Journal of the Korea Academia-Industrial cooperation Society
Vol. 14, No. 4 pp. 1942-1950, 2013

http://dx.doi.org/10.5762/KAIS.2013.14.4.1942

다중 클래스 SVM을 이용한 트래픽의 이상패턴 검출

박영재[1], 김계영[1], 장석우[2*]
[1]숭실대학교 컴퓨터학부, [2]안양대학교 디지털미디어학과

Traffic Anomaly Identification Using Multi-Class Support Vector Machine

Young-Jae Park[1], Gye-Young Kim[1] and Seok-Woo Jang[2*]
[1]**School of Computing, Soongsil University**
[2]**Department of Digital Media, Anyang University**

요 약 본 논문에서는 네트워크 트래픽 데이터를 시각화하고, 시각화된 데이터에 다중 클래스 SVM을 적용함으로써 트래픽의 공격을 자동으로 탐지하는 새로운 방법을 제안한다. 본 논문에서 제안된 방법은 먼저 송신자와 수신자의 IP와 포트 정보를 2차원의 영상으로 시각화한 후, 시각화된 영상으로부터 트래픽의 공격을 의미하는 라인과 명암값이 높은 패턴을 추출한다. 그리고 송신자와 수신자 포트의 분산도 값을 구하고, ISODATA 군집화 알고리즘을 이용하여 군집의 개수와 엔트로피 특징 값을 추출한다. 그런 다음, 위에서 추출한 여러 특징 값들을 다중클래스 SVM(Support Vector Machine)에 적용하여 네트워크 트래픽의 공격이 정상 트래픽, DDoS, DoS, 인터넷 웜, 그리고 포트 스캔인지의 여부를 효과적으로 탐지 및 분류한다. 본 논문의 실험에서는 제안된 다중 클래스 SVM을 활용한 방법이 네트워크 트래픽의 공격을 보다 효과적으로 탐지하고 분류한다는 것을 보여준다.

Abstract This paper suggests a new method of detecting attacks of network traffic by visualizing original traffic data and applying multi-class SVM (support vector machine). The proposed method first generates 2D images from IP and ports of transmitters and receivers, and extracts linear patterns and high intensity values from the images, representing traffic attacks. It then obtains variance of ports of transmitters and receivers and extracts the number of clusters and entropy features using ISODATA algorithm. Finally, it determines through multi-class SVM if the traffic data contain DDoS, DoS, Internet worm, or port scans. Experimental results show that the suggested multi-class SVM-based algorithm can more effectively detect network traffic attacks.

Key Words : Learning, Morphological Operation, Multi-Class, Support Vector Machine, Traffic Anomaly

1. 서론

최근 들어, 광범위하게 확산 보급된 유무선 인터넷을 통해서 대량의 무의미한 패킷 데이터를 집중적으로 발생시켜 피해 호스트의 네트워크를 손상시키는 네트워크 트래픽 공격이 많이 나타나고 있다. 따라서 이런 네트워크 트래픽의 공격을 효과적으로 탐지하고, 악성 공격 트래픽을 대응하는 연구에 대한 관심이 지속적으로 증가하고 있다[1].

이런 기존의 트래픽 탐지 방법 중에서 최근에는 네트워크 상에서 발생하는 방대한 양의 트래픽 이벤트를 실시간적으로 시각화하는 기술이 큰 관심을 받고 있다[2-4]. 이 기술은 네트워크 관리자에게 보안과 관련된 많은 정보를 신속하게 전달할 수 있다는 장점이 있다. 뿐만 아니라, 여러 종류의 네트워크 공격에 대한 형태를 잘 표현할 수 있어 네트워크 트래픽 공격을 용이하게 파악할 수 있다. 그러므로 본 논문에서도 네트워크의 트래픽 데이터를 시각화하고, 시각화된 IP와 포트(Port) 정보에 대

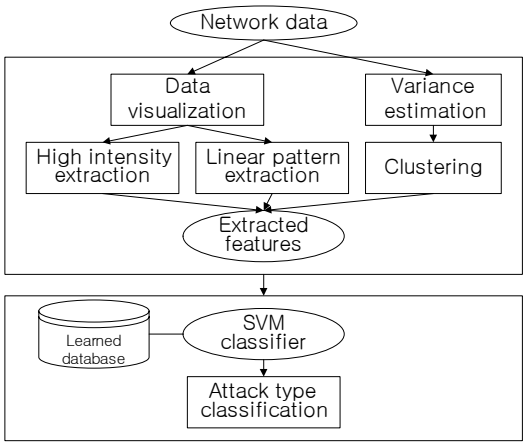
*Corresponding Author : Seok-Woo Jang(Anyang Univ.)
Tel: +82-31-467-0842 email: swjang@anyang.ac.kr
Received February 27, 2013 Revised March 18, 2013 Accepted April 11, 2013

한 특징들을 SVM(support vector machine)에 학습시켜 이상패턴을 효과적으로 검출하는 알고리즘을 제안한다.

본 논문에서는 네트워크 트래픽 데이터의 공격 중에서 서비스 거부 공격(DoS: Denial of Service), 분산 서비스 거부 공격(DDoS: Distributed DoS), 인터넷 웜(Internet Worm), 포트 스캔(Port Scan)의 4가지 종류에 대해서 각각을 탐지하고 분류하는 방법을 제안한다.

