新型的多路径匿名通信系统

周彦伟[1,2,3],杨 波[1,2,3],张文政[2]
(1. 陕西师范大学计算机科学学院,陕西西安 710062;2. 保密通信重点实验室,四川成都 610041;
3. 中国科学院信息工程研究所信息安全国家重点实验室,北京 100093)

摘 要: 本文基于信息分割理论和网络编码技术提出一种高效的多路径匿名通信模型—ACM-MP,发送者将通信消息经信息分割技术产生分片信息,并对其进行编码处理后沿编码匿名转发网络中不同的匿名通信链路发往接收者,编码匿名转发网络中各节点通过编码地址信息可获知其直接后继节点的具体位置,确保接收者可接收到所有的编码信息.理论分析与仿真结果表明,本文模型具有较高的匿名性、安全性和抗合谋攻击的能力.

关键词: 匿名通信;网络编码;信息分割;多路径传输;合谋攻击
中图分类号: TP393.08 **文献标识码:** A **文章编号:** 0372-2112 (2017)05-1234-06
电子学报 URL: http://www.ejournal.org.cn **DOI:** 10.3969/j.issn.0372-2112.2017.05.029

A New Anonymous Communication System with Multi-Path

ZHOU Yan-wei[1,2,3], YANG Bo[1,2,3], ZHANG Wen-zheng[2]
(1. *School of Computer Science, Shaanxi Normal University, Xi' an, Shaanxi* 710062, *China*;
2. *Science and Technology on Communication Security Laboratory, Chengdu, Sichuan* 610041, *China*;
3. *State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, Beijing* 100093, *China*)

Abstract: This paper proposed a new anonymous communication model with multi-path (ACM-MP) based on information slicing theory and network coding technology. In ACM-MP, the messages are transmitted through different anonymous links after being sliced by information slicing technology. Each node of the transmitted network can obtain address information of its immediate successor node location by coding operation, which can ensure that the receiver receives all the coded information about the message. The results of analysis and simulation suggest that the proposed model has better security, anonymity and better ability to resist conspiracy attack.

Key words: anonymous communication; network coding; information slicing; multi-path transmission; conspiracy attack

1 引言

传统匿名通信系统一般基于可信密钥分发中心进行事先分配密钥或密钥参数,或者基于密钥基础设施进行会话密钥的协商[1~5],因此传统匿名通信系统必须有可信的第三方服务器参与会话密钥的协商或者事先分发密钥等,这将带来密钥管理的复杂性,也增加了匿名通信系统的实施代价.针对上述不足,文献[6]提出一种具有高匿名性的编码匿名通信系统,转发网络节点的编码传输机制在提高系统匿名性、安全性和保密性的同时增强了抗合谋攻击的能力;然而中继节点对通信消息的编码操作却增加了系统的通信时延;并且每个中继节点需存储 m 个后继节点(每条链路各一个)的位置信息,增加了节点的存储开销.因此,本文将设计新型的编码多路径匿名通信系统 ACM-MP,在未降低系统匿名性的前提下,提高系统的执行效率(如节点的存储效率),降低通信时延.

文献[7]详细分析了洋葱路由[2]、Tor[3]、混淆模型[4]、Crowds[5]、DC-Net[8]、Sherwood[9]和 Tarzan[10]等传统匿名通信系统的优缺点,篇幅所限,本文不再赘述;对

收稿日期:2015-10-31;修回日期:2016-04-16;责任编辑:李勇锋
基金项目:国家自然科学基金(No.61572303, No.61272436, No.61402275);中国科学院信息工程研究所信息安全国家重点实验室基金(No.2015-MS-10);保密通信重点实验室基金(No.9140C110206140C11050);陕西省重点科技创新团队项目(No.2014KTC-18);陕西师范大学优秀博士论文基金(No.X2014YB01)

网络编码[11,12]、信息分割等基础技术同样不再赘述,详细介绍见文献[6,7,13].

2　高效的多路径匿名通信模型

如图 1 所示,ACM-MP 的基本思想是:首先发送者选择若干可信节点构造编码匿名转发网络(可信计算技术[14]可确保节点的可信性);随后基于信息分割技术将通信消息分割成多个分片信息,通过可逆的编码系数矩阵对分片信息编码生成编码信息;接着将编码系数矩阵以列为单位划分成编码系数分量,发送者将编码信息和编码系数分量分别沿编码匿名转发网络中不同的匿名通信链路进行传输;各中继节点通过编码地址信息可获知直接后继节点的位置信息,并将收到的匿名通信消息转发至直接后继节点;最后,所有的编码信息和编码系数分量在接收者处会合,通过解码编码信息还原发送者的通信消息.

