文章编号:1001 - 9081(2010)03 - 0719 - 04

匿名通信综述

刘 鑫,王 能
(华东师范大学 计算机科学与技术系,上海 200062)
(xliu620@gmail.com)

摘 要:匿名通信是网络与通信领域的热门课题。首先描述了匿名通信的起源,并从匿名属性、对手能力和网络类型三个方面对匿名通信的基本框架进行了阐述。然后阐述了匿名通信的研究现状,并对若干具有代表性的匿名通信系统进行了简要描述,匿名通信系统包括 Anonymizer、Tor、Mixminion、Crowds 和 Tarzan。最后提出了匿名通信发展面临的挑战,包括匿名通信系统的用户体验、中继节点的信誉评价体系和匿名通信系统的滥用行为。

关键词:匿名通信;基本框架;中继节点
中图分类号: TP393.08 **文献标志码**:A

Survey of anonymity communication

LIU Xin, WANG Neng
(*Department of Computer Science and Technology, East China Normal University, Shanghai* 200062, *China*)

Abstract: Anonymity communication is a hot topic in the area of network and communication. The origin of anonymity communication was first outlined, and the framework of anonymity communication in terms of anonymity property, adversary capability and network type were described. Then the research of anonymity communication was investigated, and a brief description of several major anonymity communication systems including Anonymizer, Tor, Mixminion, Crowds and Tarzan was provided. Finally the challenges confronted in the development of anonymity communication were proposed, including the user experience of anonymity communication, credibility evaluation system of relay node and misbehavior of anonymity communication.

Key words: anonymity communication; basic framework; relay node

0 引言

匿名起源于古希腊语,其含义主要是指无法识别一个人的身份信息。

匿名是人类社会很多活动的基本需求之一。在现实社会中,可以看到很多活动是以匿名方式进行的,如选举活动、慈善捐款、违法行为举报以及大量的商业活动等,对于除参与者之外的第三方是匿名的。

随着 Internet 应用的发展,特别是基于网络的电子商务应用越来越多,如电子投票、网络银行、电子证券交易和电子商务等,人们在保护传输数据秘密性、完整性和真实性的基础上,越来越关注如何保护通信用户的身份信息,如何保护提供网络服务的用户身份信息,以及如何抵御对用户通信的流量分析,而这正是匿名通信的研究范围。

Internet 最初设计目的是为了信息共享,每一个 Internet 的成员都具有唯一的标识 IP 地址,每一个传输的 IP 报文都包括报文数据源的地址和报文目的地址。因此每一个用户的行为都可以唯一识别,通信中的每一个报文也都可以识别发送者和接收者。也就是说用户的通信匿名无法得到保证。

1 匿名通信的起源

匿名通信是由 Chaum[1] 提出的,他提出了基于 Mix 节点的匿名通信算法,Mix 节点接收多个发送者的消息,并对这些消息进行混合处理,然后传输给接收者,因此掩盖了发送者和接收者的身份信息,实现了匿名。

目前匿名通信是网络与通信领域的热门课题,美国军方和大学研究机构对此都有较深入的研究。图 1 显示了从 1981 年至 2008 年被 EI 检索的和被 Freehaven 收录的属于匿名通信研究领域的文章数量。

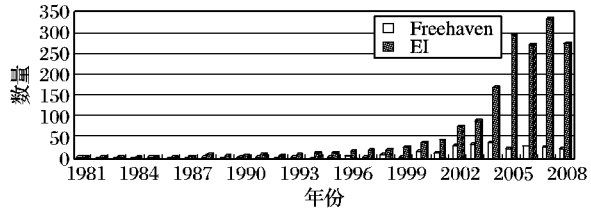

图 1 匿名通信研究文章数量示意图

2 匿名通信的基本框架

匿名通信的基本框架[2] 可以从三个方面加以阐述:匿名属性(anonymity property)、对手能力(adversary capability)和网络类型(network type)。

2.1 匿名属性

匿名属性包括不可辨识性(unidentifiability)和不可联系性(unlinkability)。不可辨识性是指对手无法识别用户的身份和行为;不可联系性是指对手无法通过观察系统将消息、行为和用户相关联。

