

## 基于粗糙集理论的主机安全评估方法

陈秀真, 郑庆华, 管晓宏, 林晨光

(西安交通大学电子与信息工程学院, 710049, 西安)

**摘要:** 针对大多数安全评估系统不能评估漏洞的组合对网络安全危害的缺陷, 提出采用粗糙集理论进行主机安全评估的方法. 该方法利用历史评估记录, 把漏洞作为安全要素, 在基于粗糙集理论的属性约简能力上, 建立了安全评估模型以及具有安全要素、服务和主机 3 个层次的安全风险度量模型, 再结合安全要素和服务重要性因子进行加权, 计算主机的安全风险, 进而评估、分析系统的安全态势. 与其他方法相比, 该方法能够自动建立基于规则的安全评估模型, 评估单个安全要素和安全要素的组合对系统安全的威胁, 且能够监控因系统配置改变引起的系统安全状态的变化. 通过仿真实验建立了系统安全态势曲线, 从 7 天的实验记录中还发现了 9 条有用的评估规则, 这表明采用该方法的评估结果更加准确、直观.

**关键词:** 网络安全; 安全评估; 粗糙集理论; 漏洞的组合; 安全态势; 数据挖掘

**中图分类号:** TP393 **文献标识码:** A **文章编号:** 0253 - 987X(2004)12 - 1228 - 04

### Approach to Security Evaluation Based on Rough Set Theory for Host Computer

Chen Xiuzhen, Zheng Qinghua, Guan Xiaohong, Lin Chenguang

(School of Electronics and Information Engineering, Xi an Jiaotong University, Xi an 710049, China)

**Abstract:** Aiming at the weakness of being unable to evaluate the threat of combination of vulnerabilities on network security for most systems of security evaluation, a novel model of security evaluation based on rough set theory was put forward. This model considered the vulnerability as a security factor and the security evaluation model was built from historical evaluation records by using attribute reduction. Further, the measurement model of hierarchical security risk with three levels: security factor, service and host computer, was built, which calculated the security risk of the host by weighting the importance of service and security factor, then the security situation of the system were analyzed and evaluated. Compared with other evaluation methods, this method can create a rule-based model of security evaluation automatically and has advantage of evaluating threat of isolated security factor and combination of security factors on the same host. It also can monitor the impact of changes of the system configuration on system security. Nine useful rules for security evaluation were discovered from 7 days' evaluation records and security situation curves were established through simulation experiment, which shows that evaluation results by our method are more accurate and intuitive than others.

**Key words:** network security; security evaluation; rough set theory; vulnerabilities combination; security situation; data mining

为保证网络安全运行, 人们采用了入侵检测、防火墙、虚拟专用网等技术, 然而这些技术属于被动防

卫手段. 为此, 网络研究人员基于防患于未然的思想提出了主动安全评估技术, 通过事先检查是否存在

收稿日期: 2004 - 03 - 11. 作者简介: 陈秀真(1977~), 女, 博士生; 管晓宏(联系人), 男, 教授, 博士生导师. 基金项目: 国家高技术研究发展计划资助项目(2001AA140213); 国家杰出青年科学基金资助项目(6970025); 国家自然科学基金资助项目(60243001).

可被黑客利用的漏洞来评估系统安全状况,并对发现的问题提出解决建议.目前的安全评估系统由扫描器发展而来,但大多简单罗列单个漏洞的危害信息及相应修补措施,未考虑漏洞组合的危害,不能提供安全态势曲线,且基于图形发现系统中的攻击路径集的方法<sup>[1]</sup>依赖于专家经验建立的攻击知识库.

鉴此,本文把安全漏洞作为安全要素,引入粗糙集理论(RST)来分析漏洞扫描器记录的海量历史信息,对系统安全要素进行约简和重要性度量,自动建立基于规则的安全评估模型,进而建立主机定量安全风险度量模型,以分析系统安全态势.

