情報保護學會論文誌
第19卷 第3號, 2009. 6

협업 기반의 중앙집중형 봇넷 탐지 및 관제 시스템 설계*

권 종 훈,[1†] 임 채 태,[2] 최 현 상,[1] 지 승 구,[2] 오 주 형,[2] 정 현 철,[2] 이 희 조[1‡]
[1]고려대학교 컴퓨터·전파통신공학과, [2]한국정보보호진흥원

Cooperative Architecture for Centralized Botnet Detection and Management*

Jong Hoon Kwon,[1†] Chaetae Im,[2] Hyunsang Choi,[1] SeungGoo Ji,[2]
JooHyung Oh,[2] HyunCheol Jeong,[2] Heejo Lee[1‡]
[1]Div. of Computer & Communication Engineering, Korea University,
[2]Korea Information Security Agency

요 약

최근의 사이버 공격은 경쟁사에 대한 DDoS공격과 기밀정보 유출, 일반 사용자들의 금융정보 유출, 광고성 스팸메일의 대량 발송 등 불법 행위를 통해 경제적 이득을 취하려는 형태로 바뀌어가고 있다. 그 중심에 있는 봇넷은 봇이라 불리는 감염된 호스트들의 네트워크로서 최근 발생하는 많은 사이버 공격에 이용되고 있다. 이러한 봇넷은 수많은 변종과 다양한 탐지 회피 기술로 무장하고 전 세계 네트워크 전반에 걸쳐 그 세력을 확장해 가고 있다. 하지만 현존하는 봇넷 대응 솔루션은 대부분 시그네쳐 기반의 탐지 방법을 이용하거나, 극히 제한적인 지역의 봇넷만을 탐지하고 있어, 총괄적 봇넷 대응에는 미흡한 것이 현실이다. 본 연구에서는 중앙집중형 봇넷을 주요 목표로 하며, 이를 빠르게 탐지하고 대응하기 위해 ISP 사업자들 간, 혹은 국가 간에 봇넷 정보 공유를 통한 중앙집중형 봇넷 탐지 및 관리 시스템 설계를 제안한다. 본 시스템은 특정 시스템이나 하드웨어에 특화되지 않은 유연함을 갖고 있어 설치 및 배포가 쉽고 ISP 간, 혹은 국가간의 정보 공유를 통해 넓은 지역의 네트워크를 수용할 수 있어, 기존의 솔루션보다 효율적인 봇넷 탐지 및 관리가 가능하다.

ABSTRACT

In recent years, cyber crimes were intended to get financial benefits through malicious attempts such as DDoS attacks, stealing financial information and spamming. Botnets, a network composed of large pool of infected hosts, lead such malicious attacks. The botnets have adopted several evasion techniques and variations. Therefore, it is difficult to detect and eliminate them. Current botnet solutions use a signature based detection mechanism. Furthermore, the solutions cannot cover broad areas enough to detect world-wide botnets. In this study, we suggest an architecture to detect and regulate botnets using cooperative design which includes modules of gathering network traffics and sharing botnet information between ISPs or nations. Proposed architecture is effective to reveal evasive and world-wide botnets, because it does not depend on specific systems or hardwares, and has broadband cooperative framework.

Keywords: Botnet, Detection, Cooperative Architecture, Malware, Network Security

접수일(2008년 12월 26일), 수정일(2009년 4월 22일), 게재확정일(2009년 5월 28일)
* 본 연구는 지식경제부 및 정보통신연구진흥원의 대학IT연구센터 지원사업(TA-2009-(C1090-0902-0016)) 및 IT핵심기술개발사업(2008-S-026-02, 신종 봇넷 능동형 탐지 및 대응 기술)의 일환으로 수행하였으며 본 연구에 참여한 연구자(의 일부)는 "2단계 BK21사업"의 지원비를 받았음.
† 주저자, signalnine@korea.ac.kr
‡ 교신저자, heejo@korea.ac.kr

I. 서 론

봇은 악의적 의도를 가진 소프트웨어에 감염된 PC로서 감염된 다수의 봇이 네트워크로 연결되어 봇넷(Botnet)을 형성하게 된다[1,2]. 이렇게 형성된 봇넷은 그 통제권을 가진 봇 마스터(Bot master)에 의해 원격 조종되며 DDoS 공격, 개인정보 수집, 피싱, 악성코드 배포, 스팸메일 발송 등 다양한 악성행위에 이용되고 있다[3]. 봇넷은 1993년 EggDrop을 시작으로 최근 10년간 Forbot, PBot, Toxbot, Machbot, PHP Bot, Strom Bot[4] 등 점차 진화된 형태를 보이고 있다. 시만텍의 발표에 따르면 2006년 하반기에만 600만대의 새로운 봇이 생성된 것으로 나타났으며 이는 2006년 상반기에 비해 29%나 증가한 수치이다. 또한 2007년에는 하루 평균 52,771대의 PC가 새로이 봇에 감염되고 있다고 전했다[5]. 또한 Arbor Networks의 2007년 조사 결과에 따르면 봇넷이 DDoS 공격을 제치고 사이버 상의 가장 위협적인 요소로 새로이 등극하였다[6]. 더불어, 최근 발생하는 DDoS 공격의 거의 대부분이 봇넷에 의해 이루어지는 것을 감안하면 봇넷이 사이버 상에 현존하는 가장 위협적인 존재임이 분명하다.

국내 상황도 봇넷의 위협에 대해 안전하지 않다. 한국정보보호진흥원(KISA)의 발표자료에 따르면 국내 봇 감염률은 2007년 한해 평균 전 세계 감염률의 11.3%에 달하는 것으로 보고되었다.[7] 이는 초고속 인터넷 인프라가 잘 갖추어져 있는 국내 지역이 봇넷 감염지로 선호되고 있기 때문이다. 또한 이 수치는 제한된 구간의 네트워크(허니넷)에서 탐지된 봇넷을 기반으로 조사된 결과이기에 실제로 탐지되지 않은 봇들이 훨씬 많은 현실을 감안하면 발표된 수치는 빙산의 일각에 지나지 않을 것이다.

