匿名通信中短距离优先分组重路由方法的研究*

王伟平+, 陈建二, 陈松乔, 王建新

(中南大学 信息科学与工程学院,湖南 长沙 410083)

Research on a Short Distance-Prior Rerouting Scheme in Anonymous Communication

WANG Wei-Ping+, CHEN Jian-Er, CHEN Song-Qiao, WANG Jian-Xin

(College of Information Science and Engineering, Central South University, Changsha 410083, China)

+ Corresponding author: Phn: +86-731-8830212, E-mail: wpwang@mail.csu.edu.cn, http://www.csu.edu.cn

Received 2003-03-10; Accepted 2003-09-26

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.

http://www.jos.org.cn/1000-9825/15/561.htm

Abstract: Rerouting is one of the main schemes used in anonymous communication systems. In some typical anonymous communication systems, data packets pass through a numbers of proxies, and then to the end receiver. So the real initiator could be concealed. However, most rerouting schemes used in some prototype systems are random schemes, which choose the forward proxy randomly among all the proxies in a system. Random rerouting scheme requires every proxy to know some information of all the proxies in the system. With systems expanding, the number of proxies in the system increases, which makes the cost for management higher. Furthermore, its expanding will increase the delay of rerouting as the distance between some of the proxies increases. The paper proposes a new rerouting scheme—short distance-prior rerouting, which implements short distance-prior forwarding, and every proxy chooses a forwarder in its nearby group. The new scheme is applied in the random probability forward rerouting and definite path length rerouting algorithms. Mathematic analysis and simulation results indicate that the new scheme can keep almost the same anonymity as the typical ones and obviously decrease the delay of forwarding. In the new scheme, each proxy only needs to know the information of proxies in its nearby group, which also gives some support to the research of scalability of anonymous systems.

Key words: rerouting; anonymous communication; short distance-prior; delay of service; scalability

摘 要: 重路由技术是匿名通信系统中采用的主要技术手段之一.目前典型的匿名系统中大多采用随机重路由的策略,即在所有中转代理中随机选择一个进行转发的策略,随机转发策略要求每个中转代理知道系统中所

* Supported by the National Natural Science Foundation of China under Grant Nos.90104028, 90304010 (国家自然科学基金)

作者简介: **王伟平**(1969－),女,江苏苏州人,博士生,副教授,主要研究领域为网络信息安全,匿名通信,网络 QoS 路由优化;**陈建二**(1954－),男,长江学者奖励计划特聘教授,博士生导师,主要研究领域为计算机理论,计算复杂性及优化,计算机网络优化算法,计算机图形理论与算法;**陈松乔**(1940－),男,教授,博士生导师,主要研究领域为软件工程;**王建新**(1969－),男,博士,副教授,主要研究领域为计算机网络理论,QoS 路由优化算法.

有其他代理.随着系统的扩大,一方面中转代理数增加使得系统维护代价增加,另一方面由于部分中转代理之间距离很远,重路由带来的延迟增加.提出了一种新的重路由策略——距离优先分组重路由,实现了短距离优先转发的策略,重路由时在近距离分组中进行随机转发.分别将距离优先分组策略应用在随机概率转发和有限路长限制的重路由算法中,数学分析和模拟测试结果表明,新的重路由策略在一定分组成员数情况下能保持与非分组重路由算法相当的匿名性能,同时明显地降低了服务延迟.新的策略中每个中转代理只需知道就近分组中的代理,这为匿名系统的扩展性研究提供了一定的基础.

关键词: 重路由;匿名通信;短距离优先;服务延迟;可扩展性

中图法分类号: TP309 **文献标识码**: A

匿名通信是指通过一定的方法将业务流中的通信关系加以隐藏,使窃听者无从直接获知或推知双方的通信关系或通信的一方.匿名通信的一个重要目的就是隐藏通信双方的身份或通信关系,从而实现网络用户的个人通信隐私及对涉密通信的更好保护.例如,在匿名投票系统、匿名支付系统以及匿名举报系统中的应用.目前,有关网络匿名通信技术的研究已越来越引起网络安全研究人员的重视,成为网络安全研究前沿的一个重要分支.

在基于因特网的匿名通信技术的研究方面,早期提出了针对电子邮件的匿名邮件系统[1],随后又出现了一些有代表性的匿名通信协议和原型系统,典型的有基于简单代理(simple-proxy)的 Anonymizer[2], ProxyMate.com[2],基于 Mix 的 Freedom[3],Onion Routing[4,5],Web Mixes[6,7],基于组群的 Crowds[8],Hordes[9],这些系统都在一定程度上保证了匿名连接.

