基于请求关键词的应用层DDoS攻击检测方法

谢柏林 蒋盛益 张倩生
(广东外语外贸大学思科信息学院 广州 510006)

摘 要 目前应用层DDoS攻击严重危害互联网的安全。现有的检测方法只针对某种特定的应用层DDoS攻击,而不能识别应用层上其它的DDoS攻击。为了能快速有效地识别出多种应用层DDoS攻击,提出一种基于请求关键词的应用层DDoS攻击检测方法,该方法以单位时间内请求关键词的频率分布差和个数作为输入,采用隐马尔可夫模型来检测应用层DDoS攻击。实验结果表明,该方法对应用层上的多种DDoS攻击都具有很高的检测率和较低的误报率。

关键词 DDoS攻击,请求关键词,隐马尔可夫模型,应用层
中图法分类号 TP393 **文献标识码** A

Application-layer DDoS Attack Detection Based on Request Keywords

XIE Bai-lin JIANG Sheng-yi ZHANG Qian-sheng
(Cisco School of Informatics, Guangdong University of Foreign Studies, Guangzhou 510006, China)

Abstract Today, the application-layer DDoS attacks may cause great harm to the security of the Internet. Existing detection methods lack the versatility, i. e., an approach only focuses on one particular application-layer DDoS attack. In order to quickly and effectively identify several different application-layer DDoS attacks, this paper presented a detection method based on request keywords. In this method, the input is the number and frequency distribution distance of request keywords per unit time. Then, the hidden markov model is used to detect application-layer DDoS attacks. The experimental results show that the proposed method is valid to discover several different application-layer DDoS attacks with relatively high detection ratio and low false positive ratio.

Keywords DDoS attack, Request keyword, Hidden markov model, Application-layer

1 引言

分布式拒绝服务(Distributed Denial of Service, DDoS)攻击是目前互联网安全面临的主要威胁之一。Arbor Network 2010年的研究报告显示:目前DDoS攻击已成为大多数企业和互联网用户的主要受攻击方式,其规模已突破100Gbit/s[1]。传统的DDoS攻击主要发生在网络层和传输层,例如SYN泛洪攻击、UDP泛洪攻击等[2]。研究者们已对这类攻击进行了深入的研究,提出了一些有效的检测方法。例如李金明等人[3]提出了一种基于VTP分析法的DDoS攻击检测方法;杨新宇等人[4]在分析DDoS攻击的网络流量特性的基础上,提出了一种基于非线性预处理网络流量预测的DDoS攻击检测方法,这两种方法都能有效识别出网络层和传输层上的DDoS攻击。

然而随着网络技术的发展,大部分攻击者都把攻击对象转向网络上的应用或服务,导致应用层攻击不断涌现。目前新出现的网络攻击绝大部分都发生在应用层[5],应用层攻击对网络安全的危害越来越大,尤其是应用层上的DDoS攻击,例如HTTP请求泛洪攻击、僵尸网络(Botnet)发起的垃圾邮件DDoS攻击[6]等。应用层DDoS攻击产生的数据流在网络层和传输层的表现与正常数据流没有显著区别[7],因此传统的DDoS攻击检测方法不能有效识别出这类攻击。

在应用层DDoS攻击的检测方面,现有的研究工作比较少,一些相关的研究工作主要有:Ranjan等人[8]基于HTTP会话(session)的3个参数(即session的建立速率、请求速率和请求消耗),提出使用统计学的方法来检测Web应用层DDoS攻击,Web应用层DDoS攻击是指基于HTTP协议的应用层DDoS攻击。肖军等人[9]根据Web应用层DDoS攻击请求的生成方式,提出了一种session异常度模型,以检测Web应用层DDoS攻击。Xie等人[10]根据单个用户的Web浏览行为,提出了一种基于隐半马尔可夫模型(Hidden semi-Markov Model, HsMM)的Web应用层DDoS攻击检测方法,该方法利用Web服务器日志中的HTTP请求记录来检测Web应用层DDoS攻击。Wen等人[11]提出了一种用于检测Web应用层DDoS攻击的框架,该框架主要根据用户的页面请求速率来检测Web应用层DDoS攻击。Hakem等人[12]提

