

基于目标图的入侵检测报警关联算法

杨晓君, 张凤斌, 晏义威

(哈尔滨理工大学 计算机科学与技术学院, 黑龙江 哈尔滨 150080)

摘要:针对入侵检测系统响应能力差以及误报率高的问题,提出了基于目标图的入侵检测报警关联算法。该算法在动作节点之间实现报警信息的前因后果链关系,并根据报警信息显式地跟踪系统状态变化以及攻击者的主观状态,监视状态的变化,推断攻击者的意图以及攻击策略。实验表明,该方法产生的关联结果能够很好地发现攻击者意图以及攻击策略,并且能够有效地降低入侵检测系统的误报率。

关键词:报警关联;入侵检测;目标图

中图分类号:TP393.08

文献标识码:A

文章编号:1673-629X(2009)02-0154-04

Intrusion Detection Alert Correlation Algorithm Based on Goal Graph

YANG Xiao-jun, ZHANG Feng-bin, YAN Yi-wei

(Computer Sci. & Tech. College, Harbin University of Sci. & Tech., Harbin 150080, China)

Abstract: In respect to the problems of the lowly ability of intrusion response and the high false-rate of intrusion detection, proposes a novel alert correlation algorithm based on goal graph, which maintains the precondition and the effect condition in the action-level, traces the system states and the attack states by the information of alert, observing the changes of the state, and infers the attack intention and method. The experiments show that this method can recognize the attack intention from the correlated alerts correctly and effectively, and is more efficient to reduce the false-rate.

Key words: alert correlation; intrusion detection; goal graph

0 引言

随着网络的发展,安全问题越来越受到大家的关注,入侵检测系统(intrusion detection system, IDS)日益成为一种检测网络安全的手段^[1,2],但是,现在的IDS普遍存在如下缺陷:

(1) 入侵检测的误报率偏高,存在大量的重复报警。IDS一天能够发出上千条虚假报警信息,对某主机的一次扫描也会产生几十上百的重复报警,让网络管理人员目不暇接,无所适从。

(2) 入侵响应能力差。目前,大部分的入侵检测系统的响应手段只限于检测到入侵后发出报警信息,将入侵响应的工作留给网络管理人员去手工完成。

因此,如何对大量的报警进行合并,关联,准确认识网络状况具有重要的意义。

国外 Valdes 等人在 2001 年提出启发式和盖然论的方法^[3]解决报警聚类关联问题。Porras 等人在 M-Correlator 中提出了威胁分析方法,实现了报警聚类^[4]。Goldman 等人在 SCYLL SRUS 系统中通过建立假设检验模型进行报警分类^[5]。Julish 在 2003 年提出了基于面向属性归纳的启发式聚类算法^[6]。Valeur 等人提出了进行任务分割的报警关联处理的观点^[7], Peng Ning 等人提出了基于报警前因与后果关系的关联技术^[8,9], Dong Yu 等人提出了基于隐着色的 Petri-Net 报警关联技术^[10]。

在国内,入侵检测报警融合与关联研究也有一定的发展。穆成坡等人提出了基于模糊综合评判的融合算法^[11],对重复报警进行融合,起到了很好的作用。龚剑等人提出了基于实时聚类的冗余消除算法^[12],对短时间的重复报警,起到了很好的效果,但不能发现入侵的行为策略与攻击意图。段海新等人提出了基于地址关联图的分布式 IDS 报警关联算法^[13]。

这些技术是入侵检测报警处理的现代技术,它们都在某一方面具有优势。存在如下缺点:在对报警信息进行融合时,造成信息“损失”严重,不利于分析入侵

收稿日期:2008-05-26

基金项目:黑龙江省普通高校毕业生学术骨干支持计划资助项目(1151G012)

作者简介:杨晓君(1983-),男,河南许昌人,硕士研究生,研究方向为网络安全与网络应用;张凤斌,博士,教授,研究方向为网络安全与网络应用。

行为;且上述的关联技术属于无状态关联技术,在处理报警信息时不考虑攻击发生的系统环境以及攻击者的主观状态,仅依靠观察到的动作来研究关联。它们忽略了系统的动态特性,仅按照特定的规则或模式来关联报警。这种方法对于由工具发起的攻击有较好的效果,但不能很好地适用于人为发起的攻击。

