분산 서비스거부 공격 탐지를 위한 데이타 마이닝 기법

(Data Mining Approaches for DDoS Attack Detection)

김 미 희 † 나 현 정 †† 채 기 준 ††† 방 효 찬 †††† 나 중 찬 ††††
(Mihui Kim) (Hyunjung Na) (Kijoon Chae) (Hyochan Bang) (Jungchan Na)

요 약 최근 분산 서비스거부 공격에 대한 피해사례가 증가하면서 빠른 탐지와 적절한 대응 메커니즘에 대한 필요성이 대두되었다. 그러나 지금까지 제안된 기존 보안 메커니즘은 이러한 공격들에 대해 충분한 대응책을 제공하지 못하고, 일부 공격에만 유효하거나 공격의 일부 변형에도 취약점을 갖고 있다. 그러므로 본 논문에서는 최신의 분산 서비스거부 공격 유형을 잘 분류해 낼 수 있고, 기존 공격의 변형이나 새로운 공격에도 탐지 가능하도록 데이타 마이닝 기법을 이용한 탐지 구조를 제안한다. 이 탐지 구조는 이미 발견된 공격을 유형별로 분류할 수 있도록 모델링하는 오용탐지모듈과, 공격의 일반적인 특성을 이용하여 새로운 유형의 공격을 발견할 수 있도록 모델링하는 이상탐지모듈로 구성되어 있다. 이렇게 오프라인으로 생성된 탐지 모델을 통해 실시간 트래픽 데이타를 이용한 탐지 구조를 갖고 있다. 본 논문에서는 실제 네트워크의 상황을 잘 반영시켜 모델링을 하고 시험하기 위해 실제 네트워크에서 사용중인 액세스 라우터에서 NetFlow 데이타를 수집하여 이용하였다. NetFlow는 많은 전처리 과정 없이 플로우 기반의 통계 정보를 제공하므로 분산 서비스거부 공격 분석에 유용한 정보를 제공한다. 또한 공격 트래픽을 수집하기 위하여 잘 알려진 공격 툴을 이용하여 실제 공격 트래픽에 대한 해당 액세스 라우터에서의 공격 NetFlow 데이타를 수집하였다. 시험 결과, 이러한 트래픽을 이용하여 두가지 데이타 마이닝 기법을 결합한 오용탐지모듈의 높은 탐지율을 얻을 수 있었고, 새로운 공격에 대한 이상탐지모듈의 탐지 가능성을 입증할 수 있었다.

키워드 : 분산 서비스거부 공격, 분산 서비스거부 공격 탐지 기법, 데이타 마이닝, 넷플로우(NetFlow), 플로우기반, 이상탐지, 오용탐지

Abstract Recently, as the serious damage caused by DDoS attacks increases, the rapid detection and the proper response mechanisms are urgent. However, existing security mechanisms do not effectively defend against these attacks, or the defense capability of some mechanisms is only limited to specific DDoS attacks. In this paper, we propose a detection architecture against DDoS attack using data mining technology that can classify the latest types of DDoS attack, and can detect the modification of existing attacks as well as the novel attacks. This architecture consists of a Misuse Detection Module modeling to classify the existing attacks, and an Anomaly Detection Module modeling to detect the novel attacks. And it utilizes the off-line generated models in order to detect the DDoS attack using the real-time traffic. We gathered the NetFlow data generated at an access router of our network in order to model the real network traffic and test it. The NetFlow provides the useful flow-based statistical information without tremendous preprocessing. Also, we mounted the well-known DDoS attack tools to gather the attack traffic. And then, our experimental results show that our approach can provide the outstanding performance against existing attacks, and provide the possibility of detection against the novel attack.

Key words : Distributed Denial of Service(DDoS) Attack, DDoS Attack Detection Mechanism, Data Mining, NetFlow, Flow-based, Anomaly Detection, Misuse Detection

· 본 연구는 한국전자통신연구원 정보보호연구단의 위탁연구과제에 의한 것임
† 비 회 원 : 이화여자대학교 컴퓨터학과
mihui@ewhain.net
†† 비 회 원 : 삼성전자 네트워크사업부 연구원
hjski@freechal.com
††† 종신회원 : 이화여자대학교 컴퓨터학과 교수
kjchae@ewha.ac.kr
†††† 비 회 원 : 한국전자통신연구원 정보보호연구단 연구원
bangs@etri.re.kr
njc@etri.re.kr
논문접수 : 2004년 1월 29일
심사완료 : 2005년 3월 3일

1. 서 론

분산 서비스거부(DDoS) 공격은 하나의 시스템 자원 뿐 아니라 네트워크 자원까지도 고갈시킬 수 있는 간단하면서도 매우 강력한 공격이다. 기존의 서비스거부(DoS, Denial of Service) 공격에 비해 더욱 강력해진 분산 서비스거부 공격은 다수의 공격 에이전트를 분산 설치해두고 동시에 공격함으로써 공격으로 인한 파급 효과를 극대화한다. 실제로 웜 바이러스와 함께 분산 서비스거부 공격으로 인한 대량의 이상 트래픽으로 인해 인터넷 사용에 있어서 연결 실패나 속도 저하 등의 문제를 일으키는 사례가 증가하고 있으며, 이로 인한 피해는 점점 더 심각해지고 있는 실정이다. 이러한 상황을 고려해 볼 때, 분산 서비스거부 공격에 대한 적절한 방어 대책이 시급하다.

이처럼 분산 서비스거부 공격에 대한 대응책의 필요성이 대두되면서, 최근 서비스거부 혹은 분산 서비스거부 공격에 대한 다각적인 연구들이 진행되어 왔다. 그러나 현재까지 제안된 보안 메커니즘들은 다양한 분산 서비스거부 공격 중 일부에 대해서만 유효하거나, 기존 공격의 일부 변형에 대해 적절히 대응하지 못하며, 강력한 분산 서비스거부 공격 툴에서 제공하는 자동화된 또는 지능적 IP 스푸핑(Spoofing)에 효과적으로 대응하지 못하고 있다.

본 논문에서는 분산 서비스거부 공격의 심각성을 인식하고, 분산 서비스거부 공격을 탐지하기 위한 실제적이고 효율적인 구조를 제안하고자 한다. 이를 위하여 분산 서비스거부 공격 유형에 대해서 알아보고, 실제 많이 사용되고 있는 대표적인 공격 툴들을 테스트해 보면서 각각의 특징을 분석하고자 한다. 그리고 분산 서비스거부 공격 탐지와 관련된 기존 연구들에 대해 살펴보고 그 한계점에 대해서 분석하며, 분산 서비스거부 공격을 탐지하기 위해서 이미 알려진 공격 유형에 대한 특성과 정상적인 트래픽과 구별되는 공격 트래픽의 패턴을 찾고자 한다. 이는 일반적인 침입 탐지(Intrusion Detection) 기법의 두 분류인 오용탐지(Misuse Detection)와 이상탐지(Anomaly Detection)를 모두 가능하게 하기 위함이다. 전자는 잘 알려진 공격 패턴이나 이미 알려진 시스템의 취약성을 시그니쳐로 정의하여 이를 이용하는 기법이고, 후자는 정상 사용 패턴을 모델링하여 이와 다른 변칙적인 행위를 감지하는 탐지 기법이다. 본 논문에서는 이러한 두 분류의 탐지, 즉 기존에 알려진 공격뿐 아니라 새로운 공격도 탐지 가능하도록 데이타 마이닝 기법을 이용한 탐지 구조를 제안하고자 한다. 이를 위해서 실제 네트워크 트래픽 정보를 모니터링하고 수집, 분석하며, 또한 대량의 트래픽 데이타를 효과적으로 분석하고 실제 공격을 구별해낼 수 있는 탐지 모델을 만들기 위해 다양한 데이타 마이닝 기법을 적용하고자 한다.

