An Efficient Lightweight Security Algorithm for Random Linear Network Coding

Hassan Noura, Steven Martin and Khaldoun Al Agha
Laboratoire de Recherche en Informatique, Université Paris-Sud- CNRS, Orsay, France

Keywords: Secure Network Coding, Data Confidentiality, Key Dependent, Flexible and Invertible Diffusion Matrix.

Abstract: Recently, several encryption schemes have been presented to Random Linear Network Coding (RLNC). The recent proposed lightweight security system for Network Coding is based upon protecting the Global Encoding Vectors (GEV) and using other vector to ensure the encoding process of RLNC at intermediate nodes. However, the current lightweight security scheme, presents several practical challenges to be deployed in real applications. Furthermore, achieving a high security level results in high computational complexity and adds some communication overhead.

In this paper, a new scheme is proposed to overcome the drawbacks of the lightweight security scheme and that can be used for RLNC real-time data exchange. First, the cryptographic primitive (AES in CTR mode) is replaced by another approach that is based on the utilization of a new flexible key-dependent invertible matrix (dynamic diffusion layer). Then, we show that this approach reduces the size of communication overhead of GEV from $2 \times h$ to h elements. In addition to that, we also demonstrate that besides the information confidentially, both the packet integrity and the source authentication are attained with minimum computational complexity and memory overhead.

Indeed, cryptographic strength of this scheme shows that the proposed scheme has sufficient security strength and good performance characteristics to ensure an efficient and simple implementation thus, facilitating the integration of this system in many applications that consider security as a principal requirement.

1 INTRODUCTION

With the evolution of the Internet and the addiction of humans to the new features provided by the developers, bandwidth becomes scarcer and not covering everyone's needs. In addition, this evolution has introduced a combination of several traffic types that can be transmitted over the internet, such as real time traffic (video, voice, etc.) and data traffic (files transfer, web browsing, messages, etc.). To respond to the need of the users in having a mixture of different types of traffic at the same time, with the guarantee of Quality of Service (QoS), especially for the applications that need a high level of security, researchers have approached security with a new technique, Network Coding.

In this context, Network Coding (NC) is an in-field principle that extends the concept of the previous traditional routing (ex: store and forward approach) by offering a new design for the packet networks, and allowing intermediate nodes to participate and take an important role by combining several input packets originated from different sources, then forwarding the resulting coded packets to the destinations. NC ensures significant improvements in network performance, especially in lossy networks and also in multicast and multipath scenarios. It achieves a maximum flow of information, shown theoretically in (Koetter et al., 2003) and experimentally in (sang Park et al., 2006). Furthermore, it has been proven that using NC enables to achieve lower energy consumption and ensures reliable communication over the networks.

In order to improve the network throughput, efficiency, scalability as well as resistance against attacks, several NC techniques have been studied and implemented. Random Linear NC denoted by (RLNC) (Ho et al., 2006) is one of the well-known methods of NC, where each transmitted packet is actually composed of independent linear combinations of previously received packets and original packets generated at this node. The coefficients of these linear combinations followed a uniform distribution in a finite field F_q, noting that the same operation is applied to each symbol existing in one packet. Decoding process, at the destination side is done by performing a Gaussian elimination on the set of receiving coded

Noura H., Martin S. and Al Agha K..
An Efficient Lightweight Security Algorithm for Random Linear Network Coding.
DOI: 10.5220/0005002701710177
In *Proceedings of the 11th International Conference on Security and Cryptography* (SECRYPT-2014), pages 171-177
ISBN: 978-989-758-045-1

packets to retrieve the original ones.

RLNC is a distributed NC scheme and it was proposed to overcome the centralized code allocation overhead of Linear NC (LNC) (yen Robert Li et al., 2003). The unreliable multi-hop transmission and willful intermediate packet mixing make the RLNC susceptible to various types of security threats, such as eavesdropping attacks and Byzantine modifications that can prevent an efficient implementation of the RLNC. The former can seriously impair the confidentiality while the latter can damage the authentication of network coded systems. Indeed, active attackers try to change, delete, or modify the packet contents by introducing a malicious code, while on the other hand, passive attackers try to extract the packet content by traffic analysis or monitoring of unprotected communications.

In addition to these attacks, the recent schemes of RLNC are not only able to introduce new attacks, but also they can make existing attacks more damaging and potentially destroying the efficiency of the most traditionally security techniques used in this domain.

