「정보시스템연구」 제15권 제4호
한국정보시스템학회
2006년 12월, pp. 211~224

데이터 마이닝의 비대칭 오류비용을 이용한 지능형 침입탐지시스템 개발*

홍태호** · 김진완***

〈목 차〉

Ⅰ. 서 론

오늘날 인터넷 사용의 폭발적 증가로 대부분의 시스템이 타 시스템과 네트워크로 연결되어 있는 상황 하에서 악의적인 해킹 또는 네트워크 침입은 그 시스템을 운영하는 조직뿐만 아니라 네트워크에 연결되어 있는 타 기관들에게도 치명적인 손해를 입힐 수 있는 구조를 갖고 있다. 따라서 정보시스템의 네트워크 환경의 급속한 발달로 인한 역기능을 줄이기 위한 네트워크 침입탐지시스템(Intrusion Detection Systems; IDS)의 필요성이 강조되고 있다. 네트워크 환경 하에서 정보시스템을 운영하는 회사, 학교, 정부 기관 등은 침입탐지시스템을 운영하여 외부의 해커 또는 공격자로부터 정보자원을 보호하고 있다.

침입탐지시스템(IDS)는 주어진 시스템에 대해 허락되지 않거나 비정상적인 활동들에 대해서 탐지하고 규명하는 기능을 가진 소프트웨어로 정의된다(Denning, 1987; Richards, 1999). 기존의 네트워크 침입탐지시스템에 사용되는 침입탐지모형은 전문가들의 지식을 이용한 네트워크 침입자 또는 해커의 행위를 탐지하여 비정상적인 정보시스템에 대한 접근을 제한하는 형태가 일반적이다. 이러한 전문가 지식에 기반에 침입탐지시스템은 알려진 네트워크 공

* 이 논문은 2003년도 한국학술진흥재단의 지원에 의하여 연구되었음(KRF - 2003 - 003-B00073)
** 부산대학교 경영학부 조교수, hongth@pusan.ac.kr
*** 부산대학교 경영학과 박사과정, kimjw@pusan.ac.kr

격 패턴 또는 비정상적인 활동에 대해서는 우수한 성과를 보이지만, 새로운 형태 또는 지식베이스에 등록되지 않은 공격 패턴에 매우 취약한 단점을 갖고 있다.

최근에 들어서는 통계적 모형 또는 인공지능 접근방법 등과 같은 데이터 마이닝 기법이 네트워크 침입탐지에 성공적으로 적용되고 있다(Heatly & Otto, 1998, Lam et al., 1996). 데이터 마이닝 기법 중 인공지능 접근방법은 기계학습기법의 범주로 분류되며 귀납적 학습방법(inductive learning), 인공신경망(artificial neural networks), 사례기반추론(case-based learning), 유전자알고리즘(generic algorithms) 등이 있다(Bose & Mahapatra, 2001). 또한 통계적 기법으로는 로짓모형, 판별분석, 연관규칙(association rules), SVM(support vector machines), 러프집합(rough set) 등이 있으며, 이중 러프집합은 데이터에 대한 사전 또는 추가적인 정보가 요구되지 않고, 데이터에 숨겨진 중요한 요인을 발견할 수 있으며 의사결정 규칙을 자연어를 통해 표현할 수 있어 실무적으로 사용하기 유리한 장점을 갖고 있다. 마찬가지로 인공지능 기법 중 귀납적 학습방법도 러프집합과 마찬가지로 데이터에 대한 가정이 없으며 의사결정 규칙을 제공한다. 러프집합과 귀납적 학습방법은 데이터마이닝 기법 중 설명력을 제공할 수 있는 데이터마이닝 기법인 반면 기법의 성과는 인공신경망의 성과에 미치지 못하는 것이 일반적이다. 하지만, 인공신경망은 타 데이터마이닝 기법과 비교해서 비록 성과는 우수하지만 설명력이 없다는 단점을 갖고 있다. 따라서, 귀납적 학습방법과 러프집합이 인공신경망에 비해서 성과가 실무적으로 사용하는 데 제한을 받을 정도가 아니라면 실무에서 적용하기에 더 적합한 방법론이다. 침입탐지시스템에서 러프집합과 귀납적 학습방법의 성과가 인공신경망의 성과와 비교해서 어떠한 차이가 있는지를 분석해 보고, 귀납적 학습방법과 러프집합의 적용가능성을 모색할 필요가 있다. Zhu(2001) 등은 네트워크 침입탐지분야에서 데이터마이닝 기법이 매우 우수한 성과를 보이는 것으로 보고했다. 데이터마이닝 기법을 이용한 침입탐지모형은 기존의 데이터를 이용하여 새로운 패턴을 발견할 수 있다는 장점이 있다.

