ELGAMAL SIGNATURE FOR CONTENT DISTRIBUTION WITH NETWORK CODING

Alireza Ghodratabadi[1] and Hashem Moradmand Ziyabar[2]

[1] EE institute, Malekashtar University of Technology, Tehran, Iran
[2]IRIB, Tehran, Iran

ABSTRACT

Network coding is a slightly new forwarding technique which receives various applications in traditional computer networks, wireless sensor networks and peer-to-peer systems. However, network coding is inherently vulnerable to pollution attacks by malicious nodes in the network. If any fake node in the network spreads polluted packets, the pollution of packets will spread quickly since the output of (even an) honest node is corrupted if at least one of the incoming packets is corrupted. There have been adapted a few ordinary signature schemes to network coding that allows nodes to check the validity of a packet without decoding. In this paper, we propose a scheme uses ElGamal signature in network coding. Our scheme makes use of the linearity property of the packets in a coded system, and allows nodes to check the integrity of the packets received easily.

KEYWORDS

network coding; crystallographic signature; homomorphic signature; ElGamal signature

1. INTRODUCTION

In recent years peer to peer networks are dynamic and each node in P2P networks is an information source. These nodes can come into and leave out the network instantly. Decentralized structure of these networks make it possible that malicious nodes disguise honest nodes and insert fake data in network. When network coding is applied, as nodes use their input to make a linear combination for each output, malicious packet spread even more. This way even honest nodes spread malicious packets. Therefore methods to detect these malicious packets seem to be necessary.

Recently there were also interests in applying network coding to distributed file systems. Many researchers have studied the usability and efficiency of network coding in peer to peer networks for distributing file and media [1], [2], [3]. Others were trying to use network coding in huge networks of computers and contribute it with peer to peer networks like freenet to update operating systems and software packages.

In traditional network model which is based on server-user model, server sends file according to user request. This model shows its inefficiency when the file size is large or the number of requests for a specific file is high, needing more bandwidth. Due to these reasons there has been a growing demand for peer to peer networks.

These networks are overcoming server-user networks; they have distributed structure and each user is considered as a server that can share network resources such as memory, processing power, and etc.

DOI : 10.5121/ijwmn.2014.6206

The best example for these kinds of networks is bitTorrent in which files are broken to smaller blocks. Whenever a node downloads a block that node is considered as its server. However these systems are being used now but due to some problems its performance is degraded considerably. For example synchronization always matters [4]. Synchronization means which block or blocks should be uploaded sooner or later. This problem will be more serious when rare blocks are placed on nodes with unreliable internet connection. To solve this problem [5] and [6] proposed to use network coding to enhance performance. Same as file is broken to blocks and every time that a user asks for the file, server send a random linear combination of its blocks to it. When the user receives the first block of the file it will be as the server for that block.

In this method each node (user) is able to restore the file only when it receives enough blocks to solve a linear set of equations. This scheme which is based on network coding omits the need for synchronization [5]. In [7] it has been shown mathematically that using this scheme, performance and reliability are also enhanced.

Here we are interested in security of distributed content schemes that use network coding. Our main considerations are simplicity and security performance. In every network that uses network coding avoiding malicious node. For example in the discussed peer to peer network with network coding, if a fake node imitate a valid node it can make a polluted block with correct coding coefficients and spread it over the network [8]. If there is not a mechanism for checking integrity of blocks it is possible for fake nodes to insert their packets into the network and other nodes who receive some polluted packet along with valid packets are not able to decode messages correctly.
In previous schemes where network coding was not used, data integrity was checked by means of hash-and-sign scheme. Source node exploited a hash function H(.) to compute hash H(X) for its data X. then it uses a digital signature method to sign it S(H(X)) and then send it. To defend against this attack many methods are proposed. [5] and [8] used homomorphic hash functions. In fact homomorphism enables nodes to sign every combination of messages without knowing their private key. In [9] exploiting SRC is suggested that is less complex than previous solutions but needs a secure channel for SRCs. Authors of [4] tried to omit secure channel so they defined a orthogonal space from messages. If received message is in this space it will be known as valid else it will be thrown away.

In this paper we have proposed a method for protecting packets from pollution attacks. We used cryptographic methods that exploit linear random coding. Data packets are considered as vector blocks. Packet contents are signed with ElGamal signature. In each receiver node, the validity of signature is being verified. If the signature is not valid that block will be omitted. Otherwise the node will use the block for making its output blocks. Our method have a larger overhead in comparison to previous signatures, but it is more robust to polluted node attacks.