네트워크 트래픽을 모니터링하기 위하여 SNMP(Simple Network Management Protocol) MIB(Management Information Base), RMON(Remote network MONitoring) MIB[1], 또는 tcpdump 등을 사용할 수 있고, 이러한 트래픽 정보를 이용한 분산 서비스거부 공격 대응 기법에 대한 다양한 연구들이 진행되어 왔다. 그러나 tcpdump 정보로부터 의미 있는 특징을 추출하기 위해서는 여러 번의 전처리 과정이 필요하며[2], SNMP MIB을 이용할 경우 역시 각 시스템의 MIB 정보를 통합하고 결합된 정보를 분석하는 과정이 필요하다[3]. 또한 RMON MIB에서는 분산 서비스거부 공격 탐지에 유용한 서브네트워크 단위의 통계 자료나 종단간 통계 정보를 제공하나 RMON MIB 정보를 제공하기 위해서는 시스템에 많은 처리 부담을 주기 때문에 실제 사용 장비에서는 이러한 기능을 대부분 사용하지 않는 실정이다. 이러한 점들을 고려하여 본 논문에서는 시스코사에서 개발한 NetFlow를 이용하고자 한다. 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)[4]는 IETF에서 표준화된 기술로 NetFlow와 유사한 기능들을 제공한다. 대부분의 분산 서비스거부 공격이 급격하게 플로우를 증가시킴으로써 스위치의 최대 플로우를 가득 차게 하여 네트워크의 인터넷 연결을 막기 때문에 NetFlow와 같이 플로우를 기반으로 트래픽을 분석하는 것은 의미 있는 공격 특성을 밝혀내는데 유리하다.

데이타 마이닝 기술은 대량의 데이타로부터 의미 있는 유용한 정보를 추출하기 위해서 데이타 베이스, 기계학습(Machine Learning), 정보 이론(Information Theory), 통계, 가시화(Visualization) 기법들을 통합한 기술이다. 군집화, 분류, 연관규칙 등의 데이타 마이닝 기법은 침입 탐지 분야에서 이상탐지와 관련하여 많이 연구되어 왔다[2]. 본 논문에서는 액세스 라우터에서 수집되는 NetFlow 데이타를 기반으로 다양한 데이타 마이닝 기법을 이용하여 공격 트래픽 패턴을 모델링하고 탐지 성능을 측정하여, 좋은 성능을 갖는 탐지 모델을 제안하고자 한다.

본 논문의 구성은 다음과 같다. 1장의 서론에 이어 2장에서는 관련 연구로서 분산 서비스거부 공격을 간단히 설명하고, 이에 대한 대응 방안으로 제안된 기존의 메커니즘들을 비교 설명하고자 한다. 3장에서는 본 논문에서 제안한 분산 서비스거부 공격 탐지 구조에 대해 설명하고, 4장에서는 3장에서 기술한 탐지 메커니즘을

표 1 공격툴별 특성

	Trinoo	TFN2k	Stacheldraht	Synk4
공격 유형	UDP Flood	UDP/SYN/ICMP Flood, Smurf	UDP/SYN/ICMP Flood, Smurf	SYN flood
소스 IP	스푸핑 안됨	스푸핑 정도 조절 가능	자동 스푸핑	스푸핑
소스 포트	지정 불가	자동선택 (랜덤/순차적)	자동선택 (랜덤/순차적)	자동선택 (랜덤)
타겟 포트	지정 불가	지정가능	범위지정	범위지정
기타		• 제어메세지의 일방향 통신 • 마스터와 에이전트간 통신 암호화	• 에이전트의 자동 업데이트 • 공격자/마스터/에이전트간 통신 암호화	

다양한 측면에서 실험한 내용과 그 결과를 기술하고자 한다. 마지막으로 5장에서는 결론으로써 본 논문을 마치고자 한다.

2. 관련 연구

2.1 분산 서비스거부 공격

침입은 자원의 무결성, 비밀성 그리고 가용성을 위협하기 위해서 시도되는 일련의 행위들로 정의된다[2]. 분산 서비스거부 공격은 이들 위협 중 시스템 자원 뿐만 아니라 네트워크 자원의 가용성을 침해하는 행위이며, 최근 들어 이러한 분산 서비스거부 공격으로 인한 피해 사례가 증가하고 있다. 분산 서비스거부 공격이 일반적으로 정상적인 프로토콜 패킷을 사용하기 때문에 이들 공격을 단지 사용자 인증 혹은, 암호화에 의한 정보보호 등의 침입 방지 기술로는 완전히 막기가 어렵다.

분산 서비스거부 공격이 위협적인 이유 중 하나는 자동화된 공격 툴 때문이다. 자동화된 공격 과정으로 인해 공격자가 일단 취약한 시스템을 발견하면, 공격 툴을 설치하고 실제 공격을 수행하는데 5초 이상이 걸리지 않는다. 대표적인 분산 서비스거부 공격 툴로 Trinoo, Synk4, TFN, TFN2k 그리고 Stacheldraht가 있다. 본 논문에서는 실험을 위해 가장 강력한 공격 툴인 TFN2와 Stacheldraht를 사용하였다. 표 1은 몇 가지 공격툴의 특성을 비교한 것이다[5]. 이들 공격 툴은 특정 포트 번호와 프로토콜을 사용하나 쉽게 변경 가능하다. 그래서 공격자와 마스터, 마스터와 에이전트 간의 제어 메시지를 모니터링하여 공격을 탐지한다는 것이 현실적으로 어렵다.

2.2 기존 대응 메커니즘

이러한 분산 서비스거부 공격에 대응하기 위한 관련 연구들은 크게 세 분류로 나뉘어 이루어지고 있다. 탐지 기법, 필터링 기법 그리고 역추적 기법이 그것이다. 기제안된 탐지 기법으로는 근원지에서의 공격 탐지 기법, 통계적 기법을 이용한 공격 탐지 기법, 데이타 마이닝 기법을 이용한 공격 탐지 기법 등이 있다.

근원지에서의 공격 탐지 기법으로 로스앤젤레스 소재 캘리포니아 대학교에서 개발한 'D-WARD'(DDoS netWork Attack Recognition and Defense)라는 기술은 보안 소프트웨어를 네트워크 게이트웨이에 설치하고, 게이트웨이를 통해 밖으로 나가는 트래픽에 대한 감시를 집중하는 방안이다[6]. D-WARD는 소스 네트워크에 위치하여 자율적으로 공격을 탐지하고 방어하는 분산 서비스거부 공격 방어 시스템이다. 소스 네트워크와 인터넷 사이의 양방향 트래픽 플로우를 감시함으로써 공격을 탐지하여 패킷율 제한(rate-limiting) 방법으로 트래픽양을 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 서비스해주지 못한다는 한계가 있다.

또다른 근원지에서의 공격 탐지 기법으로 MIB 정보를 이용한 공격 탐지 방법[3]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 도메인간의 MIB 정보를 교환하는 작업이 필요하나, 아직까지는 도메인간의 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[7]. 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 어려워 지고 있다. 그리고 다양한 유형의 공격들이 존재하기 때문에 단순히 소스 주소만을 모니터링 하는 것은 모든 공격 유형을 탐지해 내는데 충분하지 못하다.

