(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2011년06월30일
(11) 등록번호 10-1045556
(24) 등록일자 2011년06월24일

(51) Int. Cl.
H04L 12/26 (2006.01) *H04L 12/22* (2006.01)
(21) 출원번호 10-2008-0132922
(22) 출원일자 2008년12월24일
심사청구일자 2008년12월24일
(65) 공개번호 10-2010-0074470
(43) 공개일자 2010년07월02일
(56) 선행기술조사문헌
KR100615080 B1
US20080307526 A1

(73) 특허권자
고려대학교 산학협력단
한국인터넷진흥원
(72) 발명자
정현철
이희조
(뒷면에 계속)
(74) 대리인
특허법인다울

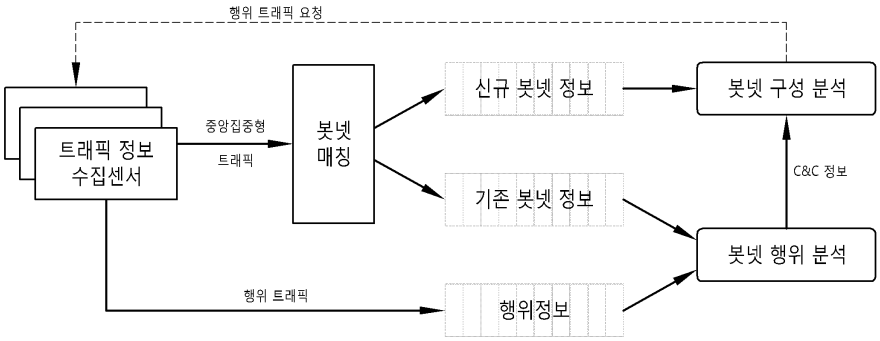
전체 청구항 수 : 총 11 항 심사관 : 천대녕

(54) 네트워크 기반의 ＩＲＣ 봇넷 탐지 방법

(57) 요 약

본 발명은 네트워크 기반의 IRC 봇넷 탐지 방법에 관한 것으로서, Domain 기반 트래픽 및 IP/Port 기반 트래픽을 포함하는 중앙집중형 접속 특성을 갖는 트래픽을 수집하고 수집된 트래픽의 봇넷 여부를 판별하는 트래픽 분류 모듈(TC), 상기 트래픽 분류 모듈(TC)에 의해 수집되어 봇넷으로 분류된 트래픽의 구성 분석을 수행하는 봇넷 구성 분석 모듈(BOA) 및 상기 트래픽 분류 모듈(TC)에 의해 수집되어 봇넷으로 분류된 트래픽의 행위 분석을 수행하는 봇넷 행위 분석 모듈(BBA)을 포함하는 봇넷 탐지 시스템에서, 다수의 트래픽 정보 수집 센서에 의해 수집된 상기 중앙집중적 접속 특성을 갖는 트래픽을 수집하는 제 1 단계 및 상기 수집된 트래픽이 IP/Port 기반 트래픽인 경우 기존에 탐지된 봇넷 정보와 비교하여 상기 트래픽을 기존 봇넷 및 신종 봇넷 중 어느 하나로 봇넷 매칭하는 제 2 단계를 포함하는 것을 특징으로 한다.

대 표 도 - 도1



(72) 발명자 **오주형**

임채태

최현상

지승구

노상균

이 발명을 지원한 국가연구개발사업
과제고유번호 2008-S-026-01
부처명 지식경제부
연구관리전문기관
연구사업명 IT성장동력기술개발사업
연구과제명 신종 봇넷 능동형 탐지 및 대응 기술 개발
기여율
주관기관 한국정보보호진흥원
연구기간 2008.03.01~2009.02.28

특허청구의 범위

청구항 1

사용자가 Domain을 해석하기 위해 외부와 통신하는 트래픽인 Domain 기반 트래픽 및 통신 주체간의 IP와 Port를 구분하여 통신 세션을 구분할 수 있는 트래픽인 IP/Port 기반 트래픽을 포함하는 트래픽이 하나의 수신지에 집중되는 중앙집중형 접속 특성을 갖는 트래픽을 수집하고 수집된 트래픽의 봇넷 여부를 판별하는 트래픽 분류 모듈(TC), 상기 트래픽 분류 모듈(TC)에 의해 수집되어 봇넷으로 분류된 트래픽에서 C&C 서버와 좀비를 분석하는 트래픽의 구성 분석을 수행하는 봇넷 구성 분석 모듈(BOA) 및 상기 트래픽 분류 모듈(TC)에 의해 수집되어 봇넷으로 분류된 트래픽에서 봇넷이 감염수를 증가시켜 그 규모를 증가시키는 행위인 봇넷의 확장 행위와 에그 다운로드를 포함하는 봇넷의 행위를 분석하는 트래픽의 행위 분석을 수행하는 봇넷 행위 분석 모듈(BBA)을 포함하는 봇넷 탐지 시스템에서, IRC 봇넷을 탐지하는 방법으로서,

다수의 트래픽 정보 수집 센서에 의해 수집된 상기 중앙집중적 접속 특성을 갖는 트래픽을 수집하는 제 1 단계; 및

상기 수집된 트래픽이 IP/Port 기반 트래픽인 경우 기존에 탐지된 봇넷 정보와 비교하여 상기 트래픽을 C&C 서버 정보와 좀비리스트를 포함하는 기존에 탐지된 봇넷 정보에 포함되는 기존 봇넷 및 상기 기존에 탐지된 봇넷 정보에 미포함되는 신종 봇넷 중 어느 하나로 봇넷 매칭하는 제 2 단계를 포함하는 것을 특징으로 하는 네트워크 기반의 IRC 봇넷 탐지 방법.

