HMM 기반 비정상 침입탐지 시스템

김주호° 공은배 조성현
충남대학교 컴퓨터공학과 홍익대학교 컴퓨터공학과
{kimjh°, keb}@ce.cnu.ac.kr scho@hongik.ac.kr

HMM Based Anomaly Intrusion Detection System

JuHo Kim° EunBae Kong SungHyun Cho
Dept. of Computer Engineering, ChungNam National University, Hongik University

요 약

인터넷 인구의 확산과 개방된 시스템 환경속에서 네트웍과 시스템에 대한 침해사고 건수가 날로 증가하고 있는 가운데 최근 국내 인터넷망 대부분이 다운되는 등 그 피해 규모도 점차 막대해지고 있다. 이에 따라 침해 사고에 대해 사고 발생 즉시 민첩하게 대응하여 피해를 최소화하고, 더 나아가서는 사고를 미연에 방지하기 위한 보안 관련 시스템들에 관한 연구가 활발히 진행되고 있다. 본 연구에서는 보안관련 솔루션 중에 하나인 침입탐지시스템(IDS: Intrusion Detection System)에 대해 살펴보고, IDS의 탐지방식 중 비정상탐지(Anomaly Detection)분야에 은닉 마르코프 모델(HMM: Hidden Markov Model)을 적용하여 사용자별로 명령어 사용 패턴을 프로파일링하는 HMM 기반 비정상 침입탐지 시스템을 제안하고자 한다. 실험결과 자신의 데이터에 대해서는 평균 93% 이상의 만족할만한 탐지 정확도를 보였고, 다른 사용자의 데이터에 대해서는 모델마다 다소 차이를 나타냈다.

1. 서 론

컴퓨터 시스템은 기본적으로 비밀성과 무결성에 대한 보장을 제공해야 한다. 그러나 인터넷으로 인한 접속자수의 급증과, 방대하고 다양한 범위의 실질적인 금융거래로 시스템은 점점 더 침입자들의 공격대상이 되어가고 있다.

시스템의 보안 메커니즘들이 시스템의 리소스와 데이터에 대한 인증되지 않은 접근을 막아내는 것은 매우 당연한 것처럼 보이지만, 실질적으로 안전성에 문제가 보이는 작은 틈까지 완벽하게 막아내는 것은 거의 불가능하다[1]. 이미 잘 알려진 침입형태에 대해서는 대부분의 상용 시스템들이 별 문제없이 잘 대처하고 있지만, 침입기법들이 점점 다양해지고 고도화되어감에 따라 보안관련 시스템들도 사고발생 이전의 이상징후들에 대해 보다 능동적으로 대처하는 기술들이 요구되어지고 있다[2][3].

현재 상용으로 개발되어 있는 대부분의 시스템들은 현재까지 알려진 공격들에 대해서만 탐지가 가능한 오용탐지(Misuse Detection)방식을 채택하고 있기 때문에 새로운 공격기법을 막아내는 데는 구조적으로 한계를 가지고 있다. 반면 사용자의 정상행위에 어긋나는 정도를 통해 침입여부를 판단하는 비정상탐지 시스템은 새로운 공격시도를 사전에 차단하여 사고를 미연에 방지하는 능동적 방어 개념이라고 볼 수 있다[4].

본 논문에서는 비정상 탐지 방식의 접근 방법 중의 하나인 HMM에 대해 살펴보고 UNIX 사용자의 쉘 명령어 데이터를 HMM으로 모델링한 후 실험을 통해 구현된 시스템의 유용성을 검증해 보려고 한다. 본 논문의 구성은 다음과 같다. 2장에서는 관련연구로써 IDS와 HMM에 대해 살펴보고, 3장에서는 실험에 사용할 데이터 SET의 가공과정과 사용자 모델링 과정을 설명한다. 4장에서는 실험결과를 통해 시스템의 유용성을 검토해보고, 마지막으로 5장에서는 결론 및 향후 연구방향을 제시한다.

