네트워크 기반의 경량 봇넷 탐지 시스템에 관한 연구

강동완, 임채태, 정현철
한국인터넷진흥원
e-mail:lupin428@kisa.or.kr

A Study on Network Based Light-weight Botnet Detection System

Dong-Wan Kang, Chae-Tae Im, Hyun-Chul Jung
Convergence Security R&D Team, Korea Internet&Security Agency

요 약

최근 봇넷은 금전적 이득을 원하는 범죄 집단에 의해 사이버 공격의 수단으로서 크게 확산되고 있다. 봇넷의 탐지는 이전부터 꾸준히 연구되었지만, 구성과 기능이 점차 진화되고 있는 봇넷을 탐지하기에는 큰 어려움이 따르고 있다. 봇넷의 탐지는 호스트 기반의 정적인 악성 코드 분석이나 네트워크 트래픽 분석등 어떠한 특정 시스템에 의존해서는 효율적인 탐지를 기대하기 어렵기 때문에 다양한 정보를 종합하여 탐지하여야 한다. 본 연구에서는 기존에 알려진 봇넷 정보와 악성 봇 바이너리 분석을 통해 알려진 정보와 네트워크 기반의 탐지 정보를 분석하여 전체적인 봇넷의 구성을 탐지할 수 있는 네트워크 기반의 경량 봇넷 탐지 시스템을 제안한다. 제안된 탐지 시스템은 대규모의 네트워크 환경에서도 단편적으로 알려진 봇넷의 부분 정보를 기반으로 전체적인 봇넷의 구성을 탐지할 수 있다.

1. 서론

컴퓨터가 등장한 이래로, 악성코드는 지속적으로 사용자를 위협해 왔다. 과거의 악성코드는 대부분 컴퓨터를 잘 알고있는 해커가 자기 실력을 과시하는 행위로 그쳤지만 정보통신망이 발달하고 인터넷이 등장함에 따라 네트워크를 활용한 악성코드는 그 피해가 단순히 개인의 자산을 파괴하는데 그치지 않고 있다. 조직적인 범죄 집단과 연계되어 금전적인 이득을 위해 개인정보를 악용하거나, 기업과 국가 기간망에 대한 공격을 통해 개인은 물론 기업과 사회 기간망에 큰 위협이 되었다.

현재의 사이버 보안의 큰 위협중 하나인 봇넷은 다수의 감염된 시스템을 이용하여, 공격자의 명령에 따라 악성 행위를 하는 감염된 시스템의 네트워크 집단이다. 범죄를 위한 공격자는 봇넷을 제어하기 위해 C&C(Command & Control)서버를 구성하고, 하위에 좀비 PC(감염된 PC) 그룹을 구성한다. 좀비 PC들은 프로그램에 따라서 주기적 혹은 비주기적으로 C&C 서버에 접속하여 명령을 받아 수행하며 공격자가 의도한 악성행위를 수행한다.

봇넷의 주된 악성 행위는 개인정보 탈취와 스팸 메일 발송, 분산 서비스 거부 공격등이 대표적이며, 특히 스팸의 경우에는 전체 스팸의 90% 이상을 봇넷에서 발송하고 있다. 스팸은 단순히 광고를 목적으로 하는 것뿐만 아니라, 새로운 감염을 위해 악성코드를 배포하는 수단으로써 사용되어 스팸으로 인한 피해는 점점 악순환이 되고 있다.

