情報保護學會論文誌
第 22 卷 第 6 號, 2012. 12

종단간의 유사 연결 패턴을 갖는 정상 서버 활동과 공격의 구분 및 탐지 방법*

장 범 환†‡
호원대학교

A Method for Detection and Classification of Normal Server Activities and Attacks Composed of Similar Connection Patterns*

Beom-Hwan Chang†‡
Howon University

요 약

보안 이벤트 시각화 기법은 기존의 시각화 기술을 네트워크 보안 분야에 적용한 형태로써 네트워크 보안과 관련있는 이벤트를 사용하여 네트워크의 트래픽 흐름과 보안 상황을 쉽고 빠르게 분석 및 탐지하는 기술이다. 특히 종단간의 연결 이벤트인 세션을 시각화하여 네트워크 이상 상황을 탐지하는 기술은 상대적으로 패킷 감시 기법에서 발생하는 오버헤드를 줄일 수 있고 알려지지 않은 공격 패턴들은 쉽게 탐지할 수 있어서 좋은 해결책이 되고 있다. 하지만, 서버들의 정상 활동과 네트워크 공격이 종단간의 유사한 연결 패턴을 가질 경우 세션 기반의 시각화 기법들은 공격 상황과 정상 상황을 구분하는 기능이 매우 취약하다. 따라서 본 논문에서는 세션 기반 시각화 기법에서 서버들의 정상 활동과 네트워크 공격 상황을 상세하게 구분할 수 있는 IP 주소 분할 표시 분석 방법 및 포트 특성 분석 방법을 제안하고자 한다. 제안하는 세션 기반의 공격 시각화 탐지 방법은 다른 공격 탐지 방법들과는 의존성이 없기 때문에 기존의 다양한 네트워크 공격 분석 및 탐지에 활용될 수 있고, 또한 네트워크 관리자에게는 현재 네트워크에서 발생되는 보안 위협을 보다 빠르게 판단할 수 있도록 도움을 준다.

ABSTRACT

Security visualization is a form of the data visualization techniques in the field of network security by using security-related events so that it is quickly and easily to understand network traffic flow and security situation. In particular, the security visualization that detects the abnormal situation of network visualizing connections between two endpoints is a novel approach to detect unknown attack patterns and to reduce monitoring overhead in packets monitoring technique. However, the session-based visualization doesn't notice a difference between normal traffic and attacks that they are composed of similar connection pattern. Therefore, in this paper, we propose an efficient session-based visualization method for analyzing and detecting between normal server activities and attacks by using the IP address splitting and port attributes analysis. The proposed method can actually be used to detect and analyze the network security with the existing security tools because there is no dependence on other security monitoring methods. And also, it is helpful for network administrator to rapidly analyze the security status of managed network.

Keywords: Network Security, Security Visualization, Network Attack Detection

접수일(2012년 1월 12일), 수정일(2012년 10월 8일), 게재확정일(2012년 10월 16일)
* 이 논문은 2012년 호원대학교 연구비 지원을 받은 것임.

† 주저자, bchang@howon.ac.kr
‡ 교신저자, bchang@howon.ac.kr

I. 서 론

네트워크 공격 탐지는 네트워크 상의 각종 서비스들을 감시하여 비정상 서비스 또는 비정상 행위들을 찾아내는 것으로써 서비스를 구성하는 IP 패킷 자체를 감시하는 방법과 종단간의 연결인 세션을 감시하는 방법이 있다[1]. 전자에서는 통신이 출발지와 목적지 간의 패킷 교환을 통해 이루어지므로 패킷을 감시하고 분석하여 어떤 악의적인 내용을 포함하고 있는지 그리고 알려진 공격 패턴들 중에 어떤 것을 포함하고 있는지를 파악할 수 있다[2,3]. 이런 패킷 감시 방법은 IDS(Intrusion Detection System) 또는 IPS (Intrusion Prevention System)와 같이 전통적인 보안시스템들과 네트워크 분석 시스템들에 많이 사용되는 방법이다. 하지만, 네트워크 전체 패킷을 검사해야함으로써 분석시스템에 많은 오버헤드가 발생하고 알려지지 않은 공격 패턴들은 검출하기 어려운 단점이 있다[1,3].

종단간의 연결인 세션을 감시하는 방법은 개별 서비스 또는 패킷에 포함된 악성 내용을 검출하기 보다는 네트워크의 이상 상황이나 네트워크 공격—서비스 거부 공격(DoS: Denial of Service), 분산 서비스 거부 공격(DDoS: Distributed DoS), 스캐닝(Scanning), 웜(Worm)—을 탐지하는데 주로 사용된다. TCP/IP 네트워크를 이용하는 서비스들이 5개의 공통 인자들로 구성되는 종단간의 연결 혹은 세션(session)으로 식별 가능함에 착안하여 현재 이용 중인 서비스들의 행태가 비정상적인 지를 찾아내는 방법이다. 여기서, 5개의 인자는 프로토콜 번호(prt: protocol number), 근원지 IP주소(sip: source IP address), 근원지 포트번호(spt: source port number), 목적지 포트번호(dpt: destination port number), 목적지 IP주소(dip: destination IP address)를 의미하는데, 공격을 포함한 대부분의 네트워크 이상 현상들은 종단간의 세션 모습들이 여러 곳으로 발산하거나 한 곳으로 수렴하는 형태를 보인다[1,3].

일례로, 분산 서비스 거부 공격이 발생하면 다수의 세션들이 특정 목적지 IP주소로 수렴하고, 반면에 인터넷 웜이 발생하면 다수의 세션들이 불특정 다수의 목적지 IP주소들로 발산하는 모습을 보인다. 따라서, 네트워크 이상 상황들은 세션 정보로 대변되는 5-tuple(〈prt, sip, spt, dpt, dip〉)을 이용하여 감시할 수 있고 패킷 분석에서는 알 수 없었던 네트워크의 전체적인 상황 정보도 알 수 있게 된다. 세션 정보를 제공할 수 있는 이벤트로는 트래픽플로우가 있다. 트래픽플로우는 일정 시간 동안 관찰 지점을 통과하는 공통 속성을 갖는 IP 패킷들(트래픽)로 정의된다. 여기서 공통 속성은 전술한 5-tuple 정보를 의미하며 대표적인 이벤트로는 Cisco의 Netflow가 있다. 다수의 IP 패킷들이 모여서 트래픽플로우로 식별되므로 패킷 감시와 분석에서 발생하는 오버헤드는 줄일 수 있으나 패킷 속에 들어있는 악성 코드는 탐지할 수 없는 단점이 있다[1].

