TCP 流的宏观平衡性

龚 俭 彭艳兵 杨 望 刘卫江

(东南大学计算机科学与工程学院 南京 210096)

摘 要 TCP 流显式的连接建立和关闭过程决定了完整的 TCP 流的不同类型 TCP 报文之间在数量关系间保持一种宏观平衡性,这种数量间的约束所表现为的宏观平衡性可以用来揭示网络流量行为规律,识别网络流量行为异常的存在,从而可以成为网络安全监测和网络管理的有效手段之一. 本文定义了 TCP 流宏观平衡性的相关测度,根据 TCP 流的到达模型和流长模型建立了测量误差的模型,并以此导出了实际测量模型和判断正常与异常的临界点. 通过实验和仿真对这些测度进行了分析,证明了这些测度和临界点的可用性.

关键词 TCP 流;报文数量的宏观平衡性;测度;测量误差;异常网络行为检测

中图法分类号 TP393

Macroscopical Quantitative Balance of TCP Packets

GONG Jian PENG Yan-Bing YANG Wang LIU Wei-Jiang

(*School of Computer Science and Engineering, Southeast University, Nanjing* 210096)

Abstract The explicit TCP connection opening and closing procedures keep the different packet quantity of a completed TCP flows at a macroscopical balance. This kind of quantitative restraint can be used to discovery the behavior rules of TCP traffics, and to detect the existence of abnormal behavior, which can be deployed as an efficient method in network security monitoring in the network management. Some metrics are proposed in this paper to describe the macroscopical TCP packets quantitative balance of TCP flows, and the measuring errors are modeled by the arrival distribution model and the length distribution model of TCP flow. The thresholds of those metrics to judge the normal and abnormal TCP behaviors are discussed under the measuring error model. The emulation and simulations shows the feasibility of these metrics as well as their thresholds.

Keywords TCP flow; macroscopical quantitative balance of packets; metric; measuring error; threshold; detection of abnormal network behavior

1 引 言

在 Internet 中,TCP 流的比重占了所有网络流量中的大部分[1],使得 TCP 流的行为在很大程度主导了网络流量的行为,因此成为网络行为学的重要研究内容. 本文所讨论的 TCP 流对应一个 TCP 连接,其报文交互过程必须遵循 TCP 协议,这意味着在 TCP 连接的建立与拆除过程中不同类型报文间存在数量的约束关系. 定义完整的 TCP 流是一个遵从 TCP 协议而正常建立和正常终止的 TCP 连接,于是不完整的 TCP 连接则表现为 TCP 协议交互过

收稿日期:2006-04-05;修改稿收到日期:2006- - . 本课题得到国家"九七三"重点基础研究发展规划项目基金(2003CB314804)资助.
龚 俭,男,1957 年生,博士,教授,博士生导师,主要研究领域为网络行为学,网络安全. 彭艳兵,男,1975 年生,博士研究生,主要研究领域为网络行为学、网络安全. E-mail: ybpeng@njnet.edu.cn. 杨 望,男,1979 年生,博士研究生,主要研究领域为入侵检测系统的评估. 刘卫江,男,1969 年生,博士后,教授,主要研究领域为网络抽样.

程的不完整,即在报文的数量平衡关系上会表现出异于完整 TCP 连接的特征.当不完整 TCP 流在全部 TCP 流里所占比例比较大的时候,通常反映出网络本身或网络应用的异常.比如 DDoS 攻击发生的时候,没有完成三次握手的短 TCP 连接的比重会明显增加.从宏观角度看,网络中 TCP 流的各种控制报文间的数量约束关系称为 TCP 报文数量的宏观平衡性,简称为 TCP 流的宏观平衡性.

在宏观尺度上评价 TCP 流行为问题的一个重要研究方面是无状态的异常检测,其在检测过程中不维护 TCP 连接和五元组的详细状态,应用报文的统计规律来检测网络里是否存在攻击行为[2,3].Bykova 等人[4,5]基于网络中异常报文的统计来检测是否存在异常和入侵,使用了关于 TCP 异常报文的数量统计,但只分析了一些异常的 TCP 报文,使得该方法只能发现少数的恶意报文,如 TCP 源端口与宿端口号相同的报文等.

Wang 等[6]使用基于 TCP 的 SYN/FIN(RST)报文对的比例来检测 SYN Flooding 攻击,基于简化的协议交互自动机,使用叶节点路由器来监视网络里的 SYN 报文与 SYN+ACK 报文的比例[3],以此推断网络的 SYN Flooding DDoS 攻击的源.Ohsita 等[7]使用了 TCP 三次握手的信息,和各类 TCP 报文的统计信息来推断 DDoS 的存在,但需要使用五元组来获得全部流的数量;同时因选取不够合理的时间粒度,使得检测结果易被 TCP 的分形行为的噪声所误导.Qiu 等[2]利用半开连接队列来检测 SYN Flooding,对 $SYN\_ONLY/SYN+ACK/RST+ACK$ 的数量关系进行了探讨.Arlitt 等人[8]以 HTTP 为依托研究 RST 报文产生的机理和行为,发现各种不同的浏览器和服务器的行为会导致各种 TCP 报文数量的不一致,Windows XP 和 IE6.0 使用 RST 中止连接的比例高达 60.4%.因此在考虑 TCP 连接完整性时,需要同时考虑这种浏览器等的异常行为导致的 TCP 连接的非正常结束.

文献[9]提出了使用利用数量比例定义的测度用于 TCP 异常检测的例子,利用了 TCP 各种控制报文数量间的约束,但是对于时间粒度的选取和异常取值的经验范围等方面也缺乏足够的考虑.

上述研究表明,TCP 各种控制报文数量间的约束关系并不是任何时间粒度下都成立的,需要更为细致的讨论.文献[10]揭示 TCP 流量行为服从多重分形,不同时间粒度下 TCP 的流量行为非常复杂.各种报文间数量关系受实际环境的影响,建立正常/异常阈值的模型是实践中需要解决的问题.

本文根据 RFC793[11],以 TCP 连接建立和关闭报文的唯一性入手,为 TCP 流的宏观平衡性提出系统的测度,并建立了测量误差模型来 TCP 宏观平衡性的测量特征.这些测度可以在大规模网络环境中,在不维护每个 TCP 连接具体状态的情况下对 TCP 连接的健康性进行宏观评估.这些测度的变化规律和趋势揭示了网络行为变化和行为异常范围,成为反映网络可用性和健康状况变化的重要指标,对于网络管理与规划和网络安全检测与攻击预警都有积极的作用.

本文第 2 节讨论了可用于描述 TCP 宏观平衡性的几种报文约束关系及其测量对象;第 3 节定义了 TCP 宏观平衡性的一般测度和面向异常检测的特殊测度,并论述了这些测度存在的必要性;第 4 节对影响测量的时间等因素进行建模,通过使用系统仿真、实测的网络流量数据、流量发生器发生的流量来验证了这些测度的可用性;第 5 节导出了判断 TCP 正常和异常的临界值;第 6 节总结了全文并对进一步的工作进行了展望.