基于以上缺点,改进了基于前因后果的报警关联技术的缺点,提出了基于目标图的入侵检测报警关联方法。该方法在动作节点之间实现报警信息的前因后果链关系,并根据报警信息显式地跟踪系统状态变化以及攻击者的主观状态,监视状态的变化,推断攻击者的意图以及攻击策略。实验表明,该方法产生的关联结果能够很好地发现攻击意图以及攻击策略,并且能够有效地降低入侵检测系统的误报率。接下来介绍报警数据处理模型以及方法的实现。

1 报警数据处理模型

1.1 报警数据处理模型分析

报警数据处理模型如图1所示,该模型包括报警信息读取模块,报警归并模块,报警关联模块。

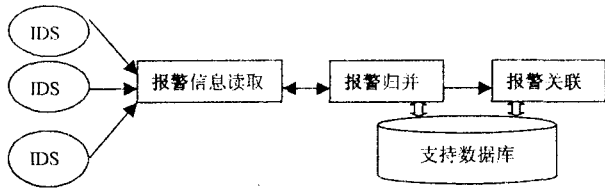


图1 报警数据处理系统结构

报警信息读取模块负责报警信息的接收和报警数据格式的统一,它可从现有的多个IDS上直接接收报警信息,也可从报警数据库中读取数据,并把IDS形成的不同报警信息转化为统一的格式,这里采用IDMEF(Intrusion Detection Message Exchange Format)。报警归并是把满足一定相似程度的报警进行融合,形成一个元报警,以降低重复警报的数量。报警关联是把可能属于同一攻击过程的报警关联起来,以发现攻击策略与攻击意图。文中基于目标图的功能与报警关联所到达的功能的相似性,形成基于目标图的入侵检测报警关联技术,对元报警进行因果关联。

1.2 基于目标图的报警关联模型

入侵检测报警关联的目的是通过分析报警信息,来发现报警之间的逻辑关系,揭示攻击者的意图与策略,同时降低入侵检测的误报率。然而,目标图的功能与报警关联分析所要达到的功能具有一致性,因此,文中在目标图的基础上,形成了基于目标图的入侵检测报警关联技术。模型如下:

定义1 一个四元组 $T = \{P, A, G, E\}$, 其中:

P 为状态节点集合,每个元素表示为 $\text{prop}(p, i)$, p 是一个状态实例,取值“true”或者“false”, i 为一个时间戳; A 为具体动作节点集合,每个元素表示为 $\text{action}(a, i)$, a 是一个具体的动作实例; G 为目标节点集合,每个元素表示为 $\text{goal}(g, i)$, g 是一个目标实例; E 为边集,有如下四种不同的边:

(1) 前提条件边: $\text{precondition_edge}(\text{prop}(p, i), \text{action}(a, i))$, 表示状态节点 $\text{prop}(p, i)$ 是动作节点 $\text{action}(a, i)$ 的前提条件;

(2) 后果状态边: $\text{effect_edge}(\text{action}(a, i), \text{prop}(p, i + 1))$, 表示动作节点 $\text{action}(a, i)$ 发生导致后果状态节点 $\text{prop}(p, i + 1)$;

(3) 状态保持边: $\text{persistence_edge}(\text{prop}(p, i), \text{prop}(p, i + 1))$, 表示状态节点 $\text{prop}(p, i)$ 保持到 $i + 1$ 新的时间片 $\text{prop}(p, i + 1)$;

(4) 目标描述边: $\text{description_edge}(\text{prop}(p, i), \text{goal}(g, i))$, 表示状态节点 $\text{prop}(p, i)$ 是目标节点 $\text{goal}(g, i)$ 的组成状态。

定义2 给定动作 a_i 与 a_j , i, j 表示时间戳,且 $i < j$, 当且仅当 a_i 造成的结果是 a_j 发生的前提条件,则 a_i 与 a_j 之间存在因果链,记作 $a_i \rightarrow a_j$ 。

文中提出的方法分两个阶段实现:第一个阶段是目标图的构建,它以前一个时间片的目标图,观察到的事件为输入,对动作节点、目标节点进行扩张,生成该时间片结束时的新的目标图;第二个阶段是目标图的分析,也就是报警的关联分析阶段。它对新的目标图中分析识别已到达的或部分到达的目标,并提取与观察事件相一致的规划和攻击目标。

第一个阶段的目标图构建分为动作扩展与目标扩展。动作扩展以第 i 个时间片的目标图,观察到事件集 A_i , 规则模式集 A 为输入,对于每一个动作,添加到动作集合 A_0 中,按照规则模式集,得到每个动作的前提条件集合、后果集合。接着构建各个动作节点的前提条件边、后果边,并添加到边集合 E 中。对于第 i 个时间片的状态节点层中所有没有改变的状态节点都保持到第 $i + 1$ 个时间片,并连接之间的保持边。