청구항 2

제 1 항에 있어서,

상기 제 2 단계가,

상기 트래픽의 중앙집중 서버가 상기 봇넷 정보의 C&C서버 정보와 상이하고 상기 트래픽의 접속 클라이언트 IP 리스트가 상기 봇넷 정보의 좀비리스트와 상이한 경우 상기 트래픽을 신종 봇넷 메시지 큐에 저장하는 제 2-1 단계를 포함하는 것을 특징으로 하는 네트워크 기반의 IRC 봇넷 탐지 방법.

청구항 3

제 2 항에 있어서,

상기 제 2 단계가,

상기 트래픽의 중앙집중 서버가 상기 봇넷 정보의 C&C 서버 정보와 일치하고 상기 트래픽의 접속 클라이언트 IP 리스트가 상기 봇넷 정보의 좀비리스트와 상이한 경우, 상기 트래픽을 봇넷 확장 행위로 구분하여 별도의 플래그를 부여하고 기존 봇넷 메시지 큐에 저장하는 제 2-2 단계;

상기 트래픽의 중앙집중 서버가 상기 봇넷 정보의 C&C 서버 정보와 일치하고 상기 트래픽의 접속 클라이언트 IP 리스트가 상기 봇넷 정보의 좀비리스트와 유사한 경우, 상기 트래픽을 C&C서버 재접속 및 에그 다운로드 행위로 구분하여 별도의 플래그를 부여하고 기존 봇넷 메시지 큐에 저장하는 제 2-3 단계; 및

상기 트래픽의 중앙집중 서버가 상기 봇넷 정보의 C&C 서버 정보와 상이하고 상기 트래픽의 접속 클라이언트 IP 리스트가 상기 봇넷 정보의 좀비리스트와 유사한 경우, 상기 트래픽을 C&C서버 이주를 포함하는 주요 봇넷 행위로 구분하여 별도의 플래그를 부여하여 기존 봇넷 메시지 큐에 저장하는 제 2-4 단계를 더 포함하는 것을 특징으로 하는 네트워크 기반의 IRC 봇넷 탐지 방법.

청구항 4

제 1 항에 있어서,

상기 트래픽 수집 관리 모듈에 의해 수집된 트래픽 중 Domain 기반 트래픽에 대하여 VDNS(Virtual DNS) 여부를 체크하는 제 3 단계를 더 포함하는 것을 특징으로 하는 네트워크 기반의 IRC 봇넷 탐지 방법.

청구항 5

제 1 항에 있어서,

상기 제 2 단계 이전에,

상기 수집된 트래픽에 대하여, 일정 대기시간동안 동일 C&C를 기준으로 유입되는 좀비 IP 리스트들을 추가적으로 열거하는 제 1-1 단계;

설정된 대기시간이 경과한 후 열거된 좀비 IP 리스트의 개수가 임계값을 초과하면 상기 수집된 트래픽을 봇넷의 트래픽으로 판단하는 제 1-2 단계; 및

설정된 대기시간이 경과한 후 좀비 IP 리스트의 개수가 임계값에 미치지 못한 경우, 관제시스템의 공유정보에 의하여 업데이트되는 C&C서버 블랙리스트와 매칭하여 봇넷의 트래픽 여부를 판단하는 제 1-3 단계를 더 포함하는 것을 특징으로 하는 네트워크 기반의 IRC 봇넷 탐지 방법.

청구항 6

제 5 항에 있어서,

상기 제 1-3 단계가,

"접속 Dst IP/Port" 와 "C&C서버 IP/Port" 를 매칭하는 제 1-4 단계; 및

"요청 도메인에 대한 응답 IP" 와 "DNS 싱크홀 IP" 를 매칭하는 제 1-5 단계를 포함하는 것을 특징으로 하는 네트워크 기반의 IRC 봇넷 탐지 방법.

청구항 7

제 1 항에 있어서,

상기 제 2 단계 이후에,

상기 수집된 트래픽 정보(Domain, Dst_IP/Port)를 수신하여 임시 구성 로그에 임시 저장하는 제 4 단계;

상기 임시 구성 로그로부터 분석에 필요한 트래픽 정보를 주기적으로 읽어 Domain 및 Dst_IP/Port 별 유사도를 분석하고, 봇넷 행위로 탐지된 트래픽을 전송하는 제 5 단계;

제 5 단계에서 탐지된 봇넷 트래픽을 전송받고 C&C를 추출 및 저장한 후, 분석을 마친 트래픽을 전송하는 제 6 단계;

제 6 단계에서 전송된 봇넷 트래픽을 전송받아 봇넷으로 탐지된 IRC Channel에 접근하는 좀비리스트를 추출하여 분석결과 로그에 저장하는 제 7 단계; 및

제 6 단계 및 제 7 단계에서 분석된 결과를 종합하여 로그 관리자에게 전송하는 제 8 단계를 포함하는 것을 특징으로 하는 네트워크 기반의 IRC 봇넷 탐지 방법.

