네트워크에서 에이전트 학습을 이용한 침입탐지시스템 모델

정종근* · 김용호** · 이윤배***

Intrusion Detection System Model using agent learning in network

Jong-Geun Jeong* · Yong-Ho Kim** · Yun-Bae Lee***

요 약

분산된 네트워크 환경에서 침입탐지시스템을 구축하는 일은 단일 환경에서보다 매우 복잡하다. 특히, 각기 다른 운영체제시스템에서 발생하는 로그데이터의 수집과 분석 등에서 많은 문제가 발생한다. 따라서 본 논문에서는 이러한 문제를 해결하고자 에이전트 학습 시스템을 적용한 침입탐지시스템 모델을 제시한다. 에이전트 학습을 위한 알고리즘으로는 데이터 마이닝 알고리즘을 적용한다.

ABSTRACT

It is very complex to construct Intrusion Detection System in distributed network environment than simple ones. Especially, In the collecting and analysis of logdata from out different operating system break out much problem. So In this paper, We present a Intrusion Detection System model applying agent learning system to solve these problem. We apply the data Mining algorithm for agent learning.

Ⅰ. 서 론

방화벽은 능동적인 방어 시스템이 아니다. 이에 반해 침입 탐지 시스템(IDS)은 방화벽이 감지하지 못하는 공격에 대해 인식할 수 있으며 더 나아가서는 이전에 경험하지 못한 공격에 대해서도 이를 감지, 방어할 수 있는 능동적인 시스템이다[1]. 기존 방화벽의 결함에 대해 설명할 때 Cold Fusion을 들 수 있다. 침입 탐지 시스템의 핵심 기술은 행위 판별(Behavior Classification)과 자료축약(Data Reduction) 기술이다. 행위 판별은 주어진 일련의 행위들에 대해 침입인지 아닌지를 판별하는 것이고, 자료축약는 시스템에서 발생하는 거대한 양의 각종 로그 데이터(log data)를 의미있는 데이터로 추출하여 변환하는 작업이다. 일반적으로 침입 탐지 규칙에서는 규칙기반 시스템(Rule-based system)과 신경망 또는 통계적 분류 시스템을 사용한다. 기존에 사용된 규칙기반 시스템, 신경망, 통계적 분류 시스템은 많은 양의 데이터가 초기 학습을 위해 필요하며, 계속적으로 시스템을 유지하는데 많은 시간과 비용을 초래하며 새로운 공격 대응 능력이 약하다는 취약점을 가지고 있다. 따라서 본 논문에서는 기존 방법들의 단점을 보완하고자 에이전트 학습시 데이터 마이닝 학습 알고리즘을 적용하여 이러한 문제점들을 해결하고자한다.

*동강대학 전자정보과
**조선대학교 전자계산학과 대학원
***조선대학교 컴퓨터공학부
접수일자 : 2002. 10. 10

II. 기존의 침입탐지 학습 방법

규칙기반 침입 탐지 시스템의 대표적인 예로 IDES 시스템을 들 수 있다[2]. 이 시스템은 대상 시스템의 취약성 및 보안정책 그리고 과거의 침입에 대한 지식을 데이터 베이스에 저장한 후 침입이 발생할 때, 탐지 시스템은 침입에 대한 규칙베이스(Rule base)에 의해 현재 시스템의 침입 여부를 결정한다. 규칙기반 IDS의 경우 초기의 규칙기반을 만들기 위해 보안 분야의 높은 지식을 가진 전문가의 지식이 필요하다. 이는 오랜 시간과 막대한 개발비용을 필요로 하는 작업이다. 게다가 전문가라 할지라도 시스템의 모든 취약성에 대해 알 수 없으며, 기존의 많은 시스템들의 약점들간의 상호작용으로 생겨나는 취약성에 대해서는 발견해 낼 수 없다는 단점이 있다. 만일 시스템의 프로파일에 중요한 변화가 발생한다면, 규칙기반 시스템의 경우 새로운 침입 가능성에 대비하여 규칙기반을 새롭게 설계해야 한다. 이는 오류 발생 가능성이 매우 높은 작업이기 때문에 새로운 규칙이 기존의 규칙들과 충돌할 수도 있다. 이러한 문제들은 시스템 운영자에게 현재 규칙의 운영 및 유지 보수를 어렵게 하여 결과적으로 침입 탐지 시스템이 과거 정보에 의존하게 만든다.

