다중척도 모델의 결합을 이용한 효과적인 침입탐지

(Effective Intrusion Detection Integrating Multiple Measure Models)

한 상 준 † 조 성 배 ††
(Sang-Jun Han) (Sung-Bae Cho)

요 약 정보통신기술이 발전함에 따라 내부자의 불법적인 시스템 사용이나 외부 침입자에 의한 중요 정보의 유출 및 조작을 알아내는 침입탐지시스템에 대한 연구가 활발히 이루어지고 있다. 이제까지는 네트워크 패킷, 시스템 호출 감사자료 등의 척도에 은닉 마르코프 모델, 인공 신경망, 통계적 방법 등의 모델링 방법을 적용하는 연구가 이루어졌다. 그러나 사용하는 척도와 모델링 방법에 따라 취약점이 있어 탐지하지 못하는 침입이 많은데 이는 침입의 형태에 따라 흔적을 남기는 척도가 다르기 때문이다. 본 논문에서는 이러한 단일척도 침입탐지시스템의 단점을 보완하기 위해 시스템 호출, 프로세스의 자원점유율, 파일 접근이벤트 등의 세 가지 척도에 대하여 은닉 마르코프 모델, 통계적 방법, 규칙기반 방법을 사용하여 모델링한 후, 그 결과를 규칙기반 방법으로 결합하는 침입탐지 방법을 제안한다. 실험결과 다양한 침입 패턴에 대하여 다중척도 결합방법이 매우 낮은 false-positive 오류율을 보여 그 가능성을 확인할 수 있었다.

키워드 : 침입탐지시스템, 비정상행위 탐지, 다중척도모델링

Abstract As the information technology grows interests in the intrusion detection system (IDS), which detects unauthorized usage, misuse by a local user and modification of important data, has been raised. In the field of anomaly-based IDS several artificial intelligence techniques such as hidden Markov model (HMM), artificial neural network, statistical techniques and expert systems are used to model network packets, system call audit data, etc. However, there are undetectable intrusion types for each measure and modeling method because each intrusion type makes anomalies at individual measure. To overcome this drawback of single-measure anomaly detector, this paper proposes a multiple-measure intrusion detection method. We measure normal behavior by systems calls, resource usage and file access events and build up profiles for normal behavior with hidden Markov model, statistical method and rule-base method, which are integrated with a rule-based approach. Experimental results with real data clearly demonstrate the effectiveness of the proposed method that has significantly low false-positive error rate against various types of intrusion.

Key words : intrusion detection systems, anomaly detection, multi-measure modeling

1. 서 론

정보통신 기술의 발전으로 네트워크 환경이 광역화, 고속화되어 이제까지는 불가능했던 다양한 서비스가 가능해졌으며 업무 효율 또한 많이 향상되었다. 하지만 점점 많은 분야에서 컴퓨터를 활용하게 됨에 따라 외부 불법 침입자의 공격에 의한 중요정보 유출 및 조작, 시스템의 불법적 사용 등에 의한 피해도 커져 시스템 보안에 대한 요구와 관심이 증대되고 있다. 실제로 한국정보보호진흥원의 보고에 따르면 2000년에는 1,943건의 국내 해킹사고가 접수되어 전년인 99년에 비해 3배 증가하였고 2001년에는 5,333건이 접수되어 폭발적인 증가세를 보이고 있다. 또한 인터넷의 대중화로 누구나 쉽게 해킹 도구를 이용할 수 있게 되어 단순한 호기심에

· 본 연구는 대학 IT 연구센터 육성/지원사업의 연구결과로 수행되었음.
† 학생회원 : 연세대학교 컴퓨터과학과
sjhan@cs.yonsei.ac.kr
†† 종신회원 : 연세대학교 컴퓨터과학과 교수
sbcho@cs.yonsei.ac.kr
논문접수 : 2003년 1월 13일
심사완료 : 2003년 3월 11일

의한 공격도 생겨나 앞으로 해킹사고는 더욱 늘어날 전망이다[1].

이렇게 해킹에 의한 인적 물적 손실이 늘어남에 따라 불법 침입에 대비하기 위한 여러 가지 도구와 장비들이 필수적인 컴퓨터 시스템의 요소가 되었는데 그중 대표적인 것이 침입탐지 시스템이다. 이것은 불법적인 시스템 사용에 대응하기 위한 중요한 도구중의 하나로 외부 침입자의 공격을 탐지할 수 있게 해주는 소프트웨어이다[2]. 침입을 탐지하기 위한 기법에는 미리 알려진 공격행위에 대한 정보를 구축하고 이를 이용해 침입을 판정하는 오용탐지 방법과 사용자나 프로그램의 정상행위에 대한 정보를 구축하고 이를 이용하는 비정상행위 탐지 방법의 두 가지가 있다[3]. 대부분의 상업용 침입탐지 시스템은 오용탐지 기법을 적용한 규칙기반 침입탐지 시스템이기 때문에 새로운 공격에 대한 탐지가 힘들다. 이런 오용탐지 기법의 단점보완을 위해 최근에는 비정상행위기반의 침입탐지 시스템에 관한 연구가 활발한데, 비정상행위기법의 경우에도 수집하는 정상행위 정보와 모델링 방법에 따라 탐지할 수 있는 침입의 종류가 제한적이라는 단점이 있다.

본 논문에서는 기존 비정상행위기반 침입탐지시스템의 단점을 극복하고 탐지 성능을 향상시키기 위해 시스템 호출, 프로세스의 자원사용량, 파일 시스템 정보 등의 척도와 각 척도에 맞는 최적의 정상행위 모델링 방법을 제시하고, 얻어진 척도별 정상행위 모델을 규칙기반 방법으로 결합하는 탐지방법을 제안한다. 이런 결합방법은 단일 척도기반 침입탐지에 비해 정상행위를 여러 가지 측면에서 모델링할 수 있기 때문에 더 좋은 결과를 기대할 수 있는데 실험을 통해 단일척도 모델링 방법에 비해 좋은 성능을 보이는지를 평가한다.

본 논문의 나머지 부분은 다음과 같이 구성된다. 2장은 비정상행위탐지 침입탐지시스템의 현황에 대해 소개하고, 3장은 제안하는 시스템의 구성과 척도별 모델링 방법에 대해 설명한다. 4장은 실험 과정과 결과를 설명하고, 5장은 논문의 결론과 향후연구에 대해 언급한다.

