doi:10.3969/j.issn.1000-436x.2014.09.005

对 Telex 互联网反监管系统的攻击

李龙海，黄城强，王万兴，慕建君
（西安电子科技大学 计算机学院，陕西 西安 710071）

摘 要：Telex 作为典型的路由器重定向型反监管系统给互联网监管者带来了新的挑战。为帮助用户逃避监管，Telex 利用路由器而非终端主机将用户的网络通信重定向到被屏蔽的目标站点。从审查者角度分析了 Telex 系统的安全性，提出了 2 类利用主动攻击破坏用户隐私的新方法。第一类为 DoS 攻击，利用了 Telex 握手协议的安全漏洞，在破坏系统可用性的同时还可能检出用户是否在使用 Telex 代理。同时给出了弥补该漏洞的改进协议。第二类称为 TCP 分组旁路攻击，利用非对称路由或 IP 隧道技术令客户端的部分 TCP 分组绕过路由器直达掩护站点，然后通过观察上行数据流的重传反应判断用户是否在使用 Telex 代理。通过一系列原型系统实验验证了旁路攻击的可行性。TCP 分组旁路攻击也适用于其他路由器重定向型反监管系统。

关键词：互联网监管；路由器重定向；用户隐私；DoS 攻击
中图分类号：TN918 **文献标识码**：A **文章编号**：1000-436X(2014)09-0040-17

Attacks on Telex Internet anticensorship system

LI Long-hai, HUANG Cheng-qiang, WANG Wan-xing, MU Jian-jun
(School of Computer Science and Technology, Xidian University, Xi'an 710071, China)

Abstract: As a typical router-redirecting based anticensorship system,Telex poses new challenges for Internet censors. To help common users evade Internet censorship, Telex employs network routers, rather than end-hosts, to relay network traffics to blocked destinations. The security of Telex from the censors' perspective is analyzed, and two kinds of active attacks aiming to break users' privacy are presented. The first is a kind of DoS attack, which exploits a security flaw of Telex handshake protocol. It can probabilistically identify the users who are using Telex, as well as break the availability of Telex. An improved handshake protocol to remedy the flaw is also proposed. The second is called TCP packets bypassing attack. Under that attacking scenario, censors make a small fraction of TCP packets from clients bypass the router and reach the cover site directly through asymmetric routing paths or IP tunnels, then determine whether a user is utilizing Telex by observing the reaction of upstream traffic. The feasibility of bypassing attack has been testified by a series of experiments in a prototype environment. The bypassing attack is also applicable to other router-redirecting based anticensorship systems.

Key words: Internet censorship; router-redirecting; user privacy; DoS attack

1 引言

互联网的广泛应用极大地促进了信息流动和传播的速度。互联网用户在足不出户的情况下可以随时掌握世界范围内政治、经济、文化、体育等各个领域的发展动态。而另一方面，互联网上的不良信息也泛滥成灾；利用网络传播谣言、色情、暴力信息，进行各种犯罪的行为屡见不鲜。在此情况下，依据法律法规对互联网实施必要的监管是各国通行的做法。目前世界上包括中国在内的 20 多个国家都实行了互联网监管(Internet censorship)制度，目的是对互联网承载和传播的内容进行监控、过滤、删除

收稿日期：2014-07-22；**修回日期**：2014-08-30
基金项目：国家自然科学基金资助项目（61101142）；中央高校基本科研基金资助项目（K50510030012）
Foundation Items: The National Natural Science Foundation of China (61101142); The Fundamental Research Funds of the Central Universities (K50510030012)

或屏蔽。这些国家在其管辖的互联网自治域的边界上部署了基于防火墙的网络过滤系统，并采用了 IP 地址封锁、DNS 域名劫持、关键字过滤等技术手段限制国内用户对部分境外 Web 网站的访问。被屏蔽的境外站点一般包含了危害国家安全和社会稳定的谣言，或者无法被本国接受的敏感政治内容，以及违反公众利益和社会道德的色情暴力信息。

由于政治观点和意识形态的不同，一些黑客、政治团体或非政府组织设计并实现了互联网反监管系统用于对抗监管系统对网络信息流动的控制。他们通过设置境外代理、建立潜信道、隐藏关键字等多种技术手段使部分境内用户可以不受限制地访问任意的境外站点，从而抵消了边界防火墙的屏蔽和过滤作用。为了抵御反监管攻击，实行网络监管的一方将这些境外代理的 IP 地址加入黑名单，有效阻断了客户端与境外代理之间的连接。与之对应，反监管方不断地征集新的可用 IP 地址与境外代理绑定，而时隔不久监管方再将新发现的代理服务器 IP 地址加入黑名单。该过程反复进行，最终演变成监管方和反监管方之间“猫捉老鼠”的游戏。

2011 年 Wustrow 等提出了一种称为 Telex[1]的对抗互联网监管、翻越边界防火墙的新方法。另外 2 个独立的研发团队几乎在同一时期也提出了 2 种新的反监管技术 Cirripede[2]和 Decoy Routing[3]，它们采用了与 Telex 非常相似的原理，即基于“路由器重定向”的代理技术。Telex 要求将反监管代理部署在境外的主干路由器上。反监管客户端利用标准的 HTTPS 加密信道伪装成访问某个被允许访问的境外非屏蔽站点，要求传输路径必须经过部署 Telex 代理的路由器。客户的上行网络数据分组被该路由器截获并被重定向到客户指定的屏蔽站点，目标 IP 地址在转发前做了相应替换。反监管代理接收到该屏蔽站点的响应数据分组后将其源 IP 地址替换为非屏蔽站点的 IP 地址，最后发送给客户端。在上述通信过程中，监管方防火墙截获的数据分组的源 IP 和目标 IP 地址分别为客户端和非屏蔽站点的 IP 地址(对于下行流量则相反)，并且通信内容被加密，因此该防火墙根本无法觉察到客户端在使用反监管代理访问屏蔽站点。即便监管方能够确定部署反监管代理的路由器的 IP 地址，也无法根据 IP 地址阻断通向该代理的所有网络流量，因为这样做会造成“过度屏蔽”现象，众多的必须经过该路由器访问的非屏蔽站点将受到牵连。根据文献[2]的分析，如果将反监管代理部署在多个一级 Internet 自治域的边界路由器上，基于目标 IP 地址的封锁策略将会造成大规模的网络断裂。而迫于经济和舆论方面的压力，绝大多数实行网络监管的国家并不愿意将自己所辖的自治域变成互联网孤岛。因此，这种基于路由器重定向技术的反监管技术是对目前互联网屏蔽技术的强有力的挑战。

本文首先对 Telex 反监管技术的原理和安全性进行了分析，发现了 Telex 系统的一个设计缺陷。为了避免使用反监管代理的用户遭受迫害，用户隐私性是 Telex 力图实现的一个关键安全特性。而 Telex 的设计缺陷使监管者能够感知到客户端是否正在使用 Telex 代理，且在特定条件下，客户端的通信内容会完全暴露给作为掩护的境外非屏蔽站点。针对该缺陷，本文也给出改进方法。站在监管者的角度提出了一种能够准确判断客户端是否在使用 Telex 代理访问屏蔽站点的攻击方法——TCP 分组旁路攻击法。利用该攻击，监管防火墙可以实时地识别出伪装的 TCP 连接并及时将其关闭，还可以准确锁定 Telex 客户端的 IP 地址并根据 IP 地址追查用户真实身份。所提的 TCP 分组旁路攻击法同样适用于另外 2 种基于路由器重定向技术的反监管代理系统 Cirripede[2]和 Decoy Routing[3]。

2　相关工作

2.1　互联网监管技术

近年来，迅速发展的 Internet 对人类社会的政治、经济、文化和生活方式的影响越来越大，各国政府对互联网监管问题日益重视。2010 年底的“阿拉伯之春”政治运动使中东多个国家的社会发生了深刻变革，Internet 在此过程中发挥的催化剂作用使各国政府对互联网监管的重要性有了更深刻的认识。在此背景下，网络监管方和反监管方之间的竞争日趋激烈，双方所使用的技术也在不断升级。

在互联网发展早期，SSL、TLS 等加密传输协议尚不流行，网络主机之间主要通过明文进行通信，因此监管方主要利用了 IP 地址封锁、DNS 域名劫持和关键字过滤 3 种技术手段限制信息的流动[4]。这 3 种技术的过滤准确度依次升高，而它们的开销也依次增大。通过 IP 地址过滤屏蔽域外 Web 网站最为简单有效，但存在过度屏蔽问题，因为将多个网站部署在 IP 地址固定的同一物理主机或集群上的现象十分普遍。基于 DNS 的过滤策略能够

解决该问题，但对于众多由用户贡献内容的 Web 2.0 站点如在线社交网络 Facebook 而言仍然太过粗糙。基于敏感关键字进行过滤可以精确地屏蔽特定网页或电子邮件，但防火墙必须实时地检查每个 IP 数据分组的内容，因此对传输性能影响较大。

2000 年之后，越来越多的互联网应用通过加密信道进行通信，能够利用 HTTPS 安全传输协议访问的网站越来越多，客户端与反屏蔽代理之间的传输信道也更加隐蔽。在此情况下，一些更加复杂的帮助审查通信内容合法性的技术被提出，其中包括带状态记忆的深度 IP 分组过滤技术[5]和主动探测技术。这些技术除了要求检查每个 IP 分组的内容，还要提取每个 TCP 连接的流量、延时等统计特征，甚至可能需要主动向域外主机发送请求以通过其响应行为判断是否为反监管代理[6]，因此对监管方防火墙带来的开销非常大。图 1 从代价和过滤准确性角度对相关的互联网监管技术做了对比。

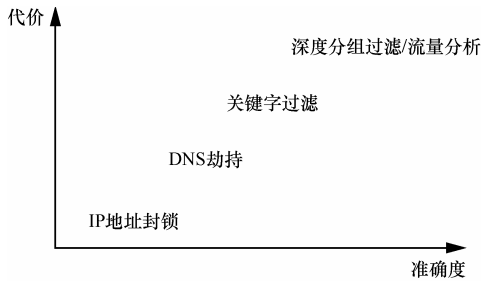


图 1 互联网监管技术对比

2.2 互联网反监管技术

要突破自治域边界防火墙的封锁最容易想到的方法是在域外设置代理，让域内用户通过代理间接访问被屏蔽网站。单一的代理服务器很容易被防火墙根据 IP 地址封锁，因此反监管方一般会征集多个具有不同 IP 地址的主机构成代理服务器池。

