제 18 회 한국정보처리학회 추계학술발표대회 논문집 제 9 권 제 2 호 (2002.11)

네트워크기반 비정상행위 탐지모델 생성을 위한 비감독 학습 알고리즘 비교분석

이효승*, 심철준**, 원일용**,이창훈**
*㈜사과의 꿈
**건국대학교 컴퓨터공학과
e-mail : *lhs@doa21.com
**clcc,chlee@konkuk.ac.kr

Comparative Analysis of Unsupervised Learning Algorithm for Generating Network based Anomaly Behaviors Detection Model

Hyo-Seong Lee*, Chul-Jun Sim**, Il-Yong Won**, Chang-Hun Lee**
*Dream of Apple, Co., Ltd
**Dept. of Computer Engineering, Kon-Kuk University

요 약

네트워크 기반 침입탐지시스템은 연속적으로 발생하는 패킷의 무손실 축소와, 패킷으로 정상 또는 비정상 행위패턴을 정확히 모델링한 모델 생성이 전체성능을 판단하는 중요한 요소가 된다. 네트워크 기반 비정상행위 판정 침입탐지시스템에서는 이러한 탐지모델 구축을 위해 주로 감독학습 알고리즘을 사용한다. 본 논문은 탐지모델 구축에 사용하는 감독 학습 방식이 가지는 문제점을 지적하고, 그에 대한 대안으로 비감독 학습방식의 학습알고리즘을 제안한다. 감독 학습을 사용하여 탐지모델을 구축하기위해서는 정상행위의 패킷을 취합해야 하는 사전 부담이 있는 반면에 비감독 학습을 사용하게 되면 이러한 사전작업 없이 탐지모델을 구축할 수 있다. 본 논문에서는 비감독학습 알고리즘을 비교 분석하기 위해서 COBWEB, k-means, Autoclass 알고리즘을 사용했으며, 성능을 평가하기 위해서 비정상행위도(Abnormal Behavior Level)를 계산하여 에러율을 구하였다.

1. 서론

침입탐지시스템(IDS: Intrusion Detection System)이란 인터넷이나 인트라넷 상에서 악의적인 사용자가 특정 시스템에 대해서 크래킹(Cracking)이나 비정상행위가 이루어지는 것을 감지하고 이에 대한 처리를 행하는 시스템이다.

비정상행위판정 침입 탐지시스템은 정상행위 학습 데이터를 바탕으로 정상행위에 관한 지식을 생성하여 탐지 모델을 구축하므로 학습 데이터의 신뢰도가 아주 중요하다. 신뢰할 수 있는 정상적인 행위만으로 이루어진 데이터를 얻기 위해서는 데이터의 검증 작업과 순수한 데이터만을 추출하는 사전작업이 필요하다. 이것은 시스템 구축에 추가적인 작업이 될 수 있으며, 데이터의 확실한 신뢰성도 보장받을 수 없는 부분이 된다. 따라서 이러한 불필요한 부분을 줄일 수 있는 연구가 필요하다고 본다.

본 논문의 구성은 다음과 같다. 2 장에서는 연구 동기와 실험에 사용되는 비감독 학습 알고리즘에 대해서 알아본다. 3 장에서는 학습 데이터의 가공과 실험 방법에 대하여 기술하고있다. 4 장에서는 학습 데이터를 바탕으로 탐지모델을 구축하고 성능을 비교한 결과를 제시한다. 5 장에서는 4 장을 바탕으로 본 논문의 결론과 향후과제에 대해서 기술한다.

2. 동기 및 관련 연구

2.1 동기

네트워크기반의 비정상행위판정 침입탐지시스템은 네트워크 상에서 정상적 행위에 의해서 발생하는 패킷을 분석하여 정상행위에 관한 모델을 생성하고, 새로이 발생하는 패킷을 이 모델과 비교하여 그 패킷이 모델에 얼마나 위배되는지를 알려주는 시스템이다. 이러한 탐지모델을 구축하는 과정은 우선 정상패킷을 수집하여, 일정 시간단위로 축소 가공하고 적당한 학습알고리즘을 적용하여 정상행위에 관한 지식을 생성하거나 규칙을 만드는 것이다[1].

