

# 匿名网络 I2P 的安全性分析

李凌燕,魏庆征,杨 云,史庭俊,贺兴亚

(扬州大学信息工程学院,江苏 扬州 225127)

[摘要] I2P(invisible Internet project)是当前应用最广泛的匿名网络之一,采用大蒜路由的方式隐藏通信双方的通信关系,使用强加密协议和构建网络层的方法隐藏用户身份,并提供隐藏服务,进而达到全匿名的目的.但其也存在许多安全问题.通过对 I2P 网络的系统结构、工作原理、大蒜路由技术等方面的分析,指出了 I2P 存在的问题,并给出解决思路和技术构想.

[关键词] 匿名网络 I2P, Kad 算法,安全性

[中图分类号] TP393 [文献标志码] A [文章编号] 1672-1292(2018)03-0010-09

## Security Analysis of Anonymous Network I2P

Li Lingyan, Wei Qingzheng, Yang Yun, Shi Tingjun, He Xingya

(College of Information Engineering, Yangzhou University, Yangzhou 225127, China)

**Abstract:** I2P(invisible Internet project) is one of the most widely used anonymous networks. It uses garlic routing to hide the relationship between both sides of communication. It uses a strong encryption protocol and a network layer to hide the user identity and provide hidden services, thus achieving the purpose of full anonymity. But it also has many security issues. Through the analysis of the system structure, working principle and garlic routing technology of I2P network, the problems of I2P are pointed out, and the solution ideas and technical ideas are given.

**Key words:** anonymous network I2P, Kad algorithm, security

传统意义的网络安全一般包括 4 个要素<sup>[1]</sup>: 秘密性、完整性、可用性和真实性. 研究表明, 匿名性(Anonymity)在某种意义上也可称为是安全性的构成要素之一. 随着互联网上应用的增多, 互联网上的安全和隐私越来越受到人们的关注, 需通过提供匿名服务保护用户隐私, 一些应用如电子投票(E2V)、电子银行(E2B)、电子商务(E2C)等已将匿名性作为一个安全性的衡量指标.

利用现代加密体制, 例如公私钥加密、电子签名、密钥协商算法等, 能够很好地解决信息的机密性、完整性、可用性和真实性. 匿名通信技术已研究发展多年, 但至今尚未形成完整的匿名理论体系.

I2P 是当前应用最广泛的匿名网络之一<sup>[2]</sup>. I2P 采用大蒜路由的方式隐藏通信双方的通信关系, 使用强加密协议和构建网络层的方法隐藏用户身份, 提供隐藏服务, 进而达到全匿名的目的. 分析 I2P 网络的拓扑结构、工作原理、大蒜路由技术等, 可发现其存在许多安全漏洞, 本文针对这些漏洞给出了解决思路和技术构想.

## 1 匿名网络

匿名通信研究的不是如何保护数据的内容, 而是如何隐藏通信实体的身份信息, 使攻击者无法通过搭线窃听和流量分析数据报头得到用户的真实身份, 或对用户通信进行跟踪.

匿名通信系统的研究目的是对网络用户的 IP 地址、通信关系等涉及用户隐私的信息进行保护, 使其不被攻击者检测和发现<sup>[3]</sup>.

收稿日期: 2018-04-18.

基金项目: 江苏省产学研前瞻性联合项目(BY2016069-16).

通讯联系人: 杨云, 博士, 教授, 研究方向: TCP/IP 协议分析、匿名通信系统. E-mail: yyang@yzu.edu.cn

## 1.1 匿名通信类型

匿名通信技术有许多分类方法,其中按底层的路由机制可分为两类:一类是单播转发机制(基于重路由技术),另一类是通过广播/组播来实现匿名。

### 1.1.1 基于单播转发机制的匿名网络

David Chaum 提出的 Mix-Net 系统是基于单播转发机制的典型代表<sup>[4]</sup>,他定义了一个经过多个中间节点转发数据的多级目标路径,通过使用中间节点变换、混杂来自多个不同用户的消息,使窃听者无法确定输入/输出消息间的对应关系,从而无法跟踪消息的传输路径。为了隐蔽接收者,发送者可选定  $k$  个连续的目标节点,仅其中之一为真正接收者。因此,窃听者在一段链路上获取真正接收者的概率为  $1/k$ ,且中间节点在传送消息时可采取重新排序、延迟和填充等手段使得获取真正目标节点的概率更低,从而加大了攻击者进行流量分析的难度<sup>[5]</sup>。Mix-Net 系统使用了很多匿名通信技术,如嵌套加密、数据封包、延时、混淆、缓存重排等。Mix-Net 网络结构如图 1 所示。

Mix-Net、Tor<sup>[6]</sup>、I2P 均属于单播转发机制的匿名网络,Mix-Net 系统是 Tor 网络的祖先,I2P 网络是基于 Tor 的思想构建的,其匿名性和数据通信安全性更高。

