情報保護學會論文誌
第19卷 第4號, 2009. 8

허니넷을 이용한 P2P 기반 Storm 봇넷의 트래픽 분석*

한 경 수,† 임 광 혁, 임 을 규‡
한양대학교

The Traffic Analysis of P2P-based Storm Botnet using Honeynet*

Kyoung Soo Han,† Kwang Hyuk Lim, Eul Gyu Im‡
Hanyang University

요 약

최근 인터넷 상에서 봇넷을 이용한 사이버 공격이 증가하고 있으며, 이러한 공격들은 금전적 이득을 목적으로 하고 있어 범죄화 양상을 보이고 있다. 봇넷을 이용하는 사이버 공격으로는 스팸 발송, 분산서비스 거부(DDoS) 공격, 악성코드 및 맬웨어(malware) 전파, 피싱, 개인정보 유출 등이 있다. IRC나 HTTP 봇넷과 같은 중앙 집중형 구조의 봇넷은 그 탐지나 완화 방법의 연구가 다수 존재하지만, P2P 봇넷에 대한 연구는 아직 초기 단계이다. 본 논문에서는 다양한 네트워크 공격의 능동적 분석에 활용되는 허니넷을 이용하여 P2P 기반 Storm 봇 중의 하나인 Peacomm 봇이 발생시키는 트래픽을 분석하였다. 그 결과 Peacomm 봇이 P2P를 통해 광범위한 외부 네트워크의 좀비를 대상으로 다량의 UDP 패킷을 발생시키는 것을 확인하였다. 또한 이를 통해 Peacomm 봇이 봇넷의 규모를 유지하거나 확장한다는 것을 알 수 있었다. 이는 P2P 봇넷을 탐지하고 완화시킬 수 있는 대응기술 마련의 기초로써 사용될 수 있을 것으로 기대된다.

ABSTRACT

Recently, the cyber-attacks using botnets are being increased. Because these attacks pursue the money, the criminal aspect is also being increased. There are spreading of spam mail, DDoS(Distributed Denial of Service) attacks, propagations of malicious codes and malwares, phishings, leaks of sensitive informations as cyber-attacks that used botnets. There are many studies about detection and mitigation techniques against centralized botnets, namely IRC and HTTP botnets. However, P2P botnets are still in an early stage of their studies. In this paper, we analyzed the traffics of the Peacomm bot that is one of P2P-based storm bot by using honeynet which is utilized in active analysis of network attacks. As a result, we could see that the Peacomm bot sends a large number of UDP packets to the zombies in wide network through P2P. Furthermore, we could know that the Peacomm bot makes the scale of botnet maintained and extended through these results. We expect that these results are used as a basis of detection and mitigation techniques against P2P botnets.

Keywords: P2P Botnet, Storm Botnet, Botnet Traffic Analysis, Honeynet

I. 서 론

최근 인터넷 상에서 봇넷을 통해 이루어지는 공격이 급증하고 있다. 봇넷(botnet)이란 악의를 가진 공격자인 봇 마스터(bot master)가 보안이 취약한 PC들에 소프트웨어적 로봇을 의미하는 봇(bot)을 감염시키고, 봇에 감염된 수많은 PC들이 네트워크를 형성한 것이다[1-3]. 또한 각 봇넷이 사용하는 프로토콜

접수일(2008년 11월 17일), 수정일(2009년 5월 13일), 게재확정일(2009년 7월 3일)
* 이 논문은 2008년도 정부재원(교육인적자원부 학술연구조성사업비)으로 한국학술진흥재단의 지원을 받아 연구되었음. (KRF-2008-331-D00573)
† 주저자, 1hanasun@hanyang.ac.kr
‡ 교신저자, imeg@hanyang.ac.kr

에 따라 IRC/HTTP/P2P 봇넷으로 분류할 수 있고, 그 구조에 따라 중앙 집중형(IRC 및 HTTP 봇넷)과 분산형(P2P 봇넷)으로 분류할 수 있다[4]. 봇 마스터는 이러한 봇넷을 이용하여 자신의 위치를 노출시키지 않고 스팸 발송, 분산 서비스 거부(DDoS: Distributed Denial of Service) 공격, 피싱 등의 공격을 수행할 수 있으며, 이러한 공격에 의해 개인정보 유출 등의 제 2차 피해로 이어질 수 있다. IRC나 HTTP 봇넷과 같은 중앙 집중형 구조의 봇넷은 그 탐지나 완화 방법에 대한 기존 연구가 다수 존재하지만 [5-7], P2P 봇넷에 대한 연구는 아직 초기 단계이다. 따라서 본 논문에서는 P2P 봇넷의 탐지 및 완화 연구에 기초를 마련하고자 다양한 네트워크 공격의 능동적 분석에 활용되는 허니넷(honeynet)을 이용하여 P2P 기반 Storm 봇 중의 하나인 Peacomm 봇이 발생시키는 트래픽을 분석하였다.

본 논문의 Ⅱ장에서는 봇넷의 기본 배경과 봇넷 분석과 관련된 연구를 기술한다. Ⅲ장에서는 허니넷을 이용한 분석 환경을 설명하고, Ⅳ장에서는 P2P 기반 Peacomm 봇의 트래픽 분석 내용을 기술한다. Ⅴ장에서는 결론과 함께 본 논문에서 분석한 결과를 바탕으로 한 향후 연구 과제에 대하여 언급한다.

