情報保護學會論文誌
第12券 第4號, 2002. 8

이종의 침입탐지센서 관련성을 이용한 통합탐지의 민감도 향상 방법*

김 용 민**, 김 민 수***, 김 홍 근****, 노 봉 남*****

An Aggregate Detection Method for Improved Sensitivity using Correlation of Heterogeneous Intrusion Detection Sensors

Yong-Min Kim**, Min-Soo Kim***, Hong-Gun Kim****, Bong-Nam Noh*****

요 약

침입행위에 대한 비정상행위 탐지방법은 탐지에 대한 오판율이 높게 나타난다. 즉, 실제 침입이 아닌데 침입으로 판정하는 과탐지와 실제 침입인데 탐지하지 못하는 미탐지에 대한 오판율의 경우이다. 침입탐지의 민감도를 향상시키기 위하여 오용행위 및 비정상행위 탐지센서들 사이의 관련성을 이용하여 오판율을 감소하는 통합탐지의 방법을 연구하였다. 정상행위 및 비정상행위에 대해 하나의 탐지센서로부터의 결과가 다른 탐지센서에 의한 결과와 어떠한 관련성을 갖고 있는지의 반영비율을 오프라인에서 생성하고, 이를 실시간에 탐지된 결과에 적용하여 오판율을 감소하도록 하였다.

ABSTRACT

In general, the intrusion detection method of anomalous behaviors has high false alarm rate which contains false-positive and false-negative. To increase the sensitivity of intrusion detection, we propose a method of aggregate detection to reduce false alarm rate by using correlation between misuse activity detection sensors and anomalous ones. For each normal behavior and anomalous one, we produce the reflection rate between the result from one sensor and another in off-line. Then, we apply this rate to the result of real-time detection to reduce false alarm rate.

Keyword : *침입탐지시스템(IDS), 탐지센서(detection sensor), 통합탐지(aggregate detection), 관련성(correlation)*

Ⅰ. 서 론

침입탐지시스템은 컴퓨터 시스템의 비인가된 사용자에 의한 비정상적인 사용(anomaly use)과 인가된 사용자에 의한 잘못된 이용(misuse)을 탐지 및 예방하고자 하는 것이다. 즉, 사용자 및 외부 침입자가 컴퓨터시스템과 네트워크의 자원을 합법적인 권한 없이 불법적으로 사용하기 위한 시도 또는 내부 사용자가 자신의 권한을 오용하여 권한이외의 자원을 사용하기 위한 시도를 사전에 탐지하여 그 피해를 예방하는 시스템이다.[1,2,3]

이러한 침입탐지시스템은 크게 호스트기반 및 네트

* 본 연구는 한국정보보호진흥원 위탁연구과제 지원으로 수행하였습니다.
** 전남대학교 대학원 전산통계학과 박사과정(ymkim@chonnam.ac.kr)
*** 전남대학교 정보보호협동과정 객원교수(phoenix@athena.chonnam.ac.kr)
**** 한국정보보호진흥원(hgkim@center.kisa.or.kr)
***** 전남대학교 컴퓨터정보학부 교수(bongnam@chonnam.ac.kr)

워크기반의 침입탐지시스템, 그리고 오용행위 및 비정상행위의 침입탐지시스템으로 분류할 수 있다. 현재 상업용 또는 공개용으로 개발되어 널리 보급된 것은 네트워크 트래픽 데이터를 분석하여 판정하는 네트워크기반 침입탐지시스템과 응용 프로그램의 취약점을 이용하여 침입하는 공격 형태에 대한 특징을 규칙이나 상태로 표현하여 탐지하는 오용행위 침입탐지시스템이다. 최근에는 시스템 로그를 분석하여 호스트에서 감시와 분석을 통하여 판단하는 호스트기반 침입탐지시스템과 오용행위 침입탐지에서 새로운 공격이 발생할 때마다 새로운 규칙을 만들어 보급해야 하는 단점을 극복하기 위해서 사용자나 세션 등의 정상행위를 바탕으로 하는 비정상행위 탐지방법을 연구하고 있다.[4,5,6]

현재의 침입탐지시스템은 탐지 결과를 실시간의 제한된 시간 및 임의의 공격에 관련된 다수의 많은 보고를 함으로써 관리자가 대응하는데 어려움이 있으며, 과탐지(false positives) 및 미탐지(false negatives)의 잘못된 탐지 결과를 보고한다. 또한, 침입탐지시스템은 단일 시스템 또는 단일 환경을 고려하여 설계한 침입탐지 기능만을 제공하므로 보안의 대상을 제한할 뿐 아니라 대규모 및 다양한 환경의 시스템으로 확장에 대한 유연성에 한계가 있으므로 다양한 형태의 침입 탐지가 불가능한 경우가 있다. 즉, 호스트 혹은 네트워크 기반의 침입탐지시스템 단독으로 탐지할 수 없는 침입의 형태가 있다. 따라서, 대규모 네트워크 환경에서 다양한 형태의 침입을 탐지하기 위해서는 호스트 혹은 네트워크 기반에서의 감시 및 탐지, 침입 여부에 대한 판정과 함께, 각 시스템이 제공하는 침입탐지 정보의 통합 분석을 통하여 침입탐지의 능력을 향상시키는 방법이 필요하다.

일반적으로, 침입행위에 대한 비정상행위 탐지방법은 오판율이 높게 나타난다. 즉, 실제 침입이 아닌데 침입으로 판정하는 과탐지와 실제 침입인데 탐지하지 못하는 미탐지에 대한 오판율의 경우이다. 침입탐지시스템의 오판율을 줄이는 방법을 위해 호스트 및 네트워크의 침입탐지 영역에서 오용행위 및 비정상행위의 침입행위에 대한 각 탐지센서의 판정결과 값을 관련성(correlation)에 의해 반영하여 판정의 신뢰도를 향상할 수 있는 통합판정 방법에 대한 연구가 필요하다.

이 논문에서는 침입탐지의 민감도를 향상시키기 위하여 계층적 침입탐지 구조에서 오용행위 및 비정상행위 탐지센서들 사이의 관련성을 이용하여 탐지능력을 향상시키는 통합판정의 방법을 제안한다.

