# Security Analysis and Evaluation of Random Linear Network coding under Wiretapping Attack

LV Xiaoxing, ZHANG Baihai
Beijing Institute of Technology
School of Automation
Beijing 100081, China
lvxx@mail.btvu.org, zhangbh@bit.edu.cn

Journal of Digital
Information Management

***ABSTRACT:*** *This paper discussed the weak security of random linear network coding in wireless network in view of network coding weak security theory where the wiretapper can successfully obtain initial data and conditions threatening transmission security by decoding. The experiment results displayed, in wireless network, the relationship between decodes success probability of multi-node co-wiretapping, the decodes time of destination node and the node number in round and network, the number of data packet which source node sends every time, Galois field size and transmission radius of each node.*

**Categories and Subject Descriptors:**
**C.2.1 [Network Architecture and Design]:** Wireless Communication; **D.4.6** Security and Protection

**General Terms:**

**Keywords:** Random Linear Network Coding, Security, Multi-node Co-wiretapping

**Received:** 8 July 2013, Revised 18 August 2013, Accepted 23 August 2013

## 1. Introduction

Network coding theory is a major development in communications field, aggregating codes, routing and information theory. It allows coded combination for different links such that intermediate node can route and code and network performance will reach the theoretic limit of maximum flow transmission. Research found that network coding has an advantage in improving network throughput, bettering load balancing, reducing transmission delay, reducing energy consumption, enhancing network robustness. Therefore, network coding is suitable for unstable, time-varying wireless network.

Being sensitive to malicious attack or transmission error, network coding cannot defense wiretapping attack and one critical error will lead to decodes failure. Since network coding theory in 2003, security problem in adopting network coding had been a limelight. Ning Cai and Haowei Yang first proposed a communication system on a wiretap network(CSWN) and proved a necessary condition of implementing network security by network coding, meaning that source node sends message to destination node without leaking out any useful information to wiretapper [1]. Based on this, Ning Cai and Weihao Yang proposed and proved the necessary and sufficient conditions of implementing security when linear network coding is adopted in multiple source network by researching the algebra structure of network coding [2]. Kamal Jain, based on Cai and Yang's wiretap network, proved that if there exists an wiretapper-free link from source node to destination node in network which contains one source node to destination node, then the network is secure and the wiretapper cannot obtain any useful information [3]. Kapil Bhattad and Kdrishna R. Narayanan proposed a weakly secure network coding [4] and defined network security as wiretapper cannot obtain "*meaningful*" information from source node. They pointed out that system security probability can approach 1 if both random network coding is used and codes field is big enough. Jianlong Tan and Muriel Medard studied the cost standard problem of secure network coding [5] and adopted random linear codes. Cai and Yeung first proposed network error correcting codes concept [6] an designed an information theory secure network coding [7] on account of certain number of wiretappers in network. Jaggi et al. [8] considered network confronting both pollution attack and wiretapping attack and proposed an adaptively secure network coding based on coset code. Terence Chan and Alex Grant set up the multicast capacity boundary of secure network coding [9]. Silva et al. proposed rank-metric codes [10-11] based on classic error correcting theory and Rank-Metric Codes. Zhang proposed a security strategy (permutation codes) to defend global wiretapper. Xikun Wu et al. put forward a global wiretapper resilient Shannon secure network error correcting codes. With theoretically proving, this secure network error correcting codes wan Shannon secure and able to defeat global eavesdropper. Moreover, the proposed codes will not suffer from any rate loss for the enhanced security [13].

