대한산업공학회/한국경영과학회 2004 춘계 학술대회
2004 년 5 월 21 일 ~22 일 전북대학교

데이터마이닝 기법을 이용한 침입탐지시스템에 관한 연구

홍태호, 김진완, 김유일
부산대학교 경영학과

Abstract

최근 들어 네트워크 침입탐지시스템은 정보시스템 보안에서 매우 중요하게 인식되고 있다. 네트워크침입시스템에 데이터마이닝 기법들을 활용하는 연구들이 활발하게 그 동안 활발하게 진행되어 왔다. 하지만 단순한 데이터마이닝 기법의 적용만으로는 침입탐지시스템의 효과를 극대화 할 수 없다. 침입탐지시스템은 오류의 종류에 따라 조직에 미치는 영향이 매우 상이한 특징을 갖는다. 따라서 본 연구에서는 침입탐지시스템의 오류의 특징에 따른 각기 다른 데이터마이닝 기법을 적용하는 방안을 제시하였다. 또한 국내에서 사용된 실제 네트워크를 통한 침입공격에 관한 데이터를 수집하고, 신경망, 귀납적 학습법, 러프집합을 적용하여 국내 데이터 특성을 고려한 네트워크 침입탐지모형을 제시하였다.

Ⅰ. 서론

오늘날 인터넷 사용의 폭발적 증가로 대부분의 시스템이 타 시스템과 네트워크로 연결되어 있는 상황 하에서 악의적인 해킹 또는 네트워크 침입은 그 시스템을 운영하는 조직뿐만 아니라 네트워크에 연결되어 있는 타 기관들에게도 치명적인 손해를 입힐 수 있는 구조를 갖고 있다. 따라서 정보시스템의 네트워크 환경의 급속한 발달로 인한 역기능을 줄이기 위한 네트워크 침입탐지시스템의 필요성이 강조되고 있다. 네트워크 환경 하에서 정보시스템을 운영하는 회사, 학교, 정부 기관 등은 침입탐지시스템을 운영하여 외부의 해커 또는 공격자로부터 정보자원을 보호하고 있다.

기존의 네트워크 침입탐지시스템에 사용되는 침입탐지모형은 전문가들의 지식을 이용한 네트워크 침입자 또는 해커의 행위를 탐지하여 비정상적인 정보시스템에 대한 접근을 제한하는 형태가 일반적이다. Zhu(2001) 등은 네트워크 침입탐지분야에서 데이터마이닝 기법이 매우 우수한 성과를 보이는 것으로 보고했다. 데이터마이닝 기법을 이용한 침입탐지모형은 기존의 데이터를 이용하여 새로운 패턴을 발견할 수 있다는 장점이 있다.

그러나 침입탐지시스템의 오류에 따른 조직에 미치는 영향은 서로 상이하다. 이러한 오류의 특성에 데이터기법의 적용을 위한 방법이 필요하며, 국내 네트워크 환경에 적합한 침입탐지시스템의 개발이 또한 필요하다. 따라서 본 연구에서는 최근 매우 중요하게 인식되고 있는 정보시스템 보안의 한 분야인 네트워크 침입탐지시스템에 데이터마이닝 기법을 적용하여 국내 특성에 맞는 침입탐지시스템을 개발하고자 한다.

Ⅱ. 이론적 배경

2.1 침입탐지시스템

침입이란 컴퓨터 시스템의 자원들에 대한 허가되지 않은 접속 또는 사용을 말한다(Esmaili et al., 1996). 정보자원을 보호하기 위한 침입탐지시스템(Intrusion detection system)에 대한 정의는 다음과 같이 다양하게 되고 있다. 침입탐지시스템은 목표 시스템에 대한 허가되지 않았거나 변칙적인 활동들을 탐지하고, 식별하고, 대응하는 기능을 가진 소프트웨어이다(Richards, 1999). 침입탐지시스템의 목적은 실시간으로 또는 일괄처리 방식으로 보안 위반을 탐지하기 위한 메커니즘을 제공하는 것이다(Debar et al., 1992). 위반은 시스템을 파괴하