봇넷은 주기적 업데이트, 실행압축기술, 코드자가변경, 명령채널의 암호화 등의 첨단기술을 사용하여 탐지 및 회피가 어렵도록 사용 기법이 고도화되고 있다. 또한 봇넷은 그 소스가 공개되어 쉽게 변경이 가능하여 수천 종의 변종이 발생하고 있으며 유저 인터페이스를 통해 쉽게 봇 코드를 생성하거나 제어할 수 있어 전문적인 지식이나 기술이 없는 사람들도 봇넷을 만들고 이용할 수 있다. 현재 이러한 봇넷의 심각성을 주지하고 많은 연구가 이루어지고 있지만 봇넷이 전 세계에 넓게 형성되어 있다는 점에서 그 탐지 및 관제에 어려움이 있는 것이 현실이다[8,9]. 현재 봇넷에 대응하기 위해 다양한 솔루션이 제시되고 있다. 하지만 아직까지 대부분의 솔루션은 시그네쳐 기반의 봇넷 탐지 방식을 채택하고 있다. 이 방법은 기존에 이미 알려진 봇넷을 탐지하기에는 효과적이지만, 하루에도 수천, 수만의 종의 변종이 발생되는 봇넷의 특성상 이를 모두 탐지하기에는 역부족이다. 봇넷을 탐지하기 위한 보다 효과적인 방법은 넓은 지역에 걸친 봇넷의 행위를 분석하고 탐지하는 방법이라 할 수 있다. 현재 Arbor Network사의 Peakflow[10]나 Damballa사의 Failsafe[11]와 같은 네트워크 기반의 행위 분석 탐지 시스템이 현재 서비스 중이지만 시스템 배포에 많은 비용이 필요하며 특정 하드웨어에 의존적이기 때문에 전 세계 네트워크 전반에 퍼져있는 봇넷을 모두 방어하기에는 한계가 있다.

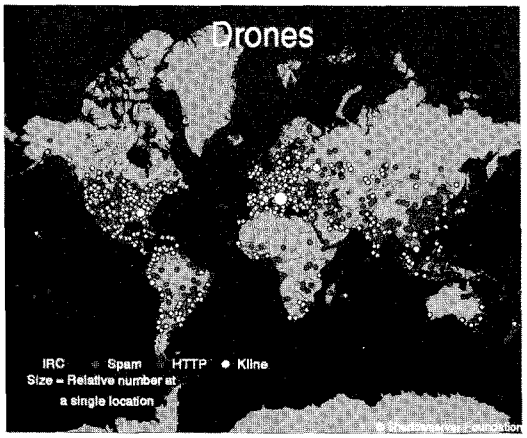
본 연구에서는 이러한 봇넷에 대처하기 위해 ISP 사업자들 간, 혹은 국가 간에 걸친 사회 전반적인 협력을 통한 보다 효과적인 중앙집중형 봇넷의 행위 기반 분석 탐지 및 관리 시스템 구조를 제안하려고 한다. 2장에서는 중앙집중형 봇넷의 특징을 분석하고 이에 대응하기 위한 탐지 및 관제 시스템의 전반적인 구조를 설명한다. 3, 4, 5장에서는 시스템의 세부적인 구조인 트래픽 수집 센서, 탐지 시스템, 관제 및 보안관리에 관한 구조를 설명하고자 한다. 마지막으로 6장에서는 결론과 함께 향후 전망에 대하여 언급하기로 한다.

II. 시스템 구조

2.1 중앙집중형 봇넷의 특성

봇넷의 구성은 크게 봇 마스터와 봇 C&C (Command and control) 서버, 봇 감염 호스트로 구분 지을 수 있다[1]. 봇 마스터에 의해 배포된 악성 봇 코드는 불특정 다수의 PC를 찾아 보안 취약점을 통해 감염시킨다. 이렇게 악성 봇 코드에 감염된 PC를 좀비 호스트 혹은 봇 호스트라고 부른다. 일단 봇에 감염되면 봇 호스트들은 미리 정의된 C&C 서버를 찾아 접속하고, 이 후 봇 마스터의 명령에 따라 각종 악성 행위를 수행하게 된다.

[그림 1]은 Shadowserver에서 2008년에 집계한 전 세계 네트워크에 펼쳐진 봇 호스트들의 분포도이다[12]. 이처럼 봇넷은 특정 지역에 집중되기 보단 무작위로 전 세계 네트워크에 분포하는 특징을 가진다. 또한 최근 봇넷에 의한 공격 행위를 분석한 연구자료들에서도 이러한 봇넷의 분포 형태를 잘 보여주고



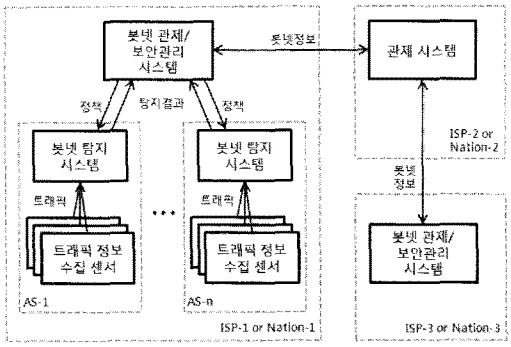
〔그림 1〕 전세계 봇넷 감염 호스트 분포 현황 (www.shadowserver.org, 2008)

있다[13,14]. 때문에 일부 지역의 봇 호스트만을 분석하여서는 봇넷 전체를 파악하기 매우 힘들며, 보다 효율적인 봇넷 파악 및 관리를 위해 본 연구에서는 넓은 범위의 호스트들을 분석이 반드시 필요하다.

이와 같이 봇넷은 소수의 C&C 서버에 대규모의 봇 호스트들이 연결되어 일대 다의 통신을 하며 봇 호스트들에게 명령을 전달하며 집단 행위 특징을 보이게 된다. 이러한 특성을 보이는 봇넷을 중앙집중형 봇넷이라고 하며 본 연구의 주요 목표이다. 중앙집중형 봇넷은 봇 마스터가 C&C서버를 통해 봇 호스트들에게 명령을 전달하기 때문에 봇 마스터와 봇 호스트의 중계지 역할을 하는 C&C서버를 찾아 무력화 시키는 것이 현실적으로 가장 효과적인 방안이다. 하지만 봇 마스터 역시 이 사실을 잘 알기 때문에 최근에는 탐지 및 대응을 피하기 위해 다양한 방법을 동원하고 있다. 다수의 C&C 서버를 이용하여 주기적으로 봇 호스트들을 이주시켜 C&C의 탐지를 어렵게 하거나, 웹 프로토콜인 HTTP(hypertext transfer protocol)을 이용하여 특정 URL을 C&C서버로 사용하기도 한다. 이와 같은 문제점을 해결하기 위해 본 연구에서는 IRC 봇넷 및 HTTP 봇넷의 효과적인 탐지를 위한 네트워크 행위 분석 기반의 봇넷 탐지 및 관제 시스템을 다음과 같이 설계하였다.

2.2 시스템 구성

중앙집중형 봇넷을 탐지 및 관리하기 위해 제안된 시스템은 가능한 넓은 네트워크 수용하도록 설계되었으며 그 구성도는 〔그림 2〕와 같다. 제안된 시스템은 크게 봇넷 관제 및 보안관리 시스템, 봇넷 탐지 시스템, 트래픽 정보 수집 센서로 구분될 수 있으며, 각 시스템 별로 3단계의 계층구조를 가지고 있다.