Last but not the least, existing research results are mostly focused on discussing and algorithm simulation, real wireless network coding experiment systems are needed home and abroad. Random linear network coding security is to be studied considering the network coding practicability. In wireless network, wiretapper can easily receive any broadcast data such that wireless network sometimes cannot satisfy the requirement of network coding weak security. Using random linear network coding information transmission system based on Matlab 2009b, this paper discussed weak security of random linear network coding in wireless network where wiretapper can successfully obtain initial data and conditions threatening transmission security by decoding. We studied the relationship between decodes time of destination node and the node number in round and network, the number of data packet which source node sends every time, Galois field size, transmission radius of each node and wiretapper's ability in wireless network by quantitative analyzing. We also make clear the relationship between success probability of multi-node co-wiretapping and network variables. All work aims to efficiently solve the wiretapping problem in wireless sensor network and wireless ad hoc networks, laying the foundations for improving global wiretapper resilient secure network coding.

## 2. Basic Model and Concept

### 2.1 Network Model

In wireline network, directed weighted graph is used to describe network topology. Incidence matrix and linear equation are used to establish theory framework of network coding.

Wireless network has a randomly varying topology and cannot be described through directed weighted graph.

Therefore, we use a random graph as the wireless network model, assuming each node has the same covering radius $R$.

In random graph $G(V, E)$,$V$ is node set, $E$ is edge set and link node. In node set $V = \{v_1, v_2,\ldots,v_n\}$, any node is distributed in $S \times S$ square field. Assuming distance between node $v_i$ and $v_j$ is $D(i, j)$, there is a direct link between $v_i$ and $v_j$, if and only if $D(i, j) \leq R$.Therefore, edge set:

$$E = \{(v_i, v_j) \mid D(i, j) \leq R, 0 < i,\ j < n \qquad (1)$$

When $R$ is relatively small, the probability of nodes linking is $P = \Pi R2$; when $R$ is relatively big, $P < \Pi R$ [16].

Assume the format of information that source node sends is as follows:

$$X_{m \times n} = \begin{bmatrix} x_{11} & x_{11} \cdots & x_{1n} \\ x_{21} & x_{22} \cdots & x_{2n} \\ \vdots & \vdots \quad \ddots & \vdots \\ x_{m1} & x_{m2} \cdots & x_{mn} \end{bmatrix} = (X_1\ \ X_2 \cdots X_n) \qquad (2)$$

$X_i\,(i = 1, 2,\ldots, m)$ is information packet and $X_{i,j} \in F_q$ is an individual data. Data number of individual information packet is $n$ (length of information packet). $F_q$ is finite field with size $q$.

### 2.2 Attack model

Single source single destination model: define source as Alice, destination as Bob, Calvin as wiretapper. Adversary Calvin eavesdrops on a few channels to obtain information sent from Alice to Bob. Assume a channel set $\Lambda = \{A_1, A_2,$ ..., $A_{|\Lambda|}\}$ where $A_i \in E$ is what Calvin can eavesdrop on in unit time and set $\wedge$ does not change over time. Row number of $A_i = k_i$ and $k = max\ \{k_1, k_2,\ldots\}$ where $k_i\,(i = 1, 2,\ldots)$ means link number Calvin can get in unit time and $k \leq |\Lambda|$.

### 2.3 Correlation Conception

**Maximum network flow:** Theoretically speaking, the number of information packet that Alice can send to Bob in unit time. Here we assume the maximum network flow between Alice and Bob is $m$.

**Multicast capacity:** The number of information packet that Alice can send to Bob in unit time when using a specific code strategy and facing an adversary.

**Information theory secure:** $M$ stands for any information set and $U$ stands for a subset of source information $X$. If $I\,(U\,; M\,) = 0$, then we say $M$ does not leak out any information about $U$. When $U = X$, if $I\,(X; M\,) = 0$, then we say Calvin does not get any information about $X$. $M$ is what Calvin gets by wiretapping. In this case, we call it information secure.

**Weakly secure:** $X_i$ stands for individual source information. If $I\,(X_i\,; M\,) = 0,\ \forall X_i \in U$, we say $M$ does not leak out any meaningful information about $U$. If $I\,(X_i\,; M\,) = 0,\ \forall X_i \in X$ , we say Calvin does not get any meaningful information about $X$. We call it weakly secure.

