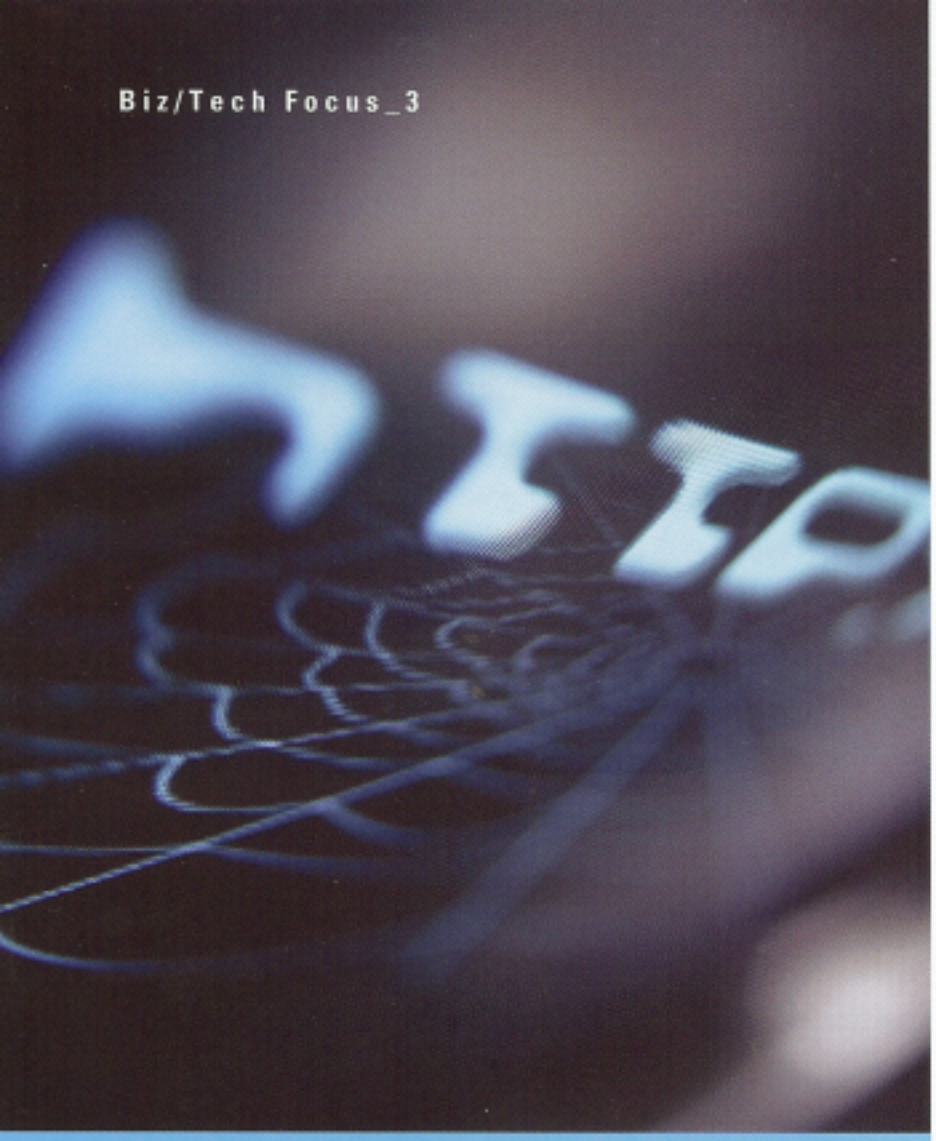


악성코드를 이용한 봇넷 공격

탐지 · 방어기술 개발되면 재빨리 우회공격 나타나

2007년2월, 미국의 2개 루트서버에 5시간 동안 서비스 장애가 발생됐다. 봇넷을 통한 DDoS 공격으로 밝혀진 이 사건에는 전 세계에 흩어져 있는 수 백 만 대의 봇에 감염된 PC가 동원됐으며, 이 중 61%가 한국에서 발송된 트래픽으로 분석됐다. 이에 미국은 "공격 진원지에 사이버 역대응 공격 또는 실제 폭탄 투하까지도 불사하겠다"고 발표했다.