세션 기반의 네트워크 상황을 분석한다 하더라도 네트워크에서 발생하는 방대한 이벤트를 처리하는 것은 쉽지 않은 일이다. 특히, 네트워크 속도 및 종단 단말들이 더욱 증가하는 추세에 있고 네트워크 규모에 따라 차이가 있겠지만 세션 이벤트들도 초당 1,000개를 넘는 경우가 대부분이어서 관리 요원이 실시간으로 감시한다는 것도 매우 힘든 실정이다. 이런 문제점과 보안 상황을 신속하게 분석하기 위한 방법으로, 최근 네트워크의 이벤트 시각화 기술에 대한 연구가 활발히 진행되고 있다[1,4,5].

네트워크 이벤트 시각화 기술은 네트워크에서 발생하는 이벤트로부터 특성 정보를 추출한 후, 정보 시각화(Information Visualization) 기법을 사용하여 2차원 또는 3차원 공간상에 이벤트의 내용과 상황을 표현하는 기법이다. 이는 그래픽 요소를 활용하여 데이터가 정보로서 의미가 생성되도록 형상화하는 것으로써 방대한 양의 이벤트들을 직관적으로 분석하는데 매우 유용한 방법이다. 또한 개별 이벤트에서는 볼 수 없었던 패턴들이 형상화되어 관리 요원이 쉽게 패턴을 인지할 수 있고, 한정된 공간에 많은 정보를 표현할 수 있으며 정보를 직관적으로 쉽게 이해할 수도 있다. 보안 이벤트 시각화 기법은 기존의 정보 시각화 기술을 네트워크 보안 분야에 적용한 형태로써 네트워크 보안과 관련있는 이벤트, 즉 트래픽정보 또는 보안경보를 사용하여 네트워크에서 발생되는 트래픽 흐름과 보안 상황을 파악하고 빠르게 분석 탐지하는 기술이다[1,4-7]. 하지만, 기존 보안시스템의 오탐율이 문제인 것과 동일하게 대부분의 세션기반 시각화 기법들은 공격 상황과 정상 서버의 활동을 구분하는 기능이 취약하다. 따라서 본 논문에서는 종단간의 세션 이벤트를 시각화하는 방법에 있어서 정상 서버 활동과 공격 상황을 구분할 수 있는 상세 분석 방법을 제안하고자 한다.

II. 관련 연구

대표적인 시각화 기반 보안상황인지 기술에는 NVisionIP[17], VisFlowConnect-IP[18], PortVis[19], VisCat[1][3] 등이 있다. NVisionIP는 B클래스 네트워크의 서브넷을 가로축으로, 호스트를 세로축으로 설정하여 호스트에서 사용하는 유일한 포트의 수를 화면상에 표현하는 Galaxy View, 네트워크 관리자에서 선택된 서브넷의 호스트의 특정 포트별 플로우 수를 표현하는 Small Multiple View, 한 호스트에서 송·수신되는 포트별 트래픽의 양을 표현하는 Machine View 화면으로 구성된다. NVisionIP는 드릴다운(drill-down) 기능을 통하여 전체 관리 네트워크를 감시하면서 이상 현상이 발견되는 경우 상세한 화면으로 이동하는 기능을 제공한다[17]. 하지만 관리 대상이 되는 B클래스 네트워크에 초점을 맞추고 있기 때문에 트래픽의 근원지 및 근원지 포트 정보를 표현하지 못한다. 또한 한 화면에서 트래픽과 관련된 모든 정보를 표현하지 못하기 때문에 원하는 정보를 얻기 위해서는 다른 화면으로의 전환이 필요하며, 이는 관리자가 이상 현상을 파악하는데 소요되는 시간을 증가시키는 요인이 된다.

VisFlowConnect-IP는 트래픽의 세션 정보에 초점을 맞추고 있으며, 데이터를 송신하는 근원지를 표현하는 축, 관리 도메인의 호스트를 표현하는 축, 데이터를 수신하는 목적지를 표현하는 축으로 구성된 평행 축(Parallel Axis)을 사용하여 사전 정의된 임계치를 초과하는 트래픽에 대해 연결선을 표현하는 방법을 사용한다. VisFlowConnect-IP는 내부 도메인과 외부의 인터넷 도메인간의 연결 정보를 보여주는 External View, 전체 외부 도메인 중 선택된 도메인과 내부 도메인간의 연결 정보를 보여주는 Domain View, 내부 도메인의 호스트 간 연결 정보를 보여주는 Internal View 등으로 구성된다. 트래픽의 흐름을 나타내는 연결선의 경우 트래픽의 양이 많아질수록 어두운 색으로 표현되며, 연결선의 색은 사전에 정의된 도메인을 의미한다[18]. VisFlowConnect-IP는 연결 정보에 초점을 맞추고 있어서 포트별 사용량, 포트의 연속성 등의 포트와 관련된 정보를 시각화하여 보여주지 못하고 한 평행선에 호스트 및 도메인을 표현하기 때문에 특정 호스트 및 도메인을 직관적으로 인지하기 어렵다.

PortVis는 시간 흐름에 따른 포트 정보를 표현하는 기술로써, 특정 시간에 사용된 포트별 트래픽의 양을 표현한다. 또한 화면에서 특정 시간의 특정 포트에 대한 트래픽의 통계 정보(세션의 수, 유일한 근원지의 수, 유일한 목적지의 수, 유일한 근원지와 목적지 쌍의 수, 유일한 근원지 국가의 수)를 선택하여 시간에 따른 흐름을 보여주는 화면을 제공함으로써, 네트워크의 비정상적인 현상을 쉽게 검출 할 수 있다[19]. 하지만 통계 정보만을 사용하여 시각화하기 때문에 트래픽의 흐름을 보여주지 못하고 정확한 근원지 및 목적지의 주소를 인지할 수 없다. 따라서 이상 현상의 상세한 분석을 위해서는 보안 이벤트의 원본 데이터에 접근해야 하는 문제점이 있다.