### 2.4 Weak security of Random network coding

Paper [16] explicitly analyzed the weak security of random network coding.

**Theorem 1.** If the number of link that Calvin can eavesdrop on is less than maximum network flow, ($k = max_{Ai \in \Lambda}\ rank$ $(A) < m$), then the source and destination code/decode algorithm above can meet the requirement of weak security and the code complexity is ($m\ 2\ n$).

**Theorem 2.** For a given network, when $k = max_{Ai \in \Lambda} rank (A) < m$ and random code is applied in intermediate nodes, the probability of Calvin obtaining meaningful information by wiretapping is less than $\frac{|\Lambda|^2 k^2 m}{q^{2(m-k)}}$.

## 3. Security Analysis of Random linear network coding

### 3.1 Information transmission system of random linear network coding

We use Matlab to implement a random linear network coding information transmission system. This system environment can be both wireline network and wireless network. There is only one source node and one destination node in this system and a variable number of intermediate nodes. Source node sends data packet processed by random linear network coding and destination node decodes according to encoding principles. Wiretapper exists in system and wiretaps on intermediate nodes. It can be classified into single-node wiretapping and multi-node wiretapping. Error can be taken into account or not. This system can verify how nodes' decoding process and wiretapper's success probability is affected by is affected by variables in network.

The system is consisting of network topology generating, random linear network encoding and decoding, data buffer and wiretapping module. We designed suitable wiretapping method for simulation, which are single-node no-error wiretapping and multi-node co-wiretapping. In multi-node co-wiretapping, error can be taken into account or not.

### 3.1.1 Wireless network topology

Design a function $NetTop = graph\ (NodeNum, scale)$.

The experiment is designed and simulated in a scale $X$ scale planar domain. All simulation nodes (including source node and destination node) are in this domain. Because in wireless network, nodes' location will change with time, thus we use random number generation function to freely generate NodeNum nodes in the domain.

### 3.1.2 Random linear network encoding and decoding

Each time source node sends DataNum data packets, it randomly chooses network coding coefficients in Galois field and codes. Design function $NetCode = RLNC\ (x, num, m, type)$. Coding principle of intermediate nodes is the same as source nodes'. Intermediate node will perform random linear network coding after receiving new information and send it out.

Decoding of random linear network coding uses Gaussian Elimination. As long as *DataNum* individual data packets are received, then initial *DataNum* data packet can be recovered.

Receiver node has to determine whether the new data packet it has received contains new information. New information is helpful to decoding. If the new information and old information are linearly dependent, then the new information does not possess any useful information. Design $function\ rrefcode = GJrref\ (codevec, len)$ and $function\ a = reform\ (rrefcode, m)$ such that we determine whether there is new useful information through checking whether new row emerges after matrix transform. Or discard the new data packet. When destination node receives $DataNum$ data packets that are linearly dependent, then it can decode the initial data packet and source node stops sending the $DataNum$ data packets.

### 3.1.3 Receiveing buffer and sending buffer designing

Receiving buffer creates an array $Buffer[i]$ in matlab format for each node. For intermediate node, when receiving new useful data packet, it will store the packet in $Buffer[i]$. Each node has to store the data packet in a buffer before sending it out, because matlab simulation cannot implement real-time data transmission. Therefore, node has to look up the buffer containing data sent by each node in last moment when receiving data. We designed a array $NCcode$ in matlab format. Each node will perform a random linear network coding for data in buffer and new data packet. In this way, the size of data packet sent each time is one code word of network coding. $NCode$ records the data packet sent by nodes but the destination node. Note that for intermediate nodes, only if they receive new useful data packet, will they perform network coding and forward packets. Thus, a number flag $NCflag$ is needed for mark the moment whether some node has sent a data packet.