Kumar와 Spaffod에 의해 제안된 패턴 매칭 기법에 기반한 접근법은 시스템상에 요구되는 유연성의 향상에 초점이 맞추어져 있지만 학습능력을 갖추지 못하였다는 단점을 가지고 있다[6,7]. 이들은 시스템상에 나타나는 현상들에 근거하여 침입을 어떻게 분류하는 지를 보여주고 있다. 여기서의 각 패턴들은 시스템 상태들간의 의존도를 인코딩하고 있는 것이다. 이러한 접근법은 침입을 탐지하는 강력한 방법이나 사전에 만들어진 패턴들에 의존적이라는 단점 또한 가지고 있다. 즉, 패턴 자체가 완전하지 못할 경우 시스템의 방어에 커다란 헛점(Hole)이 나타나게 되는 것이다. 그리고 보안 정책이나 시스템 운영상에 변화가 있을 경우 패턴들을 다시 만들어야 한다는 문제점이 있다.

현재까지 제시된 침입탐지 시스템들은 몇 가지 문제점들을 공통적으로 가지고 있는데 이 중 가장 두드러진 문제점은 시스템 부하에 관한 것이다. 이를 해결하기 위해 별도의 침입탐지모듈에 의해 네트워크 전체가 분석되도록 하고 있다. 감사 흔적(Audit Trail)을 분석하기 위해서는 시스템 커널이 시스템상에서 이루어지는 모든 행동들에 대해 감사 정보를 만들어 내야 하는 데, 그 양이 엄청나며 분석 작업에는 시스템의 디스크 용량이나 CPU Time의 엄청난 소모가 필요하다. 실제적으로 영국의 University College London(UCL)에서도 침입탐지 시스템을 개발하기 위해 기존의 신경망 기법과 유전자 알고리즘, 그리고 전문가 시스템 기술을 이용하였으나 규모문제(Scale)에 부딪쳐 사업이 중단되었다. 이는 소규모 시제품 시스템에서는 인공지능 기법이나 분류기법, 유전자 알고리즘이 효과적으로 적용되지만 실제 네트워크 시스템에서는 그 규모 문제 때문에 운영이 어려워지기 때문이다.

III. 데이터 마이닝을 이용한 감사데이터 학습

3.1 연관규칙을 이용한 감사데이터 분류

데이터 마이닝의 연관규칙 탐사 알고리즘 중 가장 대표적인 방법이 Apriori 알고리즘이다. Apriori 알고리즘은 여러 논문에서 연구되어 다양한 분야에 응용되고 있다. Apriori 알고리즘은 데이터베이스에서 후보 항목 집합을 구성하고, 구성된 후보 항목 집합에서 빈발 항목 집합을 탐사하는 과정으로 수행된다. Apriori 알고리즘은 후보 항목 생성시 모든 데이터베이스에서의 데이터 항목에 대한 생성이 아닌, 전 단계의 빈발 항목 집합을 대상으로 후보 항목을 생성한다. Apriori 알고리즘은 전 단계에서의 빈발 항목 집합에서 현재 단계의 후보 항목 집합을 구성 한 다음 데이터베이스의 스캔을 통해 후보 항목 집합의 지지도를 계산한다. 그리고 사용자가 정의한 최소 지지도를 기초로 하여 현재 단계의 빈발 항목 집합을 구성한다. Aprirori 알고리즘의 단계의 진행은 데이터 항목의 증가에 따라 반복적으로 진행된다. k단계에서의 Apriori의 빈발 항목 탐사는 k-1 단계의 빈발 항목 집합으로부터 생성된 k-후보 항목 집합에 대하여 각각의 지지도를 계산한 후 이들 중에서 지지도를 만족하는 항목의 탐사를 통해 이루어진다. Apriori는 더 이상의 후보 항목을 생성할 수 없을 때까지 반복되어 빈발 항목을 탐사하며, 빈발항목 집합