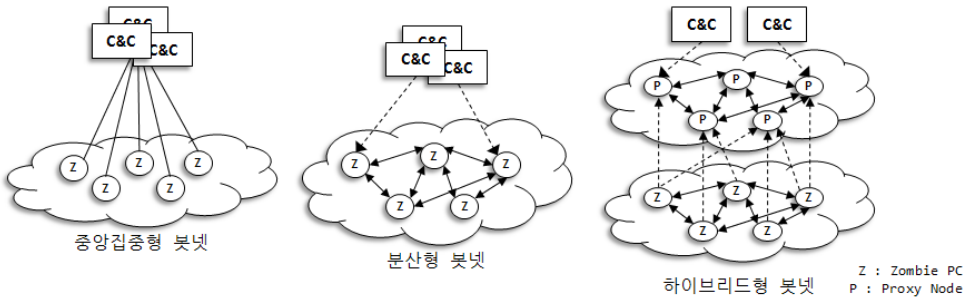
따라서 봇넷을 탐지하는 것은 현재의 사이버 보안을 위해 매우 중요한 과제이며, 이를 위한 지속적인 연구가 필요하다.

본 연구에서는 봇넷의 기본 개념과 기존 연구에 대해서 논의하고, 실제 현업에서 요구되는 봇넷 탐지 시스템의 요구사항을 분석한다. 이후 네트워크 기반의 경량 탐지 시스템의 기본 개념과 구성에 대해 제안하고 끝으로 제안된 시스템의 기대 효과와 향후 연구 방향을 제시한다.

2. 봇넷의 개요

봇넷은 공격자의 명령을 받아 수행하는 감염된 시스템의 집합으로써 일련의 네트워크로 볼 수 있다. 이러한 네트워크 구성에 따라서 정해진 통신 프로토콜을 통해 명령을 주고 받는 행위를 한다. 따라서 봇넷은 네트워크를 구성하는 토폴로지의 형태와, 사용하는 프로토콜에 따라 분류 할 수 있다[2]. 네트워크 구성에 따라 중앙집중형, 분산형, 복합형으로 분류할 수 있으며, 프로토콜에 따라 IRC, HTTP, CS, P2P, Hybrid로 분류될 수 있다.

네트워크 구성에 따른 분류는, 좀비(감염된 PC)들과 C&C 서버의 구성에 따른 분류로써, 초기 IRC 프로토콜을 사용한 중앙집중형의 Sdbot, Agobot 부터, P2P 프로토콜을 사용한 storm, nugache등과 같은 분산형, 이후 두 가지 형태를 함께 가지고 있는 waledac같은 복합형으로 볼 수 있다. 프로토콜은 봇넷이 사용하는 통신 프로토콜에 따른 분류인데, 초기 IRC 프로토콜은 현재에는 많이 사용되지 않고 방화벽에 친화적인 HTTP와 분산형 네트워크를 위한 P2P 프로토콜(eDonkey, WASTE등)이 주로 사용되고 있다[7]. 이외에 표준 프로토콜을 사용하지 않고 독자



(그림 1) 네트워크 구성에 따른 봇넷의 분류

적인 통신 프로토콜을 사용하는 Netbot과 같은 CS 형태의 봇넷도 존재한다.

이러한 분류는 기존에 많은 연구자들에 의해 분류되어 왔다[2]. 하지만 봇넷이 발전함에 따라서 그 구분이 불명확해지고, 봇넷들의 특징을 한 가지로 구분하는 것이 어렵기 때문에 봇넷 분류에 대한 기준은 명확하지 않은 상황이다.

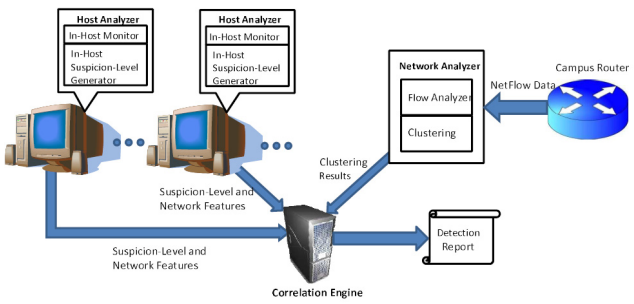
3. 기존 연구 동향

봇넷 탐지에 대한 연구는 크게 탐지하는 방법에 따라 시그니쳐 기반 탐지와 행위 기반 탐지로 분류될 수 있고, 동작하는 위치에 따라 호스트 기반과 네트워크 기반으로 분류될 수 있다.