(1)本文使用的相关定义表述如下:

消息分割函数 $f(M,L,m)$:可将长度为 L 的消息 M 分割成 m 份分片信息 $M_1,\cdots,M_m$,即

$$f(M,L,m)=(M_1,M_2,\cdots,M_m);$$

同时,存在 $f(\ )$ 的可逆函数 $f^{-1}(\ )$,能将分片信息 $M_1,\cdots,M_m$ 还原出原始消息 M,即

$$M=f^{-1}(M_1,M_2,\cdots,M_m).$$

(2)本文使用的相关符号定义如下:

中继节点 $N_{(i,j)}$ $(i=1,\cdots,m;j=1,\cdots,n)$ 表示编码匿名转发网络中第 i 条匿名通信链路上的第 j 个节点;其中 m 表示编码匿名转发网络中的链路数;n 表示编码匿名转发网络中每条链路的节点数.

E_i 表示发送者对分片信息 M_i 编码后的编码信息,$\boldsymbol{E}$ 为编码信息 E_i 组成的编码信息矩阵;G_i 表示由编码系数矩阵 $\boldsymbol{G}$ 中第 i 列元素组成的编码系数分量,即编码匿名转发网络中第 i 条匿名链路中传输的消息为 $\{E_i,G_i\}$.

函数 $H(\ )$ 为抗碰撞的哈希函数.

(3)本文相关假设条件表述如下:

假设编码匿名转发网络出现网络差错的概率可忽略,即接收者能够完全接收发送者发送的所有编码信息 E_i 和编码系数分量 G_i.

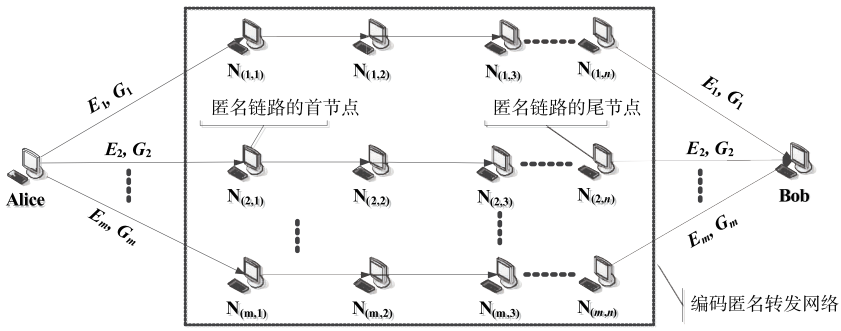


图1　编码匿名转发网络

2.1　消息编码

Alice 从节点服务器选择下载 $m\times n$ 个可信节点建立如图 1 所示的编码转发网络;生成可逆编码系数矩阵 $\boldsymbol{G}$,即 $\boldsymbol{G}=\begin{bmatrix} R_1^1 & R_1^2 & \cdots & R_1^2 \\ R_2^1 & R_2^2 & \cdots & R_2^m \\ \vdots & \vdots & \vdots & \vdots \\ R_m^1 & R_m^2 & \cdots & R_m^m \end{bmatrix}$.

Alice 通过消息分割函数 $f(M,L,m)$ 将长度为 L 的消息 M 分割成 m 份分片信息 $M_1,\cdots,M_m$,其中 M_m 的长度为 $L-(m-1)\lfloor\frac{L}{m}\rfloor$,而其他分片信息的长度均为 $\lfloor\frac{L}{m}\rfloor$;根据式(1)进行编码运算产生编码信息 E_i,得到编码信息矩阵 $\boldsymbol{E}=[E_1,\cdots,E_m]$.

$$E_i=R_i^1\oplus M_1+R_i^2\oplus M_2+\cdots+R_i^m\oplus M_m \qquad (1)$$

匿名消息生成的伪代码表示如算法 1 所示.

算法 1

Begin

1:初始化参数$(m,n,M,\cdots)$;

2:分割消息$f(M,L,m)=(M_1,M_2,\cdots,M_m)$;

3:生成可逆的编码矩阵 $\boldsymbol{G}$.

