디지털산업정보학회 논문지 제11권 제4호-2015년 12월
http://dx.doi.org/10.17662/ksdim.2015.11.4.033

의사결정트리와 인공 신경망 기법을 이용한 침입탐지 효율성 비교 연구

조 성 래* · 성 행 남** · 안 병 혁***

A Comparative Study on the Performance of Intrusion Detection using Decision Tree and Artificial Neural Network Models

Jo Seongrae · Sung Haengnam · Ahn Byunghyuk

〈Abstract〉

Currently, Internet is used an essential tool in the business area. Despite this importance, there is a risk of network attacks attempting collection of fraudulence, private information, and cyber terrorism. Firewalls and IDS(Intrusion Detection System) are tools against those attacks. IDS is used to determine whether a network data is a network attack. IDS analyzes the network data using various techniques including expert system, data mining, and state transition analysis.

This paper tries to compare the performance of two data mining models in detecting network attacks. They are decision tree (C4.5), and neural network (FANN model). I trained and tested these models with data and measured the effectiveness in terms of detection accuracy, detection rate, and false alarm rate. This paper tries to find out which model is effective in intrusion detection. In the analysis, I used KDD Cup 99 data which is a benchmark data in intrusion detection research. I used an open source Weka software for C4.5 model, and C++ code available for FANN model.

Key Words : Data Mining, C4.5, Neural Network, Decision Tree, Forward Additive Neural Network Models

Ⅰ. 서론

정보통신 기술의 발전은 우리의 삶에 다양한 변화를 가져왔으며 기업과 같은 조직적 차원에서는 업무의 효율성을 위해, 개인적 차원에서는 의사소통이나 즐거움을 위한 도구로 이용 되었다. 최근에는 스마트폰으로 대표되는 모바일 기술과 결합하여 필요한 정보를 쉽고 얻기도 하며 다양한 서비스를 제공받고 있다.

하지만 이러한 편리함은 또 다른 이면에서의 문제점으로 대두되고 있다. 인터넷을 통한 다양한 방법으로 네트워크 공격이 시도되고 있고 공격에 의한 피해가 발생하고 있다. 대표적으로 2011년 NH 금융전산

* 경상대학교 대학원 경영정보학과 석사(주저자)
** 경상대학교 경영대학 강사
*** 경상대학교 경영대학 경영정보학과 조교수, 경영경제연구소 (교신저자)

망 마비 사태와 3500만 여명의 개인정보가 유출된 네이트 개인정보유출 사태 등이 이슈화되기도 하였다.

최근에는 공격수법이 지능화되어 보안에 대한 위협은 계속 증가하고 있으며, 그 대상도 스마트폰에서부터 금융 네트워크, 지적재산권, 국가 인프라에 이른다. 공격의 목적도 개인 정보 획득, 사기, 위조, 사이버 테러 등으로 다양화되고 있고 피해의 규모도 점점 커지는 추세이다. 따라서 이에 대한 대응책이 요구되고 있다[1].

다양한 네트워크 침입에 대응하기 위한 방법으로는 방화벽, 침입탐지시스템 등이 있다. 방화벽은 외부로부터 들어오는 공격이나 바이러스로 의심되는 패킷을 차단하는 역할을 하지만 모든 공격을 차단할 수는 없다. 따라서 공격에 대응하기 위해 방화벽과 함께 침입탐지시스템의 역할이 요구된다.

본 연구에서는 기존의 침입탐지 연구에 적용된 적이 없는 FANN 모형을 이용하여 침입탐지시스템에 의해 수집된 데이터에 대한 훈련과 테스트를 통해 침입탐지에 대한 효율성을 확인하는데 그 목적이 있다. 그리고 비교를 위해 기존 연구에 적용된 C4.5를 이용한 실험을 진행한다. 본 연구의 실험에는 DARPA(Defense Advanced Research Projects Agency)의 연구를 통해 개발된 KDD Cup 99 데이터 세트를 이용한다. 그리고 침입탐지시스템의 효율성은 탐지의 정확도와 탐지율, 그리고 오경보율을 계산하고 비교한다[2].

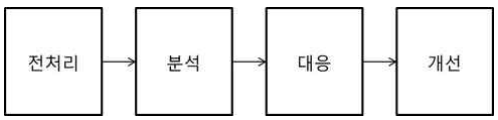
본 논문은 침입탐지와 관련된 정의와 분류, 연구에 사용된 C4.5, FANN 알고리즘을 상세하게 설명하였다. 실험 결과를 바탕으로 두 탐지 기법 간 정확도, 탐지율, 오경보율을 분석 및 평가하여 시사점을 제시하였다.

Ⅱ. 이론적 배경 및 관련연구

2.1 침입탐지 시스템과 분류

침입탐지시스템(IDS : Intrusion Detection System)은 시스템, 네트워크와 데이터베이스 등에서 발생하는 이벤트를 모니터링(monitoring)하고 침입에 대한 분석 및 탐지를 통해 이에 대응하는 소프트웨어 또는 하드웨어 장치로 구성된 자동화 시스템이다[3]. 이는 시스템에 대한 무결성을 보장하며, 관리자가 시스템의 정책을 만드는데 도움을 준다[4]. 침입탐지시스템과 관련된 기존 연구는 침입탐지시스템의 성능최적화 및 평가방법론 개발[5], 침입탐지시스템의 설계 및 구현[6] 등이 주를 이룬다.

