은닉 마르코프 모델을 이용한 네트워크 침입 탐지 시스템

이종석[O] 한상준 박찬호 조성배
연세대학교 컴퓨터과학과
{jslee[O], sjhan, cpark}@sclab.yonsei.ac.kr sbcho@cs.yonsei.ac.kr

Network Intrusion Detection System Using Hidden Markov Model

Jong-Seok Lee[O], Sang-Jun Han, Chanho Park, Sung-Bae Cho
Dept. of Computer Science, Yonsei University

요 약

최근 몇 년간 일어난 전산 네트워크의 폭발적인 확산은 전산 시스템에 대한 침입과 피해 또한 증가하는 부작용을 낳았다. 그에 따른 대책 방안으로 침입 탐지 시스템에 대한 관심과 연구가 증가하고 있다. 본 논문에서는 네트워크 상에서 이동하는 정보를 수집하여 HMM으로 모델링한 후, 외부 또는 내부 네트워크에서의 비정상적인 행위를 탐지하는 침입 탐지 시스템을 제안한다. 전처리를 거친 네트워크 패킷 시퀀스들은 forward-backward 절차와 Baum-Welch 재평가식을 이용하여 정상행위로 모델링된다. 이렇게 구축된 모델을 사용하여 forward 절차를 통해 판정하려는 시퀀스가 정상 행위에서 생성되었을 확률을 계산하며 이 값을 임계값과 비교하여 정상행위 여부를 판별한다. 실험 결과 제안한 침입 탐지시스템이 다양한 침입을 적절히 탐지하는 것을 확인할 수 있었다.

1. 서론

최근 급속한 정보통신 기반구조의 확산에 대한 부작용으로 정보보안에 대한 문제가 해를 거듭할수록 증가하고 있다. 실제로 한국정보보호진흥원의 조사에 의하면 2003년에만 13,184건의 해킹사고가 발생하였고, 2004년 7월까지 9,123건의 해킹사고가 발생하였다. 특히 얼마 전 중국에서 우리나라의 중요기관의 서버를 해킹하여 국가 기밀이 유출되는 일이 발생한 사례에서도 확인할 수 있듯이 최근 금융망이나 국방망, 전력망 등에 침입이 증가하고 있어, 불법적인 침입을 사전에 탐지하여 국가 혼란 상황을 야기할 수 있는 피해를 차단할 필요가 있다[1, 2].

침입탐지시스템은 정상적이지 않은 경로를 통한 침입을 탐지해 내는 것으로 단일 컴퓨터는 물론 여러 컴퓨터가 연결된 네트워크를 감독할 수 있다[3, 4]. 이러한 시스템은 감사기록이나 네트워크 패킷기록 등의 자료로부터 사용자가 어떠한 행위를 하였는지 분석하는 작업을 한다. 현재 상용화되어 있는 대부분의 침입탐지시스템은 오용탐지 기법을 사용하고 있는데, 이러한 기법은 알려진 공격에 대한 정보를 이용하여 사용자, 시스템, 프로그램의 행동이 공격 패턴과 일치하는지 여부를 판단한다. 공격패턴 정보를 가지고 있으므로 정상행위를 공격행위로 간주하는 오류가 적고 공격패턴만 검사하면 되므로 시스템 구축비용이 적게 드는 반면, 공격에 대한 정보를 계속 갱신해야 하고 알려지지 않은 새로운 공격은 탐지할 수가 없는 단점이 있다.

반면 본 논문에서 사용할 비정상행위 탐지 기법은 정상행위를 모델링 한 후 정상행위에 어긋나는 행위를 탐지해 내는 기법으로, 알려지지 않은 유형의 침입을 탐지할 수 있다는 장점이 있다. 본 논문의 나머지 순서는 다음과 같다. 2장에서는 이벤트 시퀀스의 모델링을 이용한 침입탐지 방법을 알아보고 3장에서는 은닉 마르코프 모델과 논문에서 제안한 탐지방법에 대해 소개한다. 4장에서는 제안한 탐지방법을 사용하여 실험한 결과를 분석한 후 5장에서 결론 및 향후 연구에 대하여 언급한다.

