

网络安全量化评估系统的研究与应用

夏阳¹ 陆余良¹ 蒋凡²

(解放军电子工程学院网络工程教研室 合肥230037)¹

(中国科学技术大学信息安全研究中心 合肥230027)²

Research and Application of a Network Security Quantitative Evaluation System

XIA Yang¹ LU Yu-Liang¹ JIANG Fan²

(Teaching and Research Office of Network Engineering of Electronic Engineering Institute Hefei 230037)¹

(Research Center for Information Network Security, University of Science and Technology of China Hefei 230037)²

Abstract The rapid development of the Network makes the comprehensive analysis as well as the quantitative evaluation of its security become more and more important. This paper illustrates the major realization process of a Network Security Quantitative Evaluation System, which, from an intruder's angle, established a Hierarchy Intrusion Relationship Graph by analyzing the credit degree fusion and relevancy of the secure information of the target network and by combining with powerful database information. At last, by applying some relative mathematics model and arithmetic, the paper analyzes and evaluates the security of this Network Hierarchy Intrusion Relationship Graph comprehensively and quantitatively.

Keywords Quantitative evaluation, Hierarchy intrusion relation graph, Trustful degree fusion

1 引言

随着计算机网络技术的不断发展,各种信息系统对计算机网络的依赖越来越强,网络安全的研究日益引起了广泛的重视,特别是在网络安全的量化评估研究领域已经取得了不少的研究成果。目前,国内外的研究者分别从不同的角度对网络安全进行研究,主要研究方向有:①从逆向工程的角度分析目标对象存在的脆弱性;②从程序代码实现的角度出发,对代码中可能出现的执行错误进行分析,然后在此基础上对软件进行错误注入^[1],以此发现目标程序的脆弱性;③以目标测试为基础,利用形式化分析技术从不同的测试角度来发现目标对象可能存在的脆弱性;④从攻击树(attack tree)和错误树(fault tree)的角度出发,以逻辑分析的方法对网络环境下的漏洞进行量化评估分析^[2];⑤以“电路理论(circuit theory)”为基础,将目标网络的拓扑结构图转变成等价的电路分析图,对目标网络的安全性进行量化评估分析^[3];⑥以图论为基础,根据目标网络的相关知识生成权限关系图,并在此基础上建立数学模型进行不同角度的安全量化分析^[4,5]。前三种方法主要是为了发现新的漏洞出发,这些方法具有较大的理论意义,而且在实际的应用中确实可以发现计算机网络系统中的安全脆弱性,但是其难度较大,不易于形成自动化系统,并且发现的漏洞数量有限。后三种方法主要是利用各种不同的理论,建立相关的模型,对目标网络的安全状况进行量化评估,但这些方法的主要不足是:模型对网络行为的描述过于理想化,没有针对网络安全的实际情况对模型的粒度进行深入分析,在实际的应用中还有待进一步改进。

本文从入侵者的角度出发,设计一种基于网络层次入侵关系图的网络安全量化评估系统,并对该系统的主要功能组成、信任度融合及关联分析的实现以及基于网络层次入侵关系图进行目标网络安全状况综合评估的方法和步骤进行了较为详细的论述。

2 网络安全量化评估系统的功能组成

本文提出的网络安全量化评估系统的主要功能组成如图1所示。

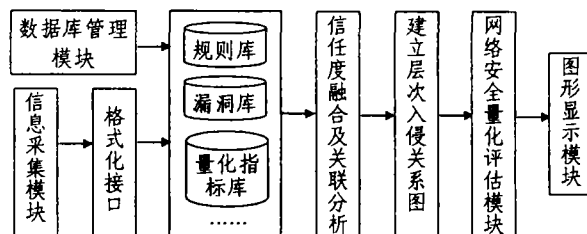


图1 系统功能组成

该系统实现了从信息采集到目标网络安全分析及相关数据库管理的全部功能,具体的工作流程可以归纳为以下几个步骤:

第一步:利用信息采集模块中内嵌的各种扫描程序针对目标网络进行扫描,它主要获取目标主机的操作平台信息、开放服务和端口信息以及各种可能的配置脆弱性信息等;