최현상 · 권종훈 · 김인환 연구원(고려대학교 정보통신대학 컴퓨터보안연구실)

이희조 교수(고려대학교 정보통신대학 BK21 소프트웨어사업단)

정보통신기술은 정치 · 경제 · 사회 · 문화 등 전 분야에 없어서는 안 될 핵심요소로 자리 잡았다. 하지만 정보통신기술의 비약적인 발전에서 오는 순기능과는 반대로, 고도화되고 비약적으로 발전한 역기능, 즉 정보기술을 악용하는 해킹, 웜 · 바이러스 등 사이버상의 각종 위협 요소 또한 첨단화 · 고속화 · 광역화되고 있다. 1 · 25 인터넷 대란 때의 피해에서도 보는 바와 같이 사이버위협의 파괴력이 단순한 서비스 마비나 경제적 손실을 넘어 국가안보까지 위협하는 심각한 단계에까지 이르렀다.

과거의 사이버 공격은 단순한 호기심과 해커들의 실력 과시용 퍼포먼스로 여겨졌지만 근래에는 경쟁사에 대한 DDoS 공격과 기밀정보 유출, 일반 사용자들의 금융정보 유출, 광고성 스팸메일의 대량 발송 등 불법 행위를 대행해주고 경제적 이득을 취하려는 목적으로 바뀌고 있다. 이번 호에서는 최근 DDoS 공격 뿐 만 아니라 다양한 사이버 범죄의 중심에 있는 봇넷(Botnet)에 대해 소개하고자 한다.

가장 위협한 사이버 위협 요소, '봇넷'

2007년10월 우리나라의 게임아이템 중개 사이트들이 수 일 동안 접속장애 현상을 일으켰다. 접속장애 현상은 DDoS 공격에 의한 것이고, 이중에서도 'UDP 플러딩(Flooding) 공격'은 망사용자가 수용할 수 있는 용량을 초과해 IDC(인터넷 데이터 센터) 및 아이템 거래를 위한 웹

서버까지 일반 사용자의 접속요청이 전달되지 못하게 되었다. 보안 전문가들은 아이템 거래 사이트의 DDoS 공격도 봇넷에 의한 것으로 분석하고 있다.

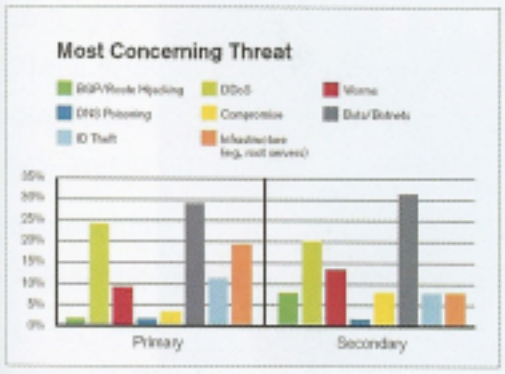
봇은 '소프트웨어 로봇(Robot)' 을 의미한다. 악의를 가진 해커가 불특정 다수의 PC를 감염시켜 자신의 마음대로 조종할 수 있는 봇으로 만들고, 이렇게 감염된 수천, 수만 대의 PC가 네트워크로 연결돼 하나의 '봇넷(Botnet)' 이 된다. 이렇게 형성된 봇넷은 봇에 감염된 PC의 통제권을 가진 봇 마스터(Bot master)에 의해 원격 조종되며, DDoS 공격이나 개인정보 수집, 스팸메일 전송, 피싱과 같은 사이버 공격 행위가 가능하기 때문에 최근 가장 위협적인 요소로 주목 받고 있다.

Arbor Networks의 2007년 조사에서는 봇넷이 DDoS 공격을 제치고 사이버상의 가장 위협적인 요소로 새로이 등극했다. 최근 발생하는 DDoS 공격의 다수가 봇넷에 의해 이루어지는 점을 감안하면 봇넷이 현재 가장 위협적인 존재라고 할 수 있다.