그러나, 탐지모델을 구축하는데 감독학습 알고리즘을 사용하면, 실제 네트워크 환경에서 발생하는 패킷을 대상으로 정상과 비정상의 분류작업을 해야 하는데 이것은, 수집된 데이터를 정제처리 하는 추가작업을 발생시키고, 또는 생성된 탐지모델의 부정확성을 야기시키는 원인이 된다.

본 논문에서는 감독학습의 이러한 문제점을 개선하기위해서 탐지모델의 구축에 비감독 학습의 기계학습(Machine Learning)을 이용할 것을 제안함과 동시에 실제로 비감독학습 알고리즘을 사용할 때 각 알고리즘별 성능을 비교한다. 비감독 학습을 이용하게 되면 실 환경에서 수집되는 데이터를 사전에 정상과 비정상으로 분류하는 작업을 거치지않고 그대로 이용할 수 있다.

2.2 비감독 학습 알고리즘

본 논문은 여러가지 비감독학습 알고리즘을 비정상행위 침입탐지시스템의 탐지모델 구축에 사용하여, 알고리즘별 성능을 비교하고, 비감독학습 알고리즘의 가능성을 제시하는데에 목표를 두고 있다. 비감독 학습은 데이터를 사전에 분류하지않기 때문에 침입탐지시스템 구축에 부담을 줄일 수 있다. 본 논문에서 실험을 한 비감독 학습 알고리즘은 <표 1>에 요약되어 있다.

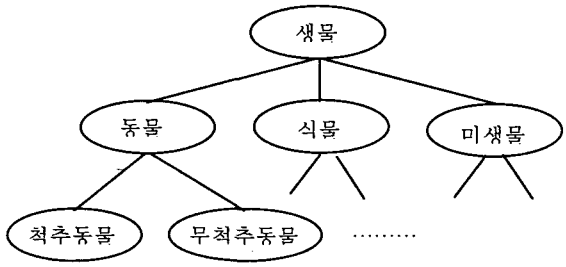
알고리즘	분류	Evaluation Funcion	선정 이유
COBWEB	◦ 점진적 개념학습 ◦ 정량적 분석 ◦ 정성적 분석	◦ Category Utility Function	◦ 클러스터 정보를 확률이나 통계값으로 저장하고 있음
K-means	◦ 기하학적 학습 ◦ 정량적 분석	◦ N 차원상의 기하학적 거리	◦ 학습 속도 빠름 ◦ 구현 용이
AutoClass	◦ 확률기반 학습 ◦ 정량적 분석 ◦ 정성적 분석	◦ Finite Mixture ◦ Bayesian 통계	◦ 수치데이터에 용이 ◦ 다양한 학습 모델 지원

<표 1> 분류별 비감독 학습 알고리즘

2.2.1 COBWEB

COBWEB 의 분류와 학습에 관한 접근은 개념 군집화(Conceptual Clustering : Michalski & Stepp,1983 ; Fisher&Langley,1986)로서 알려져 있다[2][3]. COBWEB 은 인간이 사물을 관찰하고, 사물들(things)사이에 공유되어지는 추상적 특징들을 기반으로 사물의 개념을 형성하고 분류하는 점진적 개념 형성(Incremental Concept Formation)을 모델로 하여 개발되었다. 이것의 쉬운 예가 <그림 1>에 나타나있다.

<그림 1>과 같은 클러스터링은 인간이 생물을 관찰하는 역사적인 과정을 통해 이루어진 것이다. 사물을 하향 분류하기위해서는 그 과정에서 필연적으로 분류



<그림 1> 점진적 개념 형성의 예(생물의 분류)

기준이 형성되는데 이것이 바로 부류의 추상화 된 개념정보 이다. 새로운 사물이 관찰되면 자연스럽게 이전에 형성된 분류 기준에 의해서 기존의 부류에 포함시키거나 새로운 부류를 생성하게 된다. 이러한 분류 방법을 알고리즘화한 것이 COBWEB 이다.

2.2.2 *k*-means 클러스터링

k-means 알고리즘은 전통적인 클러스터링 알고리즘으로 그 자체는 매우 단순하게 이해할 수 있다. 다음은 알고리즘의 단계이다.