### 1 基于 RST 的安全评估系统

基于 RST 的安全评估系统分为在线和离线两大部分,主要由扫描器、攻击测试、知识发现、逻辑推理和安全态势分析模块组成,如图 1 所示.在该系统中,扫描器和攻击测试用于获取系统的脆弱信息,并将结果存入漏洞库;知识发现使用 RST 对漏洞库的历史数据进行离线分析,约简系统安全要素并度量其重要性,生成基于规则的安全评估模型且存于知识库;逻辑推理利用知识库中安全评估模型对当前扫描结果进行在线推理,得到评估结果;安全态势分析利用漏洞数据库和知识发现模块的安全要素重要性信息,建立系统安全风险评估模型,分析系统安全演化状况.该系统具有自适应性能,可定期使用 RST 对历史评估记录进行知识发现,及时更新知识库,以实现对新漏洞的评估.

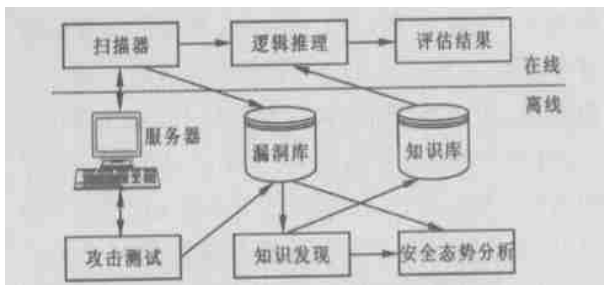


图 1 基于 RST 的安全评估系统

### 2 基于 RST 的安全评估方法

采用著名的 Nessus 扫描器<sup>[2]</sup>,其对一台主机每次扫描的记录少则有 70 条,多则达 100 多条,可见长时期的扫描结果是海量数据.由于 RST 借助于信息系统  $S = U, R = C, d, V, f, U \times R, V$  表达和处理知识,具有能从海量数据中发现有用规律并可转化这些规律为逻辑规则的优势,因此采用 RST

从长期扫描记录中发现影响系统安全的要素及要素的重要性,进一步挖掘出安全要素组合的潜在威胁规则,利用安全要素重要性建立整个主机安全风险度量模型,最终发现以及分析系统安全状态的演化是合适的.

#### 2.1 建立安全评估模型

主机扫描记录对应对象集  $U$ ,安全要素集  $F$  对应条件属性集  $C$ ,威胁评估结果集  $r$  对应决策属性集  $d$ .把主机中各个服务存在的漏洞  $v$  作为系统安全要素,威胁评估结果集的严重程度  $r = \{高, 中, 低\}$  的取值根据漏洞可能对系统造成的直接威胁来确定,参照文献[3,4]把漏洞的直接威胁分为 12 类,威胁等级的确定参见文献[3].选定条件属性集  $C = \{V_1, V_2, \dots, V_k\}$  对系统进行  $n$  次评估,建立评估信息决策表,见表 1.在评估信息决策表的基础上,充分利用 RST 的属性约简能力来度量安全要素的重要性,删除对安全评估结果无影响的系统安全要素(重要性为 0 的安全要素),使得评估信息决策表中的一个记录代表一类具有相同规律特性的样本.

表 1 评估信息决策表

U	条件属性				r
	V <sub>1</sub>	V <sub>2</sub>	...	V <sub>k</sub>	
1	V <sub>11</sub>	V <sub>12</sub>	...	V <sub>1k</sub>	r <sub>1</sub>
2	V <sub>21</sub>	V <sub>22</sub>	...	V <sub>2k</sub>	r <sub>2</sub>
...	...	...	...	...	...
n	V <sub>n1</sub>	V <sub>n2</sub>	...	V <sub>nk</sub>	r <sub>n</sub>

设  $R$  是由威胁评估结果集  $r$  的严重程度导出的分类,安全要素  $V_i$  在  $F$  中的重要性<sup>[5]</sup>为

$$I(V_i) = \frac{|P_F(R) - |P_{F \setminus \{V_i\}}(R)|}{|U|} \quad (1)$$

式中:  $P_F(R)$  为  $R$  的  $F$  正域,即根据知识  $F$  论域  $U$  中所有一定能归入集合  $R$  的元素构成的集合,表达式为

$$P_F(R) = \{Y_i | (Y_i \subseteq U | N(F) \cap Y_i \subseteq R)\} \quad (2)$$

式中:  $U | N(F)$  是不分明关系  $F$  对  $U$  的划分,即

$$U | N(F) = \{(x, y) | (x, y) \in U^2, P f \in F(f(x) = f(y))\} \quad (3)$$

评估信息决策表经过安全要素约简,就得到与安全评估决策规则相对应的结果,即就是所需要的评估模型,其形式为“ $A \rightarrow B$ ”.另外,为了对评估决

