - M[i]\, u[i]$
      else
        {admit the *modified* block into memory}
        $M[i] \leftarrow u/u[i]$
        Return True {rank increased}
  return False {discard packet}

In this case, NK multiplications are needed for each data packet, N is the number of null string in the K-dimensional matrix. When node measures linear dependency, there are operations and modifications on matrix.The information in the packet does not change. The information participates in operations in time of coding or decoding, the cost of linear dependency measure is reduced greatly in result.

4. Simulation and performance analysis

We simulated and verified the multicast added with random linear network coding, studied the comparison of

multicast based on the random linear network coding strategy. The comparison is about energy consumption, network delay and throughput under the same conditions. The Without *NC* of experiment is flood method as default.

The test model composes of nodes distributed randomly in the 600 × 600 square field. The initial number of the nodes is 10, the initial energy of each node is 3J, the actual load of data packet is 20bit, and the communication bandwidth is 1M /s. The loss rate of each information channel is 10%, and the agreement of 802.11 is used by MAC. The standard length of data packet is set as 1024 bytes.

The average delay of the network is shown in figure 4. The test records the time from sending a packet to the sink node receives K linear independent packets. The retransmission of data can be reduced effectively for the distant nodes because the coding reduced the number of sending packets.

The average energy loss of nodes is shown in figure 5. Because of the usage of network coding and the building of per-hop and multi-path, the reliability of wireless sensor network is improved greatly. This is directly reflected in the reducing of transmission path, and choosing without repetition. It is favorable to the balances of network load. The number of data transmission is reduced, especially in the situation that node location is relatively remote and signal is rejecting. The reduction of path and the fault tolerance of network greatly reduce the energy loss. Although coding and decoding need energy loss, it can be ignored in compare with the energy loss saved by network coding.

The change of throughput capacity with the increase of nodes is shown in figure 6. As shown in the figure, the transmission rate increases obviously after network coding. Especially when there are 15 to 25 nodes in the scope of 600×600, the use of throughput capacity reach to maximum. With the increase of node density, the commandeer conflicts of information channel is aggravated that the throughput capacity is reduced once more.

6. Conclusion

Network coding is proposed for improving throughput rate of wireless network multicast at first. To improve the robustnest and reliability of data transmission of the wireless sensor network, we study how to apply random linear network coding to the data tranmission. This study is to reduce the influence of link failure and repeated

