

## 网络编码数据传输的联合保障机制

朱馨培<sup>1\*</sup>, 寇应展<sup>1</sup>, 王湛昱<sup>2,3</sup>

(1. 军械工程学院 信息工程系, 石家庄 050003; 2. 哈尔滨工业大学 航天学院, 哈尔滨 150001;

3. 哈尔滨工业大学 机电工程学院, 哈尔滨 150001)

(\* 通信作者电子邮箱 zxp910224@gmail.com)

**摘要:**为了提高基于网络编码的数据传输的完整性、机密性和隐私性,提出了一种使用数字水印、栈混洗和信息认证码(MAC)等技术的安全保障机制。该机制通过异或(XOR)加密和栈混洗技术来混淆信息,提供机密性和隐私保证;通过数字水印将信息认证码随机插入混淆后的信息,进而提高机密性;中间转发节点通过验证部分信息认证码来提供完整性保证。仿真结果表明,该机制能有效降低污染信息传播跳数(低于1.5跳),即使拥有25个共谋攻击者和密钥池大小为100时,共谋概率也仅为0.1。理论分析和仿真实验证明,所提机制使得基于网络编码的数据传输能以较低的性能开销抵抗偷听攻击、流量分析攻击和污染攻击。

**关键词:**网络编码;数据传输;数字水印;栈混洗;信息认证码

**中图分类号:** TP393.08; TP309.2 **文献标志码:** A

### United protection mechanism for network-coding-based data transmission

ZHU Xinpei<sup>1\*</sup>, KOU Yingzhan<sup>1</sup>, WANG Zhanyu<sup>2,3</sup>

(1. Department of Information Engineering, Ordnance Engineering College, Shijiazhuang Hebei 050003, China;

2. School of Astronautics, Harbin Institute of Technology, Harbin Heilongjiang 150001, China;

3. School of Mechatronics Engineering, Harbin Institute of Technology, Harbin Heilongjiang 150001, China)

**Abstract:** To improve the integrity, confidentiality and privacy of network-coding-based data transmission, a secure protection mechanism combined digital watermarking, stack shuffle and Message Authentication Code (MAC) was proposed. In this mechanism, the confidentiality and privacy were provided by mixing up messages using exclusive OR (XOR) encryption and stack shuffle. Furthermore, the confidentiality was enhanced by randomly inserting MACs into mixed messages with digital watermarking technique. And the integrity was provided by checking MACs on intermediate nodes during transmitting. The simulation results show that the spread hops of polluted information were effectively reduced by using this mechanism (less than 1.5). The collusion probability was less than 0.1 even if there were 25 collusion attackers and the size of key pool was 100. Both of theoretical analysis and simulation experiment demonstrate that the proposed mechanism can defend eavesdropping attacks, flow analysis attacks and polluting attacks with low expense.

**Key words:** network coding; data transmission; digital watermarking; stack shuffle; Message Authentication Code (MAC)

## 0 引言

网络编码是一种新颖的数据传输技术,它允许中间转发节点对输入的信息进行编码(或混合)运算,并转发给下游节点编码后的信息<sup>[1]</sup>。研究证实网络编码能够有效地提高网络通信的吞吐量<sup>[2]</sup>,降低数据传输能量消耗<sup>[3]</sup>,增强网络通信的鲁棒性和减少通信延迟<sup>[4]</sup>等。网络编码的潜在优势使得它适合各种各样的网络和通信系统。目前,网络编码在无线传感器网络<sup>[5]</sup>、P2P(Peer to Peer)内容分发和存储网络<sup>[6]</sup>、无线 Mesh 网络<sup>[7]</sup>、视频流网络<sup>[8]</sup>等均有广泛应用。

在实际应用中,网络编码在中间转发节点对输入信息进行处理的特点也使得无线传感器网络容易遭受各种攻击且难以检测和抵抗。这些潜在的安全攻击严重破坏了网络编码的机密性、完整性和隐私。在无线传感器网络中,恶意攻击者可以通过监听链路通信和使用先进的分析技术来发起偷听攻

击<sup>[9]</sup>和流量分析攻击<sup>[10]</sup>,进而获取通信内容或信息收发者的隐私;此外,无线传感器网络还容易遭受污染攻击<sup>[11]</sup>等,污染攻击的发起者主动篡改通信内容,从而使得目标节点收到的信息与源节点发送的信息不一致。综上,偷听攻击、流量分析攻击和污染攻击对网络编码构成了巨大的威胁。

目前,关于网络编码的完整性,现有安全保障机制主要分为两类<sup>[12]</sup>:基于信息学理论的安全机制和基于加密技术的安全机制。关于网络编码的机密性,现有机制主要分为3个层次<sup>[13]</sup>,即 Shannon 安全、弱安全和计算安全。关于网络编码的隐私保护,Fan 等<sup>[14]</sup>使用轻量级的同态加密函数来抵抗流量分析攻击。尽管各种安全保障机制被学者们提出,但是这些工作都只针对某个攻击提出的独立防御方案,而不是从一个系统性的角度来针对上述所有攻击提出解决机制。单独考虑安全威胁的后果可能会带来更多的通信开销和计算开销。单独解决方案的简单组合可能会引入新的安全隐患,或者会

收稿日期:2014-02-12;修回日期:2014-03-12。 基金项目:国家自然科学基金资助项目(61173191)。

作者简介:朱馨培(1991-),女,吉林长春人,硕士研究生,主要研究方向:网络信息安全;寇应展(1962-),男,河北石家庄人,教授,硕士,主要研究方向:信息安全、密码学;王湛昱(1982-),男,吉林长春人,工程师,硕士,主要研究方向:计算机网络、网络化控制。