데이타 마이닝 기법은 일반적인 침입탐지를 위하여 다양한 연구가 진행되어 왔다. Wenke Lee[2]는 침입탐지를 위하여 sendmail과 tcpdump 데이타를 이용해 frequent episode 분류 기법과 연관규칙 기법을 통해 침입 모델을 생성하여 이를 시험하였고, 데이타 마이닝 적용시 많은 모델링 시간이 걸리는 점을 고려하여 학습(learning) 에이전트와 탐지 에이전트로 구성된 침입 탐지 구조를 제안하였다. 이 논문은 tcpdump 네트워크 트래픽을 이용해 서비스거부 공격을 탐지할 수 있는 기본적인 연구 과정을 소개하였나 최근 더욱 강력해진 분산 서비스거부 공격을 탐지하기 위한 다각적인 실험이 필요하며, tcpdump 데이타를 이용하면 많은 전처리 과정이 필요하므로 빠른 탐지가 어렵다는 단점이 있다. 이외에도 침입탐지의 두 분류인 이상탐지와 오용탐지를 위해 신경망을 사용한 연구[8]와 새로운 공격을 인식할 수 있도록 신경망 기반의 침입탐지 시스템[9]이 제안되었다.

또한, 이러한 분산 서비스거부 공격이 탐지되면, 공격 근원지를 찾아 그에 대응하기 위하여 다양한 역추적 기법이 제안되었다[10]. 오늘날 가장 기본적인 역추적 기법으로 사용되는 홉-바이-홉 IP 역추적 기법은 많은 연속적인 공격 패킷이 전송될 때 홉-바이-홉으로 공격 패킷의 근원지 라우터를 추적하는 방법이다. 그러나 이 방법은 모든 네트워크 도메인의 협동이 제공되지 않으면 수행 불가능하며, 상당한 노력과 기술적인 처리가 필요한 기법이다. 이에 비해 백스캐터(Backscatter) 역추적 기법은 공격이 ISP에 보고되면, 모든 라우터에 해당 공격 시스템으로 향하는 모든 패킷을 제거하도록 설정하고, 이를 통해 전송되는 목적지 도달불능(Destination Unreachable) ICMP 에러 메시지의 목적지를 보고 공격 근원지 라우터를 찾는 방법이다. 이 방법은 동시다발적으로 공격 패킷을 생성하는 분산 서비스거부 공격에 빠르고 효율적인 역추적 기법을 제공하나 이 역시 많은 연속적인 공격 패킷이 전송될 때 이용가능하며, 랜덤으로 선택된 스푸핑 IP 소스 주소가 유효한 주소를 사용하는 지능적인 공격인 경우에는 취약점을 가지고 있다. 또한 확률에 기반하여 패킷에 마킹하고 이를 수집하여 목적지에 대한 연관관계를 도출하는 역추적 기법, ICMP 역추적 메시지 기반 역추적 기법, 해쉬 기반 IP 역추적 기술 등이 제안되었다[11].

마지막으로 대부분의 분산 서비스거부 공격이 IP 패킷의 소스 주소를 스푸핑하기 때문에 이러한 스푸핑된 패킷을 라우터에서 필터링하기 위한 다양한 방법이 제안되었다. 네트워크 입출력단에서 유효하지 않은 소스 주소를 갖는 패킷을 필터링하는 입력단/출력단 필터링(Ingress/Egress Filtering)[12,13], 라우터에서 라우팅 테이블을 기반으로 잘못된 인터페이스로 수신되는 패킷을 필터링하는 패킷 기반의 필터링 기법[14], IP 헤더의 TTL(Time-To-Live) 필드를 이용해 홉카운트가 맞지 않는 패킷을 필터링하는 홉카운트 기반 필터링 기법[15], 그리고 기존에 수신되는 패킷의 소스 주소 리스트를 저장하였다가 공격 탐지 시 저장 리스트에 없는 소스 주소를 갖는 패킷을 필터링하는 히스토리 기반의 IP 필터링 기법[16]이 제안되었다. 이 모든 기법들은 어느 정도 스푸핑된 패킷을 필터링할 수 있으나 자신의 네트워크 주소를 갖고 스푸핑하거나, 같은 홉카운트 필드를 갖도록 지능적으로 스푸핑하는 공격에서 모두 취약점을 갖는 단점이 있다.

3. 분산 서비스거부 공격 탐지 구조 제안

본 장에서는 제안하는 탐지 구조를 구성하고 있는 중요 변수들을 소개하고, 이를 기반으로 제안하는 공격 탐지 구조를 설명하며, 2장에서 기술한 기존 대응 메커니즘 중에 본 논문이 속한 기법인 탐지 기법에 대한 기존 연구와 비교하고자 한다. 제안하는 공격 탐지 구조를 설계하기 위한 목표로 다음 3가지를 두었다.

- 최신의 분산 서비스거부 공격 유형에 강력한 탐지 구조
- 기존 공격의 변형이나 새로운 공격에도 탐지할 수 있는 구조

• 실제적이고 효율적인 탐지 구조

3.1 공격 탐지를 위한 중요 변수

분산 서비스거부 공격에 대한 탐지 시스템에서 성능을 좌우하는 중요 변수로서 다음의 3가지를 꼽을 수 있다; **1)데이타, 2)속성(Attribute), 3)탐지 알고리즘.** 첫째, 탐지를 위해 사용되는 데이타의 종류, 즉 실제 네트워크의 정상과 비정상 패턴을 잘 반영해야 하고 많은 전처리 과정을 요구하지 않는 데이타이어야 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 데이타와 정상 데이타가 섞여 분포되어 있는데 이러한 경우에도 높은 공격 탐지율을 가질 수 있도록 다양한 실제 데이타를 사용하여 탐지 모델을 구축해야 할 것이다. 이를 위해서 본 논문에서는 tcpdump의 정보를 사용할 때처럼 많은 전처리 과정이 필요 없는 NetFlow 데이타를 사용했다. 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)[4]는 IETF에서 표준화된 기술로 NetFlow와 유사한 기능들을 제공한다. 분산 서비스거부 공격의 가장 큰 특징중의 하나가 갑자기 많은 플로우를 발생시켜 네트워크 시스템이나 서버 시스템 자체의 리소스를 고갈시키는 것이다. 이는 그림 1에서처럼 공격 툴을 이용하여 얻은 공격 트래픽과 정상 트래픽의 5분 동안 발생된 플로우 수의 비교에서 잘 알 수 있다. 이 결과의 실험 환경에 대해서는 4.1절에서 자세히 설명할 것이다. 그래서, 플로우 별 네트워크 데이타 정보를 출력해 주는 NetFlow 데이타를 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 제안한 실제적인 탐지 결과의 신빙성을 제공하여 준다.

둘째, 사용되는 데이타의 속성 종류에 따라서도 탐지 성능이 크게 달라질 수 있다. 특히 분산 서비스거부 공격의 가장 큰 특징이 소스 IP 주소를 스푸핑하는 것이므로 통계적 방법을 이용한 탐지 메커니즘[7]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. 이후 설명될 4.2절에서 결정 트리(Decision Tree) 알고리즘이 소스 IP 주소 분산 값을 중요 속성으로 추출하지 않는데, 그 이유는 라우터에서 필터링 기능을 수행하기 때문에 인터넷 방향으로 포워딩되는 스푸핑 패킷은 모두 소스 네트워크와 같은 네트워크 주소를 가지므로 정상과 별로 차이가 없는 분산 값을 갖기 때문이다. 그래서 우선 본 논문에서는 실제 네트워크에서 얻은 정상 데이타와 강력한 분산 서비스거부 공격 툴을 구동하여 얻은 공격 데이타를 1차적으로 휴리스틱하게 분석하여 다음 13개의 후보 속성을 추출하였다.

