퍼지 및 지능시스템학회 논문지 2005, Vol. 15, No. 3, pp. 282-288

다중 클래스 SVM기반의 침입탐지 시스템
Intrusion Detection System Based on Multi-Class SVM

이한성* · 송지영* · 김은영** · 이철호** · 박대희*
Hansung Lee*, Jiyoung Song*, Eunyoung Kim**, Chulho Lee**, and Daihee Park*

*고려대학교 컴퓨터정보학과
**국가보안기술연구소

요 약

본 논문에서는 기존의 침입탐지 모델인 오용탐지 모델과 비정상 탐지 모델의 장점은 유지하되 단점은 보완하는 견지에서 새로운 침입탐지 모델을 제안한다. MMIDS로 명명된 새로운 침입탐지시스템은 다음의 평가 기준들을 모두 만족하는 차원에서 설계되었다: 1) 시스템에서 학습되지 않은 새로운 공격 유형의 신속한 발견; 2) 탐지된 공격 유형에 대한 세부적 정보의 제공; 3) 빠르고 효율적인 학습 및 갱신으로 인한 경제적인 시스템의 유지/보수; 4) 시스템의 점증성(incrementality) 및 확장성. MMIDS의 핵심 구성요소로 새롭게 제안된 다중 클래스 SVM은 빠르고 효율적인 학습 및 갱신이 가능하여 침입탐지 시스템의 유지보수 비용을 절감할 수 있다. 실험을 통해 유사한 공격 패턴에 대한 분류성능 및 각 공격 유형별 세분화 능력이 우수함을 보인다.

Abstract

In this paper, we propose a new intrusion detection model, which keeps advantages of existing misuse detection model and anomaly detection model and resolves their problems. This new intrusion detection system, named to MMIDS, was designed to satisfy all the following requirements : 1) Fast detection of new types of attack unknown to the system; 2) Provision of detail information about the detected types of attack; 3) cost-effective maintenance due to fast and efficient learning and update; 4) incrementality and scalability of system. The fast and efficient training and updating faculties of proposed novel multi-class SVM which is a core component of MMIDS provide cost-effective maintenance of intrusion detection system. According to the experimental results, our method can provide superior performance in separating similar patterns and detailed separation capability of MMIDS is relatively good.

Key words : intrusion detection, novelty detection, multi-class SVM, Kernel-ART

1. 서 론

최근 네트워크상의 컴퓨터 및 정보 자원에 대한 위협(threat)요소와 범죄 행위(criminal behavior)가 급속도로 증가하고 있으며, 일반 가정의 컴퓨터에 위해(harm)를 가할 정도로까지 보편화되고 있는 실정이다. 컴퓨터 자원에 대한 위협요소로는 스크립트(script)를 이용한 단순 침입(intrusion)으로부터 다양한 기능의 악성프로그램(malware)을 이용한 컴퓨터의 오용(misuse)에 이르기까지 그 공격 유형들이 점차 다양해지고 있으며, 그 피해의 규모 또한 급속도록 증가하고 있다[1]. 따라서 점점 다양해지는 악성 행위(malicious behavior) 및 침입을 탐지하기 위해서는 보다 효과적인 침입탐지 알고리즘의 개발이 요구된다.

침입탐지 방법론은 침입에 대한 탐지 전략에 따라 크게 오용 탐지(misuse detection) 모델과 비정상 탐지(anomaly detection) 모델로 나누어진다[2]. 오용 탐지모델은 이미 발견된 공격 유형에 대한 면밀한 분석을 통하여 규칙 베이스(rule base)화 하고 이를 기반으로 탐지를 수행하는 방법으로, 새로운 공격 유형이 발견될 시에는 수동으로 규칙 베이스를 갱신(update)해야만 새로운 공격에 대처할 수 있다는 문제점을 가지고 있다. 비정상 탐지 모델은 미리 정의된 정상 행동에 대한 프로파일(profile)로부터 크게 벗어나는 데이터를 비정상 행동으로 판단하여 공격을 탐지하는 방법으로, 새로운 공격유형을 탐지할 수 있다는 점에서는 실용적이나 탐지된 공격 유형에 대한 추가적인 세부 정보를 알 수 없기에 침입에 따른 적절한 대처를 할 수 없다는 한계점을 피할 수 없다[2-3].

