基于重路由匿名通信系统的设计

王少辉 蒋季宏 肖 甫
(南京邮电大学计算机与软件学院 南京 210003) (江苏省无线传感网高技术研究重点实验室 南京 210003)

摘 要 在分析现有匿名通信系统的基础上,设计了一种新的基于重路由的匿名通信方案。针对如何实现用户与用户之间通信匿名性的问题,提出了一种结合重路由机制中的变长策略和下一跳路由选择策略的方法来建立匿名通信路径,该方法还引入了概率转发机制和加密机制。针对如何增强匿名通信系统稳定性和抗攻击能力的问题,提出采用多服务器协调的P2P工作方式来设计匿名通信系统;并且新方案引入了分片冗余机制,在维护匿名转发路径的同时,也对通信消息提供了保护。通过理论分析和仿真分析可知,该方案设计具有较好的稳定性和匿名性。

关键词 匿名通信,重路由机制,加密机制,分片冗余
中图法分类号 TP309 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2016.10.029

New Design of Rerouting-based Anonymous Communication System

WANG Shao-hui JIANG Ji-hong XIAO Fu
(College of Computer, Nanjing University of Posts and Telecommunications, Nanjing 210003, China)
(Jiangsu High Technology Research Key Laboratory for Wireless Sensor Networks, Nanjing 210003, China)

Abstract Based on the analysis of the current anonymous communication system, a new rerouting-based anonymous communication scheme was proposed in this paper. To realize the anonymous communication among different users, a new method is presented to combine the variable length strategy and the next-hop routing selecting strategy in the rerouting mechanism to establish the anonymous communication path. Besides, this new method also introduces the probabilistic forwarding mechanism and encryption mechanism. Moreover, we applied the P2P work mode which is based on multi-server coordinating technology to enhance the stability and resistibility of the anonymous communication system. The new scheme also introduces the fragmentation-redundancy mechanism to protect the communication messages, while maintaining the anonymous forwarding path. Theoretical analysis and experimental results show that our scheme has good stability and anonymity.

Keywords Anonymous communication, Rerouting mechanism, Encryption mechanism, Fragmentation redundancy

1 引言

随着计算机网络应用的普及,互联网作为通信与信息传播的工具正快速地发展并且广为人们所接受,与此同时,信息安全的问题也备受人们的关注。在安全和隐私方面,运用最为广泛的加密技术,可保证数据的安全性,但攻击者可以通过报文分析获取通信双方的IP地址和端口信息,且通过流量分析还可以推测出一些其他有价值的信息,这些缺陷促进了匿名通信(Anonymous Communication)技术的发展。

匿名通信技术指通过一定的技术方法和手段来隐藏网络中通信实体的网络地址、实体间的通信关系等隐私信息,使攻击者无法直接窃听获取或推知双方的通信关系或通信的任何一方的身份信息,从而使网络实体参与双方的身份信息及行为信息得到更好的保护。其中有3种保护形式:发送者匿名(Sender Anonymity)、接收者匿名(Recipient Anonymity)、发送者和接收者不可连接性(Unlinkability of Sender and Recipient)。

关于匿名通信技术的发展,1981年David Chaum提出MIX技术[1],用于解决电子邮件的匿名问题;Reed等人[2]在1998年提出洋葱路由(Onion Routing)算法,在此算法的基础上,实现了洋葱路由匿名通信系统,随后Tor(第二代基于洋葱路由算法的匿名通信系统)[3]系统也相继推出。Tor是一个低时延基于TCP流的系统,主要目的是抵御攻击者对流量进行分析,保护用户的隐私;Crowds系统[4]由AT&T实验室提出,主要提供匿名Web访问,其构造了一个由很多匿名用户组成的群组,通信发送方隐藏在众多的群组成员中,任何一个成员都可以充当中继节点,将通信消息随机转发给下一个

到稿日期:2015-10-24 返修日期:2015-12-12 本文受国家自然科学基金资助项目(61572260,61373017,61572261),江苏省科技支撑计划基金项目(BE2015702),南京邮电大学校项目(NY214064,NY213036)资助。
王少辉(1977—),男,博士,副教授,主要研究领域为信息安全、密码学,E-mail:wangshaohui@njupt.edu.cn;蒋季宏(1991—),男,硕士生,主要研究领域为网络安全,E-mail:jjhfighting@sina.cn;肖 甫(1980—),男,博士,教授,主要研究领域为无线传感器网络、多媒体技术。

