순서적 이벤트에 기반한 침입탐지시스템의 성능향상을 위한 다중 HMM의 모델 결합

최종호, 조성배
연세대학교 컴퓨터과학과

Combining Multiple HMMs to Improve Intrusion Detection System with Sequential Event

Jongho Choy and Sung-Bae Cho
Computer Science Department, Yonsei University

요 약

침입탐지시스템은 침입탐지 기법에 따라 크게 오용탐지시스템과 비정상행위탐지시스템으로 나뉜다. 비정상행위 탐지시스템은 정상사용행위를 모델링한 후 현재 관찰중인 행위가 정상에서 벗어나는지를 검사한다. 시스템 사용시 발생하는 각 이벤트는 동시에 여러 가지 정보를 담고있으므로 여러 각도에서 모델링될 수 있다. 따라서 여러 결과를 종합해서 판정의 안정성을 높을 수 있다. 본 논문에서는 이벤트의 시스템호출에 평가결과와 BSM감사정보 중 시스템호출관련 정보, 파일접근관련 정보, 이 둘을 모두 고려한 정보를 통합한 평가결과를 투표방식으로 결합하여 판정하는 기법을 제안하였다. 실험결과 두 모델을 별도로 적용하는 경우보다 나아진 판정성능을 보여주었다.

1. 서 론

사회 전 분야의 컴퓨터에 대한 의존도가 높아지고 침입에 의한 컴퓨터 시스템장애에 따른 피해의 빈도와 규모가 급격히 증가하면서 침입탐지에 관한 요구와 관심이 증가되고 있다.

침입탐지 시스템은 불법적인 사용이나 오용, 남용 등에 의한 침입을 알아내는 것으로[2], 탐지기법이 공격행위의 정보를 이용하는지 정상행위의 정보를 이용하는지에 따라서 오용탐지와 비정상행위 탐지로 나눌 수 있다. 오용탐지의 경우는 잘 알려진 침입패턴에 대해서는 높은 탐지 성능을 보이지만 유사한 침입이라도 알려지지 않은 변종패턴일 경우는 탐지할 수 없는 단점이 있다. 반면, 비정상행위 탐지의 경우 잘 알려진 침입에 대해서 오용탐지 만큼의 높은 성능을 보장하지는 않지만, 알려지지 않은 침입패턴에 대해서도 탐지가 가능하다는 장점을 가지고 있다.

침입탐지 시스템은 사용자의 키입력, 시스템호출, 접속당 사용시간, 시스템의 평균부하 등 다양한 관찰심볼을 기반으로 침입여부를 결정한다. 이 일반적인 사용자파일인지 침입탐지문제는 관찰된 사용패턴을 침입행위와 정상행위로 분류하는 문제로 생각할 수 있다. 컴퓨터에서 사용자의 행위시퀀스는 일반적으로 몇 가지 정해진 패턴을 따르며 명령어 수행이나 시스템 호출 등의 이벤트로 관찰된다. 따라서 사용자의 행위시퀀스에 기반해서 사용자의 행위를 설명하고 특성을 파악할 수 있는 모델을 구축할 수 있다면 이를 기반으로 정상행위를 모델링한 후 사용자의 행위가 정상행위인지 판정할 수 있다. 본 논문에서는 음성인식 및 여러 분야에서 알려지지 않은 대상을 모델링하는데 널리 쓰이고 있는 HMM을 사용자의 정상행위 모델링과 비정상행위의 판정에 적용한다.

이때 동일한 시스템호출이 수행되고 있더라도 그 대상에 따라서 정상행위일수도 있고 침입일 수도 있다. 예로 write() 호출이 일반 사용자 파일을 대상으로 수행되고 있다면 정상적인 행위이지만 시스템 파일이 대상이라면 침입행위일 가능성이 많다. 따라서 여러 가지 변량들에서 나오는 정보를 종합적으로 고려할 수 있는 틀이 필요하다. 본 논문에서는 투표방식을 이용해서 각 변량의 평가값을 통합하는 기법을 제안한다.