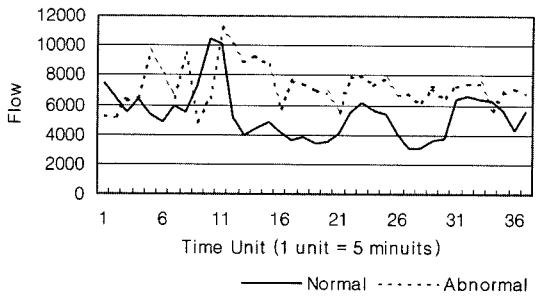


그림 1 정상/비정상 트래픽의 플로우 수 비교

- 옥텟수/플로우수(O/F), 패킷수/플로우수(P/F)
- TCP 옥텟수/플로우수(TO/F), TCP 패킷수/플로우수(TP/F)
- UDP 옥텟수/플로우수(UO/F), UDP 패킷수/플로우수(UP/F)
- TCP 트래픽의 소스 포트 분산(srcTport), 목적지 포트 분산(dstTport)
- 트래픽의 소스 포트 분산(srcUport), 목적지 포트 분산(dstUport)
- 소스 IP 주소 분산(srcVar)
- TCP 트래픽 비율(Tratio), UDP 트래픽 비율(Uratio)

우선 그림 1에서처럼 정상 트래픽보다는 공격 시 플로우 수가 증가하지만, 정상인 경우에도 플로우 수의 절대값은 크게 변화하므로 이는 후보 속성에서 제외시켰고, 대신 공격 시 한 플로우를 구성하는 옥텟 수와 패킷 수가 그림 2에서처럼 감소하므로 이를 후보속성에 추가하였다. 또한 공격 시 특정 프로토콜(TCP/UDP)을 사용할 수 있으므로 특정 공격 형태를 구별하기 위해 이를 구별하여 후보속성에 추가하였다. 여기에서 ICMP Flood 공격이 가능하지만 특정 프로토콜로 ICMP를 제외한 이유는 대부분의 대형 라우터에서는 처리 부하를 줄이기 위해 ICMP 메시지를 필터링하므로 이는 후보속

성을 위한 프로토콜로 추가하지 않았다. 또한 공격의 중요 특징이 소스 IP 주소를 변화시켜서 플로우 수를 증가하기도 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. 그러나 본 논문에서는 이미 발견된 공격에 대한 탐지 모듈을 생성하는데 있어서 이렇게 선발된 후보속성을 그대로 사용하는 것이 아니라 이미 발견된 공격을 가장 잘 분류할 수 있도록 휴리스틱하게 추출한 후보속성 중에 중요한 속성을 이론적으로 추출하기 위하여 데이타 마이닝 분류 기법 중 결정 트리 알고리즘을 사용하였다. 이에 대한 자세한 설명은 뒤에 기술할 것이다.

셋째, 탐지에 사용되는 알고리즘에 따라서도 탐지 성능이 달라질 것이다. 이를 위해 본 논문에서는 대량의 데이타로부터 의미 있는 유용한 정보를 추출하는데 유용한 데이타 마이닝 기술을 이용하였다. 데이타 마이닝 기술은 데이타 베이스, 기계 학습(Machine Learning), 정보 이론(Information Theory), 통계, 가시화(Visualization) 기법들을 통합한 기술로서. 군집화, 분류, 연관규칙 생성 등의 여러 알고리즘들이 제공되고 있다. 본 논문에서는 정상과 비정상, 그리고 비정상에 속한 공격의 타입들을

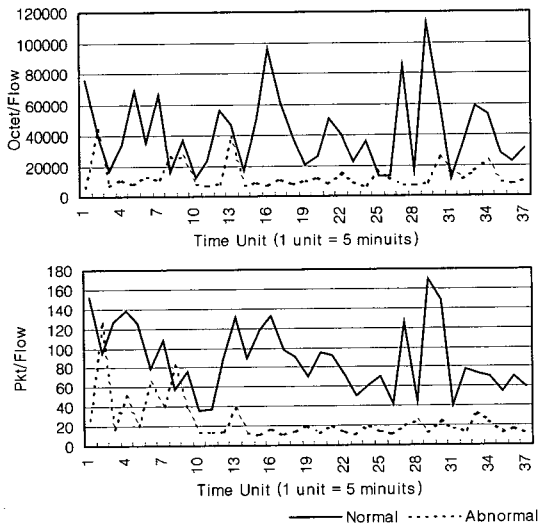


그림 2 정상/비정상 트래픽의 한 플로우당 옥텟 수와 패킷 수 비교

분류하기 위한 분류 알고리즘으로는 가장 분류 성능이 우수하다고 인정 받고 있는 다층퍼셉트론 즉 신경망 기법을 사용하였고[17], 모델 선정 기준으로는 오분류율을 사용하여 탐지 모델을 생성(training) 하였다. 또한 이미 발견된 공격 모델에 최적화된 오용탐지모듈을 만들기 위하여 다층퍼셉트론의 입력값으로 들어가는 중요 속성을 이론적으로 추출하였는데, 이를 위해 또 다른 분류 데이타 마이닝 기법인 카이제곱과 엔트로피 계산법에 의한 결정 트리 알고리즘을 사용하였다.

3.2 제안하는 공격 탐지 구조

본 논문에서 제안하는 분산 서비스거부 공격 탐지 구조는 이미 발견된 공격뿐 아니라 변형된 공격이나 새로운 공격도 탐지할 수 있기 위해 크게 오용탐지모듈(Misuse Detection Module)과 이상탐지모듈(Anomaly Detection Module)로 구성되어 있다. 오용탐지모듈은 이미 발견된 분산 서비스거부 공격을 공격별로 탐지할 수 있도록 분류 모델을 만드는 모듈이고, 이상탐지모듈은 기존 서비스거부 공격의 변형 및 새로운 유형의 분산 서비스거부 공격이 발생한 경우에도 탐지해 낼 수 있도록 데이타 마이닝을 이용하여 정상 모델을 만드는 모듈이다. 그림 3은 이러한 두 모듈로 이루어진 분산 서비스거부 공격 탐지 구조를 도시한 것이다. 이 구조가 탐지하는 대상은 감시중인 라우터에 분산 서비스거부 공격 패킷이 전송된 경우 해당 공격 발생에 대한 탐지이며, 빠른 대응을 위해 공격 유형에 대한 추가적 정보를 제공해 준다. 또한 이 탐지구조는 그림 3에서처럼 오프라인에서 주기적으로 새로운 탐지모델을 생성하는 부분과 생성된 두개의 탐지모델을 이용하여 실시간으로 라우터에 전달되는 트래픽의 플로우 정보를 감시하여 표 2에 근거하여 최종 탐지 결과를 출력하는 온라인 탐지 부분으로 구성되어 있다.

오용탐지모듈은 현재 발견된 공격들에 대한 트래픽과 정상 트래픽에 대한 NetFlow 데이타를 이용해 우선 발견된 공격과 정상 트래픽을 가장 잘 구별해 낼 수 있는 속성(Attribute)을 추출하기 위하여 자동속성추출 서브모듈(Automatic Attribute Selection Sub-Module)의 결정 트리(Decision Tree) 알고리즘을 사용하여 중요 속성을 출력한다. 이때 결정 트리 알고리즘의 입력 속성으로 사용되는 후보 속성으로는 최신 서비스거부 공격 툴을 분석하여 추출한 다음 13개의 속성을 사용하였다. 3.1절에서 설명한 1)옥텟수/플로우수, 2)패킷수/플로우수, 3)TCP 옥텟수/플로우수, 4)TCP 패킷수/플로우수, 5)UDP 옥텟수/플로우수, 6)UDP 패킷수/플로우수, 7)TCP 트래픽의 소스 포트 분산, 8)TCP 트래픽의 목적지 포트 분산, 9)UDP 트래픽의 소스 포트 분산, 10)UDP 트래픽의 목적지 포트 분포, 11)소스 IP 주소

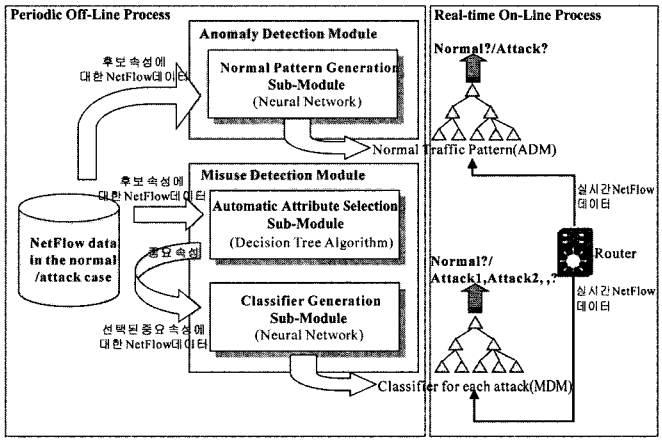


그림 3 제안한 분산 서비스거부 공격 탐지 구조

분산, 12)TCP 트래픽 비율, 13)UDP 트래픽 비율이 그것이다.

