봇넷 탐지를 위한 네트워크 세션 분석

박종민*

Network Session Analysis For BotNet Detection

Jong-Min Park*

요 약

최근의 사이버 공격은 경쟁사에 대한 DDoS(Distributed Denial of Service)공격과 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 시그네처 기반 탐지 방법을 이용하거나, 극히 제한적인 지역의 봇넷를 탐지하고 있어, 총괄적 봇넷 대응에는 미흡한 것이 현실이다. 본 논문에서는 봇넷을 제어하기 위해 사용되는 IRC(Internet Relay Chat) 통신 세션에서 서버와 연결하는 채널과의 관계 분석을 통하여 봇에 감염된 호스트와 연결된 IRC서버 채널을 탐지하는 방법을 제안한다.

ABSTRACT

In recent years, cyber crimes were intended to get financial benefits through malicious attempts such as DDoS attacks, stealing financial information and spam. 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 paper, we propose IRC (Internet Relay Chat) that is used to control the botnet communication in a session channel of IRC servers connected through the analysis of the relationship of the channel and the connection with the server bot-infected hosts and how to detect.

키워드

봇넷, 분석, 탐지, 네트워크

Key words

BotNet, Internet Relay Chat, Analysis, Detection, Network

* 정회원 : 조선이공대학교 사이버보안과 (교신저자, jmpark@cst.ac.kr)

접수일자 : 2012. 06. 20
심사완료일자 : 2012. 07. 18

Open Access http://dx.doi.org/10.6109/jkiice.2012.16.12.2689

Ⅰ. 서 론

정보화 사회로의 발전이 거듭되면서 인터넷 및 네트워크이라는 용어는 현대 사회를 대표할 수 있는 키워드로 등장하게 되었다. 네트워크는 전 세계의 정보 시스템들을 연결하여 업무의 편리성 및 생활의 편리함을 제공하고 있다. 정보시스템에는 회사나 개인의 중요한 정보를 저장하고 있고 이러한 중요한 정보는 다양한 보안시스템으로 보호를 하고 있다, 그러나 다양한 해킹방법 및 악성 프로그램들로 인해 중요한 정보는 유출되거나 손상되고 있어 금전적인 피해는 물론 사회적으로도 문제를 야기 시키고 있다.

최근 이러한 악의적인 행위를 하는 주체로서 봇넷이 등장하게 되어 네트워크의 가장 큰 위협이 되고 있다.

봇(Bot)이라는 용어는 로봇(Robot)으로부터 유래되었고, 미리 정해진 기능들을 수행하기 위하여 자동화된 방법으로 설계된 스크립트(script) 혹은 스크립트 집합을 나타내기 위하여 사용되는 일반적인 용어이다. 이런 봇들의 네트워크가 봇넷(Botnet)이며, 감염된 혹은 침해된 기계들의 네트워크를 말한다[1][2]. 봇은 웜과 같은 형태로 개발되어 취약점이 존재하는 호스트들을 감염시킨다. 이렇게 전파되어 감염된 호스트들이 하나의 네트워크를 구성하는 것을 봇넷이라 한다.

기존의 웜이나 바이러스는 특정 시스템을 공격하거나 시스템을 손상시키는 등 취약한 시스템에 감염되어 정해진 행위만을 수행한다. 그러나 봇 형태의 웜은 봇에 감염된 호스트를 제어하는 봇 마스터에 의해 수행해야 할 명령을 전달받아 봇 마스터가 의도하는 악의적인 행위를 수행한다. 악의적인 행위중 대표적인 것이 분산서비스거부 공격 및 스팸메일의 발송이다. 이러한 악의적인 행위의 목적은 고전적인 의미와 확연하게 다른 양상을 보이고 있다. 특히 금전적인 이득을 목적으로 하는 공격이 증가하고 있다.

