SVM을 이용한 침입방지시스템 오경보 최소화 기법☆

False Alarm Minimization Technology using SVM in Intrusion Prevention System

김 길 한* 이 형 우**
Kim, Gill-Han Lee, Hyung-Woo

요 약

지금까지 잘 알려진 네트워크 기반 보안 기법들은 공격에 수동적이고 우회한 공격이 가능하다는 취약점을 가지고 있어 인라인(in_line) 모드의 공격에 능동적 대응이 가능한 오용탐지 기반의 침입방지시스템의 출현이 불가피하다. 하지만 오용탐지 기반의 침입방지시스템은 탐지 규칙에 비례하여 과도한 오경보(False Alarm)를 발생시켜 정상적인 네트워크 흐름을 방해하는 잘못된 대응으로 이어질 수 있어 기존 침입탐지시스템보다 더 위험한 문제점을 갖고 있으며, 새로운 변형 공격에 대한 탐지가 미흡하다는 단점이 있다. 본 논문에서는 이러한 문제를 보완하기 위해 오용탐지 기반의 침입방지시스템과 Anomaly System 중의 하나인 서포트 벡터 머신(Support Vector Machines; 이하 SVM)을 이용한 침입방지시스템 기술을 제안한다. 침입방지시스템의 탐지 패턴을 SVM을 이용하여 진성경보만을 처리하는 기법으로 실험결과 기존 침입방지시스템과 비교하여, 약 20% 개선된 성능결과를 보였으며, 제안한 침입방지시스템 기법을 통하여 오탐지를 최소화하고 새로운 변종 공격에 대해서도 효과적으로 탐지 가능함을 보였다.☆

Abstract

The network based security techniques well-known until now have week points to be passive in attacks and susceptible to roundabout attacks so that the misuse detection based intrusion prevention system which enables positive correspondence to the attacks of inline mode are used widely. But because the Misuse detection based Intrusion prevention system is proportional to the detection rules, it causes excessive false alarm and is linked to wrong correspondence which prevents the regular network flow and is insufficient to detect transformed attacks. This study suggests an Intrusion prevention system which uses Support Vector machines(hereinafter referred to as SVM) as one of rule based Intrusion prevention system and Anomaly System in order to supplement these problems. When this compared with existing intrusion prevention system, show performance result that improve about 20% and could through intrusion prevention system that propose false positive minimize and know that can detect effectively about new variant attack.

☞ Keyword : Misuse detection, Intrusion Prevention, False Alarm, multi class Support Vector machines(SVM), Machine Learning

1. 서 론

오늘날 컴퓨터와 통신 기술의 급속한 진전으로 인터넷이용률 및 이용자 수는 매년 증가하고 있다. 하지만 이러한 인터넷 대중화와 더불어 비인가된 사용자에 의한 컴퓨터 자원의 무결성(integrity), 기밀성(confidentiality), 가용성(availability)을 저해하는 일련의 행동들과 보안 정책을 파괴하는 행위들이 나날이 증가하고 있어, 침입에 의한 네트워크에서의 시스템 자원과 데이터를 보호하기 위한 여러 가지 침입탐지/차단 시스템의 설계가 필요하다.

네트워크 보안의 1세대 솔루션인 방화벽은 허용된 룰에 대해서는 무방비한 상태이고, ID/PW 방식의 사용자 인증에 대한 취약성을 갖고 있어

* 준 회 원 : 한신대학교 컴퓨터정보소프트웨어학부 석사
gillty@empal.com (제 1저자)
** 종신회원 : 한신대학교 컴퓨터정보소프트웨어학부 부교수
hwlee@hs.ac.kr (제 2저자)
[2005/12/19 투고 - 2006/01/12 심사 - 2006/04/04 심사완료]
☆ 본 연구는 정보통신부 및 정보통신연구진흥원의 대학 IT연구센터 지원사업의 연구결과로 수행되었음(IITA-2005-(C1090-0502-0020))
☆ 학술진흥재단 지역대학우수과학자지원과제 KRF-2005-202-D00487의 지원을 통해 수행됨

백도어 등의 설치 시 암호 노출의 가능성이 큰 단점을 갖고 있다. 반면 2세대 솔루션인 침입탐지시스템(Intrusion Detection System 이하 IDS)은 일반적으로 방화벽 다음에 구축되어 방화벽을 우회한 공격에 대해 분석, 탐지 기능을 제공하지만, 단순한 CCTV의 역할을 수행하는 것으로 발견과 동시에 차단 등의 대처능력에는 한계가 있다. 따라서 이들의 취약성을 해결하고 공격에 보다 능동적인 적절한 대처를 통해 그 피해를 최소화하는 네트워크 보안기술인 침입방지시스템(Intrusion Prevention System 이하 IPS)의 등장이 불가피해졌다[1].

IPS는 공격 탐지에 기초하여 능동적으로 트래픽의 통과 여부에 대하여 결정을 내릴 수 있는 인라인 장치이며, IDS와 마찬가지로 데이터 소스에 따라 호스트 기반 IPS와 네트워크 기반의 IPS로 나눠진다. 또한 탐지 모델에 따라 오용탐지 기반 IPS와 비정상행위 기반 IPS로 구분된다. 오용탐지를 기반 IPS는 전문가에 의해서 만들어진 규칙을 이용하여 공격을 탐지하여 대응하는 것으로 탐지 비율이 높아 널리 사용되고 있지만, 변형 공격이나 규칙에 잡히지 않는 공격들에 대한 탐지 및 대응에는 취약한 단점이 있으며, 규칙 증가에 비례하여 정상행위를 공격으로 탐지하는 오탐지(false positive) 비율이 기하급수적으로 증가한다는 문제점이 있다. 이러한 문제들은 IDS에서 보다 능동적 대응이 가능한 IPS에서 정상/공격패킷에 대한 잘못된 대응으로 이어져 IPS로 인한 정상적인 서비스 흐름 방해라는 큰 위험을 가지고 있기 때문에 오경보(false alarm: false positive, false negative) 제거 문제는 IPS 성능 향상의 중요한 문제가 되고 있다[2]. 반면 비정상행위 기반 IPS는 전문가의 지식에 의존하지 않고 기계학습이나 데이터마이닝 기법들을 이용하여 사용자의 패턴을 분석하고 입력패턴과 비교하여 정해진 모듈에 벗어나는 경우를 공격으로 판단하는 것으로 알려지지 않은 공격 탐지에 유연성을 보이지만, 탐지율이 낮다는 단점을 갖는다[3]. 따라서 본 논문에서는 높은 공격 탐지율 갖는 오용탐지 기반의 IPS와 '정상' 모형화 기능을 갖는 멀티 클래스(multi class) SVM을 이용한 학습기반의 IPS의 서로 다른 특성을 상호 보완적으로 적용하여 오용탐지 기반 IPS의 공격 탐지 비율을 손상하지 않으면서 오경보를 최소화하는 방안을 제안한다.