DDoS 공격[5]은 분산 배치된 여러 대의 공격자가 피해 호스트 사이트를 동시에 공격하는 방식으로, 이 방식에서는 송신자와 수신자 IP의 관계는 N:1이다. DoS 공격[6]은 한 대의 송신자 컴퓨터가 여러 개의 포트를 이용하여 한 대의 수신자 컴퓨터에 접속하여 서비스를 방해하는 공격으로, 송신자와 수신자 IP의 관계는 1:1이고, 포트의 관계는 N:1이다. 인터넷 웜[7]은 DDoS와 반대로 한 대의 컴퓨터가 여러 대의 컴퓨터에 접속하여 감염시킴으로써 네트워크를 손상시키는 공격으로, 송신자와 수신자 IP의 관계는 1:N이다. 포트 스캔[3]은 공격자가 선택된 호스트의 열려 있는 다수의 포트를 이용하여 패킷을 전송하는 공격으로, 송신자와 수신자 포트의 관계는 1:N이다.

본 논문에서 제안된 시스템은 트래픽 데이터의 특징 추출 단계와 공격 탐지 단계의 두 단계로 구성된다. Fig. 1은 제안된 시스템의 전체적인 개요도이다.



[Fig. 1] Overall flow

첫 번째 단계에서는 네트워크 트래픽 데이터를 분석하여 특징을 추출하는데, 두 가지 방법을 통해 특징을 추출한다. 첫 번째 방법은 트래픽 데이터의 시각화 방법이고 두 번째 방법은 군집화 방법이다. 본 논문에서는 트래픽의 헤더파일에서 송신자와 수신자의 IP와 포트 정보를 이용하여 특징분석을 진행한다. 우선, 송신자와 수신자의 IP 정보를 이용하여 4개의 2차원 영상으로 시각화한다. 그런 다음, 송신자와 수신자의 포트 정보를 이용하여 하나의 2차원 영상으로 시각화한다. 이와 같이 총 5개의 2차원 영상에서 명암 값이 높은 패턴들을 분리하는 과정을 통해 명암 값 특징을 추출한다. 그리고 형태학적 연산을 이용하여 잡음제거 단계를 수행하여 얻어진 영상에 대해 허프(Hough) 변환을 이용하여 선형패턴을 추출한다. 또한, 특징 값의 정확도를 위해 포트 영상에 대해서 군집화 알고리즘을 적용하여 송신자와 수신자 포트의 분산도를 통해 보다 정확한 특징 값을 추출한다. 송신자와 수신자 포트의 분산도 값을 구하고 군집화 알고리즘을 이용하여 군집의 개수와 엔트로피의 특징을 보다 효과적으로 추출한다.

두 번째 단계에서는 첫 번째 단계에서 추출한 특징을 이용하고, 사전에 학습된 다중클래스 SVM(Support Vector Machine)을 이용하여 네트워크 트래픽의 공격을 분류하고 탐지한다.

1장에서는 본 연구를 수행하게 된 동기 및 배경, 그리고 전체적인 개요를 기술하였다. 2장에서는 트래픽 데이터를 시각화하는 방법에 대해 기술하고, 3장에서는 트래픽 영상의 잡음을 제거하는 기법에 대해 설명한다. 4장에서는 특징 추출과 군집화를 수행하는 기법에 대해 설명하며, 5장에서는 SVM을 이용하여 트래픽의 이상패턴을 검출하는 방법에 대해 설명한다. 그리고 6장에서는 제안한 방법의 성능을 비교 평가하기 위해서 수행한 실험결과를 보이며, 7장에서는 결론 및 향후 연구방향을 제시한다.

2. 트래픽 데이터의 시각화

본 논문에서는 트래픽 데이터의 IP와 포트 정보를 영상으로 시각화한다[8]. 첫째, IP의 시각화를 고려하면, 하나의 IP는 2개의 2차원 영상(256×256)으로 표현된다. 예를 들어, 만일 IP 주소가 a.b.c.d로 구성된다면 a와 b를 나타내는 해당 위치에 하나의 점이 생성되고, c와 d를 나타내는 해당 위치에 하나의 점이 생성된다. 그리고 동일한 위치에 점이 반복적으로 생성될 경우에는 명암값이 1씩 증가되어 화면에 표시된다. 둘째, 포트의 시각화는 송신자와 수신자의 포트 정보를 각각 0에서 255 사이의 숫자로 정규화한 후, 하나의 2차원 영상(256×256)으로 표현한다.

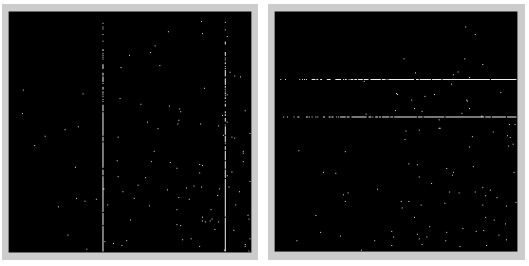
그런 다음, 시각화된 IP 영상과 포트 영상으로부터 선형 패턴과 명암도 특징을 추출한다. 첫째, DDoS 공격에 해당하는 트래픽 영상을 살펴보면, 송신자 IP의 ab와 cd 영상에서 많은 송신자들이 수신자 IP의 ab와 cd 영상의 수신자에게 모여 있다. 그러므로 송신자 IP 영상에 선형 패턴이 존재하는지 또는 명암값이 높은 패턴이 존재하는

지의 여부를 확인하여 DDoS 공격의 발생 유무를 판단할 수 있다.

DoS 공격에 해당하는 트래픽 영상에서는 선형 패턴이 존재하지 않는다. 그러나 명암값이 매우 높은 해당 송신자 IP와 수신자 IP가 트래픽의 공격자와 희생자를 의미한다. 즉, 하나의 송신자가 일반적인 전송량을 벗어난 수많은 양의 데이터를 전송하여 하나의 수신자를 접속하기 때문에 매우 높은 명암 값을 갖는 영역이 각 영상에서 모두 나타나게 된다.

인터넷 웜에 해당하는 트래픽 영상에서는 시각화된 DDoS 영상에서와는 반대로 한 대의 컴퓨터가 여러 대의 컴퓨터에 접속하기 때문에 수신자 IP 영상에 선형 패턴이 존재하게 된다.