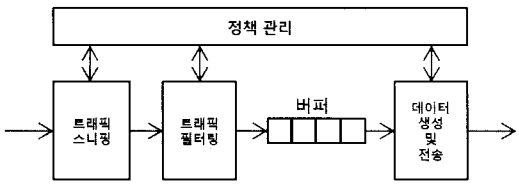
〔그림 2〕 시스템 구성도

봇넷 관제 및 보안관리 시스템은 최상위 계층으로서 하나의 ISP 망 혹은 한 국가를 관리하는 시스템이다. 하나의 봇넷 관제 및 보안관리 시스템은 n개의 봇넷 탐지 시스템을 가질 수 있으며 각 봇넷 탐지 시스템에서 분석된 정보를 종합, 관리하고 차후의 정책 결정을 수행하게 된다. 또한 각각의 봇넷 관제 및 보안관리 시스템은 서로 정보를 공유하여 협력 시스템을 구축한다.

우리는 각각의 봇넷 탐지 시스템이 관리하는 소규모 네트워크 단위를 AS(Autonomous System)로 정의하였다. 각각의 AS는 m개의 트래픽 정보 수집 센서를 가질 수 있다. 봇넷 탐지 시스템은 센서에서 수집된 네트워크 트래픽을 분석하여 신종 봇넷을 탐지하고 공격이나 이주 등의 봇넷 특성 행위를 분석한다. 비교적 작은 AS를 관리하는 봇넷 탐지 시스템의 경우 탐지된 봇넷 정보의 신뢰도가 다소 떨어질 수 있기 때문에 이와 같은 경우 봇넷 관제 및 보안관리 시스템을 통해 타 AS의 봇넷 탐지 시스템 정보와 통합하여 봇넷 탐지의 정확성을 높일 수 있도록 하였다. 봇넷 탐지 시스템에 의해 분석된 결과는 모두 봇넷 관제 및 보안관리 시스템으로 전송된다. 트래픽 정보 수집 센서는 본 시스템의 가장 하위 계층으로, 봇넷의 분석 및 탐지에 필요한 네트워크 트래픽들을 수집하고 이를 봇넷 탐지 시스템에 전송하는 역할을 한다. 하나의 봇넷 관제 및 보안관리 시스템 하에 n*m개의 트래픽 정보 수집 센서가 존재 할 수 있으며, 이는 폭 넓은 범위의 네트워크를 분석함으로써 보다 정확한 봇넷의 탐지를 꾀하고자 함이다.

III. 트래픽 정보 수집 센서

트래픽 정보 수집 센서는 봇넷 탐지 시스템에서 봇넷을 탐지하고 분석하기 위해 필요한 트래픽을 실제 네트워크에서 수집하여 전송하는 역할을 한다. 특히 봇넷의 특성상 특정 네트워크에만 집중되는 것이 아닌 서로 다른 네트워크 상에 무작위로 분포되어 있는 특성을 가지기 때문에 트래픽 정보 수집 센서는 가능한 많은 네트워크에 설치되어 동작하는 것이 유리하다. 트래픽 정보 수집 센서의 기능별 세부 모듈을 살펴보면 [그림 3]과 같이 트래픽 스니핑 모듈, 트래픽 필터링 모듈, 데이터 생성 및 전송 모듈, 그리고 정책 관리 모듈로 나눌 수 있다.



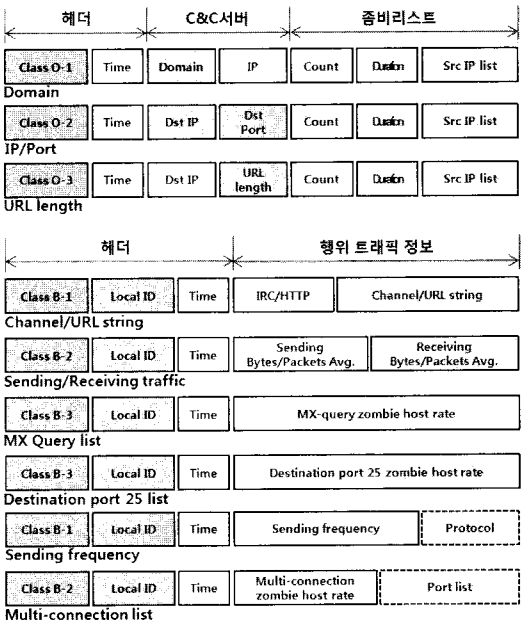
[그림 3] 트래픽 정보 수집 센서 구성도

먼저 트래픽 스니핑 모듈은 패킷 캡쳐 도구를 사용하여 모니터링하는 네트워크의 내부 DNS 서버 앞단이나 게이트웨이에서 미러링하여 트래픽 데이터를 수집한다. 수집되는 트래픽 데이터는 신규 봇넷의 도메인 기반 탐지 및 IP/Port 기반 탐지, URL 기반 탐지, 그리고 탐지된 기존 봇넷의 행위 분석을 위해 TCP/UDP 트래픽 전체를 포함한다. 이렇게 수집된 트래픽에서 분석에 필요한 정보들만을 추출하기 위해 프로토콜, 포트, 패킷 사이즈, 블랙 리스트, 화이트 리스트 등 룰 매칭에 기준하여 트래픽을 필터링한다.

[표 1] 필터링을 통한 트래픽별 분류

트래픽분류	헤더 마킹	트래픽 정보
신규 봇넷 트래픽	Class O-1	Domain 기반 트래픽
	Class O-2	IP/Port 기반 트래픽
	Class O-3	URL length 기반 트래픽
기존 봇넷 행위트래픽	Class B-1	Channel/URL string
	Class B-2	Sending/Receiving Traffic
	Class B-3	MX Query list
	Class B-4	DST Port 25 list
	Class B-5	Sending Frequency
	Class B-6	Multi-connection list

필터링 된 트래픽은 크게 두 가지 종류 트래픽으로 분류된다. 첫 번째는 탐지 시스템에서 신규 봇넷 탐지를 위한 분석에 필요한 트래픽이며, 두 번째는 이미 봇넷의 일부로 탐지된 봇 호스트들의 행위 분석에 필요한 트래픽이다. [표 1]은 필터링을 통해 분류되는 트래픽의 정보이다.