의 생성 알고리즘은 [그림 1]과 같다. [그림 2]의 알고리즘과 같이 후보 항목 집합의 생성은 전 단계의 빈발 항목 집합의 조인 연산(Join operation)과 전지 과정(Prune process)을 통해 이루어진다. 조인 연산은 두 집합의 곱집합을 구하는 것과 같으며 전지 과정은 조인을 통해 생성된 후보 항목 집합의 부분 집합이 전 단계의 빈발 항목 집합의 원소가 아닌 경우, 그 항목을 삭제하는 과정이다.

```
L1 = {large 1-itemsets}
for (k=2;Lk-1≠Ø;k++) do begin
  Ck = apriori-gen(Lk-1); //새로운 후보항목 집합
   forall transactions t∈D do begin
  Ct = subset(Ck,t); //후보항목이 빈발항목집합에 포함
        forall candidates c∈Ct do
          c.count++;
   end
Lk = {c∈Ck | c.count ≥ Smin}//최소지지도를 만족
end
Answer = ∪Lk;
```

그림 1. 빈발 항목 집합 생성 알고리즘

Fig. 1. Large item sets generation algorithm

그 이유는 전 단계에서 빈발하지 못하는 항목은 다음 단계에서도 빈발하지 못하기 때문이다. 전지 과정은 불필요한 후보 항목의 수를 줄여 데이터베이스를 읽는 횟수를 감소시키기 위하여 추가된 과정이다.

```
Algorithm Apriori-gen
insert into Ck  // 필요한 항목 추가
select a.item1, a.item2, ···,.itemk-1, b.itemk-1
from Lk-1a,  Lk-1b
where  a.item1 = b.item1, ···,a.itemk-2;
                = b.itemk-2, a.itemk-1 < b.itemk-1
// 생성된 항목이 전단계의 빈발항목원소가 아닌 경우 삭제
for all itemset c∈Ck do
  for all (k-1)-subsets s of c do
  if (s∉Lk-1) then
     delete c from Ck
```

그림 2. 조인연산과 전지과정의 알고리즘

Fig. 2. Algorithm of Join operation and Prune operation

3.2 침입패턴 분류

비정상 탐지를 위한 기계학습 방법의 어려움은 알려져 있지 않은 패턴과 알려진 패턴의 한계를 정하는 것이다. 학습 데이터에 있어서 비정상 패턴에 대한 별다른 예를 가지고 있지 않은 상태에서는 기계학습 알고리즘은 훈련 데이터에 있는 알려진 패턴에 대한 한계를 구분할 수 없다[5,6]. 일반적으로 비정상과 오용 탐지를 구분하기란 쉬운 일이 아니다. 비정상 탐지는 전형적으로 비통제된 학습 방법을 사용하는 반면에 오용 탐지에서는 통제된 분류 방법을 사용한다[2,3,7].

따라서 변형된 패턴의 공격이 발생할 경우 이를 탐지해 내지 못하므로 변형된 새로운 유형의 공격이 발생할 경우, 이 공격 패턴을 즉시 학습시킴으로써 새로운 공격에 대응하고자 한다. 이를 위해 새로운 공격 패턴이 발생할 경우 이미 분류되어 있는 침입 패턴 집합에 계속적으로 추가시킨다. [그림 3]의 알고리즘에서와 같이 H2는 새로운 침입 패턴과 정상 데이터로부터 학습된 추가된 분류자이며 알고리즘에서 결정 규칙은 출력을 위해서 평가된다. H1은 존재하는 침입 탐지 시스템 모델이고 H2는 최근에 발견된 새로운 침입 패턴을 위해 훈련된 새로운 모델이다. H1에서는 정상과 비정상 패턴만을 확인하고 새로운 침입을 확인할 수 없기 때문에 대부분의 패턴들은 비정상과 오용으로 분류한다. 그러나 H2는 새로운 침입과 정상 데이터로 분류한다. 이때 새로운 침입 패턴의 양이 적기 때문에 H2는 다른 데이터로부터 침입 패턴을 쉽게 분류할 수 있다.

```
Intrusion_learning()
{
if (H1(x)=normal) ∨ (H1(x)=anomaly) then
  //정상패턴과 비정상 패턴 분류
  if H2(x)=normal
    then output ← H1(x)(normal or anomaly)
  //존재하는 침입 패턴 모델
    else output ← new_intrusion
else output ← H1(x)
}
```

