고성능망에서 침입 탐지를 위한 효율적 트래픽 분류 기법 연구

박민우*, 이준호*, 조신영*, 정태명**
*성균관대학교 전자전기컴퓨터공학과
**성균관대학교 정보통신공학부
e-mail:mwpark@imtl.skku.ac.kr

A Study of Efficient Traffic Classification for Intrusion Detection in High Performance Network

Min-Woo Park*, Jun-Ho Lee*, Sinyoung Cho*, Tai-Myoung Chung**
*Dept. of Electrical and Computer Engineering, Sungkyunkwan University
**School of Information Communication Engineering,
Sungkyunkwan University

요 약

본 논문은 고성능망에서 침입 탐지를 위한 침입 탐지 시스템의 효율적인 병렬 구조를 제안한다. 최근 네트워크 인프라의 급속한 성장에 의해 가정까지 광 통신 인프라가 깔리는 고성능망 시대가 되었다. 급격한 인프라의 발달은 스트리밍 서비스와 같은 콘텐츠 서비스의 질을 향상시켰지만, 트래픽이 기존 보안 장치들의 허용 용량을 넘어서게 되어 단일 보안 제품이 트래픽 전체를 감당할 수 없게 되었다. 따라서 고성능망의 침입 탐지를 위해 효율적으로 기존의 침입 탐지 시스템을 연계하기 위한 연구가 진행되고 있다. 본 논문에서는 서비스와 공격 빈도를 기반으로 트래픽을 분류함으로써 효율적인 트래픽 분산 기법을 제시한다.

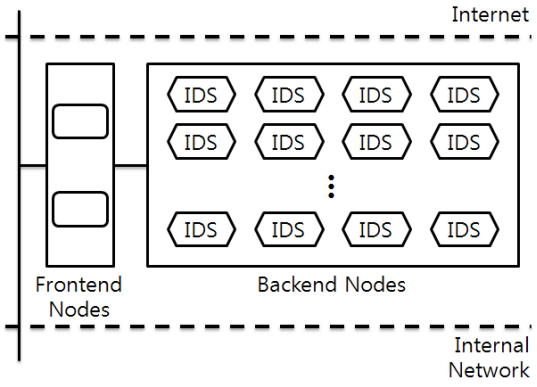
1. 서론

정보화 사회가 지속적으로 발전하면서 개인이 가정에서 광통신을 이용한 높은 대역폭을 사용할 정도로 네트워크 인프라가 크게 발달하였다. 네트워크의 성능 향상으로 인해 스트리밍 서비스의 질이 높아지고, 클라우드 컴퓨팅 서비스와 같은 다양한 서비스들을 이용할 수 있게 되면서 사람들의 생활은 보다 편리해 지고 있다. 하지만 대역폭의 증가에는 치명적인 문제점도 따른다. 네트워크의 대역폭이 커질 경우 시간당 네트워크로 유입되는 트래픽의 양이 증가하게 되는데, 트래픽을 모니터링하고 공격을 탐지하는 보안 장비들의 연산 능력보다 시간당 유입되는 트래픽의 양이 증가하게 되면, 보안 장비들의 탐지 능력이 급격하게 떨어진다. 10G 이전의 네트워크의 경우 인프라의 성장 보다 컴퓨팅 장비의 성능이 더 높아 충분히 보다 나은 보안 장치를 생산할 수 있었다. 하지만 네트워크 인프라가 더욱 발달하여 40G를 넘어서 100G 대역폭의 네트워크에 이르게 되면, 해당 대역폭을 단일 컴퓨팅 장치를 이용해서 처리하는 것이 매우 어려워진다[1, 4]. 이렇듯 인프라의 발달에 따라 매번 네트워크 인프라의 발달에 맞추어 침입 탐지 시스템에 대한 연구를 수행해야 하는 점은 매우 소모적인 부분이다. 따라서 본 논문에서는 네트워크 인프라 증가에 맞춰 확장성을 제공할 수 있는 병렬 구조의 침입 탐지 시스템을 제안한다.

