第42卷 第8期 西 安 交 通 大 学 学 报 Vol. 42 №8
2008年8月 JOURNAL OF XI′AN JIAOTONG UNIVERSITY Aug. 2008

一种面向网络行为因果关联的攻击检测方法

王志文[1,2,3]，夏秦[1]，李平均[1,3]
（1.西安交通大学电子与信息工程学院，710049，西安；2.西安交通大学智能网络与网络安全教育部重点实验室，710049，西安；3.陕西省天地网技术重点实验室，710049，西安）

摘要：为了能在攻击目标受损之前检测到攻击事件，提出了面向网络行为因果关联的攻击检测方法.该方法基于SNMP管理信息库数据，根据攻击目标的异常行为，首先利用Granger因果关联检验(GCT)从检测变量中挖掘出与异常变量存在整体行为关联的基本攻击变量，然后针对异常行为特征再次利用GCT从基本攻击变量中挖掘出与异常变量存在局部行为关联的攻击变量，最后根据攻击变量和异常变量之间的因果关系，构建面向攻击方检测的攻击关联规则.在Trin00 UDP Flood检测实验中，所提方法成功挖掘出攻击变量udpOutDatagram，取得了满意的检测效果.实验结果表明，该方法能够在攻击方检测到攻击事件，为及时阻止攻击过程向攻击目标进一步扩散提供预警.

关键词：网络行为；攻击检测；因果关联；管理信息库
中图分类号：TP309 **文献标志码**：A **文章编号**：0253-987X(2008)08-0931-05

Approach for Detecting Attack Based on Causality of Network Behavior

WANG Zhiwen, XIA Qin, LI Pingjun
(1. School of Electronics and Information Engineering, Xi′an Jiaotong University, Xi′an 710049, China; 2. Key Laboratory for Intelligent Network and Network Security of Education Ministry, Xi′an Jiaotong University, Xi′an 710049, China; 3. Shaanxi Key Laboratory of Satellite-Terrestrial Network Tech R &. D, Xi′an 710049, China)

Abstract: An SNMP MIB oriented approach based on causality in network behavior is presented in order to detect attack before the security of target is damaged. According to the behavior of an abnormal variable in target, Granger causality test (GCT) is used to find preliminary attacking variables which are causality relevant to the abnormal variable in whole network behavior. Depending on the behavior features hidden in the abnormal behavior, GCT is used again to recognize attacking variables which are causality relevant to the abnormal variable in local network behavior. The causality between attacking variables and the abnormal variable is then used to construct detecting rules, which are oriented to attacker. udpOutDatagrams acting as attacking variable are recognized successfully and detection results are acquired well in the test of Trin00 UDP Flood. The experiment results show that the approach can effectively detect attacks from attackers, which has effect on blocking the pervasion of attacking procedure to target.

Keywords: network behavior; attack detection; causality test; management information base

传统的攻击检测技术是针对攻击目标的单个行为进行的，忽略了攻击行为间的因果关联关系及攻击行为对于攻击过程的影响，从而使得漏报和虚警居高不下[1].因此，研究如何在复杂攻击场景中挖掘出影响攻击过程演进的攻击行为，并将其用于实际的攻击检测和预警分析，将有助于提前感知攻击过

收稿日期：2007-11-12. 作者简介：王志文(1973—)，男，博士，副教授. 基金项目：国家自然科学基金资助项目(60633020，60473136，60373105)；国家高技术研究发展计划资助项目(2006BAH02A24-2，2006BAK11B02，2007AA01Z475)；国家“二四二”信息安全计划资助项目(2006C26).

程,有助于采取主动防御措施.

在众多的攻击检测方法中,Cabrera 首次尝试在网络攻击检测中使用了 Granger 因果关联检验(GCT)方法[2].该方法的核心思想是利用统计工具来检验一个随机变量的滞后信息是否能够向另外一个随机变量提供统计意义上的有效预测,并已成功地应用到地震预警、股市分析和网络安全等领域.在实验中,Cabrera 将简单网络管理协议(SNMP)/管理信息库(MIB)变量作为检测变量进行了攻击检测分析,以便从攻击方挖掘出反应攻击过程的攻击变量,但由于数据序列元素之间的时间间隔设置得较大,因此难以准确反应出各数据序列间的因果关系.汪生等人[3]认为,攻击过程相对于系统行为之间的整体数据序列和局部数据序列应该存在不同的因果关联效应,据此提出了面向局部数据序列的 GCT 测试,但并没有进行具体的实验,也缺乏相应的实验数据.

