트래픽 패턴-맵을 이용한 네트워크 보안 상황 인지 기술

(Network Security Situational Awareness using Traffic Pattern-Map)

장 범 환*, 나 중 찬*, 장 종 수*
(Beom-Hwan Chang, Jung-Chan Na, Jong-Su Jang)

요 약 트래픽 패턴-맵(Pattern-Map)은 전체/세부 도메인별 보안 상황을 근원지/목적지 IP 주소 범위로 이루어진 그리드 상에 표현하여 관리자에게 네트워크 보안상황을 실시간으로 인지시키는 도구이다. 각각의 그리드는 근원지-목적지 간의 연결을 의미하며, 최다 점유를 차지하는 트래픽의 포트를 식별력을 갖는 색으로 표현한다. 이상 트래픽 현상의 검출은 가로 및 세로 열에 나타난 동일 색의 막대그래프(포트)의 개수와 그것의 합에 따라 결정되며, 그 결과로 선택된 세로 열과 가로 열을 활성화시켜 관리자에게 그 현상을 인지시킨다. 일반적으로 인터넷 웜이 발생할 경우에는 특정 근원지 열이 활성화되고, DDoS와 같은 현상은 목적지 열이 활성화되는 특징이 있다.

핵심주제어 : 보안상황인지, 트래픽 분석, 보안관리

Abstract This paper introduces a network security situation awareness tool using a traffic pattern map which facilitates recognizing a current network status by extracting and analyzing predetermined traffic features and displaying an abnormal or harmful traffic which deteriorates network performance. The traffic pattern-map consists of 26x26 intersections, on which the occupancy rate of the port having maximum occupancy is displayed as a bar graph. In general, in case of the Internet worm, the source address section on the traffic pattern map is activated. In case of DDoS, the destination address section is activated.

Key Words : Security Situational Awareness, Traffic Analysis, Security Management

1. 서 론

최근 국내외적으로 다양한 보안 이벤트 시각화를 통한 상황인지(Situational Aware- ness) 기술이 보안 관리에 있어서 화두가 되고 있다. 상황인지 기술이란 개개의 객체(이벤트, 사건)의 진위를 판단하고 규명하기 보다는 그것들의 연관성과 전체적인 패턴(동향)들을 통해 어떤 일이 발생하고 있고, 무엇을 해야 하는지를 알고자 함에서 출발한다. 사실 상황인지 기술은 인문·사회과학에 있어서 과거부터 연구되어온 학문 분야로써 보안 관리에 국한되거나 보안 분야의 새로운 기술은 아니다. 상황인지란, "당신 주변에서 진행되고 있는 또는 발생하고 있는 것을 아는 것, 그리고 당신이 알고 있는 범주 내에서 중요한 것이 무엇인가를 아는 것"이라고 정의할 수 있으며, 상황인지 단계는 식별(Perception), 이해(Comprehension), 예측(Projection)으로 구성된다[1][5][7].

* 한국전자통신연구원 정보보호연구단

식별은 다수의 센서(침입탐지시스템, 방화벽 등)들에서 발생하는 이벤트를 수신/저장하는 단계로써 발생한 이벤트의 시간과 종류, 그리고 센서에 대한 지식이 있어야 한다. 각 개별 센서들을 통해 전체 상황을 인지하기에는 아직 이르고, 이해를 준비하는 단계이다. 이해는 목표를 달성하기 위해 이벤트들을 통합 및 혼합, 재배치, 연관성분석 등을 수행하여 전체 상황을 표현한다. 예를 들면, 공격 개시 시간 및 공격자/피해자 시스템, 그리고 관리도메인의 피해 정도를 직관적으로 이해할 수 있는 형태로 표현하는 단계이다. 예측은 현재의 공격 상황 및 피해 정도를 토대로 대응 행동을 결정하도록 돕는 단계이다. 즉, 현재 진행되고 있는 의심스런 일련의 행동들을 토대로 공격자들을 파악하고, 그 공격이 계속 진행될 경우 피해 규모를 추론 및 완화시키거나 막는 방안을 관리자가 결정하도록 돕는다[5][7].