포트 스캔에 대응하는 트래픽 영상에서는 특정한 송신자에서 특정한 수신자의 많은 포트로 패킷이 전송되므로 IP 영상에서는 특정한 패턴이 생성되지 않으나, 포트 영상에서는 선형의 패턴이 생성된다. Fig. 2는 시각화된 영상 중에서 DoS와 포트 스캔 공격의 포트 영상의 한 예를 보여준다.



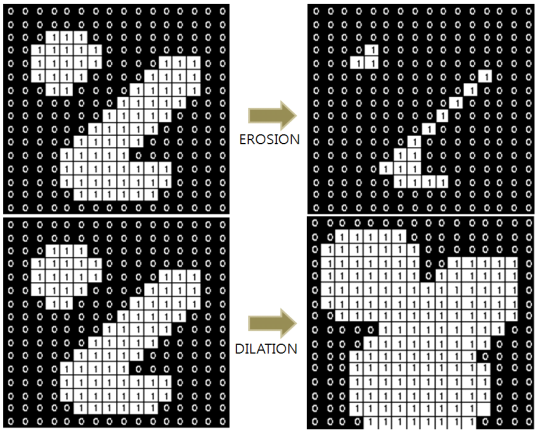
[Fig. 2] Port image for DoS and Port Scan
(a) DoS (b) Port Scan

3. 트래픽 영상의 잡음 제거

IP 영상에서는 이상현상이 발생할 경우 선형의 독특한 패턴 또는 명암값이 특별히 높은 패턴들이 나타나며, 해당 패턴의 특징을 추출하여 네트워크 트래픽의 공격을 탐지할 수 있다. 그러나 영상에는 점들로 이루어진 정상적인 트래픽들이 상당수 존재하며, 따라서 이상현상 패턴이 정상 트래픽과 섞여 있을 수 있으므로 이상 패턴에 대한 특징을 추출하기가 어렵다는 문제가 발생한다. 따라서 본 논문에서는 형태학적인 연산(morphological operation) [9]을 이용하여 주요 목표 패턴이 아닌 정상 트래픽을 제거하는 방법을 수행한다.

형태학적 연산은 영상의 내부에 존재하는 물체의 기하학적인 구조를 다루는 비선형 영상처리 기법 중의 하나이다. 형태학적 연산은 Hermann Minkowski의 집합 이론을 기초로 하였으며, Jean Serra와 George Matheron이 처음으로 이진 영상에 형태 연산의 개념을 적용하였다. 그리고 형태학적 연산 필터는 현재 영상의 잡음 제거에 많이 사용되고 있으며, 보통 에지 검출기, 영상 압축, 그리고 특징 추출 등에도 자주 사용된다.

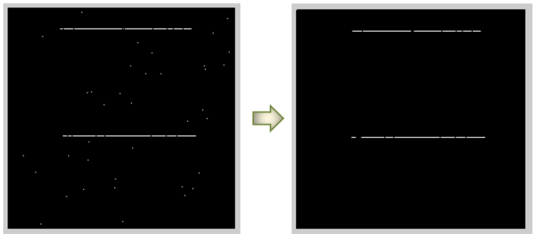
보통 네트워크 트래픽의 공격 패턴은 일반적으로 수직 또는 수평의 직선을 이루며, 정상적인 트래픽의 패턴은 대부분 점의 형태로 나타난다. 이와 같은 특성을 이용하여 수직 또는 수평의 직선에 대해 각각 열림(opening) 형태학 연산을 적용하여 직선의 형태를 보이는 패턴은 유지하고 점의 형태를 보이는 패턴은 제거한다. 열림 연산이란 침식(erosion) 연산을 적용한 후 다시 팽창(dilation) 연산을 수행하는 구조로 구성되어 있다. 침식 연산은 특정 크기의 마스크(mask) 내부에 하나라도 0이 존재한다면 해당 마스크의 영역을 모두 0으로 전환하는 연산이며, 팽창 연산은 특정 크기의 마스크 내부에 하나라도 1이 존재한다면 해당 마스크의 영역을 모두 1로 전환하는 연산이다.



[Fig. 3] Operation of erosion and dilation

Fig. 3은 3×3 크기의 마스크를 이용한 침식 연산과 팽창 연산 과정의 예를 각각 보여준다. Fig. 3의 상단 영상에서 3×3 크기 마스크 내부의 9개 모두의 값이 1이라면 0으로 변환하지 않고, 하나의 영역이라도 0이 존재한다면 모두 0으로 변환함을 확인할 수 있다. Fig. 3의 하단 영상은 팽창 연산의 예를 보여준다. 침식 연산과는 반대로 3×3의 마스크 내부에 객체가 1개라도 존재한다면 3×3 영역 모두를 1로 채워 넣는다. 일반적으로, 침식 연산 후 팽창 연산을 수행하면 일정크기 이하의 객체는 제거되고

일정크기 이상의 객체만 원래의 크기와 유사한 크기로 제거되지 않고 남아있게 된다. 따라서 본 논문에서는 5×1 크기의 마스크와 1×5 크기의 마스크 두개를 이용하여 각각 수직과 수평 직선 패턴을 제외한 잡음을 제거하는 방법을 사용하였다. Fig. 4는 제안한 방법을 이용하여 잡음을 제거한 예를 보여준다.