청구항 8

제 7 항에 있어서,

상기 제 5 단계가,

주기적으로 임시 구성 로그로부터 Domain 정보를 읽어 각 Domain별로 요청한 소스 IP들을 매트릭스에 기록하는 제 5-1 단계; 및

설정된 시간이 경과한 후 상기 매트릭스를 분석하여 Domain 유사도를 측정하고 좀비 IP 리스트를 생성하는 제 5-2 단계를 포함하는 것을 특징으로 하는 네트워크 기반의 IRC 봇넷 탐지 방법.

청구항 9

제 7 항에 있어서,

상기 제 5 단계가,

주기적으로 임시 구성 로그로부터 Dst_IP/Port 정보를 읽어 각 IP/Port 조합과 매칭되는 패킷을 전송한 소스 IP들을 매트릭스에 기록하는 제 5-3 단계; 및

설정된 시간이 경과한 후 상기 매트릭스를 분석하여 Dst_IP/Port 유사도를 측정하고 좀비 IP 리스트를 생성하는 제 5-4 단계를 포함하는 것을 특징으로 하는 네트워크 기반의 IRC 봇넷 탐지 방법.

청구항 10

제 1 항에 있어서,

상기 제 1 단계에서 수집된 트래픽 정보가 Domain 기반 트래픽인 경우 전송 데이터 포맷이,

헤더에 상기 트래픽의 발생시간인 Time 필드를 포함하고,

봇넷의 C&C 서버 필드에, C&C 서버의 DNS 쿼리 도메인명인 C&C Domain 필드 및 C&C 서버의 DNS 쿼리에 대한 응답 IP인 C&C IP 필드를 포함하고,

봇넷의 좀비리스트 필드에, 발견된 총 좀비 개체수인 Count 필드, 처음 좀비 발생시점부터 마지막 좀비 발생까지의 시간 구간인 Time Window 필드 및 접속한 총 좀비의 IP 리스트인 좀비 IP 리스트를 포함하고,

상기 C&C 서버 필드의 값이 DNS서버와의 통신 트래픽으로부터 수집되고 수집 대상 포트번호가 53번 포트인 것을 특징으로 하는 네트워크 기반의 IRC 봇넷 탐지 방법.

청구항 11

제 1 항에 있어서,

상기 제 1 단계에서 수집되는 트래픽이 IP/Port 기반 트래픽인 경우 전송 데이터 포맷이,

헤더에 상기 트래픽의 발생시간인 Time 필드를 포함하고,

봇넷의 C&C 서버 필드에, C&C 서버의 IP인 C&C IP 필드 및 접속 포트번호인 C&C Port 필드를 포함하고,

봇넷의 좀비리스트 필드에, 발견된 총 좀비 개체수인 Count 필드, 처음 좀비 발생부터 마지막 좀비 발생까지의 시간 구간인 Time Window 필드 및 접속한 총 좀비의 IP 리스트인 좀비 IP 리스트 필드를 포함하고,

상기 C&C 서버 필드의 값이 C&C 서버와 중간 전송 객체 없이 바로 통신하는 트래픽인 직접 통신 트래픽으로부터 수집되고, 수집 대상 포트번호가 모든 포트인 것을 특징으로 하는 네트워크 기반의 IRC 봇넷 탐지 방법.

명 세 서

발명의 상세한 설명

기 술 분 야

[0001] 본 발명은 봇넷 탐지 방법에 대한 것으로, 구체적으로 네트워크 기반의 IRC 봇넷을 탐지하기 위한 방법에 대한 것이다.

배 경 기 술

[0002] 현재 사이버 공간에서는 수많은 위협들이 대두되고 있다. 제3자의 개인정보를 갈취 또는 수집하여 악용하고, 불특정 다수를 향해 음란, 광고메일을 유포하고 금전적 이익을 보거나, 경쟁사의 정보화 기기의 서비스를 못하게 하는 등 인터넷상의 위협요인들은 산재해 있다. 이러한 사이버 피해가 두드러지는 가운데 새로운 위협적 요소가

인터넷을 장악해 나가고 있는데 그것이 바로 봇넷이다. 최근 Arbor Networks에서는 가장 심각한 네트워크 위협으로 봇넷과 분산서비스 거부공격(DDoS)을 선정하기도 하였다.

[0003] 봇넷(BotNet)이란 악성 소프트웨어인 봇에 감염된 다수의 컴퓨터들이 네트워크로 연결되어 있는 형태를 말한다. 즉, 봇들을 자유자재로 통제하는 권한을 가진 봇마스터에 의해 원격 조종되며 각종 악성행위를 수행할 수 있는 수천에서 수십만 대의 악성프로그램인 봇(Bot)에 감염된 컴퓨터들이 네트워크로 연결되어 있는 형태를 봇넷이라 한다.

