규칙 Set을 이용한 효율적인 실시간 침입탐지

추혜연*, 옥지웅* 김응모*
*성균관대학교 컴퓨터공학과
e-mail : haeyean@skku.edu

Efficient real time intrusion detection using a rule set

Hye-Yeon Choo*, Jee-Woong Ok*, Ung-Mo Kim*
*Dept. of Computer Science, Sung-Kyun-Kwan University

요 약

데이터 마이닝은 데이터 속에 숨겨져 있는 의미 있는 패턴을 찾아내는 것이다. 이러한 패턴들을 찾아내는 것은 데이터 마이닝에서 중요한 부분을 차지한다. 그러나 기존의 데이터 마이닝 방법들에 사용되는 데이터는 시간의 흐름에 데이터가 변하지 않는다는 특징을 가지고 있다. 시간의 흐름에 따라 변화하는 데이터의 특성을 고려해볼 때 변하지 않는 데이터에서 패턴을 찾아내는 것은 의미가 없는 일이다. 따라서 실시간으로 변하는 데이터의 특성을 고려하고 더불어 적합한 실시간 침입 탐지 방법이 필요하다. 따라서, 본 연구에서는 시간의 흐름에 따라 변하는 데이터에서 규칙을 발견하여 규칙 Set 을 생성하는 실시간 데이터 마이닝 기법을 이용하여 시간의 흐름에 따라 변하는 데이터에 대한 침입을 감시하기 위해 실시간 침입 탐지 시스템에 적용함으로써 보다 효율적으로 침입을 탐지하기 위한 방법을 제시한다.

1. 서론

최근에는 침입 탐지 시스템(Intrusion Detection System)에 데이터 마이닝 기법을 적용하여 능동적인 침입 탐지 시스템을 구축하고자 하는 연구들이 활발하다. 네트워크상에서 발생하는 다양한 형태의 대량의 데이터를 정확하고 효율적으로 분석하기 위해 설계되고 있는 마이닝 시스템들은 데이터들을 어떻게 구축하여 다룰 것인지에 대한 문제보다는 얼마나 많은 데이터 마이닝 기법을 지원하고 적용할 수 있는지 등의 기법에 초점을 두고 있다. 따라서, 에이전트화, 분산화, 자동화 및 은닉화 되고 있는 최근의 보안공격기법을 탐지하거나 차단하기 위한 방법은 미흡한 실정이다.[7] 기존의 데이터 마이닝 방법들은 마이닝을 시작하기 전에 필요한 데이터들이 모두 갖춰져 있어야만 하고 시간의 변화에 관계 없이 데이터들이 변하지 않는다는 특징을 가지고 있다. 그러나 데이터는 시간이 흐름에 따라 추가되거나 변경될 수도 있을 뿐만 아니라 데이터의 성격도 변할 수 있으므로 시간의 변화에 따른 데이터의 변화에도 초점을 맞춰야 한다.[9] 따라서 실시간 데이터 마이닝이 필요하며 그에 따른 침입 탐지 시스템도 실시간으로 침입을 탐지할 수 있어야 한다. 이러한 실시간 데이터 마이닝이 필요함에 따라 본 연구에서는 시간의 흐름에 따라 변하는 데이터를 대상으로 하는 침입에 대한 효율적인 탐지를 예측하는 방법을 제안한다.