不可辨识性由发送者匿名、接收者匿名、相互匿名和位置匿名四个部分构成。发送者匿名是指不能辨识消息发送者的

收稿日期:2009 - 09 - 06;**修回日期**:2009 - 10 - 23。

作者简介:刘鑫(1979 -),男,安徽明光人,博士研究生,主要研究方向:匿名通信、计算机网络; 王能(1942 -),男,上海人,教授,博士生导师,主要研究方向:计算机网络协议、计算机网络安全、无线传感器网络、移动 Ad Hoc 网络。

身份;接收者匿名是指不能辨识消息接收者的身份;相互匿名是指既不能辨识消息发送者的身份,也不能辨识消息接收者的身份;位置匿名是指无法辨识消息发送者和消息接收者的位置、移动、路由或拓扑信息。

不可联系性主要是通信匿名。通信匿名是指特定的消息不能和任意通信会话相关联,或者特定的通信会话不能和任意的消息相关联。通信匿名的匿名程度要低于发送者匿名和接收者匿名。

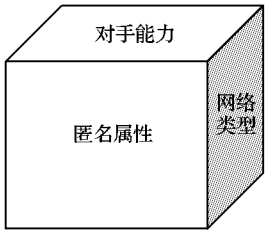

图2 匿名通信基本框架示意图

2.2 对手能力

对手是意图降低、消除通信匿名的通信网络用户或用户的集合。匿名通信系统一般通过提出威胁模型(thread mode),来表明该系统能够抵抗的对手能力。对手能力分为三个方面:可达能力(reachability)、攻击能力(attackability)和适应能力(adaptability)。

对手的可达能力分为全局(global)和本地(local)两种。具有全局能力的对手可以访问网络中所有的节点和链路,而具有本地能力的对手只能访问网络中部分的节点和链路。

攻击能力分为被动(passive)和主动(active)两种。攻击的目的是为了识别消息的发送者或接收者。被动攻击一般由匿名通信网络的外部观测者发起,其主要行为为观测网络中传输的消息、网络中数据的流量,并通过对消息和流量的分析达到攻击的目的。主动攻击一般由匿名通信网络的内部节点发起,其主要行为为通过其控制的部分通信节点修改通信消息、追溯通信行为、修改通信行为,来达到攻击的目的。

适应能力分为动态和静态两种。在匿名通信系统中,对手的适应能力一般是动态的,动态地跟踪网络的变化,实时地收集路径选择算法信息,实时地监控网络传输的消息和流量的变化。

2.3 网络类型

匿名通信系统的网络类型由以下三个因素确定,分别为:路径拓扑(path topology)、路由机制(route scheme)和路径类型(path type)。

匿名通信系统的路径拓扑有两种,分别为:瀑布型(cascade)和自由型(free)。在瀑布型的网络中,发送者从匿名通信网络中选择固定的通信路径进行消息的传输;在自由型的网络中,发送者可以选择任意长度的通信路径进行消息的传输。一般意义上,自由型的网络拓扑比瀑布型的网络拓扑具有更强的匿名。

匿名通信系统的路由机制分为单播(unicast)、组播(multicast)、广播(broadcast)和任意播(anycast)。目前基于系统效率和系统部署等实际问题的考虑,大多数实际部署的匿名通信系统的路由机制都是单播的机制。

匿名通信系统的路径类型分为简单(simple)和复杂(complex)。简单的路径类型不允许出现路径的循环,中继的节点在整个路径中只能出现一次;复杂的路径类型可以出现路径的循环,中继的节点在整个路径中可以出现多次。

3 匿名通信研究现状

现阶段匿名通信的研究主要分为三类:基于Mix算法的匿名通信系统、基于Onion Routing算法的匿名通信系统和基于泛洪算法的匿名通信系统。

基于Mix算法的匿名通信系统是在文献[1]的匿名通信算法基础上发展起来的。该类通信系统的核心思想是利用单个Mix节点或瀑布型的多个Mix节点实现匿名通信。Mix节点是指网络中向其他节点提供匿名通信服务的节点,它接收用其公钥加密的数据,并对数据进行解密、批处理、重序、增加冗余字节等处理,然后将数据传输给下一个Mix或最终接收者。基于Mix算法的匿名通信系统具有以下特点:

1)匿名通信系统网络中一部分节点为其他节点提供匿名通信服务;

2)发起者需要在发起匿名通信之前确定整个通信的传输路径,该路径在传输中不会改变;

3)发起者需要在发起匿名通信之前,得到整个传输路径中各个Mix节点的信息,包括地址、密钥信息等;

4)Mix节点对来自多个发送者的通信信息进行解密、复用、批处理、重序、增加冗余字节等处理,系统匿名较高,但通信传输的时延较高,一般不适合实时的数据通信。

基于Mix算法的匿名通信系统包括Babel、Cyberpunk(Type I)、Mixmaster(Type II)、Mixminion(Type III)等。

基于Onion Routing算法的匿名通信系统是在Reed等人[3]1998年提出的Onion Routing算法基础上发展起来的。相比于基于Mix算法的匿名通信系统,基于Onion Routing算法的匿名通信系统更注重数据通信的实时性以及系统的简单性、有效性和可实施性,其特点为:

1)基于Onion Routing算法的匿名通信系统建立在TCP传输的基础上,节点之间通常通过SSL方式传输;

2)基于Onion Routing算法的匿名通信系统在路径建立时采用非对称密钥算法加密,在数据通信时采用对称密钥算法加密,以提高数据传输效率,降低时延;

3)基于Onion Routing算法的匿名通信系统采用实时复用并转发,不对通信数据进行乱序、固定输入输出流量等批处理。

基于Onion Routing算法的匿名通信系统包括Tor、FreeNet等。Tor是目前Internet中最成功的公共匿名通信系统。目前,Tor在全球具有超过1 300个中继节点,大部分的节点位于美国和德国;正常状态下,全球有超过20 Gbps的匿名通信传输数据流量。

基于泛洪算法[4-5]的匿名通信系统是近期匿名通信传输领域新的研究热点,主要基于flooding、epidemic等类洪泛算法实现匿名通信,目前仍处于实验室研究阶段,没有实际部署的成熟的匿名通信系统。基于泛洪算法的匿名通信系统一般具有以下特点:

1)发起者在发起匿名传输之前完全不清楚匿名传输的路径,也无需得到传输中间节点的任何信息;

2)发起者的每一次匿名传输路径并不固定;

3)匿名通信网络中的任何一个中间节点都不知道匿名通信的发起者和接收者。

基于泛洪算法的匿名通信系统主要面临的挑战是系统会产生大量的网络传输流量,对于网络带宽的需求较大;同时在目前的状态下,系统算法的稳定性和可靠性还不够。

4　匿名通信系统

4.1　Anonymizer

Anonymizer[6]是建立在 HTTP 协议上的单一代理节点(Anonymizer-Server)匿名通信系统,通过过滤通信包文头部的身份鉴别信息和发送者的地址来实现匿名通信。Anonymizer 系统的优点是低延迟、实现简单,相对于非匿名通信系统提供了初级的匿名通信保障;缺点是匿名较弱、通信报文明码传输、通信匿名完全依赖于 Anonymizer-Server 节点,易受到攻击。

4.2　Tor

Tor[7-11]是第二代基于 Onion Routing 算法的匿名通信系统,目前以中继节点志愿的方式广泛地部署在 Internet 中,是 Internet 中最成功的公共匿名通信服务。Tor 网络在全球有超过 1 000 个的中继节点,大多数位于德国和美国,同时具有数以百万计的用户。

Tor 是基于通道(circuit)交换的低延迟的匿名通信服务。Tor 的设计引入了完美前向机密(perfect forward secrecy)、拥塞控制(congestion control)、目录服务(directory service)、完整性校验(integrity checking)和可配置的出口策略(configurable exit policies)等机制,解决了第一代基于 Onion Routing 算法的匿名通信系统设计的种种问题。

Tor 有两种实体,分别是 Tor 用户(Tor user)和 Tor 节点(Tor node)。Tor 用户在本地系统中运行 Onion Proxy(OP)程序,该程序负责建立通道,接收应用 TCP 数据流,并将该数据流通过已建立通道传输。