为了保证大部分用户能获得服务，反监管方征集新的 IP 地址的速度必须大于监管方搜集代理主机地址的速度。在这场“猫捉老鼠”的游戏中，反监管方也将付出巨大代价。为了更好地解决该问题，2011 年出现的 Telex[1]、Cirripede[2]和 Decoy Routing[3] 3 种新的反监管代理方案采用了全新的思路：将代理嵌入到境外网络路由器中。这 3 种方案具有相似的工作过程：客户端伪装成访问某个非屏蔽站点，要求传输路径必须经过部署代理的路由器(折射路由器)；为了不影响正常的路由功能，用户访问屏蔽站点的请求利用潜信道发送给该路由器；折射路由器识别出代理请求之后将用户的通信数据重定向到用户指定的屏蔽站点；路由器收到屏蔽站点的响应数据后再发送给客户端，并在发送之前将每个数据分组的 IP 地址替换成非屏蔽站点的 IP 地址。从监管方角度看，经过该路由器的每个非屏蔽站点都可能被用作反屏蔽代理。如果监管方切断通向该路由器的路径，相当数量的非屏蔽站点将受到影响而无法被访问，因此这种基于“路由器重定向”的反屏蔽代理技术可以有效对抗监管方基于 IP 地址黑名单的屏蔽手段。

Telex[1]、Cirripede[2]和 Decoy Routing[3] 3 种方案的主要区别在于客户端和代理之间采用了不同的潜信道方式进行通信。相比较而言 Telex 最为安全和高效。

与 Telex 思想类似的 CensorSpoofer[7]反监管方案是将非法通信流隐藏在 VoIP 的音频流中，并利用 IP 地址欺骗技术伪装成用户正在和某些被允许的主机进行音频通话。这些主机一般作为多个语言通信用户的代理，如果将它们的 IP 地址拉黑，将会影响大量合法用户的通信功能，因此提高了监管者过滤网络的代价。

2.3 针对 Telex 的攻击

站在监管方角度，对 Telex 等路由器重定向反屏蔽代理的攻击主要针对 2 类目标：一是针对系统的可用性，即如何阻断用户和 Telex 代理之间的连接；二是针对系统的不可觉察性，即如何及时判断特定的互联网用户是否正在使用 Telex 代理。

在可用性攻击方面，Schuchard 等[8]提出了基于非对称路由的绕路攻击方法，指出具有丰富网络资源的国家级监管者，如中国，很容易针对 Telex 所在位置有意制造非对称路由，使 Telex 不可用。但是 Houmansadr 等[9]对 Schuchard 提出的绕路攻击的代价做了更深入的分析，指出该攻击虽然理论上可行，但网络监管者将在传输延迟、网络连接性、路由长度等方面付出巨大代价，而且会导致相当数量的网络自治域边界路由器修改策略，因此大规模绕路攻击并不可行。

在不可觉察性攻击方面，Telex[1]和 Decoy Routing[3]方案的作者都提到了基于流量分析的客户端检测方法，即监管者通过分析 TCP 连接建立次数、每次连接的持续时间和传输的数据量等统计特征来判断客户端是否在真正访问合法站点。但这种基于流量分析的识别方法需要利用深度 IP 分组过滤技术[5,10]，

监管者防火墙运算开销大，因此识别速度会很低。另外，Telex 服务器端还可以通过插入冗余数据分组、改变传输延迟等方法伪装通信的统计特征。

Schuchard[8]等对 Telex 的安全性做了全面研究，除了绕路攻击，还提出了“TCP 分组重放”、“疯狂伊朗”和“时间分析”3 种针对不可觉察性的攻击方法。其中，前 2 种攻击都是基于非对称路由的，第 3 种攻击基于深度 IP 分组过滤技术。

本文所提的 TCP 分组旁路攻击也利用了非对称路由，但同时也给出了在不具备非对称路由时的攻击方法——IP 隧道法，适合于网络资源较少的监管者，并且对 3 种基于路由器重定向的反屏蔽代理方案都是有效的。本文提出的基于 DoS、基于 Telex 握手协议漏洞的主动攻击方法可以快速准确地识别 Telex 客户端，并且监管防火墙计算量很小。

3　Telex 方案回顾

3.1　基本工作过程

3.1.1　系统构架

Telex 方案假定系统构架如图 2 所示。该系统主要由 Telex 客户端、监管者的边界防火墙、折射路由器、Telex 代理服务器、被允许访问的站点和被禁止访问的站点等参与者构成。其中，Telex 客户端和监管防火墙位于被监管者控制的互联网自治域，任意客户端访问域外站点所产生的通信数据都会经过监管防火墙的过滤。Telex 服务器被部署在一些不受监管者控制的网络服务器提供商(ISP)的路由器上，作为路由器的扩展模块存在。该路由器被称为折射路由器。在 Telex 服务器失效的情况下，折射路由器仍能完成正常的网络路由功能。Telex 服务器通过路由器扩展接口与其进行交互，可以监听符合特定条件的通信数据，并向路由器发送命令。

3.1.2　攻击模型

Telex 假定监管者的目标是利用边界防火墙的过滤和屏蔽作用禁止自治域内用户访问某些域外站点，但迫于经济和政治舆论等方面的压力，监管者不愿意造成大规模互联网断裂，并且允许用户利用 HTTPS 等广泛存在的加密传输协议与域外主机通信。另外假设具有多项式计算能力的监管者虽然能监听域内任意的网络链路，但不能控制域内用户的主机。

3.1.3　基本工作过程

1) Telex 客户端选择一个合适的没有被列入黑名单的域外站点 NotBlocked.com，要求通向该站点的网络路径必须经过某个部署 Telex 代理的折射路由器。

2) Telex 客户端开始利用 HTTPS 协议访问 NotBlocked.com。在 TLS 握手阶段，客户端将一个用 Telex 服务器公钥加密的标签隐藏在“ClientHello”数据分组中，只有持有私钥的 Telex 服务器能够识别出该标签。除了表明获得代理服务的意图，该标签还被客户端用于和 Telex 服务器进行密钥协商。

3) 折射路由器通过扩展监听接口将所有 HTTPS 协议握手阶段的数据转发给 Telex 服务器。

4) 如果 Telex 服务器能够从某个 HTTPS 连接中识别出有效标签，则指示路由器继续向自己转发该连接的通信数据并令路由器中断与 NotBlocked.com 之间的 TCP 连接；否则，Telex 服务器将停止监听该 HTTPS 连接，并指示路由器针对该连接只完成正常的路由功能。

5) 针对识别出的网络连接，Telex 服务器将扮演 HTTP 代理的角色，即将客户访问屏蔽站点 Blocked.com 的请求转发给 Blocked.com，并将该站点的响应转发给客户端。客户端与 Telex 服务器进行通信时，所有数据都利用步骤 2)建立的共享密钥

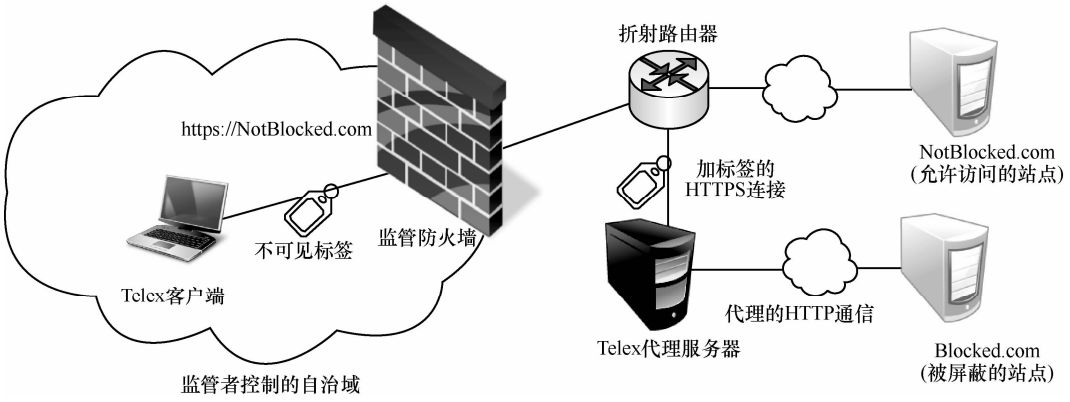


图 2　Telex 系统构架

进行加密，因此防火墙无法通过通信内容判断客户端的真实意图。

Telex 代理方案设计需要满足如下主要特性。

1) 不可屏蔽性：因为 Telex 代理服务器将被部署在互联网的主干路由器上，因此即便监管者知道 Telex 代理的具体位置，也无法屏蔽其服务，否则将招致不可接受的代价（相当数量的合法站点将无法访问）。

2) 不可察觉性（隐私性）：监管者无法察觉客户端是否正在使用 Telex 代理服务，更无法获得客户端与 Telex 代理服务器之间的通信内容。

3) 容易部署：Telex 代理服务器作为扩展模块部署在路由器上，即便 Telex 服务器失效，路由器仍能完成正常的网络路由功能，该特性使路由器的运营商更容易接受部署 Telex 代理。

3.2 Telex 握手协议

本节主要讲述客户端在 TLS 握手协议中嵌入隐藏标签、Telex 服务器识别标签以及它们利用标签进行密钥协商的具体过程。

3.2.1 参数选择

设 p 为大小合适的素数，其二进制长度为 l_p。E 和 E' 是定义在有限域 F_p 上的 2 个“孪生”的椭圆曲线，分别用 $y^2=x^3-3x+b$ 和 $-y^2=x^3-3x+b$ 定义。选择合适的 b 使 E 和 E' 的阶数都为素数。Telex 代理服务器随机选取 $r\in\{0,1\}^{l_p}$ 作为私钥，并计算相应的公钥 $\alpha_0=g_0^r$ 和 $\alpha_1=g_1^r$，其中，g_0 和 g_1 分别是 E 和 E' 的生成元。Telex 系统还需选择 2 个密码学散列函数 $H_1:\{0,1\}^*\to\{0,1\}^{l_{H_1}}$ 和 $H_2:\{0,1\}^*\to\{0,1\}^{l_{H_2}}$。代理系统对所有用户公布的参数包括：$E$ 和 E' 的描述信息、生成元 g_0 和 g_1、公钥 α_0 和 α_1 以及散列函数 H_1 和 H_2。

为了能够在保证安全性的前提下将标签嵌入到 TLS 握手阶段 224 bit 的随机数中，必须小心地选择 l_p、p、b 和 l_{H_1} 的值。它们的具体取值方法可以参考文献[1]。另外为了使标签足够短，Telex 只用 x 轴坐标表示 E 或 E' 上一个点，其 y 轴坐标可以用椭圆曲线的定义式计算出来。

3.2.2 TLS 握手协议

安全传输层协议（TLS）由 2 个子协议组构成：TLS 握手协议和 TLS 记录协议。而客户端和 Telex 服务器之间的握手协议是基于 TLS 握手协议修改而成，因此先对 TLS 握手协议作简要回顾。