2. 관련 연구

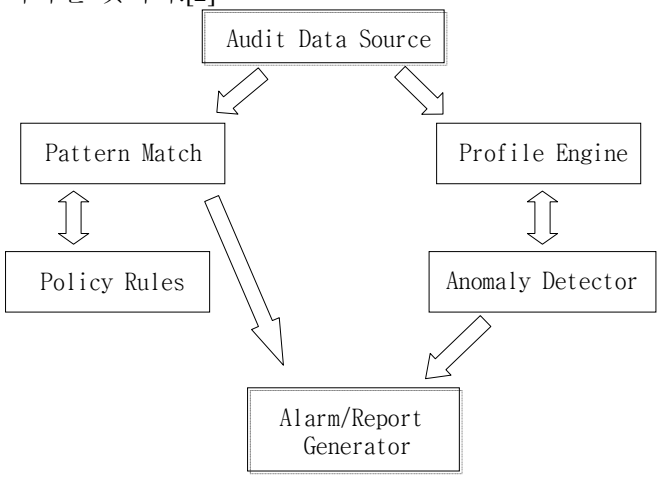
2.1 데이터 마이닝

데이터 마이닝 기법은 대량의 저장된 데이터로부터 의미 있는 패턴을 찾아내는 과정으로서, 연관성 분석을 위해 사용 가능한 다양한 방법론들이 존재한다. 그 중에서 대량의 데이터 내에서 자주 발견되는 연속적 패턴(Sequential pattern)들을 찾아내는데 사용되는 Frequent Episodes 알고리즘은 특정 네트워크에서 탐지된 침입 탐지 정보를 대상으로 적용했을 경우, 그 네트워크에서 비교적 자주 발생되는 침입 패턴을 찾아낼 수 있다. 또한 Association Rules 알고리즘은 다양한 요소들 간에 존재하는 모든 연관성을 분석하여 Rule의 형식으로 표현해주기 때문에, 네트워크 도메인에서 중요도가 높은 요소들을 찾아내는데 적용할 수 있다. [1]

2.2 침입 탐지

침입은 정보 접근, 정보 조작, 시스템 무력화 등 대상 시스템에 대한 고의적이면서도 불법적인 행위로 정의할 수 있으며, 침입 탐지 시스템은 이러한 침입을 목적으로 특정 시스템에 불법으로 접속하여 시스템을 사용, 오용, 남용하는 것을 감지하고 문제점을 처리하는 시스템이라 정의하고 있다. 즉, 침입 탐지 시스템이란 불법적인 침입 행위를 신속하게 감지하고 대응하는 소프트웨어를 말한다.[8] 침입 탐지 시스템들은 한 개의 침입 탐지 프로세스에 의해 침입 탐지를 수행하므로 시스템의 부하는 물론 한 프로세스의 결함이 전체 시스템의 성능을 떨어뜨리는 문제점을 가지

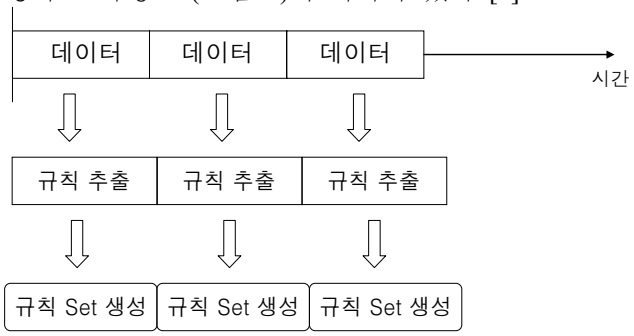
고 있다. 이에 대한 해결책은 다중 에이전트를 이용하여 분산 시스템 전체에서 시스템에 대한 감시와 자료 수집, 탐지 등의 작업을 수행토록 하는 것이다. 다음 (그림 1)은 일반적인 실시간 탐지 시스템의 구조를 나타낸 것이다.[2]



(그림 1) 실시간 탐지 시스템의 구조

3. 연구 모형

본 연구는 시간의 흐름에 따라 변화하는 데이터에 대한 침입 탐지를 하고자 한다. 실시간으로 변화하는 데이터에 대한 침입을 보다 효율적으로 하기 위해 연속적으로 발생하는 데이터에서 규칙을 추출하여 규칙 Set 을 생성한다. 생성된 규칙 Set 단위로 침입 탐지를 한다. 데이터에서 규칙 Set 을 생성하는 과정은 (그림 2)에 나타나 있다. [9]



