QALAAI ZANIST JOURNAL
A Scientific Quarterly Refereed Journal Issued by Lebanese French University – Erbil – Kurdistan – Iraq
Vol. (2), No. (2), April 2017
Speical Issue : The 1st International Conference on Information Technology (ICoIT'17)
ISSN 2518-6566 (Online) - ISSN 2518-6558 (Print)

Using RSA Digital Signature to Solve Pollution Problem in Single-Source Network Coding

Rami S. Youail
Department of Information System Engineering, Erbil Polytechnic University-Iraq
r.rami.samir@gmail.com
r.rami.samir@epu.edu.krd
r.rami.samir@lfu.edu.krd *(Part time Lecturer at Lebanese French University)*

ARTICLE INFO

Article History:

Received: 15 March 2017

Accepted: 1 April 2017

Published: 10 April 2017

DOI:

10.25212/lfu.qzj.2.2.46

Keywords: *Network coding, pollution problem, RSA encryption algorithm, digital signature, authentication.*

ABSTRACT

Pollution problem in network coding becomes a serious problem, where no receiver can be reached by a source it does not need, in other words, the intermediate nodes may send polluted or faked messages into the network. This injection will prevent the sinks nodes from recovering the original messages correctly. Additionally, a malicious node can inject garbage into the distribution network, if undetected; the garbage will pollute the whole network. The most difficult problem in pollution propagation is the expansion of pollution. If a small number of polluted packets/messages is not detected at the early stages of the network; these small polluted messages will be expanded and used by downstream nodes and will affect the entire network. Therefore, the polluted or forged messages should be detected and filtered as early as possible before it grows up and distributed overall the network. To overcome this problem, a signature-based authentication system is proposed, which can detect the polluted or forged message at each intermediate/sink node. The system only needs to transmit the private key of the source node to each intermediate node.

1. INTRODUCTION

Network coding (NC) is a new technique proposed by Ahlswede et.al.[1] to maximize the throughput of multicast networks. The main idea behind NC is to allowing intermediate node to encode the incoming data and generates a new encoded packet. This technique is different from the traditional networking which is capable of duplicating and forwarding only. NC has various applications in computer networks, such as wireless networks and P2P systems. Network coding produces new challenges of security, for example, the pollution free problem, where no receiver can be reached by a source it does not need. If the pollution does not detected as early as possible, the pollution will grow up since the polluted messages could be used by all down-stream nodes, this kind of pollution will prevent the sink nodes from recovering the source messages correctly.

QALAAI ZANIST JOURNAL
A Scientific Quarterly Refereed Journal Issued by Lebanese French University – Erbil – Kurdistan – Iraq
Vol. (2), No. (2), April 2017
Speical Issue : The 1st International Conference on Information Technology (ICoIT'17)
ISSN 2518-6566 (Online) - ISSN 2518-6558 (Print)

2. Literature review

Many attempts have been done to overcome this problem. In [2], the pollution problem of network coding is solved by using a network of tagged pipes; where each link is partitioned into tagged pipes and each tag is a subset of the source messages. The packets outflowing in each pipe can only be a function of the sources mentioned in the tag. Authors in [3] can detect or filtered the polluted messages in sink nodes only. The work in [4] presents a security scheme for network coding that reduces the cost of verifying blocks on-the-fly while preventing the probing of malicious blocks. Digital signature, which is adopted in this work, is used to prevent this problem. The source node generates its own signature and attached it to the messages, the intermediate nodes check whether this signature is valid or not. The source node needs only to transmit its secret key to the intermediate nodes in order to regenerate a signature after packets being network coded.

3. INFORMATION SECURITY

In any information security system, there are several objectives need to be achieved by the system itself. If any of these objectives is absent; there will be a defect in the system and will not function properly [5]. RSA signature, by itself, can achieve these objectives. Some of these objectives which are achieved by the proposed system are listed below:

TABLE 1: ACHIEVED OBJECTIVE IN PROPOSED SYSTEM

confidentiality	Only authorized persons can access the secret information.
data integrity	Ensuring information has not been changed during the transmission.
Message authentication	It ensures that the sender is the real one, not other person claiming.
signature	Binding information to a source and the source cannot deny.