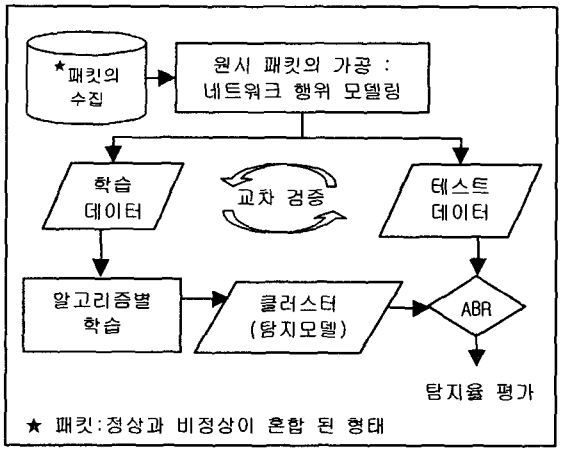
- 1 단계 : 클러스터의 수 k 를 정한다..
- 2 단계 : 초기 k 개 클러스터의 중심점을 임의로 선택한다.
- 3 단계 : 각 학습 레코드를 그 중심과 가장 가까운 거리에 있는 클러스터에 할당한다.
- 4 단계 : 각 클러스터별로 그에 속하는 학습 레코드를 이용해 새로운 중심점을 계산한다.
- 5 단계 : 3, 4 단계를 기존의 중심과 새로운 중심의 차이가 없을 때까지 반복한다.

2.2.3 AutoClass

베이지안 접근방법[4]에서 클러스터는 객체의 속성에 대한 확률분포로 표현되는데, 특히 model function 과 그것들의 파라미터들로 표현된다. 이러한 접근법은 finite mixtures 이론을 그 기반으로 하고 있다.

3. 실험 방법

실험 절차는 우선 패킷을 가공처리한 Event 를 생성하고, 제시한 알고리즘으로 탐지모델을 구축하고 성능평가 및 비교를 한다. 알고리즘별 학습 결과의 성능평가 판정 기준으로는 비정상행위도(ABR : Abnormal Behavior Ratio)를 만들어 계산하였다. 교차검증이란 기계학습 알고리즘의 성능을 평가하는데 사용하는 실험 절차상의 전통적인 방법 중의 하나이다.



<그림 2> 실험방법 절차도

3.1 Event

Event 는 네트워크 패킷을 침입탐지시스템에 적합하도록 가공 처리한 결과이다. Event 생성에서 고려해야 할 사항은, 복수개의 패킷을 일정 시간 단위로 하나의 레코드로 축약 시키는 것과, 레코드로 전환하는 과정에서 패킷에 담겨진 네트워크 정보를 요약할 탐지항목을 설정하는 것이다.[1] 본 논문에서 사용한 학습 데이터는 침입탐지시스템개발 및 평가를 위해 DARPA[5] 산하 MIT Lincoln Lab.에서 제공하는 데이터를 사용한다. DARPA 는 1998 년부터 2000 년까지, 매년 약 7 주간의 tcpdump 데이터를 제공하고있다. 본 논문은 1998 년도 7 주간의 training 데이터 중 3,4,5,6 주차의 목요일과 금요일의 총 8 일간의 tcpdump 데이터를 사용하였다.

탐지항목의 설정은 침입탐지시스템의 탐지 영역을 결정하고 탐지성능에도 영향을 미치는 중요한 요소이다. 본 논문은 전문가의 경험을 바탕으로[6][7]으로 <표 2>와 같은 탐지항목을 만들어 Event 를 생성했다. 이 탐지항목은 패킷의 헤더를 통해서 구할 수 있는 통계적 요소들이다.

