데이터 마이닝을 이용한 공격 탐지 메커니즘의 실험적 비교 연구

정회원 김 미 희*, 준회원 오 하 영*, 정회원 채 기 준*

An Empirical Comparison Study on Attack Detection Mechanisms Using Data Mining

Mihui Kim* ***Regular Member***, Hayoung Oh* ***Associate Member***,
Kijoon Chae* ***Regular Member***

요 약

본 논문에서는 최신의 공격 유형을 잘 분류해 내고, 기존 공격의 변형이나 새로운 공격에도 탐지 가능하도록 데이터 마이닝 기법을 이용한 공격 탐지 모델 생성 방법들을 소개하고, 다양한 실험을 통해 탐지율 및 탐지 시간 측면에서 이 모델들의 성능을 비교한다. 이러한 탐지 모델을 생성하는데 중요한 요소로 데이터, 속성, 탐지 알고리즘을 꼽을 수 있는데, 실제 네트워크에서 수집된 NetFlow 데이터와 대량의 KDD Cup 1999 데이터를 사용하였다. 또한 탐지 알고리즘으로서 단일 지도/비지도학습 데이터 마이닝 기법 및 결합된 방법을 이용하여 탐지 모델을 생성, 비교 실험하였다. 시험 결과, 결합된 지도학습 알고리즘을 사용한 경우 모델링 시간은 길었지만 가장 탐지율이 높았고, 모든 경우 탐지 시간이 1초 내외로 실시간 탐지 가능성을 입증할 수 있었다. 또한 새로운 공격에 대한 이상탐지 결과로도 92% 이상의 탐지율을 보임으로 탐지 가능성을 입증할 수 있었고, SOM 기법을 사용하는 경우에는 새로운 공격이 기존 어느 공격에 유사한 특성을 갖는지에 대한 부과적인 정보도 제공하였다.

Key Words : Attack Detection Mechanism, Data Mining, Empirical Comparison, Detection Rate/Time

ABSTRACT

In this paper, we introduce the creation methods of attack detection model using data mining technologies that can classify the latest attack types, and can detect the modification of existing attacks as well as the novel attacks. Also, we evaluate comparatively these attack detection models in the view of detection accuracy and detection time. As the important factors for creating detection models, there are data, attribute, and detection algorithm. Thus, we used NetFlow data gathered at the real network, and KDD Cup 1999 data for the experiment in large quantities. And for attribute selection, we used a heuristic method and a theoretical method using decision tree algorithm. We evaluate comparatively detection models using a single supervised/unsupervised data mining approach and a combined supervised data mining approach. As a result, although a combined supervised data mining approach required more modeling time, it had better detection rate. All models using data mining techniques could detect the attacks within 1 second, thus these approaches could prove the real-time detection. Also, our experimental results for anomaly detection showed that our approaches provided the detection possibility for novel attack, and especially SOM model provided the additional information about existing attack that is similar to novel attack.

※본 연구는 정보통신부 및 정보통신연구진흥원의 대학 IT연구센터 육성·지원사업의 연구결과로 수행되었습니다.
* 이화여자대학교 컴퓨터학과 ({mihui, hyoh}@ewhain.net, kjchae@ewha.ac.kr)
논문번호 : KICS2005-03-117, 접수일자 : 2005년 3월 25일

Ⅰ. 서 론

네트워크 기술의 발달과 인터넷을 이용한 사업영역의 확장으로 인해 다양한 공격 트래픽이 발생하고 그 피해의 심각성이 점점 커짐에 따라 네트워크상의 컴퓨터에 대한 보안 및 실시간 침입 탐지 기술이 절실히 요구되고 있다. 따라서 공격이 공격 대상 시스템 혹은 네트워크를 완전히 마비시키기 전에 실시간으로 공격을 탐지하는 것이 현재로선 가장 중요한 보안 방법 중의 하나다. 이처럼 다변화하는 공격에 대한 실시간 대응책의 필요성이 대두되면서, 최근 공격에 대한 다각적인 보안 메커니즘에 관한 연구들이 진행되어 왔다. 그러나 현재까지 제안된 보안 메커니즘들은 다양한 공격 중 일부에 대해서만 유효하거나, 기존 공격의 일부 변형에 대해 적절히 대응하지 못하며, 공격 툴에서 제공하는 자동화된 또는 지능적 공격에 효과적으로 대응하지 못하고 있다.

본 논문에서는 다양한 공격의 심각성을 인식하고, 이러한 최신 공격을 잘 탐지하며 기존 공격의 변형이나 새로운 공격에 대해서도 탐지 가능하도록 적응적인 탐지 모델을 만들기 위해 다양한 데이터 마이닝 기법을 적용하여 그 성능을 탐지 정확성 측면과 실시간 탐지 측면으로 비교 실험하고자 한다. 이러한 비교 실험을 통하여 각 데이터 마이닝 기법의 장단점을 비교하여, 각 성능 측면에서의 효율적인 기법을 소개하고자 한다.

다양한 공격에 대한 탐지 모델의 성능을 좌우하는 중요 변수로 1)데이터, 2)속성, 3)탐지 알고리즘을 꼽을 수 있다. 본 논문의 비교 실험에서는 다양한 모델 생성 및 시험을 위하여 실제 네트워크에서 수집한 공격 및 정상 트래픽에 대한 NetFlow 데이터와 대량 데이터 실험을 위해 KDD Cup 1999 데이터를 사용하였다. 또한 탐지 모델 성능에 사용되는 속성 선정의 중요성을 실험을 통하여 입증하고자 한다. 마지막으로 탐지 알고리즘으로서 크게 두 가지 부류로 나누어 지도학습(Supervised Learning) 데이터 마이닝 기술과 비지도학습(Unsupervised Learning) 데이터 마이닝 기술에 의한 모델 성능을 비교 실험하고자 한다.

데이터 마이닝 기술은 대량의 데이터로부터 의미있는 유용한 정보를 추출하기 위해서 데이터베이스, 기계 학습(Machine Learning), 정보 이론(Information Theory), 통계, 가시화(Visualization) 기법들을 통합한 기술이다. 이러한 기술을 이용하여 대량의 데이터에서 알려지지 않은 일정한 공격 패턴을 찾아내어 공격을 탐지할 수 있어서, 군집화, 분류, 연관규칙 등의 데이터 마이닝 기법은 침입 탐지 분야에서 이상탐지(Anomaly Detection)와 관련하여 많이 연구되어 왔다[1]. 본 논문에서의 선행 연구로서 분산 서비스거부 공격을 탐지하기 위하여 플로우 기반의 NetFlow 데이터를 실제 네트워크에서 수집하여 이용하고, 탐지 모델 생성을 위해 지도학습의 데이터 마이닝 기법을 결합하여 적용한 탐지 구조를 제안하였다[2, 3]. 본 논문에서는 다양한 공격에 대한 탐지 모델을 생성함에 있어서, [3]의 기법과 비지도학습 데이터 마이닝 기법인 자기조직화지도(SOM, Self-Organizing Map) 기술을 각각 적용하여 그 성능을 비교하고자 한다.

