봇넷의 탐지 및 관제 시스템 설계

권종훈*, 임채태**, 최현상*, 정현철**, 이희조*
*고려대학교 컴퓨터• 전파통신공학과
**한국 정보보호진흥원
e-mail : * {signalnine, realchs, heejo}@korea.ac.kr
**{chtim, hcjung}@kisa.or.kr

Cooperative Architecture for Botnet Detection and Management

Jonghoon Kwon*, Chaetae Im**, Hyunsang Choi*, Hyuncheol Jeong**, Heejo Lee*
*Div. of Computer & Communication Engineering, Korea University
** Korea Information Security Agency, Seoul, Korea

요 약

최근의 사이버 공격은 경쟁사에 대한 DDoS 공격과 기밀정보 유출, 일반 사용자들의 금융정보 유출, 광고성 스팸메일의 대량 발송 등 불법 행위를 대행해주고 경제적 이득을 취하려는 의도로 바뀌어가고 있다. 그 중심에 있는 봇넷은 봇이라 불리는 감염된 호스트들의 네트워크 집단으로서 일련의 거의 모든 사이버 공격에 이용되고 있다. 이러한 봇넷은 수 많은 변종과 다양한 탐지 회피 기술로 그 세력을 확장해 가고 있지만 마땅한 총괄적 대책은 미흡한 것이 현실이다. 이 논문에서는 날이 갈수록 위협을 더해가는 봇넷을 빠르게 탐지하고 대응하기 위해 ISP 사업자들 간, 혹은 국가 간에 걸친 사회 전반적인 협력을 통한 봇넷 탐지 및 관리 시스템 구조를 제안한다.

1. 서론

봇은 로봇(Robot)의 줄임말로서 악의적 의도를 가진 소프트웨어에 감염된 PC 를 의미한다. 이렇게 감염된 수 천, 수 만대의 봇이 네트워크로 연결되어 봇넷(Botnet)을 형성하게 된다. 이렇게 형성된 봇넷은 봇 마스터(Bot master)에 의해 원격 조종되어 DDoS 공격, 개인정보 수집, 피싱, 악성코드 배포, 스팸메일 발송 등 다양한 악성행위에 이용되고 있다.

시만텍의 발표에 따르면 2006 년 하반기에만 600 만대의 새로운 봇이 생성된 것으로 나타났으며 이는 2006 년 상반기에 비해 29%나 증가한 수치이다. 또한 2007 년에는 하루 평균 52,771 대의 PC 가 새로이 봇에 감염되고 있다고 전했다. 또한 Arbor Networks 의 2007 년 조사 결과에 따르면 봇넷이 DDoS 공격을 제치고 사이버 상의 가장 위협적인 요소로 새로이 등극하였다. 또한 최근 발생하는 DDoS 공격의 거의 대부분이 봇넷에 의해 이루어지는 것을 감안하면 봇넷이 사이버 상에 현존하는 가장 위협적인 존재임이 분명하다.

국내 상황도 크게 다르지 않다. 한국정보보호진흥원(KISA)의 발표자료에 따르면 국내 봇 감염률은 2007 년 한해 평균 전 세계 감염률의 11.3%에 달하는 것으로 보고 되었다. 하지만 이 수치 또한 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 더욱 교묘해지고 있다. 또한 봇넷은 그 소스가 공개되어 있어 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 이루어 지고 있지만 봇넷이 전 세계에 넓게 형성되어 있다는 점에서 그 탐지 및 관리에 어려움이 있는 것이 현실이다.

본 논문에서는 이러한 봇넷에 대처하기 위해 ISP 사업자들 간, 혹은 국가 간에 걸친 사회 전반적인 협력을 통한 봇넷 탐지 및 관리 시스템 구조를 제안하려고 한다.

