内网环境下基于时空事件关联的攻击检测方法

孙伟[1,2]，张鹏[2]，何永全[2,3]，邢丽超[2,3]

（1. 北京交通大学计算机与信息技术学院，北京 100044；
2. 中国科学院信息工程研究所，北京 100093；
3. 中国科学院大学网络空间安全学院，北京 100049）

摘 要：针对入侵检测系统使用单个事件作为攻击检测的特征会导致较高误报率的问题，提出了利用贝叶斯网络模型进行跨空间的事件关联和利用卡尔曼滤波器线性模型进行跨时间的事件关联的内网攻击检测方法。基于该方法实现了一个进程查询系统，该系统可以根据用户的高层过程描述来扫描和关联分布的网络事件。实验分析表明，该方法在不增加明显计算开销的情况下能够显著减少内网攻击检测的误报率。

关键词：时空事件；内网攻击检测；进程查询；入侵检测系统

中图分类号：TP393.1

文献标识码：A

doi: 10.11959/j.issn.1000−436x.2020001

Attack detection method based on spatiotemporal event correlation in intranet environment

SUN Wei[1,2], ZHANG Peng[2], HE Yongquan[2,3], XING Lichao[2,3]

1. School of Computer and Information Technology, Beijing Jiaotong University, Beijing 100044, China
2. Institute of Information Engineering, Chinese Academy of Sciences, Beijing 100093, China
3. School of Cyber Security, University of Chinese Academy of Sciences, Beijing 100049, China

Abstract: In view of the fact that a single event as an attack detection feature leads to a higher false positive rate, an intranet attack detection method using Bayesian network model for cross-space event correlation and Kalman filter linear model for cross-temporal event correlation was proposed. Based on the method, a process query system was implemented, which can scan and correlate distributed network events according to the user's high-level process description. Experimental analysis show that the proposed method can significantly reduce the false positive rate of intranet attack detection without increasing the computational overhead.

Key words: spatiotemporal event, intranet attack detection, process query, intrusion detection system

1 引言

网络空间测绘技术应用于网络空间的多个领域，对网络空间的高效管理、资源的合理分配及有效的安全监测和防护都有着十分重要的意义。美国国家安全局（NSA, National Security Agency）和英国政府通讯总部（GCHQ, the Government Communications Headquarters）联合开展的“藏宝图”计划，聚焦于逻辑层捕获路由及自治系统的数据，试图绘制出一张“近乎实时的、交互式的全球互联网地图”。Spring 等[1]的 Rocket Fuel 首次全面探测了大型互联网服务提供商（ISP, Internet service provider）

收稿日期：2019–05–20；**修回日期**：2019–07–08

基金项目：国家重点研究发展计划基金资助项目（No.2016YFB0801300）；国家自然科学基金资助项目（No. 61602474, No.61602467, No.61702552）

Foundation Items: The National Key Research and Development Program of China (No.2016YFB0801300), The National Natural Science Foundation of China (No.61602474, No.61602467, No.61702552)

拓扑，在扫描工具方面有业内常用的功能强大的 Nmap[2]。由著名安全专家 John Matherly 于 2009 年创建的 Shodan，是全球第一个网络空间安全搜索引擎。Shodan 每月在大约 5 亿个服务器上不停地搜集信息，主要针对服务器、网络摄像头、交换机、路由器等网络基础设备进行扫描。

网络空间测绘不仅在传统互联网上应用前景广阔，在内部网络中应用价值也很高。本文的内部网络主要是指某行业依托政务外网建设的系统网和与互联网逻辑隔离的大型企业网。内部网络的组件数量可以达到 40 万台，会产生大量基于网络事件的数据，包括来自防火墙和入侵检测系统的警报、各种软件系统的日志文件、来自互联网的路由信息等，这些数据可被收集用于网络安全分析。由于网络具有分布式的特性，攻击网络及其资源的证据常常嵌入在分散的事件之中。此外，对网络的攻击可能涉及多个步骤。因此，攻击的证据也通常随时间推移而分散。随着大量分布式的事件不断产生，一个关键的挑战是如何跨空间和时间去关联这些分布的事件以检测和跟踪各种攻击场景。

传统的入侵检测系统仅使用单个事件作为检测攻击的特征，产生了较高的误报率。从大量的网络事件中挖掘更多的证据以获得更高的检测精度是十分必要的。为此，本文提出利用贝叶斯网络模型进行跨空间的事件关联和利用卡尔曼滤波器线性模型进行跨时间的事件关联的方法。基于该方法实现了一个进程查询系统，该系统可以根据用户的高层过程描述来扫描和关联分布的网络事件。本文的主要贡献如下。

1) 使用结合了空间和时间的联合事件场景特征来描述和区分各种攻击，而不是使用单个事件作为特征。随着更多的证据从分布式事件中挖掘出来，该方法能够提高检测准确度，尤其是在一个有噪声的网络环境中。

2) 提出了为各种攻击场景构建场景特征的方法。该方法采用 2 种技术：一种是基于因果关系分析，使用专家知识来构建场景特征的技术；另一种是使用数据挖掘从大量训练数据中提取特征的技术。

