基于嵌入式马尔可夫链的网络防火墙性能建模与分析*

马永红[1]，高 洁[2]

(1. 南阳理工学院 计算机与信息工程学院，河南 南阳 473004；2. 郑州航空工业管理学院 计算机科学与应用系，郑州 450015)

摘 要：提出了一种基于嵌入式马尔可夫链的解析排队模型来分析和研究基于一定准则的防火墙在面对正常流量和 DoS 攻击时的性能。基于这种排队模型，得到了一组关于防火墙特征和性能的指标计算方法，这对防火墙的设计来说具有重要意义。同时还提出了一种易于实现的算法来得到这种马尔可夫链模型的状态概率，对防火墙的状态和性能也作了深入的分析。最后通过实验分析验证了提出的解析模型的有效性。

关键词：网络防火墙；性能建模；排队论；马尔可夫链

中图分类号：TP393.08 **文献标志码**：A **文章编号**：1001-3695(2014)05-1491-04

doi:10.3969/j.issn.1001-3695.2014.05.050

Performance modeling and analysis of network firewall based on embedded Markov chain

MA Yong-hong[1], GAO Jie[2]

(1. *College of Computer & Information Engineering, Nanyang Institute of Technology, Nanyang Henan* 473004, *China*; 2. *Dept. of Computer Science & Application, Zhengzhou Instiute of Aeronautical Industry Management, Zhengzhou* 450015, *China*)

Abstract: This paper presented an analytical queueing model based on the embedded Markov chain to study and analyze the performance of rule-based firewalls when subjected to normal traffic flows as well as DoS attack flows. It derived equations for key features and performance measures of engineering and design significance. Moreover, it proposed an algorithm easy to be implemented to derivate the state probability of the analytic model. In addition, it verified and validated this analytical model using simulation and real experimental measurements.

Key words: network firewall; performance modeling; queueing theory; Markov chain

0 引言

网络防火墙是抵御恶意攻击和未授权接入的第一道防线。防火墙通常位于网络的边缘或者私人网络的接入点。传入或传出的流量都需要经过防火墙的检查。基于一定的准则，防火墙能够允许或者是阻断传入或者传出的流量。为了完成这样的任务，网络防火墙会基于一定的准则，对输入数据包按照准则进行比对，直到找到正确的匹配。特别是商业防火墙，如 Cicso PIX 和 FreeBSD 等，都包含了大量的准则，每一条准则都是允许接入的条件[1~3]。如果输入流量满足一条准则的所有条件，防火墙就会采取一定的行为，如允许输入流量通过。一个数据包可能会满足多条准则，在这种情况下，第一条准则采取的行为具有优先权。

防火墙本身可能会受到恶意攻击，因为其大多数都位于网络的边缘。其中一种最严重的攻击就是 DDoS 攻击。根据 Arbor Networks 在 2010 年的报告，与 2009 年相比，DDoS 的增长率是令人吃惊的 102%。这种增长主要是由于僵尸网络的快速增长造成的。

鉴于以上这些原因，分析防火墙在遭受 DDoS 攻击时的性能成为一个非常迫切的问题[4,5]。如果网络防火墙的抗 DDoS 攻击的能力非常弱，则会危及到整个网络的安全性。因此建立一个解析模型以帮助网络设计者预测防火墙在噪声攻击时的性能就变得非常重要。除此之外，防火墙的建模和性能分析有助于理解防火墙的行为和特性。防火墙设计者和系统管理者能够了解到系统安全的瓶颈和影响其性能的关键参数，从而能够作出关键的调整。分析能够为各种设计和操作问题提供快速的答案。本文中提出一种基于排队论的解析模型，以便研究和分析基于准则的防火墙的性能。基于准则的防火墙是使用最广泛的。

本文对防火墙的关键性能指标，如吞吐量、延时、CPU 使用率、数据包丢失等进行了详细的分析，并且得到了这些指标的数学表达式。除此之外，本文提出了一种简便易行的算法以得到本文提出的 Markov chain 模型的状态概率。同时本文通过实验验证了这种解析模型的有效性。最后给出了对防火墙行为和性能的深入分析，特别是在变化的 DoS 攻击流量的影响下，防火墙的吞吐量和 CPU 的利用率是如何受到影响的。

收稿日期：2013-07-08；**修回日期**：2013-08-25 **基金项目**：河南省科技攻关计划资助项目（A13060232）；河南省科技厅计划资助项目（132400411191）

作者简介：马永红(1979-)，男，陕西榆林人，讲师，硕士，主要研究方向为计算机网络、网络信息(mayongred@163.com)；高洁(1978-)，女，河南新乡人，讲师，硕士，主要研究方向为网络信息安全、多媒体水印.