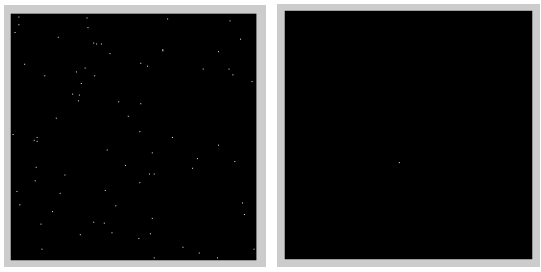


[Fig. 4] Noise elimination

네트워크 트래픽에서 일부에 집중되는 현상이 발생할 경우 IP 영상의 일부 영역에서 명암의 강도가 주변 영역에 비해서 상당히 높게 나타나는 패턴이 발생한다. 이와 같은 패턴은 강도는 높으나 형태는 정상 트래픽 패턴과 유사한 점의 형태를 띠고 있어 형태학적 연산을 이용하여 분류할 수 없다. 따라서 명암도를 측정하여 패턴을 분리하는 방법을 사용한다. 즉, 다른 영역에 비해서 월등하게 명암 값이 큰 영역은 영상에서 명암도가 높은 화소와 그렇지 않은 화소를 0과 1로 나누는 이진화 기법을 사용하여 간단하게 분리할 수 있다. 특히, 네트워크 이벤트 영상에서는 누적의 폭이 제각각이기 때문에 임계값을 동적으로 할당하여야 한다. 따라서 본 연구에서는 이러한 문제를 해결하기 위해 식 (1)을 이용하여 임계값을 결정하고, 주어진 패턴이 해당 임계값보다 높은 값을 가지면 이상패턴으로 판단한다.

$$th = \frac{1}{2N}\sum_{x=1}^{N}\sum_{y=1}^{N}\max(p) - p(x,y) \qquad (1)$$

식 (1)에서 N은 IP 영상의 한 축의 크기, 즉 256을 의미하며, p(x,y)는 영상의 각 화소의 명암 값을 의미한다. max(p)는 사용자가 미리 설정한 임계값이며, 만약 해당 임계값보다 더 큰 명암 값을 갖는 화소가 존재한다면 해당 화소의 명암 값을 임계값 max(p)로 사용한다. 식 (1)은 가장 큰 값에서 일반적인 값을 뺀 값으로 가장 큰 값과의 평균차이를 산출하며, 이를 이용하여 이진화를 수행하면 평균 차이보다 큰 값만을 구할 수 있다.



[Fig. 5] Noise elimination by intensity

Fig. 5는 이와 같은 방법을 이용하여 이진화한 결과를 보여준다. 해당 과정을 거쳐 추출된 패턴들에서 특징을 추출하여 트래픽의 공격을 분류하게 된다.

4. 특징 추출과 군집화

본 논문에서는 네트워크 트래픽의 공격을 탐지하기 위한 특징으로 이전 단계에서 생성한 IP와 포트에 대한 5개의 시각화된 영상으로부터 선형 패턴인 라인(line)과 명암값이 높은 패턴을 추출한다. 먼저, 시각화된 영상으로부터 라인을 효과적으로 검출하기 위해서는 기존에 성능이 검증되어 널리 활용되고 있는 허프 변환(Hough transform) [10]을 적용한다. 그리고 식 (2)를 이용하여 0과 1 사이의 값으로 정규화된 가장 긴 라인을 라인 특징 F_l로 사용한다. 식 (2)에서 $L(i)$는 i번째 라인을 의미한다.

$$F_l = \frac{\max(L(i))}{256} \qquad (2)$$

그런 다음, 5개의 시각화된 영상에서 식 (3)을 활용하여 명암값이 높은 패턴을 추출하여 명암값 특징 F_b로 사용한다. 식 (3)에서 N은 입력되는 패킷의 전체 개수를 나타내고, $p(x,y)$는 (x,y) 좌표에서의 화소의 명암값을 의미한다.

$$F_b = \frac{\max(p(x,y))}{N} \qquad (3)$$

보통, 트래픽 데이터를 2차원의 영상으로 시각화함으로써 네트워크의 공격을 용이하게 판단할 수 있으나, 이 방법은 트래픽 데이터의 양을 측정할 수가 없다는 단점이 존재한다. 따라서 본 논문에서는 군집화 알고리즘을 적용하여 트래픽 데이터의 양과 집중도를 측정하고, 이를 통해 보다 정확하게 네트워크 트래픽의 이상패턴을 분별

한다.

유사도를 판단하기 위한 군집화 알고리즘은 분할과 합병을 통해 군집의 개수를 유동적으로 설정하는 ISODATA 알고리즘[11]을 적용한다. 제안된 방법에서는 트래픽 데이터의 송신자와 수신자 포트에 대한 군집화를 수행하여 유사한 패턴을 가진 트래픽 데이터를 군집화한다. 트래픽 데이터를 군집화하기 위한 공간은 송신자 IP가 일정할 때 송신자 포트가 얼마나 분산되어 있는 지를 의미하는 송신자 포트의 분산도 D_x, 수신자 포트에 대한 분산도 D_y, 그리고 해당 분산도 (D_x, D_y)를 형성하는 패킷 데이터의 양을 의미하는 엔트로피(entropy)[12]의 3차원 공간으로 구성된다. 여기서, 엔트로피는 각각의 총 데이터에서 해당 분산도를 갖는 이벤트의 개수가 얼마나 많은지를 의미하며, 유사한 데이터들이 많이 존재할수록 엔트로피도 함께 증가한다.

ISODATA를 이용하여 군집화를 수행하면 x축 분산도, y축 분산도, 그리고 엔트로피를 이용하여 비슷한 유형을 가지는 데이터들을 하나의 군집으로 구성할 수 있으며, 같은 군집에 속한 데이터들은 네트워크상에서 비슷한 패턴을 가진다고 말할 수 있다.

결과적으로, 본 논문에서는 Table 1과 같이 수집된 24개의 패턴 특징들을 다음 장에서 언급할 SVM에 적용하여 네트워크 트래픽 데이터의 이상현상 검출 및 분류에 활용할 것이다.