3) 基于上述方法实现了一个进程查询系统。实验分析表明，该系统在不明显增加计算开销的情况下，能够显著减少内网攻击检测的误报率。

2 研究动机

计算机网络由许多组件组成，如路由器、交换机、Web 服务器、邮件服务器、数据库服务器、DNS（domain name server）服务器、IDS（intrusion detection system）和防火墙等。此外，计算机网络是动态的系统，每隔一段时间，这些组件就会产生大量基于事件的数据。一般而言，所有这些事件都可以由网络数据分析中心收集，而攻击的痕迹通常分散在这些事件中。如果没有高效的关联算法，在这个庞大且充满噪声的事件空间中识别攻击的踪迹基本上是很难的。与其他模式识别问题一样，需要一个攻击场景特征（或模式）来将这个攻击与其他攻击和正常网络活动区分开。检测的准确性取决于提取场景特征的准确性以及收集到事件的准确性。因此，一个关键的挑战是如何描述各种各样的攻击场景。

图 1 说明了一个攻击的证据是如何在空间和时间上分布的。基于因果关系，一个攻击可以同时影响多个观测空间的事件。例如 CodeRed 和 Nimda 等蠕虫会生成和扫描随机 IP 地址，以搜索 IP 空间中易受攻击的目标。由于许多 IP 地址未分配给网络或未被网络使用，这种主动探测过程可能在网络路由中生成大量 ICMP（Internet control message protocol）不可达的分组[3]，密集的蠕虫传播过程也会影响互联网的延迟情况。此外，蠕虫攻击可能导致不稳定的 Internet 边界网关协议路由（BGP, border gateway protocol）[4]。基于这些因果关系，至少有 3 个独立的观测空间来感知蠕虫攻击：ICMP 不可达分组的数量、网络延迟、BGP 路由稳定性。因此，本文使用这 3 个指标作为组合特征，而不是使用单个事件进行蠕虫攻击检测，从而在空间上关联事件并检测蠕虫攻击。

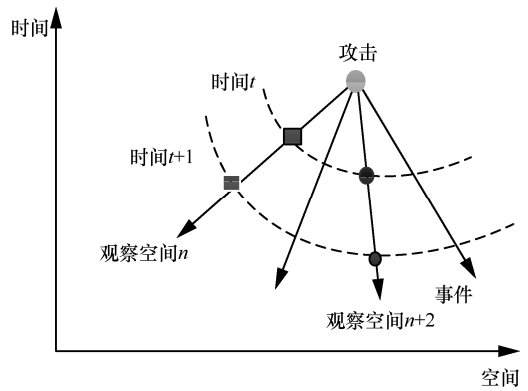


图 1 时间和空间上攻击的证据分布

此外，攻击也会影响跨时间的事件。例如，蠕虫按传染病模型在互联网上传播时，其生命周期中经历多个阶段：突破、传播、消除。在此动态过程中，ICMP 不可达分组的数量遵循时间模式，而该时间模式可用作蠕虫检测的时间特征。实际上，上面讨论的每个观测空间都可以通过其特定的时间模式独立地感知蠕虫攻击。因此，本文将使用一个过程模型来描述时间特征[5]，通过将每个观测空间中的时间事件与过程模型相关联来检测攻击。下面，分别介绍基于空间序列和基于时间序列的分布网络事件关联方法。

3　基于空间序列的事件关联

关联的速度和准确度是事件关联系统中 2 个重要的性能。事件关联的一种经典方法是基于规则的分析，即关联系统不断使用一组预定义规则集来评估到来的观察结果，直到得出结论。因此，关联能力仅取决于规则集的深度和规模。设计正确的规则集需要大量的专业知识。遵循规则集的严格路径，观察结果需要对照众多条件逻辑检查，导致基于规则的系统通常不能很好地扩展。同时，基于规则的系统本质上是无状态的，并且不能很好地处理动态数据的关联。下面将讨论如何提高事件关联的速度和准确度。

基于空间序列的关联将同时来自多个观测空间或传感器的事件关联起来，以检测攻击场景。将一个攻击场景表示为 s，假设有一组包含 m 个攻击场景的集合 $S=\{s_1,s_2,\cdots,s_m\}$和一组包含 n 个观测空间的集合 $O=\{o_1,o_2,\cdots,o_n\}$。每个观测空间都可以作为攻击场景中一个独立的指标。基于空间序列的事件关联是关于如何关联 n 个指标来检测和区分 m 个攻击场景。

3.1　确定性方法

图 2 表示具有 3 个攻击场景和 4 个观测空间的因果关系图，其中有向边表示因果关系。例如，如果发生攻击 s_1，则会导致 o_1 和 o_3 中的异常观察结果，然而这种攻击不会影响 o_2 和 o_4 中的观测结果。基于这些因果关系，构建如表 1 所示的 codebook 相关矩阵，其中 1 和 0 分别表示用特定阈值分类的“异常”和“正常”观察。因此，本文可以将来自多个观测空间的事件与相关矩阵进行比较，以检测和区分这些攻击。每个攻击场景必须在此相关矩阵中具有可区分的场景特征，需要专家知识来构建场景特征和相关矩阵。相关矩阵的大小可以减小，但是为了达到区分的目的，场景特征相距必须是最小汉明距离[6]。

