

## 重路由匿名通信跟踪技术研究

程 昊

(北京科技大学 计算机与通信工程学院, 北京 100083)

**摘 要:** 在匿名代理由于客观原因无法定位和控制的前提下,对匿名通信中重路由技术进行了分析和分类的同时,还介绍了一些针对发送者匿名、接收者匿名和通信关联关系匿名的跟踪方法和相关技术.

**关键词:** 匿名通信;重路由技术;跟踪技术

**中图分类号:** TP309

**文献标识码:** A

**文章编号:** 1000-7180(2013)06-0156-03

## Research of Anonymous Communication Tracking Based on Rerouting Technique

CHENG Hao

(College of Computer and Communication, University of Science and Technology Beijing, Beijing 100083, China)

**Abstract:** On the assumption that anonymous agents could be neither positioned nor controlled, we analyze and classify the rerouting technique used in anonymous communication. In this paper, we propose an anonymous communication tracking technique in allusion to anonymity of sender, receiver and communication relationship in order to reveal the relationship of sender and receiver in anonymous communication.

**Key words:** anonymous communication; rerouting technique; tracking technique

### 1 引言

从技术角度而言,目前一般匿名通信系统大都采用多级代理的重路由(rerouting)技术、填充包(padding)技术和加密技术,以达到匿名发送/接收(或通信关联关系匿名)的目的.现有的匿名通信系统如:DC-Net, Mixes, Anonymous, Remailer, LP-WA, Onion Routing I, Onion Routing II, Crowds 均采用重路由机制以提供匿名保护<sup>[1]</sup>.

重路由技术是由多个代理主机在应用层进行数据存储转发,形成一条由多代理主机组成的虚拟安全信道.在该安全信道上传送的 IP 数据包,跟踪者无法获得真实的发送者或接收者的 IP 地址等身份信息.因而,通信实体的身份信息被有效地隐藏.例如:Mixes 隐藏邮件发送者的身份信息;Onion Routing 隐藏实时通信中通信实体之间的通信关联关系;Crowds 则保护 Web 浏览器用户的身份信息

息<sup>[2-3]</sup>.在重路由技术中,一般采用多级匿名代理来解决匿名通信,采用传输路径上代理主机的多级加密,来解决信息传输的机密性;利用传输路径上代理主机填充包技术来防止跟踪者对报长的统计分析;利用报文的随机延迟来防止一定程度上的流量分析.

匿名通信行为根据加密和非加密传输的不同,按发送者、匿名、接收者匿名、通信关联匿名的特点,在不同监测区域的汇聚节点部署检测器,对进出该区域的所有报文进行检测,可获得发送者、接收者和通信关联关系跟踪.因为不同的匿名通信会存在不同的会话过程,不同会话过程一定利用特定协议发送特有的报文.为此,匿名通信行为确定的关键是对不同匿名通信初始会话过程进行分析和特征提取,如图 1 所示.

#### (1) 非加密传输方式

在进行匿名通信前,发送者和接收者之间要进

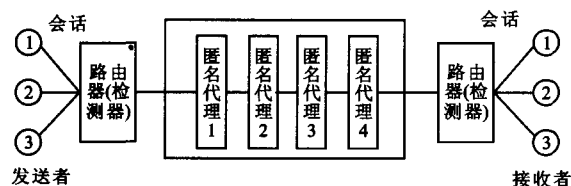


图1 匿名通信示意图

行初始会话过程,该过程一般采用特殊协议,通过在本地链路部署检测器,如在本地会聚节点(路由器)通过流量镜像技术,对进/出的数据流量进行过滤,一旦发现有特殊协议格式的报文进/出,表明有人在进行匿名通信,当然,这只能是一种怀疑。

### (2)加密传输方式

如果代理为VPN安全网关,并且各个VPN网关有可能采用IPSec的隧道应用模式。在这种情况下,IPSec协议密钥材料和加密算法的协商可以采用两种方式:自动协商和手动配置。如果采用自动协商,则初始会话过程采用一般采用IKE,而IKE协商过程一般要使用到ISAKMP、Oakley、SKEME协议,在本地汇聚节点(路由器)通过流量镜像技术,对进/出的数据流量进行过滤,一旦发现有这三种协议格式的报文进/出,则怀疑有人在进行匿名通信。如果密钥材料和加密算法的协商采用手动配置,则在数据传输阶段,通过流量过滤技术只要检测到有ESP(AH)协议格式报文进/出,同样也可以怀疑有人在进行匿名通信。

