情報保護學會論文誌
第13券 第2號, 2003. 4

베이지언 추정을 이용한 웹 서비스 공격 탐지*

조 상 현**, 김 한 성**, 이 병 희**, 차 성 덕**

SAD : Web Session Anomaly Detection based on Bayesian Estimation

Sang-hyun Cho**, Han-sung Kim**, Byung-hee Lee**, Sung-deok Cha**

요 약

본웹 서비스는 일반적으로 침입 차단 시스템에 의해 통제되지 않은 채 외부에 공개되어 있어 공격의 수단으로 이용될 수 있고, 다양한 웹 어플리케이션의 특성에 따라 많은 형태의 취약성을 내포하고 있다. 본 논문에서는 웹 서비스의 정상적인 이용 사례를 모델링하고, 이와 다른 사용례를 보이는 이상 사례를 베이지언 추정 기법을 이용하여 통계적으로 찾아내는 SAD(Session Anomaly Detection)을 제안한다. SAD의 성능을 평가하기 위하여 1개월간 수집된 웹 로그 자료를 이용하였고 침입은 웹 스캐너 프로그램(Whisker)을 이용하여 수행하였다. 기존 NIDS인 Snort를 이용한 실험 결과 평균적으로 36%의 탐지율을 보인 반면 SAD의 경우 윈도우 사이즈, 훈련데이터의 크기, 이상탐지 필터, 웹 토폴로지 정보의 이용유무에 따라 다소 차이는 있지만 전반적으로 90%가 넘는 탐지율을 보여 주었다.

ABSTRACT

As Web services are generally open for external uses and not filtered by Firewall, these result in attacker's target. Web attacks which exploit vulnerable web-applications and malicious users' requests cause economical and social problems.

In this paper, we are modelling general web service usages based on user-web-session and detect anomal usages with Bayesian estimation method. Finally we propose SAD(Session Anomaly Detection) for detection unknown web attacks. To evaluate SAD, we made an experiment on attack simulation with web vulnerability scanner, whisker. The results show that the detection rate of SAD is over 90%, which is influenced by several features such as size of window or training set, detection filter method and web topology.

Keyword : *intrusion detection, web attack detection, anomaly detection, network security, bayesian estimation*

Ⅰ. 서 론

네트워크 인프라의 급속한 발전과 더불어 인터넷 환경에서의 이상 행위들의 발생 빈도가 빠른 속도로 증가하고 있고, 그로 인해 발생하는 정치, 경제, 사회에 미치는 피해는 매우 심각한 문제로 대두되고 있다.

네트워크 오용에 대한 대응책으로 여러 보안 메커니즘이 제안되어 사용되고 있는데, 대표적으로 침입 차단 기법과 침입 탐지 기법을 들 수 있다. 침입 차단은 사전에 정해져 있는 접근 제어 목록에 따라 내, 외부의 접근 행위를 통제하며, 침입 탐지는 시스템에 따라 비정상적인 공격 행위들을 실시간으로 발견해 준다.

침입 탐지[6]는 알려진 공격 기법을 분석하고 패턴화 하여 이 패턴을 따르는 이벤트를 분석해 내는 오용 탐지 기법과 일정 기간 이벤트에 대한 프로파일링을

* 본 연구는 첨단정보기술 연구센터를 통하여 과학재단의 지원을 받았음.
** 한국과학기술원 전자전산학과 전산학 전공{shcho, bhlee, kimhs, cha}@salmosa.kaist.ac.kr

통해 기존 프로파일과 다른 유형의 이벤트의 발생을 분석하는 이상 탐지 기법이 있다.

오용 탐지 기법의 경우 알려지지 않은 새로운 공격에 대해서는 근본적으로 대응하기 어려운 면이 있고, 이상 탐지 기법에서는 false alarm 비율이 높다는 문제점을 가지고 있다. 최근의 연구에서는 오용 탐지 기법과 이상 탐지 기법을 혼용하거나 이상 탐지에서의 false alarm을 최소화하기 위해 노력하고 있다.

방화벽과 같은 접근 통제 솔루션과 공격 시그너쳐 기반의 침입 탐지 시스템을 활용하여 외부로부터의 접근을 통제함으로써 과거만큼의 침입에 대한 피해는 막을 수 있게 되었다. 그러나, 웹서비스와 같은 경우에는 누구나 외부에서 사용할 수 있도록 제공되어야 하므로 접근통제의 메커니즘을 적용할 수 없거나, 제한적으로 밖에는 사용할 수밖에 없다. 따라서 그 효과를 제대로 발휘할 수 없고, 최근에는 이와 같은 서비스의 취약점을 이용한 비정상적이거나 악의적인 공격 행위들이 많이 발생하고 있다.

웹은 기본적으로 다른 연동 어플리케이션과의 상호작용을 할 수 있는 통로 역할을 하고 클라이언트로 부터의 요청을 어플리케이션 서버 프로그램에 전달하고 수행 결과를 다시 돌려주는 매개체 역할을 하고 있다. 따라서, 다른 서비스와 같이 일반적인 IP 기반의 접근 인증은 큰 의미가 없으며, 외부에 웹 포트 접근은 허용되어 있는 상태이다.

또한 웹 어플리케이션은 클라이언트의 웹 브라우저, 웹 인터페이스, 웹 서버, 전 처리기, 데이터 베이스와 같은 요소들로 계층적으로 이루어지므로, 공격이 발생하는 지점과 각 위치에서의 취약 요소들이 상이할 수 있다. 따라서 웹 기반의 공격에 대한 탐지 기법은 보호 시스템이 제공하고 있는 웹 서비스의 특징과 어플리케이션에 많이 좌우되고, 이에 잘 대응할 수 있는 방법이 만들어져야 하므로 기존의 시그너쳐 기반의 패턴 매칭을 주로 하는 침입 탐지 기법은 효과적으로 대응하기 어렵다. 2001년 전 세계적으로 유행한 코드레드웜이나 님다웜[2]이 좋은 예이다. 기존의 시그너쳐 기반의 침입 탐지에서 코드레드웜이나 님다웜의 특징을 미리 알 수 없었기 때문에 전혀 대응을 할 수 없었으며, 공격이 발생한 후 피해 시스템의 특징 분석을 통해 시그너쳐를 만들어 대처하고 있지만, 이와 유사한 형태의 변형 공격에 대해서는 여전히 대응하기 어렵다.