본 논문의 2장에서는 오경보를 줄이기 위해 사용되고 있는 기존연구들을 소개하였고, 3장에서는 본 논문에서 학습도구로 사용하는 SVM에 대해 설명하였다. 4장에서는 본 논문에서 제안하고자 하는 SVM을 이용한 침입방지 기법을 기술하였고, 5장에서는 제안된 모델의 실험 및 결과 분석에 대해 설명하였다. 마지막으로 6장에서 결론 및 향후 연구에 대해 논의하였다.

2. 오경보 최소화를 위한 기존 연구

네트워크 보안시스템에서의 오경보(false alarm)란 정상인데 공격이라고 탐지하는 오탐지(false positive) 및 공격을 정상이라고 판단하는 미탐지(false negative)의 경우를 가리킨다. 이러한 오경보는 보안시스템에 불필요한 대응을 요구하게 되어 자원을 낭비하게 되고 보안 시스템에 대한 신뢰도를 떨어트리는 중요한 요소이기 때문에 오경보를 줄이기 위해 아래와 같은 연구들이 진행되어 오고 있다.

2.1 시스템 환경설정 변경

(1) 침입 패턴 개선

정상 패킷을 공격으로 오인하는 경우는 원초적으로 침입을 판단하는 침입 패턴이 잘못 되었기 때문이다. 따라서 오탐지를 줄이기 위해서는 시그니처를 정확하게 만들어야 한다. 모든 상황에서 정확한 침입을 탐지하는 오탐지를 만드는

것은 어렵다. 따라서 오탐지가 발생했을 경우 이를 개발자에게 알려주어 개발자들이 시그니처를 개선하여 오탐지를 줄이는 방법이다[4].

(2) 시스템 네트워크 상황과 정책에 맞게 설정

개발자가 아닌 시스템 관리자에 의해 자신의 네트워크와 상황에 맞게 시그니처들을 튜닝 함으로써 오탐지를 줄이는 방법이다. 침입에 대한 룰을 default로 설정하는 것은 정상 패킷을 해킹으로 오인할 수 있기 때문에 관리자는 네트워크 상황에 맞게 몇 번이고 시행의 오류를 겪으면서 관리하고자 하는 네트워크 설정에 맞게 시그니처를 변경해줘야 한다.

또한 기업이나 조직의 보호정책에 맞게 IPS의 환경 설정함으로써 정책에 맞지 않는 시그니처들을 무력하게 함으로써 오탐지를 줄인다.

2.2 데이터 마이닝 기법

데이터 마이닝 기법은 많은 양의 데이터로부터 알려지지 않은 유용한 지식을 추출해내는 기술로써 데이터 필터링 효과도 있으며 알려지지 않은 공격에 대한 시퀀스추출에 활용 가능하다. 이를 이용한 지식기반의 선행필터로 하여금 중요한 알람(alarm)만 취급하게 하는 것으로 실제 시스템에서의 30% 정도의 오탐지 감소가 보고되고 있다[5,6].

2.3 상관관계분석 기법

상관관계분석 기법은 다른 정보보호 시스템을 설치함으로써 이로부터 나오는 정보를 이용하여 오탐지를 줄이는 방법으로 각 보호대상 시스템의 취약점을 DB화 시켜 둠으로써 경고가 발생하면 그 경고와 자산 DB로 이동하여 취약성을 검색한 다음 동일한 취약점들이 있다면 그 경고를 출력한다. 이는 선택된 속성에 의존적이며 경보데이터 간의 인과 관계를 완벽하게 탐지하기에 적당하지 못하다는 단점을 가진다[7,8].

2.4 행동기반 분석 기법

오용탐지 기반의 IDS가 경고를 발행시켰을 때 이것을 학습 또는 클러스터링 방법을 이용하여 관리자에게 통보할 가치가 있는가를 재차 판단하는 방법으로 학습 데이터로 상용되는 패킷의 정보를 그대로 사용하기 때문에 정보 손실이 적으며 자신의 네트워크 환경에 맞는 적절한 필터링 모듈을 학습시킬 수 있는 기법이다. 지금까지 알려진 연구에서는 학습예제(instance)를 그대로 사용하는 예제기반학습(Instance Based Learning)을 학습기법으로 사용하였다. 하지만 이는 입력되는 새 데이터의 분류를 위해서는 모든 학습 데이터들과의 유사도(similarity) 계산이 필요하기 때문에 비용이 크며, 새로운 공격에 대한 대처 능력이 없다는 단점을 갖는다[9].

2.5 기존 연구의 문제점

오탐지(False positive) 및 미탐지(False negative)를 낮추기 위한 기존의 연구들은 시스템 관리 및 구조상의 원천적 취약성, 새로운 공격에 대한 대처 능력이 미흡 등 표 1과 같은 문제점들을 가지고 있다. 따라서 본 논문에서는 SVM을 이용한 행동기반 분석기법으로 학습 자료의 손실없이 새로운 공격에 대한 탐지가 가능한 IPS 모델을 제안하고자 한다. 기존의 Anomaly Detection System에서 주제에 따른(공격과 공격이 아닌 것 특정 공격(DoS, FFRR, U2Su, R2L, ... etc)과 특정 공격이 아닌 것) 분류를 위한 이진 SVM을 사용[10-12]한 연구 사례들이 있지만, 본 논문에서는 이와는 다르게 오용탐지기반 IPS의 4가지 탐지 결과(공격을 공격이라고 판단 : true positive, 정상을 공격이라고 판단 : false positive, 정상을 정상이라 판단 : true negative,

공격을 정상이라고 판단 : false negative)들 가운데 이진 SVM의 조합인 멀티 클래스 SVM 설계를 통하여 오용탐지기반 IPS의 오경보 패턴 분석 및 재판단 하고자 한다.

〈표 1〉 기존연구의 문제점

환경설정 변경	네트워크 환경이 자주 바뀌는 상황에서는 관리가 어려우며 취약성 정도가 다른 개별 시스템을 보호하는 경우 이용이 어렵다.
데이터 마이닝기법	자주 발생하지 않는 공격을 무시하는 등 그 구조상 정보 손실을 피할 수 없어 비정상인 식도의 손상 및 새로운 공격에 대한 대처 능력이 미흡하다.
상관관계 분석기법	선택된 속성에 의존적이며 경보 데이터간의 인과 관계를 완벽하게 탐사하기에 적당하지 못하다.
행동기반 분석기법	계산 비용이 많이 소요되고 데이터에 의존적이며, 새로운 공격에 대한 대처 능력이 없다.