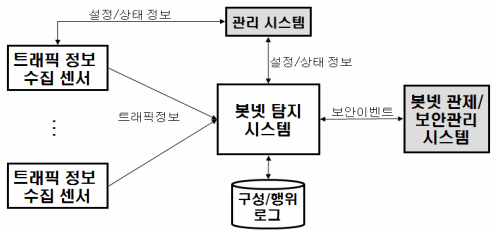
본 논문의 2 장에서는 중앙집중형 봇넷의 탐지 및 관제 시스템의 전반적인 구조를 설명하고 3, 4, 5 장에서는 시스템의 세부적인 구조인 트래픽 수집 센서, 탐지 시스템, 관제 및 보안관리에 관한 구조를 설명하고자 한다. 마지막으로 6 장에서는 결론과 함께 향후 전망에 대하여 언급하기로 한다.

• 본 연구는 지식경제부 및 정보통신연구진흥원의 IT 핵심기술개발사업의 일환으로 수행하였음. [2008-S-026-01, 신종 봇넷 능동형 탐지 및 대응 기술]
• 본 연구는 지식경제부 및 정보통신연구진흥원의 대학 IT 연구센터 지원사업의 연구결과로 수행되었음" (IITA-2008-C1090-0801-0016)

2. 시스템 구조

봇넷은 크게 봇 마스터와 봇 C&C 서버(Command and control server), 봇 호스트로 구분 지을 수 있다. 봇 마스터는 C&C 서버를 통해 봇 호스트들에게 명령을 전달하기 때문에 봇 마스터와 봇 호스트의 중계지 역할을 하는 C&C 서버를 찾아 무력화 시키는 것이 현실적으로 가장 효과적인 방안이다. 특히 C&C 서버는 대규모의 봇들과 연결되어 1:N 통신을 하며 봇들에게 명령을 전달하는데 이러한 특성을 중앙집중형 봇넷이라고 하며 본 논문에 주요 목표이다.

중앙집중형 봇넷을 탐지 및 관리하기 위한 시스템은 크게 트래픽 수집 센서, 봇넷 탐지 시스템, 관제 및 보안관리 시스템으로 구분될 수 있으며 전체 시스템 구성도는 (그림 1)과 같다.

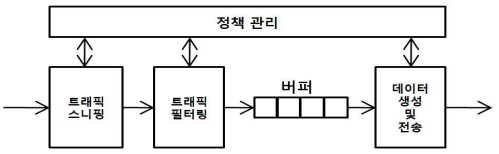


(그림 1) 시스템 구성도

트래픽 수집 센서는 봇넷 탐지에 필요한 트래픽을 수집하고 탐지 시스템에 전송하는 역할을 한다. 봇넷은 봇넷 구성의 특성 상 전 세계 네트워크에 넓게 퍼져있기 때문에 수집 센서는 가능한 다양한 네트워크에 설치되어야 한다. 봇넷 탐지 시스템은 여러 네트워크에 설치된 센서에서 보내오는 트래픽 정보를 받아 신종 봇넷을 탐지하고 공격이나 이주 등의 봇넷 특성 행위를 분석한다. 또한 이렇게 분석된 결과는 관제 및 보안관리 시스템에 의해 수집 및 기록되고 차후에 정책 결정 및 탐지된 봇넷의 관리에 사용된다.

3. 트래픽 수집 센서

트래픽 수집 센서는 탐지 시스템에서 봇넷을 탐지하고 분석하기 위해 필요한 트래픽을 실제 네트워크에서 수집하여 전송한다. 특히 봇넷의 특성상 특정 네트워크에만 집중되는 것이 아닌 무작위로 분포되어 있는 특성을 가지기 때문에 트래픽 수집 센서는 가능한 많은 네트워크에 설치되어 동작하는 것이 유리하다.



(그림 2) 트래픽 수집 센서 구성도

트래픽 수집 센서의 기능별 세부 모듈을 살펴보면 (그림 2)와 같이 트래픽 스니핑 모듈, 트래픽 필터링 모듈, 데이터 생성 및 전송 모듈로 나눌 수 있다.