### 3.1.4 Wiretapping method designing

In experiments, we assume wireless channel is no error channel. Bit error and data packet loss are beyond the scope of this text. As long as the distance between two nodes is less than transmission distance ($radius$), these two node can receive data packets from each other.

**Multi-node co-wiretapping method:** Each wiretapper co-store what has been wiretapped, increasing the probability of obtaining enough number of data packet ($DataNum$ linearly dependent data packets) and the probability of wiretapping initial data. In real wireless network, each intermediate node is no different from others, hence witetapper cannot find the node combination way that can be optimized. If the information obtained by multiple wiretapper is reorganized before decoding, then the processing time will be increased.

### 3.2 Security of multi-node co-wiretapping in wireless network

(1) When Galois field $m = 8$, the number of data source node sends each time: $DataNum = 4$, transmission radius of each node: $radius = 3 \times sqrt\ (scale) = 3 \times 10 = 30$, we test the relationship between decoding success probability and total node number *NodeNum* when three nodes are co-wiretapping.

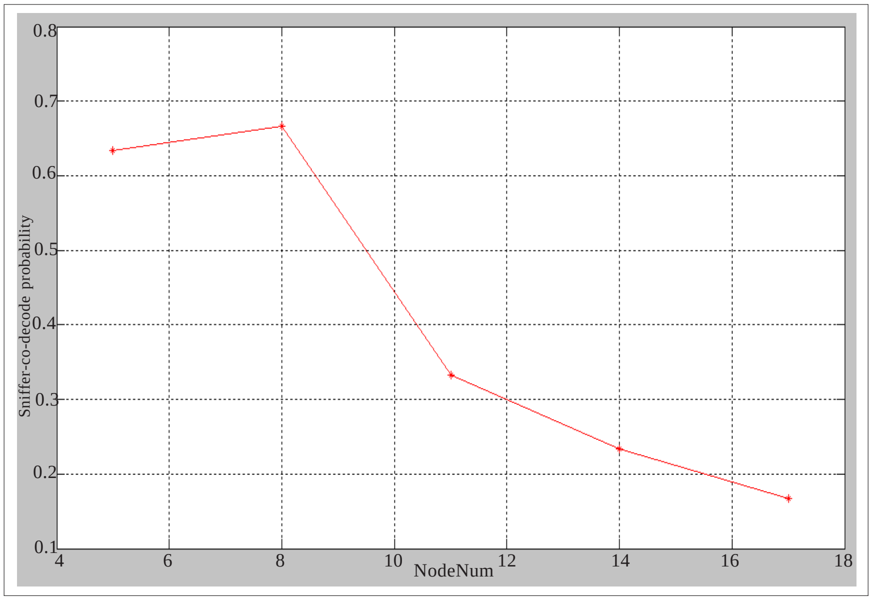

Figure 1. The relationship between decoding success probability and total node number NodeNum

| NodeNum | 5 | 8 | 11 | 14 | 17 |
|---|---|---|---|---|---|
| Decoding Probability | 0.633 | 0.667 | 0.333 | 0.233 | 0.167 |

Table 1. The relationship between decoding success probability and total node number NodeNum