These types of attacks affect the confidentiality and the authentication of the transmitted packets (Fathy et al., 2011). These two requirements are considered as the core of security. Therefore, benefits from the features introduced by RLNC cannot be assured in practice without building an efficient and fully secured scheme.

RLNC may be applied in different domains like the banking or the military systems, where transmission of sensitive information is a major concern.

To ensure the basic elements of security: confidentiality, authentication and integrity, several techniques of RLNC have been proposed in the literature (e.g. (Lima et al., 2007), (Zhang et al., 2010)) and authentication (Li et al., 2010)). These techniques are only interested in achieving security without taking into consideration in their implementation the energy consumption, and the computation and communication overhead, which are also considered as important issues that should be studied and analyzed.

As we know, there is always a trade-off between security and complexity. The existing techniques commonly agreed on the designing of network coded that fulfill Shannon security, but with low throughput, while in our paper, we are interested in building a secure scheme with being aware of achieving a good performance level.

In this paper, an efficient and robust authenticated confidentiality scheme is proposed to ensure the necessary security services for RLNC. Our solution relies on combining a Hash Message Authentication Code (HMAC) in a selective manner with a dynamic mixing cipher scheme. Additionally, our confidentiality scheme presents an efficient solution to (Lima et al., 2007), since the second GEV is not transmitted and using a dynamic diffusion layer instead of the AES block cipher that can reduce the communication overhead and computation complexity and consequently the energy consumption.

The security level achieved in our proposed scheme is similar to the Shannon security level, but with low complexity due to the use of secret encoding scheme that reduces the computational complexity and minimizes the amount of secret mixing needed to ensure the confidentiality of *RLNC*. This leads to be considered as suitable for real time (live streaming) applications.

The rest of this paper is organized as follows: In Section 2, we give a general idea about the existing scenario of RLNC, and we focus on the method used by each scenario to achieve the required security level. Then, we highlight the weakness points presented in each technique that prevent it from being utilized as a standard secured scheme. After that, in Section 3, our proposed authenticated-confidential scheme is defined, and the proposed technique used to construct the invertible dynamic matrix in integer fields is explained. The Cryptography strength is shown in Section 4 Finally, a global conclusion about the work is given in section 5.

2 PRELIMINARY

2.1 Overview of RLNC

In this section, several existing techniques concerning RLNC are discussed. This discussion allows examining the implementation of an efficient and secured scheme by taking into account the advantages of existing methods and avoiding as much as possible their vulnerabilities.

First, we start by describing the traditional RLNC in details and then explaining the important role of the set of Global Encoding Vector (GEV) that forms the Global Encoding Matrix (GEM) G to ensure the security services. G is a linear transformation represented by a matrix, and can be considered as a diffusion layer for the cipher. The encoding process of RLNC consists of two steps: the first step is resumed by the generating of GEVs, while the second step is resumed by the formation of modular vector matrix multiplication. If the coefficients are chosen randomly from a large field, then the resulting matrix is invertible with high probability, which explains why this approach is capable of achieving the multicast capacity of a network.

2.2 Related Work

Since confidentiality is a major concern, recent works are very interested to introduce methods in the favor of achieving the confidentiality issue. Using RLNC, there have been several interesting methods for countering wiretapping attacks in networks. The most of these existing schemes deal with the scenario where an adversary can only intercept to a limited number of packets from a subset of links in the network. This means that NC provides an intrinsic security. However, if the adversary intercepts k packets with $k < h$, then the probability that the adversary does not get any useful or meaningful information about the original packets can be defined by (Cai and Yeung, 2002):

$$p(k) = \prod_{i=1}^{k}(1 - hq^{i-h} + hq^{i-h-1}), (q \mapsto +\infty) \qquad (1)$$

In this case, the security characteristic of RLNC is low and the system is unreliable, and has low resistance against passive attacks (Bhattad and Narayanan, 2005). Once sufficient packets are collected ($k = h$), a leakage for all the information occurs and the whole system is broken.

With the aim of ensuring the requirement of Shannon security, and achieving the full protection of information, some existing solutions are well illustrated in Figure 1.

Later, many schemes were proposed using a new kind of encryption called Homomorphic Encryption (HE) (Clarkson, 1994) as in (Lima et al., 2007), (Zhang et al., 2010), (Najeem and Siva Ram Murthy, 2011) to be secure against passive attacks.