II. 계층적 통합탐지 시스템

2.1 관련연구

현재의 탐지영역 및 탐지방법에 따른 침입탐지시스템은 대규모 네트워크 환경의 적용에 어려움을 가지고 있어 탐지영역에 따른 탐지센서의 계층적 구조화 및 탐지결과의 통합 방법이 필요하다. 대규모 네트워크 환경에서 침입탐지의 광범위한 분석을 가능하게 하는 계층적 구조의 침입탐지시스템에 대한 연구는 DARPA/ITO에 의해 주도적으로 이루어져 왔다. DARPA의 주요 연구에서 JAM[7]은 침입패턴에 대한 빠른 분배를 위해 분산환경에서 메타학습(mata-learning)의 이식성과 확장성을 제공하는 에이전트 기반의 데이터마이닝시스템에 대한 연구, JiNao[8]는 IDS 사이의 침입탐지와 네트워크 하부구조에 대한 상호보완적 접근 및 안전한 연결 프로토콜에 대한 연구, AAFID와 NetSTAT[9]에서는 각각 경량의 다중 프로세스가 서로 협력하여 침입탐지를 수행 및 호스트와 네트워크 기반의 IDS를 자동 배치하며 공격시나리오를 자동 분석하는 연구, 그리고 GrIDS[10]는 호스트들의 행위와 호스트들 사이의 트래픽에 대한 행위 관계 그래프를 생성하는 연구를 수행하였다.

그리고, 다수의 탐지센서의 탐지결과를 통합하는 방법의 요구는 침입에 대한 탐지결과 메시지가 실시간의 제한된 시간에 많은 결과(flooding message)를 보임으로써 관리자가 대응하는데 어려움이 있기 때문이다. 임의의 공격에 대해 각 탐지센서에 의해 다수의 관련된 결과 메시지가 발생하고, 이러한 것을 논리적으로 그룹화하여 관리자에게 제공하지 못함으로써 과탐지 및 미탐지의 잘못된 탐지결과(false alarm)를 발생하고 있다. 이러한 침입탐지시스템들의 탐지결과를 통합하고 판정하기 위한 대표적인 연구로서 IBM/Tivoli의 TEC[11] 및 DARPA/ITO 계획의 일부인 EMERALD가 있다.[12,13,14]

TEC(Tivoli Enterprise Console)은 호스트 및 네트워크 기반의 침입탐지시스템의 결과를 통합(aggregation)하고 각 탐지결과 메시지의 관련성을 분석하여 관리자에게 압축된 결과를 보고하는 체계를 갖는 시스템이다. TEC은 침입 데이터의 수집, 형식화 및 분석, 진단, 그리고 탐지결과 메시지를

보이는 기능을 하는 프로브(probe)가 있으며, 통합 및 관련성 분석 콘솔(Aggregation & Correlation Console : ACC)은 프로브 결과의 수집, 분석, 압축된 탐지결과 메시지를 관리자에게 제시하는 기능을 한다. TEC은 통합탐지를 통한 탐지메시지의 보고를 줄이고 오판율을 감소하기 위해 탐지메시지들의 관련성을 호스트 주소, 목적지주소, 탐지메시지를 이용하여 7가지의 상태로 그룹화한다. 그룹화하는 단계는 탐지메시지에 대한 기본 정보를 생성하는 프로브 계층, 다중의 목적지 주소를 갖는 메시지를 분석하는 목적지 계층, 메시지 발생의 호스트 주소가 실제 주소인지 분석하는 근원지 계층, 목적지의 서비스 이용을 분석하는 상세 목적지 계층으로 이루어진다. 각 계층 그리고 계층 사이에는 중복된 관계인지 또는 이전의 사건과 연결되는 관계인지를 보이는 방법으로 분석하고 있으며, 사용자 세션에 대한 행위 프로파일 분석이 필요하다.

EMERALD는 네트워크 서비스의 실시간 보호를 제공하기 위해 통계학적인 프로파일과 서명 분석을 결합한 능률적인 침입탐지 기능을 제공하며, 광범위한 침입탐지 기능과 네트워크에서 일어날 수 있는 중요한 공격에 대한 대응 능력을 제공하기 위해 분산되어 있는 모니터의 분석을 종합하는 프레임워크 개념을 도입하여 개발하고 있다. EMERALD는 각 다중 탐지시스템들의 탐지 메시지의 교환을 위해 탐지메시지 템플리트(alert template)를 이용하고 있으며, 이 템플리트에는 탐지시스템의 유형, 위치, 공격의 목표, 그리고 정상 및 비정상의 행위 표시 필드를 포함하고 있다. 또한, 메타 탐지메시지(meta alert)에서는 공격에 대하여 다른 길이의 패턴 비교, 비교된 특징의 수량 및 질, 비교된 횟수, 탐지메시지의 특성을 비교하여 탐지메시지들 간의 유사성을 정의하기 위한 방법을 연구 진행중이다. EMERALD의 특징은 메타메시지의 통합에 의해 기존의 탐지메시지와 새로운 탐지메시지 사이의 유사가능성(expectation of similarity)을 찾아 관련 정도를 표현한다. 각 탐지메시지의 유사가능성은 비교되는 특징들의 정도에 의해 나타나게 되며, 두 개의 탐지메시지 사이의 유사가능성은 베이지안 통계에 의해 표현한다. EMERALD는 탐지메시지의 관련성을 사건의 클래스에 따른 유사가능성으로 표현하였으며, 이를 위한 메타 탐지메시지의 구성 및 사건의 분류는 호스트의 감사로그, TCP 연결 등의 저수준 사건을 통합하고, 다음 단계에서는 동일한 행위에 대해 하나의 탐지센서가 다른 탐지센서의 상태를 인지하여 다수의 탐지메시지를 하나의 메시지로 통합하며, 마지막으로 각 센서의 탐지 메시지를 통합하는 단계의 절차를 수행한다.