1 解析模型

本章将给出一个有限的排队模型来表示基于准则的网络防火墙的行为和性能。如图 1 所示,包含请求的输入数据包到达防火墙,然后排队等待处理。第一层处理包含数据连接和网络层的一些功能,接下来是包含一定准则的防火墙进行输入数据包的处理。特别地,在 Linux 和 FreeBSD 中,Rx NIC 接收到数据包,然后经过 DMA 复制到 Rx DMA Ring[6]。在 Rx DMA Ring 顺利进行了接收数据包的排队以后,会产生一个中断以通知设备驱动程序接收到了新的数据包。设备驱动程序开始执行数据连接层功能,然后触发 IP 核处理任务。IP 核处理将执行 IP 网络层的任务。

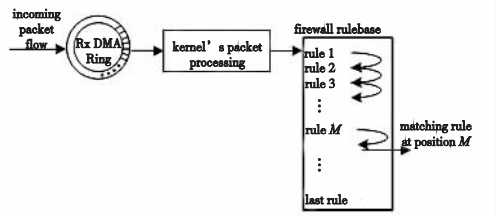

图1 防火墙对输入数据包的检查流程

图 2 为表示图 1 所示系统的行为和动态的有限排队模型。在这个模型中,输入数据包以到达率为 λ 的比例到达防火墙。排队系统的缓存器大小为 K 个数据包,排队大小为 $K-1$ 个数据包。数据包首先需要进行排队,然后进行数据包处理,平均处理时间为$\frac{1}{\mu}$。然后,防火墙会按照一定的准则对数据包进行逐一比对,直到满足其中的一条准则 M。每一准则的平均检查时间为$\frac{1}{r}$。数据包顺序进行 N 个阶段的处理,其中 $N=1+M$,M 为匹配准则的位置。在前一个数据包离开排队系统后,下一个数据包就进入到第一阶段的处理中。各个阶段的执行是相互排斥的,同一时刻只能执行一个阶段的处理。假设输入数据包满足 Possion 到达率 λ,所有的处理时间是相互独立的,并且服从参数为$\frac{1}{\mu}$和$\frac{1}{r}$的指数分布。数据包的处理服从先到先处理的原则。在实际中,$r>\mu$。

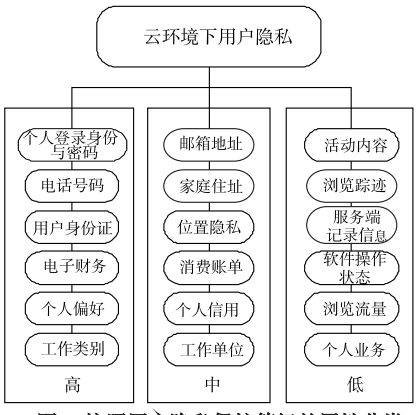

图2 按照用户隐私保护等级的属性分类

本文给出两种解析模型,其中第一种表示基于准则的防火墙在所有输入数据都只满足同一条准则的情况下的行为。第二种模型是对第一种模型的扩展,能够得到输入数据包满足不同位置的准则时的防火墙行为。

1.1 模型分析和求解

本文具有多级处理任务的有限排队模型能够用状态空间为 $S=\{(k,n),0\leqslant k\leqslant K,0\leqslant n\leqslant N\}$的嵌入式 Markov chain 表示,其中 k 表示系统中数据包的数量,n 表示数据包处理阶段的数量。排队系统的排队规模为 $K-1$。当 $n=N$ 时,CPU 进行数据包处理,当 $n=1,\cdots,N-1$ 时,CPU 进行准则比对。也就是说,状态(0,0)表示系统空闲,状态(k,n)表示 CPU 正在处理第 n 条准则,状态(k,N)表示 CPU 正在进行数据包处理。图 3 所示为状态转换图。

