제20회 한국정보처리학회 추계학술발표대회 논문집 제10권 제2호 (2003. 11)

Netflow 트래픽을 이용한 분산 서비스거부 공격 탐지 기법

나현정*, 김미희*, 채기준*, 나중찬**
*이화여자대학교 컴퓨터학과
**한국전자통신연구원
email : hjna@ewha.ac.kr

Distributed Denial of Service Attack Detection using Netflow Traffic

Hyunjung Na*, Mihui Kim*, Kijoon Chae*, Jung Chan Na**
*Dept. of Computer Science and Engineering, Ewha Womans University
**Electronics and Telecommunications Research Institute

요 약

최근 분산 서비스거부 공격에 의한 특정 서버의 기능 마비, 더 나아가 네트워크 전체를 마비시키는 사례가 증가하고 있다. 이로 인한 피해의 심각성을 고려해 볼 때 적절한 대응이 시급한 실정이지만, 공격 특성상 공격을 탐지해 내기가 어렵다는 문제점이 있다. 본 논문에서는 분산 서비스거부 공격의 패턴을 파악하여 공격을 효과적으로 탐지할 수 있는 방법을 제안하였다. 공격 패턴 파악을 위해서 시스코사에서 개발한 Netflow 데이터를 이용하여 트래픽을 분석하고, 그 결과로 분산 서비스거부 공격의 효과적인 탐지에 공헌도가 큰 속성들을 추출하였다. 실제 인터넷 망에 연결된 라우터에서 수집한 Netflow 데이터와 분석에 의해 추출된 속성을 기반으로 데이터 마이닝 기술을 이용하여 공격 탐지의 성능을 측정하였다.

1. 서론

분산된 공격 에이전트가 한 시스템의 자원을 독점하거나 또는 모두 사용해버리거나 파괴함으로써 다른 사용자들이 이 시스템의 서비스를 올바르게 사용할 수 없도록 만들고, 더 나아가 네트워크 전체를 마비시키는 분산 서비스거부 공격 사례가 증가하고 있다. 그로 인한 피해의 심각성을 고려해 볼 때 적절한 대응이 시급한 실정이다. 그러나 분산 서비스거부 공격은 그 특성상 미리 공격을 탐지해서 대응하기가 매우 어렵다. 공격 트래픽의 일반적인 특징 부재, 관리 도메인들간의 협동 대응 부족, 자동화되고 지능적인 공격 도구, 소스 IP 위조를 통한 실체 숨기기, 수많은 인터넷 호스트 상의 보안 취약점 등이 그 이유이다. 이처럼 공격에 대한 완벽한 보안책이 불가능한 상황에서 공격 대상 시스템 혹은 네트워크를 완전히 마비시키기 전에 가능한 한 미리 공격을 탐지하는 것이 현재로선 가장 최선의 보안 방법이다.

본 논문에서는 분산 서비스거부 공격을 탐지하기 위하여 트래픽 정보 분석과 데이터 마이닝 기법을 이용하였다. 네트워크 트래픽 정보는 SNMP(Simple Network Management Protocol) MIB(Management Information Base), RMON(Remote network MONitoring) MIB 정보나 tcpdump 를 통해서 얻을 수 있다. 또한 시스코사에서 개발한 Netflow 를 이용해서 네트워크 트래픽 상태를 모니터링할 수도 있다. 각각의 데이터에서 분석할 수 있는 정보는 조금씩 다르며 본 논문에서는 플로우 기반으로 데이터를 분석하는데 효과적인 Netflow 트래픽 정보를 이용하였다.

데이터 마이닝 기술은 대량의 데이터로부터 의미 있는 정보를 추출해 내는데 많이 이용되는 방법이다. 군집화, 분류, 연관규칙 등의 데이터 마이닝 기법은 침입 탐지 분야에서 비정상행위 탐지와 관련하여 많이 연구되어 왔다. 본 논문에서는 실제 분산 서비스거부 공격에 의한 액세스 라우터에서의 Netflow 데이터를 기반으로 분석을 통하여 중요한 속성을 추출해 내고, 이를 데이터 마이닝 기술을 이용해 탐지의 성능을 측정하고자 한다.