第二步:将第一步获取的各种扫描信息格式化为标准样式存储在相关数据库中;

第三步:将经过第二步格式化后的数据信息进行信任度融合及关联分析。由于扫描所获取的各种信息之间可能存在着冲突、重复或者互相支持的关系,在该系统中称处理这些关系而进行的操作为信任度融合处理;

第四步:综合目标网络所有主机的脆弱性信息,建立网络层次入侵关系图;

第五步:基于网络层次入侵关系图进行安全评估并图形化显示。

3 目标主机信任度融合

系统的信息采集模块对目标网络进行扫描所得到的扫描

信息经过标准格式化后存储在数据库中,这些信息包括目标网络中各主机的系统平台、开放服务及端口、配置脆弱性等相关信息,数据信息之间可能会有不同程度的重叠部分,甚至有些信息之间还存在冲突,因此我们必须对这些信息进行综合分析及信任度融合处理。

Dempster-Shafer 证据推理理论是由经典 Bayes 推理理论发展而来的。Dempster-Shafer 最大特点是给出了事件的信任函数和事件之间的分离程度,而不强求对不确定性信息分配概率值,避免了对缺乏信息的数据强作假设。

Dempster 组合规则是 Dempster-Shafer 证据推理理论在数据融合应用中的关键,它用来将相对同一个辨别框架 Θ 上的多个证据体进行合并以融合成一个新的证据体,并通过相容关系将证据体从一个框架转移到另一个框架,这个新的证据体强调了原多个证据体的相同点削弱了异同点。设有基于判别框架上 Θ 的两个证据体 m_1, m_2 , 并分别含有焦点 A_1, \dots, A_n 和 B_1, \dots, B_n , 则它们的组合运算为 $m = m_1 \oplus m_2$, m 为组合产生的新的证据体。

$$m(C) = \frac{\sum_{A_i \cap B_j = C} m_1(A_i) * m_2(B_j)}{1 - k}$$

$$k = \sum_{A_i \cap B_j = \emptyset} m_1(A_i) * m_2(B_j)$$

k 表示两个证据体矛盾的程度, Dempster 组合规则的结果根据 k 值进行归一化, 如果 $k=1$ 则表示几个证据体完全矛盾, 不能应用 Dempster 组合规则, 如果 $k=0$ 则不用归一化。

在本系统的规则库中定义有如下推理规则:

- ① OS = {W, U, L}, OS 代表操作系统类型, W, U, L 分别代表 Windows, Unix 和 Linux;
- ② Port = i , i 代表计算机进行通讯的端口号, $0 < i < 65536$;
- ③ Service = X, X 代表经过标准格式化以后的服务程序名称;
- ④ V_i 代表漏洞 i ;
- ⑤ 存在类似与表1所列的规则表格。

表1

数据	推理
OS=W;Port=139	$V_1; V_2; V_3$
OS=U;Service=FTP/2.5.0	$V_4; V_5$
OS=W;Port=21	$V_5; V_6$
OS=W;Port=80	$V_7; V_8; V_9$
OS=L;Port=80	$V_8; V_9; V_{10}$
Service=Apache/1.3.32	$V_9; V_{10}; V_{11}$
.....

表格左栏数据代表从目标主机获取的标准扫描信息,右栏是对应该扫描信息可能存在的漏洞列表。

下面开始依次对目标网络中的各台主机进行扫描,并获取某一台主机的扫描结果如表2所示。

表2

	证据	证据可信度
第一证据	OS=W;port=80	80%
第二证据	OS=L;port=80	60%
第三证据	Service=Service=Apache/1.3.32	90%

根据 Dempster 组合规则,首先对第一和第二证据体进行融合推理,得出如表3结论。

表3

	$m_1(\{V_7, V_8, V_9\}) = 0.8$	$m_1(\Theta) = 0.2$
$m_2(\{V_8, V_9, V_{10}\}) = 0.6$	$m'(\{V_8, V_9\}) = 0.48$	$m'(\{V_8, V_9, V_{10}\}) = 0.12$
$m_2(\Theta) = 0.4$	$m'(\{V_7, V_8, V_9\}) = 0.32$	$m'(\Theta) = 0.08$