최근의 연구문헌 조사에 의하면, 보다 지능적인 침입탐지 모델의 설계를 위하여 데이터마이닝 및 기계학습 기법을 침입탐지시스템에 적용하려는 시도가 활발히 진행 중이다. 이러한 연구 동향 중 특히 패턴 분류(pattern classification) 및 함수 근사(function approximation) 등의 문제에서 매우 우수한 성능을 보이는 SVM(Support Vector Machine)을 침입탐지에 적용하려는 연구가 주목을 받고 있다. SVM을 이용한 침입탐지 모델은 크게 세 가지 부류로 분류된다. 첫 번째 부류[4]는 이진 분류기(binary classifier)인 SVM의 성격을 이용하여 정상 데이터와 공격 데이터를 단순히 이분 분류하는 방법을 취하고 있으며, 두 번째 부류[5]는 단일 클래

접수일자 : 2004년 11월 12일
완료일자 : 2005년 6월 7일

스 SVM (one-class SVM)을 이용하여 비정상 탐지 모델을 구현하는 방법이다. 그러나 위의 두 가지 방법은 모두 정상 데이터와 비정상 데이터만을 분류할 뿐 탐지된 공격 유형에 대한 추가적 정보를 제공하지 못한다. 마지막으로 세 번째 부류[6]는 비교적 많은 개수의 이진 분류기인 SVM들을 조합하여 학습시키는 구조로 다중 클래스 SVM(multi-class SVM)을 구축하여, 정상데이터와 4개의 공격 유형들을 분류한다. 이 방법은 위의 두 개의 방법론에 비해 보다 발전적인 부류이기는 하나, 학습데이터의 질에 시스템의 성능이 전적으로 의존할 뿐만 아니라 학습 시간이 많이 소요 되는 등 여러 가지 문제점들을 가지고 있기에 현실적으로 실무에 적용할 수 있는 지능적 침입탐지시스템과는 여전히 거리가 멀다.

본 논문에서는 기존의 침입탐지 모델인 오용탐지 모델과 비정상 탐지 모델의 장점은 유지하되 단점은 보완하는 견지에서 새로운 침입탐지 모델을 제안한다. MMIDS(Multi-step Multi-class Intrusion Detection System)로 명명된 새로운 침입탐지시스템은 단일 클래스 SVM, 본 논문에서 새롭게 제안하는 다중 클래스 SVM, 그리고 점증적 갱신의 클러스터링 알고리즘인 Kernel-ART[3]를 계층적으로 결합한 구조로, 다음의 평가 기준들을 모두 만족하는 차원에서 설계되었다: 1) 시스템에서 학습되지 않은 새로운 공격 유형의 신속한 발견; 2) 탐지된 공격 유형에 대한 세부적 정보의 제공; 3) 빠르고 효율적인 학습 및 갱신으로 인한 경제적인 시스템의 유지/보수; 4) 시스템의 점증성(incrementality) 및 확장성.

본 논문의 구성은 다음과 같다. 2장에서는 본 논문에서 새롭게 제안하는 다중 클래스 SVM, 침입탐지 분야의 적용을 위해 개발된 Kernel-ART 및 침입탐지 모델 MMIDS에 대해 자세히 설명하고, 3장에서는 실험결과 및 분석을 기술한다. 마지막으로 4장에서는 결론 및 향후 연구과제에 대하여 논한다.

2. 다중 클래스 SVM 및 침입탐지 시스템

본 장에서는 우선 단일 클래스 SVM을 기반으로 하는 새로운 다중 클래스 SVM를 제안하고, Kernel-ART에 대하여 설명한다. 또한 이들을 주요 구성 요소로 하는 새로운 침입탐지 모델인 MMIDS에 대하여 자세히 설명한다.

2.1 단일 클래스 SVM기반 다중 클래스 SVM

SVM은 주어진 문제의 전역적 최적해(global optimum) 값을 보장함으로 패턴 분류 및 함수 근사 등에 적용되어 매우 우수한 성능을 보이고 있으나, 구조적으로 이진 분류기(binary classifier)라는 제약점을 가지고 있기에 다중 클래스 분류 문제인 대부분의 실제 응용 분야에 적용하기가 쉽지 않다. 따라서 이진 분류기인 SVM들을 일대다(one-against-all), 일대일 (one-against -one), 그리고 DAG(directed acyclic graph) 등과 같이 비교적 많은 개수의 SVM들을 조합하여 학습시키는 방법론 등이 제안되었다[7].

침입탐지의 경우 각 공격 유형별로 학습에 사용가능한 데이터의 크기는 차이가 크다. 따라서 학습 데이터 크기의 불균형으로 인하여 학습 시, 한 공격 유형의 학습 결과가 다른 공격 유형의 데이터로부터 큰 영향을 받을 가능성이 매우 높다. 또한 한 공격 클래스에 속하는 새로운 공격 유형들이 계속적으로 생성되기에 현재의 학습 데이터가 클래스 전체를 대표한다고 말하기도 어렵다. 따라서 이진 분류기 SVM은 관측되지 않은 영역을 포함하여 결정 경계면을 생성하여 새로운 학습 데이터에 대해서 오분류(misclassification)할 가능성이 크다. 그러므로 해당 클래스만을 독립적으로 표현하는 단일 클래스 분류기(one-class SVM)[8-9]로서 결정 경계면을 선택하는 것이 보다 유리하다. 따라서 본 논문에서는 단일 클래스 분류기를 기반으로 하여 침입탐지시스템의 여러 공격 유형들을 분류 할 수 있는 다중 클래스 SVM을 제안한다.