본 논문의 구성은 다음과 같다. 2 장에서는 분산서비스부인 공격 탐지와 관련된 연구들을 간략히 살펴보고, 3 장에서는 공격 도구들의 특성을 알아본다. 4 장에서는 본 논문에서 사용한 Netflow 데이터 분석과 데이터 마이닝 기법을 살펴보고, 5 장에서는 실험 내용 및 결과를, 마지막으로 6 장에서는 결론을 내린다.

2. 관련 연구

분산 서비스 부인 공격에 대응하기 위한 관련 연구들이 다각적으로 이루어지고 있다.

Xuan 은 고속의 네트워크 환경에서 분산 서비스거부 플러

딩 공격을 저지할 수 있는 방어 시스템을 제안하였다[1]. 게이트웨이를 기반으로 하는 접근 방법으로서 네트워크 상의 여러 위치에 게이트웨이들을 분산 설치하여 이들이 서로 협력하여 공격 탐지나 트래픽 접근 제어와 같은 공격 대응 기능을 수행하도록 하였다. 각 게이트웨이는 현재 지나가는 트래픽의 일부만을 트래픽 샘플링을 통해 선택하여 TCP-ACK 기반의 공격 탐지나 신뢰도 기반의 트래픽 접근 제어 기능을 수행한다. 그러나 여기서 제안한 ACK 을 기반으로 한 정상 트래픽의 판별은 다양한 분산 서비스거부 공격을 탐지하지 못하는 한계점을 갖는다.

Park은 라우팅 정보를 이용하는 방법으로 공격에 대응하고자 하였다[2]. 라우터에 도착한 패킷의 소스 주소를 보고 라우팅 정보에서 얻어진 소스 호스트로부터의 경로가 실제 들어온 경로와 일치하는지를 판단하여 소스 주소의 스푸핑 여부를 알 수 있다. 그 결과 스푸핑 됐다고 판단된 패킷은 필터링하게 된다. 이 방법은 경로 기반의 패킷 필터링을 갖춘 라우터를 어떤 위치에, 얼마나 많은 라우터에 설치하느냐에 따라 공격을 막는 성능이 좌우된다. 또한 올바른 경로로 들어오는 스푸핑 된 패킷에 대해서는 필터링 하지 못하는 한계점을 갖는다.

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[3]. 각 소스 IP 주소별 나타나는 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.

대부분의 관련 연구들이 특정 공격에 국한된 특성만을 고려하여 공격 탐지 및 대응 메커니즘을 제안하기 때문에 실제 공격 상황에 대처하는데 많은 한계점을 갖는다. 본 논문에서는 보다 실제 네트워크에 반영되는 공격 트래픽에서 나타나는 특성들을 고려하고자 하였다.

3. 공격 도구 분석

현재 사용되고 있는 분산 서비스거부 공격 도구에는 Trinoo, TFN, Stackeldraht, TFN2K, Mstream, Shaft 등이 있다. 이 도구들이 가지고 있는 기본적인 기능은 시스템에 대해 취약성을 검사한 다음 그 취약성을 이용하여 접근하여 마스터 또는 에이전트 도구를 대상 시스템 내에 설치하는 것이다[4]. 아래는 본 논문의 실험에서 사용된 공격 도구별 특성이다.

표 1. 공격 도구별 특성

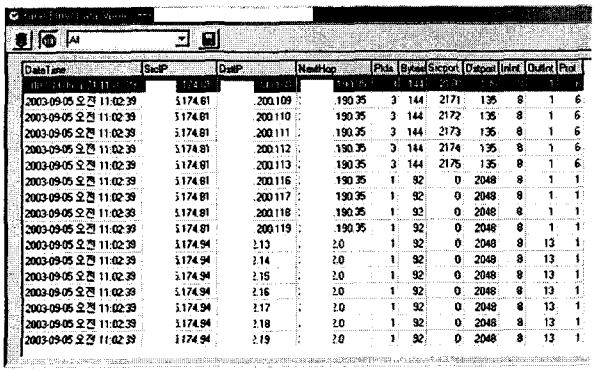
	Trinoo	TFN2k	Stacheldraht	Synk4
공격 유형	UDP flood	UDP/SYN/ ICMP flood, Smurf	UDP/SYN/ ICMP flood, Smurf	SYN flood
소스 IP	스푸핑 안됨	스푸핑 정도 조절 가능	자동 스푸핑	스푸핑
소스 포트	지정 불가	자동선택 (랜덤/순차적)	자동선택 (랜덤/순차적)	자동선택 (랜덤)
타겟 포트	지정 불가	지정가능	범위지정	범위지정
기타		· 제어메세지의 일방향 통신 · 마스터와 에이전트간 통신 암호화	· 에이전트의 자동 업데이트 ·공격자/마스터/에이전트간 통신 암호화	

4. 제안한 탐지 메커니즘

본 장에서는 분산서비스부인공격의 효과적인 탐지를 위해 사용한 Netflow 와 데이터 마이닝 기법에 대해 설명한다.