从技术上而言,目前匿名系统的研究主要是对匿名有效性的研究,这些系统借助于多个代理的重路由(rerouting)技术、填充包(padding)技术和加密技术来达到匿名发送或接收的目的.例如在 Onion Routing[4,5]系统中,采用了路径上主机的多级嵌套公开密钥加密,并且利用路径上代理填充包使得所有的输入输出具有同样的长度,这使得网络附加开销增大,同时也使得客户请求相应的延迟有所增加.

在重路由技术方面,大部分采用的是随机重路由的策略,即在所有中转主机中随机选择一个进行转发.随着系统的扩展,代理主机数增加,由于部分中转主机之间的距离很远,随机重路由带来延迟可能会很大.本文提出了一种距离优先分组重路由策略,数学分析和仿真实验都证明,在一定分组成员数的情况下,该策略能够较好地保持匿名度,同时极大地降低了重路由转发延时.

本文第 1 节介绍相关的研究工作.第 2 节是新的重路由策略及其匿名性能分析.第 3 节是模拟测试及其分析.第 4 节是结论.

1 相关的研究工作及基础

在这一节中,我们介绍一些典型匿名系统中的重路由策略,对相关的匿名特性进行分析比较.

1.1 符号及定义表示

由于本文中采用概率分析的方法,在文章中多处用到了一些符号,表 1 给出了这些符号的说明.

Table 1 Symbol description

表 1 符号说明

Symbol	Description
I	The event that the first collaborator on the path is immediately preceded on the path by the path initiator
$H_k, k\geq 1$	The event that the first collaborator on the path occupies the kth position on the path (Assume the position of the path initiator is 0)
H_{k+}	$H_k \vee H_{k+1} \vee H_{k+2} \vee \ldots$
$P(I\|H_{1+})$	Given that a collaborator is on the path, the probability that the path initiator is the first collaborator's immediate predecessor
$P=(n-c)/n$	The probability of non-collaborator in all members
L	The number of forwarding proxies on the rerouting path
$P(v_i)$	Given that the member on the ith position is non-collaborator, the probability that it is the path initiator

1.2 随机概率转发重路由(RPRR——randomized probability rerouting)

AT&T 的 Crowds 系统[9]是基于组群的思想来实现匿名的,能够提供匿名 Web 浏览服务.系统中每个成员用户有一个代理,称作 jondo.当用户发出请求时,jondo 充当请求代理,将请求以等概率随机发给组中任一 jondo.之后,路径上每个 jondo 以概率决定是转发给下一个 jondo 还是将请求直接给 Web server,即发给组中任一 jondo 的概率是 p_f,发给 Web server 的概率是 $1-p_f$.

Crowds 系统主要实现了发送者匿名.在文献[8]中利用概率分析的方法,对泄密者攻击情况下的匿名性能作了分析,用路径上有泄密者情况下,攻击者能准确判断发起者的概率 $P(I|H_{1+})$来表示匿名性能.

$$P(I \mid H_{1+}) = \frac{n - p_f(n-c-1)}{n} \tag{1}$$

其中,n 是组群中的成员数,即系统中的中转代理总数;c 是泄密者个数;p_f 是转发概率.在 Crowds 中平均每条路径经过转发的代理个数为 $K=1/(1-P_f)$.

在其他匿名系统中大部分也是采用随机转发的策略,例如在 Onion Routing 系统中,借鉴了 Crowds 的方法,采用了随机转发源选路径的方式.

1.3 有限长度重路由

有限长度重路由[10]是对随机概率重路由的改进.Crowds 中采用概率的方法决定是传给下一 jondo 还是传给 server,路径长度是没有制约的.这使得路径长度没有上界,极端情况下会达到无限长,这使得请求应答时间无限长(尽管这一概率非常小).

有限长度匿名通信协议的工作过程:当用户要发送信息时,将请求交给它的用户路由代理(route_proxy),代理以一定方式随机产生一个路径长度(path_length),然后任选一组员代理传送,该转发代理将 path_length 减 1,若 path_length 不为 0,则该代理按同样的方式将请求转交给其他代理,直到某个代理发现 path_length 降为 0,就将请求直接交给接收方(Web server).

用同样的概率分析方法[10],得到了当路径长度为 k 时的固定有限长度重路由(static path-length rerouting,简称 SKRR)的匿名性能.

$$P(I \mid H_{1+}) = \frac{1 - p + \left(1 - p^{k-1}\right)\frac{1}{n}}{1 - p^k} \tag{2}$$

由于采用固定长度的路径会带来另一种攻击,即由于 k 是路径的上界且所有中转主机都知道 k,所以只要路径上的某个泄密者获得的 path_length 是 $k-1$,就可以准确地得出前者是发送者.

