IP 动态伪装模型研究*)

何聚厚[1,2] 何秀青[2] 李 由[1] 李伟华[1]

(西北工业大学计算机科学与工程系 西安710072)[1] (陕西师范大学计算机科学学院 西安710062)[2]

摘 要 IP 伪装学有助于转移视线,迷惑攻击者并防止攻击者从 IP 地址标识和 IP 包中获取有价值的信息。本文提出了一个动态伪装模型,它可扩大网络会话通过共享和竞争这些 IP 来选择的伪装 IP 数。本文的主要贡献是:(1)通过伪装 IP 数划分网络会话数定义了计算机伪装度;(2)提出了维护伪装 IP 队列的算法,通过发送 APR 消息,改变其状态和竞争伪装 IP;使用监测数据包的 DiverSocket,Netfilter 和 IPTables,我们设计了一个基于此模型的实验 IPDM,获得了满意结果。

关键词 网络安全,IP 伪装,算法

Study of IP Dynamic Masquerading Model

HE Ju-Hou[1,2] HE Xiu-Qing[2] LI You[1] LI Wei-Hua[1]

(Department of Computer science and Engineering, Northwestern Polytechnical Uniersity, Xi'an 710072)[1]

(College of Computer Science, Shanxi Normal University, Xi'an 710062)[2]

Abstract The discipline of IP masquerading can help to puzzle the attacker's attention and prevent the attacker from getting the valuable information from the IP packets which are identified by IP address. This paper presents an IP dynamic masquerading model to extend the masqueraded IP number that the network session can chose by sharing and competing these IPs. Our main contribution is in presenting: (1)the definition of masqueraded degree for a computer in sometime by the number of masqueraded IP divide the number of the network session; (2)the algorithm for maintaining the queue of masqueraded IP, changing their states and competing the masqueraded IP by sending ARP messages; Using DivertSocket to monitor the packets, and basing Netfilter and IPTables, we design an experimental system IPDM based on this model, and satisfactory results are obtained.

Keywords Network security, IP Masquerading, Algorithm

1. 引言

目前解决网络安全问题的主要技术手段有加密技术,基于网络、主机的各种防范措施(如防火墙、入侵检测等),以及信息伪装技术等。这些技术在防御网络入侵、保护信息安全方面均有一定的作用。但不能很好地保护、隐藏网络特征信息[1],如 IP 数据流的出处等,这是因为网络中数据帧、IP 数据报的报头信息在网络中传输时是透明的,攻击者往往通过"嗅探"网络中的数据流来确定攻击目标的 IP 地址与端口号,进而扫描目标计算机,或对"嗅探"到的数据包根据 IP 地址进行分类组装,对加密数据进行破解,或对相同 IP 的数据包的相关性进行分析,从中获取有助于攻击、窥探目标计算机有用信息的目的。Linux 系统中的 IP 伪装[2](IP Masquerading)通过把基于本机 IP 的通信伪装成多个 IP 进行,这样单机的通信在攻击者看来是多个计算机在进行通信,从而在一定程度上可以达到迷惑攻击者、保护本机安全通信的目的[2,4]。但是,在后面的讨论中可以看出,Linux 系统的 IP 伪装(伪装 IP 的个数是静态的)具有一定的局限性。本文提出的基于竞争伪装 IP 地址的 IP 动态伪装模型,充分利用了网络中数据传输的突发性和随机性的特点,使多个计算机共享同一伪装 IP 池,扩大了 IP 伪装的范围,从而增加了攻击者的攻击难度,起到了主动防御的作用。

2. IP 动态伪装模型

2.1 伪装度

在本文中,一次网络会话指通信双方从开始通信到结束通信的过程。为了叙述方便,假定每个计算机有 K 个网络会话。每个会话所使用的 IP 可以从 M_P 个伪装 IP 中选取,若 $M_P \geqslant K$,则 K 个会话可以使用 K 个伪装 IP 进行;若 $M_P < K$,则 K 个会话可以使用 M_P 个伪装 IP,部分会话使用同一伪装 IP。由于伪装 IP 选取的随机性,无论哪一种情形,在一段时间后,攻击者将会把一台计算机的通信看成 M_P 台计算机的通信。显然,$M_P \geqslant K$ 比 $M_P < K$ 具有更好的伪装效果。但在实际应用中,由于网络数据传输的突发性和随机性,K 的取值不确定。因此为了表示伪装效果的好坏,本文使用伪装度,即每个会话所拥有的伪装 IP 数来衡量。