침입탐지 시스템의 일반적인 과정은 그림 1과 같이 4 단계로 구성된다[7-8]. 데이터 전처리 단계에서는 IDS 센서로부터 활동에 관한 데이터를 수집한다. 그리고 수집된 데이터를 분석에 적합한 형태로 가공한다. 분석 단계에서는 전처리 된 데이터를 정해진 패턴이나 규칙과 비교하여 침입을 탐지하게 된다. 데이터가 정상으로 탐지되면 다음 데이터에 대한 분석을 수행한다. 대응 단계에서는 침입이 감지되면 경보를 통해 이 사실을 알려준다. 대응에는 자동적으로 조치를 취하는 능동적 대응과 알림에 의해 관리자가 직접 대응을 관리하는 수동적 대응이 있다. 마지막으로 개선 단계에서는 분석을 통해 수집된 정보나 침입 데이터를 시스템에 반영한다. 이를 통해 오경보율을 낮추고 보안을 강화할 수 있다.



<그림 1> 침입탐지시스템의 침입탐지과정

Bace and Mell은 표 1과 같이 침입탐지시스템을 데이터 소스, 공격을 탐지를 위한 분석기법, 경과 시간, 제어 전략, 대응 옵션을 기준으로 분류하였다[3].

데이터 소스에 의한 분류는 분석의 대상이 되는 데이터 소스의 종류에 따라 분류되며, 호스트의 시스템에서 수집된 데이터를 분석하는 호스트 기반 침입탐지시스템과 다수의 호스트에 대한 네트워크 트래픽을 분석하는 네트워크 기반 침입탐지시스템이 있다.

<표 1> 침입탐지시스템의 분류

기 준	항 목
데이터 소스	호스트 기반(host based) 네트워크 기반(network based) 애플리케이션 기반(application based)
분석기법	오용 탐지(misuse based) 비정상 행위 탐지(anomaly based)
경과 시간	실시간(real-time) 인터벌 기반(interval-based)
제어 전략	집중화(centralized) 부분적 분산(partially distributed) 완전 분산(fully distributed)
대응 옵션	능동적 대응(active) 수동적 대응(passive)

분석기법에 의한 분류에는 공격의 서명을 대상으로 분석하는 오용탐지와 이용자의 일반적인 행동을 바탕으로 침입을 탐지하는 비정상 행위 탐지가 있다.

경과 시간에 따른 분류는 모니터링 된 이벤트와 이벤트가 분석되는 시점 사이에 시간 차이가 기준이 되며, 실시간 침입탐지시스템과 인터벌 기반 침입탐지시스템이 있다. 실시간 침입탐지시스템은 데이터의 획득과 분석이 연속적으로 이루어지며, 침입에 대한 신속한 대처가 가능하다. 하지만 시스템 자원을 많이 사용하는 단점이 있다. 인터벌 침입탐지시스템은 데이터가 모니터링 되는 시점과 이를 분석하는 시점이 분리되어 있다. 이는 다양한 데이터를 디스크에 저장하고 별도의 분석과정을 거치기 때문에 침입이나 발생한 문제에 대한 정밀한 분석이 가능하고 시스템 자원도 비교적 덜 이용한다. 그러나 더 큰 디스크 공간을 요구한다.

제어 전략에 따른 분류는 침입탐지시스템의 요소들에 대한 제어 방법과 입력 또는 출력의 관리 방법을 고려한다. 이 분류에는 모니터링, 탐지, 보고 등 모든 기능이 중앙에서 제어되는 집중화 제어 전략과 모니터링과 탐지는 분산시키고 중앙에서 보고를 제어하는 부분 분산 제어 전략, 그리고 다수의 에이전트가 개별적으로 작동하는 완전 분산 제어 전략이 있다.

대응 옵션에 따른 분류는 침입탐지시스템이 침입을 탐지한 상황에서의 대응과 관련이 있다. 능동적 대응은 침입을 탐지한 경우 해당 침입에 대한 차단과 침입자에 대한 조치, 네트워크 환경의 재설정 등의 역할을 자동적으로 수행한다. 수동적 대응은 동일한 경우에 관리자에게 침입을 알리는 역할을 하며, 대응은 관리자에 의해 수행된다.

2.2 침입탐지 시스템 연구동향

침입탐지의 개념은 1980년 Anderson에 의해 최초로 소개되었다. 1987년 Denning은 최초의 침입탐지 모형을 제안하였다. Denning의 모형은 특정 시스템, 애플리케이션 환경, 시스템의 취약점이나 침입 유형에 독립적인 다목적 침입탐지 전문가 시스템이며, 감사 데이터를 바탕으로 규칙을 생성하여 비정상 행위를 탐지하는 구조를 가진다. 위 초기 연구들을 바탕으로 침입탐지분야에서는 다양한 모형과 연구들이 진행되고 있다[9].

DARPA에서는 1998년부터 침입탐지시스템의 평가에 대한 연구를 진행하였다. 이 과정에서 네트워크간의 접속에서 발생하는 공격과 정상으로 구분되는 데

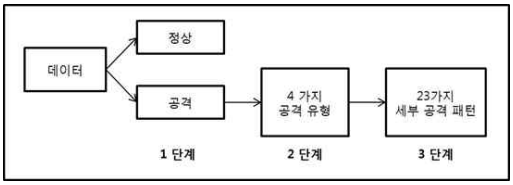
이터를 추출하고 이를 바탕으로 시뮬레이션을 통해 구축된 데이터세트를 제공하고 있다. 또한 이 데이터세트의 다른 버전인 KDD Cup 99 데이터세트는 최근의 연구에 많이 이용되고 있다.

Nguyen and Choi은 KDD Cup 99 데이터세트를 이용하여 무료 데이터 마이닝 도구인 Weka에 포함된 BayesNet, NaïveBayes, J48(C4.5), NBTree, Decision Table, JRip, OneR, MLP, SMO와 LBK 분류기를 이용하여 훈련과 테스트를 수행하고 공격유형별 정확도와 훈련시간을 계산하였다[10]. 이 과정을 통하여 DoS와 Probe 유형에는 JRip 분류기, U2R 유형에는 Decision Table 분류기가 높은 정확도를 나타내는 것으로 밝히고 있다. 또한 다중 알고리즘을 이용한 병렬 모형을 제안하고 있다.