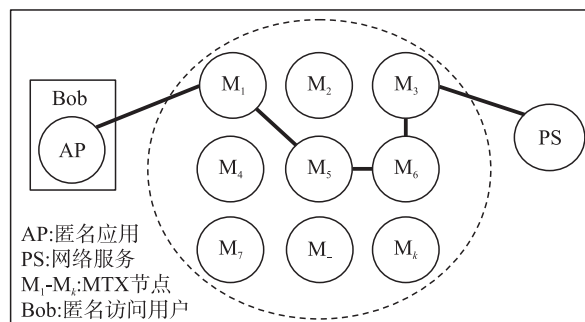


图 1 Mix-Net 网络结构

Fig. 1 Mix-Net network

### 1.1.2 基于广播/组播机制的匿名网络

基于单播转发机制的匿名是通过很多个中间路由构成匿名路径,而基于广播/组播的匿名技术是在一个广播组中隐藏真正的信息接收者。

基于广播/组播机制的匿名通信的核心思想是:在系统运行的每个周期,所有参与者(包括发送者)均向系统中各个成员广播一个报文。通过在接收到的所有报文中施以运算,各个参与者能够计算出所要广播的消息内容,而无法推断出发送者的身份。DC-Net 是其典型同时也最为著名的实现<sup>[7-8]</sup>。

基于广播/组播机制的匿名网络存在严重缺陷:一是广播容易引起信道冲突;二是报文数量过多。每次报文发送中都需要每个成员的参与,将导致效率和健壮性方面的问题。

## 1.2 匿名通信的保护类型

- (1) 匿名发送:保护发送者的身份标志(不能关联到通信发起者)。
- (2) 匿名接收:保护接收者的身份表示(接受者可能不确定)。
- (3) 匿名发送接收关系:找不到特定消息的发送与接收者,无法推测通信双方。

## 2 I2P 网络工作原理分析

I2P 是一个基于 Tor 的匿名网络,它只暴露一个简单的层,提供给应用程序之间进行匿名和安全的通讯<sup>[5]</sup>。这个网络本身是严格基于消息的(通过 IP 方式),但也存在一个库可用于在其上传输可靠的信息流(通过 TCP 方式)。所有的通讯都是端到端加密(发送一条消息一共进行四层加密),甚至是终点节点(目标)也是加密的标识(本质上是一对公钥)。

### 2.1 Tor 网络分析

Tor 网络设计目标不是提供代理服务器,而是实现匿名<sup>[9]</sup>。Tor 网络在应用过程存在如下问题:

(1) 网络速度慢。为了保证用户的匿名性,采用多层加密和多个节点(Tor 规定至少 3 个节点)转发,大大降低了连接速度,影响了用户体验。因为转发机制无法从根本上避免,故在匿名性与网络速度之间无法达到有效平衡。

(2) 高时延。基于中继的匿名系统设计有两个方向:高时延和低时延。相对较大和可变的延迟可最大化匿名性,使得高延迟的系统拥有高防御性。但网页浏览、即时通讯或 SSH 连接则无法忍受这样的高延迟。

(3) 中心化的系统结构。Tor 属于集中式设计,其正常工作完全依赖于目录服务器,如图 2 所示。Tor 始

终通过主目录获取最新中继信息,但因为主目录被屏蔽,因而需要设置网桥或其他代理.为防止目录服务器被冒充,目录服务器 IP 地址都被硬编码在客户端(主用 9 个,备用 151 个).中心化设计存在“单点故障”严重问题.

(4)单一的通信链路. Tor 网络的洋葱路由中,一条或多条数据流的上传与下载共用一条电路,且采用电路交换(circuit switching)方式进行数据交换,这种方式难以抗击“流量分析”攻击.

(5)亚匿名性.由于 Tor 随机地在节点之间传输数据包,虽保证了匿名性,但也存在缺陷.即当攻击者掌握了大量节点之后,虽然 Tor 会随机挑选节点,但当所有节点都被掌控时,攻击者虽然截获不了数据包的内容(因为多层加密的存在),却有可能替换内容,同时,用户发送数据包的源头和目标节点也被暴露,所谓的匿名性即不复存在.

(6)DHE 密钥交换. Tor 低版本使用 64 位的 AES 密钥交换.虽然 Tor 2.3 版本的 DHE 密钥交换使用了 64 位的密钥,但从丝绸之路到斯诺登事件说明:Tor 只是一个工具,可以实现匿名,但不确保用户可以安全地连接. Tor 网络的密钥交换存在问题,需要使用更高强度的密钥.

(7)洋葱路由.在洋葱路由中,入口节点、中继节点和出口节点这 3 个关键节点是随机选择算法随机选择的,可能出现两个问题:一是所选择的路径不是最佳的;二是同时多个用户选择了同一条路径,会造成同一条路径上流量过载,负载不均衡,存在安全隐患.

(8)信息泄漏.服务器间数据接收未进行身份认证,蓄意的攻击者在匿名网络中混入了恶意服务器,可能会篡改某些节点服务器上经过的流量.

## 2.2 I2P 系统构架

I2P 是一个专为隐藏服务设计和优化、完全分布式和自组织的、去中心化(无目录服务器)、使用复杂算法进行流量加密的系统,使用隐藏服务技术、强加密协议和构建网络层方法是 I2P 的重要特征. I2P 网络是一种建立在互联网之上的覆盖网络(overlay network),可称为虚拟互联网(virtual internet)<sup>[9]</sup>.

### 2.2.1 I2P 网络物理构成

I2P 的物理构成包括用户客户端、网络节点和网络数据库(NetDB)3 个部分<sup>[10]</sup>,如图 3 所示.

(1)用户客户端:用户安装在个人 PC 上的 I2P 程序.

(2)网络节点:在 I2P 网络上逻辑上平等的所有节点,既是 I2P 网络上的用户,又是 I2P 网络服务的提供者.

(3)网络数据库(NetDB):NetDB 是一个分布式数据库,包含两种类型的数据:路由器信息和隧道口地址.其中的每一条信息都是经过验证后正确的信息,信息中还含有动态成分,可将原有不相关的条目删除并将新加入的条目储存.目前 NetDB 采用了 Floodfill 技术,通过 Floodfill 路由器,维护分布式数据库<sup>[6]</sup>.