2 基于协议的报文数量约束

根据 TCP 协议,对于 N 条完整的 TCP 流,在时间 Δt 内,各类 TCP 报文的数量存在如下约束:

$$N_{SYN} = N_{SYN+ACK} \tag{1}$$

$$N_{SYN+ACK} = N_{Flow} \tag{2}$$

$$N_{RST}/N_{SYN+ACK} \approx 0 \tag{3}$$

$$N_{FIN}/N_{SYN+ACK} = 2 \tag{4}$$

其中 N_{xxx} 表示含带 xxx 标记的 TCP 报头的报文数量.式(1)表示如果在信道里 TCP 的前二次握手都是完整的,此时 TCP 第一次握手和第二次握手发生过程中报文数量是平衡的.式(2)表明,如果前两次握手不完整,则 $SYN+ACK$ 的报文数量更加接近于 TCP 流的数量.式(3)表明,由于 RST 主要是用于处理异常报文,在正常的非拥塞信道里,其数量应该是趋向于为 0 的.式(4)表示完整的 TCP 流的 FIN 报文的数量应该是 $SYN+ACK$ 报文数量的 2 倍.不同于 IP 流,TCP 流宏观平衡性测度的重点不在于与性能有关的流长和吞吐量等性质,而是关注 TCP 流的完整性,即流是否正常地建立和拆除.上述公式里没有出现 ACK 报文的数量关系,因为 ACK 报文的数量不受流数量的约束.

实际上测量时如果维护完整流的信息,其操作

的复杂性将难以接受. 对于实际网络里流的研究，通常需要取一个充分长的时间粒度 Δt 来进行，在这段时间内，完整的 TCP 连接数与不完整的 TCP 连接数的比值充分大，则可以认为 TCP 流接近于完整的. 如果使用 ΔN_x 来表示在 Δt 内某种报文的数量，由于不完整连接的存在，则上述公式可转化为

$$\Delta N_{SYN+ACK} \leq \Delta N_{SYN\_ONLY} \tag{5}$$

$$\Delta N_{SYN+ACK} \approx \Delta N_{\text{flow}} \tag{6}$$

$$\Delta N_{RST} / \Delta N_{SYN+ACK} \approx 0 \tag{7}$$

$$\Delta N_{FIN} / 2\Delta N_{SYN+ACK} \approx 1 \tag{8}$$

上述公式反映了理想情况下 TCP 连接三个阶段中 TCP 的报文类型和数量的宏观平衡性约束. 由于 TCP 协议并不禁止无效连接请求的发出，所以式(5)中使用的是“$\leq$”，这也是下面定义 TCP 流首报文的初衷.

由于实际网络里可能存在服务质量问题和恶意的用户行为，导致网络存在不定数量的不完整 TCP 连接，实际的 TCP 宏观平衡性测度会更为复杂. 但是上述理想条件下的宏观平衡性约束方程可以作为一个基准来量化实际网络的宏观平衡性. 而选择多大的时间粒度作为测量的周期，对于测量误差的影响非常大. TCP 宏观平衡性的测量误差与时间粒度的选取、TCP 流到达分布、流长分布和流内报文延时密切相关.

为了有效地获得 TCP 宏观平衡性的测度值，在主干网络不能也不必等同地处理所有观测到的报文，需要选择网络中的测量对象的参考基准. 对于完整的 TCP 流，SYN 报文、$SYN+ACK$ 报文、三次握手的 ACK 报文(本文简称 ACK_3)和任意方向的 $FIN+ACK$ 报文等报文的数量均可代表完整 TCP 流的数量. 但是由于失效连接、空连接、链路故障、病态路由以及恶意用户行为等现象的存在，使得在真实网络里上述报文的数量与完整的 TCP 流的数量产生偏差. SYN_ONLY 报文[①]只携带了单个主机的真实信息，因此有

定义 1.　定义响应 TCP 连接发起方的 SYN_ONLY 而返回的 $SYN+ACK$ 报文为 TCP 流的首报文.

TCP 流首报文由于其同时携带了连接双方的信息，如五元组、两个方向的顺序号等，所以其数量比 SYN 报文更接近 TCP 流的数量. 与 ACK_3 相比，在不维护每个 TCP 流状态的情况下，ACK_3 与后面确认数据的 ACK 报文难以区分，使得单纯计数得来的 ACK 数量难以作为 TCP 流的数量. 对于 $FIN+ACK$ 报文的数量，由于 TCP 流的长度分布模型服从重尾分布，使得在单位时间内的其波动比较大，也不能很好地代表 TCP 流数量. TCP 流的首报文将作为测量对象的参考.

3　TCP 宏观平衡性测度

3.1　理想情况下的 TCP 宏观平衡性的测度

在选定了测量的参考报文以后，就可以考虑 TCP 控制报文之间的对应关系. 对于 TCP 的第一次握手和第二次握手，对应于 SYN_ONLY 和 $SYN+ACK$ 报文，根据式(5)可定义如下的测度.

定义 2.　请求盈余度(Shaking-Response Rate，SRR)$=\Delta_{SYN\_ONLY} / \Delta N_{SYN+ACK}$.

它反映了连接请求与响应之间的比例关系. 根据上一节的分析可知其应该在 1 附近. 一般由于空连接、扫描、循环路由等的存在，SYN_ONLY 报文得不到充分的响应，使得这个值远大于 1，使得 TCP 连接建立请求显得盈余，对相应报头的进一步观察可以发现与异常的用户行为或网络状态相关的节点. 文献[6]把这个比例的倒数作为异常行为挖掘的触发器. 根据上一节的分析，$SRR=1$ 时，TCP 的头两次握手很完整.

由于在不维护每流信息的宏观条件下，ACK_3 和数据传输的 ACK 报文很难区分，所以这里没有定义 TCP 第三次握手相关的测度. 根据式(8)，有

定义 3.　连接完整度(Connection Completeness Rate，CCR)：$\Delta N_{FIN+ACK} / 2\Delta N_{SYN+ACK}$.

它反映了 TCP 正常拆链的比率，即连接结束过程与理想情况的差距. 根据上一节的讨论，理想情况下它的比值为 1. 而在存在异常行为如扫描和 DDoS 攻击的时候会偏离其理想值 1.

3.2　参考报文的时间序列测度模型

流到达模型表明 TCP 流到达率在短时间粒度内可以认为是服从泊松分布的[12~14]，因此可以为作为基准的 $SYN+ACK$ 报文定义一个测度来以其自身的历史作为自己的量度.

定义 4.　响应平稳度(Shaking Response Variable，SRV)$=\Delta N_{SYN+ACK_i} / \Delta N_{SYN+ACK_{i-1}}$ $\Delta N_{SYN+ACK_{i-1}}$ 为这个测度在上一个时间粒度内的 $SYN+ACK$ 报文数. 这个测度反映了 TCP 流的数量或者 $SYN+$

① 即不带 ACK 标志的 SYN 报文，以区别 $SYN+ACK$ 报文，RST_ONLY 和 FIN_ONLY 报文的含义类似.