Wu and Yen은 데이터 마이닝 기법 중 의사결정트리 알고리즘인 C4.5와 SVM을 KDD Cup 99 데이터세트에 적용하였다[2]. 두 기법의 분류 결과에 대한 정확도, 탐지율과 오경보율을 계산하고 비교한다. 그 결과 C4.5는 대부분의 경우에서 SVM보다 높은 정확도와 탐지율을 나타내었지만, 오경보율은 SVM에서 더 높게 나타나는 것으로 밝히고 있다.

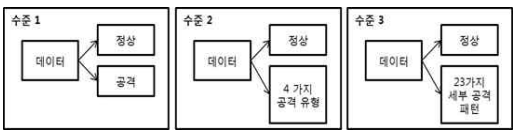
Jalil et al. 은 Osareh and Shadgar의 연구 결과를 바탕으로 C4.5, SVM, 신경망 알고리즘을 KDD Cup 99 데이터세트에 적용하여 연구를 진행하였다[11], [12]. 이 연구의 특징은 결과의 정확성을 검증하기 위해 훈련 데이터세트를 정상 데이터 비율에 따라 구성하고 cross-validation 기법을 이용하였다. 그 결과 탐지율은 정상 데이터 비율의 증가에 따라 높아지는 추세를 보였으며, 각 공격유형에 대한 탐지율은 C4.5 알고리즘에서 가장 높은 나타났다고 설명하였다.

Ibrahim et al. 은 그림 1의 단계 모형(phase-model)과 그림 2의 수준 모형(level-model)에 의한 침입에 대한 분류를 제안하였다[13]. 분류의 기법으로는 의사결정트리 기법의 알고리즘은 C5, CRT, CHAID,



<그림 2> 단계모형

Quest를 이용하였다. 단계 모형(phase-model)은 세 개의 의존적인 단계로 구성되고 이전 단계가 끝나면 다음 단계가 시작된다. 1단계는 정상 데이터와 공격 데이터를, 2단계에서는 1단계에서 공격으로 분류된 데이터를 DoS, U2R, R2L와 Probing Attack의 4가지 유형으로 분류한다. 3단계에서는 각 유형에 대한 23가지의 세부 공격 패턴으로 분류한다. 이 모형의 장점은 확실한 공격 레코드가 아니더라도 의심이 생기는 레코드를 발견할 수 있다는 점이라고 설명하고 있다. 수준 모형(level-model)은 세 개의 독립적인 탐지 수준으로 구성된다. 수준 1에서는 공격과 정상 프로파일을 탐지한다. 수준 2에서는 수준 1의 결과와는 독립적으로 정상 데이터를 탐지하고 공격 데이터에 대해서는 4가지 유형으로 분류를 실시한다. 수준 3에서는 정상 데이터를 탐지하고 공격 데이터를 23가지의 세부 공격 패턴으로 분류한다. 이 연구에서는 C5 알고리즘의 높은 탐지율, 단계 모형의 짧은 훈련 시간과 새로운 공격에 대하여 높은 탐지율을 결과로 제시하고 있다.



<그림 3> 수준모형

2.3 의사결정트리

의사결정트리(decision tree)는 데이터 마이닝의 대표적인 분석 방법 중 하나로 주어진 데이터의 분류(classification)를 위해 이용한다. 이는 단순화된 시각적 모형을 생성하는 기법으로 다른 데이터 마이닝 기법들에 비해서 수행속도가 빠르고 결과를 쉽게 이해하고 설명할 수 있는 장점이 있다. 대표적인 알고리즘은 C4.5, CART(Classification and Regression Trees), CHAID (Chi-square Automatic Interaction Detection) 등이 있고 본 연구에서는 C4.5 알고리즘을 이용하였다.

C4.5 알고리즘은 1993년 Quinlan에 의해 제안된 의사결정트리 생성 알고리즘으로 Quinlan의 이전 알고리즘은 ID3(Interactive Dichotomizer 3)의 단점을 보완하고 발전시킨 알고리즘이다[14].

C4.5 알고리즘은 분할정복기법을 반복적으로 수행하여 트리구조를 생성하는 하향식 접근법에 속한다. 어떤 속성을 최상위 노드로 선택할 것인지를 결정하고 각각의 가능한 값에 대해 하나의 가지를 만드는 것이며, 이를 반복적으로 수행하여 트리를 생성한다. C4.5에서 노드의 결정은 정보이득(information gain)이나 이득비율(gain ratio)에 의해 결정된다.

주어진 데이터 로 구성되는 집합을 라고 하고 집합의 데이터는 개의 클래스로 구분되고 각 클래스는 이라고 가정하자. 클래스 에 속하는 데이터의 수를 라고 하면 엔트로피 즉 집합 를 분류하기 위해 기대되는 정보의 양은 아래 식과 같다.

$$\mathrm{I}(s_1, s_2, \ldots, s_m) = -\sum_{i=1}^{m} p_i \log_2(p_i)$$

그리고 개의 서로 다른 값 를 가지는 속성 를 가정한다. 속성 는 집합 를 개의 부분집합 로 분리할 수 있다. 여기서 특정 부분집합 는 속성 의 값을 포함한다. 그리고 부분집합 에서 클래스 에 속하는 데이터의 수를 라고 한다. 이 때 특성 A를 노드로 선택할 경우의 엔트로피는 식 아래 식과 같이 계산된다.

$$\mathrm{E(A)} = \sum_{j=1}^{v} \frac{s_{1j} + \cdots + s_{mj}}{s} I(s_{1j}, \ldots, s_{mj})$$

정보이득은 특정 속성을 선택함으로써 감소하는 엔트로피의 양 즉 복잡도가 감소된 정도를 의미한다. 따라서 아래 식을 통해 계산할 수 있다.

$$\mathrm{Gain(A)} = \mathrm{I}(s_1, s_2, \ldots, s_m) - E(A)$$

모든 속성에 대한 정보이득을 계산하여 비교하고 이 중 가장 큰 정보이득을 나타내는 속성을 노드로 선택한다.