### 2.2.2 消息数据结构

同一个连接中的指令与数据,在 Tor 中沿着通过 TCP 协议建立的信道(circuit)流动至目的节点;而在 I2P 中,连接被消息机制(message)打散为数据包,经由不同的 TCP 或 UDP 隧道(tunnel)交叉传输后,在接收方重组为数据流.可见,Tor 基于电路交换,而 I2P 基于分组交换.

#### (1)消息

I2P 的用户在发送消息之前必须确定 2 个 I2P 隧道(路径),并获得该路径上所有中继节点的 RSA 公

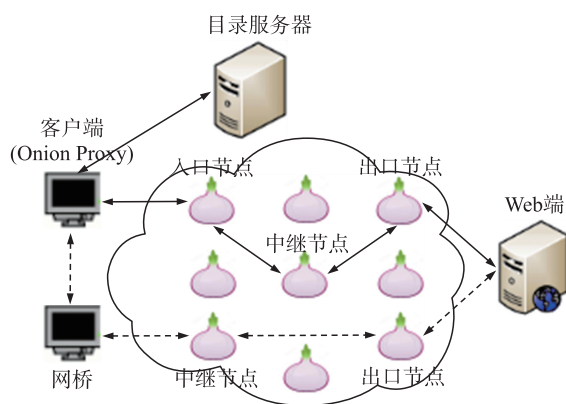


图 2 Tor 系统结构

Fig. 2 Tor system

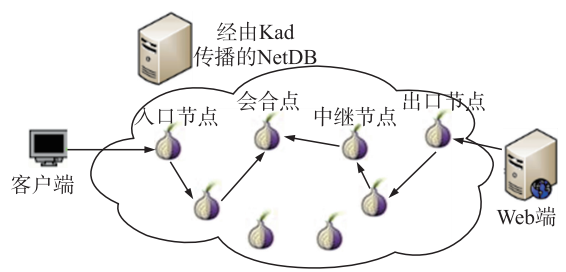


图 3 I2P 系统结构

Fig. 3 I2P system

钥,随后构造如下消息:

$$M_0 = K_0(R_0, M)$$

$$M_1 = K_1(R_1, M_0, A_0)$$

$$M_2 = K_2(R_2, M_1, A_1)$$

...

$$M_n = K_n(R_n, M_{n-1}, A_{n-1})$$

式中,  $K_i$  为 AES 对称密钥;  $R_i$  为 RSA 公钥;  $A_i$  为中继节点;  $M$  为原文;  $M_i$  为密文消息;  $M_n$  为用户最终构造出消息.  $M_n$  发送至第 1 个节点  $N$ ,  $N$  节点用自己的私钥解密得到  $M_{n-1}$ , 然后发送给节点  $N-1$ ,  $N-1$  节点用自己的私钥解密数据得到  $M_{n-2}$ , 依此类推, 最后的节点  $M_0$  得到用户发送的消息  $M$ . I2P 网络大蒜数据包帧结构如图 4 所示.

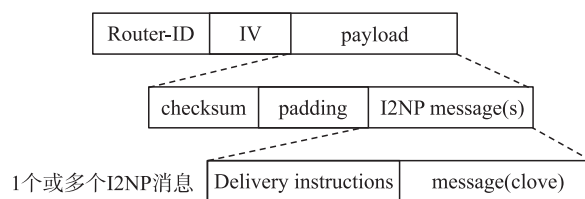


图 4 I2P 消息结构

Fig. 4 I2P message frame

## (2) 控制消息帧

控制消息帧结构如图 5 所示.



图 5 I2P 控制消息结构

Fig. 5 I2P control message frame

## (3) 传输消息帧

传输消息帧结构如图 6 所示.

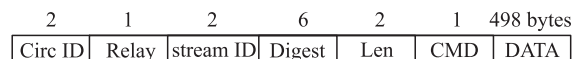


图 6 I2P 传输消息结构

Fig. 6 I2P transmission message frame

## 2.3 I2P 工作原理

匿名的发送消息时, 每个客户端程序都会用到自己的 I2P“路由器”, 这个路由器是由一系列出站和入站的“隧道”组成的, 而每一条隧道则是在一个方向(分为出站方向和入站方向)上传递信息的节点队列. 当一个客户端想要发送信息给另一个客户端时, 发送端会通过一条出站隧道将信息发出, 发向接收端的一条入站隧道, 最终到达消息的终点. 网络中的参与者可以根据需要, 通过设置隧道的长度, 在匿名性、延时、带宽之间取得平衡. 其结果就是, 大量的节点中继了每一条端到端的消息, 这样的模型将消息发送者和消息接受者的暴露风险降到了最小. 当一个客户端第 1 次试图与另一个客户端取得联系时, 它将会去查询分布式的“网络数据库”(一个自定义结构的、基于 Kademlia 算法的分布式哈希表 DHT). 这样做是为了有效地找到其他客户端的入站隧道, 但其间的后续消息中通常包含有这些数据, 因此没必要进行网络数据库的进一步查找.

## 2.4 大蒜路由(garlic routing)

大蒜路由起源于洋葱路由, 是一种扩展的洋葱路由<sup>[11-12]</sup>. I2P 采用大蒜路由的方式隐藏通信双方的通信关系, 即通过使用包含多个节点的隧道, 使得隧道中的任意单一节点都不能同时获知通信双方的身份信息, 从而达到匿名的目的.