데이터 마이닝은 추세나 패턴을 대 용량의 데이터를 통해서 발견하는 과정이다. 하지만, 지금까지의 침입탐지시스템에 대한 대부분의 연구는 침입 시도를 정확하게 분류하는 데는 한계는 갖고 있다. 즉, 침입탐지시스템은 근본적으로 오류를 가질 수 밖에 없는 치명적인 약점을 내포하고 있다. 결국 서버에 대한 치명적인 공격을 막기 위해 다양한 기법들을 사용하고 있지만, 침입탐지시스템은 오류를 생성한다. 침입탐지시스템의 목적은 정상적인 사용자와 침입자를 구분하는 것이다. 네트워크의 무수히 많은 다양성과 복잡성 때문에 모든 가능한 오류를 없애는 것은 불가능하다. 또한, 데이터 마이닝이 필터링에 따른 정보과부화를 줄이고 침입탐지시스템의 성과를 높일 수는 있지만, 두 가지 유형의 오류는 필수적으로 존재하게 된다. 첫 번째 오류는 "false positive errors"로 침입탐지시스템이 정상 패킷이나 활동을 침입으로 잘못 분류함에 따른 오류이다. "false positive errors"는 불필요한 대처방안을 수행하게 하여 정상업무를 방해하고 이로 인해 조직의 생산성이 감소하고 기회비용을 발생시킨다. 반대로,

"false negative errors"는 침입탐지시스템이 악의적인 공격이나 비정상적인 활동을 정상 사용자 또는 정상 활동으로 잘못 분류함에 발생하는 오류이다. "false negative errors"는 기업, 학교, 병원 등의 시스템이 네트워크로 연결되어 있기 때문에 한번 발생으로 엄청난 손실을 입을 수 있다. 그러나 침입탐지시스템의 오류에 따른 조직에 미치는 영향은 서로 상이하다. 이러한 오류의 특성에 데이터마이닝 기법의 적용을 위한 방법이 필요하며, 국내 네트워크 환경에 적합한 침입탐지시스템의 개발이 또한 필요하다. 따라서 본 연구에서는 최근 매우 중요하게 인식되고 있는 정보시스템 보안의 한 분야인 네트워크 침입탐지시스템에 데이터마이닝 기법을 적용하여 국내 특성에 맞는 침입탐지시스템을 개발하고자 한다.

Ⅱ. 이론적 배경

2.1 침입탐지시스템

침입이란 컴퓨터 시스템의 자원들에 대한 허가되지 않은 접속 또는 사용을 말한다(Esmaili et al., 1996). 정보자원을 보호하기 위한 침입탐지시스템(intrusion detection system)에 대한 정의는 다음과 같이 다양하게 되고 있다. 침입탐지시스템은 목표 시스템에 대한 허가되지 않았거나 변칙적인 활동들을 탐지하고, 식별하고, 대응하는 기능을 가진 소프트웨어이다(Richards, 1999). 침입탐지시스템의 목적은 실시간으로 또는 일괄처리 방식으로 보안 위반을 탐지하기 위한 메커니즘을 제공하는 것이다(Debar et al., 1992). 위반은 시스템을 파괴하려고 시도하는 외부인들에 의해 일어나거나, 권한을 오용하기 위해 시도하는 내부인들에 의해 일어난다(Weber, 1999). 침입탐지시스템은 다양한 시스템과 네트워크 자원들로부터 정보를 수집한 뒤에 침입과 오용에 관한 신호를 보내기 위해 정보를 분석한다(Lippmann and Cunningham, 2000). 침입탐지시스템에 의해 수행될 수 있는 주요한 기능으로는 사용자와 시스템 활동을 모니터링하고 분석하며, 중요한 시스템과 데이터 파일들의 무결성을 평가하며, 알려진 공격들을 반영한 활동 패턴들을 인식하며, 탐지된 활동에 자동적으로 대응하며, 그리고 탐지 프로세스의 결과를 보고한다. 이러한 침입탐지시스템은 전자상거래, 교육기관, 은행 등의 금융기관, 일반회사의 인트라넷 등에 다양하게 적용될 수 있다.

침입탐지는 탐지 방법에 따라서 크게 오용탐지(misuse detection)와 비정상적 탐지

<표 1> 침입탐지시스템의 적용영역 및 기대효과

적용영역	기대효과
· 전자상거래 영역 · 은행, 증권사 등의 금융기관 · 교육기관 · 기업 인트라넷	· 데이터 무결성(integrity)와 기밀성(confidentiality) 확보 · 사용자 프라이버시 보호 · 쇼핑몰 등의 시스템 보안성 향상 · 고객 DB 등의 보호 · 시스템에 대한 비권한자 활동(unauthorized activity) 보호 등

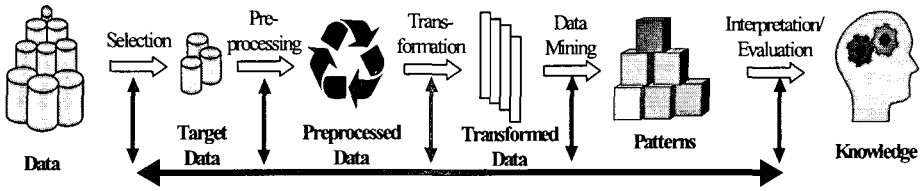
(anomaly detection)의 두 가지 범주로 구분할 수 있다(Zue et al., 2001). 첫째, 오용 탐지는 잘 알려진 공격들의 증적이나 패턴을 검색하는 방법으로 탐지한다. 오직 특징적인 증적을 남긴 잘 알려진 공격들만이 이 방법으로 탐지될 수 있다. 둘째, 비정상적 탐지는 정상적 사용자나 시스템 행동의 모델을 사용하고, 악의적인 가능성이 있는 경우에는 정상적인 사용과 편차가 발생하는 지를 탐지한다. 정상적 사용자나 시스템 행동에 관한 이 모델은 일반적으로 사용자 또는 시스템 프로파일로서 알려져 있다. 비정상적 탐지의 주요한 강점은 사전에 알려지지 않은 공격들을 탐지하기 위한 능력이 있다는 것이다.