d-차원 입력공간상에 존재하는 K-개의 데이터의 집합 $D_k = \{ x_i^k \in R^d \mid i = 1, \ldots, N_k \}$; $k = 1, \ldots, K$ 이 주어졌을 경우, 각각의 클래스를 분류하기 위한 다중 클래스 분류기는 각 클래스의 학습 데이터를 포함 하면서 체적을 최소화 하는 구체(sphere)를 구하는 문제로 정의 되며, 다음의 최적화 문제를 통하여 수식화 된다.

$$\begin{aligned} &\min L_0(R_k^2, a_k, \xi_k) = R_k^2 + C\sum_{i=1}^{N_k}\xi_i^k \\ &s.t \quad \|x_i^k - a_k\|^2 \le R_k^2 + \xi_i^k,\ \xi_i^k \ge 0,\ \forall i. \end{aligned} \tag{1}$$

여기에서, a_k는 k-번째 클래스를 표현하는 구체의 중심이며, R_k^2은 구체의 반경의 제곱, ξ_i^k는 k-번째의 클래스에 속한 i-번째 학습 데이터 x_i^k가 구체에서 벗어나는 정도를 나타내는 벌점 항이며, C는 상대적 중요성을 조정하는 상수(trade-off constant)이다.

식(1)에 관한 쌍대 문제(dual problem)를 구하기 위하여 라그랑제 함수(Lagrange function) L을 도입한다[8-9].

$$\begin{aligned} &L(R_k^2, a_k, \boldsymbol{\xi_k}, \boldsymbol{\alpha_k}, \boldsymbol{\eta_k}) \\ &= R_k^2 + C\sum_{i=1}^{N_k}\xi_i^k \\ &\quad + \sum_{i=1}^{N_k}\alpha_i^k[(x_i^k - a_k)^T(x_i^k - a_k) - R_k^2 - \xi_i^k] \\ &\quad - \sum_{i=1}^{N_k}\eta_i^k\xi_i^k \end{aligned} \tag{2}$$

단, $\alpha_i^k \ge 0$, $\eta_i^k \ge 0$, $\forall i$.

식(2)는 변수 $R_k^2, a_k, \boldsymbol{\xi_k}$에 대해서는 최소 값을 변수 $\boldsymbol{\alpha_k}, \boldsymbol{\eta_k}$에 대해서는 최대 값을 가져야 함으로[9], 아래의 조건식을 만족해야 한다.

$$\begin{aligned} &\frac{\partial L}{\partial R_k^2} = 0: \sum_{i=1}^{N_k}\alpha_i^k = 1. \\ &\frac{\partial L}{\partial \xi_i^k} = 0: C - \alpha_i^k - \eta_i^k = 0 \ \ \therefore \alpha_i^k \in [0, C], \forall i. \\ &\frac{\partial L}{\partial a_k} = 0: a_k = \sum_{i=1}^{N_k}\alpha_i^k x_i^k \end{aligned} \tag{3}$$

조건 식(3)을 라그랑제 함수 L에 대입하면, 다음의 쌍대 문제를 얻는다.

$$\begin{aligned} &\min_{\alpha_k} \sum_{i=1}^{N_k}\sum_{j=1}^{N_k}\alpha_i^k\alpha_j^k\langle x_i^k, x_j^k\rangle - \sum_{i=1}^{N_k}\alpha_i^k\langle x_i^k, x_i^k\rangle \\ &s.t. \quad \sum_{i=1}^{N_k}\alpha_i^k = 1,\ \alpha_i^k \in [0, C],\ \forall_i \end{aligned} \tag{4}$$

입력 공간위에서 정의되는 구체는 매우 간단한 형태의 영

역만을 나타낼 수 있다. 이러한 한계를 극복하기 위하여 커널 함수(kernel fuction) k를 통하여 정의되는 고차원의 특징 공간(feature space) F 위에서 정의되는 구체를 사용하는 방향으로 확장될 수 있다[8]. 각각의 클래스는 각자의 특징공간에서 자신의 경계를 보다 정확하게 표현할 수 있으므로, 시스템의 학습은 각각의 클래스들이 매핑되는 특징공간의 독립성을 고려하여 아래의 QP 문제의 해답을 얻음으로써 이루어진다.

$$\min_{a_k} \sum_{i=1}^{N_k}\sum_{j=1}^{N_k} a_i^k a_j^k k_k(x_i^k, x_j^k) - \sum_{i=1}^{N_k} a_i^k k_k(x_i^k, x_i^k)$$

$$s.t. \quad \sum_{i=1}^{N_k} a_i^k = 1, \ a_i^k \in [0, C], \ \forall_i \tag{5}$$

특히 가우시안 커널(Gaussian kernel)을 사용할 경우, $k(x,x)=1$이 성립함으로 식(5)는 아래와 같이 단순화 된다.

$$\min_{a_k} \sum_{i=1}^{N_k}\sum_{j=1}^{N_k} a_i^k a_j^k k_k(x_i^k, x_j^k)$$

$$s.t. \quad \sum_{i=1}^{N_k} a_i^k = 1, \ a_i^k \in [0, C], \ \forall_i \tag{6}$$