시그니쳐 기반 탐지는 V3, 바이로봇, Norton Antivirus 등, 전통적인 안티바이러스 업체의 상용 솔루션에 사용되는 악성 코드의 특정 시그니쳐를 사용한 호스트 기반 탐지 방법과 IDS(Intrusion detection system)와 같이 특정 통신 패턴을 이용하여 탐지하는 네트워크 기반 탐지 방법을 들 수 있다.

행위 기반 탐지는 호스트 혹은 네트워크에서의 이상행위를 탐지한 방법으로써, Norman Sandbox, CWSandbox 등이 악성 코드의 행위를 모니터링하여 이상 행위를 탐지하며, BotHunter[4]와 BotMinier[5], BotSniffer[3], 그리고 DNS 트래픽 분석을 통한 탐지 연구[6], 봇넷 트래픽 특성을 모델링하여 탐지하는 연구[1]등이 네트워크 기반의 이상 행위를 기반으로 봇넷을 탐지하는 연구로 진행되었다.

[8]는 (그림2)와 같이 기존의 호스트와 네트워크 기반의 탐지 정보를 종합해서 봇넷을 탐지하는 시스템을 제안하였는데, 호스트 기반의 분석 정보와 네트워크에서의 흐름 정보를 클러스터링하고 이상 그룹 정보를 분석하여 봇넷



(그림 2) [8]에서 제안한 호스트/네트워크 종합 분석 개념

을 탐지하는 개념이다.

이러한 기존 연구들은 봇넷을 부분적으로 탐지하는데 효과가 있지만, 전체적인 봇넷의 구성을 탐지할 수 있는데는 한계가 있다.

4. 봇넷 탐지 시스템의 주요 요구사항

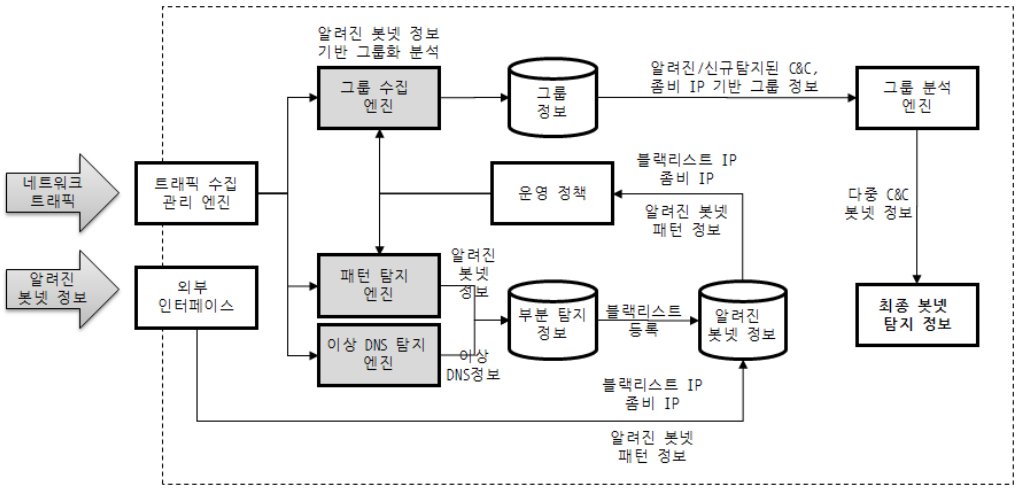
봇넷을 탐지하는 목적은 봇넷의 C&C를 탐지하여 대응하는 것에 목적이 있으며, 나아가 이러한 봇넷의 규모와 구성에 대한 분석을 통해 향후 봇넷에 대한 위협을 능동적으로 대응하는 것에 있다. 최근의 구조적, 기능적인 진화를 하고 있는 봇넷을 탐지하기 위해서는 기존의 호스트 기반과 네트워크 기반, 시그니쳐 기반과 행위 기반의 탐지 기능을 함께 고려하여 상호 보완적인 시스템이 요구된다. 따라서 네트워크 기반의 탐지 시스템은 호스트 기반의 분석 정보를 활용하여 네트워크에서의 관련 정보를 수집하고, 이를 토대로 봇넷을 탐지할 수 있는 경량화된 시스템이 되어야 한다.