HE should assure that the arithmetic operations taking place on cipher-text are reflected on the plaintext. Many homomorphic crypto-systems belong to the asymmetric crypto-system as RSA (Menezes et al., 1996), which is expensive in terms of computational time and their required parameter lengths are large and appear to be not feasible to utilize in practice. Also, the public HE operations require a heavyweight computational at each participating node and it is not scalable. All these limitations lead to consider the public HE as a non-efficient solution.

Many cryptographic schemes take advantage of this characteristic and apply their solution at G to ensure the confidentiality property of RLNC, since G generations is necessary for packets decoding process. In (Fan et al., 2009), researchers studied the potential of HE along with NC to resist against traffic analysis.

In (Lima et al., 2007), they presented a scheme, in which a set of global encoding matrix $G = \{G_1, G_2, \cdots, G_h,\}$ with $G_i, i = 1, 2, \cdots, h$ represents the i^{th} Global Encoding Vector used at the source, is encrypted, while another set of unencrypted ones is attached to maintain the standard coding processes at intermediate nodes.

Obviously, this scheme requires less data to be encrypted, but it actually needs to iterate AES for $\frac{h^2}{16}$ blocks, together with two rounds of encoding/decoding processes, and a considerable space overhead is added by the extra set of Global Encoding Vectors. Thus, making the system not efficient as much as we expected.

Additionally, the same key is used throughout the transmission, and the problem of single generation failure may occur, where an accidental key disclosure in one generation will compromise the secrecy of the following transmission.

The problems of the previously discussed methods allow us to deduce that these methods are inefficient in both computation and space (communication overhead), and they fail in achieving the required security level.

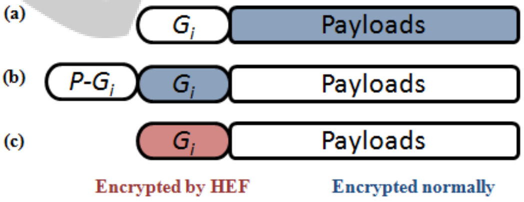


Figure 1: Existing solution against passive attacks.

For that reason, in the next section, we present our solution that aims to ensure the lightweight scheme for RLNC, by eliminating the communication overhead introduced by the secret encoding process, then sending each encoded packet with its corresponding tags rather than using a secret GEV. The first contribution presented in our scheme appears in the design of a new dynamic secret encoding secret encoding process, which requires a low computational complexity and consequently low energy consumption compared to symmetric cipher such as AES.

3 PROPOSED AUTHENTICATION-CONFIDENTIALITY SCHEME FOR RLNC

In this section, we introduce our new efficient Secure Coding Scheme (SCS) for practical RLNC. This new proposed scheme ensures the security of network coding conditions with a low computational complexity. Its main features are resumed by several points: high level of security, no need for any space overhead, and efficiency in computation. In fact, our secure solution of RLNC requires a simple implementation in order to operate with constrained resources devices such as mobile terminals and sensors. The network modeled is considered as a directed graph with one source and multiple destinations, i.e. multicast sessions. We focus on intra-flow NC, where each node mixes packets belonging to the same flow. Without loss of generality, a Key Distribution Center (KDC) is assumed, which is responsible for symmetric Master Key (MK) establishment. Then the source and destinations can get the secret key MK offline. Additionally, our solution used a secret G, that is produced in a dynamic manner and it is kept hidden from the other nodes. This can bring a considerable confusion of the adversary. Also, the process of the secret encoding scheme using dynamic G can be considered as a diffusion layer to the whole system.

In the following, we describe the proposed Authentication-Cipher Scheme (ACS) and Authentication Decipher Scheme (ADS) in details.

3.1 The Proposed Authentication-encryption Scheme

In practical RLNC scenarios, the source may need to transmit a large amount of data M. In this case, the source should first divide M into generations $\{m_1, m_2, \ldots, m_g\}$. We recommend to perform the authentication and encryption simultaneously using different keys in order to achieve a powerful GEV level of security. The different steps of the proposed scheme at the emitter side are described below in details:

3.1.1 Key Generation

This section defines two processes: updating the master key and producing the section and dynamic keys. These processes are described in detail in (Noura et al., 2013). In addition, experimental results indicate clearly the randomness of the generated dynamic key. Therefore, it can be used as a secure Key Derivation Function (KDF). The confidentiality and authentication, dynamic key (KE, KA) are calculated by flipping the even and odd bits of dynamic key Kd.