학습 종료 후 적용 과정에서, 각각 클래스의 결정함수는 다음과 같이 정의 된다.

$$f_k(x) = R_k^2 - \left[1 - 2\sum_{i=1}^{N_k} a_i^k k_k(x_i^k, x) + \sum_{i=1}^{N_k}\sum_{j=1}^{N_k} a_i^k a_j^k k_k(x_i^k, x_j^k)\right] \geq 0 \tag{7}$$

서로 다른 특징 공간상에서 정의되는 단일 SVM의 출력 $f_k(x)$값은 각 클래스의 특징 공간상의 경계로부터 해당 테스트 데이터와의 절대 거리를 의미함으로, 서로 다른 특징 공간상의 절대거리를 비교하여 소속 클래스를 결정하는 것은 바람직하지 않다. 따라서 특징 공간상의 절대거리 $f_k(x)$를 특징 공간상에서 정의되는 구형체의 반경 R_k로 나눔으로서 상대적 거리 $\hat{f}_k(x) = f_k(x)/R_k$를 계산하고, 상대거리가 가장 큰 클래스를 입력 데이터 x의 소속 클래스로 결정한다.

$$\text{Class of } x \equiv \arg\max_{k=1,\ldots,K} \hat{f}_k(x)$$

$$\equiv \arg\max_k \left[\left\{R_k^2 - \left(1 - 2\sum_{i=1}^{N_k} a_i^k k_k(x_i^k, x) + \sum_{i=1}^{N_k}\sum_{j=1}^{N_k} a_i^k a_j^k k_k(x_i^k, x_j^k)\right)\right\}/R_k\right] \tag{8}$$

알고리즘 분석 : 본 논문에서 제안된 다중 클래스 SVM과 기존의 방법론들과의 비교를 위하여, k개의 클래스와 클래스 당 학습 데이터의 개수가 n개라 가정하자. 일대다 방법의 경우, 학습해야할 SVM의 개수가 k개이고 SVM당 nk의 데이터를 학습해야 함으로 학습을 위한 총 데이터 수는 nk^2이다. 또한 일대일 방법과 DAG의 경우, $k(k-1)/2$개의 SVM을 학습하며 SVM당 $2n$개의 데이터가 학습에 참여한다. 따라서 총 학습해야하는 데이터의 수는 $nk^2 - nk$이다. 반면, 제안된 알고리즘은 k개의 SVM을 학습하며 SVM당 n개의 데이터가 참여함으로 학습에 참여하는 총 데이터의 수는 nk로써 학습 속도가 상대적으로 빠르다. 또한, 이미 학습된 다중 클래스 SVM에 새로운 클래스를 추가할 경우를 고려한다면, 일대다의 경우 전체 클래스 개수인 k+1개의 SVM을 새로 학습하여야 하며, 일대일인 경우는 k개의 SVM을 재학습하여야 한다. DAG의 경우는 k개의 SVM을 재학습해야 할 뿐만 아니라 그래프를 재구성하는 비용이 추가적으로 요구된다. 반면 제안된 알고리즘은 추가되는 단하나의 SVM만을 학습함으로 시스템 재구성 시 매우 경제적이다. 기존의 다중 클래스 SVM 방법론들과 본 논문에서 제안된 다중 클래스 SVM의 보다 구체적인 성능분석을 표1에서 정리하였다.

표 1. 다중 클래스 SVM들의 알고리즘 성능 분석
Table 1. Comparisons of multi-class SVMs

	학습SVM수	SVM 당 학습데이터의 수	클래스 추가 시 학습 SVM 수	테스트 SVM수
일대다	k	nk	$k+1$	k
일대일	$k(k-1)/2$	$2n$	k	$k(k-1)/2$
DAG	$k(k-1)/2$	$2n$	k + DAG 재구성	k
제안된 방법	k	n	1	k

2.2 Kernel-ART

본 절에서는 침입탐지 시스템을 위해 개발된 클러스터링 알고리즘인 Kernel-ART에 대해 소개하고자 한다. 이한성[3] 등에 의해 제안된 Kernel-ART는 개념 벡터(concept vector)와 SVM(support vector machine)의 머서 커널(mercer-kernel)을 온라인 클러스터링 알고리즘인 ART(adaptive resonance theory)에 접목시킨 클러스터링 알고리즘이다. Kernel-ART는 개념 벡터의 특성상 메모리 측면의 효율성이 높을 뿐만 아니라 클러스터링 수행 후, 각 클러스터의 내용 요약을 위해 별도의 대표 벡터 계산 없이 개념 벡터를 바로 사용하여 클러스터의 레이블링(labeling)을 수행할 수 있다는 장점을 가지고 있다. 한편 머서 커널의 도입은 분류성질이 좋지 않은 데이터에 대해서도 특징 공간으로 데이터를 매핑하여 분류 성질을 높임으로써 발견하기 힘든 패턴의 발견 가능성을 높여준다는 장점을 가지고 있다. Kernel-ART는 점증성과 확장성, 알고리즘의 빠른 수행속도 그리고 입력 데이터의 순서에 민감하지 않은 성질 등 침입탐지 시스템의 관점에서 요구되는 평가항목을 모두 만족할 뿐만 아니라 유사한 패턴의 세분화 능력의 향상 등 클러스터링 관점에서의 성능도 향상시킨 알고리즘이다[3]. 다음의 그림 1은 Kernel-ART 알고리즘의 의사 코드(pseudo code)이다.

