情報保護學會論文誌
第19卷 第4號, 2009. 8

개인정보유출 확신도 도출을 위한 전문가시스템개발*

김 진 형,† 이 알렉산더, 김 형 종,‡ 황 준
서울여자대학교

Rule-base Expert System for Privacy Violation Certainty Estimation*

Jinhyung Kim,† Alexander Lee, Hyungjong Kim,‡ Jun Hwang
Seoul Woman's University

요 약

개인정보 유출을 위한 공격자의 시도는 다양한 보안 시스템에 로그를 남기게 된다. 이러한 로그정보들은 개인정보 유출에 관여했다고 보고된 특정 IP 주소에 대한 확신도를 도출하기위한 요소가 될 수 있다. 본 논문에서는 보편적으로 활용 가능한 보안 시스템들의 로그정보들을 기반으로 확신도를 도출하기 위한 규칙기반 전문가 시스템의 설계 및 구현을 다루고 있다. 일반적으로 개인정보유출과 연관된 다양한 로그정보들은 개인정보 관리자에 의해서 분석되어, 의심 대상이 되는 IP 주소에 대해 정보유출에 관여한 정도를 도출하게 된다. 이러한 개인정보 관리자가 수행하는 분석절차는 전문가의 축적된 지식(Know-how)이라고 할 수 있으며, 이는 규칙 형태로 정의되어 분석절차의 자동화에 활용될 수 있다. 특히, 개인정보유출과 관련된 로그정보의 분석 범위는 다양한 해킹시도를 탐지 해내야하는 침입탐지 및 대응 분야와 비교할 때 상대적으로 넓지 않다. 따라서 도출해내야 하는 규칙의 개수가 상대적으로 많지 않다고 할 수 있다. 본 논문에서는 특히 IDS, Firewall 및 Webserver 의 로그정보들을 개인정보유출의 관점에서 상호 연관성을 도출하였고, 이러한 연관성을 기반으로 규칙을 정의하고 이들을 생성/변경/삭제 할 수 있는 시스템을 개발하였다. 본 연구의 결과에 해당하는 규칙기반 지식베이스 및 전문가 시스템은 개인정보유출에 관여 했다고 여겨지는 특정 IP 주소에 대한 낮은 수준(Low-level)의 검증을 수행하여 확신도를 도출하는데 활용이 가능하다.

ABSTRACT

Logs from various security system can reveal the attack trials for accessing private data without authorization. The logs can be a kind of confidence deriving factors that a certain IP address is involved in the trial. This paper presents a rule-based expert system for derivation of privacy violation confidence using various security systems. Generally, security manager analyzes and synthesizes the log information from various security systems about a certain IP address to find the relevance with privacy violation cases. The security managers' knowledge handling various log information can be transformed into rules for automation of the log analysis and synthesis. Especially, the coverage of log analysis for personal information leakage is not too broad when we compare with the analysis of various intrusion trials. Thus, the number of rules that we should author is relatively small. In this paper, we have derived correlation among logs from IDS, Firewall and Webserver in the view point of privacy protection and implemented a rule-based expert system based on the derived correlation. Consequently, we defined a method for calculating the score which represents the relevance between IP address and privacy violation. The UI(User Interface) expert system has a capability of managing the rule set such as insertion, deletion and update.

Keywords: privacy violation

접수일(2009년 5월 27일), 게재확정일(2009년 7월 29일)
* 본 연구는 서울시 산학연 협력사업(NT070103)의 지원을 받아 수행된 연구임.
† 주저자, jinny@swu.ac.kr
‡ 교신저자, hkim@swu.ac.kr

I. 서 론

인터넷 기술이 지속적으로 발달하면서 오프라인에서 제공되던 서비스들이 대부분 인터넷 환경 내에서도 가능하게 되었고, 이러한 인터넷 환경 내 서비스를 이용하는 사용자는 증가하였다. 이러한 환경의 변화와 함께 인터넷 상에서 개인 정보의 활용 빈도가 증가 하게 되었고, 개인의 정보에 대한 오· 남용의 위험성이 대두 되었다. 현재 사용자의 웹 환경 내 정보 제공 및 사용에 대한 기록을 남기는 로그의 양은 꾸준히 증가하고 있으며, 이러한 로그데이터는 사용자의 행위를 분석할 수 있는 증거로 활용이 가능하다. 따라서 사용자의 공격을 탐지 하거나, 예상 되는 공격에 대한 대응책을 마련하는 과정에서 로그데이터를 활용하여 분석한다.

예를 들어 민감 정보인 개인의 금융 정보를 관리하는 은행 사이트가 해킹 공격을 받아, 고객의 정보가 유출이 되고, 통장 잔고의 변동이 된 사고가 발생 하였다면,[1] 이러한 개인정보가 유출되는 사고에 대하여 로그데이터의 분석을 통해 개인정보유출에 관여했다고 보고된 특정 IP주소에 대한 확신도를 도출 할 수 있다.[1] 이에 본 논문에서는 의심되는 IP주소와 연관된 로그데이터의 분석을 하여 개인정보 관리자가 개인정보를 유출하고자 접근 하는 사용자의 IP주소를 판단하고 위험도에 대해 지수화 할 수 있는 시스템을 설계 및 구현하고 있다. 개인정보 관리자가 공격을 판단하고 지수화 하는 과정은 전문가 시스템을 활용하여 자동화할 수 있도록 설계하며, 관련 로그데이터는 공격에 대한 증거 자료로 판단 할 수 있다. 웹 페이지를 통해 개인정보를 다루는 시스템에 접근 하는 IP주소를 기반으로 각 로그 데이터의 연관성을 분석하고, 이러한 연관성을 기반으로 규칙을 생성/수정/삭제 하는 시스템을 구현한다. 규칙기반 지식베이스 및 전문가 시스템은 개인정보유출에 관여 했다고 여겨지는 특정 IP 주소에 대한 로그데이터를 사용한 낮은 수준(Low-level)의 검증을 수행하여 확신도를 도출하는데 활용이 가능하다.