Ⅱ. 관련 연구

2.1 봇넷(Botnet)

2.1.1 봇넷이란

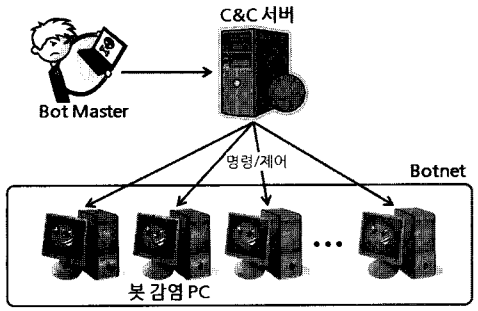
사이버 공격자들은 악성 행위를 통해 금전적인 이득을 취하기 위하여 자신의 위치가 노출되지 않으면서 자신이 통제하고 제어하는 컴퓨터 네트워크를 구성하는 기술을 고안하게 되었다. 즉, 악의를 가진 공격자인 봇 마스터(bot master)가 보안이 취약한 PC들을 마음대로 제어할 수 있도록 소프트웨어적 로봇을 의미하는 봇(bot)을 감염시키고, 이렇게 감염된 수천~수만 대 이상의 PC들이 네트워크를 통해 연결되어 제어 권한을 가진 봇 마스터에 의해 원격 조종된다. 이러한 형태를 봇넷(botnet)이라고 한다. 봇넷은 공격자인 봇 마스터, 봇 마스터로부터 명령을 전달받아 악성 행위를 수행하는 봇 감염 PC 혹은 좀비(zombie), 봇 마스터와 감염된 PC들 사이에서 명령을 전달하거나 실행결과 반환의 역할을 하는 C&C(Command and Control) 서버로 구성된다[1-3]. PC는 여러 경로를 통해서 봇에 감염된다. 스팸에 첨부된 실행 파일을 클릭해 감염되는 경우, 웹 페이지에 악성 코드를 포함시키고 취약성을 가진 PC를 감염시키는 경우, 메신저를 통한 감염 등 그 경로가 다양하다. 이렇게 구성된 봇넷은 봇 마스터가 주로 스팸 발송, 분산 서비스 거부(DDoS: Distributed Denial of Service) 공격, 피싱 등의 공격에 이용된다.

2.1.2 봇넷의 분류

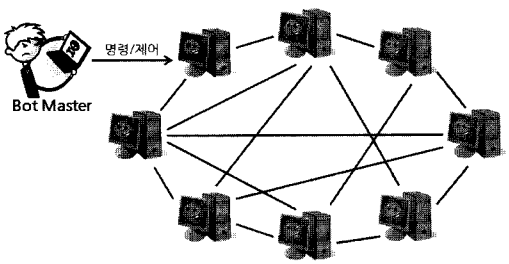
각각의 봇넷이 제어 명령을 전달하기 위해 사용하는 프로토콜에 따라 IRC/HTTP/P2P 봇넷으로 분류할 수 있다. 또한 그 네트워크 구성에 따라 [그림 1]과 같은 중앙 집중형(IRC 및 HTTP 봇넷)과 [그림 2]와 같은 분산형(P2P 봇넷)으로 분류할 수 있다[4].

1) IRC 봇넷

IRC 봇넷은 봇넷 중 가장 초기 형태인 중앙 집중형 구조를 이루고 있으며, 2004년 이후부터 그 수가 급속하게 증가하였다. IRC는 Internet Relay Chat의 약어로, 원래 인터넷 상에서 채팅을 할 수 있



[그림 1] 중앙 집중형 구조(IRC, HTTP)의 봇넷



[그림 2] 분산형 구조(P2P)의 봇넷

도록 고안된 프로토콜이다. IRC 프로토콜은 서버들끼리 직·간접적으로 연결되어 세계 어느 곳이든 한 서버에 연결하면 자동으로 전 세계의 모든 서버와 연결이 된다. 따라서 명령 전달을 위한 "전용 채팅 채널"을 쉽게 구성할 수 있고, 봇넷의 규모가 계속 증가하게 될 경우 봇 감염 PC들을 그룹별로 나누어 관리할 수 있는 채널이 존재하여 공격이나 관리에 용이하다. 봇 마스터로부터 선택된 IRC 서버는 봇에 감염된 PC들에게 명령을 내리고 악성 코드를 업데이트할 수 있는 C&C 서버로 사용된다[6].

2) HTTP 봇넷

HTTP 봇넷은 IRC 봇넷과 같이 중앙 집중형 구조를 이루고 있다. 그러나 IRC 봇넷과는 달리 인터넷 상에서 많이 사용되는 웹 서비스를 명령 전달의 매개체로 사용한다. 또한 웹 서비스를 이용함으로써 탐지와 차단을 회피할 수 있는데, 이는 자주 사용되는 80번 웹 포트를 차단하기 어렵고 정상적인 트래픽이 빈번하게 발생되기 때문이다. 즉, HTTP 봇넷은 IRC 봇넷의 구조와 비슷하지만 HTTP를 이용하여 명령 및 제어를 전달하는 차이점이 존재한다.

3) P2P 봇넷

2007년 초에 등장한 P2P 봇넷은 기존의 봇넷 구조와는 다른 구조로 구성된다. 즉, P2P 봇넷은 기존의 중앙 집중형 명령/제어 방식을 사용한 IRC 봇넷이나 HTTP 봇넷과는 달리, 각각의 봇 감염 PC들이 서로 연결되어 있어 모든 피어들이 C&C 서버 역할을 하게 된다. 따라서 이러한 P2P 봇넷의 경우 하나의 C&C 서버를 무력화하더라도 봇넷을 제거할 수 없고, C&C 서버 역할을 하는 모든 피어를 폐쇄해야만 봇넷을 제거할 수 있다[8].

2.1.3 Storm 봇넷

P2P 봇넷 중의 하나인 Storm 봇넷은 Zhelatin이나 Peacomm으로 알려져 있는 봇에 감염된 PC들의 네트워크이다[8]. 이는 주로 사회공학적 기법을 통해 사용자들이 게시판의 링크나 스팸의 첨부 파일을 클릭하도록 유도함으로써 전파되며 브라우저 익스플로잇, 악성 코드가 포함된 파일의 다운로드를 통해서도 전파된다. 이러한 전파 경로를 통해 P2P 봇넷을 형성하여 원격지로부터 여러 가지 기능을 수행하는 악성 코드를 다운로드하고 실행하며, 사용자들의 중요 정보 수집, 웹 사이트 공격, 스팸 발송 등의 도구로써 사용된다. Storm 봇넷의 크기와 각각의 감염된 PC들의 성능을 정확히 파악할 수 없지만, 그 위력은 약 2백만의 봇에 감염된 PC들이 메이저 슈퍼컴퓨터 하나의 성능을 능가할 것이라 추측되기도 한다[9].

