한국 퍼지 및 지능시스템 학회 2002, Vol. 12, No. 6, pp. 531-536

자동적인 규칙 기반 방법을 이용한 지능형 침입탐지시스템

The Intelligent Intrusion Detection Systems using Automatic Rule-Based Method

양지홍 · 한명묵
Ji-Hong Yang, Myung-Mook Han

경원대학교 전자계산대학원

요 약

본 연구는 유전자 알고리즘을 IDS에 적용된 오용 탐지 기법을 처음으로 제안하고 구현한 점에서 의미가 있다. 세계적인 대회인 KDD 콘테스트의 데이터를 사용하여 실험하였으며, 그에 따른 가능한 한 같은 환경 하에서 실험을 실시하였다. 실험은 레코드집합을 하나의 유전자로, 즉 하나의 공격패턴으로 간주하고 유전자 알고리즘을 활용하여 진화 시켜 침입 패턴, 즉 침입 규칙(Rules)을 생성한다. 데이터 마이닝 기법중 분류(Classification)에 초점을 맞추어 분석과 실험을 하였다. 이 데이터를 중심으로 침입 패턴을 생성하였다. 즉, 오용탐지(Misuse Detection) 기법을 실험하였으며, 생성된 규칙은 침입데이터를 대표하는 규칙로 비정상 사용자와 정상 사용자를 분류하게 된다. 규칙은 "Time Based Traffic Model", "Host Based Traffic Model", "Content Model" 이 세가지 모듈에서 각각 상이한 침입 규칙을 생성하게 된다. 규칙 생성의 지속적인 업데이트가 힘든 오용탐지 기법에 지속적으로 성장하며 변화해 가는 규칙을 자동적으로 생성하는 시스템으로서, 생성된 규칙은 430M Test data 집합에서 테스트한 결과 평균 약 94.3%의 탐지율을 보였다.

Abstract

In this paper, we have applied Genetic Algorithms(GAs) to Intrusion Detection System(IDS), and then proposed and simulated the misuse detection model firstly. We have implemented with the KDD contest data, and tried to simulated in the same environment. In the experiment, the set of record is regarded as a chromosome, and GAs are used to produce the intrusion patterns. That is, the intrusion rules are generated. We have concentrated on the simulation and analysis of classification among the Data Mining techniques and then the intrusion patterns are produced. The generated rules are represented by intrusion data and classified between abnormal and normal users. The different rules are generated separately from three models "Time Based Traffic Model", "Host Based Traffic Model", and "Content Model". The proposed system has generated the update and adaptive rules automatically and continuously on the misuse detection method which is difficult to update the rule generation. The generated rules are experimented on 430M test data and almost 94.3% of detection rate is shown.

Key Words : 침입탐지시스템, 유전자 알고리즘, 데이터 마이닝, 오용탐지

1. 서 론

인터넷을 통한 자료의 공유와 전자상거래의 폭넓은 사업성과 경제성을 통해 소규모 또는 대규모 전산망의 구축이 활발하게 구축되고 있다. 그러나, 그 이면에는 엄청난 위험성이 내재되어 있어, 네트웍의 세밀한 운영 및 제어가 절실히 요구된다. 그 위험성이란 경제적인 이윤을 목적으로 또는 악의적인 목적으로 불특정 다수의 네트웍에 존재하는 시스템을 공격함으로써 불법적인 이익을 취득하려는 행위를 말한다. 특히 현 시점에서 사용되어지고 있는 시스템들은 OS(Operating Systems)의 근본적인 보안 문제와 네트웍의 하드웨어 및 소프트웨어의 근본적인 보안 문제와 함께 가용 시스템의 무결성, 가용성, 기밀성을 완벽하게 보장할 수 없다. 또한 지속적인 보안성 향상을 위한 노력에도 불구하고, 중요 시스템에 대한 불법적인 접근은 매년 증가하고 있다. 침입은 "자원의 가용성 또는 은밀함, 즉 무결성을 침해하기 위한 시도들의 집합"으로 정의 되어 질 수 있다[1]. 사용자 인증(패스워드 또는 생체인식 등등)과 같은 침입 방지 기술들은 암호화 같은 정보보호 기술들과 함께 시스템을 보호하기 위한 첫 번째 단계로서 사용되어지고 있다. 시스템이 더욱더 복잡해짐에 따라, 그러한 침입 방지 기술만으로는 충분하지 않다. 또한, 시스템과 파일접근 권한에 관한 정책들이 가용 시스템의 보안성을 더욱 향상 시켜 주기는 하나, 이것으로 보안에 관한 불안을 모두 해소 할 수는 없다. 따라서 시스템의 기밀성 및 무결성을 확보하기 위한 또 하나의 방법으로 침입 탐지가 필요하게 되었다. 이러한 침입탐지 시스템은 정밀성과 적응성, 그리고 확장성을 필요로 한다. 이와 같은 조건을 포함하면서 복잡한 Network 환경에서 중요하고 기밀성이 유지되

접수일자 : 2002년 9월 18일
완료일자 : 2002년 11월 9일
본 연구는 2002년도 경원대학교 학술연구비의 지원을 받아 이루어졌음.