시만텍(Symantec)의 통계에 의하면 2006년 후반에만 600만대 이상의 봇이 생성된 것으로 보고되고 있으며, 2011년에는 하루 평균 6만대 이상의 호스트들이 새로운 봇에 감염되는 것으로 알려지고 있다. 또한 구글의 Vint Cert에 의하면 전 세계 인터넷에 연결된 호스트들 중 약 1억5000만대의 호스트들이 봇에 감염되었을 것으로 추측하고 있다[3].

그림 1은 인터넷 상에서 나타나는 악성코드를 나타낸다.

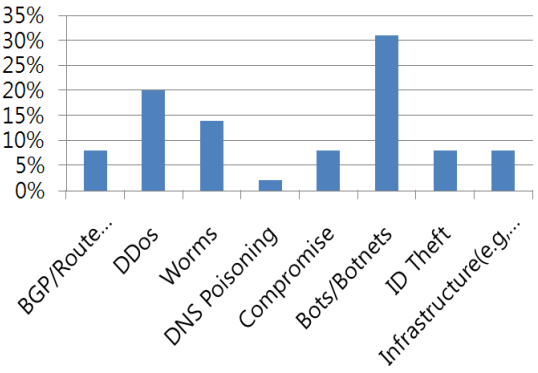

그림 1. 인터넷상의 가장 위협적인 요소
Fig. 1 The Internet coat most the element which is threatening

봇은 표준 IRC 네트워크 서비스 포트와 외부 연결을 설정하며 공격자 개인 채널에 가입한다.

IRC 네트워크는 공격자에게 수천 개의 봇을 통제할 수 있는 융통성 있게 제공한다. 합법적인 IRC 트래픽으로 가장함으로써 공격을 난해하게 만들고, 시스템 관리자에 의한 공격 소스의 추적도 어렵게 만든다.

본 논문에서는 탐지를 회피하기 위한 다양한 방법의 봇넷 제어 프로토콜이 존재하지만 여전히 봇넷을 제어하기 위한 수단으로 가장 많이 이용되고 있는 IRC 서버 채널과 이와 연결되어 명령을 수행하는 봇에 감염된 호스트간의 세션을 분석한다. 또한 봇이 IRC서버의 연결을 수행하는 과정에서 봇 생성 시 봇 마스터에 의해 지정된 채널에 참여한다는 것에 기반 하여 봇이 접속하는 서버와 채널의 관계를 통해 일반적인 대화 세션과 봇을 제어하기 위한 세션을 구분하는 방법을 제안 한다[4].

Ⅱ. 관련 연구

봇넷은 봇 마스터에 의해 제어가 되어 다양한 악의적인 행위를 수행한다. 봇 마스터가 봇을 제어하기 위한 방법으로 IRC를 주로 이용한다. 그러나 최근 IRC가 주로

사용하는 포트를 모니터링 함으로써 이를 차단하고 다른 포트를 사용하더라도 네트워크를 통해 IRC트래픽을 탐지하는 기술이 연구되었다. 그리하여 공격자들은 HTTP 또는 P2P를 이용한 봇을 개발하여 일반적인 네트워크 트래픽과 봇넷 트래픽을 구분할 수 없도록 하였다. 그러나 아직도 봇넷을 제어하기 위한 수단으로 IRC가 주로 사용되고 있다.

2.1. IRC(Internet Relay Chat)

IRC는 Internet Relay Chat의 머리글자를 딴 것으로, 인터넷 실시간 대화를 뜻한다. IRC는 IRC 클라이언트 프로그램이나 IRC 클라이언트를 제공하는 서버에 접속하기만 하면 시간이나 공간에 구애 받지 않고 전 세계의 어떤 사람과도 대화가 가능하다. 동시에 다중 대화가 가능한 채팅 구조를 갖는다[5]. IRC서버는 서버와 서버가 연결되어 서버에 개설되어 있는 채널을 통해 대화가 가능한 구조를 갖는다. IRC서버의 구조는 간단하게 그림 2와 같이 구성될 수 있다[6]. 실제 인터넷 상에서의 IRC서버의 구성은 그림 2의 확장된 구조를 갖는다.

