一种基于动态建链推理的网络攻击过程分析方法

刘 晶，伏 飞，戴江山，肖军模

(解放军理工大学通信工程学院 南京 210007)

【摘要】提出一种动态漏洞链构造推理网络攻击过程的分析方法。以漏洞间推理关系为前提，从受害主机入手，构造有色加权有向图，在多日志中查找漏洞被利用的解释信息，并由查找结果对漏洞链动态剪枝，得到主机漏洞攻击链和攻击该受害主机的嫌疑主机，对嫌疑主机迭代分析，推理出网络漏洞攻击链。实例表明该方法能够快速有效地实现网络攻击过程分析，并且具有良好的可扩展性。

关 键 词 网络取证； 攻击分析； 漏洞链； 安全
中图分类号 TP393.08 文献标识码 A

A Method of the Network Attack Process Analysis Based on Dynamic Linking Inference

LIU Jing，FU Fei，DAI Jiang-shan，XIAO Jun-mo

(Institute of Communications Engineering, PLA Univ. of Sci. & Tech. Nanjing 210007)

Abstract Based on response after the attack incidents, a method of network attack process analysis by dynamic vulnerability linking is designed. The corresponding color weighted diagraph is setup in dependence on the inference relation of the security holes in the intruded machine. With the vulnerability-log relation matrix, we searched different forensic information sources are searched. The corresponding support value and the remote suspicious host are obtaind. Then the suspicious host in the same way is analyzed. The illustration indicates that this method can get the network attack process rapidly and effectively.

Key words network forensic; attack analysis; vulnerability; security

目前，由于网络攻击技术的发展以及软件系统固有的缺陷，网络入侵事件时有发生，特别是利用安全漏洞进行攻击已成为主要的攻击手段。根据2000～2003年CERT/CC公布的漏洞数据统计，平均每天就有十几个新的安全漏洞被发现，网络攻击者利用这些安全漏洞成功地对目标主机进行攻击。

本文提出动态漏洞链构造推理网络攻击过程的分析方法，该方法依据漏洞间的推理关系以及在日志信息中查找漏洞被利用的合理解释，进行动态漏洞链构造推理整个网络的漏洞攻击过程，其主要特点是：(1) 动态建链极大地缩小了日志检查的范围，提高了整体分析效率；(2) 多源日志信息查找具有完备性、准确性和真实性；(3) 具有良好的可扩展性，对于新的漏洞可以通过增加插件，识别新的攻击过程。

1 相关定义

定义 1 被入侵网络N由六元组$N=(H, E, L, \delta, \boldsymbol{\varphi}, \boldsymbol{M})$表示，其中$H$为$N$内的主机集合；$E$为$N$内漏洞$e_i$的集合；$\delta$为$H\rightarrow E$的映射；$\delta(h_i)=E_i$为主机$h_i$所含的漏洞集合；$L$为$N$内的日志集合，包括网络流日志、主机系统日志、服务日志、文件操作日志等；漏洞日志关联矩阵$\boldsymbol{\varphi}$表示漏洞与日志之间的关联关系，其中行代表漏洞，列代表日志，矩阵项内元素φ_{ij}定义为：

$$\varphi_{ij}=\begin{cases}1 & \text{漏洞}e_i\text{在日志}L_j\in L\text{中存在攻击者利用安全漏洞}e_i\text{攻击的相关日志记录}\\ 0 & \text{漏洞}e_i\text{在日志}L_j\in L\text{中不存在攻击者利用安全漏洞}e_i\text{攻击的相关日志记录}\end{cases}$$

访问关系矩阵$\boldsymbol{M}$表示网络N内主机之间的访问关系，其中行对应主体，列对应客体，矩阵项内存放主体对客体的访问能力：1表示可以访问，0表示不能访问。

收稿日期：2005 - 01 - 06
基金项目：国家自然科学基金资助项目(69931040)
作者简介：刘 晶(1981 -)，女，博士生，主要从事网络信息安全方面的研究.

