한국컴퓨터정보학회 하계학술대회 논문집 제18권 제2호 (2010. 7)

개선된 ISODATA 알고리즘을 이용한 공격 자동탐지

김애숙°, 최재영**, 최형일*
°*숭실대학교 미디어학과
**숭실대학교 컴퓨터학과
e-mail: aishu0208@gmail.com, {choi, hic}@ssu.ac.kr

Automatic Attack Detection based on Improved ISODATA Algorithm

Ai-Shu Jin°, Jae-Young Choi**, Hyong-Il Choi*
°*Dept. of Media, Soongsil University
**Dept. of Computer, Soongsil University

● 요 약 ●

본 논문에서는 기존의 ISODATA 알고리즘을 네트워크 공격탐지에 더욱 적합하도록 개선하여 공격을 탐지하는 새로운 방법을 제안한다. 수많은 인터넷상의 트래픽 정보들을 군집화하여 유사도를 비교하는 방법을 통해 공격을 판단한다. 기본적인 절차는 송신자 IP와 Port, 수신자 IP와 Port 정보를 이용하여 송신자와 수신자 사이의 관계를 분석하고 그 특징 값들을 이용하여 개선된 군집화 알고리즘을 이용하여 군집화를 수행한다. 그리고 얻어진 패턴의 특징값을 인공신경망에 학습하여 공격유형을 분류하고 탐지하도록 한다. 기존의 공격탐지 방법과 비교했을 때, 계산양이 적고 속도가 빠르다는 장점이 있으며 제안하는 방법의 우수성을 실험을 통해 증명하였다.

키워드: 공격탐지(attack detection), 군집화(clustering)

I. 서론

네트워크 기술의 발전과 인프라의 급속한 확산으로 인해 현대 사회는 인터넷이 없는 세상을 상상할 수도 없게 되었다. 경제, 사회, 문화, 교육 등에 관련된 수많은 서비스를 인터넷을 통해 제공받을 수 있으며 인터넷을 이용한 다양한 활동들이 지속적으로 시도, 연구되고 있다. 하지만 인터넷의 이러한 장점과 더불어 단점 또한 존재하는데 네트워크 공격이 그 중 하나이다. 인터넷의 급속한 확산과 함께 네트워크 공격은 나날이 증가하고 있으며, 공격 성격도 단순 악의적 목적에서 협박을 통한 금품갈취 증의 용도로 사용되는 등 여타 다른 범죄와 혼합되어 나타나는 양상을 보이고 있다. 이와 같이 사회 전 분야의 네트워크 의존화가 극대화에 다다른 오늘날 네트워크 공격은 네트워크를 사용하는 모든 사업체에 큰 위협이 되고 있으며, 최근에는 공격 등 대상이 사업체에서 국가로 확대되어 각 국가에서도 무시할 수 없는 심각한 문제가 되었으며 네트워크 공격에 대한 방어/탐지 기술이 절실히 요구되고 있다.

따라서 이러한 피해를 막기 위해 인터넷상의 방대한 트래픽 데이터들을 효과적으로 분석하고 악성 공격 트래픽에 대해 빠르게 인지, 대응하는 것이 중요하다[1]. 그러나 수많은 정보 중에서 자신이 원하는 정보만을 빠르게 분석하고 탐지하는 것은 쉽지 않다. 이와 관련해 최근에는 정보시각화에 대한 연구가 활발히 진행되고 있으며 그 목적과 방식에 따라 다양한 기법들이 개발되고 있다[2]. 하지만 IP와 Port 정보의 특정상 시각화를 통해 분석하기에는 어려움이 있다. IP는 4개의 숫자로 구성되었기 때문에 하나의 영상으로 구현하기에는 부적합하고 Port는 하나의 0~65535사이의 숫자로 구성되었기 때문에 직접 영상으로 표현하기에는 너무 큰 숫자이기 때문에 어려움이 있다.