4:For $i=1$ to $i=m$ //生成编码信息 E_i

$$E_i=R_i^1\oplus M_1+R_i^2\oplus M_2+\cdots+R_i^m\oplus M_m;$$

End For

For $i=1$ to $i=m$ //生成编码系数分量

$$G_i=\begin{bmatrix} R_1^i \\ R_2^i \\ \vdots \\ R_m^i \end{bmatrix};$$

End For

5:$IP_{\text{源地址}}=IP_{\text{Alice}}\oplus IP_{\text{Bob}}\oplus H(ID_{\text{Bob}})$;

6:For $i=1$ to $i=m$ //匿名通信数据的封装

Init(SP_i);//初始化目的地址堆栈

SP_i. $Push(IP_{N(i,m)} \oplus IP_{Bob} \oplus H(ID_{N(i,m)}))$;

For $j=m$ to $j=2$//依次建立每条匿名链路的地址堆栈

SP_i. $Push(IP_{N(i,j-1)} \oplus IP_{N(i,j)} \oplus H(ID_{N(i,j-1)}))$;

End For

$IP^i_{\text{目的地址}} = SP_i$;//第 i 条链路的目标地址堆栈

$Data_i = SID_i \| G_i \| E_i$

$Message_i = IP_{\text{源地址}} \| IP^i_{\text{目的地址}} \| Data_i$;

End For

7:发送 $Message_i$ 给节点 $N_{(i,1)}$.

2.2 中继节点转发

各中继节点 $N_{(i,j)}$ 依次取出目的地址堆栈的栈顶元素与其 *IP* 地址 $IP_{N_{(i,j)}}$ 和自己的身份哈希值 $H(ID_{N_{(i,j)}})$ 进行异或运算后,可获知匿名通信数据下一跳节点 $N_{(i,j+1)}$ 的 *IP* 地址 $IP_{N_{(i,j+1)}}$,然后将匿名通信数据填充后转发到该中继节点;转发过程中,填充机制可保证通信消息的大小保持不变,防止敌手进行流量分析攻击.

2.3 解码策略

接收者 Bob 通过以下步骤的操作恢复发送者 Alice 的通信消息:

(1) 重构编码系数矩阵 $\boldsymbol{G}$ 和信息矩阵 $\boldsymbol{E}$;即有

$$\boldsymbol{G}=\begin{bmatrix} R_1^1 & R_1^2 & \cdots & R_1^2 \\ R_2^1 & R_2^2 & \cdots & R_2^m \\ \vdots & \vdots & \vdots & \vdots \\ R_m^1 & R_m^2 & \cdots & R_m^m \end{bmatrix} \text{和 } \boldsymbol{E}=[E_1,E_2,\cdots,E_m].$$

(2) 验证 $\boldsymbol{G}$ 是否可逆,若 $\boldsymbol{G}$ 是不可逆矩阵,拒绝接收该消息,否则进行下述操作;

(3) 计算编码系数矩阵 $\boldsymbol{G}$ 的逆矩阵 $\boldsymbol{G}^{-1}$,然后通过式(2)计算获得 m 份分片信息 $M_1,\cdots,M_m$.

$$\begin{aligned}(M_1,\cdots,M_m) &= \boldsymbol{G}^{-1} \oplus \boldsymbol{E}^{\mathrm{T}} \\ &= \begin{bmatrix} R_1^1 & R_1^2 & \cdots & R_1^2 \\ R_2^1 & R_2^2 & \cdots & R_2^m \\ \vdots & \vdots & \vdots & \vdots \\ R_m^1 & R_m^2 & \cdots & R_m^m \end{bmatrix}^{-1} \oplus \begin{bmatrix} E_1 \\ E_2 \\ \vdots \\ E_m \end{bmatrix} \end{aligned} \tag{2}$$

其中 T 表示矩阵的转置.

(4) 根据 m 份分片信息 $M_1,\cdots,M_m$ 恢复出原始通信消息 M,即 $M=f^{-1}(M_1,M_2,\cdots,M_m)$.

3 模型分析

编码的可行性与文献[6]相类似,本文不再赘述.