2. 관련연구

오용탐지기법이 알려진 오용행위만을 탐지할 수 있다는 한계가 있기 때문에 날로 다양화되어 가는 침입기법에 대응하기 위해 비정상행위 탐지기법에 관한 연구가 활발해지고 있다. 이들은 사용자의 사용패턴에 기반하거나 프로그램의 사용패턴에 기반하여 정상행위를 모델링한 후 판정하려는 순서적 이벤트를 여러 기법을 통하여 모델링된 정상행위와 비교해서 정상행위에서 얼마나 벗어나는 행위인지를 측정한다.

[3]은 사용자별로 사용패턴 벡터를 구축한 후 판정하려는 이벤트 시퀀스 벡터와의 유사성을 패턴매칭을 통해 판별하며, [1,5]는 프로그램별 사용패턴을 모델링한다. [1]은 패턴매칭, BP 신경망, Elman 신경망을 사용한 경우를 비교하고 있으며, [5]는 빈도수, 데이터 마이닝기법, HMM의 성능을 비교한다. [1,5]에서 연구결과 이벤트들간의 순서적 정보를 이용한 경우에 다른 경우보다 나은 성능을 보여주었으며 [3]의 경우에도 패턴매칭시 순서적 정보를 중요한 요인으로 고려

하고 있다.

다중 평가값에 바탕을 둔 판정방식으로는 투표방식, 신경망, 전문가시스템 등 다양한 기법들을 사용할 수 있으며, 변량의 중요도를 반영하기 위해서 변량별 가중치를 도입해서 사용하기도 한다. 또한 퍼지추론을 통해 변량별 평가값 및 판정결과를 계량화시키고 어휘적인 의미를 부여할 수도 있다.

3. HMM을 사용한 침입탐지

3.1 침입탐지 시스템 개요

본 논문에서 개발하는 침입탐지시스템은 그림 1과 같이 데이터 필터링과 데이터 축약을 담당하는 전처리 모듈과 정상행위 모델링과 추론 및 판정을 담당하는 비정상행위 판정모듈로 구성된다. 모델학습을 통해 정상행위를 프로파일 데이터베이스로 구축하여 판정시 사용한다.

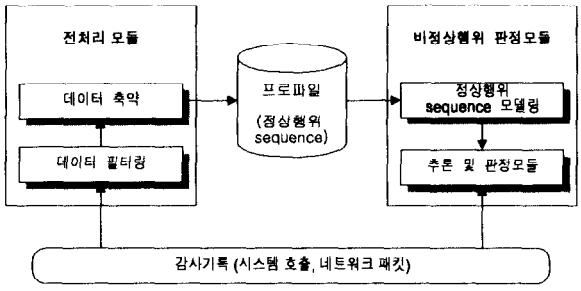

그림 1 . 침입탐지 시스템 개요

3.2 전처리

전처리에서는 크게 두 가지 변량을 추출하였다. 첫 번째는 사용자의 시스템호출을 사용하였다. 모든 시스템호출번호가 다 사용되지 않으므로 통계적으로 빈도가 높은 49개의 시스템호출에 대해 0부터 48번까지의 번호를 부여하였고 그 밖의 시스템 호출은 49번을 부여하여 총 50개의 축약된 시스템 호출번호를 사용하였다.

두 번째는 BSM을 통해 얻을 수 있는 정보 중 시스템 호출관련(시스템 호출 ID, 시스템 호출 반환 값, 시스템 호출 반환상태), 프로세스 관련(프로세스 ID, IPC ID, IPC permission, exit value, exit status), 파일접근관련(중요파일 접근 여부, 중요경로 접근여부, 파일 접근 모드, 인자길이) 등을 각각 추출한 후 SOM을 통해 축약한 정보를 사용하였다.

순서적으로 생성되는 이벤트는 일정 크기의 윈도우를 옆으로 이동시켜가면서 윈도우 크기 만한 시퀀스로 추출하였다.

3.3 침입탐지를 위한 HMM

HMM은 실제적인 생성모델을 알 수 없고 단지 생성된 시퀀스에 의해서만 확률적으로 관측할 수 있는 이중으로 확률적인 절차로서[6], 사용자의 행위시퀀스를 모델링하기에 유용한 도구이다. HMM은 관찰 시퀀스의 길이, 상태수, 심볼수와 학습에 의해 조정되는 전이확률, 관측확률, 초기상태분포로 구성이 된다. 전이확률은 한 상태에서 다음상태로 전이할 확률을 나타내며, 관측확률은 한 상태에서 특정 심볼이 관측될 확률을 나타낸다. 초기 상태 분포는 처음에 해당 상태에서 시작할 확률을 나타낸다. HMM은 다음과 같이 표현되며, 모델 λ는 간략히 (A, B, π)로 표현될 수 있다.