图 3 所示为一个基本的 TLS 握手过程。下面描述双方交换的每个数据分组的具体内容。

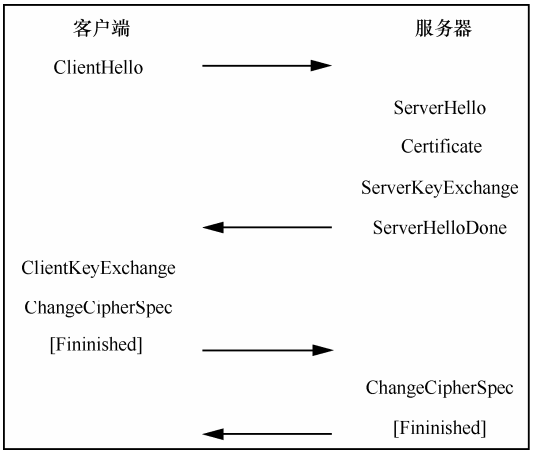


图 3 TLS 握手协议

1) ClientHello 由 4 byte 时间戳、28 byte 随机数、会话恢复标识以及客户端支持的密码学算法列表 CipherSuites 构成。

2) ServerHello 由 4 byte 时间戳、28 byte 随机数以及服务器从 CipherSuites 列表中选择的密码学算法标识组成。

3) Certificate 包含了服务器的 X.509 证书用于服务器向客户端进行认证。

4) ServerKeyExchange 包含了服务器端的密钥协商参数，与双方采用的密钥协商算法有关。如果采用 Diffie-Hellman 密钥交换算法，则该参数等于 $g^{s_{priv}}$，其中，s_{priv} 为服务器随机生成的秘密值，g 为某个循环群的生成元，在该群上离散对数问题难解。

5) ServerHelloDone 是一个空的 TLS 记录分组，用于更新接收端的 TLS 状态。

6) ClientKeyExchange 包含了客户端的密钥协商参数。如果采用 Diffie-Hellman 密钥交换算法，则该参数等于 $g^{c_{priv}}$，其中，c_{priv} 为服务器随机生成的秘密值。

7) ChangeCipherSpec 用于通知服务器端 TLS 记录将开始采用协商成功的密钥进行加密。

8) Finished 用于验证密钥协商结果，具体采取的方法是将之前双方交互的所有消息取散列值再加密，接收方用协商生成的共享密钥进行解密并验证解密结果。

3.2.3 Telex 握手协议

该协议对 TLS 标准握手协议做了修改，使客户

端能够将一个仅能被 Telex 服务器识别的标签隐藏在 TLS 握手数据分组中，并且双方开可以利用该标签生成共享的秘密信息。Telex 客户端与服务器之间的具体握手过程如下。

1) 客户端选择合适的非屏蔽网站(假设为 NotBlocked.com)，利用 DNS 域名解析获得其 IP 地址 server_ip，然后尝试与该主机的 443 端口(HTTPS 协议端口)建立 TCP 连接。

2) TCP 连接建立成功则进入 TLS 握手阶段。客户端随机选取 $s \in \{0,1\}^{l_p}$ 和 $b \in \{0,1\}$ 并计算标签

$$\tau = g_b^s \| H_1(\alpha_b^s \| \chi)$$

其中，χ=server_ip||timestamp||TLS_session_id 为上下文字符串，用于防范重放攻击。Telex 对相关参数做了巧妙选择使标签 τ 长度恰好为 224 bit，并且在不知道密钥 $r(\alpha_0 = g_0^r, \alpha_1 = g_1^r)$的条件下任意第三方能够将 τ 和均匀分布的随机字符串区分开的概率是可忽略的。客户端按照 TLS 协议规定构造 ClientHello 数据分组并用 τ 替换其中 224 bit 的随机数。最后，客户端将 ClientHello 分组发送给 NotBlocked.com。标签 τ 除了用来表明用户的特殊需求，还用于生成客户端与 Telex 服务器共享的密钥 $k_1 = H_2(\alpha_b^s \| \chi)$。

3) 客户端接收到 NotBlocked.com 的应答分组 ServerHello、Certificate 和 ServerKeyExchange 之后仍严格按照 TLS 协议规定对 NotBlocked.com 进行认证、记录对方的密钥交换参数和更新自己的 TLS 状态。

4) 客户端在构造密钥协商分组 ClientKeyExchange 时需要用到的所有随机数都用 PRG 伪随机数生成器生成，并以共享密钥 k_1 为种子。ChangeCipherSpec 和 Finished 数据分组仍严格按照 TLS 协议规定进行构造。最后客户端将 ClientKeyExchange、ChangeCipherSpec 和加密的 Finished 数据发送给服务器端。

5) Telex 服务器监听到 ClientHello 分组之后，将其中的随机数解析为 $\tau = \beta \| h$ 的形式，然后验证 $h = H_1(\beta^r \| \chi)$ 是否成立。如果不成立，则停止监听该连接，并指示路由器针对其只完成正常的路由功能。如果成立，则说明该用户使用了 Telex 客户端软件并且有访问代理服务器的需求，因此进入步骤 6)。

6) Telex 服务器利用监听到的特殊标签 $\tau = \beta \| h$ 生成共享密钥 $k_1 = H_2(\beta^r \| \chi)$。在监听到双方的密钥协商分组 ServerKeyExchange、ClientKeyExchange 之后，再用 k_1 作 PRG 种子生成客户端与 NotBlocked.com 进行密钥协商时用到的所有秘密值(例如 Diffie-Hellman 密钥交换算法中客户端的秘密指数 c_{priv})。最终 Telex 服务器可以获得客户端与 NotBlocked.com 的共享对称密钥 k_2。

7) Telex 服务器利用 k_2 对双方的 Finished 数据分组进行解密。如果解密失败，则停止监听该连接，并指示路由器针只完成正常的路由功能(该步骤主要用于防范监管方的重放攻击)。如果解密成功，则通知路由器向 NotBlocked.com 发送一个 TCP RST 分组以中断与 NotBlocked.com 的连接。之后 Telex 服务器利用密钥 k_2 代替 NotBlocked.com 与客户端进行通信，类似于在客户端与 NotBlocked.com 之间实施“中间人攻击”。

8) 上述握手完成之后，客户端通过加密信道将访问 Blocked.com 的 HTTP 请求发送给 Telex 服务器，该服务器将请求转发给 Blocked.com 并将响应数据转发给客户端。

4　利用 DoS 攻击破坏用户隐私

监管方有多种办法对 Telex 系统实施拒绝服务(DoS, denial of service)攻击，但自身也要付出一定的代价。一个比较有价值的发现是利用 DoS 攻击不仅可以破坏 Telex 的可用性，还可以破坏用户的隐私性，即能够检测客户端是否安装了 Telex 客户端软件以及是否正在请求 Telex 代理服务。这类攻击源于 Telex 握手协议的一个安全漏洞，即客户端利用隐藏标签向 Telex 服务器发出请求之后，Telex 服务器并没有向客户端确认是否能够为其提供代理服务。

下面首先讨论如何定位 Telex 代理服务器的部署位置，然后介绍对 Telex 实施 DoS 攻击的具体方法，最后重点分析如何利用 DoS 攻击破坏 Telex 用户的隐私。

4.1　如何定位 Telex 代理服务器

对 Telex 进行 DoS 攻击必须遵循一个基本原则，即不能造成严重的互联网断裂和影响普通用户对域外合法网站的访问。因此，为了有效实施 DoS 攻击，监管者必须获得 Telex 代理服务器的部署位置，这样监管者发出的 DoS 攻击分组才能准确到达 Telex 代理，避免“误伤”其他合法站点。

在原始的 Telex 方案[1]中，所有 Telex 代理服务器的部署位置都是向用户公开的，这样用户可以方便地选择合适的掩护站点使客户机通向掩护站点的路径刚好经过某个 Telex 代理。但即便反监管方不公开这些信息，监管者仍然很容易利用主动检测法枚举出绝大部分 Telex 代理的部署位置。

Schuchard 等在文献[8]中对检测 Telex 代理位置的方法做了详细分析。其主要思想是首先根据 CAIDA[11]的公开数据构建 Internet 的实际拓扑图，然后监管者以普通 Telex 用户身份尝试连接多个域外合法站点，并根据尝试结果对拓扑图进行反复修剪。Schuchard 等还对上述 Telex 代理检测算法的成功率进行了仿真分析[8]。仿真时以中国、伊朗等国家为假想的监管者，并在 Internet 拓扑图中随机选取部分节点部署 Telex 代理，然后运行检测算法。仿真结果表明在各种情况下超过 90%的代理都能被检测出来。

4.2 基于资源消耗的 DoS 攻击

实施 DoS 攻击的方法可以分为 2 类，一是大量消耗 Telex 服务器的资源，使其不能正常工作；二是破坏客户端与 Telex 服务器之间的握手协议，使 Telex 客户端不能被正确识别。

下面给出一种消耗 Telex 服务器计算资源的 DoS 攻击方法。该攻击主要利用了 Telex 系统的以下几个特点。

1) 每接收到一个客户端 TLS ClientHello 握手分组，Telex 服务器都要利用复杂的椭圆曲线群指数运算验证数据分组中是否含有隐藏标签，而客户端生成一个 ClientHello 分组几乎不需要花费任何计算量。因此，在该握手步骤上客户端与服务器端所付出的资源消耗不对等，这是能够成功实施 DoS 攻击的关键。

2) 根据原 Telex 方案的设计，为了减小对折射路由器正常网络路由功能的影响，Telex 服务器在每个 TLS 握手协议的开始阶段只是被动监听。因此，客户端与 NotBlocked.com 之间的 TLS 握手过程与 Telex 服务器检查隐藏标签的过程是同步进行的。如果 Telex 服务器没有在 TLS 握手协议结束前完成标签检验过程，则无法以“中间人”身份及时介入到客户端与 NotBlocked.com 之间的通信过程。

3) Telex 服务器作为折射路由器的扩展模块部署，在 Telex 服务器无法正常工作时，折射路由器仍能提供正常路由服务。

具体的攻击方法如下：监管者伪造大批具有不同 IP 地址的客户端（这对掌握大量网络资源的政府级监管者而言是非常容易的），并令它们利用 TLS 协议同时访问不同的域外站点，且访问路径都经过目标路由器。为了减小网络带宽开销和降低对普通用户的影响，伪造客户端在 TLS 握手阶段可以只发出 ClientHello 数据分组后而不做其他操作，服务器端的应答也可以简单丢弃。这样在短时间内 TLS 服务器会监听到大量的 ClientHello 握手分组，因此需要花费大量计算资源检查这些握手分组中是否含有隐藏标签。此时如果有真正的 Telex 客户端向折射路由器发送 ClientHello 分组，由于其标签无法被及时识别，该客户将无法获得正常的 Telex 代理服务。这些认证失败的 Telex 客户端只能被折射路由器当做普通网络用户对待，它们的应用层数据将被转发给掩护站点 NotBlocked.com，而该用户在通信开始阶段对此毫不知情。