4. DIGITAL SIGNATURE

A cryptographic process which is considered as the backbone in message authentication, authorization, and non-repudiation is the *digital signature*. The purpose behind a digital signature is to provide a method for a signer to bind its identity to a piece of information. In easy way, the digital signature of a message is a number resulted from a function of some secret information belongs to the signer, and on the content of the message [5].

A. Preliminaries or Basic Terminologies and Concepts

In the following, some terms and basic concepts are defined [5].

- *M* is a set of message space which can be signed, *M* consists of strings of symbols from an alphabet, each element belongs to *M* is called plaintext.
- *K* is a set of key space.
- *S* is a set of elements called signatures.

QALAAI ZANIST JOURNAL
A Scientific Quarterly Refereed Journal Issued by Lebanese French University – Erbil – Kurdistan – Iraq
Vol. (2), No. (2), April 2017
Speical Issue : The 1st International Conference on Information Technology (ICoIT'17)
ISSN 2518-6566 (Online) - ISSN 2518-6558 (Print)

- S_A is a transformation function from M to S, and is called a *signing transformation* for the signer. S_A is kept secret by the signer and will be used to create signatures for messages from M.
- V_A is a transformation from the set $M \times S$ to the set {true, false}. V_A is called a *verification transformation* for the signer's signatures, which is publicly known, and is used by other entities to check the originality of the signature created by the signer.

The signing procedure (phase two in the proposed scheme) consists of two steps. In the first step, the signer computes a signature s for each message $m \in M$, and then transmits the pair (m,s) at the second step.

The verifier (phase three) needs to verify that a signature s on a message m is created (signed) by the real signer, so he\she needs to perform the following:

1. Obtain the verification function V_A of the signer, which is publicly known.
2. Calculate $u=V_A(m,s)$.
3. If u=true; the signature is correct and accepted, otherwise reject it.

B. The Proposed Model

The proposed model consists of three phases:

Phase one: the source node chooses the security parameters and calculates the public and private keys, send the secret key to intermediate nodes via secure channels. The source node sends the public key via unsecure channel or attaches it with the message.

Phase two: the source node calculates the signature which will be attached with its messages.

Phase three: the intermediate/sink nodes verify the received message.

RSA digital signature scheme is adopted in this work, the source node performs the following steps at phase one [5], [7]:

1. Generate two distinct random numbers n_1 and n_2. Note that n_1 and n_2 must be prime and large enough.
2. Calculate $n= n_1 \times n_2$ and $\phi = (n_1 - 1)(n_2 - 1)$.
3. Select a random integer e, which represents the public key, such that $1 < e < \phi$ and gcd $(e, \phi) = 1$.
4. Calculate the unique integer d where $1 < d < \phi$, such that $ed \bmod (n_1 - 1)(n_2 - 1) = 1$
5. The source's public key is (n,e) and its private key is d.

Taking in consideration that n should be long enough (for example; n=1024 bit length) and the difference between the primes *n1 and n2* (which are half length of n, i.e.: each one of them is 512 bits of length) should not be small. If n_1 and n_2 is small, then n could be factorized and the secret key could be determined [5].

At the second phase, the source node calculates the signature for each $m \in M$ as follows:

1. Compute the signature $s=m^d \bmod n$, $0 \leq m < n$.
2. Attach s to m and send the packet (m,s), downstream to the nodes.

QALAAI ZANIST JOURNAL
A Scientific Quarterly Refereed Journal Issued by Lebanese French University – Erbil – Kurdistan – Iraq
Vol. (2), No. (2), April 2017
Speical Issue : The 1st International Conference on Information Technology (ICoIT'17)
ISSN 2518-6566 (Online) - ISSN 2518-6558 (Print)

At the third phase, each node will verify the received signature by using the verification transformation, if the received packet has a valid signature and it does not contain any forged message, it will be accepted, otherwise ignored. The transformation verification is given by $m=s^e \bmod n$.