분산 서비스거부 공격 시, 하나의 플로우에 해당되는 옥텟이나 패킷수가 작아지고, 포트번호와 소스 IP 주소를 변화시켜 플로우를 증가시키며, 공격에 따라 특정 전송 프로토콜(TCP,UDP 등)을 사용하므로 이에 근거하여 후보 속성을 선정하였다. 그러나 이런 후보 속성은 제안된 탐지 구조가 사용되는 시점에 발견된 최신의 공격을 잘 분별해 줄 수 있도록 추가 가능하다. 이러한 후보 속성들 중에 가장 잘 정상/공격 및 각 공격 유형을 분별해 줄 수 있는 중요 속성을 선정하는데 있어서 카이제곱과 엔트로피 기법을 이용한 결정 트리 알고리즘을 사용하게 된다. 이렇게 이론적으로 중요 속성을 선정하는 이유는 어떤 속성을 사용하느냐에 따라 분류 모델의 정확성이 좌우되기 때문이다. 마지막으로 자동으로 선택된 중요 속성에 대한 NetFlow 데이타를 이용해 분류기생성 서브모듈(Classifier Generation Sub-Module)의 신경망 기법으로 최종 분류 모델을 생성하게 된다.

이상탐지모듈은 일반적으로 분산 서비스거부 공격에 중요하다고 분석되는 후보 속성에 대한 NetFlow 데이타들을 이용, 신경망 기반의 정상 패턴을 모델링하여 결과를 출력한다. 이 모듈에서 오용탐지모듈처럼 결정 트리 알고리즘을 이용하여 중요 속성을 추출하여 사용하지 않는 이유는 공격의 일반적인 다양한 속성을 모두 포함하여 일반화된 분산 서비스거부 공격 및 정상 상태를 잘 구별할 수 있는 모델을 만들기 위함이다. 이 두 모듈 모두 최종 모델을 생성하는데 있어 신경망 기법을 이용하였는데 그 이유는 다양한 분류 데이타 마이닝 기법 중에서 일반적으로 분류 성능이 가장 뛰어나기 때문이다[17]. 사용된 신경망의 구조는 일반적으로 많이 사용되는 다층퍼셉트론을 사용하였고, 모델 선정 기준으로는 오분류율이 사용되었다. 그러나 신경망 기법이 모델 생성을 위한 학습에 많은 시간이 걸리는 단점이 있으나 이러한 작업은 오프라인으로 수행되고, 또한 최신의 공격 및 데이타에 대한 적응을 위해 주기적으로 모델 갱신을 수행하는 구조를 갖고 있다.

이렇게 생성된 두 모델을 가지고 라우터에서 제공하는 실시간 NetFlow 데이타를 이용해 실시간 탐지 기능을 수행한다. 기본적으로 오용탐지모듈에서 생성한 모델(MDM)을 이용하여 기존 공격의 유형별 탐지를 수행하는데, 공격 탐지시 제공되는 공격 유형 정보는 빠른 대응 처리를 가능하도록 추가적 정보 제공을 목적으로 한다. 이러한 추가적인 정보를 제공하는 이유는 완전히 공격 에이전트를 분리, 제거하거나 공격 트래픽의 필터링을 수행하기 전까지 기본적인 대응을 할 수 있도록 추가적인 정보를 제공하기 위함이다. 공격 유형에는 예를 들어 TCP 공격, UDP 공격, MIX 공격 등이 될 수 있다. 그리고 이상탐지모듈에 의해 생성된 모델(ADM)을 이용해 기존에 발견되지 못한 변종의 공격 또는 새로운 공격을 탐지하도록 한다.

표 2는 이러한 두 모델의 출력 결과에 따른 최종 탐지 판정 결과이다. 첫째, MDM과 ADM이 모두 정상임을 출력하면 이는 최종 "정상"으로 판정하고, 둘째 MDM은 정상 ADM은 공격으로 출력하면 이는 새로운 공격의 가능성을 말해주므로 최종 "새로운 공격"으로 판정

표 2 오용탐지모듈 모델(MDM)과 이상탐지모듈의 모델(ADM) 결과에 따른 최종탐지결과

ADM결과 / MDM결과	정상(Normal)	공격(Attack)
정상(Noraml)	정상	새로운 공격
공격유형 (Attack1,…, Attackn)	특정유형의 공격 가능성	특정유형의 공격

하며, 셋째 드문 경우겠지만 ADM은 정상, MDM은 특정 공격유형이 출력되면 "특정유형의 공격 가능성"으로 판정하여 운영자의 트래픽 분석 및 감시가 요구됨을 경고해 준다. 마지막으로 MDM과 ADM이 모두 공격으로 출력하면 "특정 공격"으로 판정하여 해당 공격에 대응할 수 있도록 한다.

3.3 기존 탐지 메커니즘과 비교

본 논문에서 제안된 구조와 2장에서 소개된 기존 탐지 기법과 비교하여 표3에 정리하였다.

4. 실험 및 결과

4.1 공격 테스트 실험 환경

실제 네트워크의 트래픽을 관리하기 위하여 설치된 NetFlow를 이용하여 네트워크 트래픽 정보를 수집하였다. 이때 수집되는 NetFlow 데이타는 Cisco7507 라우터에서 송수신 플로우에 대한 데이타로서 이 라우터는 외부 200Mbps, 내부 1Gbps 인터페이스를 갖고 있고, 5000대 정도의 호스트/서버를 인터넷에 연결하는 액세스 라우터의 역할을 한다.

데이타 마이닝을 이용한 패턴 모델을 구축하기 위한 트레이닝 데이타를 얻기 위해서는 정상 트래픽과 공격 시의 이상 트래픽이 필요하다. 공격 시의 이상 트래픽을 얻기 위해 실제 공격에 쉽게 이용될 수 있는 강력한 공격 툴을 이용해서 다양한 시나리오로 직접 공격을 수행하였고, 그 결과로 얻어진 트래픽을 이상 트래픽으로 분류하였다. 그림 4는 공격 테스트를 수행한 네트워크 환경 및 시나리오를 나타낸 것이다.

