Applied Mathematics & Information Sciences
An International Journal

http://dx.doi.org/10.12785/amis/072L33

Low-overhead Secure Network Coding based on Chaotic Sequence

Guangxian Xu[1], *Xiao Fu*[2] *and Wei Wu*[3]

[1]School of Electronic and Information Engineering, Liaoning Technical University, HuluDao 125105 China
[2]School of Electronic and Information Engineering, Liaoning Technical University, HuluDao 125105 China
[3]School of Electronic and Information Engineering, Liaoning Technical University, HuluDao 125105 China

Received: 22 Oct. 2012, Revised: 6 Feb. 2013, Accepted: 8 Feb. 2013
Published online: 1 Jun. 2013

Abstract: A minimum overhead secure network coding based on chaotic sequence is presented in this paper. And only the source need to be modified, intermediate nodes implement a classical distributed network code. So the proposed scheme is applied to all the linear network coding. Combine the chaotic sequence with original source information vector, because of the high randomness and the sensitivity to initial state of chaotic sequence, the presented network code is ”One-Time Pad (OTP)”. So the secure network coding achieves complete secrecy. This scheme requires only one noisy symbol to be embedded in the original information symbol vector to achieve complete secrecy. The theoretical analysis confirm that this scheme can achieve the information-theoretic security condition and the signalling overhead to obtain complete security is minimized, while the attacker has limited wiretapping ability.

Keywords: Network coding; chaotic sequence; OTP; applicability; complete secrecy; minimum overhead

1 Introduction

In 2000, based on the concept of network information flow, network coding was first studied by Ahlswede et al[1]. It allows intermediate nodes to encode the information received from their input links, so it can make network flow to maximize. Network coding is mainly divided into two categories. One is linear network coding and the other is non-linear network coding. As linear network coding is practical and has a simple structure, so we discuss linear network coding in this research.

The initial purpose of using network coding is to achieve maximum network flow and improve network throughput. And the deeper researches show that network coding is also a good way of secure network transmission [1,2]. In order to solve the security problems and improve the robustness and the transmission efficiency, in 2002, Cai et al[3] discussed the application of network coding in security aspects first. They proposed the eavesdropping communication network model and constructed the information-theoretic security network coding. Then there are many fruitful researches on the basis of the Cai. J. Feldman et al[4,5]discussed the alphabet and overhead of network coding. In order to ensure that the source information is secure, Zhang et al[6] combine the original source information with a same length random vector, then transmit the mixed message to network. In the above secure schemes, they all suppose the attacker can wiretap the information in any link, and the attacker's eavesdrop ability is limited. The attackers can eavesdrop k channels at most. To establish a secure network transmission scheme by using of network coding, k independent random vectors need to be chosen from a same field and be coded with k source information vectors. So lot of bandwidth must be used to transmitting the k independent random vectors. So there is a lot of overhead. The secure network coding protects the data effectively, but, at the same time, it has large overhead and high node complexity.

To improve the robustness of the network coding and reduce overhead, after analysis the characteristics of random network and mobile network, M.R. et al[7] proposed an algorithm that combine secure network coding and channel coding, at the same time, the training sequence was embedded in the source data. But the channel coding was applied to the entire data packet, so the coding complexity is high. Xu Guangxian et al[8] proposed a low-complexity secure network coding. The

[*] Corresponding author e-mail: fx09291108@163.com

source information was transformed with sparse matrix to reduce the coding complexity. But the scheme still need to add amount of redundancy into the source information to make the network secure, so the overhead is still large. Adeli et al[9] proposed a minimum overhead secure network coding based on hash function. Although he has already discussed the safety of the algorithm and added additional conditions, but it still has secure vulnerabilities (the vulnerabilities were detailed in [10]). Jain et al[11] also supposed the attacker has limited wiretap ability, and he got the necessary and sufficient conditions for single-source network. At the network coding nodes, by using of hash function, he made the network coding more secure. As add function in network coding process, so the coding process is complicated. So this scheme is suitable for the small network that has less nodes and easy to know the entire topology. Bhattad et al[12]proposed a weakly secure network coding that has simple system. In a given linear network, only the source and sink need to do linear transformation, the intermediate nodes were unchanged. So the weakly secure network coding has a low complexity and low overhead. However, the weakly secure network coding scheme did not reach the information theoretic secure condition. Zhao Hui and Zhuo Xinjian [13] proposed a security theorem, and proved that it can reach a better security by using pseudo-random function in network coding process. But this scheme needs to know the entire network topology, so it is not suitable for large scale network.