攻击可行性实验用一台性能相对较好的计算机 (Intel Quad Core i7-4700, 2.4 GHz, 8 GB 内存)模拟 Telex 服务器，另外 2 台计算机 Client0 和 Client1 分别作为普通 Telex 用户和 DoS 攻击者。3 台主机利用 100 Mbit/s 以太网互联，构成如图 4 所示的拓扑。

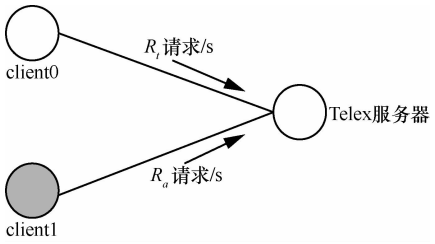


图 4 资源消耗 DoS 攻击实验

模拟的 Telex 服务器程序用 1 个线程处理客户端的 Socket 连接请求和接收 TLS ClientHello 分组，另外 3 个线程用于识别 ClientHello 分组中是否含有隐藏标签。数据接收线程和 3 个数据处理线程之间采用定长队列相连，当队列已满时，新接收的 ClientHello 分组将被丢弃。隐藏标签识别算法采用了 Telex 原型系统[12]中的相关源码。

第一次实验中，Client0 以频率 R_t 向 Telex 服务器发送 TCP 连接请求和 ClientHello 分组。在服务器端记录被实际处理的 ClientHello 分组数目和被丢弃的 ClientHello 分组数目。为了避免客户端计算能力影响实验结果，Client0 发出的 ClientHello 分组都是预生成的只包含 45 byte 的最小合法 TLS ClientHello 分组。设单位时间内能够被服务器成功处

理的 ClientHello 分组数为 R_x。在这里，无论是否被识别出包含有效标签，只要经过了识别子程序的处理，都被认为是成功处理的 ClientHello 分组。在该实验中，R_t 逐渐增大然后测试 R_x 的变化情况。发现在初始阶段 R_x=R_t，直至 R_t 增大至 930 左右；之后再增加 R_t 时，R_x 的值保持在 930 左右，利用 Linux 的 top 命令查看 CPU 使用率已超过 90%。这说明实验用 Telex 服务器的 ClientHello 分组最大处理能力 C 约为 930 分组/s。

在第 2、3、4 次实验中，分别令攻击者 Client1 以 R_a = C=930、2C=1 860 和 3C=2 790 的恒定速率向 Telex 服务器发送 ClientHello 分组，每个分组都对应一个新的 Socket 连接。这些 ClientHello 分组都是 Client1 利用伪随机算法生成的，长度都为 45 byte。在 Client1 发起 DoS 攻击的同时，逐渐增大 R_t 然后测试 R_x 的变化情况，得到如图 5 所示的实验结果。

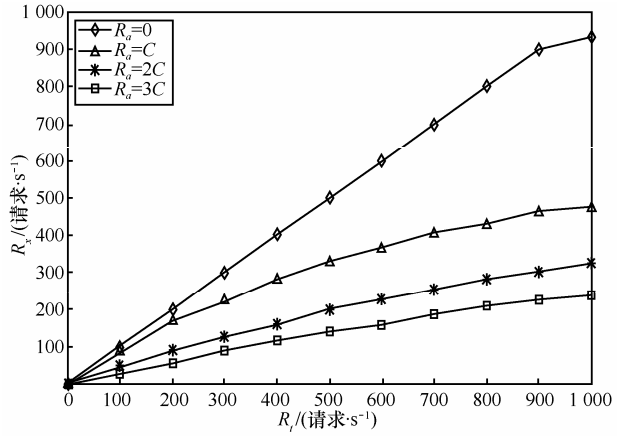


图 5　DoS 攻击下被实际处理的请求数

通过图 5 可以看出，随着 DoS 攻击强度的增加，正常用户发出的 ClientHello 分组只有部分得到处理，其他都被丢弃。当攻击速率 R_a=3C 且正常用户请求速率 R_t 等于 900 时，大量的服务器 CPU 资源都被消耗在识别攻击者伪造的 ClientHello 分组上，只有大致 1/4 的正常 ClientHello 分组被处理。在此情况下，至少有 3/4 的 Telex 用户无法获得正常的 Telex 代理服务。另外，在攻击速率 R_a 分别等于 C、2C 和 3C 时，攻击者消耗的带宽分别约为 1.17 Mbit/s、2.34 Mbit/s 和 3.51 Mbit/s（2 个 TCP 握手分组各占 40 byte，包括 IP 头和 TCP 头在内的一个 ClientHello 分组长为 85 byte），因此攻击者进行 DoS 攻击所花费的代价是非常低的。

当然，Telex 服务器对抗上述 DoS 攻击的能力和其硬件配置以及软件实现方式有关。为了提高 ClientHello 分组处理速度，关键的隐藏标签识别算法可以用纯硬件实现。当服务器最大处理速率 C 大到一定程度时，DoS 攻击者将付出巨大的带宽资源，此时还会影响普通用户对域外合法网站的访问。

4.3　基于破坏握手协议的 DoS 攻击

这类 DoS 攻击主要原理是通过破坏握手协议使 Telex 服务器无法识别 TLS 握手分组中的隐藏标签，因而也无法区分 Telex 用户和普通用户。实施该类 DoS 攻击的基本前提是不能破坏普通用户与 NotBolcked.com 之间的正常通信。与资源消耗型 DoS 攻击比，基于破坏握手协议的 DoS 攻击成功率更高，但需要在自治域外合适的位置上部署受监管者控制的物理主机。下面描述 2 种具体的攻击方法。

1) 攻击方法 1

①监管者在自治域外部署一个受自己控制的主机 A，要求客户端访问 A 的路径经过部署 Telex 的折射路由器。

②通过类似于透明代理的技术将可疑客户端发出的数据分组的目标 IP 地址都修改成域外受控主机 A 的 IP 地址，并令 A 作为 HTTP 代理将数据分组转发给 NotBolcked.com，然后将应答分组转发给客户端。对于一般用户，上述操作不会影响它与 NotBolcked.com 之间的正常通信。

③Telex 客户端与 NotBolcked.com 之间(实际是与主机 A 之间)的 TLS 握手分组被 Telex 服务器截获。客户端构造的标签 τ 中包含通信上下文 χ = server_ip||timestamp||TLS_session_id。由于 TLS 握手分组的目标 IP 地址与 χ 中的 server_ip 不符，Telex 服务器会将 τ 视为无效标签，因而拒绝为该客户端提供代理服务。

2) 攻击方法 2

文献[1]中提到了一种利用 DNS 劫持改变服务器端 IP 地址的攻击：客户端查询 NotBolcked.com 的 IP 地址时，被监管者控制的域内 DNS 服务器返回域外受控主机 A 的 IP 地址，并同样令 A 充当 HTTP 代理角色。由于 χ 中的 server_ip 与客户端握手分组的实际目标 IP 地址相符，τ 会被识别为有效标签。在该攻击下如果 Telex 服务器仍然为该用户提供代理服务，则很容易被 A 发现。文献[1]给出的应对方法是令 Telex 服务器根据服务器端 X.509 证书中的域名(即 NotBlocked.com)再做一次域名解析过程。为 Telex 服务器提供服务的 DNS 服务器不受

监管者控制，所以能够返回正确的 IP 地址。如果 Telex 服务器发现自己获得的 IP 地址和 χ 中的 server_ip 不同，则可以断定客户端遭受了 DNS 劫持攻击。为了防止监管者继续窥视该客户的隐私，Telex 服务器将其视为普通用户，不再为其提供代理服务，这样监管者看到是该客户访问 NotBlocked.com 的正常数据流。如果 Telex 服务器采用上述应对措施，那么监管者可以进一步利用 DNS 劫持实现对 Telex 系统的 DoS 攻击，其代价接近于攻击方法 1，都小于基于资源消耗的 DoS 攻击。

攻击可行性实验。为了验证攻击方法 1 描述的 DoS 攻击的可行性，利用 5 台主机进行了攻击模拟实验。这 5 台主机的连接方式如图 6 所示，其中深色的节点表示被监管者控制的网络节点。

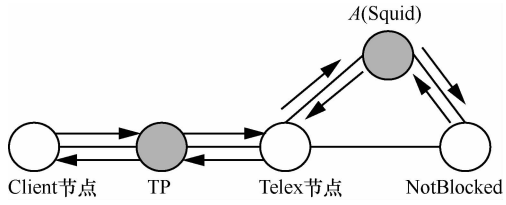


图 6 破坏握手协议 DoS 攻击实验

图 6 中 Client 节点表示普通的 Telex 客户端，TP(transparent proxy)表示被监管者控制的路由器，Telex 节点表示部署了 Telex 代理的域外路由器。A 表示被监管者控制的域外物理主机。NotBlocked 表示 HTTPS 掩护站点。TP 和 Telex 2 个路由器都用装有 Linux 系统的双网卡计算机实现，它们的 ip_forward 选项都被开启，并配置了正确的路由表。Client、TP 位于同一个子网内，TP 为网关；Telex、*A* 和 NotBlocked 位于同一个子网，网关为 Telex。TP 主机利用 Linux iptables 命令[13]的 DNAT 功能选项将 Client 发向 NotBlocked 的 IP 分组的目标地址修改为主机 *A*，因此 Client 的 HTTPS 请求都被重定向到了 *A*。*A* 中部署了 HTTP 代理软件 Squid[14]。Squid 将 Client 的 HTTPS 请求转发给 NotBlocked，然后再将 NotBlocked 的响应通过 TP 转发给 Client。

实验结果证明，即便 Client 发出的 ClientHello 分组中包含了有效标签，该数据分组经过 Telex 路由器时也无法被正确识别，因此该用户无法获得正常的 Telex 代理服务。而非 Telex 用户仍然能够访问未被屏蔽的站点 NotBlocked。

4.4 检测用户是否在使用 Telex 代理

在 4.2 节和 4.3 节描述的 DoS 攻击中，如果标签识别失败或者标签被 Telex 服务器因负载过重而丢弃，客户端对此将毫不知情，仍然发送访问 Blocked.com 的 HTTP 请求。这些请求最终会被掩护网站 NotBlocked.com 获得。现在分析 NotBlocked.com 对这些请求的反应。

假定用户在浏览器中输入了 http://Blocked.com/test.html，那么浏览器会通过 Telex 客户端本地代理进程向 NotBlocked.com 发送如下请求：