2.3 MMIDS(Multi-step Multi-class Intrusion Detection System)

본 절에서는 기존의 침입탐지 모델인 오용탐지 모델과 비정상 탐지 모델의 장점은 유지하되 단점은 보완하는 견지에서 새로운 침입탐지 모델을 제안한다. MMIDS (Multi-step Multi-class Intrusion Detection System)로 명명된 새로운

침입탐지시스템은 다음의 평가 기준들을 모두 만족하는 차원에서 설계되었다: 1) 시스템에서 학습되지 않은 새로운 공격 유형의 발견; 2) 탐지된 공격 유형에 대한 세부적 정보의 제공; 3) 효율적인 학습 및 갱신으로 인한 경제적인 시스템의 유지/보수; 4) 시스템의 점증성(incrementality) 및 확장성.

Step0. Normalize input pattern with L2 norm. Initialize Weights:

$$\underline{w}_1 = \underline{x}_1 = \underline{w}_1^R + \underline{w}_1^S = \underline{x}_1^R + \underline{x}_1^S$$

Step1. While Stopping Condition is false, do Step 2-7

Step2. For each training input, do Step 3-6

Step3. Set activation of all F2 to zero

Step4. Compute Activation Function:

$$AF(\underline{x}_i, \underline{w}_j) = \lambda \cdot \exp\left\{-\frac{1}{c}\,||\underline{x}_i^R - \underline{\hat{w}}_j^R||^2\right\} + (1-\lambda)\cdot\frac{\sum_{l=1}^{m}\delta(x_{il}^S, w_{jl}^S)}{m}$$

Step5. Find $j*$ with max activation

Step6. Test for reset:

If $AF(W_{j*}, X_i) \geq \rho$ then

$$\underline{w}_{j*}^{R(i)} = \underline{w}_{j*}^{R(i-1)} + \underline{x}_i^R$$

$$\underline{w}_{j*}^{S(i)} = \textit{Most frequent symbol}$$

else new processing element allocation: c = c + 1

$$\underline{w}_{j*}^{R(i)} = \underline{w}_{j*}^{R(i-1)} + \underline{x}_i^R$$

$$\underline{w}_{j*}^{S(i)} = \textit{Most frequent symbol}$$

Step7. Test for stopping condition

그림 1. Kernel-ART 알고리즘
Fig. 1. Kernel-ART Algorithm

그림 2에 도식화된 MMIDS의 구조는 세 개의 주요 컴포넌트인 정상데이터와 공격데이터를 분류하는 단일 클래스 SVM, 공격 데이터를 DOS(denial of service), R2L(remote to local), U2R(user to root), Probing의 네 가지 공격 유형 중 하나로 분류하는 다중 클래스 SVM, 그리고 각 공격 종류별로 보다 세부적 클러스터링을 수행하는 Kernel-ART로 구성된다. 테스트 데이터에 대한 침입탐지 절차는 다음의 세 단계에 걸쳐 수행되는 바, 각 단계에 대한 보다 상세한 설명은 다음과 같다.

제 1단계 : 정상데이터에 의해 학습된 단일 클래스 SVM은 정상 데이터와 공격 데이터에 대한 일차적인 분류를 빠른 속도로 수행한다. 학습 시 정상데이터만을 요구함으로 학습을 위한 별도의 공격 데이터를 준비할 필요가 없으며, 학습 속도 또한 매우 빠르다. 학습된 단일 클래스 SVM은 비정상 탐지모델로써 시스템에서 학습되지 않은 새로운 공격(novel attack)을 탐지하며, 시스템 운영 시 정상데이터에 대한 추가적인 처리과정은 없고, 공격데이터가 탐지되면 침입대응시스템(intrusion response system)에 1차적 경고를 발생시킴과 동시에 제 2단계 절차를 진행시킨다.