Action_Expansion($\langle P, A_0, G_R, E \rangle, A_i, i, A$)

For 每一个 $a_i \in A_i$

添加 $\text{action}(a_i, i)$ 到集合 A_0 ;

通过 A 获得 a_i 的前提条件集合 S_P 与后果集合 S_E, S_P', S_E' 是 S_P, S_E 的等价集合

For 每一个 $p_P \in S_P', p_P$ 等于 $\text{not}(p_P')$

If $\text{prop}(\text{neg}(p_P'), i) \in P$

添加 $\text{precondition_edge}(\text{prop}(\text{neg}(p_P'), i), \text{action}(a_i, i))$ 到集合 E ;

```

For 每一个  $p_P \in S_P', p_P$  不等于  $\text{not}(p_P')$ 
  If  $\text{prop}(p_P, i) \in P$ 
    添加  $\text{precondition\_edge}(\text{prop}(p_P, i), \text{action}(a_i, i))$  到
    集合  $E$ ;
  For 每一个  $p_e \in S_E$ 
    添加  $\text{prop}(p_e, i + 1)$  到集合  $P$ 
    添加  $\text{effect\_edge}(\text{action}(a_i, i), \text{prop}(p_e, i + 1))$  到集
    合  $E$ ;
  For 每一个  $\text{prop}(p, i) \in P$ 
    If  $\text{prop}(\neg p, i) \in P \&\& \text{prop}(p, i + 1) \in P$  then
      添加  $\text{prop}(p, i + 1)$  到集合  $P$ 
      添加  $\text{persistence\_edge}(\text{prop}(p, i), \text{prop}(p, i + 1))$ ;
  Return  $\langle P, A_O, G_R, E \rangle$ 

```

目标扩展是以第 i 个时间片的目标图, 攻击目标描述集 G 为输入。首先对每个攻击目标实例化, 获得攻击目标描述实例集, 添加描述边, 添加目标到目标图。当一个目标添加到第 i 个时间片的目标层, 且被第 i 个时间片的状态层满足, 则其是一个到达的目标。通过判断第 i 个时间片的状态层是满足目标的部分描述边, 还是全部描述边, 来决定这个目标是部分达到, 还是完全达到。

```

Goal_Expansion( $\langle P, A_O, G_R, E \rangle, i, G$ )
  For 每一个  $G_k \in G$ 
    For 每一个  $G_k$  的实例  $g$ 
      得到目标表述集合  $S_g$ , 实例化  $S_g$ , 得到  $S'_R$ ;
      For 每一个  $p_g \in S'_R, p_g$  等于  $\text{not}(p_g')$ 
        If 每一个  $\text{prop}(\text{neg}(p_g'), i) \in P$ 
          添加  $\text{description\_edge}(\text{prop}(\text{neg}(p_g'), i), \text{goal}(g, i))$ 
          到集合  $E$ ;
        For 每一个  $p_g \in S'_R, p_g$  不等于  $\text{not}(p_g')$ 
          If  $\text{prop}(p_g, i) \in P$ 
            添加  $\text{description\_edge}(\text{prop}(p_g, i), \text{goal}(g, i))$  到集
            合  $E$ ;
          If 目标  $g$  满足条件描述边集条件
            添加  $\text{goal}(g, i)$  到  $G_R$ 
  Return  $\langle P, A_O, G_R, E \rangle$ 

```

目标图中有效规划描述了攻击者完成某一目标所采用的策略, 也可以理解为对报警的解释, 因为它反映了一组报警之间的关系, 以及它们所对应的攻击动作要完成的任务。生成有效的规划是报警关联要完成的基本工作。

下面介绍生成有效规划的关联算法: G_R 是一个目标集合, 用三元组 $\langle g_i, \langle A_O, O, L_a \rangle, L_g \rangle$ 表示已到达的目标 g_i 和到达 g_i 的一个有效规划, L_g 是一组动作和目标 g_i 之间的因果链, GoalPlan 表示已识别的有效规划集合。算法首先找到与已完成的每个目标相关的动作集, 对每个与目标相关的动作, 通过因果链连接到该动作相关的动作, 重复这个过程, 直到没有相关