려고 시도하는 외부인들에 의해 일어나거나, 권한을 오용하기 위해 시도하는 내부인들에 의해 일어난다(Weber, 1999). 침입탐지시스템은 다양한 시스템과 네트워크 자원들로부터 정보를 수집한 뒤에 침입과 오용에 관한 신호를 보내기 위해 정보를 분석한다(Lippmann and Cunningham, 2000). 침입탐지시스템에 의해 수행될 수 있는 주요한 기능으로는 사용자와 시스템 활동을 모니터링하고 분석하며, 중요한 시스템과 데이터 파일들의 무결성을 평가하며, 알려진 공격들을 반영한 활동 패턴들을 인식하며, 탐지된 활동에 자동적으로 대응하며, 그리고 탐지 프로세스의 결과를 보고한다. 이러한 침입탐지시스템은 전자상거래, 교육기관, 은행 등의 금융기관, 일반회사의 인트라넷 등에 다양하게 적용될 수 있다.

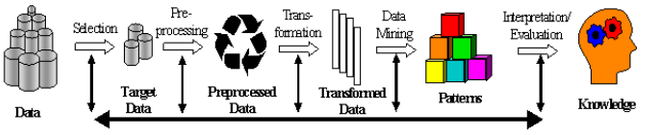
침입탐지는 탐지 방법에 따라서 크게 오용 탐지(misuse detection)와 비정상적 탐지(anomaly detection)의 두 가지 범주로 구분할 수 있다(Zue et al., 2001). 첫째, 오용 탐지는 잘 알려진 공격들의 증적이나 패턴을 검색하는 방법으로 탐지한다. 오직 특징적인 증적을 남긴 잘 알려진 공격들만이 이 방법으로 탐지되어질 수 있다. 둘째, 비정상적 탐지는 정상적 사용자나 시스템 행동의 모델을 사용하고, 악의적인 가능성이 있는 경우에는 정상적인 사용과 편차가 발생하는 지를 탐지한다. 정상적 사용자나 시스템 행동에 관한 이 모델은 일반적으로 사용자 또는 시스템 프로파일로서 알려져 있다. 비정상적 탐지의 주요한 강점은 사전에 알려지지 않은 공격들을 탐지하기 위한 능력이 있다는 것이다.

침입탐지시스템은 분석하는 감사 데이터 출처의 종류에 따라서도 분류되어질 수 있다(Joo et al., 2003). 대부분의 침입탐지시스템은 공격들을 인지하거나 피하기 위한 접근법으로 네트워크 기반 침입탐지 또는 호스트 기반 침입탐지로 구분된다. 침입탐지시스템이 네트워크 트래픽에서 패턴을 찾는 경우에는 네트워크 기반 침입탐지로 분류된다. 침입탐지시스템이 로그 파일에서 공격 흔적을 찾는 경우에는 호스트 기반 침입탐지로 분류된다.

침입탐지시스템에 대한 최근의 많은 접근법들은 데이터마이닝 기술들을 활용하고 있다(Lam et al., 1996). 이러한 접근법들은 시스템에 의해 수집된 감사 추적의 대형 데이터 셋에 데이터마이닝 기술들을 적용하는 방법으로 탐지 모델을 구축한다 (Helman and Liepins, 1993).

2.2. 지식채굴 프로세스

지식채굴은 데이터에서 유효하고, 새롭고, 유용한 그리고 궁극적으로 이해할 수 있는 패턴을 확인하는 과정이다(Fayyad et al., 1996). 지식채굴 과정에 대한 완전한 방법론은 존재하지는 않지만 지식체굴은 <그림 1>과 같은 단계를 거쳐 발견된다고 할 수 있다.