어야 할 리소스를 보호하기 위해, 우리는 더욱 구조적이며 지능적인 IDS 개발의 필요성이 요구되고 있다.

본 연구는 유전자 알고리즘을 IDS에 적용하여 오용 탐지 기법을 처음으로 제안하고 구현한 점에서 의미가 있다. 세계적인 대회인 KDD 콘테스트의 데이터를 사용하여 실험하였으며, 그에 따른 가능한 한 같은 환경 하에서 실험을 실시하였다[2]. 실험은 레코드집합을 하나의 유전자로, 즉 하나의 공격패턴으로 간주하고 유전자 알고리즘을 활용하여 진화 시켜 침입 패턴, 즉 침입 규칙(Rules)을 생성한다. 데이터 마이닝 기법중 분류(Classification)에 초점을 맞추어 분석과 실험을 하였다. 이 데이터를 중심으로 침입 패턴을 생성하였다. 즉, 오용탐지(Misuse Detection) 기법을 실험하였으며, 생성된 규칙은 침입데이터를 대표하는 규칙로 비정상 사용자와 정상 사용자를 분류하게 된다. 규칙은 "Time Based Traffic Model", "Host Based Traffic Model", "Content Model" 이 세가지 모듈에서 각각 상이한 침입 규칙을 생성하게 된다. 규칙 생성의 지속적인 업데이트가 힘든 오용탐지 기법에 지속적으로 성장하며 변화해 가는 규칙을 자동적으로 생성하는 시스템으로서, 생성된 규칙은 430M Test data 집합에서 테스트한 결과 평균 약 94.3%의 탐지율을 보였다.

본 논문의 구성은 다음과 같다. 데이터 마이닝 기법에 핵심적으로 사용된 유전자 알고리즘과 침입탐지 시스템에 관하여 2장에서 언급하며, 3장에서는 제안한 지능형 침입 탐지 시스템을 설명한다. 4장에서는 실험에 사용된 데이터에 대한 분석과 실험 환경 및 실험 결과를 기술하고, 마지막으로 5장에서 결과를 분석하며 성능을 평가한 후 향후 연구 방향에 대하여 논한다.

2. 유전자 알고리즘과 침입탐지 시스템

2.1 유전자 알고리즘(Genetic Algorithms:GAs)

GAs는 유전적 계승과 다윈적 생존 경쟁이라는 자연 현상을 모델링한 확률적인 탐색방법으로, 유전검색이 불가능할 정도로 큰 후보해 공간을 갖는 최적화문제에 적용할 수 있다[3]. 즉, 해가 될 가능성이 있는 개체집단을 유지함으로써 여러 방향의 탐색을 실행하고 이들 방향간의 정보 형성과 교환을 행한다. 개체집단은 진화과정을 모방하는데, 각 세대에서 비교적 우량한 해들이 재생산되고, 반면에 비교적 불량한 해들은 소멸된다. 또한 다른 해들간의 차이를 구별하기 위해 환경의 역할을 수행하는 목적함수를 사용한다. 이러한 유전자 알고리즘은 특정한 문제에 대해 다섯 가지의 요소를 가져야만 한다. 유전자적 표현방법, 초기 개체집단을 만들어 내는 방법, 목적함수, 유전 연산자, 그리고 여러 가지 매개변수의 값이다.

어떤 개체집단을 초기화하기 위해서는 단순히 개체집단의 염색체를 비트 단위로 임의로 설정할 수 있다. 혹은 가능한 최적 값들의 분포에 관한 지식을 가지고 있다면 초기의 해집합을 배열하는 데 그 정보를 이용할 수 있다.

알고리즘의 나머지 부분은 각 세대에서 각각의 염색체를 평가하고, 적합도 값에 기초한 확률분포에 의하여 새로운 개체집단을 선택하며, 돌연변이와 교배연산자에 의하여 새로운 개체집단의 염색체들을 변화시킨다. 여러 세대 후에 더 이상의 개선이 없으면, 그 세대의 가장 좋은 염색체가 최적해를 나타낸다. 선택과정에서는 적합도에 비례해서 가장 좋은 염색체는 더 많이 복제되고, 보통 염색체는 비슷하게 남아 있으며, 최악의 염색체는 소멸된다.