4.1. Netflow 트래픽 분석

Netflow 는 시스코사에서 개발한 프레임워크로 네트워크 서비스 이용 과금이나 모니터링에 주로 사용된다. 구조적으로 네트워크 장비에 장착된 Netflow 엔진이 데이터를 저장소로 내보내면 데이터가 적절하게 수집, 필터링, 저장되고, 이를 사용자가 원하는 목적에 따라 불러와 분석하는데 사용하게 된다. Netflow 에 저장되는 기본 단위는 플로우이다. 플로우란 주어진 근원지 IP 와 목적지 IP 사이의 단방향 패킷 스트림을 말한다. 각 플로우는 근원지와 목적지 IP 주소, 근원지/목적지 포트 번호, 3 계층 프로토콜 타입, TOS, 그리고 입력 논리 인터페이스로 정의 된다[5]. 따라서 실시간으로 각 플로우별로 위의 정보들을 확인할 수 있다. 또한 근원지 혹은 도착지 주소, AS, 포트별로 플로우수, 패킷수, 옥텟수를 각각 알 수 있으며 3 계층 프로토콜별 혹은 근원지-도착지 주소쌍별 플로우수, 패킷수, 옥텟수를 알 수도 있다. 이러한 정보들을 이용해서 현재 네트워크의 이상 여부를 탐지할 수 있다. 그림 1 은 실시간 Netflow 데이터의 한 예로, 135 번 포트로 나가는 블래스트 웜 트래픽과 비정상적인 ICMP 트래픽을 관찰할 수 있다.



DateTime	SrcIP	DstIP	NextHop	Pkts	Bytes	Srcport	Dstport	Inint	Outint	Prot
2003-09-05 오전 11:02:39	3.174.81	.200.109	.190.35	3	144	2171	135	8	1	6
2003-09-05 오전 11:02:39	3.174.81	.200.110	.190.35	3	144	2172	135	8	1	6
2003-09-05 오전 11:02:39	3.174.81	.200.111	.190.35	3	144	2173	135	8	1	6
2003-09-05 오전 11:02:39	3.174.81	.200.112	.190.35	3	144	2174	135	8	1	6
2003-09-05 오전 11:02:39	3.174.81	.200.113	.190.35	3	144	2175	135	8	1	6
2003-09-05 오전 11:02:39	3.174.81	.200.116	.190.35	1	92	0	2048	8	1	1
2003-09-05 오전 11:02:39	3.174.81	.200.117	.190.35	1	92	0	2048	8	1	1
2003-09-05 오전 11:02:39	3.174.81	.200.118	.190.35	1	92	0	2048	8	1	1
2003-09-05 오전 11:02:39	3.174.81	.200.119	.190.35	1	92	0	2048	8	1	1
2003-09-05 오전 11:02:39	3.174.94	2.13	2.0	1	92	0	2048	8	13	1
2003-09-05 오전 11:02:39	3.174.94	2.14	2.0	1	92	0	2048	8	13	1
2003-09-05 오전 11:02:39	3.174.94	2.15	2.0	1	92	0	2048	8	13	1
2003-09-05 오전 11:02:39	3.174.94	2.16	2.0	1	92	0	2048	8	13	1
2003-09-05 오전 11:02:39	3.174.94	2.17	2.0	1	92	0	2048	8	13	1
2003-09-05 오전 11:02:39	3.174.94	2.18	2.0	1	92	0	2048	8	13	1
2003-09-05 오전 11:02:39	3.174.94	2.19	2.0	1	92	0	2048	8	13	1

그림 1. 실시간 Netflow 데이터

기존의 공격 툴은 대부분 근원지 주소를 매 패킷마다 랜덤하게 스푸핑하고, 타겟 포트 역시 랜덤 혹은 순차적으로

변화시켜서 보내게 된다. 이로 인해 공격 시의 트래픽은 데이터량에 비해 플로우 수가 급증하는 현상을 보이고, 그로 인해 플로우 당 데이터량이나 패킷량은 정상일 경우보다 감소할 것이다. 이러한 공격 특성을 Netflow 에 저장된 플로우수, 패킷수, 옥텟수의 비율을 계산해 봄으로써 이상 트래픽을 탐지해 낼 수 있다.