2.2 봇넷의 분석을 위한 기술

2.2.1 봇넷 역추적 및 분석 기술

기존의 봇넷 역추적 및 분석 기술은 대표적으로 추적 클라이언트 기술, 메모리 감시 기반 역추적 기술, DNS 쿼리를 이용한 행위 기반 탐지 기술 등이 있다. 추적 클라이언트 기술은 분석자가 시스템에 취약점을 의도적으로 생성하여 이를 에뮬레이션하고 봇 프로그램이나 파일들을 수집하는 방식의 기술이다[1]. 그러나 이는 수집된 봇 프로그램이나 파일을 수동으로 정적 분석을 해야 하며, 봇넷의 구조나 그 규모를 파악할 수 없다. 메모리 감시 기반의 역추적 기술은 봇의 행위를 지속적으로 모니터링하여 추적하기 위한 기술로, 메모리 감시를 통해 봇넷의 추적 회피 기술을 처리하며 네트워크 감시를 통해 역추적 모듈이 악성 행위를 수행하려는 가능성에 대한 위험도를 낮춘다[3]. DNS 쿼리를 이용한 행위 기반 탐지 기술은 봇넷을 구성하는 봇들이 발생시키는 DNS 쿼리가 정상적인 DNS 쿼리와 구분되고, 주로 DDNS를 이용한다는 특징을 이용하여 봇넷을 탐지한다[10]. 이 기술은 IRC 봇넷에만 적용할 수 있는 기술이다. 네트워크 트래픽을 기반으로 봇의 행위를 분석한 연구로 [11]에서는 봇을 탐지하기 위해 플로우 데이터를 분석하고, 봇들이 접속하는 C&C 서버 사이의 데이터를 모니터링하였다.

2.2.2 허니넷(Honeynet)

최근 다양한 침해 행위에 대한 피해 현황을 파악하고 대응하기 위해 침입탐지시스템(IDS: Intrusion Detection System) 기반의 솔루션들이 사용되고 있으며, 허니넷 또한 능동적인 침해 현황 파악을 위해 사용되고 있다. 허니넷(honeynet)은 웜이나 바이러스 등의 악성 코드에 대하여 네트워크 자원을 제공함으로써 여러 가지 정보를 얻고, 그 정보를 가지고 안정적인 인터넷을 위해 정보를 도출하는 네트워크라고 할 수 있다[12,13]. 또한 허니넷은 실제 네트워크 및 자원과

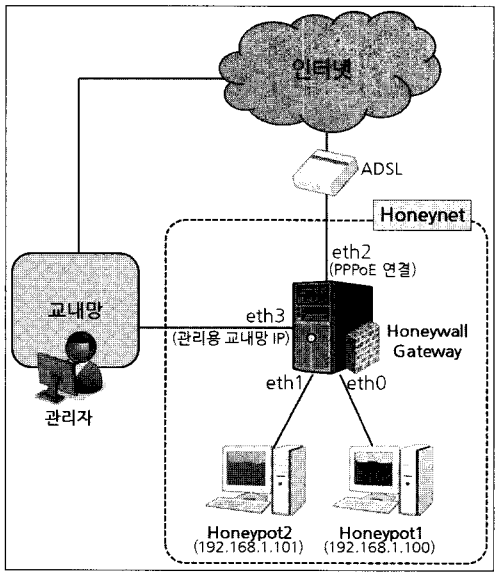
는 격리된 가상의 네트워크이며, 침입자나 해커 등의 공격자를 유인하여 찾아내기 위한 목적이 아니라, 그 행동을 모니터링하고 분석하기 위한 목적으로 사용되기 때문에 공격자는 이러한 사실을 알지 못한다.

허니넷을 구성하는 하드웨어적인 필수 요소는 허니팟(honeypot)이다. 허니팟은 공격자에게 공격을 당하는 것처럼 보이거나, 실제로 악성 코드를 유인하고 감염되어 악성 행위를 수행시킬 수 있는 컴퓨터이다[3]. 분석자는 이를 통해 공격자를 역추적하거나 공격자 및 악성 코드의 행동 분석에 필요한 정보를 수집할 수 있다.

허니넷을 구성하는 소프트웨어적인 도구로는 데이터 제어 도구, 데이터 캡처 도구, 데이터 분석 도구가 있으며, 이는 일종의 방화벽 역할을 하는 허니월(honeywall)에 구성된다. 이를 통해 허니팟 및 허니넷의 트래픽을 제어하거나 관찰 및 저장하고, 분석할 수 있다[13].

III. 분석 환경

기존의 연구[10]에서는 IRC 봇넷의 DNS 트래픽 분석을 통해 사용자들에 의한 정상적인 DNS 트래픽과 비교함으로써 봇넷을 탐지하고자 하였다. 본 논문에서는 P2P 봇넷이 가지는 행위적 특징을 찾아 P2P 봇넷을 탐지하고 완화시킬 수 있는 대응기술에 대한 기초를 마련하기 위해 [그림 3]과 같이 허니넷을 구축



[그림 3] 허니넷 구축 모델

하였으며, 이를 이용하여 각각의 허니팟에 P2P 기반의 Storm 봇 중 하나인 Peacomm 봇을 감염시킨 후, 허니팟에서 발생하는 P2P 봇의 전체적인 트래픽을 분석하였다.

허니넷을 구성하는 각 컴퓨터들의 사양은 [표 1]과 같다.

[표 1] 허니넷 구성 컴퓨터 사양

	OS	CPU	RAM
Honeywall	CentOS	Pentium D 3.0GHz	2048MB
Honeypot1	Windows XP SP2	Pentium 4 3.0GHz	1024MB
Honeypot2	Windows XP SP2	Pentium 4 3.0GHz	1024MB

허니월에는 4개의 네트워크 인터페이스가 존재한다. eth0과 eth1은 각각 허니팟이 브릿지로 연결되고, eth2는 PPPoE를 통해 인터넷에 연결된다. 또한 eth3는 허니월 서버에 SSH 접속 및 SFTP 접속을 할 수 있도록 교내망의 IP를 부여함으로써 트래픽 모니터링과 관리, 트래픽 덤프 파일 다운로드가 가능하다. 허니넷 구축 도구로는 허니넷 프로젝트(honeynet project)에서 배포하는 허니월 CD Roo 1.4 버전을 사용하였으며, 이는 CentOS 및 Snort를 기반으로 하는 데이터 제어 도구, 데이터 캡처 도구, 데이터 분석 도구를 모두 포함한다[12].