The organization of paper is as follows. In section II, ElGamal signature is described briefly. Our proposed method is presented in section III along with a detailed example. Finally section IV concludes the paper.

2. ElGamal Signature

The proposed scheme is based on well-known ElGamal signature scheme proposed in 1984. Here we briefly describe this signature scheme. In this scheme when Alice wants to sign a message and send it to Bob, she chooses a prime number p, a generator g of a product field, F_p^*, and a random number x such that $0 < x < p-1$. She computes her public key as $y = g^x \bmod p$ and spreads public key vector (p, y, g). Alice also chooses a secret random number k and computes s and r as,

$$\begin{cases} \mathrm{r} = \mathrm{g}^{\mathrm{k}} \bmod \mathrm{p} \\ (\mathrm{k}, \mathrm{p}-1) = 1 \bmod \mathrm{p} \end{cases} \begin{cases} \mathrm{s} = \mathrm{k}^{-1}(\mathrm{m} - \mathrm{xr}) \bmod \mathrm{p} - 1 \\ \mathrm{m} : \mathrm{message} \end{cases}$$

The signature of message m, is (r,s). Receiving the message and its signature, Bob want to verify it. First he checks whether $1 \le r \le p-1$. If it does not hold he rejects the message, otherwise he checks the congruence $y^r r^s = g^m \bmod p$ .He accepts the signature if this congruence holds too.
Proposed Scheme

Using ElGamal signature in peer-to-peer networks where network coding is implemented, enhances security by determining intentionally inserted errors. Just with these methods, peer-to-peer network can gain the benefits of network coding.

First, we describe linear network coding briefly. Let $G = (V, E)$ be a directed graph. Suppose a source node $s \in V$ sends a large file *F* to set $T \subseteq V$ of nodes in *V*. File *F* is consisted of k *n*-dimensional vectors $\mathbf{m}_1, ..., \mathbf{m}_k \in F^n$, where F^n is the n-dimensional vector space over a finite field F. So the file is consisted of k vectors as,

$$\begin{cases} \mathbf{m}_1 = (m_{1,1}, ..., m_{1,n}) \\ \mathbf{m}_2 = (m_{2,1}, ..., m_{2,n}) \\ \qquad ... \\ \mathbf{m}_k = (m_{k,1}, ..., m_{k,n}) \end{cases}$$

Then the source creates *augmented vectors* of $\mathbf{m}_i = (m_{i,1}, ..., m_{i,n})$ by setting

$$\bar{\mathbf{m}}_i = (\overbrace{\underbrace{0, ..., 0, 1}_{i}, 0, ..., 0}^{k}, m_{i,1}, ..., m_{i,n})$$

that is, the dimension of each augmented vector is $t = k + n$, and the first *m* entries are all zero except at *i*-th, where it is 1.

Assuming that **m**₂ is being transmitted through the network. The source chooses a big prime number p and assumes g_1 to g_n as generators of product field, F_p^*. Also it assumes x as secret key and computes y_i's as
$y_i = (g_i)^x \bmod p$

Then the source distributes (p, g_i, y_i) as public key. For signing a message, it chooses a random secret number k so as k is prime to p-1, i.e. $(k, p-1) = 1$. Now it computes r_i's as;

$r_i = (g_i)^k \bmod p$

And sets r as product of r_i's and computes s too:

$$r = \prod_{i=1}^{n} r_i$$

$$s = k^{-1}\prod_{i=1}^{n}(m_{ji} - r_i) \bmod (p-1)$$

The signature is send as (r,s). At each node, receiver can check the validity of message by $r^s.y^r$; According to ElGamal signature, this term should be equal to $(g_1,...,g_n)^m \bmod p$. If the receiver node gains this, it verify the message otherwise the message is incorrect and probably fake and will be thrown away.

Table 1. Algorithm of Signature Scheme

Step	Action
System parameters	1- Selecting big prime number 2- Selecting g_1 to g_n as generator of product field, F_p^*
Producing key	1- Selecting x as secret key 2- Computing y_1 to y_n; $y_i = (g_i)^x \bmod p$ 3- Distribute y_i's, g_i and p as public key
Producing signature	1- Selecting random number k and computing r_i's so that (k,p-1)=1 and $r_i = (g_i)^k \bmod p$ and $r = \prod_{i=1}^{n} r_i$ 2- Computing $s = k^{-1}\prod_{i=1}^{n}(m_{ji} - r_i) \bmod (p-1)$ 3- Sending (r,s) as signature
Validating signature	checking $(g_1,...,g_n)^m \bmod p \overset{?}{=} r^s.(y_1 ... y_n)^r \bmod p$

ElGamal signature is from hardness degree of discrete logarithm in product field, F_p^*. So unless discrete logarithm is broken, it will be secure.