본 논문의 구성은 다음과 같다. 1장의 서론에 이어, 2장에서는 관련 연구로서 기존에 제안된 공격 탐지 기법에 대해 비교 설명하고자 한다. 3장에서는 공격탐지의 중요 변수로서 실험에 사용된 데이터와 모델 생성에 사용된 속성 및 속성 추출 방법, 지도학습/비지도학습 데이터 마이닝 기술에 의한 탐지 모델 생성 방법에 관하여 설명하고자 한다. 4장에서는 두 공격 모델의 성능을 탐지 정확성 및 탐지 시간 측면에서 비교하고, 결론으로써 본 논문을 마치고자 한다.

Ⅱ. 관련 연구

최근, 네트워크 기술 및 서비스 발달과 함께 이에 대한 공격 또한 다변화되고 있다. 시스템 및 네트워크 자원을 순간적으로 마비시키는 분산 서비스 거부 공격(DDoS, Distributed Denial of Service) 및 각종 시스템의 취약점을 이용한 다양한 변종 웜, 바이러스 등이 이러한 공격에 포함된다. 이들 공격에 대한 예방책이 최우선이지만, 완벽한 예방이 불가능하므로 공격 발생시 실시간으로 공격을 탐지하여 차단하는 방법이 필수적이다. 이러한 취지에 맞춰 기존에 제안된 탐지 기법으로는 근원지에서의 공격 탐지 기법, 통계적 기법을 이용한 공격 탐지 기법, 데이터 마이닝 기술을 이용한 공격 탐지 기법 등이 있다.

근원지에서의 공격 탐지 기법으로 로스앤젤레스 소재 캘리포니아 대학교에서 개발한 'D-WARD' (DDoS netWork Attack Recognition and Defense)라는 기법은 보안 소프트웨어를 네트워크 게이트웨이에 설치하고, 게이트웨이를 통해 밖으로 나가는

트래픽에 대한 감시를 집중하는 방안이다[4]. D-WARD는 소스 네트워크에 위치하여 자율적으로 공격을 탐지하고 방어하는 공격 방어 시스템이다. 소스 네트워크와 인터넷 사이의 양방향 트래픽 플로우를 감시함으로써 공격을 탐지하여 패킷율 제한(rate-limiting) 방법으로 트래픽 양을 조절한다.

D-WARD는 많은 공격들을 성공적으로 탐지하고 방어하며, 공격 이전이나 후에 시작된 합법적인 연결에는 좋은 서비스를 유지시켜준다. 그러나 다음과 같은 몇 가지 단점이 있다. 첫째로 짧게 반복되는 공격에 대해서는 이전 공격을 메모리에 저장해 놓지 않기 때문에 매번 계산해야 하는 점이 비효율적이다. 둘째, TCP, ICMP 트래픽과 달리 역방향 트래픽이 없는 UDP 트래픽의 경우, 공격을 판단하는 기준 변수가 달라 탐지에 한계가 있다. 셋째, D-WARD의 위치에 따라 피어 간에 양방향 비대칭 통신 경로를 사용하는 경우, 플로우를 정확히 분석하여 탐지하기가 어렵다. 마지막으로 공격 중에 시작한 합법적인 플로우는 패킷율 제한으로 인해 제대로 서비스해 주지 못한다는 한계가 있다.

또 다른 근원지에서의 공격 탐지 기법으로 MIB 정보를 이용한 공격 탐지 방법[5]은 공격 대상이 공격을 받기 전에 미리 공격을 탐지하는 방법의 하나로 네트워크 관리 시스템(Network Management System : NMS)을 이용하여 MIB 정보를 감시함으로써 이루어진다. 공격 각 단계마다 발생하는 이벤트에 대해 MIB 정보의 중요 변수들의 변화를 살펴보고 공격을 탐지할 수 있는 방법이다.

분산 서비스거부 공격을 탐지하기 위해서 MIB 정보의 중요 변수로 공격자와 마스터, 마스터와 슬레이브 간의 명령 트래픽을 사용하나, 실제 네트워크 트래픽 중에서 명령 트래픽은 극히 일부이며, 또한 명령 트래픽의 정보가 쉽게 수정 가능하다는 점 때문에 현실적으로 이러한 정보로 큰 변화를 탐지하는 것이 불가능하다. 또한 일반적으로 공격 네트워크의 대부분 타겟 네트워크와 멀리 떨어져 있기 때문에 공격 탐지를 위해서는 도메인 간의 MIB 정보를 교환하는 작업이 필요하나, 아직까지는 도메인 간의 협동이 이루어지지 않는 실정이다.

통계적 방법을 이용한 공격 탐지 연구에서는 공격 도구에 의해 생성된 공격 트래픽들은 정상 트래픽과 구별되는 특징을 갖고 있으며, 통계적인 기준을 이용하여 중심 라우터에서 정상과 공격 트래픽을 구별할 수 있다고 가정하였다[6]. 각 소스 IP 주소별 패킷의 빈도수를 계산하고 이를 바탕으로 소스 주소의 분포 모델을 만들었다. 이 분포를 이용하여 패킷의 소스 IP 주소가 공격 도구에 의해서 랜덤하게 선택된 것인지 여부를 측정할 수 있다. 실제 정상 트래픽에서의 소스 주소의 분포와 공격 트래픽의 소스 주소의 분포가 다르다는 점을 이용하여 공격임을 탐지하였다. 여기서 사용된 통계 기법은 엔트로피 통계와 카이제곱 통계 방법이다. 그러나 갈수록 공격 도구가 지능화 되면서 스푸핑의 랜덤 정도를 조절 가능하게 되고, 이로 인해 정상과 공격의 소스 주소 분포를 구분 짓는 것이 어려워지고 있다. 그리고 다양한 유형의 공격들이 존재하기 때문에 단순히 소스 주소만을 모니터링 하는 것은 모든 공격 유형을 탐지해 내는데 충분하지 못하다.

데이터 마이닝 기법은 일반적인 침입탐지를 위하여 다양한 연구가 진행되어 왔다. Wenke Lee[7]는 침입탐지를 위하여 sendmail과 tcpdump 데이터를 이용해 frequent episode 분류 기법과 연관규칙 기법을 통해 침입 모델을 생성하여 이를 시험하였고, 데이터 마이닝 적용시 많은 모델링 시간이 걸리는 점을 고려하여 학습(learning) 에이전트와 탐지 에이전트로 구성된 침입 탐지 구조를 제안하였다. 이 논문은 tcpdump 네트워크 트래픽을 이용해 다양한 공격을 탐지할 수 있는 기본적인 연구 과정을 소개하였나 최근 더욱 강력해진 공격들을 탐지하기 위한 다각적인 실험이 필요하며, tcpdump 데이터를 이용하면 많은 전처리 과정이 필요하므로 빠른 탐지가 어렵다는 단점이 있다. 이외에도 침입탐지의 두 분류인 이상탐지와 오용탐지(Misuse Detection)를 위해 신경망을 사용한 연구[8]와 새로운 공격을 인식할 수 있도록 신경망 기반의 침입탐지 시스템[9]이 제안되었다.