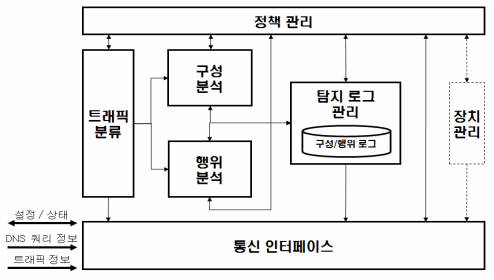
먼저 트래픽 스니핑 모듈은 패킷 캡쳐 도구를 사용하여 모니터링하는 네트워크의 내부 DNS 서버 앞단이나 Gateway 에서 미러링하여 트래픽 데이터를 수집한다. 트래픽 스니핑의 과정에서 패킷 이외에도 Netflow 와 같이 플로우 정보가 수집되어 봇넷의 탐지에 이용될 수 있으며 이를 위해서는 플로우 정보를 이용한 탐지 알고리즘에 대한 구현이 탐지시스템에서 이루어져야 한다.

이렇게 수집된 트래픽에서 분석에 필요한 정보들만을 추출하기 위해 프로토콜, 포트, 패킷 사이즈, 블랙 리스트, 화이트 리스트 등 룰 매칭에 기준하여 캡쳐된 패킷을 필터링하고 버퍼에 쌓는다.

필터링된 트래픽들을 탐지 시스템에 전송하기 위한 데이터 생성 및 전송 과정을 거치게 된다. 다수의 수집 센서에서 수집된 데이터를 탐지 시스템에 무작위로 전송하게 되면 탐지 시스템에 부하가 걸리기 때문에 데이터 전송에도 룰이 필요하다. 전송 타이머는 이런 다수의 수집 센서와 탐지 시스템간의 데이터 전송 스케쥴링을 위해 전송 여부를 결정한다. 이러한 과정은 정책관리 모듈에서 수행하게 된다. 전송 타이머에 의해 데이터 전송이 결정되면 필터링된 트래픽을 버퍼로부터 읽어와 전송 헤더를 생성하고 집약된 데이터 전송 포맷을 생성하여 UDP 로 탐지 시스템에 전송하게 된다. 센서에 각 모듈에서 사용되는 룰들은 관제 및 보안관리 모듈에 의해 관리된다.

4. 탐지 시스템

트래픽 수집 센서에서 수집된 트래픽 정보들은 탐지 시스템에 의해 종합된다. 탐지 시스템은 네트워크 전반에 걸쳐 넓게 퍼져있는 센서들로부터 트래픽 정보들을 수집하고 분석하여 봇넷의 구성 및 악성 행위들을 분석하여 신속한 탐지 및 대응이 가능하도록 하는 주요 모듈이다. 탐지 시스템의 전반적인 구성도는 (그림 3)과 같다.



(그림 3) 탐지 시스템 구성도

트래픽 분류 모듈

트래픽 분류 모듈은 수집 센서들로부터 전달되는

트래픽들을 수집하여 각 모듈에서 필요로 하는 종류별로 미리 분류하는 역할을 하는 수집관리 기능을 수행한다. 수집된 트래픽들은 일차적으로 봇넷 매칭을 통해 기존에 탐지된 봇넷과 관련이 있으면 기존 봇넷 메시지 큐에, 신종 봇넷으로 의심이 되면 신종 봇넷 메시지 큐에 저장된다. 신종 봇넷 메시지 큐의 데이터는 구성 분석 모듈에서 읽어 신종 봇넷의 C&C 서버와 봇 호스트 리스트를 추출하며, 기존 봇넷 메시지 큐의 데이터는 행위 분석 모듈에서 읽어 공격 행위나 확산, 이주 행위 등의 분석에 쓰인다.

