ISSN 1000-9825, CODEN RUXUEW
Journal of Software, Vol.21, No.9, September 2010, pp.2338−2351
doi: 10.3724/SP.J.1001.2010.03612

E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel/Fax: +86-10-62562563

一种基于多路径网络编码的匿名通信机制*

段桂华, 王伟平+, 王建新, 杨路明

(中南大学 信息科学与工程学院,湖南 长沙 410083)

Anonymous Communication Mechanism with Multi-Paths Network Coding

DUAN Gui-Hua, WANG Wei-Ping+, WANG Jian-Xin, YANG Lu-Ming

(School of Information Science and Engineering, Central South University, Changshan 410083, China)

+ Corresponding author: E-mail: wpwang@mail.csu.edu.cn

Duan GH, Wang WP, Wang JX, Yang LM. Anonymous communication mechanism with multi-paths network coding. *Journal of Software*, 2010,21(9):2338−2351. http://www.jos.org.cn/1000-9825/3612.htm

Abstract: This paper first proposes a new information slicing and transmitting method ITNC (information slicing and transmitting with multi-path network coding) with multi-paths network coding. Next, a novel anonymous communication mechanism AC-ITNC (anonymous communication mechanism based on ITNC) without key infrastructure, which is based on ITNC, is presented. In the new mechanism, the anonymous path setup information is sliced into pieces, and every piece is coded by the random coding coefficient. The coding coefficient and coded information pieces are delivered along different paths, which make the anonymous paths be set up in the case of non-cryptographic scheme. Theoretical analysis and simulation results show that AC-ITNC can significantly improve resistance against conspiracy attacks in an anonymous communication system within complete distributed environment without key infrastructure.

Key words: anonymous communication; network coding; multi-paths; conspiracy attack; network security

摘 要: 提出了基于多路径网络编码的信息分割传输策略 ITNC(information slicing and transmitting with multi-path network coding),并基于 ITNC 提出了一种无需密钥基础设施的匿名通信机制 AC-ITNC(anonymous communication mechanism based on ITNC).该机制将建路信息分割后编码传送,每个编码节点都对转发信息进行再次随机编码,编码系数与编码信息沿不同编码路径分离传输,从而可以在不需要密钥机制的情况下进行匿名建路.理论分析与仿真结果表明,AC-ITNC 与单纯依靠源节点信息分割而中间节点不编码的匿名建路机制相比,显著提高了匿名系统的抗合谋攻击能力.

关键词: 匿名通信;网络编码;多路径;合谋攻击;网络安全

中图法分类号: TP309 **文献标识码**: A

* Supported by the National Natural Science Foundation of China under Grant Nos.60873265, 60673164 (国家自然科学基金); the Hu'nan Provincial Natural Science Foundation of China under Grant No.06JJ10009 (湖南省自然科学基金); the Specialized Research Fund for the Doctoral Program of Higher Education of China under Grant No.20060533057 (高等学校博士学科点专项科研基金); the New Century Excellent Talents in University of China under Grant No.NCET-05-0683 (新世纪优秀人才支持计划)

Received 2008-09-01; Revised 2009-02-13; Accepted 2009-03-31

匿名通信技术是指在通信过程中将通信关系隐藏,使攻击者无法直接获知或推知双方的通信关系或通信一方的身份信息.Chaum 于 1981 年提出 Mix-Net 概念[1],之后,基于 Mix-Net 出现了很多改进的匿名通信协议和系统原型,如 Onion Routing[2],Tor[3],Crowds[4],P5[5],Tarzan[6],Morphmix[7]等等.

典型的匿名通信路径一般由多个串行的匿名代理组成.匿名通信过程经历两个阶段:一是路径建立阶段,即让每个匿名路径上的中间节点知道其后继节点和前驱节点,从而建好匿名转发路径;二是数据传输阶段,发送者将数据沿着建好的路径传送给接收者.

其中,要做到匿名的关键包括两个方面:一是在路径建立阶段要保证建路信息的安全性,即只有路径上的节点才能够知道其下一跳和前一跳的地址,其他节点不能知道;二是在匿名数据传输阶段,要保证信息传输的不可追踪性,即数据传输在经过匿名路径上的节点时,信息的入出关系无法关联,即窃听者不能依据节点入出的信息推断出匿名路径.

针对这两个关键问题,典型匿名通信机制一般都基于密钥基础设施,即需要信任中心为节点事先分配密钥或密钥参数.这类系统中,一般在建路阶段依靠事先已有的公钥或共享密钥加密传输下一跳路径信息和协商共享会话密钥.在数据传输阶段,依靠已协商好的共享会话密钥对收到的信息进行加解密操作,从而做到入出信息的无法关联.

在洋葱路由(onion routing)[2]中,建路阶段采用源路由的方式,利用公钥嵌套加密机制,将下一跳地址和与源节点间的共享密钥封装在节点的公钥加密结构中,使得路径上任意节点可以通过私钥解密来获得下一跳地址和共享密钥,同时无法解密用其他节点公钥加密的路径信息,从而保证建路信息的安全性.但这种机制中,匿名通信需要公钥机制,增加了系统的复杂性,同时,建路过程中嵌套公钥加密的计算代价也比较大.

针对公钥加密代价大的问题,Tor[3]使用了 Diffie-Hellman 密钥交换协议来协商源与路径上的节点之间的密钥,并利用协商好的密钥加密传输包含下一跳节点的建路信息.为了防止 Diffie-Hellman 密钥交换中的中间人攻击,Tor 在传输密钥协商参数时仍然需要采用了路径上节点的公钥加密以完成中间节点的认证.文献[8]中,Kate 等人提出了基于对密钥的 Diffie-Hellman 密钥协商方案.这种方案不需要使用公钥认证机制,但需要可信的服务器来分配对密钥.在文献[9]中,Øverlier 等人提出利用预先分配的 Diffie-Hellman 密钥参数来建立会话密钥.在这种机制中,需要一个服务器来存储和更新所有节点的密钥参数,节点可以从服务器上获取其他节点的由服务器签名的密钥参数.后两者尽管可以避免公钥机制,利用预分配技术来完成密钥的生成,但需要可信的第三方服务器来管理密钥参数的分配和更新,同样不可避免地带来了密钥管理的开销.

Crowds 系统[4]采用下一跳路由的方式,路径上的任意节点知道通信的目的节点.在建路过程中不需要已有密钥,而由中间节点以一定概率来选择传输给下一跳节点或者接收者.这种机制采用下一跳路由的方式,中间节点需要知道接收者,不能做到接收者匿名;同时在 Crowds 中,仍然需要一个可靠的密钥分发中心来分配所有节点之间的通信密钥,数据传输阶段采用路径上节点之间的通信密钥加密.

DC-Net[10]假设任意节点间都存在共享密钥,发送者将信息与所有共享密钥异或后发送,其他节点直接发送共享密钥的异或值,由网关收集所有信息后还原成发送信息,从而来获得发送者匿名.但该机制中任意时刻只能有一个节点发送,需要大量的额外带宽.

Sherwood 等人在 P5 协议[5]中采用了公钥机制,假设任意通信成员之间可以通过带外方式获得对方的公钥,依据公钥来映射广播信道,基于局部共享广播信道实现了发送者匿名和接收者匿名.但这种机制中需要预先分配好的公钥,而且局部广播也会带来较大的传输代价.

Tarzan[6]是 2002 年由 Freedman 和 Morris 提出的一个基于 P2P 的 IP 层匿名网络.数据包的转发路径是基于 IP 层的传输虚电路,由匿名通信发送者来建立.发送者首先确定转发节点个数 l,并利用节点的公钥以逐跳方式交换转发对称密钥,每个转发节点保存前一跳节点和后一跳节点的地址以及对称密钥、流标识符,形成整个转发路径.该机制同样需要公钥机制来完成匿名路径的建立.

从上述典型的匿名通信机制可以看出,利用已有密钥基础设施的方法在安全性方面依赖于已分配密钥本身的安全性,具有很好的信息保密性.但需要预先存在的密钥,系统中必须要有一个可信的密钥分发中心,这带

来了密钥管理的复杂性,代价比较大.同时,依赖于密钥基础设施的匿名机制不能适用于纯分布式的无信任中心的系统结构.随着无线网络与 P2P 技术的发展,越来越多的网络应用都基于无信任中心的分布式环境,因此,研究在没有密钥基础设施环境下的匿名通信机制具有很重要的意义.

