SDN 环境下基于 Renyi 熵的低速率分布式拒绝攻击的检测

王文涛,王玲霞,黄 烨
(中南民族大学 计算机科学学院,武汉 430074)

摘 要 针对现在对低速率分布式拒绝攻击的研究不足,提出了一种在软件定义网络(SDN)环境下,利用 Renyi 熵来检测 L-DDoS 的方法. 该方法首先在控制器上收集 PACKET_IN 数据包,然后基于目的 IP 来计算 Renyi 熵,最后通过设定一定的阈值来检测异常流量. 实验结果表明: 相比于利用香农熵的检测方法,该方法通过调整一定目的 IP 熵的阶数可以检测 L-DDoS 攻击流量从而降低误警率.

关键词 低速率分布式拒绝攻击; 因特网安全; 软件定义网络; Renyi 熵; 误警率
中图分类号 TP393 **文献标识码** A **文章编号** 1672-4321(2017)03-0131-06

Detection of Low Rate DDoS Attacks Based on Renyi Entropy in SDN Environment

Wang Wentao, *Wang Lingxia*, *Huang Ye*
(College of Computer Science, South-Central University for Nationalities, Wuhan 430074, China)

Abstract At present, since the research of the L-DDoS attack is not too much, a method of detecting L-DDoS using Renyi entropy based on a software defined network (SDN) was proposed. Firstly, PACKET_IN data packets were collected on controller, and then the Renyi entropy was calculated based on destination IP. Finally setting a threshold was used to detect abnormal traffic. The experimental results showed that compared with the Shannon entropy detection method, this method can detect the L-DDoS attack traffic and reduce the false alarm rate by adjusting the number of orders.

Keywords L-DDoS; Internet security; software defined network; Renyi entropy; false alarm rate

近年来,由于分布式拒绝(DDoS)攻击的发生频率、危害性与设计复杂度逐渐增加,已经成为当今因特网安全的最主要威胁之一. 拒绝服务的攻击方式具有带宽消耗型攻击和资源消耗型攻击这两种形式,两种形式都是透过大量合法或伪造的请求占用大量网络资源,从而达到使网络以及系统瘫痪的目的. 随着攻击方式的演变,低速率分布式拒绝(L-DDoS)攻击作为一种和正常流量非常相似的 DDoS 攻击在网络中产生. 它具有很强的隐蔽性,能躲避一些异常检测装置,因此目前针对 L-DDoS 的检测是困难的[1]. 误警率是将正常流量当作异常流量的概率. 如果在一个稳定网络中,设置检测阈值时一味地追求高检测准确率,会造成误警率过高,一些正常用户的访问请求就会被处理,从而造成他们获取不良的服务体验. 因此,误警率在攻击检测中也是评价检测方法的一个重要性能指标.

目前针对传统网络架构的缺陷,具有大量的针对 DDoS 攻击的检测方法. 目前 DDoS 攻击检测方法主要分为误用检测和异常检测. 目前大多数的 DDoS 攻击检测都属于异常检测. 文献[2,3]利用数据流量之间具有线性关系,提出在流量分布服从任何分布的情况下,都可以利用等级相关系数的方法来区分正常流量和异常流量. 但这种方法没有在真实网络环境下实现. 文献[4]提出新的信息度量方法以

收稿日期 2017-05-04
作者简介 王文涛(1967-),男,副教授,博士,研究方向: 计算机网络与控制,E-mail: wangwt@mail.scuec.edu.cn
基金项目 国家民委教改基金资助项目(15013)

及联合检测,可以提前检测出异常流量以及追溯至攻击网络. 文献[5]利用拥塞控制特性,通过检测丢包率来检测. 这些方法都是假设网络具有完全控制能力. 但由于在传统网络不能集中控制所有路由器,不能及时反映流量信息,这些检测方法易造成检测率低. 在针对SDN网络中存在的DDoS攻击,文献[6]提出了基于目的IP的信息熵的方法来进行早期检测,目的是防止控制器瘫痪,降低控制器负担,但这种基于香农熵的检测在大流量攻击的前提下有效,对低速率检测效果不明显并且误警率高. 文献[7,8]提出基于神经网络方法来对OpenFlow流量进行检测,由于神经网络方法训练时间可能很长、收敛慢,可能导致漏检率高. 然而在一些要求时延的网络,对时间要求是非常严格的. 文献[9]利用序贯概率比检测方法来检测哪些北向接口容易受到攻击,这种检测对于把流量分成两类即正常流量和低速率异常流量,不能区分出正常突发流量和低速率异常流量,导致抗背景流量弱,而且容易受到攻击样本的影响,如果攻击样本大,这种方法检测效率低.