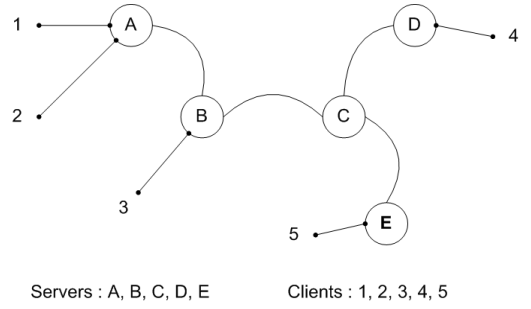

그림 2. IRC 네트워크 기본 구조
Fig. 2 IRC Network Basic structure

IRC서버는 일반적으로 6667과 6668번 포트와 TCP프로토콜을 사용한다. IRC는 다양한 명령어를 통해 대화채널을 생성, 참여, 대화의 과정을 수행할 수 있다. IRC 프로토콜은 IETF(The Internet Engineering Task Force)의 RFC문서로도 정의가 되어 있다. IRC서버에 접속하여 대화채널에 참여하기 위해서는 일반적으로 IRC 클라이언트 프로그램을 사용한다[7].

2.2. 봇넷(BotNet)

봇넷은 봇에 감염된 호스트들로 구성된 네트워크를 말한다. 봇(Bot)은 운영체제의 취약점이나 비밀번호 취약성, 웜/바이러스의 백도어 등을 이용하여 전파되며, 명령을 전달하는 서버와 연결되어 분산서비스거부 공격이나 스팸메일 발송에 악용 가능한 프로그램이다[8].

봇 마스터는 IRC서버에 대화채널을 만들고 봇에 감염된 호스트가 접속을 하면 봇이 수행할 명령을 전달하고 봇은 전달받은 명령을 호스트에서 수행한다. 또한 감염된 봇을 업데이트할 수 있는 명령을 수행하기도 한다. 봇에 감염된 호스트가 수행하는 기능은 특정 서버로의 분산서비스거부공격, 스팸메일 발송, 감염된 호스트 내에서 중요한 파일의 전송 등을 통한 정보유출이 있다[9].

그림 3은 IRC 봇넷의 기본적인 구조를 나타낸다.

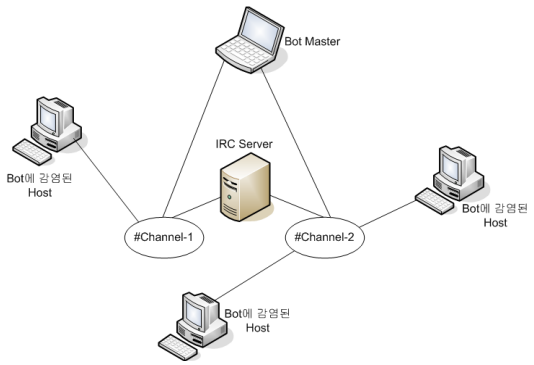

그림 3. IRC BotNet 기본 구조
Fig. 3 IRC BotNet Basic structure

봇넷을 탐지하기 위한 연구도 활발하게 이루어지고 있지만 현실적으로 적용가능하고 효과적인 방법은 제시되지 못하고 있다. 봇넷을 탐지하기 위한 기법으로는 안티바이러스 제품에서 파일기반의 진단으로 적용하고 있는 시그니처 탐지기법(Signature Detection Tech)과 시그니처 탐기기법의 한계를 보완하기 위해 개발된 휴리스틱 탐지기법(Heuristic Detection Tech)등이 제안되었다[9][10].

Ⅲ. IRC Channel 분석을 위한 설계

IRC봇은 봇 마스터가 봇을 생성할 때 특정 서버 도메인 또는 IP로 접속하여 명령을 수신하도록 설계가 되어 있다. 봇들은 봇 마스터가 미리 정의한 특정 명령어에 반응하도록 설정이 되어 있어서, 표준화된 IRC명령어 이외의 다른 명령전달은 일반 대화의 형식으로 전달된다.

그림 4는 IRC 접속과 Chat 과정을 나타낸다.