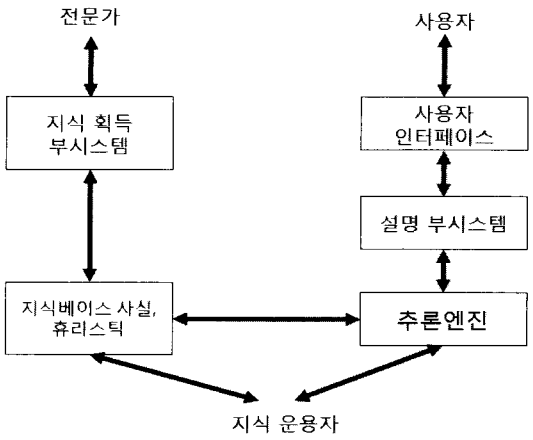
II. 관련 연구

웹을 통해 개인정보의 사용이 증가함에 따라 사용자의 개인정보 유출 및 기업 홈페이지의 변조, 금융사고 등 웹 공격의 사례가 늘어나고 있다. 이에 본 장에서는 개인정보 관리 시스템에 접속하는 사용자에 대해 개인정보를 유출 하고자 하는 의도가 있는지 판단 할 수 있는 시스템의 필요성과, 시스템 구현에 사용되는 이론을 관련 연구로 정리 하였다. 개인정보 유출을 탐지할 수 있는 규칙 도출을 위한 규칙 기반 전문가시스템과 연관성 분석 방법에 대해 알아본다.

2.1 규칙기반 전문가 시스템

사실과 규칙을 이용하는 인공지능 추론 기술을 가장 성공적으로 적용한 예는 의학, 공학, 업무와 같이 특정 전문 분야의 지식을 집약한 전문가 시스템(expert system)이다. 전문가 시스템은 지식을 사용하여 전문가 수준의 문제 해결 능력을 가진 인공지능 프로그램을 지칭한다.

전문가 시스템의 기본 구조는 [그림 1]과 같다.



[그림 1] 전문가 시스템의 기본적인 구조[2,14]

전문가 시스템의 주요부분은 지식베이스와 추론엔진이다. 지식베이스는 주어진 분야에 대한 사실과 규칙으로 이루어진다. 추론 엔진은 사용자가 요구한 지식을 추론하기 위하여 지식베이스를 조작하는 모든 과정을 의미한다. 사용자 인터페이스는 제한적이긴 하지만 사용자가 자연언어로 시스템과 대화할 수 있는 자연어처리 시스템으로 구성되기도 하고, 또한 사용자 인터페이스로서 메뉴를 사용한 GUI가 사용되기도 한다. 설명 부시스템은 시스템이 수행한 추론 구조를 분석하고 이것을 사용자에게 알려주는 기능을 수행한다.

1) 임성현, "인터넷해킹 빈발…당신 돈은 안전한가", 매일경제, 2009.04.17

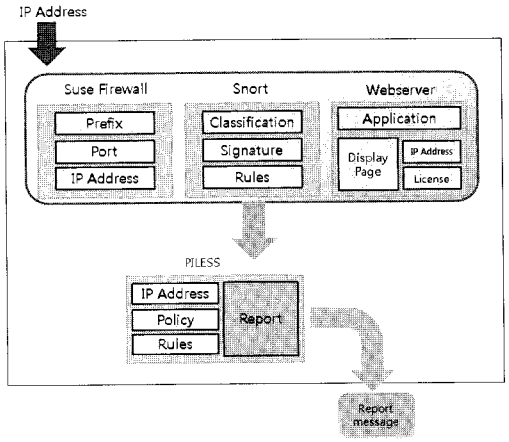
본 논문에서는 이러한 지식베이스를 활용한 전문가 시스템의 처리 과정을 통해 규칙을 기반으로 하는 확신도 도출 시스템을 설계하고 구현한다.

2.2 연관성 분석 방법

본 장에서는 연관성을 통해서 하나의 사건에 포함되어 있는 둘 이상의 항목들의 상호관련성을 발견 하는 연관성 분석에 대한 기존 연구를 살펴본다. 규칙을 기반으로 하여 연관성에 따라 현상을 분석하고자 하는 접근은 사전에 정의된 규칙에 의해 연관성 분석을 수행하는 형태로 stanford대학에서 개발한 CIDF Correlator와 Planing Process Model을 들 수 있다[2]. 이 모델은 시나리오에 기반을 두어 탐지되는 침입정보의 패턴이 일치하는가 여부를 통해 연관성을 가지는 침입탐지 정보들을 찾아낸다. 이러한 규칙 기반의 연관성 분석 방법은 완벽한 시나리오에 대해서는 향상된 성능을 기대할 수 있으나, 기존에 정의되지 않은 공격에 대해서는 정보들을 찾을 수 없다는 단점이 있다. 데이터 마이닝 기법을 통한 연관성 분석방법은 대량의 저장된 데이터를 통해 의미 있는 패턴을 찾아내는 과정으로 데이터 중에서 자주 발견되는 연속적 패턴을 찾아내는 방법을 말한다. 발견된 연속적 패턴은 침입을 탐지하기 위한 지식으로 활용된다. 이 방법은 특정 네트워크 내의 유입된 데이터를 기반으로 만들어진 침입패턴과 비교 하여 공격여부를 판단할 수 있다는 장점이 있으나, 대량의 데이터를 실시간으로 처리하기 어렵다는 단점을 가지고 있다[3]. 확률론적으로 연관성을 분석하는 방법은 발생 가능한 현상의 근원지, 목표, 시간정보 등의 요소들 간의 유사도를 0과 1사이의 값으로 부여하여 연관성을 분석하는 방법이다. 특정한 상황 내에서 완전히 일치하지 않더라도 최소한의 유사도를 발견하여 분석할 수 있는 유연성이 있다는 장점으로 종합적인 유사도 분석하는데 사용이 가능하다[4,5].

III. 로그데이터 분석을 활용한 규칙기반 전문가 시스템 설계

본 논문에서 제안하는 규칙기반 전문가 시스템의 구조는 [그림 2]와 같다. [그림 2]는 웹 페이지에서 개인정보를 다루는 시스템을 나타낸다. 사용자는 특정 IP주소를 가지고 Webserver에서 제공하는 페이지에 접속한다. Firewall과 IDS 등의 네트워크 시스템



[그림 2] PILESS시스템 구조도

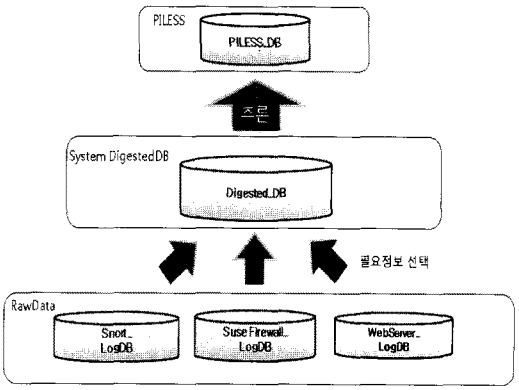
을 거쳐 가는 경우 각 시스템에 로그데이터가 남게 된다. 각 시스템을 거쳐 최종적으로 웹 페이지에 접속한 이상 사용자 IP주소에 대해 위험도를 지수화 하여 점수를 부여하며, IP주소를 기반으로 개인정보 유출 확신도를 계산하게 된다.