Katti 等人在文献[11,12]中提出利用信息分割传输的方式将路由信息和共享密钥信息分割后沿着不相交的路径发送给路径上的每一个节点,从而使每个中间点可以得到下一跳地址和共享密钥信息.在下文中,我们称这种方法为源信息分割策略 SIS(source information slicing).这种方案实现了在没有密钥基础设施环境下的匿名传输路径的建立,但是只要每条建路信息转发路径上任意存在一个泄密节点,这些泄密点合谋就可以恢复出相应的信息内容,获得匿名路径信息和共享密钥信息.针对该方案抗合谋攻击能力弱的缺陷,Katti 等人提出路径上的每一个节点改变转发数据包的流标识,以避免攻击者依据流标识识别同一数据流.但尽管进行了流标记的修改,泄密节点仍然可以用信息片尝试合并的方法来恢复出转发的建路信息.由于该方法中仅仅是由源节点对要传送的信息进行了分割,尝试合并十分容易,因此仍然无法提高该机制抗合谋攻击的能力.

针对上述情况,本文首先提出了基于多路径网络编码的信息分割传输策略 ITNC(information slicing and transmitting with multi-path network coding).在该策略中,信息被分割并编码后传输,每个中间转发节点对信息进行随机异或运算,同时将编码后的信息和编码系数分离在不同路径上传输,从而实现了在无须加密机制情况下的节点间传输信息的安全性,增强了信息传输的抗合谋攻击能力.然后,我们基于 ITNC 提出了一种新的匿名通信机制 AC-ITNC(anonymous communication mechanism based on ITNC),该机制利用 ITNC 信息分割传输方式向匿名路径上的节点传输下一跳路由和共享密钥信息,编码路径上的任意单个中间点无法依据转发的信息解码出建路信息.只有编码转发网络中路径上符合某些位置关系的多个节点同时泄密,才可能导致匿名性被破坏.与单纯依靠源节点信息分割而中间节点不编码的匿名建路机制 SIS 相比,该策略显著提高了匿名系统的抗合谋攻击能力.

本文的主要创新点在于:

(1) 将网络编码引入匿名通信的建路阶段.由于中间节点独立选取编码系数,使得传输的建路信息经过一个节点后信息包发生很大变化,无法依据建路阶段的信息包进行匿名路径的追踪;

(2) 将编码后的建路信息片与编码系数分离传输.该方法使得只有多个泄密节点具有特殊位置时才能解码得到建路信息,提高了无密钥匿名建路的抗合谋攻击能力,从而增强了匿名性.

本文第 1 节描述基于多路径网络编码的信息分割传输策略 ITNC 的设计思想,阐述两个节点之间通信内容的信息分割、编码、传输和解码的过程.第 2 节对编码信息的可解性和信息内容的安全性进行理论分析和证明.第 3 节将 ITNC 应用到源路由的匿名通信中,提出一种新的基于多路径网络编码的匿名通信机制,并详细描述编码转发网络中节点的编码与转发过程.第 4 节对该机制的匿名性能进行分析.最后给出结论.

1 基于多路径网络编码的信息分割传输策略 ITNC

在传统网络中,中间节点一般仅具有转发数据的作用.Ahlswede 等人于 2000 年提出了网络编码(network coding)的思想[13],改变了这一状况,中间节点能够对需要转发的信息进行编码.在网络编码实际应用方面,Chou 等人提出了第一个实用的网络编码方案[14],由各节点分布式随机产生编码系数,将编码系数矩阵携带在数据包中一起进行传输,目的节点依据接收到的数据包解码信息的内容.近年来,网络编码的研究目标主要在于提高网络吞吐率、能量利用效率和安全性等方面[15−17].

利用网络编码中节点能够对转发信息进行重新编码的特点,我们首先提出了一种基于网络编码的信息分割编码传输策略 ITNC,旨在实现无加密情况下信息传输内容的保密性.ITNC 的基本思想是:首先选择若干个节点构成多路径转发网络;发送者将发送信息分割成信息片,分别选择随机的编码系数对信息片进行编码,将编码后的信息片分别沿着不同编码路径进行传输,同时将编码系数按位分割成编码系数分量,分解到不同编码路径传输;传输路径上的中间节点随机选择编码系数对收到的信息片进行编码后转发,同时也将编码系数分量分解到不同路径进行传输;最后,所有的信息片和编码系数分量在目的节点会合,目的点可以解码出该信息.在该策

略中,由于单个中间节点只能获得部分编码信息和编码系数分量,无法解码还原出信息内容.

1.1 ITNC的编码转发网络

假设源节点(发送者)为 S,目的节点(接收者)为 X,S 要将信息 M 安全地传送给 X.

如图 1 所示,S 选择 $m\times n$ 个中间节点,建立 m 条长度为 n 的路径,即每条路径包含 n 个节点,第 i 条路径上的第 j 个点记作 $O_{i,j}$.源节点 S 将要发送的信息 M 分割成 m 份信息片,对每个信息片选择一个随机编码系数进行异或编码,并将编码后的消息和拆分后的编码系数分量分别沿着 m 条不同的路径传送至 $O_{i,1}(i=1,2,\ldots,m)$.任意中间节点 $O_{i,j}(i=1,2,\ldots,m,j=1,2,\ldots,n)$随机选择编码系数对收到的信息进行编码,将编码后的信息发送给 $O_{i,j+1}$,并将编码系数拆分成 m 个分量后分别沿着 m 条不同的路径继续传输.路径上最后的节点 $O_{i,n}$ 将编码后的信息和编码系数传送给目的节点 X.我们将由 $m\times n$ 个中间节点 $O_{i,j}$ 以及这些中间节点之间的转发路径所组成的网络称为编码转发网络(coding and forwarding network).

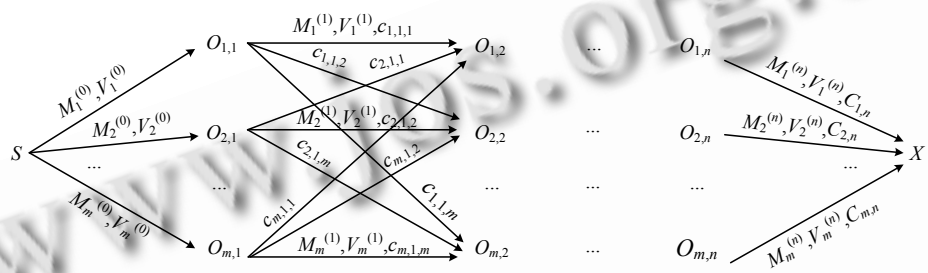

Fig.1 Coding and forwarding network with $m\times n$ nodes

图 1 $m\times n$ 个节点组成的编码转发网络

1.2 信息编码与传输过程

为了便于描述,先对编码过程中的运算进行定义,并对编码转发网络中的符号进行说明.

定义 1. 累积异或运算 $\bigoplus_{i=1}^{m} x_i$ 定义为 m 个数据 $x_1,x_2,x_3,\ldots,x_m$ 进行异或运算,即 $\bigoplus_{i=1}^{m} x_i = x_1 \oplus x_2 \oplus \ldots \oplus x_m$.

定义 2. 拼接运算 $x_1 \cup x_2$ 定义为数据 x_1 和 x_2 的二进制位串连接运算,即将数据 x_1 的二进制位作为高位、x_2 的二进制位作为低位形成一个新的二进制位串.例如 x_1='110110',x_2='001101',$x_1 \cup x_2$='110110 001101'.

定义 3. 累积拼接运算 $\bigcup_{i=1}^{m} x_i$ 定义为 m 个数据 $x_1,x_2,x_3,\ldots,x_m$ 进行拼接运算,即 $\bigcup_{i=1}^{m} x_i = x_1 \cup x_2 \cup \ldots \cup x_m$.

定义 4. 向量 C 与向量 V 的异或运算 $C\oplus V$.若有向量 $C=(c_1,c_2,c_3,\ldots,c_m)$和 $V=(v_1,v_2,v_3,\ldots,v_m)$,定义两者的异或运算为 $C\oplus V=(c_1\oplus v_1,c_2\oplus v_2,\ldots,c_m\oplus v_m)$.