改进的做法是采用随机产生路径长度的方法——随机路长重路由(randomized path-length rerouting,简称 RKRR),但若采用均匀分布,在$[k_1,k_2]$间随机取值作为路径长度 k,则仍然有 $1/(k_2-k_1+1)$的概率取到 $k=k_2$,这种情况下,同样会导致固定长度情况下的问题.因此,我们设法采用其他分布随机取值,设法使取到 k_2 的概率降低,因此构造了一个取值权值函数 $W(k)$.

$$W(k) = \begin{cases} \dfrac{1}{(mid - k + 1)^2}, & k \le mid \\ \dfrac{1}{(k + 1 - mid)^2}, & k > mid \end{cases},$$

其中,$mid \in [k_1,k_2]$,代表权值最大的 k.

因此,取到路径长度 k 的概率是

$$\rho(k) = \frac{W(k)}{\sum_{k=k_1}^{k_2} W(k)} \tag{3}$$

在采用随机路长方法之后,使得路径长度取到 k_2 的概率降低.例如,当 k 在[3,20]间取值,取 mid=5,6,7,8 时,对应的 k 取到上界 20 的概率分别是 0.0020,0.0022,0.0025,0.0029.

2 短距离优先分组重路由策略

重路由将原来的单个源与目的之间的直接连接变成了多个主机之间的转接,从而使得通信双方的通信关系得以隐藏或者发送方、接收方的身份得以隐藏.重路由付出的代价是显而易见的,体现为网络的代价和服务的代价两方面,网络方面的开销体现在网络资源的占用上,多个代理对信包作了转发,占用了更多的信道资源和主机资源.服务方面的代价体现为增加了服务的延迟时间,原来单个连接的延迟变成了多个连接的延迟之和.在以往的研究中,集中于获得较好的匿名性能而忽略了协议的开销,尤其是对延迟时间的增加.本文提出的短距离优先重路由策略在考虑匿名性能的同时兼顾了性能方面的影响.

2.1 短距离优先重路由策略的基本思想

相关的重路由策略研究都是将代理主机作为一个大组,当选择转发代理时采用随机选择转发的策略.而我们可以观察到,当这些代理主机分布很广且数目很多的情况下,代理主机间的连接延迟差别就会很大,这主要是由于代理之间的物理距离以及代理连接上的拥挤程度不同造成的,因而造成了选择不同转发代理所付出的代价有较大的差异.

因此,我们的基本出发点是优先选择代价小的代理主机进行转发.我们主要基于用户延迟方面的考虑,提出了短距离优先的重路由策略.基本思想是,在 n 个代理主机的环境下,每个主机代理通过测试获得自己周围的近距离的 m 个主机代理,在转发时采取随机策略在 m 个邻近主机代理中间进行选择.

从实现的角度来看,这里的距离可以用时延来表示,由于主机之间的往返时延可以通过底层的 ICMP 报文测知,所以距离的测试是可行的.距离可以采取定期测试的方法,因此每个代理主机(router_proxy)都定期保留它的 m 个邻近主机(nearby route_proxy)的信息.

以下是分组之后的随机概率转发重路由算法的描述,设随机转发概率是 p_f:

```
Receive client.request
If (client==local_procedure)
   Randomly choose the next route_proxy from its m nearby route_proxys
   Forward_request (next route_proxy)
Else
   Prior route_proxy=client.id       /*记下来的路径,用作返回*/
    Randomly generate value p between 0 and 1
    If (p>pf)
        Submit_request (receiver)    /*将请求传给接收者*/
    Else
      Randomly choose the next route_proxy from its m nearby route_proxys
      Forward_request (next route_proxy)

    Procedure Forward_request (route_proxy id)
    Send packet to route_proxy[id]
    Reply=Await_reply()
    If (reply='next route_proxy failed')
      Randomly choose the next route_proxy from other nearby route_proxy
      Forward_request (next route_proxy)
    Else
       Send reply to prior route_proxy

    Procedure Submit_request (receiver)
     Send request to receiver
     Reply=Await_reply (timeout)
     Send reply to prior route_proxy
```

2.2 距离优先分组重路由策略的匿名性分析

我们将分组策略引入到第 1.2 节和第 1.3 节所描述的重路由算法中.在匿名性能的分析上,为了便于比较,我们采用与 Crowds 系统相同的攻击模型,即中转主机代理可能成为泄密者,位于路径上的第 1 个泄密者猜测前

者是发送者.用 $P(I|H_{1+})$表示泄密者猜测准确的概率,也就是在被猜测的信包中发送者被猜中的概率.并假设泄密点在物理位置上是均匀分布的(即若非泄密者比例是 p,每个中转主机周围选 m 台主机,其中 $m\times p$ 是非泄密主机).以下分别给出了概率转发、固定路长和随机路长情况下的距离优先分组重路由策略的匿名性能分析结果(具体数学推导过程详见附录).