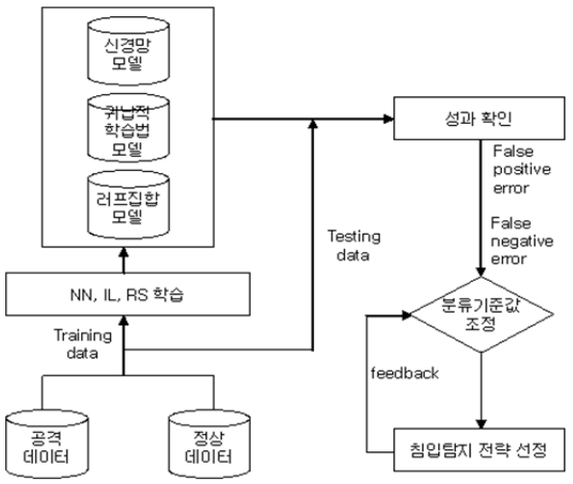


<그림 1> 지식채굴 프로세스의 개요

지식채굴을 위해서는 먼저 데이터와 업무를 파악하고 분석을 위한 적당한 환경으로 옮겨서 데이터를 사용되어질 형태로 통합시키고 검토하게 된다. 이때 데이터에서 이상치와 에러를 제거하여 데이터를 정리하고 적용하는 기법에 따른 모형을 세우고 이를 검증하기 위한 기본 가정을 세운다. 이 가정을 검증하기 위해 데이터마이닝 기법을 적용하여 패턴이나 지식을 채굴해 나간다. 여기서 채굴된 지식의 유용성을 검증하고, 새로 발견된 지식의 사용을 위해 해석을 하게 된다. 지식채굴을 수행하기 위해서는 통계적 방법론과 인공지능기법을 이용하는 방법 등이 사용되고 있는데, 본 연구에서는 인공지능기법들 중에서 신경망, 귀납적 학습법, 러프집합을 적용하도록 한다.

Ⅲ. 연구모형

본 연구에서 제안하는 연구모형은 <그림 2>와 같이 인공신경망, 귀납적 학습법, 러프집합을 이용한 침입탐지모형을 개발하였다. 각 모델의 성과로 정상적인 사용자를 잘못 탐지하여 사용제한을 시키고 이에 대한 대응책을 실행하게 하여 일어나는 기회비용인 False positive error 와 악의적인 침입자를 정상 사용자로 분류하여 시스템에 대한 접근을 허용함으로써 발생되는 정보 시스템 자산의 피해인 False negative error 를 사용하였다. 추가적으로 기업의 상황에 맞는 침입탐지 전략 선정을 위해서 분류기준값(Treshold)을 0.3~0.7 로 조정하면서 False positive error 와 False negative error 의 변화에 따른 성과분석을 제시하였다. 이를 통해 보안담당자는 기업 환경에 적합한 최적의 비용으로 침입탐지 전략을 설정할 수 있다.



<그림 2> 지능형 침입탐지시스템 연구모형

3.1 실험 설계

실험에 사용된 데이터는 통합 보안 서비스를 제공하는 Cyber-PATROL 사의 IDS 센스로부터 무작위로 정상 패턴과 공격 패턴을 포함한 500 개의 사례를 수집하였다. 이 표본은 다시 Training data 와 Testing data 로 분할하였다. 침입탐지를 위한 신경망, 귀납적 학습방법, 러프집합의 입력 변수들은 문헌 연구 및 전문가와의 인터뷰를 통해 <표 1>과 같이 선정하였다.

<표 1> 모델에 사용된 변수

변수	설명
Event Data	기록된 사건의 날짜와 시간
Protocol ID	사건과 관련된 프로토콜
Source Port	출처의 포트 숫자
Destination Port	도착지의 포트 숫자
Source IP Address	출처의 IP 주소
Destination IP Address	도착지의 IP 주소

본 연구에서는 침입탐지를 위한 데이터마이닝 기법으로 인공신경망, 귀납적 학습법, 러프집합의 3 가지 실험을 진행하였다. 3 가지 실험은 모두 10-fold cross validation 으로 수행하기 위해서 무작위로 서로 다른 데이터를 생성하였다. k-fold cross-validation 은 전체 표본을 k 개의 테스트 집합으로 분할하여 실험하는 것이다.