以上检测方法都是针对高速率检测有效,对低速率检测则误警率高,从而它们对低速率的检测不适用. 本文结合SDN控制器集中控制和全面管控全网信息的特点,以及Renyi熵比香农熵更能区分两个概率分布之间差异的特性来检测L-DDoS攻击,从而降低误警率.

1 L-DDoS

文献[4]是从数据包角度来定义L-DDoS的,为每秒攻击包个数低于1000,高速率攻击定义为每秒攻击包超过10000. 因此L-DDoS可能和正常流量不管在速率上还是在分布上非常相似,没有高速率攻击差距大. 比如,利用TCP协议的自适应机制而形成的攻击. 超时重传机制是TCP拥塞控制中一种自适应机制来避免网络拥塞,攻击者可通过扫描获得设置的RTO时间,这期间内发送一些虚假的数据包使网络拥塞程度加剧,利用这种机制来使网络不断处于启动超时重传机制,每次超时就会启动慢启动算法,将发送端的阈值设置为当前拥塞窗口的一半,将窗口值设置为1,同时重新发送此报文. 当启动RTO机制时,就停止攻击. 其攻击过程如图1所示,这时攻击时间短,但平均和正常的速率差不多.

尽管L-DDoS特性是攻击速率低,但可能造成网络性能低下,如果攻击周期长,可能造成巨大的威胁. 因此,检测效率对网络性能起到关键作用. 由于在真实网络攻击中,L-DDoS攻击流量和正常流量混在一起,没有完全消耗网络资源,最高也只占完全攻击的60%,若要在正常流量中区分出异常流量比在完全攻击情况下困难. 因为完全攻击中只存在攻击流量,也就是高速率攻击和正常流量在速率、特性分布上都有较大差距. 低速率攻击检测相比高速率攻击检测更加困难. 因此需要一些检测方法来增加正常流量和L-DDoS攻击流量之间的区分度.

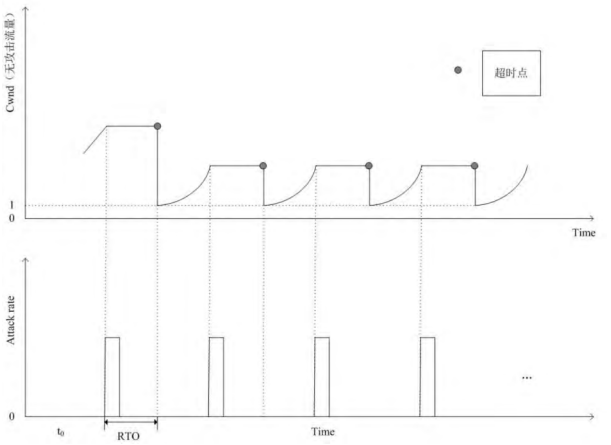

图1 基于RTO机制的攻击

Fig. 1 An attack based on RTO

2 Renyi熵

在信息论中,信息熵是用来衡量随机变量的多样性、不确定性和随机性的指标. 随机变量的随机性越高,熵值越大;相反,随机变量的确定度越大,熵值越小. 信息熵度量已经被运用在各个领域,反映某种分布变化,如基于香农熵、Tsallis熵的应用,但针对小流量的度量,这些信息熵不适用. 从而体现出Renyi熵在小流量检测上具有重大意义. 文献[10]给出Renyi熵的一般定义,定义α阶Renyi熵为:

$$H_\alpha(x) = \frac{1}{1-\alpha}\log_2\left(\sum_{i=1}^{n} p_i^\alpha\right), \quad (1)$$

满足$p_i \in \{p_{x_1}, p_{x_2}, \cdots, p_{x_n}\}$($p_i$是随机变量$x_i$的概率),$p_i \geqslant 0$,$\sum_{i=1}^{n} p_i^\alpha = 1$,$\alpha \geqslant 0$,$\alpha \neq 1$.

当$\alpha = 0$或者$p_{x_1} = p_{x_2} = \cdots = p_{x_n}$时获得最大值:

$$\max(H_\alpha(x)) = \log_2(n). \quad (2)$$

获得最大值时表明概率密度达到最大集中度.