2.2.1 随机概率转发分组重路由(DGRPRR——randomized probability rerouting based on distributed group)策略

将分组策略引入随机概率转发重路由算法中,每次转发在就近组主机中随机选择,其他维持原策略不变,经推导可得

$$P(I\mid H_{1+})=\frac{p_f+n-np_fp}{n\left[1-p_f\left(\frac{1}{m}-\frac{1}{n}\right)\right]}.$$

当 $m=n$ 时,即不分组时,上式即为$(p_f+n-npp_f)/n$,以 $p=(n-c)/n$ 代入,即为$[n-p_f(n-c-1)]/n$,与公式(1)是一致的.因为 $m<n$,所以显然分组后 $P(I|H_{1+})$比未分组时有所增加,即匿名性能有所下降.

2.2.2 固定路长分组重路由(DGSKRR——static path-length rerouting based on distributed group)策略

同样,将分组策略引入到有限固定长度的重路由策略中.当路径长度固定为 k 时,推导可得

$$P(I\mid H_{1+})=\frac{\frac{1}{n}}{p+\frac{1}{n}-\frac{1}{m}}+\frac{p-\frac{1}{m}}{\left(p+\frac{1}{n}-\frac{1}{m}\right)\sum_{i=0}^{k-1}p^i}\left(1+\sum_{i=1}^{k-1}\left(\frac{1}{m}-\frac{1}{n}\right)^i\right) \tag{4}$$

上式中 $m>=2$,因为当 m 为 1 时,最靠近自己的点是自己,每次都会选择自己,即发送者每次都传给自己.

当 $m=n$ 时,公式(4)可以简化为

$$P(I\mid H_{1+})=\frac{1-p+\frac{1}{n}(1-p^{k-1})}{1-p^k}.$$

这与公式(2)一致,即这时匿名性能等同于不分组时固定路长的情况.

2.2.3 随机路长分组重路由(DGRKRR——randomized path-length rerouting based on distributed group)策略

由于采用固定路长会带来与有限固定路长重路由(SKRR)同样的问题,因此考虑采用与公式(3)相同的概率函数$\rho(k)$来随机取得路长 k,$P(I|H_{1+},k)$表示在路长为 k 的情况下,固定路长分组重路由策略下的 $P(I|H_{1+})$值.这种情况下,可以按照下式求得 $P(I|H_{1+})$的期望值.

$$\overline{P(I\mid H_{1+})}=\sum_{k=k_1}^{k_2}\rho(k)\cdot P(I\mid H_{1+},k)=\sum_{k=k_1}^{k_2}\rho(k)\cdot P(I\mid H_{1+}).$$

2.3 匿名性能分析与比较

根据以上概率分析结果,我们对多种重路由策略的在不同路径长度期望、不同泄密者数以及不同分组大小下的匿名性能进行了比较.

2.3.1 路径长度对匿名性的影响

图 1 计算了不同路径长度期望情况下,取 n=1000,m=100,c=100,各种重路由策略下的匿名性能计算值比较.因为 $P(I|H_{1+})$表示被猜中的概率,所以用 $1-P(I|H_{1+})$可以作为衡量匿名度的指标.

可以看出,当路径长度增加时,各种重路由策略的匿名性能有所提高,而这种增长趋势随着路径长度的增加而放缓.显然,引入了短距离分组优先策略后,各种策略与不分组时相比,匿名性能略有下降,这可以解释为由于采用短距离分组重路由,每个转发代理出现在自己的近距离小组中,使得发送者出现在路径上的概率提高,从而增加了泄密者猜测准确的可能性,降低了匿名性.

无论是否分组,几种路由策略相比,有限路长重路由策略获得的匿名性要好于随机概率转发重路由策略,在有限路长中采用固定路长较随机路长方法获得的匿名性能理论值要好一些.这可以解释为当应用随机概率转发算法(RPRR)时,路径长度取到较小值的可能性较大.