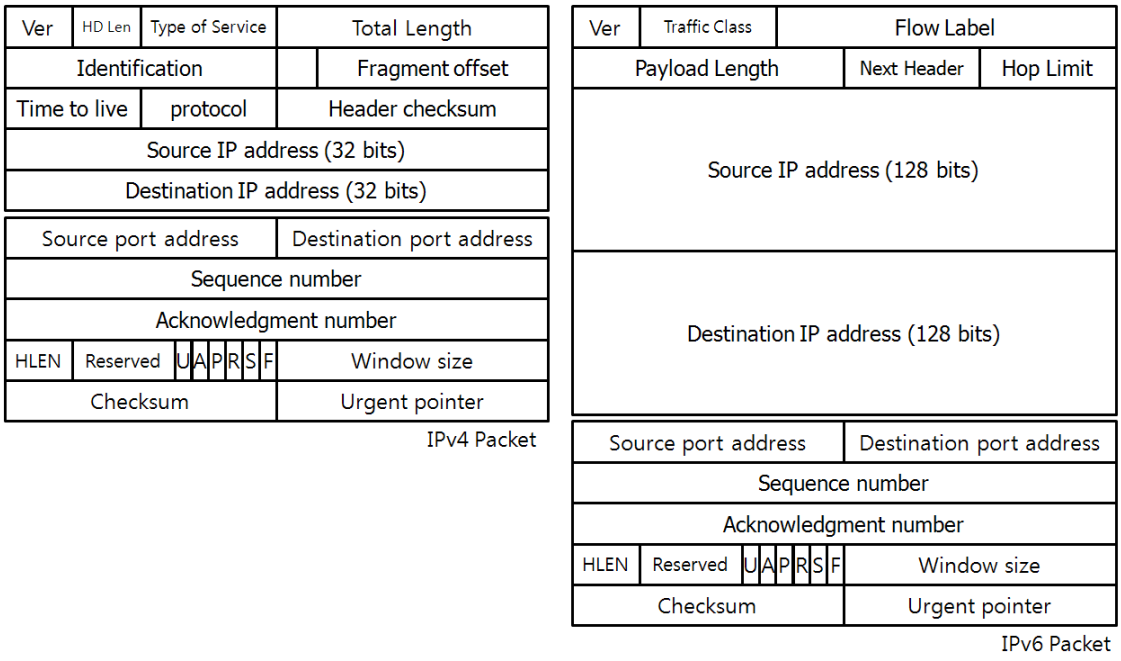
병렬 구조의 침입 탐지 시스템은 동일한 버전의 침입 탐지 시스템이나, 혹은 다른 버전의 침입 탐지 시스템을 이용하여 유입되는 트래픽을 분산하여 전달함으로써 개개의 침입 탐지 시스템들은 적은 량의 트래픽에 대해 탐지 과정을 수행할 수 있는 구조로 (그림 1)과 같다[2].

본 논문에서는 효율적인 병렬 구조의 침입 탐지 시스템을 구축하기 위한 트래픽 분류 방안에 대해 제시한다.

본 논문의 구조는 다음과 같다. 2 장에서 고성능망에서 병렬 구조 침입 탐지 시스템을 구축하기 위한 고려 사항에 대해 서술한다. 3 장에서는 본 논문에서 제시하는 서비스와 공격 빈도에 기반을 둔 트래픽 분류 방법을 제시한다. 마지막으로 4 장에서 결론을 맺는다.



(그림 1) 병렬 구조의 침입 탐지 시스템



(그림 2) IPv4 패킷과 IPv6 패킷 비교

2. 고성능망을 위한 병렬 구조 침입 탐지 시스템의 고려 사항

고성능망을 위한 병렬 구조 침입 탐지 시스템을 설계할 경우 패킷들을 어떠한 침입 탐지 시스템으로 전달할 것인지를 결정하는 프론트엔드 노드의 역할이 중요하다. 프론트엔드 노드는 단시간에 많은 량의 트래픽을 처리할 수 있어야한다. 따라서 프론트엔드 노드는 그간 트래픽들의 패턴을 저장하고 이를 패킷 분류에 적용하기는 매우 어렵다. 따라서 프론테엔드 노드는 적은 규모의 테이블을 이용하여 패킷에 포함된 기본 정보들을 이용하여 패킷을 분류할 수 있어야 한다. (그림 2)는 각각 IPv4 패킷과 IPv6 패킷의 헤더에 담긴 정보와 헤더의 위치를 나타낸 그림이다.

3. 병렬 구조 침입 탐지 시스템을 위한 트래픽 분류 방법

트래픽 분류는 세 가지 방법이 있다. 첫 번째 방법은 균등한 순서로 패킷을 분류 하는 방법이며, 두 번째 방법은 플로우(flow)를 기준으로 패킷을 분류 하는 방법이다. 마지막 세 번째 방법은 패킷의 서비스를 기준으로 패킷을 분류하는 방법이다.