当$\alpha \to 1$时,$H_x(\alpha)$收敛于香农熵.

$$\lim_{\alpha\to 1}H_x(\alpha) = \lim_{\alpha\to 1}\frac{1}{1-\alpha}\log_2\left(\sum_{i=1}^{n}p_i^{\alpha}\right) = \lim_{\alpha\to 1}\frac{\sum_{i=1}^{n}p_i^{\alpha}\ln p_i}{\sum_{i=1}^{n}p_i^{\alpha}\ln 2},$$

由于 $\sum_{i=1}^{n}p_i^{\alpha}=1$,因此:

$$\lim_{\alpha\to 1}H_x(\alpha) = -\sum_{i=1}^{n}p_i\cdot\log_2 p_i. \tag{3}$$

(3) 式表明阶数趋近于 1 时,等价于香农熵. 它们的本质都是衡量一个变量随机性的程度.

当 $\alpha\to 0$,

$$\frac{\partial H_x(\alpha)}{\partial\alpha}\leqslant 0. \tag{4}$$

(4) 式表明 $H_x(\alpha)$ 是一个随着 α 增大而减小的非递增函数.

用熵来衡量网络的随机性: 随机性越高,熵值越大; 反之亦然. 当 $\alpha>1$ 时,高概率事件对 Renyi 熵的影响比香农熵要大. 文献 [11] 证明了 Renyi 熵比香农熵更能增大两个不同分布之间的差异. 因此 Renyi 熵能区分合法流量和异常流量之间的分布变化,最终使其能区分出 L-DDoS 攻击.

3　滑动窗口模型

滑动窗口模型是为减少采样数据的存储成本和计算复杂度,同时保持数据之间的连续性而被提出来的. 假设定义滑动窗口模型为 $W(b,w,a)$,b 是基于滑动窗口的数据包,w 是滑动窗口的宽度,a 是滑动窗口的滑动步长,w 和 a 均以一个采样点作为单位长度.

$b(i,j)$ 表示第 $i\sim j$ 个采样点所组成的数据包,其中 $i<j$; $i=1,2,3,\cdots$; $j=1,2,3\cdots$. 具体实施时,数据包进入大小为 w 的当前滑动窗口内,假设当前滑动窗口第一个数据是 $b(i)$,那么当 w 被填满时窗口内的数据流为 $b(i,i+w-1)$,滑动窗口向前滑动 a 个单位长度,当前滑动窗口就被更新为 $b(i+a,i+a+w-1)$,从而随后收集的数据包流入新的滑动窗口,直至被填满后再一次滑动.

4　基于 Renyi 熵的 L-DDoS 检测

L-DDoS 攻击的检测利用了 SDN 的特性,并在此网络中实施. 随着云计算、大数据等互联网新技术的出现,SDN 是为了解决传统网络架构的逻辑功能和转发功能耦合强度被提出来的. SDN 被划分成 3 层: 基础设施层、控制层、应用层. 利用 SDN 的特性不仅可以降低成本,而且可以及时反映全网信息. 文献 [12] 提出 OpenFlow 技术作为 SDN 架构的一种实现方式被应用在很多方面,如负载均衡、流量管理、路由,等. 大多数情况下,控制器和交换机之间通过 OpenFlow 协议来进行通信. OpenFlow 协议中的消息可以解决检测启动问题和流量收集复杂问题,分别为 packet_in 消息、ofp_stats_request、ofp_stats_reply.

检测算法开始之前遵循算法启动机制,本文的检测机制是利用 SDN 网络中的 packet_in 来开始启动. 在 SDN 网络中,对于每一新连接,控制器将在交换机流表里安装流表项,使到达的数据包能被正确转发到目的主机. 但当新的数据包到达交换机时,由于交换机无此流表项,因此不能匹配交换机的流表项,这时交换机会产生 packet_in 消息并发送给控制器,控制器接受这个消息后,决策是丢弃该数据包或者添加相应的流表项. 一般情况下的动作是添加相应的流表项. 在攻击产生时,往往会在交换机中产生这个消息. 因此,我们可以利用 packet_in 特性来定义检测开始周期,通过这个消息的到达来启动检测算法(算法 1).