이런 문제점을 해결하기 위해 본 논문에서는 개선된 군집화 방법을 이용하여 송신자와 수신자 사이의 관계를 정확하게 분석하고, 유사도 비교를 통하여 공격분류를 하는 방법을 제안한다. 우선 송신자와 수신자 의 IP와 Port 각각에 대해 분산도를 측정하고, 그 분산도를 구성하는 데이터의 양을 계산하여 특징 값을 추출하도록 한다. 예를 들어 여러 개의 송신자 IP가 하나의 수신자 IP에 접근하였다면, 송신자 IP의 분산도는 1에 가까운 값을 갖게 되고 수신자 IP의 분산도는 0에 가까운 값을 갖게 된다. 그리고 데이터의 양을 나타내는 엔트로피 값을 구하여 군집화 과정을 거쳐 트래픽의 패턴특징을 추출한다. 추출된 패턴의 특징 값을 이용하여 인공신경망에 학습을 하는 방법으로 공격분류와 탐지를 하는 새로운 방법을 제안한다. 그림 1은 전체 시스템의 개요도이다.

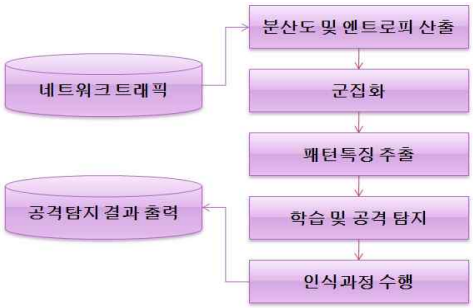


그림 1. 시스템 개요도
Fig. 1. Flow chart of proposed method

II. 공격 특징 분석

본 논문에서는 인터넷 공격 중 가장 흔히 발생하는 DDoS(분산 서비스 거부공격), DoS(서비스 거부공격), Worm, 포트스캔 4가지 종류의 공격에 대해 자동분류 및 탐지를 하는 방법을 제안한다. 우선 각각의 공격특징에 대해 설명하면 아래와 같다.

DDoS 공격은 다수의 감염된 호스트가 피해 호스트에게 다량의 무의미한 패킷을 전송하여, 피해 호스트와 인터넷 사이의 자원 불균형을 초래한다. 여러 대의 공격자를 분산 배치하여 동시에 동작하게 함으로써 특정 사이트를 공격하는 공격방식이다. 분산 서비스 공격은 여러 개의 부동한 송신자 IP가 동시에 하나의 수신자 IP를 공격하기 때문에 송신자 IP와 수신자 IP의 관계는 N:1 이다.

DoS 공격은 피해 호스트가 인터넷에 정상적인 서비스를 제공하거나 서비스를 받는 것을 방해하는 공격이다. 시스템을 악의적으로 공격해 해당 시스템의 자원을 부족하게 하여 원래 의도된 용도로 사용하지 못하게 하는 공격이다. 특정 서버에게 수많은 접속 시도를 만들어 다른 이용자가 정상적으로 서비스 이용을 하지 못하게 하거나, 서버의 TCP 연결을 바닥내는 등의 공격이 이 범위에 포함된다. DoS 공격은 한 대의 송신자 컴퓨터가 여러 Port를 이용하여 한 대의 수신자 컴퓨터에 접속한다는 것이다. 따라서 송신자와 수신자 IP의 관계는 1:1이고, Port의 관계는 N:1 이다.

Worm 공격은 스스로를 복제하는 컴퓨터 프로그램으로 컴퓨터 바이러스와 비슷하다. 단 바이러스가 다른 실행 프로그램에 기생하여 실행되는 데 반해 Worm 공격은 독자적으로 실행되며 다른 실행 프로그램이 필요하지 않다. 일반적으로 Worm 공격은 네트워크를 손상시키고 대역폭을 잠식한다. Worm 공격의 특징을 보면 DDoS와 반대로 한대의 컴퓨터가 여러 대의 컴퓨터에 접속하여 감염시킨다. 즉 하나의 송신자 IP를 통해 여러 개의 수신자 IP를 방문한다는 것이다. 때문에 송신자 IP와 수신자 IP의 관계는 1:N 이다.

포트 스캔은 몇몇 네트워크 공격의 준비과정으로 사용된다. 공격을 실행하기 앞서 공격자는 목표 호스트의 열려있는 port를 찾기 위해서 포트 스캔을 수행한다. 공격자는 선택된 호스트의 열려있는 port를 이용하여 네트워크 공격을 수행한다. 포트 스캔인 경우, 특정한 송신자에서 특정한 수신자를 향하여 다양한 수신자 port로 패킷이 전송된다. 때문에 송신자 Port와 수신자 Port의 관계는 1:N 이다.

