一种可控路长的 P2P 匿名通信协议

邓 琳,谢 鲲,李仁发,练 琪
DENG Lin,XIE Kun,LI Ren-fa,LIAN Qi

湖南大学 计算机与通信学院,长沙 410082
School of Computer and Communication,Hunan University,Changsha 410082,China
E-mail:denglin0820@126.com

DENG Lin,XIE Kun,LI Ren-fa,et al.Length controllable protocol for P2P anonymous communication.Computer Engineering and Applications,2010,46(7):110-114.

Abstract: P2P anonymous communication system brings well scalability;however,it is still a difficult problem to trade off anonymity and efficiency.Anonymous technology makes users obtain anonymous performance but contemporary often increases communication delay and member payload,sacrifices efficiency.Based on analyzing existing anonymous communication technology, this paper proposes LCPACP,a new length controllable protocol for P2P anonymous communication.It uses nested encryption to guarantee strong anonymity and adopts forwarding probability decreasing strategy to control path length to obtain high efficiency. The theory analysis and calculation results show that this protocol can not only shorten path,ensure well transmission performance,but also can provide good anonymous protection.

Key words: anonymous communication;P2P;nested encryption

摘 要:P2P 匿名通信系统带来了良好的可扩展性,然而要兼顾匿名和效率仍然是个难题。匿名技术在使用户获得良好匿名性能的同时,往往增加了通信延时及成员负载,牺牲了效率。在分析现有匿名通信技术的基础上,提出了一种新的可控路长的 P2P 匿名通信协议 LCPACP(Length Controllable Protocol for P2P Anonymous Communication)。LCPACP 采用嵌套加密来保证强匿名性,利用转发概率递减的策略来有效控制重路由路径长度,以取得高效率。理论分析与计算结果表明,新的协议能显著缩短路长,保证良好的传输性能,同时能提供良好的匿名保护。

关键词:匿名通信;P2P;嵌套加密
DOI:10.3778/j.issn.1002-8331.2010.07.033 **文章编号:**1002-8331(2010)07-0110-05 **文献标识码:**A **中图分类号:**TP393.08

1 引言

随着计算机网络广泛应用,互联网上的安全和隐私越来越受到人们的关注。虽然密码技术很好地保护了通信内容本身,但不能阻止攻击者通过流量分析推出有价值的信息。作为保护隐私的一种方法, 匿名已成为许多实时因特网应用的基本需求。匿名技术通过一定的方法将通信关系加以隐藏,以达到保护通信双方的目的。自从 David Chaum[1]提出针对电子邮件的匿名邮件系统以来, 出现了许多匿名通信协议和原型系统,匿名通信协议的设计也逐渐引起了研究人员的重视。

由于传统的 C/S 模式的匿名通信系统存在单点失效、可扩展性差等问题,近年来,随着对等网 P2P 的兴起,人们开始关注 P2P 匿名通信协议的研究。P2P 网络为解决匿名通信问题提供了新的技术手段,带来了良好的可扩展性,但同时也面临着新的挑战。P2P 匿名系统设计中,增加匿名性往往会带来额外的消息传输和维护的代价,需要在强匿名性和系统高效性之间进行平衡。现有的 P2P 匿名通信协议,典型的有 Crowds[2]、P5[3]、Tarzan[4]、MorphMix[5]、Cashmere[6]。这些协议提供了不同程度的匿名性。Crowds[2]采用随机转发取得匿名,虽然效率很高,但它只保证了发送者匿名且匿名度较低,没有解决通信双方均匿名的问题,而且 Crowds 中路径长度没有限制,重路由路径过长会导致通信延时等性能开销过大,甚至可能使建路失败;P5[3]采用分层广播的思想建立匿名网络,但基于广播的系统使得网络传输的带宽增加,效率较低;Tarzan[4]和 MorphMix[5]都是基于 Mix 技术,采用嵌套加密获得了强匿名,但是在使得用户获得匿名性能的同时,却增加了消息头部开销和服务的延迟,牺牲了系统的效率;Cashmere[6]采用前缀路由的策略,它的中继由前缀相同的一组节点组成,但组内无论任何节点收到消息包都必须把消息包广播给组内其他节点,这样消耗了大量的网络带宽,效率较低,且它基于结构化的 P2P 网络,缺少非结构化系统的灵活性。

