一种改进的网络安全态势量化评估方法

席荣荣 云晓春 张永铮 郝志宇
(中国科学院信息工程研究所 北京 100093)

摘 要 在基于隐马尔可夫模型的网络安全态势评估中,观测序列的获取和状态转移矩阵的确立是影响评估准确性的关键.目前观测序列多以随机方式获取,不能有效表征网络的安全性;而状态转移矩阵往往依据经验给出,具有很强的主观性.该文提出改进方法:首先,基于警报的统计特性提出警报质量的概念,依据警报质量获取的观测序列,可改进数据源的有效性;其次,基于安全事件和防护措施的博弈过程,提出确定状态转移矩阵的方法,并结合攻击成功的概率对其进行修正,提高状态转移矩阵的有效性.对比实验证明,基于改进算法生成的风险值对网络安全态势的量化更加合理.

关键词 观测序列;状态转移矩阵;警报质量;博弈矩阵;攻击成功的概率
中图法分类号 TP393 **DOI号** 10.3724/SP.J.1016.2015.00749

An Improved Quantitative Evaluation Method for Network Security

XI Rong-Rong YUN Xiao-Chun ZHANG Yong-Zheng HAO Zhi-Yu
(*Institute of Information Engineering, Chinese Academy of Sciences, Beijing* 100093)

Abstract Obtaining high-quality observation sequence and establishing correct state transition matrix are important to assess network security situation based on Hidden Markov Models. Currently observation sequence is obtained at random, it can't ensure the effectiveness of data source; and state transition matrix is established based on experience, it is subjective. An improved method is presented in this paper. Firstly, it obtains observation sequence based on quality of alert, which can improve the effectiveness of data source. Secondly, it determines state transition matrix based on the game of attack and defense, and adopts the successful probability of attack to amend it, which can improve the effectiveness of the matrix. The experiment demonstrates the improved method is more accurate, and can reflect the trend of network security more reasonably.

Keywords observation sequence; state transition matrix; quality of alert; game matrix; success probability of attack

1 引 言

面对当前日益复杂的网络安全问题,传统的安全检测和防护措施为网络管理者提供了海量多源数据,但由于大量误报或不相关信息的存在,使得数据信息具有不确定性.基于不确定信息评估网络的安全态势,准确性受到严重影响.为解决警报信息的不

收稿日期:2014-03-06;最终修改稿收到日期:2014-11-12.本课题得到国家"八六三"高技术研究发展计划项目基金(2012AA012803,2013AA014703)、国家科技支撑计划基金(2012BAH46B02)、国家自然科学基金(61100188)及中国科学院知识创新基金(XDA06030200)资助.**席荣荣**,女,1979年生,博士,助理研究员,中国计算机学会(CCF)会员,主要研究方向为网络安全、网络安全态势感知、网络测量.E-mail:xirongrong@iie.ac.cn.**云晓春**,男,1971年生,博士,研究员,中国计算机学会(CCF)会员,主要研究领域为信息安全、计算机网络.**张永铮**,男,1978年生,博士,研究员,中国计算机学会(CCF)会员,主要研究领域为网络安全.**郝志宇**,男,1980年生,博士,副研究员,中国计算机学会(CCF)会员,主要研究方向为网络安全、网络安全测量、网络模拟.

确定性问题，Årnes 等人[1]提出采用隐马尔可夫模型(Hidden Markov Model，HMM)对网络安全状态建模，基于统计模型特性解决警报信息的不确定性问题.但模型中的观测序列随机获取，无法保证数据源的有效性；状态转移矩阵根据经验值设定，不能客观反映网络安全状态的变化过程.后续研究不断对该方法进行改进，但仍存在局限性.如 Haslum 等人[2]基于安全状态的持续时间确定状态转移矩阵，但由于网络攻击行为的随机性，使得安全状态的持续时间难以确定.李伟明等人[3]利用遗传算法自动求解状态转移矩阵和观测向量分布矩阵，该方法可量化评估网络的安全态势.但方法中风险描述规则是针对特定攻击和风险的形式化描述，面对网络攻击行为的灵活性和多样性，构造通用的风险描述规则库存在很大难度.

针对观测序列的获取和状态转移矩阵的确定，本文提出改进方法，主要工作包括两点：第一，基于警报统计特性提出警报质量的概念，并依据警报质量获取观测序列，提高数据源的有效性；第二，基于安全事件和防护措施的博弈过程，提出确定状态转移矩阵的方法，并结合攻击成功的概率对其进行修正，提高状态转移矩阵的有效性.

2 基于 HMM 的网络安全态势评估方法

在实际网络中，网络安全状态是不可见的，但安全状态产生的警报信息是可见的，而且各安全状态与可能产生的警报信息之间存在特定的概率函数，警报序列能够揭露安全状态的变化信息，这与隐马尔可夫模型的核心思想是一致的[4].因此本文采用隐马尔可夫模型刻画网络安全状态的变化过程.

隐马尔可夫模型由五元组构成 $\boldsymbol{\lambda}=\{\boldsymbol{S},\boldsymbol{V},\boldsymbol{P},\boldsymbol{Q},\boldsymbol{\pi}\}$[5]，为了将其应用于网络安全态势评估，首先对五元组进行说明：

(1) $\boldsymbol{S}$ 为状态集合空间，$\boldsymbol{S}=\{S_1,S_2,\cdots,S_N\}$，其中 S_i 表示一个独立的状态，N 表示状态的数目.不同的安全事件导致网络进入不同的安全状态，依据安全事件的划分，将网络的安全状态划分为安全状态 G、探测状态 R、入侵状态 B 和攻陷状态 C，即 $\boldsymbol{S}=\{G,R,B,C\}$.

安全状态 G(Good)表示网络中不存在任何攻击行为，处于安全状态；

探测状态 R(Reconnaissance)表示网络中存在扫描类的行为，攻击者采集信息阶段；

入侵状态 B(Break-in)表示网络中存在破坏系统权限的行为，攻击者入侵网络阶段；

攻陷阶段 C(Compromised)表示网络已被攻陷，攻击者获得系统权限.