为某个未知攻击留有后门。因此,有必要综合考虑网络编码的机密性、完整性与隐私问题,提出一个能够抵抗偷听攻击、流量分析攻击和污染攻击的联合防御机制。

为了解决上述问题,本文针对污染攻击、偷听攻击和流量分析攻击提出了一个安全、高效的联合保障 PIC (Privacy Integrity Confidentiality) 机制。该机制通过使用 XOR 编码、栈混洗<sup>[15]</sup>和数字水印<sup>[16]</sup>等技术为网络编码同时提供完整性、机密性和隐私保证。XOR 加密和栈混洗技术的引入为网络编码提供机密性和隐私保护;数字水印技术和同态信息认证码(Message Authentication Code, MAC)<sup>[17]</sup>则将用于完整性验证,也增强了机密性和隐私保护。

## 1 问题描述

### 1.1 系统模型

本文主要考虑基于网络编码数据传输的单会话多播网络(Single-session Multicast Network, SMN)。该模型可以很好地概括无线传感器网络分布式数据存储、视频流网络、P2P 存储与分发网络等多种应用。单会话多播网络由一个源节点  $S$ , 多个中间转发节点  $F_1, F_2, \dots, F_f$  和多个目标节点  $D_1, D_2, \dots, D_s$  组成,如图 1 所示。

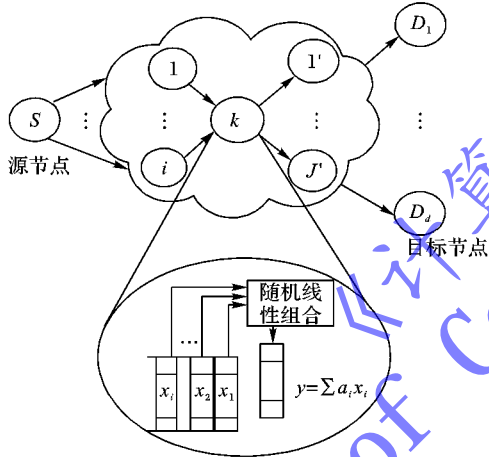


图1 基于网络编码数据传输的单会话多播网络

在通信阶段,源节点  $S$  给目标节点发送  $n$  个信息  $M_1, M_2, \dots, M_n$ 。每个信息  $M_i$  包括  $m$  个码字,每个码字都是有限域  $\mathbb{F}_q$  上的元素,其中  $q$  是预先设定的系统参数。则每个信息  $M_i (i = 1, 2, \dots, n)$  可以表示为:

$$M_i = (m_{i,1}, m_{i,2}, \dots, m_{i,m}) \quad (1)$$

其中  $m_{i,j} \in \mathbb{F}_q (j = 1, 2, \dots, m)$  表示码字。实际应用中,每个信息前需要加上编码向量构成编码信息。例如,源信息  $M_i$  的编码信息  $\tilde{M}_i$  可以表示为:

$$\tilde{M}_i = (\beta_i, M_i) = (\underbrace{0, \dots, 0}_{i-1}, \underbrace{1, 0, \dots, 0}_{n-i}, m_{i,1}, m_{i,2}, \dots, m_{i,m}) \quad (2)$$

其中  $\beta_i = (\beta_{i,1}, \beta_{i,2}, \dots, \beta_{i,n})$  为全局编码向量(Global Encoding Vector, GEV)。对于中间节点的每次转发,输出的编码信息可以表示为:

$$\tilde{E} = (\beta, E) = (\beta_1, \beta_2, \dots, \beta_n, e_1, e_2, \dots, e_m) \quad (3)$$

其中  $e_j \in \mathbb{F}_q (j = 1, 2, \dots, m)$  表示信息  $E$  的码字。输出的编码信息是输入信息  $\tilde{X}_i$  的线性组合,即:

$$E = \left( \sum_i \alpha_i \tilde{X}_i \right) \bmod q \quad (4)$$

其中  $\alpha_i$  为局部编码系数(Local Encoding Coefficient, LEC),它是从有限域  $\mathbb{F}_q$  上随机选取的元素。以此类推,可以把编码信息  $E$  表示成源信息的线性组合,即:

$$E = \sum_{i=1}^n \beta_i M_i \bmod q \quad (5)$$

当目标节点收集到足够多的编码信息后,把每个信息的编码向量按序排列,组成全局编码矩阵(Global Encoding Matrix, GEM)  $G$ 。目标节点通过执行下列运算恢复源信息:

$$[M_1 \ M_2 \ \dots \ M_n]^T = G^{-1} [E_1 \ E_2 \ \dots \ E_n]^T \quad (6)$$

其中矩阵  $G^{-1}$  是  $G$  的逆矩阵。

### 1.2 攻击模型

为方便讨论,假设源节点和目标节点都是安全的,而中间节点是不可信且容易遭受攻击的。本文主要考虑两类的潜在攻击,即主动攻击和被动攻击。

1) 主动攻击。其目标在于完全控制中间转发节点并发起污染攻击,即恶意篡改输出的数据使得目标节点无法恢复源数据。如果传输的编码信息满足式(7),则该信息是污染信息。

$$(e_1, e_2, \dots, e_m) \neq (\beta_1, \beta_2, \dots, \beta_n) \times (M_1 \ M_2 \ \dots \ M_n)^T \bmod q \quad (7)$$

其中  $e_i (i = 1, 2, \dots, m)$  和  $\beta_j (j = 1, 2, \dots, n)$  分别表示码字和编码信息的全局编码向量。由于网络编码在中间转发节点具有混淆输入信息的特点,一个污染信息很容易引起数据污染并在下游节点处大规模泛滥,使得污染信息的数量以指数形式递增。