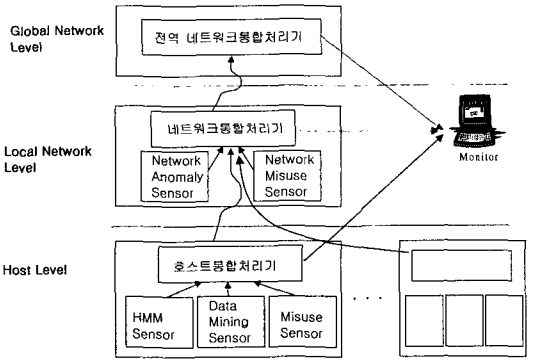
이 논문에서는 침입탐지의 과탐지 및 미탐지에 대한 오판율의 감소를 위하여 각 탐지센서의 결과를 통합하는 방법으로 각 탐지센서의 탐지결과를 비정상행위 정도의 값으로 표현하였으며, 동일 행위에 대해 각 탐지센서가 나타내는 값으로서 탐지센서 사이의 관련성을 표현하였다. 이러한 방법은 각 탐지센서의 모든 탐지결과를 축약 및 간결하게 보고하며, 정상행위 및 비정상행위에 대하여 비정상행위 탐지센서가 오프라인에서 학습한 각 탐지센서의 판정 결과를 실시간에 반영하여 오판율을 감소하고자 한다. 그리고 호스트 수준에서는 사용자의 식별자에 의한 비교 및 통합을 하고, 네트워크에서는 호스트 주소 및 서비스에 기반한 방법을 적용하여 탐지범위에 따라 탐지정보를 선택할 수 있도록 한다. 이러한 방법은 현재의 단일 환경에서 동작하는 침입탐지시스템의 제한 범위를 대규모의 네트워크 환경에서도 적용할 수 있도록 한다.

2.2 통합탐지시스템의 기반환경

통합탐지시스템의 기반환경은 침입행위에 대한 탐지를 목표로 구성된 시스템이며, 침입행위에 따라 오용행위 탐지센서와 비정상행위 탐지센서로 구분한다. 비정상행위 탐지센서는 탐지영역에 따라, 순서기반에 따른 비정상행위 탐지센서, 연관 규칙 및 분포기반에 의한 비정상행위 탐지센서로 구성된다.

이 논문에서는 침입에 대한 통합탐지 및 탐지율 향상의 목표를 위해 이미 개발된 호스트 기반의 탐지센서를 이용하여 연구하였다.[15] 각 탐지센서들은 비정상행위 침입에 대해 은닉마르코프 모델(Hidden Markov Model : HMM)을 이용한 순서기반의 비정상행위 탐지센서(시스템호출, 파일접근, 시스템호출-파일접근), 데이터마이닝을 이용한 분포기반의 비정상행위 탐지센서(연관규칙 탐지센서, 클러스터링 탐지센서), 그리고 Petri Net을 이용한 오용행위 탐지센서로 구성되어 있다. [그림 1]은 계층적 통합 침입탐지의 개념을 보인다.

호스트 통합탐지 처리기는 호스트에서의 오용행위 및 비정상행위 탐지센서들의 결과를 통합 조율하며, 네트워크 통합탐지 처리기는 각 지역 호스트의 통합탐지 처리기와 네트워크 오용탐지 센서, 네트워크



〔그림 1〕 계층적 통합 침입탐지 구성도

비정상행위 탐지센서의 결과를 통합 조율하는 위치에 있다. 따라서, 각 수준에서의 이종의 탐지센서 결과를 통합 및 판정하는 알고리즘이 필요하다.

호스트의 각 탐지센서는 서로 다른 영역을 탐지처리하며 동일한 영역도 탐지할 수 있다. 오용행위 탐지센서에서는 여러 종류의 버퍼 오버플로우 공격을 탐지하고 경쟁상태를 유발하여 임의파일을 생성하거나 파일의 허가권을 변경하는 공격을 탐지한다. 또한, 잘못된 파일 링크, 잘못된 파일 접근, 잘못된 로그인, 많은 프로세스나 파일 생성을 탐지한다. 그리고, 여러 공격을 조합하거나 여러 세션을 통해 이루어지는 다중 공격을 탐지할 수 있다. 비정상행위 탐지센서에서는 사용자별로 프로파일링을 거쳐 사용자에 대한 비정상행위를 탐지하게 된다. 이러한 비정상행위 탐지센서의 주요 탐지 대상은 사용자 ID를 도용한 침입, 비정상적인 명령어 사용, 그리고 불법적인 파일/디렉토리 접근 등이다.

Ⅲ. 이종의 탐지센서 통합탐지 방법

3.1 통합탐지 성능 고려사항

현재의 침입탐지시스템의 평가에 대한 관심은 실제 침입이 아닌데 침입으로 판정하는 과탐지의 경우와 실제 침입인데 탐지하지 못하는 미탐지의 경우를 어떻게 처리하는 가에 있다. 이 논문에서는 과탐지율 및 미탐지율, 침입탐지의 실시간성을 통합탐지의 성능 향상을 위한 요소로 고려한다.

과탐지를 측정하는 방법은 정상적인 수행 내용을 자동으로 발생하여 주는 시스템 및 네트워크 정상행위 발생기에서 정상행위를 수행하고, 이와 동시에 주어진 공격을 연속으로 수행하는 방법으로 배경잡음 시험을 이용하여 수행한다. 과탐지는 통합탐지시스템에서 탐지 및 판정한 목록과 실제 공격한 목록을 비교하여 공격에 해당하지 않는 내용을 탐지한 경우이다. 만일, 임의의 호스트에 총 공격한 횟수가 n이고 탐지메시지가 나타난 것이 m이라고 할 때, 탐지메시지 중에서 정확히 탐지한 횟수가 k라고 하자. 그러면, 과탐지율(false positive ratio : FP)은 (m-k)/m이다.

통합탐지시스템의 미탐지 평가는 호스트나 네트워크에 대한 공격 패턴을 보유하고 이를 수행하며 과탐지의 측정 방법과 동일하게 수행한다. 즉, 어느 호스트에 대한 공격 횟수가 n이었을 때 탐지한 횟수가 k라면 미탐지율(false negative ratio : FN)은 (n-k)/n이다.

통합탐지의 실시간은 침입 중의(hard) 실시간과 침입 후의(soft) 실시간으로 구분할 수 있다. 침입탐지에서 침입중의 실시간은 침입이 완료되기 전에 탐지하여 대응하는 것이고, 침입후의 실시간은 침입이 완료된 후에 다른 행위가 일어나기 전에 탐지하여 대응하는 것이다. 일반적으로 시스템이나 네트워크에 대한 침입 행위는 침입탐지시스템과 경쟁을 하는 상황이므로 침입중의 실시간을 적용한다. 따라서, 공격 시작시간과 공격 완료시간 사이에 탐지 메시지를 비교하여 근사한 시간(수초~일분) 이내에 탐지하는 방법을 사용한다.