[0004] 봇넷은 1993년에 EggDrop으로 처음 나온 이후로 최근 10년간 Forbot, PBot, Toxbot, Machbot,PHP Bot, Storm Bot 등으로 진화한 봇이 출현하였으며, 최근에는 너무 많은 변종 봇이 출현하면서 대응을 매우 어렵게 하고 있다(매일 5,000개의 신규 악성코드 출현, TechNewsWorld, 2007). 특히, 전 세계적으로 C&C 서버(Command & Control, 봇 좀비들에게 명령을 내리고 제어하기 위한 서버)와 악성 봇이 광범위하게 분포하고 있고 특정 지역에 밀집되는 양상을 보이고 있다. 이는 초고속 인터넷이 잘 갖추어진 환경에서는 기존에 비해 1/10의 PC들만 이용해도 더욱 강력한 DDoS와 같은 공격이 가능하기 때문이며, 특히 초고속의 인터넷 인프라가 잘 갖추어져 있는 국내 지역은 봇넷 감염지로 선호되고 있다. 또한, 세계적으로 봇에 감염되어 좀비 PC로 바뀌는 PC의 수가 지속적으로 증가하고 있으며 봇넷의 규모 또한 커지고 있다. TCP/IP 프로토콜 공동 창시자인 Vint Cerf는 전 세계 컴퓨터의 약 11% 정도인 1억~1억 5천 컴퓨터가 봇 악성코드에 감염되어 공격 수행에 사용될 것으로 예상하였으며, 현재까지 알려진 가장 큰 봇넷은 Storm 봇넷으로 230,000개의 좀비들로 연결되어 있는 것으로 알려져 있다.

[0005] 봇넷으로 인한 공격이 더욱 심각해지는 이유는 범죄화 양상을 띠고 있기 때문이다. 2007년 발생한 아이템 거래업체 서비스 장애 유발 및 현금요구 협박 사고에서와 같이 서비스 장애유발을 빌미로 서비스 업체에 협박하여 금품 갈취하거나, 개인/금융 정보 수집 및 스팸 발송 등을 통하여 대가를 받는 사고가 빈번하게 발생하고 있다.

[0006] 봇은 웜/바이러스, 백도어, 스파이웨어, 루트킷 등 다양한 악성코드들의 특성을 복합적으로 지니며, 봇넷을 통해 DDoS, Ad-ware, Spyware, 스팸발송, 정보불법 수집 등 대부분의 사이버 공격이 가능하다.

[0007] 초기의 봇넷은 구조가 유연하고 널리 사용되는 IRC의 특성을 이용한 IRC 봇넷이 주를 이루었었다. 하지만, 탐지 및 대응을 보다 어렵게 하기 위해 웹 프로토콜인 HTTP를 기반으로 하거나, C&C라는 중앙집중형 명령/제어 방식(IRC, HTTP 봇넷)에서 탈피하여, 모든 좀비들이 C&C가 될 수 있는 분산형 명령/제어 방식(P2P 봇넷)의 봇넷으로 진화하고 있다.

[0008] 더 나아가, 명령/제어를 위해 2가지 이상의 프로토콜을 사용하는 하이브리드 형태로 진화하고 있다. 최근 등장한 MayDay 웜의 경우, 하이퍼텍스트 프랜스퍼 프로토콜(hypertext transfer protocol, HTTP) 및 인터넷 제어 메시지 프로토콜(Internet Control Message Protocol, ICMP)과 같이 두 가지 프로토콜을 사용하거나, 다수의 중앙 집중 포인트(C&C서버)가 존재하고 중앙 집중 포인트는 P2P 방식으로 연결되는 하이브리드 형태로 진화하고 있다.

[0009] 명령/제어 방식의 진화와 함께, 악성코드도 빠른 속도록 진화하여 탐지/대응을 매우 어렵게 하고 있다. Packing/압축/암호화 기술, VM(Virtual Machine)/디버거/샌드박스 탐지 우회 기술, 악성코드를 숨기기 위한 RootKit 기술 등이 적용되어 봇에 대한 탐지 및 분석을 매우 어렵게 하고 있으며, 감염경로 또한 기존의 시스템 취약점을 악용한 방식에서, 웹, 이메일, 메신저 등 다양한 수단으로 다양해지고 있다.

[0010] 이러한 봇넷의 대응 기술은 적용 대상에 따라, PC상에서 악성 봇 프로그램 설치 및 행동을 기반으로 탐지/분석하는 호스트 기반과 봇 좀비 및 C&C로부터의 네트워크 트래픽을 기반으로 탐지/분석하는 네트워크 기반으로 구분되며, 기술 특성에 따라 시그니처 기반과 행위기반으로 구분될 수 있다.

[0011] 최근 들어, 봇넷의 심각성이 부각되면서 국제적으로 연구가 활성화되고 있으나 진화하고 있는 봇넷에 대한 대응이 아직도 미흡한 수준이다.

발명의 내용

해결 하고자하는 과제

[0012] 본 발명은 상기와 같은 문제점을 해결하기 위하여 제시된 것으로서, 본 발명의 목적은 네트워크 기반의 IRC 봇넷을 효과적으로 탐지할 수 있는 네트워크 기반의 IRC 봇넷 탐지 방법을 제공하는 것이다.