算法 1: 基于 Renyi 熵的 L-DDoS 攻击检测算法

```
INPUT: Packet _ In, Window Size, Threshold, Threshold Count
OUTPUT: Attack Detected
  WHILE Number of Packet_In < sliding Window Size DO
    Receive a new Packet_In message
    get D_IP by parsing Packet_In
  Calculate Renyi entropy in sliding Window Size
  IF Renyi entropy < Threshold THEN
    count = count + 1
    IF count = = Threshold Count THEN
      RETURN Attack Detected
    ELSE
      GOTO 1
ELSE
    count = 0
    GOTO 1
```

假设窗口大小为 n,则检测计算复杂度就为 $O(n)$,该过程在线性时间内完成,空间复杂度为 $O(n)$. 检测算法需要对流量进行收集,流量收集通

过发送 ofp_stats_request 消息以及回复 ofp_stats_reply 消息来获得流量信息,据此有效地监测网络流量.

我们开始预先定义窗口值、检测阈值以及超过检测阈值的次数. 当网络中具有 packet_in 消息时,就解析 packet_in 消息中的目的 IP. 达到一定窗口下的数据包就统计目的 IP 地址并求得熵,当 Renyi 熵值超过阈值,就增加统计次数,达到一定的统计次数则丢弃该数据包,继续下一窗口的检测. 具体检测流程如图 2 所示.

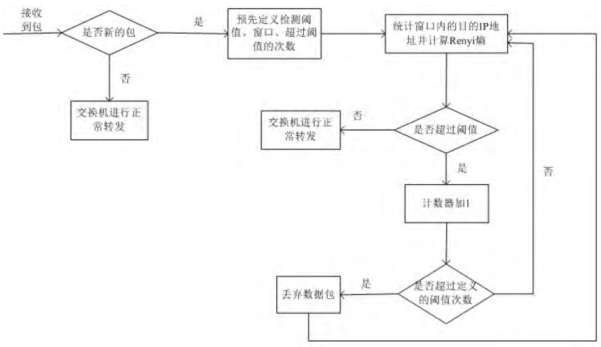

图 2 检测流程图

Fig. 2 A flowchart of detection

5 实验及分析

本文实验采用 POX 控制器和 Mininet 搭建网络仿真环境. POX 控制器基于组件化设计,为软件组件提供定义良好的 API,网络管理者可以有效地利用 API 创建新的网络管理和控制应用程序. Mininet 是由 Stanford 大学开发的一套进程虚拟化网络模拟器,它使用轻量级的虚拟化技术在单个系统中模拟出拥有多个主机、交换机和链路的网络环境. Mininet 是一个轻量级的软件定义网络和测试平台,采用了轻量级的虚拟化技术. Mininet 支持 OpenFlow、OpenvSwitch 等各种协议,也可以在一台计算机上模拟一个包含主机、链路和交换机的完整网络. OVS 被用来当作网络交换机. Mininet 运行在 Ubuntu 操作系统上,控制器监听端口为 6633. POX 控制器来实现全网管控功能. 使用 Mininet 仿真一个具有 3 层、9 个交换机、64 台主机的树形拓扑. IP 地址为 10.0.0.1 至 10.0.0.64,选择目的 IP 地址为 10.0.0.64 的主机作为受害者,IP 为 10.0.0.1、10.0.0.2、10.0.0.3 作为攻击者来发动攻击. 使用 Scapy 生成虚假的 IP 地址数据包来产生低速率异常流量,模拟 L-DDoS 攻击,其中主要产生的数据包是 Tcp 包. 本实验在不同的滑动窗口下以及在不同强度的低速率流量下进行. 实验中滑动窗口的滑动步长设置为窗口宽度的一半,窗口大小是数据包个数的大小.

在不同的窗口下的实验如图 3 ~ 5 所示,α 表示熵的阶数,在每个窗口内随着阶数 α 的增大,熵值逐渐减少,但正常流量和异常流量的香农熵非常接近. 图 3 表明在窗口为 400 个数据包时,在 $t = 20$s 加入低速率流量,正常流量和异常流量的香农熵基本不变;但在 $\alpha = 10$ 时,正常流量和异常流量 Renyi 熵之间的差距变大,从 0.0093 增大到 0.0923. 图 4 表明在窗口为 500 个数据包时,同样在 $t = 20$s 加入低速率流量,正常流量和异常流量的香农熵非常接近,但在 $\alpha = 10$ 时,正常流量和异常流量 Renyi 熵之间的差距变大,从 0.0088 增大到 0.1564. 图 5 表明在窗口为 600 个数据包时,正常流量和异常流量 Renyi 熵之间的差距从 0.0167 增加到 0.0836. 表 1 ~ 3 显示不同窗口下的正常流量和异常流量平均熵值,在每个窗口下,都是随着 α 的增大,正常流量和异常流量的平均熵值在减小,但其差距变大.