到稿日期:2012-09-16 返修日期:2013-01-13 本文受国家自然科学基金项目(61202271,61070154),广东省自然科学基金项目(S2012040007184),教育部人文社会科学研究青年基金项目(12YJCZH281),广州市哲学社会科学规划项目(2012GJ31)资助。
谢柏林(1982—),男,博士,讲师,主要研究方向为网络应用层异常检测、微博虚假信息检测,E-mail:xiebailin96@126.com;蒋盛益(1963—),男,博士,教授,主要研究方向为数据挖掘、自然语言处理;张倩生(1975—),男,博士,副教授,主要研究方向为模糊推理。

出利用 TCP 连接(Connection)的一些属性,并基于统计学的方法来检测 Web 应用层 DDoS 攻击。另外,Nagamalai 等人[13]提出了一种用于检测 Botnet 发起的垃圾邮件 DDoS 攻击的方法,该方法主要依靠邮件源地址和内容来检测垃圾邮件 DDoS 攻击。

上述检测方法只针对某种特定的应用层 DDoS 攻击,而不能识别应用层上其它的 DDoS 攻击。为了全面检测应用层 DDoS 攻击,需要在网络中部署多种检测方法。由于这些方法的检测原理和参数设置基本上都不相同,会导致网络管理复杂化,即给网络管理员带来极大的不便。另外,同时部署多种检测方法也会导致网络性能的下降。所以研究一种能快速有效地识别出多种应用层 DDoS 攻击的方法就显得十分必要。

我们前期的研究表明:可以用应用层协议关键词和关键词之间的时间间隔构成观测序列,用 HsMM 来描述单个正常用户在使用某种应用层协议时的行为。基于此研究,我们提出了一种应用层异常检测方法[14]。由于这种方法是把用户产生的每个关键词作为输入,另外由于 HsMM 的时间复杂度比较高,如果采用该方法来检测应用层 DDoS 攻击,则其对 CPU 的消耗会比较大,因此该方法不太适合于应用层 DDoS 攻击的快速检测。

本文提出一种基于请求关键词的应用层 DDoS 攻击检测方法,该方法以单位时间内请求关键词的频率分布差和个数作为观测量,采用隐马尔可夫模型(Hidden Markov Model,HMM)来刻画大量正常用户在同时使用某种应用层协议时的整体行为。这种方法是从全网的角度来检测应用层 DDoS 攻击,该方法首先对大量正常用户在同时使用每种应用层协议时的整体行为进行建模,然后利用这些模型来检测应用层 DDoS 攻击。

2 基于请求关键词的应用层 DDoS 攻击检测方法

应用层 DDoS 攻击是由大量傀儡机(Bot)同时使用某种应用层协议向某个或某组服务器发起的。当大量用户(或主机)在同时使用某种应用层协议时,在网关处可以使用应用层协议关键词(keyword)来描述他们与服务器的通信过程。应用层协议关键词是指能反映用户与服务器通信过程的协议请求命令和服务器响应状态码。例如图 1 为用 HTTP 协议关键词表示的 4 个用户在同时使用 HTTP 协议时的情况,其中 HTTP 协议的关键词为:协议请求命令"GET"、"HEAD"、"POST"、"PUT"、"DELETE"、"TRACE",以及服务器响应状态码"100"、"200"、"304"、"404"等。

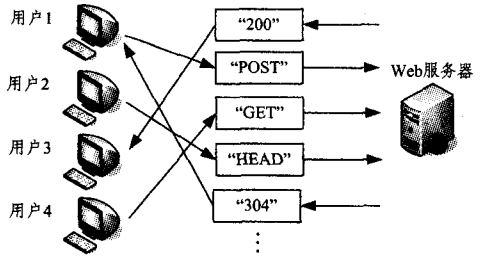

图 1 用户与 Web 服务器之间的通信过程

当大量用户在同时使用某种应用层协议时,他们在一段时间内共同产生的协议关键词按照时间先后可以构成一个关键词序列,该关键词序列中的关键词可分成两类:一类是用户发送的关键词,即请求关键词;另一类是服务器发送的响应关键词。例如在图 1 中,用户共同产生的关键词序列为:"200","POST","GET","HEAD","304",…,其中用户发送的请求关键词序列为:"POST","GET","HEAD",…;服务器发送的响应关键词序列为:"200","304",…。