2. 관련연구

2.1 침입탐지 시스템(IDS)

IDS는 대상 시스템에서 허가되지 않거나 비정상적인 행위에 대하여 탐지, 식별, 보고하는 기능의 소프트웨어로 정의할 수 있다[1]. IDS는 크게 보호하고자 하는 대상 시스템에 따라 네트워크 기반(Network-based) IDS와 호스트 기반(Host-based) IDS 그리고 탐지방식에 따라 비정상 탐지(Anomaly Detection) 시스템과 오용탐지(Misuse Detection) 시스템으로 분류할 수 있다[12].

탐지 대상에 있어서는 네트워크 기반 IDS가 좀 더 일반적인데, 네트워크를 통해 전송되는 트래픽을 검사하여 침입을 탐지하는 반면 호스트 기반 IDS는 로컬 호스트에서 사용자의 행위나 프로세스(process) 등을 검사하여 침입을 탐지하게 된다.

탐지방식에 있어서 비정상 탐지는 시스템 가동 전에 정상적인 사용자 행동에 대한 프로파일을 작성해 두고, 가동 후에 실제 행위들을 정상적인 프로파일과 비교하여 공격을 탐지하는 방식이다. 비교 과정에서 기존의 프로파일을 수정하거나 새로운 프로파일을 추가하기도 한다.

오용탐지는 알려진 취약성에 대한 시그널(패턴) 정보를 가지고 실제적인 공격이 시도될 때 이를 탐지하는 방식이다[10]. 비정상 행위 탐지가 침입으로 여겨지는 행위를 탐지한다면 오용탐지는 패턴과 100% 일치하는 경우, 즉 명백한 침입을 탐지하게 된다. 오용탐지는 비교적 구현하기 쉽고 비용이 저렴하여 대부분의 상용 시스템들이 채택하고 있다[3].

비정상 탐지방식은 사용자에 대한 프로파일을 구축하고 정확도를 높이기 위해 대량의 데이터 분석이 필요하고, 복잡한 계산과정과 False-Positive 오류 등 상용 제품에 실제로 도입하기에는 여러 가지 문제점들이 제기되고 있다.

2.2 은닉마르코프 모델(HMM)

HMM은 우리가 알 수 없는 미지의 상태(Hidden States)들의 확률론적 전이과정을 그 상태들이 발생시키는 현실적으로 관찰 가능한 상태(Observable Symbol)의 확률적 전이과정을 통해 모델링하는 이중의 확률론적 과정이다[5]. 예를 들면 해초의 마

른 정도를 보고 날씨의 상태변화를 모델링한다거나, 음파나 공기의 압력, 또는 조음기관의 위치 등을 이용하여 두뇌의 활동을 모델링하는 것 등을 예로 들 수 있다.

HMM은 그동안 음성인식이나 Classification 분야에서 많이 사용되어 왔는데, 비정상 탐지 시스템 자체가 사용자의 의도가 은닉된 상태에서 사용자가 만들어 내는 순서 데이터의 정상 여부를 판정하는 문제로 볼 수 있기 때문에 비정상 탐지분야에도 효과적으로 적용될 수 있다[6][7]. HMM은 기호로 $\lambda=(\Pi,A,B)$로 표현되며 은닉 상태수를 N, 관찰심볼의 개수를 M이라고 할때 다음과 같은 구성요소들을 가진다[5].

1) $Q = \{q_1, q_2, \ldots\ldots, q_n\}$: 은닉 상태 집합
2) $V = \{v_1, v_2, \ldots\ldots, v_m\}$: 관찰심볼 집합
3) $O = (o_1, o_2, \ldots\ldots, o_{t-1}, o_T)$: 관찰심볼로 이루어진 관측열
4) π_i : 은닉 상태들의 초기 확률 백터
5) A_{ij} : 은닉 상태들간의 전이 행렬
6) B_{ij} : 특정 은닉상태에서 특정 관찰 가능 상태가 나타날 확률

우리가 풀고자 하는 문제에 대해서 HMM으로 정의하고 나면 실제로 모델을 사용하기 위해서는 다음과 같은 세 가지 문제를 해결해야만 한다.