과제 해결수단

[0013] 상기의 목적을 달성하기 위하여, 본 발명에 따른 네트워크 기반의 IRC 봇넷 탐지 방법은, 사용자가 Domain을 해석하기 위해 외부와 통신하는 트래픽인 Domain 기반 트래픽 및 통신 주체간의 IP와 Port를 구분하여 통신 세션을 구분할 수 있는 트래픽인 IP/Port 기반 트래픽을 포함하는 트래픽이 하나의 수신지에 집중되는 중앙집중형 접속 특성을 갖는 트래픽을 수집하고 수집된 트래픽의 봇넷 여부를 판별하는 트래픽 분류 모듈(TC), 상기 트래픽 분류 모듈(TC)에 의해 수집되어 봇넷으로 분류된 트래픽에서 C&C 서버와 좀비를 분석하는 트래픽의 구성 분석을 수행하는 봇넷 구성 분석 모듈(BOA) 및 상기 트래픽 분류 모듈(TC)에 의해 수집되어 봇넷으로 분류된 트래픽에서 봇넷이 감염수를 증가시켜 그 규모를 증가시키는 행위인 봇넷의 확장 행위와 에그 다운로드를 포함하는 봇넷의 행위를 분석하는 트래픽의 행위 분석을 수행하는 봇넷 행위 분석 모듈(BBA)을 포함하는 봇넷 탐지 시스템에서, IRC 봇넷을 탐지하는 방법으로서, 다수의 트래픽 정보 수집 센서에 의해 수집된 상기 중앙집중적 접속 특성을 갖는 트래픽을 수집하는 제 1 단계 및 상기 수집된 트래픽이 IP/Port 기반 트래픽인 경우 기존에 탐지된 봇넷 정보와 비교하여 상기 트래픽을 C&C 서버 정보와 좀비리스트를 포함하는 기존에 탐지된 봇넷 정보에 포함되는 기존 봇넷 및 상기 기존에 탐지된 봇넷 정보에 미포함되는 신종 봇넷 중 어느 하나로 봇넷 매칭하는 제 2 단계를 포함하는 것을 특징으로 한다.

[0014] 바람직하게는, 제 2 단계가, 상기 트래픽의 중앙집중 서버가 상기 봇넷 정보의 C&C서버 정보와 상이하고 상기 트래픽의 접속 클라이언트 IP 리스트가 상기 봇넷 정보의 좀비리스트와 상이한 경우 상기 트래픽을 신종 봇넷 메시지 큐에 저장하는 제 2-1 단계, 상기 트래픽의 중앙집중 서버가 상기 봇넷 정보의 C&C 서버 정보와 일치하고 상기 트래픽의 접속 클라이언트 IP 리스트가 상기 봇넷 정보의 좀비리스트와 상이한 경우, 상기 트래픽을 봇넷 확장 행위로 구분하여 별도의 플래그를 부여하고 기존 봇넷 메시지 큐에 저장하는 제 2-2 단계, 상기 트래픽의 중앙집중 서버가 상기 봇넷 정보의 C&C 서버 정보와 일치하고 상기 트래픽의 접속 클라이언트 IP 리스트가 상기 봇넷 정보의 좀비리스트와 유사한 경우, 상기 트래픽을 C&C서버 재접속 및 에그 다운로드 행위로 구분하여 별도의 플래그를 부여하고 기존 봇넷 메시지 큐에 저장하는 제 2-3 단계, 그리고 상기 트래픽의 중앙집중 서버가 상기 봇넷 정보의 C&C 서버 정보와 상이하고 상기 트래픽의 접속 클라이언트 IP 리스트가 상기 봇넷 정보의 좀비리스트와 유사한 경우, 상기 트래픽을 C&C서버 이주를 포함하는 주요 봇넷 행위로 구분하여 별도의 플래그를 부여하여 기존 봇넷 메시지 큐에 저장하는 제 2-4 단계를 포함하도록 한다.

[0015] 더욱 바람직하게는, 상기 트래픽 수집 관리 모듈에 의해 수집된 트래픽 중 Domain 기반 트래픽에 대하여 VDNS(Virtual DNS) 여부를 체크하는 제 3 단계를 더 포함하도록 한다. 또한, 제 2 단계 이전에, 상기 수집된 트래픽에 대하여, 일정 대기시간동안 동일 C&C를 기준으로 유입되는 좀비 IP 리스트들을 추가적으로 열거하는 제 1-1 단계, 설정된 대기시간이 경과한 후 열거된 좀비 IP 리스트의 개수가 임계값을 초과하면 상기 수집된 트래픽을 봇넷의 트래픽으로 판단하는 제 1-2 단계 및 설정된 대기시간이 경과한 후 좀비 IP 리스트의 개수가 임계값에 미치지 못한 경우, 관제시스템의 공유정보에 의하여 업데이트되는 C&C서버 블랙리스트와 매칭하여 봇넷의 트래픽 여부를 판단하는 제 1-3 단계를 더 포함하는 것을 특징으로 하며, 제 1-3 단계가, "접속 Dst IP/Port”와 “C&C서버 IP/Port”를 매칭하는 제 1-4 단계 및 "요청 도메인에 대한 응답 IP”와 “DNS 싱크홀 IP”를 매칭하는 제 1-5 단계를 포함하도록 한다.