Peacomm 봇을 감염시킨 허니팟의 운영체제로 Windows XP 서비스팩2를 설치하고, 추가적인 보안 업데이트는 제외하였다. 이는 봇넷에 감염된 PC들의 약 70%가 Windows 환경에서 동작하고, 거의 대부분의 일반 사용자들이 주로 사용하는 운영체제이기 때문이다[14].

허니월에서 트래픽을 관찰하기 위해 Tshark를 사용하였다. Tshark는 윈도우 환경에서 널리 사용되는 네트워크 트래픽 캡처 도구인 Wireshark의 콘솔용 커맨드 라인 버전이다[18]. 다음의 명령어를 통해 트래픽 덤프를 할 수 있으며, 브릿지로 연결된 각각의 허니팟에 대한 모든 인바운드 및 아웃바운드 트래픽을 30분 단위의 파일로 나누어 저장한다.

```
tshark -i br0 -b duration:1800 -w filename1
tshark -i br1 -b duration:1800 -w filename2
```

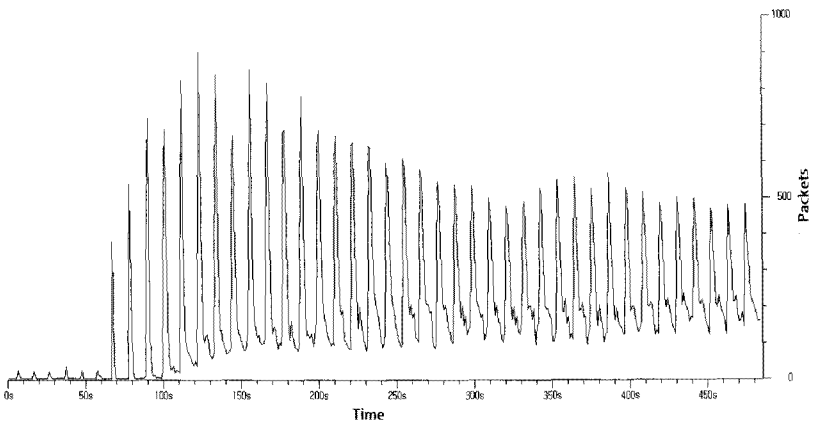
IV. 분석 결과

4.1 트래픽의 양과 특징

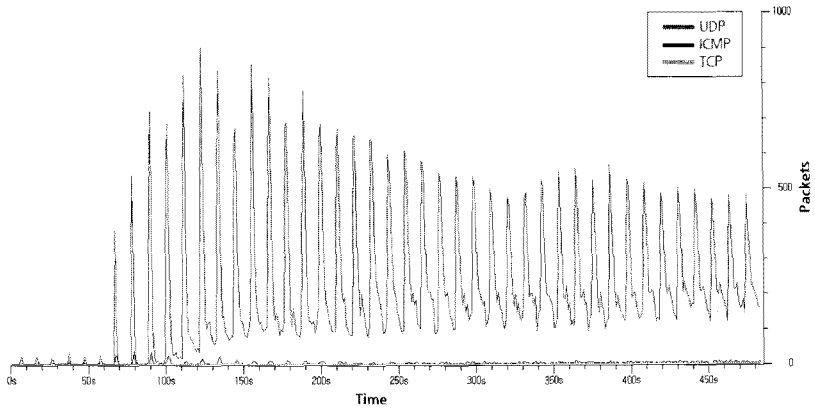
2008년 8월부터 10월까지 3개월간 Peacomm 봇을 감염시킨 허니팟에서 발생되는 모든 트래픽에 대하여 샘플링을 하지 않고 저장도록 하였으며, 30분마다 생성되는 덤프 파일의 크기는 평균 약 55MB이었다.

먼저 Peacomm 봇이 허니팟에서 동작을 시작하면, 봇넷에 속해 있는 다른 감염된 PC들과의 통신을 시도한다. 이때 허니팟은 자신의 존재를 알리고 감염된 PC들에 대한 정보를 얻어오는 과정에서 다량의 UDP 패킷이 발생한다. 이는 IRC 봇넷이나 HTTP 봇넷이 C&C와의 통신을 위해 UDP 패킷보다 TCP 패킷이 더 많이 발생하는 특징과는 반대이다. 즉, 모든 피어들이 C&C가 될 수 있는 P2P 봇넷은 소수의 C&C만을 포함하는 기존의 봇넷과 다르게 피어들 간의 존재 여부 및 정보를 알리기 위하여 UDP 패킷을 발생시킨다. [그림 4]는 전체 패킷의 트래픽 양을, [그림 5]는 UDP/TCP/ICMP 패킷의 트래픽 양을 나타낸 것이다.

[그림 4]와 [그림 5]의 트래픽 변화를 비교해보면 그래프가 거의 유사한 패턴을 보이는 것을 알 수 있다. 이는 대부분의 패킷이 UDP라는 것을 의미한다. Peacomm 봇이 감염된 직후 1분 정도는 트래픽이 거의 발생하지 않다가 1분 후에는 트래픽이 갑자기 증가하여 발생하는 초당 패킷 수가 최대 900개 정도이며, 그 이후로는 발생하는 초당 패킷 수의 최대치는 점차 감소하고 최소치는 점차 증가하는 패턴을 보인다. 또한 7분 정도가 지나면 발생하는 초당 패킷 수의 최대치와 최소치는 각각 약 500개 정도와 150개 정도로 거의 일정하게 유지된다. [15]에 따르면, 또다른



[그림 4] 전체 패킷의 트래픽



[그림 5] UDP/TCP/ICMP 패킷의 트래픽

P2P 봇인 SpamThru와 Nugache는 약 20분이 경과하더라도 패킷의 발생량이 지속적으로 증가한 후 일정하게 유지되는 양상을 보이나, Strom 봇의 경우 10분 이내에 패킷의 발생량이 급증하여 최대치로 유지되는 독특한 형태의 특징을 보인바 있다.