현재 보안이벤트 시각화 기술은 크게 모든 데이터를 일정한 규칙에 따라 좌표평면 이나 기학학적인 도형으로 표현하는 방법과 관심 대상이 되는 데이터만을 동일한 방식으로 표현하는 방법이 있다. 전자의 예로는 The Spinning Cube of Potential Doom[1], RainStorm[2], VISUAL[3], Visual Finger- printing[4], PortVis[9] 등과 같은 대부분의 도구가 여기에 속하며 이는 알려지지 않은 공격이나 패턴을 검출하는데 유리한 반면, 데이터량이 너무 방대하여 속도저하와 패턴이 중첩되는 단점이 있다. 후자는 관심대상을 어떻게 선정하느냐에 따라 전자의 단점을 보완할 수 있는데 ETRI에서 개발한 VisualScope이 여기에 속한다.

VisualScope/PatternMap는 전체/세부 도메인별 보안 상황을 근원지/목적지 IP 주소 범위로 이루어진 그리드 상에 표현하는 도구이다. 각 그리드는 근원지-목적지 간의 연결을 의미하며, 최다 점유를 차지하는 트래픽의 포트를 식별력을 갖는 색으로 표현한다. 이상 트래픽 현상의 검출은 가로 및 세로 열에 나타난 동일 색의 막대그래프(포트)의 개수와 그것의 합에 따라 결정되며 그 결과로 선택된 세로 열과 가로 열을 활성화시켜 관리자에게 그 현상을 인지시킨다. 일반적으로 인터넷 웜이 발생할 경우에는 특정 근원지에서 불특정 다수의 목적지로 동일 트래픽을 전송하기 때문에 특정 근원지 열이 활성화되고, DDoS와 같은 현상은 여러 근원지에서 특정 목적지로 트래픽을 전송하기 때문에 해당 목적지 열이 활성화되는 특징이 있다.

2. 보안 이벤트 및 특성 인자 선정

보안 상황 시각화 기술에 있어서 첫 번째 단계는 보안이벤트 및 특성 인자의 선정이다. 보안이벤트란, 네트워크의 보안상황과 관련있는 모든 이벤트들의 집합으로 정의할 수 있는데, 대표적인 것으로는 트래픽정보(Netflow, tcpdump, SNMP MIB 등)와 보안경보(IDS, Firewall, ESM 등)로 분류할 수 있다[2][6][9]. 현재 개발한 네트워크 보안상황 인지 시스템(VisualScope)에는 여러 종류의 보안이벤트가 입력 데이터로 활용 가능하지만, 본 논문에서는 트래픽 정보 중에서 IETF 및 산업계 표준으로 자리잡고 있는 트래픽 플로우 개념을 이용하여 설명하고자 한다.

2.1 트래픽 플로우

플로우(flow)란, 일정 시간 동안 관찰 지점을 통과하는 IP 패킷들을 공통 속성 갖는 트래픽으로 정의할 수 있으며, 공통 속성으로는 근원지주소(src: source IP), 목적지주소(dst: destination IP), 근원지 포트(spt: source port), 목적지 포트(dpt: destination port), 프로토콜 식별자(prt: protocol number) 등을 취할 수 있다. 현재 트래픽플로우로 활용 가능한 이벤트에는 Cisco의 Netflow가 있으며 tcpdump 데이터 역시 tcptrace나 sanitize를 이용하면 쉽게 플로우 정보로 변환하여 사용할 수 있다. 이제 트래픽 플로우를 구성하는 여러 가지 요소들(플로우의 시작/종료시간, 패킷개수, 볼륨, TCP 플래그 등) 중에서 시각화에 이용할 요소들 즉 특성 인자들을 선정해야 한다.

2.2 군집화(Aggregation)

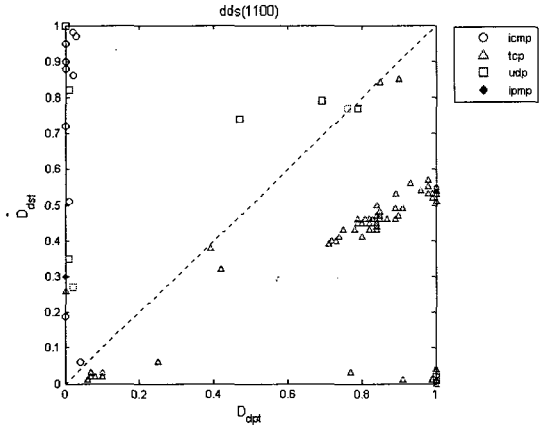
지금까지 개발된 모든 시각화 도구들[3][8][10]의 공통점으로는 src, spt, dpt, dst, prt 등의 군집(집중) 현상을 이용하여 현재의 네트워크 보안상황을 표현한다. 그리고 시각화의 대상이 되는 이