定义1 伪装度=主机所拥有的伪装 IP 数/网络会话个数

在 Linux 系统的 IP 伪装技术中,如果把局域网内的 M 个伪装 IP 平均分给 N 个计算机,则每个计算机的伪装 IP 数

*)基金项目:国家863高科技研究发展计划资助项目(2001AA142100);国家教育部博士学科点基金资助项目(20010699018)。**何聚厚** 博士生,讲师,主要研究领域为计算机网络安全、多媒体通讯技术。**何秀青** 硕士,讲师,主要研究领域为 Internet 与管理信息系统。**李 由** 主要研究领域为 Internet 及其应用。**李伟华** 教授,博士生导师,主要研究领域为计算机网络安全、多媒体通信技术、智能决策支持系统等。

$M_P=M/N$，每个计算机的伪装度为 $C_P=M_P/K$。由于 M_P 是常数，如果 K 增大，则伪装度 C_P 会降低，从而影响伪装效果。为了达到比较恒定的伪装度，在本文提出的模型中，通过共享伪装 IP 池，使 M_P 能够随着 K 的增大而增大。

2.2 网络会话标识

在 IP 动态伪装中，即使目标主机是同一台计算机，不同的网络会话也选择不同的伪装 IP，因此，需要唯一地标识每次网络会话。

定义2 一次网络会话使用一个四元组来标识：$S_E=(P_r, S_P, D_A, D_P)$，其中 P_r 为网络会话所使用的协议，S_P 为本机的端口号，D_A 为目标计算机的 IP 地址，D_P 为目标计算机的端口号。

由 TCP/IP 协议的定义可知，对于通过 TCP、UDP 协议进行的网络会话，通信双方使用的协议是一致的；在本机方，其端口号是唯一的；针对不同的目标主机 IP 及其端口号，本机选择不同的伪装 IP。因此上述四元组可以唯一的标识一次网络会话及其使用的伪装 IP。对于其它协议，如 Requst/Echo ICMP，不需要关心端口号，只需要目标主机的 IP 地址即可标识该网络会话。

2.3 伪装 IP 描述及其状态变化

通过伪装 IP 的状态变化，实现对伪装 IP 的竞争及共享。

2.3.1 *伪装 IP 描述* 为了便于对伪装 IP 进行动态调度管理，每个伪装 IP 使用下面四元组进行描述。

定义3 伪装 IP 描述为：$P_M=\{P,S,U,N\}$，其中 P 为伪装 IP 地址。S 为该伪装 IP 的当前状态，$S=\{Used/Ready|Dead\}$，对于当前主机，若该伪装 IP 正应用于某一网络会话，则该伪装 IP 的状态为 *Used*；若该伪装 IP 可以应用于下一次网络会话，则该伪装 IP 的状态为 *Ready*；若该伪装 IP 正应用于其它计算机、或处于空闲状态，则该伪装 IP 的状态为 *Dead*。N 为应用于该伪装 IP 的网络会话的个数，对于 $M_P<K$ 的情况，部分网络会话使用同一伪装 IP，相应的 N 值则会大于1。U 用于描述竞争该伪装 IP 的历史，是伪装 IP 状态转换的一个条件，U 定义为：

$$U_{new}=\alpha\times U_{old}+(1-\alpha)\times f$$

其中 $\alpha(0<\alpha<1)$ 为地址状态变化的速度，f 为每一伪装 IP 状态更新周期的采样值，如果有其他计算机竞争该伪装 IP，则 f 的采样值为1，否则为零。缺省情况 U 的初始值为0。很明显，如果该伪装 IP 长时间没有被竞争，则 U 其值会逐渐减小。