策规则  $A \rightarrow B$  推出的正确结论概率进行估计,引入评估决策规则  $A \rightarrow B$  的可信度

$$K_{A \rightarrow B} = \frac{|X \cap Y|}{|X|} \quad (4)$$

式中:  $X$  为安全要素值满足  $A$  的实例集合;  $Y$  为威胁评估结果的严重程度值满足  $B$  的实例集合.

### 2.2 主机安全风险评估

一台主机的安全取决于运行服务的安全,同时服务存在的漏洞影响服务的安全. 为了对主机安全作出整体量化评估,分析系统配置改变对系统安全的影响,提出了基于安全要素的重要性度量,从构成主机安全的安全要素层、服务层和主机层进行分析,建立了层次化主机安全风险评估模型,如图 2 所示. 主机层和服务层的安全风险度量可由它的下层各个子节点的安全风险指数加权和得到,具体分析如下.

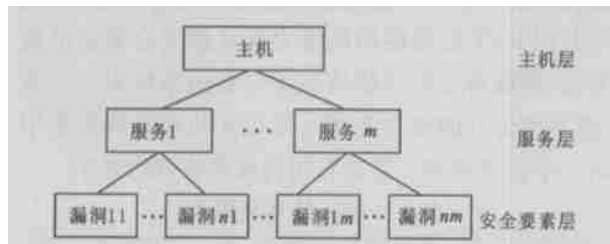


图 2 系统安全风险评估模型

主机  $H$  的安全风险  $Q_H$  定义为系统中服务所占的比重向量 与服务层安全风险向量  $Q_S$  的内积

$$Q_H = \sum_{i=1}^m w_i \cdot Q_{S_i} \quad (5)$$

其中  $w = (w_1, w_2, \dots, w_m)$ ,  $m$  为主机  $H$  运行的服务数,元素  $w_i$  为系统中各项服务所占的重要性比重,可由管理员根据各个服务在系统中的重要性来确定;  $Q_S = (Q_{S_1}, Q_{S_2}, \dots, Q_{S_m})$  是服务  $S_i$  的安全风险值,是各安全要素在系统安全中所占比重向量  $W_i$  与安全风险等级向量  $V_i$  的内积,即

$$Q_{S_i} = \sum_{j=1}^n W_{ij} \cdot V_{ij} \quad (6)$$

其中  $W_i = (w_{i1}, w_{i2}, \dots, w_{in_i})$  应来自于样本的客观信息,即利用安全要素的重要性信息并进行归一化处理,得到服务  $S_i$  的第  $j$  个安全要素在系统安全中所占的比重

$$w_{ij} = \frac{I(v_{ij})}{\sum_{k=1}^m \sum_{t=1}^{n_k} I(v_{kt})} \quad (7)$$

$V_i = (v_{i1}, v_{i2}, \dots, v_{in_i})$  是服务  $S_i$  的第  $j$  个安全要素的风险等级,按照漏洞本身的直接威胁等级进行赋值,对严重性为高、中、低的漏洞风险分别赋值为 7、5、3. 主机  $H$  的安全风险  $Q_H$  的意义在于,计算出一段时期内的安全风险态势,以便清晰地看出系统配

置变化对系统安全状况的影响.

## 3 仿真实验

为测试如图 3 所示的实验环境,在网关配置 Nessus<sup>[2]</sup>扫描器来探测、分析主机的脆弱性,管理员可根据实验需要作安全设置. 每天的 8:00、15:00 各扫描一次,测试时间达 3 个月,且每台主机每次扫描 100 多条记录,3 个月的记录可达 18 000 多条. 下面以局域网中配置有 RedHat Linux 710 操作系统的主机为例进行分析.