当大量正常用户在同时使用某种应用层协议时,他们共同产生的请求关键词序列能够反映出这些用户在同时使用该协议时的整体行为。当网络中存在应用层 DDoS 攻击时,会造成请求关键词序列一些统计特征的变化。例如当网络中存在 Botnet 发起的垃圾邮件 DDoS 攻击时,由于攻击者快速发送大量的垃圾邮件,导致单位时间内用户发送的请求关键词明显增多。目前有些应用层 DDoS 攻击为了躲避检测,会降低请求关键词的发送速率,使得单位时间内请求关键词的个数与正常情况相同。由于攻击者很难模拟产生正常的请求关键词序列,因此这类 DDoS 攻击会使请求关键词的频率分布发生变化,请求关键词的频率分布是指每个请求关键词在请求关键词序列中所占的比例。例如当网络中存在速率较低的 HTTP 请求泛洪攻击时,在单位时间内用户共同产生的请求关键词序列中,关键词"GET"所占的比例会非常高,而其它关键词所占的比例会非常低,这使得请求关键词的频率分布与标准分布差距很大,我们把这个差距称为:单位时间内请求关键词的频率分布差。假设某种应用层协议具有 M 个请求关键词,分别表示为:$K_1, K_2, \cdots, K_M$,并假设关键词 K_m ($1 \leqslant m \leqslant M$)在大量正常的请求关键词序列中所占比例的平均值为 μ_m(μ_m 的值在模型训练中得到),则把($\mu_1, \mu_2, \cdots, \mu_M$)作为请求关键词的标准分布。本文使用单位时间内请求关键词的频率分布差和个数,来刻画大量正常用户在同时使用某种应用层协议时的整体行为,以便识别应用层 DDoS 攻击。

当大量正常用户在同时使用某种应用层协议时,他们的整体行为一般都会发生变化。例如大量正常用户在同时使用 SMTP 协议时,当大部分用户分别处于在线阅读邮件、发送邮件时,其整体行为是不一样的。用户整体行为的变化会导致单位时间内请求关键词的频率分布差和个数发生变化。可以把大量正常用户在同时使用某种应用层协议时,其整体行为的不同表现称为状态。假设大量正常用户在同时使用某种应用层协议时,其整体行为具有 N 个离散状态,分别表示为 S_1, $S_2, \cdots, S_N$。而状态的变化就表现为单位时间内请求关键词的频率分布差和个数的变化。假设状态的变化过程可以看作是一个马尔可夫过程,即当前状态只与前一个状态有关,则大量正常用户在同时使用某种应用层协议时,其整体行为的状态转移关系可以用一个具有 N 个状态的马尔可夫链来描述。令 p_{ij} 表示大量正常用户在同时使用某种应用层协议时,其整体行为的状态从 S_i 跳转到 S_j 的概率,其中 $1 \leqslant i, j \leqslant N$。

当大量用户在同时使用某种应用层协议时,令 $x_t = (d_t, q_t)$ 表示在网关处观测到的第 t 个观测量,其中 d_t 表示第 t 个单位时间内请求关键词的频率分布差,q_t 表示第 t 个单位时间内请求关键词的总个数。在本文中,1 个单位时间就是 10 秒钟。假设关键词 K_m ($1 \leqslant m \leqslant M$)在大量正常的请求关键词序列中所占比例的标准差为 σ_m(σ_m 的值在模型训练中得到),则采用简化的马氏距离[15]来求 d_t 的值,d_t 的计算公式如式

(1)所示。其中 c_t^m 表示第 t 个单位时间内关键词 K_m 的总个数，另外 d_t 取的是离散化的整数值。

$$d_t=\sum_{m=1}^{M}\left|\frac{c_t^m/q_t-\mu_m}{\sigma_m}\right| \tag{1}$$

令 $X=x_1x_2\cdots x_T=x_{1\to T}$ 表示大量用户在同时使用某种应用层协议时产生的一个长度为 T 的二维观测序列。由于 d_t 主要与用户的请求方式有关，而 q_t 主要与用户的请求速率和网络时延有关，因此可以近似假定：在给定的状态下，d_t 和 q_t 是相互统计独立的。当大量正常用户在同时使用某种应用层协议时，由于 d_t、q_t 与用户整体行为的状态都不具有一一对应的关系，因此不能直接从他们产生的观测序列中得到用户整体行为的具体状态。所以大量正常用户在同时使用某种应用层协议时，其整体行为的状态转移可以看作是一个隐马尔可夫模型(Hidden Markov Model，HMM)[16]。