本文在上述研究的基础上,充分利用了已有的 SNMP/MIB 数据,对面向网络行为因果关联的攻击检测进行了深入研究,并提出了"先整体,后局部"的分级 GCT 检测方法.

1 网络攻击时序模型

典型的网络攻击跨越了空间和时间 2 个维度,空间维度是指攻击过程涉及到的网络实体在物理位置上分布的任意性,而时间维度是指参与攻击的网络实体间由于存在相互交互而产生的前后时序关系.图 1 描述了网络攻击的时序关系,一个完整的攻击过程需要经历如下 4 个阶段.

(1)准备攻击(T_0):攻击方执行漏洞扫描、系统识别等动作,以确定攻击目标.

(2)实施攻击(T_1):攻击方发动攻击命令,如 TCP 半连接攻击、ICMP Flood 攻击等.

(3)攻击生效(T_2):攻击命令到达目标,并导致出现异常行为.

(4)安全受损(T_3):持续的攻击命令使得目标的安全受损.

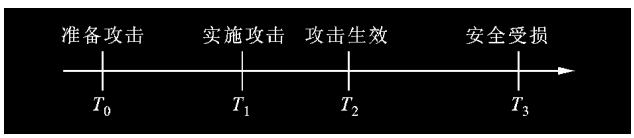

图 1 网络攻击时间序列

网络攻击在空间分布上的随意性和在时间滞后上的不确定性,加剧了网络攻击检测的复杂度,因此所有的检测方法都是基于时间或空间维度来记录数据,并利用规则推理、有限状态机、模式匹配以及统计分析等技术来抽取网络攻击特征,以避免攻击过程步入到 T_3 或 T_2 阶段.

2 术语定义

为了准确描述基于分级 GCT 的网络攻击检测方法,先作如下术语定义.

(1)网络行为.指描述网络运行状态的检测变量,如 CPU 占用、可用网络带宽、内存消耗等.在一段连续检测时间内的观测数值序列,记作 $B=\{v_k\}$ $(k=1,2,3,\cdots,N)$,其中 v_1 和 v_N 分别表示检测变量 V 在起始时刻 t_{init} 和结束时刻 t_{end} 的观测数值,则定义 $t_{\text{intervat}}=(t_{\text{init}}-t_{\text{end}})/N$ 为观测时间间隔,其大小会影响到网络行为的描述精度.

(2)时间窗口.与网络行为相对应的检测时间中的局部时间范围,记作 W $(t_{\text{low}},\ t_{\text{upper}})$,其中 t_{low} 和 t_{upper} 分别表示时间窗口的下沿和上沿,上沿与下沿的差值定义为时间窗口的大小 t_{win}.

(3)行为特征.指在时间窗口内或时间窗口之间,一个网络行为被观测到的数值是呈现一定规律性的序列,记作 $F=\{v_\lambda\}\subseteq B$ $(\lambda=1,2,3,\cdots,n)$,其中 v_1 和 v_n 分别表示时间窗口下沿 t_{low} 和上沿 t_{upper} 的观测值.

(4)局部行为.指一个行为特征对应的时间窗口下沿 t_{low} 在网络行为时间序列上回退一个时间窗口大小所得到的观测值序列.

(5)异常行为.在攻击目标的 T_3 或 T_2 阶段,安全受损的网络实体在运行过程中表现出的网络行为.用于构建异常行为的检测变量称为异常变量.

(6)攻击行为.在攻击方的 T_0 或 T_1 阶段,导致单个或多个网络实体的安全受损的网络行为.

(7)基本攻击变量.指与异常变量存在网络行为整体因果关联的检测变量.

(8)攻击变量.指与异常变量存在网络行为局部因果关联的检测变量,用于构建攻击行为.

(9)攻击关联.利用 GCT 确定与异常变量存在因果关联的检测变量的处理过程,分为整体关联和局部关联 2 个级别,前者确定基本攻击变量,后者确定攻击变量.

3 网络攻击检测模型

面向网络行为因果关联的分级 GCT 攻击检测模型如图 2 所示.该模型主要包含 4 个部分:网络行

为构建、异常检测、攻击变量挖掘以及攻击预防.