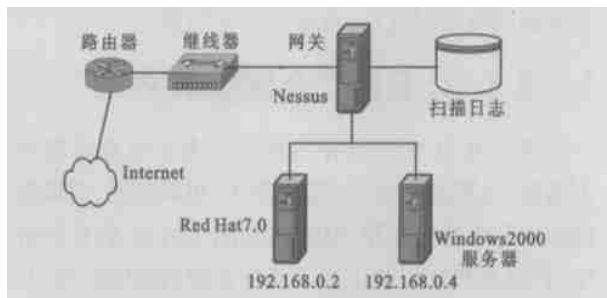


图 3 实验环境

### 3.1 安全评估建模

取 Linux 系统常见服务的漏洞作为安全要素,取  $F = \{ V_1(\text{opensslRBOF}), V_2(\text{chargenDoS}), V_3(\text{mysqlLBOF}), V_4(\text{fingerInfoLeak}), V_5(\text{lpdForStr}), V_6(\text{rloginAuth}), V_7(\text{ftpRBOF}) \}$ , 且  $F$  中元素为二进制变量,为 1 时表示存在,为 0 时表示不存在. 通过查询<sup>[6]</sup>,得到  $F$  中的 7 个元素,分别为导致黑客获取普通用户权限的 https 远程缓冲区溢出漏洞  $V_1(\text{opensslRBOF})$ ,chargen 服务的远程拒绝服务漏洞  $V_2(\text{chargenDoS})$ ,导致黑客获取管理员权限的 mysql 本地缓冲区溢出漏洞  $V_3(\text{mysqlL2BOF})$ ,finger 服务的信息泄漏漏洞  $V_4(\text{fingerIn2foLeak})$ ,导致黑客获取远程管理员权限的 lpd 格式化字符串漏洞  $V_5(\text{lpdForStr})$ ,rlogin 认证欺骗漏洞  $V_6(\text{rloginAuth})$ ,导致黑客获取远程管理员权限的 ftp 缓冲区溢出漏洞  $V_7(\text{ftpRBOF})$ . 对每次扫描结果进行预处理,提取  $F$  中各个元素的对应信息,每次扫描形成一条记录. 基于 RST 的安全评估方法描述如下(取 14 条记录为例).

就对对象集  $U = \{1, 2, \dots, 14\}$ ,根据专家知识对各个安全要素及安全要素组合的威胁进行赋值,建立评估决策表,见表 2.

使用 RST 属性重要性计算方法分析表 2,得到安全要素的重要性依次为  $\{1/7, 0, 1/7, 0, 1/7, 0, 1/14\}$ . 可见,安全要素  $V_1(\text{opensslRBOF})$ 、

$V_3$ (mysqlLBOF)、 $V_5$ (lpdForStr)和 $V_7$ (ftpRBOF)对系统安全影响比较大.通过安全要素约简,得到约简结果如表 3 所示.从表中可以得到 9 条评估决策规则,这就是所建立的安全评估模型.比如第 4 条规则“opensslRBOF(1) and mysqlLBOF(1) and lpdForStr(0) and ftpRBOF(0)  $r$ (高) ( $K=1$ )”,表示如果系统存在  $V_1$ (opensslRBOF)和 $V_3$ (mysqlLBOF),根据此规则判定威胁评估等级为“高”.这两个要素的组合使得远程攻击者利用远程溢出漏洞  $V_1$ (opensslRBOF)获得系统普通用户权限,然后利用本地溢出漏洞  $V_3$ (mysqlLBOF)获得超级用户权限,给系统造成潜在的、致命的威胁.可以看出,利用本文方法得出的实验结果比较理想,而且在样本足够大的情况下,可以发现更多的漏洞组合对应的系

统安全评估的规则.