침입탐지시스템은 분석하는 감사 데이터 출처의 종류에 따라서도 분류되어질 수 있다(Joo et al., 2003). 대부분의 침입탐지시스템은 공격들을 인지하거나 피하기 위한 접근법으로 네트워크 기반 침입탐지 또는 호스트 기반 침입탐지로 구분된다. 침입탐지시스템이 네트워크 트래픽에서 패턴을 찾는 경우에는 네트워크 기반 침입탐지로 분류된다. 침입탐지시스템이 로그 파일에서 공격 흔적을 찾는 경우에는 호스트 기반 침입탐지로 분류된다.

침입탐지시스템에 대한 최근의 많은 접근법들은 데이터마이닝 기술들을 활용하고 있다(Lam et al., 1996). 이러한 접근법들은 시스템에 의해 수집된 감사 추적의 대형 데이터 셋에 데이터마이닝 기술들을 적용하는 방법으로 탐지 모델을 구축한다 (Helman and Liepins, 1993). 데이터마이닝 기반 침입탐지시스템은 시스템의 일부분을 감시하는 감지장치(sensor)로부터 데이터를 수집한다. 감지장치들은 네트워크 활동, 사용자 프로세스에 의해 사용되는

<표 2> 침입탐지시스템에서의 데이터마이닝 응용사례

분류기준	침입탐지시스템 유형	데이터마이닝 기법
탐지 방법	오용탐지	CBR(Esmaili et al., 1996) NN(Endler, 1998) NN(Cannady, 1999) GA(Balajinath and Raghavan, 2001)
	비정상적 탐지	NN(Debar et al., 1992) NN(Bonifacio et al., 1998) NN(Endler, 1998) GA(Balajinath and Raghavan, 2001)
감사 자료 출처	네트워크 기반 탐지	NN(Kumar and Venkataram, 1997) NN(Endler, 1998) NN(Bonifacio et al., 1998) GA(Sinclair et al., 1999) NN(Lippmann and Cunningham, 2000) SOM(Lichodzijewski, 2002)
	호스트 기반 탐지	CBR(Esmaili et al., 1996) NN(Heatley and Otto, 1998) GA(Balajinath and Raghavan, 2001) SOM(Lichodzijewski et al, 2002)



<그림 1> 지식채굴 프로세스의 개요

시스템 호출, 그리고 파일 시스템 접속을 감시한다. 감지장치들은 탐지를 위해 사용될 수 있는 형식화된 데이터를 만들기 위해 수집된 원본 데이터로부터 예측적인 특징들을 추출한다. 감지장치에 의해 수집된 데이터들은 탐지기술을 사용하는 탐지기에 의해 평가된다. <표 2>는 침입탐지시스템을 위한 데이터마이닝 응용연구를 보여주고 있다.

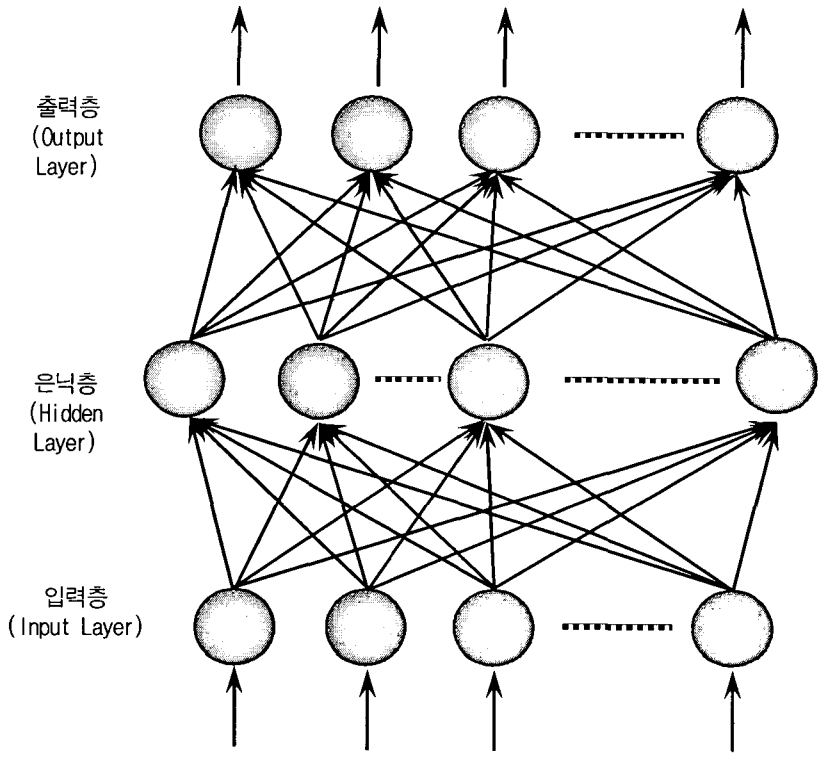
2.2 데이터마이닝 기법

데이터마이닝은 지식채굴과정(knowledge discovery process in database)의 한 부분으로 볼 수 있다. 지식채굴은 데이터에서 유효하고, 새롭고, 유용한 그리고 궁극적으로 이해할 수 있는 패턴을 확인하는 과정이다(Fayyad et al., 1996). 지식채굴 과정에 대한 완전한 방법론은 존재하지는 않지만 지식채굴은 <그림 1>과 같은 단계를 거쳐 발견된다고 할 수 있다.