GET http://Blocked.com/test.html HTTP/1.1\r\n

HOST: Blocked.com\r\n

……[其他部分被省略]……

NotBlocked.com 对上述请求很可能会返回带有错误状态代码和错误信息的 HTTP 响应分组。具体地讲，会触发如下 2 类错误。

1) Host 识别错误。如果 NotBlocked.com 上的 Web 服务器软件工作在多虚拟主机模式，即允许同一物理主机部署多个具有不同域名的虚拟主机，那么该 Web 服务器会根据 HTTP 请求 URL 和 HOST 字段中的域名区分客户正在请求的虚拟主机。对于无法识别的域名“Blocked.com”，Web 服务器会返回特定的错误状态代码。根据 W3C 的标准，Web 服务器一般会返回带有错误状态码“400 Bad Request”HTTP 响应分组。

2) 页面缺失错误。如果 NotBlocked.com 没有工作在多虚拟主机模式，那么 HTTP 请求中的 HOST 字段将被忽略。但由于本地并不包含网页资源 test.html，因此一般会返回带有错误状态码“404 Not Found”的 HTTP 响应分组。

另外如果用户想利用 Telex 代理访问域外的 HTTPS 站点，例如在浏览器中输入了 https://Blocked.com/test.html，那么浏览器会向 Telex 客户端本地代理进程发送如下请求，要求该 HTTP 代理与主机 Blocked.com 首先建立 TCP 通道：

CONNECT Blocked.com:443 HTTP/1.1\r\n

Proxy-Connection: keep-alive\r\n

……[其他部分被省略]……

在正常情况下，Telex 客户端本地代理会将上述 CONNECT 请求利用 TLS 加密通道发送给 Telex 服务，Telex 服务器再建立与 Blocked.com 之间的 TCP 连接。但在 Telex 服务器被 DoS 攻击的情况下，CONNECT 请求实际上被发送给了 NotBlocked.com，因此必然会被 NotBlocked.com 拒绝。根据以上分析，Telex 客户端的错误请求还会触发第 3 类错误，

表 1　　**常用站点对 3 类错误的响应**

站点	HOST 错误		页面缺失		CONNECT 请求	
	响应码	分组长度	响应码	分组长度	响应码	分组长度
bing.com	400	403	301	306*	400	426
google.com	400	462	404	1 172*	405	1 141
amazon.com	忽略	—	404	15 298	400	176
walmart.com	400	403	404	57 892*	400	426
wikipedia.org	400	200*	404	3 314*	400	3 256*
usatoday.com	400	403	301	226	400	426
msn.com	301	295*	404	41 195*	400	472
yahoo.com	302	545	404	1 021	302	545
aol.com	忽略	—	404	23 141*	503	188
apple.com	400	403	404	10 201*	400	426
hotmail.com	忽略	—	302	137*	400	503
bbc.co.uk	301	688	404	18 053	301	632
lycos.com	301	485	404	2 190	400	415
flickr.com	404	210	404	17 287*	400	66
imdb.com	301	1 018*	404	478	400	176

即 CONNECT 请求方法错误。

选取 15 个常用站点并利用测试程序向它们发送上述 3 类错误请求。表 1 给出了这些站点的响应情况。其中，“分组长度”表示站点返回的 TCP 分组的长度。长度后加“*”表示该长度不固定，主要和请求的 URL 字符串内容有关，但基本上接近于表中给出的数字。

由以上测试可以看出，Internet 中大部分专业网站都不会忽略 HTTP 请求中的 HOST 字段，一般以会返回带有错误状态码“400 Bad Request”的响应分组，或利用 301 和 302 响应码通知浏览器重定向到指定的错误信息页面。即便 HOST 字段被忽略，NotBlocked.com 上恰巧包含 test.html 的几率也较小。在 Telex 代理服务暂时不可用并且客户端软件对此没有觉察的情况下，用户会反复多次重试，或尝试访问其他的屏蔽网站。因此，在一个较短的时间段内，Telex 客户端很有可能会频繁触发来自 NotBlocked.com 的上述 3 类错误响应，而普通客户端不会。在这里可以利用该特征区分 Telex 客户端和普通客户端。

虽然监听 TLS 加密信道的监管者无法获得 NotBlocked.com 返回的错误响应分组的具体内容，但它可以提取响应分组的长度。对于一个特定网站，针对上述 3 类错误其响应分组的长度规律性很强。例如 Google 针对包含 HOST 错误的请求，所返回的 TCP 应答分组长度都是 462 byte。而发生页面缺失错误时，Google 返回的 TCP 应答分组的长度一般在 1 172 byte 左右。因此，监管者可以利用响应分组长度特征判断客户端触发了服务器端的哪类错误。在没有发生请求错误的正常情况下，专业网站返回的响应分组要长得多，因为其中一般承载了包含多媒体信息的 HTML 页面。

基于以上分析，下面给出利用 DoS 攻击检测客户端是否正在使用 Telex 代理软件的具体过程。

1) 对可能被作为掩护的网站，如 NotBlock-ed.com，用定时扫描的方法分别为它们建立错误响应特征库，即记录这些网站在发生 HOST 错误、资源缺失错误和 CONNECT 请求错误时响应数据分组的长度规律。

2) 针对可疑的客户端，监管者利用 DoS 攻击使 Telex 代理服务对其不可用，并观察掩护站点 NotBlocked.com 收到客户请求后的反应。如果发现 NotBlocked.com 多次返回上述 3 类错误响应分组，则猜测该客户端正在使用 Telex 代理软件。

为了提高准确度，还可以进行多次测试，例如以间歇方式阻断 Telex 服务器向该客户提供代理服务，并观察掩护网站的响应。

4.5　与 NotBlocked.com 合谋破坏用户隐私

对于特定的 Telex 服务器，能够充当掩护者角色的网站有很多，并且它们被不同的组织或商业公司所控制。因此，理论上很难避免其中一些网站的管理者被监管方暗中收买。

如果掩护网站 NotBlocked.com 成为监管方的合谋者，那么判断客户端是否正在使用 Telex 代理

服务将变得非常容易。监管者只需对比边界防火墙截获的关于某客户端的通信数据和NotBlocked.com实际接收到的数据是否相同就可判断是否为 Telex 客户端。如果不同，则说明该客户端发出的 HTTP 请求被转发路径中的某个路由器“重定向”到了其他主机，即该客户端正在使用 Telex 代理服务。

在使用 Telex 代理服务的过程中，掩护网站是客户端软件随机选取的，并且为了避免引起监管者的怀疑，掩护网站还应该定时更换。虽然被收买的网站只占少数，但是一旦某个恶意网站被用户选定作为掩护者，那么不仅该用户正在使用 Telex 代理服务的事实被暴露，而且还会导致更严重的隐私泄露。监管者可以利用如下过程与 NotBlocked.com 合谋并利用 DoS 攻击破坏 Telex 用户隐私。

1) 如果发现某个客户端正在利用 HTTPS 协议访问合谋网站 NotBlocked.com，那么首先利用前面所述的对比 Telex 折射路由器两端通信内容的方法检测该客户是否正在使用 Telex 代理。如果是普通客户端，则放弃对该客户的监视；否则继续进行下面的攻击过程。

2) 利用DoS攻击使Telex服务对该客户暂时不可用。此时，由于 Telex 服务器没有介入，客户端与 NotBlocked.com 之间的 TLS 连接不会中断，客户端访问屏蔽网站 Blocked.com 的 HTTP 请求都被发送给 NotBlocked.com。由于 NotBlocked.com 持有 TLS 会话密钥，所以它能够看到这些 HTTP 请求的明文。

3) NotBlocked.com 自己充当 HTTP 代理角色，将该客户端的 HTTP 请求转发给 Blocked.com，并将 Blocked.com 的响应利用 TLS 加密通道转发给客户端。

在上述攻击中，落入陷阱的 Telex 用户与多个屏蔽网站交互通信的详细过程及具体内容都被监管方的合谋者全部截获，而客户端对此毫不知情。

另外，上述漏洞也可能被一些“好奇”的掩护网站利用。它们在 Telex 代理服务暂时不可用时自己冒充 HTTP 代理，其目的是窥探用户隐私。

4.6 对 Telex 握手协议的改进

以上 2 类隐私攻击能够成功的关键是 Telex 代理服务不可用时客户端未及时得到通知。为此，给出关于 Telex 握手协议的一种改进，使 Telex 客户端在 TLS 握手结束后可以得到 Telex 服务器的确认反馈，其具体改进如下。

1) 在原协议参数设置基础上，改进协议还需要一个安全散列函数 $H_3:\{0,1\}^*\to\{0,1\}^{l_{H_3}}$，其中，$l_{H_3}$ 等于 TLS 加密通信阶段所采用的对称加密算法的密钥长度。例如，采用 DES 或三重 DES 算法时，l_{H_3} 分别等于 56 和 112 bit。如果想令系统支持多种对称加密算法，则需要额外定义多种具有不同输出长度的散列函数，并且这些函数定义都要作为公共参数传播给所有 Telex 用户。

2) 在 TLS 握手的开始阶段，客户端仍然按照 3.2 节 Telex 握手协议规定构造并发送包含隐藏标签 $\tau=g_b^s\|H_1(\alpha_b^s\|\chi)$ 的 ClientHello 消息；然后接收并处理服务器端的 ServerHello、Certificate 和 ServerKey Exchange 消息；最后将 ClientKeyExchange、Change CipherSpec 和加密的 Finished 消息发送给服务器端。

3) Telex 服务器从 ClientHello 分组中识别出有效标签 $\tau=\beta\|h$ 后，继续监听双方的密钥交换信息，但与原协议不同的是，Telex 服务器命令折射路由器阻断 NotBlocked.com 发往客户端的加密的 Finished 消息，而其他消息仍然放行。

4) Telex 服务器根据标签 $\tau=\beta\|h$ 计算原协议规定的密钥 k_1 和 k_2，另外还需计算与客户端共享的密钥 $k_3=H_3(\beta^r\|\chi)$。注意 k_3 与 k_2 长度相同。等待 NotBlocked.com 发送完 ChangeCipherSpec 消息后，Telex 服务器计算之前双方交互的所有消息的散列值，并将该散列值用 k_3 加密，即计算利用 k_3 加密的 Finished 消息。最后，Telex 服务器将修改过的 Finished 消息分组发送给客户端，并要求折射路由器中断与 NotBlocked.com 的 TCP 连接。

5) 客户端收到 Finished 消息后，尝试用 k_3 对 Finished 消息进行解密。如果解密成功，则说明 Telex 服务器可以为该客户端提供正常的代理服务了，客户端继续访问被屏蔽站点；否则，说明 Telex 代理服务不可用，客户端终止通信过程。

6) 利用改进协议握手成功后，客户端和 Telex 服务器之间的 TLS 通信也利用 k_3 进行加解密(而不是利用 NotBlocked.com 也知道的密钥 k_2)，这样“好奇的”NotBlocked.com 站点也无法获得客户端与 Telex 服务器之间的通信内容。

改进 Telex 握手协议分析。首先分析安全性。与原协议相比，改进协议只是在加解密 Finished 消息和应用层消息时使用了新的对称密钥 k_3，其他步骤未做改变。只有 Telex 客户端和 Telex 代理 2 个参与者知道密钥 k_3，因此利用 k_3 加密的 Finished 消息