[Table 1] Pattern features

Transmitter IP1	Maximum Intensity	Maximum line length	Average line length	Number of linear patterns
Transmitter IP2	Maximum Intensity	Maximum line length	Average line length	Number of linear patterns
Receiver IP1	Maximum Intensity	Maximum line length	Average line length	Number of linear patterns
Receiver IP2	Maximum Intensity	Maximum line length	Average line length	Number of linear patterns
Port image	Maximum Intensity	Maximum line length	Average line length	Number of linear patterns
	Dx	Dy	Maximum entropy	Number of clusters

5. SVM을 이용한 이상패턴 분류

5.1 표준 SVM

SVM은 모형학습용 데이터들을 두 집단으로 분류할 때 기본이 되는 분리 경계면(hyperplane)을 학습 알고리즘을 통해서 탐색하는 형태로 작동된다[13-18]. 좀 더 구체적으로, SVM은 학습용 입력벡터 x를 고차원의 특징공간으로 사상시킨 다음, 두 집단 사이의 폭(margin)을 최대화 시키는 분리 경계면을 찾는 것을 목적으로 한다. 이러한 최대 폭 분리 경계면(maximum margin hyperplane)은 두 집단 사이의 거리를 최대로 분리시킨다. 이 때 최대 마진 분리 경계면에 가장 근접한 데이터들로 모형의 학습에 반영되는 모형학습용 데이터를 서포트 벡터(support vector)라고 부른다.

위의 과정을 간단한 수식으로 살펴보면 다음과 같다. 우선 선형분리 문제에서 독립변수가 3개인 경우 분리 경계면은 다음의 식 (4)와 같이 표현될 수 있다.

$$y = \omega_0 + \omega_1 x_1 + \omega_2 x_2 + \omega_3 x_3 \quad (4)$$

여기서 y는 출력 값이고, x_i는 변수 값, 그리고 w_i는 학습 알고리즘에 의해 학습된 가중치이다. 상기 식에서 가중치 w_i는 분리 경계면을 결정하는 파라미터이다. 이 때 최대 마진 분리 경계면은 서포트 벡터를 사용해서 다음의 식 (5)와 같이 나타낼 수 있다.

$$y = b + \sum \alpha_i y_i x(i) \cdot x \quad (5)$$

여기서 y_i는 모형학습용 데이터 $x(i)$의 분류값이고, $\cdot$는 내적(dot product)이다. 한편, 벡터 x는 모형검증용 데이터를 나타내고, 벡터 $x(i)$는 서포트 벡터를 나타낸다. 식 (5)에서 b와 α_i는 분리 경계면을 결정하는 파라미터이다. 이때 서포트 벡터를 찾아내고 파라미터 b와 α_i를 결정하는 것은 선형적으로 제약된 이차 계획문제(linearly constrained quadratic programming)를 해결하는 것과 같다.

앞서 말했듯이 SVM은 저차원의 입력변수를 고차원의 특징 공간으로 이동시킴으로써 비선형 문제를 선형 분류기로 분류하도록 설계되어 있다. 그리고 비선형 분류 문제에서 사용하게 될 식 (5)의 고차원적인 형태는 다음 식 (6)과 같이 나타낼 수 있다.

$$y = b + \sum \alpha_i y_i K(x(i), x) \quad (6)$$

식 (6)에서 함수 $K(x(i),x)$는 커널함수(kernel function)라고 정의된다. 커널 함수는 원래 데이터를 고차원 공간으로 사상시킴으로써 특징 공간 내에 선형으로 분리 가능한 입력 데이터 셋을 형성한다. 이 때 사용될 수 있는 커널함수는 여러 가지가 있으며, 어떤 커널함수를 선택하는 것이 바람직한가는 주어진 문제에 따라 상이하고, 이

는 SVM을 적용하는 데 있어서 가장 중요한 요소 중의 하나이다. 일반적으로 많이 사용되는 커널함수로는 선형 함수(linear function)와 다항식 함수(polynomial function), 그리고 가우시안 RBF 함수(Gaussian radial basis function)를 들 수 있으며, 각 함수식은 아래와 같다.

$$\text{선형함수: } K(x,y) = xy \tag{7}$$

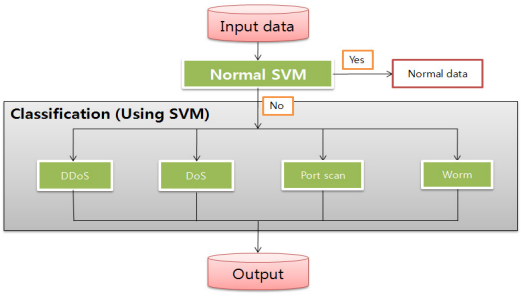
$$\text{다항식함수: } K(x,y) = (xy+1)^d \tag{8}$$

$$\text{가우시안함수: } K(x,y) = \exp(-1/\sigma^2 \cdot (x-y)^2) \tag{9}$$

여기서 d는 함수의 차수이고, σ^2은 가우시안 RBF 함수의 대역폭이다. 그리고 분리 가능한 문제에 있어 상기 식의 계수 α_i의 하한은 0이다. 분리가 불가능한 문제에서 SVM은 계수 α_i의 하한 이외에 상한 C를 추가함으로써 일반화된 결과를 얻을 수 있다.

5.2 다중 클래스 SVM의 구조

본 논문에서는 실제 네트워크 환경에 적용 가능한 SVM 기반의 네트워크 트래픽 공격 탐지와 세분화된 트래픽 공격 유형별 분류에 대하여 설명한다. 다음의 Fig. 6과 같이 여러 개의 단일 클래스 SVM을 이용하여 다중 클래스 SVM을 만들고, 이를 이용하여 여러 가지 트래픽 공격을 분류하도록 설계하였다. 그리고 일차 학습 시에는 한 개의 단일 클래스 SVM을 이용하여 정상 트래픽만을 학습시킨 후, 만일 입력된 데이터가 정상 트래픽 데이터가 아닌 경우에는 공격이 발생했다는 경보를 보낸다. 그런 다음, 사전에 각각의 공격 유형 별로 학습된 다중 클래스 SVM을 이용하여 입력된 트래픽 데이터에 대해 어떤 공격에 속하는지를 판단하게 한다.