[0016] 더욱 바람직하게는, 제 2 단계 이후에, 상기 수집된 트래픽 정보(Domain, Dst_IP/Port)를 수신하여 임시 구성 로그에 임시 저장하는 제 4 단계, 상기 임시 구성 로그로부터 분석에 필요한 트래픽 정보를 주기적으로 읽어 Domain 및 Dst_IP/Port 별 유사도를 분석하고, 봇넷 행위로 탐지된 트래픽을 전송하는 제 5 단계, 제 5 단계에서 탐지된 봇넷 트래픽을 전송받고 C&C를 추출 및 저장한 후, 분석을 마친 트래픽을 전송하는 제 6 단계, 제 6 단계에서 전송된 봇넷 트래픽을 전송받아 봇넷으로 탐지된 IRC Channel에 접근하는 좀비리스트를 추출하여 분석결과 로그에 저장하는 제 7 단계 및 제 6 단계 및 제 7 단계에서 분석된 결과를 종합하여 로그 관리자에게 전송하는 제 8 단계를 포함하도록 한다.

[0017] 이때, 제 5 단계는, 주기적으로 임시 구성 로그로부터 Domain 정보를 읽어 각 Domain별로 요청한 소스 IP들을 매트릭스에 기록하는 제 5-1 단계 및 설정된 시간이 경과한 후 상기 매트릭스를 분석하여 Domain 유사도를 측정하고 좀비 IP 리스트를 생성하는 제 5-2 단계를 포함하고, 주기적으로 임시 구성 로그로부터 Dst_IP/Port 정보를 읽어 각 IP/Port 조합과 매칭되는 패킷을 전송한 소스 IP들을 매트릭스에 기록하는 제 5-3 단계 및 설정된 시간이 경과한 후 상기 매트릭스를 분석하여 Dst_IP/Port 유사도를 측정하고 좀비 IP 리스트를 생성하는 제 5-4 단계를 포함하도록 한다.

[0018] 한편, 제 1 단계에서 수집된 트래픽 정보가 Domain 기반 트래픽인 경우 전송 데이터 포맷은, 헤더에 상기 트래픽의 발생시간인 Time 필드를 포함하고, 봇넷의 C&C 서버 필드에, C&C 서버의 DNS 쿼리 도메인명인 C&C Domain 필드 및 C&C 서버의 DNS 쿼리에 대한 응답 IP인 C&C IP 필드를 포함하고, 봇넷의 좀비리스트 필드에, 발견된 총

좀비 개체수인 Count 필드, 처음 좀비 발생시점부터 마지막 좀비 발생까지의 시간 구간인 Time Window 필드 및 접속한 총 좀비의 IP 리스트인 좀비 IP 리스트를 포함하고, 상기 C&C 서버 필드의 값이 DNS서버와의 통신 트래픽으로부터 수집되고 수집 대상 포트번호가 53번 포트가 되도록 한다.

[0019] 또한, 제 1 단계에서 수집되는 트래픽이 IP/Port 기반 트래픽인 경우 전송 데이터 포맷은, 헤더에 상기 트래픽의 발생시간인 Time 필드를 포함하고, 봇넷의 C&C 서버 필드에, C&C 서버의 IP인 C&C IP 필드 및 접속 포트번호인 C&C Port 필드를 포함하고, 봇넷의 좀비리스트 필드에, 발견된 총 좀비 개체수인 Count 필드, 처음 좀비 발생부터 마지막 좀비 발생까지의 시간 구간인 Time Window 필드 및 접속한 총 좀비의 IP 리스트인 좀비 IP 리스트 필드를 포함하고, 상기 C&C 서버 필드의 값이 C&C 서버와 중간 전송 객체 없이 바로 통신하는 트래픽인 직접 통신 트래픽으로부터 수집되고, 수집 대상 포트번호가 모든 포트가 되도록 한다.

효 과

[0020] 이상 설명한 바대로, 본 발명에 따른 네트워크 기반의 IRC 봇넷 탐지 방법은 IP/Port 기반 트래픽에 대하여 봇넷 매칭 모듈에 의해 신규 봇넷 및 기존 봇넷으로 분류하고, 기존 봇넷에 대하여 그 유형에 따라 분류하여 메시지 큐에 저장함으로써 봇넷을 효과적으로 탐지하고 봇넷 행위 분석 모듈이 봇넷의 유형을 빠르게 인식할 수 있도록 한다.

발명의 실시를 위한 구체적인 내용

[0021] 이하에서는, 첨부한 도면을 참조하여 본 발명의 장점, 특징 및 바람직한 실시예에 대하여 상세히 설명하도록 한다.

