자기 조직화 지도(SOM) 기반 실시간 DoS 트래픽 탐지 시스템

신무곤, 심규석, 구영훈, Huru Hasanova, 김명섭
고려대학교

{tm0309, kusuk007, gyh0808, hhuru, tmskim} @ korea.ac.kr

Real Time DoS Traffic Detection System Based on Self Organizing Map(SOM)

Mu-Gon Shin, Kyu-Seok Shim, Young-Hoon Goo, Huru Hasanova, Myung-Sup Kim
Korea Univ.

요 약

오늘날 네트워크 기술의 발달로 인해 인터넷을 사용하는 컴퓨터의 수가 폭발적으로 늘어나고 있다. 이에 따라 다양한 형태의 네트워크상의 공격이 발생하고 그 피해는 매우 커지고 있는 실정이다. 따라서 네트워크상의 컴퓨터에 대한 보안 및 침입탐지 기술이 요구되고 있다. 네트워크 기반의 공격은 그 위험성과 피해의 규모가 크기 때문에 공격 초기에 빨리 탐지하는 것이 중요하다. 네트워크상의 비정상 트래픽을 탐지하는 것은 방대한 양의 데이터 전처리와 관리자의 분석이 요구된다. 또한 관리자의 분석이 정확하다는 보장이 없을 뿐만 아니라 각 네트워크의 실시간 특성을 고려해야하기 때문에 정확한 탐지의 어려움이 크다. 이러한 한계를 극복하기 위해 본 논문에서는 데이터 마이닝 기법 중 하나인 자기 조직화 지도(SOM)를 활용한 비정상 트래픽(DoS) 분류 시스템을 제안한다. 제안된 시스템은 분류 과정에서 네트워크의 실시간 특성을 반영한 탐지에 대한 기대 효과를 나타낼 수 있다.

Ⅰ. 서 론

네트워크 기반의 공격에 속하는 DoS 및 DDoS 공격은 타겟 시스템에 대량의 트래픽을 보냄으로써 다른 정상적인 이용자들이 이 시스템의 서비스를 제공받지 못하고, 네트워크 전체를 마비시킬 수 있는 위험한 공격이다. 이러한 공격들은 웜 바이러스 등 지능적인 공격툴을 이용하여 빠르게 확산되고, 그 피해가 매우 커지고 있는 실정이다. 따라서 이러한 공격 트래픽이 확산되기 전에 빨리 탐지하는 것이 중요하다 [2].

본 논문에서는 데이터 마이닝 기법 중 하나인 자기 조작화 지도(Self Organizing Map)을 활용한 실시간 DoS 트래픽 탐지 시스템을 제안한다. 이 방법은 별도의 관리자 분석이 필요없이 수집된 데이터를 사용하여 초기 학습을 통해 클러스터 맵을 형성한다. 이후의 학습에서는 학습과 함께 탐지가 이루어 진다. 이를 통해 분 단위 실시간 탐지가 가능하며 학습 데이터의 점진적인 갱신이 가능하는 점에서 장점을 가진다.

본 논문에서 제안하는 시스템은 다음의 세 단계로 이루어져 있다. 첫 번째 단계는 데이터 전처리 단계로 MS Net Monitor 프로그램을 통해 정상 트래픽과 DoS 트래픽을 수집 한 뒤 레이블 된 데이터로 만들어준다. 두 번째 단계는 탐지에 앞서 초기 학습 모듈을 구성하는 단계로 자기 조직화 지도를 활용하여 특징이 비슷한 트래픽끼리 클러스터링된 맵을 생성한다. 앞선 단계에서 만들어진 전처리 된 데이터를 가지고 각 맵에 비정상/정상 클러스터를 레이블링 한다. 마지막으로 레이블링된 맵을 활용하여 실시간 탐지와 함께 점진적인 학습이 이루어지게 된다.

본 논문은 본 장 서론에 이어, 2 장에서 자기 조직화 지도 기반 탐지 시스템 구조에 대해 제안하고, 마지막 3 장에서 결론 및 향후 연구에 대해 언급한 후 논문을 마친다.