3. Support Vector Machine(SVM)

SVM(Support Vector Machine)은 1995년 Vapnik에 의하여 개발되고 제안된 학습 알고리즘이다. 전통적인 학습 알고리즘들이 학습 집단의 학습오류(empirical error)를 최소화하기위한 경험적 위험을 최소화(Empirical Risk Minimization: ERM)하는 것에 기초한 반면, SVM은 고정되어 있지만 알려지지 않은 확률분포를 갖는 데이터에 대해 잘못 분류하는 확률을 최소화하기위해 구조적 위험을 최소화(Structural Risk Minimization: SRM)하는 것에 기초하고 있다[13].

3.1 선형 SVM -분리가 가능한 경우

SVM 의 목적은 학습 자료로 주어지는 두개의 부류를 구분하는 함수를 추정하는 것으로, SVM 학습은 두 부류간에 모든 점들 사이의 거리를 최대화하도록 제한을 두는 선형 평면 분류 경계(optimal separating hyperplane: OSH)를 찾는 과정이라 할 수 있다.

두 부류에 속하는 학습 벡터의 집합을 선형적으로 분리 가능하도록 하는 문제를 생각해 보면, 가중치 벡터 w와 바이어스 b로 구성되는 $(w_o^T \cdot x)+b_o=0$의 초월면(hyperplane)을 가지도록 훈련 데이터 셋(training data set) $\{(x_i,d_i)\}_{i=1}^N$를 학습시키는 것을 나타내며, 여기서 x_i는 입력 패턴이고, d_i는 목표값이 된다. 초월면 $(w_o^T \cdot x)+b_o=0$는 식 (1)의 조건을 만족하게 된다.

$$\exists w,b \ s.t. \begin{cases} w^T x_i + b > 0 \ \text{for} \ d_i = +1 \\ w^T x_i + b < 0 \ \text{for} \ d_i = -1 \end{cases} \quad (1)$$

식 (1)에서 등호의 조건을 만족하는 입력패턴들 중에서 결정 표면(decision surface)에 가장 가까이 위치한 패턴들을 support vector라고 하며, 개념적으로 이 벡터들은 초월면에 가장 가까이 위치하여 분류하기가 어려운 벡터들이다. 따라서 분류를 위한 학습은 제약조건 식 (2)을 만족하는 최적의 초월면을 찾는 것이다. 이것은 제약조건을 가지는 최적화 문제로 훈련 데이터 셋 $\{(x_i,d_i)\}_{i=1}^N$이 주어질 때 최적의 초월면을 위한 최적의 파라미터 w와 b를 찾는 Quadratic 문제이다.

$$\begin{cases} \text{minimize} \ \Phi(\mathrm{w}) = \frac{1}{2}\|\mathrm{w}\|^2 \\ s.t. \ d_i(w^T x_i + b) \geq 1 \quad \text{for} \quad i = 1,...,N \end{cases} \quad (2)$$

여기서 최적은 최대 마진(margin)을 가지는 것이며, 최대 마진 초월면은 최적으로 두 개의 클래스를 분리할 수 있는 초월면이다. 결국 최적의 선형 분리 경계면을 $g(x)=w_o^T \cdot x+b_o$로 놓으면, support vector와 g(x)의 거리를 $1/\|w\|$로 나타낼 수 있으며, 입력패턴을 최적으로 분류하는 초월면은 식 (3)과 같이 비용함수 $\Phi(w)$를

최소화한다.

$$\Phi(w) = \frac{1}{2}\|w\|^2 \tag{3}$$

식 (3)의 비용함수는 w의 블록함수이며, 제약조건 식 (2)는 w에 선형임을 확인할 수 있다. 지금까지 서술된 분류를 위한 SVM을 정리하면, 학습 패턴이 주어질 때 제약조건 식 (2)를 만족하는 가중치 벡터 w와 바이어스 b를 찾는 최적화 문제로 생각할 수 있으며, 이때 $\|w\|^2$을 최소화하여 분리 간격을 최대화하도록 하여 최적 분리면을 찾아낸다. 이 최적화 문제를 해결하기 위하여 라그랑제(Lagrange) 계수법을 이용하면 식 (4)과 같은 라그랑제 배수로써 쌍대화시키면 아래의 Quadratic 문제가 된다.

$$\Theta(\alpha) = \sum_{i=1}^{N}\alpha_i - \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j d_i d_j < x_i, x_j > \tag{4}$$
$$\text{s.t. } \alpha_i \geq 0,\ i = 1, \cdots, N \text{ and } \sum_{i=1}^{N}\alpha_i d_i = 0$$

3.2 선형 SVM - 분리 불가능 경우

학습표본이 선형 분리 초월면에 의해 완전히 두 개의 범주로 분리되지 않을 경우 오분류를 허용하는 것이 불가피하다. 이를 위한 최적화 문제는 여유변수(slack variable: ξ)이 추가되어 다음과 같이 표현된다.

$$\text{minimize } \tau(w,\xi) = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N}\xi_i \tag{5}$$
$$\text{s.t. } d_i\,(<w, x_i> + b) \geq 1 - \xi_i,\ \xi_i \geq 0,\ i = 1, \cdots, N$$

즉 두 개의 평행한 초월면의 마진에 객체가 놓이는 것을 허용하되 이에 대하여 페널티 파라메타 : C값을 주는 것이다.

C는 non-separable 데이터에 대한 페널티로 작용하는 변수로서 모델 복잡성과 trade off 관계에 있다. 즉, C가 커지면 학습된 분류기는 최적의 초월면을 구성하는 해결책을 제공하는 경향이 있으며, C가 0으로 수렴하는 값일 경우 마진 maximization term을 최적화하려는 효과를 제공하게 되며, 그 결과 misclassification error를 최소화하는 term에는 그다지 큰 중점을 두지 않음으로 인해 마진 width가 아주 큰 SVM 분류기를 생성해 내게 된다.

3.3 비선형 SVM

일반적인 경우 대부분의 패턴은 선형적으로 분리가 가능하지 않다. 따라서 비선형 패턴을 분리하기 위하여 비선형 패턴의 입력 공간을 선형 패턴의 특정 공간으로 전환한다.

$$\theta(\alpha) = \sum_i \alpha_i^N - \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j d_i d_j K(x_i, x_j) \tag{6}$$
$$\text{s.t.} \sum_i^N \alpha_i d_i = 0,\ 0 \leq \alpha_i \leq C,\ \forall i.$$

위의 모델에서 라그랑제 배수 i를 구하면 특징 공간에서 가장 평평한 함수인 아래의 식(7)을 구할 수 있다.

$$f(x) = \text{sgn}(<w, \phi(x)> + b) \tag{7}$$
$$= \text{sgn}\left(\sum_{i=1}^{N}\alpha_i d_i K(x_i, x) + b\right)$$

SVM에는 비선형 mapping function을 지원하기 위한 표 2와 같은 커널이 지원된다.