符号说明:

- $M_i^{(j)}$ 表示节点 $O_{i,j}$ 编码后的信息块,即信息片 M_i 经过源节点 S、中间节点 $O_{i,1},\ldots,O_{i,j}$ 编码后的信息;
- $V_i^{(j)}$ 表示节点 $O_{i,j}$ 计算得到的编码系数的累积异或向量,即前 $m\times(j-1)$编码转发网络中,节点所选的编码系数的第 i 个分量与源节点的编码系数向量 $V_i^{(0)}$ 进行异或得到的累积结果;
- $C_{i,j}$ 表示节点 $O_{i,j}$ 随机选择的对转发信息片进行编码的编码系数;
- $c_{i,j,k}$ 表示编码系数分量,即编码系数 $C_{i,j}$ 按二进制位等分成 m 个分量,$c_{i,j,k}$ 表示其中的第 k 个分量.

以下是源节点和编码转发网络各中间节点对信息进行分割和编码的具体策略.

(1) 源节点 S 的信息分割和编码策略

首先,源节点将要发送的消息 M 分割成长度为 d 的 m 份信息片 $M_1,M_2,M_3,\ldots,M_m$.这里,令分割函数 $f(x)$是可逆运算,即存在 $f^{-1}(x)$,依据分割后的信息片可以还原成信息 M,即有 $M=f^{-1}(M_1,M_2,M_3,\ldots,M_m)$.只有获得所有的 m 份信息 $M_1,M_2,M_3,\ldots,M_m$,才能恢复信息 M.

然后,源节点随机选取 m 个长度为 d 的编码系数 $C_{i,0}$ 分别对分割后的信息片 M_i 进行异或编码运算,得到

$$M_i^{(0)}=C_{i,0}\oplus M_i.$$

将编码系数 $C_{i,0}$ 按二进制位等分成 m 个编码系数分量 $c_{i,0,1},c_{i,0,2},\ldots,c_{i,0,m}$,即 $C_{i,0}=c_{i,0,1}\cup c_{i,0,2}\cup\ldots\cup c_{i,0,m}$.最后一个系数分量 $c_{i,0,m}$ 的长度为 $d-(m-1)\times\lfloor d/m\rfloor$,其他各个系数分量的长度均为$\lfloor d/m\rfloor$,所以各信息分量的编码计算可以表示成以下形式:

$$\begin{pmatrix}M_1^{(0)}\\M_2^{(0)}\\\ldots\\M_m^{(0)}\end{pmatrix}=\begin{pmatrix}C_{1,0}\oplus M_1\\C_{2,0}\oplus M_2\\\ldots\\C_{m,0}\oplus M_m\end{pmatrix}=\begin{pmatrix}(c_{1,0,1}\cup c_{1,0,2}\cup\ldots\cup c_{1,0,m})\oplus M_1\\(c_{2,0,1}\cup c_{2,0,2}\cup\ldots\cup c_{2,0,m})\oplus M_2\\\ldots\\(c_{m,0,1}\cup c_{m,0,2}\cup\ldots\cup c_{m,0,m})\oplus M_m\end{pmatrix}=\begin{pmatrix}\bigcup_{k=1}^{m}c_{1,0,k}\oplus M_1\\\bigcup_{k=1}^{m}c_{2,0,k}\oplus M_2\\\ldots\\\bigcup_{k=1}^{m}c_{m,0,k}\oplus M_m\end{pmatrix}.$$

令 V_S 是由编码系数分量组成的 $m\times m$ 的编码系数矩阵,即有

$$V_S=\begin{bmatrix}C_{1,0}\\C_{2,0}\\\ldots\\C_{m,0}\end{bmatrix}=\begin{bmatrix}c_{1,0,1}&c_{1,0,2}&\cdots&c_{1,0,m}\\c_{2,0,1}&c_{2,0,2}&\cdots&c_{2,0,m}\\\ldots&\ldots&\ldots&\ldots\\c_{m,0,1}&c_{m,0,2}&\cdots&c_{m,0,m}\end{bmatrix}.$$

如图 1 所示,在分路传输时,为了避免中间节点解码,将编码后的信息片和编码系数分量分路传输.具体方法是将编码后的信息片 $M_i^{(0)}$ 沿第 i 条路径传输,同时将对应的编码系数 $C_{i,0}$ 的分量 $C_{i,0,j}$ 分别沿第 j 条路径传输.即源节点将 V_S 中的第 i 列的系数分量组成向量 $V_i^{(0)}=(c_{k,0,i}|k=1,2,\ldots,m)$,与 $M_i^{(0)}$ 一起在第 i 条路径上传输.由于 $M_i^{(0)}$ 是由 M_i 与 V_S 的第 i 行系数分量进行累积拼接和异或运算得到的,这样,第 i 条路径上的第 1 个节点 $O_{i,1}$ 收到 $\{M_i^{(0)},V_i^{(0)}\}$ 后是无法直接解码得到 M_i 的.

(2) 编码转发网络节点的转发策略

如图 1 所示,编码转发网络任意节点 $O_{i,j}(i=1,2,\ldots,m,j=1,2,\ldots,n)$中,各路径上的第 1 个节点 $O_{i,1}$ 直接从源节点接收到 $\{M_i^{(0)},V_i^{(0)}\}$,而其他节点会从它的前一个节点 $O_{i,j-1}$ 收到$\{M_i^{(j-1)},V_i^{(j-1)},c_{i,j-1,i}\}$,同时会分别收到来自于其他路径第 $j-1$ 个节点 $O_{k,j-1}(k=1,2,\ldots,m,k\neq i)$的编码系数分量 $c_{k,j-1,i}$.对信息的处理过程包含 3 个步骤:

① 随机选择长度为 d 的编码系数 $C_{i,j}$,对信息块 $M_i^{(j-1)}$ 进行编码:

$$M_i^{(j)}=C_{i,j}\oplus M_i^{(j-1)}=\bigoplus_{t=0}^{j}C_{i,t}\oplus M_i.$$

② 计算编码系数的累积异或向量 $V_i^{(j)}$:

$O_{i,1}$ 前只有一个源节点,编码系数已包含在 $V_i^{(0)}$ 中,所以 $V_i^{(1)}=V_i^{(0)}$;其他点 $O_{i,j}$ 将所有收到的编码系数组成向量 $C'_{j-1,i}=(c_{1,j-1,i},c_{2,j-1,i},\ldots,c_{m,j-1,i})$,计算

$$V_i^{(j)}=C'_{j-1,i}\oplus V_i^{(j-1)}=\left(\bigoplus_{t=0}^{j-1}c_{1,t,i},\bigoplus_{t=0}^{j-1}c_{2,t,i},\ldots,\bigoplus_{t=0}^{j-1}c_{m,t,i}\right).$$

③ 将 $C_{i,j}$ 按二进制位等分成 m 个编码系数分量$(c_{i,j,1},c_{i,j,2},\ldots,c_{i,j,m})$,将 $M_i^{(j)},V_i^{(j)}$ 和编码系数分量 $c_{i,j,i}$ 一起传给第 i 条路径上的下一跳节点 $O_{i,j+1}$,编码系数分量 $c_{i,j,k}$ 分别传送给第 k 路径上的第 $j+1$ 个节点 $O_{k,j+1}(k=1,2,\ldots,m,k\neq i)$.由于 $O_{i,n}$ 的后继节点只有一个,所以直接将 $M_i^{(n)},V_i^{(n)}$ 和编码系数 $C_{i,n}$ 发送给 X.

(3) 目的节点 X 的解码策略

目标节点 X 能够收到来自于所有路径上的编码信息片 $M_i^{(n)}$、编码系数的累积异或向量 $V_i^{(n)}$ 和编码系数 $C_{i,n}(i=1,2,\ldots,m)$,由于目的节点得到了所有的编码系数和编码信息片,所以可以解码得到信息 M.具体的解码方法见第 2 节定理 1 的证明.

2 ITNC 编码可解性分析

利用 ITNC 策略可以实现源节点 S 将信息 M 安全地传送给目的节点 X.下面,我们首先给出目的节点对信息 M 的可解性分析.

定理 1. 源节点 S 运用 ITNC 策略将信息 M 传送给 X,在不考虑网络差错的情况下,目的节点 X 能够正确解出源节点 S 发出的信息.