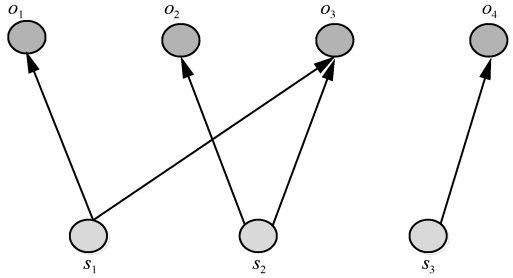


图 2　事件的因果关系

表 1　　事件的相关矩阵

观测空间	攻击场景		
	s_1	s_2	s_3
o_1	1	0	0
o_2	0	1	0
o_3	1	1	0
o_4	0	0	1

将相关矩阵定义为 **OS**={os$_{ij}$}是该矩阵的元素，其中 $1\leqslant i\leqslant n$ 且 $1\leqslant j\leqslant m$。在 codebook 方法中，os$_{ij}$ = 1 或 0，即观察结果分为“异常”或“正常”。这种二进制表示没有提供太多关于“异常”观察结果强度的信息。相反，os$_{ij}$ 可能是一个实值，例如 ICMP 不可达分组的数量或特定系统调用 os$_{ij}$ 的数量。在这种情况下，认为关联问题可以表示为如式(1)所示的整数规划问题。

$$\begin{aligned}&\min_{\boldsymbol{H}} \|\boldsymbol{V}_O-\mathbf{OS}\boldsymbol{H}\|\\&\text{s.t. }\forall 1\leqslant i\leqslant n,\sum_{j=1}^{m}\text{os}_{ij}h_j\leqslant v_i;\\&h_j\geqslant 0\end{aligned}\tag{1}$$

其中，$\boldsymbol{H}=(h_1,h_2,\cdots,h_m)^{\mathrm{T}}$ 是假设向量，$\boldsymbol{V}_O=(v_1,v_2,\cdots,v_m)^{\mathrm{T}}$ 是来自 m 个观察空间的观测向量。上面的整数规划问题是关于如何组合攻击场景，以便解释真实的观察。例如，$\boldsymbol{H}=(0,1,1,0,\cdots,0)^{\mathrm{T}}$ 表示攻击场景 s_2 和 s_3 同时发生，码本方法无法检测这种攻击场景的组合问题。如果本文将一个观测空间视为一条信号通道，则观察到的事件通常包括来自攻击的信号和来自网络环境的噪声信号。本文的整数规划方法可以同时检测多个攻击场景实例，

并且可以在观察的信噪比强的环境中工作。尽管如此，仍需要专家知识来获取 os_{ij} 值，并且这些值必须在各种攻击场景进行标准化。

3.2 概率方法

如上所述，确定性方法在有噪声的环境中不能很好地工作，网络噪声源自正常的网络活动。例如，主要路由器故障可能会生成许多 ICMP 不可达的分组，忘记密码可能导致多次登录失败的警报。问题是如何根据有所偏倚的观察结果来检测攻击情景。将在观测空间 o_p 中的观测值表示为 $v_i(1\leqslant i\leqslant n)$，$o_p$ 属于观测空间集合 O。根据专家知识和统计数据，假设先验概率已知，如式(2)所示。

$$p\left(o_i \mid s_j\right)=\Pr\left(o_p=v_i \mid s=s_j\right) \tag{2}$$

其中，$1\leqslant j\leqslant m$，$1\leqslant i\leqslant n$。也就是说，由攻击引起的观察值的分布是已知的。根据贝叶斯定理，可以计算出后验概率，如式(3)所示。

$$p\left(s_j \mid v_i\right)=\frac{p\left(v_i \mid s_j\right)p\left(s_j\right)}{p\left(v_i\right)} \tag{3}$$

现在的问题是如何关联来自多个观测空间的观测结果。本文使用 o_p 和 o_q 分别表示 2 个观测空间，它们均属于观测空间集合 O，并且有来自 o_p 和 o_q 的观察，可以得到联合后验概率，如式(4)所示。

$$p\left(s_j \mid v_{pi},v_{qk}\right)=\frac{p\left(v_{pi},v_{qk} \mid s_j\right)p\left(s_j\right)}{p\left(v_{pi},v_{qk}\right)} \tag{4}$$

如果观测空间 o_p 和 o_q 是相互独立的，也就是说，一个观测空间中的事件不会导致另一个观测空间中的事件，反之亦然，则式(4)可以写成式(5)所示形式。

$$p\left(s_j \mid v_{pi},v_{qk}\right)=\frac{p\left(s_j \mid v_{pi}\right)p\left(s_j \mid v_{qk}\right)}{p\left(s_j\right)} \tag{5}$$

实际上，如果只想识别导致当前观测的最可能的攻击，可以使用式(6)来比较不同攻击场景的可能性。

$$\frac{p\left(s_j \mid v_{pi},v_{qk}\right)}{p\left(s_l \mid v_{pi},v_{qk}\right)}=\frac{p\left(v_{pi},v_{qk} \mid s_j\right)p\left(s_j\right)}{p\left(v_{pi},v_{qk} \mid s_l\right)p\left(s_l\right)} \tag{6}$$

然而，在大多数情况下，$p\left(s_j\right)$ 和 $p\left(s_l\right)$ 这样的概率是未知的，并且必须假设它们具有相同的分布。在该假设下，将选定的阈值与式(6)中的先验概率的比值相评估，以确定攻击场景。基于此阈值，Neyman-Pearson 检测理论[5]可用于得出相关的误报率和误检率。如果增加观察空间的数量，本文方法可以使攻击场景更加容易区别。