(2) $\boldsymbol{V}$ 是观测向量集合空间，$\boldsymbol{V}=\{\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_M\}$，其中 $\boldsymbol{v}_i$ 表示观测向量，M 表示各状态观测向量值的数目.网络安全状态是不可见的，可直接观测到的只有安全防护设备产生的警报信息.若直接将海量警报种类应用于 HMM 模型，会导致模型规模过大，严重影响运算效率.为解决上述问题，将原始警报进行分类.根据警报所代表的安全事件，将警报分为 4 类 $\boldsymbol{V}=\{g,r,b,c\}$.

g 表示在采样周期内，没有采集到任何警报信息；

r 表示扫描类的警报信息，如 ICMP ping，ICMP destination unreachable 等；

b 表示入侵类的警报信息，如 web cgi redirect access，bad traffic loopback 等；

c 表示获取 root 权限的警报信息，如 services rsh root 等.

(3) $\boldsymbol{P}$ 为状态转移矩阵，表示从一个状态转移到另一个状态的概率分布，$\boldsymbol{P}=\{p_{ij}\}$，其中 $p_{ij}=P(q_{t+1}=S_j\,|\,q_t=S_i)$，$1\leq i,j\leq N$，表示在 T 时刻网络处于 S_i 状态，在 $T+1$ 时刻处于 S_j 状态的概率.本文采用的状态转移矩阵如图 1 所示.

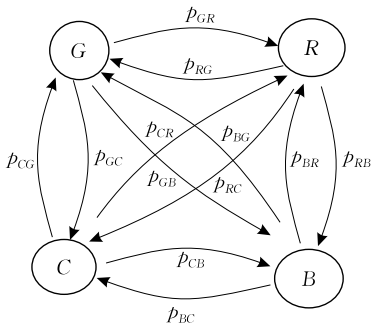


图 1 网络安全状态转移矩阵

(4) $\boldsymbol{Q}$ 为观测向量概率分布矩阵 $\boldsymbol{Q}=\{q_i(v_k)\}$，其中 $q_i(v_k)=p(o=v_k\,|\,q_t=S_i)$，$1\leq i\leq N$，$1\leq k\leq M$，表示 T 时刻，网络处于状态 S_i 且观测到的警报信息为 v_k 的概率.

(5) $\boldsymbol{\pi}$ 为初始状态概率分布矩阵，$\boldsymbol{\pi}=\{\pi_i\}$，其中 $\pi_i=p(q_1=S_i)$，$1\leq i\leq N$，表示在最初时刻，网络处于状态 S_i 的概率.

确定 HMM 五元组之后，根据每个采样周期获取的观测向量可实时更新网络 T 时刻处于状态 S_i

的概率$\lambda_t(i)$，再引入一个风险损失向量$\boldsymbol{C}(i)$，可求解任意 T 时刻网络总风险$\boldsymbol{R}_T$[6]：

$$\boldsymbol{R}_T=\sum_{i=1}^{N}\boldsymbol{R}_t(i)=\sum_{i=1}^{N}\lambda_t(i)\boldsymbol{C}(i) \tag{1}$$

其中：$\lambda_t(i)$表示 T 时刻网络处于状态 S_i的概率，$\boldsymbol{C}(i)$表示状态 S_i相应的风险损失，N 表示状态数目.

在安全态势评估过程中，观测序列的获取和状态转移矩阵的确立是影响评估准确性的关键. 目前观测向量多以随机方式获取，不具有代表性，不能表征网络的安全特性. 本文通过选取质量最高的警报作为 HMM 观测向量，可提高警报数据源的有效性. 对于状态转移矩阵的确立，往往根据专家经验设置，但手工设置主观性强，准确性不高，效果的好坏往往依赖于专家水平的高低. 为解决上述问题，基于安全事件和防护措施的博弈过程，提出确定状态转移矩阵的方法，并结合攻击成功的概率对其进行修正，提高状态转移矩阵的有效性.

3　算法改进

3.1　获取观测序列

入侵检测系统通常会产生海量警报信息，其中很大比例为误报或者不相关警报. 报警量大、不相关警报多，使得 HMM 观测向量的获取存在很大难度. 为了更加明确地定义警报表征网络安全特性的有效程度，参考质量因子[7]的思想，本文引入一个新的概念——警报质量.

定义 1.　警报质量 QoA(Quality of Alert)，指警报表征网络安全特性的有效程度. 警报的质量越高，越能有效表征网络的安全特性[8].

为了量化警报质量，首先将警报模型化为笛卡尔积的形式：

$$Alert=D_{A_1}\times D_{A_2}\times\cdots\times D_{A_n} \tag{2}$$

其中$(A_1,A_2,\cdots,A_n)$表示警报的属性[9]，D_{A_i}表示属性 A_i的取值范围. 警报的属性 A_i既包括警报的基本属性，如警报的源 IP，目的 IP，类型以及产生时间等，也包括警报的统计特性[10-13]，如警报出现频次 AlF(Alert Frequency)，警报关键程度 AlC(Alert Criticality)以及警报严重程度 AlS(Alert Severity)等. 基本属性用于描述警报的内在属性，与网络安全特性无关，因此不属于讨论范畴，主要讨论警报的统计属性：

AlF 为警报出现频次，表示单位时间内警报出现的相对次数. 安全设备的特性决定针对某一攻击行为，在短时间内会产生大量同类型的警报信息，即出现频次越高的警报信息越能刻画当前的网络攻击行为. 因此将警报出现频次作为警报质量的一个统计特征，定义为

$$AlF=\frac{\text{第 } i \text{ 条警报的数目}}{\text{所有警报的数目}} \tag{3}$$

AlC 为警报关键程度，指警报表征网络安全状态发生转变的强弱程度. 警报的关键程度越高表示网络安全状态发生转变的可能性越大. 监测过程中，若出现新的警报，表明网络中存在新的攻击行为，网络安全状态发生转变的可能性增加. 因此将警报的出现节点作为警报关键程度的指征. 根据警报的出现情况，将其分为 3 类：本采样周期内已出现过的警报；在之前的 N 个周期内出现过的警报；在之前的 N 个周期内未出现过的警报. 其相应的优先级别分别设为 1，2，3. 其中周期 N 的选择不仅与周期间隔有关，而且与周期的持续时间相关. 根据文献[14]提出的攻击事件持续时间的阈值及本文的采样周期，将周期 N 设置为 3.

AlS 为警报严重程度，表示警报产生的影响. 严重程度越高表明其对安全状态的影响越大. 参考 Snort 用户手册定义的攻击类别，将警报严重程度划分为高、中、低 3 个等级，分别设置为 1，2，3.

上述分析表明，警报出现频次、关键程度和严重程度是决定警报质量的关键因素，结合 3 个关键因素，给出警报质量的量化模型如图 2 所示.