교배연산자는 교배연산확률을 토대로 두개의 염색체에 적용되서 새로운 두 개의 자손을 생산하며, 마지막으로 돌연변이 연산자가 돌연변이 확률에 의해 비트별로 적용된다. 이러한 선택, 교배, 그리고 돌연변이를 한 후에 새로운 개체집단은 평가를 받는다.

대부분의 "데이터 마이닝 시스템"은 전통적인 기계학습(Machine Learning) 알고리즘의 변형을 사용해 왔다. 기계학습에서는 복잡한 시스템을 대상으로 하여 그 대상시스템을 학습시킬 뿐만 아니라 시스템에 대한 적절한 출력을 만들어 내는 두 가지의 목적을 가진다. 기계학습에서 유전자 알고리즘의 기법을 이용한 것을 GA기계학습 또는 GBML(Genetic Based Machine Learning)이라고 한다.

기계학습이 최적화 문제와 근본적으로 다른 점은 패턴의 집합를 구하지 않으면 안 되는 점이다. 최적화 문제에서는 최적화에 가까운 우수한 해를 구하는 것이 목적이기 때문에 최후의 한 종류만이 개체에 수렴하면 되지만 기계학습에서는 가장 좋은 패턴 하나만 구하는 것이 아니라 서로 협조하는 패턴의 집합을 구하는 것이 필요하다. GBML에서는 일반적으로 두 가지 접근 방법이 있다. 전체 패턴 집합을 하나의 개체로 표현하고, 후보 패턴 집합들의 개체 집단을 유지하고, 그리고 패턴 집합들의 새로운 세대를 생성하기 위한 선택과 유전 연산자를 사용하는 것이 자연스러운 방법으로 여겨질 것이다. 즉, 전통적인 유전자 알고리즘을 사용하며, 집단안에서의 각 실재(entity)는 학습 문제에 대한 완전한 해를 표현하는 패턴의 집합이다. 이러한 접근 방법을 Pitt 접근 방법이라 한다. 또한 같은 시기에 Holland는 개체 집단의 소속원들이 각각의 패턴들이고, 하나의 패턴 집합이 전체 개체 집단에 의해 표현되는 분류 시스템을 발전시켰다. 이러한 방법은 Michigan 접근 방법이라 불리워진다. 이러한 Michigan 접근 방법은 상당히 다른진화 방법을 사용하는데, 집단은 개인적인 패턴들로 구성되었으며 각 패턴은 전반적인 학습 임무에 대한 부분 해를 표현한다. 즉 Pitt 접근 방법은 진화 연산과 유사하지만, Michigan 접근 방법은 매우 다른 새로운 방법을 사용한다. 본 연구에서는 Pitt 접근 방법을 사용한다.

2.2. 침입탐지 시스템(Intrusion Detection Systems)

현재 사용되어지고있는 침입 탐지를 위한 접근 방법에는 두 가지가 있다. 첫 번째 접근방법은 동적인 시스템 사용자의 행위를 and/or의 형태로 특성화시킨 정상적인 패턴-보통은 통계적인 방법을 사용하여, 사용자 행위들간의 관계를-을 정의하여, 그 정의된 것으로 비정상적인 사용자를 탐지하는 방법이다[4]. 이것을 비정상행위 탐지(anomaly detection)라 한다. 이러한 방법들은 시스템의 사용자의 이벤트를 정의하는 것에 의존하여, 정의된 정상 사용자의 행위와 비정상적인 사용자의 행위를 분리해 낸다.

이러한 비정상행위 탐지는, 정상적인 사용자들의 행위를 얼마나 잘 표현하여 정의하는 가에 그 성능이 좌우된다고 해도 과언이 아니다. 그런데, 수많은 사용자들의 행동 양식을 정의한다는 것이 그리 쉬운 것이 아니기 때문에 정상적인 행위를 정의하여 침입을 탐지한다는 것이 매우 어렵다.(즉, 실제로 시스템 사용자의 행위가 침입인지 정상적인 행위 인지의 구분이 모호할 때가 많다.)

두 번째 접근방법은 오용 탐지(misuse detection) 방법이다. 이것은 이미 알려진 공격방법, 비정상적인 사용자들의 행위, 일정한 코드들의 수행 등을 정형화하여 비정상적인 사용자들을 탐지한다.