2) 被动攻击。被动攻击者监听通信链路并截获监听的消息,通过分析获取诸如信息内容或收发者的隐私信息等;或者破坏转发节点,读取其内存,从而破坏点到点(Link to Link)的加密机制。

本文认为存在一个安全且匿名的路由协议来协助中间节点决定转发路径。由于有安全路由机制的保障,攻击者不能获取信息的时间戳来进行后续分析。

## 2 联合保障机制设计

### 2.1 机制概述

PIC 机制首先使用 XOR 加密技术对源信息的每个码字进行加密,并使用栈混洗技术打乱编码向量和信息码字的排列顺序,有效地保证网络编码的隐私和机密性。对于完整性保证,本文使用数字水印对上述信息进行运算,产生一系列同态信息认证码(MACs)。任何一个信息认证码都可以用来验证整个信息的完整性。与其他安全保障机制不同,本文所提 PIC 机制不是将信息认证码简单地放置于信息的开头或结尾,而是使用水印将这些认证码隐藏在混淆后的信息中。只有知道某个信息认证码的具体位置才能够进行动态完整性验证。由于目标节点知道所有 MACs 的位置,因此它可以剔除所有信息认证码并恢复源数据。通过上述运算, PIC 机制不仅可以为网络编码提供完整性和机密性保证,还可以防止攻击者识别出源节点和目的节点的隐私信息。PIC 机制的思路如图 2 所示。

PIC 机制由 4 部分组成,即参数设置、传输信息生成、完整性验证和源信息恢复。

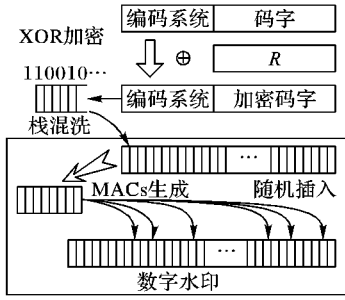


图2 PIC 机制框架

## 2.2 参数设置

1) 针对目标网络,设置大小为  $K$  的密钥池。网络中每个节点从密钥池中随机选取  $\omega$  个密钥,并将它们安全地保存在存储器中。

2) 选择两个公开的哈希函数  $H(\cdot)$ ,  $H_r(\cdot): 2^r \rightarrow \mathbb{F}_q$ 。给定一个种子  $x$ ,使用上述哈希函数可以计算出哈希链。定义符号  $H^i(\cdot)$  表示对种子连续执行  $i$  次哈希运算。

3) 选择一个伪随机函数(Pseudo Random Function, PRF)  $PRF(\cdot)$ ,该函数仅对源节点和目标节点公开。函数  $PRF(\cdot)$  的作用是从  $[1, n + m + \omega]$  中随机选择一个元素不同、大小为  $\omega$  的随机序列。

4) 在通信阶段,源节点和目标节点共享一个安全密钥(Secret Key, SK)  $SK$ 。

参数设置阶段的所有工作都是离线的,因此无需考虑效率问题。

## 2.3 传输信息生成

在数据传输前,源节点将生成包括编码向量、XOR 加密后的码字、 $\omega$  个信息认证码(MACs)和  $\omega$  个位置标签的传输信息,具体生成过程如下:

1) 由于源节点和目标节点共享密钥  $SK \in \mathbb{F}_q$ ,它们都是使用该密钥计算本次通信的会话密钥  $SK_i = H(SK_{i-1})$ ,其中  $SK_0 = SK$ 。

2) 源节点产生 XOR 加密密钥  $R_i = H_r(SK_i)$ ,将  $R_i$  分解成  $m$  个密钥  $R_{i,1}, R_{i,2}, \dots, R_{i,m}$ ,并对源信息进行一次一密式 XOR 加密,即  $\hat{M}_j^i = (m_{j,1} \oplus R_{i,1}, m_{j,2} \oplus R_{i,2}, \dots, m_{j,m} \oplus R_{i,m})$  ( $j = 1, 2, \dots, n$ )。

3) 基于密钥  $R_i$ ,使用栈混洗技术打乱上述信息并得到混洗信息  $\tilde{M}_i^2$  ( $i = 1, 2, \dots, n$ )。

4) 源节点对密钥  $SK_i$  使用伪随机函数  $PRF(\cdot)$ ,得到序列  $p_j$  ( $j = 1, 2, \dots, \omega$ )。符号  $p_j$  表示  $MAC_j$  在传输信息中的具体位置。

5) 源节点从有限域  $\mathbb{F}_q$  中随机选取  $\omega$  个元素  $r_1, r_2, \dots, r_\omega$ 。对于每个元素  $r_i$ ,使用哈希函数  $H(\cdot)$  计算一个长为  $m + n + \omega$ 、以  $H(r_i)$  为起始的哈希链。

6) 通过下面有  $\omega$  个未知数的方程组求解信息认证码(MACs)。

$$\begin{bmatrix} r_{1,1} & \dots & r_{1,p_i} & \dots & r_{1,m+n+\omega} \\ r_{2,1} & \dots & r_{2,p_i} & \dots & r_{2,m+n+\omega} \\ \vdots & & \vdots & & \vdots \\ r_{\omega,1} & \dots & r_{\omega,p_i} & \dots & r_{\omega,m+n+\omega} \end{bmatrix} \begin{bmatrix} m_{1,1} \\ \vdots \\ m_{i,1} \\ \vdots \\ m_{\omega,1} \end{bmatrix} = \begin{bmatrix} MAC_1 \\ \vdots \\ MAC_i \\ \vdots \\ MAC_\omega \end{bmatrix} \quad (8)$$