지식채굴을 위해서는 먼저 데이터와 업무를 파악하고 분석을 위한 적당한 환경으로 옮겨서 데이터를 사용되어질 형태로 통합시키고 검토하게 된다. 이때 데이터에서 이상치와 에러를 제거하여 데이터를 정리하고 적용하는 기법에 따른 모형을 세우고 이를 검증하기 위한 기본 가정을 세운다. 이 가정을 검증하기 위해 데이터마이닝 기법을 적용하여 패턴이나 지식을 채굴해 나간다. 여기서 채굴된 지식의 유용성을 검증하고, 새로 발견된 지식의 사용을 위해 해석을 하게 된다. 지식채굴을 수행하기 위해서는 통계적 방법론과 인공지능기법을 이용하는 방법 등이 사용되고 있는데, 본 연구에서는 설명력을 제공할 수 있는 규칙생성이 가능한 귀납적 학습법, 러프집합을 적용해보고, 인공지능기법 중 성과가 가장 좋은 것으로 보고되고 있는 인공신경망과의 성과 비교를 통하여 귀납적 학습방법과 러프집합의 침입탐지시스템에서의 적용가능성을 제시하도록 한다.

일반적인 인공신경망은 다층퍼셉트론 (Multi layer perceptron)이라 불리우며, 다층퍼셉트론은 입력층과 출력층 사이에 하나 이상의 중간층이 존재하는 신경망을 지칭하는 것이다. 이때, 입력층과 출력층 사이의 중간층을 은닉층(hidden layer)이라 하며 network는 입력층, 은닉층, 출력층으로 연결되어 있다. 다층 퍼셉트론에서의 가중치는 지속적으로 전체 신경망이 만족할 만한 목표에 도달할 때까지 변하게 된다. 즉 인공신경망을 통해 계산된 출력값과 목표출력값(output)을 비교하여 그 차이(오차합수)를 최소화시킬 수 있도록 지속적으로 가중치를 조정하는 것이다.

이러한 신경망의 가중치의 조절은 역전파 알고리즘에 의한 학습과정을 통해 이루어진다. 역전파 학습 알고리즘의 기본 원리는 다음과



<그림 2> 인공신경망 모형

같다. 입력층의 각 유니트에 입력패턴을 주면, 이 신호는 각 유니트에서 변환되어 중간층에 전달되고 최후에 출력층에서 신호를 출력하게 된다. 이 출력값과 기대값을 비교하여 차이를 줄여 나가는 방향으로 연결강도를 조절하고, 상위층에서 역전파하여 하위층에서는 이를 근거로 다시 자기층의 연결강도를 조정해 나가게 된다. 인공신경망에 대한 자세한 내용은 Rumelhart & McClelland(1986)을 참고한다.

귀납적 학습(Inductive Learning)방법은 지식 구조를 나무 가지처럼 생성해서 예로부터 일반화된 개념을 추론하는 방식이다. 의사결정나무의 각 마디(node)는 속성을 부여 받고 잎은 개체로 부여받는다. ID3 알고리즘(Quinlan, 1986)과 이후에 발전된 C4.5(Quinlan, 1993)는 간단하지만 예로부터 학습을 하는 강력한 알고리즘이다. 귀납적 학습방법은 의사결정나무(decision tree)라고도 불리며 의사결정규칙(decision rule)을 나무구조로 도표화하여 분류 및 예측을 수행하는 분석방법이다. 이 방법은 분류 또는 예측의 과정이 나무구조에 의한 추론규칙(induction rule)에 의해서 표현되기 때문에, 다른 데이터마이닝 기법에 비해서 높은 설명력을 제공한다.

러프집합 이론은 1982년에 Pawlak에 의해서 부정확한 것, 모호한 것 그리고 부확실한 것에 대한 새로운 수학적 접근법으로 제안되었다(Pawlak, 1999). 러프집합 이론은 세상에 모든 개체들은 그들이 가진 어떤 정보로서 집합을 지울 수 있다는 가정에서 설립되었다. 동일한