"비정상행위 탐지"와 "오용탐지"의 가장 중요한 기술적 차

이점은 패턴 즉, 규칙과 데이터를 "어떻게 표현하는가?" 또는 "어떻게 침입자의 행위를 잘 표현한 패턴을 구성하는가?" 라고 할 수 있다. 오용탐지 시스템은 침입행위의 표시를 위한 이벤트(events)를 표현하기 위해 규칙이 사용된다. 이 패턴은 보안 관리자가 시스템 안에서 찾게 된다. 다양하고 큰 숫자들의 패턴은 번역하기가 어렵다. 만약 이 패턴이 침입 시나리오에 의해 모여진 게 아니라면 이러한 패턴은 정형화 시키기 어렵게 된다. 이러한 어려움을 극복하기 위해서는 침입 시나리오에 구체적인 패턴들을 발견하는 방법을 개발하여야 한다. 최적화된 패턴들의 집합을 만들어 가는 것이 성능을 향상시키는데 즉, 침입을 탐지 율을 높이거나, 정상적인 행위를 침입으로 탐지하는 에러 율을 줄이는데 매우 중요하다. 또한 이러한 패턴들은 시스템 관리자들이 이해하기 쉬운 표현으로 메시지를 전달해 주어야만 적절한 대응을 할 수 있게 된다. 이러한 패턴들을 생성하기 위한 자원으로 시스템 감사 레코드(Audit records)를 빈번하게 사용한다. 이유는 모든 시스템은 O/S를 사용하여 구동되고 있으며, 이 O/S는 시스템 사용자의 행위를 즉각적으로 감사 레코드에 기록함으로써 가장 빠르게 사용자의 행위를 알 수 있으며, 이로써 오용 탐지 시스템은 패턴 생성을 위한 자료를 얻게 된다. 아울러 이미 알고 있는 패턴과 비교하여 침입자를 탐지하게 된다. 침입시나리오가 구체화될 수 있게 된 이후로 오용탐지 시스템은 사용자 행위(시스템 이벤트)들을 모니터 할 수 있게 된다. 이러한 시스템 이벤트들은 침입 시도(마우스 이벤트, 커멘드 이벤트 등등의 연속)의 흔적을 남기게 되어 순차적인 이벤트들이 진행되는 동안, 오용 탐지 시스템은 다음 침입 시나리오를 사용하는 침입의 단계를 예측 할 수 있다. 이러한 정보가 제공되므로 침입 탐지 시스템은 더욱 깊이 있게 다음 단계의 audit record들의 분석 할 수 있다.

침입탐지 시스템은 비정상행위 탐지와 오용탐지로 접근해 각각 따로 구현되어 사용되어 지거나, 이 둘이 결합하여 침입 탐지에 사용되어 진다. 이들이 갖는 장점들을 모두 갖추기 위해 현재는 두 가지 방법을 모두 구현한 IDS가 널리 사용되어 지고 있다.

3. 지능형 침입탐지 시스템(Intelligent Intrusion Detection Systems:IIDS)

유전자 알고리즘을 IDS에 적용하는 것은 여러 가지로 어려움이 많은 연구 분야이다. 데이터의 표현이나 유전형질의 변환 그리고 그 연산 방법들이 구체적으로 적용되어지는 기본적인 개념부터 고심해야 하기 때문이다.

유전자 알고리즘을 사용한 분류 시스템(Classifier System)으로 대표되는 시스템에는 Holland의 Michigan 접근방법과 Smith의 LS-1 시스템으로 대표되는 Pittsburgh 접근방법이 있다[5][6]. 이중에서 Pittsburgh 접근방법을 사용한 시스템을 통하여 인증되지 않은 사용자들(그것이 외부든 내부든)로부터의 침입에 어떠한 컴퓨터 네트워크를 보호하는 침입탐지 시스템을 개발하는데 그 목적이 있다.

본 시스템에서는 41개 각각의 features value의 개수와 범위를 전처리기를 통하여 산출해 냈으며, Features를 각각의 Value Type에 따라 Features를 정의 하였다.