证明:显然,在没有网络差错的情况下,目的节点 X 能够收到来自于 m 条路径上不同的编码信息片 $M_i^{(n)}$、编码系数的累积异或向量 $V_i^{(n)}$ 以及 $O_{i,n}$ 的编码系数 $C_{i,n}(i=1,2,\ldots,m)$.其中,

$$M_i^{(n)}=\bigoplus_{t=0}^{n}C_{i,t}\oplus M_i,\ V_i^{(n)}=\left(\bigoplus_{t=0}^{n-1}c_{1,t,i},\bigoplus_{t=0}^{n-1}c_{2,t,i},\ldots,\bigoplus_{t=0}^{n-1}c_{m,t,i}\right),$$

则由所有的 $V_i^{(n)}$ 可以组成编码系数矩阵 $V^{(n)}$:

$$V^{(n)}=\begin{bmatrix}V_1^{(n)}\\V_2^{(n)}\\\cdots\\V_m^{(n)}\end{bmatrix}=\begin{bmatrix}\bigoplus_{t=0}^{n-1}c_{1,t,1} & \bigoplus_{t=0}^{n-1}c_{2,t,1} & \cdots & \bigoplus_{t=0}^{n-1}c_{m,t,1}\\ \bigoplus_{t=0}^{n-1}c_{1,t,2} & \bigoplus_{t=0}^{n-1}c_{2,t,2} & \cdots & \bigoplus_{t=0}^{n-1}c_{m,t,2}\\ \cdots & \cdots & \cdots & \cdots\\ \bigoplus_{t=0}^{n-1}c_{1,t,m} & \bigoplus_{t=0}^{n-1}c_{2,t,m} & \cdots & \bigoplus_{t=0}^{n-1}c_{m,t,m}\end{bmatrix}.$$

将编码系数矩阵 $V^{(n)}$的第 i 列进行累积拼接运算后再与 $C_{i,n}, M_i^{(n)}$ 进行异或操作即可解码出信息 M_i $(i=1,2,\ldots,m)$.即

$$\left(\left(\bigoplus_{t=0}^{n-1}c_{i,t,1}\right)\cup\left(\bigoplus_{t=0}^{n-1}c_{i,t,2}\right)\cup\ldots\cup\left(\bigoplus_{t=0}^{n-1}c_{i,t,m}\right)\right)\oplus C_{i,n}\oplus M_i^{(n)}=\left(\bigoplus_{t=0}^{n}C_{i,t}\right)\oplus M_i^{(n)}=M_i.$$

然后根据 m 片消息 $M_1,M_2,\ldots,M_m$ 恢复出信息 $M=f^{-1}(M_1,M_2,\ldots,M_m)$.证毕. □

传输路径上的任意点 $O_{i,j}$ 参与信息的编码与转发,下面将分析其对传送信息 M_i 的不可解性.

定理 2. 运用 ITNC 策略传送信息 M_i 的过程中,除了目标节点之外,传输路径上的任意点 $O_{i,j}$ 无法解码 M_i.

证明:显然,任意路径上的第 1 个节点 $O_{i,1}$ 收到 $\{M_i^{(0)},V_i^{(0)}\}$ 后是无法直接解码得到 M_i 的.

传输路径上的任意其他点 $O_{i,j}(j>1)$收到的信息为:从 $O_{i,j-1}$ 收到的$\{M_i^{(j-1)},V_i^{(j-1)},c_{i,j-1,i}\}$,从其他的前驱节点分别收到 $c_{1,j-1,i},c_{2,j-1,i},\ldots,c_{m,j-1,i}$,其中,

$$M_i^{(j-1)}=\bigoplus_{t=0}^{j-1}C_{i,t}\oplus M_i=\bigoplus_{t=0}^{j-1}\bigcup_{k=1}^{m}c_{i,t,k}\oplus M_i,$$

$$V_i^{(j-1)}=\left(\bigoplus_{t=0}^{j-2}c_{1,t,i},\bigoplus_{t=0}^{j-2}c_{2,t,i},\ldots,\bigoplus_{t=0}^{j-2}c_{m,t,i}\right).$$

显然,要解码信息 M_i,需要知道编码系数 $\bigoplus_{t=0}^{j-1}\bigcup_{k=1}^{m}c_{i,t,k}$,而节点 $O_{i,j}$ 只有部分编码系数分量,即只有 $V_i^{(j-1)}$ 中的第 i 项参数 $\bigoplus_{t=0}^{j-2}c_{i,t,i}$ 和 $c_{i,j-1,i}$,所以无法直接求解 M_i.证毕. □

推论 1. 源节点 S 运用 ITNC 策略将在信息 M 传送给 X 的过程中,除了目的节点之外,对于传输路径上的任意一个点,都无法解出完整信息 M.

由于需要知道所有 M_i 才能得到完整信息 M,推论 1 显然成立.

因此,除目的节点外,路径上的任意单个编码转发网络节点无法通过解码方式获得完整信息.下面将进一步分析多个编码转发网络节点解码的充要条件.

定理 3. 源节点 S 在运用 ITNC 策略将信息 M 传送给 X 的过程中,如果编码转发网络有 m 条路径,每条路径上有 n 个点,假设只有存在泄密点才能获得该泄密点转发的数据包,则信息 M_i 可解的充要条件是在第 i 条路径上存在一个或者多个连续的泄密点 $O_{i,x},x\in[p,p+r],r\geqslant 0$,且在其他每条路径上都存在至少一个泄密点,并且这些

泄密点在转发路径上的位置为[p,p+r+2].

在给出定理 3 证明之前,我们给出满足和不满足充要条件的示意图.如图 2 所示,图中呈矩阵排列的点表示编码路径上的节点,第 i 行第 j 列的点表示编码转发网络节点 $O_{i,j}$.图 2(a)表示一种 M_1 可解的泄密节点位置示意图,即第 1 条路径有两个泄密点,其他每条路径上在灰色点位置至少有一个泄密点.图 2(b)表示 M_1 不可解的情况,即第 2 条和第 4 条路径在符合条件的位置上没有泄密点.

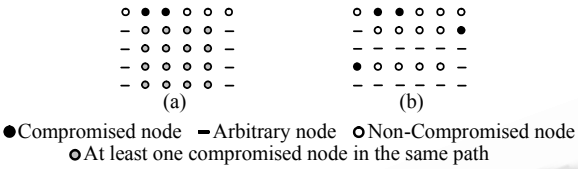

Fig.2 Positions of compromised nodes

图 2 泄密节点位置

以下给出定理 3 的证明.

证明:我们分别从充分性和必要性两个方面进行证明.

(1) 充分性:我们证明由这些位置上的泄密点收到的编码系数分量可以计算得到求解 M_i 的编码系数.

第 i 条路径上的泄密点 $O_{i,p}$ 收到 $M_i^{(p-1)}$,$V_i^{(p-1)}$ 和 $c_{1,p-1,i},c_{2,p-1,i},\ldots,c_{m,p-1,i}$,其中,

$$M_i^{(p-1)}=\bigoplus_{t=0}^{p-1}C_{i,t}\oplus M_i=\bigoplus_{t=0}^{p-1}\bigcup_{k=1}^{m}c_{i,t,k}\oplus M_i=R_{i,p}\oplus M_i,$$

$$V_i^{(p-1)}=\left(\bigoplus_{t=0}^{p-2}c_{1,t,i},\bigoplus_{t=0}^{p-2}c_{2,t,i},\ldots,\bigoplus_{t=0}^{p-2}c_{m,t,i}\right),$$

$$R_{i,p}=\bigoplus_{t=0}^{p-1}\bigcup_{k=1}^{m}c_{i,t,k}=\bigcup_{k=1}^{m}\bigoplus_{t=0}^{p-1}c_{i,t,k}=\bigcup_{k=1}^{m}R_{i,p}(k).$$

显然,$O_{i,p}$ 要求解 M_i 必须知道其编码系数 $R_{i,p}$.

$O_{i,p}$ 自身依据收到的信息可以得到编码系数 $R_{i,p}$ 的在路径 i 的分量 $R_{i,p}(i)=\bigoplus_{t=0}^{p-1}c_{i,t,i}$.

若其他路径 k 上的泄密点的位置为 p,则泄密点 $O_{k,p}$(k=1,2,…,m,$k\neq i$)依据更新得到的编码系数累积异或向量 $V_k^{(p)}$ 中的第 i 项元素就可以获得编码系数 $R_{i,p}$ 在路径 k 上的分量 $R_{i,p}(k)=\bigoplus_{t=0}^{p-1}c_{i,t,k}$.

若路径 k 上的泄密点的位置为 p+1,则泄密点 $O_{k,p+1}$(k=1,2,…,m,$k\neq i$)收到的向量 $V_k^{(p)}$ 中的第 i 项元素即为 $R_{i,p}(k)=\bigoplus_{t=0}^{p-1}c_{i,t,k}$.