根据重路由路径长度是否确定,可以将重路由技术分为随机概率转发重路由技术和有限长度重路由技术。下面分情况对不同的重路由策略的跟踪问题进行讨论。

## 2 随机概率转发重路由(RPRR randomized probability rerouting)

Crowds系统是基于随机概率转发策略来提供匿名Web客户端的浏览服务。系统中每个成员用户有一个代理,当用户发出请求到代理时,代理将请求以等概率( $P_f$ )随机发给匿名通信网络中任一代理,同时将请求以概率 $(1-P_f)$ 直接给Web server,依次类推,直到Web server接收到该请求<sup>[4]</sup>。

如图2所示,发送源以等概率事件与1,2,3,4,5,6,7,8,9中的任一个代理进行通信,然后代理1以一定的概率( $P_f$ )发送数据给2,3,4,5,6,7,8,9和接收者 $(1-P_f)$ ;依次类推,直到接收者接收到该消息。

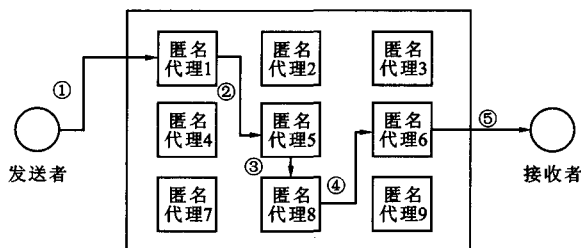


图2 随机概率转发重路由示意图

### 2.1 非加密传输模式下跟踪技术

#### (1)发送者匿名跟踪

如图1所示,在本地链路①检测到发送源有匿名会话过程,或在数据传输阶段,有特殊格式的IP报文,则直接解析IP头中源IP即可获取发送者的IP地址信息。

#### (2)接收者匿名跟踪

如图1所示,在本地链路⑤检测到接收者有匿名会话过程,或在数据传输阶段,有特殊格式的IP报文,则直接解析IP头中目的IP即可获取接收者的IP地址信息。

#### (3)通信关联匿名(发送者和接收者同时匿名)跟踪

数据传输阶段,在本地链路①以概率 $P_f$ 将检测到的匿名通信报文x送到分析中心,在本地链路⑤以概率 $P_f$ 将检测到的报文y送到分析中心,如果报文x和报文y的内容相同,利用报文x可跟踪到源IP地址,利用报文y可跟踪到目的IP地址,同时可建立发送者和接收者之间的对应关系。也可以采用一种简化的方法,首先分析在该匿名通信过程中,IP报文头中哪些字段在整个通信过程中网络互连设备不对其进行操作,如标记字段(紧急标记,PUSH标记)等。在本地链路①对检测到的匿名数据报文头中标记字段打一个特殊标记,相当于一个水印,只要比较报文x和报文y是否有相同的标记,然后根据报文y中的目的IP地址和报文x中源IP地址就可以建立匿名通信关联关系。

当然,如果条件允许,还可以在链路②③④部署检测器,对匿名通信的整个路径进行跟踪,目的是在Internet网络上伪装一些代理服务,以达到对整个匿名通信网络进行控制的目的。

### 2.2 加密传输模式下跟踪技术

在这种应用模式下,代理主机由VPN安全网关代替;VPN网关处理除具有匿名通信的功能外,同时实现了对数据的加密,这使得匿名跟踪相比较较难。如图3所示,发送者与代理,代理与代理,代理

与接收者之间均采用隧道应用模式,数据通信真正的源 IP 地址和目的 IP 地址被加密封装在内部 IP 头中,而外部 IP 头的 IP 地址仅仅为隧道的起始节点 IP 地址和终止节点 IP 地址。