[Fig. 6] Multi-class SVM

5.3 다중 클래스 SVM을 이용한 공격 탐지

본 논문에서는 다중 클래스 SVM을 이용하여 네트워크 트래픽 데이터를 정상 트래픽, DDoS, DoS, 인터넷 웜, 포트 스캔의 5가지 유형으로 분류하는데, 아래 Table 2에서 정의한 것과 같이 해당되는 출력노드의 값에 따라서 트래픽 공격의 종류를 판단한다.

[Table 2] Output node and attack types

Output node	Attacks
0000	Normal traffic
1000	DDoS
0100	DoS
0010	Internet worm
0001	Port scan

6. 실험결과

본 논문의 실험을 위한 컴퓨터는 인텔 Pentium Core 2 Duo의 2.66GHz CPU와 8GB의 메모리를 사용하였고, 운영체제로는 마이크로소프트사의 Windows 7을 사용하였다. 그리고 본 논문에서 제안된 방법을 구현하기 위해서 Visual C++ 2008 통합 개발 환경을 이용하였다. 학습에 사용된 데이터는 인위적으로 공격 특징에 맞게 매개의 데이터마다 10,000개의 IP와 포트 정보가 들어가 있도록 종류별로 나누어서 200개씩 만들어 사용하였다. 실험은 인위적으로 만들어진 공격특징별로 각각 100개의 데이터에 대해 이루어졌다.

Table 3에서부터 Table 7까지는 정상 트래픽과 4가지 트래픽 공격의 패턴 특징 값들의 실제 예를 보여준다. 이는 여러 단계의 전처리(pre-processing) 과정을 거쳐서 최종적으로 추출된 특징 값이며, 이 값들을 이용하여 네트워크 트래픽 공격의 분류 및 탐지를 수행한다.

[Table 3] Pattern(normal traffic)

Transmitter IP1	0.0250	0.0000	0.0000	0.0000
Transmitter IP2	0.0100	0.0000	0.0000	0.0000
Receiver IP1	0.0390	0.0000	0.0000	0.0000
Receiver IP2	0.0320	0.0000	0.0000	0.0000
Port image	0.0120	0.0000	0.0000	0.0000
	0.8790	0.8132	0.0523	0.5000

[Table 4] Pattern(DDoS)

Transmitter IP1	0.4850	0.0000	0.0000	0.0000
Transmitter IP2	0.0100	0.5390	0.3252	0.4000
Receiver IP1	0.4390	0.0000	0.0000	0.0000
Receiver IP2	0.4920	0.0000	0.0000	0.0000
Port image	0.0120	0.0000	0.0000	0.0000
	0.4790	0.1132	0.8523	0.4000

[Table 5] Pattern(DoS)

Transmitter IP1	0.1250	0.0000	0.0000	0.0000
Transmitter IP2	0.2100	0.0000	0.0000	0.0000
Receiver IP1	0.1390	0.0000	0.0000	0.0000
Receiver IP2	0.1320	0.0000	0.0000	0.0000
Port image	0.4120	0.0000	0.0000	0.0000
	0.7680	0.1528	0.7523	0.3000

[Table 6] Pattern(Internet Worm)

Transmitter IP1	0.4850	0.0000	0.0000	0.0000
Transmitter IP2	0.4390	0.0000	0.0000	0.0000
Receiver IP1	0.1290	0.0000	0.0000	0.0000
Receiver IP2	0.1100	0.5390	0.3252	0.4000
Port image	0.0520	0.0000	0.0000	0.0000
	0.3790	0.4132	0.0523	0.5000

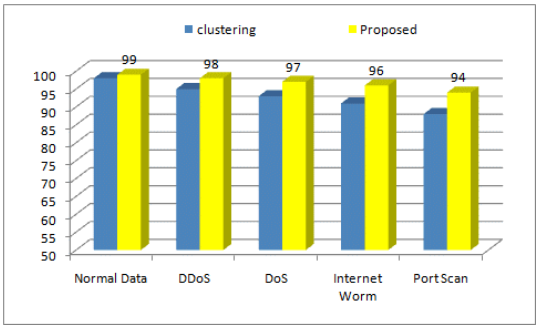
[Table 7] Pattern(port scan)

Transmitter IP1	0.1250	0.0000	0.0000	0.0000
Transmitter IP2	0.2100	0.0000	0.0000	0.0000
Receiver IP1	0.1390	0.0000	0.0000	0.0000
Receiver IP2	0.1320	0.0000	0.0000	0.0000
Port image	0.4120	0.0000	0.0000	0.0000
	0.1820	0.7570	0.6523	0.5000

본 논문에서는 다중 클래스 SVM의 학습 과정에서 감마(gamma) 값의 범위는 $2^{-15} \sim 2^{3}$이고, 비용 함수(cost function) 값의 범위는 $2^{-5} \sim 2^{13}$으로 설정하였다. 이와 같이 총 19×19개의 경우가 발생하며, 그중에서 파라미터 값이 최대가 될 때의 감마와 비용 함수의 값을 선택하도록 하였다. 실험을 통해 얻어진 값마의 최적화 값은 8.0이고, 비용 함수의 최적화 값은 1.0이다. 본 논문에서는 제안된 네트워크 트래픽 공격 탐지 알고리즘의 성능을 정량적으로 측정하기 위해서 식 (10)과 같은 정확도 척도를 사용하였다.

$$Rate = \frac{correctly\ detected\ attack\ data}{number\ of\ total\ traffic\ data} \times 100 \quad (10)$$

식 (10)에서 공격 탐지율 *Rate*은 전체 트래픽의 실험 데이터 중에서 공격의 유형을 정확하게 인식한 비율을 백분율(%)로 표시하였다.