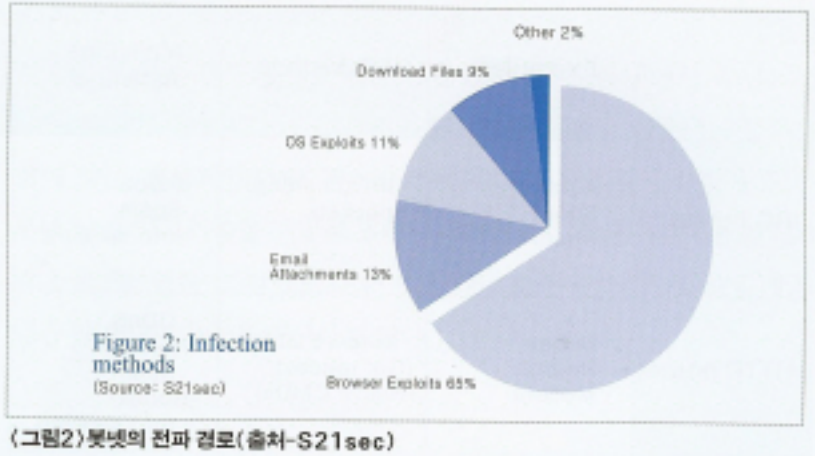
이렇게 위협적인 봇은 다양한 경로를 통해 감염, 전파 된다. 스팸메일에 실행코드를 첨부하여 전송하거나 웹에 감염코드를 싣고 취약성을 가진 PC를 찾아 감염시키는 방법, 메신저나 파일 공유 사이트를 통해 감염시키는 방법 등 그 수법이 매우 다양하다. S21sec의 조사 결과에 따르면 브라우저를 통해 전파되는 경우가 65%로 가장 큰 비중을 보였으며 다음으로 메일을 통해 전파되는 경우가 13%, 운영체제의 취약점을 찾아 전파되는 경우 11%, 파일 다운로드 형태로 전파되는 경우 9%, 기타 2%로 각각 집계 됐다.

이렇게 감염된 봇 코드는 루트 킷(Root kit)을 이용해 자신을 은폐하거나 안티 바이러스 프로그램을 종료시켜 사용자로 하여금 봇에 감염된 사실을 인지하지 못 하도록 한다. 또한 악성 봇 코드의 소스파일이 공개돼 하루에도 수천, 수만 종의 변종이 발생되고 있으며 한번 봇 코드에 감염되면 주기적인 업데이트, 패킹(Packing) 기법, 암호화 통신(SSL) 기법 이용 등 첨단 기술을 이용하기 때문에 그 탐지가 매우 어렵다. 또한 설사 탐지돼도 봇에 감염된 PC는 거의 대부분 선량한 일반 사용자의 PC이며, 봇 마스터를 추적하기는 매우 어려운 일이기 때문에 봇넷을 근본적으로 제거하는 것은 현실적으로 어렵다.

시만텍의 통계에 따르면 2006년 하반기에만 600만대의 새로운 봇이 생성된 것으로 나타난다. 이는 2006년 상반기에 비해 약 29% 증가한 수치다. 2007년에는 하루에 평균 5만2771대의 새로운 봇 감염 PC가 발견된 것으로 알려졌다. 인터넷의 아버지라고 불리는 빈트 서프(Vint Cerf) 구글 부사장은 전 세계적으로 인터넷에 연결된 PC의 4분의 1, 약 1억 대에서 1억5000 대 정도의 컴퓨터가 봇에 감염됐을 것이라고 추측한 바 있다. 실제로 봇넷이 14만대 이상의 봇으로 연결돼 있는 사례가 발견되기도 했다. 봇넷은 이미 우리 주변 깊숙한 곳까지 침투해 있어 자신도 모르는 사이에 악성 행위에 동참하고 있다.