的动作被找到。

```

Alert_Correlation( $\langle P, A_O, G_R, E \rangle, t$ )
  For 每一个  $g_i \in G_R$ 
    初始化  $A_O' = \emptyset, A = \emptyset, L_a = \emptyset, L_g = \emptyset$ ;
    For 每一个  $a_i \in A_O$  且与  $g_i$  相连
      添加  $a_i \rightarrow g_i$  到集合  $L_g, a_i$  到集合  $A_O', a_i$  集合  $A$ 
    Do { If  $A = \emptyset$ , and  $a_i \in A_O, a_i \in A_O'$ 
      获得时间约束集合  $O$ , 添加  $\langle g_i, \langle A_O, O, L_a \rangle, L_g \rangle$ 
       $>$  到  $\text{GoalPlan}$ 
      If  $A \neq \emptyset$ 
        删除集合  $A$  中的  $a_i$ 
      For 每一个  $a_i \in A_O$  且与  $a_i$  相连
        添加  $a_i \rightarrow g_i$  到集合  $L_a$ 
      If  $a_i$  不属于  $A_O'$ 
        添加  $a_i$  到集合  $A_O', a_i$  集合  $A$  } while (只要有相关的
    动作)
  Return  $\text{GoalPlan}$ ;

```

1.3 算法分析

对于这两个阶段的算法, 它们的空间与时间复杂度都是多项式级规模的。目标图构建阶段, 给定一个 t 时间片的目标图, 动作节点 s 个, 初始状态 p 个, 目标节点 m 个, 动作节点的最大效果边 l_1 , 目标节点的最大描述边 l_2 , 最大实体数 n 个, 则生成 $t + 1$ 时间片的目标图, 创建状态节点为 $O(p + l_2 s)$, 生成的边集以及节点集分别是 $O(l_2 mn), O(nm)$; 可见, 目标图构建过程的时空复杂度是多项式规模的。

报警关联分析阶段, 给定一个 t 时间片的目标图, 部分达到或完全达到的目标节点 l_1 , 一个目标的最大描述边 m_1 , 动作节点 l_2 , 一个动作节点的最大前提条件边 m_2 。对于一个目标节点, 它访问动作节点的路径与该目标节点的因果链的数目有关, 是 $O(m_1)$ 。对于每一个与该目标相连的动作节点, 被访问的次数与该动作节点所具有的因果链有关, 是 $O(m_2)$, 因此, 可能的因果链是 $O(l_1(m_1 + l_2 m_2))$ 。可见, 报警关联分析阶段的时空复杂度为多项式规模的。

2 实验结果

用 2000 年 DARPA 入侵检测系统测试数据集来验证该算法。这个数据集是 MIT 的 Lincoln 实验室在实际网络环境中进行攻击实验而产生的, 以 MIT 日常网络流量作为背景流量, 用于评估入侵检测系统的性能。2000 年 DARPA 数据集包括两个攻击实例: LLDOS 1.0 和 LLDOS 2.0.2。LLDOS 1.0 攻击是利用了 solaris 系统的 Sadmin 漏洞, 然后利用远程缓冲区溢出攻击该漏洞, 再通过 Rsh 登录到攻击源主机上拷贝并安装拒绝服务攻击代理。攻击者一共控制了 3 台主机, 然后发起分布式拒绝服务攻击。LLDOS 2.0.2 的

表 1 入侵检测结果

数据集		可观察数量	用的方法	报警数量	检测到攻击数	正确报警数量	检测率	误报率
LLDOS 1.0	DMZ	89	RealSecure	891	51	57	57.30%	93.60%
			Our method	57	50	54	56.18%	5.26%
	Inside	60	RealSecure	922	37	44	61.67%	95.23%
			Our method	44	36	41	60%	6.82%
LLDOS 2.0.2	DMZ	7	RealSecure	425	4	6	57.14%	98.59%
			Our method	5	3	4	42.86%	20%
	Inside	15	RealSecure	489	12	16	80.00%	96.73%
			Our method	13	10	12	66.67%	7.69%

攻击过程与 LLDOS 1.0 类似,区别是漏洞探测的方式采用了更加隐蔽的方式,而拷贝攻击代理软件采用了 ftp 上载的方法。DARPA 数据集还包括用:tcpdump 分别在 DMZ(非军事区)和内部网中监听到的全部数据包。文中用 RealSecure Network Service 6.0 对 4 个数据集进行分析,再用文中的关联算法来处理报警。由于关联结果生成的图较多,在此给出其中一个进行说明。入侵检测结果见表 1,关联结果见图 2。