2.1. An example

Here we give an example to make the method clear. Let m=(1,0,1,1) in binary format which is equal to 11 in decimal.

Assume p=11, so primitive roots g_i's are {2,6,7,8} and private key x=9 and

$$
\begin{aligned}
y_1 &= (g_1)^9 \bmod 11 = 6\\
y_2 &= (g_2)^9 \bmod 11 = 2\\
y_3 &= (g_3)^9 \bmod 11 = 8\\
y_4 &= (g_4)^9 \bmod 11 = 7
\end{aligned}
$$

Therefore the public key is $\left\{11,\overbrace{2,6,7,8}^{g_i\text{'s}},\overbrace{6,2,8,7}^{y_i\text{'s}}\right\}$. Since K have to prime to $p\text{-}1$, $\ k,p-1\ =1$, it should be selected form set {3,5,7,9}. Assume K= 7. Computing r_i's:

$$
\begin{aligned}
r_1 &= (g_1)^7 \bmod 11 = 7\\
r_2 &= (g_2)^7 \bmod 11 = 8\\
r_3 &= (g_3)^7 \bmod 11 = 6\\
r_4 &= (g_4)^7 \bmod 11 = 2
\end{aligned}
$$

So $r = r_1'.r_2'.r_3'.r_4' = 672$ and $r \bmod 11 = 1$. Now to find k^{-1} we should find x so that $K.x \bmod 10 = 1$ which concludes to $k^{-1} = 3$.

S is;

$$
\begin{aligned}
&s = k^{-1}\prod_{i=1}^{n}(m_{ji} - r_i)\bmod(p-1) =\\
&= 3(0-9\times 7)(1-9\times 8)(0-9\times 6)(1-9\times 2)\bmod 10\\
&= 3(-62)(-71)(-54)(-17)\bmod 10 = 12123108 \bmod 10\\
&= 8
\end{aligned}
$$

So the signature is (672,8). To check the signature, this equation is checked;

$$
\begin{aligned}
&(g_1,\dots,g_n)^m \bmod p \overset{?}{=} r^s.(y_1\dots y_n)^r \bmod p\\
&(2\times 6\times 7\times 8)^{11} \bmod 11 \overset{?}{=} 672^8.672^{672} \bmod 11\\
&\overset{\text{OK}}{1 = 1}
\end{aligned}
$$

Thus the equality means a valid signature.

3. CONCLUSIONS

Security problem is a main obstacle in the implementation of content distribution networks using random linear network coding. To tackle this problem, instead of trying to fit an existing signature scheme to network coding based systems, in this paper, we proposed a new signature scheme that is made specifically for such systems. We introduced a signature vector for each file distributed, and the signature can be used to easily check the integrity of all the packets received for this file. We have shown that the proposed scheme is as hard as the Discrete Logarithm problem, and the overhead of this scheme is negligible for a large file.

ACKNOWLEDGEMENTS

The authors would like to thank everyone, just everyone!

REFERENCES

[1] T. Ho, M. Medard, M. Effros and D. Karger, "The benefits of coding over routing in a randomized setting," IEEE Symposium on Information Theory, 2003.

[2] Z. Li and B. Li, "Network coding: the case of multiple unicast sessions," 42th Annual Allerton Conference on Communication, control and Computing, 2004.

[3] D. S. Lun, M. Medard and R. Koetter, "Network coding for efficient wireless unicast," International Zurich Seminar on Communications, 2006.

[4] F. Zhao, "Signature for Content Distribution with Network Coding," 2009.

[5] S. Acedanski, S. Deb, M. Medard and R. Koetter, "How good is random linear coding based distributed network storage?," netcode, 2005.

[6] P. Rodriguez and C. Gkantsidis, "Network Coding for large scale content distribution," IEEE INFOCOM, 2005.

[7] C. Gkantsidis, J. Miller and P. Rodriguez, "Comprehensive view of a live network coding P2P system," ACM SIGCOMM/USENIX Internet Measurement Conference, 2006.

[8] C. Rodriguez and G. P., "Cooperative security for network coding file distribution," IEEE INFOCOM, 2006.

[9] M. Krohn, M. Freedman and D. Mazieres, "On-the-fly verification of rateless erasure codes for efficient content distribution," IEEE Symposium on Security and Privacy, 2004.