然后将上表推出的结果与第三证据体进行规则组合,结果如表4。

表4

	$m'(\{V_8, V_9\}) = 0.48$	$m'(\{V_8, V_9, V_{10}\}) = 0.12$	$m'(\{V_7, V_8, V_9\}) = 0.32$	$m'(\Theta) = 0.08$
$m_3(\{V_9, V_{10}, V_{11}\}) = 0.9$	$m''(\{V_9\}) = 0.432$	$m''(\{V_9, V_{10}\}) = 0.108$	$m''(\{V_9\}) = 0.288$	$m''(\{V_9, V_{10}, V_{11}\}) = 0.072$
$m_3(\Theta) = 0.1$	$m''(\{V_8, V_9\}) = 0.048$	$m''(\{V_8, V_9, V_{10}\}) = 0.012$	$m''(\{V_7, V_8, V_9\}) = 0.032$	$m''(\Theta) = 0.008$

计算推理结果的置信函数 $BEL(A) = \sum_{B \subset A} m(B)$, $BEL(A, B)$ 代表结果是 A 或 B 的信任度:

$$BEL(V_9) = 0.432 + 0.288 = 0.72$$

$$BEL(V_9, V_{10}) = 0.432 + 0.288 + 0.108 = 0.828$$

$$BEL(V_8, V_9) = 0.432 + 0.288 + 0.048 = 0.768$$

从上述置信函数的结果中明显可以看出, V_9 最有可能是目标主机上存在的漏洞,其次是 V_{10} , 最小存在可能性的漏洞是 V_8 , 至于 V_7 和 V_{11} 存在的可能性太小, 可以忽略不计。

假设 $P(V_9)$ 代表 V_9 漏洞存在的概率, $P(V_9, V_{10})$ 代表 V_9 或 V_{10} 漏洞存在的概率, 根据 $P(V_8, V_9) = 1 - (1 - P(V_8)) \times (1 - P(V_9))$, $P(V_9) = 0.72$, $P(V_9, V_8) = 0.768$ 及 $P(V_9, V_{10}) = 0.828$ 可以得出: $P(V_8) = 0.171$; $P(V_{10}) = 0.386$ 。因此该主机上漏洞存在可能性的比例为: $V_9:V_{10}:V_8 = 0.768:0.386:0.171 \approx 4:2:1$ 。

4 目标主机脆弱性关联分析

通过上述基于目标主机扫描信息的信任度融合得出了目标主机上的漏洞存在情况及漏洞存在可能性。该系统中的漏洞数据库已经存储了大量的漏洞信息, 包括漏洞发现时间、影响平台、影响系统服务程序、CVE 号、危害等级、攻击类型、攻击技术难度等, 为了更好地对漏洞的攻击成功率进行量化, 首先我们必须提取合理的漏洞量化指标, 指标的提取要具有科学性和有效性, 指标之间存在着相互依存和制约关系, 大量指标要素的集合才能对某一漏洞进行精确的量化。为了更加清晰地论述漏洞间的关联性问题, 表5中提取的若干漏洞量化指标只是作为案例数据, 并不具有实际应用价值, 假设上述目标主机中的漏洞 V_8 对应于表5中的各指标参数。

表5

量化指标	权重	指标成功率
$a_1 =$ 漏洞发现时间	w_1	$t_1 = f_1(a_1)$
$a_2 =$ 漏洞攻击技术难度	w_2	$t_2 = f_2(a_2)$
$a_3 =$ 攻击隐蔽性	w_3	$t_3 = f_3(a_3)$
$a_4 =$ 目标主机系统平台类型	w_4	$t_4 = f_4(a_4)$
.....