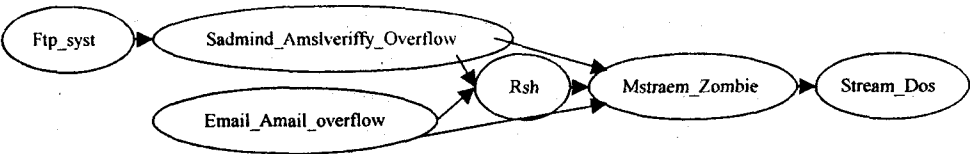


图 2 报警关联结果

从结果中可以看到,文中的方法在发现攻击意图时,能够把属于同一攻击的报警关联起来,揭示攻击者的意图与策略,这为下一步的入侵响应奠定了基础。同时能够在保证检测率基本不变的情况下,使误报率降低。

3 结束语

文中基于目标图的入侵检测报警关联算法能够从大量底层入侵报警信息中识别攻击意图,能够对攻击过程的状态进行跟踪,从而能够有效地对报警进行关联并清晰地体现攻击流程;同时能够保证报警关联的完备性。对于孤立的节点,它们有可能是由于攻击者发起的独立的攻击而产生正确的报警,也有可能是没有被我们的算法关联而产生的孤立节点,接下来将引入不确定知识表示方法和推理机制对其进行推理;同时刻画更加详细的抽象攻击模式,进一步验证本算法,提高其实用性。

参考文献:

[1] Wang Y, Behera S R, Wong J, et al. Towards the Automatic Generation of Mobile Agents for Distributed Intrusion Detec-

tion System[J]. Journal of Systems and Software, 2006, 79 (1):1-14.

[2] Rawat S, Arun K, Pujari A K, et al. On the Use of Singular Value Decomposition for a Fast Intrusion Detection System [J]. Electronic Notes in Theoretical Computer Science, 2006, 142(3):215-228.

[3] Valdes A, Skinner K. Probabilistic Alert Correlation[C]//Fourth International Symposium on Recent Advances in Intrusion Detection. RAID 2001. Berlin: Springer Press, 2001:54-68.

[4] Porras P A, Fong M W, Valdes A. A Mission-Impact based Approach to INFOSEC Alarm Correlation[C]//Fifth International Symposium on Recent Advances in Intrusion Detection. Berlin: Springer Press, 2002:95-114.

[5] Goldman R P, Heimerdinger W, Harp S A, et al. Information Modeling for Intrusion Report Aggregation[C]//DARPA Information Survivability Conference Exposition DISCEX. New York: IEEE Press, 2000:329-342.

[6] Julish K. Clustering Intrusion Detection Alarms to Support Root Cause Analysis[J]. ACM Transactions on Information and System Security, 2003, 6(4):443-471.

[7] Valeur F, Vigna G, Kruegel C, et al. Comprehensive Approach to Intrusion Detection Alert Correlation[J]. IEEE Transactions on Dependable and Secure Computing, 2004, 1(3):146-149.

[8] Ning Peng, Cui Yun, Reeves D S. Constructing attack scenarios through correlation of intrusion alerts[C]//In Proceedings of the 9th ACM conference on Computer and communications security. Washington D. C.: [s. n.], 2002:245-254.

[9] Ning Peng, Xu Dingbang. Learning attack strategies from intrusion alerts[C]//In Proceedings of the 10th ACM conference on Computer and communications security. Washington D. C.: [s. n.], 2003:200-209.

[10] Yu Dong, Frincke P D. Improving the quality of alerts and pre-

4 基于身份认证安全机制问题与解决方法

基于身份认证中的私钥是由 TA 生成,所以并不是只有用户才知道自己的私钥,密钥托管的问题与生俱来,这样就存在一些安全隐患,比如说主密钥泄漏将引起整个系统崩溃等,恶意的 TA 可以伪造任何用户的签名。

另外私钥与证书一样需要一定的有效期,这里对私钥的更新也是需要研究的一个问题。

4.1 密钥托管问题

因为私钥是集中产生后分发给用户,TA 知道所有用户的私钥。为了解决这个问题,可以构造一种新的实体 PKG (Private Key Generator) 用来生成并分发用户私钥。