3.1 匿名性分析

3.1.1 位置匿名

匿名链路中各中继节点用其 IP 地址和身份哈希值与目标地址堆栈的栈顶元素进行异或运算后,仅能获知下一跳节点的 IP 地址,并不能获知其他节点的地址信息,更无法获知发送者和接收者的地址信息;即便链路节点从目的地址堆栈中取出多个元素,由于其无法确定其他的链路节点,也无法获知具体的节点地址信息;即使外部观察者侦听到通信数据,而攻击者所看到数据报中的地址并不是发送者和接收者的地址,保证了发送者和接收者的位置匿名性.

3.1.2 身份匿名

匿名通信数据不包含发送者和接收者的身份等相关信息,并且匿名通信消息的安全传输保证了发送者和接收者的身份匿名性. 即使编码匿名转发网络中存在泄密节点,该泄密节点仅能获悉其直接后继节点的身份哈希值,无法获知发送者和接收者的任何身份信息.

3.1.3 通信匿名

匿名通信链路的节点 $N_{(i,j)}$ 仅能获知节点 $N_{(i,j+1)}$ 的地址信息;由于无法获知完整的编码系数矩阵 $\boldsymbol{G}$,因此第三方难以推断发送者与接收者间的通信模式. 在 ACM-MP 中,编码系数矩阵 $\boldsymbol{G}$ 的安全性及匿名链路节点的可信性保证了发送者和接收者间通信的匿名性.

3.1.4 匿名度仿真

定义 匿名度 D 为编码消息 E_i 传输时被泄密者合谋解码的概率. 即合谋解码的概率越低,系统匿名性越强.

假设由 $m\times n$ 个中继节点组成的编码匿名转发网络中存在泄密节点的个数为 k;则编码匿名转发网络中转发节点为攻击节点的概率为 $P=\dfrac{k}{m\times n}$. 本节基于文献[5]提出的针对共谋攻击的匿名度分析方法对 ACM-MP 的匿名性进行分析.

(1) $k<m$

发送者对 m 份分片信息 M_i 分别进行编码产生 m 个不同的编码信息 E_i 并沿 m 条匿名链路发送给接收者,攻击者要解码所有的编码信息 E_i,则每条链路上至少有一个泄密者. 因此当 $k<m$ 时,$D=0$,ACM-MP 具有绝对匿名性.

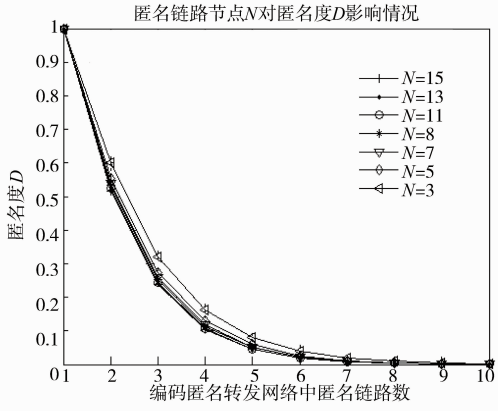
(2) $k=m$

编码匿名转发网络由 m 条长度为 n 的匿名通信链路组成,当泄密节点数 $k=m$ 时,泄密者合谋获得通信息的概率为$\dfrac{n^m}{C_{mn}^m}$,即 $D=\dfrac{n^m}{C_{mn}^m}$.

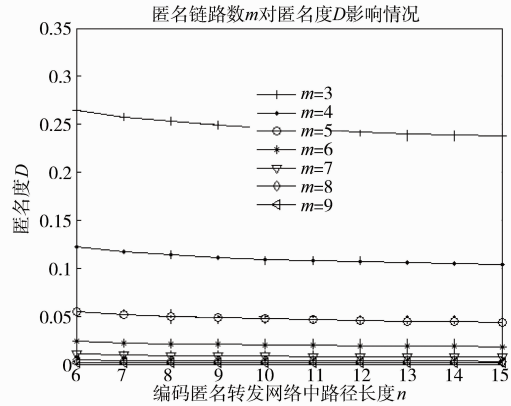
当泄密节点数与分片信息数目相同时,攻击者若能解码编码信息 E_i 还原通信消息 M,则各泄密者在编码网络中的位置关系满足下列条件:

(a)每一条路径上有且只有一个泄密者;

(b)这 m 个泄密节点间必须相互通信且进行数据共享.