通道建立过程如下。

OP 首先访问目录服务,得到网络中 Tor 节点的信息,包括 IP 地址、公钥、出口策略、带宽和在线时间等。然后 OP 随机选择三个 Tor 节点作为中继节点,分别为入口节点、中间节点和出口节点。中继节点中只有入口节点知道通信发起者的身份,因此如何选择入口节点对于保护通信发起者的匿名十分重要。中间节点知道通道中入口节点和出口节点的身份,但是不知道匿名通信发起者和接收者的身份。出口节点作为网关负责 Tor 网络和外部 Internet 网络的应用层连接,并充当加密的 Tor 网络传输流量和非加密的 Internet 传输流量之间的中继。出口节点知道匿名通信接收者的身份。当 OP 在构建通道时,OP 和每一个中继节点协商共享的会话密钥。在这种设计下,通道中没有一个单一的节点知道匿名通信发起者和接收者的身份,因此实现了通信的匿名。

通信消息传输过程如下。

一旦建立完通道,OP 可以开始传输应用数据。OP 通过 SOCKS 协议接收应用程序数据,然后选择最新建立的通道进行传输。通信传输时,OP 将应用消息分割为 512 B 的 Cell,每一个 Cell 依次使用 OP 和中继节点共享的会话密钥进行加密,顺序如下:出口节点、中间节点和入口节点。当数据经过通道传输时,中继节点使用会话密钥进行解密,解密后传输给下一个中继节点。出口节点将消息还原为明文消息并传输给接收者。

Tor 适合于既有数据传输匿名要求,也有数据传输实时性要求的低延迟匿名通信,如 Web 访问和即时消息传输等。

4.3　Mixminion

Mixminion[12]是第三代基于 Mix 算法的匿名通信系统,在第二代基于 Mix 算法的匿名通信系统 Mixmaster 基础上增加了如下的机制:应答机制、综合目录服务期、哑元数据传输、前向匿名、基于密钥循环的重播预防机制、出口策略和接收者匿名。Mixminion 的通信路径是自由型的路径,路径中节点之间的通信采用 TLS 实现。应答机制是 Mixminion 增强的最显著功能。系统中 Mix 节点并不区分前向的数据消息和反向的数据应答消息,所以反向的数据应答消息和前向的数据消息具有相同等级的匿名。

由于 Mixminion 具有批处理、哑元数据传输等特性,相比于 Tor,Mixminion 更适合对数据传输匿名要求较高,对数据传输实时性要求较低的高延迟匿名通信,如邮件传输等。

4.4　Crowds

Crowds[13]是由 AT&T 实验室提出,和 Anonymizer 一样,是面向网络 Web 访问的匿名通信系统。Crowds 的匿名通信过程如下:

1)用户首先将通信消息随机传输给一个 Crowds 的中继节点(Crowds 称之为 jondos);

2)接收到通信消息的中继节点自主确定是将消息提交给最终的 Web 服务器,还是将消息随机传输给下一给 Crowds 的中继节点;

3)每一个接收到通信消息的中继节点循环执行第 2)步操作,直到该消息传输到最终 Web 服务器。

由于中继节点具有和匿名通信发起者同样的行为,因此 Crowds 确保了发送匿名和通信匿名。但由于最后一个中继节点知道 Web 服务器的身份,因此 Crowds 没有实现接收者匿名。

4.5　Tarzan

Tarzan[14]是 P2P 架构的匿名通信系统,其基本原理和 Tor 类似,都是通过分层消息加密和多跳路由实现匿名通信。发送者首先选择中继节点,基于中继节点构建静态传输通道,并产生哑元数据流量来提供匿名。Tarzan 能够提供发送者匿名、接收者匿名和通信匿名,但是通道的构建需要较大的计算开销和延迟。