3.2 탐지센서 관련성

3.2.1 관련성 접근방법

탐지센서의 관련성은 침입탐지의 가능성을 높이기 위하여 저수준의 사건을 통합하여 탐지의 민감도를 향상하고 잘못된 판단의 억제를 위해 각 탐지센서의 결과를 조정하고 통합하여 판정하는 방법이다. 탐지센서의 결과를 통합하는 방법으로서 기계학습(machine learning)에 의한 방법의 예는 결정트리의 방법이다. 각 탐지센서 A, B, C가 있을 때, 임의의 비정상행위 X에 대하여 탐지센서 A가 탐지하거나 만일 탐지하지 못한다면 탐지센서 B와 C의 AND 연산에 의해 탐지하는 방법으로 수행한다. 탐지센서는 '비정상행위' 또는 '정상행위'의 판단 결과를 갖기 때문에 만일 B 센서는 비정상행위 X에 대하여 탐지하고 C는 탐지하지 못한다고 하였을 때, 탐지센서 C가 정상행위로 판단하였거나 또는 판단 기준치에 근사한 값으로 인해 공격행위 임에도 불구하고 탐지하지 못

하였는지에 대한 의문을 갖게 한다. 따라서 탐지센서 B와 C에 의한 연산의 결과 탐지수준의 결정 기준에 대한 의문이 있게 된다. 결정트리에 의한 방법은 각 탐지센서 사이에 판단을 위한 신뢰 인자(confidence factor)를 공유하는 경우는 유효하지만 각 탐지센서의 특성에 따라 탐지하는 영역의 차이 및 탐지하는 방법의 차이로 인해 신뢰인자를 공유하기 어려워 부적절한 결과를 가질 수 있다.[14] 따라서 이 논문에서는 확률적 접근 방법에 의해 각 탐지센서의 관련성을 생성한다.

확률적 접근 방법에 따른 하나의 예는 통계적 방법이다. 하나의 방법으로서 동일한 비정상행위에 대해 각 탐지센서가 탐지한 결과를 더한 후 평균의 값에 의해 판정하는 방법을 생각할 수 있다. 이러한 경우 각 탐지센서의 탐지 특성에 따른 결과를 정확하게 반영하지 못하는 경우를 보일 수 있으며, 판정의 기준을 낮추었을 때는 과탐지가 많이 발생하고, 판정의 기준을 높였을 때는 미탐지가 많이 발생하는 결과를 보인다. 즉, 각 탐지센서 특성의 결과를 반영하지 못하여 판정기준의 모호함을 갖게 된다. 따라서 탐지센서의 특성을 반영하기 위하여 조건부 확률적 접근 방법에 의해 관련성을 분석하고자 한다.

3.2.2 조건부 확률적 접근 방법

각 탐지센서의 탐지 영역 및 탐지방법에 따른 결과에 대한 관련성을 나타내는 방법은 동일한 침입 행위들에 대하여 각 탐지센서가 반응하는 정도의 관계를 보이는 것이며, 각 탐지센서의 탐지영역이 얼마나 관련성이 있는지를 알 수 있는 방법이 된다. 즉, 하나의 탐지센서가 임의의 행위에 대해 '침입' 가능성이 높다고 판단하였을 때 다른 탐지센서는 어떻게 판단하는 지가 탐지센서 사이의 관련성으로 표현한다.

탐지센서의 관련성은 일반적으로 같은 행위에 대하여 비슷한 결과 값을 나타내는 경우 관련성이 높다고 정의한다. 이러한 관련성은 동일한 행위에 대해 탐지센서 A의 결과 값이 탐지센서 B의 결과 값보다 클 때, 탐지센서 B의 결과 값은 탐지센서 A의 반영비율로 표현한다. 이러한 결과 값에 대해 i 번째 센서가 가장 높은 결과 값을 도출하였을 때 j 번째의 센서가 나타내는 값의 반영비율(η)을 표현한 식은 다음과 같다.

탐지센서의 수를 M, 각 탐지센서의 결과 값을 v_k, v_k가 나타날 확률을 P_i라 하면

$$P_i = P(v_i) = \frac{v_i}{\sum_{k=1}^{M} v_k},$$

$P_j \le P_i$ $(i = \max_{k=1}^{M} P_k)$ 이므로

$$\begin{aligned} \eta'_{i,j} &= P(P_j \mid P_i = \max_{k=1}^{M} P_k) \\ &= P\left(\frac{v_j}{\sum_{k=1}^{M} v_k} \;\middle|\; \frac{\max_{k=1}^{M}(v_k)}{\sum_{k=1}^{M} v_k} \right) \\ &= \frac{v_j}{\max_{k=1}^{M}(v_k)}, 1 \le i \le M, 1 \le j \le M \end{aligned} \quad (1)$$

이다.

호스트 반영비율 테이블의 생성은 각 탐지센서의 로그파일에서 지정된 시간 간격과 사용자를 비교하여 탐지결과 중의 최대값을 읽어와 통합로그에 기록한다. 통합로그에서 사용자 및 지정된 시간간격의 조건에 맞는 탐지센서의 결과는 가장 최대값을 보인 탐지센서의 행에 각각 기록한다. 반영비율 테이블의 각 행의 탐지센서에 기록시 횟수에 따른 평균 및 최대값의 탐지센서에 대한 비율로써 계산한다. 반영비율은 한번의 계산으로 얻어지는 것이 아니라 많은 학습 데이터를 통한 결과이어야 하며, 호스트와 지역 네트워크에서 반영비율 테이블 생성을 위한 로그파일은 정상행위 및 비정상행위가 포함된 명령 순서를 이용한다. 반영비율 테이블 생성을 위해서는 비정상행위가 포함된 로그를 사용하며, 각 탐지센서의 임계값 또는 필요시 사용자별 임계값을 구하기 위해서는 순수 정상행위 로그를 가지고 사용한다. 반영비율을 구하기 위한 학습용 로그 파일을 적용하여 각 센서별로 결과를 도출하고 평균값을 반영비율로 취한다. 즉, η를 η'의 평균으로 다음과 같이 보일 수 있다.

$$\eta_{i,j} = \frac{\sum_{t=0}^{R} \eta_{i,j}(t)}{R} \quad (2)$$

R : 시간 개념으로 일정한 기간의 레코드 크기

반영비율의 역할은 다중의 탐지센서가 동일한 행위에 대하여 각 결과 값을 보였을 때, 그 결과 값의 의미를 반영한 하나의 종합된 결과로 만드는 것이다. 반영비율은 하나의 탐지센서가 강한 '침입'의 가능성을 보였을 때, 그것을 얼마만큼 믿을 수 있느냐는 다른 탐지센서의 지지율이므로 탐지센서의 관련성이 잘 반영되어 있어야 한다. 반영비율은 탐지센

서 수에 따라 2차원 배열의 테이블로 구성된다. 반영테이블에서 행은 동일한 행위에 대해 '침입'의 가능성이 가장 크다고 결과 값을 갖는 탐지센서이며 각 열은 '침입'의 가능성에 대한 다른 탐지센서의 지지율이 된다.