구성 분석 모듈

구성 분석 모듈은 신종 봇넷 메시지 큐에 저장된 트래픽을 읽어 봇넷의 구성 정보를 분석한다. 구성 분석 모듈에 의해 탐지되는 봇넷의 구성은 크게 C&C 서버와 봇 호스트들로 구분될 수 있다.

앞에서 언급한 바와 같이 중앙집중형 봇넷은 다수의 봇 호스트들이 소수의 C&C 서버에 의해 동작하려는 특성을 가지고 있다. C&C 서버에 접속하는 과정에서 발생하는 DNS 쿼리나 접속을 유지하기 위한 Ping/Pong 트래픽, 이주, 코드 다운로드, 공격 행위 등에서 이런 특성을 자주 확인 할 수 있으며 이러한 그룹 행위를 기반으로 봇넷 트래픽으로부터 유사도를 측정할 수 있다 [3].

봇넷 구성 분류 모듈은 분석에 필요한 트래픽 정보를 주기적으로 읽어 Domain, IP/Port, URL 별로 유사도를 측정하고 프로토콜 별 C&C 서버를 추출한다. C&C 서버를 추출하고 나면 이 정보를 바탕으로 해당 C&C 서버에 접속하려는 봇 호스트 정보 또한 추출할 수 있다. 분석된 결과는 로그에 기록하여 관제 및 보안관리 모듈에서 차후 정책 결정 및 탐지된 봇넷의 관리에 사용될 수 있도록 한다.

행위 분석 모듈

행위 분석 모듈은 기존 봇넷 메시지 큐의 트래픽을 읽어 봇넷의 행위 특성을 분석하며 분석 결과를 행위 로그로 기록하여 차후 정책에 반영시킨다. 행위 분석 모듈에서 분석되는 행위 모델은 크게 공격 행위와 확산/이주 행위로 구분될 수 있다.

공격 행위 분석은 봇넷에 의해 발생 할 수 있는 다양한 공격 행위를 탐지하기 위함으로, Egg download, DDoS, 스팸메일 발송, 개인정보 탈취 등이 공격행위에 해당된다. Egg download 는 세부 공격 수행 모듈 다운로드나 스팸메일 발송을 위한 메일주소 및 메일 컨텐츠 템플릿 다운로드, 자가 업데이트 등이 있으며 일정시간 동안 발생하는 패킷당 평균 수신 트래픽량을 비교 검사해 알 수 있다. DDoS 공격은 동일 전송 트래픽의 발생 주기를 관찰하며 스팸메일 발송은 SMTP 서버로 발송되는 메일 전송 트래픽을 관찰하거나 MX 쿼리 트래픽을 관찰하는 분석방법을 사용한다. 또한 개인정보 탈취 공격은 일정한 시간 동안 발생하는 패킷당 평균 송신 트래픽 량을 비교 검사하거나 업로드 트래픽을 감지하는 분석방법을 사용한다.

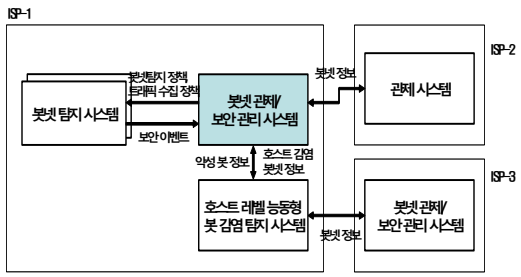
확산/이주 행위 분석은 봇넷의 운영에 필요한 확산, C&C 서버 이주, 재 접속 및 접속유지, 봇 코드 전파와 같은 행위를 분석한다. 기존에 알려진 C&C 서버로 접속하려는 신규 봇 호스트의 IP 가 발견되는 경우 확산 행위에 해당되며, 반대로 기존 봇 호스트들이 신규 IP 에 접속하려는 행위가 발견되면 C&C 서버의 이주 행위에 해당된다. 간혹 네트워크 이상이나 시스템이 다시 부팅되는 경우 봇 호스트들은 C&C 서버와의 연결을 유지하기 위해 재접속 할 수 있는데 이때 C&C 서버에 대한 DNS 쿼리를 다시 발생시키기 때문에 이를 이용해 재접속 및 접속유지 행위를 분석할 수 있다. 또한 봇 코드의 전파를 위해 무작위 IP 에 대한 취약 Port 스캐닝이 발생하는 경우 단일 봇 호스트에서 다중 목적 IP 들로 전송되는 트래픽 유무를 검사함으로써 코드 전파 행위를 분석할 수 있다.