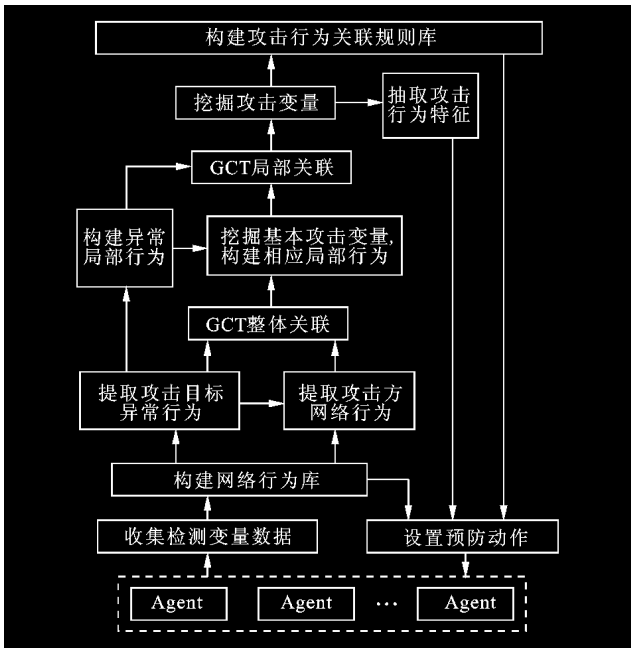

图 2　基于分级 GCT 的网络攻击检测模型

3.1　网络行为构建

根据网络行为的定义，SNMP/MIB 完全符合检测变量的要求，但要构建合理的用于攻击检测的网络行为，还必须解决下述 3 个技术问题.

(1)选择检测变量. 为了减少网络行为的数量，改善对攻击变量的挖掘效果，有必要对 SNMP/MIB 变量集合进行过滤. 多数情况下，攻击方和攻击目标均属于网络末端实体，其相互之间的数据传输处在网络层或在更高的层次上进行，可以选择 IP、ICMP、TCP 和 UDP 变量组中的 32 个变量作为检测变量.

(2)确定检测变量值的收集方式. 构建网络行为需按观察间隔定期监测、检测变量值，采用轮询方式较为合适.

(3)确定网络行为的时间长度. 网络行为可以用小时、天、星期或月来衡量，文献[4]通过观测和实验发现，正常网络流量均以天为周期，所以网络行为的时间长度可用天来衡量.

3.2　异常检测

异常检测的首要问题就是从检测变量中确定异常变量，基本界定方法有 2 种，即基于领域知识的分析方法和统计分析方法. 前者适用于攻击流程清晰的攻击类型，利用领域知识可以直接确定充当异常变量的 MIB 变量；后者适用于无法直接确定异常变量的攻击类型，需要逐个统计攻击场景下和正常场景下各检测变量的网络行为偏差，将存在较大差异的检测变量确定为异常变量.

在确定异常变量之后，对网络目标发动攻击，并根据行为特征来检测异常行为，从中提取异常行为特征.

3.3　攻击变量挖掘

基于网络攻击的时间特性，可以利用逆向回溯法在攻击方与攻击目标之间建立时序上的网络行为关联，即根据攻击目标方已确定的异常行为，先利用 GCT 从检测变量中找出与异常变量存在整体行为关联的基本攻击变量，然后根据异常行为特征，再次利用 GCT 从基本攻击变量中找出与异常变量存在局部行为关联的攻击变量，2 次 GCT 测试结果都超过设定门限的检测变量方可确定为攻击变量.

网络行为整体关联的执行流程如下.

(1)从网络行为库中提取攻击目标方的异常行为 b_{abnorm}.

(2)从网络行为库中提取所有与 b_{abnorm} 处于同一检测时间的攻击方的网络行为 $h_{\text{attack}}(j)$.

(3)计算所有输入/输出对($h_{\text{attack}}(j)$，b_{abnorm})的 GCT 检验统计量 g_{whole}.

(4)如果某一 $h_{\text{attack}}(j)$ 的 g_{whole} 值大于显著性水平 α 下 F 分布的临界值 F'，则表明它与 b_{abnorm} 存在整体关联，其对应的检测变量被标识为基本攻击变量[5].

由于 GCT 是一种统计学方法，仅仅执行一次整体关联所求得的基本攻击变量会存在一定的偶然性，因此攻击检测模型需要多次执行整体关联，以便找出准确的基本攻击变量.

网络局部行为关联的执行流程如下.

(1)抽取异常行为 b_{abnorm} 的全部行为特征 $f(i)$，$i=1,2,3,\cdots,M$.

(2)构建与异常行为特征相对应的异常局部行为 $\text{local\_}b_{\text{abnorm}}(i)$，$i=1,2,3,\cdots,M$.

(3)针对与 b_{abnorm} 存在整体关联的 $h_{\text{attack}}(j)$，提取与异常局部行为 $\text{local\_}b_{\text{abnorm}}(i)$ 处于同一检测时间的局部行为 $\text{local\_}h_{\text{attack}}(i,j)$.