再观察在同一个窗口内不同强度的低速率攻击流量. 本实验攻击强度使攻击流量比例从 1 增加到 5,结果如图 4、图 6 所示. 图 4 是攻击比例为 1 时的熵值,图 6 是攻击比例为 5 的熵值;随着攻击强度的增大,差距逐渐增大. 在攻击比例为 5 时,差距从 0.0767 增加到 0.6850,表明检测率更高. 攻击强度越大,正常流量和异常流量香熵值之间的差值增大,这是由于攻击流量占总流量的比例在增加. 但其增加的比例小,在短时间内是无法识别的. 因为不管强度多大,每次攻击还是处于低速率攻击,每秒攻击包数都小于 1000. 然而,Renyi 熵不管在哪种强度下,都可以"放大"它们之间的差距,从而可以更容易地检测出流量异常. 因此表明 Renyi 熵对于 L-DDoS 攻击的检测是有效的.

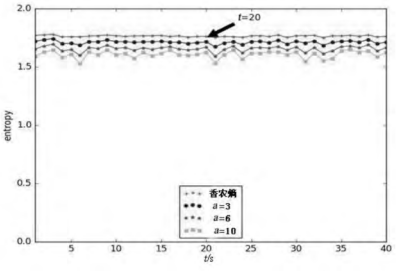

图 3 窗口 400

Fig. 3 Window 400

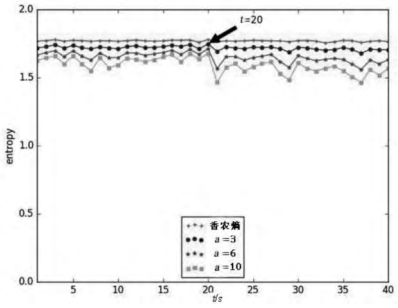

图 4　窗口 500-强度 1

Fig. 4　Window 500-intensity 1

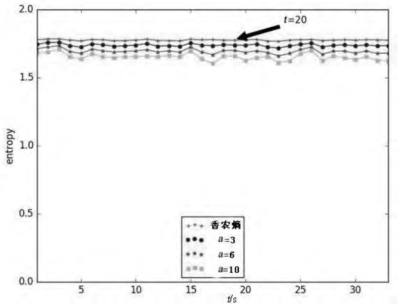

图 5　窗口 600

Fig. 5　Window 600

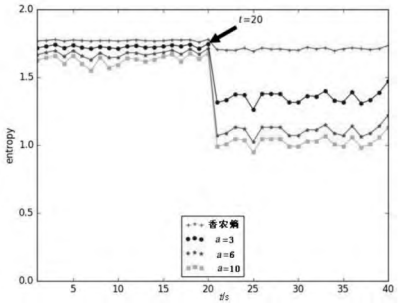

图 6　窗口 500-强度 5

Fig. 6　Window 500-intensity 5

表 1　窗口 400 不同阶的 Renyi 熵

Tab. 1　Different order Renyi entropy in 400 window

阶数	正常流量平均熵值	异常流量平均熵值	平均熵值差	拐点差
香农熵	1.7660	1.7637	0.0023	0.0093
$\alpha=3$	1.7150	1.7103	0.0047	0.0428
$\alpha=6$	1.6585	1.6539	0.0045	0.0786
$\alpha=10$	1.6091	1.6066	0.0025	0.0923

表 2　窗口 500 不同阶的 Renyi 熵

Tab. 2　Different order Renyi entropy in 500 window

阶数	正常流量平均熵值	异常流量平均熵值	平均熵值差	拐点差
香农熵	1.7703	1.7685	0.0017	0.0088
$\alpha=3$	1.7233	1.7233	0.0025	0.0350
$\alpha=6$	1.6718	1.6718	0.0068	0.1003
$\alpha=10$	1.6279	1.6279	0.0129	0.1564