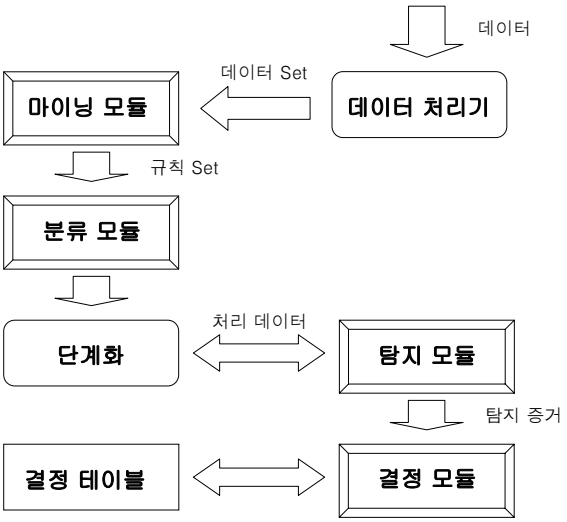
(그림 2) 규칙 Set 추출 과정

(그림 2)는 데이터 Set 으로부터 규칙을 추출하여 규칙 Set 을 생성하는 과정을 나타내고 있다. 이렇게 생성된 규칙 Set 중에서는 상대적으로 중요도가 높은 규칙 Set 과 그렇지 않은 규칙 Set 이 존재한다. 생성된 규칙 Set 들은 중요도에 따라 신뢰도가 부여되고, 이 규칙 Set 들은 침입 탐지에서 신뢰도에 따른 효율적인 침입 탐지를 가능하게 한다. 따라서 중요도가 높은 규칙 Set 을 판단할 필요가 있다. 중요도를 판단하는 방법은 각각의 데이터 Set 에서 추출된 규칙 Set 을 특정 기준과 결합한다.[9] 이렇게 생성된 규칙 Set 에 가중치를 부여한 후 중요도에 따라 중요도가 높은 순으로 단계화한다. 단계화를 하는 과정은 다음에 제시하는 알고리즘을 통해 단계화된다.

```
Algorithm

Database D          //D에는 이전의 침입 Pattern들이 미리 저장

탐지 Pattern A

A' ^ A              //A와 D안의 가장 유사한 패턴 A' 와 비교


Switch    두 패턴의 유사도 K
   Case 1 : 유사도 100 ≥ k ≥90
            A Grade
            break;
   Case 2 : 유사도 90 > k ≥ 85
            B Grade
            break;
   Case 3 : 유사도 85 > k ≥ 80
            C Grade
            break;
   Case 4 : 유사도 80 > k
            break;
```

(그림 3) 단계화 알고리즘

(그림 3)은 규칙 Set 을 단계화하는 과정이다. 데이터베이스에는 이전에 발생한 침입에 대한 패턴들을 저장한다. 기존의 탐지 패턴 A 와 데이터베이스 안의 가장 유사한 패턴 A'와 비교하여 유사도에 따라 등급을 나눈다. 이와 같이 단계화된 규칙 Set 들은 침입 탐지 모듈과 결정 모듈에 통해 침입이 확인되면 단계마다 다른 경고를 받는다. 본 연구에서 제시하는 전체적인 과정은 다음 (그림 3)과 같다.



(그림 3) 규칙 Set 을 이용한 침입 탐지

(그림 3)에서 마이닝 모듈은 데이터 처리기를 통해 들어온 데이터 Set 들에서 규칙을 추출하여 규칙 Set 을 생성하는 작업을 한다. 분류 모듈에서는 마이닝 모듈을 통해 생성된 규칙 Set 들을 중요도가 높은 규칙 Set 과 그렇지 않은 규칙 Set 들을 분류해서 가중치를 부여하는 작업을 한다. 분류 모델을 통해 나뉘어진 규칙 Set 들은 단계화를 거치게 되고 단계화된 규칙 Set 들은 탐지 모듈에서 침입이 확인되면 결정 모듈에서 각 단계에 따라 경고를 받게 된다. 본 연구에서 제안하는 방법을 이용하면 시간의 흐름에 따라 변화하는 데이터에 대해 보다 효율적으로 대응할 수 있다.