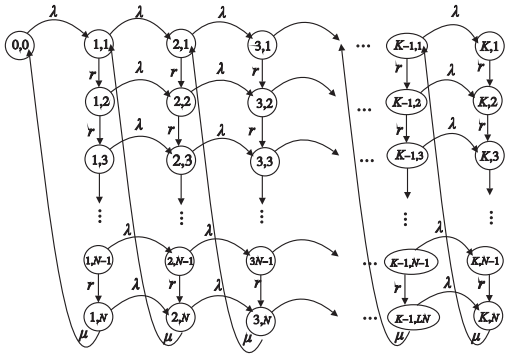

图3 基于准则的具有有限状态的网络防火墙的状态转换

假设 $p_{k,n}$表示状态(k,n)的稳态概率。每个状态的稳态平衡方程如下所示:

$$\text{state}(0,0)\quad 0=-\lambda p_{0,0}+\mu p_{1,N} \tag{1}$$

$$\text{state}(1,N)\quad 0=-(\lambda+\mu)p_{1,N}+rp_{1,N-1} \tag{2}$$

$$\text{state}(1,n)\quad 0=-(\lambda+r)p_{1,n}+rp_{1,n-1}\ 2\leqslant n\leqslant N-1 \tag{3}$$

$$\text{state}(1,1)\quad 0=-(\lambda+r)p_{1,1}+\lambda p_{0,0}+\mu p_{2,N} \tag{4}$$

$$\text{state}(k,N)\quad 0=-(\lambda+\mu)p_{k,N}+\lambda p_{k-1,N}+rp_{k,N-1}\ 2\leqslant k\leqslant K-1 \tag{5}$$

$$\text{state}(k,n)\quad 0=-(\lambda+r)p_{k,n}+\lambda p_{k-1,n}+rp_{k,n-1}\ 2\leqslant k\leqslant K-1;2\leqslant n\leqslant N-1 \tag{6}$$

$$\text{state}(k,1)\quad 0=-(\lambda+r)p_{k,1}+\lambda p_{k-1,1}+\mu p_{k+1,N}\ 2\leqslant k\leqslant K-1 \tag{7}$$

$$\text{state}(K,N)\quad 0=-\mu p_{k,N}+\lambda p_{K-1,N}+rp_{K,N-1} \tag{8}$$

$$\text{state}(K,n)\quad 0=-rp_{K,n}+\lambda p_{K-1,n}+rp_{K,n-1}\ 2\leqslant n\leqslant N-1 \tag{9}$$

$$\text{state}(K,1)\quad 0=-rp_{K,1}+\lambda p_{k-1,1} \tag{10}$$

由式(1)~(10)可知,状态概率 $p_{k,n}$可以由 $p_{0,0}$递归地表示为

$$p_{1,N}=\left(\frac{\lambda}{\mu}\right)p_{0,0} \tag{11}$$

$$p_{1,N-1}=\left(\frac{\lambda+\mu}{r}\right)p_{1,N} \tag{12}$$

$$p_{1,n-1}=\left(\frac{\lambda+r}{r}\right)p_{1,n}\quad 2\leqslant n\leqslant N-1 \tag{13}$$

$$p_{2,N}=\left(\frac{\lambda+r}{\mu}\right)p_{1,1}-\left(\frac{\lambda}{\mu}\right)p_{0,0} \tag{14}$$

$$p_{k,N-1}=\left(\frac{\lambda+\mu}{r}\right)p_{k,N}-\left(\frac{\lambda}{r}\right)p_{k-1,N}\quad 2\leqslant k\leqslant K-1 \tag{15}$$

$$p_{k,n-1}=\left(\frac{\lambda+r}{r}\right)p_{k,n}-\left(\frac{\lambda}{r}\right)p_{k-1,n}\ 2\leqslant k\leqslant K-1;2\leqslant n\leqslant N-1 \tag{16}$$

$$p_{k+1,N}=\left(\frac{\lambda+r}{\mu}\right)p_{k,1}-\left(\frac{\lambda}{\mu}\right)p_{k-1,1}\quad 2\leqslant k\leqslant K-1 \tag{17}$$

$$p_{K,N-1}=\left(\frac{\mu}{r}\right)p_{K,N}+\left(\frac{\lambda}{r}\right)p_{K-1,N} \tag{18}$$

$$p_{K,n-1}=p_{K,n}-\left(\frac{\lambda}{r}\right)p_{K-1,n}\quad 2\leqslant n\leqslant N-1 \tag{19}$$

$$p_{K,1}=\left(\frac{\lambda}{r}\right)p_{K-1,1} \tag{20}$$

需要注意的是,状态概率 $p_{K,1}$既能由式(19)得到,也能由

式(10)得到,是等效的。采用正则化条件 $p_{0,0}$ 可以表示为

$$p_{0,0}+\sum_{k=1}^{K}\sum_{n=1}^{N}p_{k,n}=1 \tag{21}$$

将式(21)的两边同时除以 $p_{0,0}$,可以得到

$$p_{0,0}=\frac{1}{1+\sum_{k=1}^{K}\sum_{n=1}^{N}\frac{p_{k,n}}{p_{0,0}}=1} \tag{22}$$