群组成员，从而实现发送者匿名。

匿名通信系统按实现技术可以分为两大类：1）基于广播的匿名通信系统（Broadcast-based Anonymous Communication System，BACS）；2）基于重路由的匿名通信系统（Rerouting-based Anonymous Communication System，RACS）。其中重路由机制是目前最常见的匿名通信技术，重路由匿名通信系统主要采用重路由机制进行保护，现有的实时匿名通信系统（如 Tor，Crowds 和 Hordes[5]等）都属于重路由匿名通信系统。

随着研究工作的不断深入，研究者发现现有的重路由匿名通信系统存在着很多不足，如 Tor 系统虽然很好地实现了发送者和接收者匿名，但其通过单路径进行消息转发和采用源路由方式构建匿名路径，路径中的首尾节点直接与发送者和接收者进行通信，这样导致系统存在一定的弱安全性。Crowds 系统采用下一跳路由选择方式来构建匿名通信路径，只能提供发送方的匿名服务，其路径上的每一个节点都能获知数据的内容和接收者的地址信息，从而对于前驱攻击抵御能力较弱。

本文在深入分析 Tor 和 Crowds 两类系统实现技术的基础上，针对现有匿名通信技术存在的不足，提出了采用分片冗余机制、重路由机制中变长策略和下一跳路由选择策略以及加密技术相结合的匿名通信系统的构建方案，较好地解决了发送者和接收者的同时匿名以及其相互关系的匿名问题，分析表明新方案具有较好的安全性和抗攻击能力。

2 相关技术介绍

2.1 重路由机制

重路由机制[6]是一种应用层路由机制，作为当前匿名通信系统中最为常见的一种技术，其基本实现原理是：在匿名通信过程中，发送者所发送的消息不是直接发送给接收者，而是先通过一个或多个中继节点进行转发，然后由路径中最后一个中继节点将消息传输给接收者。数据包经过路径上的中继节点时，其中的相关信息会被中继节点进行改写，在通信过程中，数据包所经过的多个中继节点形成了一条重路由路径。这样，网络中攻击者使用常规的监测手段将很难对数据包进行解析，因而可以有效地保护通信双方的身份信息。图 1 展示了一个基于重路由匿名通信系统的通信过程，系统中发送者和接收者进行通信时，通信消息都需经过多个中继节点进行转发，最后一个中继节点将消息转发给接收者。

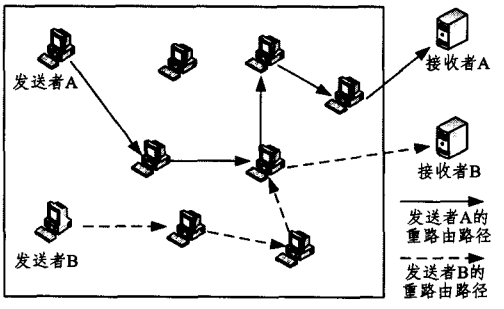

图 1 基于重路由机制的匿名通信系统

重路由路径的建立需要解决两个问题：1）如何选择路径上的中继节点；2）如何确定路径长度。关于这两个问题分别对应两种策略：成员选择策略和路长控制策略。其中成员选择策略有两种方法：随机策略和非随机策略。随机策略就是通过一定的算法从系统的 N 个成员中随机选择某一节点作为下一跳转发节点，而非随机策略会根据对某些参数（如节点的负载情况、运行的可靠性、节点的信誉度）的权衡来选择满足用户或系统所需要的节点作为路径上的中继节点。路长控制策略也包含两种方法：定长策略和变长策略。其中定长策略指路径上中继节点的个数为定值，此类策略都是通过源节点采用源路由方式一次性选取路径上所有中继节点结合嵌套加密机制来实现的，如 Tor 系统。而变长策略的路径长度则满足某一概率分布的离散随机变量，其采用下一跳路由选择方式来建立匿名通信路径，典型代表有 Crowds 系统。

根据重路由路径建立方式的不同，重路由匿名通信系统可分为 3 类：1）基于源路由方式的匿名通信系统，匿名路径上的所有中继节点由初始消息的发送者决定；2）基于下一跳路由的匿名通信系统，系统中中继节点由路径上的前驱节点确定；3）将源路由方式与下一跳路由方式相结合的方法[7]。

2.2 分片冗余机制

分片冗余机制是在数据存储和共享系统中用来提高数据可用性的一种机制，本文引入此机制对通信消息进行处理以提高系统的稳定性和抗攻击能力，其实现原理为：首先将消息分割成 k 片，再将此 k 片消息编码形成 n 片冗余消息（$n>k$），然后通过多路径将 n 片消息进行转发，最终发送给接收者，接收者在接收到任意 k' 片（$k' \geqslant k$）消息的情况下，均可通过译码重构成原始信息。分片冗余机制的过程如图 2 所示。