表 3　窗口 600 不同阶的 Renyi 熵

Tab. 3　Different order Renyi entropy in 600 window

阶数	正常流量平均熵值	异常流量平均熵值	平均熵值差	拐点差
香农熵	1.7647	1.7564	0.0083	0.0167
$\alpha=3$	1.7144	1.6969	0.0175	0.0368
$\alpha=6$	1.6620	1.6368	0.0252	0.0634
$\alpha=10$	1.6176	1.5860	0.0316	0.0836

我们需要计算误警率,利用香农熵和 Renyi 熵,减少的误警率定义如下:

$$\beta' = \frac{H_x - H_r}{H_x}. \tag{5}$$

H_x 为当取香农熵正常流量和异常流量之间差距的中值为阈值时,正常流量被判断为异常流量的个数,H_r 为当取 Renyi 熵正常流量和异常流量之间差距的中值为阈值时,正常流量被判断为异常流量的个数. 如表 4 所示,和香农熵相比,随着阶数的增大,Renyi 熵不同阶数降低的误警率越大. 在 $\alpha = 10$ 的情况下,降低的误警率为 100%,表明此阶数下的误警率最低,因此更能区分异常流量和低速率异常流量.

表 4　Renyi 熵减少的误警率

Tab. 4　False alarm rate reduced by Renyi entropy

阶数	误警的次数/次	减少的误警率/%
$\alpha=3$	8	20
$\alpha=6$	1	90
$\alpha=10$	0	100

本文的实验是在单一受害结点上进行的,下一步可能会针对多受害结点攻击进行研究.

6　结语

分布式拒绝攻击已经成为当今因特网安全的最主要威胁之一. 为了获得可信赖的网络,攻击检测和攻击防御是必不可少的,但攻击检测是攻击防御的前提,因此攻击检测在高可靠网络中必不可少. 及时检测隐蔽性强的 L-DDoS 攻击对今后因特网安全具有深远的意义. 本文充分利用 SDN 全网管控的特性提高流量收集的及时性,并利用 Renyi 熵的特性来区分正常流量和攻击流量. 实验结果表明: 本文方法

在针对L-DDoS的检测时误警率低,可以有效地检测在网络中的L-DDoS攻击流量.

参 考 文 献

[1] 文 坤,杨家海,张 宾. 低速率拒绝服务攻击研究与进展综述[J]. 软件学报,2014,25(3):591-605.

[2] Ain A, Bhuyan M H, Bhattacharyya D K, et al. Rank correlation for low-rate DDoS attack detection: An empirical evaluation[J]. International Journal of Network Security, 2016, 18(3):474-480.

[3] Wei W, Chen F, Xia Y, et al. A rank correlation based detection against distributed reflection DoS attacks[J]. IEEE Communications Letters, 2014,17(17):173-175.

[4] Xiang Y, Li K, Zhou W. Low-rate DDoS attacks detection and traceback by using new information metrics[J]. IEEE Transactions on Information Forensics & Security, 2011, 6(2):426-437.

[5] Zhang C, Cai Z, Chen W, et al. Flow level detection and filtering of low-rate DDoS[J]. Computer Networks the International Journal of Computer & Telecommunications Networking, 2012, 56(15): 3417-3431.

[6] Mousavi S M, Sthilaire M. Early detection of DDoS attacks against SDN controllers[J]. International Conference on Computing, 2015,17(17):77-81.

[7] Braga R, Mota E, Passito A. Lightweight DDoS flooding attack detection using NOX/OpenFlow[C]//IEEE. Conference on Local Computer Networks. Washington D C: IEEE Computer Society, 2010:408-415.

[8] Cui Y, Yan L, Li S, et al. SD-anti-DDoS: fast and efficient DDoS defense in software-defined networks[J]. Journal of Network & Computer Applications, 2016, 68: 65-79.

[9] Dong P, Du X, Zhang H, et al. A detection method for a novel DDoS attack against SDN controllers by vast new low-traffic flows[C]//IEEE. International Conference on Communications. Kuala Lumpur: IEEE, 2016:1-6.

[10] Yan R, Zheng Q, Peng W. Multi-scale entropy and Renyi cross entropy based traffic anomaly detection[C]//IEEE. International Conference on Communication Systems (ICCS). Singapore: IEEE, 2008: 554-558.

[11] Zyczkowski K. Renyi extrapolation of Shannon entropy[J]. Physics, 2003, 10(3):297-310.

[12] 张朝昆,崔 勇,唐骞祎,等. 软件定义网络(SDN)研究进展[J]. 软件学报,2015,26(1):62-81.