还具有认证功能。Finished 消息解密失败表明 Telex 代理并没有作为“中间人”介入通信，Telex 客户将及时中断通信过程，这样就避免了 4.4 节和 4.5 节的隐私泄露问题。另外，新密钥 k_3 只有 Telex 客户端和 Telex 代理 2 个参与方知道，而原密钥 k_2 有 3 个参与方知道，因此改进协议缩小了信息泄露范围，提高了安全性。

从效率角度看，改进握手协议只是使通信双方增加了一次散列运算；在应用数据通信阶段，仍然采用高效的对称加密算法，只是更换了密钥，因此对系统整体性能影响很小。

5　TCP 分组旁路攻击

TCP 分组旁路攻击的基本思想是通过干预使客户端发送的部分 TCP 数据分组绕过折射路由器直达掩护站点 NotBlocked.com。对于使用 Telex 代理的客户端而言，该旁路干扰使实际的数据接收者 Blocked.com 服务器无法收到完整的客户端 HTTP 请求。因此 Blocked.com 会利用下行的 ACK 信号通知客户端重发丢失的上行 TCP 数据分组。而对于真正访问 NotBlocked.com 的普通客户端而言，这种旁路干扰不会影响两者之间的通信过程。网络监管者可以通过检测上行路径被旁路干扰之后某些 TCP 数据分组是否被重传来判断客户端是否正在使用 Telex 代理，进而破坏用户的隐私。如果使同一 Telex 客户端的批量 TCP 分组绕过折射路由器，那么将会中断通信过程，因此旁路攻击还可以破坏 Telex 系统的可用性。

下面提出 2 种能够使 TCP 分组绕过折射路由器的方法：非对称路由法和 IP 隧道法，并分别设计了对应的 2 类旁路攻击。

5.1　基于非对称路由的旁路攻击

在客户端与 NotBlocked.com 之间制造非对称路由，使从服务器到客户端的下行路径经过折射路由器，而上行路径绕过折射路由器。基于非对称路由的旁路攻击过程如图 7 所示。实际的服务提供者 Blocked.com 由于发现某些客户端 TCP 分组丢失，将不断地利用下行 ACK 信号通知客户端重发丢失的 TCP 分组。另一种可能是客户端由于收不到 ACK，造成被旁路的 TCP 分组所对应的定时器超时，这些 TCP 分组将被客户端自动重发。监管者的边界防火墙很容易检测到某些上行 TCP 数据分组被不断重传的异常情况，因此可以准确判断客户端是否在使用 Telex 代理。

能够成功实施上述旁路攻击必须满足 2 个条件。

条件 1 是在物理上从客户端到 NotBlocked.com 服务器之间存在 2 条以上的路径，并且至少有一条路径不经过折射路由器。对于国家级监管者而言该条件很容易满足，所控制的网络区域对外一般有多个出口，连接了不同的境外 ISP 的边界路由器。关于目前互联网拓扑结构对 Telex 等反监管系统的影响，文献[8]已做了深入研究，根据其结论，针对规

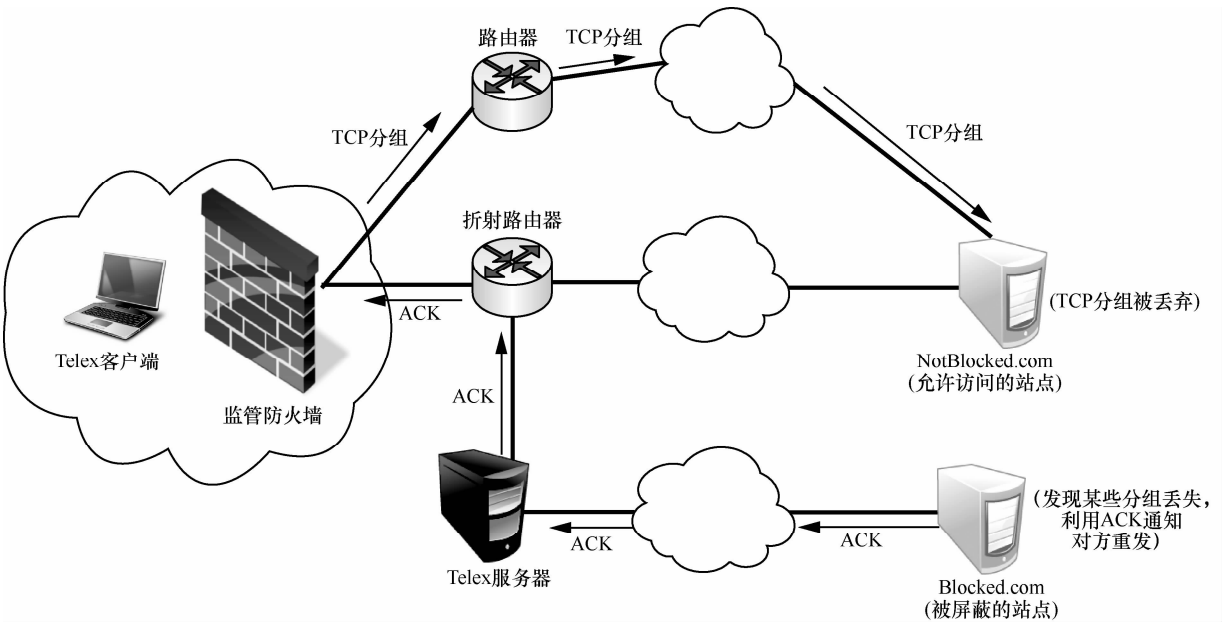


图 7　基于非对称路由的 ACK 旁路攻击

模较大的国家级监管者，要在各个可能的路径上都部署 Telex，反监管者将付出非常高的代价。

条件 2 是监管者能够实时地调整 IP 分组的路由策略。为了检测使用 Telex 的非正常客户端，监管者强制产生非对称路由的时机必须发生在客户端和 Telex 服务器握手成功之后并且与 Blocked.com 通信期间。如果采用固定的非对称路由，隐藏标签无法被 Telex 服务器识别，客户端无法使用 Telex 代理服务，及时中断了通信过程，也不会暴露自己。

5.2 基于 IP 隧道的旁路攻击

对于网络出口较少甚至只有单一出口的监管者，如一些公司、组织或规模较小的国家、地区，无法实现基于非对称路由的旁路攻击。可以采用如图 8 所示的 IP 隧道和 IP 地址欺骗技术人为制造通信旁路。实施该攻击的具体过程如下。

1) 监管者在域外部署一台完全受自己控制的间谍机。间谍机部署位置必须满足条件：间谍机通向掩护站点 NotBlocked.com 的路径不经过折射路由器。关于间谍机的部署问题在 5.3 节还会做详细讨论。

2) 在监管防火墙和间谍机之间建立并保持一个 TCP 连接作为 IP 隧道，即利用 TCP 的点对点传输服务将一些原始 IP 数据分组由防火墙直接发送给间谍机。该 IP 隧道是否经过折射路由器对攻击是否成功没有影响。折射路由器不会关心隧道中原始 IP 数据分组的内容，这是因为路由器不会关心 TCP 数据分组中承载的“应用层”数据。如果折射路由器采取相应的过滤措施，防火墙和间谍机之间则可以采用加密通道。从 IP 隧道穿梭的 IP 分组逃过了折射路由器的监听，因此它们不会被折射路由器重定向到 Telex 服务器。

3) 为了检测客户端是否正在使用 Telex 代理，等待客户端和服务器 NotBlocked.com 之间完成 TLS 握手并进入到应用数据通信阶段，监管防火墙将客户端发出的上行 IP 分组(包含了 TCP 载荷)通过 IP 隧道传送给间谍机。间谍机将这些原始 IP 数据分组提取出来，然后直接发送给服务器 NotBlocked.com。在这里，间谍机相当于实施了 IP 地址欺骗，因为该主机发出的这些 IP 分组的源地址都等于客户端的 IP 地址，而不等于其自身的 IP 地址。

4) 对于使用 Telex 代理的客户端，由于实质上并没有与 NotBlocked.com 建立 TCP 连接，由间谍机转发的 TCP 分组必然被 NotBlocked.com 丢弃，而真正参与 HTTP 通信的服务器 Blocked.com 由于发现部分客户端 TCP 分组丢失，将利用下行 ACK 信号通知客户端重发丢失的 TCP 分组。另一种可能是客户端由于收不到 ACK 自动重发这些被旁路的 TCP 分组。对于普通客户端，这种旁路干扰不会破坏通信双方的 TCP 传输状态，因而不会中断 TCP 连接，至多只会造成网络传输延时抖动。

5) 为了减小对间谍机的压力和对正常通信的影响，监管防火墙只需从一个 TLS 连接的上行数据中截取部分 IP 分组由间谍机转发。对于 Telex 用户，该转发会立即引起对应 TCP 分组的重传；而对于普通用户，只会引起轻微的传输抖动。监管防火墙很容易区

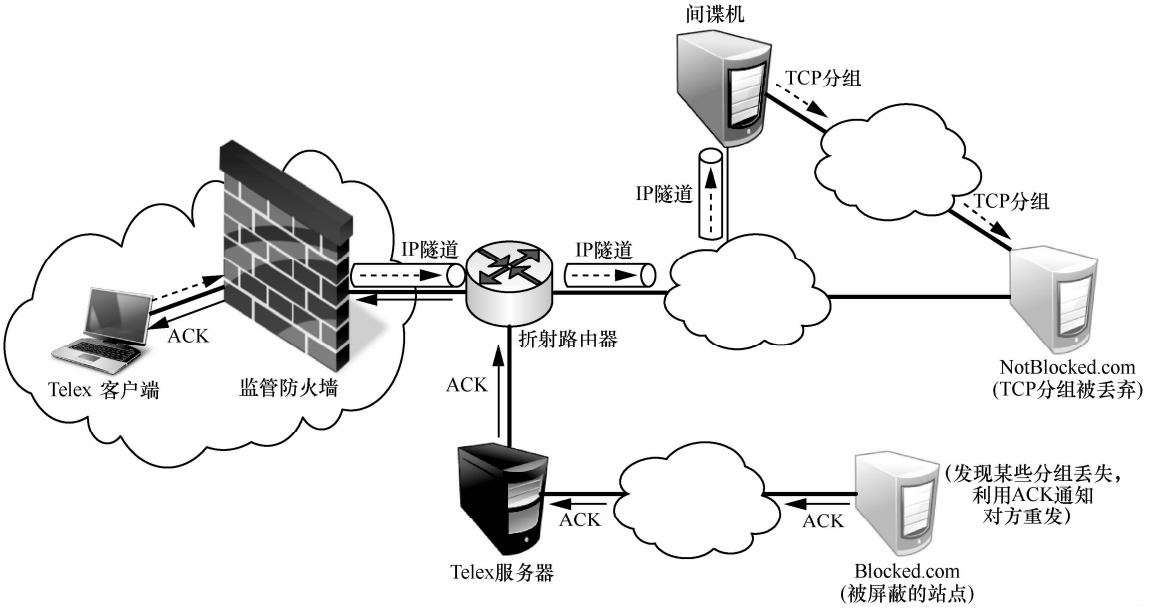


图 8 基于 IP 隧道的 ACK 旁路攻击

分这 2 种情况。为了提高准确度，防火墙还可以对上行数据反复抽样转发，并观察数据被重传的情况。

间谍机部署问题　在部署间谍机时要求该主机通向掩护站点 NotBlocked.com 的路径不能经过折射路由器。为了增加监管者直接屏蔽折射路由器的代价，Telex 必须选取大的子网或者网络自治域的边界路由器作为折射路由器。子网规模越大，网内可利用的 IP 地址也越多，其边界出口也越多。再考虑到目前互联网复杂的拓扑结构，可以说找到符合条件的间谍机部署位置是比较容易的。

