doi:10.11959/j.issn.1000-436x.2017021

面向多步攻击的网络安全态势评估方法

杨豪璞，邱辉，王坤
（信息工程大学三院，河南 郑州 450001）

摘 要：为了分析多步攻击对网络系统的影响，准确、全面地反映系统的安全态势，提出一种面向多步攻击的网络安全态势评估方法。首先对网络中的安全事件进行场景聚类以识别攻击者；对每个攻击场景因果关联，识别出相应的攻击轨迹与攻击阶段；建立态势量化标准，结合攻击阶段及其威胁指数，实现对网络安全态势的评估。通过对 2 个网络攻防实验的测评分析表明，所提出的多步攻击分析方法符合实际应用，评估结果准确、有效。
关键词：场景聚类；多步攻击；安全态势；量化分析
中图分类号：TP393.08 **文献标识码**：A

Network security situation evaluation method for multi-step attack

YANG Hao-pu, QIU Hui, WANG Kun
(The Third Institute, Information Engineering University, Zhengzhou 450001, China)

Abstract: Aiming at analyzing the influence of multi-step attack, as well as reflecting the system's security situation accurately and comprehensively, a network security situation evaluation method for multi-step attack was proposed. This method firstly clustered security events into several attack scenes, which was used to identify the attacker. Then the attack path and the attack phase were identified by causal correlation of every scene. Finally, combined with the attack phase as well as the threat index, the quantitative standard was established to evaluate the network security situation. The proposed method is assessed by two network attack-defense experiments, and the results illustrate accuracy and effectiveness of the method.
Key words: scene clustering, multi-step attack, security situation, quantification analysis

1 引言

随着计算机网络与通信技术的迅猛发展，互联网已经渗透到人类社会的各行各业，并影响着经济的发展和人类的生活方式。网络在提供各种便利的同时，也带来了诸多安全问题。目前，网络攻击行为日益猖獗，对国家、政府、企业、个人造成了极大的经济损失。因此，在复杂多变的网络环境中认知、理解并预测网络的安全状态及其发展趋势，有助于管理人员及时掌握网络安全状况，并对未来可能出现的威胁提前做出防护，减小攻击对网络的危害。

为获取网络的安全状态，研究者起初分别从攻击威胁[1, 2]、网络脆弱性[3, 4]等方面进行识别与评估。该方面的研究均比较成熟，但这些研究仅从单一的要素进行评估，已经无法满足管理人员对掌握网络整体安全状态的需求。网络安全态势感知作为一种新型主动防御技术，通过融合入侵检测系统（IDS, intrusion detection system）、防火墙、病毒检测系统（VDS，virus detection system）等网络设备的安全防护信息，以提高网络攻击的实时检测能力与网络安全的整体感知能力，对网络安全状况进行实时准确的评估与预测。因此，网络安全态势感知技术近年来已逐渐成为网络安全领域的研究热点。态势感知的概念由 Endsley[5]首次提出，该技术最先应用于

收稿日期：2016-03-22；**修回日期**：2016-08-03
基金项目：国家自然科学基金资助项目（No.61303074, No.61309013）；国家重点基础研究发展计划（“973”计划）基金资助项目（No.2012CB315900）
Foundation Items: The National Natural Science Foundation of China (No.61303074, No.61309013), The National Basic Research Program of China (973 Program)(No.2012CB315900)

航天飞行、军事、交通监管以及医疗应急调试等领域。1999 年，Bass[6]将态势感知的概念引入到网络安全领域。目前，对网络安全态势感知的研究主要集中在对当前网络态势的理解，即网络态势评估方面，以及对未来一段时间内网络态势的预测。本文研究重点在于对网络安全态势的评估。

目前，许多学者已经提出了相应的网络安全态势评估方法，如基于数学模型[7, 8]、知识推理[9, 10]、模式识别[11, 12]和博弈模型[13, 14]的评估方法，但以上方法均以单个攻击事件作为评估基础，没有挖掘安全事件之间的因果关系。而近年来网络攻击行为逐渐呈现出大规模、协同、多阶段等特点，网络攻击不再是一个个孤立事件。根据国家计算机网络应急技术处理协调中心的年度网络安全工作报告中把攻击按类型进行统计的结果，大部分攻击尤其是危害巨大的攻击几乎都是多步攻击。

面对这样的问题，吕慧颖等[15]基于多角度安全信息，挖掘攻击阶段之间的因果关联；Cyril 等[16]则通过还原攻击路径对网络安全态势进行评估。上述方法均以攻击图为基础，无法实现对攻击阶段的识别评估，同时，评估时对所有攻击信息无差别处理，缺乏对攻击者的识别。

通过以上分析，目前，网络安全态势评估方法均能在一定程度上对当前安全状态有所认识，但面向多步攻击时仍存在以下亟待解决的问题：1）缺乏对多步攻击的阶段识别，无法全面评估各攻击轨迹对网络资产所造成的影响；2）缺乏对攻击者的识别，无法准确评估各攻击者对网络造成的影响；3）尚未给出一个合理有效的安全态势量化标准。

同时，部分研究机构或企业厂商对信息系统的脆弱性进行研究，并提出相应的评估标准，如 ISO13335(《信息安全管理方针》)、ISO15408（《信息技术安全性通用评估准则》）和 GB17859（《安全保护等级划分准则》）等，然而，这些标准大部分以定性评估为主，其评估结果的可理解程度有限。由美国基础设施顾问委员会（NIAC）开发的通用安全漏洞评分系统（CVSS），为所有安全漏洞（包括已知的和未知的）的威胁程度提供一个开放、通用、合理的量化评级标准，具有较好的参考价值。

因此，在已有研究成果的基础上，本文从攻击者的角度出发，提出一种面向多步攻击的网络安全态势评估方法。首先对攻击行为进行攻击场景聚类，以识别各攻击者的攻击轨迹。建立攻击模式库，通过对攻击行为进行因果分析，识别攻击者所处的阶段，并将其作为态势评估要素，结合基于 CVSS[17]的量化标准，经过态势要素融合和节点态势融合得到网络整体安全态势。

2 网络安全态势评估基础

网络安全状态基于实时确认的一组攻击轨迹以及各条攻击轨迹对受攻击的网络资产所造成的影响。且攻击者入侵不同的网络环境其入侵方式与影响也不同。因此，对态势评估安全要素的选取包括攻击信息以及网络环境信息。首先对一些相关术语进行定义，其次对网络安全态势评估流程进行描述。

2.1 基本术语定义

定义 1 主机信息。包括网络中各主机以及网络设备，如防火墙等。由于网络设备也有可能成为攻击的对象或跳板，因此，对网络设备也要进行全面分析。描述主机信息使用一个四元组 (*HostIP*, *Services*,*Vuls*,*Weight*) 来表示。其中，*HostIP* 表示主机的 IP 地址，*Services* 表示主机所运行的服务信息（如 SSHD、SQL、HTTP、Ms-office），*Vuls* 表示主机上的脆弱性与漏洞列表，*Weight* 表示该主机在网络中的重要程度。