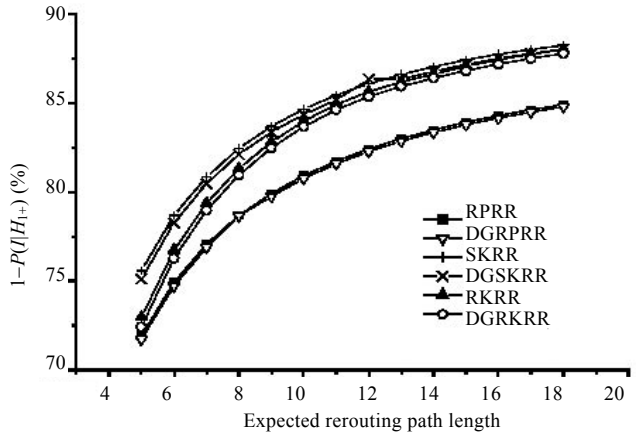


Fig.1 Anonymity of different rerouting scheme VS expected path length

图 1 不同路径长度期望下的多种重路由方法匿名性能比较

2.3.2 路径长度对匿名性的影响

图2计算了在不同泄密者个数情况下,多种重路由策略的匿名性能.可以看出,泄密者个数增加,各种重路由算法下匿名性能都下降.

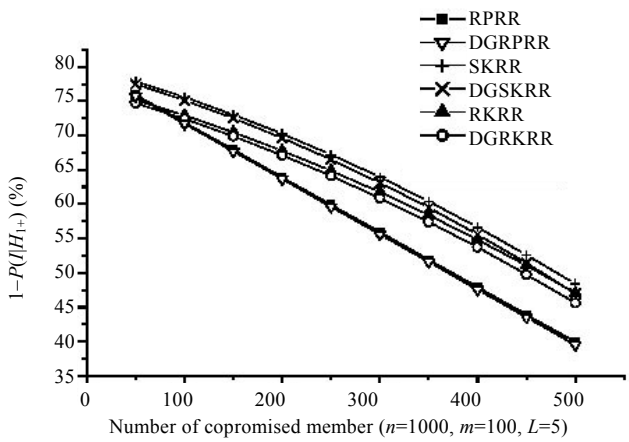
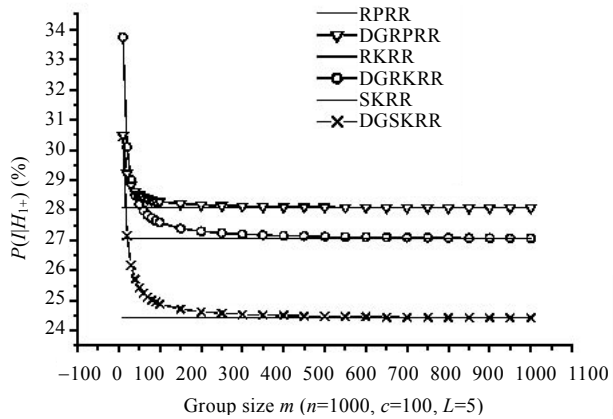


Fig.2 Anonymity performance VS numbers of compromised members

图 2 不同泄密者数量情况下的匿名性能

2.3.3 分组大小对匿名性的影响

图 3 是当 n=1000 和 n=10000,泄密者比例是 10%,L=5(或路长期望是 5)的情况下,分组大小变化时,各种匿名重路由方法下的 $P(I|H_{1+})$.可以看出,分组重路由的 $P(I|H_{1+})$与分组大小有关系,这是因为分组成员越少,发送者再次出现在路径上的概率越大,从而路径上一旦有发送者,猜中的可能性就越大,即 $P(I|H_{1+})$增加.而我们也可以看到这种变化不是线性的,当分组成员数小于 150 时,随着分组成员数的增加,$P(I|H_{1+})$下降较快(匿名性变好);当分组成员数大于 150 时,$P(I|H_{1+})$变化趋缓,且不断接近相同路长期望下的相应非分组重路由算法.同时可以看到,n 增大,即系统主机数增多,对曲线的影响不大.选择分组大于 150,可以获得接近于非分组路由情况的匿名性.



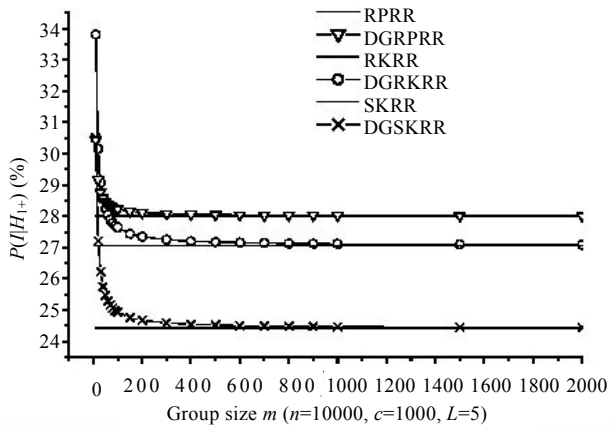


Fig.3 $P(I|H_{1+})$ with same path length and different group size

图 3 同样路长不同分组大小下的 $P(I|H_{1+})$值

3 模拟测试

在这一节中,我们将通过模拟的方法测试上述多种重路由策略的匿名性能、造成的附加开销以及对服务延迟的影响.