### 3.2 主机系统安全风险态势

对所提的风险评估模型,服务数  $m=7$ ,管理员可根据这 7 个服务的重要性确定比重  $w = \{5/17, 1/17, 4/17, 1/17, 2/17, 1/17, 3/17\}$ ,各个服务的安全要素的威胁等级分别为{中,中,中,低,高,中,高},量化赋值为{5,5,5,3,7,5,7},则根据式(7)计算出各安全要素在系统安全中所占比重为  $w_i = \{2/7, 0, 2/7, 0, 2/7, 0, 1/7\}$ .利用式(5)得到系统在某一时刻的安全风险值,最终得到 1 个月内主机的安全态势,结果如图 4 所示.图中前 14 天分析结果与表 2 的 14 条记录对应,可以看出第 3 天的安全风险值比较高.从表 2 中的第 3 条记录也可看出,它存在安全要素  $V_1$ (opensslRBOF)、 $V_3$ (mysqlLBOF)和  $V_7$ (ftpRBOF),且这 3 个安全要素的重要性都比较大,同时还存在 finger 和 chargen 服务,威胁系统安全的要素也比较多,此时安全风险值高是合理的.安全风险值波动比较大是由于每天开启、关闭一些服务造成的.风险值为 0 对应于系统仅存在安全比重为 0 的安全要素的情况.

表 2 评估决策表

U	条件属性							r
	V <sub>1</sub>	V <sub>2</sub>	V <sub>3</sub>	V <sub>4</sub>	V <sub>5</sub>	V <sub>6</sub>	V <sub>7</sub>	
1	0	0	1	1	0	1	0	中
2	1	0	0	0	1	1	0	高
3	1	1	1	1	0	0	1	高
4	1	0	0	0	0	1	1	高
5	1	0	1	1	0	1	0	高
6	0	0	0	1	0	1	1	中
7	0	0	0	0	0	1	0	低
8	1	1	1	1	0	1	0	高
9	0	0	0	0	1	1	0	高
10	0	0	0	1	0	1	1	高
11	0	0	1	1	0	0	1	高
12	0	0	0	1	0	1	0	低
13	1	0	0	0	1	1	0	高
14	1	0	0	0	1	1	0	高

表 3 安全要素约简结果

U	条件属性				r	K
	V <sub>1</sub>	V <sub>3</sub>	V <sub>5</sub>	V <sub>7</sub>		
1	0	1	0	0	中	110
2	0	1	0	1	高	110
3	1	1	0	1	高	110
4	1	1	0	0	高	110
5	0	0	0	1	中/高	015/015
6	0	0	0	0	低	110
7	1	0	0	1	高	110
8	0	0	1	0	高	110
9	1	0	1	0	高	110

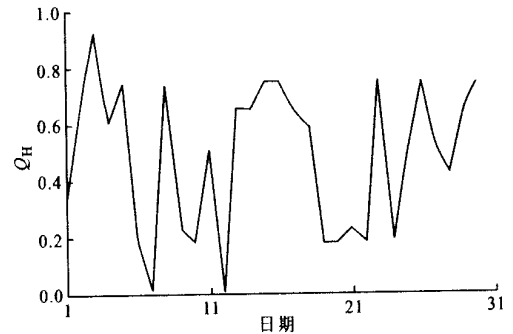


图 4 主机系统安全态势

## 4 结 论

本文提出的利用 RST 进行安全评估的方法与已有方法相比,具有度量安全要素重要性、发现和评估多个安全要素组合的威胁以及评估分析一段时间内安全态势的特点,能够反映出系统配置对系统安全的影响,使得安全评估结果更加准确、直观.该方法适用范围广,只要更改相应的条件和决策属性,就能够应用于各种操作系统.同时,还可以发现主机间漏洞的组合对系统威胁的评估规则,这一部分工作有待于进一步研究.

(下转第 1255 页)