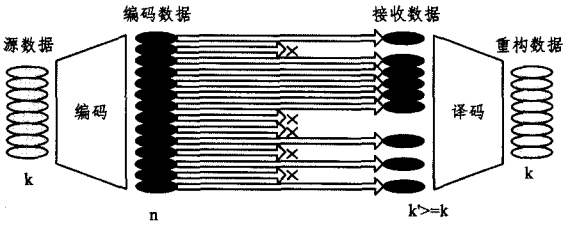

图 2 分片冗余机制的过程

分片冗余机制不仅对发送者消息提供了保护，而且具有维护匿名转发路径的功能。网络中攻击者截获的信息都经过分片编码，所以其既不完整也和原始消息毫无关联，其次，即使某条或某几条路径上的分片消息由于网络问题或节点的单点失效问题在转发过程中丢失，接收者仍然可以根据接收到的分片消息通过译码重构出原始信息，这样有效地减少了路径重构的概率，增强了系统的稳定性和抗攻击能力。

2.3 加密技术

加密技术是匿名通信系统中最为常用的一种技术，主要用来保护通信流的具体内容。匿名通信系统中报文信息或相关消息如果不进行加密，则会以明文的形式在通信过程中传递，这样显然存在巨大的安全隐患，网络中的攻击者可以轻易截获通信消息进行分析，从而获取通信双方的隐秘信息，所以加密技术的运用在匿名通信技术中不可或缺。

运用最为广泛的两类加密技术为对称加密和非对称加密。对称加密是指算法使用同一密钥进行加密和解密，其常

用的算法有 AES、RC4、DES 等。非对称加密是指算法使用不同的密钥进行加密和解密，其代表算法有 RSA、ECC 等。

加密技术的应用不仅体现在对通信流的内容保护上，而且在匿名通信系统中运用加密技术也可以很好地避免网络攻击者通过流量分析来识别通信双方的关系和身份信息，从而增加了系统的安全性。

3 一种重路由匿名通信系统的方案设计

3.1 设计思想

首先一个匿名通信系统需要建立自己的匿名通信网络且对其具有维护功能，匿名通信网络通常是指由匿名代理所构成的通信网络，新方案中匿名网络由不同网络中具有匿名需求的用户注册登录系统服务器所形成，网络中用户必须为其它用户提供相关服务，如提供消息转发功能。

匿名通信网络建立后，方案采用转发机制来实现匿名功能，采用重路由机制的变长策略和概率转发机制来控制匿名路径上中继节点的个数，采用类似 Crowds 系统的下一跳路由选择方式来建立匿名通信路径。不同的是系统中的下一跳节点都是由服务器根据网络在线用户选取，而不是由用户根据自身邻居节点的情况来选取，在通信过程中只需要服务器知道接收者，路径中节点只具有转发或接收的功能。

对于匿名通信系统来说，在建立了一条到接收者的匿名通信路径后，应该采用相关策略对路径进行维护，现有匿名通信系统对于中继节点失效问题，均采用路径重构的方法，这很容易受到恶意节点的攻击。因此，新方案引入分片冗余机制，通过多路径转发冗余信息，系统中个别节点的失效并不会导致系统对路径进行重构。

此外，匿名通信系统中各个节点还需具有数据处理功能，方案中每一个节点在发送消息前，首先对消息进行分片，再通过冗余编码形成多片冗余信息，在接收到信息后，可以通过对接收到的多个信息片采用相关译码重构出原始信息。转发消息时，由于该消息是对原始消息进行过分片编码的片信息，因此只需要单纯地转发给下一跳节点。

3.2 系统架构

方案采用多服务器下协调的杂 P2P 工作方式，其基本架构如图 3 所示，服务器保存匿名网络的完全拓扑信息，系统则利用服务器来随机生成中继节点，虽然整个过程中服务器充当中介的角色，但实际网络通信中各个用户终端之间进行直接的点对点交互。

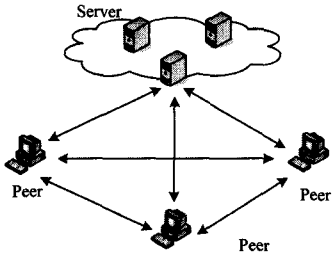

图 3 多服务器协调的 P2P 工作方式

方案中下一跳节点信息通过节点动态地向服务器请求获取，对于系统来说服务器是一个很大的安全隐患，当攻击者攻破服务器后，就会得到匿名通信路径的相关信息，从而使用户失去匿名功能。针对此问题，方案提出多服务器协调的杂 P2P 工作方式，服务器间没有联系，一条匿名通信路径由多个服务器共同构建完成，这样有效地阻止了攻击者通过攻破服务器的方式来获取通信路径的全部信息，此外也降低了单一服务器的负载。