Ⅲ. 공격 탐지의 중요 변수

공격에 대한 탐지 시스템에서 성능을 좌우하는 중요 변수로서 **1)데이터, 2)속성(Attribute), 3)탐지 알고리즘**을 꼽을 수 있다. 본 장에서는 비교 실험에 사용되는 이들 요소를 자세히 설명하고자 한다.

3.1 실험 데이터

공격 탐지 시스템의 성능을 좌우하는 첫번째 요소로서 탐지 모델 생성 및 탐지 시 사용되는 데이터의 종류는 그 시스템의 탐지 정확성 및 탐지 시간에 큰 영향을 미친다. 즉, 실제 네트워크의 정상과 비정상 패턴을 잘 반영해야 하고 많은 전처리

과정을 요구하지 않는 데이터이어야 한다. 예를 들어, 정상 패턴인데도 특정 시간대에는 특정 서비스가 증가할 수 있는데 이러한 경우를 공격으로 오판하지 않도록 실제 네트워크의 특수 정상 패턴도 잘 훈련되어야 실제 탐지 시 작은 오판률(False Positive)의 결과를 얻을 수 있다. 이와는 반대로 실제 공격 시 공격 데이터와 정상 데이터가 섞여 분포되어 있는데 이러한 경우에도 높은 공격 탐지율을 가질 수 있도록 다양한 실제 데이터를 사용하여 탐지 모델을 구축해야 할 것이다.

이를 위해서 본 논문에서 사용한 첫번째 실험 데이터는 tcpdump의 정보를 사용할 때처럼 많은 전처리 과정이 필요 없는 NetFlow 데이터이다. 여기에서 한 플로우(Flow)는 소스/목적지 IP 주소, 소스/목적지 Port, 3계층 헤더의 Protocol 타입, ToS(Type of Service) 값, 입력 논리 인터페이스에 의해 유일하게 정의된다. NetFlow는 본래 네트워크 서비스 이용 과금을 위해 개발된 프레임워크로 플로우 단위로 정보를 제공한다. NetFlow는 시스코 시스템에서만 제공되긴 하지만, NetFlow 대신 플로우 기반 다양한 통계 정보를 제공하는 sFlow를 사용할 수도 있다. sFlow(RFC 3176)[10]는 IETF에서 표준화된 기술로 NetFlow와 유사한 기능들을 제공한다. 최근 가장 위협적인 공격 중에 하나인 분산 서비스거부 공격의 가장 큰 특징 중의 하나가 갑자기 많은 플로우를 발생시켜 네트워크 시스템이나 서버 시스템 자체의 리소스를 고갈시키는 것이다. 또한 개인용 컴퓨터를 공격하는 최신의 웜들도 많은 플로우를 발생시켜 네트워크를 마비시키는 서비스거부 공격의 형태를 보이고 있다.

그러므로 플로우별 네트워크 데이터 정보를 출력해 주는 NetFlow 데이터를 이용하는 것은 큰 의미가 있고, 또한 5000대 이상의 호스트를 외부에 연결하는 라우터에서 수집한 충분한 정상 트래픽과 최신 분산 서비스거부 공격 툴을 사용하여 얻은 정상 트래픽 속의 공격 트래픽은 적절한 탐지 모델 생성과 본 논문에서 수행하는 실제적인 탐지 결과의 신빙성을 제공하여 준다. 그림 1은 공격 트래픽을 수집하기 위하여 수행한 공격 시나리오의 그림이다. 공격에는 널리 알려진 분산 서비스거부 공격 툴인 TFN2k, Stacheldraht, Synk4를 사용하였으며, 공격 타입은 실제 공격에서 가장 큰 비율을 차지하고 있는 TCP SYN Flood 공격을 비롯하여 UDP Flood 공격, ICMP Flood 공격, 그리고 TCP SYN Flood와 UDP Flood 공격이 동시에 일어나는 MIX 공격을 수행하였다. 그리고 소스 주소를 위조하는 스푸핑 정도는 공격 툴별로 제공해 주는 옵션에 따라 다양하게 변화시켜 주었다. 공격은 NetFlow 데이터를 제공하는 라우터를 기준으로 학교망 외부와 내부에 각각 에이전트들을 설치해 두고, 내부에서 외부로, 외부에서 내부로의 양방향 공격을 모두 실험해 보았다. 현실적으로 심각한 피해를 초래하는 공격은 직접 수행해 볼 수 없기 때문에 단시간 동안만 그리고 3~5개의 정도의 공격 에이전트를 이용하여 공격을 수행하였다. 따라서 실제 공격 시에는 실험 결과보다 더 큰 폭의 증감이 나타날 것으로 예상된다.

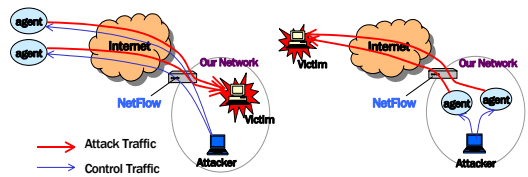

그림 1. 공격 시나리오

실험에 사용된 두 번째 데이터는 KDD Cup 1999 데이터이다[11]. 이 데이터의 사용 목적은 첫 번째 실험 데이터로 사용된 NetFlow 통계 데이터가 실제 네트워크에 공격을 수행하여 수집한 소량의 데이터이므로 대량의 데이터에 대한 실험을 위해 KDD Cup 1999 데이터를 실험 데이터로 선정하였다. KDD Cup 1999 데이터는 미국 군사 네트워크 상에서 시뮬레이션을 통해 만들어진 총 41개 속성의 레이블링된 실험 데이터이다. 이 데이터의 집합 중 kdd_10% data와 kdd_corrected data를 이용하였다. 각 데이터는 단일 TCP 연결에 대한 기본 속성, Content 속성, 2초 타임 윈도우에 의해 계산된 트래픽 속성으로 이루어져 있다. 실험에 사용된 KDD 공격 데이터로는 네트워크 기반의 서비스거부(DoS) 공격과 probing 공격 트래픽을 사용하였다.