또한 평가함수(Fitness function) 에서는 각 generation에서 모든 염색체들은 그들의 적합도에 의해 평가되고, 새로운 개체 집단은 좀더 좋은 염색체들로 구성이 된다. 여기에, 연산자들은 새로운 개체 집단에 적용되고, 다시 반복된다. 이러한 모델링을 기본으로 삼아 본 실험에서는 전체 Training Data에 맞게 예측된 데이터 개수로 나눈 것을 평가함수로 표현하였다.

$$F = \frac{Number\_of\_Corrected\_recoard}{Number\_of\_total\_record}$$

(F: Fitness value)

그림 1. 평가함수

Fig 1. Fitness Function

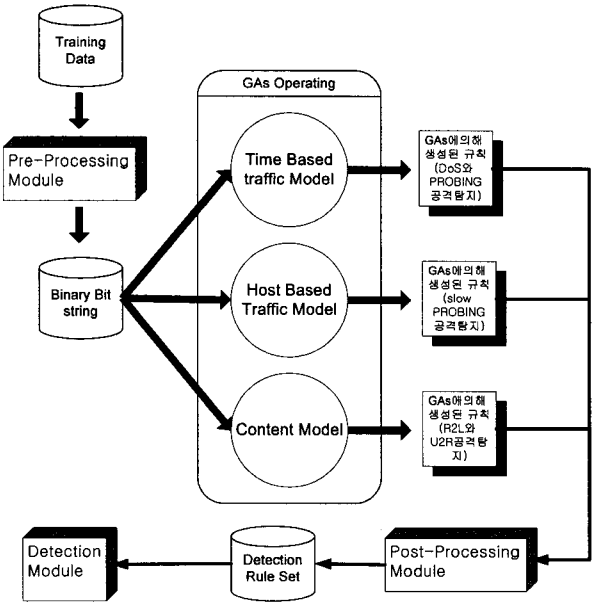


그림 2. GAs를 이용한 침입규칙 생성기

Fig 2. Intrusion Rule Generator using GAs

GA연산부분은 3부분으로 나누어지고 각 내부 루틴은 각각 데이터 타입이 틀리게 된다. 처음 랜덤으로 2진 population을 생성하고 그 data와 함께 규칙을 선택하기위해 임의의 공격 규칙의 집단을 training data로 입력 시킨다. 그렇게 각 모듈별로 입력이 있은 후 연산에 의해 생성된 최적해가 나오게 되고 GA를 통하여 하나의 새로운 rule을 만들게 된다.

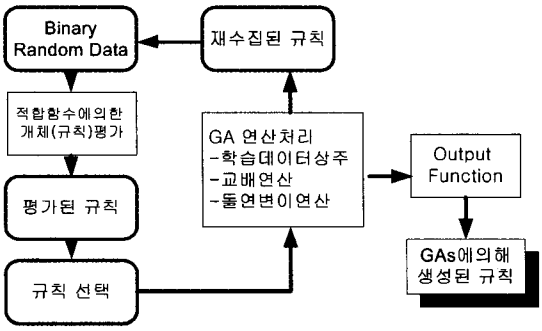


그림 3. 유전자 연산 시스템 구조도

Fig 3. Structure of GAs Operating System

위 그림에서, GA 동작부분의 세부 구조를 나타내었고 루틴이 한번 순환하게 되면 1 generation이 이루어져 한세대 진화 연산이 수행된 집단이 생성된다.

4. 실험 및 분석

본 장에서는 실험에 사용된 데이터 분석과 제안된 시스템, 그리고 실험의 진행 및 결과, 성능 평가를 기술한다.

4.1 KDD DATA 분석

실험에서 사용된 데이터는 1999년 "KDD'99 Competition: Knowledge Discovery Contest"에서 제공된 것을 활용하였다. 침입과 정상데이터로 라벨(labels)되어있는 데이터를 Training Data로 사용하였으며 라벨이 없는 데이터를 Test Data로 사용하였다[2].

침입 탐지 모듈의 학습을 시키는 작업은 "bad connections" 즉 침입 또는 공격 행위들과 "good connections" 즉 정상사용자들의 행위를 구분 가능한 예측모델(즉 분류기 또는 분류자)을 설계하는 것이다.

1998년 DARPA 침입탐지 개발 프로그램은 MIT Lincoln Labs에서 준비되었고 관리되어져 왔다. 여기서 제공되어진 데이터는 군사 네트워크 환경에서 실험되어진 방대하고 다양한 침입들을 포함하고 있는 표준 감사 데이터집합(data set)들이다. 이후 1999년 KDD Intrusion Detection contest는 바로 이 데이터집합을 활용하여 진행되었다. Lincoln Labs에서 미공군의 LAN에서 9주간의 raw TCP dump data를 얻기 위해 실험 환경을 조성하였다.

Raw Training Data는 7주간의 네트워크 트래픽에서 압축된 Binary TCP dump data 약 4기가바이트를 사용하였다. 이 데이터는 약 5백만 connection records를 포함하고 있으며 유사하게 2주간의 Test Data는 약 2백만 connection records를 포함하고 있다.