〈그림1〉사이버 상의 가장 위협적인 요소 (출처-Arbor Networks)



〈그림2〉봇넷의 전파 경로(출처-S21sec)

공격자의 위치 숨기기 위해 봇넷 사용

근래 사이버 공격자들은 금전적 이득을 얻고 불법행위를 대행하기 위해 자신의 존재와 위치 의 추적이 쉽지 않도록, 자신들의 통제 하에 자유자재로 움직일 수 있는 수많은 컴퓨터를 거느리는 기술을 고안하게 됐다. 이러한 목적으로 생겨난 공격기법이 악성 봇을 이용해 원격에서 조정이 가능한 봇넷의 운영이다. 즉 공격자는 수천에서 수만 대의 컴퓨터를 봇 코드에 감염시켜 이들을 네트워크로 연결하여 동시에 제어가 가능하게 함으로써 자신들의 목적을 충실히 이행하는 엄청난 규모의 봇을 거느리게 되는 것이다.

이러한 봇을 차단하기 위한 여러 가지 기술이 대두되고 있지만 현실적으로 악성 봇 관련 프로그램의 소스 파일이 공개돼 수천 종의 변종의 봇들이 탄생되고 있어 근본적인 대책이 마련되지 않고 있다.

봇넷은 봇 마스터가 원하는 어떠한 사이버상의 모든 공격이 가능하다. DDoS 공격뿐만 아니라 봇 감염 PC로부터 개인정보 수집, 피싱, 악성 코드 배포, 스팸 메일 전송 등 그야말로 다재다능한 면모를 보이고 있다. 현존하는 스팸메일의 70~80% 이상이 봇넷에 의해 발송되고 있다고 조사된 바 있으며 피싱 사이트의 상당수가 봇넷에 의해 운영되고 있다.

국내현황도 크게 다를 바 없다. 한국 정보 보호 진흥원(KISA)의 발표 자료에 따르면 2007년 한해 평균 국내 봇넷 감염률은 전 세계 감염률의 11.3%에 달하는 것으로 집계 됐다. 하지만 이 수치는 탐지된 봇넷을 기반으로 조사된 결과이며, 실제로 탐지되지 않은 봇들이 훨씬 많다는 현실을 감안하면 발표된 수치는 빙산의 일각에 지나지 않을 것이다.

봇넷의 동작 방식

봇넷은 ▲전파(Propagation) ▲통신(Communication) ▲공격

	Examples	C&C Method	Malicious Activities	Persistent Connection	DNS Usage	Pros	Cons
IRC botnet	· Agobot · Sdbot · Rbot	· IRC(chatting) · packets	· DDoS · Spam · Info stealing	· PING · PONG	· Rally/Recruitment · Cloning/Migration · Egg download · Update/Attacks	· Easy to use · Robust · Synchronized · One point failure	· One point failure
HTTP botnet	· Bobax · Netbot · Mocbot	· Http variables (Get request/ reveive CMDs)	· DDoS (Netbot) · Spam (Bobox) · Info stealing	· No	· Rally/Recruitment · Cloning/Migration · Egg download · Update/Attacks	· Bypass IDS,F/W · Synchronized · Not persistent	· One point failure
P2P botnet	· Peacomm · Spamthru · Phatbot	· P2P search, publish CMDs	· Spam · Info stealing	· Publicize · Publicize ACK · Connect · Connect reply	· Update · Download from published URI · Attack preparation	· Decentralized · Hard to detect	· Hard to use · Not good for time sensitive attacks

