2009 International Conference on Multimedia Information Networking and Security

A Hidden Markov Model based framework for tracking and predicting of attack intention

Xin Zan, Feng Gao
Department of Automation
Xi'an Jiaotong University
Xi'an, P.R.China
{zanxin, fengg @mail.xjtu.edu.cn}

Jiuqiang Han, Yu Sun
Department of Automation
Xi'an Jiaotong University
Xi'an, P.R.China
{jqhan, ysun@mail.xjtu.edu.cn}

***Abstract*—Recently, several approaches for intrusion correlation and attack scenario analysis have been proposed. However, these approaches all focus on the flooding alert reduction or high-level alert correlation. In this paper, we study the problem of tracking and predicting attack intentions. We use hidden markov models to represent the typical attack scenarios and design a complete framework named HMM-AIP composed of online tracking and prediction module and offline model training module. A novel and effective tracking and predicting attack intention algorithm is presented. We perform experiments to validate our algorithm and the results show that our approach can identify false alert and give the creditable prediction result when the alert observation sequence fits the typical attack scenarios nicely.**

Keywords-HMM; Intrusion detection; Intrusion alert correlation; attack intention prediction

I. INTRODUCTION

Intrusion alert correlation has gradually become an active research field of Intrusion detection technique. Alert correlation allows the administrator to aggregate the outputs of multiples, filter out spurious or duplicate alerts, and provide a succinct, high-level view of the security state of the protected network.

However, some attack processes are executed in a very short time. Even though IDSs could detect these malicious activities afterwards, the damages to the information systems have been generated and could hardly be restored in some cases. So it's necessary to develop algorithms and tools for security administrator to track and predict attack steps in advance so that administrators can take an appropriate response to decrease the damages. However, the accurate prediction of attack behavior is almost impossible and meaningless for thousands of different attack types. However, the different attack activities implied the same attack intention, i.e. the same malicious goal of hackers. If we can predict the next attack intention of hackers in advance, then we could have more time to take responses.

In this paper, we propose a novel framework named Hidden Markov Model for Attack Intention Prediction (HMM-AIP) to track and predict the attack intention based on the current attack alerts.

The remainder of this paper is organized as follows. Section 2 briefly reviews related work for intrusion alert correlation. Section 3 introduces the basic theory and applications of Hidden Markov Model (HMM). Section 4 describes our correlation framework and tracking and predicting algorithm for attack intention. We report our experiment and results in Section 5. Section 6 concludes this paper and discusses some future research directions.

II. RELATED WORK

Recently, several approaches have been proposed to correlate security alerts and analyze the attack scenarios. Existing correlation approaches generally can be categorized into three types. The first type of correlation approaches is based on the similarity of alerts [1] to aggregate the large number of similar alerts to a few high-level comprehensive alerts. In this case, some effective clustering algorithms [2,3] are successfully introduced to reduce the number of flooding alerts. The key problem of these approaches is that how to reasonably define the distance of alerts or alert properties to distinguish alerts. The second type of correlation approach is based on the causality analysis between attack alerts. The assumption is that when an attacker launches an attack, prior attack steps are preparing for later ones. So the correlation engine searches for alert pairs that have a consequence and prerequisite matching. Peng [4] and Cuppens[5] respectively present their own correlation approach in the almost same time. The challenge is that how to build the causality relationships for thousands of attack alerts. The third type of correlation approach is based on the known attack scenario pattern. Goldman et al [6] built a correlation system based on Bayesian reasoning. Porras et al [7] designed a "mission-impact-based" correlation system which focuses on the impact analysis based on the mission goals of protected networks.

III. HIDDEN MARKOV MODEL THEORY

Hidden Markov model (HMM) is a statistical model in which the system being modeled is assumed to be a Markov process with unobserved state. Therefore, there are two stochastic processes in HMM: the process of moving between states and the process of emitting an output sequence. The sequence of state transitions is a hidden

978-0-7695-3843-3/09 $26.00 © 2009 IEEE
DOI 10.1109/MINES.2009.277

process and is observed through the sequence of emitted symbols.

Let us formalize the definition of an HMM taken from an HMM tutorial by Lawrence Rabiner [8]. A HMM can be characterized by a five tuple $\{S,X,A,B,\pi\}$, usually simplified by a triplet: $\lambda=(A,B,\pi)$. S represents a set of finite states S_i, $1 \le S_i \le N$ and X represents a set of finite observations denoted as $X=\{x_1,x_2,\ldots,x_m\}$. The state transition probabilities $A = a_{ij}$ where a_{ij} is the probability of moving from state i to state j.

$$a_{ij}=P(q_{t+1} = S_j \mid q_t = S_i) \quad 1 \le i, j \le N.$$

The emission probability distributions for observation symbols in state j, $B=\{b_j(k)\}$, where

$$b_j(k)=P(x_k \mid q_t = S_j); \quad 1 \le j \le N, \ 1 \le k \le M$$

The initial state distributions

$$\pi_i = P(q_1 = S_i), \quad 1 \le i \le N.$$

Hidden Markov models are especially known for their application in temporal pattern recognition such as speech, handwriting, gesture recognition, part-of-speech tagging, musical score following, partial discharges and bioinformatics because of their great adaptability and versatility in handling sequential signals.