- T : 관찰 시퀀스의 길이
- N : 모델의 상태수
- M : 관찰 심볼의 수
- $Q = q_1, q_2, \ldots, q_N$: 상태들
- $V = v_1, v_2, \ldots, v_M$: 가능한 관찰심볼의 이산적인 집합
- $A = \{a_{ij}\}, a_{ij} = \Pr(q_j \text{ at } t+1 | q_i \text{ at } t)$: 상태전이 확률분포
- $B = \{b_j(k)\}, b_j(k) = \Pr(v_k \text{ at } t | q_j \text{ at } t)$: 관측 심볼 확률분포
- $\pi = \{\pi_i\}, \pi_i = \Pr(q_i \text{ at } t=1)$: 초기 상태 분포

가) 정상행위 모델링

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 않고 반복적으로 λ를 결정하는 방법으로 Baum-Welch의 재추정식이 있다[4].

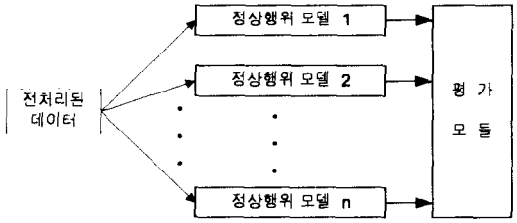

그림 2. 정상행위 모델링

나) 비정상행위 판정

비정상행위 판정에서는 이미 구축되어 있는 정상행위별 HMM에 사용자행위시퀀스를 입력으로 넣어 각 정상행위에서 현재 행위가 생성되었을 확률을 구한다. 확률을 구하는 방법으로는 forward-backward procedure나 Viterbi 알고리즘을 사용할 수 있다[4]. 각 모델별로 구해진 확률은 판정모듈에 전달되어 비정상행위인지 판정한다.

3.4 다중모델의 결합방법 및 비정상행위 판정

하나의 이벤트가 각 모델별 평가루틴을 통과하면 여러 개의 평가값이 생성된다. 이 평가값들을 토대로 침입여부를 판정하기 위해서 다중모델을 결합하기 위한 방법이 필요하다. 본 시스템에서는 모델별로 침입여부를 결정한 후 투표방식을 사용하여서 모델들의 결정사항들을 결합하고 있다. 투표방식을 적용할 때는 모델들의 신뢰도에 따라 투표가중치 W_m이 부여되고, 원하는 결과에 따라 투표의 집계방식을 결정한다. 표결방식으로는 대표적으로 만장일치 방식, 다수결 방식, OR 방식이 사용된다. OR방식은 하나의 투표자라도 가로 투표하는 경우 결과가 가로 결정되는 방식이다. 투표방식은 집계결과 R이

표결방식에 따른 임계값 T를 넘는지를 결정한다.

$$R=\sum W_m * V_m \quad \begin{pmatrix} W_m : \text{모델별 가중치} \\ V_m : \text{모델별 투표결과} \end{pmatrix}$$

$$\text{가결 if} \begin{cases} R=1 & (\text{만장일치}) \\ R\geq 0.5 & (\text{다수결}) \\ R>0 & (OR\text{방식}) \end{cases}$$

일반적으로 OR방식으로 갈수록 판정율이 높아지지만 오류율도 함께 높아지며 만장일치로 갈수록 오류율을 줄일 수 있지만 판정율도 낮아진다.

4. 실험결과

실험 데이터로는 한 명의 사용자가 1주일간 발생시킨 데이터를 사용하였다. 주사용 프로그램은 문서편집기와 컴파일러, 그리고 사용자가 작성한 프로그램이었다. 학습 데이터와 테스트 데이터 모두 10,000개를 사용했으며, 테스트 데이터에는 침입을 17차례가 포함되어있다.

그림 3은 침입행위가 발생한 경우 시퀀스 평가값의 변화를 보여준다. 침입이 시작된 시간 11에서부터 침입행위가 종료된 시간 32사이에 있는 시퀀스들의 평가값이 급격히 낮아져 HMM이 효과적으로 침입을 탐지하고 있는 것을 보여주고 있다.