4. SVM을 이용한 침입방지시스템

4.1 SVM 기반 침입방지시스템 구조

본 논문에서는 제안하고자 하는 침입방지시스템은 오용탐지 기반 IPS(Snort_inline)의 오탐지(false positive) 및 미탐지(false negative)를

〈표 2〉 커널함수 종류

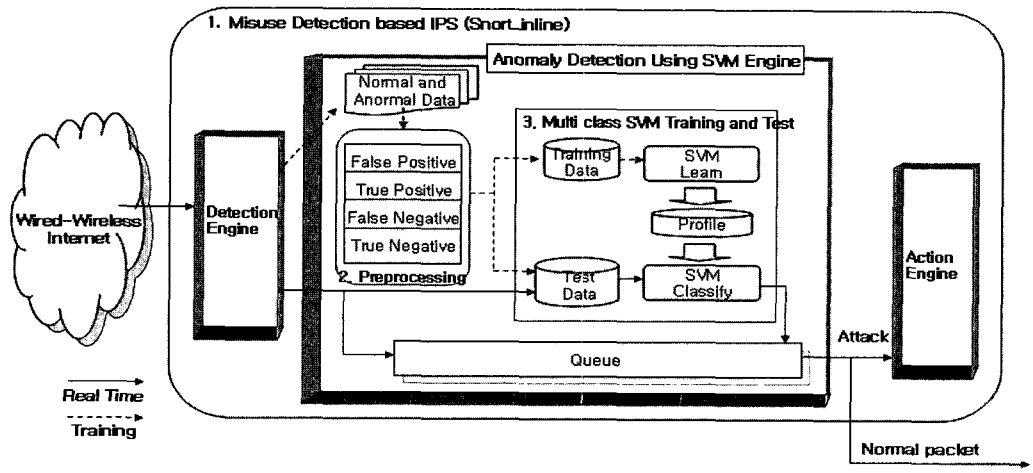
커널 종류	커 널	비 고
Dot kernel	$x \bullet y$	x와 y의 내적
Polynomial kernel	$(x \bullet y+1)^d$	d=1,2,3...
RBF(Radial Basis Function) kernel	$\exp(-g \parallel x - y \parallel^2)$	g는 커널형태를 결정하는 모수
Perceptron kernel	tanh(ax * y + b)	a,b는 Mercer 조건 만족하는 상수

최소화하기위해 규칙 기반으로 탐지된 결과들을 4가지 클래스(공격을 공격이라고 판단 : true positive, 정상을 공격이라고 판단 : false positive, 정상을 정상이라 판단 : true negative, 공격을 정상이라고 판단 : false negative)로 분류하여 학습시킨 멀티 클래스 SVM 모듈로 구성된다. 본 시스템 구조는 아래 그림 1과 같다.

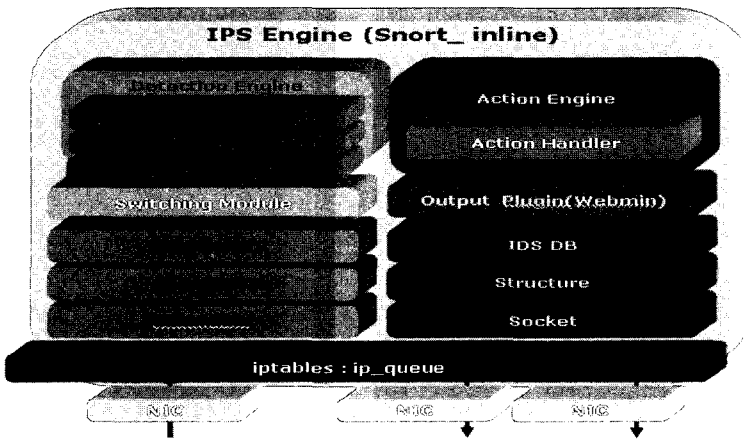
IPS의 네트워크 패킷 수집은 Promiscuous mode의 Snort_inline에서 이뤄지며 탐지엔진에서 룰에 의해 이상트래픽인지 아닌지 판단된다. 이렇게 탐지된 패킷들에는 오탐지 및 미탐지 패킷이 포함되어 있기 때문에 멀티 클래스 분류를 위한 SVM 모듈에서 오프라인으로 학습된 4가지 클래스의 탐지 유형별로 다시 여과되어 True Positive, False Negative로 판단된 공격 패킷만 이 액션 엔진에서 공격에 맞는 액션으로 보다 안전하게 처리된다.

4.2 오용탐지 기반의 IPS(Snort_inline)

Snort_inline은 침입탐지시스템으로 잘 알려진 Snort의 패치 버전 중의 하나로 IDS 게이트웨이 또는 NIPS(네트워크 침입 방지 시스템)으로 사용되고 있다. Snort_inline은 공격 패킷에 대한 룰 파일을 가지고 이상 패킷을 검사하며, 패킷을 계속 라우팅하게 하는 약간의 메커니즘들이 필요하기 때문에 IPTables와 더불어 사용하게 된다. 패킷들은 IPTables에 의해 커널 공간으로부터 얻을 수 있으며, ip_queue 기능을 통해 Snort_inline으로 넘겨지게 된다[14]. 이는 네트



〈그림 1〉 SVM 기반 침입방지시스템 구조



〈그림 2〉 Snort_inline Architecture

워크상에 지나다니는 패킷을 랜덤하게 캡쳐하여 침입을 탐지하던 Snort에서의 단점을 보안한 것으로 실시간으로 패킷을 유지 및 제어할 수 있다는 장점을 갖는다. 본 논문에서의 시스템 구조는 아래 그림 2와 같다.

4.3 Preprocess Data Set

(1) DARPA 1998년 Data Set

본 논문에서 사용하는 DARPA 1998년 Data는 MIT Lincoln Labs에서 침입탐지 개발 프로그램을 계획하면서 준비된 데이터로 인위적으로 실험용 군사 네트워크를 구성하여 실험되어진 방대하고 다양한 침입들을 포함하고 있는 표준 감사 데이터집합(data set)이다.

