文章编号：1007-5321(2017)增-0015-05　　DOI:10.13190/j.jbupt.2017.s.004

基于改进隐马尔可夫模型的复合攻击预测方法

饶志宏[1]，徐　锐[2]，刘　方[2]，杨春亮[1]，方恩博[1]

(1. 中国电子科技集团公司第三十研究所，成都 610041；2. 中国电子科技网络信息安全有限公司，成都 610041)

摘要：提出了一种基于改进的隐马尔可夫模型和维特比算法的复合攻击预测方法. 在训练数据较少时，采用最大似然估计得到的隐马尔可夫模型可能存在较大误差，针对这种情况，采用修正的概率矩阵计算方法以降低误差. 针对告警事件序列中存在误报的情况，在维特比算法中引入了一个判决门限，用于在告警事件存在误报的情况下对预测结果进行修正. 基于 DARPA2000 数据集对提出的方法进行了仿真和实验验证，实验结果表明该方法能有效地提高攻击预测的正确率.

关　键　词：隐马尔可夫模型；复合攻击；维特比算法；攻击意图；告警序列

中图分类号：TP393　　**文献标志码**：A

A Method of Predicting Multi-Step Attacks Based on Improved HMM Model

RAO Zhi-hong[1], XU Rui[2], LIU Fang[2], YANG Chun-liang[1], FANG En-bo[1]

(1. No. 30 Institute of China Electronic Technology Group Corporation, Chengdu 610041, China;
2. China Electronics Technology Cyber Security Compang Limited, Chengdu 610041, China)

Abstract: An approach of predicting multi-step attacks based on improved hidden Markov model (HMM) and Viterbi algorithm was proposed. When the training data was sparse, poor probability estimates of the HMM were obtained by using maximum likelihood estimation. Thus, a modified calculation method of probability matrix was used to reduce error. When there existed false alerts in the alert sequence, a decision threshold was introduced in the Viterbi algorithm for correcting the forecast results. From the simulation and the experimental results based on the DARPA2000 data set, it is concluded that the proposed method can effectively improve the predicting accuracy.

Key words: hidden Markov model; multi-step attacks; Viterbi algorithm; attack intent; alert sequence

目前，对于复合攻击预测的研究是近几年发展起来的一个前沿课题，张松红等[1-3]将隐马尔可夫模型应用于复合攻击预测，通过从报警信息中分析每个状态的可能概率. 这些方法都是在假设入侵检测系统的报警信息正确的情况下对攻击进行预测. 针对入侵检测系统存在误报的情况下，提出了一种新的基于改进隐马尔可夫模型的复合攻击预测方法，并利用 DARPA2000 数据集进行仿真实验验证，证明了提出方法具有较好的预测性能.

1　改进的隐马尔可夫模型

隐马尔可夫模型最早是由 Baum 和 Petrie 在 1966 年提出来的，主要用于描述随机过程统计的概率模型. 近年来，隐马尔可夫模型被广泛应用到网络攻击检测和预测上面[4-6]. 隐马尔可夫模型包含 2 个随机过程：隐状态之间的转移过程和随机输出序

收稿日期：2016-10-31
基金项目：国家高技术研究发展计划(军口 863 计划)项目(2015AA7111006)
作者简介：饶志宏(1970—)，男，高级工程师(研究员级)，博士生，E-mail：charao@tom.com.

列集.

针对复合攻击行为,攻击者使用不同的攻击手段来达到其攻击意图,但其攻击意图往往隐藏于各种复杂的网络攻击行为当中,是不可见的. 现有的入侵检测系统仅能对各种攻击行为产生不同的告警信息,而攻击者的攻击意图却淹没于大量的告警信息中. 复合攻击包括若干个攻击步骤,每一个攻击步骤由该步骤所对应的告警信息折射出来,利用隐马尔可夫模型能够较好地刻画攻击步骤之间及攻击步骤与告警信息间的关系. 通过提取相应的攻击意图后,也就确定了隐马尔可夫模型隐含层中的状态.