在大、中型网络中 q_t 的取值范围会非常大，这会使 HMM 的模型参数变得极度复杂。为了简化 HMM 的模型参数，我们使用 q_t 的对数来表示第 t 个单位时间内请求关键词的总个数。此时观测量 x_t 则变为 $\bar{x}_t=(d_t,\bar{q}_t)$，其中 $\bar{q}_t=\log(q_t+1)$，$\bar{q}_t$ 取的是离散化的整数值。令 $b_i(v,r)$ 表示在给定状态 S_i 下，当 $d_t=v,\bar{q}_t=r$ 时的概率，其中 $0\leqslant v\leqslant V,0\leqslant r\leqslant R$，$V$、$R$ 分别为 v 和 r 的最大可能取值。$b_i(v,r)$ 的定义如式(2)所示，其中 y_t 表示用户整体行为在第 t 个单位时间内所处的状态；$b_i^1(v)$ 表示在给定状态 S_i 下，当 $d_t=v$ 时的概率；$b_i^2(r)$ 表示在给定状态 S_i 下，当 $\bar{q}_t=r$ 时的概率。令 $\pi(i)$ 表示大量正常用户在同时使用某种应用层协议时，其整体行为在第 1 个单位时间内处于状态 S_i 的概率。

$$\begin{aligned}b_i(v,r)&=P[d_t=v,\bar{q}_t=r|y_t=S_i]\\&=P[d_t=v|y_t=S_i]\times P[\bar{q}_t=r|y_t=S_i]\\&=b_i^1(v)\times b_i^2(r)\end{aligned} \tag{2}$$

2.1 模型训练

本文提出的检测方法分为模型训练和攻击检测两个阶段，该方法主要应用于大、中型网络的网关处。在网关处，采集大量正常用户在同时使用某种应用层协议时产生的大量的请求关键词序列作为模型训练的数据集。在模型训练过程中，首先训练每个关键词 $K_m(1\leqslant m\leqslant M)$ 在大量正常的请求关键词序列中所占比例的平均值 μ_m 和标准差 σ_m，然后训练 HMM 的模型参数。假设训练数据集包含 H 个不同的请求关键词序列，即 $Q^{(1)},Q^{(2)},\cdots,Q^{(H)}$，则 μ_m 和 σ_m 通过以下两个步骤得到。

第 1 步　在每个序列 $Q^{(h)}(1\leqslant h\leqslant H)$ 中计算出每个关键词 $K_m(1\leqslant m\leqslant M)$ 在该序列中所占的比例 $f_m^{(h)}$。$f_m^{(h)}$ 的计算公式如式(3)所示，其中 $sum(K_m|Q^{(h)})$ 表示序列 $Q^{(h)}$ 中关键词 K_m 的总个数。

$$f_m^{(h)}=\frac{sum(K_m|Q^{(h)})}{\sum_{m=1}^{M}sum(K_m|Q^{(h)})} \tag{3}$$

第 2 步　求 $f_m^{(h)}$ 的平均值 μ_m 和标准差 σ_m，μ_m、σ_m 的计算公式分别如式(4)、式(5)所示。

$$\mu_m=\frac{\sum_{h=1}^{H}f_m^{(h)}}{H} \tag{4}$$

$$\sigma_m=\sqrt{\frac{\sum_{h=1}^{H}(f_m^{(h)}-\mu_m)^2}{H}} \tag{5}$$

在训练得到每个关键词 $K_m(1\leqslant m\leqslant M)$ 的 μ_m 和 σ_m 后，利用式(1)从训练数据集中提取观测序列。假设从训练数据集中提取出 L 个不同的观测序列，即 $\{\bar{x}_{1\to T_l}^{(l)};1\leqslant l\leqslant L\}$，其中 $\bar{x}_{1\to T_l}^{(l)}$ 为第 l 个观测序列，T_l 为对应序列的长度，并且假定各个观测序列相互独立。我们首先利用文献[16]中的方法给 HMM 的模型参数赋初值，然后运用 Baum-Welch 算法[16]，并采用多个观测序列来训练 HMM 的模型参数。在训练过程中，HMM 的模型参数根据式(6)－式(9)进行更新。

$$\bar{p}_{ij}=\frac{\sum_{l=1}^{L}\sum_{t=1}^{T_l-1}\left(\frac{1}{\omega_l}\chi_t^{(l)}(i,j)\right)}{\sum_{l=1}^{L}\sum_{t=1}^{T_l-1}\left(\frac{1}{\omega_l}\psi_t^{(l)}(i)\right)} \tag{6}$$