〔표 2〕는 10시간동안 저장된 전체 패킷 중 UDP 패킷이 차지하는 비율을 30분 간격으로 나누어 나타낸 것이다. 10시간동안 저장된 전체 패킷 수는 12,107,062개였으며, UDP 패킷은 11,987,673개로 99.01%를 차지하였고 그 외에는 ICMP 패킷이 0.93%, TCP 패킷이 0.06%를 차지하였다. 또한 30분간 발생하는 평균 패킷 수는 약 605,353개이고, 그 중 UDP 패킷은 599,383개였다.

〔표 2〕 전체 패킷 수와 UDP 패킷 수

시간 (분)	전체 패킷 수	UDP 패킷 수	비율 (%)
30	418,254	411,566	98.40
60	488,334	481,597	98.62
90	528,063	521,625	98.78
120	541,336	535,512	98.92
150	537,731	531,951	98.93
180	537,955	531,567	98.81
210	558,750	552,252	98.84
240	583,856	577,930	98.99
270	615,446	609,054	98.96
300	621,118	614,612	98.95
330	648,998	642,373	98.98
360	670,910	665,236	98.15
390	684,297	679,442	99.29
420	681,693	676,414	99.23
450	691,422	686,250	99.25
480	682,391	677,291	99.25
510	673,966	668,607	99.20
540	661,343	655,754	99.15
570	648,186	642,107	99.06
600	633,016	626,533	98.98

4.2 16-Bit Prefix에 따른 분류

봇의 여러 가지 전파 방법 중 한 가지는 네트워크 서브넷 스캐닝을 통해 취약점이 발견된 PC를 감염시키는 것이다[5]. 즉, 봇이 전파 대상을 스캔할 때 IP 주소의 특정 상위 비트는 고정하고, 하위 비트들을 변경해서 생성된 IP 주소를 스캔하여 취약한 PC를 찾아낸다. 따라서 이러한 과정에서 네트워크 주소가 동일한 IP 주소를 가진 취약 PC들이 봇에 감염된다. 여기서 네트워크 주소는 〔그림 6〕과 같이 IP 주소를 구성하는 요소로, 장치가 연결된 네트워크를 지정하는 것을 말한다.

Network Address			Host Address		
172	.	16	122	.	104

〔그림 6〕 IP 주소 구성의 예

본 논문에서는 Storm 봇넷의 통신을 파악하기 위해 허니팟이 같은 봇넷에 포함된 다른 감염된 PC들에게 자신을 알리고, 통신 대상을 찾는 과정에서 실제로 네트워크 주소가 얼마나 많은 IP 주소들과 연관되는지 알아보기 위한 분석을 수행하였다. 또한 분석에 앞서 IP 주소의 상위 16 비트를 고정하고 이에 속하는 IP 주소들을 분류하도록 기준을 정하여 이를 '16-Bit Prefix'로 지칭하였다.

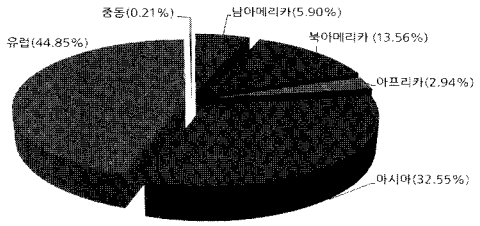
분석 과정은 허니팟에 Peacomm 봇을 감염시킨 직후 10시간 동안의 패킷에 저장된 IP 주소를 16-Bit Prefix 별로 분류하였으며, 〔표 3〕은 이러한 분류 결과를 나타낸 것으로 총 4,385개의 16-Bit Prefix로 분류되었다. 전체의 60.86%에 해당하는 2,669개의 16-Bit Prefix는 단 1개의 IP 주소와 연관되었고, 10개 이상의 IP 주소와 연관된 16-Bit Prefix는 90개(2.05%)였으며, 최대 94개의 연관 IP 주소를 포함하는 것도 있었다. 비록 2.05%의 낮은 비율이지만 90개의 16-Bit Prefix와 연관된 IP 주소는 1,994개로, 이는 전체 9,532개 IP 주소의 20.91%를 차지하는 비율이다. 이렇게 동일한 16-Bit Prefix를 가지는 IP 주소에 해당하는 PC들은 실제 지리적으로도 가까운 거리에 위치해 있을 수 있다.

〔표 3〕 16-Bit Prefix와 연관 IP 주소 비율 (단위: 개)

16-Bit Prefix	전체	연관 IP 주소		
		1개	2~9개	10개 이상
	4,385	2,669 (60.86%)	1,626 (37.09%)	90 (2.05%)

또한 허니팟이 통신을 하는 피어들의 IP 주소는 whois IP 검색[19,20] 및 국가별 IP 대역 도구[21]를 통해 〔그림 7〕과 같이 세계 전역에 걸쳐서 분포되

어 있는 것을 확인하였으며, 그중에서도 유럽이 가장 큰 비율을 보였다. [그림 8]은 IP 주소의 상위 8 비트에 속하는 IP 주소들을 분류하여 나타낸 것으로, 이를 통해 Storm 봇넷을 형성하는 피어들은 특정 IP 주소 대역에 집중해 있다는 것을 알 수 있다.



[그림 7] 허니팟이 통신하는 피어의 분포

Peacomm 봇은 초기화 단계에서 spooldr.ini 파일을 생성하고, 초기화 단계가 완료되면 좀 더 다양한 조합의 다른 감염된 피어들과 연결하기 위하여 spooldr.ini 파일을 호스트의 리스트로써 사용한다. 만약 피어 리스트의 호스트가 응답이 없을 경우 Peacomm 봇은 매 10분마다 이러한 통신을 시작하기 위하여 재시도하고, 응답이 있을 경우에는 새로운 스팸의 템플릿 및 업데이트된 실행 파일을 다운로드하거나 지속적인 피어 리스트 교환, 명령 전달 등이 이루어진다[16,17].