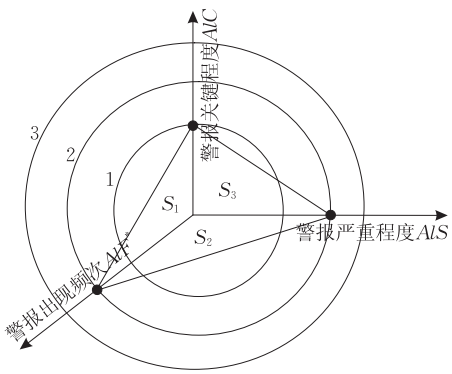


图 2　警报质量量化模型

模型将各属性值作为警报质量的量化因子，将警报质量定义为各属性值连线所包围的面积之和：

$$QoA=\sum_{i=1}^{3}S_i \tag{4}$$

$$S_1=\frac{1}{2}AlF^*\times AlC$$

$$S_2=\frac{1}{2}AlF^*\times AlS$$

$$S_3 = \frac{1}{2}AlC \times AlS$$

其中：AlC，AlS 分别表示警报关键程度和严重程度的属性值；AlF^* 表示警报出现频次 AlF 的标准化值. 因为 $AlF \in [0,1]$，与其他属性的取值范围不同，为均衡各属性对警报质量的影响，应保证各属性具有相同的值域. 因此采用离差标准化的反函数将警报出现频次 AlF 标准化为 AlF^*：

$$AlF^* = AlF(Max - Min) + Min = 2AlF + 1 \quad (5)$$

标准化值 $AlF^* \in [1,3]$，与其他属性具有相同的值域.

当警报的出现频次、关键程度和严重程度的属性值确定时，可根据式(4)求解警报质量. 如图 2 所示，当 $AlF=0.4$，$AlC=1$，$AlS=2$ 时，$AlF^*=1.8$，$QoA=(1.8,1,2)=3.7$.

在每个采样周期，通过选取质量最高的警报，获取 HMM 的观测向量. 依据警报质量获取的观测向量可有效改进数据源，提高评估准确性.

3.2 确定状态转移矩阵

状态转移矩阵是影响安全态势评估准确性的关键. 传统的状态转移矩阵往往依据经验给出，具有很强的主观性. 本文将基于网络安全状态转移过程确定状态转移矩阵.

事件驱动的网络安全状态转移过程将网络视作一个状态机[15]，安全事件则是网络状态转移的依据. 事件驱动的网络状态转移模型如图 3 所示.

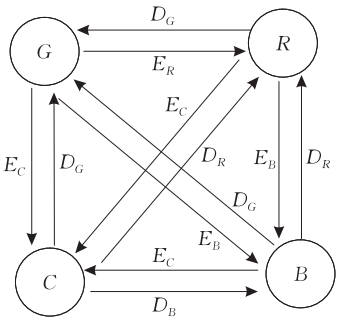


图 3 事件驱动的安全状态转移模型

其中 G,R,B,C 表示网络状态；E_i 表示安全事件；D_i 表示防护措施. 若 T 时刻网络处于状态 S_i，在此状态下，发现异常情况形成安全事件 E_j，或者执行了安全防护措施 D_j，则导致网络 $T+1$ 时刻进入状态 S_j. 该过程表示为

$$S_i \xrightarrow{E_j \cup R_j} S_j, \quad S_i, S_j \in \{G,R,B,C\}.$$

上述分析表明网络状态转移的实质为安全事件与防护措施的博弈. 因此基于博弈过程提出了确定状态转移矩阵的方法，该方法根据博弈过程确定状态转移矩阵，并结合攻击成功的概率对其进行修正.

(1) 确定状态转移矩阵

基于安全事件与防护措施的博弈，网络安全状态的转移可描述为 $|\boldsymbol{E}| \times |\boldsymbol{D}|$ 矩阵，如表 1.

表 1 安全事件与防护措施的 $|\boldsymbol{E}| \times |\boldsymbol{D}|$ 矩阵

	防护措施 D_i	…	防护措施 D_m
E_1	S_j	…	S_l
…	…	…	…
E_n	S_m	…	S_n

$\boldsymbol{E}$ 为安全事件，网络中存在多种安全事件，为简化分析，定义 4 种安全事件：无安全事件 ϕ、扫描类安全事件 E_R、入侵类安全事件 E_B、获取系统权限类安全事件 E_C；定义 $\boldsymbol{E}=\{\phi,E_R,E_B,E_C\}$.

$\boldsymbol{D}$ 为网络中的防护措施，将网络中的防护措施简化为：无任何防护措施 ϕ，监测类防护措施 D_S，阻止类防护措施 D_F，修复类防护措施 D_R；定义 $\boldsymbol{D}=\{\phi,D_S,D_F,D_R\}$.

$|\boldsymbol{E}| \times |\boldsymbol{D}|$ 矩阵直观给出了安全状态转移的分布情况. 由矩阵可知，状态 S_i 转移到 S_j 的概率 $p_{ij} = \frac{N_{S_j}}{\sum_{i \in \{G,R,B,C\}} N_{S_i}}$. 其中 p_{ij} 表示从状态 S_i 转移到 S_j 的概率，即从 S_i 转移到 S_j 的状态数目占从 S_i 转出的所有可能状态数目的比重. N_{S_j} 表示从状态 S_i 转移到状态 S_j 的数目，$\sum_{i \in \{G,R,B,C\}} N_{S_i}$ 表示从状态 S_i 转出的所有可能状态数目之和. 根据各状态的概率分布，可确定状态转移矩阵.

例如当网络处于安全状态 G 时，安全事件与防护措施的博弈过程[16]如下：

当网络中无安全事件时，则无论采取何种防护措施，网络均处于安全状态.

当网络中发生扫描类安全事件时，若网络中无任何防护措施，则网络转移到探测状态 R；若网络中存在监测类防护措施，表明网络可检测到扫描行为，但并不阻止，则网络转移到探测状态 R；若网络中存在阻止类或者修复类防护措施，则网络保持在安全状态 G.

当网络中发生入侵类安全事件时，若网络中无任何防护措施，则网络转移到入侵状态 B；若网络中存在监测类防护措施，则网络安全状态的转移取决于管理人员采取的策略. 本文对 50 名网络安全管理人员进行了调查，发现其中 27 名安全管理人员实时处理安全事件，另外 23 名安全管理人员周期性对系统进行维护. 根据不同的管理策略，将网络保持在安

全状态 G 和转移到入侵状态 B 的概率均设为 1/2. 若网络具有阻止类或者修复类防护措施，则网络保持在安全状态 G.