事件的多层因果关系可以用贝叶斯网络表示[7]。贝叶斯网络是有向无环图，其中，节点是随机变量，边表示源点对终点施加的直接因果影响。在贝叶斯网络中，联合概率作为因素计入一组条件概率集中，该条件概率可以沿网络中的因果关系路径顺序地计算。例如 Benferhat 等[8]已经使用贝叶斯网络模型来检测分布式拒绝服务（DDoS, distributed denial of service）攻击。空间关联的另一种方法是使用 Dempster-Shafer 理论，该理论可以结合来自多个观测空间的可信度。

基于概率的关联方法可以在有噪声的环境中很好地工作。然而，获得先验概率和条件概率存在困难，这使该方法实际上不如确定性关联方法可行。

4 基于时间序列的事件关联

在时间上，许多攻击涉及多个步骤，攻击的证据往往分散在事件中。计算机网络本身是动态系统，网络事件是其动态活动的可观测量。攻击或正常网络行为的时间特征可以被描述为确定的或随机的动态过程。过程模型描述了对象的状态转换，该状态转换根据特定的已知定律随时间演变。例如，可以用状态转移方程、马尔可夫模型、有限状态机等来描述过程模型。“状态”是基于时间关联中一个重要概念。

基于时间序列的关联致力于及时关联观察到的事件以检测攻击，并且可以将其形式化为一个目标跟踪问题。来自雷达和声纳信号处理的目标跟踪算法可以应用于基于时间序列的事件关联。如果已知攻击的动态过程，则基于时间序列的关联可以通过跟踪事件是否遵循该攻击过程来检测此攻击。如果已知正常网络行为的过程，则基于时间序列的关联可以通过跟踪事件是否遵循正常网络行为的过程来检测未知攻击，也被称为“攻击检测”。

4.1 确定性方法

以前的许多工作都使用有限状态机来描述攻

击或软件行为的确定性过程，根据状态转换序列评估事件以检测攻击。文献[9]使用状态转换图来精确识别侵入阶段，并仅显示成功侵入所必须发生的关键事件。文献[10]使用有色 Petri-Nets 来描述攻击的时间特征。这些方法都模拟了攻击的时间特征或侵入过程。

目前，许多异常检测工作已经模拟了正常软件行为或网络行为的过程以检测未知攻击。文献[11]使用运行程序执行的短序列系统调用作为时间特征来检测异常软件行为。文献[12]使用审计日志来捕获程序的行为，并将该规范用作检查行为的 oracle。众所周知，程序执行的 80%通常只发生在其代码的 20%中，程序中的热路径通常代表该程序的主要行为。

已知攻击的动态过程的检测方法需要攻击的专家知识来构建时间特征。已知正常网络行为的过程的检测方法可以基于训练过程自动构建软件行为的时间特征，但是这种方法无法检测攻击的具体类型。

4.2　概率方法

在确定性相关中，动态过程的状态是在没有噪声的环境下被观察和跟踪的。在有噪声的环境中，观测结果通常会被网络噪声污染。将动态过程的状态表示为 x，观察结果表示为 r。将从开始到时间 t 的状态 x 表示为 $x_{1:t}=x_1,x_2,\cdots,x_t$，相关的观察结果 r 表示为 $r_{1:t}=r_1,r_2,\cdots,r_t$。由于动态过程中的几个状态可能导致相同的观察结果并且存在噪声，状态本身也是不可观察的，因此只能基于观察结果来估计状态。在时间 t，基于时间序列的关联的一个任务是将直到 t 的观察结果相关联来估计当前状态 x_t，即 $p(x_t\mid r_{1:t})$。用贝叶斯滤波器[13]递归地计算这个后验概率，有

$$p(x_t\mid r_{1:t-1})=\int_0^{+\infty}p(x_t\mid x_{t-1})p(x_{t-1}\mid r_{1:t-1})\mathrm{d}x_{t-1}$$

$$p(r_t\mid r_{1:t-1})=\int_0^{+\infty}p(r_t\mid x_t)p(x_t\mid r_{1:t-1})\mathrm{d}x_t$$

$$p(x_t\mid r_{1:t})=\frac{p(r_t\mid x_t)p(x_t\mid r_{1:t-1})}{p(r_t\mid r_{1:t-1})} \tag{7}$$

如果关于该过程的以下假设成立：1)过程模型的状态转换具有马尔可夫性质，即当前状态 x_t 仅依赖于先前状态 x_{t-1} 而不依赖于任何更早的状态；2)观察结果 r_t 仅取决于当前状态 x_t，但不依赖于任何早期状态和观察。

线性卡尔曼滤波器[14]模型和隐马尔可夫模型（HMM, hidden Markov model）[15]是满足这 2 个假设的强大模型。对于这些特定模型，诸如卡尔曼滤波器和维特比算法等有效的关联算法可以从式(7)导出。卡尔曼滤波器中使用的线性模型的描述如式(8)和式(9)所示。

$$x_{t+1}=\boldsymbol{D}x_t+w \tag{8}$$

$$r_t=\boldsymbol{L}x_t+z \tag{9}$$

其中，w 和 z 是高斯噪声，$\boldsymbol{D}$ 和 $\boldsymbol{L}$ 是常数矩阵。卡尔曼滤波器使用观察到的 $r_{1:t}$ 来估计潜在的未知 x_t。在离散情况下，隐马尔可夫模型使用一个状态转移矩阵和一个发射矩阵分别代替式(8)和式(9)。