DARPA 데이터 집합은 1998년부터 2000년까지, 매년 약 7 주간, 그리고 매주 5 일간(월요일 ~ 금요일)을 일별로 Solaris 기반 BSM 감사 데이터와 tcpdump 데이터를 침입탐지시스템의 개발과 평가를 위한 데이터로 제공하고 있다. tcpdump 데이터 자체는 단순하게 네트워크상의 지나다니는 패킷을 저장한 것으로 tcpdump 데이터만으로는 공격 및 정상 패킷을 구분하는 것이 불가능 하다. 따라서 각각의 tcpdump 데이터에 포함된 공격여부를 표시한 별도의 list 파일이 함께 제공된다. 또한 1998년 Data Set에서 포함되어 있는 공격의 종류는 27종류로 크게 4가지(DoS, Probing, R2L, U2R) 공격 유형으로 분류된다.

(2) SVM 학습 Data Set 변환

1998년 DARPA에서 제공되는 tcpdump 원시 데이터를 SVM 학습 알고리즘에 training/test data로 적용하고, 보다 정확한 분류 결과값을 도출하기 위해서는 학습데이터로 사용되는 Packet의 Feature 선택이 가장 중요하다. 본 논문에서는 패킷 단위의 행동모형을 적용하기 위해서 tcpdump형식(-vv 옵션 사용)의 패킷 하나에 학습 자료 하나를 생성하는 방법을 사용하였으며 한 개의 학습 자료를 구성하는 feature 집합은 아래 표 3과 같이 패킷의 헤더범위 내의 모든 값을 추출한 26개의 feature를 선택하였다.

4.4 SVM 학습모델

오용탐지 기반 IPS의 오탐지 및 미탐지 최소화를 위한 SVM 학습모델은 Snort_inline에 의한 4가지 탐지유형(TP, FP, TN, FN)을 분류하

〈표 3〉 SVM 학습을 위한 Packet Features 구성

Index	Feature	Type value
label		+1, -1
1	ToS	Real
2	TTL	Real
3	IP_ID	Real
4	Offset	Real
5	Fragment	DF, MF etc
6	Protocol	TCP, UDP, ICMP, IGMP, EGP, OSPF, IGRP, GRE, etc
7	IP_Length	Real
8	Source IP Address	DARPA 네트워크 고려
9	Source Port	Echo(7), Discard(9), Users(11), Quote(17), Nameserver(53), Bootps(67), Bootpc(68), TFTP(69), RPC(111), NTP(123), SNMP(161), SNMP_trap(162), FTP_data(20), FTP(21), TELNET(23), SMTP(25), DNS(53), Finger(79), HTTP(80). Rlogin(513), Rsh(514), etc
10	Destination IP Address	DARPA 네트워크 고려
11	Destination Port	Echo(7), Discard(9), Users(11), Quote(17), Nameserver(53), Bootps(67), Bootpc(68), TFTP(69), RPC(111), NTP(123), SNMP(161), SNMP_trap(162), FTP_data(20), FTP(21), TELNET(23), SMTP(25), DNS(53), Finger(79), HTTP(80). Rlogin(513), Rsh(514), imap(143), ssh(22), etc
12	DgmLength	Real
13	TCP_Length	Real
14	Sequence number	Real
15	Acknowledgement	Real
16	UGR	0,1
17	ACK	0,1
18	EOM	0,1
19	RST	0,1
20	SYN	0,1
21	FIN	0,1
22	Window Size	Real
23	UDP_Length	Real
24	ICMP_Type	Real
25	ICMP_Code	Real
26	ICMP_Length	Real

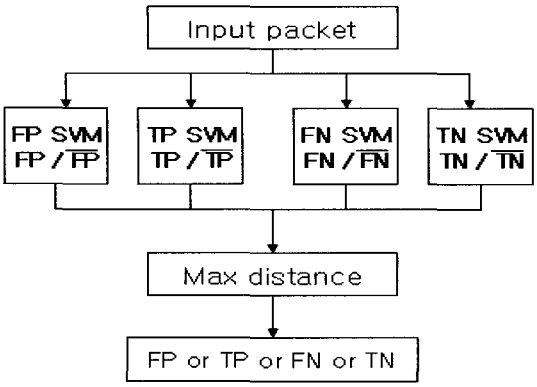
기 위하여 SVM에 적용된다. 하지만 기본적으로 이진분류를 수행하는 SVM을 이용하여 4가지 클래스를 가지는 본 연구에 적용하기 위해서는 이진분류 조합에 의한 전략들이 제시되어져야 한다. 본 논문에서는 이진 SVM의 조합으로 구성된 one-against-al[15]l, one-against-one[16] 기법을 적용한 오경보 최소화를 위한 2가지 멀티 클래스 SVM 모델을 설계하여 실험하였다.

(1) One-Against-All 기법 적용

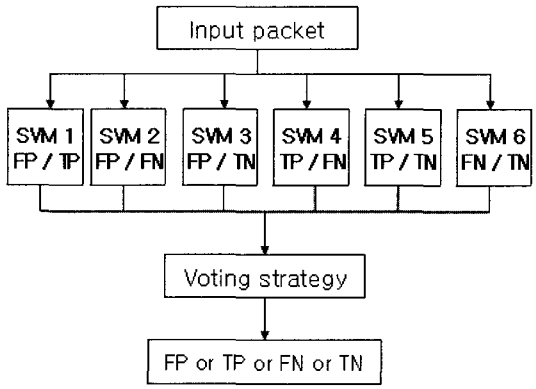
One-Against-All(OAA) 기법은 k개의 클래스를 분류하기위해 k개의 이진 SVM을 이용하는 모델로써 각 SVM에는 개별 클래스를 분류하기 위한 학습데이터가 이용된다. TP, FP, TN, FN를 구별하는 본 연구 문제에 있어서 구조도는 아래 그림 3과 같으며, 첫 번째 SVM은 FP를 구별하는 SVM으로 클래스 FP와 나머지 클래스 TP, TN, FN을 구별하기 위하여 FP에 해당하는 학습데이터는 +1값을 갖게 되고 나머지는 -1값을 갖는다. 이후 테스트 데이터에 대해서는 각 SVM에 동일한 값을 입력 데이터로 입력 받아 출력되는 출력 값의 크기를 비교하여 최대의 값을 가지는 분류기(SVM)에 해당하는 클래스 값으로 판별된다.