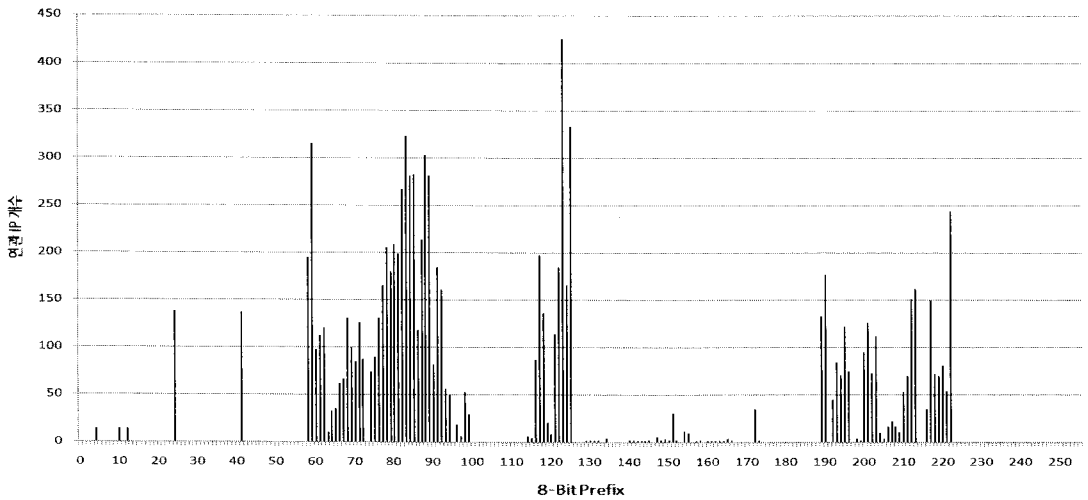
spooldr.ini 파일에는 다른 감염된 PC들의 IP 주소와 포트 번호가 16진수 형태로 저장되어 있다. 이 파일의 내용을 읽어서 10진수로 자동 변경시키도록 간단한 프로그램을 작성하였으며, 이를 통해 16진수 데이터를 10진수로 변경한 후 위에서와 마찬가지로 16-Bit Prefix별로 포함된 IP 주소의 개수를 파악하였다. [표 4]는 Peacomm 봇을 허니팟에 총 4차례 감염시킨 후 생성된 각각의 spooldr.ini 파일에 저장된 IP 주소들을 분석한 것이다.

[표 4] spooldr.ini 파일의 16-Bit Prefix 분류
(단위: 개)

	전체 IP 주소	16-Bit Prefix	연관 IP 주소	
			1개	2개 이상
첫 번째 감염	810	613	540 (88.1%)	73 (11.9%)
두 번째 감염	838	721	640 (88.8%)	81 (11.2%)
세 번째 감염	855	674	587 (87.1%)	87 (12.9%)
네 번째 감염	920	741	664 (89.6%)	77 (10.4%)

spooldr.ini 파일 내에서도 단 1개의 IP 주소와 연관된 16-Bit Prefix가 가장 많은 양을 차지하였으며, 2개 이상의 IP 주소와 연관된 16-Bit Prefix는 전체의 약 11% 내외를 차지하였다. 한편, spooldr.ini 파일 내에 존재하는 피어 리스트의 IP 주소(최대 920개)와 트래픽 덤프를 통해 추출한 IP 주소(9,532개)가 차이를 보이는 이유는 Peacomm 봇이 지속적으로 피어 리스트를 교환하고 업데이트를 함으로써 통신 범위를 확장하기 때문이다.

이와 같은 분석 결과로 미루어보면, 기존의 서브넷



[그림 8] 8-Bit Prefix일때 연관 IP 주소의 분포

스캐닝을 이용한 전파 방법을 사용하지 않거나, 사용하더라도 그 비율은 낮다고 볼 수 있다. 또한 같은 네트워크 서브넷에 속해 있어 비슷한 IP 주소를 가지는 감염된 PC들과 통신하기보다는 지속적으로 피어 리스트를 교환함으로써 더 넓은 범위의 외부 네트워크에 속해 있는 감염된 PC들과 통신을 한다. 이는 같은 네트워크 서브넷에 속한 감염된 PC들의 경우, 내부 네트워크에서 이러한 감염된 PC들 간의 통신을 차단해 버린다면 봇넷의 활동 규모 및 위력에 큰 영향을 미치기 때문에 감염된 PC들이 P2P를 이용하여 더 넓게 분산되어 있는 감염된 PC들과 통신을 함으로써 그 규모를 유지하거나 확장하고, 광범위한 네트워크를 형성하는 것이다.

4.3 UDP 패킷의 포트 번호 분석

다음으로는 허니팟에서 나가고 들어오는 UDP 패킷의 포트 번호를 분석하였다. 허니팟은 봇 감염 초기에 다른 감염된 PC들과 통신을 하기 위해 자신을 알리는 과정에서 특정 포트 번호를 임의로 선택하며, 이렇게 선택된 한 개의 특정 포트 번호를 통해 UDP 패킷을 보내고 받는다. 또한 선택되는 포트 번호는 봇 감염 시마다 달라진다. 분석 결과 허니팟이 통신하는 대상들에 대한 UDP 패킷의 포트 번호는 다양하였다. 이는 허니팟과 통신하는 각각의 감염된 PC들도 봇 감염 초기에 포트 번호가 임의로 선택되기 때문이다.

〔표 5〕는 Peacomm 봇을 4차례 감염시킨 결과 생성된 각각의 spooldr.ini 파일을 통해 허니팟이 통신하는 다른 감염된 PC들에서 임의로 선택된 포트 번호들이 얼마나 중복되는지를 나타낸 것이다. spooldr.ini 파일에 저장된 포트 번호의 전체 개수는 각각 810/838/855/920개였으며, 중복된 포트 번호를 하나로 취급할 경우 각각 722/795/766/845개로, 그 수가 크게 감소하지 않았다. 이는 중복된 포트 번호가 많지 않다는 것을 의미하며 중복 횟수도 대부분 10번 미만이었다. 또한 〔표 6〕은 각 spooldr.ini 파일에 저장된 중복성이 큰 상위 5개의 포트 번호를 나타낸 것이다.