하지만, 만약 보호하고 있는 사이트의 일반적인 웹 서비스 특성을 분석하고 있었다면 코드레드웜나 님다웜의 전파 방법중 하나인 웹서버 취약점을 이용한 접근 시도 탐지는 가능했을 것이며, 공격이 시도하는 특징적인 작업 순서도 발견할 수 있었을 것이다.

본 논문에서는 접근 통제가 상대적으로 미약한 웹 서비스의 보호를 목표로, 기존의 침입 탐지 기법들이 탐지하기 어려운 변형 혹은 새로운 형태의 공격에 대응할 수 있는 웹 사용자 세션 기반의 이상 탐지 기법을 제안한다.

웹 사용자 세션 기반의 이상 탐지(이하 SAD: Session Anomaly Detection)는 웹 서비스를 이용하는 사용자들이 접근하는 페이지 목록을 자료화하여 과거의 접근 페이지 목록과 상이한 접근 요청을 찾아 비정상적인 사용자 요청을 탐지한다.

본 논문에서 제안하고 있는 SAD-Bayes는 웹서버의 접근 로그(access log)로부터 웹 페이지에 접근하는 사용자 별 세션을 구분해 내고 이 세션 정보를 바탕으로 사이트의 페이지 접근 순서를 목록화한다. 제안 기법에서는 현재의 웹 페이지 접근 순서가 과거 사이트의 접근 순서 목록과 비교해 볼 때 확률적으로 얼마나 낮은지 베이지언 추정 기법을 이용해 판단한다.

II. 관련 연구

기존의 침입 탐지 기법 중 웹 서비스에 초점을 둔 연구는 많지 않으며, 대부분 일반적인 침입 탐지 기법에 웹 공격과 관련된 공격 시그너쳐를 활용하였다.

네트워크 기반의 오용 탐지 기법의 경우 공격의 탐지는 http 패킷의 내용(content)에 알려져 있는 공격 시그너쳐를 담고 있는지 여부로 판단하고 있다. 대표적인 네트워크 기반 IDS인 SNORT[15]는 웹 공격에 대한 1000여 개의 시그너쳐를 가지고 있다. [그림 1]은 초기 웹서버 설치 시 설치되었던 phf cgi 취약성을 이용하려는 시도를 탐지해내는 SNORT의 시그너쳐이다.

```
alert tcp $EXTERNAL_NET any -> $HTTP_SERVERS 80
(msg:"WEB-CGI phf access";flags:
A+; uricontest:"/php"; nocase; reference:bugtraq,629;
feference:arachnids,128;
reference:cve,CVE-1999-0067; classtype:attempted-recon;
sid:886; rev:3;)
```

〔그림 1〕 Snort IDS의 웹 공격 탐지 시그너쳐

위 시그너쳐에서는 HTTP요청 시 urlcontent에 '/phf'이라는 내용이 들어있을 경우 "WEB-CGI Phf Access 공격"으로 탐지한다. 그런데, 시그너쳐에 기술된 패턴에 약간의 변형을 가할 경우 이를 침입 탐지 시스템이 탐지 못하게 된다. 혹은 반대로 침입 탐지 시스템이 잘못된 경고를 나타내게끔 사전에 알려진 시그너쳐 대로 패킷을 생성하여 전송할 수도 있다.[17] 이 경우 침입 탐지 시스템이 잘못된 오류메시지를 과다하게 생성하여 무력화될 수 있으며, 그 과정에서 실제의 공격 행위들이 감춰질 수 있다.

일반적으로 네트워크 기반의 오용 탐지 시스템들의 경우 주로 네트워크 패킷의 모니터링을 통해 침입을 탐지하는 메커니즘을 사용하기 때문에, 특정 어플리케이션 혹은 시스템에 전달되는 암호화되는 트래픽에 대해서는 대응하지 못하는 한계를 가진다. 또한, 공격에 따라서는 비정상적인 네트워크 트래픽을 일으키지 않아 네트워크 기반 탐지 시스템에서는 탐지할 수 없는 공격이 있다. 예를 들어 시스템 내부로 접근한 후에 일어나는 버퍼 오버플로우 혹은 race condition을 이용한 공격은 특징적인 네트워크 트래픽을 만들지 않는다. 그리고 센서와 보호 시스템간의 운영체제가 동일하지 않을 경우 TCP/IP 프로토콜 구현상의 차이를 이용한 다양한 우회 공격이 가능하고, 거짓 경고(false alert)메시지를 의도적으로 많이 생성하여 실제 공격 행위에 대한 탐지 기록을 발견하기 어렵게 할 수 있다.

따라서 네트워크 기반보다는 호스트 기반의 데이터 수집을 통한 침입 탐지를 시도하였는데, 호스트에 생성되는 로그에 중점을 둔 연구가 많았다. IBM[12]의 경우에도 웹 서버의 대표적인 억세스 로그를 활용하여 알려진 패턴에 대한 탐색을 수행하였으며, 분석된 결과를 몇 가지 클래스로 그룹화 하여 비슷한 부류의 공격을 하나의 이름으로 축약함으로써 생성되는 보고 메시지의 양을 줄일 수 있었다. 그러나, 이 연구에서는 cgi프로그램의 비정상적인 활용과 일부 서비스 공격의 탐지만 가능할 뿐 로그인 절차를 무시한 페이지 접근 또는 특정 웹 어플리케이션이 갖고 있을 취약성을 이용하는 공격에는 대응하기 어렵다는 문제점을 가지고 있다.

한편 호스트 중심이 아닌 네트워크 관점에서 네트워크 트래픽의 이상 현상을 통계적으로 탐지하려는 시도도 진행되고 있는데, 주로 IP패킷의 헤더값의 비정상적인 분포를 탐지하고 있다. 대표적인 연구로서 Mahoney et al.[13]과 Krugel[4]의 연구가 있다. 물론 이러한 시도는 시그너쳐를 이용하여 비교적 정확히 어떤 공격인지 탐지하는 오용 탐지에 비해 특정 공격 현상을 명확히 설명하지는 못하지만, 알려지지 않은 새로운 현상들의 발생을 탐지해 낼 수 있다는 장점을 가지고 있다. 또한 이들의 연구는 네트워크 계층에서 네트워크 패킷의 헤더 상의 이상 유무를 파악하는데 초점을 두었고 이를 통하여 최근에 유행하고 있는 웹 기반의 다양한 변형의 분산 서비스 거부 공격을 효과적으로 탐지할 수 있는 기반을 마련해 주었다.