将攻击场景表示为 s，并假设有一组 m 个攻击过程模型的集合 $S=\{s_1,s_2,\cdots,s_m\}$。这里的检测问题是确定哪个攻击正在产生这些观察结果 $r_{1:t}=r_1,r_2,\cdots,r_t$。基于式(6)的分析，比较各种攻击情形的可能性 $p(r_{1:t}\mid s_j)$，并用式(10)所示的不等式识别攻击。

$$B<r^t=\frac{p(r_{1:t}\mid s_j)}{p(r_{1:t}\mid s_k)}<A \tag{10}$$

其中，A 和 B 是 2 个阈值。如果 r^t 大于 A，得出攻击是 s_j 的结论；如果 r^t 小于 B，得出攻击是 s_k 的结论；如果 r^t 小于 A 但大于 B，将继续接收新的观察结果，直到 r^t 超过阈值 A 或 B。虽然概率 $p(r_{1:t}\mid s_j)$ 可以递归计算并从式(7)导出，但在大多数情况下并不知道概率分布的分布公式（例如，如何计算 $p(r_{1:t}\mid s_j)$，在文献[12]中被称为 HMM 的“问题 1”）。因此，不能使用 Neyman-Pearson 检测理论来得出相关的误报率和误检率。将误报率表示为 $\alpha=p_{s=s_k}(r^t>A)$，即攻击是 s_k 但是 $r^t>A$。类似地，将误报率表示为 $\beta=p_{s=s_j}(r^t<B)$。根据顺序分析的结果[16]，得到不等式 $1-\beta\geqslant A\alpha$ 和 $\beta\leqslant(1-\alpha)B$。

卡尔曼滤波器线性模型和 HMM 都已应用于模拟动态攻击过程或正常软件行为。基于传染病模型和对快速传播的蠕虫观察数据，文献[17]使用线性模型来描述蠕虫传播的动态过程，并使用卡尔曼滤波器实时预测蠕虫传播。文献[18]使用训练数据学习 HMM，表示正常的软件行为。然而，在通常情况下，获得这些模型的准确参数是

比较困难的，因此，本文的方法是基于无参数的时间序列模型。

5 基于时空序列的事件关联

由于内网攻击的证据通常分散在跨空间和时间分布的事件上，单纯依赖基于空间序列或者时间序列的事件关联都存在不足，因此将空间和时间事件关联集成在一起进行入侵检测非常重要。假设一个攻击过程可以在3个观测空间观察。每个观察空间都可以沿着时间序列将其事件与过程模型相关联。在每个时间 t，来自这 3 个观测空间的事件应该在空间上相关，则有以下2种方法可以对时空序列的事件进行关联。

5.1 确定性方法

如图3所示，多个观察空间可以独立地沿时序关系关联事件。每个时序关联的结果可以指示该特定观察空间的“正常”或“异常”行为。利用来自多个观测空间的结果，如第3节所述，可以使用一种码本或整数规划方法在空间上将这些时间关联结果相关联。

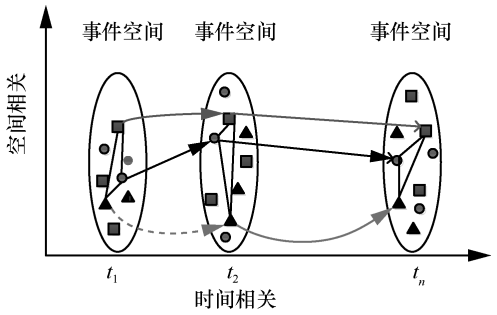


图3 事件的时空相关性

动态过程中的几个状态可能导致相同的观察结果。因此，隐藏状态是无法观察的。例如，HMM具有发射矩阵。在基于时间序列的关联中，观察序列可能源自隐藏状态序列的多个假设。通过多个观测空间，在每个时间 t，理论上本文可以使用码本方法来区分隐藏状态而不是攻击场景。只要在表1所示的相关矩阵中添加足够的具有可区分特征的观测空间，就可以使每个状态都可观察到。在这种情况下，一个时间关联过程可以直接将观察序列映射到状态序列。但是，在大多数情况下，不需要对每个观察结果区分出每个状态，因为可以根据过程模型的状态转移性质来得出隐藏状态的序列。

5.2 概率方法

将动态攻击过程的状态表示为 x，并将到时间 t 的状态 x 表示为 $x_1, x_2, \cdots, x_t$。假设有2个观测空间 o_p 和 o_q 来检测这个攻击过程，将 o_p 直至时间 t 的观测结果表示为 $\mathrm{rp}_{1:t}=(\mathrm{rp}_1, \mathrm{rp}_2, \cdots, \mathrm{rp}_t)$，并且将 o_q 直至时间 t 的观测结果表示为 $\mathrm{rq}_{1:t}=(\mathrm{rq}_1, \mathrm{rq}_2, \cdots, \mathrm{rq}_t)$。如果具有后验概率 $p(s_t \mid \mathrm{rp}_{1:t}, \mathrm{rp}_{1:t})$，则在每个时间 t，空间上分布和时间上分布的事件可以相互关联在一起。根据式(4)和式(7)，通常很难计算联合概率。但如果2个观测空间是独立的，则通过以下3个步骤计算 $p(s_t \mid \mathrm{rp}_{1:t}, \mathrm{rp}_{1:t})$。