IV. SYSTEM DESIGN FOR HMM-AIP

In this section, we design a novel framework named Hidden Markov Model for Attack Intention Prediction (HMM-AIP) for model training and attack scenario identification. Furthermore, we present an effective tracking and predicting algorithm for attack intention.

A. Basic Concepts

Some related terms and concepts for attack description have been given different meanings by researchers in various papers. To avoid the confusion, some primary attack terms are informally defined and described as follows.

Definition *Attack Scenario* is the typical attack procedures usually performed by attackers to gain their specific attack intentions under the conditions of given attack and network environment. In general, a typical attack scenario consists of several relative attack steps. In this paper, we define six kinds of typical attack scenarios with our security experience, i.e. AS = {remote_ privilege_ escalating, worm, botnet, phishing, web attack scenario, ddos}.

Definition *Attack Target* is the system objects or computer resources that hacker expects to control or occupy illegally. In this paper, we classify attack target as four types, i.e. AT = {information, privilege, resource, data}.

Definition *Attack Intentio*n is also meant "attack goal" which the attacker attempts to reach by executing a series of attack activities. Therefore the attack activities are completely determined by the internal attack intention. On the other hand, the activities of attackers could be monitored and detected by IDSs, but the attacker's intention couldn't be directly observed by detector. In this paper, we overview the seven kinds of common attack intentions: i.e. AI={GetSysinfo, GetVulinfo, PrivilegeEscalation, AccessSysObject, IllegalExecution, RemoteCotrol, AccessData}.

B. HMM models for Typical Attack Scenario

In the HMM-AIP, the observation sequence is the alert stream reported by multiple IDSs. The state variable of HMM-AIP is the real attack intention hidden behind the attack behaviors which can't be observed directly by IDS sensors.

There are some basic assumptions made in HMM-AIP. First, attack activities observed by IDSs only rely on the current hidden attack intention. This assumption is reasonable since the hacker's activity is completely determined by attacker's goals. In this paper, the attack activity is equal to the alert of reported by IDS.

Second, the next attack intention is only driven by current attack intention, i.e. the attack intention meets the markov property. From our practical security experience, this is creditable because the prior attack step provides the necessary information and rights for next attack step.

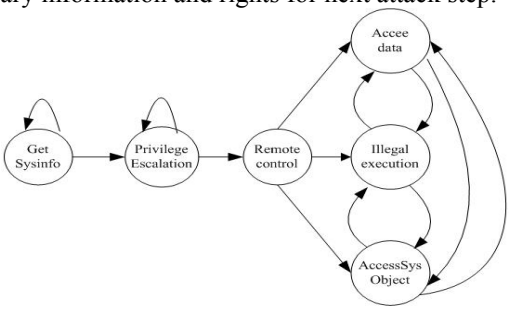

Figure 1. HMM for remote privilege escalation attack scenario

In HMM-AIP, we build HMM for each typical attack scenario. A typical attack scenario could be represented as a series of the transition of attack intentions. The figure 1 shows the remote privilege escalation scenario which describes the most common network penetration procedure.

C. System Architecture of HMM-AIP

HMM-AIP consists of two main modules: offline training module and online tracking and predicting module. The overall system architecture of HMM-AIP is described in detail in figure2.

Initially, the raw alert stream reported by multiple IDSs is dealt with the preprocess module of alerts called "PMA". PMA module filtered trivial alerts in real time and identify apparent false positive alerts.

In offline training module, the historical alert data is used to build the hidden markov models for the typical attack scenario. The model was estimated by the Baum-Welch estimation procedure, which is the standard method for maximum likelihood estimation of HMMs. Baum-Welch Algorithm in essence is an Expectation Maximization (EM) algorithm.