4.2. 데이터 마이닝을 이용한 성능 측정

Netflow 트래픽 분석을 통해 얻어진 공격 트래픽의 패턴을 검증하고 탐지의 정확성을 측정하기 위해서 데이터 마이닝 방법을 사용하였다. 4.1 의 분석 결과로 공격 탐지에 중요한 속성으로 추측된 {옥텟수/플로우수}와 {패킷수/플로우수}를 계산한 후 이들을 트레이닝 데이터의 속성에 포함시키고, 실제 데이터 마이닝 툴을 이용해서 신경 네트워크 혹은 클러스터링 기법을 적용시켜 모델을 구축한다. 대량의 트레이닝 데이터를 통해 얻어진 모델에 테스트 데이터를 입력하여 출력 결과에 따라 모델에 의한 공격 탐지 정확성을 측정하였다.

5. 실험 및 결과

5.1. 실험 환경

실제 네트워크의 트래픽을 관리하기 위해 설치된 Netflow 를 이용하여 네트워크 트래픽 정보를 수집하였다. 데이터 마이닝을 이용한 패턴 모델을 구축하기 위한 트레이닝 데이터를 얻기 위해서는 정상 트래픽과 공격시의 이상 트래픽이 필요하다. 공격 시의 이상 트래픽을 얻기 위해 실제 분산 서비스거부 공격에 쉽게 이용될 수 있는 공격 툴을 이용해서 다양한 시나리오로 직접 공격을 수행한 다음, 그 때 나타나는 트래픽을 이상 트래픽으로 분류하였다.

공격에는 널리 알려진 서비스 부인 공격 툴인 TFN2k, Stacheldraht, Synk4 를 사용하였으며 공격 타입은 실제 공격에서 가장 큰 비율을 차지하고 있는 TCP SYN 플러드 공격을 위주로 하였다. 그리고 근원지 주소를 위조하는 스푸핑 정도를 완전 랜덤에서 자신의 서브네트워크 주소를 가지는 정도까지 다양한 형태로 공격해 보았다. 공격은 학교망 외부에서 내부로의 공격과 내부에서 외부로의 공격, 두 가지 경우 모두 실험해 보았다. 현실적으로 심각한 피해를 초래하는 공격은 직접 해볼 수 없기 때문에 단시간 동안만 공격을 수행하였다. 따라서 실제 공격 시에는 실험 결과보다 더 큰 폭의 증감이 나타날 것이다.

이렇게 만들어낸 공격시의 이상 트래픽과 평소의 정상 트래픽은 Netflow 데이터베이스에 여러 개의 테이블로 저장되어 있다. 일정 시간 단위로 저장되어 있는 트래픽 정보 중에 TCP 의 플로우, 옥텟, 패킷의 수를 분석하였으며, 그 결과는 다음과 같다.

5.2. 실험 결과

공격이 일어난 경우 기본적으로 평소보다 많은 플로우수와 옥텟수, 패킷수가 관찰된다. 그러나 이 세 값들간의 비는 정상 트래픽과 공격 트래픽에서 다르게 나타난다. 3 장에서 설명했듯이 공격 도구들은 일반적으로 매 패킷마다 소스 주소나 포트를 랜덤하게 혹은 순차적으로 변경시켜서 보낸다. 이 때문에 공격이 진행되는 동안에는 플로우의 수가 급증하게 된다. 그림 2 는 정상적인 트래픽과 공격을 수행한 비정상 트래픽의 TCP 플로우의 수를 5 분을 단위시간 1 로 하여 나타낸 것이다. 각각의 트래픽은 3 시간 동안 관찰한 데이터로 대체로 공격을 수행한 트래픽에서 더 많은 플로우가 발생함을 관찰할 수 있다.

그림 3 에서 나타나는 바와 같이, 정상 트래픽에서 보여지는 데이터량과 플로우수의 비가 공격 시에는 플로우수의 급격한 증가로 {옥텟수/플로우수}가 감소하는 것을 관찰할 수 있다.