(2) One-Against-One 기법 적용

One-Against-One(OAO) 기법은 OAA기법과는 달리 k개의 클래스를 분류하기 위해 k(k-1)/2개의 이진 SVM으로 구성되는 모델로써 각 학습데이터는 2개의 클래스만을 분류하는 데이터만으로 구성된다. 본 연구 문제에 있어서 구조도는 아래 그림 4와 같으며, 첫 번째 SVM은 클래스 FP와 클래스 TP만으로 구성된 학습데이터를 가지며 테스트 데이터에 대해서도 클래스 FP와 클래스 TP에 대해서만 분류하게 된다. 이후 테스트 데이터에 대한 판별은 모든 SVM을 수행한 후 가장 많은 투표 값을 가진 클래스로 테스트 데이터의 소속을 판별하게 된다.



〈그림 3〉 one-against-all 기법 적용한 multi class SVM 구조



〈그림 4〉 one-against-one 기법 적용한 multi class SVM 구조

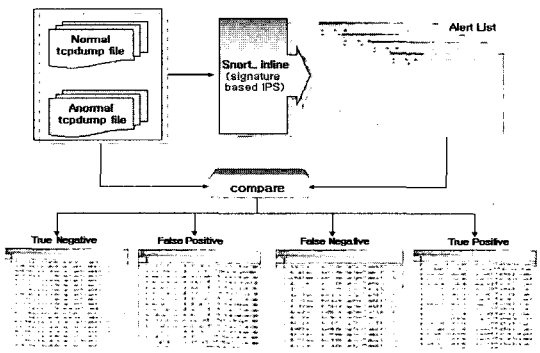
5. 실험 및 결과 분석

5.1 실험 환경 및 실험데이터 구성

본 논문의 실험은 리눅스 환경에서 이뤄졌으며, 오용탐지 기반 IPS의 탐지결과를 분석하기 위하여 Snort_inline-2.3.0-RC1 버전의 IPS를 설치하였고, 탐지 결과를 분석 및 분류하기위한 SVM 실험도구로는 SVM_light[17]을 사용하였다.

기존 오용탐지 기반 IPS의 오경보 최소화를 위해 본 논문에서 제안한 SVM을 이용한 침입방지시스템의 실험 및 성능 평가/분석을 위해서는, 우선 오용탐지 기반의 IPS(Snort_inline)의 탐지결과 분석이 선행되어져야 한다.

학습 및 실험데이터로 사용될 DARPA 데이터집합의 각 주 요일별 tcpdump file과 attack list file들을 이용하여 Duration을 포함한 공격시간, Source IP 주소, Destination IP 주소값을 비교하여 우선 정상 tcpdump data와 비정상 tcpdump data를 분류하였다. 이렇게 정상과 비정상으로 분류된 tcpdump data는 Snort_inline에서의 오경보 패턴 및 정확도를 측정하기 위한 입력값으로 사용되며, 아래 그림 5와 같이 Snort_inline의 탐지엔진에서 규칙에 의해 탐지된 경고 리스트와 비교하여 4가지 유형(TP, FP, TN, FN)의 데이터집합을 생성한다. 이는 4가지 클래스를 분류하는 SVM의 학습데이터로 사용되며 좀 더 구체적인 패킷의 feature 값을 설정하기 위해서 '-vv' 옵션을 사용한 tcpdump data로 변경하여 실험하였다.



〈그림 5〉 4 class 분류

1998년도 DARPA 데이터 집합에 2주 금요일, 3주 수요일 금요일, 4주 화요일 수요일, 6주 화요일 수요일, 7주 수요일 데이터를 이용하여 4가지 클래스로 분류된 tcpdump 형태의 패킷들 가운데 각 클래스당 1000개씩 총 4000개의 패킷을 SVM 학습데이터로 사용하였으며 학습데이터에 이용되지 않은 DARPA 1998 데이터집합에 2주 화요일, 3주 월요일 화요일, 4주 금

요일, 5주 월요일 금요일, 7주 금요일 데이터를 분류기의 성능 평가를 위한 실험데이터로 사용하였다.

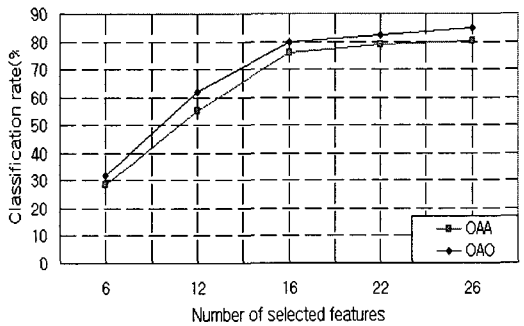
〈표 4〉 SVM에 대한 Train/Test data Set

	TP data	FP data	TN data	FN data	
Train Set	1000개	1000개	1000개	1000개	4000개
Test Set	300개	300개	300개	300개	1200개

5.2 실험결과 및 분석

(1) Feature 선택에 따른 SVM 결과

올바른 분류를 위해서는 학습데이터로 사용되는 패킷의 Feature 선택이 중요하다. 아래 그림 6은 표 3의 학습데이터 특징 파라미터 수가 6, 12, 16, 22, 26개일 때의 테스트 데이터에 대한 분류기 성능 추이에 대해 나타내고 있다. 실험결과 16개 이상의 패킷 정보를 이용한 학습 자료에 대해서 높은 정확도를 보이고 있다. 이는 TOS, TTL, 패킷의 길이, Flagment, Protocol, port number, IP address 등의 16개의 패킷 헤더 정보가 4가지 클래스 패턴 분류에 있어서 가장 중요하게 고려되었음을 알 수 있다.



〈그림 6〉 Feature 선택 수에 따른 multi class SVM 성능 변화

(2) 커널 함수에 따른 SVM 결과

앞선 실험에 의해 가장 분류성능이 높았던 학습데이터 특징 파라미터 수가 26개일 때의 멀티 클래스 SVM의 성능 평가를 위해서 각 모델의 Classification rate와 분류되어진 각 클래스의 정확도과 재현율 값을 계산하였다. 이를 위해 각 모델의 $class_i$ 에 대한 분류 결과를 아래 표 5와 같이 표시 할 때, 멀티 클래스 SVM에 의해 분류되어진 class의 정확도(Precision)와 재현율(Recall)은 식 (8)과 같이 나타낼 수 있다.

〈표 5〉 Classi에 대한 Contingency Table

Classi		True Label	
		Yes	No
분류기 판정	Yes	C_TP	C_FP
	No	C_FN	C_TN

$$P_i = \frac{C\_TP_i}{C\_TP_i + C\_FP_i}, \quad R_i = \frac{C\_TP_i}{C\_TP_i + C\_FN_i} \tag{8}$$

테스트 환경에서는 dot, polynomial 그리고 RBF(Radial Basis function) 커널 함수를 사용하였다. 다수의 실험을 통한 경험적 정보를 이용하여 가장 좋은 결과값을 보인 polynomial 커널의 parameter degree(d)값으로 정수값 1, 4를 적용한 결과값을 나타내었고, RBF 커널의 parameter gamma(g)값으로는 실수값 0.01과 0.5를 적용한 결과를 나타내었다.