当网络中存在攻陷类安全事件时，分析同入侵类安全事件.

依据上述分析，构建安全状态 G 的 $|\boldsymbol{E}|\times|\boldsymbol{D}|$ 矩阵如表 2.

表 2　安全状态 G 的 $|\boldsymbol{E}|\times|\boldsymbol{D}|$ 矩阵

安全事件	防护措施			
	ϕ	D_s	D_F	D_R
ϕ	G	G	G	G
E_R	R	R	G	G
E_B	B	$1/2G$, $1/2B$	G	G
E_C	C	$1/2G$, $1/2C$	G	G

根据状态 G 的 $|\boldsymbol{E}|\times|\boldsymbol{D}|$ 矩阵，求解状态 G 转移到其他状态的概率分别为

$$p_{GG}=\frac{\text{转移到状态 }G\text{ 的状态数}}{\text{所有可能的状态数}}=\frac{11}{16}=0.69;$$

$$p_{GR}=\frac{\text{转移到状态 }R\text{ 的状态数}}{\text{所有可能的状态数}}=\frac{2}{16}=0.13;$$

$$p_{GB}=\frac{\text{转移到状态 }B\text{ 的状态数}}{\text{所有可能的状态数}}=\frac{3/2}{16}=0.09;$$

$$p_{GC}=\frac{\text{转移到状态 }C\text{ 的状态数}}{\text{所有可能的状态数}}=\frac{3/2}{16}=0.09.$$

分别构建状态 R,B,C 的 $|\boldsymbol{E}|\times|\boldsymbol{D}|$ 矩阵，则可初步确定状态转移矩阵如下所示：

$$\boldsymbol{P}=\begin{vmatrix} p_{GG} & p_{GR} & p_{GB} & p_{GC}\\ p_{RG} & p_{RR} & p_{RB} & p_{RC}\\ p_{BG} & p_{BR} & p_{BB} & p_{BC}\\ p_{CG} & p_{CR} & p_{CB} & p_{CC}\end{vmatrix}=\begin{vmatrix} 0.69 & 0.13 & 0.09 & 0.09\\ 0.21 & 0.52 & 0.13 & 0.13\\ 0.21 & 0.04 & 0.62 & 0.13\\ 0.21 & 0.04 & 0.04 & 0.71\end{vmatrix}$$

上述分析中存在一个假设，即若安全事件发生，则攻击行为肯定成功. 但我们已知，真实网络中并非所有的攻击行为都会成功[17]. 为了使得状态转移矩阵更加合理，对攻击行为成功的概率进行分析，修正状态转移矩阵.

(2) 修正状态转移矩阵

通过进一步分析攻击行为，发现攻击成功的概率与攻击者的类型，攻击者的能力，攻击所需的资源及网络的防护措施相关，因此定义状态转移矩阵的修正函数为

$$p=\sum_{i,j,k\in S}AM(i)\frac{AC(i)}{10AR(j)+10^{D(k)}}\tag{6}$$

AM 为攻击者的类型，网络中的攻击行为大部分由初始攻击者发起，少部分由熟练攻击者发起，只有极少部分由专业人员发起，但由极少数专业人员发起的攻击却对网络安全状态的变化起决定作用，这与帕累托分析法[16]的思想是一致的. 因此运用帕累托法则，将攻击者为专业人员，熟练攻击者和初始攻击者的概率分别设定为 1/25、4/25 和 4/5.

AC 为攻击者的能力，攻击者的能力越强，攻击成功的概率越高. 按照能力的不同将攻击者划分为专业人员、熟练攻击者和初始攻击者 3 个等级，其能力分别设定为 100、10、1.

AR 为完成攻击所需要的资源，攻击需要的资源越多，攻击成功的概率越低. 按照安全事件对资源需求的不同，将其划分为需要大量资源、需要部分资源和不需要资源 3 个等级，分别设定为 100、10、1. 与攻击者的能力相比，攻击对资源的需求对攻击成功概率的影响要弱一些，为了提高函数的合理性，采用系数 10 进行调整.

D 为网络中的防护措施，防护措施级别越高，攻击成功的概率越低. 根据定义的 4 种防护措施，将无任何防护措施 ϕ，监测类的防护措施 D_S，阻止类的防护措施 D_F 和修复类的防护措施 D_R 分别量化为 0、1、2、3. 与完成攻击所需的资源相比，防护措施对于攻击成功的影响更大一些，因此采用指数形式表征防护措施.

i,j,k 的取值由划分面决定. 划分面是指进入某种网络安全状态所必须的条件组合. 例如网络若要从安全状态 G 进入探测状态 R，划分面如图 4 所示. 任何能力的攻击者都可以实现，$AC\in\{1,10,100\}$；且相应的 $AM\in\{4/5,4/25,1/25\}$；不需要任何资源，$AR\in\{1\}$；网络中不存在阻止类及更高级的防护措施，$D\in\{0,1\}$；得到的修正函数 p 如下：

$$p=\sum_{i=1}^{100}\sum_{j=1}^{1}\sum_{k=0}^{1}AM(i)\frac{AC(i)}{10AR(j)+10^{D(k)}}=0.9$$

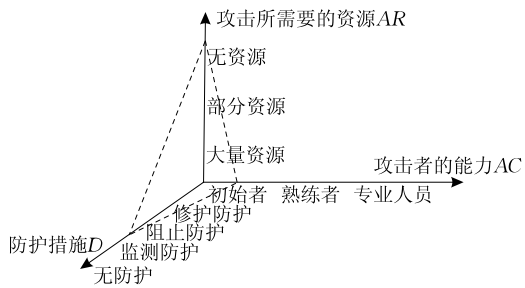


图 4　修正函数划分面

采用修正函数 p 更新状态 G 转移到状态 R 的概率 P'_{GR}：

$$P'_{GR}=P_{GR}\times p=0.13\times 0.9=0.117$$

同理可分别求得状态 B,C 的修正函数，利用修正函数更新状态转移矩阵，修正结果如下：

$$\boldsymbol{P}=\begin{vmatrix} p_{GG} & p_{GR} & p_{GB} & p_{GC} \\ p_{RG} & p_{RR} & p_{RB} & p_{RC} \\ p_{BG} & p_{BR} & p_{BB} & p_{BC} \\ p_{CG} & p_{CR} & p_{CB} & p_{CC} \end{vmatrix}=\begin{vmatrix} 0.871 & 0.117 & 0.01 & 0.002 \\ 0.003 & 0.973 & 0.021 & 0.003 \\ 0.003 & 0.004 & 0.99 & 0.003 \\ 0.003 & 0.004 & 0.02 & 0.973 \end{vmatrix}$$

3.3 网络状态更新算法

确定观测序列 O_t 和模型参数 $\lambda=\{\boldsymbol{S},\boldsymbol{V},\boldsymbol{P},\boldsymbol{Q},\boldsymbol{\pi}\}$ 之后，采用与文献[1]相同的前向算法更新 T 时刻网络处于状态 S_i 的概率 $\lambda_t=\{\lambda_t(i)\}$. 算法的复杂度为 $O(N^2)$，算法如算法 1 所示.