若路径 k 上的泄密点的位置为 $x\in[p+2,p+r+2]$,则泄密点 $O_{k,x}$(k=1,2,…,m,$k\neq i$)的 $V_k^{(x-1)}$ 中的第 i 项参数为 $\bigoplus_{t=0}^{x-2}c_{i,t,k}$,由于第 i 条路径上存在的连续泄密点 $O_{i,p},\ldots,O_{i,x-2}$ 分别有系数分量 $c_{i,p,k},\ldots,c_{i,x-2,k}$,所以,这些点合谋可以获得 $R_{i,p}(k)=\bigoplus_{t=0}^{p-1}c_{i,t,k}=\bigoplus_{t=0}^{x-2}c_{i,t,k}\oplus c_{i,p,k}\oplus\ldots\oplus c_{i,x-2,k}$.

所以,依据所有路径上泄密点提供的编码系数分量可以计算得到 $R_{i,p}=\bigcup_{k=1}^{m}R_{i,p}(k)=\bigcup_{k=1}^{m}\left(\bigoplus_{t=0}^{p-1}c_{i,t,k}\right)=\bigoplus_{t=0}^{p-1}\bigcup_{k=1}^{m}c_{i,t,k}$,从而解码出 $M_i=R_{i,p}\oplus M_i^{(p-1)}=\bigoplus_{t=0}^{p-1}\bigcup_{k=1}^{m}c_{i,t,k}\oplus M_i^{(p-1)}$.

(2) 必要性:根据信息的编码与传输策略,显然,若泄密者要解码出信息 M_i,则在每一条路径上必须存在泄密节点.

仍然假设第 i 条路径上的泄密节点是 $O_{i,x}$,而另外有一条路径 k 上的泄密点的位置为 $j,j\notin[x,x+2]$.为了求解 M_i,泄密点需要得到求解 M_i 的编码系数 $R_{i,x}=\bigoplus_{t=0}^{x-1}C_{i,t}=\bigoplus_{t=0}^{x-1}\bigcup_{k=1}^{m}c_{i,t,k}=\bigcup_{k=1}^{m}\bigoplus_{t=0}^{x-1}c_{i,t,k}$.$O_{i,x}$ 可以根据自己收到的 $V_i^{(x-1)}$ 和

$c_{i,x-1,i}$ 得到 $R_{i,x}$ 在第 i 条路径上的分量 $R_{i,x}(i)=\bigoplus_{t=0}^{x-1}c_{i,t,i}$.由于编码系数采用分路传输的策略,显然,$R_{i,x}$ 中的分量 $R_{i,x}(k)=\bigoplus_{t=0}^{x-1}c_{i,t,k}$ 只能从第 k 条路径上获得.

根据假设,第 k 条路径上的泄密点为 $O_{k,j},O_{k,j}$ 能得到的包含在 $R_{i,x}$ 中的编码系数分量有 $\bigoplus_{t=0}^{j-2}c_{i,t,k}$ 和 $c_{i,j-1,k}$.

根据假设,$j\notin[x,x+2]$,若 $j<x$, $R_{i,x}(k)=\bigoplus_{t=0}^{x-1}c_{i,t,k}=\bigoplus_{t=0}^{j-2}c_{i,t,k}\oplus c_{i,j-1,k}\oplus\left(\bigoplus_{t=j}^{x-1}c_{i,t,k}\right)$,编码系数分量 $R_{i,x}(k)$中还有 $\bigoplus_{t=j}^{x-1}c_{i,t,k}$ 无法获得,所以无法求解 M_i(图 2(b)的第 4 条路径情况).

若 $j>x+2$, $R_{i,x}(k)=\bigoplus_{t=0}^{x-1}c_{i,t,k}=\bigoplus_{t=0}^{j-2}c_{i,t,k}\oplus\left(\bigoplus_{t=x}^{j-2}c_{i,t,k}\right)$,由于无法获得分量 $\bigoplus_{t=x}^{j-2}c_{i,t,k}$,同理可证无法求解 M_i(图 2(b)的第 2 条路径情况).证毕. □

定理 4. 源节点 S 运用 ITNC 策略将信息 M 传送给 X 的编码转发网络中有 m 条路径,每条路径上有 n 个点,若 S 到 X 的编码转发网络中的泄密节点数为 m,则泄密节点合谋获得信息 M 的概率 $Pr=n/C_{mn}^{m}$.

证明:当泄密节点数与信息分割的路径数相同时,攻击者要获得解码 M 的全部信息片,要求泄密节点在编码转发网络中的位置关系满足以下条件:

(1) 每一条路径上有且只有一个泄密者;

(2) 根据定理 3 可知,这 m 个泄密节点必须处于每条路径的相同转发位置上.

在一个有 $m\times n$ 个节点的编码转发网络中,选择同时满足上述两个条件的 m 个节点的概率为 n/C_{mn}^{m},即这 m 个泄密者合谋获得完整信息内容 M 的概率为 $Pr=n/C_{mn}^{m}$.证毕. □

通过上述分析我们可以看出,利用 ITNC 传输数据可以保证单个中间转发节点无法解码出转发的信息分片,在有多个泄密节点存在的情况下,只有泄密节点位置具有某种关系时才能解码信息片.因此,编码信息和编码系数的分离传输提高了无密钥信息分割传输的抗泄密攻击的能力.

3 基于多路径网络编码的匿名通信机制 AC-ITNC

我们将 ITNC 策略用于匿名通信的建路信息传输中,提出了基于多路径网络编码的匿名通信策略 AC-ITNC.

3.1 AC-ITNC通信机制

典型的匿名系统中,通信节点是知道匿名系统中的匿名代理节点的.AC-ITNC 中采用源路由的方式,即由源节点选择匿名系统中的部分匿名代理节点构成匿名传输路径.为实现匿名建路,即要中间代理节点知道它的下一跳节点和与源节点间的共享密钥.AC-ITNC 中采用 ITNC 策略向路径上的每个中继节点传输匿名建路信息.

发送者 S 首先确定一条匿名路径,即由 S 经过若干个转发代理节点到达 R 的路径.设该路径包含了 l 个中间代理节点,即 $S\to P_1\to P_2\to P_3\to\ldots\to P_l\to R$.源节点 S 采用 ITNC 策略传输建路信息,当 S 需向路径上的任意一点 P_i 传输建路信息时,首先在匿名系统中随机选择除路径上节点之外的 $m\times n$ 个节点构成编码转发网络,然后以 S 作为源点,P_i 作为接收点,将 P_{i+1} 的地址以及 P_i 与源节点的共享密钥作为传输信息进行分割编码后传输.

为了增强安全性,对于不同的 P_i,源节点分别选择构造不同的编码转发网络来传输.同时,为了保证源节点的匿名性,编码转发网络路径上的第 1 个节点是必须安全的点,因为第 1 个节点可以依据其前驱只有一个节点来确定前一个点是发送者.因此在编码转发网络的选取过程中,源节点可以选取 m 个与源节点在同一安全区域的代理节点作为编码转发网络的 m 条路径上的第 1 跳节点,也可以虚拟出 m 个节点作为第 1 跳节点.这样,第 1 跳编码节点的匿名性就等价于源节点了.考虑到在建路阶段也可能遭到典型的前驱攻击,文献[4,18]指出,转发路径长度越长,源节点匿名性越好.为了不降低源匿名性,编码转发网络路径长度必须大于等于匿名路径长度.因此,这里源节点构造的编码路径长度至少是 l.

当建路信息传输完成后,路径上节点 P_i 得到了下一跳节点 P_{i+1} 的地址和 P_i 与源节点之间的共享密钥 Key_i.然后就进入匿名传输阶段,即由源节点利用已协商好的密钥对信息进行嵌套加密传输,每个路径上的节点利用自己与源节点间的密钥解密最外层信息后,将未解密的部分向下一节点传输,传输方式等同于洋葱路由数据传输阶段.

本文提出的方法只是在建路阶段引入多路径,在数据传输阶段仍然采用单路径,建路阶段传输的信息量仅包含下一跳地址和协商的密钥信息,传输的信息量是很少的,并且建路信息只需要传输一次.因此,不会导致匿名通信系统传输负载的大量增加.