### 2.4.1 大蒜路由原理

在大蒜路由中,客户端匿名访问暗网服务器的 Tor 电路需要经过 6 个 Tor 节点,客户端和服务端各通过一个 Tor 电路连接到一个中继 Tor 节点(称为“会合点”),这个会合点既不知道客户端地址,也不知道服务器地址,从而保证通信双方的匿名性.大蒜路由和洋葱路由的示意图分别如图 7、图 8 所示.

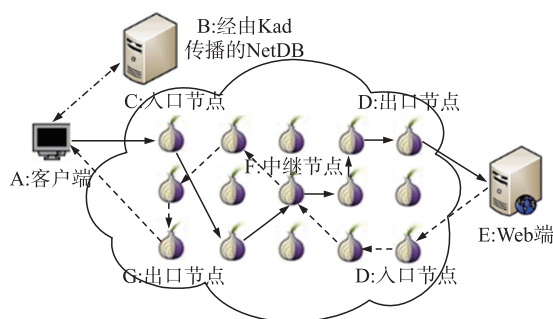


图 7 I2P 的大蒜路由

Fig. 7 Garlic Routing

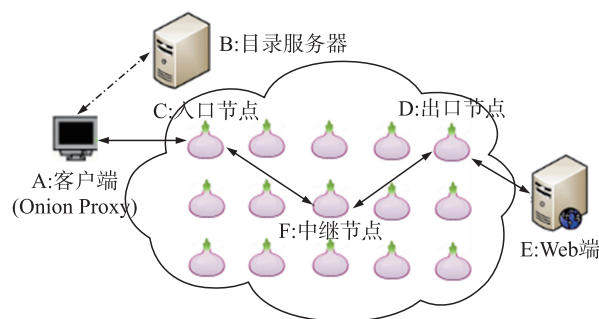


图 8 Tor 的洋葱路由

Fig. 8 Onion routing

### 2.4.2 大蒜路由与洋葱路由比较

大蒜路由与洋葱路由的主要区别:

(1) I2P 通过本地网络数据库 NetDB 得知其他节点的存在,NetDB 通过 Kad 算法在连接其他节点时获悉更多节点的存在. Tor 通过连接中央目录服务器得知所有中继、进入/退出节点的存在.

(2) I2P 接入网络后,I2P 建立 2 个隧道出站隧道 A-C 和入站隧道 A-G,分别负责数据的传出和传入,实际的隧道数更多,可能存在 A-C1、A-C2、A-G1、A-G2...当 2 个隧道建立后,I2P 客户端发出 HTTP 请求希望访问 Web 端的匿名资源. Tor 接入网络后,Tor 建立 1 个信道 A-C,负责数据的双向传输. 信道建立后,Tor 客户端发出 HTTPs 请求希望访问 Internet 的 Web 端资源.

(3) I2P 将请求拆分并加密为数据包,采用分组交换方式由出站隧道 A-C(A-C1、A-C2...)发送至对方的入站隧道网关 D(D1、D2...),通过入站隧道 D-E(D1-E、D2-E...)到达 E 的 I2P 数据包被接收并重组为 HTTP 请求. Tor 将请求加密,采用电路交换方式送入信道 A-C-F-D,通过 D 转接至 Internet 的 Web 端 E. 反之同理.

(4) I2P 的大蒜路由的中继节点之间同时使用 TCP/UDP 连接进行分组交换传输. Tor 的洋葱路由的中继节点之间使用 TCP 连接.

(5) 大蒜路由将传输的原始数据拆散为加密数据包,通过多条隧道交叉疏散传递上传与下载隧道相互独立,且两个方向上的隧道数量都可能大于 1. 洋葱路由中一条或多条数据流的上传与下载共用一条隧道.

## 2.5 节点加入与寻找

I2P 网络使用 Kad 协议、距离算法获取网络节点. 在 Kad 网络中,每个节点的路由表都可表示为一棵二叉树,将具有相同 ID 前缀的节点信息存放在 K-桶中,前缀就是该 K 桶在二叉树中的位置. 这样,每个 K 桶都覆盖了 ID 空间的一部分,全部 K 桶的信息汇总就覆盖了整个 160 bit 的 ID 空间.

### 2.5.1 Kad 协议

Kad 是 Kademlia 的简称,其以独特的异或算法(XOR)为距离度量基础,Kad 节点之间使用无连接的传输通信协议 UDP 连接,Kad 节点利用分布式散列表(distributed hash table,DHT)储存文件索引,其结构为二叉树的拓扑结构,协议追求的主要目标就是无中心服务器而实现快速定位期望的节点.

分布式散列表技术是一种分布式存储方法. 这种网络不需要中心节点服务器,而是每个客户端负责一个小范围的路由,并负责存储一小部分数据,从而实现整个 Kad 网络的寻址和存储. 和中心节点服务器不同,Kad 网络中的各节点并不需要维护整个网络的信息,而是只在节点中存储其临近的后继节点信息,在与关键字最接近的节点上复制备份冗余信息.

#### 2.5.1.1 分布式散列表 DHT

DHT 的主要思想是:将每条文件索引表示成一个(Key, Value)对,其中 Key 为关键字,可以是文件名

(或文件的其他描述信息)的哈希值; Value 是实际存储文件的节点的 IP 地址(或节点的其他描述信息). 所有的文件索引条目(即所有的 (Key, Value) 对)构成一个大的文件索引哈希表,只要输入目标文件的 Key 值,就可从这张表中查出所有存储该文件的节点地址. 将大文件哈希表分割成很多局部小块,按照特定的规则把这些小块的局部哈希表分布到系统中的所有参与节点上,使得每个节点负责维护其中的一块. 当节点查询文件时,由于各个节点维护的哈希表分块中含有要查找的 (Key, Value) 对,只要把查询报文路由到相应的节点即可.