본 시스템은 Java 및 Jsp언어로 개발된 웹페이지를 운영한다. Apache Tomcat 6.0서버와 MySQL을 연동하여 운영하며[17], OS는 Windows XP Professional이다. 개인정보 유출에 대한 판단 근거로 활용 가능한 로그데이터를 남기는 네트워크 시스템은 IDS(Intrusion Detection System), Firewall이 있다. IDS 시스템으로는 Snort를 설치하였으며, Firewall은 Suse Firewall을 Linux 플랫폼에 설치하였다. 네트워크 시스템에 남게 되는 로그데이터 중에서, 개인정보 유출에 관여하였다고 판단할 수 있는 근거로 활용이 가능한 데이터 필드를 추출하여 Digested 데이터베이스를 구성한다. 각 시스템은 접근에 대한 로그데이터를 가지고 있으며, 로그데이터를 기반으로 공격자IP를 판단할 수 있는 증거를 수집할 수 있다. 각 시스템에서 남는 로그데이터 중 규칙과 비교 분석하는 과정에 필요한 속성을 선별하여 Digested 데이터베이스를 구성 하고, Digested 데이터베이스의 정보를 기준으로 하여 도출한 규칙과 비교 분석하게 된다. 데이터베이스는 Linux Suse 11.1에 설치하였다.

3.1 Digested 데이터베이스

본 논문에서는 시스템에 접속한 IP주소를 기준으로

각 시스템에 기록된 로그 데이터 중에서 규칙과 비교 분석 수행에 필요한 데이터를 중심으로 Digested 데이터베이스를 구성한다. [그림 3]은 데이터의 계층 구조를 나타낸다. 각 시스템에 저장되는 RawData 중 본 시스템에서 규칙을 도출하는데 필요한 로그들을 IP주소를 기준으로 선별하여 Digested 데이터베이스를 구성한다. 이렇게 구성된 로그데이터를 가지고 위험도를 도출 할 수 있도록 데이터 마이닝 연관성 분석을 수행하여 규칙 기반 데이터를 만든다. 최종적으로 PILESS_DB에 분석된 결과가 저장 되게 되며, 개인정보 유출에 대한 지수화 된 근거로 활용된다.



[그림 3] 데이터 계층 구조[16]

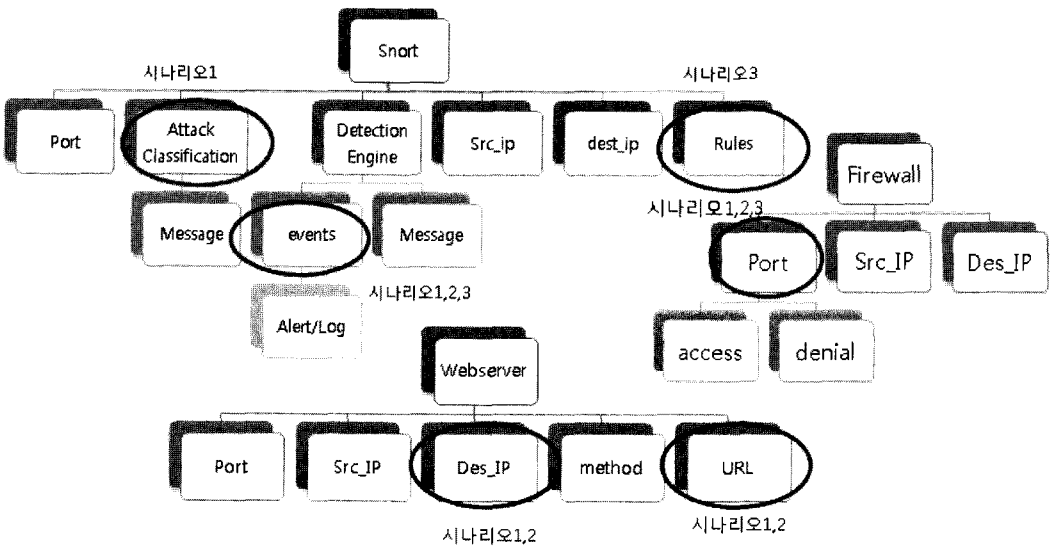
3.1.1 규칙에 따른 Scoring시스템

본 시스템에서 활용하는 로그데이터는 3.1.2장, 3.1.3장, 3.1.4장에서 제시할 각 시스템별 로그 데이터 유형을 따른다. IP주소를 가지고 접속하는 사용자에 대해 각 시스템이 구동되어 입력된 IP주소를 기반으로 로그 데이터를 남기고 이를 분석하여 선정해 놓은 규칙을 기반으로 점수를 부여할 수 있다. 아래 [그림 4]는 점수 부여시 활용하는 로그데이터를 정리한 것이다.

로그데이터 값을 활용하여 공격으로 의심되는 IP주소에 대한 판단을 수행하고자 한다. IP주소를 기반으로 로그데이터를 분류하고, 시간정보와 로그 데이터 값을 분석하여 공격 유형 및 시기에 대한 정보를 확인한다. 본 시스템 구동 후 나오는 결과를 기준으로 공격자에 대한 판단이 가능하며 공격에 대한 증거를 제공할 수 있다. 각 시스템의 로그데이터 값을 비교하여 점수를 부여함으로써 점수의 정도에 따라 공격의 위험성을 판단할 수 있는 기준으로 제시 할 수 있다.

3.1.2 Webserver 로그데이터

사용자가 브라우저를 통해 Webserver에 접속하여 서비스를 요청하게 되면, 남게 되는 로그데이터를 통해 사용자의 행동을 분석할 수 있다. 본 논문에서는



[그림 4] 각 시스템의 로그데이터 구성요소

Webserver에 남는 RawData중에서 발신자의 IP 주소와 호스트 이름, Timestamp와 URL정보만을 가지고 Digested Webserver 데이터베이스를 구성한다. 발신자의 IP주소 값을 통해 접속자의 IP주소를 알 수 있으며, Timestamp정보를 통해 접속한 시간을 알 수 있다. 접속하고자 하는 URL 정보를 통해 사용자의 접속 목적을 알 수 있으며, URL경로를 참고하여 사용자의 접근 순서를 파악할 수 있다. [표 1]에서 URL값을 참고하여 사용자가 개인정보를 제공하는 Info.php에 접속하고자 했었다는 사실을 알 수 있다. URL값이 일련의 순서로 나타나는지 확인하여 권한이 없는 사용자가 전반적으로 웹페이지 접근을 시도한 것을 파악할 수 있다[6,7].