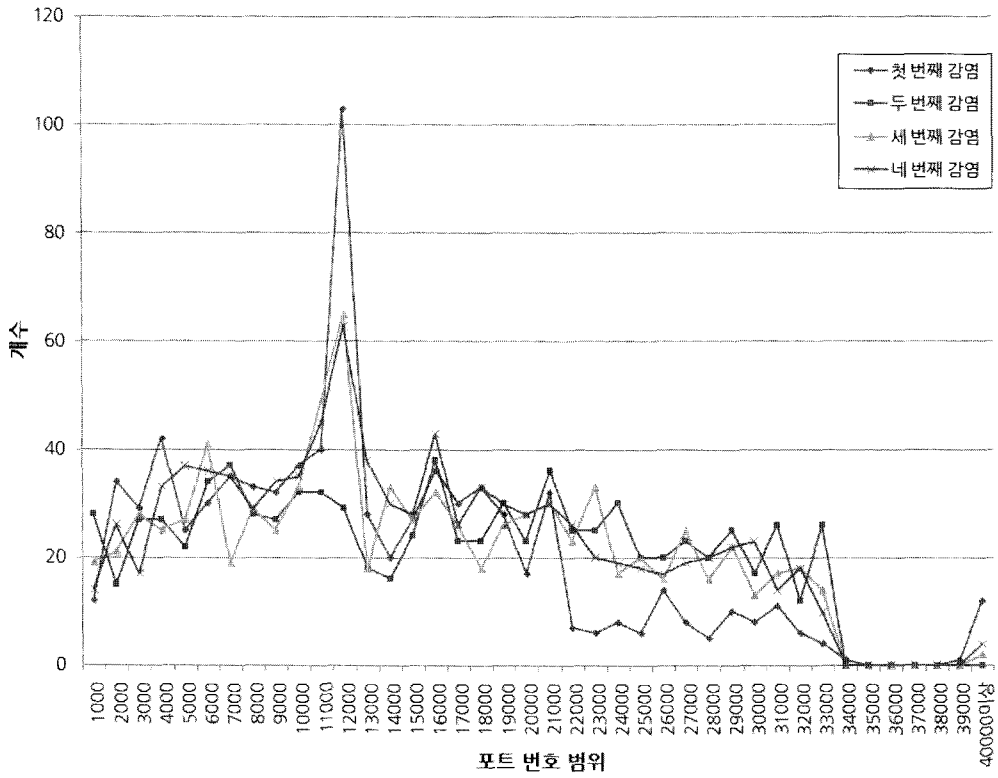
〔표 6〕을 통해 허니팟이 통신하는 다른 감염된 PC들의 포트 번호가 서로 인접해 있음을 알 수 있다. 따라서 우리는 주로 이용되는 특정 범위의 포트 번호가 존재할 것으로 예상하고, 포트 번호를 1000개씩 구분하여 해당 범위에 속하는 포트 번호의 수가 얼마나 되는지 분석하였다. 〔그림 9〕는 이러한 분석 결과를 나타낸 것이다. 허니팟이 다른 감염된 PC들과 통신을 하기 위해 이용되는 포트 번호는 주로 35000번 이하의 포트 번호에 집중해 있으며 11000번 및 12000번대의 포트 번호가 가장 많이 이용되는 것을 알 수 있다. 결론적으로 각각의 포트 번호들은 중복성이 크지 않지만 주로 이용되는 포트 번호 범위가 존재하는 것을 알 수 있다.

〔표 5〕 UDP 패킷의 포트 번호 중복성

	포트 번호 개수	중복 횟수				최소 포트 번호	최대 포트 번호
		중복 없음	2~9번	10~20번	20번 이상		
첫 번째 감염	810개	704개	14개	3개	1개	1103	65503
두 번째 감염	838개	766개	29개	0개	0개	1061	33742
세 번째 감염	855개	747개	27개	2개	0개	1069	50284
네 번째 감염	920개	818개	23개	4개	0개	1028	63976

〔표 6〕 중복성이 큰 상위 5개의 포트 번호

	1		2		3		4		5	
	포트 번호	중복 횟수	포트 번호	중복 횟수	포트 번호	중복 횟수	포트 번호	중복 횟수	포트 번호	중복 횟수
첫 번째 감염	12021	25회	12024	17회	12023	14회	12022	13회	12025	9회
두 번째 감염	11275	8회	16275	7회	10775	3회	12024	3회	17262	3회
세 번째 감염	11275	15회	12024	15회	12025	9회	16275	8회	12022	5회
네 번째 감염	12024	15회	11275	11회	12025	11회	16275	11회	11274	4회



〔그림 9〕 이용되는 포트 번호의 분포

V. 결론 및 향후 연구 과제

본 논문에서는 P2P 봇넷을 탐지 및 완화할 수 있는 대응기술 마련을 위한 기초로써 허니넷을 구축하고 각각의 허니팟에 P2P 기반의 Storm 봇 중 하나인 Peacomm 봇을 감염시켜 트래픽 분석에 대한 내용을 기술하였다. 이를 통해 Peacomm 봇이 다량의 UDP 패킷을 발생시키며, 자신이 속한 네트워크 서브넷의 감염된 PC들과 통신하기보다는 P2P를 통해 광범위한 외부 네트워크의 감염된 PC들과 통신함으로써 그 규모를 유지하거나 확장하고, 더 넓은 범위의 네트워크를 형성하는 것으로 나타났다. UDP 패킷 수는 Peacomm 봇이 활동을 함에 있어서 생성되는 패킷의 대부분을 차지하고 있었다. 이러한 UDP 패킷들은 허니팟에 감염되고 얼마 지나지 않아 급격히 증가한 후, 시간이 흐름에 따라 일정한 양을 유지하며 증가와 감소를 반복하는 패턴을 보였다. 또한 Peacomm 봇이 활동을 하면서 생성하는 UDP 패킷 및 IP 주소, spooldr.ini 파일을 분석하여 16-Bit Prefix별로 분류하고 이에 따른 관련 IP 주소와의 관계를 파악하였으며, 이용되는 포트 번호의 범위를 분석하였다. 이러한 분석을 통해 나온 결과들은 향후 트래픽 변화를 이용한 행위 기반의 프레임워크 설계 등의 대응기술을 마련하는데 활용할 수 있을 것으로 기대된다. 특히 대량의 UDP 패킷 발생에 의한 트래픽 증가 특성을 이용한 탐지 프레임워크나 알고리즘 개발에 활용할 수 있을 것으로 예상된다.