궁극적으로 탐지하고자 하는 봇넷 정보는 단순한 C&C 정보뿐 아니라, 봇넷의 구성에 따른 적절한 대응을 할 수 있도록 봇넷을 구성하는 C&C와 봇들의 정보를 포함한 전체적인 구성 정보를 탐지해야 한다. 이를 위해 네트워크 기반 탐지 시스템에서의 기능과 구성적인 요구사항으로는 아래와 같이 제시될 수 있다.

- 입력 요구사항
  - 네트워크 트래픽
  - 알려진 봇넷 C&C IP, url 정보
  - 알려진 봇넷 통신 패턴 정보
  - 알려진 감염 시스템 IP 정보

- 기능 요구사항
  - 알려진 봇넷 C&C IP를 기반으로 전체적인 봇넷을 탐지하는 기능
  - 알려진 url을 기반으로 봇넷 구성을 분석하는 기능

- 출력 요구사항
  - 다중 C&C를 가지는 전체적인 봇넷 구성 정보 (C&C IP 리스트, url 정보, 감염 PC IP 리스트)



(그림 3) 경량 탐지 시스템의 기능 구성도

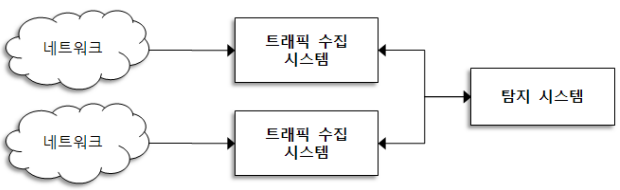
5. 네트워크 기반 경량 봇넷 탐지 시스템

경량 봇넷 탐지 시스템의 목적은 이미 알려진 봇넷의 부분적인 정보나 악성 봇의 정적/동적인 바이너리 분석을 통해 얻은 외부 통신 정보를 바탕으로 봇넷의 전체적인 구성을 탐지하는 것에 있다.

탐지 프로세스를 경량화 하기 위해 알려진 봇넷 정보를 대상으로 한 네트워크 트래픽을 필터링 하고 필터링 된 트래픽 정보에서 그룹 행위를 분석하여 전체적인 봇넷 구성을 탐지하는 과정이 필요하다. 이러한 과정은 네트워크에 따른 트래픽을 수집하는 시스템과 수집된 트래픽을 바탕으로 실질적인 탐지를 하는 시스템으로 분리되어 동작할 수 있다. (그림4)는 봇넷 탐지를 위한 시스템의 전체적인 구성도이다.

네트워크 기반 봇넷 탐지 시스템은 해당 네트워크의 규모를 감안하여 트래픽 수집 시스템과 탐지 시스템을 함께 둘 수도 있고, 따로 둘 수도 있다. 제안하는 시스템의 기능적인 요소들은 전체적인 구성도에 국한되지 않고, 네트워크의 상황에 따라 유연하게 적용하도록 해야 한다.

봇넷 탐지를 위한 기능적인 측면에서 네트워크 트래픽을 그룹 분석과 패턴 탐지, 이상 DNS 탐지를 위한 엔진에 분산하여 입력으로 주고, 알려진 봇넷 정보를 기반으로, 봇넷 정보를 분석한다. 패턴 기반과 이상 DNS 탐지에서는 운영 정책에 따라 봇넷 통신 패턴에 매칭 되는 C&C 서버, 좀비 IP를 탐지하고, 트래픽 내의 악성 url을 탐지한다. 탐지된 정보는 운영 정책에 의해 그룹 수집과 패턴, 이상 DNS 탐지 엔진에 피드백 된다. (그림3)은 경량 탐지 시스템의 전반적인 기능 구성도이다.