그림 3. 침입 탐지 분류 알고리즘

Fig. 3. Algorithm for Intrusion Detection Classification

3.3 에이전트 학습

본 논문에서는 이러한 에이전트 시스템의 특성에 기존의 학습 방법 대신 데이터 마이닝 학습 방법을 제안하였다. 데이터 마이닝 학습 방법과 기계 학습은 과거의 경험을 바탕으로 학습한다는 점은 동일하지만 다양한 패턴으로의 변화와 새로운 패턴의 추측 면에서는 데이터 마이닝 방법이 훨씬 뛰어나다. 따라서 기능은 기존의 에이전트 시스템과 동일하지만 학습 방법에서는 데이터 마이닝 방법을 채택하였다.

에이전트에 대한 학습 모듈은 [그림 4]와 같다.

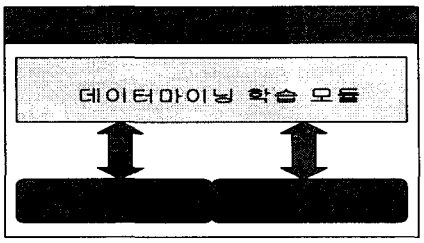

그림 4. 에이전트에서의 학습모듈
Fig. 4. Learning module in the agent

[그림 4]에서처럼 기존의 에이전트 엔진과 통신모듈로 구성되어있고 데이터마이닝 학습 모듈을 에이전트에 탑재하였다. 에이전트는 수집하여 변환된 감사데이터를 학습모듈에서 학습시켜 직접 탐지에 이용한다. 침입 탐지시 발생하는 긍정적 결함을 최소화하기 위해 임계값(threshold)을 학습단계에서 조정하여 시스템의 유연성을 크게 하였다. 다음 [그림 5]는 에이전트에 감사데이터 학습 결과를 보여주고 있다.

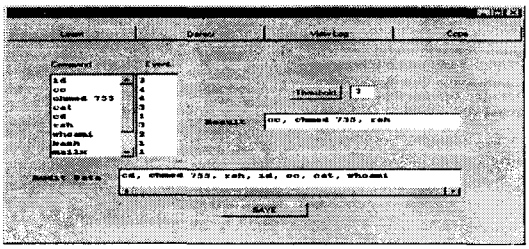

그림 5. 감사데이터 학습화면
Fig. 5. Figure of Audit Data learning stage

IV. 구현 및 성능 평가

4.1 침입 탐지 시스템 구현

각 호스트에서 수집된 표준화된 로그 데이터는 이미 침입 탐지 시스템 데이터 베이스에 저장되어 있는 침입 패턴과의 매칭을 통해 침입을 판단한다. 이때 데이터베이스에 저장되어 있는 감사 데이터는 지속적인 학습을 통해 새로운 유형의 침입 패턴을 계속 갱신(update)한다. 또한 미리 정의해 놓은 규칙들로부터 변형된 침입 패턴을 예측·생성한다.

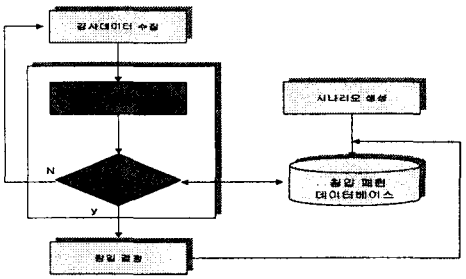

그림 6. 탐지 모듈 구성도
Fig. 6. Strusture of detection module

[그림 6]은 침입 탐지 모듈에 대한 구성도이고, [그림 7]은 침입 탐지 알고리즘을 나타낸다. 에이전트에서 수집한 사용자의 로그 데이터와 침입 패턴(PT)과 비교하여 기대치 이상이거나 일치하는 침입 유형을 찾아서 일치하면 침입 상태를 보고한다. 시나리오에는 없지만 새로운 침입 패턴이라고 판명 될 때는 감사 데이터 DB에 저장한다.

```
Intrusion_Detection() {
 event=Associate;
    forall intrusion p ∈ PT
 {
    CMP = compare(PT(p), event);
    if (CMP > MIN(E)) then report=CMP
 }
 warning = Intrusion_detection(report)
 store = new pattern }
```