벤트의 규모에 따라 모든 데이터를 특정 규칙에 따라 좌표평면 또는 기학학적인 도형/축에 표현하여 네트워크의 보안상황을 표현하는 방법과 관심대상이 되는 데이터만을 추출하여 표현하는 방법이 있다. 전자는 알려지지 않은 공격이나 패턴을 검출하는데 유리한 반면, 데이터량이 너무 방대하여 속도저하와 패턴이 중첩되는 단점이 있다. 후자는 관심대상을 어떻게 선정하느냐에 따라 전자의 단점을 보완할 수 있다[1][4]. 아래 식은 패턴-맵에서 사용하는 군집화(aggregation) 방법이다. 최초 동일 프로토콜에 따른 1차 군집을 시도한 후, 아래와 식에 의거 군집화한다. 따라서, 군집화에 의한 보안 정보는 총 14개(4C1 + 4C2 + 4C3)로 만들어진다.

$$S_{info} = Agg(ABCD)$$
$$where\ A = src, B = spt, C = dpt, D = dst$$
$$ABCD = \begin{cases} 1, & 군집화에\ 참여 \\ 0, & otherwise \end{cases}$$

[그림 1]은 Agg(1100)에 의한 보안 정보를 표현한 것이다. 즉, 플로우 개수 임계치에 따른 관심대상을 추출하고 src, spt를 기준으로 플로우를 군집화하여 각 포트들의 관계를 나타낸 것으로써 좌측상단에 나타나는 점들은 호스트스캐닝, 우측하단의 점들은 포트스캐닝, 그리고 특이 사항에 따른 포트별 특징을 갖는 몇 개의 그룹들로 좌표평면상에서 트래픽을 클러스터링할 수 있다.

이와같이 전체 이벤트 플로우를 표현하지 않더라도 군집화 방법을 이용하면 적은 이벤트이지만



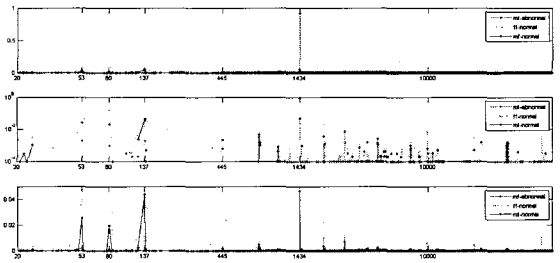
<그림 1> Agg(1100)에 의한 트래픽 군집화

보다 쉽게 보안상황을 표현할 수 있다. 본 논문에서는 src와 dst의 군집화를 사용하여 각 이벤트들의 연관관계를 표현하고, 관심대상 이벤트를 추출하기 위해서는 플로우의 연결지속시간을 사용한다.

2.3 플로우 연결지속시간

일반적으로 인터넷-웜을 포함한 자동화된 공격도구에 의해 생성되는 플로우의 연결지속시간은 매우 짧다. 따라서, 연결지속시간은 정상적인 트래픽패턴과 공격패턴을 구분지을 수 있는 중요한 특성인자가 된다. 우리는 가변적인 임계시간을 적용하여 임계시간 미만인 플로우를 마이크로플로우(mf: micro-flow), 임계시간 이상인 플로우를 매크로플로우(MF: Macro-flow)로 정의하였다.

[그림 2]는 Slammer 웜이 발생했을 때의 마이크로플로우(mf-abnormal, 붉은색)와 정상상태에서의 마이크로플로우(mf-normal, 파란색)의 포트별 점유비율을 비교하여 나타낸 것이다. 여기서 ff-normal은 정상상태일 때의 전체 트래픽 플로우를 의미하는데, 일반적으로 정상상태에서는 특정 몇몇 포트를 제외하고 매크로플로우의 비율이 전체를 차지하게 된다.



<그림 2> 플로우 연결지속시간에 의한 포트별 트래픽 점유비율

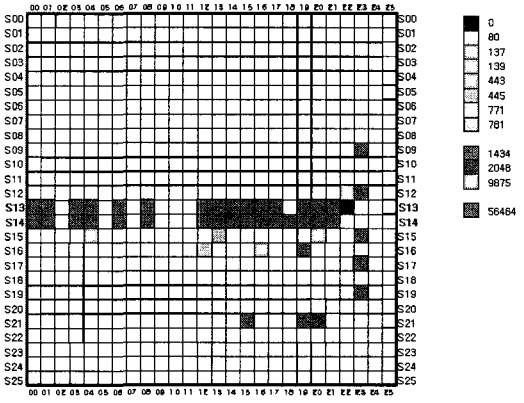
이상과 같이 플로우의 연결지속시간은 트래픽의 이상 현상을 설명할 수 있는 매우 중요한 특성 인자로 정의할 수 있다.