U 主要用来控制伪装 IP 从 *Ready* 到 *Dead* 状态的变化。如果 $U\geqslant U_{Max}$（U_{Max} 为伪装 IP 状态变化的阈值），该伪装 IP 的状态就可以转换了，U 值随着对伪装 IP 的竞争而动态变化。U_{Max} 的选取对伪装 IP 竞争的公平性至关重要，如果 $U_{Max}\to 1$，则本机不放弃该伪装 IP（状态保持为 *Ready*），而竞争该伪装 IP 的主机经过长时间的竞争也得不到该伪装 IP；如果 $U_{Max}\to 0$，则本机检测到其它主机竞争该 IP 后，立即放弃该 IP 的控制，这样有可能会使伪装 IP 的状态频繁变化，出现伪装 IP 状态变化的波动，增加主机和网络的负担。在我们的实验模型中，$U_{Max}=1-N/M$，M 为局域网内伪装 IP 的个数，N 为主机的个数，这样本机在主机个数比较多时比主机个数比较少时更容易放弃对某一伪装 IP 的控制，从而实现公平伪装的目的。

2.3.2 *伪装地址表* 每个计算机除自己固定的 IP 地址外，还需要维护针对伪装 IP 三个状态的地址表 $L_U(N_U)$，$L_R(N_R)$，$L_D(N_D)$。N_U，N_R，N_D 分别为不同状态地址表中伪装 IP 的个数。对于 $L_R(N_R)$，如果一些计算机的 N_R 取值太大，则共享的 *Dead* 态的伪装 IP 减少，这样会增加竞争伪装 IP 的次数，但如果 $N_R=1$，则在新的网络会话产生频率高时，可能会由于 $L_R(N_R)$ 为空而影响伪装的效果，一般至少 $N_R\geqslant 2$。$L_R(N_R)$ 和 $L_D(N_D)$ 中伪装 IP 的 U 值的变化规则是：如果其它计算机竞争某些伪装 IP，则这些 IP 的采样值 f 为1，其它 IP 的采样值为0，从而更新 $L_R(N_R)$ 和 $L_D(N_D)$ 中所有伪装 IP 的 U 值。伪装 IP 从不同地址表中的迁移过程，也就是其状态的变化过程。三种状态的变化过程如图1所示。

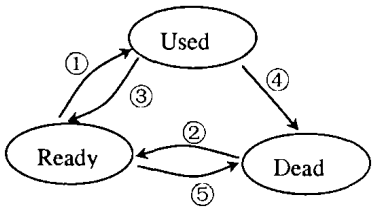

图1 伪装 IP 的状态变化过程

2.3.3 *竞争伪装 IP* 计算机 i 竞争某一伪装 IP 时，连续广播 ρ 个针对该伪装 IP 的 ARP 消息，如果在 T_i 时间内没有任何计算机响应，则表明其它计算机没有使用该伪装 IP，计算机 i 竞争到该 IP；如果计算机 j 正在使用该伪装 IP（Ready 态或 Used 态），则计算机 j 向计算机 i 回应相应的 ARP 消息，计算机 i 置该伪装 IP 的 U 值为1；如果该伪装 IP 在计算机 j 的 *Dead* 队列，则不进行任何响应，并置其 U 为1。T_i 是计算机 i 和其它计算机进行通信的往返时间的最大值与发送 ρ 个 ARP 消息所用时间之和，它可以根据每次通信的情况进行动态调整。

2.3.4 *三种状态队列的维护算法*

(1)对于 Used 队列，如果有新的会话，则从 Ready 队列中取伪装 IP 进行伪装（转换①），如果 Ready 队列为空，则使用本机的固定地址 P_0；如果某一伪装 IP 的 N 值为0，表示应用于该伪装 IP 的所有会话都结束，则该 IP 从 Used 态转换到 Dead 态（转换④）。并置其 U 值为1。U 值取最大值有利于该伪装 IP 通过竞争转移到其它计算机。