그리고 각 공격 패킷이 서로 다른 플로우로 인식되므로 그림 4 와 같이 {패킷수/플로우수} 역시 감소하는 것을 볼 수 있다.

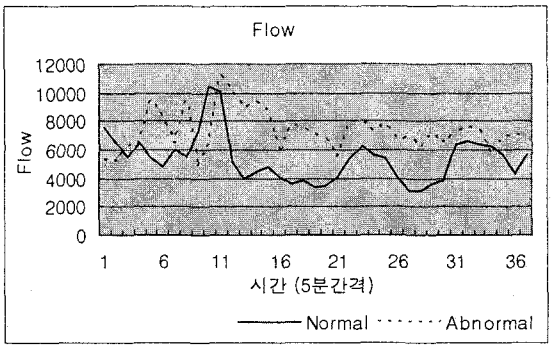


그림 2. 정상 트래픽과 비정상 트래픽의 플로우수 비교

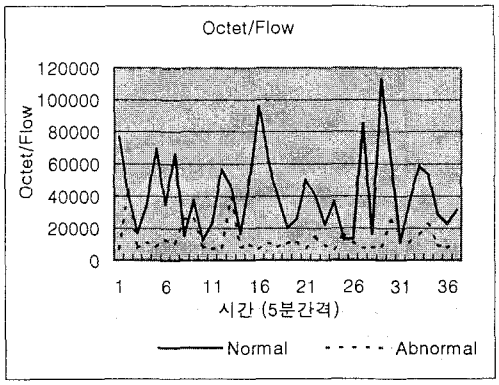


그림 3. 정상 트래픽과 비정상 트래픽의 Octet/Flow 비교

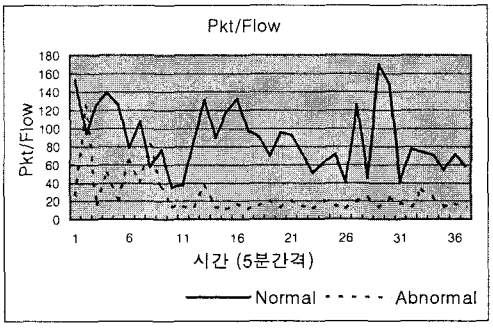


그림 4. 정상 트래픽과 비정상 트래픽의 Packet/Flow 비교

실험을 통해 수집한 공격 데이터와 정상 데이터가 혼재된 대량의 Netflow 트래픽 데이터를 그림 5 와 같은 Neural Connection 이라는 데이터 마이닝 툴을 이용해서 모델을 구축하였다[6].

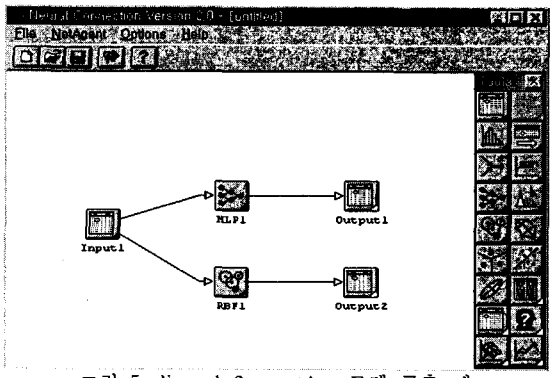


그림 5. Neural Connection 모델 구축 예

본 논문에서는 다층 퍼셉트론을 이용하여 예측 모델을 구축하였으며, 이에 테스트 데이터를 입력하여 얻은 결과는 아래 그림과 같다. 그림 6 의 첫번째 그림은 입력 데이터를 이용하여 트레이닝한 결과이고, 두번째 그림은 트레이닝 결과 생성된 모델에 테스트 데이터를 입력하여 정상인지 비정상인지 예측한 결과이다. 본 실험 결과, 트레이닝 정확도와 예측 정확도는 각각 95.6%와 90.9%로 측정되었다.