IP 地址欺骗　部署间谍机时要考虑的另一个关键问题是该主机是否能成功实施 IP 地址欺骗。如果间谍机发出的带有欺骗性源地址的 IP 分组被某些路由器过滤无法到达 NotBlocked.com，那么该旁路攻击也无法奏效。由于缺乏动力，对抗 IP 地址欺骗最有效的路由器入口过滤(ingress filtering)和出口过滤(egress filtering)并没有被全世界的网络管理者广泛实施。根据文献[15]的研究和 Spoofer 计划[16]的实测，全球约有 25%的自治域和 16%的 IP 地址可以被利用发起 IP 地址欺骗，其中一些自治域内的主机可以伪造任意的 IP 地址。

另外，由于间谍机伪造的是不同自治域的 IP 地址，一般只会受到所在自治域的出口过滤干扰。因此，监管者也可以收买一些自治域网络管理者，令其对间谍机的 IP 分组额外放行。

5.3　客户类型检测方法

在旁路攻击中，通过检测被旁路的 TCP 分组和被重传的 TCP 分组之间是否具有相关性就能够准确区分普通客户端和 Telex 客户端。为此为每一个 TLS 连接定义 2 个函数，旁路函数 $a(n)$和重传函数 $b(n)$，其中，n 表示 TCP 字节流中的字节编号。

1) 如果 TCP 流的第 n byte 被旁路，则令 $a(n)=1$；否则令 $a(n)=0$。

2) 如果检测到 TCP 流的第 n byte 被重传，则令 $b(n)=1$；否则令 $b(n)=0$。

截获一个 TCP 分组之后，监管防火墙可以通过 TCP 头部的起始序号 n_1 和 IP 头部的长度字段获得该 TCP 分组所包含字节数组的编号区间$[n_1,n_2]$。假定该 TCP 分组被旁路，则可知：对于任意 $n\in[n_1,n_2]$，$a(n)=1$。如果检测到该 TCP 分组被重传，则对于任意 $n\in[n_1,n_2]$，$b(n)=1$。

基于 $a(n)$和 $b(n)$函数，定义相关系数 R 为

$$R=\frac{\sum_{n=N_1}^{N_2}a(n)b(n)}{\sum_{n=N_1}^{N_2}a(n)}$$

其中，N_1 表示该 TCP 字节流的初始字节编号，N_2 表示目前截获的关于该 TCP 字节流的最大字节编号。显然 $0\leqslant R\leqslant 1$，R 越大客户端为 Telex 用户的可能性越大。

下面基于 R 设计判断客户端为普通用户还是 Telex 用户的检测方法。

1) 在 TLS 握手完成之后，监管防火墙按照一定规则从客户端上行流量中选取 K 个 TCP 数据分组，然后利用非对称路由或 IP 隧道将它们转发给间谍机，再由间谍机发送给 NotBlocked.com。监管者同时记录对应的 $a(n)$函数。选取被旁路的 TCP 分组时需要避开不携带任何应用层数据的 ACK 分组（一般仅包含 20 byte 的 TCP 头且 ACK 标识位有效）。

2) 检测上行 TCP 数据分组被重传的情况，并记录 $b(n)$函数。

3) 在 K 个 TCP 数据分组全部被旁路之后计算相关系数 R。如果 $R=1$，则判断客户端类型为 Telex 客户端；否则判断为普通客户端。

由于 TCP 协议采用了基于滑动窗口的流量控制，而且 ACK 分组对发送者的确认具有累积效应，所以只旁路小部分不携带应用层数据的 ACK 分组未必能引起对应数据分组的重传。因此在第 1 步中不选择旁路“纯粹”的 ACK 分组。

当上述检测方法获得的相关系数 $R<1$ 时，可以 100%断定客户端为普通用户，即不存在假阴性(false negative)情况。这是因为如果是 Telex 客户端，则被旁路的 TCP 分组一定会被重传，除非客户端意外终止了 TLS 连接（这里忽略这种小概率事件），因此一定满足 $R=1$。

当获得的相关系数 $R=1$ 时，则存在假阳性(false positive)的可能性。对于非 Telex 用户，如果被旁路的 TCP 分组在发往 NotBlocked.com 的途中意外丢失、也会造成 TCP 分组的重传。具体地讲，发生假阳性的概率为$(P_{\text{Loss}}+P_{\text{ACK}})^K$。其中，$K$ 表示被旁路的 TCP 分组数目；P_{Loss} 表示从间谍机到 NotBlocked.com 之间网络路径的分组丢失概率；P_{ACK} 表示从 NotBlocked.com 到客户端的下行 ACK 信号连续丢失使对应的上行 TCP 数据分组被重传的概率。

根据文献[17]的数据，Internet 上大部分测量点之间网络路径的分组丢失率接近 0%，小部分为 1%，极少的网络路径分组丢失率超过 5%。当 P_{Loss}≤5%时，P_{ACK} 几乎可以忽略不计。为了减轻间谍机的压力，K 越小越好。如果令 K=1，则假阳性概率≤5%；如果令 K=2，则假阳性概率≤0.25%。因此，监管者只需选取 2 个 TCP 分组进行旁路测试，就可以获得超过 99.75%检测成功率。

5.4 攻击可行性实验

为了验证基于 IP 隧道旁路攻击的可行性，搭建 Telex 原型系统和攻击测试环境如图 9 所示。该系统中，客户端、监管者、折射路由器（包含了 Telex 代理服务器）、间谍机、掩护站点 NotBlocked.com 和被屏蔽站点 Blocked.com，分别用 6 台装有 Linux 的物理主机代替，其他不重要的路由器用 Linux 虚拟机代替。

1) 原型系统的实现方法

客户端部署了 Telex 原作者 Wustrow 等设计的 Telex 客户端本地代理软件，HTTP 客户端采用了专门设计的测试工具 HTTPClient。HTTPClient 支持通过本地 HTTP 代理（127.0.0.1）访问 Web 网站。

监管者防火墙的正常网络路由功能利用 Linux 的 iptables 模块实现，防火墙功能通过在 Linux 的 NetFilter 框架[13]中插入自己设计的过滤函数实现。监管者和间谍机之间建立普通的 TCP socket 连接作为 IP 隧道。

折射路由器和 Telex 代理服务器被部署在同一台物理主机上。正常路由功能利用 iptables 实现，Telex 的 HTTP 代理功能利用开源软件 Squid[14]实现。通过在 NetFilter 框架中插入过滤函数实现抓取 TLS 的 ClientHello 握手分组，并利用 Wustrow 等设计的子程序识别其中是否含有合法的隐藏标签。标签识别成功后，通过在 Linux 内核中直接创建相关数据结构和状态信息建立客户端和 Telex 代理服务器(Squid)之间的 Socket 连接。基于该 Socket 连接，再利用 OpenSSL 构建 TLS 连接。TLS 对称密钥生成方法也做了相应修改。

间谍机从 IP 隧道中提取出原始 IP 分组后，不加修改地直接利用 Libnet 库发送给 NotBlocked.com。

NotBlocked.com 和 Blocked.com 对应的主机上都部署了 Web 服务器软件 nginx。

另外用 Linux 的 tc(traffic control)命令[18]模拟产生折射路由器到 Blocked.com 之间路径的分组丢失率 Q_{Loss}、从间谍机到 NotBlocked.com 之间网络路径的分组丢失率 P_{Loss}。

2) $a(n)$和 $b(n)$相关性实验

为了测试客户类型对 $a(n)$和 $b(n)$函数之间相关性的影响，在第 1 次实验中令 HTTPClient 作为 Telex 客户通过 Telex 本地代理软件访问 Blocked.com。HTTPClient 随机生成长度为 1~100 的 HTTP 请求，Blocked.com 随机生成长度为 100~10 000 的 HTTP 响应。这样做是为了模拟交互式 Web 访问和 Web 服务中上下行流量不对称的特征。为了加强实验结果的可演示性，令 Q_{Loss} 和 P_{Loss} 都等于 10%。

通信开始后，监管者每接收 10 个 TCP 分组就用旁路方式向间谍机发送 1 个 TCP 分组。监管者实时记录 $a(n)$和 $b(n)$函数，得到如图 10 所示的结果。

图 10 的上半部分为旁路函数 $a(n)$，下半部分为重传函数 $b(n)$，横轴都表示字节序号 n。为了便于分析，TCP 字节序号做了平移，使其编号从 0 开始。图中每个方柱都对应 1 个被旁路或重发的 TCP 分组，方柱宽度对应 TCP 分组长度。