2.4 신경망

인공 신경망은 노드(node)와 아크(arc)로 구성되며, 각각의 노드는 아크를 통해 입력(input)을 받고 이 입력을 변형하여 출력(output)을 생성한다. 다양한 신경망 모형은 네트워크 구조, 노드의 활성화 함수(activation function)와 학습 또는 훈련 규칙에서 차이를 보인다. 네트워크 구조는 노드의 수와 각 노드의 연결을 포함한다. 활성화 함수는 입력을 출력으로 변형시키는 역할을 한다. 훈련 규칙은 각 아크의 가중치(weight)를 어떻게 조정할 것인지를 결정한다. 모든 신경망은 훈련에 의해 형성되며, 지도 학습(supervised learning)은 훈련을 위해 데이터의 각 패

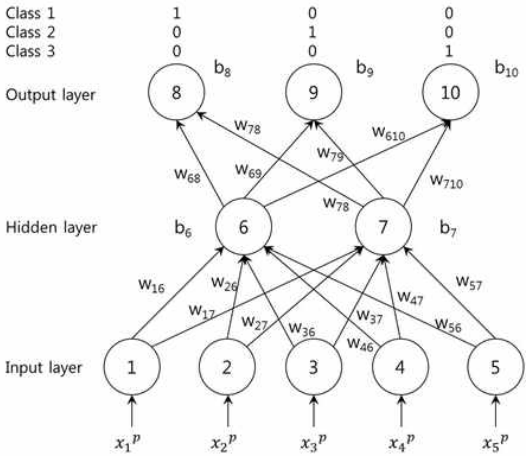
턴에 대해 목표로 설정된 클래스를 레이블의 형태로 포함한다. 반면 자율 학습(unsupervised learning)은 레이블이 없는 훈련 데이터를 이용한다.

최초의 신경망 모형은 McCulloch and Pitt에 의해 개발되었으며, 그 이후의 대부분의 모형은 이 모형을 토대로 발전하였다[15]. Rosenblatt의 단층 퍼셉트론(single- layer perceptron)과 Widrow and Hoff의 Adaline/ Madaline은 McCulloch and Pitt의 뉴런 노드(neuron node)를 여러 개 일렬로 사용하였으며, 노드가 여러 계층으로 분리된 다층 퍼셉트론(multi-layer perceptron)은 Rosenblatt와 Minsky and Papert에 의해 제안되었다[16-18]. 그러나 두 연구에서는 다층 퍼셉트론에 적합한 학습 규칙은 설명되지 않았다. Rumelhart et al. 이 다층 퍼셉트론의 학습에 적합한 알고리즘인 델타룰을 소개한 이후 신경망 모형에 대한 연구는 폭발적으로 증가하였다[19].

Feedforward 네트워크는 네트워트 상의 정보가 한 방향으로만 전송되는 네트워크를 의미하고 가장 널리 이용되는 신경망 모형이다. 그림 4는 3개의 계층으로 구성된 feedforward 네트워크의 구조를 나타내고 있다.

Feedforward 네트워크의 가장 대표적인 훈련 방법은 역전파 알고리즘(back propagation algorithm)이며, 일반화된 델타 규칙(generalized delta rule)이라고도 부른다.

역전파 알고리즘은 최급하강법(steepest descent method)의 특별한 형태로 볼 수 있다. 하지만 최급하강법은 위해서는 점진적인 이동의 반복을 통해 해에 접근하기 때문에 수행 속도가 느리다. 이 뿐만 아니라 역전파 알고리즘이나 gradient 기반 알고리즘들은 지역 최소치(local minima)에 수렴하는 경우가 있어 네트워크의 훈련에 어려움이 주기도 한다. 지역 최소치(local minimum)는 이웃하는 다른 해보다 더 작은 함수값을 나타내는 해 지점(solution point)이다.



<그림 4> 계층 feedforward 네트워크

본 연구에서 사용하는 FANN 모형은 역전파 알고리즘의 단점을 보강한 모형으로 효율적인 비선형 최소화 기법을 사용하며, 통계적 검증에 의해 네트워크의 크기를 최소화 할 수 있는 모형이다. FANN 모형의 특징은 다음과 같다[20].

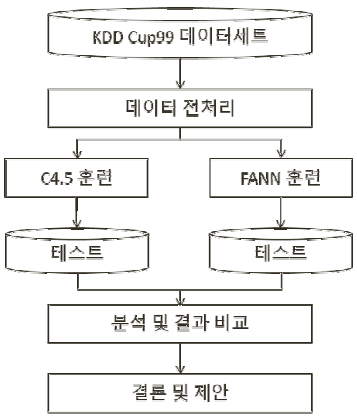
- FANN 모형은 효율적인 지역 탐색을 위해 비선형 최소화 기법인 limited memory BFGS(Broyden–Fletcher–Goldfarb–Shanno) 기법을 이용한다.
- FANN 모형은 훈련의 성과에 큰 영향을 줄 수 있는 역전파 알고리즘의 학습률이나 momentum 같은 조정을 위한 파라미터(parameter) 값은 가지지 않는다.
- FANN 모형은 네트워크 아크의 가중치 의 최소값(initial solution)을 난수(random number)로 정하지 않고 활성화 함수로 1차 함수를 사용함으로써 결과적으로 선형 회귀분석의 결과를 초기값으로 사용한다.
- FANN 모형은 네트워크의 은닉 노드 추가에 따른 오차제곱합(sum of squared errors)의 감소를 보장한다.
- FANN 모형은 모형의 결정을 위해 통계적 검증을

이용하여 최소한의 네트워크 크기를 결정하며, 훈련된 네트워크는 실제 문제에 적용될 때 통계적 신뢰를 가질 수 있다.