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 근원지 주소를 위조하는 스푸핑 정도는 공격 툴별로 제공해주는 옵션에 따라 다양하게 변화시켜 주었다. 완전 랜덤하게 변하는 스푸핑 옵션부터 자신의 서브 네트워크 주소까지는 고정 상태로 하고, 그 나머지가 랜덤하게 변하는 스푸핑 옵션까지 다양한 형태로 공격해 보았다. 공격은 NetFlow 데이타를 제공하는 라우터를 기준으로 학교망 외부와 내부에 각각 에이전트들을 설치해 두고, 내부에서 외부로, 외부에서 내부로의 양방향 공격을 모두 실험해 보았다. 현실적으로 심각한 피해를 초래하는 공격은 직접해 볼 수 없기 때문에 단시간 동안만 그리고 3~5개의 정도의 공격 에이전트를 이용

표 3 기존 탐지 기법과의 비교

기존 방법	기존 모델의 단점	제안된 모델의 상대적 장점
D-WARD[6]	- 비연결형 공격 트래픽(UDP), 비대칭 통신 경로 사용하는 공격에 탐지 한계	- MDM에서 공격 연결 형태(연결형/비연결형)에 무관하게 다양한 공격에 대한 시그니처를 생성하므로 탐지시, 공격의 연결 형태 및 공격 경로와 상관없음
MIB을 이용한 기법[3]	- 탐지시 사용하는 "공격 명령 트래픽"은 쉽게 수정 가능하고, 공격 명령어 자체는 실제 트래픽에서 큰 변화 주지 못함 - 도메인간 협동이 필수적	- 탐지시 "데이타 트래픽" 변화 자체를 사용하므로 정상 트래픽과 큰 차이를 나타내어 실제적으로 높은 탐지율을 나타냄 - 소스 네트워크에서의 공격 탐지이므로 도메인간 협동에 무관
통계를 이용한 기법[7]	- 지능화된 스푸핑 IP 공격시 탐지가 제한적 - 탐지시 소스 주소만을 이용하는 경우 다양한 공격 탐지에 제한적 - 통계시 사용하는 속성 수에 따라 시스템 로드가 급격히 증가	- 탐지 및 트레이닝시 소스 IP 주소뿐 아니라 다양한 속성을 사용하므로 다양한 공격 탐지가 가능 - 감시 속성 추가가 오프라인 트레이닝에는 시간적 증가에 영향을 주지만, 실제 탐지시에는 큰 영향을 주지 않음
데이타 마이닝을 이용한 방법[2]	- 데이타 마이닝을 위해 많은 전처리 과정이 수반되는 tcpdump 데이타 사용	- 많은 전처리 과정이 필요없는 NetFlow 데이타를 사용

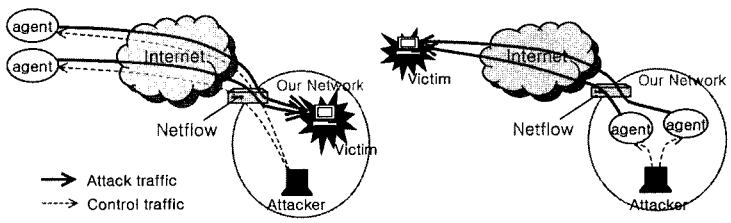
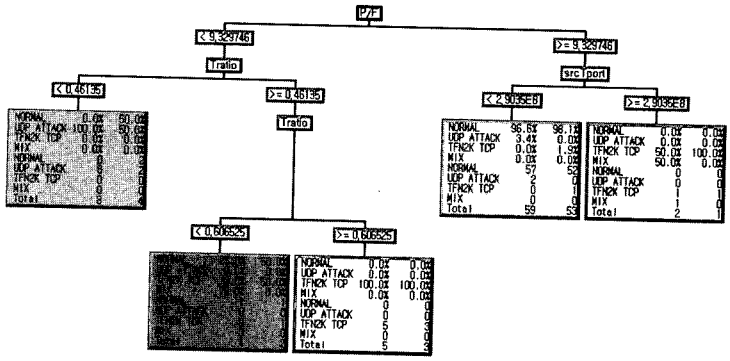


그림 4 공격 시나리오



Leaves	Training	Validation
1	0.7436	0.8730
2	0.8333	0.8571
3	0.8974	0.9206
4	0.9359	0.9048
5	0.9487	0.9206
6	0.9487	0.9206

Name	Importance	Role	Rules	Variable Label
P_F	1.0000	input	1	P/F
TRATIO	0.7259	input	2	Tratio
SRCTPORT	0.3973	input	1	srcTport
O_F	0.0000	rejected	0	O/F
SRCVAR	0.0000	rejected	0	SrcVar
SRCUPORT	0.0000	rejected	0	srcUport
DSTUPORT	0.0000	rejected	0	dstUport
URATIO	0.0000	rejected	0	Uratio
UO_F	0.0000	rejected	0	Uo/f
TO_F	0.0000	rejected	0	To/f
TP_F	0.0000	rejected	0	Tp/f
UP_F	0.0000	rejected	0	Up/f
DSTTPORT	0.0000	rejected	0	dstTport

그림 5 엔트로피에 의한 의사 결정 트리

하여 공격을 수행하였다. 따라서 실제 공격 시에는 실험 결과보다 더 큰 폭의 증감이 나타날 것으로 예상된다.

위와 같은 실험 환경 하에 공격 테스트를 반복 수행하여 트래픽 분석과 모델 생성에 필요한 데이타를 수집하였다. 수집한 데이타는 NetFlow 데이타베이스에 저장된 여러 종류의 테이블로부터 추출한 것으로, 5분 단위의 플로우 통계정보를 이용하여 트래픽 정보를 분석하였다. 추출된 각 데이타는 적절한 비율로 나누어 모델학습(training), 모델평가(evaluation), 시험(test)을 위해 사용하였다. 즉, 수집된 데이타의 일부는 공격 실험 동안 발생된 이상 트래픽(UDP Attack, TCP Attack, MIX Attack)과 평소의 정상 트래픽(Normal)으로 각각 레이블 하여 분석 및 탐지 모델 생성에 이용하였고, 또 다른 일부는 생성된 모델을 평가하기 위한 데이타로, 그 나머지는 생성된 모델의 분류 정확성 시험을 위해 사용하였다.

4.2 오용탐지모듈 실험

본 실험에서는 알려진 공격에 대해서 여러 데이타 마이닝 기법으로 생성된 탐지 모델들의 성능 비교를 통해서, 본 논문에서 제안한 결합된 데이타 마이닝 기법을 이용한 오용탐지 모델의 성능을 검증하고자 한다. 데이타 마이닝 기법 중 대표적인 분류 기법인 의사 결정 트리와 신경망 기법을 사용하여 모델들을 생성하고 이들의 탐지 성능을 오분류율을 기준으로 비교하였다. 각각의 실험에서 쓰인 입력 데이타의 상태 정보는 "Normal"과 세 가지 공격 유형인 "UDP Attack", "TCP Attack", "MIX Attack"으로, 이러한 공격들을 분류해낼 수 있는 탐지 모델을 생성하고자 하였다. 각 비교 모델의 입력 데이타로 사용된 기본 속성은 3.1절에서 설명한 것처럼 실제 공격 툴을 수행하여 수집된 공격 트래픽을 분석하여 얻어진 13개의 후보 속성을 사용하였다.

기본 속성들을 가진 입력 데이타를 이용하여 다음 네 가지 경우의 탐지 모델을 생성하고 성능을 비교해 보았다. 대조군으로써 첫번째 비교 모델은 기본 속성들로 정의된 데이타를 입력 데이타로 하여 신경망 기법을 이용한 탐지 모델이다. 두 번째, 세 번째는 각각 엔트로피와 카이제곱에 의한 의사 결정 트리 기법을 이용한 모델이다. 그림 5와 그림 6은 그림 3의 자동속성추출 서브모듈의 수행결과로서 각각 엔트로피와 카이제곱에 의해 생성된 트리와 기본 속성들의 중요도를 나타낸다. 그림 5에서 생성된 트리는 입력된 13개의 기본 속성 중 가장 각 클래스를 잘 구분해 주는 중요 속성으로 순서화 한 후, 순서화된 중요 속성부터 트리를 생성하게 된다. 이렇게 생성된 기본 트리를 평가(validation) 데이타로 평가한 후, 학습(training) 데이타와 평가 데이타의 분류 정확도가 가장 높으면서 단말 노드의 수가 적은 레벨을 선정하게 된다. 그림 5에서는 이러한 기준에 의해 왼쪽 표와 같이 5개의 단말 노드가 선택이 되어 트리가 구성되었고, 중요 속성으로 P/F, Tratio, srcTport가 선택되었음을 알 수 있다. 각 단말 노드는 분류 결과인 각 클래스가 되는데, 그림 5에서는 단말 노드의 왼쪽부터 UDP attack, MIX attack, TFN2k TCP attack,