탐지 로그 관리 모듈

구성 분석 모듈과 행위 분석 모듈에서 분석된 데이터를 구성/행위 로그에 저장하고 관제 및 보안관리 시스템에 전송하기 위해 전송 주기 설정에 따라 탐지 로그를 전송한다

5. 관제 및 보안관리

봇넷 관제 및 보안관리 시스템의 네트워크 구성은 (그림 4)과 같다. 시스템은 여러 ISP 사업자 망에 존재하여 각 시스템은 서로 정보공유를 통해 보안관리를 수행하게 된다. 전체 네트워크 구성도 상에서 각각의 시스템 요소는 다음과 같이 요약될 수 있다



(그림 4) 봇넷 탐지 및 관제 시스템 관계도

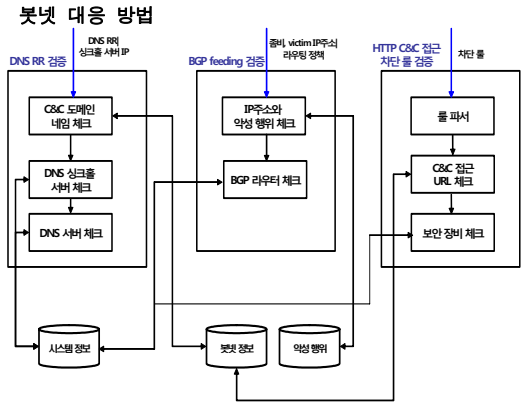
트래픽 수집 센서는 봇넷 탐지를 위한 트래픽을 수집하는 역할을 한다. 트래픽 수집 센서는 ISP 망에 n*m 개 존재할 수 있으며, 봇넷 관제 및 보안관리 시스템에서 설정한 수집 정책에 따라 DNS 트래픽, 트래픽 정보 등을 수집한다. 이렇게 수집된 트래픽 정보는 주기적으로 봇넷 탐지 시스템으로 전송된다.

봇넷 탐지 시스템은 ISP 망에 m 개 존재할 수 있으며, 수집된 트래픽 정보를 이용하여 봇넷을 탐지하고, 악성행위를 분석한다. 탐지된 봇넷 정보는 봇넷 관제 및 보안 관리시스템으로 전송한다.

봇넷 관제 및 보안관리 시스템은 일반적으로 ISP 망에 1 개 존재하며 ISP 망 내의 봇넷 정보를 시각화하여 보여주고, 대응 정책을 설정할 수 있는 기능을 제공한다.

호스트레벨 능동형 봇 감염 탐지 시스템은 독립적으로 설치된 시스템으로, 능동적으로 감염된 악성 봇을 분석하여 봇넷이 사용하는 봇 정보를 제공한다.

타 ISP 망의 관제 시스템 또는 봇넷 관제 및 보안관리 시스템은 ISP 간 또는 국가간 봇넷 대응을 위해 탐지된 봇넷 정보를 유기적으로 공유한다. 이러한 정보 공유를 통해서 잘 클러스터링된 봇넷 (전 세계 네트워크에 넓게 분포하고 있는 봇넷)의 경우에도 탐지가 가능하게 되며 빠르게 대응할 수 있게 된다.