표 1은 실험에 사용된 데이터의 종류와 그 구성을 정리한 것이다. NetFlow 데이터는 각각 100개씩 4가지 유형으로 레이블링된 총 400개의 데이터이고, 각 데이터는 5분간의 트래픽에 대한 플로우 정보이다. DDoS 공격 유형인 TCP SYN Flood, UDP Flood, MIX 공격과 정상 데이터인 normal을 포함하고 있다. KDD Cup 1999 데이터는 성능과 모델링 시 사용된 데이터의 양 사이의 관계를 알아보기 위해, 각각 100개씩 10가지 유형으로 레이블링된 총 1000개의 데이터와 각각 5000개씩 10가지 유형

표 1. 실험 데이터의 종류와 구성

비교항목		NetFlow	KDD Cup 1999
출처		실제 수집	미국 군사 네트워크
사용목적		DDoS 공격 탐지	DoS 및 probing 공격 탐지
데이터 포함유형 (각100개)	공격	TCP SYN Flood, UDP Flood, MIX	-DoS: back, Neptune, pod, smurf, teardrop -Probing: ipsweep, nmap, portsweep, satan
	정상	normal	normal
총 데이터 사이즈		400행(각 100개) : NetFlow_400	-1000행(각 100개) : KDD_1000 -50000행(각 5000개) : KDD_50000
실험비율	학습	70%	70%
	테스트	30%	-30% -kdd_corrected data 100%

으로 레이블링된 총 50000개의 데이터를 사용하였다. 이 데이터에는 DoS 공격 유형인 back, Neptune, pod, smurf, teardrop과 probing 공격 유형인 ipsweep, nmap, portseep, satan과 normal 데이터를 포함하고 있다. 각 공격 탐지모델을 생성하는데 각 데이터의 70%를 사용하였고, 공격 탐지 모델의 성능을 테스트하는데 나머지 30%를 사용하였다.

3.2 속성

공격 탐지 시스템의 성능을 좌우하는 두 번째 요소는 데이터를 구성하는 속성이다. 예를 들어 분산 서비스거부 공격의 가장 큰 특징이 소스 IP 주소를 스푸핑하는 것이므로 통계적 방법을 이용한 탐지 메커니즘[6]에서처럼 주로 소스 IP 주소의 분포를 탐지를 위한 중요 속성으로 사용한다. 그러나 실제적으로 많은 소스 네트워크의 라우터에서는 대부분의 스푸핑 패킷을 필터링할 수 있도록 입/출력단(Ingress/Egress) 필터링 기능을 사용하고 있어서 소스 IP 주소의 분포가 중요 속성이 되지 않을 수 있다. 이러한 이유로 선행연구[3]의 의사결정트리에 의한 중요 속성 자동 추출 결과는 소스 IP 주소 분산 값을 중요 속성으로 주출되지 않았다. 그러므로 각 네트워크 정상 패턴과 그 네트워크에서의 공격 패턴에 따라 공격 탐지 모델이 생성되어야 하며, 이를 잘 수행하기 위한 속성의 종류 또한 네트워크 특성에 따라 달라질 수 있다.

본 논문의 실험에서 사용되는 데이터의 기본 속성은 표 2, 표 3과 같다. 비교 실험을 위하여 사용된 첫 번째 실험 데이터인 NetFlow 데이터로는 5분 단위의 통계 정보를 이용하였고, 이는 표 2와 같이 22개의 속성을 제공해 주는데, 우선 휴리스틱하게 분석하여 공격의 특징이 잘 나타나는 13개의 속성을 후보속성으로 추출하여 사용하였다.

표 2. NetFlow 데이터의 속성

	NetFlow 데이터의 속성	후보 속성
1	플로우수 (Flow)	X
2	옥텟수 (Octets)	X
3	패킷수 (Pkts)	X
4	옥텟수/플로우수 (O/F)	O
5	패킷수/플로우수 (P/F)	O
6	TCP 트래픽의 소스 포트 분산 (srcTport)	O
7	UDP 트래픽의 소스 포트 분산 (srcUport)	O
8	TCP 트래픽의 목적지 포트 분산 (dstTport)	O
9	UDP 트래픽의 목적지 포트 분산 (dstUport)	O
10	소스 IP 주소 분산 (srcVar)	O
11	TCP 전체 플로우수 (Tflow)	X
12	TCP 전체 옥텟수 (Toctets)	X
13	TCP 전체 패킷수 (Tpkt)	X
14	TCP 트래픽의 옥텟수/플로우수 (To/f)	O
15	TCP 트래픽의 패킷수/플로우수 (Tp/f)	O
16	UDP 전체 플로우수 (Uflow)	X
17	UDP 전체 옥텟수 (Uoctets)	X
18	UDP 전체 패킷수 (Upkt)	X
19	UDP 옥텟수/플로우수 (Uo/f)	O
20	UDP 패킷수/플로우수 (Up/f)	O
21	TCP 트래픽 비율 (Tratio)	O
22	UDP 트래픽 비율 (Uratio)	O

13개 후보속성의 추출 기준은 다음과 같다. 정상 트래픽보다는 분산 서비스거부 공격 시 플로우수가 증가하지만, 정상인 경우에도 플로우수의 절대값은 크게 변화하므로 이는 후보 속성에서 제외시켰고, 대신 공격 시 한 플로우를 구성하는 옥텟 수와 패킷 수가 감소하므로 이를 후보속성에 추가하였다. 또한 공격 시 특정 프로토콜(TCP/UDP)을 사용할 수 있으므로 특정 공격 형태를 구별하기 위해 이를 구별하여 후보속성에 추가하였다. 또한 공격의 중요 특징이 소스 IP 주소를 변화시켜서 플로우수를 증

가하기도 하지만, 짧은 시간 동안 많은 플로우 수 증가를 위해 소스 포트와 목적지 포트 번호를 변화시켜 사용한다. 이를 탐지하기 위해 TCP, UDP 트래픽의 소스, 목적지 포트 분산 값을 후보속성으로 추출하였고, 공격 시 대부분 소스 IP 주소 스푸핑을 사용하므로 소스 IP 주소 분산 값을 후보속성으로 추출하였다. 마지막으로 공격 시 특정 프로토콜(TCP/UDP)을 사용하므로 이로 인해 전체적인 프로토콜 비율이 공격 시에는 달라지므로 이를 마지막 두 후보속성으로 추가하였다.