该文设计 P2P 匿名通信协议以取得良好系统匿名性的同

基金项目:教育部重点项目(the Education Ministry Grand Projects of China under Grant No.108168);湖南省自然科学基金(the Natural Science Foundation of Hunan Province of China under Grant No.06JJ2090)。

作者简介:邓琳(1983-),女,硕士生,主要研究领域为网络安全;谢鲲(1978-),女,博士,主要研究领域为可信系统与网络、网络测试等;李仁发(1957-),男,教授,博士生导师,主要研究领域为无线通信与网络、信息网络;练琪(1982-),女,硕士生,主要研究领域为 P2P 网络。

收稿日期:2008-09-17 **修回日期:**2008-11-17

时,尽量减少代价和提高效率为目的,提出了一种新的可控路长的 P2P 匿名通信协议 LCPACP。协议除了采用基于 Mix 的嵌套加密技术,还在随机转发中采用了转发概率等比递减的策略来有效控制重路由路径长度获得高效率。LCPACP 协议可提供双方匿名且无中心节点管理信息,完全分布式属性具有较好的可扩展性。理论分析和实验结果表明新协议在具有较高效率的同时提供了良好的匿名性。

2　LCPACP 协议

鉴于 Mix 嵌套加密的强匿名性及随机转发策略头部开销少、效率高的优点,为了增强匿名和提高效率,提出一种基于 Mix 和随机转发的 P2P 匿名通信协议 LCPACP。协议主要包括两个阶段:节点发现和路径建立。第一阶段的主要任务是在网络中发现节点,通过节点发现协议得到网络中其他节点的信息,为路径建立中的节点选择作准备。第二阶段的主要任务是选择节点构建匿名路径,它是匿名协议设计中很关键的一步,其设计的好坏关系到协议的效率和匿名性。这一阶段包括了节点的选择及加密操作。

下面将详细介绍协议的设计,表 1 给出了协议中一些符号的含义。

表 1　符号定义

符号	含义
M_0	原始消息
S	发送者
R	接收者
G_i 或 H_i	除发送者和接收者之外的节点 i
IP(·)	节点的 IP 地址
PK_i	第 i 个中间节点的公钥
$PK_i(\cdot)$	用节点 i 的公钥对消息进行加密
$SK_{i,j}$	节点 i 与 j 的共享密钥
$nonce_i$	随机填充数

图 1 给出了 LCPACP 协议的一个简单实例,其中标识了节点 S 与 G_1 各自发送的加密消息,S–H_1–H_2–G_1–H_3–R 是一条 LCPACP 路径,其中 H_1、H_2、H_3 是在发送的过程中随机选择的节点。在不存在随机选择的节点的情况下,S–G_1–R 则成为一条 Mix 路径。协议的基本思想是:

S 首先选择节点 G_1(其中 G_1 是在消息发送之前选择的,称其为固定节点)。然后利用 G_1 的公钥对 M_0 等参数进行加密,再用 S 与 H_1 的共享密钥进行对称加密后以概率 p_f 发送给 H_1,H_1 抛币决定是把该消息发给 G_1 还是发给一个随机选择的节点 H_2,结果 H_1 把消息发给了 H_2 但转发概率变成了 $p_f q$,并告诉其下一跳是 G_1。以后每个节点收到消息后也进行抛币决定是把该消息发给下一个节点还是发送给一个随机选择的节点。其中随机选择的节点 H_1、H_2、H_3 不进形嵌套加密,而是简单的对称