5　匿名通信面临的挑战

匿名通信的发展目前仍面临诸多挑战[15-16],主要为:匿名通信系统的用户体验、中继节点的信誉评价体系和匿名信通信系统的滥用行为。

5.1　匿名通信系统的用户体验

目前匿名通信系统的设计的基本关注点是系统匿名和系统性能,往往忽视了系统的用户体验。

用户体验包括系统的易用性和易部署性,一个具有良好易用性和易部署性的匿名通信系统能够吸引更多的用户和中继节点,而更多的用户和中继节点能够为系统带来更强的匿名和抗攻击能力。反过来说,一个用户很少的匿名通信系统即使具备很好的通信性能和匿名算法设计,其提供的匿名也是不够的,因为攻击者很容易在很少的用户中遍历识别出通信的发起者和接收者。

Tor 是匿名通信系统中对用户体验认识较深的系统,也是实际部署中做得较好的系统,因此它也是目前最成功的 Internet 公共匿名通信服务,吸引了大量的用户和中继节点,即使中继节点需要自发自愿地贡献出自己的带宽和系统资源。

通过上述分析,本文认为系统用户体验和系统匿名、系统性能一样,是评估匿名通信系统优劣的重要元素,是匿名通信系统设计中必须考虑的重要环节,是决定匿名通信在实际部

署中成败的重要因素。因此用户体验是匿名通信系统进一步发展所面临的挑战,是系统设计者必须要解决的问题。

5.2 中继节点的信誉评价体系

目前,匿名通信系统的路由框架往往基于中继节点的自发行为实现。如在Tor系统中,中继节点自发地在目录服务中公告自己的信息,包括在线时间、IP地址和带宽等,用户根据这些信息选择所需的中继节点建立通道。Tor系统的目录服务对中继节点公告的信息不做任何的校验和确认。对手可以利用中继节点的自发行为发起路由攻击。对手依据用户选择中继节点的算法,公告适合的中继节点信息,使这些节点更加容易被用户选择为匿名通信的节点,降低甚至消除匿名通信系统提供的匿名,包括发送者匿名、接收者匿名和通信匿名,从而达到对用户匿名通信进行攻击的目的。

目前,匿名通信系统缺乏对中继节点既往行为的评价。一个为系统提供了较好服务(包括带宽、计算性能和匿名)的中继节点和一个为系统提供了较差服务(包括带宽、计算性能和匿名)的中继节点具有相同的系统地位,因此制约了中继节点提供更好服务的积极性,同时系统用户也无法区分系统内提供较好服务和较差服务的中继节点。

基于上述的分析,本文认为匿名通信系统需要建立中继节点的信誉评价体系,根据中继节点的既往行为进行中继节点评价,并基于信誉评价体系的结果构建匿名通信系统的路由框架。但是如果信誉评价体系构建不当,将会给系统带来新的匿名问题,并影响用户对匿名系统的使用及中继节点的部署,因此构建中继节点的信誉评价体系是匿名通信系统进一步发展所面临的挑战,系统的设计者必须考虑提出适当、简洁的信誉评价体系,促进匿名通信系统发展。目前构建中继节点的信誉评价体系是匿名通信系统的研究热点之一。

5.3 匿名通信系统的滥用行为

随着匿名通信系统发展,使用匿名通信服务的用户也越来越多,但是不幸的是,在这些用户的匿名通信中,既有合理合法的网络行为,也有大量的网络滥用行为,包括网络破坏、网络攻击和垃圾邮件等。滥用行为对于匿名通信系统的发展具有相当大的制约作用。

匿名通信系统是一个技术系统,是为人类社会发展服务的,因此要充分考虑系统对社会的影响。匿名传输应该在网络社会及人类社会的发展中起到积极的影响,而不是消极的影响,如果系统中充斥着大量的违规和非法行为,社会和国家是不会接受、支持的,其发展也必然受到制约。

Internet中,服务的提供者往往采用黑名单的方式来抵抗破坏者的攻击,即一旦发现某个网络节点具有大量的非法行为或产生了大量的非法流量,则将该节点的IP地址放入黑名单中,不接受来自该网络节点的任何数据流量。对于匿名通信系统来说,服务提供者面对的是系统的中继节点,当系统出现针对该服务提供者的滥用行为时,服务提供者将中继节点的IP地址放入黑名单,从而导致后续用户利用该中继节点的合法行为也不被接受。在这种机制下,最严重的后果是匿名通信系统的大量中继节点被放入Internet各种服务提供者的黑名单中,使匿名通信系统不能再为用户提供匿名通信服务。