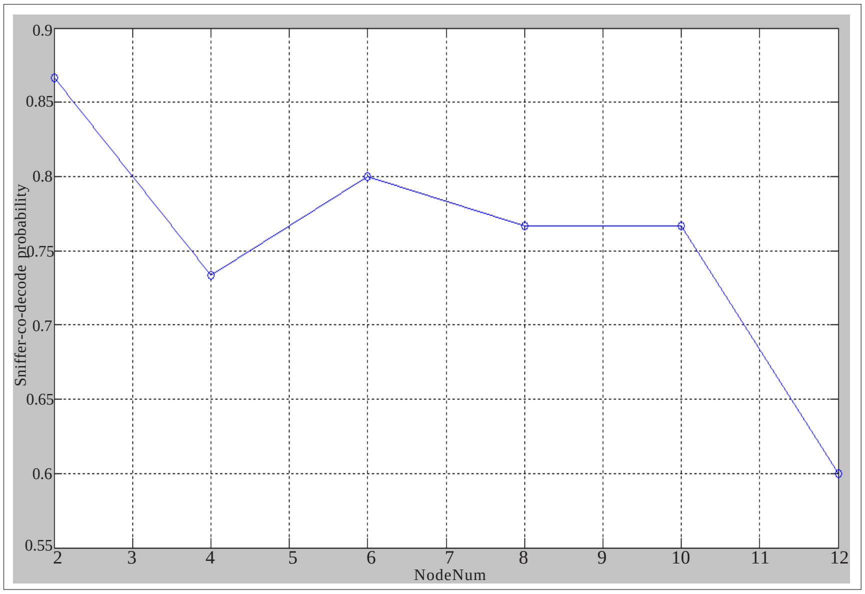

Figure 2. The relationship between decoding success probability and number DataNum of data sent by source node one time

| DataNum | 2 | 4 | 6 | 8 | 10 | 12 |
|---|---|---|---|---|---|---|
| Decoding probability | 0.867 | 0.733 | 0.800 | 0.767 | 0.767 | 0.600 |

Table 2. The relationship between decoding success probability and number DataNum of data sent by source node one time

Set minimum of $NodeNum$ is $5$ and wiretappers wiretap on node 2, 3 and 4 (There is no difference between nodes in wireless network). Figure 1 of the relationship between decoding success probability and total node number $NodeNum$ when three nodes are co-wiretapping. Table 1 of the relationship between decoding success probability and total node number $NodeNum$ when three nodes are co-wiretapping.

When $NodeNum = 5$, meaning wiretapper has been wiretapping all intermediate nodes in network, resulting in a higher decoding success probability of wiretapping. However, as intermediate nodes are increasing and wiretapping ratio decreasing and data loss, wiretapper's success probability will be reduced.

(2) When Galois field $m = 8$, total node number in network: $NodeNum = 8$, transmission radius of each node: $radius = 3 \times sqrt\,(scale) = 3 \times 10 = 30$, we test the relationship between decoding success probability and number $DataNum$ of data sent by source node each time when three nodes are co-wiretapping. Wiretappers wiretap on node 2, 4 and 6. Figure 2 of the relationship between decoding success probability and number $DataNum$ of data sent by source node one time when three nodes are co-wiretapping. Table 2 of the relationship between decoding success probability and number DataNum of data sent by source node one time when three nodes are co-wiretapping.

From experiment results, the decoding success probability of three-node co-wiretapping reduces as the number DataNum of data source node transmits each time increases. Because DataNum has increased, wiretapper needs more linealy independent data packet to successfully decode, resulting in reducing probability.

(3) When the number $DataNum$ of data source node transmits each time is 5, total node number $NodeNum = 8$ and transmission radius of each node: $radius = 3 \times sqrt\,(scale) = 3 \times 10 = 30$, we test we test the relationship between decoding success probability and Galois field m when three nodes are co-wiretapping. Wiretappers wiretap on node 2, 4 and 6. Figure 3 of the relationship between decoding success probability and Galois field m when three nodes are co-wiretapping. Table 3 of the relationship between decoding success probability and Galois field $m$ when three nodes are co-wiretapping.

(4) When the number DataNum of data source node transmits each time is 5 total node number NodeNum=8 and Galois field $m = 8$, we test the relationship between decoding success probability and transmission radius of each node when three nodes are co-wiretapping. Wiretappers wiretap on node 2, 4 and 6. Figure 4 of the relationship between decoding success probability and transmission radius of each node when three nodes are co-wiretapping. Table 4 of relationship between decoding success probability and transmission radius of each node when three nodes are co-wiretapping.