3.2 建路信息传输的具体实现

当发送者 S 要告知匿名路径上代理节点 $P_x(x=1,2,\ldots,l)$建路信息时,S 分别为不同的 P_x 建立不同的编码转发网络.具体实现过程如下:

首先,由 S 选取 m 个与源节点在同一安全区域的代理节点,或者虚拟出 m 个节点作为编码转发网络的 m 条路径上的第 1 跳节点,然后,在匿名系统的其他已知代理中随机选择 $m\times(n-1)$个节点,确定这些节点在 $m\times n$ 编码转发传输网络中的位置.如图 1 所示,任意节点 $O_{i,j}$ 需要知道其 m 个下一跳节点 $O_{t,j+1}(t=1,2,\ldots,m)$.用 A_i 表示节点 $O_{t,i+1}(t=1,2,\ldots,m)$的地址信息,即编码转发网络路径上第 j 个节点的 m 个下一跳节点地址.

为了建立起这些节点间的编码转发网络路径,并将建路信息传输给 P_i,源节点 S 需要告诉这些被选择的节点在编码转发路径上的下一跳节点.一种简单直观的处理方法是:源节点 S 首先沿 m 条路径将 A_1 发送给距离 S 为 1 跳的 m 个节点 $O_{t,1}(t=1,2,\ldots,m)$,然后将这 m 个节点 $O_{t,1}(t=1,2,\ldots,m)$看作一个 $m\times1$ 的编码转发网络,在此编码转发网络中,利用 ITNC 策略将 A_2 分别发送到目的节点 $O_{t,2}(t=1,2,\ldots,m)$.此时,可以进一步由节点 $O_{t,1},O_{t,2}$ $(t=1,2,\ldots,m)$构造一个 $m\times2$ 的编码转发网络,并利用该编码转发网络将 A_3 分别发送到目的节点 $O_{t,3}(t=1,2,\ldots,m)$.以此方法,我们可以构造出一个 $m\times n$ 编码转发网络,利用 ITNC 策略在此网络中可以将匿名路径建路信息 M 发送给匿名路径上代理节点 P_i.源节点通过此方法建立不同的编码转发网络,并将匿名路径建路信息发送给匿名路径上的代理节点,从而实现在无密钥基础设施环境下匿名传输路径的建立.

在这种方法中,源节点要传输的建路信息包括 $A_1,A_2,A_3,\ldots,A_n,M_x$,其中,$M_x$ 是传输给匿名代理 P_x 的匿名建路信息,包括 P_x 在匿名路径的下一跳 P_{x+1} 以及 P_x 与 S 之间的共享密钥等.为了减少编码转发网络建立的开销,本文采用边建编码转发网络边传输匿名建路信息的方式,将所有的建路信息合成一个数据包来传输.

任意编码转发网络节点收到建路信息包后,首先从建路数据包的第 1 层信息中解码得到自己在编码转发网络中的 m 个下一跳节点,然后对去除第 1 层后的建路信息继续按照 ITNC 的方式编码再传送.即从源节点 S 采用逐跳形成编码转发网络的方式,将不同层次信息以 ITNC 方式传送给编码转发网络中不同位置的节点.例如,A_k 的编码传输是以 S 为源节点,$O_{i,k}$ 为目标节点,$O_{i,k}$ 的前面 $m\times(k-1)$个节点组成编码转发网络进行传输的.

图 3 给出了任意编码转发网络节点对接收到的信息的处理与转发过程 $Code\_forward(s_1,s_2,s_3,\ldots,s_m)$,其中,$s_i$ 为节点从第 i 条路径接收到的信息,ns_i 为节点发送给下一跳节点中第 i 个节点的信息.编码转发网络节点从 m 个上一跳节点分别接收到数据包 $s_1,s_2,s_3,\ldots,s_m$,依据这 m 个数据包的第 1 层信息解码得到 m 个下一跳地址.在编码转发网络节点转发过程中,每次转发数据包层次减 1,过程 $Code_forward()$利用 $Layercount(x)$计算数据包 x 的层次数目.在编码转发网络节点上,接收到的数据包 $s_1,s_2,s_3,\ldots,s_m$ 的层次数目是一样的,并且由于任何一个数据包至少包含目的地址和匿名建路信息两层,所以数据包的层数不小于 2.编码转发网络节点提取 $s_1,s_2,s_3,\ldots,s_m$ 的任意层次 j,分别进行再次编码,$Code(s[1,j],s[2,j],\ldots,s[m,j])$表示对提取出的各数据包的第 j 层信息进行编码(节点编码的过程见第 1.2 节的 3 个步骤),得到转发给 m 个下一跳节点的数据包的第 $j-1$ 层信息 $ns[1\ldots m,j-1]$.编码转发网络节点在对各层信息编码完成后,将转发给同一个下一跳节点的多个层次信息进行合并后发送.

```
Code_forward(s1,s2,s3,…,sm)                    //Algorithm of coding and forwarding in each coding node
  For i=1 to m
    s[i,1]= Layer(si,1);                       //Layer(si,k) is to extract the kth layer's message from si
    nexthop[]←Decode(s[1,1],s[2,1],…,s[m,1]);  //Decode and get the addresses of m next hops
    countoflayer=Layercount(s1);               //Layercount(x) calculates the number of layers of packet x
    For j=2 to countoflayer
      For k=1 to m
        s[k,j]=Layer(sk,j);                    //Extract the jth layer message of input packet s
      ns[1…m,j−1]←Code(s[1,j],s[2,j],...,s[m,j]);  //Code the jth layer's message of input and get
                                                   the (j−1)th layer's message of output
  For i=1 to m
    nsi=Union_packet(ns[i,1],ns[i,2],…,ns[i,countoflayer−1]);
                                               //All messages to the same next node are combined to nsi
    Forward(nsi,nexthop[i]);                   //Forward nsi to next hop node nexthop[i]
```

Fig.3 Process of coding and forwarding

图 3 编码转发处理过程

4 AC-ITNC 匿名安全性分析

4.1 发送者匿名性

首先,我们分析 AC-ITNC 的发送者(源节点)匿名性.

在匿名路径建立过程中,由于采用了 ITNC 策略,编码转发网络中的第 1 个点因为其前驱节点只有一个,所以可以直接推断出前驱是源节点.因此,编码转发网络中的第 1 个点必须是可信的点.与其他匿名机制一样,源可以选择与自己在同一安全区域的代理节点或虚拟出 m 个节点作为第 1 跳节点.同时,由于编码路径长度至少是 l,因此不会导致建路阶段发送者匿名性的下降.

在数据传输阶段,由于采用与洋葱路由数据传输一样的重路由匿名通信机制,发送者的匿名性能可以由匿名转发路径长度保证.因此,发送者匿名性与其他定长的转发策略[18,19]相当,等价于同样路径长度的洋葱路由.

4.2 接收者匿名性

接着,我们分析采用 ITNC 传输建路信息对接收者匿名性的影响.

因为匿名路径上的下一跳地址和密钥信息是包含在建路信息包中进行编码传送的,因此下一跳地址和密钥的安全性等价于建路信息内容的编码传输的安全性.

(1) 泄密者解码建路信息中的密钥对接收者匿名性的影响.

当泄密者解码得到路径上某个节点与源节点的共享密钥后,在数据传输阶段,泄密者如果能窃听到传输给该节点的信息就可以解密该信息的最外层.同时,通过窃听该节点的输出可以关联输入与输出信息的内容,从而知道路径上的下一跳节点.如果泄密者能够获得匿名路径上所有节点的密钥,就可以确定完整的匿名路径,找到接收者,破坏匿名性.

(2) 泄密者解码建路信息中的下一跳地址对接收者匿名性的影响.

由于接收者的地址是被作为下一跳地址传给路径上的最后一个节点 P_l 的,因此泄密节点获得接收者地址的概率等价于建路信息安全传输的概率.但由于匿名网络中传输了多个建路信息,即使泄密者破解了某个建路信息,也只能肯定信息中的地址是匿名路径上的节点,而无法确定其在匿名路径上的位置.因此,即使泄密者破解了 P_l 的下一跳地址,也无法肯定该地址就是接收者.

通过以上的分析可知,泄密者只有共谋获得所有的建路信息后,才可能确定完整的匿名路径,从而确定接收者,否则只能确定路径上的部分节点,而无法确定接收者.

下面我们分析单个建路信息采用 ITNC 策略传输时,编码转发网络中存在泄密节点时的安全性.我们采用建路信息保密性来评价网络的安全性能,下面首先给出其定义.

定义 5. 建路信息保密性 S,即单个建路信息传输时不被泄密者合谋解码的概率.当源节点选择的编码转发

网络路径上存在泄密者时,若泄密点合谋成功解码建路信息的概率为 P,则 $S=1-P$.