每个用户节点有两类通信操作：1）节点和服务器之间的通信采用面向连接的 TCP 协议，用户通过此连接登录服务器，同时用户将其相关信息（如地址、公钥等信息）通过服务器公钥加密后发送给服务器，服务器通过解密后得到信息，在通信过程中接收者的地址信息都采用服务器密钥加密，多服务器采用同一个密钥，服务器将下一跳地址发送给用户时，会包含一个命令项，命令项包含 FORWORD_ADDR（转发消息）和 RECEIVE_ADDR（接收消息）两类；2）节点与节点之间的通信采用 UDP 协议，通过解析命令项来区分下一跳节点是中继节点还是接收节点。

此系统架构下，方案采用重路由机制中的变长策略和下一跳路由选择策略的方法来构建匿名通信路径，从而保证了系统的匿名度，同时可以对用户隐私进行保护。

3.3 匿名通信过程

在详细讲述匿名通信过程的不同阶段之前，首先在表 1 中列出通信系统的常用符号表示。

表 1 系统常用符号表示

符号	含义	符号	含义
M	用户名	C	用户名的密文
K	服务器密钥	Pub_i	第 i 个节点的公钥
Sec_i	第 i 个节点的私钥	IP_i，$Port_i$	第 i 节点的地址信息
ID	不同匿名路径的标示	CMD	命令项
Info	通信内容	S	密文

在基于重路由匿名通信系统中，最为关键的步骤是匿名通信路径的建立，匿名通信路径指的是信息在从发送者到接收者所经过的一条路径上，怎样建立路径直接关系到该系统的匿名性。图 4 展示了本文系统架构下用户 A 和用户 B 之间的一条通信路径，通信路径 $Path=\{A, I_1, I_2, \cdots, I_n, B\}$，新方案采用下一跳路由选择方式构建匿名通信路径，由图 4 可知路径上的每一个节点都会执行路径成员选择阶段和消息转发阶段。

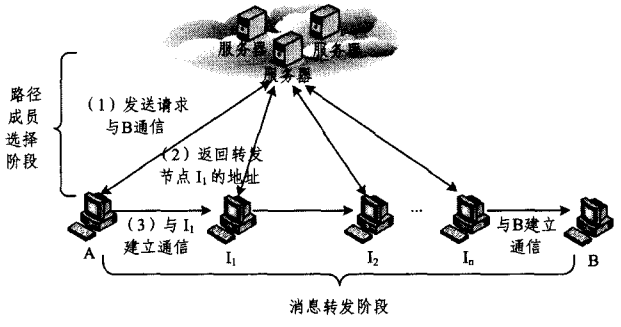

图 4 方案的通信框架图

（1）路径成员选择阶段

此阶段，方案通过服务器的介入来完成匿名通信路径中节点的选择，路径中的成员由用户向服务器请求而来，其中接收节点采用服务器密钥加密后在路径中被传输，这样实现了

接收者匿名。其具体步骤如下。

步骤 1　用户首先向服务器发送通信请求消息，以获取下一跳节点的地址信息。当用户为源节点时，服务器接收到请求后，采用自身密钥对接收节点的用户名 M 进行加密：

$$M \xrightarrow{K} C$$

得到该用户名的密文 C，然后服务器从在线用户中选出中继节点作为转发节点，且中继节点不能是接收节点，否则需重新选择。当用户是中继节点时，则对接收到的密文 C 采用解密操作：

$$C \xrightarrow{K} M$$

从而得到接收节点的用户名，服务器从在线用户中以概率 $P_f \in [0,1]$ 选择中继节点作为转发节点，或者以概率 $(1-P_f)$ 终止转发，直接选择接收节点。

步骤 2　在随机选出下一跳节点后，服务器将进行一次比较，看选出的节点是否为接收节点，然后采用命令项 FORWORD_ADDR（转发消息）和 RECEIVE_ADDR（接收消息）来区分下一跳节点是中继节点还是接收节点，同时对命令项 CMD 利用下一跳节点的公钥进行加密：

$$CMD \xrightarrow{Pub_i} S$$

密文 S 只有采用此节点的私钥才可以被解密。服务器的响应消息所包含的信息为 $\langle IP_i, Port_i, CMD, ID, C\rangle$，其中 ID 用来标识不同匿名路径，且其采用服务器密钥加密，加密操作和对用户名操作一样，得出标志信息的密文 ID'。最后服务器将响应消息 $\langle IP_i, Port_i, S, ID', C\rangle$ 发送给用户。