表5中 t_i 代表第 i 个量化指标单独作用于该漏洞的相对成功率, f_i 是由量化指标要素向 t_i 转变的映射函数, w_i 代表该量化指标要素在所有指标中的重要性。攻击该漏洞的成功率用 S_1 表示, 则 $S_1 = \sum_{0 \leq i \leq n} w_i * t_i$ 。同理, 我们可以得出漏洞 V_9 和 V_{10} 的攻击成功率分别为: $S_2 = \sum_{0 \leq i \leq n} w_i * t'_i$ 和 $S_3 = \sum_{0 \leq i \leq n} w_i * t''_i$ 。需要注意的一点是, 针对某一漏洞的权重并不是一成不变的, 它需要根据目标主机的相关信息作调整, 例如目标主机是一台数据库服务器, 一般很少有工作人员在上面操作, 因此攻击是否隐蔽将不是攻击目标服务器成功与否的主要因素, 故表5中“ $a_3 =$ 攻击隐蔽性”指标的权重 w_3 应做相应调整, 同时也牵扯到其他指标的变动, 使满足 $\sum_{1 \leq i \leq n} w_i = 1$, 在具体的实现过程中运用了线性规划的相关知识, 使漏洞量化指标的权重能够在合理的范围内动态取值, 提高了漏洞攻击成功率的可信度。

目前已经得到了目标主机可能存在的漏洞、漏洞存在可能性以及漏洞攻击的成功率, 因此针对目标主机可以生成下面的数据表格(表6)。

表6

漏洞	权重(漏洞存在可能性)	漏洞攻击成功率
V_8	w'_1	S_1
V_9	w'_2	S_2
V_{10}	w'_3	S_3

通过前面的信任度融合计算可知: $w'_2 : w'_3 : w'_1 \approx 4 : 2 : 1$, 且 S_1, S_2 和 S_3 也已经计算得出。假设“ $S =$ 目标主机的攻击成功率”, 则:

$$S = 1 - (1 - w'_1 * S_1) * (1 - w'_2 * S_2) * (1 - w'_3 * S_3)$$

$$= 1 - (1 - \frac{1}{7} * S_1) * (1 - \frac{4}{7} * S_2) * (1 - \frac{2}{7} * S_3)$$

在上述针对目标主机的关联性分析中, 我们提取了漏洞的量化指标及相应的权重, 然后按照一定的算法计算单一漏洞的攻击成功率, 最后结合漏洞存在的可能性等信息进行关联分析, 评估目标主机的总体攻击成功率, 该评估模型已经在本网络安全量化评估系统中得到实际应用并取得了预期效果。

5 基于层次入侵关系图的网络安全量化评估

大多数网络入侵事件的发生是一个层次入侵的过程, 真正的目标主机可能是难以直接攻破的, 但是由于和它相关的其他机器可能存在安全脆弱性, 这往往是由于配置不当或存在不同程度的信任关系所造成, 因此网络黑客可以从薄弱环节入手, 基于层次入侵的思想逐步提高自己的权限, 最终达到控制目标机器的目的。在网络安全评估系统中, 可以利用信息采集模块收集的各种网络安全信息, 结合漏洞库、规则库及量化指标库等逐步构造出网络层次入侵关系图。

首先给出网络层次入侵关系图的若干数学定义:

定义1 图中的每个顶点代表网络中的每一台主机, 顶点可以表示为一个多维向量集合 $U = (\{p_1, q_1\}, \{p_2, q_2\}, \{p_3, q_3\}, \dots, \{p_n, q_n\})$, 其中 p_i 代表一些关键属性名, q_i 表示相应属性的值。

定义2 有向弧 $\textcircled{X} \xrightarrow{w} \textcircled{Y}$ 代表存在从结点 X 入侵结点 Y 的途径, w 代表入侵成功的可能性或入侵成功的所需的时间等权值。

定义3 图中弧代表三种类型的入侵方法:

(1) 入侵方法由于网络中存在的一个安全脆弱性而引起, 如文件的访问权限控制不当或易猜口令;

(2) 入侵方法可能是由于系统提供的一些信任关系得到不当的扩展;