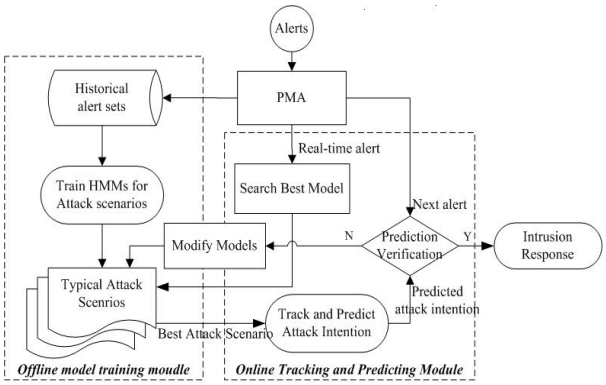

Figure 2. System Architecture of HMM-AIP

D. Tracking and Predicting Module

There are three stages in online tracking and predicting module. In the first stage, we should find the best attack scenario to describe the current alert sequence. From the HMM's point view, this is the evaluation problem, i.e. given an observation sequence and multiple HMMs, compute the probability of the given observation sequence driven by different HMMs. This problem can be solved by the forward algorithm or backward algorithm based on dynamic programming technique. Whenever a new alert reaches, the likelihood of alert sequence is recomputed for all HMMs. If the maximum likelihood of observation sequence $P(O \mid \lambda^*)$ is larger than a given threshold, we can say that the model λ^* fits the current alert sequence quite adequately.

In the second stage, we should find the best state sequence $Q^*=\{q_1,q_2,...,q_k\}$ to maximize the likelihood of given observation sequence and HMM. This is the decoding problem from the HMM's point of view. In general, Viterbi algorithm which is a dynamic programming algorithm similar to the forward algorithm except for using maximization instead of summing at the recursion is used to solve this problem. The likelihood of best state sequence can be calculated by the following formula:

$$\arg\max_{Q}\{P(Q \mid O,\lambda)\} = \arg\max_{Q}\{P(Q,O \mid \lambda)\}$$

Let $\delta_t(i)$ be the maximal probability of state sequences of the length k that ends in state i and produces the k first observations for the given model, denoted by the following formula:

$$\delta_t(i) = \max_{q} P(q_1,q_2,...,q_t = i, o_1,o_2,...o_t \mid \lambda)$$

The Viterbi algorithm is described as follows:

Initialization $\delta_1(i) = \pi_i b_i(o_1)$

$$\psi_1(i) = 0$$

Recursion $\delta_{t+1}(j) = \max_{1\le i\le N}[\delta_t(i)a_{ij}]b_j(o_t)$

$$\psi_t(j) = \arg\max_{1\le i\le N}(\delta_{t-1}(i)a_{ij})$$

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

$$q_T^* = \arg\max_{1\le i\le N}\delta_T(i)$$

In the third stage, given the current alert sequence $X=\{x_1,x_2,...,x_k\}$, the tracking and predicting algorithm is described as follows:

step1: compute the best state sequence $\{q_1,q_2,...,q_k\}$ and specify the best state i at the time k for current observation sequence using Viterbi algorithm.

step2: if the maximum state transition probability a_{ij} is larger than 0.5 then we could predict the next state as j at the time $k+1$ or else return to step 1 to continue to track. So the principle of our prediction for the next attack intention is

$$q_{k+1}(j) = \underset{j}{\text{argmax}}(a_{ij})$$

This means that the most likelihood state transition implies the maximum probability of next state to appear with knowing the current state.

step3: When the next observation $x(k+1)$ reaches, compute the emission observation probability $b_j(x_{k+1})$ which emits from the predicted state j. If the value $a_{ij}b_j(x_{k+1})$ is larger than a given threshold, then we can draw the conclusion that prediction result is fairly reasonable and the corresponding security response should be taken, or else we couldn't make sure that whether the reason of prediction failure is the inaccurate parameters of HMM or the false alert observation.

step4: As the information at a single time isn't enough to find the reason of prediction failure. So we should use the observation at the time k+2 to verify the prediction result. A basic assumption holds in this case, it's impossible that both two continuous alerts are false alerts in most case. So we could assume that the next alert $x(k+2)$ is always a real attack alert while the alert at the time $k+1$ is uncertain. Figure 3 shows the state and observation transition procedure to identify the false observation at the time k+1. Suppose the corresponding best state j' is calculated by Viterbi algorithm for the given observation $x(k+1)$. If the following inequation is met

$$a_{ij}b_j(x_{k+2}) \;>\; a_{ij'}b_{j'}(x_{k+1})$$

then we can say that the observation $x(k+1)$ is a false alert. Consider if the x(k+1) is a false observation then the calculated state j' is also a false state. Thus the real observation $x(k+2)$ should be emitted by the predicted state j and the likelihood of true state emission should be larger than the false emission transition. On the other hand, if the inequation isn't met that means the x(k+1) is a real observation and the predict result is incorrect for the inaccurate parameter of HMM. So the results of prediction and real alerts need to be modified by model modification module.