从式(22)可知,计算 $p_{0,0}$ 可以首先计算 $\frac{p_{k,n}}{p_{0,0}}$,而 $\frac{p_{k,n}}{p_{0,0}}$ 的计算只需要知道参数 λ、μ、r。在得到 $p_{0,0}$ 之后,利用式(11)~(20)可以将其应用到所有其他状态概率的计算。

在得到状态概率之后,就可以得到防火墙的关键特征和性能的度量。系统的平均吞吐量 γ 可以表示为

$$\gamma=\mu\sum_{k-1}^{K}p_{k,N} \tag{23}$$

等效的,系统的平均吞吐量 γ 也可以表示为

$$\gamma=(1-p_0)/\hat{X} \tag{24}$$

其中:p_0 如式(21)所示;$\hat{X}$ 为平均服务时间,且可以表示为

$$\hat{X}=\frac{1}{\mu}+\frac{N-1}{r} \tag{25}$$

参数 γ 也可以用有效到达率 $\lambda'=\lambda(1-P_{\text{loss}})$ 表示。因此

$$\gamma=(1-p_0)/\hat{X}=\lambda(1-P_{\text{loss}}) \tag{26}$$

其中:P_{loss} 为丢失概率,且可以表示为

$$P_{\text{loss}}=1-\frac{1-p_0}{\rho}=\frac{p_0+\rho-1}{\rho} \tag{27}$$

其中:$\rho=\lambda\hat{X}$ 表示流量强度。等效的 P_{loss} 也可以表示为

$$P_{\text{loss}}=\sum_{n=1}^{N}p_{K,n} \tag{28}$$

系统中的数据包数量 $\hat{K}$ 可以表示为

$$\overline{K}=\sum_{k=1}^{K}\sum_{n=1}^{N}kp_{k,n} \tag{29}$$

利用以上结果,可以得到系统中后续的任务进入队列的平均时间为

$$W=\frac{\overline{K}}{\gamma}=\frac{1}{\gamma}\sum_{k=1}^{K}\sum_{n=1}^{N}kp_{k,n} \tag{30}$$

因此,队列的平均等待时间可以表示为

$$W_q=W-\overline{X}=\frac{1}{\gamma}\sum_{k=1}^{K}\sum_{n=1}^{N}kp_{k,n}-\frac{(N-1)\mu+r}{\mu r} \tag{31}$$

防火墙的另一个关键性能指标为 CPU 的利用率。CPU 的利用率可以表示为

$$U_{\text{util}}=\gamma\overline{X} \tag{32}$$

其中:γ 如式(22)所示。

1.2　多流量

在实际应用中,防火墙可能面临多个输入数据包,并且每一个数据包所遵循的准则不一样。这种情况在僵尸网络发起的 DDoS 攻击中非常普遍。本节将对这种情况进行建模和分析。为了分析的简便,假设每一个数据包只触发一条准则。如果一个数据包触发了多条准则,则定义为多数据流。在 1.1 节定义的单流量解析模型也能够用于研究多流量情况下的防火墙性能。每个数据流的到达率分别为 $\{\lambda_i:1\leqslant i\leqslant S\}$,每一个数据量触发一条特定的准则 $\{R_j:1\leqslant j\leqslant L\}$。其中 S 表示输入数据量的数量,L 表示防火墙的规则总数。不同的数据流可能触发相同的准则。

一种解决的方法就是将所有数据量加在一起得到一个总的数据流,然后确定其平均匹配准则的位置 $\overline{M}$。总的数据流的到达率可以表示为

$$\hat{\lambda}=\sum_{i=1}^{S}\lambda_i \tag{33}$$

总的数据流的平均匹配准则的位置 $\overline{M}$ 可以表示为

$$\overline{M}=\left[\sum_{i=1}^{S}\left(\frac{\lambda_i}{\lambda}\times M_i\right)\right] \tag{34}$$