## 4. Conclusion

We have implemented random linear network coding in wireless network with matlab simulation and designed response measures for each link according to matlab features, including network topology designing, random

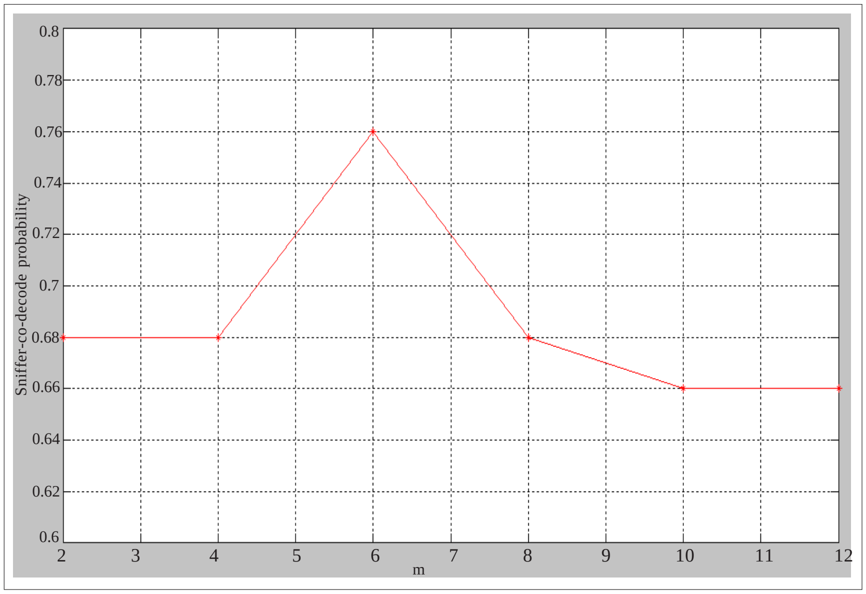

Figure 3. The relationship between decoding success probability and Galois field m

| m | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|
| Decoding probability | 0.680 | 0.680 | 0.760 | 0.680 | 0.660 | 0.660 |

Table 3. The relationship between decoding success probability and Galois field m

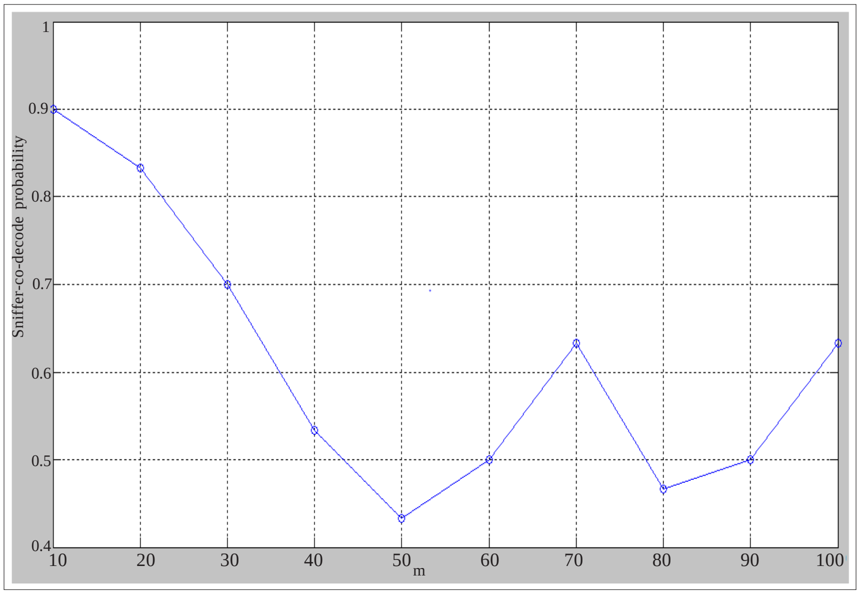

Figure 4. The relationship between decoding success probability and transmission radius of each node

| Radius | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 90 | 100 |
|---|---|---|---|---|---|---|---|---|---|---|
| Decoding probability | 0.900 | 0.833 | 0.700 | 0.533 | 0.433 | 0.500 | 0.633 | 0.467 | 0.500 | 0.633 |

Table 4. The relationship between decoding success probability and transmission radius of each node

linear network coding and decoding, buffer problem of receiving and sending.