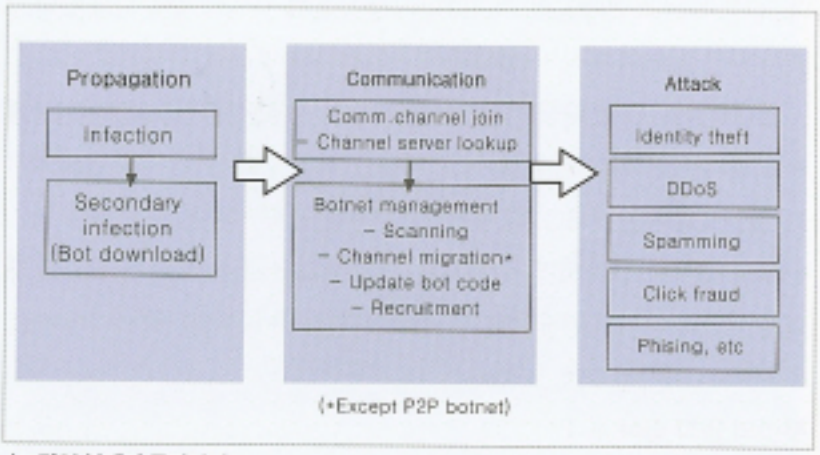
〈표1〉봇넷이 사용하는 서로 다른 프로토콜과 그 차이점

(Attack)의 3 단계를 거쳐 동작한다. 전파 단계에서는 다양한 방법으로 PC의 취약성을 이용해 악성코드를 설치한다. 감염된 PC는 실제 봇코드를 인터넷을 통해 다운로드하고 자신이 감염된 사실을 봇 마스터에게 알리기 위해 명령 및 제어 채널에 접속한다. 이후 채널을 통해 봇 마스터로부터 명령을 전달 받고 공격을 수행하게 된다. 이렇게 채널에 접속하여 봇 마스터로 부터 명령을 받아 악성행위를 수행하는 감염 PC의 네트워크가 봇넷이다. 통신단계에서 감염 PC가 접속하는 채널은 명령 및 제어(Command and Control, C&C)에 이용되는데, 이때 채널과 감염 PC 간의 통신 프로토콜에따라 중앙 집중형 봇넷과 분산형 봇넷으로 나뉘게 된다.

중앙 집중형은 봇넷과 통신을 위한 서버가 1대 혹은 다수의 서버로 구성되어, 감염된 PC에서 발생하는 봇들이 한 장소에 모여서 봇마스터로부터 명령을 받는 방식을 말한다. 대표적으로 IRC(Internet Relay Chatting) 프로토콜을 이용한 방식과 HTTP 프로토콜을 이용하는 웹 기반 방식이 있다. 한편 분산형은 통신을 위한 서버가 존재하지 않는 Peer-to-Peer (P2P) 프로토콜을 이용하는 방식으로 중앙집중형 보다는 복잡하고 사용하기 어려운 점이 있지만 탐지가 어렵고 비록 탐지가 되더라도 중앙서버가 없기 때문에 봇넷이 동작을 못하게 할 수 없다.

최신의 봇은 유저 인터페이스를 통해 쉽게 봇코드를 생성할 수 있고, 제어 할 수 있기 때문에 특별한 지식이나 기술이 없는 사람도 쉽게 봇넷을 만들고 이용할 수 있다.

실행압축기술인 패킹(Packing)기법 등을 이용하고, 자기 자신의 코드를 끊임없이 변경, 업데이트하기 때문에 바이러스 백신에 탐지되지 않는다. 또한 명령제어 채널과의 통신 과정에서 SSL과 같은 암호화 통신 기법을 이용하기 때문에 IDS, IPS, 방화벽 등 기존 보안 시스템에서 쉽게 탐지 할 수 없다. 명령제어 서버가 존재하는 중앙 집중형 봇넷의 경우에도 그 명령제어 서버가 주기적으로 이동하고, 유명한 포털사이트 등



〈그림3〉봇넷의 동작방식

〈그림4〉봇코드 생성 GUI: 유저 인터페이스를 이용해 간단한 설정값 변경으로 쉽게 봇코드를 생성할 수 있다.

〈그림5〉NetBot (HTTP 봇넷의 일종) 제어 GUI: 유저 인터페이스를 통해 쉽게 봇넷을 제어할 수 있다.

을 이용하는 경우도 있어 설사 탐지가 되더라도 동작하지 못하도록 막는 것이 쉽지 않다.