[표 1] Webserver 로그데이터

URL	Time stamp	Host name	src IP
http://211.106.28.32/~user/info.php	2009-02-24 11:17:20.0	linux-tp 0q.site	211.106 .28.84
http://211.106.28.32/phpMyAdmin/navigation.php?server=1&token=e0e0a68b0acb9fc99b7b5498ae47b3e7&db=apachelogs&table=ulog&lang=ko-utf-8	2009-02-24 12:10:02.0	linux-tp 0q.site	211.106 .28.84

3.1.3 Snort 로그데이터

[표 2]에 있는 Snort 로그데이터는 RawData 중에서 공격을 판단하는데 필요한 src_ip와 dest_ip, signature값과 classification값을 선정하여 Digested 데이터베이스를 구성한다. 여러 가지 로그데이터 값 중에서 Classification값을 중점적으로 참조하여 공격여부를 판단할 수 있다. classification값은 총 7개를 갖게 되며, 그 중 misc-attack값과 web-application-attack값을 가지게 되면 이때 접속한 사용자는 공격을 시도했다는 것을 확인할 수 있다. src_ip값과 dest_ip값을 통해 접속자 IP와 Webserver의 IP주소를 알 수 있어 의심 IP를 추출할 수 있다[8].

[표 2] Snort 로그데이터

src_ip	dest_ip	Signature	classification
210.10. 223.10	210.255. 255.250	MISC UPnP malformed advertisement	misc-attack
174.xxx .xx.xx	174.255. 255.250	MISC UPnP malformed advertisement	misc-attack

3.1.4 Suse Firewall 로그데이터

기본적으로 방화벽은 접근 시도에 대하여 Access와 Denial 여부에 따라 페이지 접속을 승인하거나, 거부하는 기능을 수행하는 것이며, 본 논문에서는 승인과 거부의 의미 외에 접속에 대한 횟수를 확인할 수 있다는 의미를 추가 하여 활용하고자 한다[4]. [표 3]에서 Prefix값을 중점적으로 분석하여 승인/거부 여부를 확인하고, 로그데이터 분석 시 Count를 산정하여 접속 시도 횟수를 확인 한다[9].

[표 3] Suse Firewall 로그데이터

src ip	Tcp port	Dest ip	Tcp port	Prefix	Date
210.10. 223.10	80	211.106. xxx.xxx	64372	SFW2-INdmz-DROP-DEFLT	2009-02-24 07:24:33
174.xx x.xx.xx	80	211.106. xxx.xxx	64372	SFW2-INdmz-DROP-DEFLT	2009-02-24 07:24:36

3.2 로그 분석을 통한 규칙 도출 및 위험도 계산

본 논문에서 제안하는 시스템은 공격으로 의심되는 IP를 기준으로 하여 일련의 과정으로 수행이 되며, 각 단계별 모듈에 의한 결과물이 존재한다. 각 결과물은 다음 단계의 입력 값 또는 비교분석 시 필요한 기준값으로 사용이 된다. 본 시스템에서는 Snort, Suse Firewall, Webserver에서 나오는 Log를 각각 수집하여 수집된 데이터와 비교 분석하는 작업을 수행하며, 이렇게 수집된 DB와 Rule을 기반으로 위험도를 계산하게 된다.

3.2.1 로그데이터 분석을 통한 규칙 도출

위험도를 계산 할 수 있는 규칙은 독립적으로 수행되며, 각 시스템별 접속 횟수와 각 로그데이터 값의 비교 등으로 점수를 부여할 수 있는 규칙을 도출 하며, 부여한 점수에 따라 위험도를 판단 할 수 있도록 한다. 각 시스템의 규칙을 도출 하여 규칙의 연관성에 따른 추론 과정을 거쳐 위험도를 계산 할 수 있다. 위험도 계산을 위한 규칙의 구성 요소는 [표 4]와 같다.

〔표 4〕 규칙의 구성요소

구성요소	설명
Rule ID	Rule을 대표하는 고유 번호 이다. 이 값에 의해 Rule이 구분되며, 각 시스템에 따라 일련번호로 부여 한다.
System name	Log 정보를 제공하는 시스템 이름을 의미하며 IDS, Firewall, Webserver 3가지로 구분한다.
Condition	Rule이 참이 되기 위하여 갖는 속성은 조건문 형태로 표현한다. 각 조건에 따라 입력된 IP가 조건에 부합하는지 판단 해 주는 역할을 수행한다.
Action	Rule이 참인 경우 실행된다. 본 논문에서는 참인 경우 부여하게 되는 Score지수를 의미하며 지수에 따라 위험도를 표시할 수 있다.
Hypothesis	규칙의 체인을 구성하는 것을 의미한다. 임의의 조건에 대해 그 결과 값이 참일 때 Rule에 의해 수행되는 구문으로 본 논문에서는 Rule_ID가 그 역할을 수행한다.
Rule description	Rule이 의미하는 것이 무엇인지를 말해준다. Rule의 의미에 따라 각각 다른 시스템에서의 적용 방법이 달라지므로 최종적으로 descript를 참조하여 판단하게 된다.

〔표 5〕에서 제시하는 규칙을 기반으로 하여 만족하는 규칙에 대한 연산 알고리즘을 활용하여 scoring시스템을 정리 할 수 있다. 3개의 시스템에서 나오는 로그데이터를 기반으로 하여 로그데이터 인자의 값을 비교 분석하여 점수를 부여할 수 있고, 부여된 점수는 위험도를 지수화 한 것과 같은 의미를 지닌다. 3.2.2 장에서 위험도 계산에 대해 살펴본다.