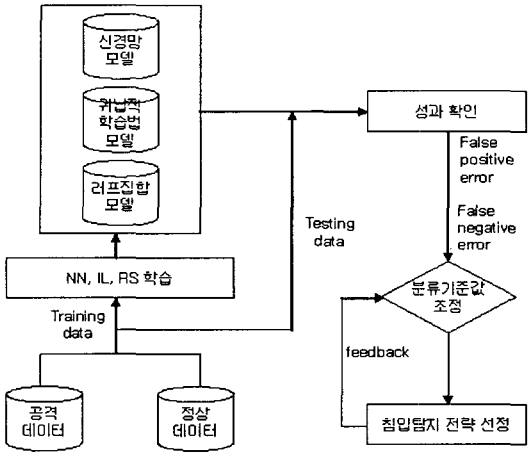
정보에 의해 특징지어지는 개체들은 그들의 정보로 인해 동일한 것으로 취급된다. 이 방법에서 생성된 동질성 관계가 러프집합 이론의 수학적 기초가 된다. 어떤 동질성을 가진 집합을 기본집합(elementary set)이라 하고, 이들이 어떤 전체집합에 대한 지식의 기본 단위가 된다. 어떤 기본집합의 합집합이 되는 집합을 일반집합(crisp set)이라 하고 그 외의 경우를 러프집합(rough set)이라고 한다. 결과적으로 각 러프집합은 개체들이 집합의 구성요소로써 명확하게 분류되지 않는 경우(상위근사) 또는 완전하게 분류되는 경우(하위근사) 등과 같은 경계영역개체(boundary-line cases)를 가진다. 러프집합 이론은 지식표현시스템(knowledge representation systems), 동질성 관계(indiscernible relations), 집합의 근사(approximation of set), 속성의 종속성(dependency of attributes), 속성의 축소(reduction of attributes), 그리고 의사결정 규칙(decision rules)으로 특징 지워질 수 있다(Slowinski and Zopounidis, 1995; Pawlak, 1999; Zue, et al., 2001).

러프집합의 주요한 장점은 데이터에 대한 사전 또는 추가적인 정보가 요구되지 않는다는 것이다. 이 방법론은 데이터에 숨겨진 중요한 요인을 발견할 수 있고, 의사결정 규칙을 자연어를 통해 표현할 수 있다. 또한 거대한 양의 정성적 및 정량적 데이터 모두를 다루는 능력을 제공하고, 비선형적이거나 비연속적인 기능적 관계를 모델화하는 능력은 복잡하고 다차원적인 패턴을 위한 강력한 방법론을 제공한다. 따라서 러프집합은 지식 획득, 예상과 예측 모델링, 그리고 의사결정 지원에서 성공적으로 적용되었다(Slowinski and Zopounidis, 1995; Pawlak, 1999; Zue, et al., 2001).

Ⅲ. 연구모형

본 연구에서 제안하는 연구모형은 <그림 2>와 같이 인공신경망, 귀납적 학습방법, 러프집합을 이용한 침입탐지모형을 개발하였다. 사용자에게 규칙을 제공함으로써 규칙결과를 쉽게 이해할 수 있고 설명 가능한 귀납적 학습방법과 이와 유사한 성격을 갖는 러프집합을 침입탐지시스템에 적용한다. 이 두 가지 데이터마이닝 기법과 데이터마이닝 기법 중 가장 성과가 높다고 보고되고 있는 인공신경망을 이용하여 침입탐지시스템을 개발한다.

또한, 침입탐지모형의 성과로 정상적인 사용자를 잘못 탐지하여 사용제한을 시키고 이에 대한 대응책을 실행하게 하여 일어나는 기회비용인 false positive error와 악의적인 침입자를 정상 사용자로 분류하여 시스템에 대한 접근을 허용함으로써 발생되는 정보 시스템 자산의 피해인 false negative error를 사용하였다. 침입탐지모형의 성과도 중요하지만, 오류의 종류에 따라 기업, 학교, 정부, 병원 등 조직에 미치는 영향도 큰 차이가 발생한다. 따라서, 본 연구에서는 false positive error와 false negative error에 따른 데이터마이닝 기법의 차이를 분석하도록 한다. 이와 더불어, 기업의 상황에 맞는 침입탐지 전략 선정을 위해서 분류기준값(treshold)을 0.3~0.7로 조정하면서 false positive error와 false negative error 의 변화에 따른 성과분석을 제시하였다. 이를 통해 보



<그림 3> 지능형 침입탐지시스템 연구모형

안담당자는 기업 환경에 적합한 최적의 비용으로 침입탐지 전략을 설정할 수 있다.

Ⅳ. 실험 및 결과

4.1 데이터

실험에 사용된 데이터는 통합 보안 서비스를 제공하는 Cyber-PATROL사의 IDS 센스로부터 무작위로 정상 패턴과 공격 패턴을 포함한 500개의 사례를 수집하였다. 이 표본은 다시 Training data와 Testing data로 분할하였다. 침입탐지를 위한 신경망, 귀납적 학습방법, 러프집합의 입력 변수들은 문헌 연구 및 전문가와의 인터뷰를 통해 <표 3>과 같이 선정하였다.