[Fig. 7] Performance evaluation graph

Fig. 7은 식 (10)을 이용하여 기존의 방법 중 군집화까지의 단계만을 이용한 방법과 제안한 SVM 기반의 방법이 네트워크 트래픽 데이터 공격의 유형을 얼마나 잘 인식하는지를 그래프로 보여준다. Fig. 7에서 확인할 수 있듯이 기존의 방법에 비해 제안된 방법이 보다 정확하게 트래픽의 공격을 검출함을 확인할 수 있다.

Fig. 7에서 정상 트래픽 데이터는 공격 트래픽 데이터와 아주 선명한 차이가 있기 때문에 탐지율이 매우 높다. 제안된 방법에서는 트래픽의 공격 탐지를 시작할 때 우선 정상 데이터와 공격 데이터의 분류를 진행한다. 그런 다음, 공격의 유형 4가지에 대해 다시 분류를 진행한다. 그리고 특징 추출 단계에서 잡음이 제거된 특징 데이터를 사용하기 때문에 최적화된 상태의 데이터를 이용할 수 있고, 따라서 좋은 공격 탐지를 산출할 수 있다.

7. 결론

본 논문에서는 네트워크 이벤트를 시각화하는 것과 분산도를 이용하여 네트워크 이상현상의 독특한 패턴을 자동으로 찾고 패턴의 특징을 추출 및 학습하여 자동으로 이상현상의 공격 형태를 인식하는 방법을 제시하였다. 본 논문에서는 자동 인식 및 분류 분야에서 가장 많이 사용되고 있는 기계학습 방법 중의 하나인 다중 클래스 SVM을 이용하여 자동으로 공격을 탐지하는 방법을 시도하였으며, 실험결과 우수한 성능을 보였다. 다중 클래스 SVM을 사용하면 변종 공격 또는 전혀 새로운 공격이 발생한 경우에도 추가적인 학습과 단일 SVM 클래스의 추가만으로 탐지할 수 있으며, 학습시간은 다른 학습방법에 비해 빠르다는 장점을 갖고 있다. 그리고 여러 이상현상의 탐지를 위해 이상현상 각각에 대해 연산할 필요없이 한 번의 연산으로 모두 탐지가 가능하다는 장점이 있다. 그러나 SVM을 비롯한 대부분의 기계학습 방법은 자동탐지를

위한 학습과정이 선행되어야 하며, 이 학습과정이 충분하지 않거나 잘못된 데이터를 사용하여 학습할 경우에는 좋은 결과를 기대하기 어렵다는 단점이 있다. 본 논문에서 생성한 시스템의 실험 시 적은 양의 학습데이터를 이용하여 학습하고 실험하였으나 SVM은 특징상 적은 양의 학습 데이터만으로도 신속하게 분별 학습을 수행할 수 있다는 장점이 있다. 전반적인 이상현상의 검출결과가 비교적 높게 산출된 것이 이를 증명한다.

본 논문에서 제안하는 방법을 이용하여 실험한 결과 정상 트래픽과 공격 사이의 분류는 매우 효과적으로 수행되나 공격들 사이의 탐지율은 오탐지 경우가 일부 발생하였다. 이는 학습데이터의 양이 적어서 나타나는 현상이라고 볼 수 있으나 패턴 특징의 산출 방법에 문제가 있어서 오탐지 현상이 발생하는 경우도 간과할 수 없다. 따라서 지속적으로 학습을 수행하고 많은 데이터를 이용하여 실험을 진행하여 특징의 산출 방법을 개선하거나 추가하는 방법을 지속적으로 연구해야 할 것이다.

네트워크 트래픽의 시각화 영상은 사용자가 이상현상을 즉각적으로 인지하기에 알맞다. 그러나 이상현상이 발생할 것을 미리 예측하는 부분은 아직 연구가 필요하다. 현재 제안한 방법은 공격이 발생하였을 경우에 공격을 탐지하지만 사람이 조치할 무렵에는 이미 공격이 발생한 후라고 볼 수 있다. 따라서 예측 알고리즘을 적용하거나 새로운 예측 알고리즘을 생성하여 본 논문에서 제안하는 방법에 적용시키는 방법을 연구하면 이상현상 발생 초기에 탐지하여 큰 피해가 발생하지 않게 조치할 수 있을 것으로 본다.

References

[1] Y. Hai, "Study on Distributed Denial of Service Attack Detection Model Based on PCA and GA-Artificial Neural Network," Lecture Notes in Electrical Engineering, Vol. 113, No. 2, pp. 1181-1188, 2012.
DOI: http://dx.doi.org/10.1007/978-94-007-2169-2_139

[2] X. Yin, W. Yurcik, and A. Slagell, "The Design of VisFlowConnect-IP: A Link Analysis System for IP Security Situational Awareness," In Proc. of the IEEE International Information Assurance Workshop, pp. 23-24, Mar. 2005.
DOI: http://dx.doi.org/10.1109/IWIA.2005.17

[3] A.-S. Jin, J.-Y. Choi, H.-I. Choi, "Automatic Attack Detection based on Improved ISODATA Algori- thm," In Proc. of the Summer Conference of the Korea Society of Computer and Information, Vol. 18, No. 2, pp. 169-172, Jul. 2010.