（a）n对匿名度D的影响（k=m）



（b）m对匿名度D的影响（k=m）

图2　m和n对匿名度D的影响（$k=m$）

m 个泄密者合谋获得完整通信消息 M 的概率为 $\frac{n^m}{C_{mn}^m}$，即 $D=\frac{n^m}{C_{mn}^m}$.

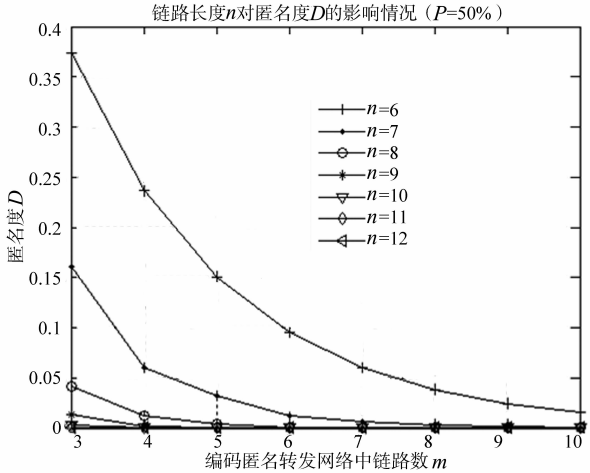
由图 2(a)可知，匿名度 D 随 n 和 m 的增大而增加，受匿名链路数 m 的影响较大，基本不受链路长度 n 值变化的影响. 由图 2(b)可知，当 $m=5$，$n=8$ 时匿名度大小趋近于 0，接近绝对匿名的等级，且此时匿名度受 m 和 n 的影响较小.

(3) $k>m$

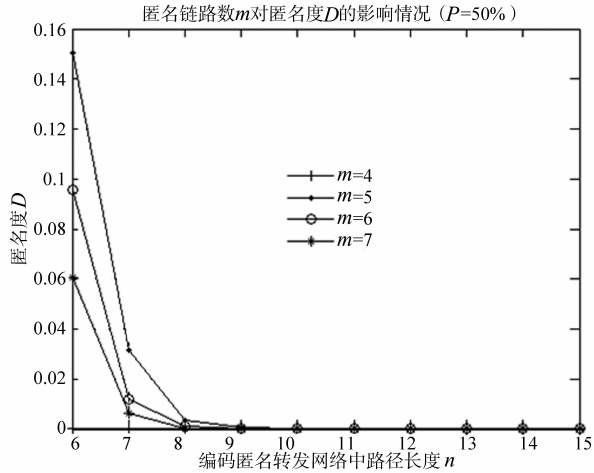
当分处在 m 条路径上的 m 个攻击者进行合谋攻击时，才可解码编码信息 E_i 还原通信消息 M. 则当 $k>m$ 时，泄密者合谋获得通信消息 M 的概率小于 $\frac{C_k^m n^m}{C_{mn}^k}$（小于是因为包含了重复情况），因此 $D<\frac{C_k^m n^m}{C_{mn}^k}$. 在链路数 m 和链路长度 n 已知时，k 值越大，匿名度 D 越大，则匿名性越弱.

图 3(a)所示为当 $P(P=50\%)$ 确定时，在不同的匿名链路长度 m 情况下，D 与 n 间的关系，当 m 和 P 一定时，n 值增加 D 变小，系统匿名性增强. 图 3(b)所示为当 $P(P=50\%)$ 确定时，在不同的匿名链路长度 n 情况下，D 与 m 间的关系，当 n 和 P 一定时，m 值增加 D 变小，系统匿名性增强.

如图 4(a)所示为当 $m(m=5)$ 确定时，在不同的匿名链路长度 n 情况下，D 与 P 间的关系，在 m 和 P 一定时，n 值增加 D 变小，系统匿名性增强. 如图 4(b)所示，当 $m=5$，$n=8$ 时，即使编码转发网路中 50% 的节点进行合谋攻击，ACM-MP 依然能够保持 D 大小趋近于 0，接近绝对匿名的等级.