$$\bar{b}_i^1(v)=\frac{\sum_{l=1}^{L}\sum_{t=1}^{T_l}\left(\frac{1}{\omega_l}\psi_t^{(l)}(i)\delta(d_t^{(l)}-v)\right)}{\sum_{l=1}^{L}\sum_{t=1}^{T_l}\left(\frac{1}{\omega_l}\psi_t^{(l)}(i)\right)} \tag{7}$$

$$\bar{b}_i^2(r)=\frac{\sum_{l=1}^{L}\sum_{t=1}^{T_l}\left(\frac{1}{\omega_l}\psi_t^{(l)}(i)\delta(\bar{q}_t^{(l)}-r)\right)}{\sum_{l=1}^{L}\sum_{t=1}^{T_l}\left(\frac{1}{\omega_l}\psi_t^{(l)}(i)\right)} \tag{8}$$

$$\bar{\pi}(i)=\frac{\sum_{l=1}^{L}\left(\frac{1}{\omega_l}\psi_1^{(l)}(i)\right)}{\sum_{l=1}^{L}\sum_{i=1}^{N}\left(\frac{1}{\omega_l}\psi_1^{(l)}(i)\right)} \tag{9}$$

在上述公式中，$\chi_t^{(l)}(i,j)$、$\psi_t^{(l)}(i)$、ω_l 的定义分别如式(10)－式(12)所示，其中 Ω 表示隐马尔可夫模型，ω_l 表示观测序列 $\bar{x}_{1\to T_l}^{(l)}$ 相对于模型的概率。另外在式(7)中，如果 $d_t^{(l)}=v$，则 $\delta(d_t^{(l)}-v)=1$，否则 $\delta(d_t^{(l)}-v)=0$；在式(8)中，如果 $\bar{q}_t^{(l)}=r$，则 $\delta(\bar{q}_t^{(l)}-r)=1$，否则 $\delta(\bar{q}_t^{(l)}-r)=0$。

$$\chi_t^{(l)}(i,j)=P[y_t=S_i,y_{t+1}=S_j|\bar{x}_{1\to T_l}^{(l)},\Omega] \tag{10}$$

$$\psi_t^{(l)}(i)=P[y_t=S_i|\bar{x}_{1\to T_l}^{(l)},\Omega] \tag{11}$$

$$\omega_l=P[\bar{x}_{1\to T_l}^{(l)}|\Omega] \tag{12}$$

在 HMM 的模型参数训练结束后，由于 ω_l 的最终取值一般都非常小，并且 ω_l 的值会随着 T_l 的增大而迅速减小[16]，因此我们使用平均对数或然概率 ALL_l(Average Log Likelihood，ALL)来表示观测序列 $\bar{x}_{1\to T_l}^{(l)}(1\leqslant l\leqslant L)$ 相对于训练后的模型的概率，ALL_l 的计算公式如式(13)所示。最后，计算出 ALL_l 的平均值 $\bar{\mu}$ 和标准差 $\bar{\sigma}$，$\bar{\mu}$ 的计算公式如式(4)所示，$\bar{\sigma}$ 的计算公式如式(5)所示。

$$ALL_l=\frac{\log(\omega_l)}{T_l} \tag{13}$$

2.2 攻击检测

在模型训练结束后，当大量用户在同时使用某种应用层协议时，在网关处按照以下步骤来检测基于该协议的应用层 DDoS 攻击。

第 1 步　记录下第 1 个单位时间内所有经过网关的请求关键词，然后计算出 d_1 和 $\bar{q}_1$ 的值，最后根据式(14)计算出前向变量 $\alpha_1(i)$ 的值，其中 $1\leqslant i\leqslant N$。

$$\alpha_1(i)=\pi(i)b_i^1(d_1)b_i^2(\bar{q}_1) \tag{14}$$

第 2 步　记录下当前第 t $(1<t)$ 个单位时间内所有经过网关的请求关键词，然后计算出 d_t 和 $\bar{q}_t$ 的值。

第 3 步　计算前向变量 $\alpha_t(i)$ 的值，$\alpha_t(i)$ 的定义和计算公式分别如式(15)、式(16)所示。然后根据式(17)和式(13)计算出观测序列 $\bar{x}_{1\to t}$ 的 ALL。

$$\alpha_t(i)=P[\bar{x}_{1\to t}, y_t=S_i \mid \Omega] \tag{15}$$

$$\alpha_t(i)=[\sum_{j=1}^{N}\alpha_{t-1}(j)p_{ji}]b_i^1(d_t)b_i^2(\bar{q}_t) \tag{16}$$