Ⅲ. 연구실험

3.1 연구과정

본 연구의 연구 과정은 그림 5와 같다. 본 연구에서 이용되는 데이터세트는 KDD Cup 99 데이터세트이며, 이를 연구의 목적에 맞게 샘플링하고 적절한 형태로 변형하는 데이터 전처리의 과정을 거친다. C4.5 알고리즘의 훈련과 테스트를 위해 Weka 데이터 마이닝 도구를 이용하였으며, FANN 모델의 훈련과 테스트는 별도의 작성된 프로그램을 이용하여 연구를 진행하였다.



<그림 5> 연구 과정

따라서 전처리를 거친 데이터는 Weka에서 이용하는 ARFF 파일과 FANN 프로그램에 적합한 MS-Access 데이터베이스 파일 형태로 저장하였다. ARFF 파일은 데이터와 함께 데이터 속성의 이름 데이터형 등을 헤더(header)로 포함하는 텍스트 형식의 파일이다.

두 알고리즘의 효율성을 확인하는 분석 및 결과 비교 단계에서는 개별 테스트에 대한 정확도(accuracy), 탐지율(detection rate), 오경보율(false alarm rate)를 계산하여 비교한다. 위 효율성에 대한 상세 설명은 분석 및 결과 비교 단계에서 언급하였다. 이를 바탕으로 결론을 도출하고 예상 가능한 제안을 서술하였다.

3.2 KDD Cup 99 데이터세트

본 연구에서는 각 기법을 이용한 분석을 위해 KDD Cup 99 데이터세트를 이용하였다. KDD Cup 99 데이터세트는 침입탐지 분야에서 가장 널리 이용되는 데이터이다[2, 13, 21].

1998년 DARPA(Defense Advanced Research Projects Agency)에서는 침입탐지 분야에 대한 연구를 목적으로 침입탐지 평가 프로그램(Intrusion Detection Evaluation Program)을 진행하였다. 이 프로그램에 사용된 표준 데이터세트는 군사 네트워크 환경에서 시뮬레이션을 통해 얻어진 네트워크 트래픽의 TCP dump 데이터이다. 본 연구에서 이용된 KDD Cup 99 데이터세트는 위 데이터의 1999년 KDD Cup 대회를 위한 버전이다. KDD Cup 99 데이터세트가 발표된 시점부터 현재까지는 10여년의 시간적 차이가 있다. 그 동안 다양한 공격유형이 등장하고 네트워크 환경 또한 크게 변화하였음에도 불구하고 KDD Cup 99 데이터세트는 여전히 침입탐지분야에서 널리 이용되는 대표적인 데이터세트이다[2, 22].

KDD Cup 99 데이터세트는 정상 데이터와 Probe, DoS, U2R, R2L의 4가지의 공격유형으로 구성되며, 상세 설명은 표 2와 같다.

<표 2> 네트워크 공격유형

공격유형	설 명
Probe	Probe 공격은 다른 공격을 준비하는 단계로서 네트워크상에 존재하는 IP 주소, 제공되는 서비스의 콘텐츠 또는 운영체제의 종류 등과 같은 정보를 획득하거나 특정 시스템의 취약점을 찾는데 중점을 둔다. 특히 Satan, Saint, Mscan과 같은 도구들은 숙련도가 낮은 공격자(attacker)들에게도 네트워크에 존재하는 장비들의 취약점을 신속하게 발견할 수 있도록 도와준다.
DoS	DoS(Denial of Service) 공격은 공격자가 특정 시스템 자원 전체를 점유하거나 대역폭 또는 시스템 자원에 장애를 발생시켜 정당한 사용자들의 접근을 거부하는 형태의 공격이다[11].
U2R	U2R(User to Root) 공격은 공격자가 특정 시스템에 일반 사용자 권한으로 접근한 다음 buffer-overflow와 같은 공격으로 취약점을 이용해 루트 권한을 획득하는 형태의 공격이다.
R2L	R2L(Remote to Local) 공격은 공격자가 호스트 장비의 취약점을 이용하여 인증되지 않은 접근권한을 다양한 방법으로 획득하여, 호스트 장비에 불법적으로 접근하는 공격 방법이다.

3.3 데이터 전처리

본 연구에서 이용하는 KDD Cup 99의 전체 데이터세트는 대용량의 데이터세트이다. 모든 데이터를 훈련과 테스트에 사용하기 위해서는 고성능의 장비가 요구된다. 따라서 본 연구에서는 일반적인 컴퓨터 환경과 연구의 목적을 고려하여 데이터세트의 전처리 과정을 거쳐 연구를 진행하였다.

훈련 표본 추출을 위해 10% KDD Cup 99 데이터세트를 이용하고, 테스트 표본 추출을 위해 Corrected 10% 테스트 데이터세트를 이용한다. 두 데이터세트는 41개의 속성 외에 세부 공격 패턴을 포함하고 있어 연구의 목적인 각 알고리즘의 탐지의 효율성을 평가하는데 유용하다. 두 데이터세트는 텍스트 파일의 형태로 구성되어 있다. 데이터세트의 표본 추출을 포함한 전처리를 위해 MS-Access를 이용하여 데이터베이스로 변환 하였다. 그리고 데이터세트의 공격 데이터에 대한 레이블은 37개의 세부 공격 패턴 중 하나로 저장되어 있으나, 이를 연구의 목적에 맞게 DoS, Probe, R2L, U2R의 공격유형으로 변환하였다[23]. 상세한 데이터세트의 용량과 레코드 수는 표 3과 같다.