실험에 사용된 두 번째 데이터인 KDD 데이터는 네트워크 기반의 서비스거부 공격과 probing 공격

표 3. KDD 데이터의 속성

	KDD Cup 1999 데이터의 속성
1	연결지속시간 (Duration)
2	프로토콜 (Protocol)
3	서비스 종류 (Service)
4	정상 또는 에러 플래그 (Flag)
5	소스로부터의 데이터 크기 (Src_byte)
6	목적지로부터의 데이터 크기 (Dst_byte)
7	소스 주소와 목적지 주소가 같으면 1, 다르면 0 (Land)
8	플래그먼트 오류 개수 (Wrong_fragment)
9	Urgent 패킷 개수 (Urgent)
10	두 호스트 간 2초 이상 연결을 지속한 접속 수 (Count)
11	두 호스트 간 한 서비스로 2초 이상 연결을 지속한 접속 수 (Srv_count)
12	"SYN" 에러율 (Serror_rate)
13	서비스 "SYN" 에러율 (Srv_serror_rate)
14	"REJ" 에러율 (Rerror_rate)
15	서비스 "REJ" 에러율 (Srv_rerror_rate)
16	접속중 같은 서비스 요청율 (Same_srv_rate)
17	접속중 다른 서비스 요청율 (Diff_srv_rate)
18	다른 호스트 접속율 (Srv_diff_host_rate)
19	목적지 호스트 개수 (Dst_host_count)
20	목적지 서비스 개수 (Dst_host_srv_count)
21	목적지 호스트상 같은 서비스 비율 (Dst_host_same_srv_rate)
22	목적지 호스트상 다른 서비스 비율 (Dst_host_diff_srv_rate)
23	목적지 호스트상 같은 소스 포트 비율 (Dst_host_same_src_port_rate)
24	목적지 호스트상 다른 호스트율 (Dst_host_srv_diff_host_rate)
25	목적지 호스트 "SYN" 에러율 (Dst_host_serror_rate)
26	목적지 호스트 서비스 "SYN" 에러율 (Dst_host_srv_serror_rate)
27	목적지 호스트 "REJ" 에러율 (Dst_host_rerror_rate)
28	목적지 호스트 서비스 "REJ" 에러율 (Dst_host_srv_rerror_rate)

트래픽을 사용하였기 때문에, 이러한 공격 탐지를 위해 휴리스틱하게 분석한 결과 Content 기반의 공격 탐지에 효과 있는 Content 속성을 배제한 나머지 기본 속성과 트래픽 속성인 28개를 후보속성으로 결정하였다. 표 3은 이러한 KDD 데이터의 기본 속성(1번-9번)과 트래픽 속성(나머지)을 보여준다.

3.3 탐지 알고리즘-지도학습 데이터 마이닝 기법

지도학습 데이터 마이닝 기법 중 의사결정트리(Decision Tree)는 어떤 집합을 여러 개의 집합으로 쪼개어 각각을 어떤 특정한 성질을 가지는 클래스로 구분하는 분류화 방법의 한 부류이다. 또한 신경망 기법(Neural Network)은 인간 두뇌의 신경 세포를 계산적으로 모델링한 것으로 특정한 작업을 수행하도록 훈련하고, 주어진 사례들을 분류하거나 예측하는데 쓰이는 지도학습 데이터 마이닝 기법 중 하나다. 이 신경망 기법은 분류 기법 중 가장 분류 성능이 우수하다고 인정받고 있다[12].

기존 선행 연구[3]로서 수행했던 분산 서비스거부 공격 탐지를 위한 결합된 데이터 마이닝 기법에서는 모델 생성 및 탐지 시험에 사용된 1)데이터로는 분산 서비스거부 공격의 특징을 잘 나타내며 많은 전처리 과정을 요구하지 않는 NetFlow 데이터를 사용하였고, 사용된 2)속성은 NetFlow 데이터를 휴리스틱하게 분석하여 추출한 표 1의 13개 기본속성을 기본으로 의사결정트리를 이용하여 중요속성을 추출하였으며(P/F, srcTport가 선정됨), 탐지 모델 생성을 위한 3)알고리즘으로는 다층퍼셉트론 신경망 기법을 이용하였고, 모델 선정 기준으로는 오분류율을 사용하여 탐지 모델을 생성하였다. 이 과정을 그림 2에 도시하였다. 제안된 방법의 탐지 성능을 시험하기 위하여 SAS Enterprise Miner[13]를 이용하여 단일 데이터 마이닝 기법을 이용한 경우와 비교 실험한 결과, 신경망 기법에 의한 오분류율은 0.0434, 의사결정트리에 의한 오분류율은 0.0869, 결합된 데

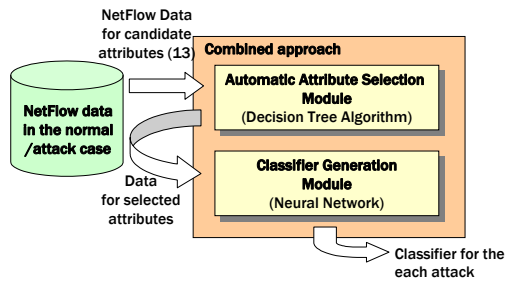

그림 2. 결합된 지도학습 기법을 활용한 탐지모델

이터 마이닝 기법에 의한 오분류율은 0.0289으로 제안된 방법의 우수한 탐지 성능을 입증하였다.

3.4 탐지 알고리즘-비지도학습 데이터마이닝 기법

지도학습 데이터 마이닝은 공격 트래픽 학습 과정과 공격 탐지 과정이 확연히 구분되어 주기적인 학습에 의해 모델을 갱신하여도 실시간으로 네트워크 상태를 반영하는 점진적 학습의 수행이 어렵고, 각 공격에 대한 트래픽과 해당 공격으로 레이블링 된 데이터가 필요하다는 제한이 있다. 따라서 본 논문의 비교 실험에 사용하는 알고리즘으로서 비지도 학습 데이터 마이닝 기법 중 자기조직화지도(SOM) 기법을 이용하여 탐지 모델을 생성하였다. 자기조직화지도 기법은 핀랜드의 Kohonen 교수 등에 의하여 개발된 비지도 신경망(unsupervised neural network)으로 위상적 순서화 성질을 갖는다. 이 기법은 패턴인식과 자료 압축/재생 등 여러 공학적 분야에서 매우 유용한 것으로 알려져 왔으며 최근 데이터 마이닝에서 다차원자료의 시각화 방법으로도 관심을 받고 있다[14].

SOM 기법을 이용한 탐지 모델 생성과정은 그림 3과 같으며, 툴로는 Matlab 상에서 somtoolbox [15, 16]를 사용하였다. ①SOM은 학습데이터가 들어오면 맵의 각 뉴런 값을 초기화 시킨 후 입력 데이터와 가장 유사한 뉴런인 BMU(Best Matching Unit)를 선택하고 그 이웃뉴런 값을 갱신한다. 모든 입력 데이터가 학습될 때까지 이 과정을 반복하면 첫 번째인 학습 단계가 끝나게 된다. 두 번째는 ②학습된 맵의 레이블링 단계가 필요하며, 이를 위해서 각 공격별로 분류된 데이터를 이용하였다. 이 단계까지 마치면 공격 탐지에 필요한 맵이 형성되므로 ③마지막으로 실시간 탐지와 점진적 학습이 이루어지는

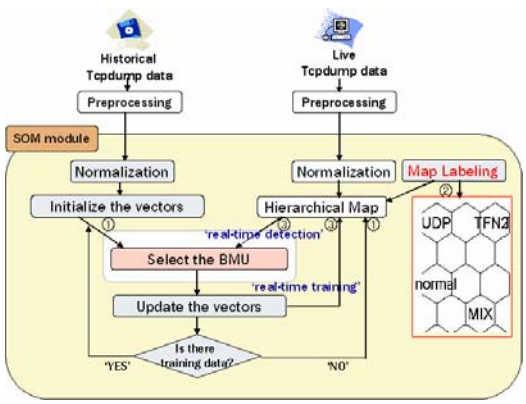

그림 3. 비지도학습 기법을 활용한 탐지모델

단계가 가능하다. 실시간으로 입력되는 데이터는 이미 훈련된 맵에서 가장 유사한 뉴런과 일치하게 되고 그 부분이 공격 클러스터이면 비정상으로, 정상 클러스터이면 정상 트래픽으로 탐지 가능하게 된다. 특히, 공격 클러스터 근처의 빈 클러스터로 매핑되면 해당 공격에 대한 새로운 변종 공격의 가능성을 알려줄 수 있다. 또한 이런 실시간 탐지와 동시에 스스로 점진적 학습이 이루어지므로 계속해서 맵이 갱신되며 결국 실시간 네트워크 트래픽의 특성을 반영하게 된다.