[13]에서는 IP패킷의 헤더 값을 프로파일링하기 위해 패킷마다 패킷 이상 점수를 산출한다. 패킷 이상 점수는 패킷 필드의 각 값을 1~4 bytes로 나타낸 후 이들에 특정한 해쉬값을 적용하거나 클러스터링을 통해 얻은 값을 이용하여 패킷 내의 필드 값의 발생 확률이 낮고, 최근 발생기간이 오래될수록 즉, 최근에 발생한 적이 없을 수록 이상 점수가 높도록 설계하였다. 공격이 없는 상태의 데이터를 가지고 일정 기간 트래픽을 분석한 후 공격이 있는 일정 기간의 데이터의 트래픽을 대조하여 공격을 탐지하는 방법으로 실험을 진행하였다. 이들은 이러한 식으로 패킷 이상 점수를 산출하여 DARPA의 IDS평가 데이터[10]로 실험한 결과, 우수한 성능을 보여주었으나, IP 패킷 필드의 값만 이용하기 때문에 패킷의 데이터(payload)에 들어있게 되는 HTTP나 SMTP과 같이 어플리케이션 레벨에서 이루어지는 공격의 탐지가 불가능하다는 단점을 가지고 있다.

Krugel은 빈도가 낮은 서비스 요청일수록, 그리고 요청의 길이가 평균보다 클 경우에 침입으로 판단할 확률을 높게 평가하였다. 또한 요청의 데이터 부분의 값의 분포 역시 빈도순으로 정렬하면, 일반적인 상태에서는 그 감소폭이 비교적 완만하나, 비정상적일 경우 특정 값의 분포가 급격히 늘어났기 때문에 감소폭이 상당히 급함을 볼 수 있으므로 분포의 유사성을 이상 점수에 반영하였다. 이는 님다 웜 공격의 예에서도 확인할 수 있는데, 공격은 쉘 코드를 서비스 요청 시 데이터 부분에 넣어서 보내기 때문에 공격 코드의 특정 문자값들이 일시적으로 많이 증가하는 것을 볼 수 있었다. Krugel은 DNS(Domain Name Server)에 관련한 공격 몇 가지를 수행한 후, 이들 패킷들의 이상 점수가 모두 임계치를 넘어 이상으로 탐지되어 이러한 접근 방법이 RtoL(Remote to Local) 공격에 효과적임을 보였다. 하지만, 패킷 내의 데이터 크기 및 문자 분포만을 이용하므로, 웹 공격 시에

나타나는 쉘 코드 등의 비정상적인 인자들의 발견은 가능하지만, 특정 CGI 취약점 스캐닝이나 접근이 허용되지 않는 페이지로의 접근 또는 웹 어플리케이션으로 전달되는 인자값의 악의적인 조작 등과 같은 공격에 대한 탐지는 어렵다.. 특히 내용(Content)의 단일 문자 빈도를 활용한 이상 점수가 얼마나 타당한 지에 대해서는 좀더 다양한 실험 결과가 요구되고 있다.

끝으로, 통계적 기법을 이용한 침입 탐지 연구로써 윈도우즈 운영체제의 레지스트리에 대한 이상 접근을 탐지하는 RAD(Registry Anomaly Detection)[19] 연구가 있다. 이 연구에서는 베이지언 추정 기법을 이용하여 바이러스나 트로이목마 프로그램을 탐지하였다. 본 논문에서 제안하는 웹 세션 이상 탐지 기법은 RAD에 많은 영향을 받았다.

RAD는 윈도우즈 응용 프로그램이 접근하는 레지스트리 키가 비교적 일반화할 수 있도록 일정한 분포를 보이고 있고, 바이러스나 트로이목마와 같은 악의적인 프로그램은 비록 동일한 프로세스 이름을 하고 있더라도 다른 레지스트리 키를 참조한다는 가정에 기반 하였다. RAD에서는 훈련 기간동안의 프로세스들이 참조하는 레지스트리의 빈도를 모델링하여, 시험 기간동안 발견되는 프로세스의 레지스트리 참조가 베이지언 추정에 의한 확률값으로 볼 때 어느 정도 낮은 사건인지로 이상 유무를 판단하였다. 프로세스 이름, 쿼리 형태, 접근키, 성공/실패 유무, 결과값의 5개 속성을 활용하여 베이지언 추정을 하였다. 실험 결과도 좋았으나, 프로세스 이름이 key로 활용되기 때문에 공격자가 공격 프로세스의 이름을 변경하거나, 새로운 형태의 프로세스가 나타날 경우 탐지가 불가능하다.

Ⅲ. SAD(Session Anomaly Detection)

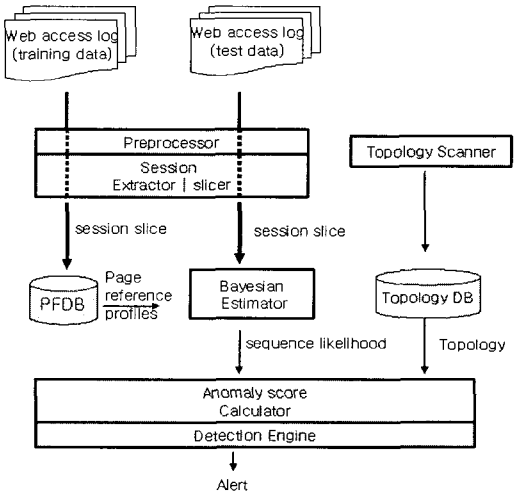
논문에서 제안하는 기법의 주된 가정은 웹 페이지를 방문하는 사용자들이 요청하는 페이지들 간에는 일정한 순서가 있다는 가정을 전제로 한다. 예를 들어 과목 페이지를 방문하는 사용자들은 그 과목, 담당 교수의 홈페이지에서부터 출발하여 과목 소개 페이지, 과제 페이지 등으로 이어질 것이다. 반면 특정 URL이나 CGI등의 자원만을 요청한다거나, 인증 페이지 등 반드시 지나야 할 페이지를 거치지 않는 요청들은 비정상적인 사용자의 행위라고 볼 수 있을 것이다. 웹 사이트의 취약한 CGI를 찾아주는 CGI스캐너 프로그램들이 전자에 해당된다.

본 논문에서는 제안하는 SAD는 웹 서버의 접근 기록인 억세스 로그로부터 사용자의 세션을 분리하고 이 세션에서 접근되어지는 일련의 요청 페이지 순서를 프로파일로 완성하게 된다. 그리고 특정 순서의 페이지 요청이 있을 때 얼마나 프로파일과 유사한지를 이상 점수(anomaly score)로 정량화하여 비정상적인 요청을 탐지하게 된다.