본 연구에서는 10% KDD Cup 99 데이터세트로부터 9개의 훈련 표본을 추출하였고, Corrected 10% 테스트 데이터세트로부터 1개의 테스트 표본을 추출하였다. 각 표본은 10, 000개의 데이터를 포함한다.

<표 3> 연구에 이용된 데이터세트의 용량과 레코드 수

	용량(MB)	레코드 수	비 고
10% KDD Cup 99	71.4	494,021	-
Corrected 10% 테스트	45.0	311,029	공격패턴 레이블 포함

훈련 표본은 정상 데이터를 각각 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% 비율로 포함하고, 나머지 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 10%는 공격 데이터를 포함하는데, 이는 10% KDD Cup 99 데이터세트에서 나타나는 공격유형별 비율에 맞추어 랜덤으로 추출하였다.

테스트 표본은 Corrected 10% 테스트 데이터세트에서 나타나는 정상 데이터와 공격 데이터의 비율에 근거하여 랜덤으로 추출하였다.

C4.5 알고리즘의 훈련 및 테스트 과정에서 이용되는 Weka의 데이터세트는 ARFF 파일 형식으로 구성되어야 한다. 따라서 데이터베이스에 저장되어있는 표본 테이블을 ARFF 파일로 변환하였다. ARFF 파일은 콤마로 구분된 데이터 외에 데이터세트의 이름, 각 속성의 이름과 자료형, 데이터세트에 대한 설명 등을 헤더에 포함한다.

3.4 훈련 및 테스트

본 과정에서는 전처리 된 데이터세트를 바탕으로

C4.5와 FANN 알고리즘을 적용하여 침입에 대한 패턴을 생성하고 이를 바탕으로 테스트 과정을 수행하였다. 정상 데이터의 비율에 따라 생성된 9개의 훈련 데이터세트를 사용하여 데이터 훈련을 진행하였고, 이를 바탕으로 9회에 걸쳐 테스트를 진행하였다. 이 중 C4.5 기법은 Weka를 이용하였고, FANN 기법은 Ahn(1996)의 C++로 작성된 프로그램을 이용하였다[20].

Weka는 뉴질랜드의 와이카토 대학(Unversity of Waikato)에서 개발된 무료 데이터 마이닝 도구이다[24]. 이 도구는 데이터 마이닝을 위한 기계 학습 알고리즘으로 데이터 전처리, 분류, 회귀분석, 군집화 분석, 연관 규칙 분석 등의 기능을 제공한다. 본 연구에서 이용된 Weka의 버전은 3.6. 8이다.

C4.5는 Weka에서 기본적으로 제공되는 분류 항목의 J48 분류기를 이용하여 테스트를 진행하였다. J48 분류기의 모든 파라미터(parameter)는 초기값을 사용하였다.

Ⅳ. 분석 및 평가

분석 및 평가 단계에서는 테스트된 9개의 결과를 바탕으로 두 알고리즘의 효율성을 비교하기 위해 정확성, 탐지율, 오경보율을 계산하고 비교한다.

<표 4> 공격과 비공격 행동의 탐지와 식별

	공격으로 분류	정상으로 분류
공격 데이터	TP	FN
정상 데이터	FP	TN

그리고 일반적으로 침입탐지에서 효율성의 계산을 위해 테스트를 통해 분류된 각 클래스의 데이터의 수를 표 4의 기준에 의해 재분류한다. TP(True Positive)는 실제로 공격인 데이터가 공격으로 분류되는 경우, TN(True Negative)는 실제로 정상인 데이터가 정상으로 분류되는 경우, FN(False Negative)는 실제로 공격인 데이터가 정상으로 분류되는 경우를 의미한다. 그리고 FP(False Positive)는 실제로 정상인 데이터가 공격으로 분류되는 경우를 의미하고 이를 오경보(false alarm)라고 한다.

4.1 정확도(accuracy) 비교

정확도는 전체 데이터 중 올바르게 분류된 데이터의 비율 즉 전체 데이터 중 실제로 공격인 데이터가 공격으로 분류되었거나 실제로 정상인 데이터가 정상으로 분류된 경우의 비율을 의미한다. 그 식은 다음과 같다

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100\%$$

표 5는 C4.5와 FANN의 정확도를 나타내고 있다. 정상데이터 비율이 10%와 20%인 경우에는 FANN 기법이 높은 정확도를 나타내었고 정상데이터 비율이 30% 이상인 테스트에서는 C4.5가 높은 정확도를 나타내었다.

<표 5> C4.5, FANN의 정확도 비교

정상 데이터 비율	C4.5(%)	FANN(%)
10%	92.44	92.70
20%	92.54	92.57
30%	92.71	92.28
40%	92.71	92.31
50%	92.63	92.36
60%	92.34	92.06
70%	92.36	91.62
80%	92.27	92.16
90%	92.12	91.37
평균	92.46	92.16

두 기법의 정확도의 평균은 C4.5는 92.46%, FANN은 92.16%로 나타났으며, 0. 30%의 근소한 차이를 나타내었다. 또한 9번의 테스트 모두에서 두 기법의 정확도는 1.00% 이내의 차이를 나타내었다.

4.2 탐지율(detection rate) 비교

탐지율은 실제로 공격인 데이터 중 정확하게 공격으로 분류된 데이터의 비율을 의미한다. 그 식은 다음과 같다.

$$Detection\ Rate = \frac{TP}{TP+FN} \times 100\%$$

표 6은 C4.5와 FANN의 탐지율을 나타내고 있다. 정상 데이터 비율이 80% 인 테스트에서는 FANN이 90.60%로 높은 탐지율을 나타내었지만 나머지 8개의 테스트에서는 C4.5가 높은 탐지율을 나타내었다. 정확도의 경우와 같이 모든 테스트에서 두 기법의 탐지율은 1.00% 이내의 차이를 나타내었다. 두 기법의 탐지율의 평균은 C4.5는 90.85%, FANN은 90.60%로 나타났다.