2. 비정상행위탐지 침입탐지시스템

대표적인 비정상행위기반의 침입탐지 방법에는 전문가시스템 방법, 통계적 방법, 신경망, 은닉 마르코프 모델(Hidden Markov Model, HMM) 등이 있다. 각 탐지 방법의 대표적인 연구들을 정리하면 다음 표 1과 같다.

통계적 방법은 비정상행위기법 중 널리 사용되는 방법인데, 사용자나 시스템의 행동을 시간에 따라 샘플링된 여러 가지 변수들에 의해 측정한다. 각 세션의 로그인과 로그아웃시간, 자원 지속성, 프로세스 당 소비된 프로세서-메모리-디스크 양 등이 변수의 예가 될 수 있다. 샘플링 기간은 작게는 몇 분에서 길게는 여러 달까지 다양하다. 정상적인 사용자의 자원 사용량, 명령어 패턴, 로그인 시간의 정보 등을 통계적으로 분석한 후 입력된 정보와 비교하여 침입을 탐지한다. 대표적인 침입탐지 시스템으로 NIDES(Next-generation Intrusion Detection Expert Systems)가 있는데 이 시스템은 장기적인 시스템의 프로파일과 단기간의 프로파일간의 유사도를 측정하여 침임탐지의 기준으로 삼는다[8]. 다양한 통계적인 방법들을 적용할 수 있으며 과거의 경험적인 자료를 토대로 처리하기 때문에 탐지율이 상당히 높다. 그러나 어떤 행위의 발생 순서에는 민감하지 못하며 모델링할 수 있는 침입 행위의 종류가 제한적이라는 단점이 있다.

신경망 모델은 통계적 기법과 부분적으로 비슷한 점이 있으나, 통계적 기법에 비해 변수들 간의 비선형적 관계를 표현하기가 쉽고 자동적으로 학습할 수 있다는 장점이 있다. 신경망을 이용한 대표적 침입탐지 시스템은 프랑스의 CS Telecom에서 개발한 Hyperview를 들 수 있는데 전문가 시스템과 신경망 2개의 모듈로 구성되어 있다[10]. 신경망을 통한 사용자 행위 학습의 경우

표 1 대표적인 침입탐지시스템 연구

기관	이름	기간	방 법			
			전문가 시스템	신경망	통계적 방법	HMM
AT&T	ComputerWatch[4]	1987-1990	X			
UCDavis	NSM[5]	1989-1995			X	
	GrIDS[6]	1995	X			
SRI International	IDES[7]	1983-1992			X	
	NIDES[8]	1992-1995			X	
	EMERALD[9]	1996-			X	
CS Telecom	Hyperview[10]	1990-1995		X	X	
Univ. of New Mexico	C.Warender et. al [11]	1999			X	X
Yonsei Univ.	Park and Cho [12]	2002				X

시간의 흐름에 관한 데이타의 윈도우를 신경망 입력 값으로 하여 사용하는데, 입력 값으로 명령어이름, CPU 사용량, 메모리 사용량 등 60개의 감사 자료를 이용한다. 신경망은 통계적 기법에 비해 비선형적 관계를 잘 표현하고 자동적으로 학습하는 장점이 있지만 계산량이 많고 입력과 출력간의 관계를 알 수 없는 단점이 있다.

HMM모델은 생성경위를 알 수 없는 이벤트시퀀스를 모델링하고 평가하는 도구로 뉴멕시코 대학의 C. Warender 등이 제안한 시스템 호출 이벤트를 척도로 사용한 침입탐지시스템이 대표적이다[11]. HMM은 다른 방법보다 시스템 호출 이벤트를 모델링하는데 있어서 좋은 성능을 보여주지만 정상행위 모델링과 침입 탐지 과정에서 매우 많은 시간이 소요되는 문제 때문에 실시간 침입탐지에 사용되기는 어렵다. 이런 단점을 극복하기 위한 방법으로는 시스템의 성능을 향상시키거나 데이타를 축약하는 등의 방법이 있는데 권한이동 이벤트 전후의 이벤트만을 추출하여 정상행위 모델링을 위한 데이타를 축약하는 방법을 사용하면 모델링 시간을 획기적으로 줄이면서도 좋은 성능을 유지할 수 있다[12].

3. 다중척도결합기반 침입탐지

여러 가지 유형의 침입이 발생하고 있지만 그중 호스트 기반 침입이 가능한 침입유형은 버퍼오버플로우(Buffer Overflow), S/W보안오류, 구성설정오류, 서비스 거부 공격 등이 있다. 한국정보보호진흥원의 2002년도 6월, 7월의 해킹 동향 조사를 살펴보면 다음 표 2와 같은데 버퍼오버플로우 침입이 대부분임을 알 수 있다. 특히 최근에 서비스 거부 공격은 인터넷사용자들의 집단행동으로 문제가 되고 있는 공격 유형이다. 본 논문에서는 이 두 가지 유형의 침입을 분석하고 그에 따른 침입탐지시스템을 구성하였다.

침입탐지에 사용되는 여러 가지 척도 중 본 논문에서는 호스트 기반 침입탐지에 주로 쓰이는 시스템 호출 이벤트, 파일 시스템 정보, 프로세스 자원사용량을 사용하였다. 각 척도를 모델링하는 방법은 여러 가지가 있지만 침입행위가 각 척도에 남기는 흔적의 특징들과 어떤 모델링 방법이 잘 구별해 줄 수 있는 지를 고려하여 척도별로 가장 적합한 방법을 선택하였다. 그러나 침입 유형에 따라 각 척도에 남기는 흔적이 다르기 때문에 각 방법별로 탐지에 적합한 침입 유형이 한정되어 있다. 그러므로 이를 개선하기 위하여 탐지방법별 결과를 결합하여 각 척도의 단점을 극복하는 과정이 필요하다. 이를 위해 척도별 관계와 침입의 특성을 고려하고 경험을 바탕으로 결합 규칙을 생성하였고 최종적으로 이 규칙에 의해 각 방법별 결과를 종합하여 침입을 판정하게 하였다. 제안하는 방법의 구조는 그림 1과 같다.