（a）n对匿名度D的影响（K>m,P=50%）



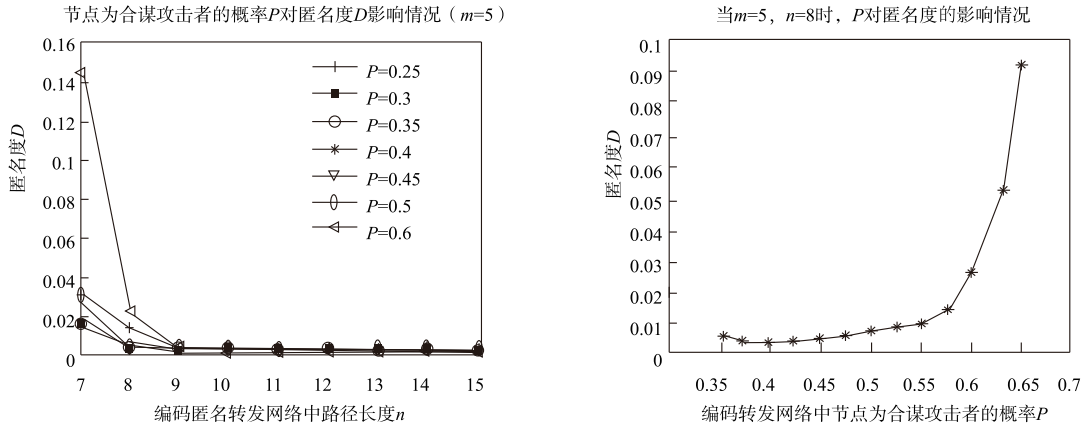
（b）m对匿名度D的影响（K>m,P=50%）

图3　m和 n 对匿名度 D 的影响（$K>m,P=50\%$）

3.2　安全性分析

(1) 消息的编码传输

发送者将通信消息分割为多个信息片后，由随机产生的编码系数矩阵对其进行编码处理，生成编码信



(*a*) *P*对匿名度*D*的影响（$k>m$，$m=5$） (*b*) *P*对匿名度*D*的影响（$k>m$，$m=5$，$n=8$）

图4 *P*对匿名度*D*的影响($k>m$)

息,采用消息分割和编码处理的匿名通信方式,可以有效的增强通信消息的安全性及匿名性. 若通信链路中某个节点被攻击者所控制,由于该节点仅知道部分编码系数,攻击者也无法解码编码信息,更无法获知原始的通信消息及其消息的源地址和目标地址,除非攻击者在每条匿名链路中至少控制一个节点,而这在现实中是很难实现的.

(2) 抗合谋攻击

编码匿名转发网络中,各中继节点仅掌握直接后继节点和前驱节点,不可能掌握转发网络中的其他节点,外部观察者即使侦听到通信数据,而他们所看到的地址信息并不是发送者和接收者的地址,因此攻击者在现实网络环境中很难通过控制链路节点对 ACM-MP 进行攻击;由匿名性分析可知,即使编码匿名转发网络中的中继节点进行合谋攻击,其成功的概率是可忽略的.

3.3 效率分析

3.3.1 编码代价

为减少发送者的编码代价,提高编码效率,发送者在进行匿名通信数据封装时,主要以计算代价较低的异或运算为主;由匿名性分析可知,建立编码匿名转发网络时,选取参数 $m=5$ 和 $n=8$,发送者即可以最低的编码代价实现接近绝对匿名的通信需求.

3.3.2 转发网络的重复使用性

发送者完成编码匿名转发网络的建立后,后续无需重复建立;在每次匿名通信时,只需根据接收者的地址封装匿名通信数据,使用现有的转发网络进行匿名通信. 为防止编码匿名转发网络中合谋节点数量的增多(合谋节点的增多会导致 $P>50\%$),削弱匿名通信的效果,发送者可定期建立新的转发网络;更新时间间隔越短,匿名性越强,但会增加执行负载;更新时间间隔越长,匿名性越弱,但执行负载较低. 因此发送者可根据实际环境的匿名性需求,设置转发网络的更新间隔.

3.3.3 通信前的协商开销

由于 ACM-MP 无需进行密钥协商,因此在匿名通信之前无需与中继节点进行协商;匿名通信数据封装时,根据匿名链路中中继节点的顺序,依次从后至前基于相关信息(如:身份哈希值、地址等)即可完成匿名通信数据的封装.

3.3.4 通信的传输开销

匿名通信过程的传输开销主要由中继节点的解码开销所决定. 每个节点只需进行一次解码操作,且解码操作以运算代价较低的异或运算为主;同时不同的通信数据解码过程基本类似,并且解码操作的分散执行,大幅度降低了接收者的解码开销.