已有的研究仅考虑攻击步骤之间的转态转移,以及攻击步骤和告警信息之间的关系,而未考虑告警信息之间的关系. 传统的隐马尔可夫模型没有考虑告警信息之间的转移概率,因此不能完全描述攻击步骤之间、步骤与告警信息之间和告警信息之间的所有关系. 将告警信息之间的转移关系融合到传统隐马尔可夫模型中,并将改进的模型应用到攻击预测中,能将告警信息(即观测值)、攻击意图(即状态)的关系描述得更加清晰,最终使攻击预测的结果更准确. 在传统隐马尔可夫模型中增加了一个告警信息之间的转移概率,改进的隐马尔可夫模型为 $\boldsymbol{\lambda}=(\boldsymbol{O},\boldsymbol{S},\boldsymbol{\pi},\boldsymbol{A},\boldsymbol{B},\boldsymbol{C})$,其中:

$\boldsymbol{S}=\{s_n,1\leqslant n\leqslant N\}$,$s_n$ 为第 n 个状态,即第 n 个攻击意图,N 为隐马尔可夫模型状态个数,也即对应的攻击意图数;

$\boldsymbol{O}=\{O_m,1\leqslant m\leqslant M\}$,$O_m$ 为第 m 个观测值,即第 m 个告警信息,M 为隐马尔可夫模型观测值的个数,即告警信息总的个数;

$\boldsymbol{\pi}=\{\pi_n,1\leqslant n\leqslant N\}$表示初始状态概率矩阵,$\pi_n=p(q_0=s_n)$为,即初始时刻模型处于第 n 个状态 s_n 的概率;

$\boldsymbol{A}=\{a_{ij},1\leqslant i,j\leqslant N\}$表示状态转移概率矩阵,$a_{ij}=p[q_{t+1}=s_j\mid q_t=s_i]$,即由 t 时刻状态 s_i 转移到 $t+1$时刻状态 s_j 的概率;

$\boldsymbol{B}=\{b_n^t(O_m),1\leqslant n\leqslant N,1\leqslant m\leqslant M\}$表示观测概率分布矩阵,$b_n^t(O_m)=p[x_m=O_m\mid q_t=s_n]$表示 t 时刻在状态 s_n 下输出观测值 O_m 的概率,即复合攻击中攻击意图为 s_n 时产生告警信息 O_m 的概率,M 为观测值的个数;

$\boldsymbol{C}=\{c_{ij},1\leqslant i\leqslant M,1\leqslant j\leqslant M\}$,表示观测值之间的转移概率矩阵,$c_{ij}=p[x_{t+1}=O_j\mid x_t=O_i]$,即在 t 时刻产生告警信息为 O_i 的条件下,$t+1$ 时刻产生告警信息 O_j 的概率.

针对隐马尔可夫模型参数的确定,已有的计算方法是采用最大似然概率,利用训练数据通过统计计算得到,但在实际应用中会受限于训练数据的规模,在告警信息集中未出现的告警信息并不代表其真实概率等于0,而出现数量相同的告警信息概率并不一定一样. 针对该问题,对最大似然概率进行修正以应对训练数据较少的情况,通过从非零告警信息产生概率和告警信息转移概率中抽取小部分概率平均分配到其他本该是非零的零概率上,使零概率增加,非零概率降低,以提高模型的准确率. 采用如下修正的概率矩阵计算方法,实现对隐马尔可夫模型的修正.

$$p[x_m=O_m\mid q_t=s_n]=\begin{cases}p[O_m\mid s_n]_c-p_c, & \text{若 } p[O_m\mid s_n]_c>0\\ vp_c/(N_m-v), & \text{其他}\end{cases}\tag{1}$$

$$p[x_{t+1}=O_j\mid x_t=O_i]=\begin{cases}p[O_j\mid O_i]_d-p_d, & \text{若 } p[O_j\mid O_i]_d>0\\ up_d/(N_m-u), & \text{其他}\end{cases}\tag{2}$$

其中:$p[O_m\mid s_n]_c$ 表示利用最大似然概率计算得到的在状态 s_n 下产生告警信息 O_m 的概率;$p_c=1/(T_s+v)$,v 是 s_n 状态下对应非零告警信息产生概率的输出告警信息种类数,T_s 是状态 s_n 输出的告警信息总条数;N_m 为总的告警信息种类数. $p[O_j\mid O_i]_d$ 表示利用最大似然概率计算得到的 t 时刻产生告警信息为 O_i 的条件下,$t+1$ 时刻产生告警信息 O_j 的概率;$p_d=1/(T_i+u)$,u 是 t 时刻产生告警信息为 O_i 的条件下,$t+1$ 时刻可能产生的告警信息种类数,T_i 表示 t 时刻产生告警信息为 O_i 的条件下,$t+1$ 时刻可能产生的告警信息总条数.