**参考文献:**

- [1] Gader P D, Mohamed M A, Keller J M. Fusion of handwritten word classifiers [J]. Pattern Recognition Letters, 1996, 17(6): 577 - 584.
- [2] Tahani H, Keller J M. Information fusion in computer vision using the fuzzy integral [J]. IEEE Trans on Systems, Man, and Cybernetics, 1990, 20(3): 733 - 741.
- [3] 杨 焯, 裴继红, 杨万海. 基于模糊积分的融合图像评价方法 [J]. 计算机学报, 2000, 24(8): 5 - 8.
- [4] Auephanwiriyaikul S, Keller J M, Gader P D. Generalized Choquet fuzzy integral fusion [J]. Information Fusion, 2002, 3(1): 69 - 85.
- [5] Sugeno M. Fuzzy measures and fuzzy integrals: a survey [A]. Fuzzy Automata and Decision Process [C]. New York: North-Holland, 1977. 89 - 102
- [6] Murofushi T, Sugeno M. A theory of fuzzy measures: representations, the Choquet integral, and null sets [J]. J Math Anal Appl, 1991, 159(2): 532 - 549.
- [7] Pawlak Z, Peters J F, Skowron A, et al. Rough measures: theory and applications [J]. Bulletin of International Rough Set Society, 5(1 - 2): 177 - 183.
- [8] Chiang J H. Aggregating membership values by a Choquet fuzzy integral based operator [J]. Fuzzy Sets and Systems, 2000, 114(3): 367 - 375.
- [9] Grabisch M, Sugeno M. Multiattribute classification using fuzzy integral [A]. The First IEEE Conference on Fuzzy Systems, San Diego, USA, 1992.
- [10] Keller J M, Osborn J. Training the fuzzy integral [J]. International Journal of Approximate Reasoning, 1996, 15(1): 1 - 24.
- [11] Borkowski M, Peters J F. Approximating sensor signals: a rough set approach [A]. Canadian Conference on Electrical and Computer Engineering, Winnipeg, Canada, 2002.
- [12] Grabisch M, Dispot E. A comparison of some methods of fuzzy classification on real data [A]. Second International Conference on Fuzzy Logic and Neural Networks, Lizaoka, Japan, 1992.
- [13] Moore R E. Interval analysis [M]. New York: Prentice Hall, 1966.
- [14] Lin T Y, Liu Q. First order rough logic: approximate reasoning via rough sets [J]. Fundam Inform, 1996, 27(2 - 3): 137 - 153. (编辑 苗 凌)

**(上接第 1231 页)****参考文献:**

- [1] Ammann P, Wijesekera D, Kaushik S. Scalable, graph-based network vulnerability analysis [A]. 9th ACM Conference on Computer and Communications Security, Washington DC, 2002.
- [2] Deraison R. Nessus [EB/OL]. <http://www.nessus.org>, 2003 - 02 - 15.
- [3] Stardust. 计算机网络系统安全漏洞分类研究 [EB/OL]. <http://www.xfocus.net/articles/200103/126.html>, 2003 - 03 - 03/2003 - 11 - 13.
- [4] Bishop M, Bailey D. A critical analysis of vulnerability taxonomies [R]. Technical Report 96 - 11. Davis, USA: Department of Computer Science, University of California, 1996.
- [5] Pawlak Z. Rough sets [J]. International Journal of Computer and Information Science, 1982 (11): 341 - 356.
- [6] Mell P. ICAT metabase: your CVE vulnerability search engine [EB/OL]. <http://icat.nist.gov/icat.cfm>, 2003 - 08 - 16. (编辑 苗 凌)

**(上接第 1250 页)**

- [3] Cabrera J B D, Lewis L, Qin Xinzhou, et al. Proactive detection of distributed denial of service attacks using MIB traffic variables: a feasibility study [A]. Proceedings of 2001 International Symposium on Integrated Network Management [C]. Piscataway, USA: IEEE Press, 2001. 6092622.
- [4] Kulkarni A B, Bush S F, Evans S C. Detecting distributed denial of service attacks using Kolmogorov complexity metrics [R]. Technical Report 1CRD176. New York: Research and Development Center, General Electric Company, 2001.
- [5] Tao Peng, Leckie C, Ramamohanarao K. Protection from distributed denial of service attacks using history-based IP filtering [A]. IEEE International Conference on Communications, Anchorage, USA, 2003.
- [6] Cohen L. Time-frequency analysis [M]. Englewood Cliffs, USA: Prentice Hall, 1995.
- [7] Duda R O, Hart P E, Stork D G. Pattern classification [M]. 2nd ed. New York: Wiley, 2001.
- [8] MIT Lincoln Laboratory. 2000 DARPA intrusion detection scenario specific data sets [EB/OL]. <http://www.ll.mit.edu/IST/>, 2003 - 10 - 21. (编辑 苗 凌)