또한 중요도에 따라 데이터를 분류하고 침입 탐지를 하게 됨으로써 모든 데이터들에 대한 침입을 탐지하는 것이 아니라 상대적으로 중요도가 높은 데이터에 대한 침입 탐지를 함으로써 침입 탐지에 대한 시간을 줄이고 중요한 데이터를 보호할 수 있는 보다 효율적인 효과를 가져 올 수 있다.

4. 결과

본 연구는 시간의 흐름에 따라 변하는 데이터에 대한 침입을 감시하기 위해 실시간 데이터 마이닝 기법에 실시간 침입 탐지 시스템을 적용함으로써 보다 효율적인 침입 탐지 기법을 제시하였다. 데이터는 시간의 흐름과 함께 갱신되고 변할 수 있다. 그러나 기존의 데이터 마이닝에 사용되었던 데이터들에 대해서는 이러한 점이 고려되지 않았다. 따라서 기존의 데이터 마이닝의 기법을 개선하기 위하여 실시간 데이터 마이닝이 기법이 필요하였다. 실시간 데이터 마이닝을 하기 위해 본 연구에서는 연속적으로 발생하는 데이터에 대해 규칙을 추출하고 이러한 규칙을 기반으로 규칙 Set 을 생성하였다. 이렇게 생성된 규칙 Set 들을 중요도에 따라 단계화하여 침입을 탐지하였다. 결과적으로 본 연구에서 제안하는 방법을 사용하면 데이터의 변화에 따른 침입을 보다 효과적으로 탐지할 수 있다. 모든 데이터에 대한 침입을 탐지한다면 비효율적인 면이 발생한다. 하지만 제안한 방법에서는 데이터를 중요도에 따라 단계화를 함으로써 상대적으로 중요도가 낮은 데이터보다는 중요도가 높은 데이터에 대한 침입을 우선시한다. 따라서 침입 탐지에 대한 시간을 절약 할 수 있고 중요한 데이터를 보호할 수 있는 효과를 가져 올 수 있다.

참고문헌

[1] Wenke Lee,. “A Framework for Constructing Features and Models for Intrusion Detection System” PhD thesis, Columbia University, Jun 1999.

[2] Wenke Lee, Salvatore J.Stolfo, Philip K.Chan, “Real Time Data Mining-based Intrusion Detection”. In proceedings of IEEE symposium on research in security and privacy, 2000.

[3] A. Ghosh and A. Schwartsbard. “A study in using neural networks for anomaly and misuse detection.” In proceecings of the Eighth USENIX security Symposium, 1999.

[4] Ayan, N.F., A.U. Tansel, and M.E. Arkun, “A efficient algorithm to update large item sets with early pruning,” Proceeding of the Fifth CM SIGKDD, Internationl Conference on Knowledge Discovery and Data Mining, 1999, pp.287-291

[5] Cheung, D.W., J. Hand, V. Ng, and C.Y. Wong, “Maintenance of discovered association rules in large databases : An incremental updating technique,” Proceedings of the Twelfth International Conference on Data Engineering, 1996, pp.106-114.

[6] D. Schnackenberg. K. Djahandari. and D.Sterne. Infrastructure for Intrusion Detection and Response. Proceedings of the DARPA Information Survivability Conference and Exposition. SC. Jan. 2000

[7] 최윤정, 박승수. “이상탐지(Anomaly Detection) 및 오용탐지(Misuse Detection)분석의 정확도 향상을 위한 개선된 데이터 마이닝 방법 연구”, 2006 한국컴퓨터종합 학술 대회 논문집 Vol.33, No 1(B)

[8] 김병구, 정태명. “침입탐지 기술의 현황과 전망”, 2000.1 정보과학회지 제 18 권 1 호

[9] 김진화, 민진영. “연속발생 데이터를 위한 실시간 데이터 마이닝 기법”, 한국경영과학회지 제 29 권 제 4 호 2004 년 12 월

[10] 한국정보보호센터, 실시간 네트워크 침입탐지 시스템 개발에 대한 연구, Dec., 1998

[11] 편석범, 정종근, 이윤배, “데이터 마이닝 기법을 적용한 최적 침입 탐지 모듈 설계,”, 1999. 춘계정보 과학회 논문집