4.2 실험결과 및 분석

본 연구에서는 침입탐지를 위한 데이터마이닝 기법으로 인공신경망, 귀납적 학습법, 러프집합의 3가지 실험을 진행하였다. 3가지 실험은 모두 10-fold cross validation으로 수행하기 위해서 무작위로 서로 다른 데이터를 생성하였다. 10-fold cross-validation은 전체 표본을 10

<표 3> 모델에 사용된 변수

변수명	설명
Event Data	기록된 사건의 날짜와 시간
Protocol ID	사건과 관련된 프로토콜
Source Port	출처의 포트 숫자
Destination Port	도착지의 포트 숫자
Source IP Address	출처의 IP 주소
Destination IP Address	도착지의 IP 주소

<표 4> 신경망, 귀납적 학습법, 러프집합에 따른 성과표(오류 %)

모델	데이터셋	S1[1]	S2	S3	S4	S5	S6	S7	S8	S9	S10	Avg.
NN[2]	학습용	13.78	15.11	15.33	15.56	16.22	15.33	14.22	15.56	15.56	15.56	15.38
	검증용	18.00	26.00	6.00	20.00	14.00	20.00	14.00	22.00	22.00	10.00	15.38
IL[3]	학습용	16.00	14.22	17.11	17.33	12.89	16.44	15.56	16.22	15.56	16.89	15.82
	검증용	18.00	26.00	8.00	6.00	18.00	14.00	22.00	16.00	22.00	10.00	16.00
RS[4]	학습용	15.56	14.22	16.00	16.22	14.67	16.22	14.67	15.11	15.33	16.22	15.42
	검증용	14.00	26.00	8.00	6.00	22.00	12.00	20.00	24.00	22.00	10.00	16.40

1) S1 - S10, 10-fold 데이터 셋, 2)신경망, 3)귀납적 학습방법, 4)러프집합

개의 테스트 집합으로 분할하여 실험결과의 타당성을 확보하기 위한 방안이다. 데이터마이닝 기법중 학습을 하게되는 경우는 특정 데이터셋의 학습용 표본과 검증용 표본의 분포에 따라 성과가 상이하게 나올 수 있으므로, 실험결과의 타당성을 확보하는 전략으로 모드 데이터셋이 검증용 셋이 되도록 실험을 10회 반복하였다.

침입탐지를 위한 3가지 데이터마이닝 기법을 10-fold cross validation으로 실험한 결과는 <표 4>과 같다. 각 데이터 셋마다 성과가 조금씩 차이를 보이고 있지만 전체 평균으로 보면 3가지 기법은 성과에 큰 차이가 없는 것으로 보인다. 3가지 실험에 대해 변화의 유의성을 실제 통계적으로 검증하기 위해 멕네마르 검정(McNemar test)을 수행하였다. 멕네마르 검정은 사전사후(before and after) 형태의 실험에서 변화의 유의성을 검정하는 비모수적 방법이다. 이러한 변화의 유의성을 검정하기 위해 자료는 동일한 개체에서 두 가지 처리를 적용하여 나타나며, 2×2 분할표로 표현이 된다. 침입탐지를 데이터마이닝 기법간의 성과를 검정한 결과는 <표 5>과 같다. 멕네마르 검정의 결과를 살펴보면 모두가 유의수준 0.05에서 유의하지 않으므로 신경망, 귀납적 학습법, 러프집합의 성과차이는 있다고 할 수 없다.

<표 5> 멕네마르 검정 결과표

	IL	RS
NN	통계량(S) 0.1111 Pr > S 0.7389	통계량(S) 0.6923 Pr > S 0.4054
IL	.	통계량(S) 0.2857 Pr > S 0.5930

일반적으로 데이터마이닝 기법에서 이진 의사결정을 위한 분류기준값(threshold)으로 0.5를 사용하고 있다. 분류기준값의 결정은 학습용 셋의 분포에 가장 많은 영향을 받는다. 본 연구에서처럼 학습용 셋의 분포가 침입인 경우와 정상인 경우가 50:50인 경우에는 분류기준값을 0.5로 설정하는 것이 합당하다. 그러나 이러한 획일적인 분류기준값보다는 각 분류기준값별로 성과 분석을 통해 최대 이익을 가져올 수 있는 특정 분류기준값을 발견하는 것이 필요하다.

분류기준값의 변화에 따른 false positive error와 false negative error 결과의 변화는 <표 6>과 같다. 본 결과에서 귀납적 학습방법은 제외하였다. 그 이유는 귀납적 학습방법의

<표 6> 신경망과 러프집합의 성과변화비교

분류 기준값	신경망			러프집합			맥네마르(p)
	FPE[1]	FNE[2]	전체	FPE	FNE	전체	
0.3	26.00	6.40	16.20	29.60	5.60	17.60	0.07*
0.4	25.20	6.80	16.00	27.20	6.00	16.60	0.37
0.5	24.40	7.20	15.80	26.80	6.00	16.40	0.41
0.6	22.40	8.40	15.40	26.80	7.20	17.00	0.07*
0.7	20.40	24.00	22.20	22.80	20.40	21.60	0.75