其中:M_i 为数据流 i 的匹配准则位置。

可以采用 1.1 节中的算法估计 p_0,只需要将 λ 替换为 $\hat{\lambda}$,将 N 替换为 $\overline{M}+1$,其他的输入参数 μ、r、K 相同。式(22)~(31)就可以计算总的数据流的性能指标。对每一个数据流也可以计算其性能指标。如对单独的吞吐量 γ_i 可以表示为

$$\gamma_i=\frac{\lambda_i}{\hat{\lambda}}\times\hat{\gamma} \tag{35}$$

其中:$\hat{\gamma}$ 为式(22)表示的总的吞吐量。在得到 γ_i 以后,可以将其用来计算其他的性能指标。CPU 的利用率可以表示为

$$U_{\text{util},i}=\gamma_i\overline{X} \tag{36}$$

每一个数据流的平均数据包延迟 W_i 与总的数据包的延迟 W 相同:

$$W=W_i=\frac{\overline{K}}{\hat{\gamma}} \tag{37}$$

每个数据包的丢失率 $P_{\text{loss},i}$ 和总的数据包的丢失率 P_{loss} 相等。

2　模型验证

为了验证本文模型的有效性,进行了离散事件的仿真,各种假设条件与前面模型中的假设条件一样。仿真的步骤与文献[7]所描述的一致。本文采用 PMMLCG 作为随机数发生器。通过自动化仿真来产生基于不同初始种子的独立复制。在仿真运行过程中,本文检查了随机数据流的重叠问题以确保这样的情况不会发生。具体数据的产生方法如文献[7]所介绍的一样。

为了验证本文的解析模型,将本文的分析结果与文献[9]报道的实际实验结果相比较。图 4 所示为实验设置和测试平台。在文献[9]中报道的测量值包含了防火墙的吞吐量、数据包丢失率、CPU 利用率、数据包延迟。这些测量值是在防火墙面临两种数据流量的情况下得到的正常流量和 DDoS 攻击流量。整个实验设备如图 4 所示。

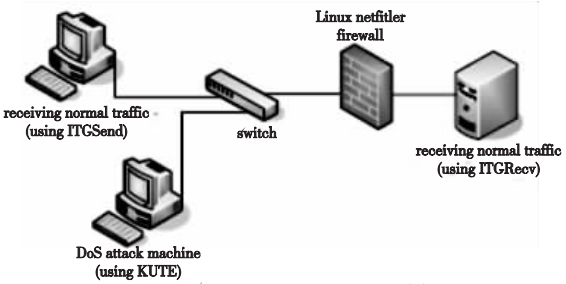

图4　实验设置和测试平台

为了使正常流量通过防火墙,采用开源 D-TIG 2.4.4 发生器[10]。为了产生一个无方向性的数据流,D-TIG 需要进行配置,以便采用 ITGSend agent 发送 UDP 数据流到 ITGRecv。采用 NTP 协议以使得 ITGSend 和 ITGRecv 同步。这对策量准确的数据包延迟非常必要。为了产生 DoS 流量,本文采用 KUTE[11]。KUTE 是开源 UDP 流量发生器。在实验中,CPU 的利用率通过 sar Linux 单元进行测量。然而,吞吐量、数据包丢失率、双向时

间通过 D-ITG 测量。本文采用 64 Byte 的最小数据包。

将防火墙与 Linux netfilter 配置在一起,用 iptables 命令产生 D-ITG 流量准则以及其他 10 000 条准则。通过配置让 Linux Netfilter 接收和允许通过 D-ITG 流量,但是丢掉探测数据包。与 D-ITG 流量或者普通流量相关的准则在防火墙准则的起始部分进行标注。D-ITG 流量被用来进行防火墙性能参数的计算,而其他由 iptables 命令的 shell 脚本产生的没有无链准则,除了 MAC 地址源以外都具有自己独立的条件。所有这些准则都由 UDP 进行编码。对源 MAC 地址使用了一个条件,因为源 MAC 地址消耗的计算资源更多。这是因为 MAC 地址是 netfilter 需要检测的最后条件之一。