系统可以分成不同的安全域,每个域内会有一个 PKG_i , 用来为用户生成部分私钥 $s_i Q_{ID}$, s_i 为 PKG_i 的主密钥。域内的通信定义为安全的,域间的通信可以用发送方和接收方的 PKG 协同生成用户私钥 $d_{ID} = s_i Q_{ID} + s_j Q_{ID}$, 只有两个 PKG 共谋时才能知道用户的私钥,增加了密钥的机密性,一个 PKG 受到攻击时也不会使整个系统崩溃^[8]。

采用多个 PKG 共同生成用户私钥的方法,安全性的提高和计算复杂度的增加是成正比的,这里采用了一种折中的办法。简单的证书由 TA 颁发只有在向 PKG 申请的部分私钥是才会用到,用来证明用户的合法身份^[9]。

4.2 密钥管理

对于私钥的管理,在 PKI 机制中注销和更新证书过程还是相对复杂,工作量很大的。同一个私钥用的时间过久会降低安全性,而且对于 SIP 系统中的用户代理对服务的使用权需要有一定的有效期,所以密钥更新是系统需要解决的一个问题。

在文中提到的系统中密钥是根据用户身份和主密钥共同生成的,如果重新生成主密钥将会导致整个系统私钥的变化,但是用户的身份也是广播到各个用户,变化起来也会比较麻烦。在文献[8]中描述了一种解决方法,可以把有效期写在 ID 后面用来生成公钥,

$Q_{ID} = h((ID, \text{current} - \text{year}))$, 再根据公钥生成私钥,超过有效期的私钥就不起作用,用户需要重新向 PKG 申请新的私钥。应用在 SIP 中可以在 SIP 头中添加一个表示有效期的头字段,在发送请求消息的同时把有效期传递给目标服务器。这种方案使私钥的管理简单方便,减轻了密钥管理的负担,同时增强了系统的安全性。

5 结束语

介绍了 SIP 协议的安全需要,基于身份认证的原理和特点,提出了一种基于身份加密的 SIP 认证方案,保持了基于身份认证的结构简单、维护容易等特点,不需要复杂的证书管理和维护。对于身份认证本身固有的密钥托管问题,提出了一种简单的缓解方案。提高了 SIP 的安全性能,同时又能保持系统的高效性。

参考文献:

- [1] Roseberg J, Schulz I H, Camar I G. SIP: Session Initiation Protocol[S]. IETF RFC 3261, 2002.
- [2] 刘 华, 王 琨. 基于 PKI 的 SIP 协议安全的研究[J]. 电子科技, 2005(2): 37-40.
- [3] 李士达, 胡 玢, 王兴秋, 等. 一种基于 ECC 的 SIP 认证方案的提出与实现[J]. 计算机应用, 2007, 27(2): 311-313.
- [4] 庞红玲, 安 可, 戎锋洪. 基于身份加密系统的 SIP 认证机制[J]. 信息安全与通信保密, 2007(5): 133-135.
- [5] 金康双, 王泽兵, 冯 雁, 等. SIP 协议的认证机制及其性能分析[J]. 计算机应用研究, 2004(8): 110-112.
- [6] Shamir A. Identity - based Crypto Systems and Signature Schemes[C]//Advance in Cryptology - crypto'84. Germany: Springer - Verlag, 1984: 47-53.
- [7] Boneh D, Franklin M. Identity - based Encryption from the Weil Paring in Advance[C]//Cryptology - Crypto'01. Germany: Springer - Verlag, 2001: 213-229.
- [8] 徐茂智, 游 林. 信息安全与密码学[M]. 北京: 清华大学出版社, 2007.
- [9] 李新国, 葛建华, 赵春明. IBE 公钥加密系统的用户私钥分发方案[J]. 西安电子科技大学学报: 自然科学版, 2004, 31(4): 569-573.

(上接第 157 页)

- dicting intruder's next goal with Hidden Colored Petri - Net Computer Networks[J]. The International Journal of Computer and Telecommunications Networking, 2007, 51(3): 632-654.
- [11] 穆成坡, 黄厚宽, 田盛丰, 等. 基于模糊综合评判的入侵检测报警信息处理[J]. 计算机研究与发展, 2005, 42(10):

1679-1685.

- [12] 龚 俭, 梅海彬, 丁 勇, 等. 多特征关联的入侵事件冗余消除[J]. 东南大学学报, 2005, 35(3): 366-371.
- [13] 段海新, 于雪利, 王兰佳. 基于地址关联图的分布式 IDS 关联算法[J]. 大连理工大学学报, 2005, 45(10): 126-131.