ACK 报文数量的变化是否平稳，可用来发现 TCP 流的突发现象，特别是大规模 TCP 异常，如基于 $SYN+ACK$ 的 DDOS 攻击. 对于泊松分布的流到达来说，有 $(SRV)=1$，$VAR(SRV)=VAR(flow)$，其中 $flow$ 代表时间粒度 Δt 内流的数量. 这个测度在发生大规模 TCP 流异常的时候，可以作为综合判定异常类型的参考. 上述几个测度期望为 1.

3.3 异常报文的测度模型

定义 5. 异常报文比例(异常报文比例，Abnormal Packet Rate，APR)$=\Delta N_{RST}/\Delta N_{syn+ack}$.

根据 RFC793 和式(7)，在非拥塞、正常交互的情况下，RST 报文在单流应答中出现的概率很小，因此 $APR\approx 0$，也可以认为理想情况下 $APR\to 0$.

对于异常到达的报文，TCP 协议会以 3 种不同的行为方式对该报文进行应答：

(1) 若该报文带有 RST 标志，不进行任何应答；

(2) 如果该报文带有 ACK，且已有 TCB 对应，协议栈应答一个带有 ACK 的 RST 报文；如果没有 TCB 对应，应答一个不带 ACK 标志的 RST 报文；

(3) 如果该异常报文不带 ACK 标志，则应答一个带有 ACK 标志的 RST 报文.

因此异常报文比例 APR 可以分解为两种不同的行为，对于区别异常是 DDoS 还是扫描非常有用.

定义 6. ACK 拒绝率(ACK rejected Rate，ARR)：$\Delta N_{RST\_ONLY}/\Delta N_{SYN+ACK}$.

RST_ONLY 报文即不带 ACK 标志的 RST 报文. 这个测度主要反映了 TCP 连接建立失败的情况. 正常情况下，ACK 报文异常的数量很小，因此 ARR 趋近于 0；但在网络出现异常或者有攻击行为的时候，可能导致该报文数量的异常. 因此这个测度可以用来对网络的异常进行量度.

定义 7. 非 ACK 拒绝率(Non-ACK Rejected Rate，NARR)：$\Delta N_{RST+ACK}/\Delta N_{SYN+ACK}$.

反映了不带 ACK 的异常到达报文的比例，主要集中反映在连接建立过程中的异常情况.

上述两种测度的定义是相辅相成的，在基于非 RST 报文的扫描过程中，回应的 RST 报文会大量出现，带有 ACK 的扫描报文的回应是 RST_ONLY 报文，不带 ACK 标字的扫描报文的回复是 $RST+ACK$ 报文；基于 RST 的扫描过程中，少量的 $ICMP$ 报文是回应的报文主流；而在大规模 DDoS 攻击中，两种报文的比例会比较接近，而且报文的数量会比平常多很多倍. 综合考虑这两种报文的测度可以很好地确定异常的种类，特别是其他报文的数量也异常的时候. 这两种测度的值越小越好.

RFC793 规定了对于有 TCB 的带 ACK 的 TCP 报文需以带有 ACK 的报文进行应答，根据在网络里观察的情况来看，无 ACK 标志的 FIN 报文很少，因此为这种不应在正常交互中出现的 FIN_ONLY 报文也定义一个测度，用于检测这种异常.

定义 8. FIN 鬼影测度(FIN Ghost Rate，FGR)：$\Delta N_{FIN\_ONLY}/\Delta N_{SYN+ACK}$.

它反映正常交互中不可能出现的 FIN_ONLY 报文出现在网络里的情况，因此 $FGR=0$. 它的比值越大，反映网络里存在的异常规模越大. FGR 的取值为 $FGR=0$，大于 1 即发生了异常.

4 宏观平衡性测度可测性

4.1 时间粒度对上述平衡性测度的影响

确定时间粒度的大小对于宏观平衡性测度的测量误差影响密切；而时间粒度的选择，又与 TCP 流到达率、流长的分布及流内报文到达速率 D 相关. 然而无论怎样选择时间粒度的大小，总有流被测量边界截断，因此时间粒度的确定要使得观测到的完整流的数量比重充分大，需要对被截断的流对测量的影响建模.

对于整个 TCP 交互过程，n_3 代表在第 i 个时间粒度 Δt_i 内产生的流的数量，等价于 Δt_i 内的流头标志的数量，则在 Δt_i 内讨论流的情况应当分成无始有终(对应于图 1 的粗线段的数量 n_1)，有始有终(对应于图 1 的黑色线段的数量 n_3-n_2)，有始无终(对应于图 1 的双线段的数量 n_2)，无始无终(对应于图 1 的黑色虚线段的数量 n_4)等 4 种情况. 无始无终的流 n_4 对于误差没有影响，因此不予考虑.

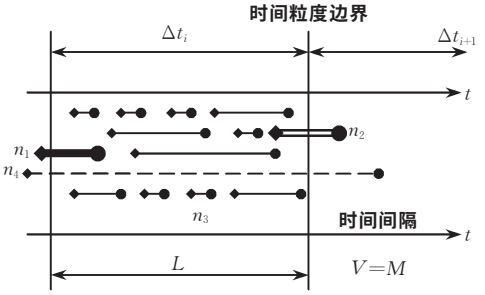

图 1 切分点附近 TCP 流的分布与被切分的情况

若以流的头/尾标志的数量来表示上述变量，则有 n_1 表示没有头的尾标志的数量，n_2 表示没有结束的尾标志的数量，n_3 表示全部的头标志. 再引入几个

中间变量，n_0 表示完整流的尾标志的数量，n_0' 表示完整流的头标志的数量；n_2' 表示没有尾的头标志的数量，这些中间变量在推导过程中有助于理解测量误差，但是在推导出测量误差后，这些中间变量将消失. 显然有 $n_0=n_0$，$n_2=n_2'$ 和 $n_3=n_0'+n_2'$. 由于没有考虑头标志报文和尾标志报文在信道里的方向，因此上述模型本身是没有方向的，但是已能保证检测的精度；如果考虑方向能够更精确，但需要维护的信息量加倍.

第 2 节提出的测度一般为两个报文的比例，如 SRR 表示 SYN_ONLY 和 $SYN+ACK$ 报文的比例，而实际上测量的时候，测量时间的边界会切断一些流，使得在本时间粒度 Δt_i 里面会有一些不完整的连接，导致了测量的误差. 分析如下. 测度 A 的期望为 Δt_i 里面完整流的头尾标志之比：

$$A=(n_0+n_2)/(n_0'+n_2')=1 \tag{9}$$

但实际上测量的测度 B 是 Δt_i 里面所有流的头尾标志之比：

$$B=(n_1+n_0)/(n_0'+n_2') \tag{10}$$

B 里含有不完整连接的数量 n_1，同时还缺少不完整连接数 n_2. 因此，这种差别会导致测量值与期望值之间存在误差 ε：

$$\varepsilon=|A-B|=|(n_0+n_2)-(n_1+n_0)|/(n_0'+n_2')$$
$$=|n_1-n_2|/(n_0'+n_2')=|n_1-n_2|/n_3 \tag{11}$$

因为 n_1，n_2 和流长分布和流内报文延时有关，而 n_3 和 Δt 的大小有关. 由于最终的测量误差 ε 与 n_1，n_2 和 n_3 有关，与 n_0，n_0' 和 n_2' 没有显式关系，因此图 1 只需标注了 n_1，n_2 和 n_3.