2 改进的维特比预测算法

攻击者在发动复合攻击时,每个攻击步骤都可能采用不同的攻击手段,而每种攻击手段可能有不同的攻击意图. 相应地,相同的攻击步骤可能产生不同的告警信息,不同的攻击步骤可能产生相同的告警信息. 如何从大量的告警信息序列中找到攻击者的攻击意图并预测攻击者的下一步攻击意图是目前的研究难点. 在复合攻击还未完成时,将这些复

合攻击已经完成的攻击步骤所对应的告警信息形成告警信息序列$\{O_1, O_2, \cdots, O_T\}$，代入事先已建立好的复合攻击隐马尔可夫模型来推测该告警信息序列$\{O_1, O_2, \cdots, O_T\}$隐藏的攻击意图序列$\{S_1, S_2, \cdots, S_T\}$以及预测攻击者下一步可能的攻击意图.

张等[1-2]使用经典的维特比算法求出隐含的攻击意图序列，但都在假设告警信息序列未出现误报时进行攻击意图预测. 而在实际应用中，入侵检测系统常会出现误报的情况. 当给出的告警信息序列出现误报时，因对应于该误报的告警信息的观测值转移概率和告警信息产生概率可能很低甚至为零，使得后续计算得到的攻击意图序列产生偏差，从而得到错误的攻击意图序列. 针对该问题，采用改进的维特比算法，在每一个攻击状态转移过程中引入了一个阈值，当产生告警信息序列$\boldsymbol{O}=\{O_1, O_2, \cdots, O_T\}$且$t$时刻处于攻击意图$s_j$的概率低于对应的阈值时，则判定此时告警信息存在误报，同时丢弃该告警信息序列，等待下一个告警信息的到来. 通过设置合适的阈值，在预测复合攻击步骤时，可用于判断当前告警信息是否为误报，阈值的选取通过训练数据计算得到. 改进的维特比算法如下.

输入：告警信息序列$\boldsymbol{O}=\{O_1,O_2,\cdots,O_T\}$，改进的隐马尔可夫模型$\boldsymbol{\lambda}=(\boldsymbol{O},\boldsymbol{S},\boldsymbol{\pi},\boldsymbol{A},\boldsymbol{B},\boldsymbol{C})$，阈值$\boldsymbol{V}=\{v_1,v_2,\cdots,v_{N-1}\}$

输出：攻击意图序列$\boldsymbol{Q}^*=\{q_1^*,q_2^*,\cdots,q_T^*\}$

第 1 步　初始化

$$\delta_1(i)=\pi_i b_i^1(O_1),1\leqslant i\leqslant N;$$

$$\theta_1(i)=0,1\leqslant i\leqslant N$$

第 2 步　递归计算

产生告警信息序列$\boldsymbol{O}=\{O_1,O_2,\cdots,O_t\}$且$t$时刻处于攻击意图$s_j$的概率：

$$\delta_t(j)=[\max_{1\leqslant i\leqslant N}[\delta_{t-1}(i)a_{ij}]c_{t-1,t}]b_j^t(O_t),$$

$$2\leqslant t\leqslant T,\quad 1\leqslant j\leqslant N,$$

$$\theta_t(i)=\operatorname{argmax}_{1\leqslant i\leqslant N}\{\delta_{t-1}(i)a_{ij}\}$$

$$\lambda_{\max}=\operatorname{argmax}_{1\leqslant j\leqslant N}\{\delta_t(j)\}$$

如果$\lambda_{\max}<V_{t-1}$且$t<T$，那么转向第 3 步.

如果$\lambda_{\max}<V_{t-1}$且$t=T$，则

$$P^*=\max_{1\leqslant i\leqslant N}\{\delta_{T-1}(i)\}$$

$$q_T^*=\operatorname{argmax}_{1\leqslant i\leqslant N}\{\delta_{T-1}(i)\}$$

否则转向第 4 步.