3.1 模拟测试环境

模拟环境中共有 n 个中转代理,代理间相互是可以连接的,代理间的距离在$[L_1,L_2]$之间均匀分布,每个代理周围最近的 m 个代理形成该代理的转发代理组,c 个泄密者在 n 个成员中随机产生.随机产生测试请求,请求的发送者在 n 个成员中随机产生.

3.2 测试结果及其分析

共测试了 10 000 个请求,取代理主机总数 n=1000,分组成员数 m=100,泄密者数 c=100,距离分布范围值 L_1=2,L_2=20,L_1,L_2 的值以延时单位计.表 2 列出了多种重路由算法的性能测试值,并且列出了相应的理论计算值作为对照(由于固定路长会带来不利影响,这里只测试了随机有限路长的情况).其中,$P(I|H_{1+})$是指攻击准确的概率;平均路段数是指每个请求平均经过的路段数,这反映了重路由中多少个代理主机参与了转发;平均转发延时单位测试了每个请求平均经过的延时单位,即路径上的延迟之和,这反映了重路由对服务延时的影响.

为了便于比较,我们在相同路径长度期望下进行测试,在对 RKRR 和 DGRKRR 算法的测试中,取路径长度范围值区间 k_1=3,k_2=20,通过变化 mid 值使路径长度期望变化;在 RPRR 和 DGRPRR 算法测试中,通过改变 p_f 值,

使路径长度期望变化.

Table 2 Simulated results for performance of different anonymous rerouting algorithms

表 2 多种匿名重路由算法模拟测试性能比较

Expected path length (Theory result)		$P(I\|H_{1+})$ (%) (Theory result)	$P(I\|H_{1+})$ (%) (Simulated result)	Expected path length (Simulated result)	Average forwarding delay
L=6.268 5 (mid=5.5, p_f=0.840 47)	RKRR	22.580 7	20.83	6.255 3	68.634 9
	DGRKRR	23.016	22.05	6.273 6	17.905 2
	RPRR	24.441 7	24.030 6	6.658 0	73.152 6
	DGRPRR	24.627	24.134 6	6.603 9	19.229 1
L=7.101 41 (mid=6.5, p_f=0.859 18)	RKRR	20.436	19.95	7.112 4	78.311 6
	DGRKRR	20.829 91	19.70	7.099 8	20.249 1
	RPRR	22.759 7	22.150 1	7.043 4	77.395 0
	DGRPRR	22.937 1	23.186 7	7.281 6	20.773 6
L=7.101 41 (mid=7.5, p_f=0.874 37)	RKRR	18.766 1	17.94	7.968 1	87.580 0
	DGRKRR	19.127 8	18.32	7.958 2	22.737 9
	RPRR	21.394 1	20.35	8.445 4	92.821 0
	DGRPRR	21.563 8	20.584 5	8.411 8	24.027 2

从表 2 中可以看出,$P(I|H_{1+})$模拟测试结果值与概率分析理论计算结果基本是一致的,在相同路段期望情况下,采用了距离优先分组策略后匿名性能略有降低,同时,从对服务延迟的影响来看,距离优先分组重路由策略使请求转发延迟时间有明显的降低.例如,在路径期望 L 为 6.268 5 时,RKRR 与 DGRKRR 比较,$P(I|H_{1+})$测试值从 20.83 增加到 22.05,而转发延迟单位由 68.634 9 降为 17.905 2;RPRR 与 DGRPRR 比较,$P(I|H_{1+})$测试值从 24.030 6 增加到 24.134 6,而转发延迟单位由 73.152 6 降为 19.229 1.

4 结 论

本文提出了一种用于匿名通信的重路由算法——距离优先分组重路由算法.我们利用与 Crowds 系统分析中同样的攻击模型,分别在随机概率转发和有限路长限制下研究了该算法的匿名性能、网络开销以及转发造成的延迟.数学分析和模拟测试结果表明,与非分组重路由算法相比,在一定分组大小情况下,距离优先重路由算法在保持了较好的匿名性能的同时降低了服务延迟.

进一步的研究将针对泄密者集中分布情况下的匿名性能进行研究,并且将这种新的重路由策略用到匿名通信系统中,研究在实际系统中的可实现性以及可扩展性.

References:

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

[2] Joris C, Bart P, Joos V. Solution for anonymous communication on the Internet. In: Sanson LD, ed. Proc. of the IEEE 33rd Annual 1999 Int'l Carnahan Conf. on Security Technolgy. Madrid, 1999. 298~303.