定义 2 漏洞e_i采用五元组$e_i=(e_i^{\text{pre}},e_i^{\text{post}},\lambda,\delta^{-1},\partial)$表示，其中$e_i^{\text{pre}}=\{c_1^{\text{pre}},c_2^{\text{pre}}\cdots,c_n^{\text{pre}}\}$表示$e_i$被成功利用必须满足的前提条件，$e_i^{\text{post}}=\{c_1^{\text{post}},c_2^{\text{post}},\cdots,c_m^{\text{post}}\}$表示$e_i$被成功利用后产生的后置条件；$\lambda$表示漏洞$e_i$的风险等级值；$\delta^{-1}$表示$e_i\rightarrow h_i$的映射，即$e_i$所属主机；$\partial$表示漏洞$e_i$到漏洞类型的映射，漏洞类型分为远程进入系统漏洞、本地越权访问漏洞、数据库推理泄漏漏洞等，用T_j表示，$\partial(e_i)=T_j$表示e_i属于T_j类型，逆映射∂^{-1}表示有限漏洞集E中，T_j到E的映射，$\partial^{-1}(T_j)=\{e_i|\partial(e_i)=T_j$，且$e_i\in E\}$。

定义 3 设漏洞序列为$s_i=(e_1,e_2,\cdots,e_n)$：(1) 对于$\forall e_i$且$1<i\leq n$，有e_{i-1}^{post}满足e_i^{pre}，记作$e_{i-1}\prec e_i$；(2) $(e_1,e_2,\cdots,e_n)$中任意〈e_i,e_j〉满足偏序关系，即当满足$e_{i-1}\prec e_i$，不可能存在$e_i\prec e_j$且$e_j\prec e_{i-1}$；(3) 当$(e_1,e_2,\cdots,e_n)$为有限序列，存在极大值与极小值；(4) $(e_1,e_2,\cdots,e_n)$构造的有向图是哈斯图。

定义 4 日志信息支持度矩阵$\boldsymbol{D}$是依据漏洞日志关联矩阵$\boldsymbol{\varphi}$在各日志中查找漏洞是否被利用的合理解释的结果，其行代表日志，列代表漏洞。$\boldsymbol{\varphi}$中行向量$\boldsymbol{\varphi}_i$代表漏洞e_i可能在哪些日志中存在相关信息，$\boldsymbol{D}$中列向量$\boldsymbol{d}_i$相应地表示漏洞e_i在各日志中查找得到的合理解释支持度值。矩阵项内元素d_{ij}表示在日志L_i中找到的信息对漏洞e_j被利用的支持程度，用百分比表示。设所有日志对漏洞e_i的总体支持度为$\chi_i=|\boldsymbol{\varphi}_i\times\boldsymbol{d}_i|$，其中$\boldsymbol{\varphi}_i$表示矩阵$\boldsymbol{\varphi}$的第$i$行行向量，$\boldsymbol{d}_i$表示矩阵$\boldsymbol{D}$第$i$列列向量。根据训练数据给出平均支持度阈值$\overline{\chi}$，当$\chi_i<\overline{\chi}$表明没有足够信息证明该漏洞被利用。

2 动态建链推理的网络攻击过程分析方法

2.1 动态建链推理网络攻击过程分析方法描述

动态创建漏洞链推理网络攻击过程分析方法以漏洞间推理关系为前提，从受害主机入手，构造其有色加权有向图，依据漏洞日志关联矩阵在多日志中查找漏洞被利用的解释信息，计算支持度值，对漏洞链动态剪枝，由剪枝后的攻击链中远程漏洞取证信息判断攻击该机的嫌疑主机。嫌疑主机属于外网则是攻击机，推理结束；属于内网则可能是攻击机或是跳板主机，同理依据该主机漏洞攻击链中存在具有合理支持度值的远程漏洞，证明为跳板主机，否则判定其为攻击机。每一次迭代分析都对漏洞攻击链进行动态扩展，直至推理出全网的漏洞攻击过程。

步骤 1：设被攻击主机为h_i，采用网络安全扫描系统对被攻击主机h_i进行扫描，获得漏洞集合$E_I=\{e_i|\delta^{-1}(e_j)=h_i$，$e_i\in E$，$i=1,2,\cdots,n\}$。