当有限域  $\mathbb{F}_q$  足够大时,系数矩阵是可逆的,信息认证码(MACs)也可以通过上述方程求解得到。此时,源节点可以构建扩展信息。扩展信息中  $p_j$  ( $j = 1, 2, \dots, \omega$ ) 位置的内容是信

息认证码  $MAC_j$ ,而其他位置的内容是混淆信息,且排列顺序与混淆信息相同。

7) 对于每个位置  $p_i$ ,源节点构建三元组  $\{p_i, r_i, k_i\}$ ,并使用密钥  $k_i$  进行加密,即  $\{p_i, r_i, k_i\}_{k_i}$ ;其次,源节点在扩展信息的后面附加上述位置标签;最后,源节点发送。

$$\{\hat{M}_i = (m_{i,1}, \dots, MAC_j, \dots, m_{i,m+n}), \{p_i, r_i, k_i\}_{k_i}, \dots, \{p_\omega, r_\omega, k_\omega\}_{k_\omega}\} \quad (9)$$

源节点生成传输信息的具体算法如下:

```

 $SK_i = H(SK_{i-1});$  // Produce the key in this generation
 $R_i = H_r(SK_i);$ 
 $M_j^1 = (m_{j,1} \oplus R_{i,1}, m_{j,2} \oplus R_{i,2}, \dots, m_{j,m} \oplus R_{i,m});$ 
 $\hat{M}_j^2 = \text{StackShuffle}(\hat{M}_j^1, R_i);$ 
 $[p_0, p_1, \dots, p_\omega] = PRF(SK_i);$ 
// Generate unequal random numbers in  $[1, m + n + \omega]$ 
for each  $j \in [1, \omega]$  do
     $r_j = \text{Random}();$ 
end for
for each  $j \in [1, \omega]$  do
    for each  $k \in [1, m + n + \omega]$  do
         $r_{j,k} = H^k(r_j);$ 
    end for
end for
Solve(Eq. (10));
Construct  $Msg(\hat{M}_j)$ ;
// size of  $\hat{M}_j = m + n + \omega$ , and the position of  $MAC_{i,j}$  is  $p_j$ 
for each  $j \in [1, \omega]$  do
    Encrypt( $p_j, r_j, k_j$ );
    Append(); // Append the encrypted data to the message
end for

```

当中间转发节点收到多个输入信息时,它对输入信息进行网络编码操作,即按照式(4)对输入信息进行线性随机组合。由于本文的信息认证码具有同态性,因此也可以进行网络

编码运算。输出的信息为  $E = \sum_{i=1}^n \beta_i \hat{M}_i \bmod q$ ,输出信息  $E$  的

信息认证码也为  $MAC_j = \sum_{i=1}^n \beta_i MAC_{i,j} \bmod q$ ,其中  $MAC_{i,j}$  是

输入信息  $\hat{M}_i$  第  $j$  个信息认证码。证明上述操作的正确性:

$$\begin{aligned} MAC_j &= r_1 \times e'_{j,1} + r_2 \times e'_{j,2} + \dots + r_{p_j} MAC_j + \dots + r_{n+m+\omega} \times \\ e'_{j,n+m} &= r_1 \times \sum_{i=1}^n \beta_i m'_{i,1} + r_{p_j} \sum_{i=1}^n \beta_i MAC_{i,j} + \dots + r_{n+m+\omega} \times \\ \sum_{i=1}^n \beta_i m'_{i,m+n} &= \sum_{i=1}^n \beta_i (r_1 \times m'_{i,1} + r_2 \times m'_{i,2} + \dots + \\ r_{p_j} MAC_{i,j} &+ \dots + r_{n+m+\omega} \times m'_{i,n+m}) = \sum_{i=1}^n \beta_i \times MAC_{i,j} \quad (10) \end{aligned}$$

由式(10)得,中间转发节点无需知道源节点和目标节点的共享私钥就可以计算输出信息。除了信息认证码(MACs),中间节点需要附加一些位置标签来标识 MAC 的具体位置。由于信息  $\hat{M}_i$  ( $i = 1, 2, \dots, n$ ) 中加密的位置都是相同的,因此同一批传输信息之需要传一次位置标签即可。

## 2.4 完整性验证

为了防止污染信息在网络中大规模传播,中间转发节点需要验证输入信息的正确性,并过滤掉错误的传输信息。当中间转发节点收到一个输入信息时,先检查输入信息的位置标签,判断是否具有解密位置标签的密钥。如果有密钥  $k_i$ ,则解密对应的位置标签并获得  $p_i, r_i, k_i$ 。随后,中间转发节点对随



机数  $r_i$  使用公开哈希函数  $H(\cdot)$  来计算长度为  $m+n+\omega$ 、起始为  $H(r_i)$  的哈希链。该哈希链上元素使用符号  $r_{i,1} \sim r_{i,m+n+\omega}$  来表示。由于第  $p_i$  个元素是  $MAC_i$ , 因此可以使用式(11)来验证数据的正确性:

$$e_1 \times r_{i,1} + e_2 \times r_{i,2} + \dots + MAC_i \times r_{i,p_i} + \dots + e_{m+n} \times r_{i,m+n+\omega} = MAC_i \quad (11)$$