(2) Ready 队列未满时，从 Dead 队列中取 U 值最小的伪装 IP $P_{M.0}$，并竞争 $P_{M.0}$，如果 $P_{M.0}$ 可用，则 $P_{M.0}$ 的状态从 Dead 转换到 Ready 态（转换②）；否则再取 U 值最小值、次小值的伪装 IP 进行竞争，依次类推，逐渐增加竞争伪装 IP 的个数。如果竞争到一个伪装 IP，则选取该 IP，并停止竞争过程。这样逐渐增加竞争伪装 IP 个数的过程，可以在更大程度上确保主机在少量通信的情况下尽快竞争到伪装 IP。如果探测到所有的伪装 IP 都在使用，Ready 对列空且 Used 队列中有伪装 IP，则从 Used 对列中取一 N 值最小的伪装 IP 加入到 Ready 对列（转换③）。如果 Ready 对列中某一伪装 IP 的 U 值由于其它计算机的竞争达到了阈值 U_{Max}，并且该 IP 没有出现在 Used 对列，为了伪装 IP 的公平分配，则该伪装 IP 从 Ready 态转换到 Dead 态，并置 U 其值为1（转换⑤）。

(3)Used 和 Ready 中的伪装 IP 是独占的，而 Dead 队列中的是共享 IP，如果其它计算机竞争的 IP 在 Ready 队列或 Dead 对列中，则每次需更新该伪装 IP 的 U 值。

2.4 伪装网络会话队列

定义4 每个计算机都维护一个伪装网络会话队列 L_C，L_C 定义为：$L_C=\{(S_{E.i}, C_T, T)|i\in M)\}$，其中 $S_{E.i}$ 为第 i 个会话的标识；C_T 用于跟踪 $S_{E.i}$ 的会话过程，如 TCP 中建立/结束

会话的三次握手过程等，从而确定是否需要将 $S_{E,I}$ 从 L_C 中删除；T 为 $S_{E,I}$ 没有进行任何数据传输的计时器，每次检测到 $S_{E,I}$ 的一次通信，T 置零，如果 T 到达某一最大值 T_{MAX}，则从 L_C 中删除 $S_{E,I}$，该计数器对通过 UDP、ICMP 进行通信的网络会话的终止控制特别有用。

3. 实现技术

基于上述 IP 动态伪装模型，我们使用 DivertSocket、Netfilter 和 IPTables，并通过动态构造 ARP 数据包实现了一个 IP 动态伪装系统 IPDM（IP Dynamic Masquerading）。IPDM 的体系结构如图2所示。DivertSocket[3] 提供了一种便利的在用户空间控制通过协议栈数据包的途径。Netfilter[2,5] 提供了一个抽象、通用化的框架，和 IPTables 构成了 Linux 的防火墙系统。

3.1 IPDM 体系结构

Netfilter 在 IP 层的五个 HOOK 点分别是：(1)NF_IP_PRE_ROUTING：对进入网络层的数据包进行版本号、校验和检测、源地址转换等；(2)NF_IP_LOCAL_IN：送往本机的 IP 数据包通过此监测点，INPUT 包过滤在此进行；(3)NF_IP_FORWARD：转发 IP 数据包通过此检测点，FORWORD 包过滤在此进行；(4)NF_IP_POST_ROUTING：要通过网络设备出去的 IP 数据包通过此检测点，目的地址转换（包括 IP 伪装）在此点进行；(5)NF_IP_LOCAL_OUT：本机进程发出的 IP 数据包通过此检测点，OUTPUT 包过滤在此进行。