(4)计算所有输入/输出对($\text{local\_}h_{\text{attack}}(i,j)$，$\text{local\_}b_{\text{abnorm}}(i)$)的 GCT 检验统计量 $g_{\text{local}}(i,j)$.

(5)若 $g_{\text{local}}(i,j)$ 低于显著性水平 α 下 F 分布的临界值 F'，则表明 $h_{\text{attack}}(j)$ 与 b_{abnorm} 之间不存在局部行为关联.

(6)对于存在局部行为关联的 $h_{\text{attack}}(j)$，统计各局部行为的 $g_{\text{local}}(i,j)$ 总和 $g_{\text{local}}(j)$. $g_{\text{local}}(j)$ 的值越

大，表明 $h_{attack}(j)$ 作为攻击行为的可能程度越高，$g_{local}(j)$ 的定义式为

$$g_{local}(j)=\frac{\sum_{i=1}^{M}(g_{local}(i,j)t_{win}(i))}{\sum_{i=1}^{M}t_{win}(i)}$$

式中：$t_{win}(i)$ 表示异常行为的第 i 个行为特征对应的时间窗口大小.

(7)根据已确定的攻击变量构建攻击行为关联规则.

3.4 攻击预防

攻击关联规则表明了攻击变量与异常变量之间的因果关系，但要想利用该规则来检测攻击事件，还必须提取攻击变量在攻击过程中所内含的行为特征，且在时间上必须超前于异常行为特征. 本文设计的攻击行为特征提取方法较为简单，即对于攻击行为关联规则($\{V_{attack}\}$, V_{abnorm})，只需扫描 $\{V_{attack}\}$ 中每一个攻击变量在攻击过程中与异常行为特征处于同一时间范围的局部行为，并抽取出局部行为前半时间段的行为特征.

4 实验仿真

为了验证基于分级 GCT 的检测效果，本文拟从攻击变量的确定性以及攻击检测效果 2 个方面进行测试.

实验环境：一台攻击主机，一台目标主机，安全管理主机通过局域网连接，其中攻击主机和目标主机部署了 SNMP Agent，而安全管理主机负责实施攻击检测. 攻击类型选择 Trin00 UDP Flood[6]，根据该攻击的工作机理可以确定异常变量是 udpInDatagram. 实验以天作为构建网络行为的时间单位，攻击主机每次攻击持续时间为 1 h，并按 10 s 和 1 min 的间隔独立采集攻击主机上作为检测变量的 32 个 MIB 变量值及目标主机的 udpInDatagram 变量值. 所有测试都是针对攻击主机进行的，主机的运行配置有 3 种：①仅运行攻击程序；②运行攻击程序和 FTP 下载程序；③运行攻击程序和 Netflow 采集程序.

4.1 攻击变量的确定性测试

攻击变量的确定性测试是对检测变量和异常变量的数据分别执行整体 GCT 检测和分级 GCT 检测，并比较各自在不同环境下挖掘出来的攻击变量是否一致.

表 1 给出了 GCT 关联统计量 g_{whole} 和 g_{local} 在显著性水平 α 为 0.05 的 F 分布临界值 $F'(P,T-2P-1)$，其中 P 表示数据序列中的前置元素个数，T 表示数据序列的元素个数. 整体 GCT 检测需要逐个计算 32 个检测变量的网络行为与 udpInDatagram 网络行为之间的 GCT 整体行为关联统计量 g_{whole}，并将超过临界值 F' 且具有最大 g_{whole} 的检测变量确定为攻击变量，表 2 给出了采样间隔为 1 min 的测试结果.

表 1 GCT 的临界值 F' 分布

统计量	采样间隔/s	采样次数	P	T	$1-\alpha$/%
g_{whole}	60	1 440	200	1 240	1.19
g_{local}	10	720	100	620	1.28

分级 GCT 检测只是将 g_{whole} 超过临界值 F' 的检测变量认定为基本攻击变量，相对于原始的 32 个检测变量而言，基本攻击变量的数量大幅减少，这有利于缩短执行网络局部行为的时间，降低 GCT 开销. 通过异常行为分析，发现存在 3 次数值单调递增的行为特征，这与攻击主机启动的 3 次攻击动作相一致，但各自的持续时间有所差异，与攻击动作的持续时间并不严格等同. 为了保持检测时间的一致性，保留单调递增持续时间的前 60 min 作为异常行为特征的时间窗口大小，相应地，局部行为的检测时间则设置为 120 min. 为了挖掘出攻击变量，需要计算基本攻击变量与异常变量之间的 GCT 局部行为关联统计量 g_{local}，并将其超过临界值 F' 且具有最大 g_{local} 的基本攻击变量确定为攻击变量. 表 3 给出了采样间隔为 10 min 的分级 GCT 方法挖掘出的攻击变量.