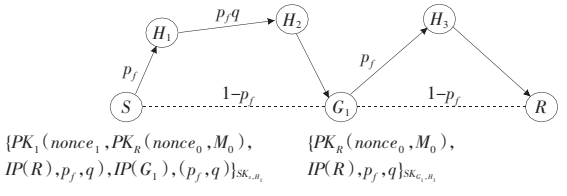

图 1　LCPACP 协议的简单实例

加密。

消息在随机节点之间传递时采用转发概率等比递减的策略来有效地控制重路由路径长度,降低传输延时,提高效率。

2.1　节点发现

在一个自组织的非结构化 P2P 网络中,每个节点开始仅仅知道局部的一些节点的信息,需要一个方法来获知其他活动节点。如果直接采用洪泛(flooding)方式发现对等节点,会产生大量的广播消息,发现过程占用大量的传输带宽,因此,该协议中采用类似 Tarzan[4]的基于闲谈(gossip)的协议来发现对等点,每个节点 a 都有一个缓存节点列表 Memlist(包含了 a 自己及其邻居的信息),数据格式为{$IP(\cdot)$,$port$,PK},其中 $IP(\cdot)$ 是节点的 IP 地址,$port$ 是节点的端口号,PK 是节点的公钥(每个节点首次加入网络时,都会通过 RSA 算法产生自己的公钥和私钥)。加入网络的每个节点 a 通过与已知节点交换邻居信息来得知网络中其他节点的信息,并将这些节点的信息加入到 Memlist 中。这个列表里既包含邻居节点(一个节点知道另一个节点的存在,并且建立了连接,则它们具有邻居关系),也包含无效邻居节点(一个节点知道另一个节点的存在但并没有建立连接)。过程简述如下:

(1)起初,新加入的节点 a 发送请求到已知的一些节点 b,以此来发现新的无效节点地址集合。

(2)当 a 收到来自 b 的回复时,说明连接建立成功,则 b 成为 a 的邻居节点,节点 a 才把节点 b 添加到自己所维护的邻居节点集中,a 与 b 之间便自发地不定期地交换 Memlist 列表,并将交换得来的信息加入 Memlist 列表中。

(3)节点 a 在重新连接已知的邻居节点之前继续连接新的无效邻居节点。

通过节点发现协议,每个新加入的节点通过与其他节点交换邻居信息就可以获得整个网络的拓扑信息。且这种方式完全依赖于节点之间的互相通信和协作,不需要中心节点的支持。

2.2　路径建立

在获知网络中所有节点信息之后,协议进入路径建立阶段,该阶段包括了节点的选择及加密操作,步骤如下:

步骤 1　由发送者 S 决定匿名路径上进行嵌套加密的固定节点数 m(发送者和接收者不包括在内)。S 首先从自己的 Memlist 列表中随机选择第 1 个固定节点 G_1,再通过 G_1 来选择第 2 个固定节点。G_1 收到来自 S 的 extend 命令(包含了参数 V)后,便延伸通道,采用基于 IP 地址前缀的选择算法(即选择具有不同 IP 地址前缀的节点,从而使匿名路径穿过多个域,从而可降低路径被攻击者观察到的概率)从 G_1 自己的 Memlist 列表中选择 V 个节点传给 S。S 从中挑选出一个作为第 2 个固定节点 G_2。接下来的每个固定节点都通过其上一个节点按照这种方法进行选择,从而构建一条固定的路径 S–G_1–G_2,…,G_i,…,R。

步骤 2　利用选出的固定节点的公钥进行加密,创建消息 M_1 并把它发送给 G_1,并用 G_{m+1} 表示接收者 R。用 M_i 表示对等点 G_i 收到的消息,M_R 表示接收者 R 收到的消息。

$$M_R = PK_R(nouce_0, M_0)$$

$$\cdots$$

$$M_i = PK_i(nouce_i, M_{i+1}, IP(G_{i+1}), p_f, q)$$

$$\cdots$$

$$M_1 = PK_1(nouce_1, M_2, IP(G_2), p_f, q)$$

其中,(p_f,q)是一个随机转发的概率参数,见步骤 3 说明。

步骤 3 在 LCPACP 中,在两固定节点之间采用转发概率递减的随机转发策略。眭鸿飞等人[7]在 Crowds 模型中采取此策略来避免 Crowds 路径无限长,这种递减概率转发的方式在此同样适用。递减方法采用等比数列递减,即若此时概率为 p_f,则在信息随机转发给随机选择的下一跳(此节点在 Memlist 邻居列表中选择的)而不是固定节点时,概率变为 p_fq,其中 q 是一个等比因子。具体如下:

发送者 S 首先确定好初始转发概率 p_f 及等比因子 q,并将 (p_f,q)包含在创建的消息中。然后 S 进行抛掷硬币来判断将消息提交给 G_1 还是转发到随机选择的节点,即 S 进行一次有偏的抛硬币操作,得到一个概率值 μ,S 将这个值与设定的转发概率门限值 p_f 进行比较,若 $\mu>p_f$,则 S 将消息直接提交给 G_1,若 $\mu<p_f$,则 S 从 Memlist 邻居列表中随机选择一节点,设为 H_1,将消息转发给 H_1。消息在 S 与 H_1 之间传递前,S 与 H_1 先协商好共享会话密钥 SK_{S,H_1},S 使用 SK_{S,H_1} 对消息及 (p_f,q)进行对称加密,然后将它们发给 H_1。节点 H_1 用会话密钥解密后继续进行抛硬币操作,这时转发给随机选择的节点的概率变为 p_fq。每随机转发一次概率就变为原先的 q 倍,假设消息被转发到 H_i,则 H_i 创建新消息以概率 p_fq^i 将其继续转发给随机选择的节点 H_{i+1},或以概率 $(1-p_fq^i)$将消息 M_1 提交给 G_1。

$S->H_1:\{M_1,IP(G_1),(p_f,q)\}_{SK_{S,H_1}}$

$H_i->H_{i+1}:\{M_1,IP(G_1),(p_fq^i,q)\}\}_{SK_{H_i,H_{i+1}}}$

采用的转发概率等比递减的选路算法如下(以 S 与 G_1 之间的路径为例):

Symbolic definition:H_0-S(Sender),H_R-G_1(Receiver),$H_1,\cdots,H_n$-randomly selective node,p_f-forwarding probability,q-geometric proportion factor,V-node sequence set on path,

H_0 choose p_f and q

//algorithmic initializing

$V=H_0$;

Sender=H_0;

μ=coin.flip(); //obtain a probability from flip coin

while(0<μ<1)

if($\mu\geqslant p_f$)

Sender.SendMSG(H_R,MSG) //send message"MSG"to H_R directly

exit Loop;

else

H_i=GetRandomNode() //choose a node H_i randomly from the Memlist table

MSG=encrypt(MSG,SK(Sender,H_i))

//Sender use Sender and H_i'share-key

SK(Sender,H_i)to encrypt MSG

Sender.SendMSG(H_i,MSG)//send MSG to H_i

MSG=decrypt(MSG,SK(Sender,H_i))

//H_i use SK(Sender,H_i)to decrypt MSG

Sender=H_i //Let node H_i be the Sender

$V=V\parallel H_i$

$p_f=p_f\times q$

μ=coin.flip();

end while

$V=V\parallel H_R$;

return V.

步骤 4 任意两个固定节点之间的路径都采用上面所讲的转发概率等比递减的算法来进行选择,假设消息传到固定节点 G_i,G_i 收到的消息 M_i 对其解密后,获得下一个固定节点 G_{i+1} 的地址 $IP(G_{i+1})$,然后重复步骤 3 的抛币操作来决定将消息直接发送给 G_{i+1} 还是随机转发给其他节点。即 G_i 根据发送者 S 在每一层设定的概率 p_f,以概率 $1-p_f$ 把 M_{i+1} 直接提交给 G_{i+1},或创建新消息以概率 p_f 将其转发给 H_j(此节点是由 G_i 从它的 Memlist 邻居列表中随机选择的)。H_j 收到此消息后,陆续用共享密钥和私钥解密后,得到下一个固定节点 G_{i+1} 的地址。H_j 继续与 G_i 类似的操作,抛币来决定将消息发送给 G_{i+1} 还是继续随机转发,若随机转发则采用上面所讲的转发概率等比递减算法进行操作。

$G_i->H_j:\{M_{i+1},IP(G_{i+1}),(p_f,q)\}_{SK_{G_i,H_j}}$

步骤 5 经过多个固定节点 $G_1,G_2,\cdots,G_i,\cdots$及随机选择的节点,消息 M_0 到达接收者 R。发送者 S 与接收者 R 之间建立的匿名路径的一个实例如图 2 所示。

图中,H_j^k $(j=1,2,\cdots;k=0,1,\cdots)$是随机选择的节点,$S$ 与 R 之间建立的一条 LCPACP 路径为 $SH_1^0H_2^0G_1H_1^1H_2^1H_3^1H_4^1G_2,\cdots,G_{m-1}H_1^{m-1}G_mH_1^mH_2^mH_3^mR$。此路径可看成 $m+1$ 个分段,两个相邻的固定节点之间的路径为一个分段,它们都是通过转发概率等比递减策略建立的,分别设为 $r_0,r_1,\cdots,r_{m-1},r_m$。