Secure network coding against wiretapping has been extensively studied. The researches show that coding source message to reach the security condition is a very effective method. This method not only can reduce the complexity of network coding processing, but also can make the overhead lower. It can choose symbols from the same field when choose the random vector added in source message or do linear transform to source message. The field is different from the network coding finite field, so the method has a general applicability.

In this paper, we suppose one network system that only wiretapping attack exist in it, then discuss and select appropriate chaotic map, and use it as a pseudo-random sequence generator. Then mix the generated pseudo-random sequence with plaintext. Chaotic sequence has its unique characteristics. It is useful to encrypting files. The theoretical analysis and simulations both confirm that the algorithm can improve network security effectively and reduce the overhead of the network coding.

2 Network model

To simplify the analysis of secure network coding, we consider an acyclic and delay-free communication network represented by a directed multigraph with unit capacity edges, a single source node, and multiple destination nodes. The pair $G=(V,E)$ is called a directed multigraph, where V and E are the nodes set and the edge set of G, respectively. For constructing network coding, let $GF(q)$ be a finite field (q is a big prime).

Definition A linear coding of dimension n based on $G=(V,E)$ assign a vector $u(e)$ called the ”global encoding kernel” of channel e to every edge e in $G(e \in E)$. And we call it network coding if the following conditions are satisfied.

(1) The assigned vector space of the information source X is $z(X)=GF_n(q)$. That is, $z(X)$ is a vector space of dimension n.

(2) The vector assigned to edge $e(e \in E)$ should be a linear combination of vectors assigned to the incoming channels of $tail(e)$ ($tail(e)$ is the tail node of edge e). The vector of e should choose from $z(X)$ if the tail node of e is information source X.

(3) In every information sink, there must be a function operation that can recover the information from source node.

We use $z(S)$ to stand for the vector space of node S. And as we know, to every intermediate node S, $z(S)$ is the set of every possible linear combination of vectors assigned to the incoming channels of node S. So it's easy for us to know that $u(e) \in z(tail(e))$ for any $e \in E$.

Source A: A produce non-compressed data message, denoted by m. These messages form a vector of length n, denoted by $m=(m_1,m_2,\cdots,m_n)^T$. By using of the encoding algorithm, A encodes the messages and sends the encoded data to sink nodes(the encoding algorithm will be introduced in next section).

Receiver T: There are t sink nodes in the network. The secure network coding ensure the source message be sent to sink nodes safely.

Adversary C: C has limited eavesdropping capabilities, he can eavesdrop k channels. The set of channels be eavesdropped denote by $\Omega=\{\alpha_1,\alpha_2,\cdots,\alpha_k\}$.

Construct the linear network coding according to the above definitions and network model. Each channel has a encoding vector v_i . And i is the number of links in the network, that is, i from 1 to $N=\parallel E \parallel$. Then

$$v_i=(v_{i1},v_{i2},\cdots,v_{in})^T, \tag{1}$$

All these encoding vectors are selected from the same coding domain. The data transmitted in channels denoted by x_i, x_i is the linear combination of all messages, and then we have

$$x_i=v_i^T m, i=1,2,\cdots,N, \tag{2}$$

The existing secure network coding schemes mainly add random number in the encoding vectors and then do linear combination with messages. This make the secure network coding have a large encoding overhead. Furthermore, the adversary can destroy security of the encoded message by obtain data in more independent links. The basic approach of the minimum overhead secure network coding is using chaotic sequence and network coding to construct one-time pad encryption system.

3 Choose the chaotic sequence

Our encoding scheme is based on using chaotic sequence. Chaotic sequence has the following characteristics.

(1) A chaotic map can output fixed sequence with an arbitrary input. The generation of chaotic sequence can be controlled by user, that is, we can have a fixed length sequence.

(2) Given a chaotic map and an input, the chaotic sequence can be calculated. But given an output sequence, it can not find the corresponding input. The chaotic sequence is pseudorandom, so it is impossible that two data in a certain length is same.