이러한 기법은 기존의 특정 시그너쳐, 예를 들어 요청하는 페이지에 특정 취약 CGI나 잘못된 요청의 형태를 시그너쳐화 하여 이를 탐지하는 IDS와는 달리 훈련 기간동안의 일반적인 사용자들의 접근 시퀀스를 모델링하게 되므로, 특정한 시그너쳐 없이도 이상 현상을 탐지할 수 있는 장점을 갖는다. 특히 각 사이트의 독특한 접근 패턴을 반영할 수 있어 사이트에 최적화된 이상 탐지 시스템을 만들 수 있다.

제안하는 기법을 사용자 명령어 중심의 이상 탐지 기법과 비교한다면, 사용자 명령어의 경우 시작과 끝이 명확하며, 각 명령어 사이에는 특정한 일련의 순서관계가 상대적으로 미약하다. 반면 웹 세션의 경우에는 시작과 끝이 상대적으로 불분명하나, 요청하는 페이지간에 순서관계 및 연관관계가 강하게 나타난다. 예를 들어 특정 페이지를 요청하기 위해서는 반드시 거쳐가야 하는 페이지들이 있을 수 있다.

SAD(Session Anomaly Detection)-Bayes는 [그림 2]와 같이 사이트 토폴로지 분석, 전처리, 이상 점수 산출, 보고 모듈로 구성되어 있다.



[그림 2] SAD-Bayes의 구조도

3.1 사이트 토폴로지 분석

웹 사이트의 전체 연결 구조를 활용하기 위해 사이트 토폴로지 스캐너(Site Topology Scanner)를 구현하였다. 일반적으로 사이트 구조 정보는 WUM(Web Usage Mining)에서 사용자의 세션을 분리하는데 도움을 주는데, 본 논문에서는 페이지간의 연관성을 찾는데 이용된다.

구현된 사이트 토폴로지 정보는 웹 스트럭쳐 관리 등 다양한 목적으로 활용할 수 있도록 XML 문서 형식으로 저장하였다. 사이트 토폴로지를 구성하는데 있어 외부 도메인으로 연결되는 링크 정보나 jpg, exe등과 같이 웹페이지가 아닌 데이터 엔터티는 나타내지 않도록 하였다.

3.2 전처리

웹서버의 접근 로그는 아파치 웹 서버의 combined 형식으로 남겼다. [그림 3]은 웹 로그 형식이다. 웹 접근 로그(Web Access Log)에서는 어떤 IP로 부터 어떠한 페이지가 요청되었는지 알 수 있으며 이 기록은 웹서버가 외부로부터의 요청을 수행한 후에 남기는 로그이다. 따라서, 이 로그를 이용한 침입 탐지는 실제 공격이 발생한 시점과의 어느 정도 시간적인 차이가 있을 수 있다. 즉, 악의적인 내용을 요청, 수행한 후 그 결과에서 공격을 발견하게 되므로, 공격의 시도가 있는 시점과는 차이가 날 수 있다. 이점은 대부분의 어플리케이션 기반 로깅이 갖는 한계이다.

본 연구에서 필요한 세션 정보는 사용자가 요청한 페이지들의 조합으로 구성된다. 페이지에 있는 그림 등의 멀티미디어 데이터 엔터티를 포함할 경우 세션의 정보가 너무 방대하여 이들 데이터는 삭제하였다.

세션은 IP와 요청 시간을 기준으로 정렬한 후 동일 IP에서 마지막 페이지 요청이 있은 후 30분 이내

```
%h : Source IP
%l : Client's Identity
%u : User Name in authentication
%t : Time
%r : Request
%s : Return status code
%b : Size of Transferred data
%{Referer}i : Referrer (previous visiting page)
%{User-agent}i : Agent (OS, Web browser information)
```

〔그림 3〕 웹 서버 로그 포맷

```
3 142 322 75
1 555
7 1 2 3 4 13 14 107
5 1 2 3 8 82
2 13 15
4 96 75553 1 1
```

〔그림 4〕 세션정보

에는 같은 세션으로 처리하였다. 그러나 이 알고리즘을 사용할 경우 Dynamic IP Address를 할당받거나 Proxy서버를 통해 웹 페이지를 요청한 경우는 다른 사용자의 행위가 동일 세션으로 오인 받을 수 있으며, 동일 사용자의 다중 세션들을 구별해 내기 어려운 한계를 가지고 있다.

본 논문에서는 세션을 어떻게 분리해 내는 것이 더 효과적인가 보다는 세션 정보를 활용하여 어떻게 비정상적인 행위들을 발견할 수 있는가에 더 초점을 두고 있다. 세션 정보는 아래와 같이 페이지 ID로 구성된다. 각 페이지 ID는 웹 토폴로지 스캐닝에서 얻어진 ID를 부여했으며, 웹 토폴로지에 없는 페이지의 경우 토폴로지의 마지막 ID이후로 순차적으로 ID를 부여하였다. 각 세션 정보의 처음에는 세션을 구성하는 페이지를 나타낸다.

아래 데이터의 첫번째 세션은 3개의 페이지를 요청했으며, 요청 페이지는 144, 322, 75번 페이지임을 나타낸다.

이렇게 얻어진 세션 정보의 경우 실제 사용자의 트랜잭션과는 조금 다를 수 있다. 즉, 사용자가 웹 브라우저를 이용해서 이전 페이지로 돌아갈 경우, 동일 페이지는 클라이언트가 서버쪽에 페이지를 요청하지 않게 될 수 있다. 따라서 실제 사용자는 이전 페이지를 보고 있지만, 웹 로그에는 남지 않게 된다.

이 때문에 좀더 정확하게 사용자의 페이지 접근 순서를 알기 위해서는 웹 토폴로지를 이용하여 경로를 찾는 작업도 요구되나 현재는 이를 고려하지 않고, 사용자가 방문하는 페이지의 물리적 접근 순서만을 이용한다.

[정의 1]

논리적 접근 순서 : 사용자가 하나의 세션에서 이전 방문 페이지로 돌아와서 다른 페이지를 방문하는 경우가 많이 발생하는데, 이렇게 사용자가 실제로 방문한 페이지의 순서를 말한다.

[정의 2]

물리적 접근 순서 : 사용자가 이전 페이지를 방문하는 경우 웹 서버쪽에 이전 페이지를 요구하지 않고, 사용자 컴퓨터의 캐쉬 정보를 활용하는 경우가 있는데, 이처럼 사용자가 실제로 웹 서버쪽에 요청하는 페이지들의 순서를 말한다.