We also designed suitable wiretapping method for our simulation and in multi-node co-wiretapping, error can be taken into account or not. Experiment showed that, in wireless network, decreasing total node number *NodeNum*, and number *DataNum* of data source node transmits each time will increase the probability of wiretapper's successfully decoding when multi-node are co-wiretapping. Therefore, this paper has provided the security relationship between variables and random linear network coding, we can improve the security of random linear network coding by adjusting parameters.

## References

[1] Cai, N., Yeung, R. W. (2002). Secure network coding, *In*: IEEE International Symposium on Information Theory, Lausanne, Switzerland, 7, p. 323.

[2] Cai, N., Yeung R. W. (2007). A Security Conditional for Multi-Source Linear Network coding, *In*: IEEE International Symposium on Information Theory, 7, p. 561-565.

[3] Jain, K. (2004). Security based on network topology against the wiretapping attack, *IEEE Wireless Communications*, 11 (1) 68-71.

[4] Bhattad Kapil, Narayanan Krishna, R. (2005). Weakly Secure Network coding, *In*: Proc. First Workshop on Network coding Theory, and Applications (Netcod), April. (In Italy)

[5] Tan Jianlong, Medard Muriel. (2006). Secure Network coding with a Cost Criterion, *In*: International Symposium on IEEE Modeling and Optimization in Mobile, Ad Hoc and Wireless Network, 4, p. 1-6.

[6] Cai, N., Yeung, R. W. (2002). Network coding and Error Correcting, *IEEE Inform Theory Workshop*. Bangalore: IEEE Press, p. 119-122.

[7] Cai, N., Yeung, R. W. (2002). Secure Network coding, *IEEE Intl Symp Inf Theory*. Lausanne: IEEE Press, p. 323.

[8] Jaggi, S., Langberg, M., Katti, S. et al. (2008). Resilient network coding in the presence of Byzantine adversaries *J. IEEE Trans Inform Theory*, 54 (6) 2596-2603.

[9] Chan Terence, Grant, A. (2008). Capacity Bounds for Secure Network coding, *IEEE Communications Theory Workshop*, AusCTW 2008, p. 95-100.

[10] Silva, D., Kschischang, F. R. (2008). Security for Wiretap Networks Via Rank-Metric Codes, *IEEE Intl Symp Inf Theory*.Toronto: IEEE Press, p. 176-180.

[11] Silva, D., Kschischang, F. R. (2009). On metrics for error correcting in network coding *J. IEEE Trans Inform Theory*, 55 (2) 5479-5490.

[12] Zhang, P., Jiang, Y., Lin, C. et al. (2010). P-codes:Secure network coding against wiretapping attacks. C.San Diego, CA, USA: *Proc IEEE INFOCOM*, p. 1-9.

[13] Wu Xi-kun, XIA Shu-tao. (2012). Hannon secure error-correcting network coding against global wiretapper. *Computer Engineering and Design*, 33 (4) 1261-1265.

[14] Ho, T., Medard, M., Koetter, R., Effros, M., Shi, J., Karger, D. R. (2006). A random linear codes approach to multicast, *IEEE Transactions on Information Theory*, 52, p. 4413-4430.

[15] Ramamoorthy, A., Shi, J., Wesel, R. (2005). On the capacity of network coding for randomnetworks. *IEEE Transactions on Information Theory*, 51 (8) 2878-2885.

[16] Ye-jun, Zhou., Hui, LI., MA Jian-fen. (2009). Random network coding against the wiretapping adversaries. *Journal of Xidian0University*, 36 (4) 696-701.