算法 1. 网络状态概率更新算法.

输入：T 时刻的观测向量 $\boldsymbol{O}_t$，HMM 的模型参数 λ

输出：T 时刻网络处于状态 S_i 的概率 $\lambda_t(i)$

1. If $T=1$ then
2. for $i=1$ to N do
3. $\alpha_1(i)=\pi_i q_i(O_1)$
4. $\lambda_1(i)=\dfrac{\alpha_1(i)}{\sum_{i=1}^{N}\alpha_1(i)}$
5. End for
6. Else
7. For $i=1$ to N do
8. $\alpha_t(i)=q_i(O_t)\sum_{j=1}^{N}\alpha_{t-1}(j)p_{ji}$
9. $\lambda_t(i)=\dfrac{\alpha_t(i)}{\sum_{i=1}^{N}\alpha_t(i)}$
10. End for

利用该算法可实时更新 T 时刻网络处于状态 S_i 的概率 $\lambda_t(i)$，结合风险损失向量 $\boldsymbol{C}(i)$，即可求解网络安全风险值 R_T[18].

4 实验验证

为了验证方法的有效性，采用林肯实验室提供的经典数据源 LLDOS1.0① 进行验证. LLDOS1.0 是一个 DDOS 攻击场景测试集，提供了 DDOS 攻击的网络流量. 攻击场景分为 5 个阶段，每个阶段的安全事件导致网络处于不同的安全状态.

根据说明文档，各阶段的信息如表 3 所示. 为了再现攻击场景，采用流量重放技术，将原始流量导向 snort，由其充当采集器对数据进行分析. 采集信息见表 3 最右侧，相对于攻击说明文档，各阶段的警报数目略有差别，主要是由于不同的 snort 规则集导致的，但警报数目总体趋势是一致的，且所有算法均采用相同数据集，并不影响实验结果.

表 3 LLDOS 1.0 的攻击说明

	描述	开始时间	结束时间	数据包	警报数	Snort 警报
1	获得活动主机列表；	9:51:36	9:52:00	40	31	40
2	找到薄弱 soliar 主机；	10:08:07	10:18:10	158	32	72
3	由 sadmind 缓冲区溢出漏洞侵入系统；	10:33:10	10:35:01	225	35	11
4	在控制主机上安装 DDOS 攻击木马；	10:50:01	10:50:54	520	22	21
5	由控制主机攻击远程服务器；	11:26:15	11:34:21	73924	33754	1404
总计	LLDOS1.0 所有数据流	9:21:36	12:35:48	649787	—	3262

设定采样周期 Δt 为 5 min，在每个采样周期内，从 snort 产生的警报流中选取质量最高的警报作为观测向量. 获取的观测序列如附表 1 所示.

LLDOS1.0 作为一个网络流量的基准评测数据集，并不提供网络防护措施的说明，依据对安全事件及警报信息的分析[19]，分别构建各状态的 $|\boldsymbol{E}|\times|\boldsymbol{D}|$ 矩阵，其中安全状态 G 的 $|\boldsymbol{E}|\times|\boldsymbol{D}|$ 矩阵如表 4.

表 4 安全状态 G 的 $|\boldsymbol{E}|\times|\boldsymbol{D}|$ 矩阵

安全事件	防护措施		
	ϕ	D_S	D_F
ϕ	G	G	G
E_R	R	R	G
E_B	B	G	G
E_C	C	G	G

根据 $|\boldsymbol{E}|\times|\boldsymbol{D}|$ 矩阵求解状态转移的概率分布，确定状态转移矩阵如下：

$$\boldsymbol{P}=\begin{vmatrix} p_{GG} & p_{GR} & p_{GB} & p_{GC} \\ p_{RG} & p_{RR} & p_{RB} & p_{RC} \\ p_{BG} & p_{BR} & p_{BB} & p_{BC} \\ p_{CG} & p_{CR} & p_{CB} & p_{CC} \end{vmatrix}=\begin{vmatrix} 0.667 & 0.167 & 0.083 & 0.083 \\ 0.333 & 0.417 & 0.125 & 0.125 \\ 0.25 & 0.167 & 0.417 & 0.167 \\ 0.25 & 0.167 & 0.25 & 0.333 \end{vmatrix}$$

① Lincoln Laboratory. Lincoln laboratory scenario (DDoS) http://www.ll.mit.edu/mission/communications/cyber/CSTcorpora/ideval/data/1999data.html, 2013.03.16

修正后的状态转移矩阵：

$$\boldsymbol{P}=\begin{vmatrix} p_{GG} & p_{GR} & p_{GB} & p_{GC} \\ p_{RG} & p_{RR} & p_{RB} & p_{RC} \\ p_{BG} & p_{BR} & p_{BB} & p_{BC} \\ p_{CG} & p_{CR} & p_{CB} & p_{CC} \end{vmatrix} = \begin{vmatrix} 0.839 & 0.15 & 0.009 & 0.002 \\ 0.005 & 0.972 & 0.02 & 0.003 \\ 0.004 & 0.017 & 0.975 & 0.004 \\ 0.004 & 0.017 & 0.125 & 0.854 \end{vmatrix}$$

为了保证与 Årnes 方法的可比性，采用与 Årnes 方法相同的观测向量概率分布矩阵 $\boldsymbol{Q}$，初始状态概率分布矩阵 $\boldsymbol{\pi}$ 及风险损失向量 $\boldsymbol{C}(i)$.