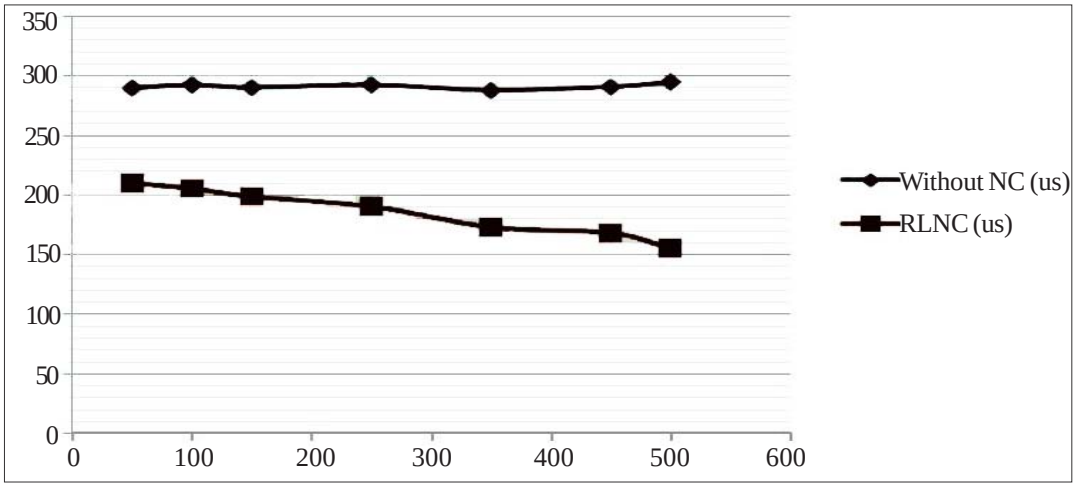

Figure 4. The comparison of average network delay

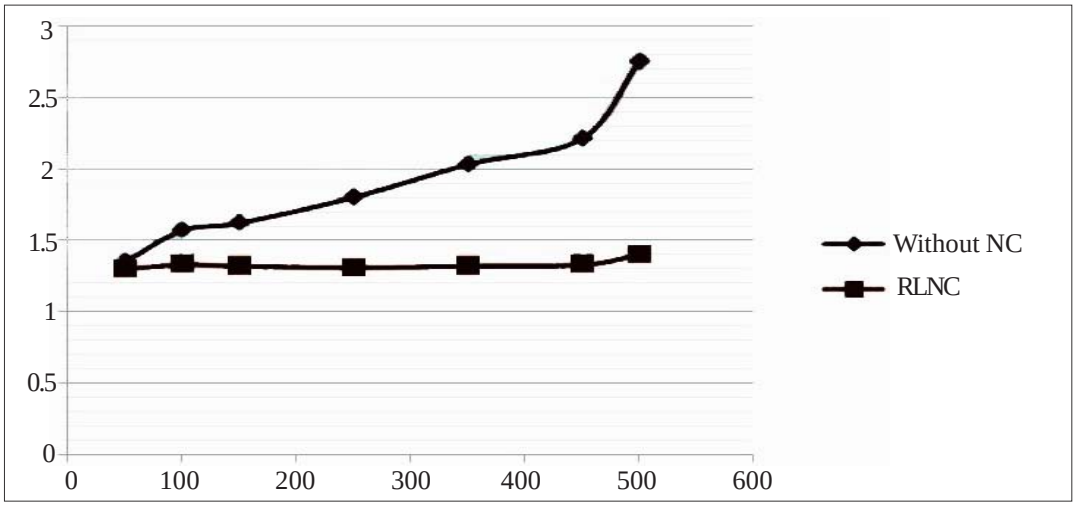

Figure 5. The comparison of average energy loss

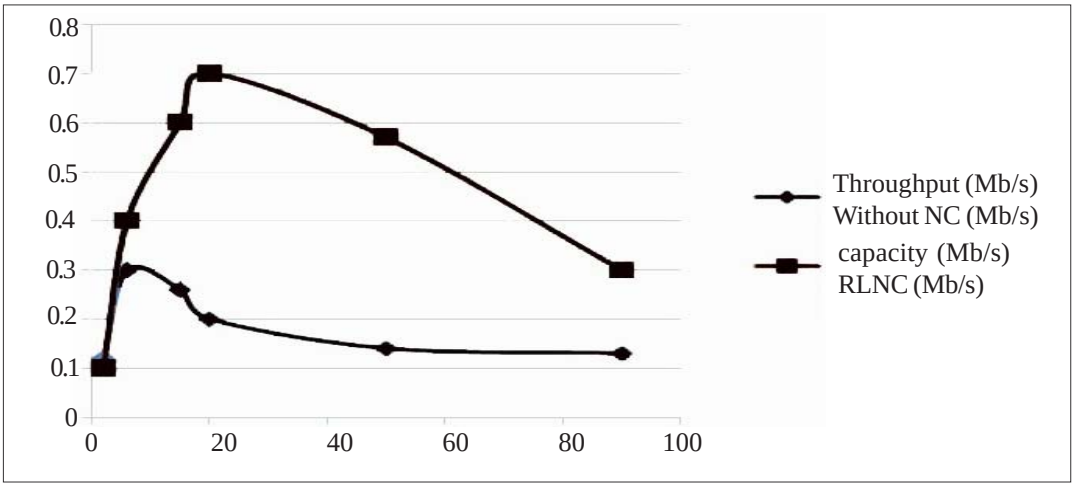

Figure 6. The comparison of throughput capacity

sending of same packets. Network coding not only code in source node and decode in target node, but check and recover data and code once more in intermediate node. Experimental results show that, network is well adapted to the features of sensor network, such as dynamic and large scale. In the future, we will study the routing algorithm that integrates random linear network coding into m-paths, to control the evaluation parameters of Date Flow Control and Quos.sending of same packets. Network coding not only code in source node and decode in target node, but check and recover data and code once more in intermediate node. Experimental results show that, network is well adapted to the features of sensor network, such as dynamic and large scale. In the future, we will study the routing algorithm that integrates random linear network coding into m-paths, to control the evaluation parameters of Date Flow Control and Quos.

References

[1] Rudolf Ahlswede, Ning Cai, Shuo-Yen Robert Li, et al. (2000). Network information flow. *IEEE Trans. Inform. Theory*, 46 (4) 1204-1216.

[2] Patrick, P. (1994). Minimization Method for Training Feedforward Neural Networks. *Neural Networks*, 7, p. 1-11.

[3] Ho, Medard, M., Shi, J. (2003). On Randomi zed Network Coding. 41st Annual Allerton Conference on Communication Control and Computing, p. 442- 445.

[4] Shah-Mansouri, Wong, V., VWS. (2012). Link-by-Link Feedback Mechanism for Intra-Session Random Linear Network Coding in Wireless Sensor Networks. *IEEE International Conference on Communications* (ICC), 10-15.

[5] Chen, J. Y., Tseng, Y. Y. (2012). Distributed Intrusion Detection of Byzantine Attacks in Wireless Networks with Random Linear Network Coding. *International Journal of Distributed Sensor Networks* Theory.

[6] Fragouli, Christina., Le Boudec, Jean-Yves Widmer, Jorg. (2002). Network coding: an instant Primer. *IEEE Transactions on Information Theory*, 48, p. 271-277.