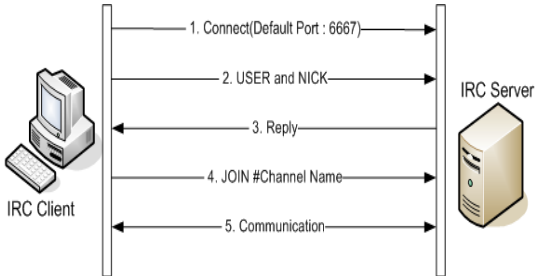

그림 4. IRC 접속 및 Chat 프로세스
Fig. 4 IRC Contact and Chat Prosess

IRC봇이나 일반 Chat의 과정은 공통적으로 사용자와 대화명(Nick)을 통해 접속클라이언트의 정보와 대화명을 서버로 전송한다. 서버와의 접속이 이루어진 후에는 기본적으로 클라이언트가 의도하는 채널에 연결명령을 통해 참여하게 된다. 이 과정에서 봇은 봇 마스터가 생성한 채널에 참여해야 수행할 명령을 전달받을 수 있기 때문에 봇에 감염된 호스트들은 특정 서버의 특정 채널에 참여하게 된다. 기존에 봇이 IRC서버에 접속하여 채널에 참여할 때 사용하는 Nickname의 형태를 분석하여 봇넷 세션을 판단하는 연구가 이루어졌다[10][11]. 하지만 Nickname은 다양한 생성규칙에 의해 다르게 생성될 수 있기 때문에 우회가 가능하다.

IRC 봇들이 참여하는 채널의 경우 봇이 업데이트되어 다른 특성, 즉, 접속 서버 및 송수신 명령 메시지 등이 변경되더라도 봇 마스터로부터 명령을 수신하기 위해 접속해야할 채널의 이름은 변경이 되지 않는다. 만약 업데이트를 통해 채널명이 변경된다 해도 봇에 감염된 호스트들은 공통적인 채널에 참여를 해야 하기 때문에 접속대상 서버와 채널 명을 모니터링 및 분석함으로써 IRC통신에 있어서 봇넷 제어용으로 사용되는 서버 및 채널, 봇에 감염된 호스트의 정보를 판별할 수 있다.

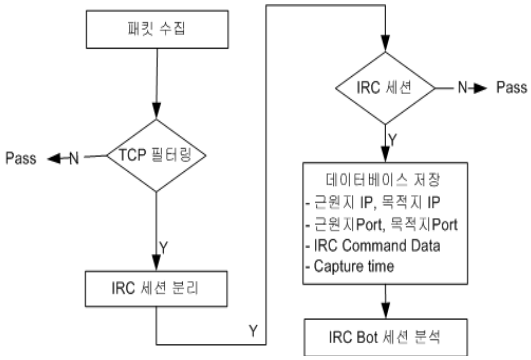

그림 5. IRC세션 패킷 수집 프로세스
Fig. 5 IRC Session packet collection process

그림 5는 패킷수집 단계에서 수집된 트래픽 데이터는 탐지된 기존 봇넷의 행위 분석을 위해 TCP/UDP 트래픽 전체가 포함되는데 트래픽에서 분석에 필요한 정보들만 추출하기 위해 프로토콜, IP, 포트, 갈무리 시간, 패킷 트래픽을 필터링하는 과정을 나타낸다. 그림 6은 IRC세션을 분석하기 위해서는 TCP IRC세션만 필요하기 때문에 UDP패킷의 수집은 제외하고, 채널명, 채널에 연결하는 클라이언트 IP, 채널개설 서버IP, 서버 포트를 분석하는 과정을 나타낸다.