[0022] 도 1은 본 발명에 따른 봇넷 탐지 시스템의 구성도이다. 트래픽 정보 수집 센서로부터 IRC 봇넷 탐지를 위하여 중앙집중형 접속 특성을 가지는 트래픽의 테이터 포맷을 전달받고, 이 과정을 통해 스타형 토폴로지를 구성하는 중앙집중형 네트워크를 발견하며, 봇넷 매칭 모듈에서 기존 봇넷 여부를 판별한 후 각종 행위를 재분류한다. 봇넷 매칭 모듈은 발견된 중앙집중형 트래픽을 유발하는 네트워크가 이미 탐지된 기존 봇넷과 매칭되는지의 여부를 체크한다. 봇넷 매칭 모듈에 의해 신규 봇넷으로 판별될 경우 봇넷 구성 분석 모듈(BOA)은 발견된 좀비들의 추가적인 행위 트래픽을 트래픽 정보 수집 센서로 요청한다.

[0023] 구체적으로, 봇넷 매칭 모듈은 중앙집중 서버와 접속 클라이언트 IP 리스트의 유사성을 식별함으로써 발견된 중앙집중 트래픽의 봇넷 매칭을 수행한다.

[0024] 먼저, 봇넷 매칭 모듈은 탐지된 기존 봇넷 정보들과 매칭하여 중앙집중 서버가 기존 C&C서버 정보와 같지 않고 접속 클라이언트 IP 리스트 또한 기존 좀비리스트와 유사하지 않을 때 신규 봇넷으로 규정하여 봇넷 여부를 탐지한다. 봇넷 구성 분석 모듈(BOA)은 탐지된 신규 봇넷의 좀비들이 발생시키는 추가적인 행위 트래픽을 수집할 수 있도록 트래픽 정보 수집 센서에게 요청하여 이후 봇넷의 행위를 모니터링하고 분석한다.

[0025] 한편, 중앙집중 서버가 기존 C&C 서버 정보와 같거나 접속 클라이언트 IP 리스트가 기존 좀비리스트와 유사할 경우 봇넷 매칭 모듈은 기존 봇넷으로 규정하는데, 기존 봇넷의 매칭 결과를 이후 봇넷 행위 분석 모듈(BBA)이 빠르게 인식할 수 있도록 하기 위하여 봇넷 매칭 모듈은 세가지 행위기반(A, B, C)에 의하여 선분류하고 플래그를 부여하도록 한다. 각 행위기반에 따른 플래그는 이하와 같다.

[0026] - A 플래그(봇넷 확장 행위): 중앙집중 서버가 C&C 서버 정보와 일치하지만 접속 클라이언트 IP 리스트가 기존 좀비리스트와 차이를 보이는 경우

[0027] - B 플래그(C&C서버 재접속 및 에그 다운로드 행위): 중앙집중 서버가 C&C 서버 정보와 일치하고 접속 클라이언트 IP 리스트도 기존 좀비리스트와 유사한 경우

[0028] - C 플래그(C&C서버 이주 및 기타 주요 봇넷 행위): 중앙집중 서버가 C&C 서버 정보와 다르지만 접속 클라이언트 IP 리스트가 기존 좀비리스트와 유사한 경우

[0029] 상기한 바와 같이, 중앙집중형 트래픽(DNS 요청 Domain, 중앙집중 서버의 IP/Port 정보)을 모니터링하여 탐지된 봇넷에 대하여 봇넷 구성 분석 모듈(BOA)은 트래픽 정보 수집 센서에게 추가적인 행위 트래픽을 요청하고, 이때 전달받은 행위 트래픽을 기반으로 봇넷 행위 분석 모듈(BBA)은 다양한 봇넷 행위들을 재분류한다. IRC Channel은 봇넷 구성 분석 모듈(BOA)로 다시 전달되어 봇넷 프로토콜 판단의 정확한 기준이 되도록 한다. 중앙집중형 트래픽을 기반으로 대분류된 A/B/C 세가지 그룹의 행위들은 행위 트래픽을 기반으로 다시 각 구체적인 행위들로

세분류된다.

[0030] 도 2는 본 발명에 따른 트래픽 분류 모듈(TC)의 블럭 구성도이다. 도 1에 도시한 바와 같이, 본 발명에 따른 트래픽 분류 모듈(TC)은, 필터, 수집/분류 관리 정책 모듈, 트래픽 수집관리 모듈, 봇넷 매칭 모듈 및 VDNS 체크 모듈을 포함한다.

[0031] 필터는 다중의 트래픽 정보 수집 센서로부터 전달되는 유효 트래픽만을 수신하며, 자신이 담당하는 센서들의 IP에 대한 화이트 리스트 방화벽 기능을 제공한다.

[0032] 수집/분류 관리 정책 모듈은 관리 시스템으로부터 전달되는 정책을 설정하여 상태값을 조정하고, 봇넷 관제/보안관리 시스템의 실시간 보안 이벤트를 참고하여 봇넷 탐지 시스템이 신속히 대응할 수 있도록 관리 정책에 반영한다. 따라서, 알려지지 않은 신종 봇넷에 빠르게 대응하기 위하여 관제시스템들의 공유정보를 기반으로 좀비 리스트 분포량과 관계없이 C&C 서버 블랙리스트 매칭이 가능하도록 할 수 있다.