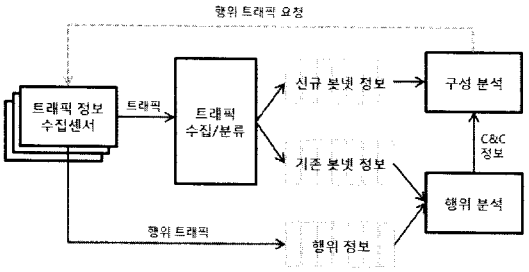
[그림 4] 필터링 룰별 송신 트래픽 송신 규격

필터링된 트래픽들을 봇넷 탐지 시스템에 전송하기 위한 데이터 생성 및 전송 과정을 거치게 된다. 다수의 트래픽 정보 수집 센서에서 수집된 데이터를 봇넷 탐지 시스템에 무작위로 전송하게 되면 봇넷 탐지 시스템에 부하가 걸리기 때문에 데이터 전송에도 룰이 필요하다. 전송 타이머는 이런 다수의 트래픽 정보 수집 센서와 봇넷 탐지 시스템간의 데이터 전송 스케쥴링을 위해 전송 여부를 결정한다. 이러한 과정은 정책 관리 모듈에서 수행하게 된다. 트래픽 정보 수집 센서의 각 모듈에서 사용되는 룰들은 봇넷 관제 및 보안관리 시스템을 통해 관리되며 각 AS의 환경에 따라 서로 다른 룰을 작성하여 적용한다.

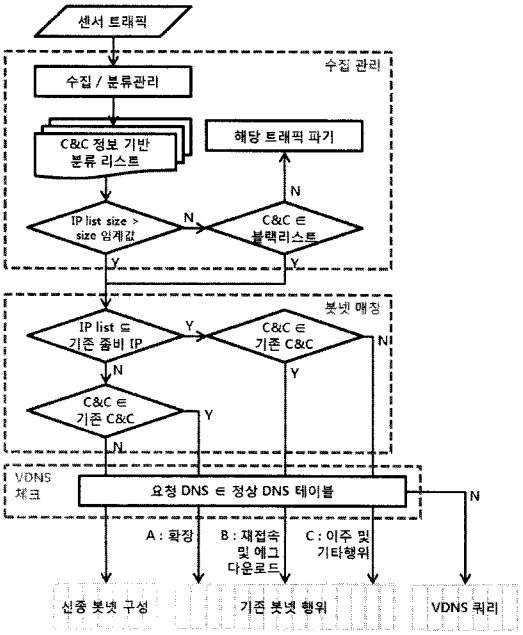
전송 타이머에 의해 데이터 전송이 결정되면 필터링된 트래픽을 버퍼로부터 읽어와 전송 헤더를 생성하고 집약된 데이터 전송 포맷을 생성하여 봇넷 탐지 시스템에 전송하게 된다. [그림 4]는 트래픽 정보 수집 센서로부터 생성된 데이터의 전송 규격이다.

IV. 봇넷 탐지 시스템

트래픽 정보 수집 센서에서 수집된 트래픽 정보들은 봇넷 탐지 시스템에 의해 취합된다. 봇넷 탐지 시스템은 네트워크 전반에 걸쳐 넓게 퍼져있는 트래픽 정보 수집 센서들로부터 트래픽 정보들을 수집하고 봇넷의 구성 및 악성 행위들을 분석하여 신속한 탐지 및 대응이 가능하도록 하는 시스템이다. 봇넷 탐지 시스템의 전반적인 구성도는 [그림 5]와 같다.



[그림 5] 탐지 시스템 구성도



[그림 6] 트래픽 수집 및 분류 모듈 순서도

본 시스템은 크게 중앙집중형 특성을 갖는 트래픽을 수집하고 수집된 트래픽의 봇넷 여부를 판별하는 트래픽 분류 모듈, 상기 트래픽 분류 모듈에 의해 수집되어 봇넷으로 분류된 트래픽의 분석을 통해 C&C 및 봇 호스트의 구성 분석을 수행하는 봇넷 구성 분석 모듈, 상기 트래픽 분류 모듈에 의해 수집되어 봇넷으로 분류된 트래픽의 분석을 통해 해당 봇넷의 확장, 이주, 공격 등의 행위 분석을 수행하는 봇넷 행위 분석 모듈로 구분된다.

4.1 트래픽 수집 및 분류 모듈

트래픽 수집 및 분류 모듈은 각각의 트래픽 정보 수집 센서들로부터 전달되는 트래픽들을 취합하고 기존의 봇넷 탐지정보를 바탕으로 비교를 통해 트래픽을 처리할 모듈별로 분류하는 역할을 하는 수집관리 기능을 수행한다. [그림 6]은 트래픽 수집 및 분류 모듈의 순서도이며, 크게 세단계로 구성된다.

4.1.1 트래픽 수집 분류

첫 번째 단계는 트래픽 정보 수집 센서에 의해 수집된 중앙집중형 트래픽들을 수집하는 단계이다. 수집된 트래픽에 대하여, 일정 대기시간동안 동일 C&C를 기준으로 유입되는 좀비 IP 리스트들을 추가적으로 열거하고, 설정된 대기 시간이 경과한 후 열거된 좀비 IP 리스트의 개수가 임계값을 초과하면 상기 수집된 트래픽을 봇넷 트래픽으로 판단한다. 만일 임계값에 미치지 못한 경우라도, 관제 시스템의 공유정보에 의하여 업데이트되는 C&C서버 블랙리스트와 비교하여 봇넷 트래픽 여부를 판단한다. 상기 단계에서 봇넷 트래픽 여부를 판별할 수 없는 경우, 해당 트래픽은 무효 처리한다.

4.1.2 봇넷 매칭

본 단계에서는 트래픽 수집 단계에서 수집된 트래픽을 기존에 탐지된 봇넷 정보와 매칭하는 과정을 수행하며, 본 단계를 통해 크게 4가지 형태로 트래픽이 분류된다. 그 첫 번째로, 수집된 트래픽들의 중앙집중 서버가 기존 봇넷 정보의 C&C 서버 정보와 상이하고 접속 클라이언트 IP리스트 역시 기존의 봇넷 좀비리스트와 상이한 경우 해당 트래픽을 신종 봇넷 메시지 큐에 저장한다. 둘째로 상기 트래픽의 중앙집중 서버가 기존 봇넷 정보의 C&C 서버 정보와 일치하고 상기 트래픽의 접속 클라이언트 IP 리스트가 기존 봇넷 정보의 좀비리스트와 다른 경우, 해당 트래픽을 봇넷 확장 행위로 구분하여 A 행위 플래그를 부여하고 기

존 봇넷 메시지 큐에 저장한다. 셋째는, 상기 트래픽의 중앙집중 서버가 기존 봇넷 정보의 C&C 서버 정보와 일치하고 상기 트래픽의 접속 클라이언트 IP 리스트가 기존 봇넷 정보의 좀비리스트와 유사한 경우, C&C 서버 재접속 혹은 egg download 행위로 구분하여 B 플래그를 부여하고 기존 봇넷 메시지 큐에 저장한다. 마지막으로 상기 트래픽의 중앙집중 서버가 상기 봇넷 정보의 C&C서버 정보와 다르지만 접속 클라이언트 IP리스트가 기존 봇넷 정보의 좀비리스트와 유사한 경우, 이는 C&C서버 이주 및 기타 주요 봇넷 행위로 구분하여 C 플래그를 부여하고 기존 봇넷 메시지 큐에 저장한다.