#### 2.5.1.2 Kad 的基本元素

Node ID: 节点位置标志符,由节点 IP 地址应用 SHA1 (Secure Hash Algorithm, 安全哈希算法)生成的摘要与随机数发生器产生的随机数“凑合”组成,长度 160 bit.

Key: 查询关键字字符串、关键词字符串的 SHA1 散列或所需下载文件的 SHA1 校验值,长度 160 bit.

Value: 列表值,长度 160 bit,格式为 (File-name, File-lens, File-SHA1).

(Key, Value): 字典条目列表对,用于节点数据存储和查询.

K-bucket: K-桶,格式为 (IP address, UDP port, Node ID), 用于存储节点特征数据.

K 与  $\alpha$ : K 表示子树的个数,参数  $\alpha$  表示本节点知道的每个 K-桶中节点个数(种子数).

#### 2.5.1.3 Node ID、Key、Value 三者的关系

Node ID、Key、Value 的长度均为 160 bit,其设计目的在于保证三者是同构的.

(1) Node ID. 可以将一个 IPv4/IPv6 地址转换为十进制数:若地址为 IPv4,即 IP = A.B.C.D,则 ID number =  $A \times 2^{24} + B \times 2^{16} + C \times 2^8 + D \times 2^0$ ;若地址为 IPv6,即 IP = A:B:C:D:E:F:G:H,则 ID number =  $A \times 2^{112} + B \times 2^{96} + C \times 2^{80} + D \times 2^{64} + E \times 2^{48} + F \times 2^{32} + G \times 2^{16} + H \times 2^0$ . 当地址为 IPv4 时,由于一个 A 类网络中 IP 地址数为  $2^{24} - 2 = 167\,772\,164 < 167\,772\,166$ ,这表明:表示 ID number 的整数不会超过 32 bit;当地址为 IPv6 时,任何一个网络中 IPv6 的地址数为  $2^{112} - 2 = 5\,192\,296\,858\,534\,827\,628\,530\,496\,329\,220\,094$ ,这表明:表示 ID number 的整数不会超过 128 bit. 而 32 bit + 128 bit = 160 bit,即采用 160 bit 长度表示 Node ID,可以保证不同节点的 Node ID 相同的概率几乎为零.

(2) Key. 考虑到 DHT 需要承载的数据量通常较大,散列函数产生的“散列值范围”(keyspace)要足够大,以防止太多的碰撞. 若 keyspace“大到一定程度”,使得“随机碰撞”的概率可以忽略不计,就有助于简化 DHT 系统设计. 因而 DHT 中的 Key 采用 160 bit 的散列值(因为  $2^{160}$  比“地球上所有电子文档总数”还要大“很多数量级”).

(3) Value. 由于 Value 是存储文件的节点的 IP 地址或节点的其他描述信息,其格式为 (File-name, File-lens, File-SHA1), File-SHA1 与 Key 对应,故 Value 长度必须为 160 bit.

#### 2.5.2 距离算法

距离算法用于计算节点之间距离、数据之间距离、节点与数据的距离,距离属于逻辑层面,其与地理位置无关,也与互联网的拓扑结构无关.

采用 XOR(按比特异或操作)算法计算节点之间的距离,其具备类似于几何距离的特性(用  $\oplus$  表示 XOR):

$(A \oplus B) = (B \oplus A)$ , 交换律,具备对称性;

$(A \oplus A) = 0$ , 反身性,自身距离为零;

$(A \oplus B) > 0$ , 不同的两个 key 之间的距离必大于零;

$(A \oplus B) + (B \oplus C) \geq (A \oplus C)$ , 三角不等式.

采用距离算法可以在 I2P 网络中快速地搜索到距离最近的节点.

#### 2.5.3 二叉树拓扑与路由查询

##### 2.5.3.1 二叉树拓扑

在 Kad 网络中,所有节点都被当作一颗二叉树叶子,且每一个节点位置都由其 ID 值的最短前缀唯一确定,Kad 协议确保每个节点知道其各子树的至少一个节点,只要这些子树非空,每棵子树用 K-桶存储. 节点 0011 的子树划分如图 9 所示.

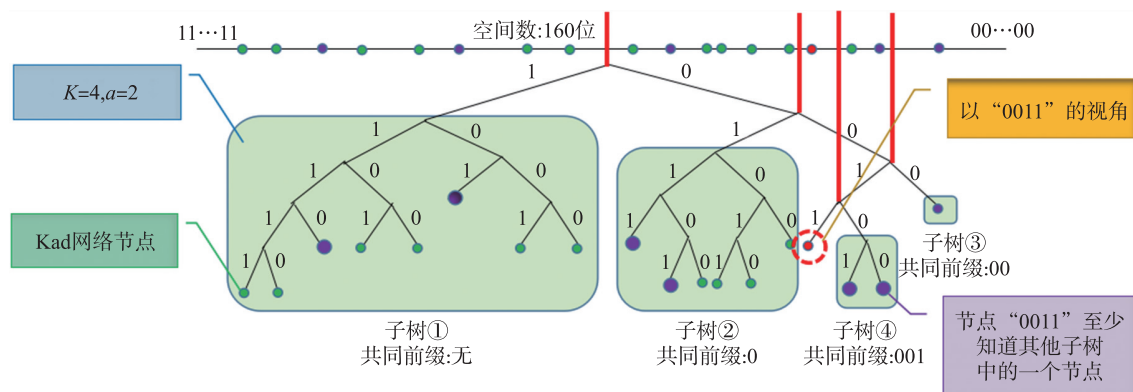


图 9 节点 0011 的子树划分

Fig. 9 Subtree division of node 0011

### 2.5.3.2 路由查询

路由查询机制中,通过 Node ID 值定位目标节点,路由查询递归过程的数学表达式如下:

$n_0 = x$  (即查询操作的发起者), Node ID 为目标节点 ID,

$n_1 = \text{find\_node}_{n_0}(\text{IP address}, \text{UDP prot}, \text{Node ID})$ ,

$n_2 = \text{find\_node}_{n_1}(\text{IP address}, \text{UDP prot}, \text{Node ID})$ ,

...

$n_l = \text{find\_node}_{n_{l-1}}(\text{IP address}, \text{UDP prot}, \text{Node ID})$ .

该递归过程一直持续到  $n_l = \text{Node ID}$ , 或者  $n_l$  的路由表中没有任何关于 Node ID 的信息, 则查询失败. 节点 0011 通过 Node ID 值搜索目标节点 1110 如图 10 所示.

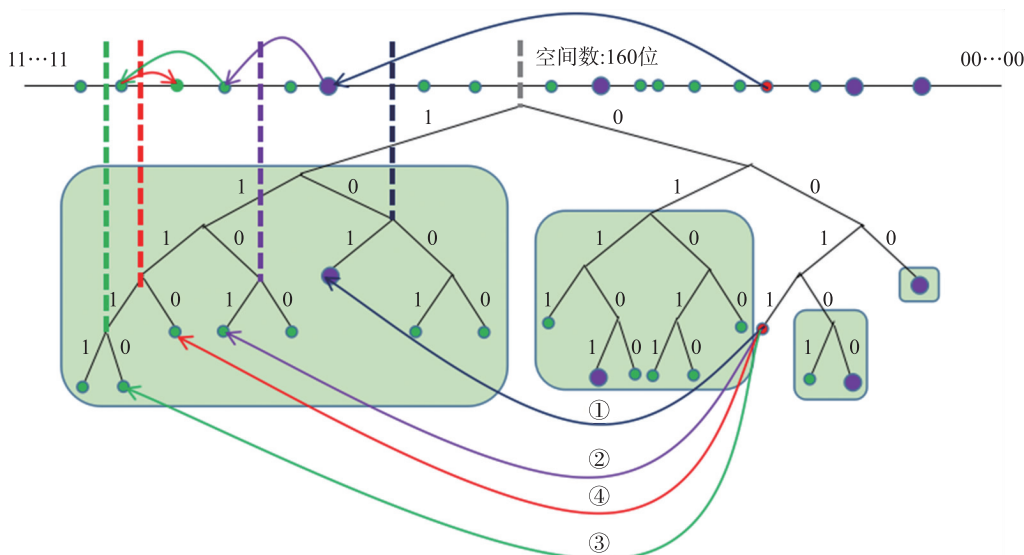


图 10 节点 0011 通过 Node ID 值搜索目标节点 1110

Fig. 10 Node 0011 searches for target node 1110 by Net ID value

### 2.5.4 Kad 数据结构

Class Kad {

IDS et=lookup(Net ID, k) % 返回与目标 Net ID 距离最近的 k 个节点;

Put(Key, Value) % 将 value 值放在离 Key 最近的节点上;

Value=Get(Key) % 将 key 值取回}.

## 3 I2P 的安全性

设计 I2P 的目的并不是为了实现所谓的“完美匿名”, 而是使得针对网络匿名性的攻击更加昂贵以至

于难于进行. 由于 I2P 属于低时延的混淆网络,这类网络系统自身存在一定的局限性.

### 3.1 网络结构

I2P 是依靠 Kad 算法通过不断扩散获取越来越多的网络节点信息. 在其运行过程中并不需要节点服务器, Kad 算法得到的节点信息只是整个 I2P 网络的一小部分. 因为 I2P 节点会动态变化,同时由于出口代理数量的限制,出口节点在安全上可能存在隐患. 从出口节点服务器的角度来看,其安全性不如 Tor.

与 Tor 不同,“出口节点”或“出口代理”并不是 I2P 网络固有的一部分,可能不在线. 对于 I2P 系统新用户来说,首次启动需要连接网络 Net DB 去下载 RouterInfo 文件信息,若网络 NetDB 被封堵或管控,则 I2P 客户端将无法获得 I2P 网络中的路由信息,即无法找到 I2P 系统中的 Peer 提供服务,无法使用 I2P 网络.

### 3.2 随机选择节点

I2P 网络节点众多,在建立隧道选择节点时,有可能会造成隧道跳数过多,导致请求应答的时间很长,严重影响服务质量.

I2P 网络采用随机选择算法选择节点,这是传统的地址空间随机加载技术. 地址空间随机加载技术是让代码在其运行的内存中进行转换,而 Selfrando<sup>[13]</sup>的工作方式是将不同功能的代码分开,并将其所运行的内存地址进行随机分布. 若攻击者不能准确猜到每个代码执行所在的内存地址,则其不能触发到内存中所存在的漏洞,也就不能让 Tor 浏览器运行他们的恶意代码,可以更好地防止黑客对 I2P 用户去匿名化的攻击. Tor 项目决定将 Selfrando 的解决方案纳入 Tor 浏览器的强化版本,该浏览器目前正在进行现场测试.