(그림 4) 봇넷 탐지 시스템 구성도

5.1 트래픽 수집 엔진

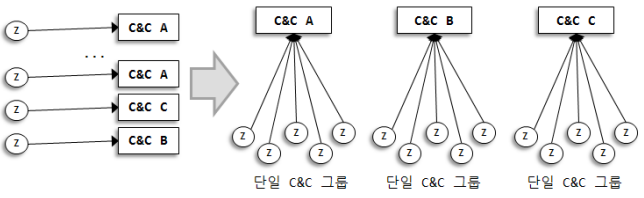
트래픽 수집 엔진은 네트워크 트래픽을 수집하는 엔진으로, 수집된 트래픽을 운영 정책에 따라서 DNS 트래픽 및 수집 대상 트래픽으로 분류한다. 분류된 트래픽은 이상 DNS 탐지 엔진이나 그룹 수집 엔진으로 전송된다. 패턴 탐지 엔진은 알려진 트래픽 패턴을 탐지해야 하므로, 수집 트래픽을 필터링 하지 않고 Low-packet을 그대로 전송해 주어야 한다.

5.2 외부 인터페이스

알려진 봇넷에 대한 부분 정보를 입력으로 줄 수 있는 인터페이스로써, 봇넷의 C&C정보, 악성 URL등을 입력으로 운영 정책에 반영될 수 있는 기능이다. 외부 인터페이스는 관리자에 의해 수동으로 입력될 수도 있으며, 악성 봇 바이너리를 동적으로 분석하는 시스템과 실시간으로 연결되어 자동화 될 수 있다.

5.3. 그룹 수집 엔진

필터링된 트래픽은 C&C IP에 접속하는 트래픽이나, Zombie PC가 활동하는 트래픽이다. 이를 통해 각각의 통신을 세션별로 분석하고, 이에 따른 통신 구성을 그룹화 하는 기능이다. 통신 그룹을 그룹화 하게 되면 특정 시간 구간에 대한 네트워크에서의 그룹행위를 볼 수 있으며, 이를 통해 봇넷의 개별 C&C가 활동하는 구성을 판단할 수 있다.



(그림 5) 네트워크에서의 그룹 행위

5.4 패턴 탐지 엔진

최근 봇넷은 다양한 변종이 등장하고 있다. 변종은 일련의 봇넷 제작 툴에 의해서 만들어 지는데, C&C 서버만 다를뿐 비슷한 행동을 보이는 경우가 대부분이다. 따라서 이러한 봇넷의 통신 패턴을 기반으로 하여 유사한 변종의 봇넷을 탐지하는 것은 필수적이다.

봇넷의 통신 패턴은 실질적으로 봇넷에 따른 악성 봇의 행위를 분석해야 한다. 패턴 양식은 L7 레이어에서 세션을 이루고 난 후의 특정 통신에 대한 컨텐츠나 필드값이 될 수 있다.

5.5 이상 DNS 탐지 엔진

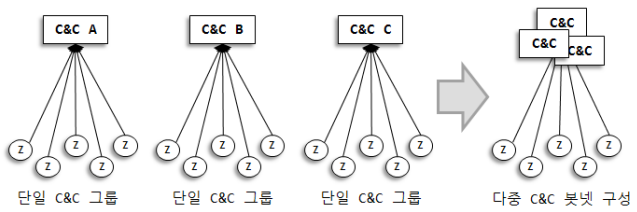
DNS는 봇넷이 자기의 C&C 서버를 URL을 통해 접속하기 위한 통신이다. 몇몇 봇넷들은 URL없이 하드코딩된 IP에 접속하기도 하지만, url을 통해서 접속하기도 한다. url에 대한 쿼리 결과는 보통은 고정적이지만, 봇넷은 url에 대한 ip 주소를 짧은 시간 주기를 가지며 계속 변경하는 fastflux 기능을 가질 수 있다. 이러한 경우에는 ip기반에서의 그룹행위는 알기 어려우며, url에 대한 쿼리 요청 그룹으로써만 그룹 행위를 판단할 수 있다. DNS 쿼리는 정상적인 네트워크 행위에서도 관찰되는 것으로써 화이트리스트 처리가 중요하며, 잘 알려지지 않은 url에 대한 비정상적인 요청이 의심될 경우, fastflux를 Active Monitoring으로 판단하여 봇넷 url을 탐지할 수 있다.