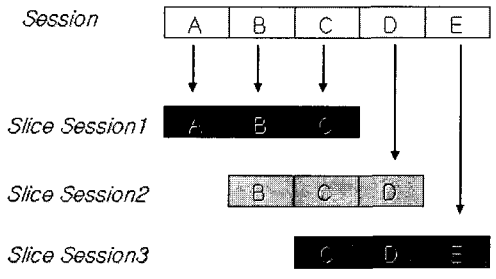
제안하는 기법에서는 사용자 세션에서 연속되어 요청되는 페이지를 일정한 윈도우 사이즈 단위로 분할하여 슬라이스 세션을 생성하며, 이 슬라이스 세션으로 각 페이지 순서의 빈도를 계산하게 된다.

윈도우 사이즈 크기에 따라 아이템의 수가 달라지는데, 본 연구에서는 윈도우 크기 3으로 했을 때 최적의 결과를 보여주었다. 한편 하나의 세션에서의 이상 접근 페이지의 가중치의 높여 정상적인 순서와의 구별이 명확하게 하기 위해서 [그림 5]와 같이 슬라이싱을 수행한다.

3.3 빈도 높은 패턴 생성기

SAD-Bayes엔진에서 발생 확률을 산출하기 위해서는 이전 이벤트 즉, 이전에 본 페이지들에 대한 빈도를 구하여야 한다. 이를 위하여 Pattern Generator 모듈은 슬라이스 세션 각각의 Page Frequency Database에 기록하며 슬라이스의 빈도를 계산한다.

[그림 6]는 SAD-Bayes에서 만든 빈도 데이터 베이스의 예이며, 첫번째 줄을 보면 페이지 1을 보고 다음으로 페이지 2을 본 사용자가 연속해서 페이지 4을 보는 빈도는 23임을 나타낸다. 따라서, 표에 의하면 페이지 1, 2를 보고 페이지 5를 보는 사용자의 빈도가 페이지 4를 보는 사용자보다 많다는 것을 의미하게 된다.



[그림 5] 세션 슬라이스의 생성

Page a	Page b	Page c	Frequency
1	2	4	23
1	2	5	112
.....	.....	.....	.....

[그림 6] 페이지 빈도 DB

3.4 베이지언 추정

본 연구에서는 세션이 훈련 기간동안의 세션들과의 유사성을 평가하기 위해 이전 데이터의 표본 분포를 통해 특정 세션의 요청 확률을 계산하기 Friedman이 제안하는 베이지언 추정 기법[7]을 이용하였다.

$$C(D,L)=\sum_{k=k^0}^{L}\frac{k^0\alpha+N}{k\alpha+N}P(k|D)$$

$$P(X^{N+1}=i|D)=\begin{cases}\frac{\alpha+N_i}{K^0\alpha+N}C(D,L) & \text{if } i\in\Sigma^0\\ \frac{1}{n-k_0}(1-C(D,L)) & \text{if } i\notin\Sigma^0\end{cases}$$

베이지언 추정에서는 기존에 발생한 이벤트 혹은 발생하지 않은 이벤트에 따라서 해당 이벤트에 대한 발생 확률을 이전 분포에 근거하여 추정할 수 있다. 특히 Friedman의 기법은 무한의 알파벳(이벤트의 종류가 한정되지 않은)의 경우에도 잘 활용될 수 있다고 알려져 있다. 이전의 페이지 요청 순서의 빈도를 프로파일링하고 현재의 페이지 요청이 이전의 프로파일과 비교할 때 발생할 확률이 어느 정도 되는지 계산하였다. 이 계산을 통하여 확률적인 낮은 페이지 요청 순서라면 비정상적인 요청으로 판단한다.

위 수식에서 C(D,L)은 이전에 발생한 이벤트가 발생할 확률을 의미한다. N_i는 특정 이벤트가 발생한 빈도를 N는 전체 관찰 표본수를, L은 발생 가능한 이벤트의 종류를, 그리고 k_0는 관찰된 다른 종류의 이벤트 수를 의미한다. 윈도우 크기 3으로 분할하는 SAD-Bayes-3은 P(페이지1, 페이지 2, 페이지 3)를 구한다. 이 확률은 페이지1과 페이지2를 방문한 다음 페이지3을 방문할 조건부 확률을 의미하며, Friedman의 식을 이용하여 추정한다. SAD-Bayes-3에서는 N_i는 페이지1과 페이지2를 방문한 후 페이지3를 방문한 빈도를, N은 페이지 1과 페이지 2를 방문한 빈도를, L은 방문할 수 있는 페이지의 수를, 그리고 k_0는 이전에 훈

련 데이터에서 관찰된 페이지1과 페이지2를 연속해서 방문한 빈도를 의미한다.

3.5 이상 점수

SAD-Bayes에서 구한 슬라이스 세션의 발생 확률에 따라 슬라이스 세션의 이상 점수를 산출한다. 이상 점수는 확률이 낮을수록 높고, 확률이 높을 수록 점수가 낮게 나오도록 다음과 같이 산출한다.

$$AS_{sliceS} = -\log(P(page\ \gamma | page\ \alpha, page\ \beta))$$

$$AS_{session} = \frac{\sum AS_{slice\ session}}{the\ number\ of\ slice\ sessions}$$

사용자 세션의 이상 점수는 슬라이스 세션 각각의 이상점수의 평균을 구하는 방법과 최소값, 최대값을 구하는 방법을 쓸 수 있다. 슬라이스 세션의 이상 점수의 최대값으로 세션의 이상 점수를 산출하는 방식은 슬라이스 세션의 일부라도 확률적으로 낮은 빈도를 보이는, 즉 훈련기간에 관찰할 수 없었던 페이지 요청 순서라면 높은 이상 점수로 나타난다. 따라서, 조금이라도 이상한 페이지 순서를 찾게 해주지만, 훈련기간에 없던 요청 페이지 순서라는 이유로 정상적인 요청도 이상으로 탐지될 수 있다.(false positive). 반면 평균값을 이용하는 이상 점수 산출 방법에서는 일부 슬라이스 세션의 경우 비정상적이더라도 평균값이 임계값(threshold)보다 작을 경우 정상적으로 받아들여 질 수 있다. 특히 비교적 길이가 긴 세션의 경우 이상 행위가 발생하더라도 은닉될 수 있는 오류의 가능성이 있다. (false negative)

본 논문에서는 False Positive 오류의 발생 비율을 줄이고자 사이트 토폴로지 정보를 활용하였다. 사이트 토폴로지 정보에 의거하여 이 사이트에서 요청 가능한 페이지 순서라면 비록 이전 훈련 기간의 프로파일에는 존재하지 않더라도 상대적으로 낮은 이상 점수를 부여하고 토폴로지에 존재하지 않는 페이지들의 요청이라면 높은 이상 점수를 부여하였다.