Connection은 잘 정의된 일정한 시간동안 그 처음과 끝이 TCP packets의 연속으로 구성되어있다. 이것은 신뢰할 만한 프로토콜을 통하여, Packet들의 출발지 IP와 목적지 IP까지의 packets, 그리고 Data의 그 flows를 포함한다. 각각의 connection은 label들이 표시되어있는데, 정상사용자 인지 비정상사용자(네트워크에 대한 공격, 정확한 공격유형중 하나)를 표시하는 라벨이다. 각 connection record는 대략 100 bytes로 표현된다. 이 데이터에서 주요한 4개의 공격들의 유형은 다음과 같이 4가지로 분류된다.

- DOS : denial-of-service, 예) syn flood 등등
- R2L : unauthorized access from a remote machine, 예) guessing password 등등
- U2R : unauthorized access to local superuser (root) privileges, 예) various, buffer overflow 등등
- probing : surveillance and other probing, 예) port scanning 등등

여기서 중요하게 생각해야 할 것은, Test data는 Training data와 같이 동일한-공격 유형들의- 확률적 분포를 나타내지는 않으며, Training data에는 없는 상세한 공격 유형들을 포함한다. 이러한 것들이 본 실험을 더욱 실제적으로 유용하게 만들어준다. 이러한 Training 데이터집합은 24개의 Training Attack Types을 포함 하고 있으며, 테스트 데이터에는 14개의 공격 유형(types)을 더 포함하고 있다.

Training Attack Types은 다음과 같다.

표 1. 학습 공격 유형
Table 1. Training Attack Types

	Attack
DoS	back, land, neptune, pod, smurf, teardrop
PROBING	ipsweep,nmap, portsweep, satan
R2L	ftp_write, guess_passwd, imap, multihop, phf, spy, warezclient, warezmaster
U2R	loadmodule, buffer_overflow, perl, rootkit

위 공격유형은 "4.1.2 탐지 모델"에 제시된 모델에서 각각 탐지하게 된다. 먼저, Dos와 PROBING 공격들은 "Time based traffic model"에서, Slow PROBING 공격은 "host based traffic model"에서, R2L과 U2R은 " content model"에서 각각 생성된 규칙을 사용하여 침입을 탐지 한다.

4.1.1 속성 분석 및 선택

참고문헌 [7]에서 공격자들의 connections으로부터 정상적인 사용자들의 connections을 구별하기에 도움을 주는 higher-level features를 정의하였다. Derived Features의 몇몇 카테고리(categories)가 다음과 같다.

표 2. 도출된 속성
Table 2. Derived Features

	connections 만을 살펴본 것이다. 2초 동안, 현재 connecton과 동일한 목적지 host를 가지고 있는지, service 등등과 같은 프로토콜 행동들과 관련된 통계량을 계산한 것이다.
	2초간 현재 Connection에서 같은 service를 가지고 있는 connection을 살펴 본 것이다.

"same host" 와 "same service" features 모두 connection records의 time-based traffic features 라고 부른다.

호스트(또는 ports)를 스캔하는 PROBING 공격은 매우 오랜 시간을 필요로 한다.(2초이상) 그러므로 connection records는 목적지 호스트로 정렬이 되었으며, features는 time window 대신 동일한 호스트에 100개의 connection windows를 사용하여 구성되었다. 이것은 "host-based traffic features"라 불리는 것들의 집합을 산출해 낸다.

대부분의 DOS 와 PROBING 공격은 연속적인 패턴들이 발생하며 매우 짧은 시간동안 여러 호스트(들)에 많은 connections를 수반하지만, , R2L과 U2L공격들의 records에는 연속적인 패턴이 발생하지 않으며 R2L과 U2L 공격들은 packets의 데이터 부분에 포함되어있으며 정상적인 사용자는 오직 하나의 connection만을 가지고 있다.

packets의 구조화되지 않은 데이터 부분들을 자동적으로 탐색하는 것에 대한 유용한 알고리즘을 연구하는 것은 이미 잘 알려진 연구 분야이다. Stolfo et al.은 데이터 일부에서 로그인을 시도했을 때 실패한 횟수와 같은 행위를 한 의심이 가는 사용자를 찾는 features를 추가하는 기법(또는 지식 분야)를 사용하였다. 그들 feartures를 "content" feartures라 한다.

Features는 각각 KDD Data에 각각의 표현형 대로 나열되어 있으며 총 41개의 Features가 있다. Training Data 에는 오른쪽 마지막에 라벨이 있다.