* 유의수준 10%

1)FPE: false positive error, 2)FNE: false negative error

경우에 예측값들이 0과 1에 가깝게 동일하므로 분류기준값을 0.3~07로 조정하더라도 결과가 모두 똑같이 나타나기 때문이다. 신경망과 러프집합의 분류기준값에 따른 멕네마르 검정 결과를 살펴보면 0.3과 0.6이 유의수준 10%에서 통계적으로 유의한 결과를 나타내고 있다. 이것은 신경망과 러프집합의 성과에 차이가 있음을 보여주는 것이다. 분류기준값 0.3과 0.6에서 신경망이 러프집합 보다는 false positive error를 더욱 잘 탐지하고 있기 때문에 false negative error에서 러프집합에 비해 성과가 조금 낮지만 전체 성과는 높은 것으로 나타나고 있다. 하지만, false positive error는 기업, 학교, 정보, 병원 등의 네트워크침입 경보에 대한 대처에 따른 기회비용을 유발하는 오류이다. 반대로, false negative error는 실제 침입을 발견하지 못하는 오류로, 만일 침입을 막지 못했을 경우에 조직내에 미치는 영향은 false positive error로 야기되는 비용과 비교할 수 없을 정도로 막대하다. <표 6>에서 러프집합의 false negative error가 분류기준값 0.3에서 5.60%로 신경망의 오류보다 0.8% 포인트 적다. 즉, 오분류 비용이 false negative error가 false positive error보다 중요하다는 것을 감안하면, 오히려 신경망 모형보다 러프집합이 false negative error가 적고, 설명력을 제공하기 때문에 실무적으로 침입탐지모형에 더 적합하다고 할 수 있다. 이렇게 분류기준값을 통한 변화를 제공함으로써 보안담당자들은 자사의 환경에 맞추어 false positive error와 false negative error에 대한 최적의 비율로 유연하게 침입탐지 전략을 설정할 수 있을 것이다.

V. 결 론

오늘날 정보시스템 네트워크 환경의 급속한 발달로 인한 역기능을 줄이기 위한 네트워크 침입탐지시스템(intrusion detection system)의 필요성이 강조되고 있다. 기존의 네트워크 침입탐지시스템에 사용되는 침입탐지모형은 전문가들의 지식을 이용한 네트워크 침입자 또는 해커의 행위를 탐지하여 비정상적인 정보시스템에 대한 접근을 제한하는 형태가 일반적이다. 그러나 이러한 네트워크 침입탐지시스템은 규칙기반(rule-based)으로 구성되어 나날이 발전

되고 있는 네트워크 공격기술에 적절한 대응력이 부족한 실정이다. 반면에, 최근에 들어서 데이터마이닝 기법을 네트워크 침입탐지시스템에 적용하여 매우 우수한 성과를 보이는 것으로 보고되고 있다. 데이터마이닝 기법을 이용한 침입탐지모형은 기존의 데이터를 이용하여 새로운 패턴을 발견할 수 있다는 장점이 있다. 따라서 본 연구에서는 국내에서 사용된 실제 네트워크를 통한 침입공격에 관한 데이터를 수집하고, 3가지 데이터마이닝 방법론(신경망, 귀납적 학습법, 러프집합)을 적용하여 국내 데이터 특성을 고려한 네트워크 침입탐지모형을 설계하였다.

침입탐지를 위한 3가지 데이터마이닝 기법을 10-fold cross validation으로 실험한 결과로 도출된 성과는 유사한 것으로 나타났다. 그러나 일반적으로 데이터마이닝 기법에서 이진 의사결정(binary decision)을 위한 분류기준값(threshold)을 획일적으로 사용하지 않고 오분류 오류의 성격에 따라 최대의 이익을 가져올 수 있는 데이터마이닝 기법을 이용한 침입탐지 시스템 개발방안을 제시했다. 분류기준값을 변경함에 따라 데이터마이닝 기법간의 성과차이가 통계적으로 유의하게 나오고, 침입을 탐지함으로써 발생하는 비용을 감안한 침입탐지 모형의 개발 전략을 제시할 수 있었다. 또한, 본 연구에서 설계한 침입탐지시스템은 분류기준값을 통한 변화를 제공함으로써 기업의 보안담당자들이 자사의 환경에 따라 유연하게 침입탐지 전략을 설정할 수 있도록 의사결정을 지원할 것이다.

본 논문은 침입탐지시스템 개발을 위해 설명력이 높은 러프집합과 귀납적 학습방법의 적용타당성을 검증하고 인공신경망과 성과비교를 수행함으로써 인공신경망의 대안으로 러프집합과 귀납적학습방법을 제시하였다. 그러나, 최근 support vector machine 등과 같은 다른 데이터마이닝기법의 적용도 고려해보아야 할 것이며, 침입의 오분류에 따른 기회비용을 정확히 산정할 수 있는 위험분석방법도 연구되어야 할 것이다. 또한, 실무적으로는 실시간으로 대응 가능한 침입탐지시스템의 설계와 개발이 기대된다.

참고문헌

박기남, 이훈영, 박상국, "러프집합을 이용한 통합형 채권등급 평가모형 구축에 관한 연구", 한국경영과학회지, 제25권, 제3호, 2000, pp. 125-135.

Balajinath, B., Raghavan, S.V., "Intrusion Detection through Behavior Model," *Computer Communications*, Vol. 24, 2001, pp. 1202-1212.

Barber, R., "The Evolution of Intrusion Detection Systems-The Next Step," *Computer & Security*, Vol. 20, 2001, pp. 132-145.

Bonifaicio, J. M., Jr, Cansian, A.M., Carvalho, A.C.P.L.F., & Moreira, E. S., "Neural Networks Applied in Intrusion Detection Systems," *Proceedings of the IEEE International Joint Conference*, 1998, pp. 205-210.