표 2 침입 유형 동향

유 형	2002년 5월	2002년 6월	2002년 7월
S/W 보안오류	0	1	1
버퍼오버플로우	34	25	10
구성,설정 오류	3	4	0
서비스 거부공격	0	5	0
계	36	30	10

3.1 시스템 호출 척도를 이용한 HMM방법

솔라리스 운영체제의 BSM(Basic Security Module)에서 제공하는 감사자료는 운영체제에서 발생한 시스템 호출 이벤트들과 그에 관련된 사용자 및 프로세스 정보를 포함하는 자료로 호스트기반 침입탐지 시스템에서 널리 사용되는 척도이다. 프로그램의 오작동을 유도하여 루트권한을 획득하는 버퍼오버플로우 등의 침입유형은 정상행위의 것과는 다른 시스템 호출 이벤트를 발생시킨다. 따라서 정상행위의 시스템 호출 이벤트 시퀀스를 모델링한 후 이와 비교하여 다른 시퀀스를 발생시키는 프로세스나 사용자를 찾아내면 효과적으로 침입을 탐지할 수 있다. 본 논문에서는 시스템 호출 감사자료를 모델링하기 위하여 음성인식과 영상인식 등의 분야에서 널리 쓰이는 HMM (Hidden Markov Model)을 사용하였다. HMM은 실제적인 생성경위를 알기 힘든 이벤트 시퀀스를 잘 모델링할 수 있는 방법으로 시스템 호출 이벤트 시퀀스를 모델링하는데 매우 유용한 도구이다[13].

HMM은 상태의 집합 Q와 관찰기호의 집합 V, 관찰기호의 수 M에 기반한 초기상태 확률분포 π, 상태전이 확률분포 A 및 관측기호 확률분포 B의 세가지 확률분포로 구성된다. 모델의 상태 수 N과 관찰기호 시퀀스 길이 T에 따라 다른 형태로 설정되며 일반적으로 기호로는 $\lambda=(A, B, \pi)$로 나타낸다.

여기서 관찰기호 집합 V는 감사자료에 존재하는 모든 시스템 호출의 집합에 해당하며 M은 그 개수이다. 관찰기호 시퀀스의 길이 T는 HMM에 입력될 시스템 호출 시퀀스의 길이에 해당하는데 방대한 양의 모든 감사자료를 한번에 학습시킬 수는 없으므로 전체 감사자료를 윈도우단위로 분할해 적용한다. 정해진 크기의 윈도우를 한 단계씩 이동시키며 얻어진 자료를 사용하였는데 T는 정해진 윈도우의 크기에 해당된다. 본 논문에서는 일반적으로 시간이나 순서적인 신호를 더 잘 모델링한다고 알려져 있는 left-to-right 형태의 HMM을 사용하였다[14].

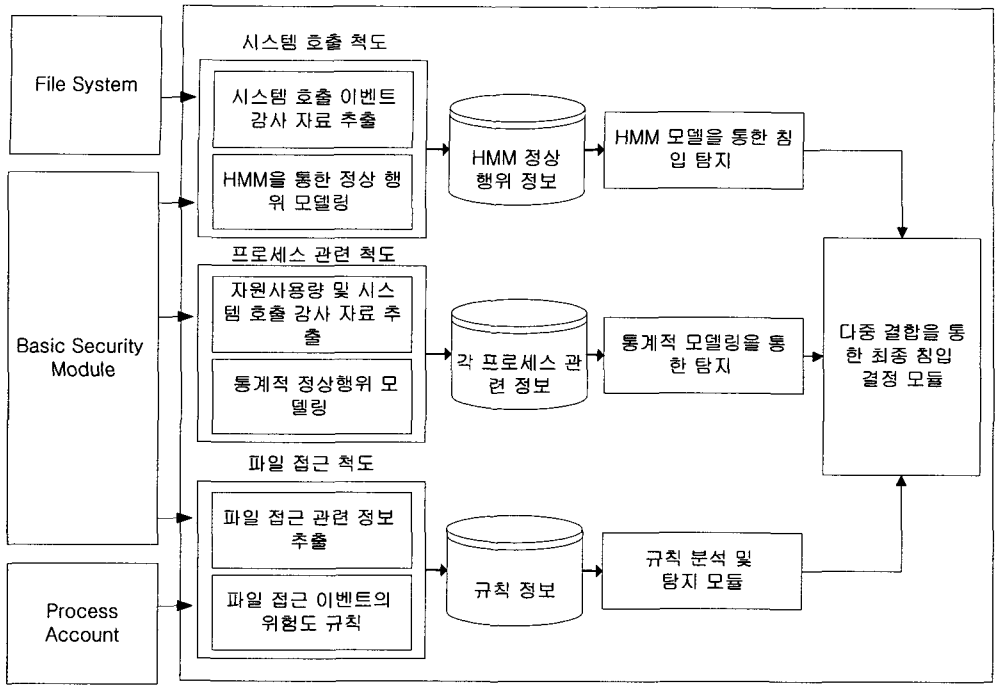

그림 1 제안하는 침입탐지기법의 구조

HMM을 이용한 침입탐지는 정상행위 모델링과 침입탐지의 두가지 과정으로 나뉠 수 있다. 정상행위 모델링은 구축된 모델 λ로부터 주어진 정상행위 시스템 호출 시퀀스 O가 나올 수 있는 확률 값인 P(O|λ)가 최대가 되도록 HMM의 구성요소들을 조정해 나가는 과정으로서 Baum-Welch 재추정식을 사용하여 정상행위 감사자료를 모델링한다. 침입탐지 과정은 정상모델과 입력된 시퀀스를 비교하여 침입을 판별하는 과정이다. 생성된 정상행위 모델에 침입이 들어있는 시스템 호출 감사자료를 입력하고 정상행위 모델에서 입력된 행위가 나올 확률 값을 구해 평가한다. 이 확률값은 forward-backward procedure을 사용해 구해지는데 확률 값이 정상행위 모델링에서 정한 임계값보다 낮게 나올 경우 침입으로 판정한다.

3.2 자원사용량과 시스템 호출 척도를 이용한 통계적 방법

대부분의 유닉스기반 운영체제에서 감사자료로 쓰이고 있는 PACCT(Process Account)감사자료는 프로세스의 CPU점유율, 메모리 사용량, I/O사용량 등의 정보를 제공한다. 서비스 거부 공격(Denial of Service)을 받는 프로세스는 정상적인 프로세스와는 달리 시스템자원 사용량이 급격하게 상승하게 된다. 이런 유형의 공격은 PACCT 감사자료를 이용하여 정상프로세스와 다른 양상으로 자원을 사용하는 프로세스를 찾아내는 방법으로 침입을 탐지할 수 있다.