如果式(11)成立,则该传输信息是正确的;否则,丢弃该错误信息。动态完整性验证算法如下:

```

 $r_{i,0} = r_i;$ 
for each  $j \in [1, m+n+\omega]$  do
     $r_{i,j} = H(r_{i,j-1});$  //The position of MAC
end for
 $A \leftarrow e_1 \times r_{i,1} + e_2 \times r_{i,2} + \dots + MAC_i \times r_{i,p_i} + \dots + e_{m+n} \times r_{i,m+n+\omega};$ 
 $B \leftarrow$ The  $p_i$ -th codeword of the message  $E$ ;
if  $A = B$  then
    return TRUE;
else
    return FALSE;
end if

```

### 2.5 源信息恢复

当目标节点收集到足够多的传输信息时,则恢复源信息。首先,目标节点根据式(6)求解源节点处的传输信息。其次,由于源节点和目标节点之间共享密钥  $SK$ ,目标节点使用公开的哈希函数计算  $SK_i$  和  $R_i$ 。目标节点按照2.2节中第4)步计算出信息认证码的具体位置,按照式(11)验证信息的完整性,并删除所有信息认证码。最后,目标节点通过执行栈混洗的逆运算与 XOR 解密操作来获取源信息。

## 3 安全评价

### 3.1 完整性评价

考虑两种情况分别评价 PIC 机制在完整性保护方面的性能。

情况1 攻击者不知道 MACs 的具体位置,只会盲目篡改传输信息  $\hat{E}$  中的任何部分,并将该污染信息发送给下游转发节点。对于接收到该污染信息的第一个下游转发节点,它和该信息源节点共享  $u$  个密钥的概率为:

$$Pr_k = \frac{C_u^{m+n+\omega} C_{K-\omega}^{m+n+\omega}}{C_K^{m+n+\omega}} \approx C_u^u \left( \frac{\omega}{K} \right)^u \left( \frac{K-\omega}{K} \right)^{\omega-u} \quad (12)$$

若  $u \geq 1$ ,则该节点可以解密  $u$  个位置标签,并获得对应的信息认证码进行动态完整性验证。假设该转发节点知道的信息认证码是  $MAC_j$ ,则有:

$$m_{i,1}r_{j,1} + m_{i,2}r_{j,2} + \dots + MAC_j r_{j,p_j} + \dots + m_{i,m+n}r_{j,m+n} = MAC_j \quad (13)$$

其中,部分  $m_{i,k}$  或信息认证码(MACs)可能被篡改,但公式仍成立的概率是  $Pr_p = 1/p$ 。由于完整性验证节点有  $u$  个密钥,则可进行  $u$  次完整性验证。污染信息能够成功逃脱完整性验证的概率为:

$$Pr_{\text{success1}} = \sum_{u=0}^{\omega} Pr_k \cdot (Pr_p)^u = \sum_{u=0}^{\omega} C_u^u \left( \frac{\omega}{K} \right)^u \left( \frac{K-\omega}{K} \right)^{\omega-u} \quad (14)$$

情况2 攻击者知道部分信息认证码(MACs)的具体位置,能够篡改所知道的信息认证码从而使得式(13)在动态完整性验证时成立。当攻击者与源节点之间没有共享密钥时,污染信息能够成功逃脱完整性验证的概率为:

$$Pr_{dk} = Pr_k(0) = C_{K-\omega}^{\omega} / C_K^{\omega} \approx (1 - \omega/K)^{\omega} \quad (15)$$

若下游第一个节点与源节点共享  $u$  个密钥,其中  $t$  个与攻击者拥有的不同,则该污染信息能够成功逃脱完整性验证的概率为:

$$Pr_s = \sum_{u=1}^{\omega} \left( Pr_k(u) \sum_{s=0}^u \frac{C_s^{m+n+\omega} C_{K-u}^{m+n+\omega}}{C_K^{m+n+\omega}} (1/q)^{u-s} \right) \quad (16)$$

综上,污染信息能够成功逃脱完整性验证的概率为:

$$Pr_{\text{success2}} = Pr_{dk} + Pr_s = \frac{C_{K-\omega}^{\omega}}{C_K^{\omega}} + \sum_{u=1}^{\omega} \left( Pr_k(u) \frac{C_{K-u}^{m+n+\omega}}{C_K^{m+n+\omega}} + Pr_k(u) \sum_{s=0}^{u-1} \frac{C_s^{m+n+\omega} C_{K-u}^{m+n+\omega}}{C_K^{m+n+\omega}} (1/q)^{u-s} \right) \quad (17)$$

### 3.2 机密性评价

PIC 基于 XOR 加密技术,利用栈混洗和数字水印技术来保证网络编码的机密性。对被动攻击者,需要执行下列操作才能获得源信息的内容:1)从截获的传输信息中剔除所有信息认证码;2)重新排列混淆信息,获得编码向量和 XOR 加密的码字;3)猜测 XOR 加密密钥并对加密信息进行异或运算,从而最终获得源信息。

对于第1)个操作,恶意攻击者能够成功剔除所有信息认证码(MACs)的概率是:

$$Pr_1 = 1/C_{m+n+\omega}^{m+n+\omega} \quad (18)$$

当攻击者成功剔除所有 MACs 后,需要对混淆信息进行重新排列。攻击者能够成功获得正确排序的概率是:

$$Pr_2 = (m+n+\omega)/C_{m+n}^{2(m+n)} \quad (19)$$

攻击者为了获取原信息  $M_j^1$ ,还需要知道 XOR 加密密钥。设  $H_E(\cdot)$  表示信息论中的熵函数,则有如下结论:

定理1 由于每个码字都是有限域  $F_q$  中的元素,则有:  
 $H_E(m_j) = H_E(m_j | m_1 \oplus R_i, \dots, m_m \oplus R_i)$  (20)