3.1.2 Construction of the Secret Matrix G

This technique is simple, flexible and it is efficiently implemented in hardware. Additionally, our approache of invertible encoding is defined in the following. Indeed, to construct the invertible secret matrix (integer or binary), a new method based on a special rule of algebra, and provides a key dependent invertible matrix with determinant equal -1 (nonsingular matrix) is used. It is used to form the secret matrix. Then, a sub matrix A is obtained from the produced binary key-stream, with length equal to $= q \times h^2/4$, where each element consists of q bits. However, the invertiblity probability of a $h \times h$ matrix over field q is calculated as follows:

$$\frac{\prod_{k=1}^{h}(q^h - q^{k-1})}{q^{h^2}} = \prod_{k=1}^{h}(1 - q^{k-1-h}) < 1 - \frac{1}{q} \qquad (2)$$

In a real implementation, the invertibility probability decreases as h grows. In fact, our proposed solution can overcome this problem and preserves the invertibility property, even with higher values of h. Further, the matrix form used to construct the invertible dynamic integer secret matrix G, is given below:

$$G = \begin{bmatrix} A & A + I_m \\ A - I_m & A \end{bmatrix} \qquad (3)$$

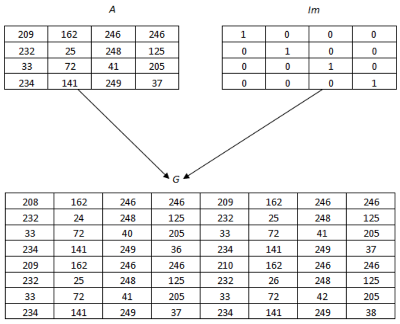
I_m and A are the identity matrix and a non-zero matrix of size $\frac{h}{2}$, respectively. The elements of A can be freely chosen from any Galois field such that G is of full rank. In our simulation, the elements of this sub-matrix are chosen between 0 and 255. Then, the Least Significant Bit (LSB) for each byte is fixed to zeros to preserve the limitation field ($\leq$255). Therefore, the necessary condition to possess an inverse is attained and the receivers can calculate the inverse secret matrix G^{-1} to get the original data. An example to construct the secret matrix G is shown in Figure 2 for $h = 8$.

The calculation of the inverse matrix G in integer field is obtained as follows:

$$G^{-1} = \begin{bmatrix} A & -(A + I_m) \\ -(A - I_m) & A \end{bmatrix} \qquad (4)$$

3.1.3 Encryption: Secret Encoding Packets (Generation) X

The buffering model of RLNC divides the packets stream into generations of size g, such that the



209	162	246	246
232	25	248	125
33	72	41	205
234	141	249	37

1	0	0	0
0	1	0	0
0	0	1	0
0	0	0	1

208	162	246	246	209	162	246	246
232	24	248	125	232	25	248	125
33	72	40	205	33	72	41	205
234	141	249	36	234	141	249	37
209	162	246	246	210	162	246	246
232	25	248	125	232	26	248	125
33	72	41	205	33	72	42	205
234	141	249	37	234	141	249	38

Figure 2: An example of construction secret matrix G for $h = 8$.

packets of the same generation are tagged with a common generation number NG. To complete in Encryption scheme, our proposed scheme performs a secret encoding (mixing) RLNC on a series of source packets $\{p_1, p_2, ..., p_g\}$, an invertible secret dynamic GEM G is constructed as described before, and used for encoding the source packets, then the obtained encoded packets are sent to the intended destinations. So, the encryption process includes two steps:

3.1.4 Construction of Encoded Packets Y

Each mixed (encrypted) payload is concatenated with the header of RLNC (here the tagged u_i is put instead of its corresponding GEV to form, the encrypted coded packet $y_i = [u_i, x_i]$, where u_i is the $i-th$ tag that contains all the elements equal to zero except the $I-th$ column whose values are equal to 1. As seen our scheme work in contrast of (Lima et al., 2007) that combined the source GEV with payload.

3.1.5 Authentication of the h Encrypted Packets

Different steps to obtain the MAC value are well designed. The overall cost of authenticating the stream data is closer to double that of hashing this data, especially when dealing with a huge data size. In order to reduce the complexity, the contents of the h encrypted packets are XORed together to form a unique payload called $(temp)$. $HMAC$ is used with SHA-512 to avoid hash collision. The input of HMAC is composed of the vector $temp$, the extension header $NG||GS$ and Ka which used as an authentication key. Then, the output of $HMAC$ is represented by a matrix with 4 lines and 128 columns. The 4 lines are XORed together to obtain the MAC value MAC with a size of 128 bits.