假设在编码转发网络中存在泄密节点的个数为 k,我们分析泄密节点数 k、编码转发网络的路径数 m 以及编码路径长度 n 与建路信息保密性之间的关系.

由于在 AC-ITNC 中采用了 ITNC 策略来进行信息 M 的传送,根据定理 3,只有泄密节点位置之间存在特殊的关系时,才能解码得到信息内容.下面将依据编码转发网络中存在泄密节点的个数 k 与编码转发网络的路径数 m 之间的关系讨论建路信息保密性 S:

1) 由于建路信息被分割成了 m 份,在 m 条路径上分别传输不同的分割信息片,攻击者要获得解码的全部信息片,则每一条路径上都至少有一个泄密者.因此,当 $k<m$ 时,$P=0$,$S=1$,可以保证建路信息的保密性;

2) 当泄密节点数 $k=m$ 时,根据定理 4 和定义 5 可知,系统建路信息内容保密性为

$$S_{AC\text{-}ITNC}=1-(n-1)/C_{m(n-1)}^{m}.$$

为了对 AC-ITNC 策略在传输建路信息时的安全性进行分析,我们选择了文献[12]中的 SIS 策略进行比较.SIS 策略也是一种不需要公钥机制进行匿名建路信息传输的策略,但 SIS 只是由源节点对信息分割,中间节点不参与编码,尽管进行了流标记的修改,泄密节点仍然可以将收到的信息尝试进行合并.因此,每条路径的任意位置只要存在一个泄密点,它们合谋就可以恢复出信息 M.所以,当 $k=m$ 时,系统的建路信息保密性 $S_{SIS}=1-(n-1)^m/C_{m(n-1)}^{m}$.显然,在相同的 m 和 n 条件下,AC-ITNC 策略的建路信息保密性 $S_{AC\text{-}ITNC}$ 比 SIS 策略的建路信息保密性 S_{SIS} 要好.表 1 给出了当 $k=m$,$n=11$ 时,AC-ITNC 和 SIS 两种策略下建路信息保密性的比较.

Table 1 Path setup information confidentiality of SIS and AC-ITNC (n=11)

表 1 建路信息内容保密性 S_{SIS} 和 $S_{AC\text{-}ITNC}$(n=11)

m	1	2	3	4	5	6	7	8	9	10
S_{SIS}	0	0.473 7	0.753 7	0.890 6	0.952 8	0.980 0	0.991 7	0.996 6	0.998 6	0.999 4
$S_{AC\text{-}ITNC}$	0	0.947 4	0.997 5	0.999 9	1	1	1	1	1	1

从表 1 中可以看出,当 $m=1$ 时,信息没有进行分割,以明文方式进行传送,两种策略下任意泄密者都可以获得信息 M,因此建路信息安全性 $S=0$.当 n 一定,随着转发路径条数 m 的增大,两种策略的建路信息保密性 S 都增加;AC-ITNC 获得了优于 SIS 的建路信息保密性.当 $m=2$ 时,$S_{AC\text{-}ITNC}=0.947\,4$,而 $S_{SIS}=0.473\,7$.

(3) 当泄密节点数大于转发路径条数,即 $k>m$ 时,我们模拟测试了 SIS 和 AC-ITNC 两种策略的建路信息保密性.

模拟测试中选定了如下一些参数:m 是编码转发网络的路径数,n 是编码转发网络的路径长度,k 为编码转发网络中的泄密节点数,f 为编码转发网络中的泄密节点比例.我们根据测试要求进行参数设置,构建 $m\times n$ 规模的编码转发网络,并随机确定 k 个泄密节点在编码转发网络中的位置,在每种规模条件下进行了 100 000 轮的测试.

图 4 给出了当路径数 m 和路径长度 n 固定时,SIS 和 AC-ITNC 两种策略下建路信息保密性 S 和泄密节点数 k 之间的关系.从图 4 可以看出,随着泄密节点数的增加,两者的建路信息保密性有所下降,但 $S_{AC\text{-}ITNC}$ 比 S_{SIS} 的下降速度要慢.当同样的编码转发网络条件下,即同样的 m 和 n 情况下,泄密节点数 k 相同时,$S_{AC\text{-}ITNC}>S_{SIS}$,表明在同样的多路径条件下,AC-ITNC 方法较 SIS 方法具有更好的抗泄密者合谋攻击的能力.

从图 4 可以发现:当编码转发网络中除去路径上的第 1 列安全节点外的转发节点数相同,即 $m\times(n-1)$相同时,m 越大,建路信息保密性越好.

同样,我们考虑在匿名系统中泄密者比例 f 一定的时候,两种策略中路径条数 m 对建路信息保密性 S 的影响,如图 5 所示.显然,f 越大,建路信息保密性越弱.在同样的 f 和 n 情况下,m 越大,建路信息保密性越好,这与定理 3 中的解码条件是一致的.解码要求至少每条路径上存在一个泄密点,编码路径条数 m 越多,解码要求的泄密者数也越多.同时,在相同的编码转发网络和泄密者比例条件下,AC-ITNC 策略的建路信息保密性大大高于 SIS 策略.当 $m>5$ 时,AC-ITNC 策略在 30%泄密者情况下仍能保持 0.85 以上的建路信息保密性.

图 6 所示的是在编码路径条数 m 和泄密者比例 f 一定的情况下,编码路径长度 n 对 S 的影响.可以看出,当 n 较小时,建路信息保密性较好;随着 n 的增长,建路信息保密性是减弱的,AC-ITNC 策略较 SIS 下降要慢.但由于

是源节点利用编码转发网络发送建路信息,文献[4,19]证明了针对常见的发送者匿名攻击方式前驱攻击而言,路径长度越长,发送者匿名性越好.因此,为了保证发送者匿名性,编码路径长度不能太短.为了在建路阶段不会降低发送者匿名性,需保证编码路径长度 n 不小于匿名路径长度 l.从图 6 可以看到当 m=5,f<20%,n<15 时,AC-ITNC 策略中的编码路径长度对建路信息保密性影响不大,S 一直可以保持在 0.95 以上.当泄密节点比例增加到 30% 时,编码路径长度为 10 时仍能保持在 0.9 左右.

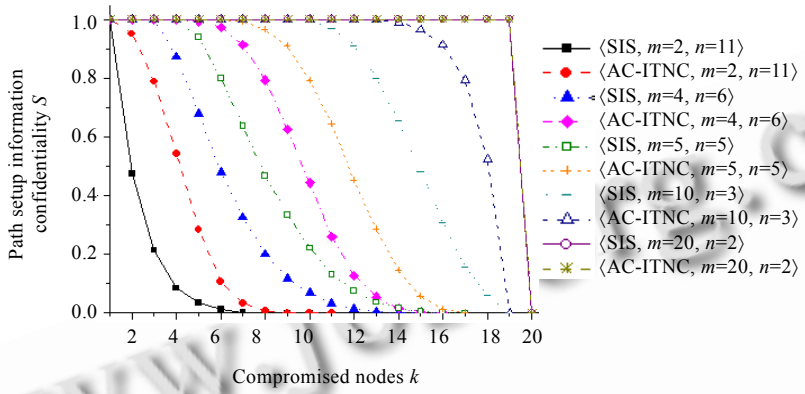

Fig.4 Path setup information confidentiality S vs. number of compromised nodes k (m×(n−1)=20)

图 4 建路信息保密性 S 与泄密节点数 k (m×(n−1)=20)

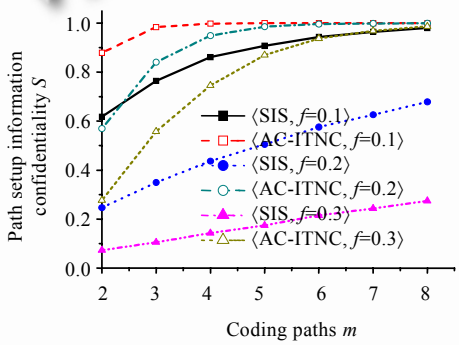

Fig.5 S vs. coding paths m (n=10)

图 5 S 与路径数 m (n=10)

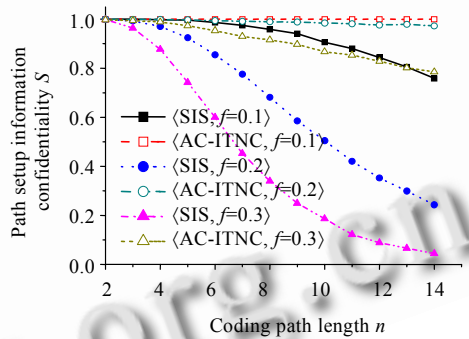

Fig.6 S vs. coding path length n (m=5)

图 6 S 与路径长度 n (m=5)

图 7 表明了当 m=5 时不同编码路径长度情况下,建路信息保密性与泄密节点比例的关系.显然,泄密节点比例 f 增加,建路信息保密性 S 下降.编码路径增长,S 下降,这主要是由于编码路径长度增长,路径上可能的泄密节点数增加,同时解码条件中泄密节点位置关系较易满足造成的.