3. 트래픽 패턴-맵 설계

3.1 기본 설계

트래픽 패턴-맵의 기본 개념은 세로축에 근원지 주소를 할당하고 가로축에는 목적지 주소를 할당하여 최다 점유를 보이고 있는 포트의 마이크로 플로우를 그 교차면에 표현하는 것이다. [그림 3]은 트래픽 패턴-맵을 2차원으로 표현한 것으로써, 현재 인터넷(IPv4) 주소 범위는 0.0.0.0-255.255.255 .255로 구성되므로 첫 번째 프리픽스를 10으로 나누고 그 몫을 취하면 세로축과 가로축은 각각 26개의 구간으로 구분된다. 예를 들면, 목적지 주소의 범위 0.0.0.0-9.255.255.255(D00), 10.0.0.0-19. 255.255.255 (D01), ... , 250.0.0.0-255 .255.255.255 (D25)으로 표현되므로 트래픽 패턴-맵은 26x26개의 교차면들로 이루어진 그리드 형태로 표시된다. 여기서, 각 교차면은 근원지와 목적지간의 연결을 의미하는데 이는 근원지와 목적지를 기준으로 군집화한 것과 동일한 의미를 갖는다.

보안상황 인지를 목적으로 트래픽 패턴-맵의 각 교차면에는 근원지-목적지 연결에 해당되는 플로우 개수가 임계치 이상이고, 특정 포트의 점유비율이 임계비율을 넘을 경우 구분되는 색으로 그 교차면에 표시한다. 임계치와 비율은 관리 대상 네트워크에 따라 가변적인 값을 취한다. [그림 3]에서 가로축 S13, S14가 동일 색의 직선 형태를 취하는 것은 불특정 목적지 주소로 마이크로 플로우가 많이 생성되었기 때문인데, 이는 slammer 웜이 발생했을 때의 현상이다. 또한 세로축 D23

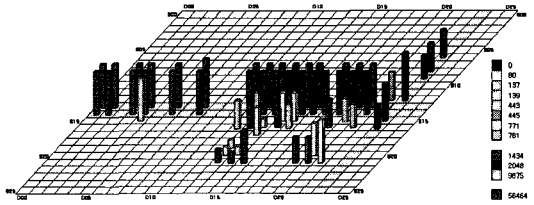


<그림 3> 트래픽 패턴-맵 (2D)

이 특정 색들로 변한 것은 DDoS와 같은 공격이 특정 목적지로 발생했기 때문이다.

3.2 트래픽 패턴-맵 (3D)

기본 설계된 트래픽 패턴-맵의 정확성과 식별력을 높이기 위해서는 이상 상태의 정도와 그 유형을 표시하는 방법이 개선되어야 한다. 앞서 가로축과 세로축이 동일 색의 직선 형태로 나타나는 현상을 설명했는데, 동일한 목적지 또는 근원지 네트워크 주소 구간 즉, 가로 및 세로 교차면들에서 동일 포트의 개수가 소정의 임계치를 넘어서게 되면 해당 근원지 네트워크 주소구간 또는 해당 목적지 네트워크 주소 구간을 해당 포트에 할당된 색으로 표시하여 그 유형을 표시하고, 각 포트별 점유비율을 막대그래프 형태로 각 교차면들에 표시하는 방법을 통해 이상 정도를 표현한다. [그림 4]는 개선된 트래픽 패턴-맵을 나타낸 것이다.

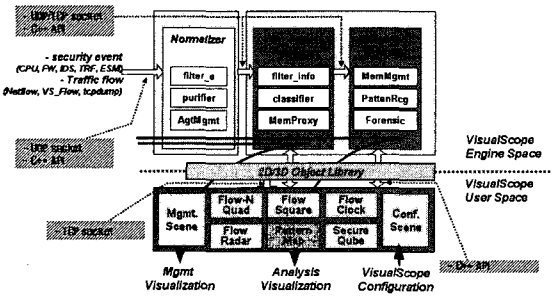


<그림 4> 트래픽 패턴-맵 (3D)