(3) Moreover, the main reason of using chaotic map is that it can generate different random signal with only one redundancy β [14]. Its security is equivalent to using $n-1$ independent vectors encode $n-1$ messages in m.

There are three main chaotic maps, logistic chaotic map, the improved logistic chaotic map and Chebyshev chaotic map. Their expressions are shown in Table 1.

Table 1 Several Chaotic Maps.

Chaotic maps	expressions
Logistic chaotic map	$x_{n+1} = \lambda \cdot x_n(1-x_n)$, $x_n \in (0,1)$
Improved logistic chaotic map	$x_{n+1} = 1-2\cdot x_n^2$, $-1 < x_n < 1$
Chebyshev chaotic map	$x_{n+1} = \cos(g \cdot \arccos x_n)$, $-1 < x_n < 1$

λ is a parameter in logistic chaotic map, when $\lambda \in (3.5699, 4]$ the system is in a chaotic state. g is also a parameter in Chebyshev chaotic map, when $g > 2$ the system is in a chaotic state.

To construct the minimum overhead secure network coding and avoid increasing the parameters needed to transfer in the network, we choose the improved logistic chaotic map to generate chaotic sequence. Figure 1 shows the sensitivity simulation of the improved logistic chaotic map to the initial input.

In figure 1, blue line stands for the chaotic sequence with original input 0.1234. And red line stands for the chaotic sequence with original input 0.1235. The sequence length is 1024. The statistics show that the 0/1 change rate of improved logistic chaotic sequence is 50.68%. So the chaotic sequence generated by improved logistic chaotic map has a strong sensitivity to the initial input and a great difficulty to decipher. It is really suitable for one-time pad stream cipher.

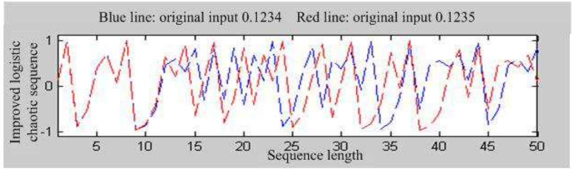

Fig. 1 Chaotic sequence generated by Improved Logistic mapping.

4 Information source encoding algorithm

In the single - source and multi - sink multicast network $G = (V,E)$. The source want to send $n-1$ characters to sink, the characters denoted by $x_1, x_2, \ldots, x_{n-1} \in F_q$. $Y(\cdot)$ stands for the chaotic sequence $y_1, y_2, \ldots, y_n$. And $y_n = 1-2y_{n-1}^2, (-1 < y_n < 1)$. Using the chaotic sequence $Y(\cdot)$ generated by the improved logistic chaotic map and a random number β to form a message vector denoted by following formula.

$$\overline{m} = (x_1 + Y(\beta), x_2 + Y(x_1, \beta), x_3 + Y(x_1, x_2, \beta), \\ \cdots, x_{n-1} + Y(x_1, x_2, \cdots, x_{n-2}, \beta), \beta)^T, \quad (3)$$

Where the comma-separation of independent variables in $Y(\cdot)$ denotes concatenation, which means the individual components are put next to each other in a bit-wise manner and then applied to the improved logistic chaotic map as a single input. $\overline{m}$ can also be written as

$$\overline{m} = (e_1, e_2, \cdots, e_{n-1}, \beta)^T, \quad (4)$$

And $e_1, e_2, \ldots, e_{n-1}$ is uniform distribution, $q > \max\{t, |\Omega|\}$.

Note that the initial input of the chaotic map in formula (3) is different to each other (similar to one-time-pad stream ciphers). So, if the random number β is not obtained, the adversary can not obtain any information of $x_1, x_2, \ldots, x_{n-1}$ in polynomial time even though he gets the first $n-1$ data in $\overline{m}$. This minimum overhead secure network coding scheme only use one random signal to achieve the purpose of hiding information. The following encoding process is sending the linear combination of $\overline{m}$ to each sink nodes (that is linear network coding).

5 Information sink decoding algorithm

The improved logistic chaotic map in this scheme is public to every part including the adversary. The feasibility of recovering $\overline{m}$ is depends on the complexity of the standard linear network code. And since the sink nodes know the improved logistic chaotic map, they get all original source message after recovering $\overline{m}$ and having β which is the last data of $\overline{m}$.