5.6 그룹 분석 엔진

수집된 그룹에 대해서 각 그룹간의 유사도 분석을 통해 전체적인 봇넷 그룹을 분석하는 엔진이다. 그룹간의 유사도는 다변량 데이터를 이용하여 분석을 할 수 있지만, 탐지 시스템의 경량화를 위해 구성원들간의 일치정도를 위해 jaccard coefficient를 사용한다.

$$J(A,B) = \frac{A \cap B}{A \cup B}$$

다중 C&C 봇넷 구성을 통해 현 시점에서의 봇넷에 대한 구성을 탐지할 수 있다. 이후, URL에 따른 군집과의 유사도 분석을 통해 어떠한 봇넷이 URL을 사용하고, fastflux를 사용하는지에 대한 정보를 함께 통합하여 최종 봇넷 탐지 정보를 출력한다.



(그림 6) 다중 C&C 봇넷 구성 탐지

6. 결론 및 향후 연구 방향

제안된 경량 봇넷 탐지 시스템의 목적은 이미 알려진 봇넷의 부분적인 정보와 악성 봇의 바이너리 분석을 통해 얻은 외부 통신 정보, 네트워크 정보를 종합 분석 하여 봇넷의 전체적인 구성을 탐지하는 것에 있다. 제안 시스템은 악성 봇의 정적/동적 분석 시스템으로부터 C&C 정보 및 악성 url, 네트워크 행위에 대한 정보를 종합하여 봇넷으로 동작하기 위한 일련의 그룹행위를 네트워크에서 탐지하는 시스템이다. 제안 시스템은 구조적인 경량화를 통해 대용량 네트워크 구간에서의 봇넷 탐지를 기대할 수 있다.

Acknowledgement

본 연구는 지식경제부 및 한국산업기술평가관리원의 IT산업원천기술개발사업의 일환으로 수행하였음. [KI001863, 신종 봇넷 능동형 탐지 및 대응 기술]

참고문헌

[1] A. Karasaridis, B. Rexroad, and D. Hoeflin, "Wide-Scale Botnet Detection and Characterization", 1st Workshop on Hot Topics in Understanding Botnets, April 2007.

[2] D. Dagon, G. Gu, C. Lee, and W. Lee, "A taxonomy of botnet structures", The 23 Annual Computer Security Applications Conference (ACSAC'07), Dec 2007.

[3] G. Gu, J. Zhang, and W. Lee, "BotSniffer: Detecting botnet command and control channels in network traffic", The 15th Annual Network and Distributed System Security Symposium (NDSS'08), February 2008.

[4] G. Gu, P. Porras, V. Yegneswaran, M. Fong, W. Lee., "BotHunter: Detecting Malware Infection Through IDS-Driven Dialog Correlation", In USENIX Security Symposium, Aug 2007.

[5] G. Gu, R. Perdisci, J. Zhang, and W. Lee., "Bot-Miner: Clustering Analysis of Network Traffic for Protocol and Structure Independent Botnet Detection", In USENIX Security Symposium, July 2008.

[6] Hyunsang Choi, Hanwoo Lee, Heejo Lee, Hyogon Kim, "Botnet Detection by Monitoring Group Activities in DNS Traffic", IEEE Computer and Information Technology (CIT), Oct 2007.

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

[8] Yuanyuan Zeng, Xin Hu, K.G. Shin, "Detection of Botnets Using Combined Host- and Network-Level Information", IEEE Dependable Systems and Networks (DSN), June 2010.