当 ε 充分小，或者误差 ε 是可以接受的，那么就可以用 B 代替 A 进行测量，并可保证第 2 节里式(1)～式(4)的推广为可以测量的式(5)～式(8).

对于 SRR 等测度，n_3 就是 SYN_ACK 报文. 下面先讨论长流的误差模型，主要是对完整的 TCP 流进行建模.

设定 TCP 流的到达服从泊淞分布，而流的长度服从 Pareto 分布[12～14]，考虑完整的 TCP 流的所有报文，由于 TCP 流的持续长度服从重尾分布，参数为 $P(x)=ra^rx^{-(r+1)}I_{[a,\infty]}(x)$，$I_{[a,\infty]}(x)$ 为符号函数；设流的到达率的均值为 M，Δt 内新产生的流的数量为 $C=M\times\Delta t$，未结束的流的数量可以通过重尾分布推导出来的. 假设流内报文间延时的均值为 D，长度为 L 的流刚好持续了 Δt 秒，则有 $\Delta t=L\times D$. 我们可以通过流长的重尾分布推导出流长度在 $[L,\infty)$ 内的流数量分布：

由于长度大于 L 的流在 Δt 时间内是不能结束的，作为一种简化模型，这里把误差估计模型理想化为，所有该时间粒度内产生的 TCP 流都在测量开始的时刻产生，在 Δt 时间内产生的长度大于 L 的流的数量的极大似然估计为

$$\hat{n}_1=\hat{n}_2=C\int_L^{\infty}ra^rx^{-(r+1)}I_{[a,\infty)}(x)\mathrm{d}x \tag{12}$$

根据式(11)，有

$$\hat{\varepsilon}\leq\frac{\hat{n}_2+\hat{n}_1}{\hat{n}_3}=2C\int_L^{\infty}ra^rx^{-(r+1)}\mathrm{d}x/C$$
$$=2\left(\frac{a}{L}\right)^r=2\left(\frac{aD}{\Delta t}\right)^r=\mu \tag{13}$$

这样的简化模型会产生误差，导致理论测量误差比实际测量误差小，但非常接近，后面将会有仿真分析. 这个模型会对 CCR，APR，ARR 和 NARR 参数影响.

为保证测量误差 μ 是可以接受的，$\Delta t\geq aD/(\mu/2)^{1/r}$. 如果知道网络里 TCP 流的平均报文延时 D，就可以推导出在给定测量误差 μ 时 Δt 取值的下限. ElAarag，Bassiouni 等[15] 指出，非理想情况下的，即使含有无线连接，网络里各种 TCP 实现的报文延迟 D 小于 0.46s. 因此可以得出，在允许的测量误差 $\mu=0.01$，$r=1$ 的情况下只需 $\Delta t>200$s 即可. 而 Internet 里的流长度服从 Pareto 分布，其指数参数 r 一般大于 0.5，在 1.1 左右[16,21]，达到指定的测量精度的可以选取得更短.

如果只考虑 TCP 流三次握手部分，则模型可简化为短间隔模型，报文延时 D 为 RTT，且 $\hat{n}_1=RTT_{\mathrm{mean}}\times M$，$\hat{n}_2=RTT_{\mathrm{mean}}\times M$，$\hat{n}_2=\Delta t\times M$，在 Δt 内的误差与流到达率 M 没有关系：

$$\hat{\varepsilon}\leq\frac{\hat{n}_2+\hat{n}_1}{\hat{n}_3}=\frac{2RTT_{\mathrm{mean}}}{\Delta t}=\sigma \tag{14}$$

显然只需要 $\Delta t>40\times RTT$ 即可满足 $\sigma=0.05$，本文就以 $\sigma=0.05$ 作为可以接受的测量误差的上限. Jiang 等在文献[17]提出使用被动测量方法估计 RTT 的值，他们对 RTT 进行估计，发现 95% 的连接的 RTT 小于 0.5s. Lan 等人[18] 发现，DDoS 和蠕虫等异常行为的 RTT 也符合这个规律. 因此在短间隔模型中选取 $\Delta t=20$s 即可达到 $=0.05$ 的精度，选取 $\Delta t=300$s 应可以达到比 $\sigma=0.01$ 的更好的精度.

这样测量误差 μ 和 σ 就与流内报文延时均值 D，RTT，流长分布参数、Δt 等参数建立了联系. 上述分析表明，上节定义的测度的计算是可行的，在选

取合适的测量时间粒度后具备可操作性和相应的精度.

4.2 时间粒度选择的验证

下面对上节测量误差模型进行仿真检验,由于短间隔模型较简单,这里只对长流模型进行仿真.

4.2.1 长流的测量误差仿真分析

根据第 4.1 节的分析,流的到达模型为泊淞分布,流的长度符合 Pareto 分布,流内报文到达率服从均匀分布.以此模型,我们可以对完整 TCP 流的 TCP 宏观平衡性进行理论仿真.

下面以 $M=100$flow/s, $D=0.5$s, Pareto 的指数 $r=0.5\sim3$, $a=1$ 时间粒度 10～600s,对 TCP 宏观平衡性的测量误差公式(13)进行仿真.由于仿真过程只需记录流的首报文产生的时间和流结束的时间,非常简单,这里使用 GNU Scientific Library (GSL v1.4)①作为随机函数发生器,开发了一个程序来进行,没有使用 Opnet 和 NS2 之类的仿真工具.

对 Δt 与 μ 的关系进行仿真,模拟 20 次连续测量,不同时间粒度的 CCR 测度的曲线如图 2 所示.

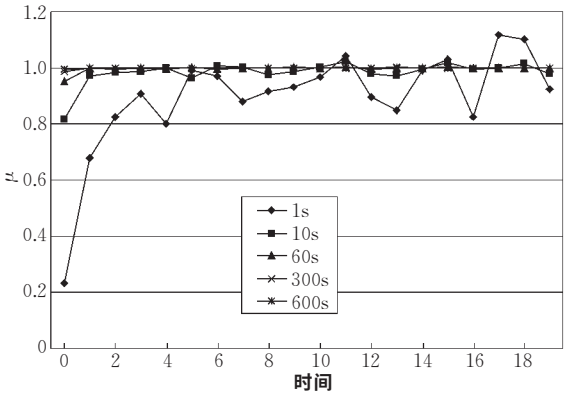

图 2 不同时间粒度下 20 个连续时间粒度的 CCR

从图 2 可以看出不同的时间粒度 Δt 下与 CCR 的关系,CCR 的期望值为 1,而图中 CCR 的波动的程度随着时间粒度的增加而减小,测量的误差 μ 也随之减小.