步骤 2：对主机h_i内的漏洞构建有色加权有向图G_i。提前建立有漏洞推理规则库，对已公布漏洞前置条件、后置条件进行细化，并可根据最新公布的漏洞采用增加插件的方式进行实时更新。分析时用漏洞链关联引擎依据漏洞推理规则库对漏洞进行关联，创建漏洞序列s_i，s_i满足定义3中所有条件，因此可用有色加权有向图G_i表示。

G_i中节点表示漏洞，有向边〈e_i,e_j〉表示e_i的后置条件满足e_j的前提条件，有向边权值为漏洞e_j风险等级值$\boldsymbol{\lambda}_i$。有向图G_i有色指的是属于相同类型的漏洞具有同种颜色，同种颜色节点之间不存在有向边，这是因为如果攻击者发现同一主机存在多个漏洞，他为达到目的会选择最快捷的方式，同一类型的漏洞用以达到相同的目的，因此攻击者仅利用其中一个。有向图G_i的有色性有助于推理过程的简化。

步骤 3：对G_i中节点进行日志信息查找，得到日志信息支持度矩阵$\boldsymbol{D}$。采用遍历节点日志信息查找分析算法forensic，该算法采用递归调用遍历G_i中所有节点，对每个节点调用搜索评估函数search，传递参数e_i及$\boldsymbol{\varphi}_i$，得到查找信息结果$\boldsymbol{d}_i$，即漏洞e_i在$\boldsymbol{D}$中对应的向量。由定义3可知图G_i是哈斯图且存在极值，从任一极值节点e_i开始对图G_i进行遍历可减少搜索冗余度。设初始节点为e_0。

Input: $E=\{e_1,e_2,\cdots,e_n\}$; $V=\{\langle e_i,e_j\rangle|e_i^{\text{post}}\subseteq e_j^{\text{pre}}\}$; $M=\varnothing$; $P=\varnothing$; matrix $\boldsymbol{\varphi}$; 初始参数e_0

Output: matrix $\boldsymbol{D}$

Forensic(e_i) {

$c=e_i$;

$\boldsymbol{M}=\boldsymbol{M}\cup\{c\}$;

Search in set V，if $\exists$ 〈e_i c〉 $\in V$ or 〈c,e_i〉 $\in V$ then $P=P\cup\{e_i\}$；

While($P\neq\varnothing$) {e=rand e_i in set P; $P=P-\{e_k\}$;

if $e\notin\boldsymbol{M}$, then {$\boldsymbol{M}=\boldsymbol{M}\cup\{e\}$; $\boldsymbol{d}_i$=search(e_i, φ_i); forensic(e);}}}

函数forensic(e_i)最多进行$|E|=n$次迭代，每次迭代执行时间为$O(n)$，故时间复杂度最大为$O(n^2)$。

步骤 4：计算漏洞支持度值并对比阀值进行剪枝。依据步骤3得到的日志信息支持度矩阵$\boldsymbol{D}$，分别计算漏洞e_i总体支持度值 $\chi_i=|\boldsymbol{\varphi}_i\times\boldsymbol{d}_i|$，当 $\chi_i<\overline{\chi}'$ 对有向图G_i进行剪枝，即小于阈值 $\overline{\chi}'$ 的漏洞判定为没有足够的日志信息证明其被利用，在图G_i中将该漏洞表示的节点相连的边去除。

步骤 5：剪枝后有向图G'_i表示主机h_i中的漏洞攻击序列，它仍满足定义3的条件，因此漏洞攻击链具有有序性，检查其中极小值节点，若其中存在远程类型漏洞，转向步骤6，否则终止。

步骤 6：深入分析远程漏洞网络操作相关日志，如ftp、telnet等活动记录，利用时间约束性条件及空间约束性条件，查找网络嫌疑主机h_j。如果$h_j\in N$，赋值$h_i=h_j$，转向步骤2，否则终止。