PACCT 감사자료는 서비스 거부 공격을 탐지하는데 좋은 자료이지만 이 자료에 기록되지 않는 자원을 대상으로하는 유형의 공격은 탐지하지 못하는 단점이 있다. 예를 들면 운영체제가 처리할 수 있는 프로세스의 수 이상의 프로세스를 생성해 프로세스 테이블을 고갈시켜 시스템이 올바로 기능할 수 없게 만드는 형태의 공격은 PACCT 자료에는 별다른 흔적을 남기지 않는다. 하지만 이런 유형의 공격은 정상행위와는 달리 과다한 양의 시스템 호출 이벤트를 만들어낸다. 이 경우에는 프로세스가 발생시킨 시스템 호출 이벤트 수의 양상을 정상행위의 것과 비교하여 다른 프로세스를 찾아내는 방법이 유용하다.

본 논문에서는 정상적인 프로세스가 남긴 감사자료를 모델링하여 정상행위 프로파일을 만들고 입력된 감사자료의 침입을 판정하는데 통계적인 방법을 사용하였다. 기존연구로 PACCT자료에 HMM을 이용하여 정상행위를 모델링한 사례가 있었지만 좋은 성능을 보여주지는

못하였다[15]. 통계적 방법은 장기 데이타와 단기 데이타사이의 차이를 비교하여 정량화하는 방법을 사용하였다. 장기 데이타와 단기 데이타 변화의 계산은 Q통계치에 의하여 실행되는데 장기적인 사용자의 행위와 최근의 단기 행위를 비교하여 차이가 클수록 더 큰 Q통계치를 가지게 되며 적정 수준을 넘는 값을 가지는 경우 침입으로 간주한다. 정상적인 사용자의 행위에서 Q통계치의 분포를 구한 후 Chi-square와 유사한 방법으로 테스트 데이타의 Q통계치와 비교하여 침입을 탐지한다. Chi-square란 교차분석에서 이론적인 빈도와 실제빈도의 차이를 이용하여 독립성 혹은 관련성이 있는지의 여부를 판단하는 방법 중 가장 많이 이용되는 방법인데 본 논문에서는 정상행위의 Q통계치의 빈도와 테스트 데이타의 차이를 이용하여 침입을 탐지한다. 본 논문에서는 기존의 NIDES에서 사용된 방법을 기반으로 서비스 거부 공격을 탐지하기에 적합하도록 수정하여 사용하였다.

D_k를 k번째 자료와 k+1번째 자료 사이의 변화량이라 하고, t_k는 마지막에 들어온 자료와 k번째 자료 사이의 시간, r을 최근 자료의 가중치라고 할 때 Q통계치는 다음 공식으로 얻을 수 있다[16].

$$Q = \sum_{k\geq 1}^{\infty} D_k \times 2^{-rt_k}$$

이 공식에서 오래전의 레코드일수록 t_k값이 커져 Q값에 미치는 영향은 작아지게 되고 r값이 클수록 작아지는 비율이 커지게 되어 일정 t_k값을 넘어선 레코드는 더 이상 통계값에 영향을 주지 못하게 된다. 따라서 적당한 r값을 선정하여 현재 레코드부터 얼마만큼 떨어진 레코드까지 반영할지를 결정하는 것이 Q통계치를 계산하는데 있어서 중요한 요소이다.

t_k를 계산하는 방법에는 감사 레코드가 발생한 시간을 기준으로 하는 방법과 발생한 순서를 기준으로 하는 방법이 있다. 시간을 기준으로 할 경우 시스템의 사용시간에 따라 통계치가 안정되지 않은 단점이 있지만 최근 값에 가중치를 확실히 둘 수 있는 장점이 있다. 발생한 순서로 계산하는 경우 시스템의 주 사용시간에 관계없이 통계치가 안정되는 장점이 있지만 통계치에 반영되는 레코드들의 시간대가 고르지 않은 단점이 있다. 본 논문에서는 레코드가 발생한 순서를 기준으로 하는 방법을 사용하였는데 이는 정상행위 모델 구축에 사용한 시스템이 특정시간에 많이 사용되는 것이기 때문이다.

이러한 공식에 의하여 정상행위 자료에 대한 Q통계치를 얻어낸 후 이 값을 미리 정해놓은 구간에 대응시켜 정상행위의 Q통계치가 각 구간에 속할 확률을 구한다. 이 확률을 누적 정규분포표에 대응시켜 각 구간에 대한 최종 평가 값을 구하는데 이 값을 S통계치라고 한다. 실험에서 S통계치를 구하는데 사용된 Q 통계치의 구간은 표 3과 같다.

구간의 처음은 전체 정상행위 평가값의 누적분포를 구한 후 각각 50%, 60%, 70%, 80%, 90%의 값으로 다섯 개의 구간을 설정하고 그 이후의 구간은 다섯 번째 구간 값의 배수를 취하여 설정하였다.

이렇게 정상행위 자료에서 얻어진 각 구간별 S통계치는 침입 판단의 기준이 되는데, 테스트 자료의 Q통계치가 속한 구간의 S통계치가 그 자료의 평가 값이 된다. PACCT레코드의 경우 프로세스의 자원사용량을 종합적으로 평가하기 위해 각 척도 별로 구해진 S통계치에 가중치를 두어 결합하였다. 이렇게 구해진 하나의 레코드에 대한 최종 평가 값을 T 통계치라하고 이를 이용해 침입여부를 결정한다. T값은 s_n을 각 척도별 평가 값, a_n을 각 척도별 가중치라고 했을 때 다음 식에 의하여 구해진다.

$$T = a_1 s_1 + a_2 s_2 + \ldots + a_n s_n$$

3.3 파일접근 정보를 이용한 규칙기반 방법

일반적으로 공격행위는 공격을 성공시키기 위하여 보안레벨이 높은 파일에 접근한다. 예를 들면 SETUID권한이 설정된 파일을 실행하거나 루트 권한을 얻은 후 시스템을 파괴하거나 중요한 정보를 얻기 위하여 관리자나 루트권한의 파일에 읽기를 시도하거나 속성을 변경한다. 때문에 파일 접근 이벤트를 이용하여 낮은 권한의 사용자나 프로세스가 높은 보안 레벨의 파일에 접근하는 것을 관찰하는 방법으로 침입을 탐지할 수 있다.