<표 6> C4.5, FANN의 탐지율 비교

정상 데이터 비율	C4.5(%)	FANN(%)
10%	91.37	91.23
20%	91.52	91.18
30%	91.06	90.78
40%	91.05	90.82
50%	90.92	90.80
60%	90.52	90.46
70%	90.56	89.94
80%	90.41	90.60
90%	90.24	89.59
평균	90.85	90.60

4.3 오경보율(false alarm) 비교

오경보율은 실제로 정상인 데이터 중 공격으로 잘못 분류된 경우의 비율을 의미한다. 그 식은 다음과 같다.

$$False\ Alarm = \frac{FP}{FP+TN} \times 100\%$$

표 7과 그림 12는 C4.5와 FANN의 오경보율을 나타내고 있다. C4.5는 정상 데이터 비율이 10%와 20%일 때 각각 3.13%, 3.23%로 매우 높은 오경보율을 나타내었으나 30% 부터는 평균보다 낮은 오경보율을 나타내었다. C4.5의 오경보율 평균은 0. 90%, FANN의 오경보율 평균은 1.40%로 C4.5의 오경보율 평균이 C4.5에 비해 0. 50% 높게 나타났다.

<표 7> C4.5, FANN의 오경보율 비교

정상 데이터 비율	C4.5(%)	FANN(%)
10%	3.13	1.23
20%	3.23	1.69
30%	0.46	1.54
40%	0.41	1.54
50%	0.31	1.18
60%	0.15	1.33
70%	0.21	1.44
80%	0.05	1.39
90%	0.10	1.28
평균	0.90	1.40

V. 결론

현재의 침입탐지시스템은 높은 탐지율, 정확도, 그리고 낮은 오경보율을 요구한다. 본 연구에서는 데이터 마이닝 기법인 C4.5와 FANN을 이용한 침입탐지

모형의 정확도, 탐지율, 오경보율을 측정하고 비교하였다. 두 기법의 훈련과 테스트를 위해 침입탐지 분야의 대표적은 데이터인 KDD Cup 99 데이터세트를 바탕으로 연구의 목적에 맞게 전처리 과정을 수행한 후 연구에 이용하였다. 정상 데이터의 비율에 따른 9개의 훈련 데이터세트와 1개의 테스트 데이터세트를 구성하여 실험을 진행하였다. 특히 본 연구에서는 기존의 침입탐지 연구에 이용된 적이 없는 FANN 모형을 이용하여 C4.5와 효율성을 비교 분석하였다.

본 연구의 실험에서 C4.5가 상대적으로 높은 정확도를 나타내었으며, FANN도 비슷한 수준의 정확도를 보여주었다. 공격 데이터를 공격으로 탐지하는 비율인 탐지율은 두 기법이 거의 흡사한 수치를 나타내었으나 C4.5의 정확도가 다소 높게 나타났다. 두 기법은 정도의 차이는 있지만 정상 데이터 비율의 증가에 따라 탐지율이 낮아지는 추세를 보여주었다. 정상 데이터를 공격으로 탐지하는 비율인 오경보율은 FANN이 상대적으로 높은 수치를 나타내었다. 정상 데이터 비율이 10%와 20%인 경우에는 C4.5의 오경보율이 높게 나타났으나 그 이후에는 낮은 수치를 나타내었다.

두 기법 모두 높은 정확도와 탐지율을 나타내었으며, 오경보율 또한 높은 수치를 나타내었다. 실험의 과정에서 FANN은 하나의 은닉 노드를 사용하여 테스트되었음에도 불구하고 높은 정확도와 탐지율을 나타내는 것으로 보아 본 연구의 문제는 선형에 가까운 문제로 판단되며, 이것은 의사결정트리 기법인 C4.5에서 정확도와 탐지율이 높게 나타나는 것과도 관련이 있는 것으로 보인다.

본 연구의 결과를 바탕으로 한 시사점 및 향후 연구의 과제는 다음과 같다.

첫째 본 연구의 목적 중의 하나가 FANN 모형이 기존 연구에 사용된 기법과 비교하여 어느 정도 효율적인지를 파악하는 것이었다. FANN 모형은 탐지의 정확도와 탐지율에서 C4.5와 거의 유사한 성과를 보였다. 다만 오경보율이 높은 것은 어느 정도 과적합(overfitting)이 일어난 것으로 파악된다.

둘째 본 연구는 제한적인 수의 훈련 및 테스트 데이터를 이용하였다. 따라서 본 연구의 결과를 일반화하기 위해서는 KDD Cup 99 전체의 데이터세트를 이용하여 연구를 진행할 필요가 있다. 그리고 KDD Cup 99 데이터세트 외에 현재의 네트워크 환경을 반영하는 다양한 네트워크 데이터를 이용하여 연구를 진행하는 것도 중요한 의미가 있을 것이다.

셋째 본 연구에서는 각 모형을 보다 정교하게 구성하지 못한 점이 있다. 연구에서 사용된 세 모형은 모두 해당 소프트웨어에서 정한 default 값을 기초로 하였다. 향후 보다 나은 탐지 성과를 위해 각 모형을 보다 정교하게 작성할 필요가 있다.

넷째 높은 정확도와 탐지율뿐만 아니라 동시에 낮은 오경보율을 보이는 침입탐지시스템에 대한 연구가 필요하다. 본 연구에서 높은 정확도와 탐지율을 나타내는 기법은 오경보율이 높게 나타났다. 따라서 이 두 가지 평가 기준을 동시에 만족할 수 있는 방안을 모색할 필요가 있다.

참고문헌

[1] 박대우, “국가사이버보안정책에서 해킹에 대한 소고,” 한국정보보호학회논문지, 제21권, 제6호, 2011, pp. 24-41.