VisCat은 한국전자통신연구원에서 개발한 이벤트 시각화 도구로써 네트워크 상황 정보를 제공하는 VisNet과 네트워크 공격을 상세하게 분석하는 VisMon으로 구성된다. VisNet의 IPGrid는 전체 IP주소 공간에서 근원지와 목적지 사이에 발생하는 트래픽의 흐름을 시각화하여 각종 네트워크 서비스들과 공격들을 직관적으로 표시한다. 근원지와 목적지 IP주소로부터 추출된 소속 국가, 소속 기관의 정보를 동일 화면상에 표시함으로써 네트워크 공격자와 피해자에 대한 상세 정보를 빠르게 인지할 수 있다. VisNet의 Center는 이상 현상이 발생한 IP주소뿐만 아니라 실제 호스트가 존재하는 전자지도 상의 실제 위치를 표현함으로써 이상 현상이 발생한 호스트의 논리적 위치와 물리적 위치를 빠르게 인지할 수 있다[3]. VisMon은 2차원 쿼드(Quad)와 3차원 큐브(Cube)를 이용하여 네트워크 공격을 상세 분석하는 도구이다. 공격을 포함한 대부분의 네트워크 이상 현상들은 종단간의 세션 패턴들이 발산 또는 수렴 형태를 보이는데, VisMon은 5-tuple의 다양한 변화 모습을 2차원 평면과 3차원 공간상에 효과적으로 시각화하여 해당 공격을 표시 및 탐지한다. 대부분의 시각화 도구들이 단일 공격이 아닌 복수 개의 공격이 진행되거나 세션 개수가 많은 웜 공격 또는 DDoS 공격이 진행될 경우 소수의 공격들은 은닉되어 표시 안되는 경우가 많은데 VisMon은 이와같은 경우에도 공격 표시와 탐지를 수행한다[1].

III. 보안이벤트 시각화 기반의 공격 탐지

VisMon[1]에 따르면, 공격을 포함한 대부분의 네트워크 이상 현상들은 종단간의 세션 모습들이 여러 곳으로 발산하거나 한 곳으로 수렴되는 형태를 보인다. 하지만, 정상적인 네트워크 서비스를 제공하는 서버들의 세션들 역시 네트워크 공격과 유사한 패턴 모습을 갖기 때문에 이를 구분하기가 어려운 실정이다.

일례로, DNS 서버와 Web 서버의 경우 다수의 클라이언트 호스트들이 서버와 접속하기 때문에 전체 세션 모습은 서버로 수렴하는 DDoS 공격과 매우 유사하다. 따라서, 본 논문에서는 VisMon을 비롯한 세션 기반의 시각화 공격 탐지 방법에 있어서 정상 서버의 활동과 공격 상황을 구분하여 탐지하기 위한 IP 주소 분할 표시 방법과 포트 특성 분석 방법을 제안한다. 이를 위해 분석 대상이 되는 세션 그룹을 생성하는 방법과 그룹화된 세션이 공격으로 인해 인위적으로 생성되었는지를 IP주소와 포트를 평가하여 탐지하고자 한다.

3.1 세션 이벤트 군집화

분석 대상이 되는 세션 그룹을 생성하기 위해서 세션 이벤트의 5-tuple을 이용하여 프로토콜별로 1차 군집화하고, 남은 4개의 요소(근원지주소:sip, 근원지포트:spt, 목적지포트:dpt, 목적지주소:dip) 중에서 2개 또는 3개의 요소를 선택 및 조합(combination)하여 이벤트 그룹을 생성한다. 예를 들면, 근원지주소와 근원지포트를 선택하여 조합한다는 것은 전체 세션 이벤트 집합에서 동일한 근원지주소와 동일한 근원지포트를 갖는 세션들을 그룹화한다는 의미이다. 2개의 요소를 선택하여 이벤트 그룹을 생성할 경우, 이벤트 그룹은 6가지 형태로 군집화되고, 3개의 요소를 선택하여 이벤트 그룹을 생성할 경우는 4가지 형태로 군집화된다.

$_4C_2$: ⟨sip,spt,*,*⟩, ⟨sip,*,dpt,*⟩, ⟨sip,*,*,dip⟩,
⟨*,spt,dpt,*⟩, ⟨*,spt,*,dip⟩, ⟨*,*,dpt,dip⟩
$_4C_3$: ⟨sip,spt,dpt,*⟩, ⟨sip,spt,*,dip⟩,
⟨sip,*,dpt,dip⟩, ⟨*,spt,dpt,dip⟩

예를들면 ⟨sip, spt, *, *⟩는 근원지주소와 근원지포트를 이용하여 군집화한 이벤트 그룹으로써, 식(1)의 Agg(sip,spt) 또는 Agg(1100)과 같이 동일하게 표현한다.

$$Agg(sip,spt) = Agg(1100) = <sip, spt, *, *> \quad (1)$$

분석 대상이 되는 세션 그룹은 군집화에 참여하지 않은 나머지 요소들에 대한 수렴과 발산 정도에 따라 선택적으로 선정 가능한데, 수렴과 발산 정도는 식(2)를 이용한 고유분산도(distinct dispersion)를 통해 계산할 수 있다. 고유분산도가 의미하는 것은 특정 목적지 포트 또는 목적지주소로 세션이 수렴 또는 발산하는 지를 나타내는 값이다.

$$Dx = \frac{Distinct(x)}{n(event)}, \quad x = \{sip \mid spt \mid dpt \mid dip\} \quad (2)$$