步骤 1 在每个时间 t，对于每个观测空间，根据式(7)，关联时间事件并分别计算 $p(s_t \mid \mathrm{rp}_{1:t})$ 和 $p(s_t \mid \mathrm{rq}_{1:t})$。

步骤 2 根据式(5)，本文关联空间事件，并用步骤1得出的 $p(s_t \mid \mathrm{rp}_{1:t})$ 和 $p(s_t \mid \mathrm{rq}_{1:t})$ 计算 $p(s_t \mid \mathrm{rp}_{1:t}, \mathrm{rq}_{1:t})$，其中 $p(s_t)$ 可以递归计算。

步骤 3 $p(s_t \mid \mathrm{rp}_{1:t}, \mathrm{rq}_{1:t})$ 代替 $p(s_t \mid \mathrm{rq}_{1:t})$ 和 $p(s_t \mid \mathrm{rq}_{1:t})$。令 $t = t+1$ 并转到步骤1。

理论上，即使观测空间 o_p 和 o_q 是依赖的，也可以在式(9)所示的测量方程中增加另一个维度，并通过一个时间关联过程来关联事件，即

$$\begin{pmatrix} r_t \\ f_t \end{pmatrix} = \boldsymbol{H}x_t + \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \tag{11}$$

其中，$\boldsymbol{H}$ 如式(1)所示。在离散的情况下，正如5.1节所述，多个观测空间可以帮助区分隐藏状态并导出HMM中的稀疏发射矩阵。对于检测问题，可以使用式(10)中讨论的方法。

随着从分布式事件中挖掘出更多的证据，本文利用时间和空间模式的联合特征提升检测准确性并降低误报率。这种方法需要足够的知识来建立联合特征，为此本文实现了一个进程查询系统来扫描和关联分布的事件。进程查询系统允许用户在高抽象级别上精细过程特征，并将特征作为查询提交给关联系统。系统根据特征实时扫描和关联分布式事件，当前的进程查询系统仅支持基于时间序列的关联。如图4所示，进程查询系统由以下3个主要组件组成：用户界面、关联引擎和面向消息的中间件（MOM, message oriented middleware）。主题包括事件发布主题和事件订阅主题，其中事件以主题发布到MOM中，例如网络延迟。通过前端用户界面，用户可以定义具有高级抽象的过程特征，例如隐马尔可夫模型。

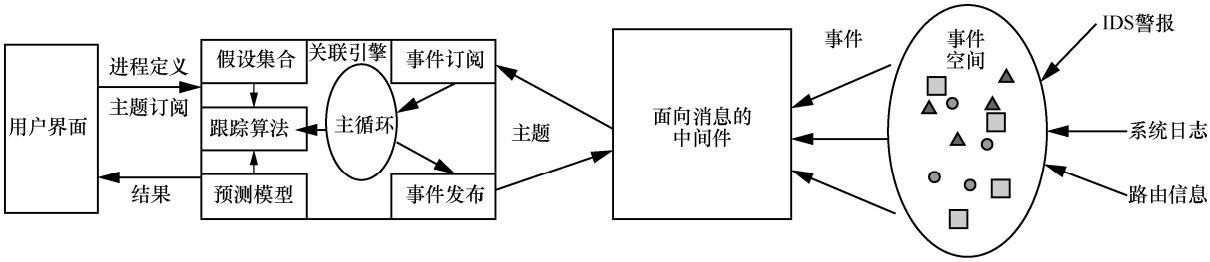


图 4　进程查询系统架构

事件订阅以主题方式将 MOM 消息提交给后端关联引擎。关联引擎解析查询并使用用户在 MOM 中订阅的事件。用户定义的过程模型调用多假设跟踪算法以扫描和关联传入事件。在事件关联期间，该算法递归地计算新的事件与现有事件假设相关的概率。新的事件被添加到具有最大似然的假设中，并且更新假设集。提交的过程模型用于计算新的事件与现有事件假设相关的条件概率。

6　实验分析

6.1　实验环境

在由异构的服务器和工作站组成的测试床上对该系统进行实验。测试床是一个隔离区（DMZ, demilitarized zone），由 4 个服务器（$host_1$~$host_4$）组成，每个服务器运行 RedHat Linux 系统和 Apache HTTP server。在另一个网络上的第五台服务器充当外部攻击主机，它发起一个拒绝服务（DoS, denial of service）攻击。每个服务器都配备了一个连续运行的用户空间传感器。该传感器除了收集系统状态外，它还监视 httpd、bash 和 sh 进程。传感器的状态提取模块每 4 s 轮询一次所有被监控的进程和系统指标。传感器的评估模块在 5 个样本的窗口上平滑这些连续的指标数据。

系统运行在一台本地 Sun Fire 210 服务器上，服务器上有双核 1 GHz UltraSPARC III 处理器和 Solaris 9，并且通过 TCP socket 监听主机的观察结果。进程查询引擎还监听来自网络入侵检测传感器（Snort）的观察结果，Snort 监视 DMZ 上的网络流量。因此，实验中有 2 个独立的观测空间。

系统向 Web 用户界面实时报告结果，该界面显示观察结果和相关的分数的轨迹。分数是一种定量度量，它介于 0.0~1.0 之间，反映了一个提交的进程模型被攻击的概率。因此，分数是对与事件关联模型相关的每个主机状态间接的和实时的度量。通过提交多个流程模型，可以有效地监控多个场景中的每个主机。