아래 표 6, 표 7에서는 26개의 모든 특징 파라미터로 구성된 학습 자료를 이용한 2가지 커널 함수에 따른 OAA, OAO 기법의 멀티클래스 SVM 모델의 실험결과를 나타내고 있다.

실험결과에 의하면 OAA기법보다 OAO 기법을 적용한 SVM에서 좀 더 좋은 결과값이 도출되는데, 이는 이진 분류를 목적으로 하는 SVM의 특성상 동일한 비율로 구성된 각 클래스끼리의 분류 패턴이 더 정확하게 계산되어졌기 때문인 것으로 생각된다. 또한 polynomial 커널에 degree 값으로 4를 적용했을 때 4가지 클래스에 대한 Classification 비율이 84.91%로 가장 우수

〈표 6〉 OAA(One-Against-All) 실험결과

Kernel	Parameter	C	C_TP	C_FP	C_FN	C_TN	P(%)	R(%)	A(%)
Dot	-	FN	95	2	205	898	97.93	31.66	82.75
		FP	300	0	0	900	100	100	
		TN	298	204	2	696	59.36	99.33	
		TP	300	1	0	899	99.66	100	
Polynomial	d=1	FN	72	2	228	898	97.29	24	80.83
		FP	300	2	0	898	99.33	100	
		TN	298	226	2	674	56.87	99.33	
		TP	300	0	0	900	100	100	
	d=4	FN	88	3	212	897	96.70	29.33	82.08
		FP	300	4	0	896	98.68	100	
		TN	297	203	3	697	59.4	99	
		TP	300	5	0	895	98.36	100	
RBF	g=0.01	FN	72	2	228	898	97.29	24	80.83
		FP	300	1	0	899	99.66	100	
		TN	298	225	2	685	56.97	99.33	
		TP	300	2	0	898	99.33	100	
	g=0.5	FN	85	3	215	897	96.59	28.33	81.83
		FP	300	0	0	900	100	100	
		TN	297	213	3	687	58.23	99	
		TP	300	2	0	898	99.33	100	

C=class, P=precision, R=recall, A=classification rate=4 class hit number(C_TP) / total test data

〈표 7〉 OAO(One-Against-One) 실험결과

kernel	Parameter	C	C_TP	C_FP	C_FN	C_TN	P(%)	R(%)	A(%)
Dot	-	FN	96	6	204	894	94.11	32	82.5
		FP	300	2	0	898	99.33	100	
		TN	294	202	6	698	59.27	98	
		TP	300	0	0	900	100	100	
Polynomial	d=1	FN	96	6	204	894	94.11	32	82.5
		FP	300	2	0	898	99.33	100	
		TN	294	202	6	698	59.27	98	
		TP	300	0	0	900	100	100	
	d=4	FN	122	3	178	897	97.6	40.66	84.91
		FP	300	2	0	898	99.33	100	
		TN	297	176	3	724	62.79	99	
		TP	300	0	0	900	100	100	
RBF	g=0.01	FN	94	5	206	895	94.94	31.33	82.41
		FP	300	2	0	898	99.33	100	
		TN	295	204	5	696	59.11	98.33	
		TP	300	0	0	900	100	100	
	g=0.5	FN	87	3	213	897	96.66	29	82
		FP	300	0	0	900	100	100	
		TN	297	213	3	687	58.23	99	
		TP	300	0	0	900	100	100	

C=class, P=precision, R=recall, A=classification rate=4 class hit number(C_TP) / total test data

한 분류 성능을 보였다. 그리고 각 클래스 별 정확도와 재현율 또한 평균 90%씩의 높은 성능을 보였으나, FN(flase negative)과 TN(true negative) 즉, 공격인데도 불구하고 Snort_inline에서 해당 공격에 대해 탐지하지 못한 공격 패킷과 정상 패킷과의 명확한 분류가 일어나지 못했음을 알 수 있었다. 이는 FN class의 실험데이터 구축에 있어서 변형 공격 및 새로운 공격에 대한 탐지 능력을 테스트하기 위하여 학습데이터에 사용되지 않은 새로운 공격 데이터를 실험데이터로 적용했기 때문인 것으로 분석된다.

(3) 성능 비교

IPS 오경보 최소화를 위해 본 논문에서 제안한 SVM 기반의 IPS의 성능 비교를 위해 멀티 클래스 SVM 모델에 사용된 테스트 데이터를 k-NN 알고리즘을 이용한 사례기반 학습기법에 적용하여 실험하였다. 비교 실험은 TiMBL (Tilburg Memory Based Learner, version 5.1)[18] 도구를 사용하였으며, 테스트 데이터에 대한 Snort_inline 및 SVM, k-NN의 성능 평가는 앞서 소개한 정확도와 재현율을 결합한 조화평균(F-measure) 값을 이용하였다. 조화평균은 아래 식 (9)와 같다.

$$F=\frac{2}{\frac{1}{P}+\frac{1}{R}} \tag{9}$$

실험결과 조화평균값을 비교하면 아래 표 8와 같이 기존 IPS에서 오탐지 및 미탐지 비율이 크게 줄어든 것을 알 수 있다. 여기서 공격 및 정상 패킷의 탐지율은 오탐지 및 미탐지에 대해 고려되지 않은 척도이기 때문에 성능 평가에 큰 의미가 없으며, 많은 오경보에도 불구하고 Snort_inline의 탐지율이 높은 이유는 실험에 사용된 패킷의 양이 비교 실험 데이터보다 훨씬 컸기 때문인 것으로 분석된다. 따라서 조화평균값을 이용하여 각 모델의 성능을 평가할 때, 본 논문에서 제안한 SVM기반 IPS의 성능이 가장 뛰어난 것으로 검증되었다.

〈표 8〉 성능 비교를 위한 테스트 분류 결과

	Rate(%)	FP(%)	FN(%)	P(%)	R(%)	F(%)
Snort_inline	89	6.8	20.36	52.36	79.56	63.15
OAA	82.75	0.33	34	99.49	65.94	79.31
OAO	84.91	0.5	29.6	99.29	70.33	82.33
1-NN	79	0	42	100	58	73.41

Rate=(attack detection + Normal detection)/total test data
FP=false positive/normal, FN=false negative/attack, P=precision, R=recall, F=F-measure

6. 결론 및 향후 연구

본 논문에서는 유무선 네트워크 환경에 적용 가능한 오용탐지(Misuse Detection)기반 IPS와 학습기반의 anomaly 탐지기반 IPS의 결합 시스템으로 오용탐지기반 IPS의 오탐지(false positive) 및 미탐지(false negative)를 최소화하여 진성경보만(true positive, false negative)을 IPS의 Action Engine에서 처리하도록 하는 SVM 기반 IPS 기법을 제안하였다.