여기서, n(event)은 그룹 내의 이벤트 전체 개수를 의미하고, Distinct(x)는 x의 고유 개수를 의미한다. 예들 들면 dpt의 집합이 {21, 23, 53, 53, 80, 80, 80, 80, 80, 80}일 경우, Distinct(dpt)는 4가 되므로 D_{dpt} = 0.4 이다. 상기와 같이 Dx의 값은 이벤트가 한 곳으로 수렴할수록 0에 가까워지고 발산할수록 1에 가까워지는 특성이 있다. 그런데, Dx 값만으로 분석 대상 이벤트 그룹을 결정하기에는 부족한 점이 있다. 왜냐하면, 위와 같이 10개의 이벤트로 구성되는 그룹의 Distinct(x)값이 4인 것과 100개의 이벤트로 구성되는 그룹에서 Distinct(x)가 40인 것은 동일한 Dx값인 0.4가 되므로 Dx 값만을 의존해서 이벤트 그룹 특성을 대표하기에는 부적절하다. 따라서, 식(3)과 같이 이벤트 그룹의 무질서도를 반영한 엔트로피(Entropy) 혹은 보정-엔트로피(E^2)를 계산하여 사용한다[1][20]. Dx 값이 동일한 이벤트 그룹이라도 이벤트의 개수가 많아지면 엔트로피는 증가하는 특성이 있는데, 이는 세션이 많이 발산할 경우 엔트로피가 증가하고 많이 수렴할수록 감소하는 특성을 가진다. 보정-엔트로피(E^2)는 기존 엔트로피 값을 0~1로 정규화한 값으로써, 시스템적으로 분석할 경우 또는 해당 값을 그래픽화하여 구분할 경우 유용하게 사용할 수 있다[1].

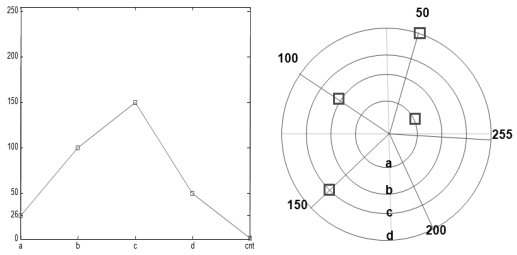
$$H = -\sum_{i=1}^{n} p_i \log_2 p_i, \quad E = H \times \sqrt{\frac{dn}{n}}, \quad E^2 = \sqrt{Ex \times Ey} \quad (3)$$

우선적으로 분석 고려 대상이 되는 이벤트 그룹은 엔트로피값이 크고, 고유분산도 값이 0 또는 1에 근접한 이벤트 그룹이다. 이것은 엔트로피 값이 클수록 그 그룹의 세션이 많이 존재하기 때문에 네트워크에 이상 상황을 야기할 확률이 높고, 고유분산도에 따라 네트워크 공격과 유사한 수렴과 발산 현상을 보여주기 때문이다. 물론, 엔트로피값이 작다고 해서 공격이 아니라는 것은 아니지만, 엔트로피가 큰 이벤트 그룹에 비해 상대적으로 공격 영향이 미비하다는 의미이다.

3.2 IP 주소 분할 표시 분석

분석 대상 이벤트 그룹이 선정된 후에는 그것이 정

상적인 세션 그룹인지 아니면, 비정상 세션 그룹인지에 대한 분석을 수행한다. 분석은 그룹 내 각 세션들의 IP주소가 비정상적인 모습을 보이고 있는 지를 찾아내는 방법인데, 〔그림 1〕과 같이 병렬좌표 분할 표시와 원형좌표 분할 표시를 통해 이루어진다. 이것은 인터넷 규약에 명시된 32bit 양의 정수인 IP주소를 x-y축을 기준으로 분할하여 표시하느냐 원형좌표를 기준으로 분할하여 표시하느냐에 따른 구분이고, 전체적인 의미와 해석은 동일하다. IP주소 체계는 점 형태의 십진수(dotted decimal) 값으로 표기할 경우 a.b.c.d 가 되고, a, b, c, d 각각의 범위는 0~255 값을 갖는데, 병렬좌표 분할 표시에서는 각각의 식별자 a, b, c, d를 X축 상에 각각의 병렬 Y축으로 지정하고, 마지막 병렬 Y축에는 그룹 내의 IP주소로 이루어진 세션 개수(count) 값을 지정한다. 이때 각각의 병렬 Y 축에 각각의 a, b, c, d 값이 나타나도록 점 데이터 형태로 표시 및 연결하여 IP주소를 분할 표시한다. 원형좌표 분할 표시는 각각의 식별자 a, b, c, d를 4개의 원형축 상에 표시하여 그 값들이 연속적인지를 분석하는 방법이다.