(그림 5) 봇넷 대응 기법 및 정책 검증

탐지된 봇넷에 대응하기 위해 (그림 5)와 같이 다양한 대응 기술이 사용되는데, 그 첫 번째 방법은 블랙리스트를 공유하는 방법이다. 특정 AS(탐지 시스템이 관리하는 영역) 짧은 시간에 다수의 좀비가 새로운 C&C 에 접근하는 것이 발견될 경우 C&C 에 대한 정보를 다른 AS 의 탐지 시스템에게 공유해서 블랙리스트를 통한 대응을 수행하는 것이 가능하다. 블랙리스트가 있는 경우에는 봇넷의 C&C 서버를 비롯한 감염 호스트들에 직접적인 통제나 제한을 가하는 것이 가능하다.

두 번째 기법은 이미 널리 알려진 DNS 싱크홀 기법을 활용하는 것이다. DNS 싱크홀 기법은 주로 IRC 기반 봇넷 C&C 접근 차단을 위해 사용되는 대응 정책으로, 신규로 발견된 IRC 봇넷에 대한 접근 차단을 위해 DNS RR 을 생성하여, DNS 서버로 전송하는 기법이다. 싱크홀 리스트에 추가된 도메인들에 대해서 싱크홀 기법을 적용한 도메인 서버로 질의를 하는 호스트들은 기존의 IP 정보가 아닌 싱크홀 처리된 IP 주소를 DNS 서버로부터 전달받게 되고 이 때문에 봇넷의 C&C 서버가 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 HTTP 기반 봇넷 C&C 접근 차단을 위해 사용될 수 있는 대응 정책으로, 공개 웹 방화벽의 룰 설정을 통해 좀비가 HTTP 봇넷 C&C URL 에 접근 하는 것을 차단할 수 있다.

네 번째로 BGP feeding 기법을 이용하여 봇넷을 무력화할 수 있다. BGP feeding 기법은 주로 DDoS 등 봇넷을 이용한 공격 행위 차단을 위해 사용되는 대응 정책으로, 공격을 받는 호스트로 가는 DDoS 트래픽 등을 null routing 의 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.

6. 결론

본 논문에서는 최근 거의 모든 사이버 공격의 주요 원인이 되고 있는 봇넷에 효과적으로 대응하기 위한 탐지 및 관제 시스템의 구조를 제시하였다. 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. 또한 본 논문에서 제안한 설계 구조는 다양한 탐지 알고리즘을 수용할 수 있어 차후에 발전되는 봇넷 탐지 연구들을 적용하기 쉽다는 장점을 가지고 있다. 또한 특정 하드웨어나 시스템에 특화되어 있지 않는 유연한 시스템 구조를 갖고 있어 그 적용 범위가 매우 넓다.

최근에는 IRC 나 HTTP 를 이용한 중앙집중형 봇넷 뿐만 아니라 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 또한 차후 풀어야 할 문제가 될 것으로 예측한다.

참고문헌

[1] Dean Turner, Marc Fossl, Eric Johnson, Trevor Mack, Joseph Blackbird, Stephen Entwisle, Mo King Low, David McKinney, Candid Wueest. “Symantec Global Internet Security Threat Report”, 2008

[2] Arbor Networks. “Worldwide Infrastructure Security Report”, 2007

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

[4] 전용희, “봇넷 기술 개요 및 분석”, 정보보호학회지, 제 18 권 제 3 호, pp. 101-108, Jun 2008

[5] Hyundo Park, Peng Li, Debin Gao, Heejo Lee, Robert H. Deng, "Distinguishing between FE and DDoS Using Randomness check", Information Security Conference(ISC2008), LNCS, Vol. 5222, pp. 131-145, Sep. 15. 2008

[6] 최현상, 권종훈, 김인환, 이희조, "악성코드를 이용한 봇넷 공격", 경영과 컴퓨터, pp. 144-147, Jul. 2008

[7] 최현상, 이희조, "행위기반의 봇넷 탐지기술", 정보보호 21C, Vol. 86, pp. 80-83, Oct. 1. 2007