第 3 步　计算最佳意图序列

$$\delta_t(j)=[\max_{1\leqslant i\leqslant N}[\delta_{t-1}(i)a_{ij}]c_{t-1,t+1}]b_j^{t+1}(O_{t+1}),$$

$$2\leqslant t\leqslant T-1,1\leqslant j\leqslant N$$

$$\theta_t(i)=\operatorname{argmax}_{1\leqslant i\leqslant N}\{\delta_{t-1}(i)a_{ij}\}$$

第 4 步　结束

$$P^*=\max_{1\leqslant i\leqslant N}\{\delta_T(i)\}$$

$$q_T^*=\operatorname{argmax}_{1\leqslant i\leqslant N}\{\delta_T(i)\}$$

攻击预测结果：

$$q_t^*=\theta_{t+1}(q_{t+1}^*),t=T-1,T-2,\cdots,1$$

3　实验和结果

为验证基于改进隐马尔可夫模型的复合攻击预测方法的性能，采用美国麻省理工学院林肯实验室 DARPA2000 提供的攻击场景测试数据集 LLDOS1.0 (inside)作为告警信息来源. LLDOS1.0 中包含了一个复合攻击 DDoS 的攻击过程，完整的攻击序列分为 5 个阶段：1)通过 IPSweep 进行活动主机探测；2)使用 Sadmind Ping 进行 Sadmind Daemon 服务端口扫描，探测可能存在 Sadmind 漏洞的主机；3)利用主机上的 Sadmind 漏洞进行系统入侵，获得活动主机的 root 权限；4)在被攻破的主机上安装用于 DDoS 攻击的木马软件；5)利用被控主机对目标发起 DDoS 攻击. 利用 LLDOS1.0 数据集，建立了一个针对 DDoS 攻击的隐马尔可夫模型. 表 1 给出了告警信息类型及其对应的攻击意图.

表 1　告警信息及其对应意图

告警信息编号	告警信息	对应状态序列(攻击意图)
A	ICMP PING NMAP	IPSweep(S_1)
B	ICMP Echo Reply	IPSweep
C	RPC portmap request sadmind UDP	SadmindPing(S_2)
D	RPC sadmind UDP PING	SadmindPing
E	RPC sadmind UDP NETMGT_PROC_SERVICE CLIENT_DOMAIN overflow attempt	SadmindExploir(S_3)
F	RPC sadmind query with root credentials attempt UDP	DaemonInstalled(S_4)

选取针对主机 172.16.112.50 的 DDoS 攻击告警信息序列作为实验数据，构建的模型参数如表 2

~表6所示.

表2 观察概率分布矩阵 *B*

	A	B	C	D	E	F
S_1	0.498 1	0.498 1	0.000 9	0.000 9	0.000 9	0.000 9
S_2	0.000 9	0.000 9	0.748 1	0.248 1	0.000 9	0.000 9
S_3	0.000 7	0.000 7	0.000 7	0.000 7	0.996 3	0.000 7
S_4	0.000 7	0.000 7	0.000 7	0.000 7	0.000 7	0.996 3

表3 状态转移概率矩阵 *A*

	S_1	S_2	S_3	S_4
S_1	0.5	0.5	0	0
S_2	0	0.5	0.5	0
S_3	0	0	0	1
S_4	0	1	0	0

表4 初始状态概率矩阵 π

S_1	S_2	S_3	S_4
0.2	0.4	0.2	0.2

表5 告警信息转移概率 *C*

	A	B	C	D	E	F
A	0.003 2	0.983 6	0.003 2	0.003 2	0.003 2	0.003 2
B	0.003 2	0.003 2	0.983 6	0.003 2	0.003 2	0.003 2
C	0.004 0	0.004 0	0.004 0	0.321 9	0.661 9	0.004 0
D	0.003 2	0.003 2	0.983 6	0.003 2	0.003 2	0.003 2
E	0.003 2	0.003 2	0.003 2	0.003 2	0.003 2	0.983 6
F	0.003 2	0.003 2	0.983 6	0.003 2	0.003 2	0.003 2

在攻击意图序列识别过程中,存在一个阈值的设置问题,通过对大量训练数据进行分析计算,阈值 v_1、v_2、v_3 和 v_4 分别选取输入为1~4个正确告警事件序列时所计算得到的最小概率值,即 $\{\min_j\delta_1(j), \min_j\delta_2(j), \min_j\delta_3(j), \min_j\delta_4(j)\}\alpha$,输入的告警事件序列组合包括了攻击者所有可能采用的攻击手段,其中 α 为调节因子,取值为0~1,根据具体情况进行设置,取 α 值为0.9,计算得到的阈值如表6所示.

表6 阈值

v_1	v_2	v_3	v_4
0.05	0.009	0.000 04	0.000 01

当收到报警信息序列为{A, B, C}时,可根据隐马尔可夫模型中的维特比算法识别攻击意图序列:现已经完成的攻击意图序列为:$\{S_1, S_2\}$,即IP-Sweep、SadmingPing已经完成,下面要进行的意图为SadmindExploir,实现了对攻击意图的预测.