(a) 병렬좌표 분할 표시 (b) 원형좌표 분할 표시

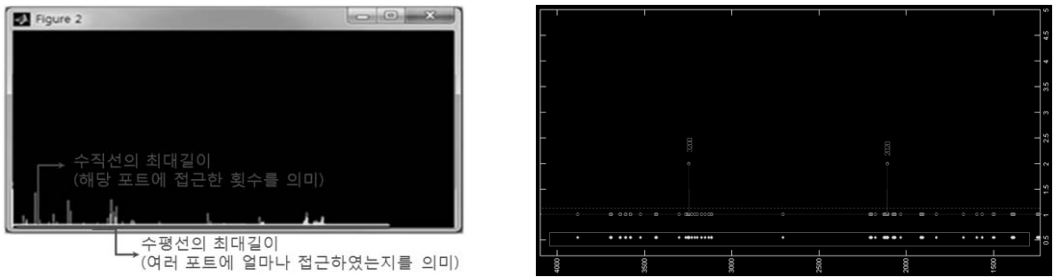
〔그림 1〕 IP 주소 분할 표시 방법

분석 대상 이벤트 그룹의 군집화가 〈sip,spt,*,*〉, 〈sip,*,dpt,*〉, 〈sip,spt,dpt,*〉와 같이 이루어졌다면 sip는 특정 IP주소를 갖게 되므로 IP 주소 분할 표시 분석은 dip에 대해서만 수행하고, 반대로 이벤트 그룹의 군집화가 〈*,spt,*,dip〉, 〈*,*,dpt,dip〉, 〈sip,spt,dpt,*〉와 같이 이루어졌다면 dip는 특정 IP주소를 갖게 되므로 IP 주소 분할 표시 분석은 sip에 대해서만 수행한다. 〈*,spt,dpt,*〉의 경우는 sip와 dip에 대해 모두 IP 주소 분할 표시 분석을 수행해야 한다. 해당 이벤트 그룹의 IP주소를 분할하여 표시한다는 것은 이미 그 그룹이 수렴 또는 발산 형태를 갖기 때문이고, 여기서는 해당 이벤트 그룹이 정상 세션 그룹인지 아니면 비정상 세션 그룹인지를 검출하기 위함이다. 분할된 IP주소 이외의 마지막 병렬 Y축인 IP주소 세션 개수 값(cnt)은 정상 서버일 경우에는 접속횟수가 다양하여 랜덤하지만, 호스트 스캔과 네트워크 스캔 공격일 경우에는 스캐닝을 반복 수행한 특정 횟수 값으로 수렴하는 형태를 보이고, 웜이나 DDoS일 경우에는 1(종단간의 연결세션이 1회)과 같이 매우 작은 값으로 수렴하는 형태를 보인다.

IP주소를 분할 표시하는데 있어서 가독성과 직관성을 높이기 위한 방법으로 IP주소의 a에 값에 따라 순차적으로 일정한 색깔을 하나씩 매핑하여 표시하면, 네트워크 공격 이벤트 그룹은 구분되는 컬러 스펙트럼 패턴들을 갖는다. 이는 자연 발생적으로 접속하는 정상 클라인언트들과는 다르게 공격을 발생시키는 세션들은 일정한 기계적인 메커니즘을 갖고 생성되는 현상 때문이다. 예를 들면, 호스트 및 네트워크 스캐닝 공격은 특정 호스트가 단일 네트워크의 다수 IP주소 또는 단일 호스트의 다수 포트로 연결하는 세션을 갖기 때문에 단일 컬러 스펙트럼을, DDoS 공격은 다수의 호스트가 특정 IP주소로 연결되는 세션을 갖기 때문에 몇 개의 독립적인 컬러 스펙트럼을, 웜 공격과 정상 서버들은 불특정 다수로 연결되는 세션을 갖기 때문에 랜덤 컬러 스펙트럼 모습을 보인다. 하지만, 정상 서버들에 비해 웜 공격의 랜덤 분포가 매우 다양함을 알 수 있다. 또한, 호스트 스캐닝 공격은 dip 값이 병렬축 d에 연속적으로 나타나고, 네트워크 스캐닝 공격은 dip 값이 병렬축 c에 연속적으로 나타난다. 반면에, DDoS는 sip 값이 랜덤하게 표시되는 것 같지만 대게는 병렬축 a에 몇 개의 점 데이터 그룹들로 나타나고 다른 병렬축 b, c, d는 랜덤한 점 데이터들로 가득차며, 정상 서버들을 sip와 dip값이 a, b, c, d 각 축에 랜덤 값들로 나타난다. IP 주소 분할 표시의 마지막 정상 유무 구분 방법은 전체 이벤트 그룹들을 하나의 병렬좌표에 표시한 후, 전체 대비 상위 5~10% 이상의 이벤트 개수를 갖는 그룹만을 필터링하여 나타낼 경우 네트워크 자원을 고갈시키는 인터넷 웜이나 DDoS 공격만이 분할 표시에서 검출되는데 이는 다른 네트워크 공격들이나 정상 서버들에 비해 세션 개수가 상대적으로 많기 때문이다.

3.3 포트 특성 분석

IP 주소 분할 표시 분석과 함께 이벤트 그룹이 정상 세션 그룹인지 아닌지를 검출하기 위해 포트 특성



〔그림 2〕 포트 특성 시각화 분석

〔그림 3〕 끊어진 직선에서의 허프변환 수행 결과

분석을 사용한다. 포트 특성 분석은 이벤트 그룹 내의 spt와 dpt의 구성이 정상적인지를 판별하는 것으로써 포트 번호의 연속성(continuety)과 발생빈도의 균등성(uniformity)을 검사하는 방법이다. 일반적으로 정상적인 세션들의 포트 번호 구성비와 발생 빈도수는 매우 랜덤하게 이루어지지만, 네트워크 공격들은 그렇지가 않다. 예를 들면, 포트 스캐닝 공격은 dpt의 번호가 연속적이고 빈도수는 균등한 모습이지만, 호스트 스캐닝 공격과 DDoS 공격은 특정 dpt 포트 번호로 수렴하고 빈도수는 매우 큰 모습이다. 따라서, 공격을 검출하기 위해서는 특정 포트 번호에 얼마나 많은 세션이 집중되었는지 또는 포트 번호의 구성이 연속적이고 균등한 지를 검사하면 된다.

〔그림 2〕는 이벤트 그룹의 포트 특성을 분석하기 위한 이벤트 시각화 방법으로써 2차원 X-Y 좌표축을 갖는 좌표 데이터 표시 방법이다. 여기서 X축은 포트 번호를, Y축은 발생 빈도수를 의미한다. 먼저 이벤트 그룹 내의 존재하는 포트번호와 빈도수가 (x, y) 값이 되는 좌표를 점 데이터로 표시하고, 다시 발생빈도수와는 무관하게 y값이 0.5를 갖는 X축 선상에 즉, (x, 0.5) 좌표 값을 갖는 점 데이터로 표시한다. 이것은 포트 균등성과 연속성을 빠르게 인지하기 위한 방법으로써 포트 번호의 빈도수가 균등하게 분포하는 지, 그리고 포트 번호들이 연속적으로 연결되어 있는 지를 동시에 쉽게 검사할 수 있다. 포트 균등성 검사에 있어서, 발생빈도가 평균값으로부터 많이 벗어나는(이격되는) 포트 번호들이 표시될 수 있는데 발생빈도의 평균값을 점선으로 표시하여 조금씩 이격되는 점 좌표들은 무시하여 분석하고, 크게 이격되는 점 좌표들은 많은 세션들이 해당 포트로 집중되고 있기 때문에 네트워크 이상 현상으로 분석할 수 있다.