[3] Roger D, Michael JF, David M. The free haven project: Distributed anonymous storage service. In: Federrath H, ed. Proc. of the Workshop on Design Issues in Anonymity and Unobservability 2000. LNCS 2009, Heidelberg: Springer-Verlag, 2000. 67~95.

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

[5] Reed M, Syverson P, Goldschlag D. Anonymous connection and onion routing. IEEE Journal on Selected Areas in Communications, 1998,16(4):482~492.

[6] Berthold O, Federrath H, Köpsell S. Web MIXes: A system for anonymous and unobservable Internet access. In: Federrath H, ed. Proc. of the Workshop on Design Issues in Anonymity and Unobservability 2000. LNCS 2009, Heidelberg: Springer-Verlag, 2000. 115~129.

[7] Berthold O, Federrath H, Köhntopp M. Project Anonymity and unobservability in the Internet. In: Cranor L, ed. Proc. of the Computers Freedom and Privacy Conf. 2000 (CFP 2000) Workshop on Freedom and Privacy by Design. Toronto: ACM Press, 2000. 4~7.

[8] Reiter MK, Rubin AD. Crowds: Anonymity for web transactions. ACM Trans. on Information and System Security, 1998,1(1): 62~92.

[9] Shields C, Levine BN. A protocol for anonymous communication over the Internet. In: Jajodia S, ed. Proc of the 7th ACM Conf. on Computer and Communication Security. Athens: ACM Press, 2000. 33~42.

[10] Wang WP, Chen JE, Wang JX, Sui HF. An anonymous communication protocol based on groups with definite route length. Journal of Computer Research and Development, 2003,40(4):607~614 (in Chinese with English abstract).

附中文参考文献:

[10] 王伟平,陈建二,王建新,眭鸿飞.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.

附录:关于短距离优先分组重路由算法匿名性能的分析与推导.

假设泄密点在物理位置上是均匀分布的,非泄密者比例是 p(即每个中转主机周围选 m 台主机,其中 $m\times p$ 是非泄密主机).

令位于路径的第 1 个泄密者在路径的第 i 个位置的概率,即路径的前 $i-1$ 次取到了非泄密者,第 i 个选到了泄密者的概率为 $P(H_i)$.第 1 个泄密者位于路径上第 1 个位置或之后的概率为 $P(H_{1+})$.

下面我们来推导随机概率转发分组重路由(DGRPRR)下的 $P(H_i)$,$P(H_{1+})$和 $P(I|H_{1+})$值.

路径上第 1 个转发主机为泄密者的概率是 $P(H_1)=1-p$,路径上第 1 个泄密者位于路径的第 i 个位置上的概率是

$$P(H_i)=(p_f p)^{i-1}(1-p) \tag{A-1}$$

所以,路径上有泄密者的概率是

$$P(H_{1+})=\sum_{i=1}^{\infty}P(H_i)=\sum_{i=1}^{\infty}(p_f p)^{i-1}(1-p)=\frac{1-p}{1-p_f p} \tag{A-2}$$

令 $P(v_i)$为路径上第 i 个位置不是泄密节点的情况下,是发起者的概率.

当第 1 个泄密者位于路径上第 1 个位置时,它的前一个肯定是发起者,事件 I 是成立的,即 $H_1\rightarrow I$.所以条件概率 $P(I|H_1)=P(v_0)=1$.

但当路径上第 1 个泄密者位于第 2 个位置时,这时猜测前者为发起者的准确概率是 $P(I|H_2)=P(v_1)=1/mp$,这个概率即是路径上第 1 个位置在排除泄密者的情况下取到发起者的概率.这是由于路径上第 1 个泄密者在第 2 个位置就可以肯定路径上的第 1 个转发者不是泄密者,即发起者(第 0 位置)选择了就近组中的非泄密者 mp 中的一个,由于发起者本身是其中的一个,所以猜中的概率是 $1/mp$.

当第 1 个泄密者位于路径上的第 3 个位置时,有两种情况,第 1 种情况是,当第 1 个位置是发起者时,则发起者会落在第 1 个主机的就近组中,即是就近组 m 个成员中的一个,这时发起者位于第 2 个位置的概率是 $1/mp$;第 2 种情况是,当第 1 个位置不是发起者时,任何非泄密主机出现在它的就近组内的概率相等,即发起者位于第 2 个位置的概率是 $1/np$.即

$$P(I\mid H_3)=P(v_2)=P(v_1)\frac{1}{mp}+(1-P(v_1))\frac{1}{np}=\frac{1}{(mp)^2}+\left(1-\frac{1}{mp}\right)\frac{1}{np}.$$