从上述定理可以看出,攻击者从 XOR 加密后的信息中猜出源信息的概率和从有限域  $F_q$  上直接猜测源信息的概率是一样的。也就是说,通过 XOR 加密后的信息来猜出源信息基本是不可能的。综上,攻击者能够成功获取源信息的概率是:

$$Pr = \frac{m+n+1}{C_{m+n+\omega}^{m+n+\omega} \cdot C_{2(m+n)}^{m+n} \cdot q} \quad (21)$$

对于一般的应用背景,可以设置参数  $\omega = 10, m = 1500, n = 25$  和  $q = 256$ 。根据式(21)可得  $Pr = 1.5329 \times 10^{-941}$ 。Zhang 等<sup>[9]</sup>证明了传统的密码分析技术和启发式搜索技术不能战胜混淆加密技术,攻击者只有使用穷举搜索来猜测源信息,而穷举搜索对于本文所提 PIC 机制是不可能成功的。因此,本文提出的 PIC 机制能提供较好的机密性保证。

### 3.3 隐私评价

PIC 机制能够有效抵抗先进的信息大小相关、时间相关和内容相关等流量分析技术。由于经 PIC 机制处理后的传输信息大小相等,故信息大小相关攻击对 PIC 机制失效。

对时间相关攻击时,每个中间转发节点为了执行网络编码操作都设置一个用于缓冲输入信息的缓存。在有缓冲技术的情况下,在时间  $T_b$  内收到两个及以上输入信息的都会使时间相关攻击失效。假设输入信息的到达服从泊松分布,则网络编码遭受时间相关攻击的概率是:

$$Pr_{tr} = \lambda \cdot T_b \cdot e^{-\lambda T_b} \quad (22)$$

其中参数  $\lambda$  是传输信息到达缓存的平均速率。

对内容相关攻击前,攻击者需要获取足够多的同一批信息。批号通过安全的路由协议进行了隐藏,所以获取信息的批号本身就是一项艰巨的任务。暂时假设攻击者能够获取信息的批号,设在一段时间内经过攻击者的信息都属于批号为  $g$

的同一批信息,且信息的数量为 $l$ 。当攻击者收到一个信息时,它需要判定新截获的信息是否是之前信息的线性组合。上述判定的计算开销是 $O(C_{m+n+\omega}^n(n^3 + nl))$ 次乘法操作。如果攻击者不能获取信息批号,计算开销则变成 $O(C_{\omega n}^n C_{m+n+\omega}^n(n^3 + nl))$ 次乘法运算。可以看出,本文 PIC 机制与显性的 GEV 机制和文献[14]相比,明显增强了网络编码的隐私。

## 4 性能评价

### 4.1 计算开销

**源节点的计算开销** 在生成传输信息阶段,源节点需要执行 $2 + \omega(m + n + \omega)$ 次哈希运算来获得密钥或者哈希链, $2m$ 次入栈或出栈操作来混淆信息和增强机制的机密性和隐私。对于每个传输信息,源节点还需要执行 XOR 异或运算来加密源信息,通过高斯消去求解式(8)来获得所有信息认证码(MACs)。这个过程计算开销是 $O(\omega^3)$ 次乘法运算。另外,源节点要加密 $[p_i, r_i, k_i] (i = 1, 2, \dots, \omega)$ ,需要执行 $\omega$ 次加密运算。综上,每个传输信息的产生需要 $O(\omega^2/n)$ 次哈希运算、 $O(\omega^3)$ 次乘法运算和 $O(\omega/n)$ 次加密运算。

**中间转发节点的计算开销** 如果中间转发节点对收到的 $r$ 个信息只有一个共享密钥,则它可以使用批验证来检查这些信息的完整性。这个过程需要执行1次解密操作, $r + \omega(m + n + \omega)$ 次乘法运算和 $m + n + \omega$ 次哈希运算。如果这些信息通过动态完整性验证,则它们会被随机编码成一个新的传输信息。网络编码操作需要 $r(m + n + \omega)$ 次乘法运算。通过上述分析可以看出中间节点的计算开销较小。

**目标节点的计算开销** 目标节点首先判断输入信息的完整性,当目标节点获取足够多的正确信息后,它开始恢复源信息。具体来说,目标节点需要执行 $\omega$ 次随机函数 $PRF(\cdot)$ 来获取所有信息认证码的位置。并要求求解式(6)来获得源信息,此过程需要执行 $O(n^3)$ 次乘法操作。与乘法运算相比,XOR 解密的异或运算和栈混洗的逆运算等计算开销可忽略不计。

### 4.2 通信开销

PIC 机制引入的通信开销主要由两部分组成:用于完整性验证的信息认证码和用于识别信息认证码位置的标签。对于前者,每个传输信息需要插入 $\omega$ 个信息认证码,这部分的通信开销是 $w \lg q$ 位。对于后者,每个传输信息需要插入 $\omega$ 个位置标签来标识 MACs 的具体位置。每个位置标签由 $p_i, r_i$ 和 $k_i$ 三部分组成。如果假设在加密的过程中,数据大小未发生变化,则这部分的通信开销是 $w(\lg(m + n + \omega) + \lg q + \lg K)$ 位。由于位置标签对于同批传输信息是一样的,因此这些标签只要传输一次,则这部分通信开销可忽略不计。综上,本文所提 PIC 机制的通信开销约为 $w \lg q$ 位,通信开销与源信息比是 $\omega/(m + n)$ 。