향후 연구 과제로는 봇넷의 탐지 및 완화를 위한 대응기술로써 트래픽을 모니터링하여 정상적으로 발생하는 트래픽과 다르게 UDP 패킷의 양이 순간적으로 증가할 경우 그 증가량을 임계치(threshold)와 비교하는 방법, 일반적인 P2P 파일 공유 프로그램이 특정 포트를 이용하여 데이터를 주고받는 방식과 다르게 다양한 IP 및 포트 번호에 대하여 다량의 UDP 패킷이 발생하는지와 각 통신 대상과 매우 짧은 시간동안 통신을 하는지 판단하는 방법, 일정한 시간 간격을 두고 트래픽량이 증가와 감소를 빠르게 반복하는지 파악하는 방법 등을 이용하여 프레임워크의 설계에 대한 연

구를 진행할 예정이다. 또한 Peacomm 외에 다른 P2P 봇넷의 트래픽을 분석함으로써 이들 사이에 공통된 특징을 도출하고 엔트로피를 이용한 행위 기반의 봇넷 탐지 프레임워크 개발을 통해 수시로 변화하고 등장하는 봇넷에 빠르게 대처하기 위한 방법의 연구를 진행할 계획이다.

참 고 문 헌

[1] F. Freiling, T. Holz, and G. Wicherski, "Botnet Tracking-Exploring a Root-Cause Methodology," ESORICS 2005, LNCS 3679, pp. 319-335, 2005.

[2] D. Barroso, "Botnets-The Silent Threat," ENISA Position Paper, no. 3, pp. 1-9, Nov. 2007.

[3] 박찬호, 강권학, 권영찬, 장희진, 김철호, "메모리 감시를 이용한 허니팟 기반의 봇넷 역추적," 한국정보과학회 한국컴퓨터종합학술대회논문집, pp. 25-28, 2007년 6월.

[4] 전용희, 오진태, "봇넷 분류법 및 진화된 봇넷 구조," 정보보호학회지, 18(4), pp. 76-86, 2008년 8월.

[5] 전용희, "봇넷 기술 개요 및 분석," 정보보호학회지, 18(3), pp. 101-108, 2008년 6월.

[6] R. Puri, "Bots & Botnet: An Overview," GSEC Practical Assignment Version 1.4b, SANS Institute, Aug. 2003.

[7] J. Canavan, "The evolution of malicious IRC bots," Proceedings of Virus Bulletin Conference 2005, pp. 104-114, Oct. 2005.

[8] J.B. Grizzard, V. Sharma, C. Nunnery, B.B. Kang, and D. Dagon, "Peer-to-peer botnets: Overview and case study," Proceedings of 1st Workshop on Hot Topics in Understanding Botnets (HotBots'07), Apr. 2007.

[9] 이한우, 최현상, 이희조, "DNS 기반의 봇넷 탐지 시스템," 한국정보처리학회 추계학술발표대회논문집, pp. 1379-1382, 2006년 11월.

[10] T. Holz, M. Steiner, F. Dahl, E. Biersack, and F. Freiling, "Measurements and Mitigation of Peer-to-Peer Based Botnets: A Case Study on Storm Worm," Proceedings of the 1st USENIX Workshop on Large-scale Exploits and Emergent Threats (LEET'08), pp. 1-9, Apr. 2008.

[11] A. Karasaridis, B. Rexroad, and D. Hoeflin, "Wide-Scale Botnet Detection and Caracterization," Proceedings of 1st Workshop on Hot Topics in Understanding Botnets (HotBots'07), Apr. 2007.

[12] P. Bacher, T. Holz, M. Kotter, and G. Wicherski, "Know Your Enemy: Tracking Botnets," The Honeynet Project & Research Alliance, http://www.honeynet.org/, Mar. 2005.

[13] 최효식, "꿀단지 네트워크, 허니넷," 마이크로소프트웨어, pp. 209-215, 2005년 6월.

[14] A. Ramachandran and N. Feamster, "Understanding the Network-Level Behavior of Spammers," Proceedings of 2006 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications, pp. 291-302, Aug. 2006.

[15] S.K. Noh, J.H. Oh, J.S. Lee, B.N. Noh, and H.C. Jeong, "Detecting P2P Botnets using a Multi-Phased Flow Model," Proceedings of 3rd International Conference on Digital Society 2009 (ICDS'09), pp. 247-253, Feb. 2009.

[16] P. Porras, H. Saidi, and V. Yegneswaran, "A Multi-perspective Analysis of the Storm (Peacomm) Worm," Computer Science Laboratory, SRI International, Oct. 2007.

[17] S. Stover, D. Dittrich, J. Hernandez, and S. Dietrich, "Analysis of the Storm and Nugache Trojans: P2P is here," In USENIX :LOGIN, vol. 32, no. 6, pp. 18-27, Dec. 2007.

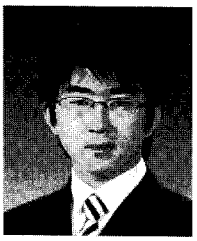
[18] http://www.wireshark.org/docs/man-pages/tshark.html

[19] http://whois.domaintools.com/

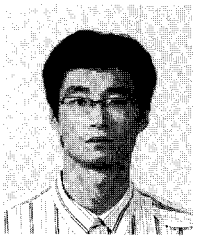
[20] http://www.ipligence.com/

[21] http://www.maxmind.com/app/ip-location

〈著者紹介〉



한 경 수 (Kyoung Soo Han) 학생회원
2008년 2월: 상지대학교 컴퓨터공학과 학사
2008년 3월 ~ 현재: 한양대학교 전자컴퓨터통신공학과 석사과정
〈관심분야〉 정보보호, 봇넷 탐지, 네트워크 보안



임 광 혁 (Kwang Hyuk Lim) 학생회원
2008년 2월: 세종대학교 컴퓨터공학과 학사
2008년 9월 ~ 현재: 한양대학교 전자컴퓨터통신공학과 석사과정
〈관심분야〉 정보보호, 네트워크 보안

임 을 규 (Eul Gyu Im) 종신회원
2002년 5월: University of Southern California 컴퓨터과학 박사
2000년 ~ 2002년: WiseNut Inc. Sr. SW Engineer
2002년 ~ 2005년: 국가보안기술연구소 선임연구원
2005년 ~ 현재: 한양대학교 공과대학 컴퓨터공학부 조교수
〈관심분야〉 유무선 네트워크 보안, 정보보안