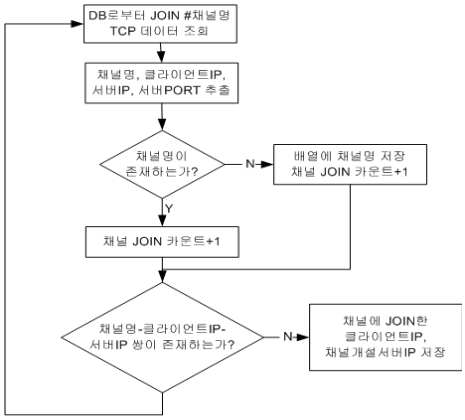

그림 6. IRC채널 분석 프로세스
Fig. 6 IRC Channel analytical process

그림 7은 IRC채널을 분석하는 과정에서는 TCP IRC 세션 중 JOIN #Channel을 포함하는 패킷에서 채널명, 채널에 연결하는 클라이언트, 채널개설 서버IP주소를 추출하여 채널명 집합, 채널별 연결 IP집합, 채널별 연결 빈도수 등의 채널 정보 분석을 실행하였다.

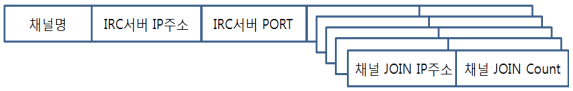

그림 7. IRC 채널 분석 정보 구조
Fig. 7 IRC Channel analytical information structure

Ⅳ. 실험 및 결과 분석

제안하는 실험을 위하여 IRC 봇넷 구조의 네트워크를 구성하였다. 일반 대화 및 봇넷 제어를 위해 개설한 채널을 구성하고, 공개된 봇 프로그램을 사용하여 실험 대상 시스템에 봇을 감염시켜 실험을 실행하였다.

실험대상 호스트 수 : 69 대
패킷 수집 시간 : 16 시간
수집된 패킷 수(TCP) : 222,669 개

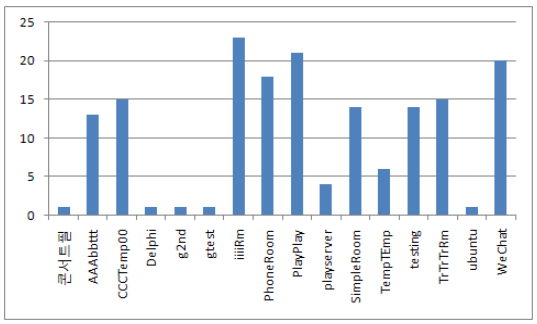

그림 8. IRC 채널별 연결 IP 주소의 수
Fig. 8 Possibility of IRC Channel by Join Ip Addresses

제안하는 실험결과를 살펴보면 호스트들이 iiiiRm, playplay, wechat의 채널에 참여하는 수가 현저히 많이 나타난다. 그림 8 은 봇에 감염된 호스트들이 봇 마스터의 명령을 전달받기 위해 iiiiRm, playplay, wechat의 특정한 채널에 많이 접속하기 때문에 봇넷 채널의 가능성이 매우 높다. 그림 9는 각 채널별로 실험 시간동안 접속하는 회수에서도 phoneRoom, simpleRoom, AAAbbttt 채널에 호스트들이 접속하는 빈도수가 현저하게 높게 나타나고 있다.

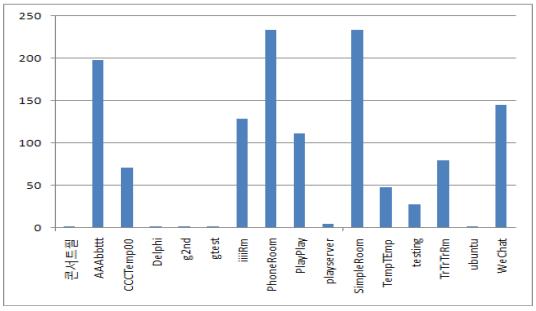

그림 9. IRC 채널별 연결 빈도수
Fig. 9 IRC Channel by Join frequency possibility

또한 실험을 통하여 대부분의 네트워크에서 외부에서 내부 망으로 연결하기 위한 요청 중에서 목적지 포트가 IRC의 기본포트로 사용되는 6667과 6668은 방화벽으로 차단하고 있음도 확인하였다. 기존에는 봇넷 위협에 효과적으로 대응하기 위하여 첫 번째 방법은 블랙리스트를 공유하는 방법이다. 두 번째 방법은 DNS 싱크홀 기법을 활용하는 것이다. 세 번째 방법은 HTTP 봇넷 C&C URL 접근 차단을 통한 HTTP 봇넷의 대응이다. 네 번째 방법으로는 BGP feeding 기법을 이용하여 무력화할 수 있다. 기존의 봇넷에 대한 탐지, 대응 및 완화방안이 제시 되었는데 제안하는 방법 또한 봇넷 위협에 효과적으로 대처할 수 있었다.