定义 2 脆弱性集 V。即网络中所有出现的配置错误与漏洞的集合。对于任一脆弱性 $v \in V$，使用一个五元组(*id*, *type*, *IP*, *impact*, *info*)来表示。其中，*id* 表示该脆弱性的唯一标识；*type* 表示该脆弱性的类型，$type \in \{C\_Error, Vulnerability\}$，其中，*C_Error* 表示配置错误类型，如非安全策略、防火墙配置错误、设备接入权限设置错误，*Vulnerability* 表示漏洞类型，其在 BugTraq、CERT/CC 等漏洞库中均有统计；*IP* 表示该漏洞出现的主机地址；*impact* 表示该漏洞对资产造成的危害性；*info* 表示该脆弱性的详细描述信息。

定义 3 拓扑结构。是网络中主机之间的物理连接结构，使用一个无向图 $G(N, E)$来表示。其中，N 表示网络中物理主机节点集合，E 表示连接节点间的边。

定义 4 网络连通性。$conn \subseteq Host \times Host$，即主机与主机之间的通信关系。为保护网络中重要的资产，管理者会设置防火墙访问策略，使外部主机无法访问内部网络，或仅允许通过部分协议与端口进行通信。使用一个三元组 $(host_i, host_j, protocol/port)$ 来描

述网络的通连关系，其中，$host_i$、$host_j$ 表示相连主机，$protocol/port$ 表示双方可以通信的协议与端口。

定义 5　原子攻击事件。是指攻击者在网络中实施的单个攻击动作，其可能是对主机服务的扫描或对主机某个漏洞的利用，使用一个八元组(id, $time$, Sip, Dip, $Sport$, $Dport$, $AttackType$, $p(a)$)来表示。其中，id 是该事件的唯一标识符；$time$ 是该事件的发生时间；Sip 是攻击者的源地址；Dip 是攻击的目标地址；$Sport$ 是攻击者的源端口；$Dport$ 是攻击的目的端口；$AttackType$ 是本次安全事件使用的攻击类型；$p(a)$是经过融合后该攻击事件的发生概率。

定义 6　攻击状态转移图，使用一个四元组 (S,τ,A,ε) 来表示。

1) S 表示状态节点集合。

2) $\tau \subseteq (S_{\text{pre}}, S_{\text{post}})$，$S_{\text{pre}}$、$S_{\text{post}} \in S$。对于 $S_i \in S$，$Pa(S_i)=\{S_j \in S \,|\, (S_j,S_i)\in \tau\}$ 表示状态 S_i 的父节点，$Ch(S_i)=\{S_j \in S \,|\, (S_i,S_j)\in \tau\}$ 表示状态 S_i 的子节点。

3) A 用一个二元组 (τ_i, At_i) 表示，其中，$\tau_i \in \tau$，At_i 表示完成状态转移 τ_i 所必需的原子攻击事件。

4) ε 用一个二元组 (S_i,d_i) 表示，表示攻击间的依赖关系，由攻击类型集合的有序对唯一确定。若 $S_j=\text{true} \Rightarrow \forall S_i \in Pa(S_j)$，$S_i=\text{true}$，则 $d_j=1$，表示该攻击状态的父节点必须全部成功，该攻击阶段才可能实现，该依赖关系为并列关系；若 $S_j=\text{true} \Rightarrow \exists S_i \in Pa(S_j)$，$S_i=\text{true}$，则 $d_j=0$，表示该攻击状态中任意一个父节点成功，该攻击状态即可实现，则该依赖关系为选择关系。

依据该状态转移模型即可对已知的网络攻击模式进行建模，得到攻击模式库。图 1 即为对一个攻击实例建立的攻击状态转移图，在该实例中，状态节点集合为 S={地址嗅探，端口扫描，…，登录}，原子攻击事件包括文件列表 Ping、端到端 Ping、网络探测扫描、“瑞士军刀”工具、登录操作等；以前 2 个状态节点为例，从图 1 中可以看出，地址嗅探是端口扫描的父节点，即若出现状态端口扫描，则状态地址嗅探必然已经成功，两者之间是并列关系。

2.2　网络安全态势评估流程

网络安全态势评估的整体流程如下。

1) 网络安全态势要素收集。通过检测与收集报警数据与网络环境自身运维信息，并对收集到的信息进行标准化规范，从而得到网络安全态势评估所需的要素集。网络安全态势要素集包括攻击方信息、环境信息这 2 类，其中，攻击方信息的来源主要依靠网络中入侵检测系统、防火墙、系统日志等传感器的报警信息，对报警信息进行数据融合得到原子攻击事件；环境信息包括主机信息、拓扑结构、网络连通性。环境信息的收集依靠对网络信息的统计与漏洞扫描系统的结果，其中，拓扑结构依据对网络结构的统计，网络连通性依据网络中防火墙的过滤规则，主机信息依据对运维系统、软件的统计和对主机的漏洞扫描。

2) 网络攻击阶段识别。依据收集到的攻击方信息，对攻击行为进行因果分析，首先对攻击行为进行攻击场景聚类，将已拥有的攻击信息融合为多个安全事件，并依据其关联关系进行场景划分，以便识别各攻击者的攻击轨迹；其次将实时生成的攻击场景与攻击模式库进行关联分析，发现攻击者的攻击阶段。

3) 网络安全态势量化。以网络攻击阶段识别结果为基础，结合网络中各资产信息，并基于 CVSS 设计了量化指标，完成对网络安全态势的评估。

3　实时攻击阶段识别

3.1　攻击场景聚类

防护网络在同一时间段内可能遭受多个攻击者的入侵，而网络被多个攻击者攻击要比被一个攻击者攻击更加危险，如防护网络中某个主机的 root 权限被多个攻击者获得比仅被一个攻击者获得的威胁程度更大。因此，需要将融合得到的安全事件聚类到不同的攻击场景中，并识别攻击者。

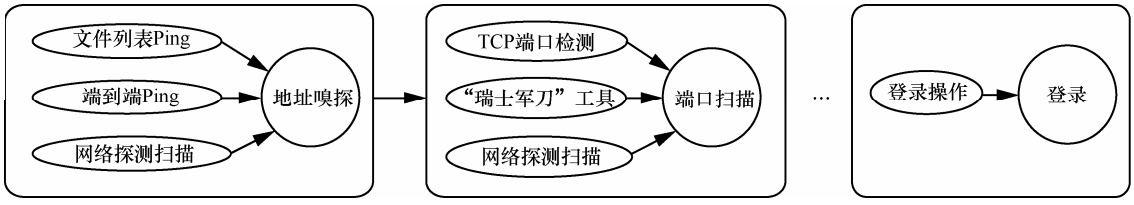


图 1　攻击状态转移

在同一次多步攻击中，因为攻击者的攻击意图和目标明确，其前后攻击步骤引发的安全报警在属性上具有一定的关联关系。如攻击者对攻击进行漏洞扫描的前提攻击事件就是对网络的 IP 地址扫描，且这 2 个攻击事件所引发的报警的源 IP 地址、目的 IP 地址均相同，且在发生时间上也存在前后因果关联。因此，为识别出每一个攻击者的攻击场景，对报警属性间的关联度进行计算。首先对攻击关联度进行定义。