Ⅳ. 성능 비교

성능 비교에 사용된 탐지모델은 3장에서 설명하였듯이, 단일 지도학습 데이터 마이닝 기법(신경망 기법)을 사용한 모델, 결합된 지도학습 데이터 마이닝 기법(의사결정트리와 신경망 기법의 결합)을 적용한 모델, 비지도학습 데이터 마이닝 기법(자기조직화지도)을 적용하여 생성한 모델이다.

각 모델의 성능을 비교하기 위하여 표 1의 데이터를 사용하여 실험하였고, 공격 탐지 실험은 기존에 발견된 공격에 대한 탐지의 정확성을 알아보기 위한 오용탐지 측면과 새로운 공격에 대한 공격 탐지 가능성 및 정확성을 알아보기 위한 이상 탐지 측면에서 실험하였다. 전자는 공격 탐지 모델 생성 시, 각 공격에 대한 데이터와 정상 데이터(70%)를 포함시켜 학습시켰고, 성능 시험 시 모델링에 사용되지 않은 각 공격에 대한 트래픽과 정상 트래픽(30%)을 사용하였다. 후자는 여러 공격 타입 중 일부분만 가지고 탐지 모델을 형성 시킨 뒤 탐지 모델 형성에 포함되지 않는 공격을 새로운 공격이라 가정하여 탐지 모델에 테스트 데이터로 입력하여 정확하게 탐지하는지 시험하였다.

탐지 성능의 척도로는 시간 측면의 모델링/탐지 시간과 탐지 정확성 측면의 탐지율(Detection Rate)을 측정하였다. 모델링 시간이란 각 메커니즘이 기존의 학습데이터로 충분한 학습을 시켜 공격 모델을 만드는데 까지 걸린 시간을 의미하며, 탐지 시간이란 각 공격 모델이 완성된 후 실제로 새로운 데이터가 들어왔을 때 정확하게 탐지하는데 걸리는 시간을 의미한다. 탐지율은 테스트 데이터 중에 해당 데이터 유형으로 정확히 분류하는 비율로 1의 최대값을 갖는다.

먼저 오용탐지 측면에서 모델 생성 및 탐지에 사용되는 타겟 레이블로 <Normal과 각 공격 타입별>

을 사용한 것과 공격을 모두 Abnormal로 레이블링 하여 <Normal과Abnormal>의 레이블을 이용하여 실험하였다. 표 4는 이에 대한 실험 결과로, 모델링 시간은 비지도 학습 모델이 가장 적은 시간이 소요되었고, 각 모델간 모델링 시간 차이는 모델링에 사용된 데이터의 양이 많아질수록 더욱 커졌으며 (NetFlow_400< KDD_1000< KDD_50000), 타겟 레이블로서 <Normal과 각 공격 타입별>을 사용한 경우보다 <Normal과Abnormal>을 사용한 경우가 더 적은 시간이 소요되었다. 탐지 시간은 모두다 1초 내외의 시간이 소요되어 데이터 마이닝에 의한 탐지 모델이 실시간 탐지를 수행함을 검증할 수 있었다.

또한, 오용탐지를 위한 탐지 정확성 측면의 비교 결과로서 그림 4와 그림 5의 결과를 얻을 수 있었다. 그림 4는 모델 생성 및 탐지에 사용되는 타겟 레이블로 <Normal과 각 공격 타입별>을 사용한 것이고, 그림 5는 <Normal과 Abnormal>을 사용한 것으로, 모두 결합된 지도학습 모델이 가장 좋은 성능을 보였고, 비지도 학습이 조금 낮은 성능을 보였으나, 모델링 데이터의 양이 많아질수록(KDD_50000) 비슷한 결과를 얻을 수 있었다. 또한 전체적인 탐지율이 <Normal과 Abnormal>로 레이블링한 것이 더 높은 값을 얻었지만, <Normal과 각 공격 타입별>로 레이블링하여 탐지하는 경우에는 해당 공격에 대한 적절한 대응을 할 수 있도록 공격의 종류에 대한 추가적 정보를 제공할 수 있다.

표 4. 오용탐지를 위한 탐지 모델의 성능비교(시간 측면)

비교항목(평균값)			지도 학습	결합된 지도 학습	비지도 학습
<Normal과 공격 타입별> 레이블링	모델링 시간	NetFlow_400	15초	20초	12초
		KDD_1000	18초	24초	13초
		KDD_50000	5분 11초	7분 5초	3분 34초
	탐지 시간	NetFlow_400	1초	1초	1초
		KDD_1000	1초	1초	1초
		KDD_50000	1초	1초	1초
<Normal과 Abnormal> 레이블링	모델링 시간	NetFlow_400	12초	16초	8초
		KDD_1000	14초	17초	7초
		KDD_50000	4분 41초	5분 58초	2분 4초
	탐지 시간	NetFlow_400	1초	1초	1초
		KDD_1000	1초	1초	1초
		KDD_50000	1초	1초	1초

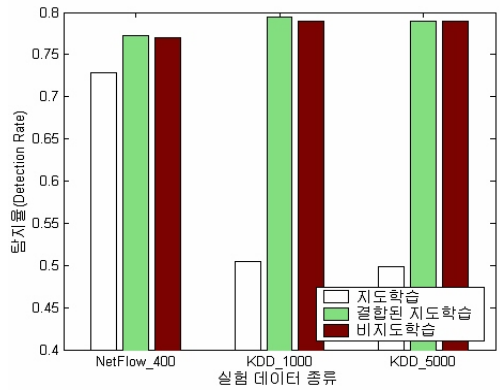

그림 4. 오용탐지를 위한 탐지 모델의 성능비교(탐지 정확성 측면) : <Normal과 각 공격 타입별> 레이블링

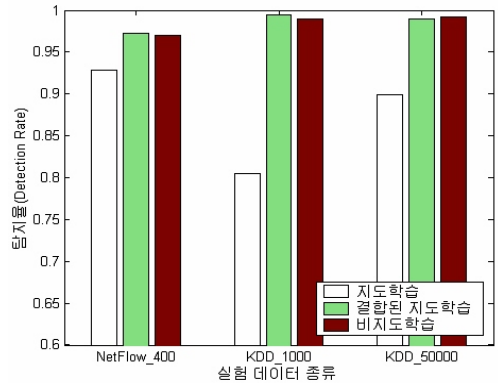

그림 5. 오용탐지를 위한 탐지 모델의 성능비교(탐지 정확성 측면) : <Normal과 Abnormal> 레이블링

오용탐지를 위한 시험 결과를 종합해 보면, 결합된 지도학습이 모델링 시간이 가장 많이 소요되지만 탐지 정확성 측면에서는 가장 성능이 좋았고, 비지도학습이 탐지 정확성은 조금 떨어지지만 가장 적은 모델링 시간이 소요되고 점진적인 학습 수행으로 시간이 지날수록, 즉 모델링에 사용되는 데이터의 양이 많아질수록 더 나은 성능을 보이고 있어 유연한 탐지가 가능함을 알 수 있었다.