$$P[\bar{x}_{1\to t} \mid \Omega]=\sum_{i=1}^{N}\alpha_t(i) \tag{17}$$

第4步　根据式(18)计算出 φ，跳转到第2步。在式(18)中，$\bar{\mu}$、$\bar{\sigma}$ 分别为模型训练阶段得到的 ALL_t 的平均值及标准差。

$$\varphi=\frac{|ALL-\bar{\mu}|}{\bar{\sigma}} \tag{18}$$

在上述检测过程中，φ 的取值反映了大量用户在同时使用这种协议时其整体行为的异常程度。当 φ 的取值大于某个阈值时，则认为出现与该协议相关的应用层 DDoS 攻击，反之则认为用户整体行为正常。当应用层发生 DDoS 攻击时，可以通过限制用户的带宽来抑制应用层 DDoS 攻击。

3　实验测试及结果分析

为了验证本文提出的应用层 DDoS 攻击检测方法的性能，我们对该方法进行了测试。实验测试分为 HTTP 测试、POP3 测试、SMTP 测试和对比测试，其中 HTTP 测试是为了验证该方法对 Web 应用层 DDoS 攻击的检测性能，POP3 测试和 SMTP 测试是为了验证该方法对垃圾邮件 DDoS 攻击的检测性能。在测试过程中，HTTP 协议的请求关键词选为："GET"、"HEAD"、"POST"、"PUT"、"DELETE"、"TRACE"；POP3 协议的请求关键词选为："USER"、"PASS"、"APOP"、"STAT"、"UIDL"、"LIST"、"RETR"、"DELE"、"RSET"、"TOP"、"NOOP"、"QUIT"；SMTP 协议的请求关键词选为："HELO"、"MAIL FROM"、"RCPT TO"、"DATA"、"REST"、"NOOP"、"QUIT"、"VRFY"、"EXPN"、"HELP"。另外，令大量正常用户分别在同时使用这3种应用层协议时其整体行为的初始状态数都为30，然后在模型训练中删除那些很少出现的状态，最终得到大量正常用户分别在同时使用这3种应用层协议时其整体行为的最终状态数。

3.1　HTTP 测试

在该测试实验中，模型训练所用的数据为广州市某个教育网进出口处采集的数据，该数据集持续的时间为一个星期(2008年1月8日－2008年1月15日)，其内网用户数在4000～5000之间。在实验测试过程中，首先需要训练得到每个 HTTP 请求关键词在大量正常的请求关键词序列中所占比例的平均值和标准差。训练得到每个请求关键词在大量正常的请求关键词序列中所占比例的平均值和标准差后，从教育网数据集中提取出6134个正常的 HTTP 观测序列，然后把这6134个观测序列随机地等分成两份，一份用于模型训练，另一份用于测试模型的误报率(False Positive Ratio, FPR)。另外，让20台主机同时运行攻击软件 DoSHTTP[17]，产生的一些真实的 Web 应用层 DDoS 攻击数据作为模型的异常测试数据，用于测试模型的检测率(Detection Ratio, DR)。图2为 HTTP 协议 φ 的阈值与模型检测率(Detection Ratio, DR)、误报率(False Positive Ratio, FPR)的关系。从图2可知，当 φ 的阈值选为2.11时，模型对 Web 应用层 DDoS 攻击的检测率为96.4%，模型对正常观测序列的误报率为0.9%。

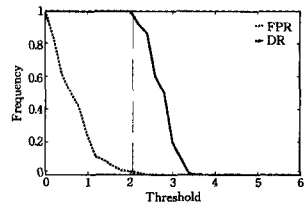

图2　HTTP 的阈值与检测率、误报率

3.2　POP3 测试

在该测试实验中，我们从上述教育网数据集中提取出2576个正常的 POP3 观测序列，并把这2576个观测序列随机地等分成两份，分别用于模型训练和模型测试。为了测试模型对 DDoS 攻击的检测率，让30台主机同时使用 POP3 协议向某个邮件服务器发送比较多的垃圾邮件(每台主机发送邮件的速率为每分钟20封)，它们共同产生的数据则作为模型的异常测试集。图3为 POP3 协议 φ 的阈值与模型检测率、误报率的关系。从图3可知，当 φ 的阈值选为2.23时，模型对基于 POP3 协议的 DDoS 攻击的检测率为97%，模型对正常观测序列的误报率为0.8%。