对测量误差 μ 的仿真分析如图 3 所示.图中的理论值是根据测量误差公式(13)计算得来的,而仿真值则为实验中计算得来.

显然从图 3 可以看出,$r=1.0$ 的时候,误差与接近于对数线性关系,对其拟合,得到下面的公式:

$$\log(\mu) = -0.93 \times \log(\Delta t) - 0.65$$

显然 $r=0.93$,非常接近于 1,表明仿真的结果与式(13)的结果吻合.

上述结果表明在 $r=1.0$ 的时候,不同的时间粒度下的测量误差的间的对数线性关系是成立的,在 $\Delta t=60\sim300$s 的时候,仿真的测量误差和理论值均低于 0.01.

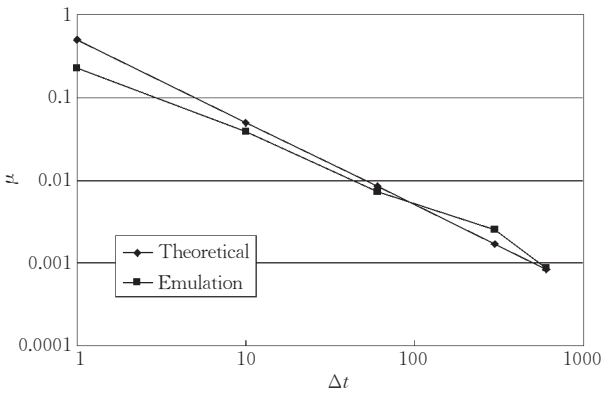

图 3 仿真时间粒度 Δt 下与 μ 的关系

下面对 $\Delta t=10$s 的时候,不同 r 的关系进行模拟.图 4 是 r 取值变化时连续 20 个时间粒度下 CCR 波动曲线,$\Delta t=10$s.图 4 表明随着 r 的变大,CCR 的测量值的方差越小,并且越贴近其期望值 1.

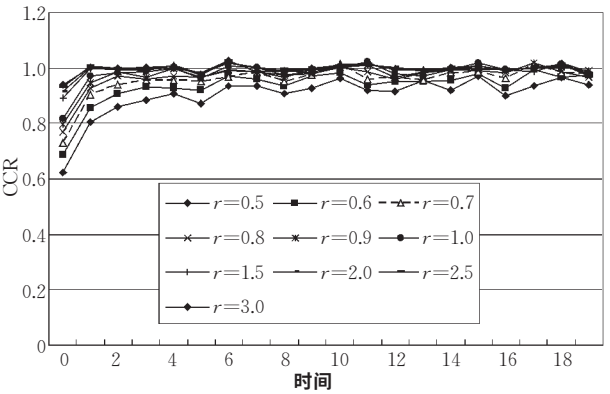

图 4 不同 r 取值时连续 20 个时间粒度的 CCR

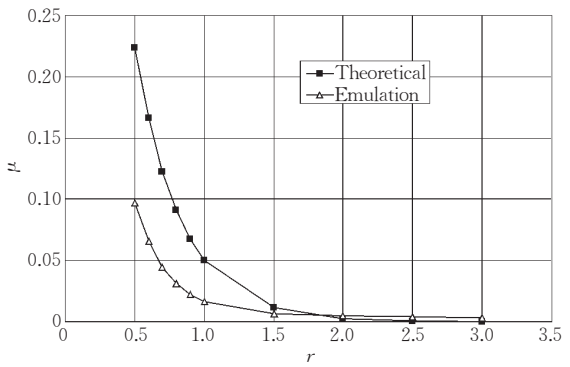

图 5 测量误差 μ 与 Pareto 分布参数 r 的关系

下面对 r 与 μ 间的关系进行讨论,$\Delta t=10$s, $a=1$, $D=0.5$,如图 5 所示.从图中可以看出,随着 r 的

① http://www.gnu.org/software/gsl/

增大，不论是理论测量误差还是仿真的测量误差 μ 均逐渐减小，在 $r>0.8$ 以后，理论测量误差 μ 与仿真的测量误差都小于 0.1，并且彼此比较接近. 而由于真实流量的 r 一般在 1.1 左右[16]，因此使用第 4.1 节模型的测量误差也是可以接受的.

Theoretical：在第 4.1 节中的模型；Emulation：仿真结果

通过对 Δt 与 μ 的关系和 r 与 μ 关系的仿真表明，第 4.1 节中的测量误差模型符合全流的到达模型.

4.2.2　短间隔模型和测量误差

为了验证上节的模型，下面使用五元组和 TCP 交互严格匹配的方法，对网络里的 TCP 流进行了过滤，略掉不完整的 TCP 流，只保留完整 TCP 流的控制报文，如三次握手的报文和 TCP 连接关闭的报文. 原始报文流采集自 CERNET 华东北地区网络里的江苏省网边界和国家主干的互联信道中 2004 年某一天的全天数据，总报文数为 1.52×10^{10} 个，字节数为 7.01×10^{12} B，TCP 报文数为 1.40×10^{10} 个，平均带宽 587Mbps. 将过滤掉不完整 TCP 连接报文的数据称为 Trace1，可以用于时间粒度和测量误差的关系验证.

图 6 显示了以请求盈余度 SRR 为例的分析结果. 讨论的时间粒度从 1～600s. 为了清晰地显示细节，图 6 只显示了两个小时的数据. 显然测量值围绕期望值波动，时间粒度越小则波动得越厉害，而在测量的时间粒度大于 60s 以后，测量的 SRR 只是在 1 附近轻微波动，幅度在 0.05 以内.

当选用 1s 作为时间粒度的时候，此时时间粒度 Δt 与 RTT 的均值接近，会导致较大的误差. 从图 6 可以看出，1s 作为时间粒度的时候，其测量值波动得很厉害，表明此时的测量模型已经失效. 此时若使用 $SYN/SYN+ACK$ 比例作为异常检测比如 DDoS 检测的阈值[7]，可能有 40%的误差.

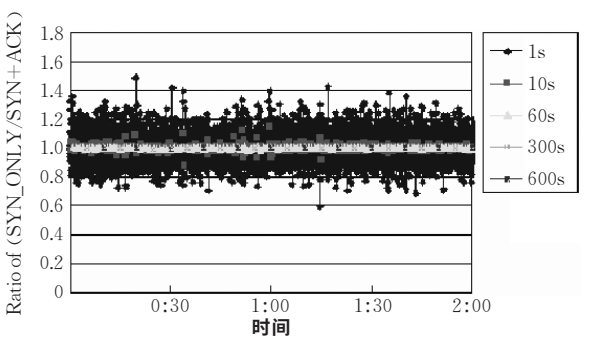

图 6　时间粒度的选择与 SRR 的测量波动

从图 6 还可以看出，随着时间粒度的增大，$SYN/SYN+ACK$ 的比例也越来越接近其期望值 1，当时间粒度在 60s 以上时，这个比例基本上只是贴近 1 附近轻微波动了. 为了描述上述波动的偏差，对各时间粒度下的方差进行了比较，这里的方差表示测量值波动的程度，如图 7 所示.