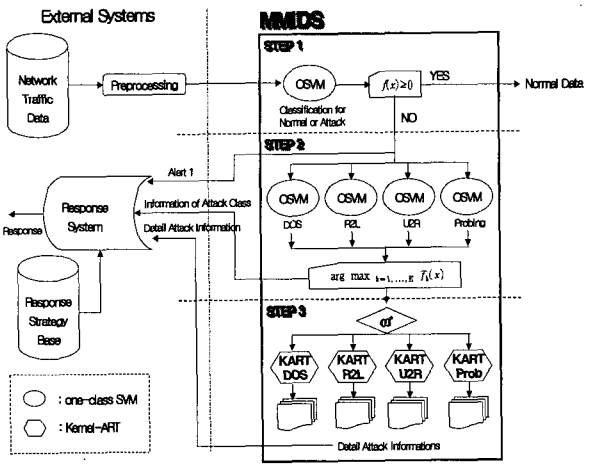


그림 2. MMIDS(Multi-step Multi-class Intrusion Detection System) 구조도
Fig. 2. The Architecture of MMIDS(Multi-step Multi-class IDS)

제 2단계 : 1단계에서 공격으로 분류된 데이터들을 다중 클래스 SVM에 의해 DOS, U2R, R2L, Probing중 하나의 공격 유형으로 분류하고 침입대응시스템에 공격 유형에 대한 추가 정보를 제공한다. 또한 새롭게 발견된 공격 데이터를 시스템에 반영할 경우, 시스템 전체를 재학습 시킬 필요 없이 해당 클래스의 분류기만을 점증적 갱신의 학습방법[8]으로 재학습한다. 따라서 실제 시스템 운용 시 시스템의 유지/보수 비용을 줄일 수 있다.

제 3단계 : 해당 공격에 대한 보다 세부적인 정보가 요구될 경우, 각 공격 종류별 클러스터링이 Kernel-ART에 의해 수행된다. 세부 공격별 데이터는 SVM과 같은 분류기를 학습시킬 정도로 많지 않으며, 그 종류를 예측하기도 어렵다. 따라서 비교사학습인 클러스터링을 이용하여 해당 공격별 분류를 수행하는 것이 보다 바람직하다. Kernel-ART는 각 클러스터의 내용 요약을 위해 별도의 대표 벡터 계산 없이 개념 벡터를 사용함으로 공격 별 레이블링(labeling)을 제공한다[3].

3. 실험 및 결과 분석

본 논문에서 제안된 침입탐지시스템인 MMIDS의 성능을 실험적으로 검증하기 위해서, 침입탐지 분야의 대표적 실험 데이터인 KDD CUP 99[10]를 사용하여 기존의 연구방법들과의 성능을 비교 평가하였다.

3.1 실험 조건

KDD CUP 99은 1998 DARPA Intrusion Detection Evaluation Program에서 침입탐지 분야의 표준 데이터 집합을 얻기 위하여 미국의 군사 네트워크(military network) 상에서 시뮬레이션(simulation)을 통해 만든 데이터로써, 본 실험에서는 실험 결과의 정확한 분석을 위하여 KDD CUP 99 데이터 중 Corrected Labeled된 데이터 집합만을 이용하였다. 사용된 총 데이터의 수는 311,029개이며 9개의 기호형

(symbolic) 속성(attribute)과 32개의 숫자형(numeric) 속성으로 구성되어있다. 데이터는 DOS, R2L, U2R, probing 등 크게 4개의 공격 유형으로 구분되며, 각 공격별 세부적인 공격 유형은 Probing 공격의 ipsweep, saint 등을 비롯한 6개를 포함하여 총 37개이다. KDD CUP 99데이터의 각 속성에 대한 자세한 설명은 [10]에 기술되어있다. KDD CUP 99 데이터의 모든 속성을 사용하기 위하여 기호형 속성에 대하여 다음과 같은 변환을 수행하였다. Σ는 한 속성에 나타날 수 있는 기호형 데이터 값들의 집합이며, $|\Sigma|$는 집합의 구성요소(element)의 수를 의미한다. 한 기호형 데이터의 값이 i 번째 구성요소라고 가정하면 해당 기호형 데이터의 값은 아래의 방법과 같이 변환된다[11].

$$\underbrace{0\,0\,0\,0\,0\;\;1/|\Sigma_i|\;\ldots\;0\,0\,0}_{|\Sigma|} \qquad i$$

3.2 실험 결과 비교 및 분석

정상 데이터와 4개의 공격 유형들에 관한 각 연구들의 분류 능력을 표 2에 비교, 정리하였다. 표 2에 의하면 본 논문에서 제안한 다중 클래스 분류기가 기존 연구 결과들에 비해 전반적으로 높은 분류 능력을 보이고 있음을 알 수 있다. 여기서 한 가지 주목할 점은 모든 연구결과들이 유독 R2L과 U2R에 대한 분류능력에서만 유난히 성능이 떨어진다는 것이다. 특히 본 실험과 동일한 데이터를 대상으로 실험을 수행한 일대일 방법의 다중 클래스 SVM을 사용한 Ambwani[6]과의 비교를 살펴보면, 본 방법론의 실험결과가 매우 향상됐음을 알 수 있다. 이는 다음과 같이 설명된다.