두 번째 실험은 이상탐지 측면에서 새로운 공격에 대한 탐지 가능성 및 성능을 비교해 보기 위해 학습에 포함되지 않은 공격 유형을 새로운 공격이라 가정하고 탐지 성능을 시험하는데 사용하였다. 즉, NetFlow 데이터에서는 “UDP”, “TCP”, “MIX” 공격 각각을 새로운 공격이라 가정하여 나머지 두 공격 데이터만으로 모델링하고, 새로운 공격이라 가정한 트래픽을 테스트 데이터로 사용했다. 같은 방식으로 KDD 데이터에서는 DoS 공격 유형인 back, Neptune, pod, smurf, teardrop과 probing 공격 유

형인 ipsweep, nmap, portseep, satan을 비롯한 9가지 공격 각각을 새로운 공격이라 가정하여 나머지 8가지 공격 및 정상데이터로 모델링하고, 새로운 공격이라 가정한 트래픽과 kdd_corrected data를 테스트 데이터로 사용했다.

표 5는 이에 대한 실험 결과로, 모델링 시간은 역시 비지도 학습 모델이 가장 적은 시간이 소요되었고, 각 모델간 모델링 시간 차이는 모델링에 사용된 데이터의 양이 많아질수록 더욱 커졌으며(NetFlow_400<KDD_1000< KDD_50000), 탐지 시간은 모두다 1초 내외의 시간이 소요되어 이상탐지를 위한 데이터 마이닝 탐지 모델이 실시간 탐지를 수행함을 검증할 수 있었다. 탐지 정확성 측면에서는 새로운 공격으로 가정한 공격의 종류에 따라 조금 다른 결과를 보였지만, 대체로 결합된 지도학습 모델이 가장 좋은 성능을 보였고, 다음으로 비지도학습 모델이 좋은 성능을 보였다. 표 5의 결과는 모델 생성 및 탐지에 사용되는 타겟 레이블로 <Normal과 Abnormal>을 사용한 결과인데, 이 레이블을 사용한 이유는 이상탐지의 목적이 새로운 공격에 대한 탐지 자체가 더욱 중요하기 때문이다.

표 5. 이상탐지를 위한 탐지 모델의 성능비교

비교항목(평균값)				지도 학습	결합된 지도 학습	비지도 학습
<Normal과 Abnormal> 레이블링	모델링 시간	NetFlow_400		11초	15초	6초
		KDD_1000		13초	18초	8초
		KDD_50000		4분 31초	5분 3초	1분 59초
	탐지 시간	NetFlow_400		1초	1초	1초
		KDD_1000		1초	1초	1초
		KDD_50000		1초	1초	1초
	탐지율	NetFlow_400	UDP	0.852	0.428	0.980
			TCP	0.360	0.533	0.970
			MIX	1.000	1.000	0.950
		KDD_1000	Back	0.819	0.909	1.000
			Neptune	0.997	0.999	0.950
			Pod	0.890	0.924	0.900
			Smurf	0.988	0.899	0.500
			Teardrop	0.899	0.983	0.810
			Ipsweep	0.897	0.990	0.790
			Nmap	0.996	0.999	0.880
			Portsweep	0.900	0.909	0.990
			satan	0.935	0.930	0.920
		KDD_50000	Back	0.899	0.999	1.000
			Neptune	0.997	0.999	0.979
			Pod	0.900	0.994	0.999
			Smurf	0.987	0.999	0.509
			Teardrop	0.890	0.993	0.899
			Ipsweep	0.899	0.997	0.832
			Nmap	0.996	0.999	0.860
			Portsweep	0.901	0.929	0.989
			satan	0.935	0.950	0.918

그림 6은 비지도학습 모델에서 제공하는 시각화된 결과로서 학습 후, 레이블링된 맵의 결과를 보여주고 있는데, 특히 이상탐지에서 새로운 공격이 탐지된 경우 매핑된 클러스터의 주변 정보를 통해 새로운 공격에 유사한 기존 공격을 가시화해 주므로, 공격에 대응하는데 추가적인 정보를 제공해 주는 장점이 있다.

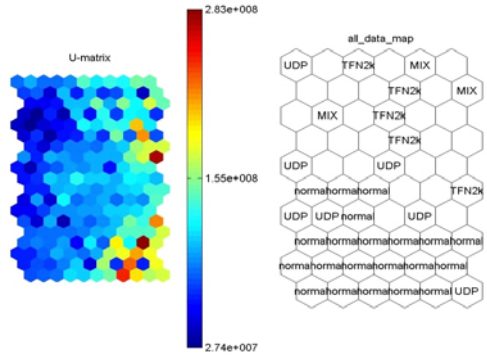

그림 6. 비지도학습 모델의 학습 후 레이블링된 맵

이상탐지를 위한 시험 결과를 종합해 보면, 마찬가지로 결합된 지도학습이 모델링 시간은 가장 많이 소요되지만 탐지 정확성 측면에서는 대체로 가장 성능이 좋았고, 비지도학습이 탐지 정확성은 조금 떨어지지만, 새로운 공격에 대한 추가적인 유형 정보를 제공해 주었다. 또한 탐지 정확성은 각 공격 유형에 대해 가장 좋은 모델의 결과인 경우 92% 이상의 결과를 얻어서 새로운 공격에도 공격 탐지 가능성을 검증할 수 있었다.

Ⅴ. 결론 및 향후 연구과제

현재 공격의 피해사례 증가와 그 심각성으로 인해 적절한 대응책이 시급하며, 이미 발견된 공격 유형과 정상 트래픽에 대한 정확한 모델링을 통한 빠른 탐지가 가장 중요한 방안이라 할 수 있다. 본 논문에서는 다양한 데이터 마이닝 기법을 활용해 다각적으로 변화하는 공격들에 유연하게 탐지가 가능한 탐지 모델의 생성 방법들을 소개하였고, 이 모델

들의 빠르고 정확한 탐지에 대한 성능 비교를 수행하였다. 탐지 모델의 성능에 중요한 영향을 미치는 요소, 즉 데이터, 속성, 탐지 알고리즘의 요소를 다양화하여 실험하였으며, 기존에 발견된 공격 탐지에 대한 오용탐지 측면과 새로운 공격에 대한 탐지 가능성 시험을 위한 이상탐지 측면에서 탐지 정확성 및 모델링/탐지 시간을 성능척도로 시험하였다.