신종 봇넷 메시지 큐의 데이터는 구성 분석 모듈에서 읽어 신종 봇넷의 C&C서버와 봇 호스트 리스트를 추출하며, 기존 봇넷 메시지 큐의 데이터는 행위 분석 모듈에서 읽어 각 플래그에 따라 공격 행위나 확산, 이주 행위 등의 분석에 사용된다.

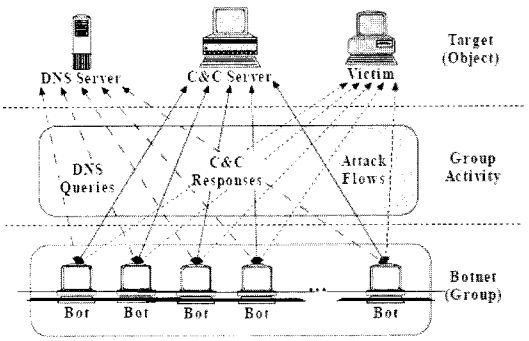
4.1.3 VDNS(Virtual DNS) 체크

VDNS는 봇 마스터가 임의로 운영하는 DNS로서 이를 이용하면 Dynamic DNS를 이용하지 않고도 C&C 서버 도메인의 IP를 쉽게 변경할 수 있다. 이 때 요청되는 도메인은 대부분 정상 DNS 쿼리에는 리턴되지 않는 경우가 많으므로, VDNS 체크 모듈은 정상 DNS 테이블과 매칭하여 VDNS 여부를 체크하고 새로이 탐지된 VDNS 쿼리 정보를 VDNS 메시지 큐에 저장하여 봇넷 구성 분석 모듈 및 봇넷 행위 분석 모듈이 참고할 수 있도록 한다.

4.2 구성 분석 모듈

구성 분석 모듈은 신종 봇넷 메시지 큐에 저장된 중앙집중형 트래픽 정보를 주기적으로 읽어 봇넷의 구성 정보를 분석하고 해당 트래픽이 실제 봇넷에 의해 발생된 트래픽인지를 판단하는 모듈이다. 구성 분석 모듈에서 정상 트래픽과 봇넷 트래픽을 구분하기 위해 중앙집중 서버별, 시간별 접근 소스 IP 그룹들의 유사도를 분석한다[15].

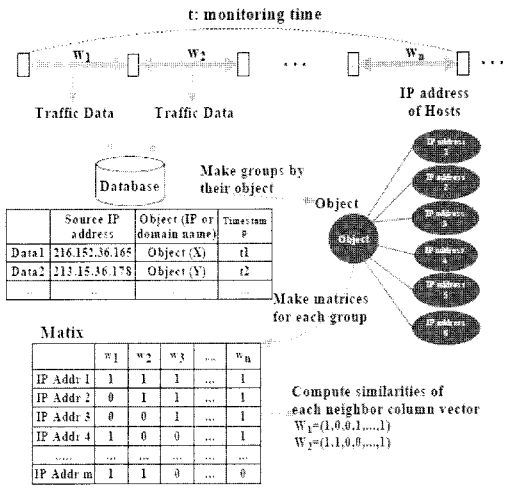
일반적인 인터넷 서비스의 경우에도 중앙집중형 구조의 트래픽이 발생할 수 있다. 하지만 이 경우, 사용자들의 성향에 따라 접속 주기, 접속 유지 시간, 서비스 이용 행위 등이 모두 다르다. 따라서 정상적인 인터넷 서비스의 중앙집중형 트래픽은 시간의 흐름에 따라 접근 소스 IP 그룹이 큰 변화를 보인다.



[그림 7] 중앙집중형 봇넷의 집단행위 특성

반면, 중앙집중형 봇넷은 [그림 7]과 같이 다수의 봇 호스트들이 소수의 C&C 서버에 의해 동작하면서 봇넷만의 고유한 행위 특성을 보이게 된다. 우선 봇넷의 C&C 서버로 사용되는 도메인이나 IP/Port, URL의 경우 일반 사용자는 알 수 없으며 오직 봇 호스트만이 상기 C&C에 접근하는 특징을 보이게 된다. 또한 C&C 서버에 접속하는 과정에서 발생하는 DNS 쿼리나 접속을 유지하기 위한 Ping/Pong 트래픽, 이주, 코드 다운로드, 공격 행위 시, C&C 서버로부터 동시에 명령 받아 수행하는 이른바 집단 행위 특성을 보이며, 시간이 지나도 접속 좀비 IP 들의 변화가 적은 특성을 가진다. 이러한 집단 행위 특성을 "Group Activity"라 정의하며, 이런 특성을 기반으로 봇넷을 탐지한다[15].

[그림 8]에서 보는 바와 같이 시간대 별로 수집된



[그림 8] 봇넷 유사도 분석 방법

데이터는 중앙집중 서버 별 소스 IP들의 그룹으로 이루어져 있으며, 이는 각 시간별 그룹 유사도 분석을 위해 matrix에 저장된다. 각 시간 동안 matrix에 기록된 그룹들은 cosine similarity 등과 같은 유사도 분석 기법을 이용하여 소스 IP들의 유사도 값을 계산하고, 일정 시간동안 기록 후 평균 유사도 값을 계산한다. 계산된 평균 유사도 값은 봇넷 그룹과 정상 그룹의 구분에 이용된다.

또한 각 탐지 시스템이 관리하는 네트워크의 사이즈가 서로 다를 수 있는 점을 감안하여 신뢰도를 계산하였다. 비교적 작은 네트워크를 담당하는 탐지 시스템의 경우, 해당 네트워크 내에 트래픽만으로는 충분히 신뢰할 수 있는 유사도 값을 기대하기 어렵다. 따라서 유사도 값이 높더라도 일정 수 이상의 소스 IP 그룹이 형성되지 않은 경우, 이를 비정성 구성로그로 작성하고 상위 관제 시스템으로 전달하여 타 탐지 시스템의 정보와 다시 분석하는 방법을 이용하였다.

봇넷의 탐지에 이용되는 정상 그룹과 봇넷 그룹과의 차이 및 유사도, 신뢰도 값을 이용한 판단에 필요한 임계값 등은 실험을 통해 도출된 값을 이용한다. [15] 구성 분석 모듈에 의해 탐지되는 봇넷의 구성은 크게 C&C 서버와 봇 호스트들로 구분되며, 분석된 결과는 관제 및 보안관리 모듈에서 차후 정책 결정 및 탐지된 봇넷의 관리에 사용될 수 있도록 한다.