4. 트래픽 패턴-맵의 구조 및 성능

4.1 시스템 구조 및 보안상황 시각화

[그림 5]는 VisualScope의 전체 시스템 구조를 나타낸 것이다. PatternMap은 VisualScope의 GUI들 중에서 트래픽 패턴을 그리드 형태로 표현한 도구이다. VisualScope의 입력 데이터에는 보안 이벤트로써 FW, IDS, ESM 등의 경보이벤트가 있고, 트래픽 이벤트에는 Netflow, VS_Flow, tcpdump/ sanitize 등이 있다. [그림 5]에서 보는 바와 같이 분석기는 Frontend와 Backend로 구성되며, 전자는 실시간 데이터 처리를 필요로 하거나 저수준 데이터를 GUI쪽으로 분석/전달하는 역할을 수행한다. Backend 분석기는 메모리와 자동 패턴 인지를 수행하고 향후 포렌직에 이용될 정보들을 관리하는 역할을 수행한다.



<그림 5> VisualScope의 구조

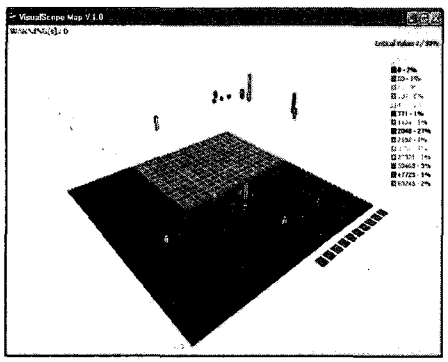
VisualScope의 시스템 사양은 [표 1]과 같다. 분석기는 SUN 계열이나 Linux(x86) 계열에서 동작하며, GUI는 Windows 운영체제에서 동작한다. 하드웨어적으로는 분석기와 GUI 모두 랩탑 환경에서도 무리없이 동작하며 관리도메인 개수나 이벤트 개수에 따라 사양을 높이면 된다.

<표 1> VisualScope 사양

	분석기		GUI
OS	Solaris 10	Linux	Windows XP
CPU	UltraSPARC II, 360 MHz	Pentium IV, 1.86GHz	Pentium IV, 1.86GHz
RAM	1GByte 이상	1GByte이상	1GByte 이상

VisualScope의 GUI 중에서 본 논문에서 설명하고 있는 패턴-맵의 구현 모습은 [그림 6]과 같으며, 2층 형태의 패턴-맵으로 표현된다. 맵의 하부는 전체 네트워크의 트래픽 패턴을 표현하고 있으며, 상부는 관리 도메인별 트래픽 패턴을 표현하고 있다.

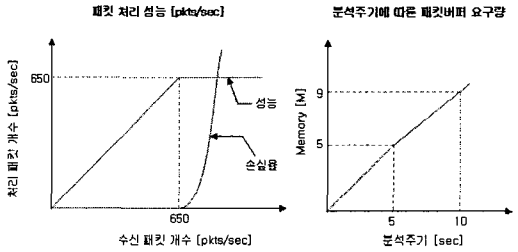
현재 분석은 최소분석주기(5초)마다 이루어지며, 표현은 요구(주기와 기간)에 따라 슬라이딩-윈도우 방식(120개의 윈도우를 가짐)으로 최대 최소분석주기*120 시간까지 표현할 수 있다.



<그림 6> VisualScope/PatternMap

4.2 시스템 성능

VisualScope의 패킷 처리 성능과 버퍼요구량은 [그림 7]과 같다. 일반적으로 이벤트 수신에서 표현까지 Netflow 이벤트는 초당 19,500 플로우를 처리할 수 있고, VS_FLW의 경우는 초당 22,750 플로우를 처리할 수 있다. 분석에 필요한 패킷 버퍼(메모리)는 (최소분석주기* 800K) + 1M 필요로 하기 때문에, 최소분석주기를 5초로 가정할 경우 5M 정도의 메모리가 필요하다.



<그림 7> 패킷 처리 성능과 버퍼 요구량

5. 결 론

본 논문에서는 트래픽 데이터의 패턴-맵을 이용하여 네트워크의 보안 상황을 표현하고 인지하는 기술에 대해 설명하였다. 보안 상황 시각화 기술에 있어서 첫 번째 단계는 보안이벤트 및 특성인자를 선정하고 추출하는 것이며, 두 번째 단계는 추출된 데이터를 관리자에게 직관적으로 인지시키기 위한 시각화 방법이다. 본 논문에서는 보안이벤트로써 트래픽플로우, 특성인자로써는 군집화와 연결지속시간을 정의하였고 시각화 방법은 그리드를 이용한 패턴-맵을 개발하였다. 개발한 트래픽 패턴-맵은 대규모 네트워크의 보안 상황을 실시간으로 처리 및 표현할 수 있으며 메모리 요구량도 매우 작다는 것을 보였다. 따라서, 트래픽 패턴-맵은 최근 문제로 부각되고 있는 대용량 이벤트의 처리, Zero-day 공격의 실시간 검출, 알려지지 않은 공격 패턴의 검출 등에 유용하게 이용될 수 있다.