각 탐지센서의 탐지영역의 결과를 보이는 반영비율 테이블이 생성되면, 실시간에는 각 탐지센서의 판정값(v)을 반영비율 테이블에 적용한다. 또한, 실시간의 각 탐지센서의 값을 반영비율에 적용시 반영비율 테이블의 각 행에 곱의 연산후 최대값을 최종 판정값으로 결정한다. 이것은 실시간의 탐지값이 각 탐지센서별로 판정 임계값 이하의 작은 비정상행위도를 나타내며 탐지값들이 유사한 값을 갖는 경우, 반영비율 테이블의 비정상행위에 따른 탐지결과 즉, 각 행의 탐지센서의 결과에 행렬의 곱 연산을 한 후 얻어진 값들 중에서 최대값을 취하도록 하여 임계치 이하의 비정상행위도를 갖는 비정상행위를 정상행위로 판정하는 오판율을 최소화하도록 한다. 즉, 최종 판정값(V)은 다음과 같다.

$$V = \max_{i=1}^{M} \left[\begin{pmatrix} \eta_{11}, \eta_{12}, \ldots, \eta_{1M} \\ \eta_{21}, \eta_{22}, \ldots, \eta_{2M} \\ \cdots \\ \cdots \\ \cdots \\ \eta_{M1}, \eta_{M2}, \ldots, \eta_{MM} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \cdot \\ \cdot \\ \cdot \\ v_M \end{pmatrix} \right]$$

$$= \max_{i=1}^{M} \left[\sum_{j=1}^{M} (\eta'_{i,j} \times v_j) \right] \quad (3)$$

이다.

〔표 1(a)〕과 같이 오프라인에서 탐지센서의 반영비율이 생성되고, (b)와 같이 실시간에 각 탐지센서의 결과 값이 있는 경우를 살펴보자. (b)의 탐지횟수 1, 3은 비정상행위, 탐지횟수 2는 정상행위에 대한 탐지결과이며, 판정 기준값은 0.5라고 한다. (b)의 각 탐지횟수 1, 2, 3의 평균에 의한 탐지결과는 각각 0.4, 0.4, 0.4이며, 탐지횟수 1은 탐지센서 A가 비정상행위임을 보인 의견이 반영이 되지 않았다. 또한 탐지횟수 3은 비정상행위이지만 각 탐지센서의 탐지결과가 판정기준값 이하의 행위도를 보임으로써 정상행위로 판정되었다.

〔표 1〕 탐지센서 탐지결과의 예

(a) 반영비율 테이블

탐지센서	A	B	C
A	0.7	0.5	0.1
B	0.1	0.5	0.4
C	0.2	0.2	0.6

(b) 실시간의 탐지결과

탐지횟수	탐지센서 A B C
1	1.000 0.200 0.000
2	0.200 0.500 0.500
3	0.400 0.400 0.400

이러한 탐지결과의 판정 방법은 각 탐지센서의 결과의 반영비율 테이블을 이용하여 다음과 같은 결과를 보일 수 있다. (b)의 탐지횟수 1은

$$\max_{i=1}^{3} \left[\begin{pmatrix} 0.7, 0.5, 0.1 \\ 0.1, 0.5, 0.4 \\ 0.2, 0.2, 0.6 \end{pmatrix} \begin{pmatrix} 1.0 \\ 0.2 \\ 0.0 \end{pmatrix} \right]$$

$$= \max_{i=1}^{3} [0.80, 0.20, 0.14] = 0.80\text{의 최종값}$$

을 갖게됨으로써 비정상행위로 판정될 수 있다. 이와 같은 방법으로 탐지횟수 2는 최종값 0.47에 의해 정상행위로 판정되며, 탐지횟수 3은 0.52의 값으로 비정상행위로서 판정될 수 있다. 따라서 정상행위 및 비정상행위에 따른 과탐지 및 미탐지의 오판율을 줄일 수 있다.

3.3 호스트 통합탐지

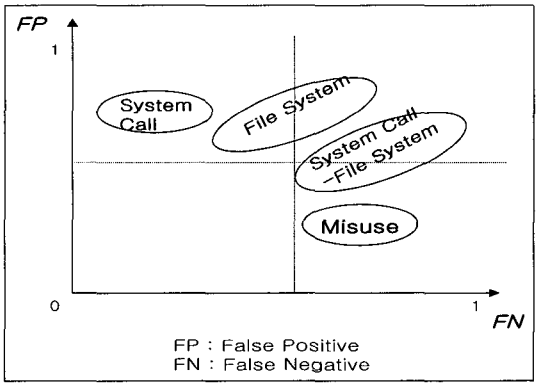
실시간으로 보내오는 호스트 탐지센서들의 로그에 대해서 통합탐지 처리는 탐지메시지의 수, 과탐지 및 미탐지의 오판율을 줄이기 위하여 각 탐지센서의 결과를 침입가능성이 높은 *High*와 판정이 모호한 부분으로 통합탐지 처리기에서 다른 탐지센서의 의견을 반영한 후 판정하도록 하는 *Medium* 값으로 판단한다. *High* 이상의 결과인 경우 오용행위 탐지센서는 즉시 세션을 종료하고 탐지메시지를 출력하고, 비정상행위 탐지센서의 경우는 탐지메시지를 출력한다. 만일 *High* 이하이고 *Medium* 이상의 결과인 경우 통합판정을 위한 단계를 수행한다. *Medium* 이하의 경우는 정상적인 경우로 판단한다.