定义 7 攻击关联度 $cor(a,b)$。是指 2 个攻击间的关联程度，用于确定 2 个攻击属于同一攻击场景的可能性。本文抽取源 IP 地址、目的 IP 地址、源端口号、目的端口号、时间与攻击类型这 6 个属性，作为确定攻击关联度的依据。并定义攻击关联度函数为

$$cor(a,b)=\frac{\sum_{k=1}^{n}\alpha_k Feature_k(a,b)}{\sum_{k=1}^{n}\alpha_k} \tag{1}$$

其中，$Feature_k(a,b)$ 和 α_k 分别表示第 k 个特征属性之间的关联度和相应的权重。特征值与相应权重的选取参考文献[18]。

当系统收到一条新的安全事件时，将其与已保存的各个攻击场景进行匹配，计算该安全事件和它们之间的关联度，若存在与多个攻击场景的关联度超过事先设定的关联度阈值，则将该安全事件加入到关联度最大的攻击场景中；若所有关联度均未超过阈值，则将该安全事件另存为新的攻击场景。

3.2 实时攻击阶段识别算法

定义 8 实时攻击场景（ATree）。用于保存攻击者的实时入侵轨迹，使用一个三元组 (S,F,Q) 来表示。其中，S 表示已经发生的攻击状态节点集合，F 表示为状态节点间转移有向边集合，Q 表示状态间的依赖关系，$Q_i \in$ (and,or)。其中，Q_i = and 表示只有状态 S_i 的全部父节点全部入侵成功，S_i 才有可能成功；Q_i = or 表示只要状态 S_i 的任一父节点被成功入侵，S_i 就有可能成功。

定义 9 状态发生函数 bool(s)。该函数用于标识攻击状态发生情况。若该攻击状态已经发生，bool(s)=true；否则，bool(s)=false。

定义 10 转移等待窗口 $\partial\tau$。攻击者对网络的入侵一般有一个攻击周期，若其在长时间内仍未发起后续攻击，则该攻击者的能力无法利用网络中出现的漏洞，入侵失败。为提高对有效攻击的识别，设置一个转移等待窗口来衡量攻击者的成功与否。已知大部分攻击的一个攻击周期 2 h，因此设置 $\partial\tau$ =2 h。

将实时安全事件(Alert)利用攻击关联度进行聚类，得到不同攻击场景的报警集合。对每一个攻击场景的报警信息与生成的攻击模式库进行关联分析，可能出现以下 4 种典型情景，如图 2 所示（其中，实线表示攻击已经发生、状态已经发生或漏洞存在；虚线表示状态未发生或漏洞不存在）。

图 2(a)中攻击的前提状态为真，且目标主机中存在该攻击所利用的漏洞，该情景属于成功状态转移。图 2(b)、图 2(c)为失败状态转移情景，图 2(b)中被攻击目标主机不存在所利用的漏洞；图 2(c)中该攻击的前提状态不为真。图 2(d)为与节点状态转移情景。图 2(e)为或节点状态转移情景。基于上述分析，实时攻击阶段状态识别算法的基本步骤如算法 1 所示。

算法 1 实时攻击阶段识别算法

输入：融合的安全事件（Alert）

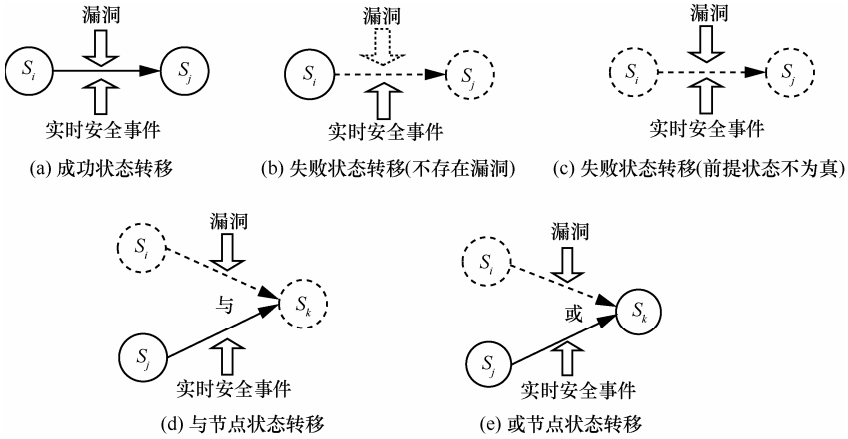


图 2 状态转移典型情景

输出：实时攻击场景（ATree）

1) 等待实时报警信息的发生，若收到新的Alert，则计算其与已生成各实时攻击场景之间的关联度，并利用攻击关联度将报警信息聚类到不同的攻击场景中。

2) 对每一个攻击场景中报警信息集与攻击模式库进行关联分析，搜索得到该攻击信息的前提阶段状态 s_x、后续阶段状态 s_y、后续阶段状态的依赖关系 e 以及攻击利用漏洞 *Vuln*，并记录当前时间 t。

3) 若 $Vuln \notin HostInf$，表明所入侵的目标主机中不存在该攻击利用的漏洞，则该攻击无法成功入侵，状态转移失败。该情景不对实时攻击场景作任何改动，转到步骤 8)。

4) 若 bool(s_x)=flase，表明实时攻击场景中不存在该攻击的前提状态，则该攻击无法成功入侵，状态转移失败。该情景不对实时攻击场景作任何改动，转到步骤 8)。

5) 若 bool(s_x)=true，bool(s_y)=true，且 $path(s_x \to s_y) \in F$，表明该攻击路径已经存在于实时攻击场景中，则该情景属于重复状态转移。该情景不对攻击场景图作任何改动。转到步骤 8)。

6) 若 e=0，表明节点 s_y 为或节点，将阶段状态 s_y 与路径 $path(s_x \to s_y)$ 加入到实时攻击场景，并设置所对应 Q_f = or，bool(s_y)=true，更新状态发生时间 $t_{\text{current}} = t$。转到步骤 8)。

7) 若 e=1，表明节点 s_y 为与节点，将阶段状态 s_y 与路径 $path(s_x \to s_y)$ 加入到实时攻击场景，并设置所对应 Q_f = and，更新状态发生时间 $t_{\text{current}} = t$。依据攻击模式库判断实时攻击场景 ATree 中状态 s_y 的前提条件是否全部满足，若已全部成功转移，则设置 bool(s_y)=true。转到步骤 8)。

8) 判断所有攻击场景的状态转移时间是否超时，如果 $t - t_{\text{current}} > \partial\tau$，则将该攻击场景删去。转到步骤 1)。