因此在选择编码转发网络时,一方面为了保证发送者匿名性,应该使编码路径长度 n 不低于匿名路径长度 l;另一方面,为了提高建路信息保密性,应尽量增多编码路径条数,减短编码路径长度.同时考虑两者,可以取 n=l,m≥5.

根据前面的分析,只有泄密者能够解码所有的建路信息时,才能确定完整的匿名路径,确定接收者;否则只能确定解码地址是匿名路径上的节点,而无法最终确定接收者.因此,解码单个建路信息并不能完全破坏接收者匿名性.显然,解码所有建路信息的概率是 P^l,因此可以定义接收者匿名性 A=1−P^l.图 8 给出了对应于图 7 条件的匿名性与泄密者比例的关系,可以看出当 m=5,n<14,f<40%的情况下,接收者匿名性可以保持在 0.99 以上.可见,采用网络编码分割策略传输建路信息可以保持较好的匿名性.

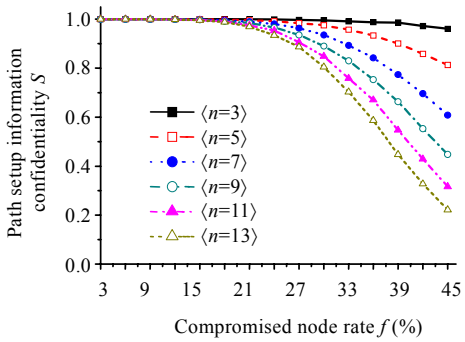

Fig.7 $S_{AC\text{-}ITNC}$ vs. compromised node rate $f(m=5)$

图 7 $S_{AC\text{-}ITNC}$ 与泄密节点比率 $f(m=5)$

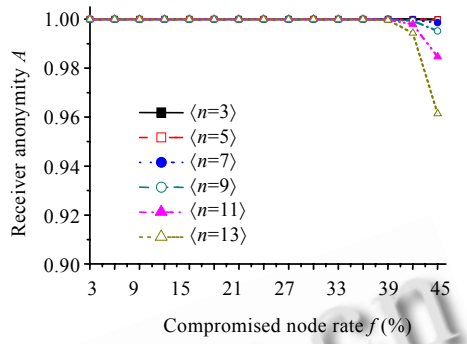

Fig.8 A vs. compromised node rate $f(m=5)$

图 8 A 与泄密节点比率 $f(m=5)$

5 结 论

本文基于信息分割机制,综合网络编码和源路由的思想,提出了一种基于多路径网络编码的信息分割传输策略 ITNC,将该策略应用到匿名通信的建路机制中,提出了一种新的匿名通信策略 AC-ITNC.AC-ITNC 实现了无密钥基础设施下匿名传输路径的建立,通过中间节点参与编码、编码信息片和编码系数分离传送的策略,极大地提高了系统的抗合谋攻击能力,优化了匿名性能.本文的研究工作为在无密钥基础设施的分布式环境中实现匿名通信提供了新的思路.

References:

[1] Chaum D. Untraceable electronic mail, return addresses and digital pseudonyms. Communications of the ACM, 1981,24(2):84−88. [doi: 10.1145/358549.358563]

[2] Goldschlag D, Reed M, Syverson P. Onion routing for anonymous and private Internet connections. Communications of the ACM, 1999,42(2):39−41. [doi: 10.1145/293411.293443]

[3] Dingledine R, Mathewson N, Syverson P. Tor: The second-generation onion router. In: Proc. of the 13th USENIX Security Symp. Berkeley: USENIX Association, 2004. 303−320.

[4] Reiter M, Rubin A. Crowds: Anonymity for Web transactions. ACM Trans. on Information and System Security, 1998,1(1):66−92. [doi: 10.1145/290163.290168]

[5] Sherwood R, Bhattacharjee B, Srinivasan A. P5: A protocol for scalable anonymous communication. Journal of Computer Security, 2005,13(6):839−876.

[6] Freedman MJ, Morris R. Tarzan: A peer-to-peer anonymizing network layer. In: Proc. of the 9th ACM Conf. on Computer and Communications Security (CCS 2002). Washington: ACM Press, 2002. 193−206.

[7] Rennhard M, Plattner B. Introducing MorphMix: Peer-to-Peer based anonymous Internet usage with collusion detection. In: Proc. of the ACM Workshop on Privacy in the Electronic Society (WPES 2002). Washington: ACM Press, 2002. 91−102.

[8] Kate A, Zaverucha GM, Goldberg I. Pairing-Based onion routing. In: Borisov N, Golle P, eds. Proc. of the 7th Privacy Enhancing Technologies (PET 2007). LNCS 4776, Berlin: Springer-Verlag, 2007. 95−112.

[9] Øverlier L, Syverson P. Improving efficiency and simplicity of Tor circuit establishment and hidden services. In: Borisov N, Golle P, eds. Proc. of the 7th Privacy Enhancing Technologies (PET 2007). LNCS 4776, Berlin: Springer-Verlag, 2007. 134−152.

[10] Chaum D. The dining cryptographers problem: Unconditional sender and recipient untraceability. Journal of Cryptology, 1988,1(1): 65−75.

[11] Katti S, Katabi D, Puchala K. Slicing the onion: Anonymous routing without PKI. Technical Report, MIT-CSAIL-TR-2005-053, Cambridge: Massachusetts Institute of Technology, 2005.

[12] Katti S, Cohen J, Katabi D. Information slicing: Anonymity using unreliable overlays. In: Proc. of the 4th USENIX Symp. on Network Systems Design and Implementation (NSDI 2007). Berkeley: USENIX Association, 2007. 43−56.

[13] Ahlswede R, Cai N, Li SY, Yeung R. Network information flow. IEEE Trans. on Information Theory, 2000,46(4):1204−1216. [doi: 10.1109/18.850663]

[14] Chou PA, Wu Y, Jain K. Practical network coding. In: Proc. of the 41st Annual Allerton Conf. on Communication, Control, and Computing. Monticello, 2003. http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.11.697

[15] Widmer J, Fragouli C, Le Boudec JY. Low-Complexity energy efficient broadcasting in wireless ad-hoc networks using network coding. In: Proc. of the Workshop on Network Coding, Theory, and Applications. 2005. http://arni.epfl.ch/contents/pubs/pubs_network/network_8.pdf

[16] Gkantsidis C, Rodriguez P. Cooperative security for network coding file distribution. In: Proc. of the IEEE Int'l. Conf. on Computer Communications (INFOCOM 2006). Washington: IEEE Press, 2006. 1−13.

[17] Charles D, Jain K, Lauter K. Signatures for network coding. In: Proc. of the IEEE Conf. on Information Sciences and Systems (CISS 2006). Washington: IEEE Press, 2006. 857−863.

[18] Wang WP, Chen JE, Chen SQ, Wang JX. Research on a short distance-prior rerouting scheme in anonymous communication. Journal of Software, 2004,15(4):561−570 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/15/561.htm

[19] Wang WP, Chen JE, Wang JX. An anonymous communication protocol based on groups with definite route length. Journal of Computer Research and Development, 2003,40(4):609−614 (in Chinese with English abstract).

附中文参考文献:

[18] 王伟平,陈建二,陈松乔,王建新.匿名通信中短距离优先分组重路由方法的研究.软件学报,2004,15(4):561−570. http://www.jos.org.cn/1000-9825/15/561.htm

[19] 王伟平,陈建二,王建新.基于组群的有限路长匿名通信协议.计算机研究与发展,2003,40(4):609−614.

段桂华(1972－),女,湖南新化人,副教授,CCF 高级会员,主要研究领域为匿名通信,网络安全.

王建新(1969－),男,博士,教授,博士生导师,主要研究领域为计算机算法,网络优化,生物信息学.

王伟平(1969－),女,博士,教授,博士生导师,主要研究领域为匿名通信,网络编码,信息安全.

杨路明(1947－),男,教授,博士生导师,主要研究领域为网络安全与计算机理论.