### 3.3 HTTP、HTTPS 与 Socks5 代理

Tor 采用 C++编写,出口中继为 Socks5 代理类型,中继节点之间使用 TCP 连接. I2P 采用 Java 编写,出口路由可以是任何使用主动连接的单连接 TCP 协议,包括 SSH 服务器、HTTP 代理服务器、HTTPS 代理服务器、Socks5 代理服务器等,也可以通过 Tor 客户端提供的网桥或 Socks5 代理接入 Tor 网络. 由于 I2P 仅支持使用主动连接的单连接 TCP 的协议,因而不支持 FTP 协议, I2P 的路由之间既使用 TCP 连接又使用 UDP 进行传输数据.

### 3.4 大蒜路由

I2P 系统中隧道的建立是基于重路由机制的. 重路由机制采取了组群思想,该思想可实现发送者匿名,其采用概率方法决定将信息继续转发给其他的节点或直接传送给接收者. 在该过程中有可能会造成隧道跳数过多,导致请求应答的时间很长,严重影响服务质量;也有可能形成很多 0 跳隧道,从而影响安全性和匿名性.

I2P 匿名通信系统 Net DB 中的节点分布在地理位置上,具有很大的不均衡性,会大大影响用户体验. 当用户把本机的 I2P 节点声明为 Floodfill NetDB 节点,从而获取整个 I2P 网络上的节点信息,则易给攻击者提供可乘之机.

### 3.5 密钥交换

分层加密是 I2P 和 Tor 的一个基本功能,也是覆盖网络保护匿名的基石.

Tor2.3 版本 DHE 密钥交换使用了 1 024 bit 的密钥, Tor 被建议使用 2.4 版本,其使用了 ECDHE 算法. I2P 网络在早期版本中就使用了 1 024 bit 的 ECDHE 进行密钥交换,但其 Kad 协议中匿名 IP 地址的 Net ID 使用了 SHA1.

DHE、ECDHE 算法由 DH、ECDH 演变而来. DH 是基于求解“离散对数问题”的困难, ECDH 是基于求解“椭圆曲线离散对数问题”的困难. DH 与 ECDH 的密钥是持久的(静态的),即通讯双方生成各自的密钥之后,可长期使用,但是存在安全隐患,无法做到“前向安全”<sup>[14]</sup>. 为了保证“前向安全”,采用“临时密钥”的方式对 DH 和 ECDH 进行改良,于是得到两种新的算法 DHE 和 ECDHE<sup>[15]</sup>,其基本思想是,即使有人记录流量并破坏服务器以获取其私钥,也无法破译该流量,因为其缺少不会保存的短暂 DH 参数. 算法中对每个会话都要重新协商一次密钥,且密钥用完就丢弃. 但 DHE/ECDHE 算法本身也有缺点,其不支持认证,也即:其虽然可以对抗“偷窥”,却无法对抗“篡改”,自然也就无法对抗“中间人攻击”. I2P 在最新的 0.9.35 版本也开始废除 SHA1 密钥,使用 ECDSA-SHA-512.

### 3.6 网络数据库 NetDB

Tor 网络采用目录服务器管理整个网络运行,而 I2P 网络基于分布式网络数据库 NetDB 管理整个网



络. NetDB 中存储着 I2P 网络中所有的信息,并向 I2P 节点提供查询服务. 但 NetDB 设计规定只有某些合格的(带宽最大的)对等体可以加入网络,这就使得 NetDB 的用户群、NetDB 的副本群规模较小,影响了 I2P 网络的运行效率. 取消带宽最大的限制,即可增加 NetDB 的用户群、NetDB 的副本群规模. 同时由于 NetDB 空间更加密集,意味着攻击者需要生成更多虚假路由密钥,攻击的成本将与网络的规模成正比<sup>[16]</sup>.

## 4 结语

匿名系统性能和安全性是一对矛盾结合体,提高了系统安全性自然要牺牲部分系统性能. Tor 系统具有集中性节点的好处:系统结构简单,用户只要从中心就可以及时获得节点的更新信息,但这个中心就是最大的安全隐患,因为其即为攻击的目标,一旦中心受到攻击,整个系统就陷于瘫痪. 而对于 I2P 系统来说,其优点在于基于 P2P 的分布式网络,整个系统中没有集中节点,因此攻击者对于局部的攻击不会影响到整个 I2P 系统的使用,系统安全性大大提高. 但其网络结构复杂,节点信息更新延迟大,验证和校验工作复杂,导致相应速度很慢.

I2P 目前面临的最大问题是网络规模太小,其流量特征很容易被识别. 对 I2P 网络进行安全性分析,发现其网络技术上存在的问题或安全漏洞,本文针对这些问题提出了技术建议,目的是从技术上探讨如何加强 I2P 网络的安全性,使得网络性能和安全性间两者协调,找到网络性能和用户隐私保障之间的平衡点.