因此如何抵抗滥用行为是匿名通信系统进一步发展所面临的挑战,系统的设计者必须考虑提出合理有效的抗滥用行为系统,既要保护合理合法通信的匿名,也要使那些非法的数据通信不能披上匿名的外衣而逃脱应有的监管和处罚,从而促进匿名通信系统的发展。目前各种抗滥用行为的算法和体系是匿名通信系统的研究热点之一。

参考文献:

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

[2] DOUGLAS K. A taxonomy for and analysis of anonymous communications networks [D]. Ohio: Air Force Institute of Technology, 2009.

[3] REED M G, SYVERSON P F, GOLDSCHLAG D M. Anonymous connections and onion routing [J]. IEEE Journal on Selected Areas in Communications, 1998, 16(4): 482-494.

[4] BANSOD N, MALGI A, CHOI B K, *et al*. MuON: Epidemic based mutual anonymity [C]// Proceedings of the 13th IEEE International Conference on Network Protocol. Boston, MA: IEEE Computer Society, 2005: 99-109.

[5] HAN JIN-SONG, LIU YUN-HAO. Rumor riding: Anonymizing unstructured peer-to-peer systems [C]// Proceedings of the 14th IEEE International Conference on Network Protocol. Santa Barbara, CA: IEEE Computer Society, 2006: 22-31.

[6] SERJANTOV A. Anonymizing censorship resistant systems [C]// IPTPS 2002: Proceedings of the 1st International Peer to Peer Systems Workshop. London: Springer, 2002: 111-120.

[7] DINGLEDINE R, MATHEWSON N, SYVERSON P. Tor: The second-generation onion router [C]// Security'04: Proceedings of the 13th USENIX Security Symposium. San Diego, CA: USENIX Press, 2004: 303-320.

[8] DINGLEDINE R, MATHEWSON N. Tor specification [EB/OL]. (2004-10-24) [2009-06-11]. http://www.freehaven.net/tor/cvs/doc/tor-spec.txt.

[9] DINGLEDINE R, MATHEWSON N. Tor path specification [EB/OL]. [2009-06-12]. http://www.freehaven.net/tor/cvs/doc/path-spec.txt.

[10] ØVERLIER L, SYVERSON P. Improving efficiency and simplicity of Tor circuit establishment and hidden services [C]// PET 2007: Proceedings of the 7th International Symposium on Privacy Enhancing Technologies. Berlin: Springer, 2007: 134-152.

[11] MCCOY D, BAUER K, GRUNWALD D, *et al*. Shining light in dark places: Understanding the Tor network [C]// PET 2008: Proceedings of the 8th International Symposium on Privacy Enhancing Technologies. Leuven. Belgium: Springer, 2008: 63-76.

[12] DANEZIS G, DINGLEDINE R, MATHEWSON N. Mixminion: Design of a type III anonymous remailer protocol [C]// SP'03: Proceedings of the 2003 IEEE Symposium on Security and Privacy. Berkeley, CA: IEEE Computer Society, 2003: 2-15.

[13] REITER M K, RUBIN A D. Crowds: Anonymity for Web transactions [J]. ACM Transactions on Information and System Security, 1998, 1(1): 66-92.

[14] FREEDMAN M J, MORRIS R. Tarzan: A peer-to-peer anonymizing network layer [C]// CCS 2002: Proceedings of the 9th ACM Conference on Computer and Communications Security. Washington, DC: ACM Press, 2002: 193-206.

[15] DINGLEDINE R, MATHEWSON N, SYVERSON P. Deploying low-latency anonymity: Design challenges and social factors [J]. IEEE Security and Privacy, 2007, 5(5): 83-87.

[16] DINGLEDINE R, MATHEWSON N. Anonymity loves company: Usability and the network effect [M]// Designing Security Systems That People Can Use. Sebastopol, CA: O'Reilly Media, 2005: 547-560.