若发送者 S 要求接收者 R 回复消息,则 S 在发送消息时构造一个包含应答地址的信息放在消息中一起发给 R,如果路径建立成功,那么从接收者返回给发送者的消息反向沿着原路径发送。

通过 Mix 嵌套加密和随机转发概率等比递减策略相结合的思想建立的 LCPACP 匿名路径既提供了强匿名性,保证了发送者匿名和接收者匿名,又解决了传统的基于 Mix 的 P2P 匿名通信协议消息头部开销过大的问题,并且还有效地控制了重路由路径长度,提高了效率。

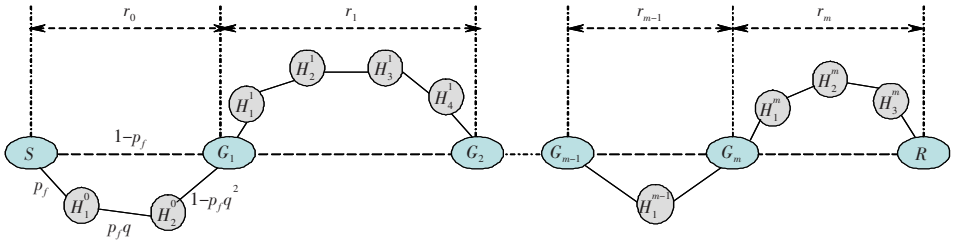

图 2 LCPACP 匿名路径

3　性能分析与评价

3.1　与其他协议比较

与完全基于 Mix 的 Tarzan、MorphMix 等 P2P 匿名协议比较，在路径长度相同的情况下,LCPACP 协议不使用所有中间节点的密钥进行嵌套加密,而只用在发送消息前选取的少数固定节点的公钥对消息进行嵌套加密，减少了嵌套加密层数,从而明显降低了消息头部开销。在两固定节点间的随机选取的节点之间的数据转发只用共享密钥进行简单的对称加密,从而减少了复杂的加密解密等资源开销。

与节点完全随机选择的协议如 Crowds 比较,LCPACP 协议中采用了 Mix 的嵌套加密思想来提供强匿名，抗攻击性更强，并且既实现了发送者匿名又保证了接收者匿名。LCPACP 还采用了转发概率递减策略,重路由匿名路径长度得到有效控制,降低了成员负载和通信延时,提高了效率。LCPACP 协议是完全分布式的,无中心节点,解决了 Crowds 协议的管理开销问题,并且所有节点都是对等的,可扩展性好。

3.2　匿名性分析

为了分析协议获得的匿名度,采用 Reiter and Rubin[2]提出的对匿名度的分析方法。攻击者由多个恶意节点组成,每个恶意节点占据了不同的位置。在某一个间隔时间内,网络中成员数目 N 固定为一个常数。这里所考虑的是匿名路径由非恶意者发起,攻击者能否确定一条路径的发送者是谁。从攻击者来看第 1 个恶意者的前驱节点比其他节点是发送者的可能性要大。针对一条从发送者到达接收者的 LCPACP 路径,发送者是位于第 1 个恶意者之前的,因此针对第 1 个恶意者进行研究。

令发送者位于第 0 个位置。$H_k(k>1)$表示第 1 个恶意者在路径的第 k 位置的事件;$H_{k+}=H_k \vee H_{k+1} \vee \cdots$表示第 1 个恶意者在路径的第 k 位置及之后的事件;$P(H_k)$表示路径的第 k 个位置取到了恶意者,而在此位置之前的节点选择的都是非恶意者的概率;I 表示路径上第 1 个恶意者的前驱是发送者,$P(I|H_{1+})$表示攻击者猜测正确的概率，即路径上有恶意者的情况下,第 1 个恶意者的前驱节点是发送者的概率。