PIC 机制与其他经典机制的通信开销对比如表1所示。表1中所有的通信开销的数值结果都是在下列参数设置下计算得到的: $\omega = 1, m = 1500, n = 25, q = 256, l = 12 \sim 22, b = 2, x = 6, t = 5 \sim 10, h = 50, lp = 1500 \times 8$ 。

从表1可以看出,只有文献[11]所提机制在通信开销方面优于本文 PIC 机制。这是因为该机制采用时间同步技术来协助数据完整性验证。在大量与时间同步无关的安全保障机制中,本文 PIC 机制的通信效率最高。此外,表1也定性比较了不同安全机制所关注的安全问题,表明只有本文 PIC 机制能够以较小的通信开销为网络编码同时提供完整性、机密

性和隐私保证。

表1 PIC 与经典安全保障机制的比较

机制	机密性	完整性	隐私性	通信开销	数值结果
PIC	✓	✓	✓	$\omega/(m + n)$	0.065%
文献[11]	✓	×	✓	—	0%
文献[12]	✓	×	✓	$(2h \log n)/lp$	0.17%
文献[13]	×	✓	×	$\frac{l+1}{m+n} + \frac{32l}{q(m+n)}$	0.95% ~ 1.69%
文献[14]	×	✓	×	$b/(m + n)$	0.13%
文献[15]	×	✓	×	$xt/m$	2% ~ 4%

## 5 仿真实验

本章使用 C 语言实现了 PIC 机制的原型系统,并讨论了 PIC 机制在过滤污染信息方面的性能与抵抗共谋攻击<sup>[18]</sup>的能力。实验的参数设置 $m = 1500, n = 25, q = 256$ 。根据文献[19],源节点与目标节点之间的跳数最大为26跳。本章结果均为1000次实验的平均结果。

密钥数 $\omega$ 和密钥池大小 $K$ 对 PIC 机制至关重要,它们直接影响 PIC 机制的计算开销、通信开销和污染信息可以传播的跳数。因此,本章首先讨论 $\omega$ 和 $K$ 之间的关系与影响。不难证明,当 $2\omega > K$ 时,污染信息可被下游第一跳转发节点立即检测出来并从网络中过滤掉;否则,污染信息就会在网络中传播和扩散,并继续参与网络编码和产生污染信息。在图3中,污染信息传播的跳数是参数 $\omega/K$ 的递减函数。当 $\omega/K$ 较小时,污染信息继续传播并污染其他传输机芯的概率较大。此外,增加参数 $\omega$ 将有效减少污染信息继续传播的危险。但是,参数 $\omega$ 负面影响机制 PIC 的通信开销。因此,参数 $\omega$ 的选择是机制性能和开销的权衡。需要强调的是,图3也证明了本文 PIC 机制的优越性。即使 $\omega = 5$ 和 $\omega/K = 0.1$ 时,污染信息传播的跳数都低于1.5跳。通过合理选择参数 $\omega$ 和 $K$ ,PIC 机制可以在一跳范围内有效地检测出污染信息,进而为基于网络编码的数据传输提供完整性保护。

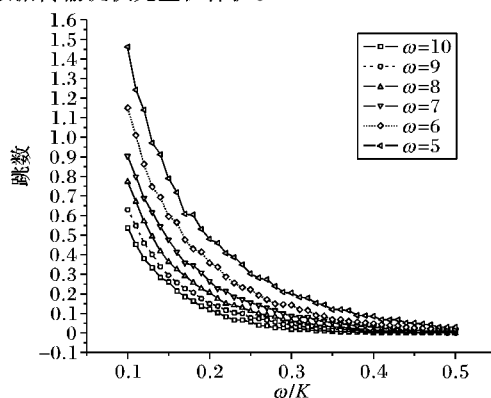
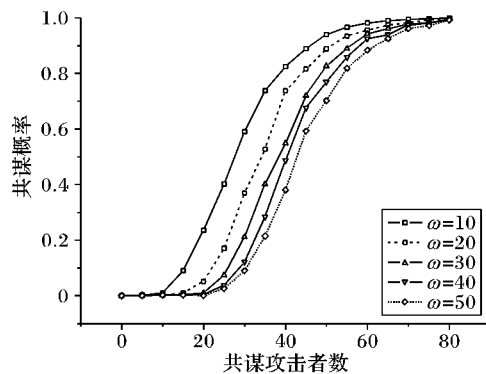
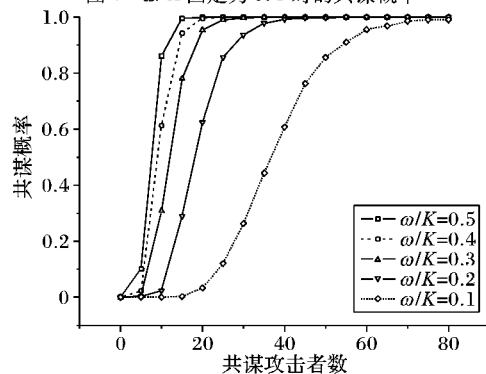


图3 污染检测的性能分析

共谋概率是衡量 PIC 机制在共谋攻击下的鲁棒性的指标,表示一个恶意攻击者能够获取到源节点所有密钥的概率。图4和图5分别分析了共谋概率在 $\omega/K = 0.1$ 和 $\omega = 25$ 时的结果。可以看出共谋概率在参数 $\omega/K$ 固定时是 $\omega$ 的递减函数,可以通过减小参数 $\omega$ 或增大参数 $K$ 来减小共谋概率,进而减小共谋攻击的负面影响。值得关注的是,即使在拥有25个共谋攻击者和密钥池大小为 $K = 100$ 时,共谋概率也仅为0.1。因此,本文提出的 PIC 机制能够有效防御共谋攻击。