참 고 문 헌

[1] 장범환, 나중찬, 장종수, "네트워크 보안 상황 인지 기술", 정보통신기술, Vol.19, No.02, pp.15-32, 2005.

[2] K. Abdullah, C. Lee, G. Conti, J. Copeland, and J. Stasko, "IDS RainStorm: Visualizing IDS Alarms", Proc. of VizSEC'05, IEEE, pp.1-7, 2005.

[3] R. Ball, G. Fink, and C. North, "Home-Centric Visualization of Network Traffic for Security Administration", Proc. of VizSEC' 04, ACM Press, pp.55-64, 2004.

[4] G. Conti and K. Abdullah, "Passive Visual Fingerprinting of Network Attack Tools", Proc. of VizSEC'04, ACM Press, pp.45-54, 2004.

[5] A. D'Amico and M. Kocka, "Information Assurance Visualizations for Specific Stages of Situational Awareness and Intended Uses: Lessons Learned", Proc. of VizSEC'05, IEEE, pp.107-112, 2005.

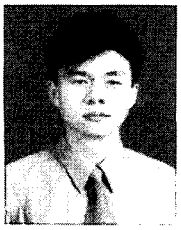
[6] S. Krasser, G. Conti, J. Grizzard, J. Gribschaw, and H. Owen, "Real-Time and Forensic Network Data Analysis Using Animated and Coordinated Visualization ", Proc. of sixth IEEE Systems, Man and Cybernetics Information Assurance Workshop, pp.42-49, 2005.

[7] K. Lakkaraju, W. Yurcik, and A. Lee, "NVisionIP: netflow visualizations of system state for security situational awareness", Proc. of VizSEC 2004, ACM Press, pp.65-72, 2004.

[8] Y. Livnat, J. Agutter, S. Moon, and S. Foresti, "Visual Correlation for Situational Awareness", Proc. of IEEE 2005 Symposium on Information Visualization (InfoVis'05), 2005.

[9] J. McPherson, K. Ma, P. Krystosek, T. Bartoletti, and M. Christensen, "PortVis: A Tool for Port-Based Detection of Security Events", Proc. of VizSEC'04, ACM Press, pp.73-81, 2004.

[10] X. Yin, W. Yurcik, M. Treaster, Y. Li, and K. Lakkaraju, "VisFlowConnect: netflow visualizations of link relationships for security situational awareness", Proc. of VizSEC'04, ACM Press, pp.26-34, 2004.

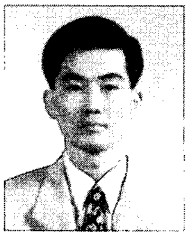


장 범 환 (Beom-Hwan Chang)

- 1997년 성균관대학교 전자공학과 공학사
- 1999년 성균관대학교 전기전자및컴퓨터공학과 공학석사
- 2003년 성균관대학교 전기전자및컴퓨터공학과 공학박사
- 2003년~현재 ETRI 네트워크보안연구그룹 능동보안기술연구팀 선임연구원
- 관심분야: 네트워크 보안, 보안 상황인지, 네트워크 트래픽 분석, 네트워크 공격상황 분석

나 중 찬 (Jung-Chan Na)

- 1986년 충남대학교 계산통계학과 졸업
- 1989년 숭실대학교 전자계산학과 석사
- 2004년 충남대학교 컴퓨터과학과 박사
- 1989년~현재 ETRI 네트워크보안연구그룹 능동보안기술연구팀 팀장
- 관심분야: 네트워크 보안 관리, 보안 상황인지, 네트워크 트래픽 분석



장 종 수 (Jong-Soo Jang)

- 1984년 경북대학교 전자공학과 공학사
- 1986년 경북대학교 전자공학과 공학석사
- 2000년 충북대학교 컴퓨터공학과 공학박사
- 1989년~현재 ETRI 네트워크보안그룹 그룹장
- 관심분야: 네트워크보안, 정책기반보안관리, 비정상트래픽탐지, 유해정보차단