연속성 검사에 있어서, y=0.5 축 상에 나타나는 점 좌표들은 선형 패턴을 갖는데 선형 패턴의 가장 큰 특징은 표시된 직선들의 길이이다. 직선의 길이가 점과 유사할 정도로 짧다면 해당 직선은 정상적인 세션 패턴일 가능성이 높고, 반대로 직선의 길이가 길다면 해당 세션들은 이상 현상 또는 공격이 발생했다는 의미이므로 직선의 길이를 산출하면 공격 발생 여부를 간단하게 판단할 수 있다. 하지만, 세션 정보를 발생시키는 대부분의 장비들이 세션들을 샘플링하여 전송하기 때문에 직선 모양이 일정하지 않고 군데군데 끊어지는 현상이 발생한다. 따라서, 허프변환(Hough Transform)을 사용하여 직선의 길이를 측정한다. 〔그림 3〕은 끊어진 직선에서의 허프변환 수행 결과를 나타낸 것이다.

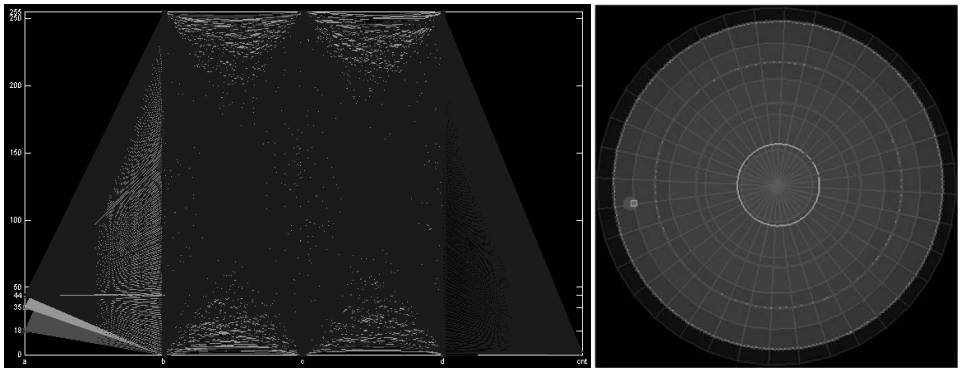
허프변환은 직선방정식을 극좌표방정식으로 변형한

후 직선을 검출하는 방식으로써 직선의 방향은 탐지할 수 있지만, 직선의 시작과 끝을 알 수 없어 정확한 패턴의 길이를 검출할 수가 없다. 따라서 검출된 직선에 대해 각도가 -45°~45°일 때와 45°~135° 일 때를 구분하여 각 열별로 허프공간으로 변형한 뒤, 해당 직선의 기울기와 절편 값에 해당하는 점에서 만나는 직선의 위치가 시작좌표이고, 그 점과 만나는 직선이 끝나는 위치가 직선의 끝 좌표가 되므로 정확하게 패턴을 검출 할 수 있다.

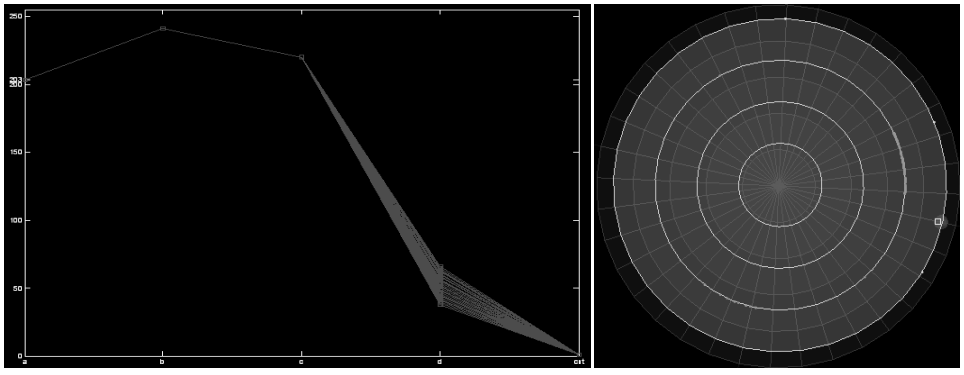
IV. 실험 결과

네트워크 공격 분석 실험을 위해서 실제 운용되고 있는 한국과학기술정보연구원의 라우터(KREONET과 미국의 StarTap 구간)에서 수집한 netflow 이벤트를 사용하였다. 실험 데이터에는 UDP 1434 포트를 이용한 Slammer Worm 공격을 비롯하여 각종 스캐닝 공격들이 포함되어 있으며, 본 논문의 식 (1)과 (2)에 따른 〈sip, spt, *, *〉 군집화 방법을 사용하여 분석 대상 세션들을 선정하였다. IP 주소 분할 표시 분석에 의해 검출된 Slammer Worm 공격과 Host Scan 공격은 〔그림 4〕와 같다〔1〕.

〔그림 5〕는 IP 주소 분할 표시 분석 및 포트 특성 분석을 수행하여 정상적인 서버와 네트워크 공격을 검출한 모습이다. 〔그림 5(a)〕는 정상적으로 운영되고 있는 웹 서버의 모습이지만, 다수의 웹 클라이언트가 접속하여 클라이언트 방향으로는 세션들이 발산하고 서버 방향으로는 세션들이 수렴하는 모습을 보인다. 이것은 세션들의 수렴과 발산 관점에서만 살펴보면 〔그림 5(c)〕와 동일한 형태를 이루기 때문에 기존 공격 분석 시스템들과 시각화 도구들에서는 정상 서버를 공격으로 오탐하는 문제점이 있었다. 하지만, 본 논문에서 제안하는 IP 주소 분할 표시 분석 및 포트 특성 분석을 수행하면, 네트워크 공격과 정상 서버를 구분하여 탐지할 수 있다.