실험 결과, 오용탐지/이상탐지 두 측면 모두 결합된 지도학습이 모델링 시간이 가장 많이 소요되었지만 탐지 정확성 측면에서는 가장 성능이 좋았고, 비지도학습이 탐지 정확성은 조금 떨어지지만 가장 적은 모델링 시간이 소요되었으며, 모델링에 사용되는 데이터의 양이 많아질수록 더 나은 성능을 보이고 있어 학습과 탐지를 동시에 수행함으로 인한 유연한 탐지 가능성을 입증할 수 있었다. 또한 모든 경우 1초 내외의 탐지 시간이 소요되어 데이터 마이닝에 의한 탐지 모델이 실시간 탐지를 수행함을 검증할 수 있었다. 특히, 이상탐지 측면에서 비지도 학습이 탐지 정확성은 조금 떨어지지만, 새로운 공격에 대한 추가적인 유형 정보를 제공해 주었고, 탐지 정확성은 각 공격 유형에 대해 가장 좋은 모델의 결과가 92% 이상의 결과를 얻어서 새로운 공격에도 탐지가 가능함을 검증할 수 있었다.

결론적으로 이러한 비교 실험 결과를 통해 다양한 공격 탐지에 대한 데이터 마이닝 기법의 특징과 장점을 이해하고, 각 상황에 맞게 적절하고 의미 있는 데이터 마이닝 기법들을 결합하여 적용할 경우 높은 탐지율을 제공하며 실시간 탐지가 가능한 보안 관리 시스템을 구축할 수 있을 것으로 예상된다. 향후 연구과제로는 데이터 마이닝 기법을 이용한 것과 기존에 제안된 다른 기법들의 비교 연구를 통해 각 성능 척도에 따른 향상된 탐지 모델을 찾는 것이 필요하다.

참 고 문 헌

〔1〕 Wenke Lee, Salvatore J. Stolfo, “Data Mining Approaches for Intrusion Detection,” *Proc. of the 7th USENIX Security Symposium*, pp.79-94, Jan. 1998.

〔2〕 나현정, 김미희, 채기준, 나중찬, “NetFlow 트래픽을 이용한 분산 서비스거부 공격 탐지 기법”, *한국정보처리학회 추계학술발표대회*, 제10권 제2호, pp.1957-1960, 2003년 11월.

〔3〕 Mihui Kim, Hyunjung Na, Kijoon Chae, Hyochan Bang, Jungchan Na, “A Combined Data Mining Approach for DDos Attack Detection,” ICOIN 2004, *LNCS 3090*, pp.943-950, Feb. 2004.

〔4〕 Jelena Mirkovic, Gregory Prier, Peter Reiher, “Attacking DDoS at the Source,” *Proc. of ICNP 2002*.

〔5〕 Joao B. D. Cabrera, Lundy Lewis, Xinzhou Qin, Wenke Lee, Ravi K. Prasanth, B. Ravichandran, Raman K. Mehra, “Proactive Detection of Distributed Denial of Service Attacks using MIB Traffic Variables,” *Proc. of ICNP 2002*.

〔6〕 Laura Feinstein, Dan Schnackenberg, Ravindra Balupari, Darrell Kindred, “Statistical Approaches to DDoS Attack Detection and Response,” *Proc. of The DARPA Information Survivability Conference and Exposition*, 2003.

〔7〕 Wenke Lee, Salvatore J. Stolfo, “Data Mining Approaches for Intrusion Detection,” *Proc. of the 7th USENIX Security Symposium*, pp.79-94, Jan. 1998.

〔8〕 Anup K. Ghosh, Aaron Schwartzbard, “A Study in using Neural Networks for Anomaly and Misuse Detection,” *Proc. of the 8th USENIX Security Symposium*, Washington, D.C., USA, Aug. 1999.

〔9〕 Susan C. Lee, David V. Heinbuch, “Training a Neural-Network Based Intrusion Detector to Recognize Novel Attacks,” *Proc. of the 2000 IEEE Workshop on Information Assurance and Security United States Military Academy*, West Point, NY, 6-7 Jun. 2000.

〔10〕 P. Phaal, S. Panchen, N. McKee, “InMon Corporation’s sFlow: A Method for Monitoring Traffic in Switched and Routed Networks,” RFC 3176, 2001.

〔11〕 KDD Cup 1999 data, http://kdd.ics.uci.edu

〔12〕 Jiawei Han, Micheline Kamber, “Data Mining: Concepts and Techniques,” *Morgan Kaufmann Publishers*.

〔13〕 최종후, 한상태, 강현철, 김은석, 심미경, 이성건, “SAS Enterprise Miner 4.0을 이용한 데이터 마이닝-기능과 사용법”, 자유아카데미.

〔14〕 허명회, "SOM(자기조직화지도)의 이론과 응용", 2004년 *한국통계학회 추계학술대회*, 2004년 11월.
〔15〕 SOM Toolbox for Matlab, http://www.cis.hut.fi
〔16〕 Juha Vesanto, John Himberg, Esa Alhoniemi, and Juha Parhankangas, "SOM Toolbox for Matlab5," SOM Toolbox Team, *Helsinki University of Technology*, 2000.

김 미 희 (Mihui Kim) 정회원

1997년 2월 이화여자대학교 전산학과 이학사
1999년 2월 이화여자대학교 컴퓨터학과 공학석사
1999년 2월~1999년 7월 (주)인티 연구원
1999년 8월~2003년 10월 한국전자통신연구원 네트워크연구소 연구원
2003년 3월~현재 이화여자대학교 컴퓨터학과 박사과정
<관심분야> 네트워크 보안, 센서 네트워크 보안, 이동 네트워크 보안, 통합망 보안

오 하 영 (Hayoung Oh) 준회원

2002년 2월 덕성여자대학교 전산학과 이학사
2001년 11월~2004년 1월 신한금융지주회사 e-신한
2004년 3월~현재 이화여자대학교 컴퓨터학과 석사과정
<관심분야> 네트워크 보안, 센서 네트워크, 홈 네트워크, 유비쿼터스 컴퓨팅, DDoS

채 기 준 (Kijoon Chae) 정회원

1982년 2월 연세대학교 수학과 이학사
1984년 5월 미국 Syracuse University 컴퓨터학과 이학석사
1990년 5월 미국 North Carolina State University 컴퓨터공학과 공학박사
1990년~1992년 미국 해군사관학교 컴퓨터학과 조교수
1992년~현재 이화여자대학교 컴퓨터학과 교수
<관심분야> 네트워크 보안, 액티브 네트워크 보안 및 관리, 인터넷/무선통신망/고속통신망 프로토콜 설계 및 성능분석