분류	탐지 항목 필드
IP	-정상정도 : 비정상 0<= 정상정도 <=1 정상 -수집 단위 시간당 IP 총 패킷 수 -수집 단위 시간당 TCP 총 패킷 수 -수집 단위 시간당 UDP 총 패킷 수
	-수집 단위 시간당 ICmP 총 패킷 수 -패킷 데이터 사이즈 -인바운드 패킷데이터 사이즈/패킷데이터 사이즈 -아웃바운드패킷데이터 사이즈/패킷데이터 사이즈
TCP	-인바운드 TCP 패킷 수 -아웃바운드 TCP 패킷 수 -인바운드 TCP 패킷수/수집단위시간당 TCP 총패킷 수 -아웃바운드 TCP 패킷 수/수집단위시간당 TCP 총패킷 수 -인바운드 SYN 패킷수/수집단위시간당 TCP 총패킷 수 -아웃바운드 SYNACK 패킷 수/수집단위시간당 TCP 총패킷 수 -아웃바운드 ACK 패킷 수/수집단위시간당 TCP 총패킷 수 -인바운드 FIN 패킷 수/수집 단위 시간당 TCP 총 패킷 수 -아웃바운드 FIN 패킷 수/수집 단위 시간당 TCP 총 패킷 수 -인바운드 RESET 패킷 수/수집 단위 시간당 TCP 총 패킷 수
UDP	-인바운드 UDP 패킷 수 -아웃바운드 UDP 패킷 수 -인바운드 UDP 패킷 수/수집 단위 시간당 UDP 총 패킷 수 -아웃바운드 UDP 패킷 수/수집 단위 시간당 UDP 총 패킷 수
ICMP	-인바운드 ICMP 패킷 수 -아웃바운드 ICMP 패킷 수 -인바운드 ICMP 패킷 수/수집 단위 시간당 ICmP 총 패킷 수 -아웃바운드 ICMP 패킷 수/수집 단위 시간당 ICmP 총 패킷 수 -인바운드 ICMP Destination unreachible 패킷 수/수집 단위 시간당 ICmP 총 패킷 수 -아웃바운드 ICMP Destination unreachible 패킷 수/수집 단위 시간당 ICmP 총 패킷 수 -아웃바운드 ICMP ICMP Source Quench packets/수집 단위 시간당 ICmP 총 패킷 수 -인바운드 ICMP 이고 Echo Request packets/수집 단위 시간당 ICmP 총 패킷 수

<표 2> Event 항목

정상정도 필드는 Event 레코드에 포함된 정상 패킷의 비율로, 1 에 가까울수록 정상 패킷이 많은 것이다. 본 논문은 위에서 언급한 비정상행위도를 구하기 위해서 DARPA 에서 제공하는 tcpdump 의 리스트 파일을 이용하여 정상 패킷과 비정상 패킷을 구분하였다.

3.2 알고리즘별 탐지모델의 생성 및 비정상행위도

생성된 Event 를 대상으로 알고리즘별로 탐지모델을 구축한다. 각 알고리즘은 구현되어져있는 표준 툴에 비정상행위도를 계산하기위한 모듈을 추가하였다. 따라서 생성된 모델의 각 클러스터에는 누적된 정상 비율 값이 있다. 이 값은 클러스터의 비정상행위도를 계산하는데 사용된다. ABR 값은 각 알고리즘으로 침입탐지시스템을 구축하여 운영한다면 침입 판정의 기준이 될 것이다. 각 클러스터의 ABR 을 구하는 계산식은 다음과 같다

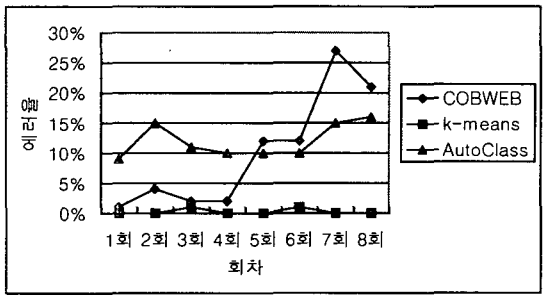
ABR = 1 - (정상비율누적합/클러스터의 총 Event 수)

정상비율누적합은 Event 의 정상정도 필드값의 누적합이고, 클러스터의 총 Event 수는 클러스터에 포함된 Event 의 수이다. ABR 은 다음과 같은 값의 범위를 가지게 된다.

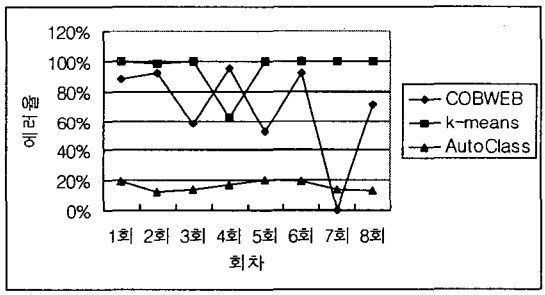
$$0 <= ABR <= 1$$

ABR 이 0 에 가까울수록 그 클러스터는 정상 Event 가 많이 포함된 클러스터이고 1 에 가까울수록 비정상 Event 가 많이 포함된 것이다.