그림 7. 침입 탐지 알고리즘
Fig. 7. Intrusion detection algorithm

[그림 8]은 에이전트에서 수집한 사용자들의 로그 데이터와 침입 패턴 감사데이터를 비교하여 패턴의 일치여부를 백분율(%)로 표시하였다. 이 탐지 화면에는 기본적으로 사용자들의 IP주소, 계정, 시간, 공격 호스트 IP 등이 표시되어 관리자가 쉽게 확인할 수 있다.

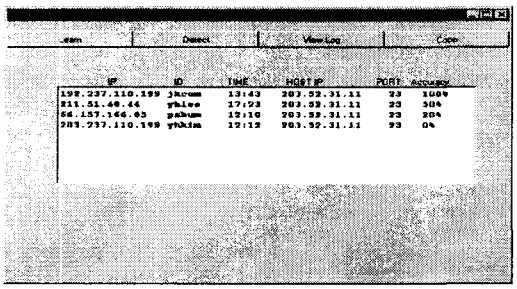

그림 8. 사용자별 침입 상황

Fig. 8. Status of individual Intrusion

4.2 성능 평가

침입 탐지 시스템의 핵심이 탐지의 정확도와 높은 탐지율이라면 가장 큰 문제점은 탐지 오판율을 최소화시키는 일이다. 침입 탐지 오판의 대부분은 긍정적 결함(false positive)과 부정적 결함(false negative)으로써, 이와같은 결함들을 최소화시키는 것이 오판율을 줄이는 것이다.

본 논문에서는 긍정적 결함을 최소화하는 것에 초점을 두었다. 긍정적 결함의 발생원인은 침입 패턴을 감사 데이터화하는 과정에서 침입 패턴에 대한 감사 데이터 범위를 결정하는 과정에서 발생한다. 이를 해결하기 위해서 본 논문에서는 감사 데이터를 학습하는 과정에서 데이터 마이닝 기법을 적용하여 하나의 침입 패턴에서 발생할 수 있는 여러 가지 변형 형태에 대한 예측 학습이 가능하도록 하여 긍정적 결함의 발생을 최소화하였다. [그림 9]는 임계값(threshold)이 증가함에 따라 긍정적 결함이 감소하는 실험 결과를 보여주고 있다. [그림 9]에서와 같이 임계값을 증가시킬수록 긍정적 결함의 비율이 작아지는 것들 볼 수 있다. 하지만 9 이상의 임계값을 주었을 때는 긍정적 결함의 비율의 오차가 별 차이가 나타나지 않는다.

이것은 본 침입 탐지 시스템에서도 완벽하게 긍정적 결함을 없앨 수는 없다는 것이다. 그 이유는 감사 데이터 학습과정에서 정상인 침입 패턴의 일부가 학습되어지기 때문에 완전하게 긍정적 결함을 제거하는 것은 불가능하였다.