파일접근에서 주가 되는 3가지 요소는 주체(Subject)인 프로세스 또는 사용자와 객체(Object)가 되는 파일 그리고 둘 사이에 일어나는 접근(Access)이벤트이다. 파일 접근을 활용한 대표적인 규칙은 Bell과 LaPadula의 BLP모델을

표 3 각 구간별 평가값(S값) 분포

	1	2	3	4	5	6	7	8
CPU	0.298	0.597	3.896	8.146	15.616	30.0	60.0	90.0
Memory	0.28	0.827	7.028	13.4223	35.422	70.0	140.0	210.0
I/O	32.712	60.859	175.492	259.629	276.488	350.0	500.0	700.0
System Call	0.52	0.637	5.632	24.53	40.122	80.0	120.0	160.0

들 수 있는데 그림 2에서와 같이 낮은 레벨을 가지는 주체가 높은 레벨의 객체에 접근하는 것을 제한하는 방법으로 간단히 No Read Up rule이라고도 한다[17].

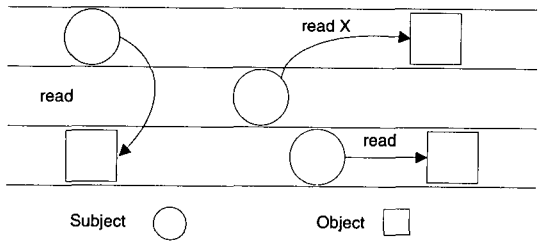

그림 2 BLP모델 규칙

본 논문에서는 이를 응용하여 파일 접근 이벤트를 관찰하고 이벤트의 내용을 분석하여 위험도를 평가하였다. 주체와 객체의 보안레벨은 루트, 시스템 관리자, 일반 사용자, 방문자의 4단계로 나누었고 주체와 객체사이의 레벨 차이와 객체가 가지는 권한설정에 따라 0부터 21까지의 위험도로 평가하였다. 이렇게 얻어진 파일 접근 이벤트의 위험도는 표 4와 같다.

표 4 파일접근이벤트의 위험도 평가의 예

주체	객체	모드	위험도
root	root	755	5
root	user	755	5
user	root	744	9
user	root	755	14
user	admin	644	4
user	guest	555	3
guest	root	711	17
guest	admin	555	6
guest	user	755	6

주체와 객체사이의 레벨의 차이가 클수록 높은 위험도를 가지며 객체의 레벨이 더 낮을 경우는 위험도를 높이지 않는다. 방문자 레벨의 주체가 루트레벨의 파일에 접근할 경우 가장 높은 위험도를 가지지만 루트 레벨의 주체가 일반사용자 레벨의 파일에 접근할 경우는 위험도를 높이지 않았다. 그리고 파일에 더 많은 권한이 설정되어 있을수록 높은 위험도를 가지는데 읽기, 쓰기, 실행 권한의 순으로 높은 위험도를 주었다. 특히 공격의 주 목표가 되는 SETUID의 권한을 가지는 파일의 경우 더 높은 위험도를 주어 위험한 파일 접근을 잘 구분 할 수 있게 하였다.

3.4 다중 모델 결합 방법

침입탐지시스템의 성능을 향상시키기 위해서는 각 탐지방법별 결과를 적절히 결합하여 탐지방법의 결과들이 서로 다르더라도 적절한 결과가 나올 수 있도록 해야 한다. 앞에서 언급한 탐지방법들의 취약점을 보완하여 탐지 가능 범위를 넓히고 오류율을 줄이기 위해 본 논문에서는 규칙기반의 다중 모델 결합방법을 제안하였다. 결합에 사용되는 규칙은 각 척도와 모델링 방법의 특징과 공격 유형을 고려하여 경험적으로 생성하였다. 논문에서 사용한 규칙은 다음과 같다.

IF ((시스템호출 평가값 < 임계값) AND (파일접근 위험도 > 임계값))
THEN Buffer Overflow 공격

IF ((이벤트수 평가값 > 임계값) AND (자원사용량 평가값 < 임계값))
THEN Process 채우기 DoS공격

IF ((자원사용량 평가값 > 임계값) AND (이벤트수 평가값 < 임계값))
THEN Memory, Disk 채우기 DoS공격

첫 번째 규칙에서는 시스템 호출 척도를 이용한 HMM방법과 파일 접근 이벤트를 이용한 규칙기반 방법을 결합하였다. 시스템 호출 시퀀스의 HMM평가 값이 임계값보다 낮아 침입으로 판정되더라도 파일 접근 위험도가 낮을 경우 침입으로 판정하지 않는 방법이다. 대부분의 버퍼오버플로우 침입은 루트 소유이거나 SETUID 권한이 있는 파일에 접근하기 때문에 이 규칙을 적용하여 오류율을 낮출 수 있다.

두 번째는 시스템 호출 척도를 이용한 통계적 방법과 자원사용량을 이용한 통계적 방법을 결합하는 규칙이다. 시스템 호출을 이용한 통계적 방법은 많은 시스템 호출을 사용하는 정상적인 프로세스도 침입으로 간주하는 오류를 범하기 쉽다. 과도한 양의 시스템 호출로 시스템을 마비시키는 유형의 공격은 많은 시스템 호출 이벤트 수에 비해 자원사용량이 매우 적은 경향을 보이는데 이에 착안하여 시스템 호출 수 평가값이 높지만 자원사용량이 낮은 프로세스만을 침입으로 판단하는 방법으로 오류율을 줄이도록 하였다.

마지막 규칙에서도 시스템 호출 척도를 이용한 통계적 방법과 자원사용량을 이용한 통계적 방법을 결합하였다. 자원사용량을 이용한 방법의 경우 자원을 많이 요구하는 정상 프로세스도 침입으로 간주할 위험이 있다. 정상적인 프로세스는 많은 작업을 하므로 많은 수의 시스템 호출 이벤트를 발생시키지만 그와는 달리 서비스 거부 공격을 위한 프로세스는 높은 자원사용량에 비해 시스템 호출수

가 적다. 이를 이용하여 높은 자원 사용량이지만 비교적 시스템 호출 이벤트의 평가값이 낮은 경우만을 침입으로 판단하여 오류율을 줄일 수 있도록 하였다.

4. 실험 및 결과

4.1 실험 환경

실험에 사용한 정상행위 감사자료는 솔라리스7 운영체제 환경에서 6명의 사용자가 보름 동안에 16,470개의 명령어를 입력하여 발생시킨 160,448개의 이벤트가 기록된 13메가바이트의 BSM감사자료와 840킬로바이트의 PACCT감사자료를 사용하였다. 테스트 자료는 같은 시스템 환경에서 최근에 가장 빈번하게 일어나고 있는 유형의 침입들을 수행하여 만들었다. 9번의 버퍼오버플로우(Buffer Overflow)침입과 4번의 서비스 거부 공격을 수행하여 만들어진 자료를 사용하였는데 테스트자료에 사용된 자세한 공격유형은 표 5와 같다.