사용자 세션의 이상 점수가 일정 임계값 보다 높을 경우 이상으로 탐지한다. 앞 절의 이상 점수 산출 방법에서 알 수 있듯이 과거에 발견되지 않았던 페이지 요청 순서는 비교적 높은 이상 점수를 갖게 되고 특히 토폴로지에 존재하지 않는 요청 페이지는 가장 높은 이상 점수를 갖게끔 되어 이상 세션으로 발견된다. 따라서, 이러한 접근은 시그너처가 알려지지 않은 새로운 형태의 웹을 이용한 공격 또는 웹 어플리케이션에 대한 비정상적인 연결을 발견할 수 있게 해준다.

Ⅵ. 실험 및 성능 평가

실험은 약 40여명의 사용자 계정을 갖고있는 대학교의 한 실험실 웹 서버에 대해 웹 스캐닝 공격 도구를 이용하여 공격을 수행하고 탐지 여부를 확인하였다.

훈련 데이터로는 사전에 알고 있는 님다웜의 접근 시도를 제외한 약 4주간의 데이터를 활용하였다. 훈련 데이터는 19643개의 다른 소스 IP주소로 부터의 34154개의 세션을 가지고 있으며, 시험 데이터는 1주간의 로그로 5427개의 IP주소로부터 13415개의 세션으로 구성되었으며 웹 스캐닝 도구 Whisker[20]를 활용한 공격 시도 4367개의 세션과 공격중에 발생했던 님다웜 공격 시도 8개의 세션이 들어있다.

실험에 사용된 Whisker[1]는 취약 CGI 스캐닝과 패스워드 추측을 하는 도구로 네트워크 기반 IDS를 피해갈 수 있는 방법으로 웹 페이지 요청을 변형, 수행한다.

실험의 목표는 제안하는 기법이 기존의 시그너처 IDS와 비교해 볼 때 특정한 패턴 정보없이 새로운 형태의 공격을 탐지할 수 있는지의 확인과 일반적으로 이상 탐지 기법이 갖는 오류 탐지율과 비교해 볼 때 베이지언 추정 기법을 활용한 SAD-Bayes의 오류 탐지율은 어느 정도이고 이 비율을 줄이기 위해 영향을 줄 수 있는 속성은 어떠한 것이 있는지 발견하는데 있다.

이를 위해 일반적인 시그너처 기반 IDS인 SNORT와의 탐지율을 비교한 결과와 SAD-Bayes의 윈도우 크기, 훈련 데이터 셋의 크기, 탐지 필터의 종류, 토폴로지 정보의 활용에 따른 오류 탐지 비율 비교 결과를 소개한다.

4.1 NIDS와의 탐지율 비교

Snort는 공개 network IDS로 가장 많이 사용되고 있

120) Whisker의 IDS 우회를 위한 10가지 Evasive Mode는 URL encoding, Reverse traversal, Self-reference directories, Premare request ending, Parameter hiding, HTTP mis-formatn g, DOS/Win directory syntanx, Null method processing, Case Sensitivity, Session splicing이다.

다. Snort를 이용하여 whisker를 사용했을 때 얼마만큼 정확히 탐지를 하고 있는지를 확인해 보았다. 실험 환경에 사용된 Snort는 version 1.8.7을 사용하였고 whisker는 v1.4를 사용하였다. 실험상의 편의를 위하여 웹과 관련된 6개 그룹의 시그너쳐 모두를 활성화 시켰다.

[그림 7]은 동일한 공격에 대한 Snort IDS와 SAD-Bayes의 탐지율을 보여준다. SNORT는 평균적으로 36%정도의 탐지를 보였는데, Whisker에서 이는 기존에 알려진 형태의 패턴과는 다르게 여러 옵션을 이용하여 변형된 형태의 요청을 수행했고, 따라서 단순히 요청 URL의 텍스트 비교를 수행하는 IDS에서는 이들에 대한 시그너쳐가 없어 탐지율이 낮을 수 밖에 없다. 반면 제안 기법의 경우는 91%의 우수한 탐지 결과를 보였다. SNORT의 경우 시그너쳐 기반의 탐지 시스템인 관계로 충분히 많은 시그너쳐를 가지고 있을 경우 공격 행위에 대한 탐지율이 높아졌을 것이다. 이 비교를 통해 프로파일을 통한 이상 탐지 기법이 기존의 시그너쳐 기반 기법과는 달리 알려진 공격에 대한 패턴정보 없이도 이상 현상을 발견할 수 있어 변형되거나 새롭게 나타나는 공격 양상을 탐지 할 수 있음을 보여주고 있다.

한편 CASE #9와 #11에서는 상대적으로 낮은 탐지율을 보였는데, 이것은 Whisker의 변형 공격 중 웹서버의 운영체제 특성을 고려한 것이 있어 이들의 경우 실험 서버에는 영향을 미치지 않아 접근 로그로 남지 않은 부분이 있기 때문이다. CASE #9와 #11을 제외한 실험 결과는 약 99%이상의 탐지율을 보이는데, 이는 Whisker공격 도구에서 수행하는 대부분의 요청이 훈련 기간 동안에는 관찰되지 못했기 때문이다.

4.2 윈도우 크기의 영향

SAD-Bayes의 윈도우 크기가 탐지율에 미치는 영향에 대해서 실험해 보았다. [그림 8]에 나타나듯이 탐지율에는 변화가 없으나, 윈도우 크기가 3일 때 결과가 좋았다. 윈도우 크기가 작을수록 생성되는 슬라이스 세션수가 많고 따라서 훈련 시간이 비교적 많이 소요되었다.

반면, 윈도우 크기가 클수록 SAD-Bayes에서는 누락되는 세션이 발생한다. 즉 일정 윈도우 크기 미만의 페이지를 요청하는 세션은 빈도 DB에 반영되지 않기 때문이다. 실험에서는 전체 훈련 세션의 47.6%가 페이지 하나 이하를 요청했으며, 74.9%가 3페이지 이하의 요청이었다. 윈도우 크기 4 이상으로 훈련할 경우 전체 훈련 기간의 25%의 요청들만 반영되므로, 탐지오류 비율이 높아진다.