1) 확률추정(probability estimation)의 문제
   관찰되어진 관측열 $O=(o_1, o_2, \ldots\ldots, o_{t-1}, o_T)$이 주어진 모델 $\lambda=(\Pi,A,B)$에서 관찰되어졌을 확률 $P(O|\lambda)$를 계산하는 문제
2) 최적 상태 순서(optimal sequence)의 결정 문제
   모델 λ에서 관찰된 관측열 $O=(o_1, o_2, \ldots\ldots, o_{t-1}, o_T)$를 생성할 확률이 가장 높은 은닉된 상태들간의 순서 $q=(q_1, q_2, \ldots\ldots, q_n)$을 찾는 문제
3) 매개변수 추정(parameter estimation) 문제
   관찰된 관측열 $O=(o_1, o_2, \ldots\ldots, o_{t-1}, o_T)$이 모델 $\lambda=(\Pi,A,B)$에서 관찰되었을 확률을 최대로 하는 모델 λ의 매개변수 π_i, a_{ij}, b_{ij} 각각의 확률값을 재추정하는 문제

위에서 제기된 3 가지 문제에 대한 해결방안으로 몇 가지 알고리즘들이 널리 사용되어왔다. 확률추정 문제의 경우 전·후향(Forward/Backward) 알고리즘을 이용하여 해결이 가능하며, 최적 상태순서의 결정 문제는 동적 프로그래밍 기법중의 하나인 Viterbi 알고리즘을 사용한다. 마지막으로 매개변수 추정은 EM 알고리즘이라 불리는 Baum-Welch 알고리즘을 사용하여 풀 수 있다[9].

3. HMM 기반 비정상 침입탐지 시스템

본 연구에서는 사용자가 내리는 유닉스 쉘 명령어 데이터를 관찰심볼로 사용하여 각 사용자별로 HMM 모델을 구성해 보았다. 명령어 열만으로는 사용자의 의도를 규정하기 어렵기 때문에 은닉 상태에 대해서는 특별히 의미를 부여하지 않고 개수만을 정의했다. 실험에 사용한 데이터는 미 퍼듀 대학의 COAST 실험실에서 2년 동안 8명(USER0~USER7)의 학생들을 대상으로 수집한 UNIX 쉘 명령어 데이터이다. 표 1은 USER0 데이터를 HMM에서 행렬로 표현하기 위해 수치화하는 과정을 보인 것이며 나머지 7명의 데이터도 동일한 과정을 거친다. 첫 번째 단계는 필터링 과정으로 원본 데이터의 분석결과 일반적인 유닉스 시스템에서 사용하지 않는 명령어와 타이핑이 잘못된 명령어 등, 데이터를 직접 사용하기에는 무리가 있다고 판단되어 유닉스 시스템의 매뉴얼 페이지에 등록된 명령어만을 골라 필터링하였다. 이 필터링 데이터를 중복없이 줄이면 각 사용자마다 모델링을 할 경우 바로 관찰심볼의 집합이 되는데 이 과정이 두 번째 단계이다. 세 번째 단계는 문자열로 되어있는 필터링 데이터를 관찰심볼 집합의 라인번호로 수치화하는 과정이다.

표 1 USER0 데이터 가공 과정

가공 과정	Raw Data	① →	Filtering Data	② →	Symbol Data	③ →	Digit Data
라인수	8974		4556		70개		4556
명령어 데이터	**SOF**		whoami		1 whoami		1
	whoami		pwd		2 pwd		2
	pwd		ls		3 ls		3
	ls		dir		4 dir		4
	dir		vi		5 vi		5
	vi		source		6 source		6
	source		source		7 exit		6
	<1>		exit		.		7
	source		.		.		.
	<1>		.		.		.
	exit		.		.		.
	EOF		.		.		.
	.....		.		70 set		.

데이터의 가공 과정이 끝나면, 우선 사용자의 Digit Data를 입력열로 하여 초기모델을 구성하고 2000개의 데이터에 대해서 학습과정을 거쳐 모델을 최적화한다. 다시 1000개 데이터의 모델 확률값으로 임계값을 결정한 다음 학습과 임계값 결정에 사용되지 않은 데이터 1000개로 완성된 모델을 테스트한다. 본 실험에서는 은닉상태수와 입력열길이가 실험결과에 미치는 영향을 살펴보기 위해 USER0를 대상으로 상태수와 열길이를 변화시키면서 서로 다른 모델을 구성해 보았고, 각 모델에 테스트 데이터로 USER4 데이터를 입력해 보았다. 위 실험에서 최적의 결과를 보인 상태수와 입력열로 나머지 사용자에 대해서도 동일한 조건으로 초기모델을 구성하고 학습과정을 거쳐 모델을 테스트해 보았다.