KDD 99 데이터는 네트워크 패킷(network packet)을 기반으로 만들어진 데이터로써, 호스트 기반(host-based) 공격인 R2L과 U2R은 매우 비슷한 성격을 가지며 특히 정상 데이터와의 분류가 매우 어렵다. 또한 각 공격 유형별로 학습에 사용 가능한 데이터 크기의 차이가 매우 크다. 즉 U2R이나 R2L공격의 경우는 데이터의 크기가 상대적으로 다른 공격 유형들에 비해 작다. 따라서 모든 클래스를 동일한 특징 공간으로 매핑한 후 각각의 클래스에 대한 결정 경계를 학습하는 기존의 다중 클래스 SVM은 학습 시 데이터의 크기가 큰 클래스가 학습 결과에 보다 많은 영향을 미칠 수 있다. 그러므로 각각의 클래스에 대하여 자신의 경계를 가장 잘 표현할 수 있는 분류기를 독립적으로 학습함으로서, 자신의 클래스에 대한 분류 성능을 높이며 학습 데이터 크기의 불균형에 대한 영향력을 최소화한다는 본 논문의 다중 클래스 SVM의 전략(식 6 참조)이 매우 효과적이었다고 판단된다.

3.3 공격 데이터의 세분화 실험

추가적으로 MMIDS의 제 3단계에 해당하는 공격 유형별 세분화 능력을 검증하고자, Kernel-ART을 대상으로 4개의 공격 유형별 세분화 분류를 수행하였다. 모든 실험 결과 중 Probing 공격에 대한 결과만을 대표적으로 표 3에 요약 정리하였다. 실험에서 결정 경계 변수인 ρ값은 0.7, 커널 함수인 가우시안 함수의 변수 c는 0.1로 설정하였으며, 알고리즘 수행 후 생성된 클러스터의 수는 19개이다. 표 3의 실험 결과에 의하면 MMIDS의 세부 분류 능력이 비교적 우수함을 알 수 있다.

표 2. 각 연구들의 분류 성능 비교
Table 2. Comparison with Other Intrusion Detection Algorithms

DR(%) / Class	Dr. Bernhard [12]	W. Lee [13]	Y. Liu [14]	Kayacik [15]	Ambwani [6]	MMIDS
Normal	99.5	-	-	95.4	99.6	**96.74**
Dos	97.1	79.9	56	95.1	96.8	**98.24**
R2L	**8.4**	**60.0**	**78**	**9.9**	**5.3**	**35.00**
U2R	**13.2**	**75.0**	**66**	**10.0**	**4.2**	**85.23**
Probing	83.3	97.0	44	64.3	75	**98.27**

표 3. Probing 공격 유형의 세분화 결과
Table 3. Experimental Results of Probing Attack
경계변수 ρ=0.7, 커널함수 변수 $c=0.1$, 생성된 클러스터 수 : 19

공 격	탐지율	공 격	탐지율
ipsweep	95.09%	satan	91.73%
saint	17.26%	mscan	89.17%
portsweep	96.33%	nmap	24%

4. 결 론

본 논문의 주된 공헌은 시스템에 알려지지 않은 새로운 공격유형의 탐지와 동시에 탐지된 공격 유형에 대한 자세한 정보를 제공할 수 있는 강건하고 효율적인 침입탐지 모델의 제안에 있다. 본 논문에서 제안하는 방법은 빠르고 효율적인 학습 및 갱신으로 시스템의 유지 보수 측면에서 비용 절감을 가져올 수 있을 뿐만 아니라, 발견되지 않은 침입 유형에 대한 침입 사고를 미연에 방지할 수 있는 새로운 방법론이라 하겠다.

본 논문에서 제안된 다중 클래스 SVM은 점증성과 확장성, 빠르고 효율적인 학습 및 갱신 등 기존의 다중 클래스 SVM에 비해 알고리즘의 복잡도에서 향상을 가져왔을 뿐만 아니라 실험에서 보여준 바와 같이 유사한 패턴의 세분화 능력의 향상 등 패턴 분류(classification) 관점에서의 성능도 향상시킨 새로운 방법론이다.

결론적으로 본 논문에서 제안된 침입탐지 모델인 MMIDS는 시스템에 의해 학습되지 않은 침입유형에 대한 탐지가 어렵다는 오용탐지 방법의 한계점 및 탐지된 공격유형에 대한 세부적인 정보를 제공하지 못한다는 비정상 탐지 모델의 한계점을 상당부분 극복하고 있으며, 침입탐지 알고리즘의 견지에서 보았을 때 기존의 침입탐지 방법론에 비해 다음과 같은 장점을 갖는다. 1) 학습 및 갱신의 효율성; 2) 각 공격유형에 대한 세부적인 정보 제공; 3) 학습되지 않은 침입유형의 발견.

본 연구의 향후 연구 과제로는 제안된 다중 클래스 SVM의 데이터마이닝 분야에 대한 다각적인 적용 사례에 대한 연구가 필요하다. 또한 침입 탐지 시스템에 의해 탐지된 공격들의 세부 정보를 침입 방지 시스템의 정책 수립에 적극적으로 활용할 수 있는 통합 시스템에 관한 연구가 향후 연구과

제로 요구된다.

참 고 문 헌

[1] 이장현, 김성옥, "신경회로망을 이용한 비정상적인 패킷탐지", *정보보호학회 논문지*, 제 11권, 제 5호, pp. 105-117, 2001.