III. 개선된 ISODATA 알고리즘을 이용한 군집화

ISODATA 알고리즘은 K-Means 알고리즘의 단점인 군집 개수를 정해주어야 하는 문제점을 해결하고, 그림 2와 같이 분할 또는 병합을 통해 군집 개수를 유동적으로 설정해 주기 때문에 네트워크 패킷 데이터의 분산도와 같이 군집 개수가 일정하지 않은 경우에 사용하기 알맞은 알고리즘이라 할 수 있다.

ISODATA 알고리즘의 군집화 과정은 아래와 같다.

1단계: 클러스터링 할 데이터들의 공간에 초기 클러스터 중심을 생성한다. 2단계: 클러스터 중심과 데이터의 거리를 각각 산출한다. 3단계: 데이터양이 일정 이하인 초기 클러스터 중심을 제거한다. 4단계: 클러스터 중심에 해당하는 데이터들의 위치값을 가지고 중심을 갱신한다. 5단계: 각 클러스터 중심과 데이터들의 위치값의 평균거리 D_j를 산출한다. 6단계: 산출한 모든 평균거리 D_j를 이용하여 전체 평균거리 D를 산출한다. 7단계: 클러스터의 개수가 임계치보다 작은지 확인한다. 8단계: 각 클러스터의 분산도를 측정한다. 9단계: 클러스터의 분산도가 클 경우, 중심에서 가장 먼 데이터와 해당 데이터에서 가장 먼 데이터를 이용하여 새로운 중심을 2개 생성한다. 10단계: 클러스터의 개수가 임계치보다 큰지 확인한다. 11단계: 클러스터간의 거리를 산출한다. 12단계: 클러스터의 개수가 임계치보다 작을 때까지 클러스터간의 거리가 가까운 클러스터들을 병합한다.

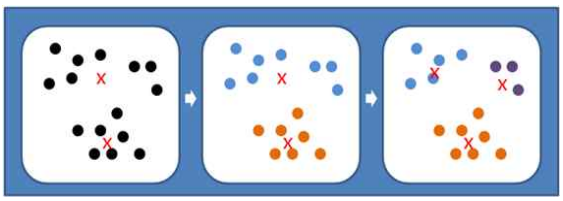


그림 2. ISODATA 알고리즘 분할의 예
Fig. 2. The Split Example of ISODATA Algorithm

본 논문에서 공격을 인식하기 위하여 송신자 IP, 송신자 Port, 수신자 Port, 수진자 IP가 얼마나 분산되어 있는지를 의미하는 분산도 Dx와 Dy, 그리고 해당 분산도를 이루는 패킷 데이터들의 양을 의미하는 엔트로피로 구성되어 있다. 우선 분산도는 총 6개로 분류되는데, 각각 동일한 송신자 IP와 송신자 Port를 가지고 있는 데이터의 수신자 Port와 수신자 IP의 분산도를 의미하는 Src-Spt-*-*, 동일한 송신자 IP주소와 수신자 Port를 가지고 있는 데이터의 송신자 Port와 수신자 IP의 분산도를 의미하는 Src-*-Dpt-*, 동일한 송신자 IP주소와 수신자 IP 주소를 가지고 있는 데이터의 송신자 포트 번호와 수신자 포트 번호의 분산도를 의미하는 Src-*-*-Dst, 동일한 송신자 포트 번호와 수신자 포트 번호를 가지고 있는 데이터의 송신자 IP 주소와 수신자 IP 주소의 분산도를 의미하는 *-Spt-Dpt-*, 동일한 송신자 포트 번호와 수신자 IP 주소를 가지고 있는 데이터의 송신자 IP 주소와 수신자 포트 번호의 분산도를 의미하는 *-Spt-*-Dst, 동일한 수신자 포트 번호와 수신자 IP 주소를 가지고 있는 데이터의 송신자 IP 주소와 송신자 포트 번호의 분산도를 의미하는 *-*-Dpt-Dst로 나뉜다.