监控过程会记录每个被监控主机的分数和相应进度条的变化，其根本目标是评估在多阶段攻击中系统减少误报的能力，方法是将来自双方空间的观察结果关联起来，并计算反映每个主机相对状态的跟踪分数。

实验利用 Apache Web 服务器版本 2.0.40~2.0.52 中的一个漏洞，该漏洞允许攻击者通过向服务器大量发送专门制作的 HTTP GET 请求来引发 DoS。这会导致大量 CPU 和内存的消耗，并可能使整个服务器瘫痪，最终导致 DoS。因为 Apache 服务器在受到攻击时从不关闭连接，所以如果及时阻止父服务器进程 httpd，并使用备份的 Web 服务器，那么可以通过短时间的中断来阻止攻击。模拟攻击序列如下。

1) 攻击者使用 nmap 从外部攻击主机发起对 4 台 DMZ 服务器的非对称隐身端口扫描，以识别潜在的易受攻击的目标。

2) 攻击者等待所有扫描结果，并注意到每个服务器都在监听端口 8 000。

3) 攻击者等待 30 s，并对一台服务器（$host_3$）发起 Apache DoS 攻击，监听端口 8 000。

4) 攻击者等待 30 s 并终止攻击。

在这个实验中，系统只加载时空事件关联模型，场景特征的构建方法在上述模拟攻击序列下，并且未进行任何干涉，模拟攻击 200 次，收集 200 台受攻击主机和 600 台未受攻击而正常运行的计算机随时间演化的状态序列集合，人工分析这个演化序列集合来构建正常主机和受到攻击主机的场景特征。

6.2　实验结果

通过传感器的数据收集和在线分析得到场景特征后，本实验使用进程查询系统来进行攻击检测实验。图 5 中的时间序列显示了在大约 10 min 内，每个主机的最高分数的变化。纵坐标是受攻击的可

能性，由事件关联模型为每个服务器确定。

图 5 显示了模型在模拟人工 DoS 攻击前、攻击中和攻击后的分类和识别性能。本文将服务器状态区分为平常状态、可疑状态和受攻击状态。其中，平常状态下服务器正常运行；可疑状态下服务器仍在正常运行但疑似已经受到攻击，即其受攻击可能性开始增长，且增长速度较快；受攻击状态表示服务器受到攻击并且无法正常运行，受攻击可能性超过了 0.9。最初，每个服务器的状态估计是平常状态。扫描启动后不久（实验开始 280 s），每个服务器都转换到可疑状态。当 396 s 向 $host_3$ 发起 DoS 攻击时，其分数突然增加，迅速将其转移到受攻击状态。其他 3 个服务器继续正常运行，并保持在可疑状态。在 415 s，也就是 DoS 攻击开始 19 s 后，$host_3$ 被攻击的可能性增加到 0.95。在此阶段，管理员或自主修复机制可以通过终止相关的 httpd 进程来降低 $host_3$ 被攻击的程度。

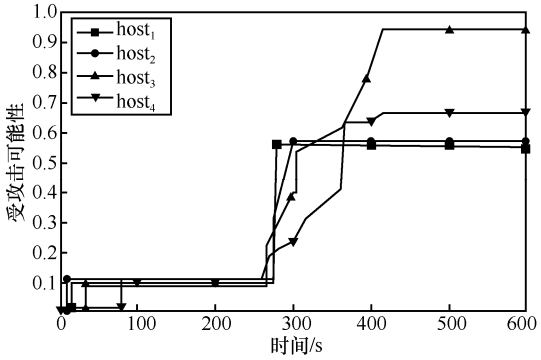


图 5　进程查询系统评估 4 台服务器受攻击的可能性

图 6 显示了同一时间段内报告给进程查询引擎的主机观测结果的累计增长情况。从图 6 可以看出，在实际 DoS 攻击开始之前，这 4 个服务器如何生成类似级别的主机观察结果。最明显的是，攻击发起后不久，受攻击服务器（$host_3$）的累计的主机观察值突然显著增加。图 5 对这一趋势的识别显示了简单的、短期的主机行为异常与可疑的网络侦察活动之间存在明显的相关性。在这次实验中，传感器提供了重要的指标，这些指标驱动受攻击服务器的状态评估为受攻击状态，从而能够在攻击仍在进行时将其快速检测出来。

关于场景特征的构建问题，系统提供了基于专家知识和基于数据挖掘这 2 种方法，并分别使用这 2 种场景特征来进行攻击检测实验，模拟攻击方法如上所述。

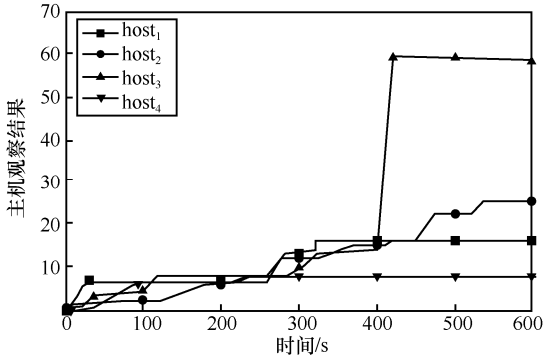


图 6　由每个被监控主机的传感器生成的主机观察结果，并报告给进程查询引擎

对于基于数据挖掘的方法，系统不仅使用上述收集到的状态序列集合，还使用了这 600 台主机和对应网络的资源随时间变化的情况，其中收集方法为从开始到 700 s 内，每 10 s 收集一次。对于某次攻击模拟在 400 s 时收集的资源情况如表 2 所示。通过使用数据挖掘系统，得到具有时间和空间模式的联合特征。