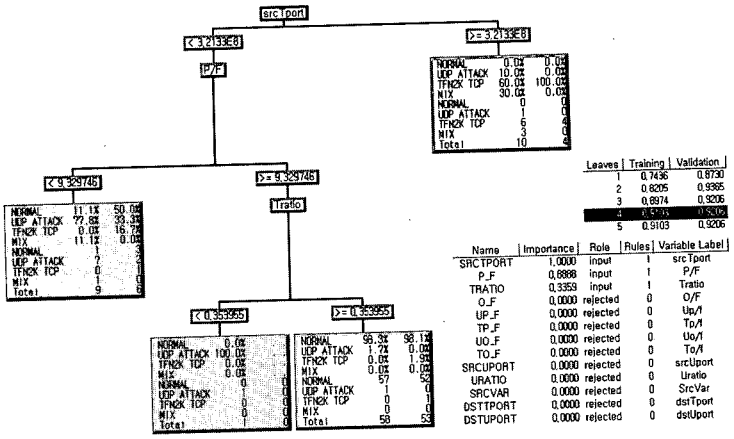


그림 6 카이 제곱에 의한 의사 결정 트리

Normal, TFN2k TCP attack의 클래스를 나타낸다. 각 단말 노드의 내용은 두 번째 컬럼은 학습 데이타에 대한 분류 결과의 비율과 개수를, 세 번째 컬럼은 평가 데이타의 분류 결과의 비율과 개수를 나타낸다. 그림 5의 왼쪽에서 첫 번째 단말 노드는 UDP attack 클래스 인데, 학습 데이타 중 UDP attack 클래스라고 레이블 된 8개의 데이타가 이 단말 노드로 분류되었고, 평가 데이타 중 Normal 클래스라고 레이블 된 2개의 데이타와 UDP attack 클래스라고 레이블 된 2개의 데이타가 이 단말 노드로 분류된 것을 나타낸다. 그림 6에서는 이전 결과와 조금 다른 트리를 출력하였지만, 중요 속성으로 이전 결과와 같은 P/F, Tratio, srcTport가 선택되었음을 알 수 있다. 즉, 분산 서비스거부 공격과 정상 트래픽을 가장 잘 분류하는 중요 속성은 한 플로우 당 패킷 수이고, 공격 별 분류를 위해 TCP 트래픽 비율과 TCP 트래픽의 소스 포트 분산 값이 중요 속성으로 추출되었다. 앞서 설명했던 것처럼 소스 IP 주소 분산 값이 추출되지 않은 이유는 모델 생성에 사용된 데이타가 수집된 라우터에서 필터링 기능을 수행하기 때문에 인터넷 방향으로 포워딩되는 스푸핑 패킷은 모두 소스 네트워크와 같은 네트워크 주소를 가지므로 정상과 별로 차이가 없는 분산 값을 갖기 때문이다.

네 번째는 의사 결정 트리를 통해 자동 추출된 속성을 이용한 신경망 모델이다. 입력 데이타를 위의 선택된 세 가지 속성-P/F, srcTport, Tratio-으로만 정의하고 이를 이용하여 신경망 기법을 적용하였다. 그 결과, 공격 종류별 상태 정보에 따라 분류되는 신경망 모델이 생성된다.

앞의 네 가지 모델 생성 방법별로 테스트 데이타를 이용하여 탐지 성능을 비교하였다. 표 4는 비교 결과로서 탐지 성능의 척도인 오분류율을 나타낸 것이다. 표에

표 4 탐지 모델별 성능 비교

탐지모델 생성 방법	오분류율	입력 데이타 속성
(1) 신경망	0.0434782609	모든 기본 속성
(2) 엔트로피 의사결정트리	0.0724637681	모든 기본 속성
(3) 카이제곱 의사결정트리	0.0869565217	모든 기본 속성
(4) 제안한 결합된 데이타 마이닝 기법 (의사결정트리 + 신경망)	0.0289855072	P/F, srcTport, Tratio

서처럼 의사 결정 트리를 통해서 선택된 속성을 이용하여 신경망 기법으로 구축한 모델이 가장 좋은 탐지 정확성을 보임을 알 수 있다.

결과적으로 신경망이나 의사 결정 트리 같은 하나의 데이타 마이닝 기법만을 사용하여 생성한 모델보다 본 논문에서 제안한 의사 결정 트리와 신경망 기법을 결합하여 생성된 모델이 오용탐지 모델로서 더 좋은 탐지 성능을 내는 것을 알 수 있다.

4.3 이상탐지모듈 실험

본 실험에서는 데이타 마이닝을 이용하여 정상 모델을 구축하고 새로운 공격에 대한 탐지 가능성 및 그 성능을 측정하였다. 모델 구축을 위해 의사 결정 트리와 신경망 기법을 각각 이용하여 비교 실험을 하였으며, UDP, TCP, MIX 공격 각각을 새로운 공격이라고 가정하여 나머지 두 공격 데이타 만으로 모델링 하였고, 새로운 공격이라 가정한 트래픽을 테스트 데이타로 사용하였다. 이때 생성된 탐지모델은 그림 3에서와 같이 "정상"과 "공격" 중 하나의 결과를 출력해 준다. 표 5는 새로운 공격별로 이상탐지모듈을 시험한 결과, 측정된 오분류율을 나타낸다. 여기에서 오분류율은 정상 트래픽으로 인식되어 들어오는 새로운 공격이 정상이 아니라고 판단되는 비율이므로, 오분류율이 높을수록 새로운 공격에 대한 탐지 성능이 높은 것이다.

표 5 이상탐지모듈 성능 측정

모델 기법	새로운 공격	오분류율
의사결정트리	UDP 공격	0.142857
	TCP 공격	0.4
	MIX 공격	1
신경망	UDP 공격	0.428571
	TCP 공격	0.533333
	MIX 공격	1

모델 구축 기법별로 살펴보면, 대체적으로 신경망 기법을 이용한 모델이 더 나은 성능을 보임을 알 수 있다. 의사 결정 트리는 정상 모델을 생성할 때 기존의 알려진 공격 특성을 가장 잘 반영하는 속성만을 선택하여 모델 생성에 사용하는 반면, 신경망 기법은 모든 기본 속성을 사용하여 모델을 생성하므로 의사 결정 트리에 비해 알려진 공격 형태에 덜 국한된 모델이 생성된다. 새로운 공격 형태별로는 MIX Flood 공격을 새로운 공격이라 가정한 경우, 기존 UDP, TCP Flood 공격과 유사한 형태이므로 정상이 아님을 100% 탐지했으며, UDP Flood 공격을 새로운 공격이라 가정한 경우, 기존 TCP, MIX Flood 공격과 성격이 많이 다르므로 다른 공격에 비해 비교적 낮은 탐지 성능을 나타냈다. 그러나 오용탐지모듈에서 나타나는 오분류율에 비해 월등히 높은 오분류율을 보임으로써 데이타 마이닝을 이용한 이상탐지모듈, 특히 신경망 기법을 이용한 이상탐지모듈이 새로운 공격을 탐지할 수 있는 가능성과 그 높은 성능을 증명할 수 있었다.