图4  $\omega/K$  固定为0.1时的共谋概率图5  $\omega$  固定为25时的共谋概率

## 6 结语

本文提出了一种保证数据完整性、机密性与隐私性的联合保障机制——PIC机制。该机制将数字水印的思想用于完整性验证的同态信息认证码,并将这些信息认证码以安全、保密的方式隐藏在信息码字和编码向量中,从而为网络编码提供完整性、机密性和隐私保护。除了数字水印技术,PIC机制还使用XOR加密技术和栈混洗技术来增强机密性和隐私,在数据传输的过程中,中间转发节点使用部分信息认证码来动态验证传输数据的完整性。仿真结果证明PIC机制能够以较低的计算开销和通信开销抵抗污染攻击、流量攻击和偷听攻击。

### 参考文献:

- [1] YANG L, ZHENG G, HU X. Research on network coding: a survey [J]. Journal of Computer Research and Development, 2008, 45(3): 400–407. (杨林, 郑刚, 胡晓慧. 网络编码的研究进展[J]. 计算机研究与发展, 2008, 45(3): 400–407.)
- [2] GABIDULIN E M, PILIPCHUK N I, HONARY B, *et al.* Information security in a random network coding network [J]. Problems of Information Transmission, 2013, 49(2): 179–191.
- [3] MOHAMMED A H, DAI B, HUANG B, *et al.* A survey and tutorial of wireless relay network protocols based on network coding [J]. Journal of Network and Computer Applications, 2013, 36(2): 593–610.
- [4] SATTARI P, MARKOPOULOU A, FRAGOULI C, *et al.* A network coding approach to loss tomography [J]. IEEE Transactions on Information Theory, 2013, 59(3): 1532–1562.
- [5] ALIREZA S, FARAMARZ H, FARAMARZ F. Network coding for multiple unicast sessions in multi-channel/interface wireless networks [J]. Wireless Networks, 2013, 19(5): 891–911.
- [6] LEI Y, CHENG S, WU C, *et al.* P2P content distribution with network coding [J]. Journal of Computer Research and Development, 2009, 46(1): 108–119. (雷迎春, 程实, 吴产乐, 等. 应用网络编码的P2P内容分发[J]. 计算机研究与发展, 2009, 46(1): 108–119.)
- [7] CHACHULSKI S, JENNINGS M, KATTI S, *et al.* Trading structure for randomness in wireless opportunistic routing [C]// Proceedings of the 2007 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications. New York: ACM Press, 2007. 169–180.
- [8] LIU Z, WU C, LI B, *et al.* UUSEE: large-scale operational on-demand streaming with random network coding [C]// INFOCOM 2010: Proceedings of the 29th IEEE International Conference on Computer Communications. Piscataway: IEEE Press, 2010: 2070–2078.
- [9] ZHANG P, JIANG Y, LIN C, *et al.* P-Coding: secure network coding against eavesdropping attacks [C]// INFOCOM 2010: Proceedings of the 29th IEEE International Conference on Computer Communications, Joint Conference of the IEEE Computer and Communications Societies. Piscataway: IEEE Press, 2010: 2249–2257.
- [10] FAN Y, JIANG Y, ZHU H, *et al.* An efficient privacy-preserving scheme against traffic analysis attacks in network coding [C]// INFOCOM 2009: Proceedings of the 28th IEEE International Conference on Computer Communications, Joint Conference of the IEEE Computer and Communications Societies. Piscataway: IEEE Press, 2009: 2213–2221.
- [11] YU Z, WEI Y, RAMKUMAR B, *et al.* An efficient signature-based scheme for securing network coding against pollution attacks [C]// INFOCOM 2008: Proceedings of the 27th IEEE International Conference on Computer Communications, Joint Conference of the IEEE Computer and Communications Societies. Piscataway: IEEE Press, 2008: 1409–1417.
- [12] ADELI M, LIU H P. On the inherent security of linear network coding [J]. Communications Letters, 2013, 17(8): 1668–1671.
- [13] BHATTAD K, NARAYANAN K. Weakly secure network coding [C]// Proceedings of the 1st Workshop on Network Coding, Theory, and Applications. Riva del Garda: [s. n.], 2005. 1–6.
- [14] FAN Y, JIANG Y, ZHU H, *et al.* Network coding based privacy preservation against traffic analysis in multi-hop wireless networks [J]. IEEE Transactions on Wireless Communication, 2011, 10(3): 834–842.
- [15] ATKINSON M. Generalized stack permutations [J]. Journal of Combinatorics, Probability, and Computing, 1998, 7(3): 239–246.
- [16] ZHAO X, HAO L. Overview of digital watermark [J]. Computer Engineering and Design, 2006, 27(11): 1946–1950. (赵翔, 郝林. 数字水印综述[J]. 计算机工程与设计, 2006, 27(11): 1946–1950.)
- [17] ZHANG Y, CAI Y, LI Z, *et al.* Homomorphic MAC-based scheme against pollution attacks in network coding [J]. Wuhan University Journal of Natural Sciences, 2013, 18(5): 435–442.
- [18] YU Z, WEI Y, RAMKUMAR B, *et al.* An efficient scheme for securing XOR network coding against pollution attacks [C]// INFOCOM 2009: Proceedings of the 28th IEEE International Conference on Computer Communications, Joint Conference of the IEEE Computer and Communications Societies. Piscataway: IEEE Press, 2009: 406–414.
- [19] STEFAN S, WETHERALL D, KARLIN A, *et al.* Network support for IP traceback [J]. IEEE/ACM Transactions on Networking, 2001, 9(3): 226–237.