(2)消息转发阶段

为了实现消息传送的高效性和保护报文消息的头部信息，方案结合概率转发机制和加密机制来实现消息转发，其具体步骤如下。

步骤 1　用户在得到服务器发来的响应报文后，通过解析报文消息获得下一跳节点的 IP 地址和端口信息，将响应报文的剩余消息和通信信息一起发送给下一跳节点，其中消息包含的信息有 $\langle S, C, Info\rangle$。

步骤 2　节点在收到用户发来的信息后，首先对报文消息进行解析，获得信息 S，之后利用自己的私钥对信息 S 进行解密操作：

$$S \xrightarrow{Sec_i} CMD$$

解密操作后节点获得了命令信息 CMD，如果命令信息为 FORWORD_ADDR，则该节点进入路径成员选择阶段，获取路径上的下一跳节点；如果命令信息为 RECEIVE_ADDR，则该节点为接收节点，路径建立完毕。如建立一条通信路径 $Path=\{I_0, I_1, I_2, \cdots, I_n\}$，路径上任意节点 I_i 都满足如下三元函数关系：

$$p(I_i)=\begin{cases} p(I_0)=1, & A \rightarrow I_1 \\ 0<p(I_i)<p(I_{i-1}), & I_i \rightarrow I_{i-1} \\ p(I_{i-1})<p(I_i)<1, & I_i \rightarrow B \end{cases}$$

其中，A 代表发送者即 I_0，B 代表接收者即 I_n，$P(I_i)$ 为服务器接收到节点 I_i 的通信请求消息后随即生成的一个位于 0～$P(I_{i-1})$ 之间的值，当 $P(I_i) \in [0, P(I_{i-1})]$ 时，继续向后继节点 I_{i+1} 转发，否则节点 I_i 就直接发送给接收者 B。

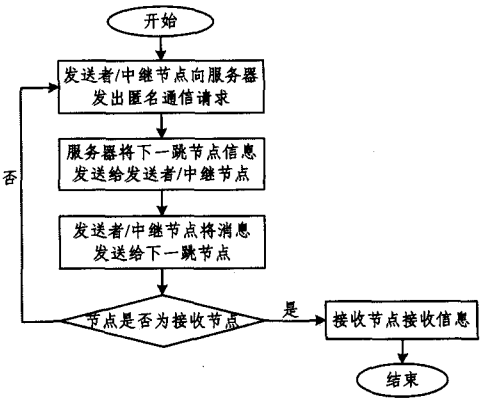

图 5　匿名通信流程

在匿名通信路径构建过程中，方案结合了概率转发机制和加密机制，其中概率转发机制采用概率递减转发算法，加密机制分别采用了对称加密和非对称加密两种加密方法，路径中所有节点都经过路径成员选择阶段和消息的转发阶段，最终将通信信息发送给接收节点。具体通信过程如图 5 所示。

4　方案分析

4.1　安全性分析

随着匿名通信技术的出现及不断发展，针对不同匿名系统的各种攻击方法也层出不穷。当前比较典型的攻击方法有合谋攻击[8]、前驱攻击[9]、流量分析攻击[10]、时间攻击[11]等。下面分别针对这些主流的攻击方式来分析第 3 节基于重路由匿名通信系统的新方案的抗攻击能力。

(1)合谋攻击。该攻击是攻击者利用系统存在的漏洞，加入到系统中成为系统用户，多个攻击者通过串通联合，再通过各节点所享有的权限和系统信息来推断系统中用户的匿名隐藏的关系。采用该攻击方式，攻击者需要控制系统的目录服务器，在建立匿名通信路径时，诱使系统选择合谋节点作为转发节点。对于此类攻击，最好的办法就是尽可能将合谋节点挡在匿名通信路径之外。Tor 系统不能很好地抵御该攻击，一旦服务器被控制，攻击者就很容易得到系统中通信用户的信息。新方案提出构建多服务器的方法，在建立匿名通信路径的过程中，每一跳中继节点都是由当前节点随机向某一服务器请求而来，即单一服务器只构建了一条路径中的部分路径，这样有效地防止在单服务器被控制的情况下出现攻击者采用合谋攻击致使路径中大部分节点为合谋节点的情况，从而增强了系统的抗攻击能力。