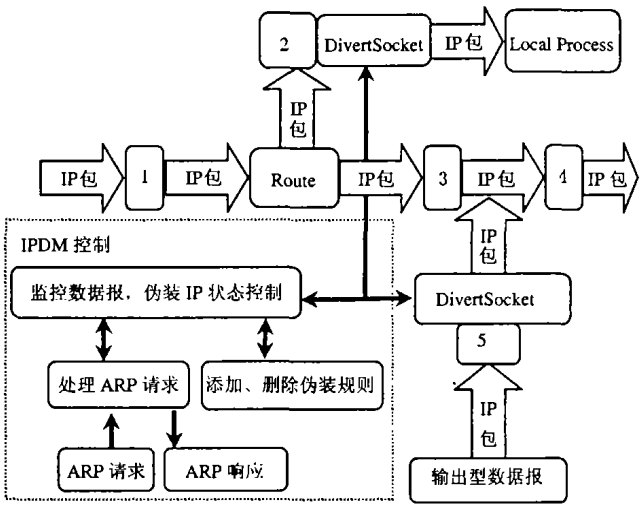

图2 IP 动态伪装系统 IPDM 体系结构

IPDM 通过 DivertSocket 在 NF_IP_LOCAL_IN、NF_IP_LOCAL_OUT 处监控 INPUT、OUTPUT 数据包，从而确定哪些网络会话是本机发起的（称为输出型会话）；哪些网络会话是其它计算机发起的（称为输入型会话）。IPDM 主要对输出型网络会话数据报进行伪装。处理 ARP 请求包括竞争伪装 IP 的 ARP 消息和正常通信时的 ARP 消息。由地址队列维护算法可知，如果本机使用了某一伪装 IP，则会响应相应的 ARP 请求。这样，无需在该计算机的 IP 地址列表中添加伪装 IP，即可进行数据通信。

3.2 伪装过程

IPDM 伪装过程如下：

(1)IPDM 控制程序对 INPUT、OUTPUT 数据包进行分析，得出协议类型 P_r、目标 IP D_A 和端口号 D_P、本地端口号 S_P，如果该数据报属于输入型会话，或该数据报对应的网络会话在伪装网络会话队列 L_C 中，则转入(3)

(2)进入伪装过程，从 Ready 队列中取一伪装 IP P_M，添加如下的防火墙规则（P_O 为本机的固定 IP 地址）：

iptables -t nat -A POSTROUTING -p P_r-d D_A-s P_O--sport S_P--dport D_P-o eth0 -j SNAT --to P_M

iptables -t nat -A PREROUTING -p P_r-d D_A-s P_M--sport D_P--dport S_P-i eth0 -j DNAT --to P_O

上述第一条规则工作在 NetFilter 的(4)处，第二条规则工作在 NetFilter 的(1)处。然后通过地址队列维护算法对地址队列进行更新，并在 L_C 中添加该次会话标识。

(3)如果该数据报对应的网络会话 $S_{E,I}$ 在 L_C 中，更新 $S_{E,I}$ 中的 C_T 和 T。如果要从 L_C 中删除 $S_{E,I}$，则需要删除(2)中添加的防火墙规则。对于输入型会话的数据报，不进行任何处理。

总结与展望 本文提出的 IP 动态伪装模型中，通过伪装度来衡量伪装的优劣，通过对伪装 IP 的竞争来实现对伪装 IP 的共享，通过伪装 IP 状态的变化和对不同网络会话的标识，以及通过 DivertSocket 对数据报进行监测、对网络会话进行跟踪、使用 ARP 消息实现探测、竞争伪装 IP，通过动态添加、删除防火墙规则，实现 IP 地址动态伪装，从而提高了伪装 IP 的利用率。在我们的实验网中，伪装 IP 共有200个，共有20台计算机，网络会话个数单台计算机最高可达50个。经过对 IPDM 系统长期运行的观察结果可以看出，IP 动态伪装模型对于突发性的、短时间的网络会话的伪装特别适用，而且伪装 IP 由于竞争，很容易从一个计算机转移到另一个计算机，伪装度最低为0.72，平均保持在0.97以上。

在后续工作中，我们将对动态 IP 伪装的响应时间、延迟、以及如何满足实时系统应用的需求等进一步进行探讨。

参 考 文 献

1 Yong G, Xinwen F, et al. NetCamo: Camouflaging Network traffic for Qos-Guaranteed Mission Critical Applications. IEEE Transactions on Systems, 2001, 31(4): 253～265

2 netfilter/iptables project, http://www.netfilter.org/

3 DivertSocket For Linux, http://www.anr.mcnc.org/～divert/index.shtml

4 张钰，唐宁久. Linux 系统 IP 伪装设定以及应用. 计算机应用研究，1999, 16(12): 70～72

5 李方敏，周祖德，刘泉. 基于 Linux 的网络信息内容分析与监管技术. 计算机科学，2003, 30(4): 67～69