(3) 入侵方法本身就是系统可以利用的权力。

定义4 割集是网络中部分主机的集合, 当所有这些主机都被攻破后, 则整个网络入侵所达到的目标才可以实现。

定义5 路集是网络中部分主机的集合, 当所有这些主机中任一主机被攻破, 则整个网络入侵所达到的目标才可以实现。

构造网络层次入侵关系图是一个逐步积累的过程, 该过程基于以下三条假设:

假设1 入侵者并不知道整个网络的拓扑结构, 他是在逐步提升权限的过程中逐步建立起层次入侵关系图;

假设2 入侵者不会从不同的路径入侵同一个主机结点;

假设3 入侵者可以记住他每次走过的路径及入侵过的结点。

基于以上三点假设, 利用前面章节讲述的信任度融合及关联性分析等相关算法, 结合数据库中预先存储的安全规则依赖关系, 针对目标网络进行扫描分析, 假设得到的目标网络层次入侵关系图如图2所示。

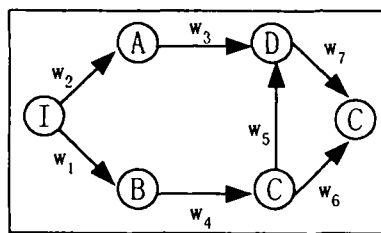


图2 网络层次入侵关系图

在该网络层次入侵关系图的基础上, 我们可以利用图论的相关知识从不同侧面对目标网络的安全状况进行评估, 下面给出了各种评估方法及步骤:

(1) 威胁路径分析:

步骤1: 选择入侵起始点和目标点;

步骤2: 对入侵关系图进行搜索, 找出所有可能通向入侵目标点的路径;

步骤3: 对所有路径进行价值分析, 并将其标记;

步骤4: 以图形和表格的方式进行结果显示;

(2) 入侵成功率评估:

步骤1: 选择入侵起始点和目标点;

步骤2: 将网络层次入侵关系图转换为网络入侵过程状态图;

步骤3: 基于网络入侵过程状态图, 按照一定的数学算法量化评估出起始点攻击目标点的可能成功率或攻击成功的可能代价值;

(3) 威胁结点分析评估:

步骤1: 找出入侵成功率相对较高的若干路径, 记录各条边在入侵路径中出现的次数;

步骤2: 找出出现次数较多的边, 并记录下这些边上的两端结点;

步骤3: 按照结点出现在威胁路径上的次数和结点本身的重要性, 对结点的威胁等级进行排序并指出相应结点的配置错误或脆弱性信息;

(4) 总体态势评估:

有如下优点:①该算法所得出的评估结果与人的正常推理和经验值都较为吻合;②该算法并不是基于网络中某一条入侵路径进行量化评估,而是基于起始点和目标点中的整个网络可达路径集进行量化评估;③该算法的评估结果可以实时反映当前网络的安全状态值。无论是路径或结点的增加或删除,还是结点配置的改变,网络的安全状况都会通过该算法有所体现。

该算法的缺点是:当网络中主机数目达到上百台时,在生成网络入侵状态图时可能会产生状态爆炸问题,影响了网络安全评估的正确性及有效性,目前正在寻找状态爆炸问题的新的解决方案。

7 系统的进一步研究方向

该系统得到了充足的预研经费支持,目前在实验室环境下能够初步进行网络安全信息的收集、分析及评估,基本上取得了预期的结果。它大大节省网络安全管理员的时间与精力,对于发现潜在的网络脆弱性也是十分有益的。本系统的各个子模块以及相关的数学算法还在不断的完善中,一些功能实现还存在不足之处,系统的智能功能、自动化功能以及显示功能也需要进一步加强,为了进一步推动该课题的进展,以下几个问题将是我们下一步急需研究和解决的:①进一步研究分析如何提取网络安全量化评估中的各种量化指标要素,使其具有科学性和有效性;②进一步研究网络安全脆弱性之间的依赖关系,从条件概率的角度结合证据推理理论建立数