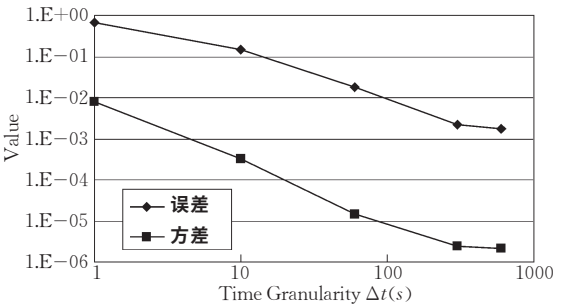

图 7　时间粒度的选取与 SRR 测量波动、测量方差

从图 7 可以看出，测量值波动的方差与时间粒度间存在近似的对数线性关系，拟合如下

$$\ln(\Delta t) = -0.7078\times\ln(VAR) - 3.43767 \tag{15}$$

测量误差 σ 和时间粒度的选择也有相似的关系式：

$$\ln(\Delta t) = -1.00149\times\ln(\sigma) - 0.40997 \tag{16}$$

这个对数线性关系的斜率很接近 1，与式(14)很接近，表明上述模型是符合实际情况的，前面的理论分析是正确的.

从图 6 可以看出，对于短的交互，比如 TCP 的三次握手而言，时间粒度的选择一般可以在 60s 以上，测量模型的抖动幅度已经小于 0.05，因而建议实际测量的过程中采用 60s 以上的时间粒度. 但是如果测量结果还要为其他活动提供依据时，比如作为大规模异常检测的触发机制时，时间粒度不能选择得太大，一般不超过 600s；作为一个折衷，可以选择 300s 作为一个合适的时间粒度.

4.2.3　长流的时间粒度选择和测量精度的关系

对于 TCP 流的全部报文，由于其流到达模型、流持续模型和流内报文到达模型的影响，使得分析更加复杂. 从整体上看，由于流的平均报文数比短间隔模型如三次握手的报文数要大，因此在相同时间粒度下的测量误差要比三次握手的大很多，并且测量均值也与期望值有较大的偏离. 本小节和上小节的实验条件一致.

从图 8 中可以看出，选择大的时间粒度明显有助于减少测量值的波动. 从图 9 还可以看出，测量时间粒度的选择和测量误差存在着明显的对数线性关系. 根据方差的数据可以拟合出数据的线性关系，有

$$\ln(\Delta t) = -1.566 \times \ln(VAR) - 3.662 \quad (17)$$

对于测量误差与时间粒度有

$$\ln(\Delta t) = -1.445 \times \ln(\mu) + 2.731 \quad (18)$$

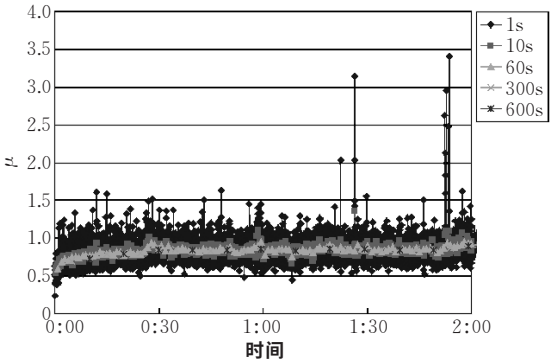

图 8 时间粒度选择和 *CCR* 的测量波动

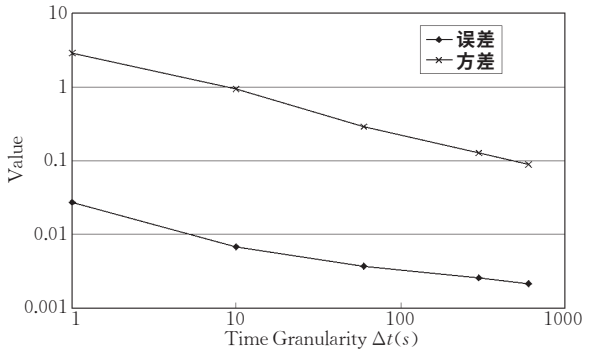

图 9 时间粒度的选取与 *CCR* 的测量误差、方差

显然,由此推出的 Pareto 分布的指数 $r=0.692$.

在选择 300s 的时间粒度时,实际上测量误差还有 9.5%,这与图中曲线整体的波动有关,由于 TCP 流是分形的[10],这种分形噪声的干扰使得其测量误差较大,特别是突发流的存在[19]. 图 8 中 60s 的波动还比较明显,300s 与 600s 的波动已经不明显了,表明 300s 已经是一个可以接受的时间粒度了. 但是由于发生异常的 TCP 流如 DDoS、扫描、病态路由都是一些很短的交互,因此这个误差不会影响对这些 TCP 宏观异常的检测.

由于流到达模型、持续模型和流内报文到达率的复杂性,也使得全流的测度与期望值有差距. 不同时间粒度的测量均值在 0.67～0.68 之间,而期望值为 1,表明还存在着其他没有考虑进来的因素,需要进行更深入的研究. 后面的理论分析表明,实际正常的流 $FIN+ACK/SYN+ACK$ 测度的最大偏离范围为[0.142,1.858],现在的波动,只要误差在这个范围内即可,300s 显然可以满足这个要求.

4.3 对于流到达模型和流长分布的模拟

下面使用 CERNET 华东北地区网络中心开发的 TCP 流量发生器 AOLES,按照指定的流到达模型和持续模型来生成所需要的报文 trace2,再对 Trace 进行相应的分析,得到测量误差模型. TCP 流的到达服从泊淞分布,其流到达强度为 20flows/s;TCP 的流内报文到达率服从指数分布;流长度的分布符合重尾分布,分布参数将在一定的范围里进行考查;RTT/D 均值为 0.02s,trace2 包含几个不同 Pareto 分布参数的实验数据集,其持续时间都为 10min. 由于 AOLES 系统的限制,低于 0.6 的 Pareto 参数的 TCP 流长度分布中长流往往会超过文件系统的限制,因此这里选择 Pareto 参数从 0.7～3.0.

对于 Pareto 参数 $r=1$ 的完整 TCP 流量而言,图 10 的 SRR 的波动幅度与测量时间粒度显然存在联系:时间粒度越小,测量误差越大. 图 11 给出了 $r=1.0$ 的完整 TCP 流的 *CCR* 波动曲线,也有同样的规律. 而对于 Pareto 参数 r 不同的完整 TCP 流量而言,1s 的时间粒度导致 *CCR* 的波动太大,基本没有正确性可言,如图 12 所示.