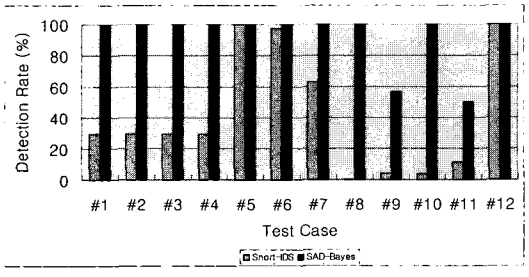
따라서, 윈도우 크기 설정에 있어서는 사이트의 세션 길이를 고려하여야 한다.

4.3 훈련 데이터 크기에 따른 오판율 비교

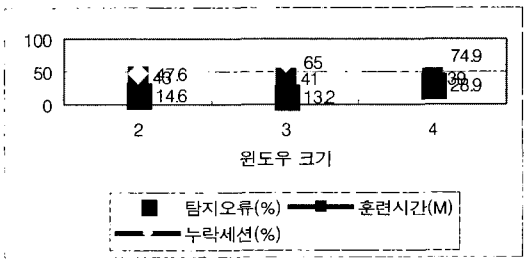
훈련 데이터의 크기가 오판율에 미치는 영향을 실험하였다. [표 1]은 훈련 데이터의 크기를 2주로 했을 때와 4주로 했을 때의 결과이다. 2주로 했을때 SAD-Bayes-3, AVE-11(세션 분석을 위한 윈도우 사이즈를 3개의 페이지로 했으며, 탐지 필터는 평균값이고 임계값은 11)의 false positive는 14.2%인데 4주간의 데이터로 시험할 경우 13.2%로 1%가량 줄었다. 따라서, 훈련기간이 길수록 false positive의 비율이 적어질 수 있음을 보여주고 있다.

4.4 이상 탐지 필터

이상을 결정하는 탐지 필터를 어떻게 정하는 것이 좋은지 실험하였다. [그림 9]를 보면 동일한 탐지율을 보이는데, 그 이유는 수행된 공격 자체가 제안 기법에서는 모두 이상으로 탐지할 만큼 과거에 볼 수 없었던 이상 사용이었기 때문이다. 반면 false positive



〔그림 7〕 SNORT와 SAD-Bayes 탐지 비율 비교



〔그림 8〕 윈도우 크기 비교

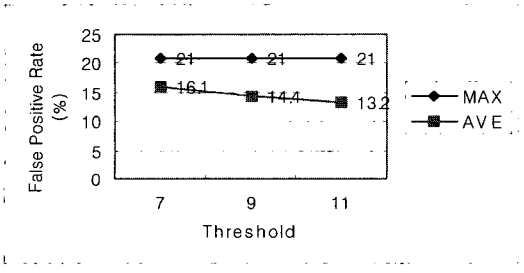
〔표 1〕 훈련 데이터 크기에 따른 오류 비율 비교

	2Weeks	4Weeks
AVE-7	17.2	16.1
AVE-11	14.2	13.2

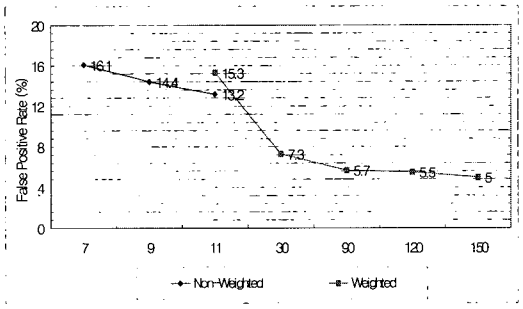
의 비율은 탐지 필터에 따라 차이를 보이는 데, 최대값을 이상 점수로 활용하는 기법보다 평균값을 이상 점수로 활용하는 기법이 더 좋은 결과를 보이고 있다. SAD-Bayes-3, AVE-11의 false positive 비율은 13.2%인데 반해 MAX-11은 21.0%로 최대값 만을 이용하는 필터의 경우 웹 세션의 일부라도 훈련 기간에 탐지되지 않을 경우 높은 이상 점수를 받기 때문에 오류 비율이 상대적으로 높다. 웹사이트의 경우 훈련기간 이후에 변경되는 부분이 존재하므로 일부 슬라이스 세션의 이상 점수만을 극대화할 수 있는 MAX 필터는 좋은 선택이 아님을 확인할 수 있다.

4.5 웹 토폴로지 정보를 이용한 탐지

앞서의 실험 결과에서는 비교적 높은 탐지 오류률을 보이는데, 이것은 훈련 기간이 4주로 짧아 이 기간중에 발견되지 않은 페이지 요청들이 시험 데이터에서 많이 발견되고 있음을 보여준다. 특히 가장 우수한 AVE-11의 false positive 비율도 13%를 넘는다는 것은 제안 기법의 오류률이 비교적 높다고 볼 수 있다. 웹 토폴로지 정보를 반영하여 이상 점수에 가중치를 부여하여 실험을 수행하였다. [그림 10]에서 보여지듯이 SAD-Bayes-3-Weighted(가중치를 부여한 SAD-Bayes-3버전), AVE-150의 오류율이 제일 낮은 5.0%를 기록하였다. 이는 가중치를 부여 하지 않은 AVE-11의 오류률을 50% 가량 개선한 것이다. 임계치 11에서 가중치를 부여한 실험 결과의 탐지 오류율이 비가중치 실험보다 높은 이유는 가중치 부여로



〔그림 9〕 필터 종류에 따른 탐지 오류 비교



〔그림 10〕 웹 토폴로지 가중치 부여

인해 정상적인 일부 세션의 이상 점수가 상승하여 이상으로 탐지되었기 때문이다.

앞서의 실험 결과를 종합하여 보면 단순히 과거 훈련기간의 웹 세션 요청 빈도를 비교하는 것만으로도 현재의 비정상적인 웹 페이지 요청을 발견하는데 도움을 줄 수 있음을 확인하였다. 좀더 정확한 이상 탐지를 위해서는 보다 긴 기간 동안의 충분한 크기의 훈련 데이터가 필요하고, 웹 토폴로지 정보를 활용하여 빈도 뿐만 아니라 요청 가능성을 활용하는 것이 필요함을 확인하였다.

V. 결 론

본 논문에서는 상대적으로 접근 통제가 미약한 웹서비스 공격을 탐지하기 위하여 웹 사용자 프로파일링을 통한 이상 탐지 기법인 SAD-Bayes을 제안하였다.