表 2　某次攻击模拟在 400 s 时的资源分布情况

资源	$host_1$	$host_2$	$host_3$	$host_4$
PROC.MEM	0	0	22	0
PROC.CPU	0	1	1	0
SYS.MEM	0	0	7	0
SYS.CPU	12	20	11	2
Snort	3	2	3	4

系统分别使用上述 2 种特征，设置报警阈值为 0.95（即受攻击的可能性超过 0.95），误报率和漏报率如表 3 所示。

表 3　　误报率和漏报率

方法	误报率	漏报率
基于专家知识的方法	0.008 3	0.025 0
基于数据挖掘的方法	0.006 6	0.003 3

由表 3 可知，使用这 2 种特征构建方法，进程查询系统在误报率上相差不是很大。而对于漏报率，基于数据挖掘的方法比基于专家知识的方法低很多。因此，可以证明随着从分布式事件中挖掘出更多的证据，使用具有时间和空间模式的联合特征可以降低误报率和漏报率。

7　结束语

本文提出了基于时空事件关联的内网攻击检

测方法，并基于该方法实现了一个可以根据用户对动态进程的高层描述对大量分布的事件进行关联查询的进程查询系统。实验表明，该系统在不增加明显计算开销的情况下，能够显著减少内网攻击检测的误报率。未来会在更大的网络范围来评估该系统运行时的稳定性和检测性能。

参考文献：

[1] SPRING N, MAHAJAN R, WETHERALL D. Measuring ISP topologies with rocketfuel[J]. ACM Sigcomm Computer Communication Review, 2002, 32(4): 133-145.

[2] DUARTE, FELIPE S L G, SIKANSI, et al. Nmap: a novel neighborhood preservation space-filling algorithm[J]. IEEE Transactions on Visualization & Computer Graphics, 2014, 20 (12):2063-2071.

[3] NORWAWI N M, GHAZALI O, FAAEQ M, et al. Detection algorithm for Internet worms scanning that used user datagram protocol[J]. International Journal of Information and Computer Security, 2019, 11(1): 17-32.

[4] TUNG T M, WANG C, WANG J. Understanding the behaviors of BGP-based DDoS protection services[C]//International Conference on Network and System Security. Springer, Cham, 2018: 463-473.

[5] LAROSE D T, LAROSE C D. Discovering knowledge in data: an introduction to data mining[M]. John Wiley & Sons, 2014.

[6] POOR H V. An introduction to signal detection and estimation[M]. Springer Science & Business Media, 2013.

[7] HOWSON C, URBACH P. Scientific reasoning: the Bayesian approach[M]. Open Court Publishing, 2006.

[8] BENFERHAT S, KENAZA T, MOKHTARI A. A naive bayes approach for detecting coordinated attacks[C]//32nd Annual IEEE International Computer Software and Applications Conference. IEEE, 2008: 704-709.

[9] CHANDOLA V, BANERJEE A, KUMAR V. Anomaly detection: a survey[J]. ACM Computing Surveys (CSUR), 2009, 41(3): 15.

[10] PATCHA A, PARK J M. An overview of anomaly detection techniques: existing solutions and latest technological trends[J]. Computer Networks, 2007, 51(12): 3448-3470.

[11] TEMPLETON S J. Detection and analysis of cyber attacks using bio-based concepts[D]. Pro Quest Dissertations Publishing, 2018.

[12] STONE L D, STREIT R L, CORWIN T L, et al. Bayesian multiple target tracking[M]. Artech House, 2013.

[13] KALMAN R E. A new approach to linear filtering and prediction problems[J]. Journal of basic Engineering, 1960, 82(1): 35-45.

[14] WITTEN I H, FRANK E, HALL M A, et al. Data mining: practical machine learning tools and techniques[M]. Morgan Kaufmann, 2016.

[15] COHEN P, WEST S G, AIKEN L S. Applied multiple regression/correlation analysis for the behavioral sciences[M]. Psychology Press, 2014.

[16] HO J W, WRIGHT M. Distributed detection of sensor worms using sequential analysis and remote software attestations[J]. IEEE Access, 2017, 5: 680-695.

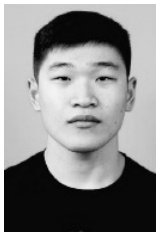
[17] SUN X, DAI J, LIU P, et al. Using Bayesian networks for probabilistic identification of zero-day attack paths[J]. IEEE Transactions on Information Forensics and Security, 2018, 13(10): 2506-2521.

[18] JIANG G, CHEN H, UNGUREANU C, et al. Multiresolution abnormal trace detection using varied-Length $ n $-grams and automata[J]. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 2006, 37(1): 86-97.

[作者简介]

孙伟（1980− ），男，山西忻州人，北京交通大学博士生，主要研究方向为计算机网络、信息安全和网络测量。

张鹏（1984− ），男，安徽淮南人，中国科学院副研究员、硕士生导师，主要研究方向为数据挖掘、网络安全。



何永全（1997− ），男，辽宁葫芦岛人，中国科学院大学硕士生，主要研究方向为并行计算与分布式系统。

邢丽超（1993− ），男，黑龙江哈尔滨人，中国科学院大学硕士生，主要研究方向为信息过滤与内容计算。