3.1 균등한 분류 방법

균등한 패킷 분류 방법은 패킷을 들어오는 순서에 따라 병렬 구조를 구성하는 침입 탐지 시스템에게 전달하는 방법이다. 균등한 분류 방법의 경우 모든 침입 탐지 시스템에 균등한 처리량을 할당할 수 있어 처리 속도 면에서는 가장 효율적인 방법이다. 하지만 침입 탐지 시스템의 탐지 과정은 때때로 단일 패킷을 대상으로 하기 보다는 패킷들의 집합으로 구성된 하나의 플로우 안에서 공격 패턴을 찾아내야 하는 경우가 있는데, 이러한 경우 균등한 분류 방법을 적용하는 경우에 하나의 플로우를 구성하는 패킷들이 각기 다른 침입 탐지 시스템으로 전달 될 경우 공격 패턴의 탐지가 불가능 하다. 따라서 균등한 분류 방법은 침입 탐지 시스템에 적합하지 못하다.

또한 균등한 분류 방법의 경우 개별 침입 탐지 시스템들을 동일한 패턴으로 동작하도록 하기 때문에 각 노드를 전문화하기 어렵다.

3.2 플로우 기반 분류 방법

플로우 기반 분류 방법은 균등한 분류 방법과 달리 분류 기준을 플로우로 삼는다. 플로우는 주로 목적지 주소, 출발지 주소, 목적지 포트 번호, 출발지 포트 번호, IP 헤더의 프로토콜 필드를 포함한 5 개의 필드를 통해 정의된다. 플로우 기반 분류 방법은 하나의 침입 탐지 시스템이 플로우에 속한 모든 패킷을 검사할 수 있기 때문에 균등한 분류 방법에서 발생하던 플로우 내에 포함된 공격 패턴을 놓치는 문제를 해결할 수 있다. 실제로 많은 병렬 구조 침입 탐지 시스템에서 해당 방법을 사용한다[2, 3]. 플로우 기반 분류 방법의 경우 플로우를 나누는 기준에 따라 2-tuple 분류 방법과 5-tuple 분류 방법이 존재한다.

- 2-tuple 분류 방법: 출발지 IP 주소와 목적지 IP 주소만을 이용하여 플로우를 구분하는 기법이다. 2-tuple 분류 방법은 5-tuple 기법에 비해 프론트엔드 노드의 처리 과정이 간단하여 트래픽을 분산하는 대에 드는 비용이 매우 적다. 하지만 2-tuple 분류 방법에 따라 분류된 플로우내에는 다양한 포트 번호가 뒤섞여 있으며, 하나의 플로우의

크기가 커지기 때문에 각각의 침입 탐지 시스템에게 균일한 트래픽을 전달하기 어렵다.

• 5-tuple 분류 방법: 출발지 IP 주소와 목적지 IP 주소, 출발지 포트 번호, 목적지 포트 번호, IP 헤더의 프로토콜 필드를 이용하여 플로우를 구분하는 기법이다. 2-tuple 분류 방법에 비해 프론트엔드 노드의 처리 과정이 복잡하다. (그림 2)와 같이 IPv4와 IPv6의 패킷에서 포트 번호와 프로토콜 필드의 위치가 각기 다르며, 각 패킷의 옵션에 따라 포트 번호와 프로토콜 필드의 위치가 달라지기 때문에 해당 정보의 빠른 추출이 어렵다. 하지만 5-tuple 분류 방법은 플로우를 세밀하게 분류할 수 있어 각각의 침입 탐지 시스템에게 균일한 트래픽을 전달하기 용이하다. 또한 비록 5-tuple 분류 방법은 포트 번호를 추출하여 분류에 사용하기 때문에 2-tuple 분류 방법에 비해 개별 침입 탐지 시스템의 전문화가 용이하다.

플로우 기반 분류 방법은 플로우 내에 포함된 공격 패턴을 발견할 수 있기 때문에 침입 탐지 시스템에서 이 방법을 많이 적용한다. 하지만 플로우 기반 분류 방법은 여전히 개별 침입 탐지 시스템을 전문화하지 못하기 때문에 분업의 효율을 극대화 할 수 없다. 그리고 트래픽을 분류하여 처리하기 때문에 스캔 공격이나 DDoS 공격과 같이 다양한 플로우에 대해 걸쳐서 발생하는 공격들을 쉽게 판단하지 못하는 한계가 있다.