学模型,更加准确地评估漏洞存在的可能性;③结合人工智能的相关技术,进一步提高系统的智能性;④加快研发系统软件,增强系统的自动化和图形显示功能。

参考文献

- 1 Thorhuus R. Software fault injection testing:[master thesis]. Ericsson Telecom,Stockholm,Sweden,2000
- 2 Sample C, et al. Quantifying Vulnerabilities In The Networked Environment:Methods and Uses. TISC,2000
- 3 Shake T,Hazzard B,Marquis D. Assessing Network Infrastructure Vulnerabilities to Physical Layer Attacks. MIT Lincoln Laboratory. In: 22nd National Information Systems Security Conf. 1999
- 4 Ortalo R, et al. Experimenting with Quantitative Evaluation Tools for Monitoring Operational Security. LAAS Report 96369,1997
- 5 Dacier M, et al. Quantitative Assessment of Operational Security: Models and Tools. LAAS Research Report 96493,1996
- 6 Shafer L R. Implementing Dempster's rule for hierarchical evidence. Artificial Intelligence,1987,33:271~298
- 7 Saaty T L. How to make a decision:the analytic hierarchy process. European Journal of Operational Research,1990,1 (48):9~26
- 8 Barker R,Kelly G. The Vulnerability Instantiation Methodology prototype. Information Security technical Report, 1998,13:79~86

(上接第88页)

关类似的输入值或者通过使用启发式的方法实现其预测。但由于在计算网格这样一个高度动态的环境中计算网格资源信息提供资源管理系统的资源信息可能已经是过时的信息,因此资源管理系统根据此信息做出的资源预留和 QoS 性能严重下降,并且很有可能产生结点的拥塞。主动式网络中的 QoS 监控和 QoS 适应性能够预料到过载的情况并及时做出分流从而导致全局 QoS 和资源管理系统性能的进一步改善。

参考文献

- 1 Baker M, Buyya R,Laforenza D. The Grid: International Efforts in Global Computing. In: Intl. Conf on Advances in Infrastructure for Electronic Business, Science, and Education on the Internet (SSGRR 2000), l'Aquila, Rome, Italy, July 31 - August 6. 2000
- 2 Foster I,Kesselman C,eds. The Grid: Blueprint for a New Computing Infrastructure. Morgan Kaufman, 1999
- 3 Foster I,Roy A,Sander V. A Quality of Service Architecture that Combines Resource Reservation and Application Adaptation. In: 8th Intl. Workshop on Quality of Service (IWQOS 2000), June 2000
- 4 Buyya R, Abramson D,Giddy J. An Economy Grid Architecture for Service-Oriented Grid Computing. In: 10th Intl. Heterogeneous Computing Workshop (HCW 2001) (In conjunction with IPDPS 2001), San Francisco, California, USA

- 5 Kon F,Campbell R,Mickunas M,Nahrstedt K. 2K: A Distributed Operating System for Dynamic Heterogeneous Environments. In: 9th IEEE Intl. Symposium on High Performance Distributed computing (HPDC'9) August 2000
- 6 Buyya R,Chapin S,DiNucci D. Architectural Models for Resource Management in the Grid. In: First IEEE/ACM International Workshop on Grid Computing (GRID 2000), Springer Verlag LNCS Series, Germany, Dec. 17, 2000, Bangalore, India
- 7 Gehring J,Streit A. Robust Resource Management for Metacomputers. In: 9th IEEE Intl Symposium on High Performance Distributed Computing, 2000
- 8 Maheswaran M,Krauter K. A Parameter-based Approach to Resource Discovery in Grid Computing Systems. In: 1st IEEE/ACM Int'l Workshop on Grid Computing (Grid'00), Dec. 2000
- 9 Krauter K,Buyya R,Maheswaran M. A Taxonomy and Survey of Grid Resource Management Systems. Software Practice and Experience,2002,32(2):135~164
- 10 Krauter L, Maheswaran M. Architecture for a Grid Operating System. <http://www.cs.umanitoba.ca/~anrl/PUBS/grid2000>
- 11 Krauter K,Maheswaran M. Towards a High Performance Extensible Grid Architecture. <http://citeseer.nj.nec.com/410299.html>
- 12 Czajkowski K,Sander V. GGF Scheduling Working Group Sched-WD 12.1. Grid Resource Management Protocol: Requirements. Sept. 11, 2001