上述步骤终止时的主机即被认定为攻击嫌疑机。它可能是内网主机，也可能是外网主机。

2.2　动态建链推理的网络攻击过程分析方法性能分析

假设网络中有N台主机，每台主机有m_i个漏洞，采用静态漏洞链分析的方法，需建立 $\prod_{i=1}^{N}m_i$ 条漏洞链，对每条漏洞链进行取证分析，共需进行 $\prod_{i=1}^{N}m_i\times N$ 次日志查找操作。设平均每台主机有m个漏洞，则需进行约 $m^N\times N$ 次日志查找操作。而采用本文提出的动态漏洞链构造分析方法，对一台主机需进行m_i次日志查找操作，通过日志支持度决定相关网络嫌疑主机。换言之，由于攻击通常会采取最迅速快捷的方式进行，普遍带有目的性，被攻击网络中受到攻击或作为攻击跳板的主机数目较少，设为$\sigma(N)$。在网络主机数目较多的情况下，仅对具有网络嫌疑的主机进行漏洞链构造与分析，$\sigma(N)/N\approx 0$。同样假设每台主机平均有m个漏洞，则日志查找工作总量为 $m\times\sigma(N)$。两种漏洞链日志查找总量比较的结果为 $\dfrac{m\times\sigma(N)}{m^N\times N}=\dfrac{\sigma(N)}{m^{N-1}\times N}<<1$，显然，本文提出的动态漏洞链构造分析方法更加快速有效。

3　实例分析

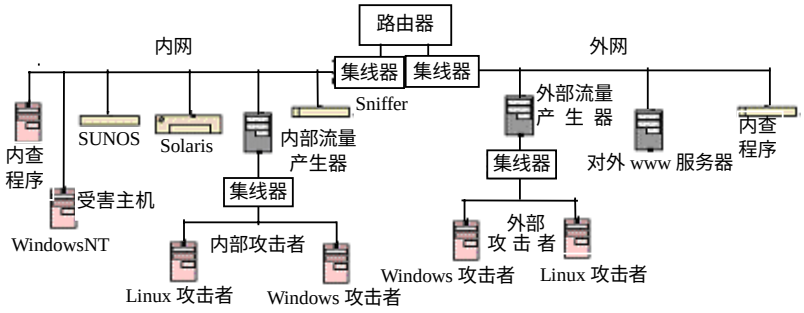

图1　网络结构图

模拟Lincoln实验室提供的2000DARPA入侵检测评估数据集[1]，搭建网络环境，见图1。网络模型$N=(H, E, L, \delta, \boldsymbol{\varphi}, \boldsymbol{M})$，其中主机集合$H$包括内网4台受害主机、2台攻击机、1个流量产生器和1个sniffer；外网2台攻击机、1台流量产生器和1个sniffer。L={netflow,syslog,telnet log,FTP log,web log, oracle log,FileAcess log}，网络访问关系矩阵$\boldsymbol{M}$如表1所示。

表**1**　网络N访问关系矩阵$\boldsymbol{M}$

主机	A	B	other
A	1	1	0
B	1	1	1
other	1	1	1

	L_1	L_2	L_3	L_4	L_5	L_6	L_7
e_1	1	1	1	1	0	0	1
e_2	1	1	1	0	0	1	1
e_3	0	1	0	0	0	0	1
e_4	0	1	0	0	0	1	1
e_5	0	1	0	0	0	1	1
e_6	0	1	0	0	0	0	1
e_7	0	1	0	0	0	0	1

图2　漏洞日志关联矩阵

实验结果表明：(1) 检查内网sniffer，检测到大量发往同一主机的数据包。源IP包括SunOS、Solaris和WinNT，对它们依次进行分析。SunOS主机中漏洞集合E_A={e_1,e_2,e_3,e_4,e_5,e_6,e_7}，依据漏洞类型分类，利用漏

洞间推理关系建立该主机的有色加权有向图G_A，如图3所示。依据E_A的漏洞日志关联矩阵$\boldsymbol{M}$，利用forensic算法遍历G_A中的漏洞并进行日志信息查找，得日志信息支持度矩阵$\boldsymbol{D}$，计算 $\chi_i = |\boldsymbol{\varphi}_i \times \boldsymbol{d}_i|$，与根据经验值设置的阈值比较后对图$G_A$进行剪枝得图 C'_A，如图4所示。从 C'_A 得漏洞攻击链(e_1,e_5,e_6,e_7)，其中e_1属于远程操作类型，查找e_1相关网络日志，可知内网Linux攻击者主机存在远程操作嫌疑。分析SUNOS和Solaris，可得相同结论。(2) 检查内网Linux攻击者，构建其有色加权有向图并分析得其漏洞攻击链，可有效地分析出外网Linux攻击者主机存在远程攻击嫌疑，且该主机即为攻击机。实验结果符合实验预期。