이상으로 3장에서 언급한 제안된 탐지 구조의 3가지 목표의 달성을 다음과 같이 입증할 수 있었다. 즉, 오용탐지모듈에 의해 최신 분산 서비스거부 공격에 대한 높은 탐지율의 결과를 얻을 수 있었고, 이상탐지모듈에 의해 기존 공격의 변형이나 새로운 공격에도 탐지할 수 있는 가능성을 검증할 수 있었다. 또한 이에 사용되는 데이타가 기본적으로 많은 전처리 과정이 요구되지 않고, 탐지 구조가 새로운 공격 패턴에 적응할 수 있도록 주기적인 모델 갱신을 수행하므로 실제적이고 효율적인 탐지 구조를 갖음을 알 수 있다.

5. 결론 및 향후 연구과제

현재 분산 서비스거부 공격의 피해사례 증가와 그 심각성으로 인해 적절한 대응책이 시급하며, 공격 특성상 이미 발견된 공격 유형과 정상 트래픽에 대한 정확한 모델링을 통한 빠른 탐지가 가장 중요한 방안이라 할 수 있다. 본 논문에서는 먼저 분산 서비스거부 공격 유형 및 툴들의 기술 동향을 분석하고, 관련 연구들의 메커니즘 및 각각의 한계점을 비교 설명하였다. 이를 바탕으로 분산 서비스거부 공격의 효과적인 탐지를 위해서 NetFlow의 플로우 기반 트래픽을 사용하여 데이타를 수집하였고, 데이타 마이닝 기법을 이용하여 다양한 접근 방법으로 공격 패턴 및 정상 패턴을 모델링하여 실시간 탐지가 가능하도록 탐지 구조를 제안하였다. 본 논문에서 제안한 탐지 구조는 크게 이미 발견된 공격들을 유형별로 분류하여 모델을 생성하는 오용탐지모듈과 분산 서비스거부의 일반적인 특징이 잘 드러나도록 NetFlow 후보 속성을 이용하여 공격과는 구분되는 정상 패턴을 모델링하는 이상탐지모듈로 구성되어 있다. 특히 전자의 모듈은 의사 결정 트리 알고리즘을 이용한 자동 중요 속성 추출 단계와 신경망 기법으로 탐지 모델을 생성하는 분류기 생성 단계로 구성된다.

본 논문에서 제안한 탐지 구조의 성능을 검증하기 위해서 두 가지 측면에서 실험하였다. 이를 위해 먼저 NetFlow를 통해 정상 트래픽을 수집하였고, 공격 트래픽을 얻기 위해 실제 공격 테스트 실험을 수행하여 수집하였다. 이를 이용한 첫 번째 실험은 오용탐지모듈의 결합된 데이타 마이닝 기법의 성능을 검증하기 위하여 다른 데이타 마이닝 기법의 단독 사용 예와 비교 실험하였고, 그 결과 제안한 결합된 기법이 가장 나은 성능을 보임을 검증하였다. 두 번째 실험에서는 이상탐지모듈의 성능을 검증하기 위하여, MIX 공격, TCP SYN Flood 공격, UDP Flood 공격 중 각각 하나의 공격을 새로운 공격으로 가정하고 나머지 두 공격 유형과 정상 트래픽만으로 모델링한 후, 성능 검증 시험에서는 새로운 공격으로 가정한 공격의 트래픽을 추가하여 분류 정확도 시험을 수행하였다. 그 결과 새롭게 추가된 공격에 대해 오분류율이 높게 측정됨으로써 새로운 공격의 탐지 가능성을 검증할 수 있었다.

향후 과제로서 더욱 다양한 공격 시나리오를 통해 수집한 데이타를 통해 신경망 모델의 구조 및 선정 기준 변화를 통한 다각적인 실험 및 기 제안된 탐지 기법과의 성능 비교 실험을 수행하고자 한다.

참 고 문 헌

[1] Remote Network Monitoring (rmonmib), http://www.ietf.org/html.charters/rmonmib-charter.html.

[2] Wenke Lee, Salvatore J. Stolfo, "Data Mining Approaches for Intrusion Detection," Proc. of the 7th USENIX Security Symposium, pp. 79-94, Jan. 1998.

[3] 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.

[4] P. Phaal, S. Panchen, N. McKee, "InMon Corporation's

sFlow : A Method for Monitoring Traffic in Switched and Routed Networks," RFC 3176, 2001.
[5] Paul J. Criscuolo, "Distributed Denial of Service - Trin00, Tribe Flood Network, Tribe Flood Network 2000, and Stacheldraht," CIAC-2319, Feb. 2000.
[6] Jelena Mirkovic, Gregory Prier, Peter Reiher, "Attacking DDoS at the Source," Proc. of ICNP 2002.
[7] 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.
[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] Howard F. Lipson, "Tracking and Tracing Cyber--Attacks: Technical Challenges and Global Policy Issues," Special Report, CMU/SEI-2002-SR-009, Nov. 2002.
[11] 이형우, "DDoS 해킹 공격 근원지 역추적 기술", 정보보호학회지, 13권 5호, 2003년 10월.
[12] P. Ferguson and D. Senie, "Network Ingress Filtering: Defeating Denial of Service Attacks which employ IP Source Address Spoofing," IETF RFC2827, May 2000.
[13] Heather L. Flanagan, "Egress filtering-keeping the Internet safe from your systems," http://www.giac.org/-practical/gsec/Heather_Flanagan_GSEC.pdf.
[14] Kihong Park, Heego Lee, "On the Effectiveness of Route-Based Packet Filtering for Distributed DoS Attack Prevention in Power-Law Internets," Proc. of ACM SGOMM, pp. 15-26, 2001.
[15] Cheng Jin, Haining Wang, Kang G. Shin, "Hop-Count Filtering: An Effective Defense Against Spoofed Traffic," Proc. of the 10th ACM Conference on Computer and Communication Security, 2003.
[16] Tao Peng, Chris Leckie, Rao Kotagiri, "Protection from Distributed Denial of Service Attacks Using History-based IP Filtering," ICC 2003.
[17] Jiawei Han, Micheline Kamber, "Data Mining: Concepts and Techniques," Morgan Kaufmann Publishers.

김 미 희
1997년 이화여자대학교 전자계산학과 학사. 1999년 이화여자대학교 컴퓨터학과 석사. 1999년 (주)인티 연구원. 1999년~2003년 한국전자통신연구원 연구원. 2003년~현재 이화여자대학교 컴퓨터학과 박사과정. 관심분야는 네트워크 보안, NEMO-(NEtwork MObility) 보안, Sensor Network 보안

나 현 정
2002년 이화여자대학교 컴퓨터학과 학사
2004년 이화여자대학교 컴퓨터학과 석사
2004년~현재 삼성전자 네트워크사업부 WLAN 선행개발팀. 관심분야는 네트워크 보안, WLAN

채 기 준
1982년 연세대학교 수학과 이학사. 1984년 미국 Syracuse University 컴퓨터학과 이학석사. 1990년 미국 North Carolina State University 컴퓨터공학과 공학박사. 1990년~1992년 미군 해군사관학교 컴퓨터학과 조교수. 1992년~현재 이화여자대학교 컴퓨터학과 교수. 관심분야는 네트워크 보안, 액티브 네트워크 보안 및 관리, 인터넷/무선통신망/고속통신망 프로토콜 설계 및 성능분석

방 효 찬
1995년 3월 홋카이도 공업대학 경영공학과 공학사. 1997년 3월 홋카이도 공업대학 기계시스템공학과 공학석사. 1997년~1999년 한국통신 운용연구단 전임연구원. 2000년~현재 ETRI 정보보호연구단 선임연구원. 관심분야는 정보공학, 시스템공학, 네트워크 보안 관리

나 중 찬
1986년 2월 충남대학교 계산통계학과 이학사. 1989년 2월 숭실대학교 전자계산학과 공학석사. 2004년3월 충남대학교 컴퓨터과학과 이학박사. 1989년~현재 ETRI 능동보안기술연구팀 팀장. 관심분야는 실시간시스템, 네트워크 관리, 네트워크보안