4.2 데이터마이닝 기법간 성과

침입탐지를 위한 3 가지 데이터마이닝 기법을 10-fold cross validation 으로 실험한 결과는 <표 2>과 같다. 각 데이터 셋마다 성과가 조금씩 차이를 보이고 있지만 전체 평균으로 보면 3 가지 기법은 성과에 큰 차이가 없는 것으로 보인다. 3 가지 실험에 대해 변화의 유의성을 실제 통계적으로 검증하기 위해 멕네마르 검정(McNemar test)을 수행하였다. 멕네마르 검정은 사전사후(before and after) 형태의 실험에서 변화의 유의성을 검정하는 비모수적 방법이다. 이러한 변화의 유의성을 검정하기 위해 자료는 동일한 개체에서 두 가지 처리를 적용하여 나타나며, 2×2 분할표로 표현이 된다. 침입탐지를 데이터마이닝 기법간의 성과를 검정한 결과는 <표 3>과 같다. 멕네마르 검정의 결과를 살펴보면 모두가 유의수준 0.05 에서 유의하지 않으므로 신경망, 귀납적 학습법, 러프집합의 성과차이는 있다고 할 수 없다.

<표 3> 멕네마르 검정 결과표

	IL	RS
NN	통계량 (S) 0.1111 Pr > S 0.7389	통계량 (S) 0.6923 Pr > S 0.4054
IL	□	통계량 (S) 0.2857 Pr > S 0.5930

4.3 오류분류에 따른 성과변화 분석

일반적으로 데이터마이닝 기법에서 이진 의사결정을 위한 분류기준값(threshold)으로 0.5 를 사용하고 있다. 그러나 이러한 획일적인 분류기준값보다는 각 분류기준값별로 성과 분석을 통해 최대 이익을 가져올 수 있는 특정 분류기준값을 발견할 수 있다. 분류기준값의 변화에 따른 False positive error(FPE)와 False negative error(FNE) 결과의 변화는 <표 4>와 같다.

<표 2> 신경망, 귀납적 학습법, 러프집합에 따른 성과표(오류 %)

모델	데이터셋	S1[1]	S2	S3	S4	S5	S6	S7	S8	S9	S10	Avg.
NN[2]	학습용	13.78	15.11	15.33	15.56	16.22	15.33	14.22	15.56	15.56	15.56	15.38
	검증용	18.00	26.00	6.00	20.00	14.00	20.00	14.00	22.00	22.00	10.00	15.38
IL[3]	학습용	16.00	14.22	17.11	17.33	12.89	16.44	15.56	16.22	15.56	16.89	15.82
	검증용	18.00	26.00	8.00	6.00	18.00	14.00	22.00	16.00	22.00	10.00	16.00
RS[4]	학습용	15.56	14.22	16.00	16.22	14.67	16.22	14.67	15.11	15.33	16.22	15.42
	검증용	14.00	26.00	8.00	6.00	22.00	12.00	20.00	24.00	22.00	10.00	16.40

1) S1 – S10, 10-fold 데이터 셋, 2)신경망, 3)귀납적 학습방법, 4)러프집합

<표 4> 신경망과 러프집합의 성과변화비교

분류기준값	NN			RS			맥네마르(p)
	FPE	FNE	전체	FPE	FNE	전체	
0.3	26.00	6.40	16.20	29.60	5.60	17.60	0.07*
0.4	25.20	6.80	16.00	27.20	6.00	16.60	0.37
0.5	24.40	7.20	15.80	26.80	6.00	16.40	0.41
0.6	22.40	8.40	15.40	26.80	7.20	17.00	0.07*
0.7	20.40	24.00	22.20	22.80	20.40	21.60	0.75