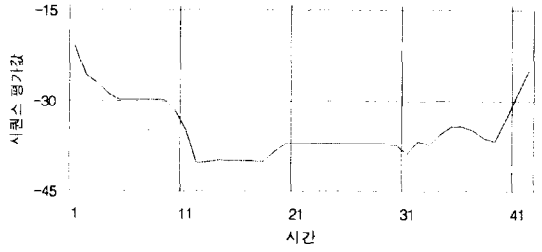

그림 3 . 침입행위 발생시 시간에 따른 평가값의 변화

다중모델 결합실험에서는 시스템호출을 사용한 경우와 SOM을 이용하여 가장 좋은 성능을 보인 3가지 변량을 축약한 경우의 평가값을 투표방식을 이용해서 결합하였다. 투표방식으로는 만장일치, 다수결, OR방식을 사용하였으며 시스템호출과 SOM축약에 동일한 가중치를 부여하였다.

	syscall	SOM 축약	투표		
			만장일치	다수결	OR투표
탐지율	100	100	100	100	100
F-P 오류율	5.33	23.53	1.18	25.75	25.75

표 1. 다중모델 결합방법의 성능

변량별로 임계값이 다를 수 있으므로 투표방식에서는 각 변량에서 가장 좋은 성능을 보인 결과만을 가지고 평가했다. 탐지율의 경우에는 모두 100%를 보였으므로 투표에 의해서 변경된 부분은 없다. False-positive 오류율의 경우, 만장일치의 경우 기존 두 방법에 비해 나은 성능을 보여주었다.

5. 결론 및 향후연구

본 논문에서는 사용자 이벤트의 여러 변량을 HMM을 사용하여 모델링한 후 각 변량에서 생성된 결과를 투표방식에 의해 결합하는 기법을 제안하였다. 실험결과에서 보듯이 다중 변량을 결합한 결과 오판율을 줄일 수 있어 더 안정적인 판정을 내릴 수 있었다.

순서 정보를 이용한 비정상행위 침입탐지 시스템을 위해서는 이후 다음과 같은 연구가 더 수행되어야 할 것이다.

• 다변량결합

다변량 결합시에 각 변량의 중요도를 반영할 수 있도록 각 변량의 가중치를 부여할 수 있어야 한다. 또한 어떠한 변량이 명백히 침입임을 나타냈을 때 전체결과가 다수의 정상을 나타내는 변량에 의해 정상으로 결정되는 것을 방지할 수 있는 장치도 필요하다. 퍼지추론등을 통해서 판정결과를 계량화시키고 어휘적으로 표현할 수 있어야 한다.

• 모델링

모델링에 있어서는 모델수와는 별도로 모델링 수준에 대한 고려도 필요하다. 현재 저수준의 BSM 감사기록을 사용하고 있지만 동일한 저수준의 이벤트정보가 고수준의 사용자 의도나 시간의 변화에 따라 다르게 해석될 수 있으므로 고수준의 문맥특성을 반영할 수 있는 모델링 방법에 대한 연구가 보충되어야 할 것이다.

• 시퀀스생성

시퀀스의 생성에 있어 중요한 것은 이전에 발생한 이벤트와의 연관성을 잃지 않고 새로운 시퀀스를 생성하는 것이다. 본 논문에서는 하나의 이벤트가 발생하면 이전 시퀀스를 슬라이딩하여 완전 중복시키는 방법을 사용하였는데 완전 중복이외의 시퀀스 생성방법과 비교검토가 수행되어야 할 것이다. 시퀀스의 길이에 있어서도 모든 행동이 동일한 길이의 시퀀스로 표현되지 않을 수 있으므로 여러 시퀀스 길이를 동시에 고려할 필요가 있다.

참고문헌

[1] A. K. Ghosh, A. Schwartzbard and M. Schatz, "Learning program behavior profiles for intrusion detection," *Proc. Workshop on Intrusion Detection and Network Monitoring*, Santa Clara, USA, April 1999.

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

[3] T. Lane and C. E. Broadly, "Temporal sequence learning and data reduction for anomaly detection," *Proc. ACCS '98*, pp. 150~158, 1997.

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

[5] C. Warrender, S. Forrest and B. Pearlmutter, "Detecting intrusions using system calls: Alternative data models," *Proc. IEEE Symposium on Security and Privacy*, May 1999.