Bose, I., Mahapatra, R., "Business Data Mining - A Machine Learning Perspective,"

Inofrmation & Management, Vol. 39, No. 3, 2001, pp. 211-225.

Debar, H., Becker, M., & Siboni, D., "A Neural Network Component for an Intrusion Detection System," *IEEE Computer Society Symposium Research in Security and Privacy*, 1992, pp. 240-250.

Denning, D.E, "An Intrusion Detection Model," IEEE Trans. S. E., Vol. 13, No. 2, 1987, pp. 222-232.

Endler, D., "Intrusion Detection. Applying Machine Learning to Solaris Audit Data," *Proceedings of Computer Security Applications conference*, 1998, pp. 268-279.

Esmaili, M., Safavi-Naini, R., Balachadran, B., & Pieprzyk, J., " Case-based Reasoning for Intrusion detection," *Proceedings of 12th Annual Computer Security Application Conference*, 1996, pp. 214-223.

Fayyad, U.M., Piatesky-Shapiro, G., & Smith, P., "The KDD Processes for Extracting Useful Knowledge and Learning from Volumes of data", *Communications of the ACM*, Vol. 39, No. 11, 1006, pp. 27-34.

Heatley, S.K., & Otto, J.R, "Data Mining Computer Audit logs to detect Computer Misuse," *International Journal of Intelligent Systems in Accounting, Finance, and Management*, Vol. 7, 1998, pp. 125-134.

Helman, P, & Liepins, G., "Statistical Foundations of Audit Trail Analysis for the Detection of Computer Misuse," *IEEE Transactions on Software Engineering*, Vol. 19, 1993, pp. 866-901.

Joo, D., Hong, T., & Han, I., "The Neural Network Models for IDS based on the Asymmetric Costs of False Negative Errors and False Positive Errors," *Expert Systems with Applications*, Vol. 25, 2003, pp. 69-75.

Lam, K.Y., Hui, L., & Chung, S. L., "A data Reduction Method for Intrusion Detection," *Systems Software*, Vol. 33, 1996, pp. 101-108.

Lichodzijewski, P., *Network Based Anomaly Detection using Self Organizing maps*, Technical Report, Nova Scotia: Dalhousie University Halifax, 2002.

Lippmann, R. P., Cunningham, R. K., "Improving Intrusion Detection Performance using Keyword Selection and Neural Network," *Computer Networks*, Vol. 34, 2000, pp. 597-603.

Pawlak, Z., "Rough Set Approach to Knowledge-based Decision Support," *European Journal of Operational Research*, Vol. 99, No.1, 1997, pp. 48-57.

Richards, K., "Network based Intrusion

Detection: A Review of Technologies," *Computer and Security*, Vol. 18, 1999, pp. 671-682.

Rumelhart, D.E., and McClelland, J.L., "*Parallel Distributing Processing: Exploration in the Microstructure of Cognition*", Vol. 1, Cambridge, MA: MIT Press, 1986.

Weber, R., *Information Systems Control and Audit*, Upper Saddle River, HJ: Princtice Hall, 1999.

Zhu, D., Premkumar, G., Zhang, X., & Hsien Chu., "Data Mining for Network Intrusion Detection : A Comparison of Alternative Methods," *Decision Sciences*, Vol. 32, No. 4, 2001, pp. 635-659.

홍태호 (Tae-Ho Hong)

현재 부산대학교 경영학부 조교수로 재직하고 있다. KAIST에사 산업공학사를 취득하였고 경영정보시스템을 전공하여 공학석사와 박사를 취득하였다. 딜로이트 컨설팅에서 컨설턴트로 재직했으며, 주요 관심분야는 인공지능 및 데이터마이닝, 지능형 의사결정지원시스템, CRM 등이다.

김진완(Jin-Wan Kim)

부산대학교 대학원 경영학과에서 경영정보전공으로 석사학위를 취득한 후, 부산시 인터넷방송국 바다 TV.com에서 e-Learning 팀장을 지냈다. 현재 부산대학교 대학원 경영학과에서 박사과정을 수료하고, 강사로 활동 중이며, 주요관심 분야는 데이터마이닝, 지식경영, 공급사슬관리 등이다.

<Abstract>

Intelligent Intrusion Detection Systems Using the Asymmetric costs of Errors in Data Mining

Tae-Ho Hong · Jin-Wan Kim

This study investigates the application of data mining techniques such as artificial neural networks, rough sets, and induction learning to the intrusion detection systems. To maximize the effectiveness of data mining for intrusion detection systems, we introduced the asymmetric costs with false positive errors and false negative errors. And we present a method for intrusion detection systems to utilize the asymmetric costs of errors in data mining. The results of our empirical experiment show our intrusion detection model provides high accuracy in intrusion detection. In addition, the approach using the asymmetric costs of errors in rough sets and neural networks is effective according to the change of threshold value. We found the threshold has most important role of intrusion detection model for decreasing the costs, which result from false negative errors.

Keywords: Intrusion Detection Systems, Data Mining, Neural Networks, Rough Sets, Inductive Learning

* 이 논문은 2006년 10월 30일 접수하여 1차 수정을 거쳐 2006년 12월 1일 게재 확정되었습니다.