4.1.2 탐지 모델

위에 제시된 Features의 구성을 통해 다음의 3가지 탐지 모델을 구성하였다[8].

- ▶ Time based traffic model
  : 9개의 Intrinsic Features와 9개의 Time based traffic Features로 구성된다. Dos와 PROBING 공격들을 탐지하는데 사용되는 규칙을 생성한다.
- ▶ host based traffic model
  : 9개의 Intrinsic Features와 10개의 host based traffic Features로 구성된다. slow PROBING 공격들을 탐지하는 규칙을 생성한다.
- ▶ content model
  : 9개의 Intrinsic Features와 13개의 content Features로 구성된다. R2L과 U2R 공격들을 탐지하는 규칙을 생성한다.

위 모델들은 MIT Lincon Lab.에서 제공된 KDD 데이터의 10%(75M) 데이터로 학습 시켰으며, 49,4021개 레코드로 구성되어있다.

4.2 결과 및 분석

본 실험에서 사용된 주요 파라미터(Parameters)는 다음과 같다. 먼저 교배율(corssover-rate)은 0.6~0.8, 돌연변이율(Mutation rate)은 0.4~0.08의 범위에서 실험 하였으며, 염색체는 2진수로 표현하였다. 염색체의 평가함수는 집단 내 개체 다양성(diversity)를 계산하기 위해 사용된다. 이 함수는 두 염색체를 비교하여 0이나 0보다 큰 실수를 반환한다.

계산된 결과는 유전자 연산에 사용된 염색체를 선택하기 위한 선택 전략에 사용된다. 염색체의 평가후 집단을 평가한다. 스케일링 전략(Scaling Sheme)으로는 선형 스케일링을 사용하였으며, 교배 연산자로는 2점 교배(2 Point Corssover) 연산자를, 돌연변이 연산자는 교환 돌연변이 연산자를 사용하였다. 비교 연산자는 "비트 비교" 연산자를 종료 연산자는 "세대 수렴에 의한 종료"연산자를 사용하였다. 선택 전략으로는 가장 많이 사용되고 있는 규칙렛휠 선택(Roulette Wheel Selection)을 적용하였다. 집단의 크기는 각각의 학습 데이터와 같게 하여 실험을 하였다. 3가지 탐지 모델들은 각각의 침입유형을 특성화하는 분류 모델들이다. 이들로서 생성된 규칙의 예는 다음과 같다.

표 3. Time Based traffic model의 규칙 예
Table 3. Rule Example of Time Based traffic model

1	duration = 011010010110	duration 속성의 속성값이 100~299, 400~499, 700 ~799, 900~999, 1000~9999의 값을 가질때 침입임.
⋮	⋮	⋮
18	src_diff_host_rate = 0000000001111	dst_host_srv_rerror_rate 속성의 속성값이 9이상, 즉 쉘 프롬프트의 횟수가 9일상 값을 가질때 침입임.

표 3에서 보듯이 18개의 속성들은 "∨(OR)"로 연결되어 있으며, 각각의 속성값들은 "∧(AND)"로 연결되어 해석되어진다. 따라서 각각의 "∨"의 개수에 따라 여러 개의 경우의 수가 발생되어, 다양한 패턴(규칙)이 생성되어 진다.

생성된 규칙들의 테스트 데이터(Test Datas)에서 평가 되어졌으며, 3가지 탐지 모델의 Performance를 평가한다. 사용된 학습 데이터는 침입과 정상이 표시된 데이터를 사용하였으며, 표시되지 않은 데이터로 테스트 되었다. 테스트 데이터는 38개의 공격 유형이 포함 되어있으며, 이중 14개는 학습되지 않은 공격 유형으로써 본 3개의 모델에서는 이것을 "새로운 공격유형"으로 분류하게 된다. 이것은 지속적인 진화 연산을 통해 생성된 규칙들이 정체되지 않고 새로운 개체(공격 유형)를 통해 새로운 규칙로 진화를 수행해 간다.