3.3 서비스 기반 분류 방법

서비스 기반 분류 방법은 분류 기준을 포트 번호로 삼는 분류 방법이다. 서비스 기반 분류 방법은 균등한 분류 방법이나 플로우 기반 분류 방법에서 개별 침입 탐지 시스템의 전문화를 이룩할 수 없었던 부분을 극복할 수 있다.

침입 탐지 시스템은 패킷내에 공격 패턴이 있는지 순차적으로 규칙들을 대조해보는 과정을 거쳐 공격을 탐지한다. 만약 대조과정에서 공격 패턴이 발견될 경우 다른 규칙들을 대조해볼 필요 없이 해당 패킷을 공격 패킷으로 판단할 수 있으며, 이에 대한 경고를 관리자에게 전달한다. 즉, 대조하는 규칙들 중 일치하는 규칙이 먼저 나올수록 침입 탐지 시스템의 탐지가 빨라진다. 이러한 침입 탐지 시스템의 특징을 이용하면 탐지 효율을 높일 수 있다. 현재 알려진 많은 공격들은 각각 포트에 따라 먼저 적용해야 하는 규칙들을 분류 할 수 있다. 이 경우 각각의 개별 침입 탐지 시스템들은 해당 포트 번호에 발생할 수 있는 공격 빈도가 높은 공격 패턴들의 규칙들을 먼저 대조함으로써 탐지 효율을 높일 수 있다. 또한 서비스 기반 분류는 실제 침입 탐지 시스템이 동작할 때 먼저 적용해야 하는 패턴들을 메모리에 미리 할당하고 들어오는 패킷들에 대해 빠르게 대조해 볼 수 있어 침입 탐지 시스템에 들어오는 패킷의 포트 번호를 예측할 수 없는 경우 보다 효율적인 탐지가 가능하다. 또한 다른 분류 방법에 비해 동일한 포트 번호에 대해 이루어지는 DDoS 공격이나 스캔 공격을 탐지하기 용이하다.

하지만 서비스 기반 분류 방법은 균등한 분류 기법이나 플로우 기반 분류 방법에 비해 균일한 트래픽 분산이 어려운 한계가 있다.

4. 결론

본 논문에서는 병렬 구조 침입 탐지 시스템에서 개별 침입 탐지 시스템의 전문화를 통한 효율적인 트래픽 분류 기법에 대해 제시하였다. 침입 탐지 시스템에서 검사해야 하는 패킷의 포트 번호를 미리 예측할 수 있는 경우 침입 탐지 시스템은 보다 효율적인 방법으로 패턴 매칭을 진행할 수 있다. 하지만 포트간의 사용 빈도가 매우 불균형하여 80번 포트와 같이 주로 사용하는 포트에 비해 잘 사용되지 않는 포트의 트래픽 차이가 매우 크기 때문에 포트 번호만을 판단 기준으로 삼을 경우 큰 불균형을 낳을 수 있다.

향후 연구 방향으로는 현재 공개된 규칙들을 통해 포트별 적용 규칙들을 분류해 보고, 실제로 포트 번호를 예측할 수 있는 병렬 구조 침입 탐지 시스템과 그렇지 않은 병렬 구조 침입 탐지 시스템의 성능 차이를 비교 분석 할 계획이다.

참고문헌

[1] Lambert Schaelicke, Kyle Wheeler, and Curt Freeland, "SPANIDS: A Scalable Network Intrusion Detection Loadbalancer", ACM, 2005.

[2] Matthias Vallentin, Robin Sommer, Jason Lee, Craig Leres, Vern Paxson, and Brian Tierney, "The NIDS Cluster: Scalable, Stateful Network Intrusion Detection on Commodity Hardware", RAID 2007, LNCS 4637, pp. 107 - .126, 2007.

[3] R. Lippmann, E. Kirda, and A. Trachtenberg "Gnort: High Performance Network Intrusion Detection Using Graphics Processors" RAID 2008, LNCS 5230, pp. 116 - -134, 2008.

[4] R. Lippmann, E. Kirda, and A. Trachtenberg "Predicting the Resource Consumption of Network Intrusion Detection Systems" RAID 2008, LNCS 5230, pp. 135 - -154, 2008.