(2)前驱攻击。该攻击是一种内部联合攻击，攻击者首先加入匿名通信系统，成为系统成员，当在通信过程中成为中继节点时，就可以收集信息来推测通信双方的身份信息，最后判断出发送者和接收者。在攻击过程中，通过路径上的恶意节点在通信过程中发生故障即单点失效，而不能转发路径中的信息，导致系统必须多次重构路径。在这里前驱攻击者认为，在多次路径重构过程中源节点出现在攻击者前驱位置的概率是最大的，因此只要在通信过程中系统路径重构次数 N 足够大，根据大数定理，任何匿名通信的方式都可以攻破。新方案针对此类攻击需要路径多次重构的前提，提出采用分片冗余机制对原始信息进行分片编码，产生多条冗余信息并通过不同路径转发，最后发送给接收者，接收者接收到其中部分分片信息就可以通过解码的方式获得原始信息，这样在通信过程

中部分路径出现失效并不需要对路径进行重构，从而攻击者很难实施此类攻击。

(3)流量分析攻击。该攻击是一种普遍有效的攻击手段，在系统通信过程中，攻击者从网络流量进行监测和分析，其主要针对通信流的特征，而不是对通信流的内容进行分析。正常攻击者从网络流量经过的时间、报文大小、报文类型等方面来发现发送者和接收者之间的关系信息或发现一些流量进出中继节点的关联性，从而推断出发送者和接收者。但由于本方案采用多服务器协调的杂 P2P 的工作方式来建立匿名通信路径，因此整个网络中包含各种报文，不仅仅是节点间的通信报文，还有节点与各服务器的通信报文等，故匿名网络中每个节点进入和出去的实时流量会比较混杂，且系统采用多路径转发，这样每个节点实时联系的节点数也会比较多，攻击者想要发起流量分析攻击会相当困难，根据报文时间信息来获取报文间的关联也会十分困难。

(4)时间攻击。该攻击主要通过网络中攻击者观察和监听匿名网络中的通信流量，从报文的时间标志入手，通过对比来分析报文间的联系。此种攻击方式比较适合用于低时延匿名通信系统，该攻击与流量分析攻击一样，由于新方案中网络里包含各类报文信息，因此攻击者想根据报文的时间标志信息来获得报文间的关联会很困难。

上述几类攻击是现今匿名通信系统中比较常见的几种攻击方法，通过对各类常见攻击方法的分析，本文的重路由匿名通信系统的设计方案具有一定稳定性和抗攻击能力。

4.2 匿名性分析

新方案采用下一跳路由选择方式来构建匿名通信路径，通信过程中将发送者的消息通过多次转发后发送给接收者，所以发送者得到了很好的匿名，在此过程中下一跳节点选取是通过节点向服务器请求获得，在通信过程中，接收者在各节点传输时都采用服务器的密钥加密，这样有效地隐藏了接收者的身份信息，且在此过程中接收者也是某节点的下一跳节点，只是在数据包发送到接收节点时，通过对报文的解析，得知是否要接收信息，所以通信路径中某节点的下一跳节点有一定的概率为接收节点，其前驱节点有一定的概率为发送节点，两者的匿名度一样。

针对衡量一个匿名通信系统提供匿名保护有效性的能力的问题，研究者们提出匿名度[12]这一概念来衡量一个系统的好坏。下面对方案提出的重路由匿名通信系统的安全性，即匿名度问题进行分析，图 6 为一个简化的基于下一跳路径选择策略的匿名通信系统模型。

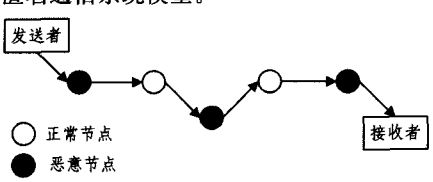

图 6 带有恶意节点的基于下一跳路径选择策略的匿名通信系统模型

从图中可以看出系统中的每个中继节点都只知道其前驱节点和后继节点，所以攻击者要想知道源发送节点和接收节点是谁，就必须同时满足两个条件：1)攻击者必须掌握源节点后的第一个中继节点，即源节点第一跳的选择就命中了系统的恶意节点，同时也必须获得路径中最后一跳中继节点；2)攻击者至少要能控制整个转发路径上一半的中继节点。

图 6 展示了一条被攻击者掌控的匿名通信路径，图中阴影节点为恶意节点，白色节点为正常节点，由于恶意节点间隔地分布在整条路径上，并且源节点的下一跳中继节点也为恶意节点，因此通过共谋攻击，攻击者可以实现对整条路径的还原，并最终发现源节点。

在路径长度为 L 的匿名通信路径上，攻击者能够知道源发送节点 S 的概率如下：

$$p(S)=\begin{cases}C^{L/2+1}, & L \bmod 2=0\\ C^{(L+1)/2}, & L \bmod 2=1\end{cases}$$