생성된 바이너리(binary) 형태의 규칙은 후처리기(Post-process)를 통해 Raw 데이터 형태로 변환되어 분류자(classifier)에 전달되어 탐지를 수행하게 된다

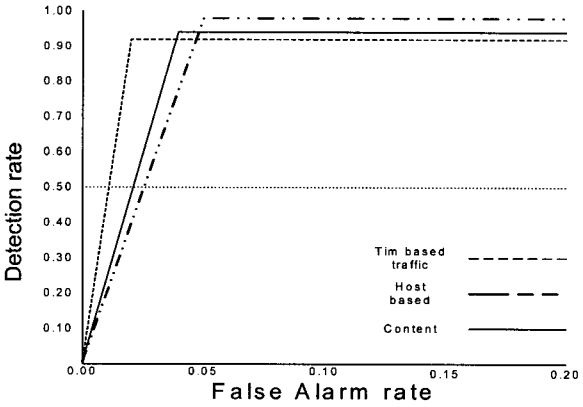


그림 4. 3가지 오용 탐지 모델의 탐지율 및 오탐지율
Fig 4. Detection Rate and False Alarm Rate for each Models

위의 그림은 본 연구에서 생성된 침입 규칙의 성능(Performance)을 보여 준다. x축은 정상적인 행위를 침입으로 오판한 비율을 전체 레코드의 개수를 기준으로 계산된 비율이며, y축은 전체 레코드에서 침입행위로 분류한 개수를 계산한 비율을 나타낸다. Time based traffic Model은 92%, Host based Model은 97%, Content Model은 94%의 탐지율을, 각각 0.02, 0.03, 0.05의 오탐지율을 산출해 냈다.

5. 결론 및 향후 연구 방향

본 연구는 유전자 알고리즘을 IDS에 적용하여 오용 탐지 기법을 처음으로 제안하고 구현한 점에서 의미가 있다. 세계적인 대회인 KDD 콘테스트의 데이터를 사용하여 실험 하였으며, 그에 따른 가능한 한 같은 환경 하에서 실험을 실시하였다. 규칙 생성의 지속적인 업데이트가 힘든 오용탐지 기법에 지속적으로 성장하며 변화해 가는 규칙을 자동적으로 생성하는 시스템으로서, 생성된 규칙은 각각의 모델들이 평균 약 94.3%의 탐지율을 보였다.

향후 보완될 점 및 지속적인 연구 방향으로는, 우선 정확한 KDD 콘테스트의 환경 하에 실험을 하여 다른 연구들과의 정확한 성능 비교가 이루어져야 하며, 데이터 마이닝의 주요 3분야중 Classification만 실험하였는데, 생성된 3가지의 침입 패턴을 결합하여 다중 분류(Meta-Classification)에 적합하도록 연구하여 성능평가를 해야 할 것이다. 또한 데이터 마이닝의 주요 2분야인 Link analysis와 Sequence

analysis를 연구하여 시스템에 추가하여 더욱 정확하고 신뢰있는 침입 규칙을 생성하는 것이다. 아울러 비정상행위 탐지에 쓰일 정상적인 사용자의 규칙 생성에 관해 연구할 것이다.

참 고 문 헌

[1] W. Lee and S. J. Stolfo, "Data mining approaches for intrusion detection," In Proceedings of the 7th USENIX Security Symposium, San Antonio, TX, Jan. 1998.

[2] The KDD data URL http://kdd.ics.edu/databases/kddcup99.html

[3] D.E.Goldberg, Genetic Algorithms in Search, Optimization, and Machine Learning, Addison wesley, 1989.

[4] D.E.Denning, "An Intrusion Detection Model," IEEE Trans. on Software Engineering, Vol.13, No.2, pp.222-232, Feb. 1987.

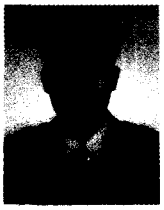
[5] J.H.Holland, "Adaptation in Natural and Artificial Systems," Ph.D. thesis, Univ. of Michigan, Ann Arbor, Mich., 1975.

[6] S.F.Smith, "A Learning System based on Genetic Adaptive Algorithms," Ph.D. thesis, Univ. of Pittsburgh, 1980.

[7] Lee,W., Stolfo,SJ., and Mok,K.W., "A data mining framework for building intrusion detection models," In Proceedings of the 1999 IEEE Symposium on Security and Privacy, May, 1999.

[8] Lee,W., and Stolfo,S.J., "A Framework for Constructing Features and Models for Intrusion Detection Systems," In Proceedings of the 5th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 1999.

저 자 소 개



양지홍 (Ji-Hong Yang)
2001년 경원대 수학과 졸업.
2001년~현재 : 경원대학 전자계산대학원 석사과정.
관심분야 : 침입탐지시스템, PKI, 유전자알고리즘
Phone : 031-750-5753
E-mail : wlghd@web.kyungwon.ac.kr

한명묵 (Myung-Mook Han)
1998년, 3~현재 : 경원대학교 소프트웨어대학 조교수
관심분야 : 데이터마이닝, 정보보호
Phone : 031-750-5522
Fax : 031-750-5522
E-mail : mmhan@kyungwon.ac.kr