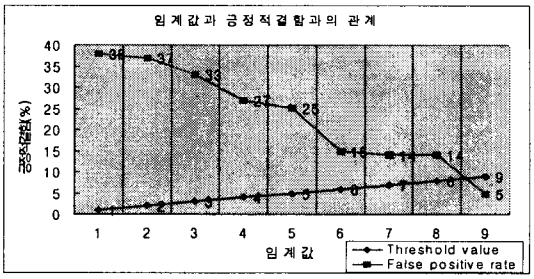

그림 9. 임계값과 긍정적 결함과의 관계

Fig. 9. Relationship threshold and false positive

V. 결론 및 향후 연구 방향

본 논문에서는 침입 탐지 시스템에 데이터마이닝 학습 기법을 도입하여 다량의 데이터 축약과 변형된 침입 패턴을 탐지할 수 있게 하였다. 또한 독립적으로 침입을 탐지할 수 있는 에이전트를 채용함으로써 분산환경에서 적합하도록 설계하였다. 에이전트는 중앙의 침입 탐지 호스트와 계속 통신하여 새로운 감사 데이터를 제공받으며 에이전트가 수집한 새로운 감사 데이터는 침입 탐지 호스트나 다른 에이전트와 서로 정보를 공유하게 된다. 특히 본 논문에서 제안한 시스템과 현재 사용되고 있는 다른 침입 탐지 방법들과 비교할 때, 탐지의 정확도를 높였고, 오판율을 줄이기 위해 임계값을 상황에 따라 조절하여 긍정적 결함을 최소화하였다. 본 시스템은 어떤 감사 데이터를 학습시키냐에 따라서 침입 탐지 범위가 결정된다. 따라서 다양한 감사 데이터 학습이나 감사 데이터 양에 따라 탐지 능력을 향상시킬 수 있다. 향후 연구 방향으로는 본 논문에서는 감사 데이터 학습 단계를 오프라인(offline)으로 처리하여 전체적인 시스템의 부하를 최소화하였으나, 온라인(online) 상태에서 수행하여 자동화된 침입 탐지 시스템을 구축하는 연구가 필요하다. 또한 감사 데이터 학습과정에서 최소 임계값을 결정하는 문제가 크게 대두되었다.

임계값을 크게 하면 수집된 데이터들에서 정확한 감사집합을 구하지 못해 부정적 결함(False negative)이 발생할 수 있다. 따라서 감사 데이터 학습시 적절한 임계값 설정에 대한 연구가 필요하다.

참고문헌

[1] R. Buschkes, M. Borning, and D. Kesdogan, "Transaction based Anomaly Detection" Proc. of the Workshop on Intrusion Detection and Network monitoring, USENIX, Apr., 1999.

[2] Anup K. Ghosh, "Learning Program Behavior Profiles for Intrusion Detection", Proc. of the Workshop on Intrusion Detection and Network Monitoring, April., 1999.

[3] Samuel I. Schaen, "Network Auditing: Issues and Recommendations", IEEE 7th Computer Security Applications Conference, pp.66-79, Dec., 1991.

[4] T. Lane, "Filtering technique for rapid user classification", In Proceedings of the AAAI98/ ICML98 Joint Workshop on AI Approaches to Time series Analysis, 1998.

[5] U. Fayyad, G. Piatetsky-Shapiro and P. Smyth, "The KDD process of extracting useful knowledge from volumes of data", Communications of the ACM, 39(11):27-34, Nov., 1996.

[6] W. Lee, S. J. Stolfo and K. W. Mok, "Mining Audit data to build Intrusion Detection Models", In proceeding of the 4th International Conference on Knowledge Discovery and Data Mining, New York, NY, Aug., 1998.

[7] 정종근, 이윤배, "새로운 침입 패턴을 위한 데이터마이닝 침입탐지시스템 설계", 대한전자공학회 논문지, 제39권 TE편 제1호, pp.77-87, 3, 2002.

[8] 한국전자통신연구원, "인터넷보안 기술/시장보고서", 12, 2001.

저자소개

정종근(Jong-Geun Jeong)

1995년 조선대학교 전자계산학과 졸업(이학사)

1997년 조선대학교 대학원 전자계산학과 졸업(이학석사)

2002년 8월 조선대학교 대학원 전자계산학과 졸업(이학박사)

1999년 3월~2002년 현재 동강대학 전자정보과 겸임교수

※관심분야: 인공지능, 검색엔진, 데이터베이스, 정보보안, 전자상거래, 바이러스

김용호(Yong-Ho Kim)

1989년 광주대학교 전자계산과 졸업(공학사)

1993년 경남대학교 전자계산과 대학원(공학석사)

1999년~2002년 현재 조선대학교 대학원 전자계산학과 박사수료

※관심분야: 멀티미디어, 전자상거래, 검색엔진

이윤배(Yun-Bae Lee)

1980년 광운대학교 전자계산학과 졸업(이학사)

1983년 광운대학교 대학원 전자계산학과 졸업(이학석사)

1993년 숭실대학교 대학원 전자계산학과 졸업(공학박사)

1988년 4월~현재 조선대학교 컴퓨터공학부 교수

1999년 7월~2000년 현재 광주광역시 시정정책자문회의 위원

※관심분야: 인공지능, 전문가시스템, 멀티미디어, 데이터베이스, 정보보안, 바이러스