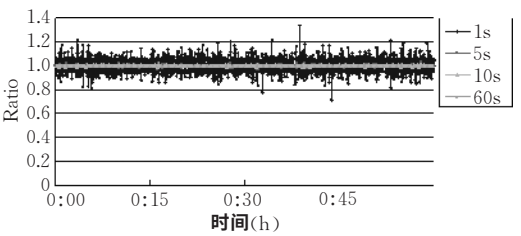

图 10 $r=1.0$ 时完整 TCP 流 *SRR* 的波动

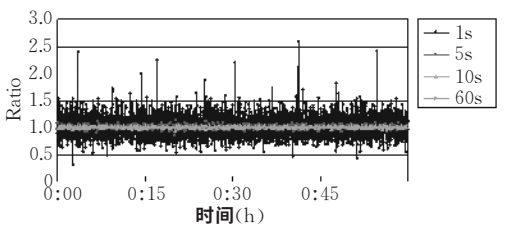

图 11 $r=1.0$ 时完整 TCP 流 *CCR* 的波动

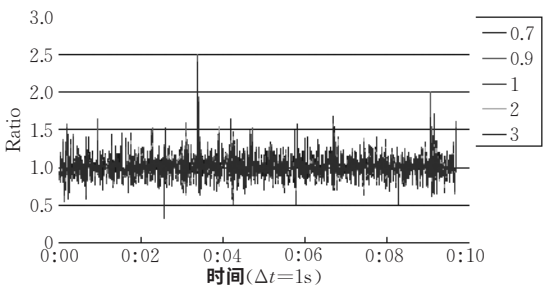

图 12 $\Delta t=1s$, $r=0.7\sim3$ 时完整 TCP 流 *CCR* 的波动

上面的分析表明,选择合理的测量时间粒度,对于测量而言会有很多好处,充分大的时间粒度,会使

得测量模型有较高的精度和稳定性；太小的时间粒度，如 1s，不管是上节的仿真还是本节的模拟实验都表明测量误差非常大，所以文献[6,7]使用太短的时间粒度测量 SRR 来检测 DDoS 攻击是一种错误的做法.

在复杂的 TCP 流持续模型下，关于流结束的测度模型应该选择稍微大一些的时间粒度，在实际使用中可以以比短间隔模型大一些的时间粒度. 如果流关闭的测度模型满足了测量精度要求，短间隔测度模型的测量精度会更好. 根据误差模型式(16)对误差 σ 和 μ 进行了描述，如图 13 所示.

对图 13 里的误差曲线进行拟合，得到了下面的两个公式. 这里 r 拟合的值为 1.013，比较接近流量发生器的 Pareto 参数 $r=1.0$.

$$-1.004\ln\Delta t - 1.894 = \ln\sigma \tag{19}$$

$$-1.013\ln\Delta t - 0.7513 = \ln\mu \tag{20}$$

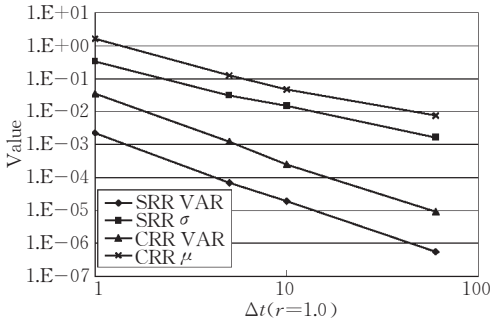

图 13　$r=1.0$ 时 SRR 和 CCR 的测量误差和方差

上述理论分析、实验和仿真还表明，时间粒度 Δt 与本文定义的 TCP 宏观平衡性的测量误差可以使用近似对数线性关系来描述；上节定义的测度的计算是可行的，适宜的时间粒度为 1～10min，并且具备可操作性和相应的精度. 有了测量精度模型，下面将以此全面定义 TCP 宏观平衡性的测度.

5　正常/异常情况的测度值

5.1　临界点取值的计算

网络异常行为的判断依赖于测度基准值以及对其的偏差要求，对 TCP 宏观平衡度测度应用也是如此. 由于超时重传机制的存在，使得 RTT 和流内报文到达率有一个上限 timeout. 文献研究的 64s 超时[20]，由于太长而一般不为各类系统采用，但可把它作为一种极限情况带入误差公式，推导理论取值范围的临界点.

如果试验里采用 5min(300s)作为时间粒度 Δt，而 $RTT \leq timeout = 64$s，令 $RTT = 64$s，对于式(14)

$$\sigma = \frac{2RTT}{\Delta t} = \frac{2\times 64}{300} = 0.427 \tag{21}$$

如果考虑这些可能的误差，则对于测度定义 2 和定义 3 有 $SRR \in [1-\sigma, 1+\sigma]$，即 $SRR \in [0.573, 1.427]$.

对于完整的 TCP 流而言，$D \leq timeout = 64$s，$a=1$，$\Delta t = 300$s. 带入式(13)，得到

$$\mu = 2\left(\frac{\Delta t}{\alpha D}\right)^{-r} = 2\left(\frac{300}{64}\right)^{-r} = 2(4.6875)^{-r} \tag{22}$$

r 取值在 0.5 以上[16,21]，可以得到 $r=0.5$ 时的误差 μ 最大，为 0.924，以其作为衡量测量的极限，可以得到 $CCR \in [1-\mu, 1+\mu]$，即 $CCR \in [0.076, 1.924]$. 对于响应平稳度 SRV，其误差模型其实与 SRR 报文对的接近，只不过这里变成了与自己比较，可以取值的范围为 $SRV \in [0.573, 1.427]$.

对于 APR，由于其一般出现在流的中间的任意位置，其误差应该不大于 CCR 具有的最大误差，由于 $APR \to 0$，$APR > 0$，可以得到 $APR \in [0, 0.924]$. 而对于 ARR 和 $NARR$，虽然这两个测度分裂了，仍然使用相同的取值范围，即 $ARR \in [0, 0.924]$，$NARR \in [0, 0.858]$.

5.2　正常/异常情况下测度的临界范围

在上面给出的测度范围的边界时，报文间延时 D 达到了与超时的最大值相同的地步，这种情况在正常交互中是绝对不可能出现的. 因而这个范围，可以作为 TCP 宏观行为是否正常与异常的分界线. 下面把各种测度的定义和取值范围作一下归纳，时间粒度为 300s，如表 1 所示.

表 1　TCP 宏观平衡性测度的临界取值范围

测度	正常范围	异常范围
SRR	$[0.573, 1.427]$	$[0, 0.573) U (1.427, \infty)$
CCR	$[0.076, 1.924]$	$[0, 0.076) U (1.924, \infty)$
SRV	$[0.573, 1.427]$	$[0, 0.573) U (1.427, \infty)$
APR	$[0, 0.924]$	$(0.924, \infty)$
ARR	$[0, 0.924]$	$(0.924, \infty)$
$NARR$	$[0, 0.924]$	$(0.924, \infty)$
FGR	0	$(0, \infty)$

使用这个表格里的值就可以进行异常行为测量和判断了. 正常情况下各取值都必须在正常范围内，信道里 TCP 的工作状况基本良好. 由于这些范围都是反映了极限情况下 TCP 宏观平衡性，当到达这个极限时一定会存在某种类型的异常，信道里已经有某些服务的功能可用性开始受到考验了，比如对于 SRR 的 RTT 等于超时，所有的 SYN 报文会被重

传,这时网络或服务已经不能正常工作.当发生大规模异常的时候,可能会影响其中的一项或多项测度,而这些大规模异常发生时往往交互的报文很少,可以用短间隔模型来对相关报文进行描述.如基于 TCP 的扫描利用了 TCP 协议的应答机制,一般只有一个交互过程,不会有完整的 TCP 连接;基于 TCP 的 DDoS 攻击一般也只有很短的交互过程,比如 Syn Flooding[3];路由循环发生的时候,SYN_ONLY 可能会因为循环而不能到达目的地,循环使得 SYN_ONLY 报文多次经过循环路径,这两个因素使得 SYN_ONLY 的数量比 $SYN+ACK$ 报文的数量多.利用上述测度和临界点就能检测这些异常行为,而不用维护每流的状态信息.