3.2.2 위험도 계산

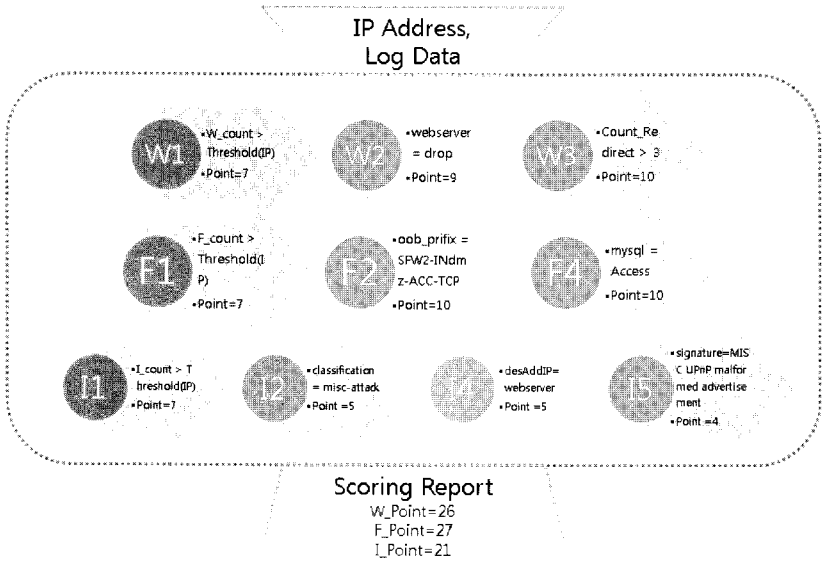
본 시스템은 하나의 의심되는 IP주소가 입력되었을 때 시작 되며, IP주소를 기준으로 하여 관련 규칙을 선정하고, 입력 시점을 기준으로 정해진 기간 동안의 로그데이터를 수집하여 비교할 수 있는 Digested 데이터베이스를 구성하게 된다. 이렇게 구성된 데이터와 앞에서 선정한 규칙의 비교 분석을 통해 시스템은 규칙별로 부여된 점수를 합산 하게 되며, Report를 작성하여 개인정보 관리자에게 위험도 계산 결과를 제공하게 된다. 〔표 5〕의 규칙 중에서 만족하는 규칙의 추론 결과에 따라 점수를 합산 하며, 합산 된 점수를 기준으로 시스템 관리자는 위험 정도를 판단할 수 있다. 각 규칙의 점수의 범위는 2점부터 10점이며 각 규칙이 모두 만족하는 경우 하나의 의심되는 IP주소를 기준으로 본 시스템을 구동한 결과, 최대 77점의 값을 확인 할 수 있다. 각 시스템 별로 IDS는 22점, Firewall는 28점, Webserver는 27점이 최대 값이다. 〔그림 5〕는 그 예를 의미한다. 입력 값으로 IP주소와 각 시스템의 로그데이터를 받아서 규칙의 추론을 통해 결과 페이지를 제시 한다.

본 시스템은 개인정보 침해 여부를 확인하기 위해 규칙에 가중치를 부여 하여 지수화 할 수 있는 알고리즘을 제안하고, 이를 구현 하였다. 본 시스템의 구현은 규칙의 가중치에 따라 점수를 부여 하여 지수화 된 결과를 제공함으로써 개인정보 유출 행위에 참여 가능성에 대한 판단을 할 수 있도록 결과를 제시 한다는

〔표 5〕 Rule Data Set

System name	Rule ID	Rule Condition	score
IDS	I1	count〉Threshold	7
IDS	I2	Classification=Misc-attack	5
IDS	I3	Classification=web-application-attack	4
IDS	I4	desAddIP=Webserver	5
IDS	I5	signature=MISC UPnP malformed advertisement	4
IDS	I6	I1==true && pr〈25	2
IDS	I7	I1==true && 25〈pr〈50	4
IDS	I8	I1==true && 50〈pr〈75	6
IDS	I9	I1==true && 75〈pr〈100	8
Firewall	F1	count 〉 Threshold	7
Firewall	F2	oob_prefix = SFW2-INdmz-ACC-TCP	10
Firewall	F3	oob_prefix =SFW2-INdmz-DROP-DEFLT	9
Firewall	F4	mysql = access	10
Firewall	F5	mysql = drop	9
Firewall	F6	F1==true && pr〈25	2
Firewall	F7	F1==true && 25〈pr〈50	4
Firewall	F8	F1==true && 50〈pr〈75	6
Firewall	F9	F1==true && 75〈pr〈100	8
Webserver	W1	count 〉Threshold	7
Webserver	W2	Webserver=access	10
Webserver	W3	Webserver=drop	9
Webserver	W4	W3==true && Redirect〉3	10

System name: IDS, Firewall, Webserver중 시스템의 이름
Rule ID: 규칙의 고유 ID
Rule condition: 입력된 IP주소의 데이터 값이 규칙과 일치하는지 판단해 주는 조건
Score: 규칙에 부여되는 점수



〔그림 5〕 PILESS 알고리즘

점에 의미가 있다. 또한 관련된 로그데이터를 개인정보에 유출에 참여하였다는 근거로써 제시 할 수 있다.

IV. 시나리오

본 논문에서 제안하는 시스템을 검증하기 위하여 공격 유형에 따른 시나리오를 제안한다. 각 시나리오의 가정은 다음과 같다[10].

1. 공격 대상 IP주소를 알고 있다.
2. 개인정보 페이지에 접근하기 위해 IDS, Firewall, Webserver를 통과 하여야 한다.

4.1 시나리오

4.1.1 Web page로의 전반적인 접근 시도

공격자는 웹 페이지를 통해 시스템에 접근 하고자 접근 시도를 한다. 원래 접근 권한이 없는 사용자로 하나의 페이지 접속 후 링크를 통해 관련된 페이지에 접속하고자 지속적으로 접근을 시도하게 된다. 권한이 없는 페이지는 볼 수 없으며 관련 페이지를 통해 권한이 없는 페이지에 접근하고자 하는 시도를 할 수 있다. 페이지에 접근하여 정보를 얻을 수 있는 단어 혹은 정보를 찾기 위해 지속적으로 접속을 시도한다.

Webserver 로그데이터를 통해 특정한 IP주소에서 지속적으로 접근을 시도하는 것을 확인 할 수 있으며, 일정한 페이지 링크에 반복적으로 접근을 시도하여 페이지 접속 횟수가 증가한다. 기존에 일반적인 사용자의 접속 횟수 기준치를 기준으로 하여 접속 횟수가 비정상적으로 증가할 경우 규칙에 의해 공격으로 판단 될 수 있다.