3 结果分析

本章将给出防火墙的各种性能指标的实验和分析结果,包括吞吐量、数据包丢失率、CPU 利用率、数据包延迟。本文给出在防火墙面临普通数据流量和 DoS 攻击流量时的这些指标值。除此之外,还给出了分析结果,以便对防火墙动态和行为有更进一步的理解。对于下面给出的实验结果,本文都进行了三次实验,最后给出的结果为平均值。对每一次实验,本文记录的结果都是在持续时间为 30 s 以后的结果。

图 5 所示为发生具有不同速率和遵循不同位置准则的 DoS 攻击时的防火墙性能指标。该性能指标是通过使 ITGSend 产生常规的 10 Kpps 的 UDP 流量而测量得到的。本文通过配置使 Linux netfilter 准则库中的第一条准则允许这样的流量通过。本文分别测量了在普通流量和 DoS 攻击流量情况下防火墙各种性能指标的下降程度。本文假设 DoS 攻击流量的目标在防火墙准则库中的位置分别为 1 000、5 000、10 000。

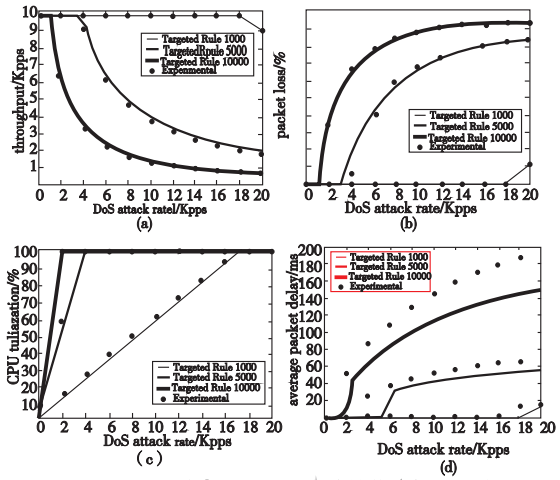

图5 DoS攻击流量对防火墙性能指标的影响

图 5 所示为在 10 Kpps 普通流量情况下的实验测量值和分析结果。关于吞吐量、数据包丢失率、CPU 利用率这三个指标,实验测量值和分析结果非常接近。但是如图 5(d)所示,对于数据包延迟,实验测量值和分析结果所得到的曲线形状非常相似,但是实验测量值更大一些。这主要是因为实验的设置造成的,实验测量值是在发送机处测量得到的,这种延迟比 Linux Netfilter 防火墙处的延迟更大。实验测得的延迟包含了 ITG-Send 发送数据包的延迟、防火墙处的延迟、接收机处的延迟以及传输和排队的延迟。

对于 CPU 的利用率,从图 5(c)中可以看出,当 DoS 攻击流量的目标准则的位置为 5 000 和 10 000 时,实验测量值和分析值非常接近;而当攻击流量的目标准则的位置为 1 000 时,分析值比实验测量值略大。这是因为当攻击流量的目标准则的位置为 1 000 时,对 CPU 的处理要求不是很紧急,这样 CPU 就能处理其他优先级较低的系统任务,从而增加了 CPU 利用率的读数。而当较高优先级的任务占用了 CPU 的大多数处理能力的时候,其他低优先级的任务就不会被处理。也就是说,在 DoS 攻击目标准则的位置较高时,CPU 主要处理防火墙准则。

图 5(a)所示为防火墙在面临 KUTE 发送的 DoS 攻击时,ITG 发送的正常流量的吞吐量降低的情况。图 5(b) ~ (d)所示为相同条件下,数据包丢失、CPU 利用率、数据包延迟的变化情况。从图中可以清楚地看到,当 DoS 攻击的目标位于较高的位置时,这些性能会有一定程度的降低;而当其攻击的目标为最底部的准则时,如位于 5 000 和 10 000 处的准则,这些性能会有比较明显的降低;DoS 攻击的速率越低,这种性能的降低越明显。然而,当攻击的目标位于顶部时,如位于 1 000 处的准则,DoS 攻击的速率越高,性能的降低越明显。

为了更进一步了解防火墙的行为和性能,图 6 给出了防火墙的吞吐量和 CPU 利用率与接收到的攻击目标准则的位置为 5 000时的接收 DoS 攻击流量的关系。图 6 所示为防火墙在面临速率为 10 Kpps 的普通 ITG 流量和速率为 0 ~ 20 Kpps 的 DoS 攻击流量时的性能分析结果和实验测试结果。图 6 中给出了单独的 ITG 和 DoS 流量以及总的流量所对应的吞吐量和 CPU 利用率指标。本文在防火墙处测量了这两个指标,测量结果和图 5(a)中的结果非常接近。单独流量对应的 CPU 无法测量。