其中，C 为恶意节点占总节点数的比例。

由于新方案提出的重路由匿名通信路径 L 的长度是不固定的，且系统采用多路径转发，因此对每一条可能路径上源节点或接收节点被发现的概率求和，就是整个匿名系统中源节点被发现的概率：

$$p(S)=\sum_{l=1}^{\infty} f^{l-1}(1-f)C^{\lceil (l+1)/2\rceil}=(1-f)\frac{fC^2+C}{1-f^2C}$$

根据 Reiter 和 Rubin[13]提出的采用概率对基于下一跳选择策略的匿名通信系统匿名度定量和定性的分析方法，在将匿名性分为 6 个等级后，采用概率 $p(x)$来表示攻击者认定发送者或接收者的概率，$1-p(x)$表示系统的匿名度，其前提条件是系统用户必须属于同一个匿名集 S，所以匿名度表示为：

$$D=1-p(x)$$

图 7 给出了转发概率 f 与匿名度间的函数关系，图中 3 条曲线代表不同恶意节点比例下的函数关系，可以看出恶意节点比例越小，匿名度越大。同时随着转发概率 f 的增大，匿名度也在增大。

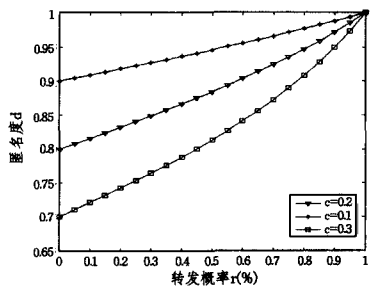

图 7 转发概率 f 与匿名度间的函数关系图

根据香农信息论，文献[14]提出了一种基于熵值的匿名度定量衡量方法，将系统的匿名度定义为：

$$D=\frac{E(x)}{E_{\max}}=\frac{-\sum p(x)\log(p(x))}{\log(N)}$$

图 8 示出恶意节点占总节点数的比例 C 与匿名度间的函数关系，图中 3 条曲线代表不同转发概率的函数关系，由图可知，转发概率越大，匿名度越大。同时随着恶意节点占总节点的比例增大，系统的匿名度降低。

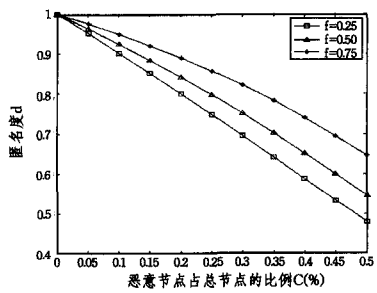

图 8 恶意节点占总节点数的比例 C 与匿名度间的函数关系图

根据上述对方案的分析，当系统存在少量的恶意节点时，只要系统用户足够多，不管转发概率 f 为多少，系统匿名度都保持在 0.7 以上，达到了 Beyond Suspicion 的等级，即无法猜测，通信实体中用户都有可能是真实的通信者，所以该方案提出的重路由匿名通信系统具有较好的匿名性，达到了良好的匿名通信效果。

4.3 性能分析

在下一跳路由方式的匿名通信系统中，如何很好地控制路径长度是备受关注的问题。Crowds 系统是一个典型的基于随机转发的匿名通信系统，其采用固定的转发概率建立重路由路径。假设转发概率为 P_f，由此可以得出其路径长度分布为：

$$p(L=k)=p_f^{k-1}\cdot(1-p_f),1<k<+\infty$$

而转发路径长度的期望值为：

$$E(L)=(1-P_f)\sum_{k=0}^{\infty}(k+1)(P_f)^k=\frac{1}{1-p_f}$$

从上式可以看出，转发概率 P_f 决定了路径的长度，显然 P_f 越大，路径则越长，这种方案会导致路径无限长而使建路失败，同时路径增长也会增加系统负载、通信延迟、性能开销等问题。

本方案采用转发概率等比递减的随机转发策略，等比递减系数设为 q。现设初始的转发概率为 $Initial\_P_f$（简称 IP_f），其取值范围为 0.5 到 1，由此可以得出其路径长度分布为：

$$p(L=k)=q^{\frac{k^2-3k+2}{2}}\cdot IP_f^{k-1}\cdot(1-IP_f\cdot q^{k-1})$$

则转发路径的长度期望值为：

$$E(L)=\sum_{k=1}^{\infty}k\cdot q^{\frac{k^2-3k+2}{2}}\cdot IP_f^{k-1}\cdot(1-IP_f\cdot q^{k-1})$$