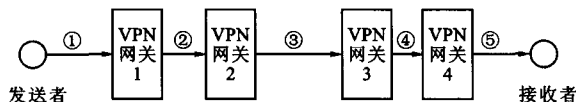


图3 VPN隧道通信应用模式示意图

#### (1) 发送者匿名跟踪

如图3所示,发送者与VPN网关1在本地链路①建立隧道时,如果隧道采用IPSec协议,则在本地链路①可以部署检测点,如果能在初始会话过程中检测到ISAKMP、Oakley、SKEME协议格式报文,或在数据传输阶段检测到ESP(AH)隧道应用模式的数据报文,说明发送者可能要进行匿名通信。即使发送者和VPN1之间采用隧道应用模式,因为该隧道的起始节点为源发送者,所以直接解析会话报文IP头中源IP地址,或ESP(AH)隧道应用模式的数据报文外部IP头部的源IP地址,即可获得发送者IP地址。

#### (2) 接收者匿名跟踪

如图3所示,在本地链路⑤部署检测器,只要检测到接收者有匿名会话过程(如IKE过程),直接解析会话报文IP头中目的IP地址,则有可能该IP地址即为接收者的IP地址。在数据传输阶段,同样只要在本地链路⑤检测到接收者接收有ESP(AH)隧道应用模式的数据,说明接收者可能进行匿名通信,即使接收者和VPN4之间采用隧道应用模式,因为该隧道的终止节点为匿名通信的接收者节点,所以直接解析ESP(AH)隧道应用模式的数据报文外部IP头部的目的IP地址,即可获得接收者IP。

#### (3) 通信关联关系匿名跟踪

在本地链路①部署检测器,如果检测到ESP(AH)隧道模式的数据,根据所检测到报文的源IP地址得到发送者IP地址。一般情况下,在本地链路①可以检测到多个发送者;同样,在本地链路⑤可以检测到多个接收者;问题是如何在多个发送者和多个接收者之间建立对应关系。可以采用试探的方法:在本地链路①将某个发送者匿名通信数据在不影响正常通信效果的前提下,暂时以一定的概率丢弃或延迟,如果本地链路⑤所部署的检测器在规定的延迟时间内没有接收到该数据,则可以在一定程度确定发送者和接收者之间的匿名通信关系,如图3所示。

一般情况下,用户(发送者或接收者)通过在IE

上设置代理,目的只是为了改变路由,如图4所示。发送者通过在IE上设置代理,表明发送者首先和代理1进行通信,代理1然后以一定的概率选择匿名通信网络中的匿名代理1,依次类推。在这种情况下,发送者发送给代理1数据报文IP头的源IP地址和目的IP地址实际上就是发送者IP地址和接收者IP地址;这样的报文到达匿名代理1后才完成匿名和加密。由此可见,只要在本链路①部署检测器一次就可以确定匿名通信的关联关系。

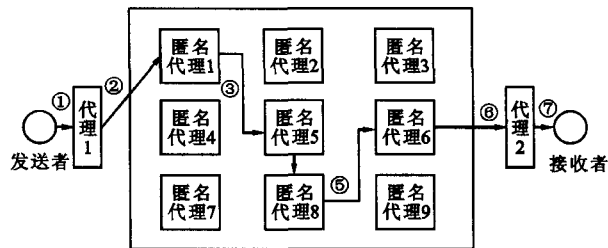


图4 通过IE配置代理示意图

### 3 有限长度重路由

有限长度重路由匿名通信协议的工作过程:当用户要发送数据时,首先将数据以概率发送给代理主机,代理主机以一定方式随机产生一个路径长度,然后以一定的概率选择代理主机,同时将路径长度减1,若路径长度不为0,则代理主机按同样的方式将数据以概率转交给其他代理主机,直到某个代理主机发现路径长度为0,就将数据直接交给接收者。

采用有限长度的重路由技术会带来一种新的跟踪,如果 $k$ (最大为匿名网络中代理主机的个数减1)是匿名通信路径的上界且所有中转代理主机都知道 $k$ ,所以只要路径上的某个跟踪者获得的路径长度是 $k-1$ ,就可以准确地得出前一个节点是发送者;同时某个跟踪者获得的 $path\_length$ 是0,就直接可以获得下一个节点即为接收者<sup>[1]</sup>。当然也可以采用2.1节中的方法来解决源匿名,接收者匿名,通信关联关系匿名跟踪问题。