4. 실험결과

실험의 결과그래프는 1000번의 테스트 중 정상 판정된 입력열의 개수를 확률로 나타낸 것인데, 이것은 바로 모델이 자신의 데이터와 다른 사용자를 구별해내는 정확도를 나타내며 자신의 데이터에 대해서는 높게 나타나야 하지만, 다른 사용자의 데이터에 대해서는 낮게 나타나야 한다.

은닉상태수와 입력열 길이는 너무 작으면 사용자들간의 변별력이 줄어들고 너무 크면 over-fit되는 현상이 있는데, 실험결과 은닉상태수는 실험결과에 미치는 영향이 거의 없는 것으로 나타났고 입력열 길이는 50개 정도에서 만족할만한 계산속도와 정확도를 보였다.

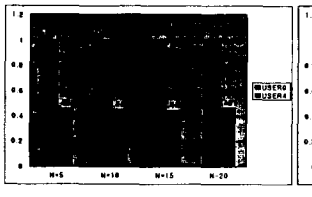

그림 1 은닉 상태수에 따른 모델 정확도

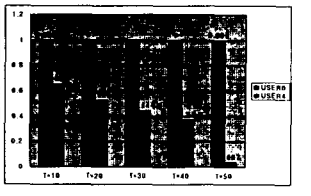

그림 2 입력열 길이에 따른 모델 정확도

다음은 위 실험결과에서 결정된 최적의 상태수(N=15)와 열길이(T=50)로 각 사용자를 모델링하고 모델링에 사용하지 않은 자신의 데이터와 다른 7명의 데이터를 테스트한 결과이다.