표 5 실험에 사용된 침입 유형

침입 유형	침입 형태
버퍼오버플로우	kcms_configure buffer overflow vulnerability
	lpset -r buffer overflow vulnerability
	xlock heap buffer overflow vulnerability
서비스 거부	디스크 채우기
	메모리 고갈
	프로세스테이블 채우기

4.2 성능 측정 기준

침입탐지시스템의 성능을 평가하는데 있어서 가장 중요한 척도는 탐지율과 false-positive 오류율이다. 본 논문에서는 척도별 침입탐지율 및 false-positive오류율의 계산기준을 통일하기위해 프로세스 단위로 침입을 탐지하였는데 계산 방법은 다음과 같다.

$$\text{침입탐지율} = \frac{\left(\begin{array}{c}\text{침입에 사용된 프로세스 중 침입으로}\\ \text{판단된 프로세스의 수}\end{array}\right)}{\text{침입에 사용된 프로세스의 수}}$$

$$\text{false}-\text{positive 오류율} = \frac{\left(\begin{array}{c}\text{정상행위 프로세스 침입으로}\\ \text{판단된 프로세스의 수}\end{array}\right)}{\text{정상행위 프로세스의 수}}$$

본 실험에서는 각 척도별 탐지방법과 다중척도결합 전후의 성능을 비교하기 위해 임계치를 조정해가며 false-positive 오류율과 침입탐지율을 구한 후 이에 대한 ROC (Receiver Operating Characteristics) 곡선을 그렸다. ROC 곡선은 임계값의 변화에 따른 가설 검정의 정확도와 측정오차에 대해 시각적이고 정량적인 결과를 얻는 방법으로서 침입탐지 시스템을 평가하는데 많이 사용되고 있다. 좋은 침입탐지 방법은 낮은 false-positive 오류율에서 높은 침입탐지율을 보여 주어야 하므로 그래프에서 곡선이 왼쪽 위에 있을수록 좋은 성능을 나타내는 방법이라고 할 수 있다.

또한 각 척도별 침입탐지 방법의 성능 차이와 다중척도 결합 전후의 성능비를 수치화하여 비교하기 위한 척도로 식별도(discriminability)와 효율도 개념을 사용하였다. 식별도는 신호탐지이론에서 나온 개념으로 침입과 정상행위를 얼마나 잘 구분하는지를 나타내는 척도로 침입탐지율이 높고 false-positive 오류율이 낮을수록 높은 값을 가진다[18]. 보통 d'으로 표기하며 다음과 같은 식에 의하여 계산된다.

$$d' = z(H) - z(F)$$

z: 정규분포함수의 역, H: 탐지율, F: false-positive 오류율

효율도는 결합 전후의 식별도의 비인데 다중척도 결합 전후의 성능차이를 비교하기 위하여 사용하였다. 결합전의 식별도를 d_A', 결합후의 식별도를 d_B',이라고 할 때 효율도 E는 다음과 같은 식으로 구하며 1보다 큰 값을 가질 경우 다중척도 결합방법으로 성능이 향상되었다고 할 수 있다.

$$E = \left(\frac{d'_A}{d'_B}\right)^2$$

4.3 실험 결과

4.3.1 단일척도침입탐지

먼저 각 척도별 탐지방법의 특징과 성능을 알아보기 위해 단일척도만을 사용하여 탐지를 수행하였다. 먼저 시스템호출척도를 이용한 HMM방법에 대하여 실험을 하였다. 최적의 결과를 얻기 위해 HMM의 상태수 S와 한번에 입력될 시스템 호출 이벤트 시퀀스의 길이 L을 여러 가지로 변화시켜 HMM의 구성을 바꾸어 가면서 실험을 반복하였다. 각 설정별 성능을 ROC곡선으로 나타내본 결과는 그림 3과 같다.

비교적 좋은 성능을 보이는 20%의 오류율까지의 구간에서 그래프를 비교해 보았을 때 상태수가 3이고 시퀀스 길이가 8설정의 ROC곡선이 가장 왼쪽 위에 위치하여 다른 설정 값에 비해 좋은 성능을 나타내었다. 하지만 77%이상의 높은 탐지율에서는 false-positive 오류율이 급격히 높아져 만족할 만한 성능을 보여주지 못하였는데 이는 시스템 호출 척도를 이용한 HMM방법이 프로그램의 오동작을 유도해 루트권한을 얻어내는 버퍼오버플로우 공격 이외의 침입형태는 잘 탐지하지 못하

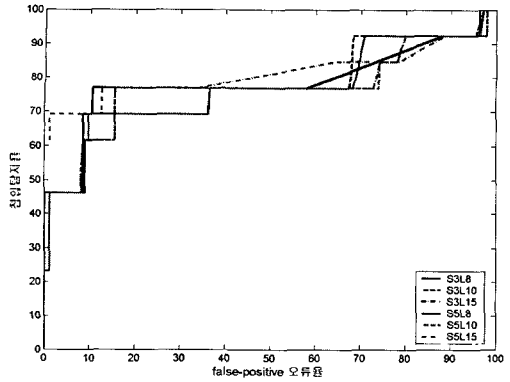

그림 3 시스템 호출 척도를 이용한 HMM방법 결과

기 때문이다. 비교적 완만하게 오류율이 증가하는 탐지율 77%에서의 각 설정별 HMM의 결과를 비교하여 본 결과는 표 6과 같다.

자원사용량을 이용한 통계적 방법의 실험에서는 각 자원사용량별 평가값의 가중치를 달리 해가며 실험해보았다. 전체 가중치의 합을 3으로 고정하고 이를 각 척도에 여러 가지 비율로 적절히 나누어 변화시키며 실험을 반복하였다. 예를 들면 각 사용량에 2:2:1의 비로 가중치를 두었을 경우 값은 각각 1.2, 1.2, 0.6이 된다. 여러 결과 중 비교적 좋은 성능을 보인 설정을 같은 침입탐지율에서 비교해 보았을 때 표 7과 같았다.