如图 6 所示,在速率为 4 Kpps 时,防火墙达到了饱和点。在这个点,如图 6(a)所示,KUTE 发送的 DoS 攻击的吞吐量开始趋于平缓,ITG 普通流量的吞吐量开始下降。除此之外,在这个点,总的流量对应的 CPU 利用率达到 100%,如图 6(b)所示。

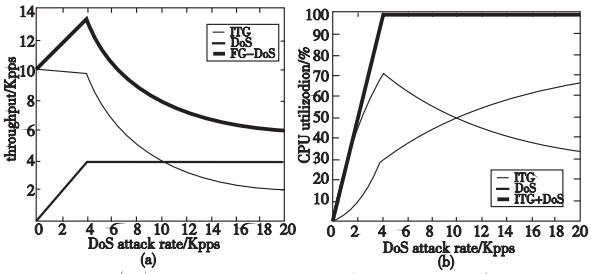

图6 防火墙吞吐量和CPU利用率在DoS攻击目标准则位于5 000时的变化情况

4 结束语

本文提出了一种基于 Markov chain 的防火墙性能的解析模型。从这个模型得到了防火墙在面临 DoS 攻击流量时关键指标的计算方式。这对网络工程师和设计者来说都具有重要作用。这种模型不仅能够计算防火墙面临普通流量时的性能指标,还能够计算面临 DoS 攻击时的性能指标。这些指标包括吞吐量、CPU 利用率、数据包丢失率、数据包延迟。通过实验发现这种模型的分析结果和实验测试结果非常接近,表明了这种模型的有效性。

参考文献:

[1] Cisco PIX firewall release notes[EB/OL]. (2004). http://www.cisco.com/en/US/docs/security/pix/pix62/release/notes/pixrn624.html.

[2] Linux netfilter[EB/OL]. http://www.netfilter.org.

(下转第 1498 页)

户隐私分类后的隐私信息进行保护。首先针对不同的属性分类,制定相应的属性访问控制策略,并根据不同的业务类型,采取相应的 ABE 属性加密方法,最后使用 ABS 属性签名方法对加密的数据进行签名,从而达到用户隐私属性高级别保护的目的。

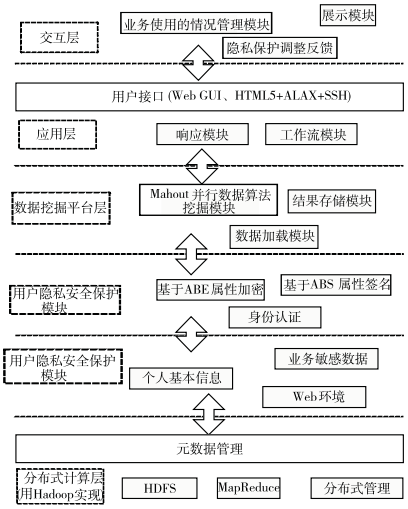

图4 云环境下电信CRM业务系统隐私保护原型系统框架

5 结束语

云计算的应用几乎涉及了信息管理与服务的各个领域,云计算中的隐私问题若得不到很好地解决,则将会直接影响到用户的信息安全和选择云计算的信心。随着电子商务、电子政务、网络教育的兴起与蓬勃发展,人们在网上的活动越来越多,如网上浏览查询、聊天、购物、学习、收发邮件等。在这些网上活动中,将涉及到大量的个人隐私信息,虽然网络导致的隐私泄露已引起社会各界的广泛关注,社会也采取和制定了许多法规、自律政策和技术等措施来加强对网络隐私的保护,但这些措施的作用非常有限,隐私保护仍然是一个亟待解决的挑战性问题,进一步将深入研究云计算环境下的用户隐私属于的形式化表示和隐私保护原型系统的开发。

参考文献:

[1] SU Mu-chun, CHOU C H. A modified version of the K-means algorithm with a distance based on cluster symmetry[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2001, 23(6): 674-680.

[2] DEAN J, GHEMAWAT S. MapReduce: simplified data processing on large clusters[J]. Communication of the ACM-50th Anniversary Issue: 1958-2008. 2008, 51(1): 107-113.