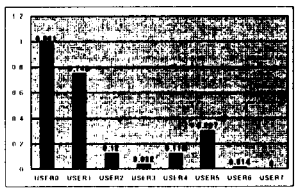

그림 3 USER0 모델정확도

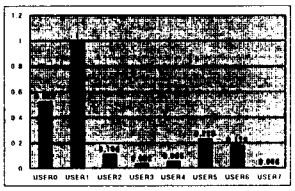

그림 4 USER1 모델정확도

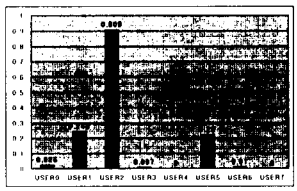

그림 5 USER2 모델정확도

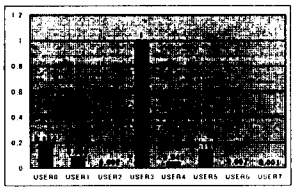

그림 6 USER3 모델정확도

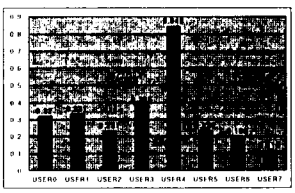

그림 7 USER4 모델정확도

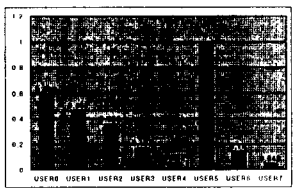

그림 8 USER5 모델정확도

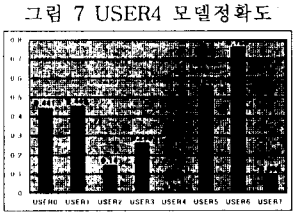

그림 9 USER6 모델정확도

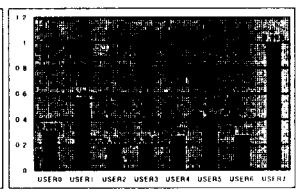

그림 10 USER7 모델정확도

실험결과 대체로 자신의 데이터에 대한 탐지 정확도는 상당히 높은 반면, 다른 사용자를 구별해내는데는 사용자마다 적지 않은 차이를 보이고 있다. 실험결과는 사용자 변별력을 기준으로 볼 때 크게 두 부류로 나눌 수 있는데, USER0~USER3은 비교적 변별력이 높은 반면, USER4~USER7의 변별력은 상대적으로 떨어진다. 이것의 원인은 사용자를 구별하는 한 요소인 명령어 사용폭, 즉 관찰심볼 개수의 차이로 볼 수 있는데, USER0~USER3의 심볼 개수가 60~77개로 비교적 적은데 반해 USER4~USER7의 심볼개수는 89~127로 USER0~USER3가 묻혀버리는 양상이 나타나기 때문이다. 또 비교적 변별력이 좋은 그룹에 속해있는 USER0와 USER1은 서로 사용하는 명령어의 차이가 적고, 명령어 사용폭도 70개와 77개로 비슷하기 때문에 서로 잘 구별하지 못하는 반면, USER2와 USER3는 명령어 사용폭은 65개와 60개로 비슷하지만 서로 사용하는 명령어가 많이 다르기 때문에 확연히 구분되고 있다.

5. 결론 및 향후 과제

본 논문에서는 UNIX 사용자의 명령어 사용 패턴과, 빈도수 그리고 사용폭 등을 주요요소로 하여 각 사용자를 HMM으로 모델링한 후 입력데이터의 모델확률값으로 정당한 사용자인지 여부를 판별하는 HMM 기반 침입탐지 시스템을 제안하였다. 실험결과 자신의 데이터에 대해서는 평균 93% 이상의 높은 정상판정율을 보이는 한편 다른 사용자의 데이터도 평균 20~30% 정도를 정상으로 판정하고 있다. 그러나 이것은 거의 100%에 가까운 자신의 데이터에 대한 정상판정율에 비하면 상당히 낮은 수치라고 볼 수 있으며, 만약 정상 사용자의 임계값으로 각 모델마다 (자신의 데이터에 대한 정상판정율-5%) 이상을 유지해야 한다고 했을 때 실험 데이터에서 비정상 사용자를 정상으로 판정할 오류율은 0이다. 즉 실험결과 그래프에서 볼 수 있듯이 8개의 모델실험 결과 정상사용자를 제외하고는 이 값을 넘는 사용자는 없다. 또한 실험 데이터의 필터링과정에서 원본 데이터 중 UNIX 시스템의 매뉴얼 페이지에 등록되어 있는 명령어들만 추출했기 때문에 X-window상에서의 어플리케이션 실행이나 네트웍 관련 툴 실행 등 사용자마다 자주 사용하는 프로그램들에 대한 실행 명령어까지 모델링에 사용한다면 사용자들간의 변별력은 상당히 커질 것으로 예상된다. 향후 연구과제로는 사용자간의 변별력을 높이는데 중요한 역할을 하는 데이터를 선별하는 작업과 대상 시스템에 무리를 주지 않는 범위내에서 실시간으로 비정상 탐지 시스템을 적용하기 위한 연구 등이 진행되어야 하겠다.

[참고문헌]

[1] Aurobindo Sundaram, "An introduction to intrusion detection"

[2] 조성배, "비정상행위 탐지기반 침입탐지 시스템의 현황 및 미래" 연세대학교 컴퓨터 · 산업공학부. 2002. 5

[3] http://www.kisa.or.kr 한국정보보호진흥원

[4] Terran Lane, "Machine learning techniques for the domain of anomaly detection for computer security" Purdue University. July 1998.

[5] 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.

[6] Terran Lane, "Hidden Markov models for human/computer interface modeling" Purdue University

[7] 최종호, 조성배, "Application of Hidden Markov Model to intrusion detection system" 정보과학회논문지:소프트웨어 및 응용 제 28 권. 제 6 호. (2001.6)

[8] 김인영, 장병탁 "A study on Hidden Markov Models for intrusion detection" 서울대학교

[9] 엄재홍, "Information extraction using Hidden Markov Models" 서울대학교, 2000년 10월

[10] Sandeep Kumar, Eugene H. Sapfford, "A pattern matching model for misuse intrusion detection" COAST Project. Department of Computer Sciences Purdue University

[11] Christina Warrender, Stephanie Forrest, Barak Pearlmutter, "Detecting intrusions using system calls: Alternative data models" Proc. IEEE Symposium on Security and Privacy. pp. 133~145(1999)

[12] Dit-Yan Yeung, Yuxin Ding "Host-based intrusion detection using dynamic and static behavioral models" Pattern Recognition 36(2003). pp. 229~243