### [参考文献] (References)

- [1] 高俊杰. I2P 匿名通信系统优化与实现[D]. 北京:北京大学,2014.  
GAO J J. The optimization and implementation of I2P anonymous communication system [D]. Beijing: Peking University, 2014. (in Chinese)
- [2] ANON. I2P[EB/OL]. (2018-07-02). <https://zh.wikipedia.org/zh-hans/I2P>.
- [3] 周勇. 基于 Tor 的匿名通信研究[D]. 西安:西安电子科技大学,2013.  
ZHOU Y. Research on anonymous communication based on TOR[D]. Xi'an: Xidian University, 2013. (in Chinese)
- [4] DAVID C. Mix-Net[EB/OL]. (2018-07-03). [https://en.wikipedia.org/wiki/Mix\\_network](https://en.wikipedia.org/wiki/Mix_network).
- [5] 黄文杰. 基于 Tor 网络的随机均匀分布路由算法[D]. 上海:上海交通大学,2012.  
HUANG W J. Uniform distribution routing algorithm based on Tor network[D]. Shanghai: Shanghai Jiao Tong University, 2012. (in Chinese)
- [6] ANON. Tor[EB/OL]. (2018-06-28) <https://zh.wikipedia.org/zh-hans/Tor>.
- [7] 周彦伟,杨启良,杨波,等. 一种安全性增强的 Tor 匿名通信系统[J]. 计算机研究与发展,2014,51(7):1538-1546.  
ZHOU Y W, YANG Q L, YANG B, et al. A Tor anonymous communication system with security enhancements [J]. Journal of computer research and development, 2014, 51(7): 1538-1546. (in Chinese)
- [8] 刘鑫. 基于 Tor 网络的匿名通信研究[D]. 上海:华东师范大学,2011.  
LIU X. Research on anonymous communication based on Tor network[D]. Shanghai: East China Normal University, 2011. (in Chinese)
- [9] 赵福祥,王育民,王常杰. 可靠洋葱路由的设计与实现[J]. 计算机学报,2001,24(5):463-467.  
ZHAO F X, WANG Y M, WANG C J. An authenticated scheme of onion routing[J]. Chinese journal of computer, 2001, 24(5): 463-467. (in Chinese)
- [10] 王伟平,陈建二,王建新,等. 基于组群的有限路长匿名通信协议[J]. 计算机研究与发展,2003,40(4):609-614.  
WANG W P, CHEN J E, WANG J X, et al. An anonymous communication protocol based on groups with definite route length[J]. Journal of computer research and development, 2003, 40(4): 609-614. (in Chinese)
- [11] 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,2004,15(4):561-570.  
WANG W P, CHEN J E, CHEN S Q, et al. Research on a short distance-prior rerouting scheme in anonymous communication[J]. Journal of software, 2004, 15(4): 561-570. (in Chinese)
- [12] 李金栓. 基于 I2P 的匿名通信协议分析与流量检测的研究[D]. 成都:电子科技大学,2015.  
LI J S. Research on the analysis of I2P anonymous communication protocol and flow identification[J]. Chengdu: University of Electronic Science and Technology of China, 2015. (in Chinese)

(下转第24页)

- [7] 王允臣,毕方明. 采用遗传算法优化点点连格棋评估函数参数[J]. 计算机工程与应用,2018,54(3):120-124.  
WANG Y C, BI F M. Using genetic algorithm to optimize parameters of evaluation function of Dots-and-Boxes[J]. Computer engineering and applications,2018,54(3):120-124.(in Chinese)
- [8] 林阳,赵欢,丁汉. 基于多种群遗传算法的一般机器人逆运动学求解[J]. 机械工程学报,2017,53(3):1-8.  
LIN Y,ZHAO H,DING H. Solution of inverse kinematics for general robot manipulators based on multiple population genetic algorithm[J]. Journal of mechanical engineering,2017,53(3):1-8. (in Chinese)
- [9] 朱剑英. 智能系统非经典数学方法[M]. 武汉:华中科技大学出版社,2001.  
ZHU J Y. Non-classical mathematics for intelligent system[M]. Wuhan: Huazhong University of Science and Technology Press,2001.(in Chinese)
- [10] 李楠,刘朋,邓人博,等. 基于改进遗传算法的无人机三维航路规划[J]. 计算机仿真,2017,34(12):22-25.  
LI N,LIU P,DENG R B,et al. Three dimensional path planning for unmanned aerial vehicles based on improved genetic algorithm[J]. Computer simulation,2017,34(12):22-25.(in Chinese)
- [11] 席光,蔡永林. 用改进遗传算法求取曲面间最小距离[J]. 计算机辅助设计与图形学学报,2002,14(3):209-213.  
XI G,CAI Y L. Calculation of minimum distance between free-form surfaces by improved genetic algorithm[J]. Journal of computer-aided design and computer graphics,2002,14(3):209-213.(in Chinese)

[责任编辑:陈 庆]

(上接第18页)

- [13] CONTI M,CRANE S,FRASSETTO T,et al. Selfrando: securing the Tor browser against de-anonymization exploits[J]. Proceedings on privacy enhancing technologies,2016(4):454-459.
- [14] 廖小平. 对前向安全代理盲签名方案的分析与改进[J]. 计算机系统应用,2018,27(3):217-220.  
LIAO X P. Analysis and improvement of forward secure proxy blind signature scheme[J]. Computer systems and applications,2018,27(3):217-220.(in Chinese)
- [15] ANON. Forward security[EB/OL]. (2018-05-25). [https://zh.wikipedia.org/wiki/Forward\\_security](https://zh.wikipedia.org/wiki/Forward_security).
- [16] JUAN P T,THIBAUT C,ISABELLE C,et al. Evaluation of the anonymous I2P network's design choices against performance and security[C]//ICISSP 2015—Proceedings of the 1st International Conference on Information Systems Security and Privacy. Angers,France,2015.

[责任编辑:严海琳]