표에서와 같이 CPU사용량, 메모리점유율, I/O사용량에 2:2:1의 비로 가중치를 두는 것이 가장 높은 식별도

표 6 HMM방법의 결과 비교

상태수/시퀀스길이	임계값	침입탐지율	f-p오류율	d'
3/8	-17.6	77%	10.699%	1.979
3/10	-9.7	77%	36.626%	1.078
3/15	-16.7	77%	15.638%	1.746
5/8	-20.5	77%	12.757%	1.874
5/10	-9.7	77%	36.626%	1.078
5/15	-16.6	77%	15.637%	1.746

표 7 자원사용량을 이용한 통계적 방법 결과

가중치비율 (C:M:I)	임계값	침입탐지율	f-p오류율	d'
2:1:1	7.0	23.0%	4.527%	0.956
1:2:1	5.2	23.0%	5.350%	0.876
1:1:2	7.1	23.0%	2.881%	1.162
1:2:2	6.0	23.0%	2.881%	1.162
2:1:2	7.4	23.0%	2.469%	1.229
2:2:1	6.0	23.0%	4.527%	0.956
1:1:1	6.5	23.0%	5.761%	0.839

(d') 값이 나와 좋은 성능을 나타내었다. 하지만 23% 이상의 침입탐지율에서는 false-positive 오류율이 급격히 증가해 만족할만한 성능을 보여주지 못하였는데 이는 탐지 가능한 침입이 디스크 채우기, 메모리 고갈 등의 서비스 거부 유형의 공격으로 제한적이기 때문이다.

프로세스의 시스템호출이벤트 수를 척도로 한 통계적 방법의 침입탐지 결과는 다음 표 8과 같다. 프로세스테이블 채우기 공격의 경우는 단시간에 수많은 fork 시스템호출 이벤트가 발생하기 때문에 이 방법으로 쉽게 탐지되었다. 하지만 다른 형태의 공격의 경우는 단순히 이벤트의 수만을 보았을 때는 정상행위의 프로세스와 큰 차이를 보이지 않기 때문에 탐지가 불가능하여 성능이 좋지 않았다.

표 9는 파일접근 이벤트를 이용한 규칙기반방법의 실험 결과이다. 가장 간단한 방법이지만 자원사용량을 이용한 통계적 방법보다는 좀더 좋은 성능을 보였다. 이는 루트권한을 얻으려는 대부분의 침입이 루트나 시스템 관리자 계정의 소유이거나 SETUID권한이 설정되어 있는 등 중요한 파일을 공격대상으로 하고 있기 때문이다. 하지만 일정 침입탐지율을 넘어서면 급격히 false-positive 오류율이 높아졌는데 이는 서비스 거부 유형의 침입은 위험도가 높은 파일에 접근하지 않고도 공격의 성공이 가능하여 파일접근 이벤트 위험도로는 탐지가 힘들기 때문이다.

단일척도 침입탐지 실험결과 각 척도와 모델링 방법에 따라 탐지 가능한 침입이 다르기 때문에 하나의 척도와 모델링 방법으로는 성능이 좋지 않음을 알 수 있었다. 그림 4에서 볼 수 있듯이 척도별로 탐지가 가능한 침입을 탐지한 후로는 침입탐지율이 완만히 상승하여 한계가 있음을 알 수 있다.

표 8 시스템호출척도를 이용한 통계적 방법 결과

임계값	침입탐지율	f-p오류율	d'
30	30.77%	4.1152%	1.235
20	53.85%	7.4074%	1.543
10	76.9%	13.1687%	1.855

표 9 파일접근 이벤트를 이용한 규칙기반 방법 결과

임계값	침입탐지율	f-p오류율	d'
13	23.0%	0.000%	2.354
11	46.2%	0.000%	2.994
9	69.2%	15.226%	1.529
7	69.2%	31.687%	0.979
5	69.2%	46.091%	0.601

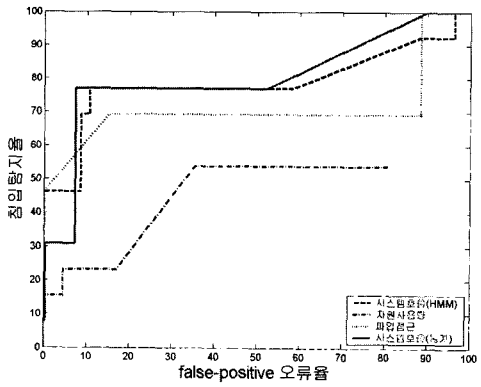

그림 4 단일척도 침입탐지 결과

4.3.2 다중척도결합 침입탐지

다음은 본 논문에서 제안한 다중척도결합 침입탐지 방법의 실험 결과이다. 각 척도별 실험 결과를 단순히 종합한 결과와 탐지 규칙을 적용한 결과를 비교하여 보았다. 결합에 사용된 모델들은 가장 좋은 침입탐지율을 보인 설정값을 사용하였는데 HMM모델에는 상태수 3, 시퀀스길이 8을 사용하였고 각 자원사용량별 가중치는 CPU사용량에 1.2, Memory점유율에 1.2, I/O사용량에 0.6을 두었다. 임계값은 시스템호출이벤트를 이용한 HMM모델은 -17.6, 자원사용량을 이용한 통계적 방법은 6.0, 시스템호출이벤트를 이용한 통계적 방법은 30, 파일접근 이벤트 위험도는 10을 사용하였다. 각 탐지방법과 다중척도 결합방법을 침입탐지율 100%에서 비교해보면 표 10과 같다.

실험 결과 단일 척도 방법에 비하여 100% 침입탐지율에서의 false-positive 오류율이 크게 개선되었다. 각 탐지방법별 다중척도 결합방법의 효율도를 비교해 보면 그림 5와 같다. 그림에서 볼 수 있듯이 효율도 값은 모두 1이상을 가져 결합전보다 성능이 향상되었음을 알 수 있다. 이는 결합규칙 적용과정에서 침입행위가 척도에 남긴 여러 흔적들을 종합하여 판단하여 각 척도별로 놓칠 수 있는 부분을 보완하였기 때문이다.

표 10 다중척도결합 결과

	침입탐지율	f-p오류율	d'
시스템 호출/HMM	100%	99.177%	1.866
자원사용량/통계적 방법	100%	80.658%	3.400
시스템 호출/통계적 방법	100%	99.177%	1.866
파일접근이벤트/규칙기반 방법	100%	88.477%	3.066
결합규칙적용	100%	5.761%	4.665

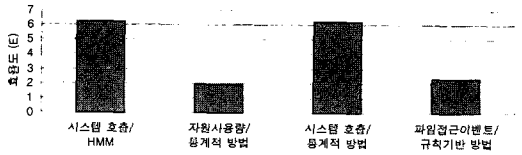

그림 5 각 방법별 효율도 비교

5. 결론 및 향후연구

본 논문에서는 비정상행위탐지기법의 침입탐지시스템에 사용될 수 있는 세 가지 감사자료에 대한 효과적인 모델링 방법을 제시하고 단일척도 침입탐지방법의 취약점을 극복하기 위하여 다중척도 결합 침입방법을 제안하였다. 제안한 침입탐지 방법은 각 단일척도방법 침입탐지 방법의 결과를 규칙기반 방법으로 결합하여 최종적으로 침입을 판단한다.