4.1.2 상위 권한 획득 시도 및 스캐닝

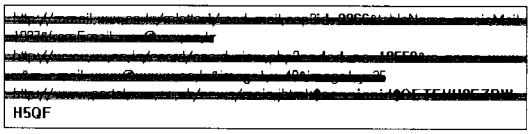
공격자는 권한이 없는데 권한을 상승 하고자 한다. 자신의 정보만 볼 수 있는 일반 사용자임에도 불구하고 타인의 정보를 보거나 활용 하고자 하며, 타인의 정보에 접근하기 위해 접근 권한을 획득 하고자 한다. 권한이 없는 사용자를 따로 구분하여 시스템에서 차단할 수 있는 기능은 없기 때문에 최종적으로 Webserver에 접속하고자 하는 사용자는 접속에 성공할 수 있다. 그러나 웹 페이지에서 권한에 따른 접근 제어를 수행

```
id=firewall time="2008-02-28 19:37:24"
fw=firew4 pri=6 proto=21/TCP
src=192.168.000.250 dst=158.216.666.1
sent=2592 msg="ALLOW SESSION"
id=firewall time="2008-02-28 19:39:24"
fw=firew4 pri=6 proto=23/TCP
src=192.168.000.250 dst=158.216.666.1
sent=2592 msg="ALLOW SESSION"
```

〔그림 6〕 Port scan후 로그데이터

할 경우 권한이 없는 사용자에 대해 접근을 제한하게 된다. 이러한 시도에 대해서도 로그 데이터가 남게 된다. 정상적인 접근이 아니므로 특정 포트 번호가 열려 있는지 시도 한 기록과 그 포트 번호로의 접근 시도 기록이 남게 된다[11].

또한 권한 획득을 위해 Brute-Force Guessing 방법을 사용한다. 이 방법은 네트워크 공격이 아닌 시스템의 비밀번호를 알아내기 위해 패스워드를 입력하여 로그인을 시도하는 공격이다. 무차별 사전 대입공격과 유사하며, 임의의 단어들을 목록화 하여 프로그램을 이용하여 로그인을 시도한다. 유추 가능한 비밀번호의 조합을 무차별 적으로 대입하여 비밀번호를 알아내는 공격이다. 80포트를 이용하여 웹상에서 이루어지는 공격이다.



[그림 7] Session Hijacking 공격시 수집 가능한 로그

권한을 획득하기 위해 Session Hijacking공격 또한 가능하다. 권한이 없는 사용자는 다른 사람의 세션 상태에 접속하고자 한다. 권한 획득을 위해 sniffing 기법을 활용하여 타인의 쿠키 정보를 통해 세션에 접속하고자 한다.

4.1.3 SQL Injection을 통한 DB로의 접근 시도

웹 페이지와 데이터베이스가 연결 되어있는 시스템은 SQL Injection공격에 취약한 특성을 보인다. 일반적인 사용자는 데이터베이스에 직접 접근 하여 SQL구문을 통해 데이터를 접근할 필요가 없으나, 공격을 하고자 하는 사용자는 권한획득이나 데이터베이스에 직접 접근하고자 하는 시도를 통해 권한이 없는 사용자가 데이터를 획득 할 수 있다. 공격자는 웹페이지의 링크가 아닌 데이터베이스에 SQL구문을 활용하여 직접 접근을 시도하였고, 스캐닝을 통해 열려있는 포트 번호도 검사 하였다. 웹 서비스의 취약점에 따라 SQL구문 명령을 통해 데이터 접근에 성공하게 된다. 개인정보를 저장 하고 있는 데이터베이스에 접근 하여 정보를 유출하고자 시도하게 된다. [그림 8]은 SQL Injection공격에 대한 로그 정보이다[12]. [표 6]은 시나리오별 공격 및 결과를 정리한 것이다. 3종류의 공격 유형에 대해 결과 및 판단 내용을 정리 하였다.

```
xxx.xxx.xxx.148 - - [22/Dec/2008:22:45:14 +0900] "GET /blog/category/?page=27'%20AnD%20(sElEcT%20ChAr(94)%2BcAsT(CoUnT(1)%20aS%20VaRcHaR(100))%2bChAr(94)%20fRoM%20[mAsTeR]..[sYs데이터베이스S]))0%20AnD%20'='  HTTP/1.1" 404 - "-" "Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.0)"
```

[그림 8] SQL Injection공격 후 web 로그데이터 변화

[표 6] 시나리오 별 공격 및 결과

	Scenario1	Scenario2	Scenario3
공격 유형	웹 페이지로의 접속시도	상위권한획득 시도 및 스캐닝	SQL Injection공격을 통한 데이터베이스로의 접근시도
공격 행위	• 페이지 URL의 반복적 접속 • Sequential한 URL접속	• 스캐닝 tool을 구동함 • 페이지 정보 획득 후 타인의 권한을 획득하고자 함(Session hijacking공격)	• ID/Password입력 창에 SQL Code입력시도 • 데이터베이스에 접속시도
결 과	• Firewall로그데이터: -access or drop • Webserver로그데이터: -count값이 기준치 이상 -sequential한 URL 접속 횟수 증가	• IDS로그데이터: Rule에 의해 스캐닝에 대한 기록 남김 • Firewall로그데이터: 잘 사용하지 않는 port 번호로의 접속 기록이 남음	• IDS로그데이터: Rule에 의해 기존 접속 패턴에 위배 *snort Rule: Mysql에 우회하여 접근을 시도하고자 함[13,15]. *snort Rule: 스캐닝 후 정보를 유출하려고 시도한 경우 • Firewall로그데이터:access • Webserver로그데이터: -접속 기록이 남음
판 단	Webserver Count와 Firewall drop메시지를 기준으로 점수를 부여함	IDS Rule에서 스캐닝으로 인한 위배되는 규칙과 Firewall의 스캐닝 후 사용된 Port번호에 대한 점수를 부여함	IDS Rule에서 기존 패턴에 위배 되는 규칙과 Firewall에서 데이터베이스로의 접근에 대한 점수를 부여함

4.2 규칙화

규칙을 만들어 규칙조건에 위배될 때 점수를 부여하는 시스템을 구현하여, 위험도를 지수화 하는 시스템에 적용 할 수 있다. 의심되는 IP주소를 입력하여 규칙의 조건에 만족하는 지에 따라 점수를 부여 하며, 부여된 점수의 합을 통해 공격자의 IP주소를 기반으로 위험도를 산정할 수 있다. 높은 점수를 획득한다면, 공격자의 IP주소로 판단할 근거를 확보 할 수 있다.