下面分析 LCPACP 协议获得的发送者匿名度。

假设 N 个成员的 LCPACP 网络中有 C 个恶意者，令转发概率中的等比因子为 q,$(N-C)/N=w$,$C/N=1-w$。

匿名路径可看成由变化路径 $r_0,r_1,\cdots,r_{m-1},r_m$ 串接而成,其长度(即边的条数)可分别设为 $l_0,l_1,\cdots,l_{m-1},l_m$。则按照协议的转发机制,分析得到:

若第 1 个恶意者位于路径 r_0 的第 i 个位置$(1\leqslant i\leqslant l_0)$,则

$$P(H_i^0)=(1-w)(1\cdot wp_fq\cdot\cdots\cdot wp_fq^{i-1})=$$
$$(1-w)(wp_f)^{i-1}\prod_{j=0}^{i-1}q^j=(1-w)(wp_f)^{i-1}q^{\frac{i(i-1)}{2}}$$

则第 1 个恶意者在 r_0 上的概率:

$$P(H_{1+}^0)=(1-w)\sum_{k=1}^{l_0}(wp_f)^{k-1}q^{\frac{k(k-1)}{2}}$$

若第 1 个恶意者位于路径 r_1 上$(1\leqslant i\leqslant l_1)$,则

$$P(H_i^1)=(wp_f)^{l_0-1}q^{\frac{(l_0-2)(l_0-1)}{2}}w(1-p_fq^{l_0-1})(1-w)(wp_f)^{i-1}q^{\frac{i(i-1)}{2}}=$$
$$w(1-w)(1-p_fq^{l_0-1})(wp_f)^{l_0-2+i}q^{\frac{(l_0-2)(l_0-1)+i(i-1)}{2}}$$

$$P(H_{1+}^1)=w(1-w)(1-p_fq^{l_0-1})q^{\frac{(l_0-2)(l_0-1)}{2}}\sum_{k=1}^{l_1}(wp_f)^{l_0-2+k}q^{\frac{k(k-1)}{2}}$$

依次类推,若第 1 个恶意者位于路径 r_m 上$(1\leqslant i\leqslant l_m)$,则

$$P(H_i^m)=(1-w)(wp_f)^{i-1}q^{\frac{i(i-1)}{2}}(wp_f)^{\sum_{j=0}^{m-1}l_j-1}q^{\sum_{j=0}^{m-1}\frac{(l_j-2)(l_j-1)}{2}}w^m\times$$
$$\prod_{j=0}^{m-1}(1-p_fq^{l_j-1})$$

$$P(H_{1+}^m)=(1-w)w^m(wp_f)^{\sum_{j=0}^{m-1}l_j-1}q^{\sum_{j=0}^{m-1}\frac{(l_j-2)(l_j-1)}{2}}\left[\prod_{j=0}^{m-1}(1-p_fq^{l_j-1})\right]\times$$
$$\sum_{k=1}^{l_m}(wp_f)^{k-1}q^{\frac{k(k-1)}{2}}$$

为简化讨论,设 $l_0=l_1\cdots=l_m=l$,则匿名路径上有恶意者的概率:

$$P(H_{1+})=\sum_{j=0}^{m}P(H_{1+}^j)=$$
$$(1-w)\left(\sum_{k=1}^{l}(wp_f)^{k-1}q^{\frac{k(k-1)}{2}}\right)\sum_{t=0}^{m}w^t(wp_f)^{t(l-1)}(1-p_fq^{l-1})^tq^{\frac{t(l-2)(l-1)}{2}}=$$
$$(1-w)\left(\sum_{k=1}^{l}(wp_f)^{k-1}q^{\frac{k(k-1)}{2}}\right)\left[\frac{1-\left(w(wp_f)^{l-1}(1-p_fq^{l-1})q^{\frac{(l-2)(l-1)}{2}}\right)^{m+1}}{1-w(wp_f)^{l-1}(1-p_fq^{l-1})q^{\frac{(l-2)(l-1)}{2}}}\right]=$$
$$(1-w)u$$

其中

$$u=\left(\sum_{k=1}^{l}(wp_f)^{k-1}q^{\frac{k(k-1)}{2}}\right)\left[\frac{1-\left(w(wp_f)^{l-1}(1-p_fq^{l-1})q^{\frac{(l-2)(l-1)}{2}}\right)^{m+1}}{1-w(wp_f)^{l-1}(1-p_fq^{l-1})q^{\frac{(l-2)(l-1)}{2}}}\right]$$