SAD-Bayes는 기존의 사용자들이 요청하는 페이지 순서와 현재의 요청 페이지 순서와의 상관 관계를 베이지언 추정 기법을 사용하여 통계적으로 추정하고, 추정된 확률이 극히 낮은 사건일 경우 이상으로 탐지한다. 제안 기법은 웹 취약점 분석 도구등을 활용하여 실험해 본 결과 높은 탐지율과 낮은 탐지 오류율을 보였다. 또한 웹 세션의 분석 기본 단위인 윈도우 사이즈의 크기 변화, 이상 탐지를 위한 점수 산출 방법, 탐지 오류율을 낮추기 위한 웹 토폴로지 활용등 실험을 통해 웹 세션 탐지에 반영할 속성들에 대해서도 살펴보았다.

요컨대, 제안 기법은 기존의 시그너쳐 기반의 IDS와는 달리 사용자의 웹 서비스 이용 성향을 분석하여 이와 다른 형태의 사용을 탐지하므로, 새로운 형태의 공격에도 잘 대응되며, 사이트 웹 토폴로지 정보를 활용하여 탐지 오류 비율도 상당히 낮아 웹

서비스를 이용하는 사용자의 비정상적인 패턴을 탐지하는데 효과적일 것으로 기대된다.

참 고 문 헌

[1] Magnus Almgren and Ult Lindqvist. Application-integrated data collection for security monitoring. In Proceeding of Recent Advances in Intrusion Detection (RAID 2001), pp. 22~36, 2001.

[2] Ryan Russell Andrew Mackie, Jensenne Roculan and Mario Van Velzen. Nimda worm analysis. Technical report, Securityfocus.com Incident Analysis Report, September 2001.

[3] CERT Coordination Center. Overview of attack trends. Technical report, CERT CC, 2002.

[4] Thomas Toth Christopher Krugel and Engin Kirda. Service specific anomaly detection for network intrusion detection. In Proceedings of Symposium on Applied Computing, March 2002.

[5] Robert Cooley, Bamshad Mobasher, and Jaideep Srivastava. Data preparation for mining world wide web browsing patterns. Knowledge and Information Systems, 1(1):5~32, 1999.

[6] Dorothy E. Denning. An intrusion-detection model. IEEE Transactions on Software Engineering, 13(2): 222~232, February 1987.

[7] N. Friedman and Y. Singer. Efficient bayesian parameter estimation in large discrete domains,1999.

[8] ISS. Network vs host-based intrusion detection. Whitepaper:http://documents.iss.net/whitepapers/nvh ids.pdf.

[9] Stefano Suin Luca Deri and Gaia Maselli. Design and implementation of an anomaly detection system: an empirical approach. Technical report, http://www.ntop.org, 2001.

[10] Richard Lippmann, et., Evaluating intrusion detection systems: The 1998 DARPA on-line intrusion detection evaluation. In Proceedings of the DARPA Information Survivability Conference and Exposition, Los Alamitos, CA, 2000. IEEE Computer Society Press.

[11] Wenke Lee and Salvatore Stolfo. Data mining approaches for intrusion detection. In Proceedings of the 7th USENIX Security Symposium, San Antonio, TX, 1998.

[12] Herve Debar Magnus Almgren and Marc Dacier. A lightweight tool for detecting web server attacks. In Proceedings of Network and Distributed System Security Symposium, 2000.

[13] Matthew V. Mahoney and Philip K. Chan. Phad: Packet header anomaly detection for indentifying hostile network traffic. Florida Tech. CS-2001-4, 2001.

[14] Thomas H. Ptacek and Timothy N. Newsham. Insertion, evasion, and denial of service : Eluding network intrusion detection. Technical report, Suite 330, 1201 5th Street S.W, Calgary, Alberta, Canada, T2R-0Y6, 1998.

[15] M. Roesch. Snort - lightweight intrusion detection for networks. In Proceedings of USENIX LISA' 99, 1999.

[16] Jeong-Seok Seo. An approaches to web attack categorization. Master' s thesis, KAIST, 2002.

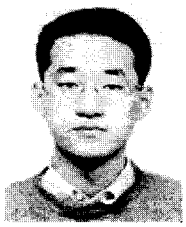
[17] William Yurcik Samuel Patton and David Dos. An achillesi heel in signature-based ids: Squealing false positives in snort. In RAID 2001, 2001.

[18] F. Gilham R. Jagnathan P. Neumann H. Javitz A. Valdes T. Lunt, A. Tamaru and T. Garvey. A real-time intrusion detection expert system (ides). Technical report, Techinal report, Computer Science Laboratory, SRI International, February 1992.

[19] Frank Apap, etc, Detecting Malicious Software by Monitoring Anomalous Windows Registry Accesses, RAID 2002, LNCS 2516, pp.36~53, 2002

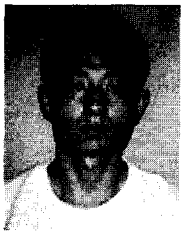
[20] Rain Forest Puppy, "A look at whisker's anti-IDS tactics", http://www.wiretrip.net/rfp

〈著者紹介〉



조 상 현 (Sang-hyun Cho) 학생회원
1997년 2월 : 고려대학교 컴퓨터학과 졸업
1999년 2월 : 한국과학기술원 전산학과 졸업 (석사)
1999년 3월~현재 : 한국과학기술원 전산학과 박사과정
<관심분야> 정보보호, 네트워크 보안, 침입 탐지

김 한 성 (Han-sung Kim) 학생회원
1990년 3월 : 육군사관학교 전산학과 졸업
1995년 9월 : 웨스턴 온타리오대학 전산학과 졸업(석사)
2001년 3월~현재 : 한국과학기술원 전산학과 박사과정
<관심분야> 정보보호, 네트워크 보안, 침입 탐지



이 병 희 (Byung-hee Lee) 학생회원
2002년 2월 : 동국대학교 컴퓨터공학과 졸업
2002년 3월~현재 : 한국과학기술원 전산학과 석사과정
<관심분야> 정보보호, 네트워크 보안, 침입 탐지

차 성 덕 (Sung-deok Cha) 정회원
1983년 : UC Irvine 전산학과 졸업
1986년 : UC Irvine 전산학과 졸업(석사)
1991년 : UC Irvine 전산학과 졸업(박사)
1994년~2001년 : 한국과학기술원 조교수
2001년~현재 : 한국과학기술원 부교수
<관심분야> 정형 기법 및 명세, 정보보호, 침입 탐지