기존의 이진 분류에만 적용되던 SVM을 오용탐지 기반 IPS의 4가지 탐지 유형으로 분류하기위해서 one-against-all(OAA), one-against-one (OAO)기법을 적용한 2가지 멀티 클래스 SVM 모델로 설계하여 실험하였으며, 실험결과 OAO 기법을 적용한 멀티 클래스 SVM에서 OAA 기법을 사용했을 때 보다 정확한 분류 성능을 보였다.

결과적으로 분류기에 사용된 테스트 데이터의 오탐지 발생비율을 약 99%, 미탐지 발생비율을 약 40.6% 줄이는 효과와, 단일 시스템으로 Snort_inline만을 사용했을 때 보다 Snort_inline 기반한 SVM모델에서 IPS의 성능을 약 20% 높이는 효과를 얻었으며, 본 시스템을 사용할 경우 진성경보의 희생 없이 오용탐지 뿐만 아니라

Anomaly 탐지 효과까지 얻을 수 있음을 보였다.

따라서 제안된 시스템은 공격에 대해 기존 IPS에서 발생시키는 과도한 알람을 최소화하여 진성경보만을 액션 엔진에서 처리하고, 규칙을 통한 탐지뿐만 아니라 지능적으로 신종 공격에 대해서도 탐지 가능하여 공격에 대한 정확한 탐지 및 대응으로 신뢰도가 높은 침입방지 시스템으로써 사용될 수 있을 것으로 기대한다.

다만, 본 모델을 실제 시스템에 적용하기 위해서 실시간 시스템에서의 성능을 정량적으로 측정해 보는 작업이 수행되어야할 것이며, Anomaly 탐지 효율을 높이기 위해 보다 많은 학습데이터 자료 구성이 필요할 것으로 보인다.

참 고 문 헌

[1] 조현정, "차세대 네트워크 보안기술 기반의 침입방지시스템" 정보과학회지, 제 23권 제 1호, p21-26, 2005.

[2] C.Kruegel and T. Toth "Using decision trees to improve signature-based detection." In6th Symposium on Recent Advances in Intrusion Detection(RAID), Lecture Notes in Computer Science. Springer Verlag. USA. September, 2003.

[3] Gary Golomb. IDS v. IPS Commentary, Linuxsecurity.com News, 6/16/2003, http://www.linuxsecurity.com/articles/forums_article- 7476.html

[4] Internet Security System. "The Truth about False Positive." White Technical Report. 2001.

[5] R. Lippman et als., "Evaluation intrusion detection system : The 1998 DARPA Off-line intrusion detection evaluation." Proc. Of DARPA Information Survivability Conference and Exposition, pp. 12-26, 2000.

[6] K. Julisch. "Mining alarm clusters to improve alarm handling efficiency," In 17th Annual Computer Security Application Conference(ACSAC), pp12-21, 2000.

[7] Cuppens, F., Miege, A. "Alert correlation in a cooperative intrusion detection framework", In Proceedings of the IEEE Symposium on Security and Privacy, 2002.

[8] H. Debar, A.Wespi, "Aggregation and Correlation of intrusion-Detection Alert", In Recent Advances in intrusion Detection, number 2212 in Lecture Notes in Computer Science, p85-103, 2001.

[9] S. Manganaris, M. Christensen, D. Zerkle and K. Hermiz, "A Data Mining Analysis of RTID Alarms," In 2nd Work-shop on Recent Advances in Intrusion Detection (RAID99), 1999.

[10] S. Mukkamala, A. H. Sung (2003) Detecting Denial of Service Attacks Using Support Vector Machines. Proceedings of IEEE International Conference on Fuzzy Systems, IEEE Computer Society Press, pp. 1231-1236.

[11] A. H. Sung, S. Mukkamala (2003) Identifying Important Features for Intrusion Detection Using Support Vector Machines and Neural Networks. Proceedings of the 2003 International Symposium on Applications and the Internet Technology, IEEE Computer Society Press, pp. 209-216.

[12] H. Deng, Q.-A. Zeng, and D. P. Agrawal, "SVM-based Intrusion Detection System for Wireless Ad Hoc Networks,"

Proceedings of the IEEE Vehicular Technology Conference (VTC'03), Orlando, October 6-9, 2003.

[13] Campbell, C and Cristianini, N. "Simple Learning Algorithms for Training Support Vector Machines", Technical report, University of Bristol, 1998.

[14] http://snort-inline.sourceforge.net

[15] Hsu, C.W. and Lin, C.J. "A Comparision of Methods for Multicalss Support Vector Machines," IEEE Transaction on Neural Networks. Vol. 13. No.2. pp 415-425, 2002.

[16] Knerr. S., Personnaz, L. and Dreyfus, G. Single-layer Learning Revisited: A Stepwise Procedure for Building and Training a Neural Network. in Neurocomuting: Algorithms. Achitectures and Applications. J. Fogelman, Ed. Springer-Verlag, New York, 1990.

[17] Christopher J.C. Burges, A Tutorial on Support Vector Machines for Pattern Recognition, 1998.

[18] Daelemans, W., Zavrel, J. van der Sloot, K, and van denBosch, A., TiMBL:Tilburg Memory Based Learner, version 5.1, Refernce Guide, Technical Report 01-04, Induction of Linguistic Knowledge, Tilburg University, 2001.

◐ 저 자 소 개 ◑

김 길 한 (Gil-Han Kim)

2004년 백석대학교(구 천안대학교) 정보통신학부 졸업(학사)
2006년 한신대학교 대학원 컴퓨터정보학과 졸업(석사)
2005년~현재 한신대학교 정보과학연구소 연구원
관심분야 : 정보보호, 네트워크 보안, 기계학습, 데이터마이닝
E-mail : gillty@hs.ac.kr

이 형 우 (Hyung-Woo Lee)

1994년 고려대학교 전산과학과 졸업(학사)
1996년 고려대학교 대학원 전산과학과 졸업(석사)
1999년 고려대학교 대학원 전산과학과 졸업(박사)
1999년~2003년 2월 천안대학교 정보통신학부 조교수
2003년~현재 한신대학교 컴퓨터정보소프트웨어학부 부교수
관심분야 : 정보보호, 네트워크 보안, 해킹/바이러스, 스테가노그래피, 컴퓨터 포렌식스
E-mail : hwlee@hs.ac.kr