3.1.6 Asymmetric Encryption of the MAC Value E_{MAC}

H is encrypted using the public-key cipher RSA, which was performed with the private key Kr of the emitter. Then, it is transmitted to the receiver in an encrypted format. Two kinds of keys are used for the encryption and the decryption processes. The use of private key provides the non-repudiation of the source, which is considered as a principal service.

3.1.7 Transmission of $[Y, E\_MAC]$

The transmitted information to the receiver is composed of the cipher E_MAC and the encrypted (secret mixing) packets $Y = \{y_1, y_2, \ldots, y_h\}$. If the opportunity of transmission at an outgoing edge is possible, the sending node first sends the encrypted generation that contains a set of the encrypted packets belonging to the current generation. Our scheme does not introduce any communication overhead per packet, since no extra GEV is added. The cipher E_H is transmitted to the receiver to allow the verification of the data integrity and the authentication of the source at the destination side.

3.2 Intermediate Recoding

Reconstructing the source packets at intermediate nodes seems to be a difficult issue, especially without the knowledge of the secret matrix that is used to mix the symbols of the coded payloads p_i.

In this subsection, we will describe in details the secure scheme at receiver side. The principal step of the proposed ADS is presented below and the details are described as follows:

3.2.1 Selection of Packets Corresponding to Each Generation

The buffering model of the receiver stores the packet stream into generations, according to their NG, such that the packets belonging to the same generation are stored in a single buffer.

3.2.2 Asymmetric Decryption of the MAC value E_MAC

At the receiver end, the recipient uses emitter public RSA key Ku to decrypt the MAC value E_MAC. The RSA algorithm is used to encrypt/decrypt a single 128 bit MAC value MAC. A tiny change in any bit of E_{Kd} leads to a different dynamic key.

3.2.3 Dynamic Key Generation R_Kd

The dynamic key for authentication (R_Ka) and decryption (inverse secret decoding) (R_Ke) is generated using the same approach that was applied at the emitter side.

3.2.4 Intermediate Decoding

To verify the source and recover the original packets, the destination needs to apply the process of intermediate decoding. When h linear independent messages are collected, the destination derives the intermediate encoding matrix between the intermediate node and the encrypted packets. Then, the process of intermediate decoding is done by using the Gaussian elimination to decode the encoded packets at intermediate nodes.

3.3 Verification of Authentication (C_MAC, R_MAC)

The efficiency of our solution appeared in the prevention of the attacker to get the opportunity of decrypting any cipher-generation unless he succeeds to verify correctly the authentication scheme, in other words, unless he gets access to the cipher key. But this is not possible, because our scheme is built with the idea that each time, an h different secret encoded packets are collected from an arbitrary generation, a new MAC is calculated at the receiver side, using the same technique that was applied at the emitter side and it is denoted by C_MAC, which was applied at the emitter side. Then this C_MAC is compared with R_MAC. If the calculated C_MAC is equal to R_MAC, the source is verified. Otherwise, the authentication process is failing.
From that, a conclusion can be provided that our authentication scheme ensures the protection of the source from any unauthorized intervention.

3.4 Decryption: Secret Decoding

Once the source is verified, the destination can start the decryption process. Once h linearly independent messages are collected, the destination produces the secret matrix C_G. The decryption of the encrypted generation R_Y is obtained by using the inverse secret matrix C_G^{-1} and it is done by: $D = C\_G^{-1} \times Z$ for the integer secret encoding process.

4 CRYPTOGRAPHY STRENGTH

A cryptographic scheme is considered secure if its scheme has the immunity against different types of attacks. The cryptographic security of our scheme supports two properties:

- use of dynamic keys.
- unpredictability and high sensitivity of dynamic keys.