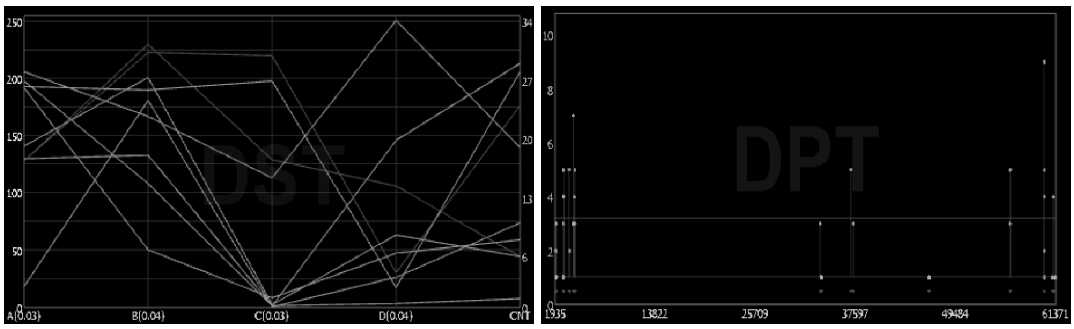


(a) Slammer Worm

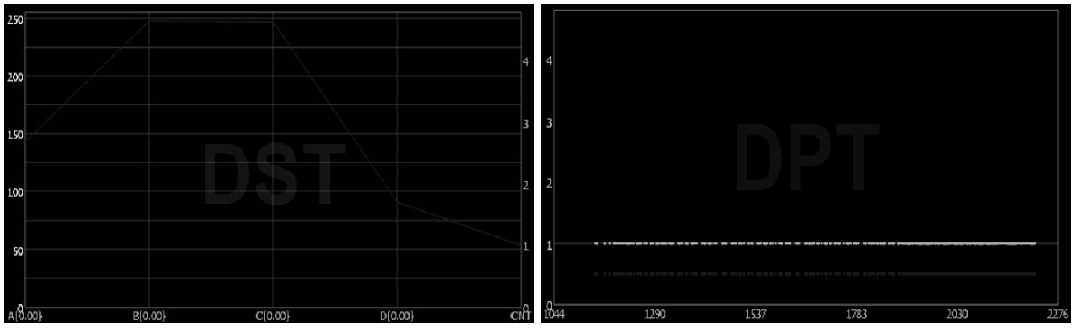


(b) Host Scan

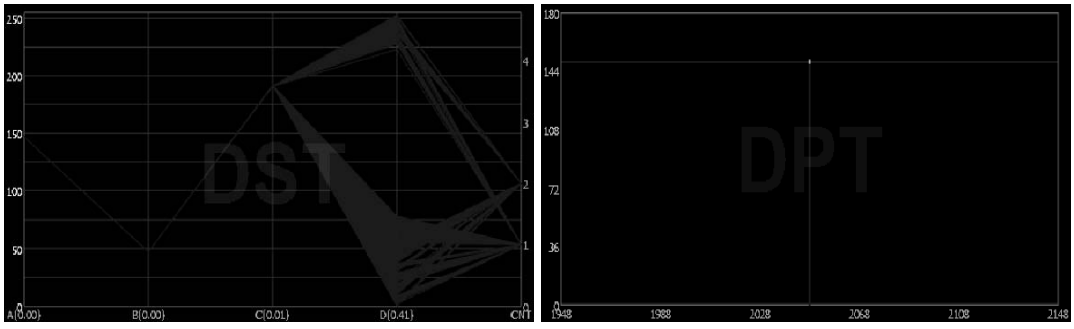
〔그림 4〕 병렬좌표 및 원형좌표를 이용한 공격 분석



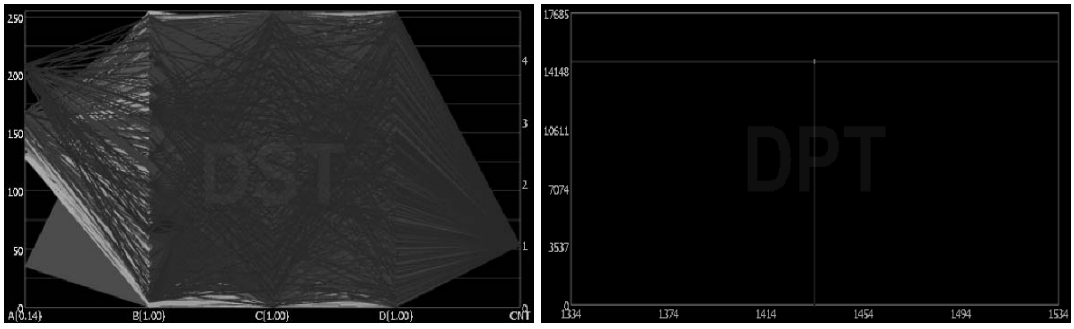
(a) 정상 서버



(b) Port Scan



(c) Host Scan



(d) Slammer Worm

〔그림 5〕 IP 주소 분할 및 포트 특성 분석을 이용한 정상 활동 및 공격의 구분

V. 결 론

네트워크 관리자는 침입탐지시스템, 방화벽 등의 보안 장비에서 발생하는 보안이벤트를 통하여 네트워크에서 이상 현상이 발생하였는지를 인지한 후, 이상 현상이 실제 네트워크의 보안 위협인지를 판단하기 위해서 보안이벤트와 관련된 트래픽을 검색하고 분석하는 등의 일련의 작업을 수행한다. 본 논문에서는 네트워크 보안 상황 분석에 소요되는 시간을 줄일 수 있는 세션 기반의 보안 이벤트 시각화 기법을 소개하였고, 정상 서버와 공격 현상을 정교하게 구분할 수 있는 IP 주소 분할 표시 분석 및 포트 특성 분석 방법에 대해 제안하였다. 이는 세션들을 군집화하여 분석 대상이 되는 세션 그룹들을 찾아내고, 해당 세션 그룹이 자연발생적인 IP주소 및 포트의 특성을 갖고 있는 지를 검사하는 방법이다.