$$\boldsymbol{Q}=\begin{vmatrix} q_G(1) & q_G(2) & q_G(3) & q_G(4) \\ q_R(1) & q_R(2) & q_R(3) & q_R(4) \\ q_B(1) & q_B(2) & q_B(3) & q_B(4) \\ q_c(1) & q_c(2) & q_c(3) & q_c(4) \end{vmatrix} = \begin{vmatrix} 0.8999 & 0.02 & 0.08 & 0.0001 \\ 0.6699 & 0.25 & 0.08 & 0.0001 \\ 0.7350 & 0.10 & 0.16 & 0.005 \\ 0.8000 & 0.04 & 0.11 & 0.050 \end{vmatrix}$$

$$\boldsymbol{\pi}=\begin{vmatrix} \pi_G & \pi_R & \pi_B & \pi_C \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 & 0 \end{vmatrix}$$

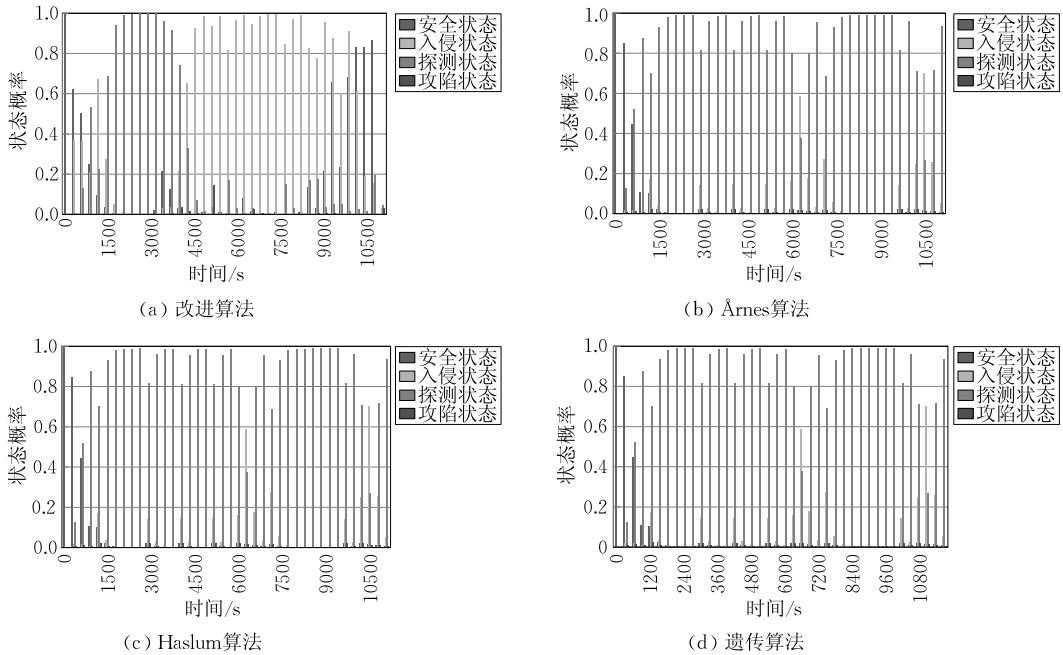
$$\boldsymbol{c}=\begin{vmatrix} c_G & c_R & c_B & c_C \end{vmatrix} = \begin{vmatrix} 0 & 25 & 50 & 100 \end{vmatrix}$$

确定观测序列 O_t 和模型参数 $\lambda=\{S,V,P,Q,\pi\}$ 之后，采用算法 1 实时更新网络在 T 时刻处于状态 S_i 的概率 $\lambda_t(i)$，实验结果如图 5 所示.

图 5(a)改进算法生成的状态概率表明在 1200 s 之前及 9600 s 之后，网络处于安全状态；在 1200 s～4200 s 之间网络处于探测状态；4200 s～9600 s 之间网络处于入侵状态. 图 5(b) Årnes 算法生成的状态概率表明在 900 s 之前，网络处于安全状态，其余时段网络处于探测状态. 图 5(c) Haslum 算法生成的状态概率表明网络交替处于探测状态和安全状态. 图 5(d)遗传算法生成的状态概率值具有均衡趋势，无明显优势状态. 上述分析表明，改进算法生成的状态概率与攻击场景的描述基本一致.

为量化评估网络的安全态势，结合风险损失向量 $\boldsymbol{C}(i)$，可求解 T 时刻各算法生成的网络风险值 R_T，结果如图 6 所示.

图 6 表明改进算法在攻击入侵阶段生成的网络风险值明显高于探测阶段的风险值，且在 5400 s 出现小波峰，表明攻击行为异常；Årnes 算法由于一直处于探测状态，其产生的网络风险值维持在 50 左右；Haslum 算法由于交替处于探测状态和安全状态，其产生的网络风险值亦处于波动状态；遗传算法生成的网络风险值无明显变化趋势. 上述分析表明，基于改进算法生成的风险值对网络安全态势的量化更加合理.



(a) 改进算法　(b) Årnes算法

(c) Haslum算法　(d) 遗传算法

图 5　各算法状态概率分布图

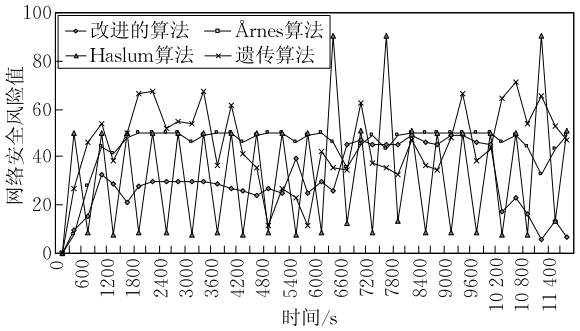


图 6 各算法生成的网络风险值

5 讨 论

(1)改进算法对于网络安全状态变化趋势的判定与攻击场景的描述基本一致,表明改进算法可以较准确地评估网络安全状态的变化趋势.

(2)Årnes 算法可准确判定网络的探测状态,但同时将入侵行为判定为探测状态,这主要是因为 Årnes 算法的观测序列 O_t 是随机采样获取,O_t 不能有效表征网络的真实状况,影响算法准确性.

(3)Haslum 算法采用的状态转移矩阵中 56% 的转移概率值为零值,而且非零值中采用互补方式描述反向的状态转移,即整个状态转移矩阵仅有 3 个有效值,使得生成的状态概率值存在误差.不合理的状态转移矩阵影响 Haslum 算法的准确性.

(4)遗传算法根据警报的严重程度,将警报划分为 10 个等级,即算法中观测向量值为 10.细化的观测向量使得算法过于敏感,导致生成的状态概率值波动频繁,无法判定网络的变化趋势.