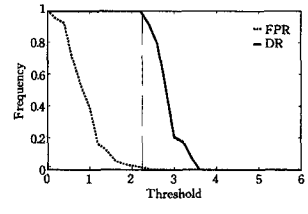

图3　POP3 的阈值与检测率、误报率

3.3　SMTP 测试

在该测试实验中，我们从上述教育网数据集中提取出4518个正常的 SMTP 观测序列，并把这4518个观测序列随机地等分成两组，分别用于模型训练和模型测试。为了模拟 Botnet 发起的垃圾邮件 DDoS 攻击，让30台主机同时使用 SMTP 协议向某个邮件服务器发送比较多的垃圾邮件(每台主机发送邮件的速率为每分钟30封)，它们产生的垃圾邮件数据则作为模型的异常测试集。图4为 SMTP 协议 φ 的阈值与模型检测率、误报率的关系。从图4可知，当 φ 的阈值选为2.46时，模型对垃圾邮件 DDoS 攻击的检测率为96.7%，模型的误报率为0.8%。

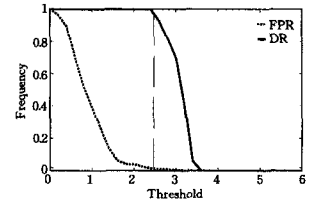

图4　SMTP 的阈值与检测率、误报率

为了进一步测试本文提出的应用层 DDoS 攻击检测方法的性能，我们使用 DARPA'1999 数据集[18]中的 SMTP 数据对该方法进行了测试。DARPA'1999 数据集是一个公开的拥有完整数据包载荷的标准测试数据集，其持续的时间为5周，其中第1周和第3周的数据不包含任何攻击，我们使用的数据为 DARPA'1999 数据集中的 inside tcpdump data。首先从 DARPA'1999 第1周和第3周的数据中提取出用户在同时使用 SMTP 协议时产生的正常的观测序列来训练模型，然

后用训练好的模型去检测 DARPA'1999 数据集中的 Mailbomb 攻击(Mailbomb 是一种应用层 DDoS 攻击)。图 5 为 φ 的阈值与模型检测率、误报率的关系。从图 5 可知,当 φ 的阈值选为 3.05 时,模型对 Mailbomb 攻击的检测率为 96%,模型的误报率为 1.2%。

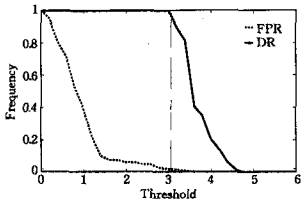

图 5 阈值与检测率、误报率

3.4 对比测试

由于 Xie 等人[10]提出的 HsMM 方法可以利用网关处的数据来检测 Web 应用层 DDoS 攻击,因此可以使用上述 HTTP 测试实验中所用的数据,对本方法和 HsMM 方法进行对比测试。对比测试所用主机的配置为:CPU:Intel Core^2 Q9550(四核心,主频:2.83GHz),内存:4G,硬盘:1T。在误报率都相同的条件下(误报率都选为 1%),测试结果如表 1 所列。HsMM 方法是基于单个用户和隐半马尔可夫模型来检测 Web 应用层 DDoS 攻击的,该方法在接收到用户的每个 GET 请求时,都要调用隐半马尔可夫模型算法来更新观测序列的平均对数或然概率,而隐半马尔可夫模型的时间复杂度比较大[16],因此该方法对 CPU 的消耗比较大。另外,HsMM 方法不能识别其它应用层 DDoS 攻击。

表 1 Web 应用层 DDoS 攻击对比测试结果

检测方法	检测率	CPU 消耗
本方法	96.5%	2.1%
HsMM 方法	98%	12.8%

另外,使用上述 SMTP 测试实验中的教育网数据对本方法和 Nagamalai 等人[13]提出的方法进行了对比测试。在误报率(False Positive Ratio,FPR)都相同的条件下(误报率都选为 1%),对比测试结果如表 2 所列。Nagamalai 等人[13]提出的检测方法主要依靠邮件源地址和内容来检测垃圾邮件 DDoS 攻击,该方法是采用多层过滤的思想来识别垃圾邮件 DDoS 攻击的,由于每层的过滤规则相对比较简单,因此该方法的检测率不是很高。