4 结束语

新型的多路径匿名通信系统 ACM-MP 基于编码技术增强攻击者的窃听难度,保证了通信消息传输的不可追踪性. 在无信任中心的网络环境下,无需无加/解密计算即可完成匿名路径的建立,通过发送者和编码匿名转发网络对通信消息的编码操作,有效地实现通信消息传输的匿名性,不仅提供发送者和接收者的匿名以及通信消息的保密性,并且克服了传统的匿名通信方法中以提高方案复杂度来换取匿名性的局限性.

参考文献

[1] Shields C, Levine B N. A protocol for anonymous communication over the internet [A]. Proceedings of the ACM Conference on Computer and Communications Security [C]. New York: ACM, 2010. 33 - 42.

[2] Goldschlag D. Onion routing for anonymous and private internet connections [J]. Communications of the ACM,

1999,42(2):39 – 41.
[3] Dingledine R, Mathewson N, Syverson P. Tor: the second-generation onion router [J]. Journal of the Franklin Institute, 2004, 239(2): 135 – 139.
[4] Chaum D. Untraceable electronic mail, return addresses, and digital pseudonyms [J]. Communications of the ACM, 1981, 24 (2): 84 – 88.
[5] Reiter M K, Rubin A D. Crowds: anonymity for Web transactions [J]. ACM Transactions on Information & System Security, 1997, 1(1): 66 – 92.
[6] 周彦伟,杨波,吴振强,等. 基于网络编码的匿名通信模型[J]. 中国科学 信息科学,2014,44(12):1560 – 1579.
ZHOU YanWei, YANG Bo, WU ZhenQiang, et al. Anonymous communication model based on network coding [J]. China Science, Information Science, 2014, 44(12): 1560 – 1579. (in Chinese)
[7] 段桂华,王伟平,王建新,等. 一种基于多路径网络编码的匿名通信机制[J]. 软件学报, 2010, 21(9): 2338 – 2351.
DUAN GuiHu, WANG WeiPing, WANG JianXin, et al. Anonymous communication mechanism with multi-paths network coding [J]. Journal of Software, 2010, 21(9): 2338 – 2351. (in Chinese)
[8] Chaum D. The dining cryptographers' problem: Unconditional sender and recipient untraceability [J]. Journal of Cryptology, 1988, 1(1): 65 – 75.
[9] Sherwood B R, Bhattacharjee B, Srinivasan A. P5: A protocol for anonymous communications [A]. Proceedings of the IEEE Symposium on Security & Privacy [C]. Berkeley: IEEE, 2002. 58 – 70.
[10] Freedman MJ, Morris R. Tarzan: A peer-to-peer anonymizing network layer [A]. Proceedings of the ACM Conference on Computer and Communications Security [C]. Washington: ACM, 2002. 193 – 206.
[11] Yu M, Sadeghi P, Aboutorab N. Performance characterization and transmission schemes for instantly decodable network coding in wireless broadcast [J]. Journal on Advances in Signal Processing, 2015, 2015(1): 1987 – 1996.
[12] Koetter R, Médard M. An algebraic approach to network coding [J]. IEEE/ACM Transactions on Networking, 2003, 11(5): 782 – 795.
[13] 吴振强,马亚蕾. 编码混淆:一种新型匿名通信模型[J]. 武汉大学学报(理学版). 2011, 57 (5): 401 – 407.
Wu ZhenQiang, Ma YaLei. A novel anonymous communication model: coding mix [J]. Journal of Wuhan University: Natural Science Edition, 2011, 57 (5): 401 – 407. (in Chinese)
[14] 冯登国,秦宇,汪丹,等. 可信计算技术研究[J]. 计算机研究与发展,2011,48(8):1332 – 1349.
Feng Dengguo, Qin Yu, Wang Dan, et al. Research on trusted computing technology [J]. Journal of Computer Research and Development, 2011, 48 (8): 1332 – 1349. (in Chinese)

作者简介

周彦伟　男,1986 年生于甘肃通渭. 陕西师范大学计算机科学学院博士生. 研究方向为密码学、匿名通信技术.
E-mail: zhouyanwei1986@163.com

杨　波(通信作者)　男,1963 年生于陕西富平. 教授,博士生导师,陕西省"百人计划"特聘教授. 研究方向为密码学、信息安全.
E-mail: byang@snnu.edu.cn

张文政　男,1966 年生. 研究员. 主要研究领为密码学、信息安全.