2. 관련연구

시간순서의 이벤트 시퀀스를 이용한 방법은 비정상행위기반 침입탐지에 매우 효과적이다. 이벤트 시퀀스를 모델링하는 방법으로는 주로 Markov 가정을 사용하는데 이는 복잡하지 않으면서도 좋은 성능을 얻을 수 있기 때문이다. 이러한 Markov 가정을 활용한 통계적 모델링 방법 중 실제로 내부에서 상태가 어떻게 변화하는지 알 수 없고 출력되는 신호만을 알 수 있는 경우에 사용하는 HMM(Hidden Markov Model)이 많이 사용된다. Warrender 등은 시스템 호출 감사자료를 HMM으로 모델링했을 때 다른 모델링 방법에 비해 좋은 성능을 보여준다고 하였고, Lane은 UNIX 쉘데이터에 HMM을 적용하는 방법을 제안하였다[5]. 이처럼 현재까지 이벤트 시퀀스 모델링을 이용한 방법들은 주로 호스트기반 침입탐지에 많이 사용되었지만 본 논문에서는 네트워크기반 침입탐지에 많이 사용되는 네트워크 패킷 감사자료에 적용하는 침입탐지 기법을 제안한다.

3. HMM을 사용한 침입탐지

3.1 침입탐지시스템 개요

일반적인 비정상행위탐지 방식 침입탐지시스템은 정상행위 감사 자료 수집과 정상행위 모델 구축, 침입을 포함한 감사자료의 테스트 과정으로 이루어진다. 제안하는 탐지 기법의 전체적인 구조는 그림 1과 같다. 전처리 모듈은 네트워크 감

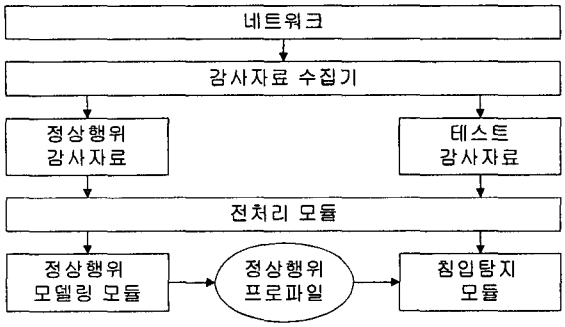

그림 1. 제안하는 탐지 방법의 구조

사 자료 중 네트워크 주소와 TCP 헤더 및 패킷의 내용을 추출하여 정상행위 자료로 변환한다. 모델링 모듈에서는 전처리 과정을 거친 정상행위 자료를 HMM을 이용하여 정상행위 모델로 구축한다. 탐지 모듈에서는 구축된 정상행위 모델에 역시 전처리 모듈을 이용하여 만든 테스트 자료를 비교하여 침입여부를 판단한다.

3.2 전처리

네트워크 감사자료 중 전처리 모듈에서 사용되는 부분은 TCP헤더와 IP헤더의 네트워크 주소 부분과 TCP 헤더의 flag, 패킷의 길이, 포트 번호, 패킷의 contents 이다. 네트워크 주소 부분은 각 연결별로 데이터를 구분하기 위해 사용되는 동시에 패킷의 source가 외부 네트워크인지 내부 네트워크인지 알아내기 위해 사용된다. TCP 헤더의 flag 부분에서는 ACK,PSH,SYN,FIN flag의 값을 추출하여 사용하고, 패킷의 길이는 최소 길이 0과 최대 길이 1500, 그리고 최소 길이와 최대 길이 사이의 길이 각 3가지로 구분하였다. 또한 FTP나 SMTP, TELNET 등의 지정된 포트를 통하였는지의 여부도 사용되었다. 그리고 패킷의 contents에 쉘이나 패스워드 파일에 접근하려는 내용이 있는지 여부도 사용하였다. 이러한 기준을 사용하여 네트워크 감사 자료의 각 패킷에 번호를 할당하여 관찰 시퀀스로 변환한 뒤 정해진 길이의 윈도우 단위로 잘라서 정상행위 자료로 추출하였다. 패킷에 번호를 할당하는 기준은 표 1과 같으며, 각 항목에서 체크된 만큼 합계를 구하여 패킷의 번호를 할당한다.

표 1. 패킷에 번호를 할당하는 기준

· 패킷의 source가 외부 네트워크이면 +1
· 사용된 포트에 FTP,SMTP 등의 지정된 포트가 사용되었을 경우 +2
· 패킷의 길이가 최소 길이일 경우 +4
· 패킷의 길이가 최소 길이와 최대 길이 사이일 경우 +8
· 패킷의 길이가 최대 길이일 경우 +16
· TCP헤더의 ACK flag가 1일 경우 +32
· TCP헤더의 PSH flag가 1일 경우 +64
· TCP헤더의 SYN flag가 1일 경우 +128
· TCP헤더의 FIN flag가 1일 경우 +256
· 패킷의 contents에 쉘이나 패스워드 파일에 접근하는 내용이 있을 경우 +512

3.3 은닉 마르코프 모델

HMM은 상태라고 불리는 N개의 노드와 상태 사이의 전이를 표현하는 edge로 구성된 그래프로 볼 수 있다. 각 상태 노드에는 초기 상태 분포값과 해당 상태에서 특정 심볼을 관측할 확률 분포값이 저장되어 있으며, 각 edge에는 한 상태에서 다른 상태로 전이될 상태 전이 확률 분포값이 저장되어 있다.