### 4 分析与验证

为了分析验证本文提出的检测方法有效性,建立了一个由25个主机代理构成的匿名代理网络,针对匿名通信中信息的加密与否,采用分类处理,分发送者匿名,接收者匿名及发送\接收者关系匿名跟踪三种情况,逐步求精的研究思路。通过建立白名单、黑名单和灰名单,分别对非加密通信和加密通信进

(下转第164页)

多目标决策灰色关联度方法[J]. 矿产综合利用, 1996(2):39-43

- [4] 李松仁,李学全,韩旭里. 选矿方案的多目标灰色局势决策[J]. 矿产综合利用,1997(1):37-40
- [5] 陈林,曹树刚. 矿产资源综合利用评定的博弈及最优化分析[J]. 矿业快报,2005(11):112-115.
- [6] Demaison G, Huizinga B J. Genetic classification of petroleum system [M]. AAPG, 1991.
- [7] Haijun Xiao, Fan Hong, Zhaoli Zhang, et al. Intrusion detection using ensemble of SVM classifiers [C]// Fourth International Conference on Fuzzy Systems and Knowledge Discovery, Haikou, Hainan, China:IEEE, 2007:45-49.
- [8] Kittler J, Hatefm, Duinr, et al. On Combining Classifiers [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(3):226-239
- [9] 杰弗瑞,杰里. 高级微观经济理论[M]. 王根蓓,译. 2版. 上海:上海财经大学出版社,2002.
- [10] 黄梦醒,潘泉,施建宇,等. 基于多分类器融合的供应链伙伴动态选择方法[J]. 计算机集成制造系统,2007(6):1131-1136.
- [11]高作峰,王友,王国成. 对策理论与经济管理决策[M]. 北京:中国林业出版社,2006.

#### 作者简介:



王玲女,(1959—),副教授. 研究方向为统计学.



余志林 男,(1981—),硕士. 研究方向为数据挖掘.



肖海军 男,(1965—),博士. 研究方向为数据挖掘.

(上接第 158 页)

行处理. 针对非加密通信方式,如果源或目的地址属于黑名单,可以确定为匿名通信行为. 所有的加密通信都属于灰名单,称为疑似匿名通信行为. 为了进一步确定加密通信的性质,通过引入 Jackson 开环排队网络模型和分析方法,对由匿名代理所组成的覆盖网络建立网络模型,利用匿名通信网络模型中等待时间和逗留时间等参数,对加密通信行为进行进一步分析,如果符合该匿名通信网络模型,则确定该加密通信为匿名通信,否则为可信通信,实验结果表明:检测结果的准确率达到 70%以上.

## 5 结束语

以上匿名通信跟踪技术分析的前提条件是代理主机由于客观原因无法定位,也无法控制或伪装,跟踪者只能在若干个本地链路上部署检测器进行跟踪,并且发送者— $n$ 个匿名代理—接收者之间采用单播通信方式. 但是,当匿名通信引入了组播技术,如何跟踪又是一个新问题,这也是我们下一步研究

的重点.

#### 参考文献:

- [1] 王伟平,陈建二,陈松乔,等. 匿名通信中短距离优先分组重路由方法的研究[J]. 软件学报. 2004,5(4):561-570.
- [2] 眭鸿飞,陈松乔,陈建二,等. 基于重路由匿名通信系统的负载分析[J]. 软件学报,2004,15(2):278-285.
- [3] 眭鸿飞,陈建二,陈松乔,等. 重路由匿名通信系统中基于秘密共享的重路由算法[J]. 计算机研究与发展, 2005,42(10):1660-1666.
- [4] Shields C,Levine B N. A protocol for anonymous communication over the internet[C]//Proc 7th ACM Conference on Computer and Communication Security New York:ACM,2000: 33-42.

#### 作者简介:

程昊 男,(1976—),硕士,高级工程师. 研究方向为网络信息安全及特种通信工程技术.