이 논문은 2015 년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업(No.2015R1D1A3A01018057)과 2017 년도 정부(과학기술정보통신부)의 재원으로 정보통신기술진흥센터의 지원을 받아 수행된 연구임(No. 2017-0-00513, Security Analytics 기반의 이기종 보안솔루션 위협 분석 및 대응 기술 개발)

Ⅱ. 본론

1) 자기 조직화 지도(Self Organizing Map)

자기 조직화 지도(SOM)는 비지도 학습을 통한 데이터 마이닝 기법으로 스스로 학습하는 인간 뇌의 학습 과정과 유사한 성격을 가진다. SOM 은 외부로부터 데이터를 받아 특정 사이즈의 노드로 구성된 맵에 데이터의 특징들을 학습한다. 여러 특징들을 가진 데이터를 반복에서 학습하면, SOM 이 가지고 있는 맵에는 비슷한 패턴을 가진 데이터를 기억할 수 있는 학습 영역이 생긴다. 충분히 많은 데이터를 학습한 맵에 새로운 데이터를 학습하게 되면, 해당 데이터가 맵의 구분된 영역 중에 어떤 부분에 속하는지에 따라 어떤 특징을 가졌는지 예측할 수 있다 [1].

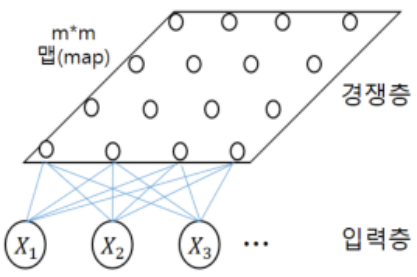

Figure 1. SOM 구조

SOM 은 입력층과 경쟁층으로 구분된 2 개의 레이어로 구성되어 있다. 입력층에서는 입력벡터를 입력 받고 경쟁층에서는 경쟁층 노드들과 입력벡터 사이의 거리를 계산하여 거리가 가까운 노드가 맵에서 선택된다. 노드가 선택된 후에는 해당 노드와 이웃노드들의 가중치 값을 수정한다. 이 과정을 반복 횟수가 끝날 때까지 반복한다[1].

2) 제안하는 자기 조직화 지도 기반 시스템 구조

자기 조직화 지도는 입력 데이터에 대해 어떠한 정보도 주지 않으므로 그 결과에 대한 해석이 힘들다는 단점이 있다. 따라서 본 논문에서는 자기 조직화 지도를 지도 학습으로 일부 보완하는 침입 탐지 시스템을 제안한다. 제안된 시스템은 비지도 학습을 사용하지만 클러스터 맵상의 구분을 위해 지도 학습의 특성을 이용한다. 이를 통해 비지도 학습만을 사용할 때의 문제점을 일부 보완할 수 있다.

전체적인 구조는 기존의 SOM 과 동일하지만, 초기 학습을 수행 할 때에 정답지가 레이블링 된 데이터를 입력한다. 클러스터링 된 맵에 정답을 같이 표기함으로써 결과에 대한 해석을 쉽게 할 수 있다. 초기 학습 과정을 거친 후에는 침입 탐지와 학습이 동시에 이루어 진다. 이와 같은 방법은 새로운 데이터가 입력되더라도 가장 유사한 클러스터에 일치시키기 때문에 유연한 탐지가 가능하며, 새로운 공격 유형이 발생하더라도 탐지가 가능하다는 장점이 있다[1].

Figure 2. 탐지 단계

학습을 통하여 만들어지는 클러스터 맵을 생성하는 단계에서 특정 속성이 맵 형성에 너무 많은 영향을 미치는 것을 방지하기 위해 모든 속성값을 0~1 사이의 값을 갖도록 하는 정규화 과정이 필요하다. 그리고 경쟁층 노드의 값을 초기화 시킨 후 입력벡터와 경쟁층 노드간의 거리가 가장 짧은 노드를 선택하고 그 노드와 이웃노드의 가중치 값과 학습율을 갱신한다. 입력벡터가 모두 학습될 때까지 이 과정을 반복한다. 이렇게 클러스터링 된 데이터를 가지고 공격별 정답지를 이용하여 레이블링 한 후 맵상의 위치를 구분할 수 있도록 한다[1].