6 One-time pad encryption system security

Theorem The one-time pad cipher scheme is perfect secrecy.

Proof u stands for the plaintext message sent by source, and $u = (u_1, u_2, \ldots, u_n)$.

d stands for the ciphertext message received by sink, and $d = (d_1, d_2, \ldots, d_n)$.

P_u is the probability that the plaintext sent by source is u.

r_d is the probability that the ciphertext received by sink is d.

q_k stands for the probability of the key that used by the source to encrypt plaintext.

So the perfect secrecy can be denoted as

$$Pr(u|d) = P_u, \tag{5}$$

As there is only one key, and it is composed of n independent random variables, so we have

$$q_k = \frac{1}{2^n}, \tag{6}$$

Then

$$r_d = \sum_{u \in P} \frac{P_u}{2^n}, \tag{7}$$

As the key selection is independent of plaintext, so we have

$$Pr\,(u \cap d) = \frac{P_u}{2^n}, \tag{8}$$

Then

$$Pr\,(u\,|d\,) = \frac{P_r\,(u \cap d)}{r_d} = \frac{\frac{P_u}{2^n}}{\frac{1}{2^n}} = P_u, \tag{9}$$

The formula (5) is proved. So the one-time pad cipher scheme is perfect secrecy.

7 Feasibility and Information theory security of coding algorithm

Theorem In the single - source and multi - sink multicast network $G = (V, E)$ and if $q > t$. Then we can construct a suitable linear network coding in polynomial time [15].

Proof h stands for the network capacity. When constructing linear network coding, finding a route from source to sink need time $O(|E|)$. Distributing coding vectors to each edges need time $O(|T|)$. And testing linear independence of the distributed coding vectors need time $O(h \cdot |T|)$. Then t sinks need time $O(|T| \cdot h^2)$ [15].

So selecting the suitable coding vector from the coding domain to construct the linear network coding can be completed in $O(|E| \cdot |T| \cdot h^2)$. The concrete process of proof is similar to [15]. The theorems and conclusions in [15] are not repeated here.

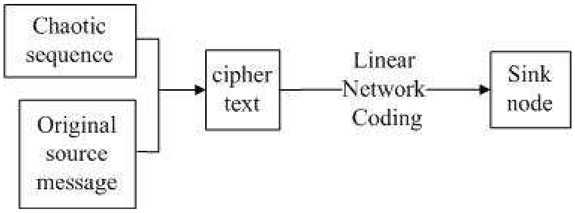

Fig. 2 Flow of the secure network coding.

Theorem In the single - source and multi - sink multicast network $G = (V, E)$. $\Omega = \{\alpha_1, \alpha_2, \ldots, \alpha_\omega\}$ stands for the set of channels be wiretapped. By using of the above minimum overhead secure network coding scheme based on chaotic sequence, and $e_1, e_2, \ldots, e_{n-1}$ is uniform distribution in F_{q^n}, $q > \max\{t, |\Omega|\}$, the secure communication multicast network is implemented.

Proof The paper [10] shows that the information adversary obtained from wiretapped channels is linearly independent vectors when $e_1, e_2, \ldots, e_{n-1}$ is uniform distribution. So the adversary can not get any information of β , and he can not get the chaotic sequence. So the encoding scheme is one-time pad cipher scheme. And we have proved that the one-time pad cipher scheme is perfect secrecy[16]. So the minimum overhead secure network coding scheme based on chaotic sequence is perfect secrecy. The second theorem shows that the suitable linear network coding can be completed in polynomial time.

So the minimum overhead secure network coding scheme based on chaotic sequence is realizable.

8 The general applicability of encoding algorithm

The low-overhead secure network coding proposed in this paper process the source information of a given linear network code. Then get the ciphertext and transport it as a new source message of a network coding system[17]. The sink nodes can recover the original source information by corresponding decoding. The process is as follows.

The flow shows that, in the entire process, the intermediate node of linear network coding is unchanged. And the network coding run on the ciphertext. So although the original input of chaotic sequence is chosen on real number field, it has no effect on the finite field of network coding. That is, the finite field of secure network coding is unchanged. So the low-overhead secure network coding based on chaotic sequence has general applicability[18].