3.3　攻击阶段识别算法的改进

通过实时攻击阶段识别算法，即可以将网络安全事件识别出攻击场景。但实际上，漏报、零日漏洞以及报警乱序等异常情景都会对攻击场景的生成造成影响，如图 3 所示。图 3(a)为漏报问题情景，因安全事件 *Alert_x* 漏报，导致状态 S_j 不为真，进而影响状态 S_k 的生成。由于攻击检测策略与实际入侵行为的特征之间会存在一定的差异，入侵检测设备在报警时常会产生漏报现象。图 3(b)为零日漏洞问题情景，因安全事件 *Alert_x* 所利用的漏洞为零日漏洞，导致检测时判断其无法成功实施，从而状态 S_j 不为真，并进一步影响状态 S_k 的生成。在网络系统中存在着大量的零日漏洞尚未被发现公布，而成熟的攻击者经常通过挖掘系统的零日漏洞对网络进行入侵，因此该情景较为常见。图 3(c)为报警乱序情景，安全事件 *Alert_y* 的检测发生时间（*Time_y*）比 *Alert_x* 的检测发生时间（*Time_x*）早，在对 *Alert_y* 进行场景生成时，因状态 S_j 不为真，从而导致状态 S_k 无法成功转移。出现该情景多为报警数据在网络传输中出现时延，或不同安全传感器的时钟出现异步。

为解决以上问题，将状态发生函数 bool(s)进行拓展，bool(s)∈{true,false,middle}，增加一个中间状态，用以保存可能发生的状态转移，并通过后续报警进行状态更正。对实时攻击阶段识别算法的改进如下。

Step 1　设置中间状态。

1) 若 $Vuln \notin HostInf$，bool(s_x)=true，表示目标主机上不存在攻击利用漏洞，且其前提状态为真，该情景有可能出现零日漏洞问题。设置 bool(s_y)= middle。

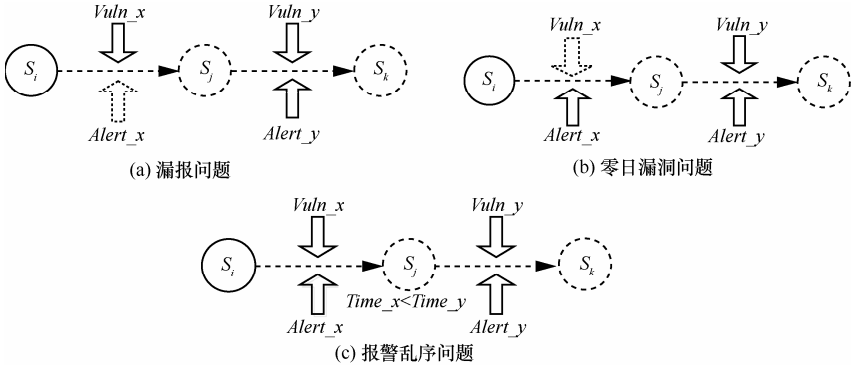


图 3　攻击阶段识别异常情景

2) 若 $Vuln \in HostInf$，bool(s_x)=false，表示目标主机上存在攻击利用漏洞，但其前提状态不为真，该情景可以为漏报或报警乱序问题。查找状态 s_x 的发生主机是否存在所需漏洞使其达到该状态，若存在，设置 bool(s_x)=middle、bool(s_y)=middle。

Step 2 状态更正。

1) 若 $Vuln \in HostInf$，bool(s_x)=middle，表示目标主机上存在攻击利用漏洞，其前提状态为中间状态。由于攻击者已经能够利用状态 s_x 为前提条件，则认为该状态已经发生，因此，设置 bool(s_x)=true、bool(s_y)=true。该情景是对漏报与零日漏洞问题的更正。

2) 若 $Vuln \in HostInf$，bool(s_x)=true，bool(s_y)=middle，表示目标主机上存在攻击利用漏洞，其前提状态为真，后续状态为中间状态。该情景是对报警乱序问题的更正。设置 bool(s_y)=true，并将状态 s_y 的所有后续节点的状态也设置为 true。

4 网络安全态势量化分析

4.1 攻击成功概率

攻击成功概率 $p(ac)$是指对于特定网络，某种攻击成功入侵的可能性。一次攻击成功与否依赖于其攻击技术与所入侵网络的环境配置与漏洞信息，仅当攻击入侵网络的环境配置与漏洞信息可以被该次攻击所利用时，本次攻击才可能成功入侵。

$$p(ac)=\begin{cases} p(a), & vlus_j \in Vlus \\ 0, & \text{其他} \end{cases} \tag{2}$$

其中，$p(a)$为该攻击发生概率，$vlus_j$ 表示该攻击成功实施所依赖漏洞，$Vlus$ 表示被入侵主机中存在的漏洞库，$vlus_j \in Vlus$ 表示被入侵主机存在该攻击所依赖漏洞。

4.2 攻击阶段实现概率

攻击阶段实现概率 $p(s)$是指攻击者已经成功入侵到某个阶段状态的可能性。攻击阶段的实现依赖于多种单个攻击的成功与否，仅当攻击阶段必需的攻击手段全部入侵成功，该攻击阶段才可能实现。

$$p(s)=\begin{cases} p_i(ac)+p_j(ac)-p_i(ac)p_j(ac), & d=0 \\ p_i(ac)p_j(ac), & d=1 \end{cases} \tag{3}$$

其中，$p_i(ac)$、$p_j(ac)$分别为攻击行为 $Alter_i$、$Alter_j$ 的成功入侵概率；d=0 表示状态节点 s 为或节点；d=1 表示状态节点 s 为与节点。

4.3 网络安全态势值

CVSS 给出了一种基于机密性、完整性、可用性 3 个指标评价的漏洞威胁得分，用于衡量单个漏洞对网络的影响。其威胁得分为

$$Impact(v)=10(1-(1-C)(1-I)(1-A)) \tag{4}$$

其中，C、I、A 分别是机密性、完整性、可用性的威胁影响得分。

每一条攻击场景需要利用多个系统漏洞，将攻击阶段实现概率 $p(s)$、攻击阶段利用的所有单个漏洞威胁得分 $Impact(v)$以及攻击阶段发生的节点权重值 $Weight$ 进行综合量化，可计算得到每一条攻击场景对网络安全态势的影响 $sa(path_i)$，即多漏洞对系统安全态势的影响值

$$sa(path_i)=\sum_{j=1}^{m} p_j(s)Impact(v)Weight \tag{5}$$

其中，m 为攻击场景 $path_i$ 已经实现的攻击阶段。$p_j(s) \leqslant 1$，$Impact(v) \leqslant 10$，$\sum Weight=1$，因此 $sa(path_i) \leqslant 10$。依据 CVSS 中对得分的威胁程度定义，设当 $sa(path_i) \in [0,4.0]$ 时，该攻击者对网络造成的危害为低风险；当 $sa(path_i) \in (4.0,7.0]$ 时，该攻击者对网络造成的危害为中度风险；当 $sa(path_i) \in (7.0,10]$ 时，该攻击者对网络造成的危害为高风险。