[2] Wu, S. and Yen, E., “Data Mining-based Intrusion Detectors,” Expert Systems with Applications, Vol. 36, No. 3, 2009, pp. 5605-5612.

[3] Bace, R. and Mell, P., NIST Special Publication on Intrusion Detection Systems, 2001.

[4] Singaraju, S., and Kalpana, P., A Precise Survey

on Intrusion Detection Systems, 2012.

[5] 신대철・김홍윤, “침입탐지 알고리즘 성능 최적화 및 평가 방법론 개발,” 디지털산업정보학회논문지, 제8권, 제1호, 2012, pp. 125-137.

[6] 양환석, “프로토콜 기반 분산 침입탐지시스템 설계 및 구현,” 디지털산업정보학회논문지, 제8권, 제1호, 2012, pp. 81-87.

[7] Beigh, B. M. and Peer, M. A. “Intrusion Detection and Prevention System: Classification and Quick Review,” ARPN Journal of Science and Technology, Vol. 2, No. 7, 2012, pp. 661–675.

[8] Kumar, Y. and Dhawan, S., “A Review on Information Flow in Intrusion Detection System,” International Journal of Computational Engineering and Management, Vol. 15, No. 1, 2012, pp.91–96.

[9] Denning, D. E., “An Intrusion-Detection Model,” IEEE Transaction on Software Engineering, Vol. 13, No. 2, 1987, pp. 222–232.

[10] Nguyen, H. A. and Choi. D., “Application of Data Mining to Network Intrusion Detection: Classifier Selection Model,” Challenges for Next Generation Network Operations and Service Management –Lecture Notes in Computer Science, Vol. 5297, 2008, pp. 399-408.

[11] Jalil, K. A., Kamarudin, M. H., and Masrek, M. N., “Comparison of Machine Learning Algorithms Performance in Detecting Network Intrusion,” Networking and Information Technology 2010 International Conference, 2010, pp. 221-226.

[12] Osareh, A. and Shadgar, B., “Intrusion Detection in Computer Networks Based on Machine Learning Algorithms,” International Journal of Computer Science and Network Security, Vol. 8, No. 11, 2008, pp. 15-23.

[13] Ibrahim, H. E., Badr, S. M., and Shaheen, M. A. “Phases vs. Levels using Decision Trees for Intrusion Detection Systems,” International Journal of Computer Science and Information Security, Vol. 10, No. 8, 2012, pp. 1-7.

[14] Quinlan, J. R., C4.5 : Programs for Machine Learning, Morgan Kaufmann Publishers, 1992.

[15] McCulloch, Warren S., and Walter Pitts., “A logical Calculus of the Ideas Immanent in Nervous Activity,” The Bulletin of Mathematical Biophysics, Vol. 5, No. 4, 1943, pp. 115-133.

[16] Rosenblatt, F., Principles of Neurodynamics. 1962.

[17] Widrow, B. and Hoff, M. E., Adaptive Switching Circuits. In: Neurocomputing: Foundations of Research. MIT Press, 1988.

[18] Minsky, M. and Papert, S., Perceptrons, MIT Press, 1969.

[19] Rumelhart, D. E., Hinton, G. E., and Williams, R. J. Learning Internal Representations by Error Propagation. Institute for Cognitive Science, University of California, San Diego, 1985.

[20] Ahn, B. H., “Forward Additive Neural Network Models,” PhD dissertation, Kent State University, Kent, OH, USA, 1996.

[21] 이한성, 임영희, 박주영, 박대희., “SVM 클러스터링 기반 적응형 침입탐지 시스템,” 퍼지 및 지능시스템학회논문지, 제13권, 제2호, 2003, pp. 237–242.

[22] Zarrabi, A. and Zarrabi, A., “Internet Intrusion Detection System Service in a Cloud,”

International Journal of Computer Science Issues, Vol. 9, Issue 5, No 2, 2012, pp. 308-315.
[23] Fares, A. H., Sharawy, M. I., and Zayed, H. H., "Intrusion Detection: Supervised Machine Learning," Journal of Computing Science and Engineering, Vol. 5, No. 4, 2011, pp. 305-313.
[24] Weka, http://www.cs.waikato.ac.nz/ml/weka/index.html

■ 저자소개 ■

조 성 래
Jo Seongrae

2013년 7월 Tai Woo Ree Engineering
2013년 2월 경상대학교 경영정보학과 (경영학석사)
2010년 8월 경상대학교 경영정보학과 (경영학학사)

관심분야 : 데이터마이닝, 빅데이터, 인공신경망
E-mail : seongraejo@gmail.com

성 행 남
Sung Haengnam

2004년~현재 경상대학교 경영대학 강사
2009년 2월 경상대학교 경영정보학과 (경영학박사)
2003년 2월 경상대학교 경영정보학과 (경영학석사)
2000년 8월 경상대학교 경영정보학과 (경영학학사)

관심분야 : 경영정보시스템, 전자상거래, e러닝, 소셜네트워크서비스
E-mail : haena@gnu.ac.kr

안 병 혁
Ahn Byunghyuk

1996년 9월~현재 경상대학교 경영정보학과 조교수, 경영경제연구소 리서치 펠로우
1996년 8월 켄트주립대학교 경영학과 (경영학박사)
1980년 2월 미시간주립대학교 경영학과 (MSinOR 석사)
1980년 2월 서울대학교 경영학과 (경영학석사)
1978년 2월 한국외국어대학교 러시아학과 (문학사)

관심분야 : 데이터베이스시스템, 데이터마이닝, 최적화모델
E-mail : bahn@gnu.ac.kr

논문접수일: 2015년 11월 16일
수 정 일: 2015년 11월 23일
게재확정일: 2015년 11월 25일