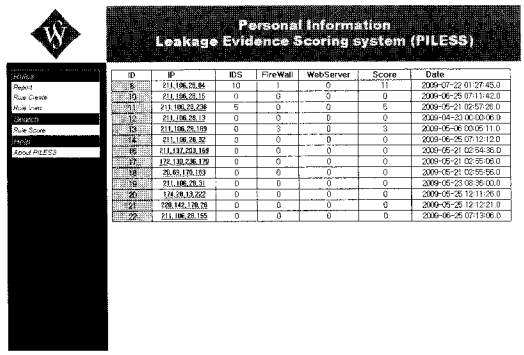
〔표 7〕 시나리오별 적용 가능한 규칙 및 지수

시나리오	규칙	Score
Scenario 1	W1:count 〉 Threshold	7
	F3:oob_prefix=SFW2-INdmz-DROP-DEFLT	9
	F3:oob_prefix=SFW2-INdmz-ACC-TCP	10
Score		26
Scenario 2	I3:Classification=web-application-attack	4
	F10:Portnum=1452	7
Score		11
Scenario 3	I5:Signature=MISC UPnP malformed advertisement	4
	F4:mysql=access	10
Score		14

〔표 7〕은 〔표 6〕에서 제시한 시나리오별 공격 및 결과에서 도출 할 수 있는 규칙을 정리한 것이다. 3가지 유형의 공격 발생시 로그데이터 분석을 통해 다음과 같은 규칙을 근거로 점수를 부여하여 Report를 작성할 수 있다. 이러한 결과는 개인정보 관리자에게 전달되며, 개인정보 관리자는 이 결과를 기준으로 위험도를 판단하게 된다. 각 시스템별 합산한 점수를 기준으로 하여 각각의 시나리오는 26점, 11점, 14점이다.

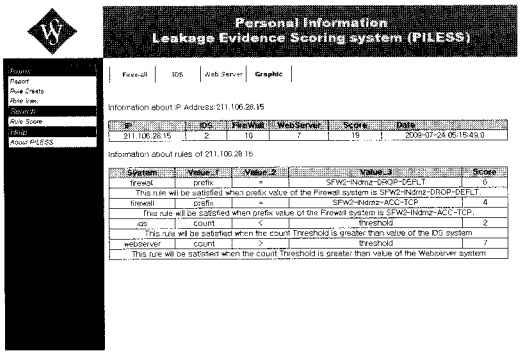
V. 구현 결과

앞에서 설계한 내용을 반영하여 위험도를 계산 할 수 있는 시스템을 구현 하였다. 본 시스템은 개인정보를 다루고 있는 시스템에 접속하는 사용자의 행위에 대해, 위험 행동으로 의심되는 행위를 한 사용자의 IP주소를 기반으로 IDS, Firewall, Webserver등 개인정보를 획득하기 위하여 거쳐 갔던 컴퓨터 시스템에 남게 되는 로그데이터를 활용하여 운영한다. 본 시스템은 개인정보보호 엔진으로부터 의심되는 IP를 전달받아 입력 값으로 하여 각각의 시스템(IDS, Firewall, Webserver등)에서 IP주소에 해당하는 로그데이터를 확인한 후 규칙에 관련된 데이터를 뽑아 규칙과 비교 분석하여 점수를 부여 하는 과정으로 이루어진다. 이를 데이터베이스 규칙들과 비교 분석 후 위험도를 지수화 하여 Report를 보내준다.



〔그림 9〕 의심 IP주소 List페이지

〔그림 9〕는 의심 되는 IP주소의 리스트를 정리 한 것이다. 의심되는 사용자 IP주소를 입력하여 시스템을 수행한 후 수행한 결과에 대해 IP주소를 기준으로 리스트를 정리한 것이다.



〔그림 10〕 의심 IP에 대한 결과 화면

〔그림 10〕은 의심되는 IP를 입력 받아 Report를 수행한 결과 화면이다. 기본적으로 각 시스템 별로 위배되는 규칙에 대한 점수를 합산하여 시스템별 점수를 보여주게 되며, 합계를 산정하여 개인정보 시스템 관리자에게 유의미한 지수를 제공한다. 그 아래 위배되는 규칙을 시스템별로 정리하여 나열해 주게 되면, 단순히 지수로 표현된 위험도 뿐 아니라 증거 자료로 활용 가능한 규칙과 로그데이터를 확인 할 수 있게 된

다. 각 시스템 이름들은 상세 로그 데이터를 확인 할 수 있도록 한다.

VI. 결론 및 향후 연구

본 논문에서는 웹 환경 내에서 개인의 정보를 다루는 페이지 내에서 일어날 수 있는 개인정보 유출에 대한 판단 및 위험도를 지수화 할 수 있는 시스템을 설계 및 구현하였다. 웹 활동 시 사용자의 활동이 기록되는 로그데이터 중 네트워크 환경을 구성 하는 3가지 시스템의 로그데이터의 연관성 분석을 통해 시스템이 공격을 받았을 때 발생하는 로그데이터의 변화를 알아보았다. 개인정보를 유출하거나 획득 하고자 하는 사용자의 IP주소를 기준으로 하여 공격 여부를 판단하고, 위험도 산정을 위해 규칙을 기반으로 하여 지수화 하는 시스템을 구현함으로써 개인정보 유출 발생 시 효과적이고 빠르게 판단하고 증거를 수집할 수 있게 되었다. 개인정보 유출 피해 발생 시 이러한 근거 자료를 통해 개인정보 시스템 관리자가 공격을 판단하고 의심 IP를 분류함으로써 2차 피해를 줄일 수 있다.

본 시스템은 기존에 정의 된 규칙에 대해서만 탐지가 가능한 시스템이다. 정의되지 않은 규칙에 대한 침입 또는 개인정보 유출 시도 발생시 탐지 할 수 없다는 한계를 가지고 있다. 규칙 기반 전문가 시스템을 기본 개념으로 사용하기 때문에 시스템을 운영하는 과정에서 추후 발생 할 수 있는 새로운 공격에 대한 규칙 추가가 용이하다. 각 시스템에서 추출이 가능한 로그데이터의 상관분석을 통해 규칙을 정의 하는 특성에 따라 공격의 패턴에 따라 새로운 공격에 대응할 수 있는 추가 규칙을 생성할 수 있다. 이는 로그데이터 적용 영역의 확대를 통해 가능하다. 이러한 규칙들의 지속적 갱신을 통해 지식베이스의 적용 범위를 확대할 수 있다. 이러한 방법으로 한계를 극복 할 수 있다. 또한 정의 하는 규칙의 의미에 대한 검증은 시스템을 운영하면서 공격의 탐지 비율이 증가를 통해 가능하다. 공격 탐지 비율이 증가 한다는 것은 규칙의 정확성이 높기 때문에 가능한 것이며, 정확도를 높이기 위해 분석하는 로그데이터의 영역을 확대 하여 규칙을 정의 하거나, 규칙을 추가 할 수 있다.