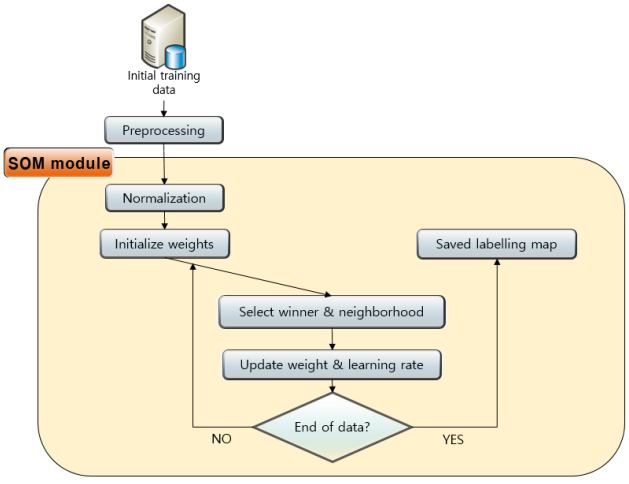

Figure 3. 초기 학습 모듈

이 단계까지 마치면 공격 탐지에 필요한 초기 학습 모듈이 형성된다. 이 모듈을 바탕으로 실시간 입력 데이터가 맵의 유사한 클러스터에 일치하게 되고 그 클러스터가 공격 클러스터라면 비정상, 정상 클러스터라면 정상 트래픽으로 탐지가 가능하게 된다. 또한 실시간 탐지와 함께 학습 또한 이루어지므로 계속해서 맵이 갱신되어 실시간 네트워크 트래픽의 특성을 반영할 수 있다.

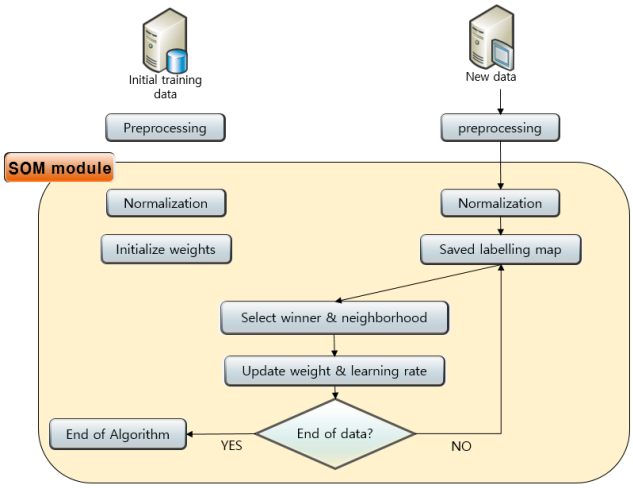

Figure 4. 탐지 & 학습

3) 실험 및 결과

MS Net Monitor 프로그램을 통해 정상 트래픽과 DoS 트래픽을 수집 한 뒤 본 연구팀이 개발한 프로그램을 통해 Flow with Packet 형태로 변환하고 레이블링 작업을 거친다. 이 데이터를 입력으로 하여 초기 SOM 학습을 진행하여 클러스터 맵을 만들어준다.

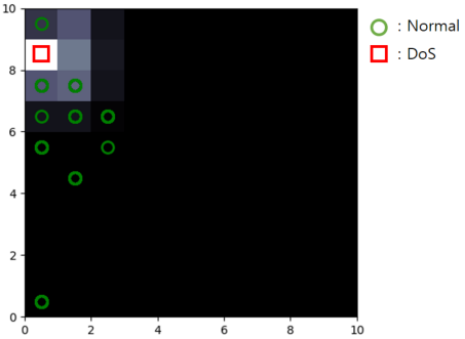

Figure 5. 초기 학습 모듈 클러스터 맵

Figure 5 에서 보이는 것처럼 정상 패킷과 DoS 패킷이 클러스터링 되어 초기 학습 모듈이 구성 된 것을 볼 수 있다.

Ⅲ. 결론 및 향후연구

기존의 자기 조직화 지도는 입력 데이터에 대한 어떠한 정보도 주지 않기 때문에 클러스터링 된 데이터 결과에 대한 해석이 힘들다. 이러한 단점 때문에 실시간 패킷 탐지와 학습에 어려움이 있다. 본 논문에서는 자기 조직화 지도 기반 실시간 DoS 트래픽 탐지 시스템을 제안하였다. 제안된 시스템은 적은 양의 학습 데이터만으로도 충분한 모듈을 구성할 수 있으며 학습에 대한 점진적 갱신이 가능하기 때문에 실시간 학습을 통한 유연한 패킷 탐지가 가능하다. 향후 연구로는 구성된 초기 학습 모듈을 바탕으로 실시간 DoS 패킷 탐지 시스템 구성 및 실험을 연구할 계획이다.

참 고 문 헌

[1] 신무곤, 심규석, 구영훈, Huru Hasanova, 김명섭, " SOM 알고리즘 기반 DoS 트래픽 탐지", KNOM Conference 2018, May, 2018.

[2] 황경애, "자기 조직화 지도(SOM)를 이용한 실시간 침입 탐지 메커니즘", 2005.

[3] 김민회, "분산 Self Organizing Map 기법을 이용한 효과적인 DoS 탐지 시스템", 2015.