最后结合每一条攻击场景对网络安全态势的影响，得到整个网络的安全态势

$$SA=\sum_{i=1}^{n} sa(path_i) \tag{6}$$

其中，n 为检测到的所有攻击场景的总和。

5 实验结果及分析

经典的林肯实验室提供的 DARPA2000[19]是一个用于测试入侵检测场景的专用数据集，2015 年举行的第 23 届世界黑客大赛中所使用的 Defcon23 CTF[20]数据集适用于对网络攻防对方的对抗过程进行解析，符合本文所提方法的应用场景，因此，为验证本文模型及算法的可行性与有效性，选用这 2 个数据集作为实验数据进行分析与对比。首先利用 TCPReplay 工具将数据集重放，并利用入侵检测系统 Snort 对流量进行检测。利用融合结果实施网

络攻击阶段识别，并利用工具 Graphviz 将状态转移图可视化。最后对网络安全态势进行量化评估。

5.1 LLDOS 数据集

DAPRA2000 入侵检测场景专用数据集包含 2 个子集：LLDOS1.0 和 LLDOS2.0.2，这 2 个子集分别包含一个完整的实施分布式拒绝服务攻击过程的场景，前者的攻击者能力较为初级（novice），而后者的攻击者能够采取更加隐蔽、更加高级（more stealthy）的技术。在数据集 LLDOS1.0 中，攻击者先是分别对网段 172.16.115/、172.16.114/、172.16.113/、172.16.112/ 进行 IP 扫描，查询有效主机；对有效主机执行 Sadmind Ping，查询运行 Sadmind 服务的主机；最终查询到有效主机分别为 172.16.115.20、172.16.112.50、172.16.112.20，并利用这 3 台主机中 Solaris 操作系统中的 Sadmind Buffer Overflow 漏洞实施 Daemon Installed 攻击；然后利用 Sadmind Exploit 获得 3 台主机的 root 权限；最后利用该 3 台主机对最终目标实施 DDoS 攻击。攻击者对目标主机实施 DDoS 攻击时采用了随机 IP 地址的方法，因此，在进行分析时选用固定的 mac 地址。利用实时攻击阶段识别算法，得到该攻击的状态转移图，如图 4(a)所示。在数据集 LLDOS 2.0.2 中，攻击过程较 LLDOS1.0 更为隐蔽，其放弃了极易被屏蔽掉的 ICMP PING，攻击者先采用正常的 DNS_Query，查询到 DNS 服务器为 172.16.115.20，并利用该主机中 Sadmind 漏洞获得 root 权限，再以该主机为跳板控制主机 172.16.112.50 的权限，最后对目标主机实施 DDoS 攻击。其攻击状态转移如图 4(b)所示。

通过图中可以看出，利用攻击阶段的识别攻击状态转移图能够清晰地表达出攻击事件之间的关联，较之报警事件之间的关联方法，能够更加明确地给出攻击间的因果关系，且状态转移图也更为简洁。在图 4(b)中，在检测 LLDOS2.0.2 数据集时，由于 DNS_Query 与 FTP upload 均是一个正常的访问方法，检测系统没有给出该事件的报警信息，然而这 2 个操作实际上是攻击实施中的原子操作，即常规检测系统在此处出现了漏报的情况。通过本文所提出的阶段识别改进算法，可以根据后续事件的发生，推测出该阶段攻击状态的发生，有效地解决了漏洞问题，更加真实地还原实际攻击过程中的状态转移过程。

在该数据集中并未给出各主机明确的脆弱性信息，本文通过攻击事件以及运行的服务信息分析得到该网络中主要被攻击的主机的脆弱性信息，如表 1 所示。

表 1　　主要受攻击主机脆弱性信息

脆弱性信息	Mill	Pascal	Locke	www.af.mil
ICMP 非法配置	√	√	√	
SunRPC 非法配置	√	√	√	
Sadmind 缓存区溢出 (CVE-1999-0977)	√	√	√	
RCP 非法配置	√	√	√	
SYN 泛洪(CVE-1999-0116)				√
HINFO Query 非法配置	√			
FTP 非法配置	√	√	√	

分析攻击者采取攻击手段的脆弱性利用关系，并结合 CVSS 得到其威胁值。结果如表 2 所示。

表 2　　脆弱性利用关系

攻击手段	利用漏洞	威胁值
IP 扫描	ICMP 非法配置	1
Sadmind Ping	SunRPC 非法配置	2
Daemon 安装	Sadmind 缓存区溢出 (CVE-1999-0977)	10
Sadmind Exploit	RCP 非法配置	4
DDoS	SYN 泛洪(CVE-1999-0116)	10

该网络分为 DMZ 和 INSIDE 共 2 个区域以及 1 个服务器，其中，DMZ 为 1 个网段，INSIDE 为 6 个网段，因此在分配权重时，分配 DMZ 的总权重为 0.1，各主机的权重平均分配；INSIDE 中 113、116、117、118 网段的权重均为 0.1，各主机的权重平均分配；115 网段的权重为 0.1，Mill 的权重为 0.05，其余主机平均分配；112 网段的权重为 0.2，Pascal 与 Locke 均为 0.05，其余主机平均分配。

本文对 LLDOS1.0 数据集的网络安全态势评估进行了验证。结合入侵检测报警数据和审计日志，依据第 3 节提出的态势评估方法，对网络中各节点依次进行数据融合、攻击阶段识别、网络安全态势评估，从而得到实时的网络安全态势值。以主机 Mill 为例，在 10:08:07 时检测到该主机受到 Sadmind Ping 攻击，经过对 DMZ、INSIDE、BSM 审计日志的数据融合得到攻击发生概率，记作 $p_2(a)=0.918$，该主机中存在其所利用漏洞，因此 $p_2(ac)=0.918$；且该攻击阶段通过 SadmindPing 攻击即可完成，因此，攻击阶段实现概率 $p_2(s)=p_2(ac)=0.918$。结合各攻击威胁与各主机的权重，可以得到当前该攻击路径对网络安全态势的影响 $sa(path)=p_1(s)impact_1(v)weight_1+p_2(s)impact_2(v)weight_2$ =0.917×1×0.5+0.918×2×0.15=0.733 9。此时该网络仅受到一条攻击路线的入侵，因此，该网络的安全态势