Ⅴ. 결 론

봇의 전파와 감염 과정이 단축되고 자동화됨에 따라서 인터넷상에 높은 대역폭을 가진 환경에서 심각한 위협을 제시하고 있다. 본 논문에서는 최근 네트워크의 가장 큰 위협이 되고 있는 봇넷의 세션을 탐지하기 위해 봇에 감염된 호스트들이 IRC서버에 접속하여 참여하는 채널 정보를 분석한다. 봇에 감염된 다수의 호스트

들이 봇 마스터의 제어를 받기 위해서 지정된 채널에 접속을 하기 때문에 호스트의 참여 빈도가 높은 채널이 봇 마스터의 제어 채널일 가능성이 매우 높음을 확인하였다. 또한 방화벽에 의해 IRC의 기본 포트가 차단되고는 있었기 때문에 서버의 포트번호를 변경한 서버를 통해 봇 마스터의 제어명령을 전달하고 있음도 확인할 수 있었다.

본 논문에서는 제한된 네트워크 환경에서 봇넷 채널을 분석하였지만 좀 더 확장된 네트워크상에서 IRC 연결 세션을 분석함으로써 일반 대화와 봇넷 제어용 채널을 분류할 수 있는 연구와 IRC봇넷 뿐만 아니라 HTTP와 P2P 프로토콜을 이용하는 봇넷에 대한 지속적인 연구가 필요하며 탐지 및 관제 시스템의 구조가 국제 표준화 제정이 필요하다고 예측된다.

참고문헌

[1] 전용희, 오진태, “봇넷 분류법 및 진화된 봇넷 구조”, 한국정보보호학회지, 제18권 제4호, pp.76-86, 2008.

[2] 권종훈, 임채태, 최현상, 지승구, 오주형, 정현철, 이희조, “협업 기반의 중앙집중형 봇넷 탐지 및 관제 시스템 설계”, 한국정보보호학회논문지, 제19권 제3호, pp.83-93, 2009.

[3] 권종훈, 임채태, 최현상, 정현철, 이희조. “봇넷의 탐지 및 관제 시스템 설계”, 한국정보처리학회, 제15권, 2호, pp.1517-1920, 2008.

[4] 박경진, 노창오, 박종민, 조범준, “IRC 봇넷 탐지를 위한 네트워크 세션분석”, 한국멀티미디어학회 춘계학술발표대회 논문집, 제12권, 1호, pp.99-102, 2009.

[5] http://www.encyber.com, keyword : IRC

[6] C. Kalt. ″Internet Relay Chat: Architecture″, April 2000. Request for Comments: RFC2810

[7] C. Kalt. ″Internet Relay Chat: Client Protocol″, April 2000. Request for Comments: RFC2812

[8] 인터넷침해사고대응지원센터, “인터넷 침해사고 동향 및 분석 월보(3월)”, 한국정보보호진흥원, pp.10, 2009.

[9] Jose Nazario, ″DDoS attack evolution″, Network Security, Vol.2008, pp7-10, 2008.

[10] 박준홍, “행동패턴기반 유사 악성코드 탐지기법에 관한 연구”, 목원대학교, 2008.

[11] Jan Geobel, ″Rishi: Identify Bot Contaminated Hosts by IRC Nickname Evaluation″, HotBots2007, 2007.

저자소개

박종민(Jong-Min Park)

1988년 조선대학교 공학석사
2005년 조선대학교 공학박사
2008년 ~ 현재. 조선이공대학교 사이버보안과 교수

※관심분야 : 바이오인식, 패턴인식, 인공지능, 정보보호 및 보안, 네트워크보안