[2] Steven Noel, Duminda Wijesekera, and Charles Youman, "Modern Intrusion Detection, Data Mining, and Degrees of Attack Guilt," in *Applications of Data Mining in Computer Security*, Kluwer Academic Publisher, pp. 1-31, 2002.

[3] 이한성, 임영희, 박주영, 박대희, "SVM과 클러스터링 기반 적응형 침입탐지 시스템", *퍼지 및 지능시스템학회 논문지*, Vol. 13, No. 2, pp. 237-242, 2003.

[4] WunHwa Chen, ShengHsun Hsu, and H. P. HwangPin Shen, "Application of SVM and ANN for intrusion detection", *Computers & Operations Research, ELSEVIER*, Vol. 32, Issue 10, pp. 2617-2634, 2005.

[5] KunLun Li, HouKuan Huang, ShengFeng Tian, and Wei Xu, "Improving one-class SVM for anomaly detection", *International Conference on Machine Learning and Cybernetics*, Vol. 5, pp. 3077-3081, 2003.

[6] Ambwani, T., "Multi class support vector machine implementation to intrusion detection", *Proceedings of the International Joint Conference on Neural Networks*, Vol. 3, pp. 2300-2305, 2003.

[7] C.W. Hsu and C.J. Lin., "A comparison of methods for multi-class support vector machines", *IEEE Transactions on Neural Networks*, Vol. 13, pp. 415-425, 2002.

[8] 박주영, 임채환, "비정상 상태 탐지 문제를 위한 서포트벡터 학습", *퍼지 및 지능시스템학회 논문지*, Vol. 13, No. 3, pp. 266-274, 2003.

[9] David M.J. Tax and Robert P.W. Duin, "Uniform Object Generation for Optimizing One-class Classifiers", *Journal of Machine Learning Research*, Vol. 2, Issue 2, pp. 155-173, 2001.

[10] KDD CUP 1999 DATA, Available in http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html and http://www-cse.ucsd.edu/users/elkan/kdresults.html

[11] Eleazar Eskin, Andrew Arnold, Michael Prerau, Leonid Portnoy and Salvatore Stolfo. "A Geometric Framework for Unsupervised Anomaly Detection: Detecting Intrusions in Unlabeled Data", in *Applications of Data Mining in Computer Security*, Kluwer Academic Publisher, pp. 77-101, 2002.

[12] Results of the KDD '99 Classifier Learning Contest, Available in http://www-cse.ucsd.edu/users/elkan/clresults.html

[13] Wenke Lee, Salvatore J. Stolfo, and Kui W. Mok, "A data mining framework for building intrusion detection models", *Proceedings of the 1999 IEEE Symposium on Security and Privacy*, pp. 120-132, 1999.

[14] Liu, Y., Chen, K., Liao, X., and Zhang, W., "A Genetic Clustering Method for Intrusion Detection", *Pattern Recognition*, Vol. 37, Issue 5, pp. 927-942. 2004.

[15] Kayacik, H.G., Zincir-Heywood, A.N., and Heywood, M.I., "On the capability of an SOM based intrusion detection system", *Proceedings of the International Joint Conference on Neural Networks*, Vol. 3, pp. 1808-1813, 2003.

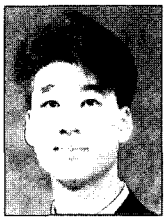
저 자 소 개

이한성 (Han-Sung Lee)

1996년 : 고려대학교 전산학과(학사)
1996년~1999년 : (주)대우엔지니어링 근무
2002년 : 고려대학교 전산학과(석사)
2002년~현재 : 고려대학교 전산학과 박사과정

관심분야 : 기계학습, 데이터마이닝, 인공지능, 인공신경망, SVM, 침입탐지, 퍼지 이론

E-mail : mohan@korea.ac.kr



송지영(Ji-Young Song)

1996년 고려대학교 전산학과(학사)
1999년 고려대학교 전산학과(석사)
2005년 8월 고려대학교 전산학과 박사학위 예정

관심분야 : 데이터 마이닝, 인공지능, 인공신경망, 생체인식, SVM

E-mail : songjy@korea.ac.kr

김은영(Eun-Young Kim)

1999년 : 충남대학교 컴퓨터공학과(학사)
2001년 : 충남대학교 컴퓨터공학과(석사)

이철호(Eun-Young Kim)

2002년 : 아주대학교 정보통신대학 정보및컴퓨터공학부 (학사)
2004년 : 아주대학교 정보통신대학 정보및컴퓨터공학부 (석사)

박대희 (Dai-Hee Park)

1982년 : 고려대학교 수학과(학사)
1984년 : 고려대학교 수학과(석사)
1989년 : 플로리다 주립대학 전산학과(석사)
1992년 : 플로리다 주립대학 전산학과(박사)
1993년~현재 : 고려대학교 컴퓨터정보학과 교수

관심분야 : 인공지능, 지능 데이터베이스, 데이터마이닝, 인공신경망, 퍼지 이론

E-mail : dhpark@korea.ac.kr