실험은 각 단일척도 침입탐지방법의 성능을 측정한 후 다중척도결합방법과 비교 평가하였다. 그 결과 단일척도 침입탐지방법은 척도와 모델링 방법의 특성에 따라 탐지할 수 있는 침입의 범위가 한정되어 있어 일정수준이상의 성능을 기대할 수 없음을 확인하였다. 각 탐지방법과 다중척도 결합방법을 비교하였을 때 100%침입탐지시 false-positive 오류율이 5.761%로 매우 낮아져 규칙기반 결합방법으로 탐지방법별 취약점이 보완되어 성능이 향상될 수 있음을 보였다.

하지만 규칙기반 결합방법은 사용자의 경험에 의해 결정되기 때문에 침입행위가 남긴 모든 흔적을 결합에 반영했는지 알 수 없어 사용된 규칙이 최적의 것이라 보장할 수 없고 규칙에 명시되어 있지 않은 침입유형은 탐지가 힘든 단점이 있다. 따라서 좀 더 체계적인 결합방법이 필요한데 의사결정나무, 인공 신경망등의 기계학습 방법을 결합에 사용하면 사용자가 결합 규칙을 지정해주는 단계를 거치지 않고도 좀더 좋은 성능을 끌어낼 수 있을 것이다. 또한 다양한 유형의 침입을 탐지할 수 있도록 본 논문에서 사용된 척도와 모델링 방법이외의 것을 이용한 침입탐지 방법의 개발을 지속적으로 할 필요가 있다. 다른 문제점으로는 침입탐지에 요구되는 계산량이 기존의 탐지 방법과 비교해 보았을 때 많기 때문에 탐지가 지연될 수 있다는 것이 있다. 이 부분은 성능향상에 따른 기회비용으로 최적화된 구현 방안등의 해결책이 향후 연구되어야 할 것이다.

참 고 문 헌

[1] CERTCC-KR, 한국정보보호진흥원, http://www.certcc.

or.kr, 2002.
[2] H.S. Vaccaro and G.E. Liepins, "Detection of anomalous computer session activity," *In Proceedings of IEEE Symposium on Research in Security and Privacy*, pp. 280-289, 1989.
[3] T. F. Lunt, "A survey of intrusion detection techniques," *Computers & Security*, vol. 12, no. 4, pp. 405-418, June 1993.
[4] C. Dowel and P. Ramstedt. "The computer watch data reduction tool," *In Proceedings of the 13th National Computer Security Conference*, pp. 99-108, Washington DC, USA, October 1990.
[5] T. Heberlein, G. Dias, K. Levitt, B. Mukherjee, J. Wood, and D. Wolber. "A network security monitor," *In Proceedings of the 1990 IEEE Symposium on Research in Security and Privacy*, pp. 296-304, Los Alamitos, CA, USA, 1990.
[6] S. Staniford-Chen, S. Cheung, R. Crawford, M. Dilger, J. Frank, J. Hoagland, K Levitt, C. Wee, R. Yip, and D. Zerkle. "GrIDS-A graph based intrusion detection system for large networks," *In Proceedings of the 19th National Information Systems Security Conference*, vol. 1, pp. 361-370, October, 1996.
[7] T. F. Lunt, A. Tamaru, F. Gilham, R. Jagannathan, C. Jalali, and P. G. Neuman. "A real-time intrusion-detection expert system (IDES)," *Technical Report Project 6784*, CSL, SRI International, Computer Science Laboratory, SRI International, February 1992.
[8] Anderson, Lunt, Javits, Tamaru and Valdes, "Detecting unusual program behavior using the statistical components of NIDES," *NIDES Technical Report*, SRI International, May 1995.
[9] P. A. Porras and P. G. Neumann. "EMERALD: Event monitoring enabling responses to anomalous live disturbances," *In Proceedings of the 20th National Information Systems Security Conference*, pp. 353-365, Baltimore, Maryland, USA, October 1997.
[10] H. Debar, M. Becker and D. Siboni, "A neural network component for an intrusion detection system," *In Proceedings of 1992 IEEE Computer Society Symposium on Research in Security and Privacy*, pp. 240-250, Oakland, CA, May 1992.
[11] C. Warrender, S. Forrest and B. Pearlmutter, "Detecting intrusion using calls: Alternative data models," *In Proceedings of IEEE Symposium on Security and Privacy*, pp. 133-145, May 1999.
[12] 박혁장, 조성배, "권한이동 모델링을 통한 은닉 마르코프 모델 기반 침입탐지 시스템의 성능 향상", *정보과학회 논문지: 정보통신*, 제29권 제6호, pp. 674-684, 2002.
[13] 최종호, 조성배, "침입탐지 시스템을 위한 은닉 마르코프 모델의 적용", *정보과학회 논문지: 소프트웨어 및 응용*, 제28권, 제6호, pp. 429-438, June 2001.
[14] L.R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257-286, 1989.
[15] J. Choy and S.-B. Cho, "Intrusion detection by combining multiple hidden Markov models," *Lecture Notes in Artificial Intelligence*, vol. 1886, pp. 829, 2000.
[16] H.S. Javitz and A. Valdes, "The SRI IDES statistical anomaly detector," *NIDES Technical Report*, SRI International, 1991.
[17] A.G. Amoroso, *Fundamentals of Computer Security Technology*, PTR Prentice Hall, New Jersy, 1994.
[18] N. A. Macmillan and C. D. Creelman, *Detection Theory : A User's Guide*, Cambridge University Press, Cambridge, 1991.

한 상 준
2002년 8월 연세대학교 컴퓨터과학과(학사). 2002년 9월~현재 연세대학교 컴퓨터과학과 석사과정 재학중. 관심분야는 인공지능, 침입탐지, 컴퓨터보안

조 성 배
1988년 연세대학교 전산과학과(학사) 1990년 한국과학기술원 전산학과(석사) 1993년 한국과학기술원 전산학과(박사) 1993년~1995년 일본 ATR 인간정보통신연구소 객원 연구원. 1998년 호주 Univ. of New South Wales 초청연구원. 1995년~현재 연세대학교 컴퓨터과학과 부교수. 관심분야는 신경망, 패턴인식, 지능정보처리