$O = O_1, O_2, \cdots, O_T$ 라는 입력 시퀀스가 주어지면 HMM은 비록 외부에서는 그 상태 전이 과정을 직접 확인할 수는 없어도 자체의 확률 매개변수를 이용하여 Markov 과정의 확률 함수로 모델링할 수 있다. 일단 모델링 과정을 통해 모델이 구축되면 임의의 입력 시퀀스가 모델로부터 생성되었을 확률을 알아낼 수도 있다. HMM에서 사용하는 매개변수는 다음과 같으며, 이 중 특정 모델 λ를 규정하는 매개변수를 묶어서 $\lambda=(A, B, \pi)$로 표현한다.

- T : 관찰 시퀀스의 길이
- N : 모델의 상태 수
- M : 관찰 심볼의 수
- $Q = q_1, q_2, \cdots, q_N$: 상태들
- $V = v_1, v_2, \cdots, v_M$: 가능한 관찰 심볼의 이산적 집합
- $A = \{a_{ij}\}, a_{ij}=\Pr(q_j \text{ at } t+1 \mid q_i \text{ at } t)$: 상태전이확률분포
- $B = \{b_j(k)\}, b_j(k)=\Pr(v_k \text{ at } t \mid q_j \text{ at } t)$: 관측심볼확률분포
- $\pi = \{\pi_i\}, \pi_i=\Pr(q_i \text{ at } t=1)$: 초기 상태 확률분포

비정상행위 판정에서는 구축되어 있는 정상행위 모델에 테스트하고자 하는 이벤트 시퀀스를 입력하여 정상행위에서 현재 행위가 생성되었을 확률을 구한다. 이 때 확률을 구하는 방법으로는 forward- backward procedure나 Viterbi 알고리즘을 사용할 수 있다[6].

forward-backward procedure는 forward 변수인 α와 backward 변수인 β를 사용해서 입력 시퀀스 O가 해당 모델 λ로부터 나왔을 확률 $\Pr(O|\lambda)$를 계산한다. forward 변수 α는 시간 t에 부분관찰 시퀀스 $O_1, O_2, \cdots, O_t$를 보고 상태 q_i 에 있을 확률로 다음과 같이 정의된다.

$\alpha_t(i) = \Pr(O_1, O_2, \cdots, O_t, i_t=q_i \mid \lambda)$

이 정의에 따르면 $\alpha_T(i)$는 입력 시퀀스 O 의 모든 심볼을 순서에 맞게 가지고 있으면서 최종 상태가 i인 확률을 나타낸다. $\alpha_T(i)$를 모든 상태 i에 대해 고려하면 $\Pr(O|\lambda)=\Pr(O_1, O_2, \cdots, O_T|\lambda)$를 구할 수 있다. $\alpha_t(i)$는 다음 절차에 의해 귀납적으로 구할 수 있다.

· 단계 1 (초기화) :

$$\alpha_1(i) = \pi_i b_i(O_1)$$

· 단계 2 (귀납) :

$$\alpha_{t+1}(j) = \left[\sum_{i=1}^{N} \alpha_t(i) a_{ij} \right] b_j(O_{t+1}) \quad \text{to } T-1$$

· 단계 3 (종료) :

$$\Pr(O \mid \lambda) = \sum_{i=1}^{N} \alpha_T(i)$$

backward 변수 $\beta_t(i)$는 다음과 같이 정의되며 forward 변수를 구하는 것과 유사한 과정에 의해서 구할 수 있다.

$$\beta_t(i) = \Pr(O_{t+1}, O_{t+2}, \cdots, O_T \mid i_t = q_i, \lambda)$$

정상행위 모델링은 전처리 단계에서 생성된 정상행위 시퀀스를 기반으로 HMM의 매개변수를 결정하는 과정이다. HMM의 매개변수 결정은 주어진 시퀀스 O가 해당 모델 λ로부터 나왔을 확률인 $\Pr(O|\lambda)$값이 최대가 되도록 $\lambda=(A, B, \pi)$를 조정한다. 이를 계산하는 해석적인 방법은 알려져 있지 않으므로, 반복적으로 λ를 결정하는 방법으로 Baum-Welch의 재추정식을 사용하였다[6].