잘 알려진 탐지 알고리즘인 BotHunter[16]나 BotMiner[17]는 IDS의 로그를 이용한 탐지 알고리즘이다. 이는 snort rule을 기반으로 이상행위로 탐지된 로그를 기반으로 탐지를 시작하기 때문에 로그에 종속적이고 이를 분석하는데에도 많은 시간이 소요된다. 또한 현재까지의 탐지 알고리즘은 특정 프로토콜에만 한정되어 있다[16,18]. 반면, 본 구성 분석 모듈에 사용된 유사도 분석 탐지 기법은 네트워크 트래픽 추이를 분석하여 봇넷을 탐지하기 때문에 빠른 탐지 시간을 보이며, 특정 프로토콜에 한정되지 않고 중앙집중형 봇넷을 탐지할 수 있는 장점을 가졌다. 알고리즘의 탐지 효율은 [15]을 참조하도록 하였다.

4.3 행위 분석 모듈

행위 분석 모듈은 기존 봇넷 메시지 큐의 트래픽을 읽어 봇넷의 행위 특성을 분석하며, 분석 결과를 행위 로그로 기록한다. 행위 로그에 기록된 기존 봇넷의 행위 분석 결과는 통합 관제 시스템으로 전송되며, 해당 봇넷의 추이 및 행위 예측, 대응 방법 결정 등에 사용된다. 분석에 사용되는 트래픽은 상기 봇넷 매칭을 통해 크게 3가지 행위 기반으로 분류되어 플래그가 부여된 형태로 행위 분석 모듈로 전달되며 각각의 의미는 [표 2]와 같다.

[표 2] 봇넷 매칭

플래그	C&C	좀비 리스트	특징
N/A	≠	≠	신종 봇넷 구성
A	=	≠	봇넷 확장 행위
B	=	=	서버 재접속 및 Egg download
C	≠	=	서버 이주, 개인정보 탈취, Egg download, 스팸메일 발송, DDoS

A 플래그가 부여된 트래픽은 봇넷의 확장 행위 분석에 사용된다. 중앙집중 서버가 기존 C&C 서버 정보와 일치하지만 접속 IP리스트가 기존 좀비리스트와 차이를 보이는 경우는 새로이 봇에 감염된 좀비 호스트들이 기존의 C&C 서버에 접속하기 위해 발생시킨 트래픽일 가능성이 높다.

B 플래그는 봇 호스트들의 재접속 혹은 Egg download 행위 분석에 사용된다. 중앙집중 서버와 접속 IP 리스트가 기존의 봇넷 정보와 모두 일치하므로 이는 네트워크 이상이나 시스템 재부팅에 의해 봇 호스트들이 C&C 서버와의 연결을 유지하기 위한 재접속 행위 혹은 접속 유지를 위한 Ping/Pong 행위로 간주할 수 있다. 또한 세부 공격 수행 모듈 다운이나 스팸메일 발송을 위한 메일주소 및 메일 컨텐츠 템플릿 다운, 자가 업데이트 등의 Egg download 역시 의심될 수 있다. 따라서 이 경우 행위 분석 모듈은 일정시간 동안 발생하는 패킷 당 평균 수신 트래픽 량을 추가로 비교 검사하여 행위 분석을 수행한다.

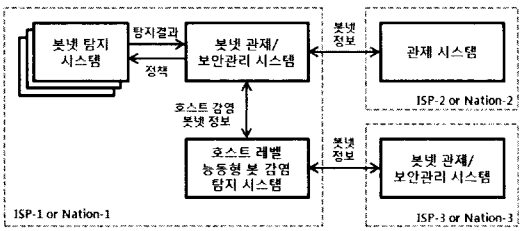
마지막으로, C 플래그가 부여된 트래픽은 이주 및 주요 공격 행위 분석에 사용된다. 중앙집중 서버가 기존 C&C 서버 정보와 다르지만 접속 IP 리스트가 기존 좀비리스트와 일치하는 경우, 신규 C&C 서버로 좀비 호스트들이 이주하는 행위로 간주 할 수 있다. 또한 봇 마스터의 명령에 의한 봇 호스트들의 집단 공격 행위로서, 상기 중앙집중 도메인, IP/Port 혹은 URL에 대한 DDoS, 스팸메일 발송, 봇넷 감염 전파를 위한 스캐닝 공격 등의 다양한 공격 행위로도 분석될 수 있다. 따라서 DDoS 공격의 경우, 동일 전송 트래픽의 발생 주기를 추가 분석하며, 스팸메일 발송으

로 의심되는 경우, SMTP 서버로 발송되는 메일 전송 트래픽을 관찰하거나 MX 쿼리 트래픽을 관찰하여 봇넷의 공격 행위를 분석한다.

V. 관제 및 보안관리

5.1 네트워크 구성도

봇넷 관제 및 보안관리 시스템의 네트워크 구성은 [그림 9]와 같다. 봇넷 관제 및 보안관리 시스템은 일반적으로 ISP 사업자 망 혹은 국가 망에 1개 존재하며 하부 AS 내의 봇넷 탐지 시스템들로부터 관할 내의 봇넷 정보를 수집한다. 또한 각 시스템은 서로 정보공유를 통해 보안관리를 수행하게 된다.

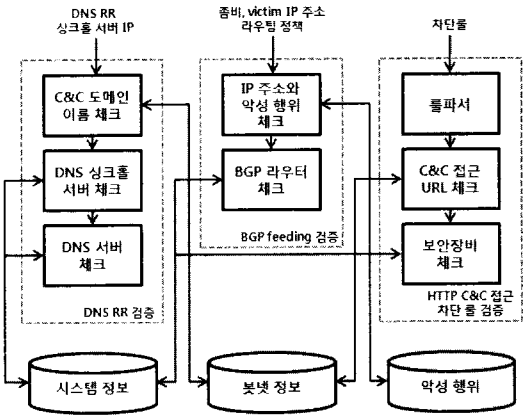


[그림 9] 봇넷 탐지 및 관제 시스템 구성도

봇넷 관제 및 보안관리 시스템의 주요 기능들은 하부 탐지 시스템들의 정보 취합, 비정상 로그 분석, 타 시스템과의 정보 공유, 대응 정책 수립, 통계 데이터 생성, 하부 탐지 시스템의 정책 적용 들을 들 수 있다. 하부 탐지 시스템들로부터 전송된 탐지, 행위 분류 로그들을 취합하여 DB에 저장하고 DNS 싱크홀 등과 같은 봇넷 대응 정책을 수립한다. 또한 비정상 로그(그룹 유사도가 임계치를 넘지만, 신뢰 임계치를 넘지 못하는 그룹 정보)들을 취합하여 봇넷 여부를 판단하고 C&C와 봇 호스트 구성 정보를 추출한다. 이렇게 추출된 정보는 다시 각 탐지 시스템으로 전송되어 이후 행위 분석 및 대응 방법등에 사용된다.