从图 10 可以看出，每个被旁路的 TCP 分组都被重传，此时相关系数 R=1，和客户端为 Telex 用

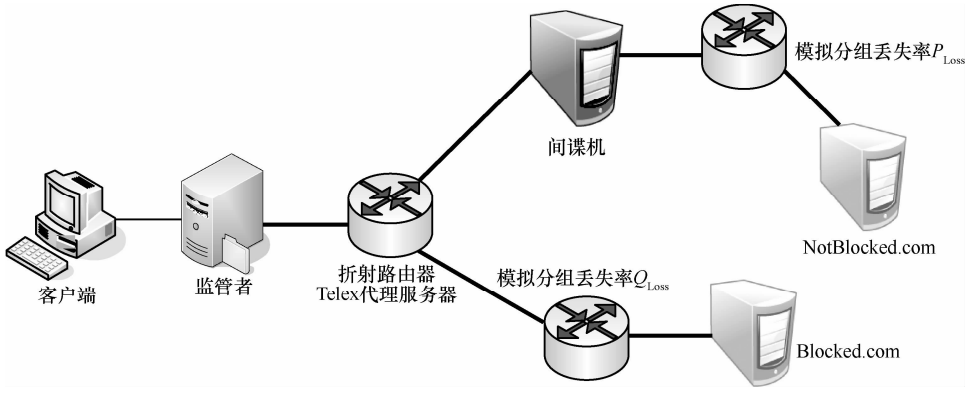


图 9 原型系统

户的情况相对应。其他没有被旁路但也被重传的 TCP 分组是网络传输分组丢失率 Q_{Loss} 造成的。

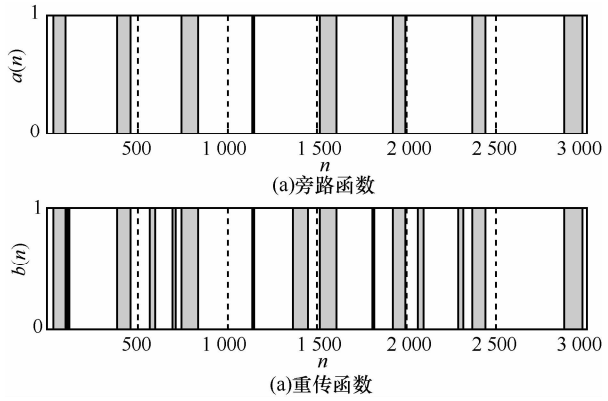


图 10　Telex 用户通信时的 $a(n)$和 $b(n)$

在第 2 次实验中，令 HTTPClient 模拟普通客户端，即令 HTTPClient 不通过代理直接访问 NotBlocked.com。HTTP 请求和响应仍然是随机生成的。其他参数与第 1 次实验相同。监管者实时记录 $a(n)$和 $b(n)$函数，得到图 11 所示的结果。

由图 11 可以看出，对于普通客户端的通信，8 个被旁路的 TCP 分组中只有 1 个被重传（图中标黑色的方块），而其他 7 个都没有检测到重传，因此 $a(n)$和 $b(n)$的相关性很小。既被旁路也被重传的 TCP 分组是网络传输分组丢失率 P_{Loss} 造成的。

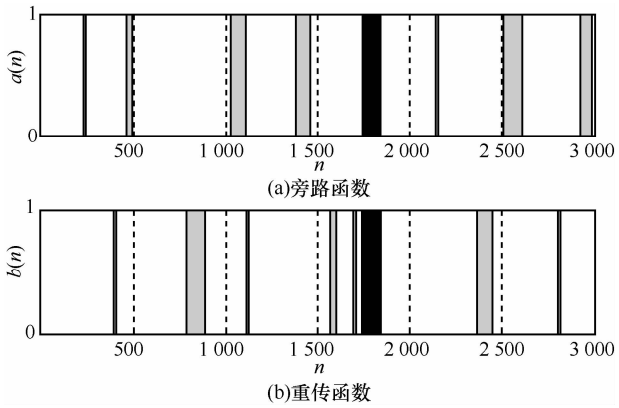


图 11　普通用户通信时的 $a(n)$和 $b(n)$

3) 客户类型检测实验

为验证 5.3 节基于相关系数 R 的客户端类型检测方法，进行了如下实验。

实验中采用了 3 组 Web 通信数据集：Berkeley HomeIP Web 流量记录[19]、WITS(waikato Internet traffic storage)流量记录[20]和随机生成的流量记录。HomeIP 数据集和 WITS 数据集主要保存了真实互联网通信流量中 HTTP 头、TCP 头和 IP 头信息，不包含完整的通信内容。客户端 HTTPClient 根据数据集中记录的 HTTP 请求长度生成相同长度的 HTTP 请求，服务器端根据数据集中记录的对应的 HTTP 响应长度生成相同长度的 HTTP 响应。Random 数据集中 HTTP 请求和响应都是随机生成的。在重复实验时，每次都从数据集中截取不同时间段的数据进行模拟。

影响检测成功率的分组丢失率 P_{Loss} 参数分别取 0%、1%和 5%这 3 种值，并且都用 Linux 的 tc 命令模拟产生。令被旁路的 TCP 分组数目 K=2。

针对上述的每种参数组合，HTTPClient 都分别扮演 Telex 客户端和普通客户端 2 种角色与服务器端进行通信。监管主机利用旁路攻击法检测 $a(n)$和 $b(n)$并计算相关系数 R。针对每种参数组合和客户端角色配置，都重复实验 100 次然后计算平均相关系数 R。实验结果如表 2 所示。

根据表 2，当客户端为 Telex 用户时，R 都等于 1；当客户端为 Telex 用户时，R 都远远小于 1。因此，基于相关系数 R 判定客户端类型是非常准确的。在上述 1 800 次重复实验中，只出现了 2 次假阳性的情况，即 R=1 但客户端为普通用户的情况。假阳性概率与理论分析值（小于等于 P_{Loss}^{K}）相符。

另外，在 5.3 节的客户端类型检测方法中，监管者只需选取 2 个非 ACK 类型的 TCP 分组通过 IP 隧道转发给间谍机，并实时记录相应时间段内被重传的 TCP 分组的起始序号和分组长度，因此该检测方法的代价是非常低的。

表 2　　**不同参数组合下的平均相关系数 R**

数据集	P_{Loss}=0%		P_{Loss}=1%		P_{Loss}=5%	
	Telex 客户	普通客户	Telex 客户	普通客户	Telex 客户	普通客户
HomeIP	1	0	1	0.009	1	0.054
WAND	1	0	1	0.013	1	0.063
Random	1	0	1	0.011	1	0.052

6 结束语

首先研究了 Telex 反监管系统的安全性，分析了对其实施 DoS 攻击的可行性，并给出了 2 种具体的 DoS 攻击方法：基于资源消耗的 DoS 攻击和破坏握手协议的 DoS 攻击。一个重要发现是利用 DoS 攻击不仅可以破坏 Telex 的可用性，还可以破坏用户隐私性。这类攻击源于 Telex 握手协议的一个安全漏洞，即客户端利用隐藏标签发出认证请求之后，Telex 服务器并没有确认认证结果。为此给出了能够进行双向确认的改进的 Telex 握手协议。

还提出了另一种利用主动攻击破坏用户隐私的方法——TCP 分组旁路攻击法，并通过原型实验验证了其可行性。构建 TCP 分组旁路的方法包括利用非对称路由和构建 IP 隧道 2 种，分别适用于大、小规模的网络自治域监管者。旁路攻击方法检测准确，对普通客户端影响很小，并且同样适用于其他基于路由器重定向技术的反监管系统，如 Cirripede[2]和 Decoy Routing[3]。要想对旁路攻击完全免疫，反监管系统必须部署大量的折射路由器以覆盖通向掩护站点的各个路径，还要提高全互联网对抗 IP 地址欺骗攻击的能力，这在短期内都难以实现。即便实现也将付出很大的代价，并且如何协调这些广泛分布的折射路由器也将是一个难题（如上下行数据分别通过不同折射路由器时它们如何相互同步认证状态）。

参考文献：

[1] ERIC W, SCOTT W, IAN G, *et al*. Telex: anticensorship in the network infrastructure[A]. Proceedings of the 20th USENIX Security Symposium[C]. San Francisco, USA, 2011.

[2] AMIR H, GIANG T K N, CAESAR M, NIKITA B. Cirripede: circumvention infrastructure using router redirection with plausible deniability[A]. Proceedings of the 18th ACM conference on Computer and Communications Security (CCS 2011)[C]. Chicago, IL, USA, 2011.187-200.

[3] JOSH K, DANIEL E, ALDEN W, *et al*. Decoy routing: toward unblockable Internet communication[A]. Proceedings of the USENIX Workshop on Free and Open Communications on the Internet (FOCI 2011)[C]. 2011.

[4] RICHARD C, STEVEN J M, ROBERT N M W. Ignoring the great firewall of China[A]. Proceedings of the Sixth Workshop on Privacy Enhancing Technologies (PET 2006)[C]. Cambridge, UK, 2006. 20-35.

[5] TARIQ E, IAN G. CORDON — A Taxonomy of Internet Censorship Resistance Strategies[R]. CACR Tech Report 2012-33, 2012.

[6] PHILIPP W, STEFAN L. How the great firewall of China is blocking[A]. Proceedings of the USENIX Workshop on Free and Open Communications on the Internet (FOCI 2012)[C]. 2012.

[7] QIYAN W, XUN G, GIANG T K, *et al*. CensorSpoofer: asymmetric communication using IP spoofing for censorship-resistant web browsing[A]. Proceedings of the 19th ACM conference on Computer and Communications Security (CCS 2012)[C]. 2012.

[8] MAX S, JOHN G, CHRISTOPHER T, *et al*. Routing around decoys[A]. Proceedings of the 19th ACM Conference on Computer and Communications Security (CCS 2012)[C]. Raleigh, USA, 2012. 85-96.

[9] HOUMANSADR A, EDMUND L W, VITALY S. No direction home: the true cost of routing around decoys[A]. Proceedings of the Network and Distributed Security Symposium (NDSS 2014)[C]. 2014.

[10] ANDREW H. Fingerprinting websites using traffic analysis[A]. Proceedings of Privacy Enhancing Technologies workshop (PET 2002)[C]. 2002.

[11] Cooperative association for Internet data analysis (CAIDA)[EB/OL]. http://www.caida.org/data/.

[12] Telex[EB/OL].https://telex.cc.

[13] Netfilter/Iptables[EB/OL]. http://www.netfilter.org.

[14] Squid[EB/OL]. http://www.squid-cache.org/.

[15] TOBY E, LI J. On the state of IP spoofing defense[J]. ACM Transactions on Internet Technology, 2009, 9(2):1-29.

[16] The MIT ANA Spoofer project[EB/OL]. http://spoofer.csail.mit.edu/.

[17] Internet Traffic Report[EB/OL]. http://www.internettrafficreport.com/.

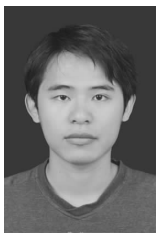
[18] Netem/TC[EB/OL]. http://www.linuxfoundation.org/collaborate/workgroups/ networking/netem.

[19] HomeIP[EB/OL]. http://ita.ee.lbl.gov/html/contrib/UCB.home-IP-HTTP.html.

[20] WITS[EB/OL].http://wand.net.nz/wits/auck/10/auckland_x.php.

作者简介：

李龙海（1976-），男，河北冀州人，博士，西安电子科技大学副教授、硕士生导师，主要研究方向为匿名通信、隐私保护技术和计算机网络安全。



黄城强（1989-），男，福建福州人，西安电子科技大学硕士生，主要研究方向为网络安全、下一代互联网和网络大数据处理技术。

王万兴（1989-），男，山东德州人，西安电子科技大学硕士生，主要研究方向为网络安全、网络协议设计与分析。

慕建君（1965-），男，陕西榆林人，西安电子科技大学教授、博士生导师，主要研究方向为计算机网络与差错控制技术、网络编码技术。