통합판정을 위한 단계는 먼저, 각 로그는 탐지센서별 큐에 삽입하며, 하나의 큐가 가득 차게되면 해당 큐의 첫번째 로그를 가져오고 그 로그를 기준으로 다른 탐지센서 큐에 있는 로그들과 비교한다. 다음으로 로그의 값이 경험적인 기준 값보다 큰 경우 이 로그는 최종 로그에 바로 기록하고 통합하는 단계를 수행한다. 기준 값보다 작다면 가득 찬 큐에서 나온 로그의 시간(date)과 사용자(user_id)를 가지고 다른 큐에 있는 로그들의 시간과 사용자를 비교하는 방식으로 이루어진다. 비교과정에서 동일한 시간과 사용자를 갖는 큐가 발견되면 탐지한 센서의 수를 증가하고 다음 큐와의 비교를 계속한다. 모든

큐와의 비교가 이루어진 다음 탐지한 센서의 수가 둘 이상인 경우, 탐지한 센서의 로그들 중에 가장 큰 값을 나타내는 로그를 찾아서 반영비율 값을 적용하고 로그파일에 기록한다.

Ⅳ. 실험 및 분석

통합탐지시스템의 탐지센서에 임계값 및 반영비율을 적용한 통합탐지의 수행 분석을 위해 각 탐지센서를 독자적으로 수행한 결과와 통합탐지를 수행한 결과를 ROC(Receiver Operating Characteristic)를 이용하여 보였다.



〔그림 2〕 호스트기반의 HMM 탐지센서와 오용행위 탐지센서의 오판율 분포 특성

〔그림 2〕는 호스트기반의 HMM 탐지센서와 오용행위 탐지센서의 과탐지율(FP)과 미탐지율(FN)의 오판율 분포 특성을 ROC에 의해 보인 것이다. 통합탐지 수행전의 각 탐지센서의 결과는 오용행위 탐지센서의 경우 규칙이 정의되지 않은(잘 알려지지 않은) 공격에 대해서는 미탐지율이 높지만 공격이 아닌 행위를 공격으로 오판하는 과탐지의 경우는 작음을 보이고 있다. 마찬가지로 HMM을 이용한 시스템호출 탐지센서는 임의 순서의 명령에 대해서 탐지를 수행함으로써 공격이 아닌 행위에 대해 공격으로 오판하는 과탐지율은 높지만 미탐지율은 낮음을 보이고 있다. 따라서 반영비율을 적용한 통합탐지를 수행함으로써 탐지결과의 분포가 과탐지율 및 미탐지율을 감소하는(0의 방향으로 이동하는) 경우에 성능을 향상시킨다고 볼 수 있다. 호스트 오용행위 및 비정상행위 센서에 대한 통합탐지 전후의 비교를 위하여 〔표 2〕와 같은 침입유형과 내용을 사용하였으며, 10회 반복하여 임의 순서로 수행하였다.

〔표 2〕 오탐율 분석을 위한 침입유형 및 내용

침입 유형		침입 내용
오용행위		-buffer overflow -race condition
관리자직관에 의한 오용 행위	인증 공격	-su fail 시도 -login fail 시도
	시스템 자원 공격	-process table 공격 : fork -file system 공격 : mkdir, create file -메모리 할당 공격 : malloc
	권한 공격	-/etc/passwd 수정 및 삭제 시도 -사용자 계정 생성시도
비정상 행위	사용자 위장	-다른 사용자 계정으로 로그인

〔표 3〕은 오프라인에서 구하여진 반영비율 테이블이다. 순서기반의 비정상행위, 오용행위 탐지센서가 있는 호스트의 경우이며, 각 행의 최대값을 1로 할 때의 다른 탐지센서들의 지지율을 표현한 것이다. 예를 들면, 시스템 호출 비정상행위 탐지센서가 1이면(항상 탐지할 때), FA는 0.52, SF는 0.35의 지지율(탐지할 확률)을 보이는 것이다.

〔표 3〕 호스트 반영비율 테이블의 예

Sensor	SA	FA	SF	MU
SA	1	0.52	0.35	0.63
FA	0.54	1	0.57	0.75
SF	0.46	0.85	1	0
MU	0.67	0.74	0.55	1

SA : System call anomaly
FA : File system anomaly
SF : System call & File system anomaly
MU : Misuse pattern

〔표 4〕는 호스트에서 순서기반 탐지센서의 통합탐지 처리 전후의 결과를 보인 것이다.

〔표 4(a)〕는 통합탐지 전에 호스트의 순서기반 비정상행위 탐지센서와 오용행위 탐지센서의 오판율을 보인 것이며, 〔표 4(b)〕는 통합탐지 처리 후의 오판율을 보인 것이다. 통합처리 후의 결과는 'ALL'의 항목만이 의미가 있지만 비교를 위하여 통합탐지 전후의 각 탐지센서의 탐지결과를 보였다. 통합처리 전후의 'ALL' 항목을 비교하면 과탐지율 및 미탐지율이 개선되었음을 알 수 있다.