관제 및 보안관리 시스템은 하위 탐지 시스템에서 분석된 봇넷 정보를 타 시스템과 유기적으로 공유한다. 이렇고 공유된 정보는 통계 데이터 및 시각화를 통해 관리자들에게 제공되므로서, 빠르고 정확한 대응 정책을 설정할 수 있도록 돕는다. 결과적으로, 봇 시스템이 설치되어 관리하는 전 영역의 봇넷 정보를 실시간으로 분석하고 이에 대응할 수 있는 시스템을 형성하게 되며, 잘 클러스터링된 봇넷 (전 세계 네트워크에 넓게 분포하고 있는 봇넷)의 경우에도 탐지가 가능하게 되며 빠르게 대응할 수 있게 된다. 호스트레벨 능동형 봇 감염 탐지 시스템은 독립적으로 설치된 시스템으로, 능동적으로 감염된 악성 봇을 분석하여 봇넷이 사용하는 봇 정보를 관제 및 보안관리 시스템으로 전달한다.

5.2 봇넷 대응 방법



[그림 10] 봇넷 대응 기법 및 정책 검증

탐지된 봇넷에 대응하기 위해 [그림 10]과 같이 다양한 대응 기술이 사용되며, 그 첫 번째 방법은 블랙리스트를 공유하는 방법이다. 특정 AS내에서 짧은 시간에 다수의 좀비가 새로운 C&C에 접근하는 것이 발견될 경우 C&C에 대한 정보를 다른 AS의 탐지 시스템과 공유해서 블랙리스트를 통한 대응을 수행하는 것이 가능하다. 블랙리스트가 있는 경우에는 C&C 서버를 비롯한 감염 호스트들에 직접적인 통제나 제한을 가하는 것이 가능하다.

두 번째 기법은 이미 널리 알려진 DNS 싱크홀 기법을 활용하는 것이다[19]. DNS 싱크홀 기법은 주로 IRC기반 봇넷 C&C 접근 차단을 위해 사용되는 대응 정책으로, 신규로 발견된 IRC 봇넷에 대한 접근 차단을 위해 DNS RR을 생성하여, DNS서버로 전송하는 기법이다. 싱크홀 리스트에 추가된 도메인들에 대해서 싱크홀 기법을 적용한 도메인 서버로 질의를 하는 호스트들은 기존의 IP정보가 아닌 싱크홀 처리된 IP 주소를 DNS 서버로부터 전달받게 되고 이 때문에 봇넷의 C&C서버가 아닌 봇넷의 대응을 하는 쪽의 서버로 접속하게 되어 봇넷의 동작을 차단하게 된다.

세 번째로 HTTP 봇넷 C&C URL 접근 차단을

통한 HTTP 봇넷의 대응이 가능하다. 이는 주로 HTTP기반 봇넷 C&C 접근 차단을 위해 사용될 수 있는 대응 정책으로, 공개 웹 방화벽의 룰 설정을 통해 좀비가 HTTP 봇넷 C&C URL에 접근 하는 것을 차단할 수 있다.

네 번째로 BGP feeding 기법을 이용하여 봇넷을 무력화할 수 있다[20]. BGP feeding 기법은 주로 DDoS등 봇넷을 이용한 공격 행위 차단을 위해 사용되는 대응 정책으로, 공격을 받는 호스트로 가는 DDoS트래픽 등을 null routing의 기법을 통해 차단 할 수 있다. 이를 활용하여 탐지된 봇 감염 호스트들이 중앙 서버와 통신하는 트래픽을 차단하는 것이 가능하며 봇넷의 정상적인 동작을 차단하는 것이 가능하다. 하지만 BGP feeding 기법의 경우 그 적용이 쉽지 않고 룰 적용 시 주의하지 않으면 네트워크의 정상적인 동작에 장애를 일으킬 수 있는 단점이 존재하므로 기법의 적용 시 각별한 주의가 필요하다.

VI. 결 론

본 연구에서는 최근 보안 이슈중 가장 위협적인 봇넷에 효과적으로 대응하기 위한 탐지 및 관제 시스템의 구조를 제시하였다. 본 시스템 구조는 첫 번째로, ISP간 혹은 국가간의 협력을 기반으로 소규모 봇넷뿐만 아니라 전 세계에 넓게 분포된 봇넷까지도 탐지하고 대응 할 수 있는 장점을 가졌다. 둘 째, 봇넷의 행위 특성을 네트워크 트래픽 분석을 통해 세분화 하여 정의하였다. 이는 차후 연구에서 보다 효과적인 봇넷의 관리 및 대응이 가능하게 할 것이다. 마지막으로, 본 연구에서 제안한 설계 구조는 다양한 탐지 알고리즘을 수용할 수 있어 차후에 발전되는 봇넷 탐지 연구들을 적용하기 쉽다는 장점을 가지고 있으며, 특정 하드웨어나 시스템에 특화되어 있지 않는 유연한 시스템 구조를 갖고 있어 실질적인 설치 및 배포가 용이하다.

본 시스템 구조는 ISP 간 혹은 국가 간의 협력을 기반으로 설계되었다. 따라서 이들을 감독하고 조율할 수 있는 상위 감독 기관이 필요할 것으로 본다. 이를 위해 현재 본 시스템 모델을 기반으로 국제 표준을 제안된 상태이다. 이는 전 세계적으로 이슈가 되고 있지만 실질적인 대책이 될 수 있는 봇넷 탐지 및 대응 시스템의 모델이 없는 현실을 감안 할 때 선도적 위치를 점유할 수 있는 사례가 될 것이다.

최근의 봇넷은 IRC나 HTTP를 이용한 중앙집중형 봇넷 뿐만 아니라 탐지 및 대응을 어렵게 하기 위해 분산형 Peer-to-Peer 방식을 사용하는 신종 봇넷이 증가하는 추세를 보이고 있다. 따라서 중앙집중형 특성을 보이지 않는 봇넷을 탐지하기 위한 연구와 함께 이를 포함 할 수 있는 시스템 구조 또한 차후 풀어야 할 과제가 될 것으로 예측한다.

참 고 문 헌

[1] E. Cooke, F. Jahanian, and D. McPherson, "The Zombie Roundup: Understanding, Detecting, and Disrupting Botnets," In Proceedings of Usenix Workshop on Steps to Reducing Unwanted Traffic on the Internet (SRUTI'05), pp. 39-44, July 2005.

[2] The Honeynet Project, "Know your enemy: Tracking botnets," http://www.honeynet.org/papers/bots, 2005.

[3] N. Ianelli and A. Hackworth, "Botnets as a vehicle for online crime," CERT, Dec. 2005.

[4] J.B. Grizzard, V. Sharma, C. Nunnery, B.B. Kang, and D. Dagon, "Peer-to-peer botnets: Overview and case study," In Usenix Workshop on Hot Topics in Understanding Botnets (HotBots'07), Apr. 2007.