当第 1 个恶意者位于路径上第 1 个位置时,它的前一个节点肯定是发送者,所以 $P(I|H_1)=1$。当第 1 个恶意者位于路径上第 2 个位置或之后时，由于出现任何非恶意者的概率是相同的,所以 $P(I|H_{2+})=1/(N-C)$。从而有:

$$P(I)=P(H_1)P(I|H_1)+P(H_{2+})P(I|H_{2+})=$$
$$(1-w)+\frac{1}{N-C}(P(H_{1+})-(1-w))=\frac{(1-w)(N-C-1+u)}{N-C}$$

在路径上有恶意者的情况下,攻击者猜测发送者正确的概率:

$$P(I|H_{1+})=\frac{P(I)}{P(H_{1+})}=\frac{N-C-1+u}{u(N-C)}$$

$P(I|H_{1+})$揭示了匿名的程度,可用来衡量匿名性,$1-P(I|H_{1+})$即为协议的匿名度。显然 $P(I|H_{1+})$越小,攻击者猜中发送者的概率越低,那么匿名性就越好。根据上面的计算结果,下面将考察攻击者猜中的概率随着(N:系统规模,C/N:恶意节点比例,p_f:初始转发概率,q:等比因子,m:固定节点数)的变化情况,对不同情况下的匿名性能进行比较。

图 3 是当 $C=100$,$p_f=0.75$,$q=0.8$,$m=9$,$l=3$ 时,攻击者猜中

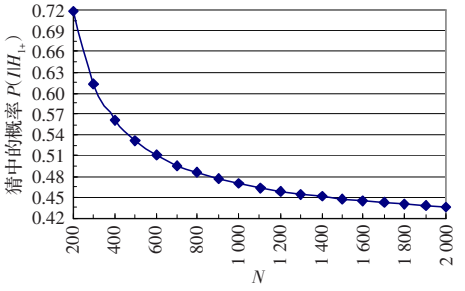

图 3　$P(I|H_{1+})$随 N 的变化趋势

的概率 $P(I|H_{1+})$随系统规模 N 的变化情况。从图中可以看出，系统规模 N 越大，猜测正确的概率开始减小越快；当 N 增大到一定规模后，概率减小的趋势变缓。$P(I|H_{1+})$越小，说明匿名度越大。图 4 显示的是当 N=5 000，p_f=0.75，q=0.8，m=9，l=3 时，攻击者猜中的概率 $P(I|H_{1+})$随 C/N 的变化趋势。猜中的概率 $P(I|H_{1+})$随着恶意节点比例 C/N 的增加而增大，匿名度逐渐变小。因为 C/N 越大，恶意节点越多，也就意味着匿名集越小，恶意节点比例必须控制在一定值内才能保证匿名性。

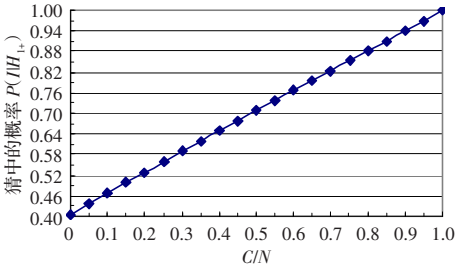

图 4 $P(I|H_{1+})$随 C/N 的变化趋势

图 5 给出了当 N=5 000，C=500，q=0.8，m=9，l=3 时，初始转发概率 p_f对概率 $P(I|H_{1+})$的影响。可见，$P(I|H_{1+})$值随 p_f增加而降低，p_f越大，$P(I|H_{1+})$值越小，匿名度越高。因为 p_f越大，消息转发给随机选择的下一跳的概率就越高，这就意味着匿名路径越长，通常情况下匿名性也会越好。图 6 是当 N=5 000，C=500，p_f=0.75，m=9，l=3 时，$P(I|H_{1+})$随等比因子 q 的变化趋势。由图可知，随着等比递减转发过程中概率的等比因子增大，$P(I|H_{1+})$值减小。这是因为随着 q 的增大，随机转发给下一节点的概率也增大，匿名路径也会增长，匿名性相对来说也就越好。用户可通过选择 p_f和 q 值来控制路径长度。