9 Conclusion

The secure network coding schemes proposed in this paper only need to modify the source and destination nodes, and the intermediate nodes implement a classical distributed network code. The low-overhead secure network coding based on chaotic sequence adds one noisy symbol in the original source messages. Its overhead is minimized. It needs not to know the network topology. The scheme can achieve the network max-flow and the information-theoretic security condition. It can apply to all linear networks.

Acknowledgement

This paper was supported by grant number 2009S051 of Liaoning Province Key Laboratory Item; its title is the establishment and its application of a new coding method for PN code. This paper was also supported by grant number LJQ2012029 of Liaoning Colleges and Universities Outstanding Youth Training Plan; its title is the secure network coding against strong attacker based on list-decoding.

References

[1] Ahlswede R., Ning Cai, Li S.Y.R., et al. Network Information Flow. IEEE Transactions on Information Theory, **46(4)**, 1204-l2l6(2000).

[2] Yeung, R.W. Distributed source coding for satellite communications. IEEE Transactions on Information Theory, **45(4)**, 1111-1120(1999).

[3] N.Cai, R.Yeung. Secure network coding. IEEE International Symposium Information Theory, (2002).

[4] Jon Feldman, Tal Malkin, Rocco A. Servedio. On the capacity of secure network coding. 42nd Annual Allerton Conf. Commun, (2004).

[5] Jon Feldman, Tal Malkin, Rocco A. Servedio. Secure network coding via filtered secret sharing. 42nd Annual Allerton Conf. Commun, (2004).

[6] Yan Zhang, Cheng-qi Xu, Feng Wang. A novel scheme for secure network coding using one-time pad. International Conference on Networks Security, Wireless Communications and Trusted Computing. (2009).

[7] R.M., Sagduyu Y.E., Honig M.L., et al. Training overhead for decoding random linear network codes in wireless networks. IEEE Journal on Selected Areas in Communications. **27(5)**, 729-737(2009).

[8] Guang-xian Xu, Xiao Fu. An Improved Secure Network Coding Algorithm based on Sparse Matrix. Journal of Information and Computational Science. **8(14)**, 3091-3098(2011).

[9] Majid Adeli, Huaping Liu. Secure Network Coding with Minimum Overhead Based on Hash Functions. IEEE Communications Letters, **13(12)**, 956-958(2009).

[10] Li Kerong. A secure communication scheme based on network coding and Hash function. Yangzhou: Yangzhou University, (2010).

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

[12] Bhattad K,Narayanan K R. Weakly Secure Network Coding .[2007-05-22].http://netcod.org/papers/06Bhattad N-final.pdf.

[13] Zhao Hui, Zhuo Xinjian, Lu Chuanfen. Analysis on an algorithm of secure network coding based on the network topology . New progress of communication theory and technology-Thirteenth National Youth Communication Conference Proceedings. (2008).

[14] Li Feifei. The study of chaotic spread spectrum sequence. HuLudao: Liao Ning Technical University, (2011).

[15] Jaggi S., Sanders P., Chou P.A., Effros M., Egner S., Jain K. Polynomial time algorithms for multicast network code construction. IEEE Transactions on Information Theory, **51(6)**, 1973-1982 (2005).

[16] Jiyoung Woo, Hwa Jae Choi, Huy Kang Kim. An automatic and proactive identity theft detection model in MMORPGs. Applied Mathematics & Information Sciences. **1**, 291S-302S(2012)

[17] Wu Xiang-hu, Qu Ming-cheng, Liu Zhi-qiang, et. al. A code automatic generation algorithm based on structured flowchart. Applied Mathematics & Information Sciences. **1**, 1S-8S(2012)

[18] Ning Zhang. Research on control routing technology in communication network. Applied Mathematics & Information Sciences. **1**, 129S-133S(2012)

Guangxian Xu received the MS degree in Liaoning technical university in 2003, and the PhD degree from Liaoning technical university in 2008. He is currently a associate professor in Liaoning technical university. His research interests are in the areas of information and coding, information systems, and network coding.

Xiao Fu is currently a postgraduate student in School of Electronic and Information Engineering, Liaoning technical university. Her research interests are in the areas of network coding, information security, and communication systems.

Wei Wu is currently a postgraduate student in School of Electronic and Information Engineering, Liaoning technical university. Her research interests are in the areas of network coding, information security, and communication systems.