본 시스템은 로그데이터 분석을 통해 위험도를 지수화 한 개인정보 시스템의 효율적인 관리를 목표로 하고 있으며, 향후 본 논문에서 개발한 시스템을 활용하여 특정 페이지에서 다루게 되는 개인정보 데이터를 안전하게 관리 및 활용할 수 있으며 민감한 개인정보를 활용하는 시스템에 대한 신뢰도 향상을 기대할 수 있다.

참 고 문 헌

[1] 허성욱, "개인정보유출소송의 현황과 법적 과제," 저스티스(한국법학원), 통권 110호, pp. 302-331, 2009년 4월.

[2] 최연정, 김일곤, "지식베이스 공유를 위한 지속적 지식기반 관리 시스템의 구현," 정보과학회논문지, 2(4), pp.357-365, 1996년 12월.

[3] 황현숙, 박규석, "연관 규칙 기반의 상품 검색 데이터베이스 최적화 연구," 멀티미디어학회논문지, 7(2), pp. 145-155, 2004년 1월.

[4] 유재학, 강복수, 이한성, 박준상, 김명섭, 박대희, "연관관계규칙을 이용한 트래픽 폭주 공격 탐지의 심층 분석," 한국정보처리학회 추계학술발표대회, pp. 1563-1566, 2008년 11월.

[5] 이수진, 정병천, 김희열, 이윤호, 윤현수, 김도환, 이은영, 박응기, "연관성을 이용한 침입탐지 정보분석 시스템의 설계 및 구현," 한국정보과학회논문지, 31(5), pp. 438-449, 2004년 10월.

[6] 이형우, "룰 기반 웹 IDS 시스템을 위한 효율적인 웹로그 전처리 기법 설계 및 구현," 한국인터넷정보학회지, 9(5), pp. 23-34, 2008년 10월.

[7] 정정기, 박대우, "로그 히스토리 분석을 사용한 웹포렌식 알고리즘 연구," 한국컴퓨터정보학회 동계학술대회발표집, pp. 245-254, 2006년 12월.

[8] 손형서, 김현성, 부기동, "암호화 기법을 적용한 침입탐지 시스템의 룰 보호 기법," 한국정보보호학회논문지, 14(6), pp. 3-13, 2004년 12월.

[9] 천준호, 신동규, 장근원, 전문석, "DDoS 공격에 대한 방화벽 로그 기록 취약점 분석," 한국정보보호학회논문지, 17(6), pp. 143-148, 2007년 12월.

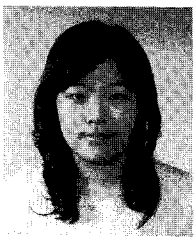
[10] 최향창, 노봉남, 이형효, "NS를 이용한 시나리오기반 공격 시뮬레이터 설계 및 구현," 한국인터넷정보학회지, 7(5), pp. 59-69, 2005년 10월.

[11] 류대희, 이세열, 김혁진, 송영덕, "퍼지인식도와 세션패턴 기반의 비정상 탐지 메커니즘," 한국 컴퓨터정보학회논문지, 10(6), pp. 9-16, 2005년 12월.

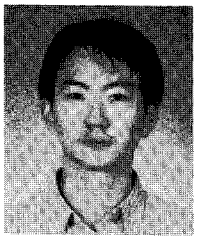
[12] M. Ahmed, D. Quercia, and S. Hailes, "A statical Matching approach to detect privacy violation for Trust-Based Collaborations," In First International

Workshop on Trust, Security and Privacy for Ubiquitous Computing(Affiliated with WOWMOM 2005), IEEE, pp. 598-602, June 2005.
[13] N.J. Nilsson, Artificial Intelligence: A New Synthesis, Morgan Kaufmann Pubishers Inc., Apr. 1998.
[14] Negnevitsky, Artificial Intelligence, 2th Ed., Addison Wesley, Nov. 2004.
[15] http://www.snort.org/vrt/
[16] R. Sadoddin and A.A. Ghorbani, "An incremental frequent structure mining framework for real-time alert correlation," computers & security, vol. 28, no. 3, pp. 153-173, Nov. 2009.
[17] http://tomcat.apache.org/tomcat-6.0-doc/index.html

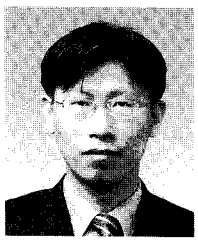
〈著者紹介〉



김 진 형 (Jin-hyung Kim) 학생회원
2006년 2월: 서울여자대학교 정보보호공학과 졸업
2008년 2월: 서울여자대학교 대학원 컴퓨터학과 석사
2008년 3월 ~ 현재: 서울여자대학교 컴퓨터학과 박사과정
〈관심분야〉 정보보호, 개인정보보호, 디지털 포렌식



이 알렉산더 (Alexander Lee) 학생회원
2002년 3월: Al-Farabi Kazakh National University, Kazah입학
2007년 2월: Al-Farabi Kazakh National University, Kazah졸업
2008년 3월 ~ 현재: 서울여자대학교 컴퓨터학과 석사과정
〈관심분야〉 개인정보보호, 디지털포렌식



김 형 종 (Hyung-Jong Kim) 종신회원
1996년 성균관대학교 정보 공학과(공학사)
1998년 성균관대학교 정보 공학과(공학석사)
2001년 성균관대학교 전기전자 및 컴퓨터공학과(공학박사)
2001년 ~ 2007년 한국정보보호진흥원 수석연구원
2004년 ~ 2006년 Carnegie Mellon University, USA Visiting Researcher
2007년 ~ 현재 서울여자대학교 컴퓨터학부 조교수
〈관심분야〉 취약점 분석 및 모델링, 이산사건 시뮬레이션 방법론, 침입감내기술

황 준 (Jun Hwang) 정회원
1985년: 중앙대학교 컴퓨터공학과 졸업(학사)
1987년: 중앙대학교 대학원 컴퓨터공학과 졸업(석사)
1991년: 중앙대학교 대학원 컴퓨터공학과 졸업(박사)
1992년 ~ 현재 서울여자대학교 정보미디어대학 미디어학부 교수
〈관심분야〉 IPTV, Convergence Computing, Digital Broadcasting, 개인정보보호