图 9 给出了递减转发概率和传统的固定转发概率下的通信路径长度期望值的对比结果。

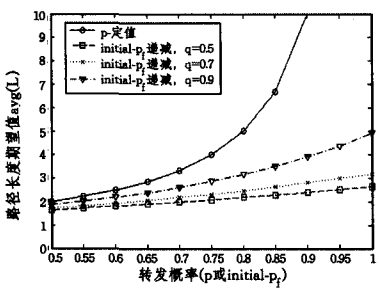

图 9 两种策略下的通信路径长度期望值的比较

从图 9 可以看出，递减转发概率算法下的通信路径长度远远小于基于固定概率随机转发策略下的通信路径长度。当 $Initial\_P_f$ 初始值为 1 时，通信路径也具有很好的收敛性，而固定转发概率时，转发概率 P_f 稍大，路径长度就急剧上升，导致路径过长而影响通信效率。所以本方案可以很好地控制通信路径长度，同时提高了系统的性能，缩短了消息的传播时延。

结束语 本文通过对现有主流匿名通信系统（如 Tor、Crowds）和相关技术进行分析研究比较，提出基于重路由的匿名通信系统的新的设计方案，该系统融合了冗余分片机制、加密机制、重路由机制、多路径等多种机制，采用多服务器协调下的杂 P2P 工作方式来建立匿名通信路径，实现了匿名通信系统的强匿名。理论分析表明新方案具有较高的稳定性和较强的抗攻击能力，通过仿真分析表明其具有较好的匿名性。下一步将侧重研究匿名通信系统如何控制匿名滥用的行为，以及如何实现系统的访问控制机制和如何撤销非法用户的匿名权利。

参 考 文 献

[1] Chaum D L. Untraceable electronic mail, return addresses, and digital pseudonyms [J]. Communications of the ACM, 1981, 24 (2): 84-90

[2] Goldschlag D, Reed M, Syverson P, et al. Onion routing for anonymous and private internet connections [J]. Communications of the ACM, 1999, 42(2): 39-41

[3] DingledineR, MathewsonN, SyversonP. Tor: Thesecond-generation onion router [C] // Proceedings of the 13th Usenix Security Symposium. 2004: 303-319

[4] Reiter M K, Rubin A D. Anonymous Web transaction with crowds [J]. Communications of ACM, 1999, 42(2): 32-48

[5] Berthold O, Federrath H, Kopsell S. Web Mixes: A System for Anonymous and Unobservable Internet Access [C] // Proceedings of Designing privacy Enhancing Technologies: Workshop on Design Issues in Anonymity and Unobservability. Springer, Heidellberg, 2000: 115-129

[6] Wright M, Adler M, Levine B, et al. An analysis of the degradation of anonymous protocols [C] // Proc of Network and Distributed System Security Symposium. California, 2002: 34-43

[7] Liu Pei-peng, Wang Li-hong, Shi Jing-qiao, et al. Towards Analysis of Security in I2P's Path Selection [J]. Journal of Computer Research and Development, 2014(7): 1555-1564(in Chinese)
刘培朋，王丽宏，时今桥，等. 匿名网络 I2P 路径选择的安全性分析[J]. 计算机研究与发展，2014(7)：1555-1564

[8] Wright M, Adler M, Levine B N, et al. Defending anonymous communication against passive logging attacks [C] // Proc of IEEE Symposium on Security and Privacy. Berkeley, 2003: 28-41

[9] Matthew W, Micah A, Brian N L, et al. Defending anonymous communication against passive logging attacks [C] // Proceedings of the 2003 IEEE Symposium on Security and Privacy (IEEE S&P 2003). IEEE Computer Society Press, 2003: 28-43

[10] Sen S, Wang Jia. Analyzing peer-to-peer traffic across large networks [J]. IEEE/ACM Transactions on Networking, 2004, 12 (2): 219-232

[11] Mitomo M, Kurosawa K. Attack for Flash MIX [C] // Proceedings of the 6th International Conference on the Theory and Application of Cryptology and Information Security. Springer-Verlag, 2000: 192-204

[12] Zhang Jia, Duan Hai-xin, Liu Wu, et al. Anonymity analysis of P2P anonymous communication systems [J]. Computer Communications, 2010, 34(3): 358-366

[13] Reiter M, Rubin A. Crowds: Anonymity for Web Transactions [J]. ACM Transactions on Information and System Security, 1998, 1(1): 66-92

[14] Diaz C. Anonymity and Privacy in Electronic Services [D]. Electrical Engineering department of Katholieke University Leuven, 2005: 23-40