测度 SRV 能够处理重大事件发生时流量的异常,而使用其它 TCP 到达模型得到的临界范围对短间隔模型的测度影响不大,对长间隔模型测度影响比较大,这里把它作为将来的工作之一.

这些正常/异常的测度临界值可以直接代替文献[9]中的经验值进行更加精确地判断.限于篇幅,对其异常判断的准确性的验证工作,这里不再给出相关的例子.

6 结论与将来的工作

本文从维护 TCP 流的完整性的角度出发,提出 TCP 宏观平衡性的理论、测度体系和测量误差模型,按照 TCP 连接建立、交互和结束三个阶段对 TCP 的宏观平衡性进行评价.对测度误差来源的仿真和模拟表明,这些测度可以用于评估网络里 TCP 连接是否完整.

对于本文定义的 8 个测度和用于判断 TCP 正常/异常工作状态的临界范围,可用于发现长期病态路由循环的存在、快速检测网络里存在的 TCP 异常行为如基于 TCP 协议的大规模扫描、DDoS 攻击行为,为及时、快速消除这些异常打下了基础.这种基于抽象端系统的宏观平衡性模型可以推广到对任意域的断面进行 Tomography 研究.

参 考 文 献

1 Moore D., Voelker G., Savage S.. Inferring Internet denial of service activity. In: Proceedings of the USENIX Security Symposium 2001, Washington DC, 2001, 9～22

2 Qiu X., Hao J., Chen M.. A mechanism to defend SYN flooding attack based on network measurement system. In: Proceedings of the ITRE 2004, London, 2004, 208～212

3 Wang H., Zhang D., Shin K. G.. Change-point monitoring for the detection of DoS attacks. IEEE Transactions on Dependable and Secure Computing, 2004, 1(4): 193～208

4 Bykova M., Ostermann S.. Statistical analysis of malformed packets and their origins in the modern Internet. In: Proceedings of the Internet Measurement Workshop, Marseille, France, 2002, 83～88

5 Bykova M., Ostermann S., Tjaden B.. Detecting network intrusions via a statistical analysis of network packet characteristics. In: Proceedings of the 33rd Southeastern Symposium on System Theory, Athens, Ohio, USA, 2001, 309～314

6 Wang H., Zhang D., Shin K. G.. Detecting SYN flooding attacks. In: Proceedings of the Infocom 2002, New York, 2002, 1530～1539

7 Ohsita Yuichi, Ata Shingo, Murata Shingo. Detecting distributed denial-of-service attacks by analyzing TCP SYN packets statistically. In: Proceedings of the IEEE GlobeCom 2004, Texas USA, 2004,2043～2049

8 Arlitt M., Williamson C.. An analysis of TCP reset behavior on the Internet. ACM SIGCOMM Computer Communications Review, 2005, 35(1): 37～44

9 Gong Jian, Peng Yan-Bing, Yang Wang, Liu Wei-Jiang. Reconstructing the parameter for massive abnormal TCP connections with bloom filter. Journal of Software, 2006, 17(3): 434～444(inChinese)
(龚俭,彭艳兵,杨望,刘卫江. 基于 Bloom Filter 的 TCP 大规模异常参数再现方法.软件学报,2006,17(3): 423～434)

10 Cáceres R., Danzig P. B., Jamin S., Danny J.. Characteristics of wide-area TCP/IP conversations. In: Proceedings of the ACM SIGCOMM'91, Zurich, Switzerland, 1991, 101～112

11 Postel J. RFC793, Transmission Control Protocol, Sep-01-1981

12 Williamson C.. Internet traffic measurement. IEEE Internet Computing, 2001, 5(6): 70～74

13 Nuzman C., Saniee I., Sweldens W., Weiss A.. A compound model for TCP connection arrivals for LAN and WAN applications. ELSEVIER, Computer Networks, 2002, 40: 319. 337

14 Park K., Willinger W.. Self Similar Network Traffic and Performance Evaluation. John Wiley & Sons, Inc. 2000, 534

15 ElAarag H., Bassiouni M.. Performance evaluation of TCP connections in ideal and non-ideal network environments. ELSEVIER, Computer Communications, 2001, 24: 1769～1779

16 Crovella M., Bestavros A.. Self-similarity in world wide web traffic: Evidence and causes. ACM Sigmetrics, 1996, 12(4): 160～169

17 Jiang H., Dovrolis C.. Passive estimation of TCP round trip times. ACM SIGCOMM Computer Communications Review, 2002, 32(3): 75～88

18 Lan K., Hussain A., Dutta D.. Effect of malicious traffic on the network. In: Proceedings of the PAM 2003

19 Leland W., Wilson D.. High time-resolution measurement and analysis of LAN traffic: Implications for LAN interconnection. In: Proceedings of the INFOCOM 1991, Florida, USA, 1991, 1360～1366

20 Claffy K.. Internet traffic characterization[Ph. D. dissertation]. Computer Science and Engineering Department, University of California, San Diego

21 Crovella M. E., Bestavros A.. Self-similarity in world wide Web traffic: Evidence and possible causes. IEEE/ACM Transactions on Networking, 1997, 5(6): 835～846

GONG Jian, born in 1957, Ph. D., professor, Ph. D. supervisor. His research topics involved network management, network security, network behavior etc.

PENG Yan-Bing, born in 1975, Ph. D. candidate. He majored in network behavior and network security.

YANG Wang, born in 1979, Ph. D. candidate. His research interests focus on evaluation of IDSs.

LIU Wei-Jiang, born in 1969, professor. He majored in packet sampling in network.

Background

The project 2003CB314804 studies the dynamic behavior in network. It is a subproject of the National Basic Research Program (also called 973 Program) 2003CB314800 which focuses on the theory for the new generation architecture of Internet. By network measuring, we are involved in disclosing the theory basement of network behavior and expanding its application. The measuring and metrics' theory and application is the major direction of our group, e. g., Reconstructing the Parameter for Massive Abnormal TCP Connections with Bloom Filter by Gong Jian, Peng Yan-Bing, *et. al.* published by Chinese Journal of Software, Vol. 17 No. 3. This paper is a theoretical basical analysis of the network packet amount balance of TCP flows, which can determine the abnormal TCP behavior in the high-speeded backbone by acceptable resource.