봇넷 탐지기술의 변천

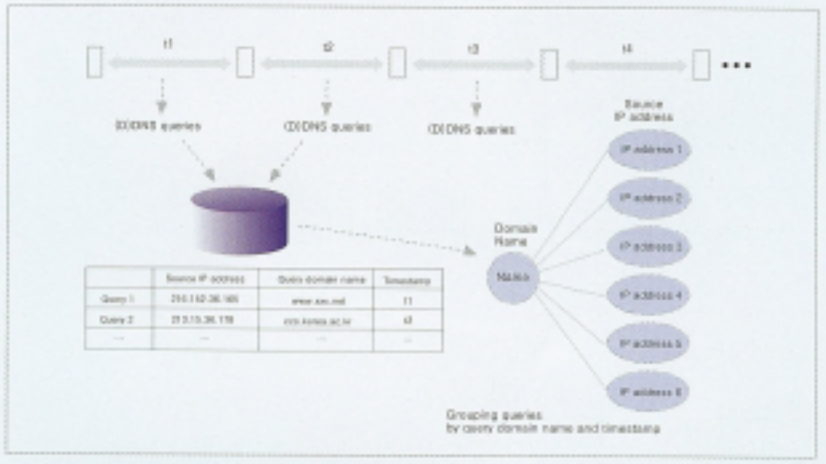
클라이언트 기반의 봇넷 탐지 기술은 크게 시그니처 탐지와 이상행위 탐지 기법으로 나누어진다. 대다수의 안티바이러스 솔루션에서 제공하는 것이 봇의 감염코드를 이용한 시그니처 탐지기법이다. 이러한 기법은 변종 봇이나 신종 봇에 대한 탐지가 느리고 소프트웨어적인 패킹기법을 통해 간단하게 회피할 수 있다는 단점이 있다. 클라이언트 기반의 이상행위 탐지 기술의 경우 스탠포드 대학의 스틴슨(Stinson)이 제안한 시스템 콜의 이상행위를 이용하여 탐지하는 기술이 제안 되었으나 오탐율이 높다는 단점을 갖는다.

네트워크 기반의 봇 탐지에 대한 연구는 주로 봇넷이 사용하는 IRC 포트인 6667번 포트의 모니터링을 통해 이루어졌다. 하지만 6667번이 아닌 다른 포트를 사용하는 봇들이 증가하자 6667번뿐만 아니라 다른 포트들에서도 트래픽의 패턴 매칭을 통해 봇넷의 IRC트래픽을 탐지하는 기술이 연구됐다. 그러나 이런 기술들은 모니터링하는 네트워크 링크의 트래픽 양이 많으면 데이터 처리가 힘들고 악성봇에 감염된 시스템과 봇넷 명령/제어 서버간 통신에 암호화 통신을 하면 패킷 모니터링을 통한 탐지가 힘들다. 실제로 현재 많은 봇들이 통신 암호화를 사용하고 있으며 포트 또한 알려지지 않은 포트를 사용한다.

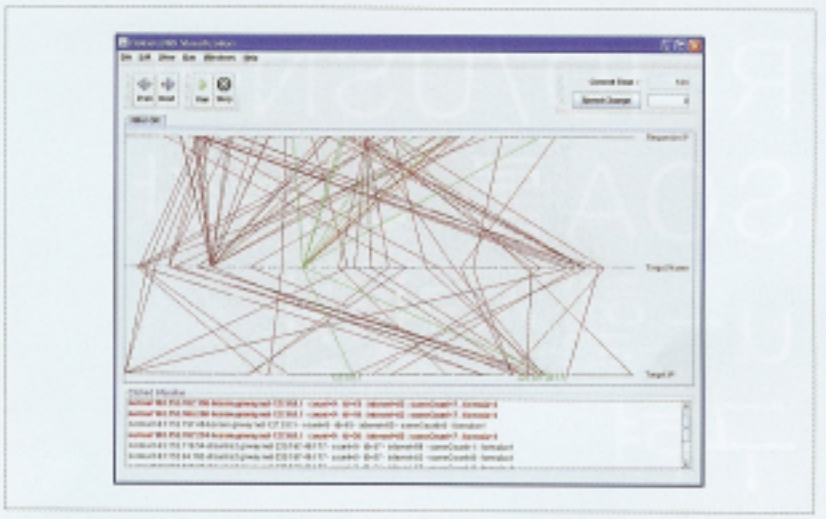
네트워크 기반의 시그니처 탐지 기술의 이러한 한계 때문에 근래에는 행위기반 탐지기술이 제안되고 있다. 조지아 공대의 데이곤(Dagon) 등이 2006년 ORAC 워크숍에서 봇넷의 도메인 네임 패턴을 이용해 명령/제어 채널서버를 찾는 기술을 발표한 바 있다. 2007년 USENIX Hotbot에서 얀 괴벨(Jan Goebel)이 봇마다 사용하는 고유의 Nickname 특성을 이용한 탐지 기법을 발표했다. 이들의 경우 봇마스터가 이 기법을 알 경우에 크게 봇넷의 구조나 감염기법의 변화 없이 단순하게 이름의 패턴만 바꿔 탐지를 회피 할 수 있는 한계를 갖고 있다.