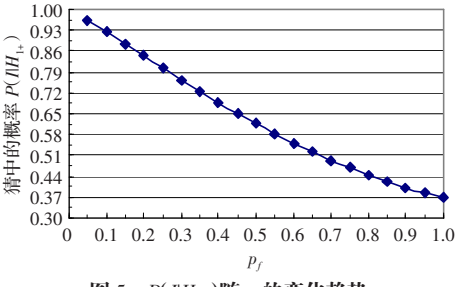

图 5 $P(I|H_{1+})$随 p_f的变化趋势

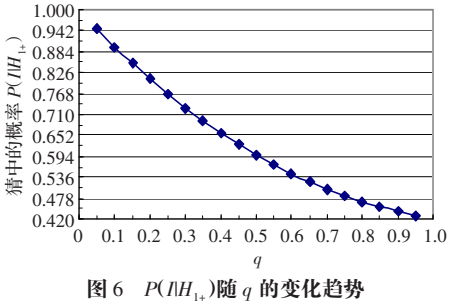

图 6 $P(I|H_{1+})$随 q 的变化趋势

当 N=5 000，C=500，p_f=0.75，q=0.8，l=3 时，$P(I|H_{1+})$随固定节点数 m 的变化趋势如图 7 所示。从图中曲线可以看出，在固定节点数小于等于 3 时，$P(I|H_{1+})$值随 m 增加而减小较快，随着 m 继续增加，$P(I|H_{1+})$减小得相当缓慢并逐渐趋向一定值。说明路径长度增加到一定长度情况下，继续增长对匿名性的影响是非常轻微的。m 越大，进行嵌套加密次数越多，从而提高了匿名性，但是消息头部开销及加解密开销也就越大，并且由图可知，m 在增长到一定值后对匿名性的影响甚微，因此，在实际应用中，m 不宜选择过大。

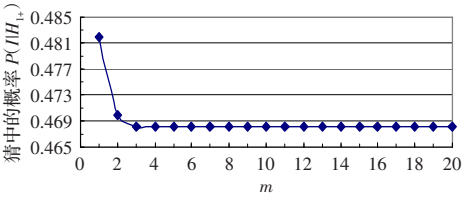

图 7 $P(I|H_{1+})$随 m 的变化趋势

4 结论

分析了现有匿名通信技术，出于加强匿名和提高效率两方面的考虑，提出了一种新的可控路长的 P2P 匿名通信协议 LCPACP。LCPACP 采用嵌套加密来保证强匿名性和转发概率递减的策略来有效控制路径长度获得高效率。证明了该协议下匿名度与系统规模 N、恶意节点比例 C/N、转发概率 p_f、等比因子 q 及固定节点数 m 之间的关系。分析结果表明，LCPACP 协议能提供良好的匿名性，N、p_f和 q 中任何一个参数的增加都会使协议的匿名度提高。

参考文献：

[1] Chaum D.Untraceable electronic mail,return addresses,and digital pseudonyms[J].Communications of the ACM,1981,24(2):84-88.

[2] Reiter M K,Rubin A D.Crowds:anonymity for web transactions[J].ACM Transactions on Information and System Security,1998,1(1):66-92.

[3] Sherwood R,Bhattacharjee B,Srinjvasan A.P5:A protocol for scalable anonymous communication[C]//Proceeding of IEEE Symposium on Security and Privacy,Berkley,California,May 2002:58-70.

[4] Freedman M J,Morris R.Tarzan:A Peer-to-Peer anonymizing network layer[C]//Proceedings of 9th ACM Conference on Computer and Communication Security(CCS2002),Washington,DC,November 2002:453-465.

[5] Rennhard M,Plattner B.Introducing MorphMix:Peer-to-Peer based anonymous internet usage with collusion detection[C]//Proceedings of the Workshop on Privacy in the Electronic Society,Washington,D.C,USA,Nov.2002.

[6] Zhuang Li,Zhou Feng,Zhao Ben Y,et a1.Cashmere:Resilient anonymous routing[C]//Proceedings of the 2nd Symposium on Networked Systems Design and Implementation,2005.

[7] 眭鸿飞，陈松乔，陈建二.Crowds 系统中基于递减转发概率的路长控制策略[J].小型微型计算机系统，2005，26(3)：387-391.