4. 실험 및 결과

4.1 실험 데이터

실험에 사용된 정상행위 감사자료는 MIT Lincoln Lab에서 1999년에 DARPA 프로젝트의 일환으로 수행한 침입탐지시스템 평가프로젝트에서 사용된 자료를 이용하였다[7]. 그 중 침입이 포함되지 않은 첫째 주의 월요일의 네트워크 감사자료를 이용하였으며 크기는 약 340MBytes 이다. 테스트 감사자료는 넷째 주와 다섯째 주의 월요일부터 금요일까지의 네트워크 감사자료 중 일부를 사용하였다. Solaris 8 운영체제 환경에서 tcpdump를 사용하여 네트워크 패킷을 감시하였다. HMM의 모델은 순서 정보를 잘 표현하는 left-to-right 모델을 사용하였다.

4.2 실험 결과

실험 결과 HMM의 상태 수가 10이고 입력 시퀀스의 길이가 10일 때, 100%의 탐지율에서 12.31%의 false-positive 오류율을 보였다. 그림 2는 실험 결과를 보여주는 ROC(Receiver Operating Characteristic) 곡선으로서, 탐지율과 false-positive 오류율의 변화를 나타낸다. ROC 곡선을 통하여 비교적 낮은 false-positive 오류에서 높은 침입탐지율을 확인하여 HMM이 효과적으로 침입을 탐지하고 있는 것을 알 수 있었다.

5. 결론 및 향후 연구

본 논문에서는 네트워크 침입 탐지를 위한 비정상행위 탐지 기법으로 HMM을 사용하는 방법을 제안하였다. TCP 프로토콜의 패킷 시퀀스를 이용하여 정상행위를 모델링하고 여러 가지 유형의 침입 행위를 탐지해보았다. 그 결과 정상행위와 비정상행위가 평가값에 있어서 큰 차이를 보이는 것을 알 수 있었으며 비교적 낮은 false-positive 오류율에서 높은 침입 탐지율을 얻을 수 있었다.

제안한 방법은 비정상행위 기반이므로 대부분의 오용탐지 기반 네트워크 탐지 방법에 비하여 예상하지 못한 프로토콜의 약점을 이용한 공격도 탐지 가능하다는 장점을 가진다. 다만 정상행위를 비정상으로 판정하는 경우가 종종 발생하는데, 이

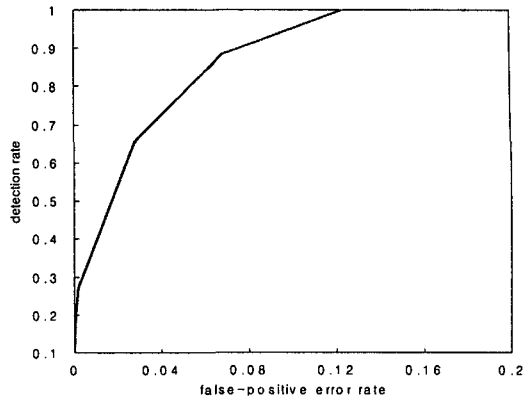

그림 2. 제안한 방법의 ROC 곡선

는 정상행위를 모델링 할 때 비록 소수이지만 학습에 제외된 정상행위 자료가 있기 때문이다. 그러므로 정상행위 모델링에 사용되는 데이터를 포괄적으로 사용하여 HMM을 통한 모델링을 한다면, 침입탐지시스템이 더욱 효과적으로 사용될 수 있을 것이다.

한편, HMM의 상태 수와 입력 시퀀스의 길이 변화에 따른 평가값의 변화도 다양한 형태로 나타날 것으로 보이므로 침입탐지에 최적의 성능을 보일 수 있는 HMM의 매개변수를 결정하기 위한 실험이 필요하다. 더 나아가 실제 침입탐지에 이용되기 위해서는 침입 여부를 판단하는 임계값의 설정에 있어서도 시스템의 상황에 맞도록 시스템 내부에서 적절하게 설정할 수 있는 장치가 마련되어야 할 것이다. 향후 이를 위하여 관련 데이터를 많이 확보하여 다양한 경우를 고려한 실험을 할 예정이다.

참고문헌

[1] 해킹바이러스 통계 및 분석 월보, *한국정보보호진흥원*, 12, 2003.

[2] 해킹바이러스 통계 및 분석 월보, *한국정보보호진흥원*, 7, 2004.

[3] S. Kumar and E. H. Spafford, "An application of pattern matching in intrusion detection," *Technical Report CSD-TR-94-013*, 1994.

[4] T. F. Lunt, "A survey of intrusion detection techniques", *Computer & Security*, vol. 12, no. 4, June, 1993.

[5] C.Warrender, S. Forrest and B. Pearlmutter, "Detecting intrusion using calls: Alternative data models", *In proc. of Symposium on Security and Privacy*, pp. 133-145, May, 1999.

[6] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition", *Proc. of IEEE*, vol. 77, no. 2, pp. 257~286, February, 1989.

[7] MIT Lincoln Labs. *1999 DARPA Intrusion Detection Evaluation*. In http://www.ll.mit.edu/IST/index.html