* 유의수준 10%

위 결과에서 귀납적 학습방법은 제외하였다. 그 이유는 귀납적 학습방법의 경우에 예측값들이 0 과 1 에 가깝게 동일하므로 분류기준값을 0.3~07 로 조정하더라도 결과가 모두 똑같이 나타나기 때문이다. 신경망과 러프집합의 분류기준값에 따른 멕네마르 검정 결과를 살펴보면 0.3 과 0.6 이 유의수준 10%에서 통계적으로 유의한 결과를 나타내고 있다. 이것은 신경망과 러프집합의 성과에 차이가 있음을 보여주는 것이다. 분류기준값 0.3 과 0.6 에서 신경망이 러프집합 보다는 False positive error 를 더욱 잘 탐지하고 있기 때문에 False negative error 에서 러프집합에 비해 성과가 조금 낮지만 전체 성과는 높은 것으로 나타나고 있다. 이렇게 분류기준값을 통한 변화를 제공함으로써 보안담당자들은 자사의 환경에 맞추어 False positive error 와 False negative error 에 대한 최적의 비율로 유연하게 침입탐지 전략을 설정할 수 있을 것이다.

V. 결론

국내 특성에 맞는 침입탐지시스템을 개발하기 위하여 국내에서 사용된 네트워크를 통한 침입공격에 관한 데이터를 수집한 후, 신경망, 귀납적 학습방법, 러프집합을 적용하였다. 그러나 일반적으로 데이터마이닝 기법에서 이진 의사결정을 위한 분류기준값으로 0.5 를 사용하고 있기 때문에 이러한 획일적인 분류기준 값보다는 각 분류기준값별로 False positive error 와 False negative error 에 따른 최대 이익을 가져올 수 있는 특정 분류기준값을 발견할 수 있게 된다.

본 연구에서 설계한 지능형 침입탐지시스템은 분류기준값을 통한 변화를 제공함으로써 기업의 보안담당자들이 자사의 환경에 맞추어 False positive error 와 False negative error 에 대한 최적의 비율로 유연하게 침입탐지 전략을 설정할 수 있도록 의사결정을 지원할 것이다.

참고문헌

1. 박기남, 이훈영, 박상국. (2000). 러프집합을 이용한 통합형 채권등급 평가모형 구축에 관한 연구. 한국경영과학회지, 제 25 권, 제 3 호, 125-135.
2. Balajinath, B., Raghavan, S.V. (2001). Intrusion detection through behavior model. Computer Communications, 24, 1202-1212.
3. Barber, R. (2001). The Evolution of Intrusion Detection Systems-The Next Step. Computer&Security, 20, 132-145.
4. Debar, H., Becker, M., & Siboni, D. (1992). A neural network component for an intrusion detection system. IEEE Computer Society Symposium Research in Security and Privacy, 240-250.
5. Fayyad, U.M., Piatesky-Shapiro, G., & Smith, P. (1996). The KDD processes for extracting useful knowledge and learning from volumes of data. Communications of the ACM, 39(11), 27-34.
6. Joo, D., Hong, T., & Han, I. (2003). The neural network models for IDS based on the asymmetric costs of false negative errors and false positive errors. Expert Systems with Applications, 25, 69-75.
7. Lippmann, R. P., Cunningham, R. K. (2000). Improving intrusion detection performance using keyword selection and neural network. Computer Networks, 34, 597-603.
8. Pawlak, Z. (1999). Rough set approach to knowledge-based decision support. European Journal of Operational Research, 48-57.
9. Zue, D., Premkumar, G., Zhang, X., & Hsien Chu. (2001). Data Mining for Network Intrusion Detection : A Comparison of Alternative Methods. Decision Sciences, 32(4), 635-659.