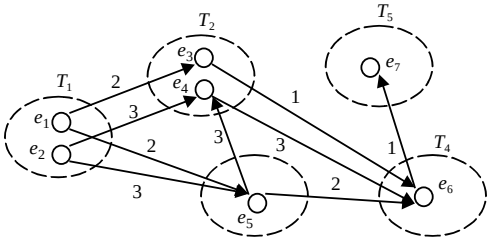

图3 主机B有色加权有向图G_A

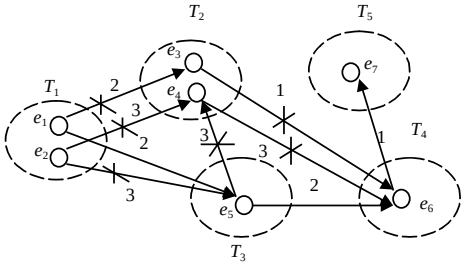

图4 剪枝后的主机B有色加权有向图 G'_A

4 结 束 语

目前，针对网络攻击过程的分析多集中于对攻击过程的某些具体问题的研究，如文献[2]采用静态漏洞链构造分析攻击过程，漏洞链数目庞大，冗余操作多，降低了分析效率；文献[3]采用因果推理入侵检测报警信息重构攻击过程，依赖存在漏报和误报的入侵检测系统，信息源缺乏准确性；文献[4]采用程序性推理，通过与攻击过程模板相匹配建立整个网络的攻击过程，缺乏对新的未知攻击的分析和识别。与以上方法相比，本文提出的动态漏洞链构造推理分析方法执行效率高，采用多信息源信息查找、分析攻击过程，准确、完整，且插件的运用使该方法具有良好的可扩展性，可对大规模网络攻击过程进行分析、识别和取证。

参 考 文 献

[1] MIT Lincoln Lab. 2000 DARPA intrusion detection scenario specific datasets[DB/OL]. http://www.11.mit.edu/IST/ideval/data/2000/2000_data_index.html, 2004-09-30.
[2] Sheyner O, Haines J, Somesh J, et al. Automated generation and analysis of attack graphs[C]// Proceddings of the 2002 IEEE symposium on security and privacy, Oakland, 2002.
[3] Peng Ning, Yun Cui, Douglas S R, et al. Constructing attack scenarios through correlation of intrusion alerts[C]// 9th ASM Conference on Computer & Communications Security, Washington, D.C., 2002.
[4] Douglas B M, Marbry T, Pauline B, et al. Diagnosis, explanation and recovery from computer break-ins[DB/OL]. http://www.ai.sri.com/~derbi/, 2004-09-30.

编 辑 熊思亮

(上接第793页)

4 结 束 语

反馈调度可以实现数据包丢失与网络时延等不确定环境的在线调度，本文利用实时控制任务周期与QoC周期的一致性特点，动态调整控制任务周期，实现控制与调度集成，同时提高了网络资源的利用率与控制系统性能。为了改善反馈调度性能，网络资源利用率的准确评估有待于进一步研究，如何将网络化控制系统中QoS回路与QoC控制回路有机地结合，还需要深入研究。

参 考 文 献

[1] Branicky M S, Phillips S M, Zhang Wei. Scheduling and feedback co-design for networked control systems[C]//Proc. IEEE on Decision and Control, Las Vegas, 2002, 2: 1211-1217.
[2] Stankovic J A, He Tian, Abdelzaher T F, et al. Feedback control scheduling in distributed real-time systems[C]// IEEE Real-Time Systems Symposium, London, UK, 2001: 712:724.
[3] Eker J, Hagander P, Årzén K E. A feedback scheduler for real-time control tasks[J]//Control Engineering Practice, 2000, 8(12): 1369-1378.
[4] Sename O, Simon D, Robert D. Feedback scheduling for real-time control of systems with communication delays[C]//9th IEEE International Conference on Emerging Technologies and Factory Automation, Lisbonne, 2003: 375-383.
[5] Cervin A. Integrated control and real-time scheduling[D]. Sweden: Lund Institute of Technology, 2003.

编 辑 漆 蓉