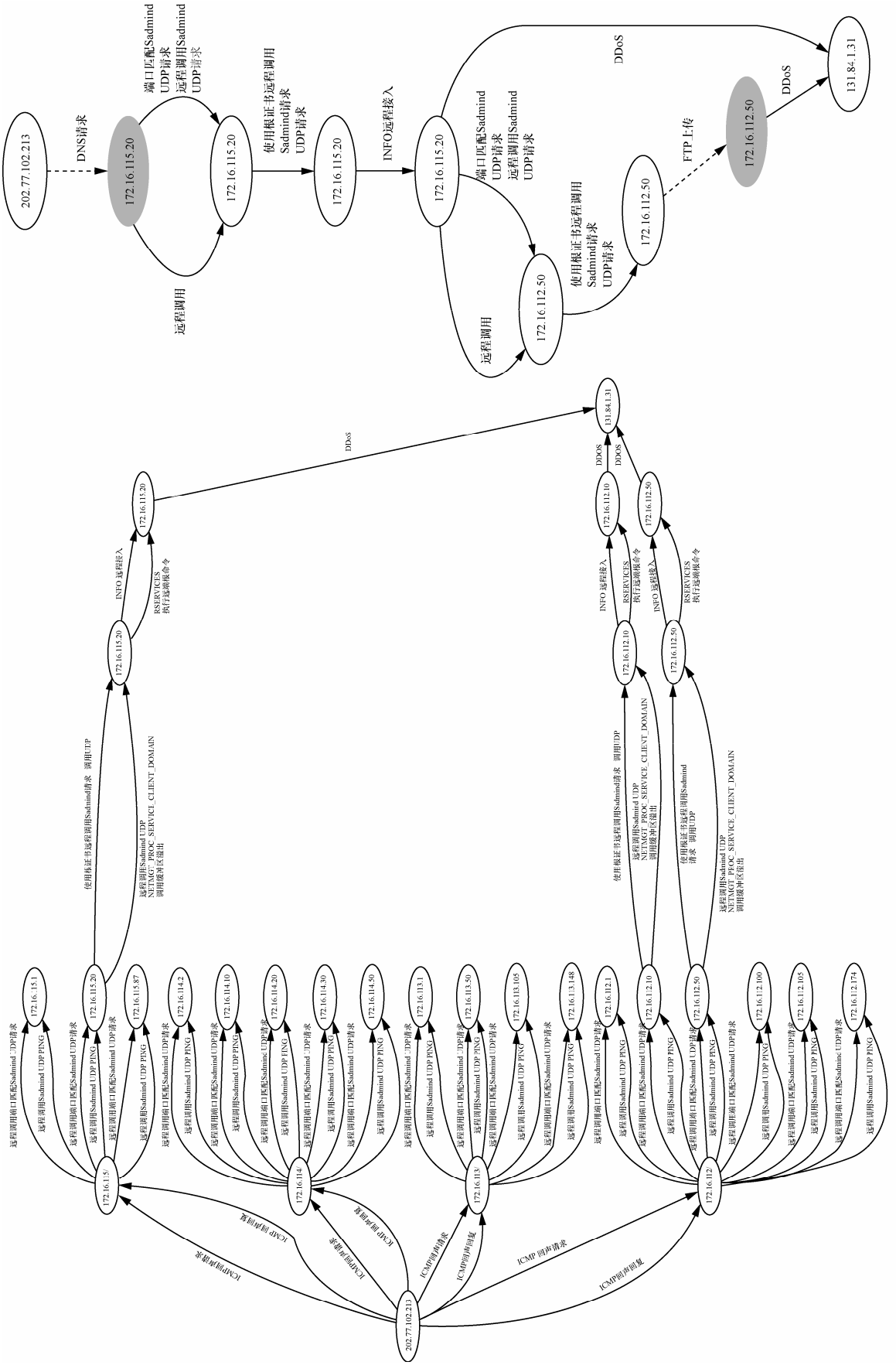


图 4　LLDOS 攻击状态转移

即该条攻击路径的影响 $SA=sa(path)=0.733\ 9$。按照此方法得到网络中受攻击主要主机的攻击阶段完成情况及网络安全态势值如表 3 所示。

表 3　　网络安全态势值计算结果

时间	攻击状态	Mill	Locke	Pascal	www.af.mil	*SA*
09:51~09:52	阶段 1	0.917	0.917	0.917	0	0.458 5
10:07~10:17	阶段 2	0.918	0.918	0.918	0	0.733 9
10:33~10:35	阶段 3	0.95	0.95	0.96	0	2.163 9
10:50	阶段 4	0.94	0.94	0.94	0	2.727 9
11:27	阶段 5	0	0	0	0.96	3.687 9

根据上述计算结果，将网络安全态势值绘制成图，以更加直观的形式展现，如图 5 所示。态势值越大表示此刻网络中受到的攻击危害程度越大。

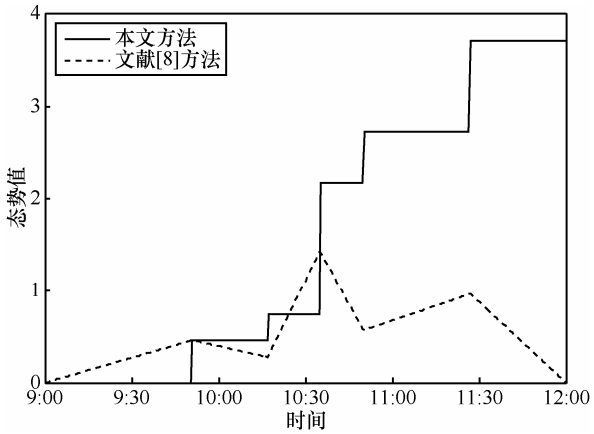


图 5　网络安全态势评估

通过图 5 可以看出，随着攻击意图的不断实现，本文方法所绘制的态势图也随之增大，能够反映出多步攻击的攻击阶段；且整个网络的安全态势值一直都处于轻度危险状态，因为对于整个防护网络而言，其中受到攻击的主机仍属于小部分，其态势值能够符合实际情况。在现有的研究中，网络安全态势值的评估方法忽略了攻击间的因果关联，如文献[8]所提模型仅考虑单个攻击的影响，无法准确反映多步攻击的阶段实现情况，因此，本文所提出的评估方法面对多步复杂攻击时更具参考价值。

5.2　CTF 数据集

Defcon 是世界上最大的黑客组织，并每年举行一次夺旗（CTF, capture the flag）比赛。有 20 个左右经过挑选的优秀网络攻防小组进行比赛，每个小组均以攻破对方网络为目标，并对本方网络进行防护。参赛队员均被公认为世界上最优秀的攻防对抗小组，因此，其攻击过程具有较大的分析意义。组织方利用抓包工具将所有小组的网络数据全部记录下来，数据集为几百吉比特大小。本实验选用 Defcon CTF23 数据集，虽然其网络结构不是特别复杂，但该数据集中的攻击者均能熟练掌握入侵工具，且有多个攻击者参与，较为适合该方法的功能测试。该实验选用来自中国的蓝莲花小组作为防护对象，并对该小组的数据分组进行分析，该数据集分为 3 天采集，本文对第 1 天采集的数据集的安全态势进行评估。

通过网络攻击阶段识别算法，可以得出该网络的攻击阶段状态转移图，如图 6 所示。通过图 6 可以看出共识别出了对该网络发起攻击的 19 个攻击者。由于该网络可以利用的脆弱性是相同的，因此，所有攻击者发起的攻击也较为相似。

在态势评估时，由于该网络没有给出脆弱性信息，因此，假设该网络中存在所有攻击的脆弱性信息。经过计算得到该网络的态势评估图，如图 7 所示。为更加直观地观察网络态势，图 7 仅对态势发生变化的时刻进行标记。