[4] E. Corchado and Á. Herrero, "Neural Visualization of Network Traffic Data for Intrusion Detection," Applied Soft Computing, Vol. 11, No. 2, pp. 2042–2056, Mar. 2011.
DOI: http://dx.doi.org/10.1016/j.asoc.2010.07.002

[5] S. M. Lee, D. S. Kim, J. H. Lee, and J. S. Park, "Detection of DDoS Attacks Using Optimized Traffic Matrix," Computers and Mathematics with Applications, Vol. 63, No. 2, pp. 501-510, Jan. 2012.
DOI: http://dx.doi.org/10.1016/j.camwa.2011.08.020

[6] Y. Xie and S.-Z. Yu, "Monitoring the Application-Layer DDoS Attacks for Popular Websites," IEEE/ACM Transactions on Networking, Vol. 17, No. 1, pp. 54-65, Feb. 2009.
DOI: http://dx.doi.org/10.1109/TNET.2008.925628

[7] T. Gamer, "Collaborative Anomaly-based Detection of Large-Scale Internet Attacks," Computer Networks, Vol. 56, No. 1, pp. 169-185, Jan. 2012.
DOI: http://dx.doi.org/10.1016/j.comnet.2011.08.015

[8] S.-W. Jang, G.-Y. Kim, and H.-S. Na, "Detecting Abnormal Patterns of Network Traffic by Analyzing Linear Patterns and Intensity Values," Journal of the Korea Society of Computer and Information, Vol. 17, No. 5, pp. 21-28, May 2012.
DOI: http://dx.doi.org/10.9708/jksci.2012.17.5.021

[9] S. Lou, X. Jiang, and P. J. Scott, "Algorithms for Morphological Profile Filters and Their Comparison," Precision Engineering, Vol. 36, No. 3, pp. 414-423, July 2012.
DOI: http://dx.doi.org/10.1016/j.precisioneng.2012.01.003

[10] B. Li, K. Peng, X. Ying, and H. Zha, "Vanishing Point Detection Using Cascaded 1D Hough Transform from Single Images," Pattern Recognition Letters, Vol. 33, No. 1, pp. 1-8, 2012
DOI: http://dx.doi.org/10.1016/j.patrec.2011.09.027

[11] Q. Liu, Z. Zhao, Y.-X. Li, and Y. Li, "Feature Selection Based on Sensitivity Analysis of Fuzzy ISODATA," Neurocomputing, Vol. 85, pp. 29-37, May 2012.
DOI: http://dx.doi.org/10.1016/j.neucom.2012.01.005

[12] B. N. Subudhi, P. K. Nanda, and A. Ghosh, "Entropy-based Region Selection for Moving Object Detection," Pattern Recognition Letters, Vol. 32, No. 15, pp. 2097-2108, Nov. 2011.
DOI: http://dx.doi.org/10.1016/j.patrec.2011.07.028

[13] X. Peng, "TPMSVM: A Novel Twin Parametric-Margin Support Vector Machine for Pattern

Recognition," Pattern Recognition, Vol. 44, No. 10 - 11, pp. 2678-2692, Oct. - Nov. 2011.
DOI: http://dx.doi.org/10.1016/j.patcog.2011.03.031

[14] S. Lee, "The Study on the Error Rate Analysis for the Occupied Bandwidth of Internet Real-time Traffic", Journal of The Institute of Webcasting, Internet and Telecommunication, Vol 12, No 4, pp. 167~172, 2012.

[15] C. Lim, "TCP Performance Improvement in Network Coding over Multipath Environments", Journal of The Institute of Webcasting, Internet and Telecommunication, Vol 11, No 6, pp. 81~86, 2011.

[16] C. Lim, "Effectiveness of DUPACK-independent TCP in Coded Wireless Mesh Networks", Journal of The Institute of Webcasting, Internet and Telecommunication, Vol 11, No 1, pp. 8~13, 2011.

[17] N. T. Tung, I. Koo, "Fuzzy-based Dynamic Packet Scheduling Algorithm for Multimedia Cognitive Radios", Journal of The Institute of Webcasting, Internet and Telecommunication, Vol 12, No 3, pp. 1~7, 2012.

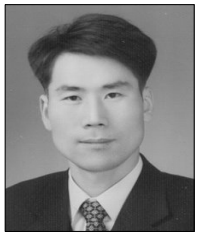
[18] H. Hwang, S.-C. Kim, "Design and Implementation of Unified Network Security System support for Traffic Management", Journal of The Institute of Webcasting, Internet and Telecommunication, Vol 11, No 6, pp. 267~273, 2011.

박 영 재(Young-Jae Park) **[정회원]**

- 2005년 2월 : 청운대학교 컴퓨터과학과 (공학사)
- 2008년 2월 : 숭실대학교 컴퓨터학과 (공학석사)
- 2008년 3월 ~ 현재 : 숭실대학교 컴퓨터학과 박사과정

<관심분야>
컴퓨터 비전, 영상처리, 패턴인식 등

김 계 영(Gye-Young Kim) **[정회원]**



- 1990년 2월 : 숭실대학교 전자계산학과(공학사)
- 1992년 2월 : 숭실대학교 컴퓨터학과(공학석사)
- 1996년 2월 : 숭실대학교 컴퓨터학과(공학박사)
- 2001년 3월 ~ 현재 : 숭실대학교 컴퓨터학부 교수

<관심분야>
컴퓨터비전, 생체인식, 증강현실, 신호처리 등

장 석 우(Seok-Woo Jang) **[정회원]**

- 1995년 2월 : 숭실대학교 전자계학과 (공학사)
- 1997년 2월 : 숭실대학교 컴퓨터학과 (공학석사)
- 2000년 8월 : 숭실대학교 컴퓨터학과 (공학박사)
- 2009년 3월 ~ 현재 : 안양대학교 디지털미디어학과

<관심분야>
로봇비전, 증강현실, HCI, 비디오 색인 및 검색, 등