〔그림 3〕은 각 탐지센서별 통합탐지 수행 전과 수행 후의 분포 결과를 보인 것이다. 〔그림 3(a)〕의

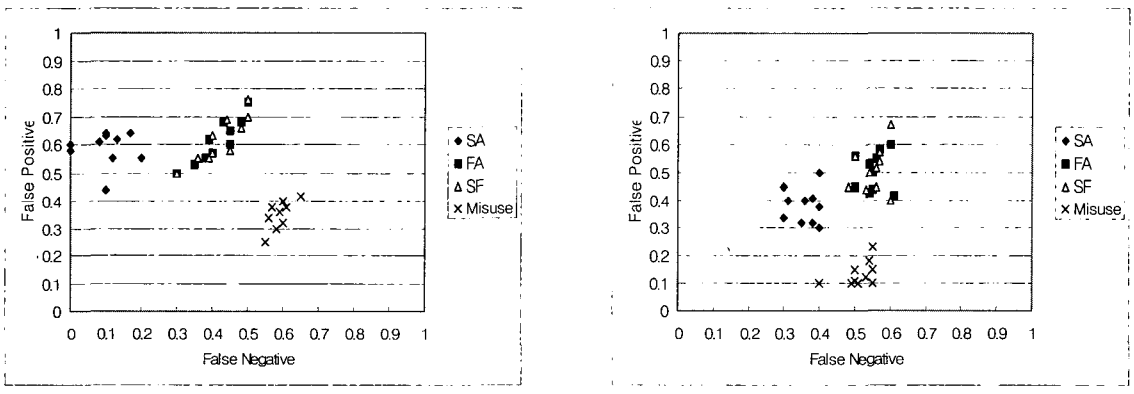
〔표 4〕 호스트 기반의 통합탐지 처리 전후의 결과

(a) 통합탐지 전

횟수	false alarm	SA	FA	SF	MU	ALL
1	FP	0.62	0.75	0.76	0.25	0.65
	FN	0	0.5	0.5	0.55	0.13
2	FP	0.64	0.57	0.57	0.4	0.58
	FN	0.2	0.4	0.4	0.6	0.00
3	FP	0.44	0.5	0.5	0.4	0.56
	FN	0	0.3	0.3	0.6	0.20
4	FP	0.66	0.69	0.55	0.3	0.62
	FN	0.1	0.35	0.36	0.58	0.12
5	FP	0.61	0.55	0.59	0.38	0.65
	FN	0.17	0.38	0.39	0.57	0.15
6	FP	0.64	0.65	0.66	0.32	0.68
	FN	0.08	0.45	0.48	0.6	0.18
7	FP	0.69	0.68	0.7	0.42	0.62
	FN	0.12	0.48	0.5	0.65	0.19
8	FP	0.6	0.6	0.58	0.38	0.65
	FN	0.1	0.45	0.45	0.61	0.18
9	FP	0.55	0.62	0.63	0.36	0.59
	FN	0.1	0.39	0.4	0.59	0.18
10	FP	0.56	0.68	0.69	0.34	0.62
	FN	0.18	0.43	0.44	0.56	0.19

(b) 통합탐지 후

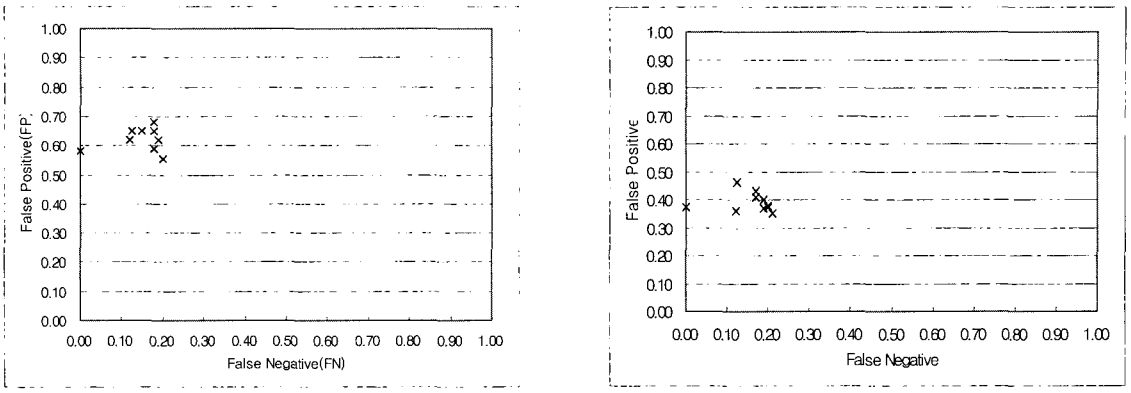
횟수	false alarm	SA	FA	SF	MU	ALL
1	FP	0.45	0.56	0.56	0.1	0.46
	FN	0.3	0.5	0.56	0.4	0.13
2	FP	0.5	0.45	0.67	0.23	0.38
	FN	0.4	0.5	0.6	0.55	0.00
3	FP	0.3	0.6	0.45	0.15	0.38
	FN	0.4	0.6	0.48	0.5	0.20
4	FP	0.4	0.5	0.52	0.15	0.36
	FN	0.31	0.55	0.56	0.55	0.12
5	FP	0.34	0.53	0.5	0.18	0.40
	FN	0.3	0.54	0.54	0.54	0.19
6	FP	0.4	0.58	0.57	0.11	0.38
	FN	0.36	0.57	0.57	0.5	0.20
7	FP	0.38	0.42	0.4	0.1	0.41
	FN	0.4	0.61	0.6	0.51	0.17
8	FP	0.41	0.43	0.44	0.1	0.37
	FN	0.38	0.54	0.53	0.55	0.19
9	FP	0.32	0.55	0.54	0.12	0.43
	FN	0.38	0.56	0.57	0.53	0.17
10	FP	0.32	0.44	0.45	0.1	0.35
	FN	0.35	0.55	0.56	0.49	0.21



(a) 통합처리전의 결과 (b) 통합처리 후의 결과

〔그림 3〕 통합처리 전후의 각 탐지센서별 분포도

각 탐지센서별 오판율의 분포도를 살펴보면 〔그림 2〕와 유사함을 알 수 있으며, 〔그림 3(b)〕의 탐지센서별 통합탐지 후의 분포도를 살펴보았을 때, 과탐지율은 조금씩 작아졌지만, 미탐지율은 오히려 커졌음을 알 수 있다. 이것은 통합시에 적용하는 탐지수준의 설정시 90% 유의 수준을 적용함으로써 실제 반영비율을 적용하여 판단하여야 하는 판정이 모호한 수준의 데이터가 정상행위로 판단되어 탐지가 되지 않았음을 추정케 한다. 따라서 관리자의 직관에 따른 탐지수준의 결정이 필요하다.



(a) 통합처리전의 결과 (b) 통합처리 후의 결과

〔그림 4〕 통합처리 전후의 총합 분포도

〔그림 4〕는 통합탐지 전의 각 탐지센서별 결과를 통합한 내용과 호스트 통합탐지 처리 후의 결과를 보인 것이다. 〔그림 4〕의 결과처럼 통합처리 전후의 총합 결과는 미탐지율의 약간의 차이와 과탐지율이 20% 정도 낮아진 결과를 보이고 있다. 현재의 과탐지율의 값이 0.4 수준에 위치하는 것은 학습오차 및 알고리즘의 오차 또한 비정상행위 발생의 순서 오류에 의한 탐지결과로 보여진다. 〔그림 3(a)〕의 통합처리전의 각 탐지센서별 결과와 〔그림 4(b)〕의 통합처리 후의 결과를 비교하였을 때 오용행위 탐지센서의 미탐지율은 비정상상 행위의 탐지센서와 관련성 분석에 의하여 현저히 개선되었음을 보이며, 시스템호출, 파일시스템, 시스템호출-파일시스템 비정상행위의 각 탐지센서의 결과도 관련성에 의해 향상되었음을 보인다. 따라서, 과탐지율 및 미탐지율이 각각 0.5이하의 영역에 위치하여 통합탐지 후에 과탐지율 및 미탐지율을 개선한 결과를 보이고 있다.