将改进的算法与文献[1-2]的方法做预测性能比较,在本研究实现的隐马尔可夫模型中采用了式(1)和式(2)中修正后的后验概率计算公式. 通过试验验证,在告警信息序列正确的时候,文献[1-2]中的方法与改进的维特比算法均能正确输出攻击意图序列;而在出现告警信息误报时,文献[1-2]中的方法无法正确识别攻击意图,而改进的维特比算法能够将误警信息进行标记,并给出正确的攻击意图序列.

为验证提出的方法在告警信息存在误报的情况下对攻击意图序列的识别和预测性能,在测试集中随机引入5%、15%、20%和25%的误警信息序列. 图1所示为不同的预测方法在数据集的测试集上得到的攻击意图序列识别准确率.

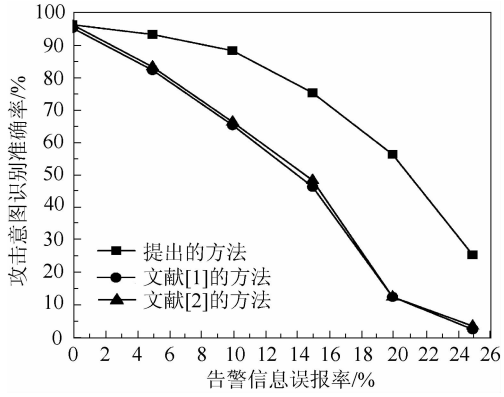

图1 不同方法的攻击意图识别准确率

从图1中可以看出,用改进的方法比文献[1-2]中的告警信息序列正确时识别准确率相差不大,而随着告警信息序列误报率增大时,改进的隐马尔可夫模型的攻击意图序列识别准确率相较文献[1-2]的方法性能下降较缓,即随着告警信息序列误报率增大时,改进的方法在攻击意图序列识别准确率明显优于已有的方法.

4 结束语

在深入分析复合攻击特征和步骤的基础上,提出了一种基于改进隐马尔可夫模型的复合攻击预测方法. 该方法针对训练样本较少的情况下,模型参数确定可能存在较大误差,提出了对最大似然概率进行修正的方法以提高模型的准确率. 同时,针对

入侵检测系统告警事件序列的观测值存在误报的情况，在改进的预测算法中引入了一个判决门限用于剔除告警事件存在误报的情况. 通过实验仿真，在告警信息序列存在误报的情况下，提出的改进隐马尔可夫模型相较已有方法具有较好的性能.

参考文献：

[1] 张松红，王亚弟，韩继红. 基于隐马尔可夫模型的复合攻击预测方法[J]. 计算机工程，2008，34(6)：131-133.
Zhang Songhong，Wang Yadi，Han Jihong. Approach to forecasting multi-step attack based on HMM[J]. Computer Engineering，2008，34(6)：131-133.

[2] Zhang Yanxue，Zhao Dangmei，Liu Jinxing. Approach to forecasting multi-step attack based on fuzzy hidden Markov model[J]. Journal of Applied Sciences，2013，13(22)：4955-4960.

[3] 张艳雪，赵冬梅，刘金星. 基于模糊-隐马尔可夫模型的复合式攻击预测方法[J]. 电光与控制，2015，22(1)：39-44.
Zhang Yanxue，Zhao Daogmei，Liu Jinxing. Approach to forecasting multi-stage attack based on fuzzy hidden Markov model[J]. Electronics Optics & Control，2015，22(1)：39-44.

[4] 杨晓峰，孙明明，胡雪蕾. 基于改进隐马尔可夫模型的网络攻击检测方法[J]. 通信学报，2010，31(3)：95-101.
Yang Xiaofeng，Sun Mingming，Hu Xuelei，et al. Improved HMM model based method for detecting cyber attacks[J]. Journal on Communications，2010，31(3)：95-101.

[5] 周东清，张海锋，张绍武等. 基于 HMM 的分布式拒绝服务攻击检测方法[J]. 计算机研究与发展，2005，42(9)：1594-1599.
Zhou Dongqing，Zhang Haifeng，Zhang Shaowu，et al. A DDos attack detection method based on hidden Markov Model[J]. Journal of Computer Research and Development，2005，42(9)：1594-1599.

[6] Yolacan E N，Dy JG，Kaeli D R. System call anomaly detection using multi-HMMs[C]// 2014 Eighth International Conference on Software Security and Reliability-Companion.[S.l.]：IEEE，2014：25-30.