예를 들어 송신자 IP가 1.1.1.1, 송신자 포트가 30, 수신자 포트가 80, 수신자 IP가 2.2.2.2의 트래픽 데이터가 10개 존재한다고 가정하고 , Src-Spt-*-* 데이터 형식으로 분산도를 산출한다면 10개 모두 동일하기 때문에 Dx와 Dy 모두 0.1이 산출된다. 반대로 송신자 IP가 1.1.1.1, 송신자 포트가 30, 수신자 포트가 80이고 수신자 IP가 10개 모두 틀리다면, Dx는 0.1이 산출되고 Dy는 1이 산출된다. 이를 보다 간단하게 식으로 표현하면 식(1)과 같다.

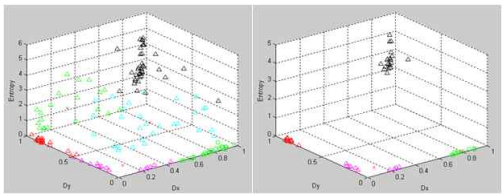
$$D_x = \frac{N_k}{N} \qquad (1)$$

식(1)에서 N은 전체 이벤트의 개수이고, N_k는 k 이벤트의 개수를 의미한다.

그리고 해당 분산도는 N의 최대값에 따라 그 중요도가 달라지게 되는데 이 중요도를 엔트로피로 사용한다. 엔트로피는 각각의 총 데이터에서 해당 분산도를 갖는 이벤트의 개수가 얼마나 많은지를 의미하며, 유사한 데이터들이 많이 존재할수록 엔트로피도 함께 증가한다. 이렇게 산출한 Dx의 엔트로피와 Dy의 엔트로피를 식(2)를 이용하여 통합 후 군집화시 함께 적용한다.

$$Ent_{(x,y)} = \sqrt{D_x \bullet D_y} \qquad (2)$$

산출된 Dx와 Dy 그리고 엔트로피를 이용하여 3차원 공간상에서 개선된 ISODATA 알고리즘을 이용하여 군집화를 수행한다. 군집화를 수행하면 x, y축의 고유 분산도, 엔트로피, Port 정보를 사용하여 비슷한 유형을 가지는 데이터들을 하나의 군집으로 구성할 수 있으며, 같은 군집에 속한 데이터들은 네트워크상에서 비슷한 패턴을 가지는 것으로 볼 수 있다.



(a) (b)

그림 3. (a)기존의 ISODATA 군집화 결과,
(b)개선된 ISODATA 군집화 결과.
Fig. 3. (a)Result of using Original ISODATA Algorithm,
(b)Result of using Improved ISODATA Algorithm.

본 논문에서 제안하는 개선된 ISODATA 알고리즘은 기존의 ISODATA 알고리즘의 군집화 과정에서 분렬 및 병합 과정을 없애고 대신 평균 분산도보다 큰 값을 갖는 데이터들을 제거시키는 과정을 추가하였다. 이런 방법을 이용하면 군집화가 더욱 선명하게 될뿐더러 군집의 특징값도 더욱 뚜렷하게 나타나게 된다. 그림 3은 기존의 ISODATA와 개선된 ISODATA의 군집화 결과를 비교한 예를 보여준다.

이렇게 군집화를 수행하면 각 이벤트 데이터들은 유사한 데이터들끼리 모여 군집을 이루게 된다. 군집의 중심점 x, y, z는 해당 군집을 이루는 데이터들의 평균 특징이라 볼 수 있으며, 이것은 이상현상을 판단하는데 가장 중요한 특징 중의 하나이다. 예를 들면 DDoS 공격 발생 시 항상 유사한 위치에 군집이 생성된다면 해당 군집은 DDoS를 의미한다고 보아도 무방하다. 따라서 Src-Spt-*-*, Src-*-Dpt-*, Src-* -*-Dst, *-Spt-Dpt-*, *-Spt-*-Dst, *-*-Dpt-Dst 이 6개의 공간에서 각각 생성된 모든 군집의 중심점 x, y, z의 값을 추출한다. 산출된 모든 군집의 중심점인 x, y, z값은 해당 위치를 의미하는 6개 공간과 프로토콜을 의미하는 패턴과 더해져 인공신경망에 학습하고 분류하여 인식하는데 사용된다. 표 1은 6개 공간에 따라 변환되는 군집의 중심점 x, y, z 앞에 붙는 패턴의 형태를 보여준다.