	Float var_0001	Float var_0002	Float var_0003	Symbol var_0004 T	Symbol MTarget M	Symbol Output O
1	18593.0	41.438014	25631.579681	Normal	Normal	Normal
2	19440.0	50.819313	33157.539529	*Normal*	*Normal*	*Normal*
3	20186.0	87.002923	62258.900376	Normal	Normal	Normal
4	17210.0	29.034689	18323.782394	Normal	Normal	Normal
5	21545.0	44.861592	31119.796983	Normal	Normal	Normal
6	20828.0	24.415978	12811.767857	Normal	Normal	Normal
7	5235.0	26.89914	6947.174976	Abnormal	Abnormal	Abnormal
8	5187.0	125.437247	43732.278581	Abnormal	Abnormal	Normal
9	16840.0	36.329335	28581.178979	Normal	Normal	Normal
10	17119.0	102.855015	55085.206087	Normal	Normal	Normal
11	16347.0	66.742644	34552.073163	Normal	Normal	Normal
12	18267.0	30.148026	16056.503476	Normal	Normal	Normal
13	6499.0	40.973534	10017.528697	Abnormal	Abnormal	Abnormal
14	9495.0	82.192944	26803.809742	Abnormal	Abnormal	Normal
15	4837.0	35.639446	25857.442836	Abnormal	Abnormal	Abnormal
16	6461.0	14.038539	8490.327349	Abnormal	Abnormal	Abnormal

	Float var_0001	Float var_0002	Float var_0003	Symbol var_0004 T	Symbol MTarget1 M	Symbol Output1 O
1	X	6715.0	44.193994	25509.105436	Normal	Normal	Abnormal
2	X	8560.0	74.532126	33447.214953	Normal	Normal	Normal
3	X	7169.0	12.776119	8569.159018	Abnormal	Abnormal	Abnormal
4	X	6481.0	24.036877	24794.056318	Abnormal	Abnormal	Abnormal
5	X	7245.0	17.294686	16622.678537	Abnormal	Abnormal	Abnormal
6	X	6437.0	138.846357	34842.907568	Normal	Normal	Normal
7	X	5382.0	125.881271	69255.265886	Normal	Normal	Normal
8	X	4822.0	78.540854	35180.176898	Normal	Normal	Abnormal
9	X	6854.0	13.548731	9346.976551	Abnormal	Abnormal	Abnormal
10	X	7133.0	16.051255	9181.661994	Abnormal	Abnormal	Abnormal
11	X	6794.0	13.115249	10563.458934	Abnormal	Abnormal	Abnormal

그림 6. 구축 모델 결과

6. 결론

본 논문에서는 분산 서비스거부 공격의 효과적인 탐지를 위해 공격 도구들의 특성을 파악하고, Netflow 를 이용한 데이터 분석을 통하여 공격 패턴을 나타내는 속성을 발견하였다. 또한, 데이터 마이닝 기법을 이용하여 공격 패턴 모델을 구축하고 성능을 검증하였다. 실험을 통하여 직접 공격 트래픽을 발생시켜 실제 공격과 유사한 공격 트래픽을 얻음으로써 좀 더 정확한 공격 패턴을 추출하고자 하였다. 본 논문의 의의는 Netflow 를 통해 얻은 실제 네트워크 데이터와 공격 툴을 이용한 공격 데이터를 토대로 분석함으로써 실제 네트워크 트래픽을 많이 반영한다는 점이다. 또한 데이터 마이닝 기법을 이용해서 트래픽 분석을 통해 추출한 공격 패턴의 성능을 검증하였다.

향후 과제로는 여러 가지 데이터 마이닝 기법에 따른 공격 탐지 성능을 비교하여 좀 더 향상된 탐지 모델을 찾는 것이 필요하다. 또한 본 논문에서 주로 다룬 TCP SYN 플러드 공격 외에 다양한 분산 서비스거부 공격 및 웜 바이러스 등의 네트워크 대역폭을 낭비하는 다른 종류의 이상 트래픽들도 탐지 가능하도록 다양한 패턴을 분석하는 연구가 필요하다.

참고문헌

[1] Dong Xuan, Riccardo Bettati, Wei Zhao, "A Gateway-based Defense System for Distributed DoS Attacks in High-Speed Networks," Proc. of The IEEE Workshop on Information Assurance and Security, 2001.

[2] 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, 2001.

[3] 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.

[4] Paul J. Criscuolo, "Distributed Denial of Service - Trin00, Tribe Flood Network, Tribe Flood Network 2000," CIAC-2319, 2000.

[5] NetFlow Services Solutions guide, http://www.cisco.com/univercd/cc/td/doc/cisintwk/intsolns/netflsol/nfwhite.pdf

[6] 조용준, "Neural connection 을 이용한 데이터 마이닝 신경망 분석," 고려정보산업, 1999.