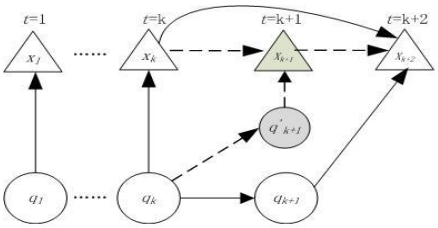

Figure 3. State and observation Graph for identification for false alert in HMM-AIP

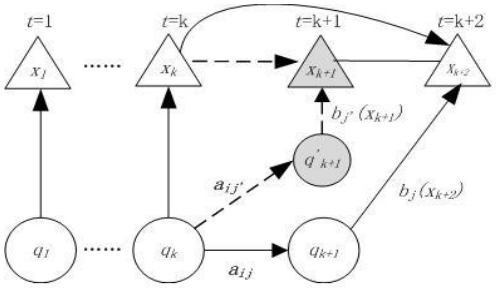

Figure 4. alerts stream of DARPA2000 evaluation dataset

V. EXPERIMENTS

To evaluate the effectiveness of our tracking and predicting algorithm, we use the DARPA 2000 intrusion detection evaluation dataset. In this dataset, a complete attack procedure was performed which contained a series of attacks: probing, break-in, installing DDoS daemon and launching DDoS attacks. Figure 4 shows the alert observation sequence reported by IDS. In the figure4, the elliptical node O_i represents an attack alert produced at the time i and the circular node is the corresponding attack intention. The gray node O_7 represents a false alert named *"Sadming_Ping"* which is deliberately inserted by us to verify our prediction algorithm and the gray node q' is the corresponding nonexistence state calculated by Viterbi algorithm. The attack procedure described in figure 4 is a good example of typical attack scenario "remote privilege escalation" showed in figure 3.

The tracking and predicting procedure in figure 4 is described as follows: the alert observation O_2 to O_5 correspond to the four kinds of variants of the attack "Sadmind_Amslverify_Overflow". When the alert O_6 reaches, the probability of alert sequence $O=\{O_1,O_2,\ldots,O_6\}$ is equal to 0.42 and larger than the threshold 0.35. So we can confirm that the *remote_privilege_escalation* attack scenario fits the alert sequence quite nicely. Meanwhile, the corresponding best state sequence $Q=\{q_1,q_2,q_2, q_2, q_2, q_3\}$ is calculated using Viterbi algorithm. As the biggest state transition probability a_{34} is 0.65 and larger than the threshold 0.5, we could give the prediction of state as q_4 at the time 7. However, the alert observation O_7 at the time 7 is the "*Sadming_Ping*" and the emission probability from state q_4 to observation O_7 is very low. So we need to compare the likelihood from the predicted state to the observation O_8 at the time 8 with the calculated state to the uncertain observation O_7. The results of formula are $a_{34}*b_4(O_8)$=0.46 and $a_{34'}*b_{4'}(O_7)$=0.23. This means the likelihood of from predicted state emitted to alert observation O_8 is larger than from the calculated state q_4' to uncertain alert observation O_7. So we can prove that the alert observation O_7 is a false alert and the prediction result of our algorithm is creditable.

VI. CONCLUSIONS

The tracking and predicting for attack intention is a very challenging problem. In this paper, a novel tracking and predicting algorithm is proposed to predict the next attack intention using HMM method. In future, we will continue to study the tracking and predicting algorithm for prediction of attack intention. We will improve the function and procedure of model modification.

REFERENCES

[1] A.Valdes and K. Skinner. "Probabilistic Alert Correlation", In Proceedings of the International Symposium on Recent Advances in Intrusion Detection, Davis, CA, October 2001,pp. 54–68.

[2] K.Julisch, "Clustering Intrusion Detection Alarms to Support Root Cause Analysis", in ACM Transactions on Information and System Security 6(4), November 2003.

[3] Tadeusz Pietraszek," Using Adaptive Alert Classification to Reduce False Positives in Intrusion Detection",In Recent Advances in Intrusion Detection (RAID2004),Sophia Antipolis, France,May 2004, pp.102-124.

[4] N.Peng, Y.Cui, and D.S.Reeves,"Constructing Attack Scenarios through Correlation of Intrusion Alerts". In 9th ACM Conference on Computer and Communications Security.Washington, D.C. November 2002, pp.245–254.

[5] F.Cuppens, and A.Miege, "Alert Correlation in a Cooperative Intrusion Detection Framework", In Proceedings of the 2002 IEEE Symposium on Security and Privacy, Oakland, CA.May 2002,pp.202-215.

[6] R.P.Goldman,W.Heimerdinger,and S.A.Harp," Information modleing for intrusion report aggregation. In DARPA Information Survivability Conference and Exposition (DISCEX II), June 2001.

[7] Phillip A. Porras, Martin W. Fong, Alfonso Valdes, "A Mission-Impact-Based Approach to INFOSEC Alarm Correlation," in Proceedings of the International Symposium on Recent Advances in Intrusion Detection, October 2002.

[8] Rabiner,L.R, a tutorial on Hidden Markov Models and selected applications in speech recognition, Proc. IEEE 77, Jan.1989 pp. 257–286.