(5)根据攻击场景描述,5400 s 时,攻击者获得了控制主机上的系统权限,网络应处于被攻陷状态.但图 5 中各算法对于攻陷状态的表征均不明显.各算法生成的攻陷状态概率值的分布情况如图 7 所示.

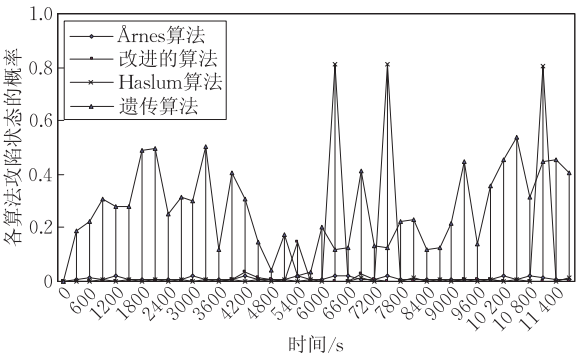


图 7 各算法攻陷状态的概率值分布

图 7 表明改进算法生成的攻陷状态概率值一直处于 0.001 的极小值,但是在 5400 s 时,出现了小幅波动,达到了 0.14.Årnes 算法生成的攻陷状态概率值一直维持在 0.003.Haslum 算法在 6300 s,7500 s 以及 11100 s 出现了 0.8 的高概率值.遗传算法在多点出现了 0.5 左右的高概率值,但根据攻击场景描述,这些点并没有出现高危险级别安全事件.分析表明,改进算法体现了攻陷状态的变化,但不明显.

5400 s 时,各算法生成的状态概率值如图 8 所示.

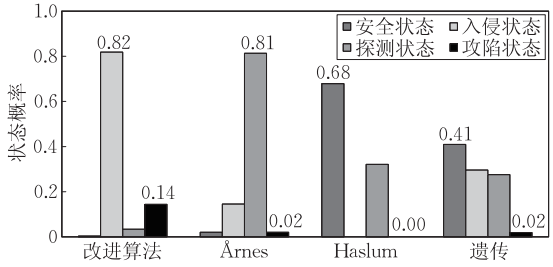


图 8 5400 s 时各算法生成的状态概率值

图 8 表明 5400 s 时,虽然改进算法攻陷状态的概率值提升到 0.14,但入侵状态的概率值为 0.82,远高于攻陷状态的概率值,所以网络处于入侵状态,与攻击场景描述的网络处于攻陷状态出现不一致.因为 T 时刻网络处于状态 S_i 的概率是由所有可能到达该状态的路径之和表示,它不仅与 T 时刻的观测向量有关,还与隐含了 $T-1$ 时刻之前观测序列信息的部分概率 $\alpha_{t-1}(i)$ 相关.在 5400 s 之前,网络中没有出现过攻陷警报,当第一次出现攻陷警报时,算法对攻陷状态的概率值有所提升,但由于累积效应,并不是最大的状态概率值.所以改进算法体现了攻陷状态的变化,但不明显.

6 结束语

本文对 Årnes 提出的基于隐马尔可夫过程的网络安全态势评估模型进行了改进,主要工作有两点:第一,基于警报统计特性提出警报质量的概念,依据警报质量获取的观测向量,可改进数据源的有效性;第二,基于安全事件和防护措施的博弈过程,提出确定状态转移矩阵的方法,并结合攻击成功的概率对其进行修正,提高转移矩阵的有效性.对比实验证明,基于改进算法生成的风险值对网络安全态势的量化更加合理.但该方法还存在局限性,例如现在的数据源主要是基于入侵检测系统产生的警报信

息，如果可以采集更多的信息完善观测序列，则可进一步提高方法的准确性；另外目前算法求解的只是下一个采样周期各安全状态的概率，如果可实现对未来多个采样周期内安全状态出现概率的求解，则可以实现对网络安全状态的预测，为网络管理员提供决策依据. 因此下一步工作主要集中在这两个方向.

参 考 文 献

[1] Årnes A, Valeur F, Vigna G, et al. Using hidden markov models to evaluate the risks of intrusions//Proceedings of the Recent Advances in Intrusion Detection. Hamburg, Germany, 2006: 145-164

[2] Haslum K, Moe M E G, Knapskog S J. Real-time intrusion prevention and security analysis of networks using HMMs// Proceedings of the 33rd IEEE Conference on Local Computer Networks. Montreal, Canada, 2008: 927-934

[3] Li Wei-Ming, Lei Jie, Dong Jing, et al. An optimized method for real time network security quantification. Chinese Journal of Computers, 2009, 32(4): 793-804(in Chinese)
(李伟明，雷杰，董静等. 一种优化的实时网络安全风险量化方法. 计算机学报，2009，32(4)：793-804)

[4] Khreich W, Granger E, Miri A, et al. Adaptive ROC-based ensembles of HMMs applied to anomaly detection. Pattern Recognition, 2012, 45(1): 208-230

[5] Rabiner L R. A tutorial on hidden Markov models and selected applications in speech recognition. Proceedings of the IEEE, 1989, 77(2): 257-286

[6] Årnes A, Sallhammar K, Haslum K, et al. Real-time risk assessment with network sensors and intrusion detection systems. Computational Intelligence and Security, 2005, 3802: 388-397

[7] Jonsson E, Olovsson T. A quantitative model of the security intrusion process based on attacker behavior. IEEE Transactions on Software Engineering, 1997, 23(4): 235-245

[8] Sendi A S, Dagenais M, Jabbarifar M, Couture M. Real time intrusion prediction based on optimized Alerts with Hidden Markov Model. Journal of Networks, 2012, 7(2): 311-321

[9] Gong Jian, Mei Hai-Bin, Ding Yong, et al. Multi feature correlation redundance elimination of intrusion event. Journal of Southeast University (Natural Science Edition), 2005, 35(3): 366-371(in Chinese)
(龚俭，梅海彬，丁勇等. 多特征关联的入侵事件冗余消除. 东南大学学报(自然科学版)，2005，35(3)：366-371)

[10] Ning P, Cui Y, Reeves D S. Constructing attack scenarios through correlation of intrusion alerts//Proceedings of the 9th ACM conference on Computer and Communications Security. New York, USA, 2002: 245-254

[11] Ning P, Xu D, Healey C G, et al. Building attack scenarios through integration of complementary alert correlation methods//Proceedings of the 11th Annual Network and Distributed System Security Symposium. San Diego, USA, 2004: 97-111

[12] Wang L, Liu A, Jajodia S. Using attack graphs for correlating, hypothesizing, and predicting intrusion alerts. Computer Communications, 2006, 29(15): 2917-2933