攻击者在 2~4 时段处于扫描阶段；在 4~6 时段进行了 SHELLCODE X86 NOOP 攻击，从图 7 中也可以看出该时段态势值随着攻击的深入逐渐增大；在 6~7 时段，攻击者未对网络发起攻击，由于转移等待窗口的限制，将长时间未发起攻击的场景删除，态势值随之不断减小。在 8~11 时段有零星攻击者对网络发起攻击，且网络不断受到 DDoS 攻击，该时段状态值的变化无明显规律。在 11~12 时段，所有攻击者均对网络实施了 SHELLCODE X86 NOOP 或 SHELLCODE X86 in ebx NOOP 攻击，该时段状态值急剧增大。因此，本文所绘制的网络安全态势图能够较好地反映攻击实情。

5.3　算法效能分析

本文所提算法旨在通过对攻击阶段与攻击者进行识别给出一个更加合理有效的安全态势量化评估值。由于现有安全防护设备的报警信息存在重复、乱序、误报、漏报等问题，是严重干扰攻击识别结果的主要原因，因此，接下来将主要围绕这 4 点问题对所提算法的效能进行分析。

1) 能够有效处理重复报警。本文对重复报警采用 2 个方法。首先对报警信息利用报警的各属性对其进行聚类，使大量重复的报警聚合成一条信息，其次利用状态转移，对重复发生的状态进行丢弃，从而减少重复报警对态势评估的影响。

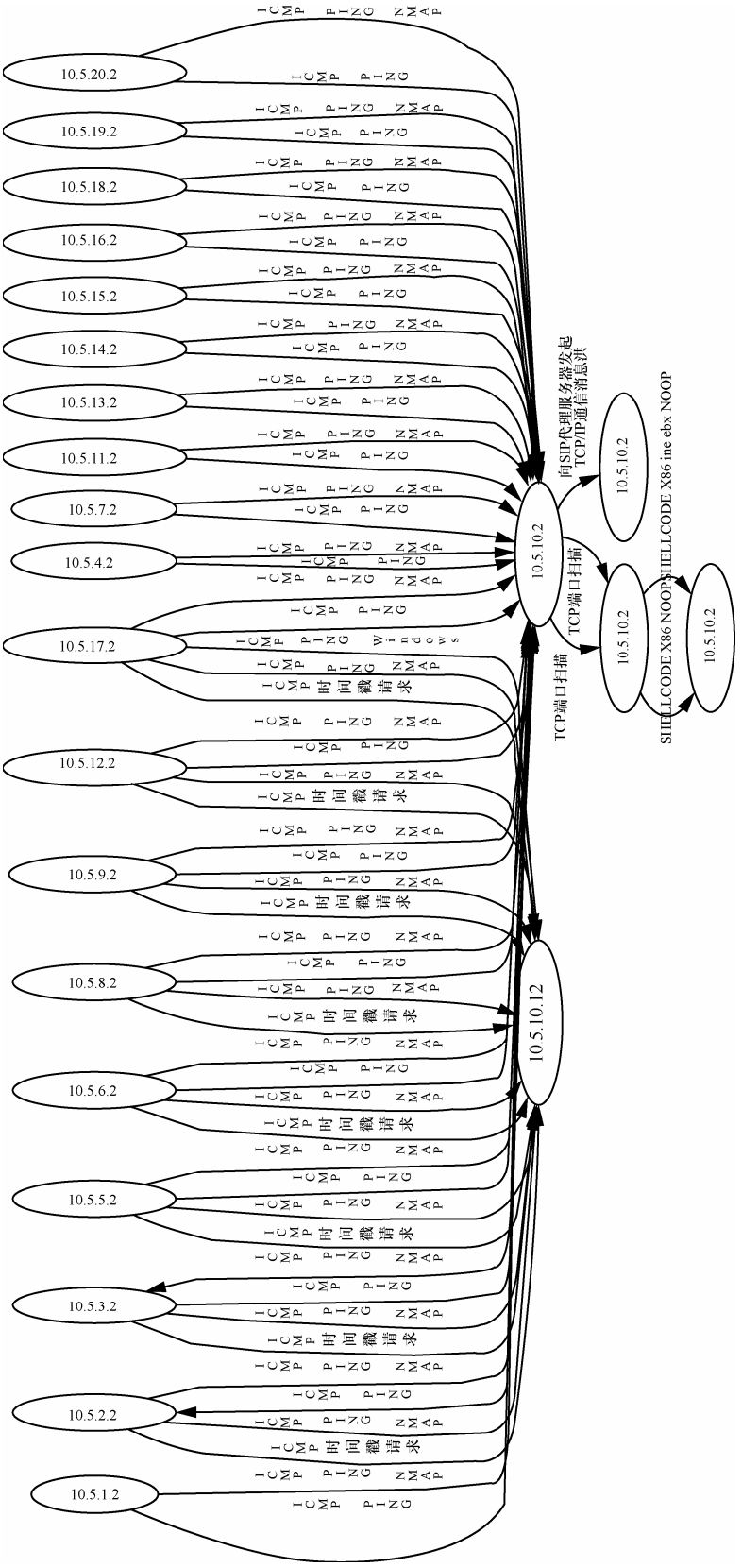

10.5.20.2
10.5.19.2
10.5.18.2
10.5.16.2
10.5.15.2
10.5.14.2
10.5.13.2
10.5.11.2
10.5.7.2
10.5.4.2
10.5.17.2
10.5.12.2
10.5.9.2
10.5.8.2
10.5.6.2
10.5.5.2
10.5.3.2
10.5.2.2
10.5.1.2
10.5.10.12
10.5.10.2
向SIP代理服务器发起TCP/IP通信消息洪
TCP端口扫描
TCP端口扫描
10.5.10.2
SHELLCODE X86 NOOP
SHELLCODE X86 ine ebx NOOP
10.5.10.2

图 6　攻击状态转移

2) 能够有效处理报警乱序问题。本文通过设置中间状态，保存先到的后续状态，当其前提状态到达，便可将后续状态更新为发生状态。通过该方法，报警乱序问题较大程度得到解决，能够更加准确识别攻击阶段。

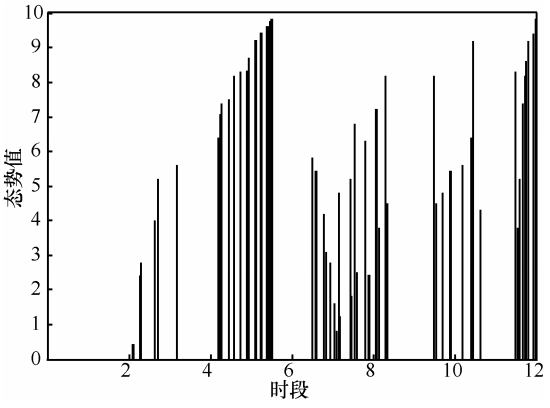


图 7　网络安全态势评估

3) 能够有效处理误报信息。由于报警系统自身设计缺陷，其会产生大量的误报信息。本文对误报的处理采用 2 个方法。首先对多源报警信息进行融合，由于不同的系统产生相同误报信息的可能性较小，因此，该方法能降低误报信息的概率，减小其对态势评估的影响。其次对报警信息进行场景重构，而误报多为一个独立的攻击，因此对真实攻击阶段识别的影响较小。