5. SYSTEM MODEL

There are $X_1,X_2,\ldots,X_m$ information source observed at source node *S*. The intermediate nodes are capable of encoding the incoming messages X_i ,i=1,..,m into encoded message. Hence, the encoded messages could be written as:

$$E = (\alpha_1, \alpha_2, \ldots, \alpha_m) \times (X_1, X_2, \ldots X_m)^T \mod \beta \qquad (1)$$

Where α_i is a coefficient picket up randomly from a finite field ß (*Galois Field*). Every intermediate node performs network coding operation on its incoming packets and generates the vector *E*, then computes the signature for *E*. If *S* denotes the signature function, then the signature will be calculated by using the, multiplicative property which is sometimes referred to as homomorphic property [5], [6], [7] as:

$$S(E) = \left(\sum_{i=1}^{r} (X_i \times \alpha_i) \bmod \beta \right)^d \bmod n \qquad (2)$$

Where *r* represents the number of incoming links (packets to be network coded) at any node , then append *S(E)* to *E* and forward it downstream, the packet format will be [*E* + *S(E)*], where the first part represents the encoded message and the second represents its signature.

The intermediate/sink node verifies the received message (packet) according to the source (signer) public key and the message signature's. Simply, the verifier check whether

$$S(E)^e \bmod n = E \qquad (3)$$

If the above equation satisfied, then the packet will be considered as valid one, otherwise it's polluted and will be discarded. Since each intermediate node or sink node will receive *r* packets from its *r* incoming links, so we need to compute (3) for *r* times in order to verify all the incoming packets which will take long time.

To increase the calculation speed of verification process; we support the batch verification [8]; which will speed up message verification. The verification process in "3" will change to be:

$$\left(\prod_{i=1}^{r} Si \right)^e \bmod n = \left(\prod_{i=1}^{r} m_i \right) \bmod n \qquad (4)$$

To see the proof of the above formula the reader may refer to [7].

QALAAI ZANIST JOURNAL
A Scientific Quarterly Refereed Journal Issued by Lebanese French University – Erbil – Kurdistan – Iraq
Vol. (2), No. (2), April 2017
Speical Issue : The 1st International Conference on Information Technology (ICoIT'17)
ISSN 2518-6566 (Online) - ISSN 2518-6558 (Print)

6. SECURITY ANALYSIS

The message is defined as polluted by the inequality (5):

$$S(E)^e \bmod n \neq E \qquad (5)$$

which means that the content of the message is incompatible with the signature attached with. To prevent unauthorized persons/nodes from polluting the encoding message; each signature should be calculated based on the secret key and the verified message.

In this system, determining the source's private key from its public key is equivalent to solve the problem of integer factorization, which is a hard problem to solve and its hardness depends on the length of *n*.

7. CONCLUSION

In this work, RSA digital signature scheme is adopted to solve the problem of pollution in single-source network coding. Each intermediate node will receive many signed packets; this will slow down the verification process. To overcome this slowness; the batch verification is used to speed up the verification process. The only disadvantage in our system is that we need to transfer the secret key to all the intermediate nodes via secure channels.

8. REFERENCES

[1] R. Ahlswede, N. Cai, S. Li, R. Yeung, "Network information flow," *IEEE Transaction on Information Theory*, vol. 46, no. 4, pp 1204-1216, 2000.
[2] Y. Wu, "On constructive multi-source network coding", in *IEEE International Symposium in Information Theory*, Seattle, July 9-14, 2006.
[3] S. Jaggi, M. Langberg, S. Katti, T. Ho, D. Katabi, M. Medard, "Resilient network coding in the presence of byzantine adversaries", *IEEE Proceeding INFOCOM*, 2007.
[4] Gkantsidis, P. R. Rodriguez, "Cooperative security for network coding file distribution," *IEEE Proceedings INFOCOM*, 2006.
[5] A. Menezes, P. V. Oorschot, S. Vnastone, *Handbook of Applied Cryptography*, 5th. ed. Florida: CRC Press, 2011.
[6] J. Talbot, D. Welsh, *Complexity and Cryptography*, 1st. ed., Cambridge, Cambridge University Press, 2006.
[7] L. Harn, "Batch verification multiple RSA digital signature," IEEE Xplore, vol. 10, Issue 12, pp. 1219-1220, 11 June, 1998.
[8] M. Bellare, J.A. Garay, T. Rabin, "Fast batch verification for modular exponentiation and digital signature", Proceedings of Advances in Cryptography, LNCS, vol. 1403, 1998