[13] Wang L, Singhal A, Jajodia S. Toward measuring network security using attack graphs//Proceedings of the 2007 ACM Workshop on Quality of Protection. Alexandria, USA, 2007: 49-54

[14] Debar H, Wespi A. Aggregation and correlation of intrusion-detection alerts//Proceedings of the Recent Advances in Intrusion Detection. Davis, USA, 2001: 85-103

[15] Han Rui-Sheng, Zhao Bin, Xu Kai-Yong. Policy-based integrative network security management system. Computer Engineering, 2009, 35(8): 201-204(in Chinese)
(韩锐生，赵彬，徐开勇. 基于策略的一体化网络安全管理系统. 计算机工程，2009，35(8)：201-204)

[16] Roger S. Software Engineering: A Practitioner's Approach. 7th Edition. Boston, Mass: McGraw-Hill, 2010

[17] Dacier M, Pouget F, Debar H. Honeypots: Practical means to validate malicious fault assumptions//Proceedings of the 10th IEEE Pacific Rim International Symposium on Dependable Computing. Papeete, French Polynesia, 2004: 383-388

[18] Holsopple J, Sudit M, Nusinov M, et al. Enhancing situation awareness via automated situation assessment. IEEE Communications Magazine, 2010, 48(3): 146-152

[19] Barford P, Kline J, Plonka D, et al. A signal analysis of network traffic anomalies//Proceedings of the 2nd ACM SIGCOMM Workshop on Internet Measurement. Marseille, France, 2002, 6(8): 71-82

XI Rong-Rong, born in 1979, Ph.D., assistant professor. Her main research interests include network security, network security situational awareness, network measurement.

YUN Xiao-Chun, born in 1971, Ph.D., professor, Ph. D. supervisor. His main research interests include information security, computer network.

ZHANG Yong-Zheng, born in 1978, Ph. D., professor. His main research interests include network security.

HAO Zhi-Yu, born in 1980, Ph. D., associate professor. His main research interests include network security, network measurement and network simulation.

Background

With issues in network security, traditional detection and prevention devices generate massive alerts. However, many of them are false positive or non-relevant alerts. These uncertain alerts impact the accuracy of assessment.

In order to solve the uncertainty problem of alerts, researchersmodel network security situation using HMM. It can solve the uncertainty of alerts based on the statistics model. But it is difficult to determine the model parameters. For example, the alert observation sequence is obtained at random, it can't ensure the validity of the data source; the state transition matrix is set based on experiences, it cannot reflect the changing process of network security situation objectively. To solve these problems, this paper presents an improved method.

The method obtains the observation sequence based on the quality of alert, which ensures the effectiveness of the data source. The method determines the state transition matrix based on the game of attack and defense, which can reflect the changing process of network security situation objectively. The experiment demonstrates the improved method is more accurate, and can reflect the change of network security situation more effectively.

This research is supported by The National High Technology Research and Development Program of China under Grant Nos. 2012AA012803 and 2013AA014703; The National Science and Technology Support Program under Grant No. 2012BAH46B02; the National Natural Science Foundation of China under Grant No. 61100188; the Knowledge Innovation Program of the Chinese Academy of Sciences under Grant No. XDA06030200.

附录 1. 每个采样周期采集的警报数及获取的观测向量.

	Start Time	End Time	Alert Num.	Candidate alert	Corresponding class
1	9:20:00	9:24:59	25	Sensitive data Email address	2 break-in
2	9:25:00	9:29:59	22	Attack response directory listing	2 break-in
3	9:30:00	9:34:59	26	Info telnet access	3 reconnaissance
4	9:35:00	9:39:59	36	Info telnet access	3 reconnaissance
5	9:40:00	9:44:59	111	TCP window closed before receiving data	2 break-in
6	9:45:00	9:49:59	58	ICMP ping/reply	3 reconnaissance
7	9:50:00	9:54:59	79	ICMP ping/reply	3 reconnaissance
8	9:55:00	9:59:59	22	Info telnet access	3 reconnaissance
9	10:00:00	10:04:59	44	ICMP ping/reply	3 reconnaissance
10	10:05:00	10:09:59	36	ICMP destination unreachable	3 reconnaissance
11	10:10:00	10:14:59	94	ICMP destination unreachable	3 reconnaissance
12	10:15:00	10:19:59	62	ICMP destination unreachable	3 reconnaissance
13	10:20:00	10:24:59	54	Netbois NT null session	2 break-in
14	10:25:00	10:29:59	21	—	4 good
15	10:30:00	10:34:59	47	telnet login incorrect	2 break-in
16	10:35:00	10:39:59	40	Attack response directory listing	2 break-in
17	10:40:00	10:44:59	34	Sensitive data Email address	2 break-in
18	10:45:00	10:49:59	53	Sensitive data Email address	2 break-in
19	10:50:00	10:54:59	78	Rservicesrsh root	1 compromised
20	10:55:00	10:59:59	64	Sensitive data Email address	2 break-in
21	11:00:00	11:04:59	50	ICMP destination unreachable	3 reconnaissance
22	11:05:00	11:09:59	64	Sensitive data Email address	2 break-in
23	11:10:00	11:14:59	24	Sensitive data Email address	2 break-in
24	11:15:00	11:19:59	13	—	4 good
25	11:20:00	11:24:59	49	Sensitive data Email address	2 break-in
26	11:25:00	11:29:59	1451	Bad traffic loopback	2 break-in
27	11:30:00	11:34:59	27	Web cgi redirect access	2 break-in
28	11:35:00	11:39:59	24	Info telnet access	3 reconnaissance
29	11:40:00	11:44:59	43	TCP windows closed before receive data	2 break-in
30	11:45:00	11:49:59	40	NetBIOS null session	2 break-in
31	11:50:00	11:54:59	73	ICMP ping/reply	3 reconnaissance
32	11:55:00	11:59:59	26	—	4 good
33	12:00:00	12:04:59	68	Consecutive TCP small segments	2 break-in
34	12:05:00	12:09:59	48	—	4 good
35	12:10:00	12:14:59	39	—	4 good
36	12:15:00	12:19:59	28	Web IIS fpcount access	2 break-in
37	12:20:00	12:24:59	99	—	4 good
38	12:25:00	12:29:59	37	—	4 good
39	12:30:00	12:34:59	47	Info telnet login	3 reconnaissance
40	12:35:00	12:39:59	6	—	4 good