The presence of passive attacks as in (Bhattad and Narayanan, 2005) makes the benefits of NC not distinguishable in addition to its weakness benefit in terms of security. Moreover, an important issue that should be mentioned here is the fact that the key is changed in a dynamic manner which ensures the protection of all information. Additionally, our proposed solution is flexible, since it gives the opportunity to control the delay in time, by choosing the adequate h value and the accurate number of encoding packets. Thus, leads to confuse the attacker and create a structure that is protected enough against timing attacks.
Furthermore, the proposed solution generates a secret encoding matrix (GEM) in a dynamic manner, and this is produced only by both the source and sink nodes, which allows to keep information secret against flow tracing attacks. Additionally, as mentioned before, our proposed scheme is based on the utilization of dynamic manner, which means that the use of a special encrypted packet fails to gain any useful information about the dynamic key as well as about the master key. Moreover, the key space of the master and the dynamic keys is 2^{256} and 2^{128}, respectively. Therefore, the key space of the master or the dynamic key of our scheme is sufficiently large to make the brute-force attack infeasible. Another benefit of using the dynamic key method is making the broken of our proposed scheme very difficult for an attacker who aims to get access to any information in the system.
All these arguments allow us to reach to a conclusion that our proposed solution is secure enough against global, flow tracing, timing and brute-force attacks. Hence, the immunity of our scheme, cryptographically talking is stronger than the existing traditional solutions.

5 CONCLUSION

The RLNC security has become very essential for a realistic, practical NC implementation, and many schemes were presented recently in this domain that deals with security issue. In this paper, a new security scheme was constructed and realized to provide a safe RLNC. Our solution is based on a new flexible and invertible secret key dependent integer matrix. This scheme provides at the same time, data confidentiality, and integrity and source authentication. The confidentiality is achieved by applying secret encoding process using invertible dynamic secret matrix G, and the authentication is realized in a selective manner. For our proposal, cryptographic strength (dynamic key in counter mode) against different types of attacks (statistical, linear, differential, Byzantine and eavesdropping). The results indicated that a satisfactory security and performance have been achieved.

REFERENCES

Bhattad, K. and Narayanan, K. R. (2005). Weakly secure network coding.

Cai, N. and Yeung, R. (2002). Secure network coding.

Clarkson, J. B. (1994). Dense probabilistic encryption. In *In Proceedings of the Workshop on Selected Areas of Cryptography*, pages 120–128.

Fan, Y., Jiang, Y., Zhu, H., and Shen, X. (2009). An efficient privacy-preserving scheme against traffic analysis attacks in network coding. In *INFOCOM*, pages 2213–2221. IEEE.

Fathy, A., ElBatt, T., and Youssef, M. (2011). A source authentication scheme using network coding. *Int. J. Secur. Netw.*, 6(2/3):101–111.

Ho, T., Mdard, M., Koetter, R., Karger, D. R., Effros, M., Shi, J., and Leong, B. (2006). A random linear network coding approach to multicast. *IEEE TRANS. INFORM. THEORY*, 52(10):4413–4430.

Koetter, R., Mdard, M., and Member, S. (2003). An algebraic approach to network coding. *IEEE/ACM Transactions on Networking*, 11:782–795.

Li, Y., Yao, H., Chen, M., Jaggi, S., and Rosen, A. (2010). Ripple authentication for network coding. In *Proceedings of the 29th conference on Information communications*, INFOCOM'10, pages 2258–2266, Piscataway, NJ, USA. IEEE Press.

Lima, L., Médard, M., and Barros, J. (2007). Random linear network coding: A free cipher? *CoRR*, abs/0705.1789.

Menezes, A. J., Vanstone, S. A., and Oorschot, P. C. V. (1996). *Handbook of Applied Cryptography*. CRC Press, Inc., Boca Raton, FL, USA, 1st edition.

Najeem, M. and Siva Ram Murthy, C. (2011). On enhancing the random linear network coding. In *Proceedings of the 2011 17th IEEE International Conference on Networks*, ICON '11, pages 246–251, Washington, DC, USA. IEEE Computer Society.

Noura, H., Martin, S., and Al Agha, K. (2013). A new efficient secure coding scheme for random linear network coding. In *Computer Communications and Networks (ICCCN), 2013 22nd International Conference on*, pages 1–7.

sang Park, J., Lun, D. S., Soldo, F., Gerla, M., and Macdard, M. (2006). Performance of network coding in ad hoc networks.

yen Robert Li, S., Member, S., Yeung, R. W., and Cai, N. (2003). Linear network coding. *IEEE Transactions on Information Theory*, 49:371–381.

Zhang, P., Jiang, Y., Lin, C., Fan, Y., and Shen, X. (2010). P-coding: secure network coding against eavesdropping attacks. In *Proceedings of the 29th conference on Information communications*, INFOCOM'10, pages 2249–2257, Piscataway, NJ, USA. IEEE Press.