[3] ARMBRUST M, FOX A, GRIFFITH R, *et al*. Above the clouds: a Berkeley view of cloud computing, UCB/ EECS-2009-28 [R]. Berkeley: University of California at Berkeley, 2009.

[4] WANG Li-zhe, TAO Jie, KUNZE M, *et al*. Scientific cloud computing: early definition and experience [C]// Proc of the 10th International Conference on High Performance Computing and Communications. Washington DC: IEEE Computer Society, 2008: 825-830.

[5] KAUFMAN L. Data security in the world of cloud computing[J]. IEEE Security and Privacy, 2009, 7(4): 61-64.

[6] 张军,熊枫. 网络隐私保护技术综述[J]. 计算机应用研究,2005, 22(7):9-11.

[7] Cloud Security Alliance. Security guidance for critical areas of focus in cloud computing v2.1[EB/OL]. (2011-12-24). http://www.cloudsecurityalliance.org/guidance/csaguide.v2.1.pdf.

[8] GODSE M, MULIK S. An approach for selecting software-as-a-service (SaaS) product cloud computing[C]//Proc of IEEE International Conference on Cloud Computing. 2009: 155-158.

[9] LIU Feng, GUO Wei-ping, ZHAO Zhi-qiong, *et al*. SaaS integration for software cloud[C]//Proc of the 3rd IEEE International Conference on Cloud Computing. 2010: 402-409.

[10] ZHANG Yong, LIU Shi-jun, MENG Xiang-xu. GridSaaS: a grid-enabled and SOA-based SaaS application platform[C]//Proc of IEEE International Conference on Services Computing. 2009: 521-523

[11] 李晓飞. 云计算环境下的用户隐私问题浅析[J]. 南昌教育学院学报,2013,28(2):194-196.

[12] SOOD S K. A combined approach to ensure data security in cloud computing[J]. Journal of Network and Computer Applications, 2012, 35(6): 1831-1838.

[13] CLINFTON C, KANTARCIOGLUGLU M, VAIDYA J. Defining privacy for data mining[C]//Proc of National Science Foundation Workshop on Next Generation Data Mining. 2002.

[14] 周水庚,李丰,陶宇飞,等. 面向数据库应用的隐私保护研究综述[J]. 计算机学报,2009,32(5):848-858.

[15] WAQAR A, RAZA A, ABBAS H, *et al*. A framework for preservation of cloud users' data privacy using dynamic reconstruction of metadata [J]. Journal of Network and Computer Applications, 2013, 36(1): 235-248.

[16] 张坤. 面向多租户应用的云数据隐私保护机制研究[D]. 济南:山东大学,2012:20-38.

[17] 兰丽辉,鞠时光,金华,等. 数据发布中的隐私保护研究综述[J]. 计算机应用研究,2010,27(8):2823-2827.

[18] Enterprise privacy authorization language (EPAL 1.2)[EB/OL]. http://www.w3.org/Submission/EPAL/..

(上接第 1494 页)

[3] MELARA A J. Performance analysis of the Linux firewall in a host [D]. [S.l.]: California Polytechnic State University, 2002.

[4] Arbor Networks Inc. Worldwide infrastructure security report[EB/OL]. http://www.arbornetworks.com/report.

[5] SALAH K. Queueing analysis of network firewalls[C]// Proc of IEEE Global Telecommunications Conference. 2010: 1-5.

[6] SALAH K, QAHTAN A. Implementation and experimental performance evaluation of a hybrid interrupt-handling scheme[J]. International Journal of Computers Communications Control, 2009, 32(1): 179-188.

[7] LAW A, KELTON W. Simulation modeling and analysis[M]. 2nd ed. [S.l.]: McGraw-Hill, 1991.

[8] LIU A X, GOUDA M G. Diverse firewall design[J]. IEEE Trans on Parallel and Distributed Systems, 2008, 19(9): 1237-1251.

[9] SALAH K, SATTAR K, SQALLI M, *et al*. A potential low-rate DoS attack against network firewalls[J]. Security and Communication Networks, 2011, 4(2): 136-146.

[10] Distributed Internet traffic generator[EB/OL]. (2008). http://www.grid.unina.it/software/ITG.

[11] ZANDER S, KENNEDY D, ARMITAGE G. KUTE: a high performance kernel-based UDP traffic engine[EB/OL]. (2005). http://caia.swin.edu.au/reports/050118A/CAIA-TR-050118A.pdf.