V. 결 론

침입탐지시스템의 성능 향상을 위하여 적용된 통합탐지 알고리즘은 비정상행위 침입탐지의 오판율을 줄이기 위하여 계층적 통합탐지 기반환경을 모델로 하였다. 각 탐지센서로부터의 탐지결과에 대해 관련성을 부여하여, 탐지하기 어려운 조화된 공격 및 알려지지 않은 공격에 대한 탐지율을 높이고 판정에 대한 신뢰도를 향상시키는 방법을 고려하였다.

통합탐지에서 오용행위 및 비정상행위 탐지 센서로부터의 결과가 서로 다른 센서의 탐지결과와 어떠한 연관성을 갖고 있는지를 적용하기 위하여 오프라인에서 반영비율을 생성하였으며, 실시간에 각 탐지센서의 결과를 반영비율에 적용하도록 하였다.

통합탐지의 방법은 최근 발생 빈도가 잦은 대규모 네트워크 구조에의 공격을 탐지하여 시스템을 보호할 수 있을 뿐만 아니라 알려지지 않은 공격을 탐지하는 능력으로 새로운 공격을 분석할 수 있다. 향후 지역 및 전역 네트워크에서도 각 호스트 및 지역 네트워크의 통합탐지처리기로부터 전송된 통합판정의 값을 가지고 호스트에서의 방법과 비슷하게 수행할 수 있다.

참 고 문 헌

〔1〕 D. E. Denning, "An Intrusion-Detection Model," *IEEE Trans. on Software Engineering*, No. 2, Feb., 1987.

〔2〕 M. Esmaili, R. Safavi-Naini, and J. Pieprzyk, "Intrusion Detection : a Survey," *International Conference in Computer Communication*, pp. 409~414, 1995.

〔3〕 T. F. Lunt, "A Survey of Intrusion Detection Techniques," *Computer & Security*, Vol. 12, No. 4, Jun., 1993.

〔4〕 R. Büschkes, M. Borning, and D. Kesdogan, "Transaction-based Anomaly Detection," *Proc. of the Workshop on Intrusion Detection and Network monitoring*, USENIX, Apr., 1999.

〔5〕 A. Mounji and B.L. Charlier, "Continuous Assessment of an Unix Configuration :

Integrating Intrusion Detection and Configuration Analysis," *Proc. of Symposium on Network and Distributed System Security*, 1997.

[6] M. Sebring, E. Shellhouse, M. Hanna, and R. Whitehurst, "Expert Systems in Intrusion Detection:A Case Study," *Proc. of 11th National Computer Security Conference*, Oct., 1988.

[7] S. J. Stolfo, et al., "JAM: Java Agents for Meta-learning over Distributed Databases," *Proc. of KDD-97 (runner up best paper, applications) and AAAI97 Workshop on AI Methods in Fraud and Risk Management*, 1997.

[8] ITO, "Survivability of Large Scale Systems," *DARPA/ITO*, http://www.darpa.mil/ito/research/lss/projects.html, 1999.

[9] M. Crosbie, E.H. Spafford, "Defending a Computer System using Autonomous Agents," *Proc. of the 18th National Information Systems Security Conference*, Baltimore, MD, pp. 549-558, Oct., 1995.

[10] S. Staniford-Chen, et al., "GrIDS - A Graph Based Intrusion Detection System for Large Networks," *In Proceedings of the 19th National Information Systems Security Conference*, pp. 361~370, Oct. 1996.

[11] H. Debar and A. Wespi, "Aggregation and Correlation of Intrusion-Detection Alerts," *IBM Research Report*, RZ 3360, Aug., 2001.

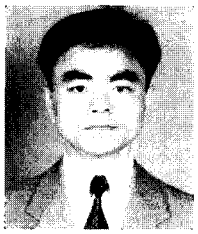
[12] P. A. Porras and P. G. Neumann, "EMERALD : Event Monitoring Enabling Responses To Anomalous Live Disturbances," *Proc. of the 20th National Information Systems Security Conference*, pp. 1~13, 1997.

[13] A. Valdes and K. Skinner, "An Approach to Sensor Correlation," *3rd International Workshop on the Recent Advances in Intrusion Detection*, Oct., 2000.

[14] A. Valdes and K. Skinner, "Probabilistic Alert Correlation", *4th International Symposium on the Recent Advances in Intrusion Detection*, pp. 54~68, Oct., 2001.

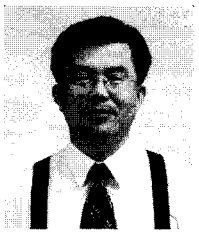
[15] 한국정보보호진흥원, *정보통신기반구조 보호기술개발*, 보고서, 2000.

〈著者紹介〉

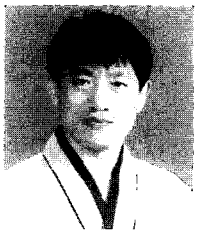


김 용 민 (Yong-Min Kim)
1989년 : 전남대학교 전산통계학과 졸업
1991년 : 전남대학교 전산통계학과(이학석사)
1996년~현재 : 전남대학교 전산통계학과 박사과정
〈관심분야〉 네트워크 관리, 시스템 및 네트워크 보안, 정보보안, 퍼지 이론 등

김 민 수 (Min-Soo Kim)
1993년 : 전남대학교 전산통계학과 졸업
1995년 : 전남대학교 전사통계학과(이학석사)
2000년 : 전남대학교 전산통계학과(이학박사)
2000년~2001년 : 한국정보보호진흥원 선임연구원
2001년~현재 :전남대학교 정보보호협동과정 객원교수
〈관심분야〉 시스템 및 네트워크 보안, 정보보안, 신경망 등



김 홍 근 (Hong-Gun Kim)
1985년 : 서울대학교 컴퓨터공학과 졸업
1987년 : 서울대학교 대학원 컴퓨터공학과(공학석사)
1994년 : 서울대학교 대학원 컴퓨터공학과(공학박사)
1994년~1996년 : 한국전산원 선임연구원
1996년~현재 : 한국정보보호진흥원 기술개발단장
〈관심분야〉 컴퓨터 보안, 병렬 알고리즘



노 봉 남 (Bong-Nam Noh) 정회원
1978년 : 전남대학교 수학교육과 졸업
1982년 : KAIST 대학원 전산학과(공학석사)
1994년 : 전북대학교 전산통계학과(이학박사)
1983년~현재 : 전남대학교 컴퓨터정보학부 교수
〈관심분야〉 통신망관리, 정보보안, 시스템 및 네트워크 보안 등