以此类推,我们可以得到

$$\begin{aligned}P(I\mid H_k)&=P(V_{k-1})\\&=P(v_{k-2})\frac{1}{mp}+(1-P(v_{k-2}))\frac{1}{np}=\frac{1}{np}+\left(\frac{1}{mp}-\frac{1}{np}\right)P(v_{k-2})\\&=\frac{1}{np}+\left(\frac{1}{mp}-\frac{1}{np}\right)\left[\frac{1}{np}+\left(\frac{1}{mp}-\frac{1}{np}\right)P(v_{k-3})\right]=\frac{1}{np}\sum_{i=0}^{k-2}\left(\frac{1}{mp}-\frac{1}{np}\right)^i+\left(\frac{1}{mp}-\frac{1}{np}\right)^{k-1}\\&=\frac{1}{np}\frac{1-\left(\frac{1}{mp}-\frac{1}{np}\right)^{k-1}}{1-\left(\frac{1}{mp}-\frac{1}{np}\right)}+\left(\frac{1}{mp}-\frac{1}{np}\right)^{k-1}.\end{aligned}$$

$$P(I)=\sum_{i=1}^{\infty}P(H_i)P(I\mid H_i).$$

$$P(I\mid H_{1+})=\frac{P(I\wedge H_{1+})}{P(H_{1+})}=\frac{P(I)}{P(H_{1+})}=\frac{\sum_{i=1}^{\infty}P(H_i)P(I\mid H_i)}{P(H_{1+})} \tag{A-3}$$

以公式(A-1),(A-2)代入公式(A-3)可以得到

$$\begin{aligned}P(I\mid H_{1+})&=\frac{P(H_1)P(I\mid H_1)+\sum_{i=2}^{\infty}P(H_i)P(I\mid H_i)}{P(H_{1+})}\\&=\frac{(1-p)+\sum_{i=2}^{\infty}(p_f p)^{i-1}(1-p)P(I\mid H_i)}{(1-p)/(1-p_f p)}\\&=(1-p_f p)\left\{1+\sum_{i=2}^{\infty}(pp_f)^{i-1}\left[\frac{1}{np}\frac{1-\left(\frac{1}{mp}-\frac{1}{np}\right)^{i-1}}{1-\left(\frac{1}{mp}-\frac{1}{np}\right)}+\left(\frac{1}{mp}-\frac{1}{np}\right)^{i-1}\right]\right\}\\&=\frac{p_f+n-np_f p}{n\left[1-p_f\left(\frac{1}{m}-\frac{1}{n}\right)\right]}.\end{aligned}$$

用同样的方法可以推出有限路长分组重路由(DGSKRR)策略下的 $P(I|H_{1+})$值(略).

第 14 届中国计算机学会网络与数据通信学术会议

征 文 通 知

中国计算机学会网络与数据通信专委会和开放系统专委会定于 2004 年 10 月在西安召开第 14 届中国计算机学会网络与数据通信学术会议,本届会议的主题是研究信息网络关键技术,营造安全可靠网络环境.

一、征文范围

开放系统及其互联技术;新一代网络结构与协议;网络智能化、网络管理;网络信息系统模型;网络计算与应用;网络环境下的信息安全;无线通信网络;电子商务系统以及光纤通信技术.

二、要求

1、欢迎围绕以上主题提交研究论文,字数一般不要超过 6000 字.录用的论文将在西北大学学报(增刊)(国内核心期刊)上发表,评选的优秀论文将推荐在国家权威期刊上发表.

2、稿件格式要求请查询网址:http://www.nwu.edu.cn/xsnh/index.htm

3、应征文章请寄两份打印稿,同时须用电子邮件的附件发来(或用软盘寄来)电子稿,电子稿件请用 WORD 文件格式(.doc 文件).

4、请随同稿件一起,用另纸写明文章题目,所属主题,作者姓名(最多 4 人),职务/职称,所属单位,详细通信地址,邮编,电话,E-mail 地址.

5、已经发表的论文请勿报送.如因一稿多投带来任何问题,责任由投稿者自负.

6、论文收寄地址:710069 西安市太白北路 229 号西北大学现代教育技术中心学术年会筹备组 宋峰 收

发送电子稿的邮件地址:cetc@nwu.edu.cn.

三、论文截止日期

2004 年 6 月 30 日(以寄出邮戳日期为准),会议组委会将于 2004 年 7 月 31 日前发出论文录用通知和参加会议邀请信.

四、联系方式

通信地址:710069 西安市太白北路 229 号 西北大学现代教育技术中心学术年会筹备组

电话:029-88302758 传真:029-88303857 电子邮件: cetc@nwu.edu.cn

联系人:宋峰 刘瑞献