표 1. 공간 분류에 따른 패턴 형태
Table 1. Pattern Forms of Different Space Classification

공간분류	Src-Spt-*-*	Src-*-Dpt-*	Src-*-*-Dst	*-Spt-Dpt-*	*-Spt-*-Dst	*-*-Dpt-Dst
패턴 형태	1100	1010	1001	0110	0101	0011

산출된 군집의 중심점 x, y, z앞에 공간 분류에 따른 패턴을 배치하여 최종 특징 패턴을 생성한다. 예를 들어 만약 Src-Spt-*-* 공간에 생성되고 군집의 중심 x, y, z가 각각 0.5, 0.3, 0.1이라면 최종적으로 생성되는 특징 패턴의 형태는 1, 1, 0, 0, 0.5, 0.3, 0.1이 된다. 이와 같은 방법으로 모든 영상에 존재하는 군집의 중심점을 이용하여 학습 또는 인식을 위한 특징 패턴을 생성한다.

Ⅳ. 인공신경망을 이용한 학습 및 분류

본 논문에서는 인공신경망을 이용하여 정상데이터와 DDoS, DoS, Worm, 포트스캔 공격을 분류하도록 하였다. 인공신경망의 입력노드는 12개, 은닉노드는 16개, 출력노드는 5개를 사용하였으며 출력노드의 결과에 따라 표 2와 같이 분류하였다. 학습데이터로는 인위적으로 생성한 네트워크 공격 패킷의 헤더정보를 모은 데이터베이스를 사용하였다. 각각의 데이터베이스는 3000개~10000개 사이의 패킷 헤더 정보로 이루어 졌다. 해당 데이터베이스는 각 종류별로 20개의 공격패턴으로 구성되었다.

출력노드 값이 0.75 이상이면 1로, 그 미만이면 0으로 간주하고 표 2의 출력노드 값과 비교하여 공격 종류를 판단하였다. 표 3은 본 논문에서 제안한 방법으로 네트워크 공격을 자동으로 탐지한 결과를 보여준다. 탐지율은 실험데이터 모두를 정확하게 탐지 하였을 때를 100%로 하여 산출하였다.

표 2. 출력 값에 따른 공격유형 분류
Table 2. Output Node Value and Attack Types

출력노드값	0000	1000	0100	0010	0001
공격종류	정상트래픽	DDoS	DoS	Worm	포트 스캔

표 3. 공격유형 분류 정확도
Table 3. Accuracy of Attack Types Classification

정상패킷	DDoS	DoS	Worm	포트 스캔
99.2%	97.1%	92.7%	93.0%	90.4%

V. 결론

본 논문에서는 송신자 IP, 송신자 Port, 수신자 Port, 수신자 IP 정보를 이용하여 인터넷상의 공격을 탐지하는 방법을 제안하였다. 이 방법은 총 6개의 3차원 군집화 영상으로 모든 인터넷 트래픽의 정보를 표현하였으며 쉽게 특징을 추출할 수 있다는 장점이 있다. 또한 공격 유형 탐지도 학습과정을 통해 아주 좋은 결과가 나왔다.

향후 개선할 점이라면 사전의 공격예측과 더욱 정확한 공격 탐지율을 보장할 수 있는 연구가 필요하다.

참고문헌

[1] H. Kim, J. H. Kim, S. Bahk, I. Kang, “Fast Classification, Calibration, and Visualization of Network Attacks on Backbone Links”, LNCS 3090, pp.837-846, 2004.

[2] C. Y. Jeong, B. H. Chang, J. C. Na, “Visualization Technology of Network Security Events”, 전자통신동향분석 제23권 제4호, 2008.

[3] R. Fontugne, T. Hirotsu, K. Fukuda, “An image processing approach to traffic anomaly detection”, Proceedings of 8th Asian Internet Engineering Conference (AINTEC 2008), pp.17-26, Bangkok, Thailand, Nov.18-20, 2008.

[4] S. Farraposo, P. Owezarski, E. Monteiro, “A Multi-Scale Tomographic Algorithm for Detecting and Classifying Traffic Anomalies”, in Proc. of the IEEE International Conference on Communications (ICC), IEEE, 2007.