4) 能够在一定程度上处理漏报问题。通过设置中间状态，类似于报警乱序问题，能够解决一部分漏报问题。但漏报的信息不会随后出现，其状态的更新依赖于其后续状态的发生，若其后续状态无法成功，本文方法则对该漏报信息无法识别。

5) 对零日漏洞与零日攻击有初步的探讨。通过设置中间状态，对尚未成功的攻击进行保存，若其后续攻击的前提状态为该中间状态，则能够推测出零日漏洞。但该方法过分依赖于后续状态的发生，且该方法难以通过实验证明，仅从理论上进行了探讨。

6）容易看出，本文所提算法主要执行判断语句而不是循环语句，根据判断条件对攻击场景进行修改与识别，主要时间耗费在对执行检验条件上，算法的时间复杂度为 $O(1)$。

6　结束语

为了分析多步攻击对网络系统的影响，准确、全面地反映系统的安全态势，从攻击者角度出发，本文提出一种面向多步攻击的网络安全态势评估方法。首先对网络中的安全事件进行场景聚类以识别攻击者；对每个攻击场景进行关联分析，识别出相应的攻击轨迹与攻击阶段；最后基于 CVSS 提出了一种态势量化指标。实验表明，该评估方法可行有效，且能较好处理重复报警、报警乱序、误报等问题。网络安全态势感知包括当前状态的评估以及未来状态的预测，本文主要针对当前状态的评估进行了研究，因此接下来将加强对状态预测的研究，进一步提高感知的全面性。

参考文献：

[1] 吴迪, 连一峰, 陈恺, 等.一种基于攻击图的安全威胁识别和分析方法[J]. 计算机学报, 2012,35(6):1938-1950.
WU D, LIAN Y F, CHEN K, et al. A security threats identification and analysis method based on attack graph[J]. Chinese Journal of Computers, 2012, 35(6): 1938-1950.

[2] 田志宏, 余翔湛, 张宏莉, 等.基于证据推理网络的实时网络入侵取证方法[J]. 计算机学报, 2014,37(5):1184-1194.
TIAN Z H, YU X Z, ZHANG H L, et al. A real-time intrusion forensics method based on evidence reasoning network [J]. Chinese Journal of Computer, 2014, 37(5): 1184-1194.

[3] ALHAZMI O H, MALAIYA Y K, RAY I. Measuring, analyzing and predicting security vulnerabilities in software systems [J]. Computers & Security, 2007, 26(3): 219-228.

[4] HANNES H, MATHIAS E, DENNIS A. Empirical analysis of system-level vulnerability metrics through actual attacks[J]. IEEE Transactions on Dependable and Secure Computing, 2012, 9(6): 825-837.

[5] ENDSLEY M R. Design and evaluation for situation awareness enhancement[C]//The Human Factors Society 32nd Annual Meeting. 1988: 97-101.

[6] BASS T. Intrusion detection systems & multisensory data fusion: creating cyberspace situational awareness[J]. Communications of the ACM, 2000, 43(4): 99-105.

[7] 陈秀真, 郑庆华, 管晓宏, 等. 层次化网络安全威胁态势量化评估方法[J].软件学报, 2006,17(4): 885-997.
CHEN X Z, ZHENG Q H, GUAN X H, et al. Quantitative hierarchical threat evaluation model for network security[J]. Journal of Software, 2006, 17(4): 885-997.

[8] 韦勇,连一峰,冯登国.基于信息融合的网络安全态势评估模型[J].计算机研究与发展,2009,46(3):353-362.

[9] MIRMOEINI F, KRISHNAMURTHY V. Reconfigurable Bayesian networks for hierarchical multi-stage situation assessment in battlespace[C]//The 39th Asilomar Conference on Signals, Systems and Computers. 2005, 104-108.

[10] 徐晓辉, 刘作良. 基于 D-S 证据理论的态势评估方法[J].电光与控制,2005, 12(5): 36-37.
XU X H, LIU Z L. A method for situation assessment based on D-S

evidence theory [J]. Electronics Optics & Control, 2005, 12(5): 36-37.

[11] ZHUO Y, ZHANG Q, GONG Z H. Network situation assessment based on RST [C]//Pacific-Asia Workshop on Computational Indulgence and Industrial Application. 2008: 502-506.

[12] ZHOU Y, ZHANG Q, GONG Z H. Research and implementation of network transmission situation awareness[C]//WRI World Congress on Computer Science and Information Engineering. 2009: 210-214.

[13] 张勇, 谭笑彬, 崔孝林, 等.基于 Markov 博弈模型的网络安全态势感知方法[J].软件学报. 2011, 22(3): 495-508.
ZHANG Y, TAN X B, CUI X L, et al. Network security situation awareness approach based on markov game model[J]. Journal of Software, 2011, 22(3): 495-508.

[14] YEE W, TANSU A, MARIMUTHU P. Security games for risk minization in automatic generation control[J]. IEEE Transactions on Power Systems, 2015, 30(1): 223-232.

[15] 吕慧颖, 彭武, 王瑞梅, 等. 基于时空关联分析的网络安全实时威胁识别与评估[J].计算机研究与发展,2014,51(5):1039-1049.
LYU H Y, PENG W, WANG R M, et al. A real-time network threat recognition and assessment method based on association analysis of time and space[J]. Journal of Computer Research and Development, 2014, 51(5): 1039-1049.

[16] CYRIL O, THOMAS O. Situational awareness in computer network defense principles, methods and applications[M]. Hershey: IGI Global Snippet, 2012:125-137.

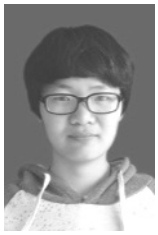
[17] SCHIFFMAN M. Common vulnerability scoring system version 2.0[EB/OL]. http://www.first.org/cvss/cvss-guide. html.

[18] FATEMEH K, BEHZAD A. Automatic learning of attack behavior patterns using Bayesian networks[C]//6th International Symposium on Telecommunications (IST'2012). 2012:999-1004

[19] MIT LINCOLN LABORATORY. 2000 DARPA intrusion detection scenario specific data sets[EB/OL]. http://ll.mit.edu/IST/ideval/data/2000/2000_data_index.html.

[20] DEFCON Capture the flag traffic dump[EB/OL]. http://www.defcon.org/html/links/dc-cft.html.

作者简介：



杨豪璞（1993-），女，福建厦门人，信息工程大学硕士生，主要研究方向为 APT 攻击防御、博奕理论。

邱辉（1991-），男，河南周口人，信息工程大学硕士生，主要研究方向为网络安全态势感知、数据挖掘。

王坤（1975-），男，河南周口人，信息工程大学副教授，主要研究方向为信息安全、数据分析。