봇넷을 통해 스팸메일을 보내기 전에 자신의 봇넷들이 블랙리스트에 추가되어 있는지 알아보는 질의 패턴의 분석을 통해 봇넷의 탐지를 수행하는 기술이 2007년 SRUTI에서 발표 됐으나, 마찬가지로 봇넷의 구조적/기술적 변화 없이 간단하게 탐지의 회피가 가능하다는 한계점을 가지고 있었다. 이처럼 대부분의 봇넷 탐지 기술들은 봇넷의 단편적인 특징이나 이름의 패턴만을 이용해 탐지를 수행하기 때문에 쉽게 회피가 가능하다는 한계점을 갖는다. Gu 등은 2007년 USENIX Security Symposium에서 IDS 기반 로그를 이용해 봇넷의 행동 주기에 따른 탐지를 수행하는 BotHunter를 제안했고, 2008년 NDSS 학술대회에서는 공간적, 시간적 유사성과 동시성을 이용해 봇넷을 탐지하는 BotSniffer를 제안하였다.

고려대학교 보안연구실의 봇넷 연구팀은 봇이 단체 행동을 하는 특징을 가진다는 점에 착안하여 기본적이면서도 쉽게 변할 수 없는 특성을 이용해 봇넷을 탐지하는 기법을 개발했다. 특히 봇넷에서 유발되는 DNS 쿼리가 정상적인 DNS 쿼리와 단체행동 특징을 이용해 구분되는 점에 착안, 정상적인 DNS 쿼리와 봇넷의 DNS 쿼리를 구분할 수 있는 봇넷 탐지 알고리즘을 개발해 2007년 IEEE CIT 학술대회에서 실험결과를 공개하였다. 그밖에도 DNS 쿼리 패턴을 이용한 시각화를 통해 봇넷을 탐지하는 기술을 연구하고 있다.



〈그림6〉그룹행위 기반 봇넷 탐지 메커니즘



〈그림7〉DNS이용한 봇넷 시각화 기법

지금까지 봇넷의 탐지를 위한 여러 연구가 진행되고 있으나 알려진 기법들이 쉽게 회피가 가능하다는 한계를 갖고 있다. 본고에서는 봇넷에서 명령/제어 과정에서 발생되는 DNS 쿼리를 이용하여 봇넷을 탐지하는 알고리즘을 소개하고, 봇넷의 탐지가 가능하다는 것을 실험을 통해 확인했다. 그러나 여기서 제안한 기법의 경우에도 회피할 수 있는 기술은 존재할 수 있으며, 이를 막기 위해서는 기술적인 보완이 필요하다.

봇넷의 탐지와 봇넷 기술의 발전은 마치 끝없는 군비확대 경쟁(arms race)과 같다. 하나의 탐지 기술이 개발되면 그것을 회피하는 기술이 개발된다. 봇넷의 피해는 사이버 기술이 발달하는 것과 비례하여 앞으로는 더욱더 커다란 위협으로 우리에게 다가올 것이다. 이러한 막강한 위협으로 다가오는 봇넷의 피해를 줄이기 위해 봇넷을 탐지하고 방어하는 매우 효과적인 기술의 개발을 위해 노력해야 할 것 이다.