제안하는 방법은 세션기반 시각화 기법의 단점인 종단간의 유사 세션 패턴일 경우 이들을 구분하여 분석할 수 없는 문제점을 개선할 수 있는 것으로써, 실험 결과에서 보는 바와 같이 정상 서버들의 활동과 네트워크 공격 상황(네트워크 스캐닝, 호스트 스캐닝, 인터넷 웜, DoS, DDoS 공격 등)을 정확하게 구분하여 탐지할 수 있다. 또한, 제안하는 방법은 독립적인 모듈 형태로 구현가능하기 때문에 기존의 다른 공격 탐지 방법들과 병행적으로 함께 운영하거나 또는 보안 관리 시스템의 공격 탐지 모듈 형태로 추가하여 운영할 수도 있다.

참고문헌

〔1〕 Beom-Hwan Chang and Chi-Yoon Jeong, "An Efficient Network Attack Visualization using Security Quad and Cube," ETRI Journal, vol. 33 no 5, pp. 770-779, Oct. 2011.

〔2〕 장범환, 나중찬, 장종수, "보안 이벤트 시각화를 이용한 보안 상황 인지 기술," 정보보호학회지, 16(2), pp. 18-25, 2006년 8월.

〔3〕 정치윤, 손선경, 장범환, 나중찬, "시각화 기반의 효율적인 네트워크 보안 상황 분석 방법," 한국정보보호학회 논문지, 19(3), pp. 107-117, 2009년 6월.

〔4〕 A. Giani, I.G.D. Souza, V. Berk, and G. CybenkoI, "Attribution and Aggregation of Network Flows for Security Analysis," Proceedings of the 2006 CERT FloCon Workshop, pp. 1-4, Oct. 2006.

〔5〕 E.W. Bethel, S. Campbell, E. Dart, K. Stockinger, and K. Wu, "Accelerating Network Traffic Analytics Using Query-Driven Visualization," Proceedings of the 2006 IEEE Symposium on Visual Analytics Science and Technology, pp. 115-122, Oct. 2006.

〔6〕 Y. Hu, "Adaptive Flow Aggregation - A New Solution for Robust Flow Monitoring under Security Attacks," Proceedings of the 10th IEEE/IFIP on Network Operations and Management Symposium, pp. 424-435, Apr. 2006.

〔7〕 E.L. Malécot, M. Kohara, Y. Hori, and K. Sakurai, "Interactively Combining 2D and 3D Visualization for Network Traffic Monitoring," Proceedings of the 3rd International Workshop on Visualization for Computer Security, pp. 123-127, Nov. 2006.

〔8〕 A. Oline and D. Reiners, "Exploring Three-Dimensional Visualization for Intrusion Detection," Proceedings of the IEEE Workshop on Visualization for Computer Security, pp. 113-120, Oct. 2005.

〔9〕 H. Koike and K. Ohno, "Snortview: Visualization system of snort logs," Proceedings of the 2004 ACM Workshop on Visualization and Data Mining for Computer Security, pp. 143-147, Oct. 2004.

〔10〕 K. Abdullah, C. Lee, G. Conti, J. Copeland, and J. Stasko, "IDS RainStorm: Visualizing IDS Alarms," Proceedings of the IEEE Workshop on Visualization for Computer Security, pp. 1-7, Oct. 2005.

〔11〕 P Ren, Y. Gao, Z. Li, Y. Chen, and B. Watson, "IDGraphs: Intrusion Detection and Analysis Using Histographs," Proceedings of the IEEE Workshop on

Visualization for Computer Security, pp. 39-46, Oct. 2005.
〔12〕 R. Erbacher, K. Christensen, and A. Sundberg, "Designing Visualization Capabilities for IDS Challenges," Proceedings of the IEEE Workshop on Visualization for Computer Security, pp. 121-128, Oct. 2005.
〔13〕 S. Lau, "The Spinning Cube of Potential Doom," Communications of the ACM, vol. 47, no. 6, pp. 25-26, Jun. 2004.
〔14〕 G. Conti, and K. Abdullah, "Passive Visual Fingerprinting of Network Attack Tools," Proceedings of the 2004 ACM Workshop on Visualization and Data Mining for Computer Security, pp. 45-54, Oct. 2004.
〔15〕 R. Ball, G.A. Fink, and C. North, "Home-Centric Visualization of Network Traffic for Security Administration," Proceedings of the 2004 ACM Workshop on Visualization and Data Mining for Computer Security, pp. 55-64, Oct. 2004.
〔16〕 S. Krasser, G. Conti, J. Grizzard, J. Gribschaw, and H. Owen, "Real-Time and Forensic Network Data Analysis Using Animated and Coordinated Visualization," Proceedings of the 2005 IEEE Workshop on Information Assurance Workshop, pp. 42-49, Jun. 2005.
〔17〕 K. Lakkaraju, W. Yurcik, and A.J. Lee, "NVisionIP: Netflow Visualizations of System State for Security Situational Awareness," Proceedings of the 2004 ACM Workshop on Visualization and Data Mining for Computer Security, pp. 65-72, Oct. 2004.
〔18〕 X. Yin, W. Yurcik, and A. Slagell, "The Design of VisFlowConnect-IP: A Link Analysis System for IP Security Situational Awareness," Proceedings of the 3rd IEEE International Workshop on Information Assurance, pp. 141-153, Mar. 2005.
〔19〕 J. McPherson, K. Ma, P. Krystosk, T. Bartoletti, and M. Christensen, "PortVis: A Tool for Port-Based Detection of Security Events," Proceedings of the 2004 ACM Workshop on Visualization and Data Mining for Computer Security, pp. 73-81, Oct. 2004.
〔20〕 A. Wagner and B. Plattner, "Entropy Based Worm and Anomaly Detection in Fast IP Networks," Proceedings of the 14th IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprise, pp. 172-177, Jun. 2005.

〈著者紹介〉

장 범 환 (Beom-Hwan Chang) 정회원
1997년 2월: 성균관대학교 전자공학과 공학사
1999년 2월: 성균관대학교 전자전자및컴퓨터공학과 공학석사
2003년 2월: 성균관대학교 전자전자및컴퓨터공학과 공학박사
2003년 3월~2012년 2월: 한국전자통신연구원 사이버융합보안연구단 선임연구원
2012년 3월~현재: 호원대학교 사이버수사경찰학부 교수
〈관심분야〉 정보보호, 네트워크 보안, 융합 보안, 보안이벤트 시각화