通过对比表 2 和表 3 可以发现，整体 GCT 检测在前 2 次运行配置中确定的攻击变量是 ipOutRequest，而第3次是 udpOutDatagram，这使

表 2 采样间隔为 1 min 的整体 GCT 检测挖掘结果

运行配置	检测变量数	g_{whole} 最大值	g_{whole} 最小值	满足 $g_{whole}\geqslant F'$ 的检测变量数	攻击变量名
①	32	4.11	1.04	8	ipOutRequest
②	32	3.67	0.91	7	ipOutRequest
③	32	3.50	0.79	11	udpOutDatagram

表 3　分级 GCT 检测挖掘结果

运行配置	基本攻击变量数	异常特征平均持续时间/min	g_{local} 最大值	g_{local} 最小值	满足 $g_{local} \geqslant F'$ 的检测变量数	攻击变量名
①	8	61.2	3.65	1.22	5	udpOutDatagram
②	7	59.4	3.41	1.02	6	udpOutDatagram
③	11	64.4	2.87	0.98	5	udpOutDatagram

得整体 GCT 方法挖掘的攻击变量具有不确定性. 相反，分级 GCT 检测在 3 种不同的运行配置中得出的攻击变量是一致的，挖掘结果具有确定性.

4.2　攻击检测效果

为了验证挖掘出的攻击变量在攻击预防中的有效性，本文连续进行了 5 天测试，每天在攻击主机上针对 3 种运行配置随机发起 10 次 Trin00 UDP Flood 攻击. 表 4 列举了以 udpOutDatagram 和 ipOutRequest 作为攻击变量的检测结果，从中可以看出，udpOutDatagram 的攻击检测成功率要明显高于 ipOutRequest. 该结果再次说明分级 GCT 方法的检测效果优于整体 GCT 方法.

表 4　不同攻击变量的检测效果

运行配置	发起攻击次数	udpOutDatagram 变量检测次数			ipOutRequest 变量检测次数		
		实检	误报	漏报	实检	误报	漏报
①	50	51	1	0	58	8	0
②	50	52	2	0	62	14	2
③	50	55	7	2	65	21	6

5　结　论

鉴于常规网络攻击检测方法主要是针对攻击过程的 T_3 阶段进行检测，难以在目标受损之前检测到攻击事件的问题，提出了一种新的面向网络行为因果关联的网络攻击检测方法：根据目标方的 SNMP/MIB 变量在攻击过程中表现出的异常行为特征，利用分级 GCT 检测理论执行 2 次反向回溯求解过程，依检测变量与异常变量之间的整体行为关联挖掘出基本攻击变量，再根据异常行为提取出行为特征，使用 GCT 从基本攻击变量中挖掘出与异常变量存在局部行为关联的攻击变量，并建立相应的攻击关联规则. 实验结果表明，通过攻击变量的行为特征检测能够在攻击方检测到攻击行为，阻止攻击过程向攻击目标进一步扩散.

参考文献：

[1] THOTTAN M, JI Chuanyi. Anomaly detection in IP networks [J]. IEEE Trans on Signal Processing, 2003, 51(8): 2191-2204.

[2] CABRERA J B D, LEWIS L, QIN Xinzhou. Proactive detection of distributed denial of service attacks using MIB traffic variables: a feasibility study [J]. IEEE Trans on Signal Processing, 2001, 49(6): 609-622.

[3] 汪生，孙乐昌，于国政. Granger 因果关系检验在攻击检测中的应用研究 [J]. 计算机应用，2005，25(6)：1282-1285.
WANG Sheng, SUN Lechang, GAN Guozheng. Application research based on Granger causality test for attack detection [J]. Computer Applications, 2005, 25(6): 1282-1285.

[4] 邹柏贤，姚志强. 一种网络流量平稳化方法 [J]. 通信学报，2004，25(8)：14-23.
ZOU Baixian, YAO Zhiqiang. A method to stabilize network traffic [J]. Journal of China Institute of Communications, 2004, 25(8): 14-23.

[5] HAMILTON J. Time series analysis [M]. Princeton, NJ, USA: Princeton University Press, 1994.

[6] CRISCUOLO P J. Distribution denial of service: trin00, tribe flood network, gribe flood network 2000, and stacheldraht, CIAC-2319 [R]. Washington DC, USA: Computer Incident Advisory Capacity, Department of Energy, 2000.

（编辑　苗凌）