[5] D. Turner, M. Fossl, E. Johnson, T. Mack, J. Blackbird, S. Entwisle, M.K. Low, D. McKinney, and C. Wueest, "Symantec Global Internet Security Threat Report Vol. XIII," Symantec, Apr. 2008.

[6] D. McPherson, C. Labovitz, and M. Hollyman, "Worldwide Infrastructure Security Report Vol III," Arbor Networks, Sep. 2007.

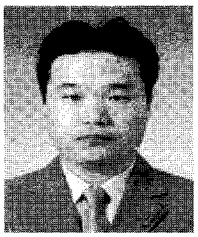
[7] 한국정보보호진흥원, "인터넷침해사고 동향 및 분석 월보," pp. 10-11, 2007년 12월.

[8] M. Ahamad, D. Amster, M. Barrett, T. Cross, G. Heron, D. Jackson, J. King, W. Lee, R. Naraine, G. Ollmann, J. Ramsey, H.A. Schmidt, and P. Traynor, "Emerging Cyber Threats Report," Georgia Tech.

Information Security Center, pp. 2-3, Oct. 2009.
[9] 전용희, "봇넷 기술 개요 및 분석," 정보보호학회지, 18(3), pp. 101-108, 2008년 6월.
[10] Arbor Networks, Peekflow-SP, http://www.arbornetworks.com/en/peakflow-sp.html
[11] Damballa, Failsafe app, http://www.damballa.com/solutions/enterprise_solutions.php
[12] Shadow server, http://www.shadowserver.org
[13] J. Jung, B. Krishnamurthy, and M. Rabinovich, "Flash crowds and denial of service attacks: Characterization and implications for CDNs and web sites," 11th Int'l WorldWideWeb Conference, pp. 252-262, May 2002.
[14] L. Zhuang, J. Dunagan, D.R. Simon, H.J. Wang, and J.D. Tygar, "Characterizing Botnets from Email Spam Records," First Usenix Workshop on Large-Scale Exploits and Emergent Threats(LEET '08), pp. 1-9, Apr. 2008.
[15] H. Choi, H. Lee, H. Lee, and H. Kim, "Botnet Detection by Monitoring Group Activities in DNS Traffic," IEEE Int'l Conf. Computer and Information Technology (CIT), pp. 715-720, Oct. 2007.
[16] G. Gu, P. Porras, V. Yegneswaran, M. Fong, and W. Lee, "BotHunter: Detecting Malware Infection Through IDS-Driven Dialog Correlation," Proceedings of the 16th USENIX Security Symposium, pp. 167-182, Aug. 2007.
[17] G. Gu, R. Perdisci, J. Zhang, and W. Lee, "BotMiner: Clustering Analysis of Network Traffic for Protocol- and Structure-Independent Botnet Detection," In Proceedings of the 17th USENIX Security Symposium (Security'08), pp. 139-154, July 2008.
[18] J. Goebel and T. Holz, "Rishi: Identify bot contaminated hosts by IRC nickname evaluation," In Proceedings of the 1st Workshop on Hot Topics in Understanding Botnets (HotBots'07), Apr. 2007.
[19] 김영백, 이동련, 최중섭, 염흥열, "DNS 싱크홀 적용을 통한 악성봇 피해방지 기법 및 효과," 정보과학회학회지, 15(1), pp. 47-55, 2009년 1월.
[20] M. Caesar and J. Rexford, "BGP Routing Policies in ISP Networks," IEEE Network, vol. 19, no. 6, pp. 5-11, Nov. 2005.

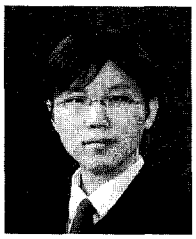
〈著者紹介〉

권 종 훈 (Jong Hoon Kwon) 학생회원
2007년 2월: 고려대학교 전산학과 학사
2008년 3월~현재: 고려대학교 컴퓨터·전파통신공학과 석사과정
〈관심분야〉 네트워크 보안, 악성코드, 봇넷 탐지

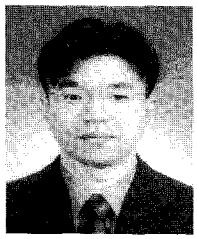


임 채 태 (Chaetae Im) 정회원
2000년 8월: 충남대학교 컴퓨터과학과 학사
2003년 2월: 포항공과대학교 컴퓨터공학과 석사
2003년 1월~현재: 한국정보보호진흥원 선임연구원
〈관심분야〉 : 정보보호, 이동통신, 네트워크

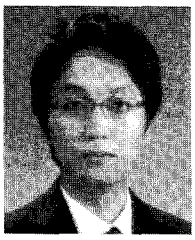
〈著者紹介〉



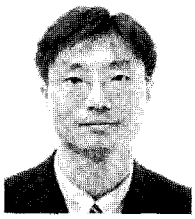
최 현 상 (Hyungsang Choi) 정회원
2004년 8월: 고려대학교 컴퓨터학과 학사
2006년 8월: 고려대학교 컴퓨터학과 석사
2007년 3월~현재: 고려대학교 컴퓨터·전파통신공학과 박사과정
〈관심분야〉 네트워크 보안, 공격시각화, 봇넷 탐지



지 승 구 (Seung Goo Ji) 정회원
1997년 2월: 중앙대학교 전자공학 학사
2002년 8월: 중앙대학교 전자전기공학 석사
2005년 6월~현재: 한국정보보호진흥원 선임연구원
〈관심분야〉 네트워크 보안, 보안시각화



오 주 형 (Joo Hyung Oh) 정회원
2005년 2월: 인제대학교 의용공학과 학사
2008년 2월: 성균관대학교 컴퓨터공학과 석사
2007년 12월~현재: 한국정보보호진흥원 연구원
〈관심분야〉 네트워크 보안, 악성코드 분석



정 현 철 (Hyun Cheol Jeong) 정회원
1996년 2월: 서울시립대학교 전산통계 학사
1998년 8월: 광운대학교 전자계산 석사
1996년 7월~현재: 한국정보보호진흥원 부장
〈관심분야〉 네트워크 보안

이 희 조 (Heejo Lee) 종신회원
1993년 2월: 포항공대 컴퓨터공학과 학사
1995년 2월: 포항공대 컴퓨터공학과 석사
2001년 2월: 포항공대 컴퓨터공학과 박사
2000년 3월~2001년 2월: Purdue University 박사후연구원
2001년 3월~2003년 10월: 안철수 연구소 CTO
2004년 3월~현재: 고려대학교 컴퓨터·전파통신공학과 부교수
〈관심분야〉 네트워크 보안, 인터넷웜/DDoS 공격 대응기술, 고가용성 시스템 설계