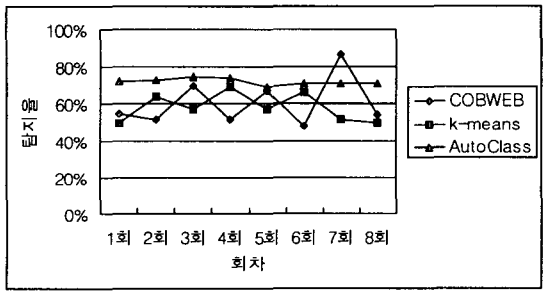
4. 결과 및 비교분석



<그림 3> 긍정 오류율 (False Positive)



<그림 4> 부정 오류율(False Negative)



<그림 5> 알고리즘별 전체 탐지율

총 학습 Event 1280 건의 데이터를 가지고 8 회 교차검증을 실행한 결과 1120 건이 학습 되었고 160 건이 테스트에 사용되었다 .

COBWEB 에서 생성된 유효한 클러스터의 수는 평균 50 개정도 였다. k-means 실험에서 클러스터의 수는 10 개를 주었다. K-means 실험에서는 ABR 을 적용한 긍정오류(false positive)와 부정오류(false negative)이 극단적 결과를 가져왔다. K-means 는 알고리즘 특성상, 학습 Event 의 필드간의 관계가 고려되지 않고 각 필드 값들을 이용하여 클러스터 중심과의 거리만을 평가하는 정량적 분석만이 적용되므로 필연적으로 이와 같은 결과가 나온 것이라 결론 지었다. AutoClass 실험 결과 생성된 클러스터 수는 8 회차 평균 30 개에서 40 개였다.

5. 결론 및 향후 과제

본 논문에서는 네트워크 기반 비정상행위판정 침입탐지시스템에서 비감독 학습 알고리즘을 사용하여 탐지모델을 구축하는 것을 제안하고 알고리즘별 성능을 비교하였다. 제안에 따르면 사전에 데이터를 분류하지 않고 그대로 사용할 수 있어 추가부담이 줄어들게 된다. 그러나 실험결과에서처럼 탐지 성능이 좋지가 않다. 이에 대한 원인으로는 다음과 같은 것을 생각할 수 있다.

① Event 생성에 있어서 탐지항목의 설정의 오류
② 비감독 학습 알고리즘 자체의 학습 능력
③ 비정상행위도 계산의 오류

이러한 문제점을 개선하기 위해서 다음과 같은 연구가 필요하다.

① 전문가의 지식을 적극적으로 활용하고, TCP/IP 프로토콜 내부의 통계적 요소 등을 반영한 탐지모델 구축에 효과적인 Event 를 제작해야 한다.
② 학습 능력이 우수하고 우리가 생성한 Event 에 적합하고 학습 능력이 우수한 비감독 학습 알고리즘을 연구하고 이를 적용한 실험
③ 탐지모델의 성능을 평가할 수 있는 추가적 평가요소를 개발해야 한다.

참고문헌

[1] 이효승, "네트워크기반 비정상행위 탐지모델 생성을 위한 비감독 학습 알고리즘 비교분석", 2002, 건국대학교 석사학위 논문

[2] Fisher, D.H., "Knowledge acquisition via incremental conceptual clustering", Doctoral dissertation, Dept. of Information & Computer Science, University of California, Irvine, 1987

[3] Fisher, D.H., "Iterative Optimization and Simplification of Hierarchical Clusterings.", Technical Report CS-95-01, Vanderbilt University, Nashville TN., 1995

[4] Peter Cheesman, James Kelly, Matthew Self, John Stutz, Will Taylor, Don Freeman, "AutoClass: A Bayesian Classification System", NASA Ames Research Center

[5] http://www.ll.mit.edu/IST/ideval/index.html

[6] 성승제, "네트워크 기반 실시간 침입탐지시스템을 위한 감사자료 수집모듈 설계 및 구현", 2000

[7] 한국정보보호진흥원, " 정보통신기반구조 보호기술개발 최종 보고서" , 한국정보보호진흥원 보고서, 2001