表 2 垃圾邮件 DDoS 攻击对比测试结果

检测方法	检测率	CPU 消耗
本方法	96.8%	1.7%
Nagamalai 等人的方法	85%	1.9%

从上面的测试实验可知,本文提出的方法在检测应用层 DDoS 攻击时具有较高的检测率和较低的误报率,并且该方法能识别出多种应用层 DDoS 攻击。由于 HMM 的时间复杂度比较低,因此该方法具有比较快的检测速度。

通常我们能很容易获得大量的正常观测序列,所以就能比较容易确定 φ 的门限值与模型误报率之间的关系。当本方法应用于实际网络中时,可以根据人们对误报率的要求来设置 φ 的门限值。另外,在实际应用中,可以采用在线训练模型来动态更新 HMM 的模型参数,即在线采集正常的观测序列,然后每隔一定的时间就重新训练 HMM 的模型参数。这样可以提高模型的准确度,使模型能更好地刻画出大量正常用户的整体行为特征,从而提高模型的检测性能。

结束语 本文提出了一种基于请求关键词的应用层 DDoS 攻击检测方法,该方法以单位时间内请求关键词的频率分布差和个数作为观测量,使用隐马尔可夫模型从全网的角度来检测应用层 DDoS 攻击。我们对本文提出的方法进行了一些测试,实验结果表明该方法能快速有效地识别出多种应用层 DDoS 攻击。今后的工作主要是:在真实的大规模网络中测试该方法的在线性能。

参 考 文 献

[1] Worldwide Infrastructure Security Report 2010 [EB/OL]. http://www.arbornetworks.com/report

[2] 孙长华,刘斌. 分布式拒绝服务攻击研究新进展综述[J]. 电子学报,2009,34(7):1562-1570

[3] 李金明,王汝传. 基于 VTP 方法的 DDoS 攻击实时检测技术研究[J]. 电子学报,2007,35(4):791-796

[4] 杨新宇,杨树森,李娟. 基于非线性预处理网络流量预测方法的泛洪型 DDoS 攻击检测算法[J]. 计算机学报,2011,34(2):395-405

[5] 谢柏林,余顺争,王宇. 应用层异常检测方法研究[J]. 计算机科学,2009,36(4):21-24

[6] Xie Y,Yu F,Achan K,et al. Spamming Botnets:Signatures and Characteristics[J]. ACM SIGCOMM Computer Communication Review,2008,38(4):171-182

[7] Yu J,Fang C,Lu L,et al. A Lightweight Mechanism to Mitigate Application Layer DDoS Attacks[J]. Scalable Information Systems,2009,18:175-191

[8] Ranjan S,Swaminathan R,Uysal M,et al. DDoS-Shield:DDoS-resilient Scheduling to Counter Application Layer Attacks[J]. IEEE/ACM Transactions on Networking,2009,17(1):26-39

[9] 肖军,云晓春,张永铮. 基于会话异常度模型的应用层分布式拒绝服务攻击过滤[J]. 计算机学报,2010,33(9):1713-1724

[10] Xie Y,Yu S Z. Monitoring the Application-Layer DDoS Attacks for Popular Websites[J]. IEEE/ACM Transactions on Networking,2009,17(1):15-25

[11] Wen S,Jia W,Zhou W,et al. CALD:Surviving Various Application-Layer DDoS Attacks That Mimic Flash Crowd[C]// The 4th International Conference on Network and System Security. 2010:247-254

[12] Hakem B,Geert D. Trackling Application-layer DDoS Attacks [J]. Procedia Computer Science,2012,10:432-441

[13] Nagamalai D,Dhinakaran C,Lee J K. Novel Mechanism to Defend DDoS Attacks Caused by Spam[J]. International Journal of Smart Home,2007,1(2):83-95

[14] 谢柏林,余顺争. 基于应用层协议关键词序列的应用层异常检测方法[J]. 计算机研究与发展,2011,48(1):159-168

[15] Wang K,Stolfo S J. Anomalous Payload-Based Network Intrusion Detection[C]// The Seventh International Symposium on Recent Advances in Intrusion Detection. 2004:203-222

[16] Rabiner L R. A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition [J]. Proceedings of the IEEE,1989,77(2):257-286

[17] DoSHTTP[EB/OL]:http://www.socketsoft.net/

[18] Mahoney M V,Chan P K. An Analysis of The 1999 DARPA/Lincoln Laboratory Evaluation Data for Network Anomaly Detection[C]// The Sixth International Symposium on Recent Advances in Intrusion Detection. 2003:220-237