[0033] 트래픽 수집 관리 모듈은 중앙집중적인 접속 특성을 가지는 트래픽, 즉 공동의 단일 서버에 연결되어 스타형 토폴로지를 구성하는 기본 트래픽을 수집 및 모니터링한다. 이는 IRC 봇넷 탐지를 위한 기본 수집트래픽으로 이를 기반으로 봇넷 매칭 모듈에서 봇넷 여부를 판별하게 된다. 또한, 트래픽 수집 관리 모듈은 분석을 위한 추가적인 요청 트래픽(봇넷 구성 분석 모듈의 행위 트래픽 요청에 따른 트래픽)을 수집함으로써, 봇넷 구성 분석 모듈이 신종 봇넷의 IRC 프로토콜(Channel 정보 등)을 분석할 수 있도록 한다. 이 경우 추가적인 요청 트래픽은 이미 탐지된 봇들 가운데 다시 샘플링된 일부 봇, 즉 봇넷 구성 분석 모듈이 행위 트래픽을 요청한 일부 봇에 대하여만 수집하도록 한다. 또한, 트래픽 수집 관리 모듈은 관제시스템의 보안관리 정책 및 공유정보관리 정책에 의하여 발견된 새로운 봇넷의 위협을 사전에 대비하기 위하여 긴급 요청된 특정한 트래픽을 전달받을 수도 있다.

[0034] 봇넷 매칭 모듈은 트래픽 수집 관리 모듈에 의해 수집된 트래픽이 IP/Port 기반 트래픽인 경우 기존 봇넷과 비교하여 기존 봇넷 또는 신종 봇넷으로 구분한다. 먼저, 봇넷 매칭 모듈은 트래픽 수집 관리 모듈에 의해 수집된 트래픽에 대하여 신종 봇넷 구성의 유사 트래픽을 탐지한다. 구체적으로, 탐지된 기존 봇넷 정보들과 매칭하여 중앙집중 서버가 기존 C&C서버 정보와 다르면서 접속 IP리스트가 기존 좀비리스트와 다른 경우 이를 탐지하여 신종 봇넷 메시지 큐에 저장한다. 하지만, 정상 트래픽일 확률 또한 배재할 수 없으므로 이후 봇넷 구성 분석(BOA) 모듈에서 세부적인 재분류가 이루어져야 하며 신종 봇넷의 구성 단계에서 발생된 트래픽인지 진위를 최종 판단하도록 한다. 다음으로, 봇넷 매칭 모듈은 기존 봇넷 탐지 및 대분류 작업을 수행한다. 구체적으로 탐지된 기존 봇넷 정보들과 매칭하여, 중앙집중 서버가 기존 C&C서버 정보와 같거나 혹은 접속 IP 리스트가 기존 좀비 리스트와 유사할 경우 이를 탐지하여 기존 봇넷 메시지 큐에 저장한다. 봇넷 매칭 모듈은 메시지 큐 저장시 대분류된 세가지(A/B/C) 행위 결과를 플래그 비트로 명시하여 이후 봇넷 행위 분석(BBA) 모듈이 빠르게 인식하도록 한다. 세가지 행위 결과 플래그의 유형은 상기한 바와 같다.

[0035] VDNS 체크 모듈은 트래픽 수집 관리 모듈에 의해 수집된 트래픽이 Domain 기반 트래픽인 경우 VDNS(Virtual DNS) 여부를 체크한다. VDNS는 공격자가 임의로 운영하는 DNS로서 이를 이용하면 Dynamic DNS를 이용하지 않고도 C&C 서버 도메인의 IP를 쉽게 변경할 수 있다. 이때 요청되는 도메인은 대부분 정상 DNS 쿼리시에는 리턴되지 않는 경우가 많으므로, VDNS 체크 모듈은 정상 DNS 테이블과 매칭하여 VDNS 여부를 체크하고 새로이 탐지된 VDNS 쿼리 정보를 메시지 큐에 저장하여 봇넷 구성 분석 모듈 및 봇넷 행위 분석 모듈이 참고할 수 있도록 한다.

[0036] 도 3은 트래픽 분류 모듈의 트래픽 분류 과정을 도시한 도이다. 도 2에 도시한 바와 같이, 트래픽 분류 모듈의 트래픽 분류 과정은 수집관리 프로시저, 봇넷 매칭 프로시저, VDNS 체크 프로시저로 이루어진다.

[0037] 1. 수집 관리 프로시저

[0038] 트래픽 정보 수집 센서로부터 수집된 트래픽 중 유효 트래픽을 수집/분류 관리 정책에 따라 필터가 필터링한 후, 트래픽 수집 관리 모듈이 상기 수집/분류 관리 정책에 따라 중앙 집중적인 접속 특성을 가지는 트래픽을 수집한다(ST100 내지 ST120).

[0039] 이후 트래픽 수집 관리 모듈은 C&C 기반 좀비리스트와 탐지시스템 요청 트래픽을 분류한다. C&C기반 좀비리스트는 동일 C&C를 기준으로 연결 요청한 좀비들의 IP 목록을 나타내고, 탐지시스템 요청 트래픽은 봇넷 구성 분석 모듈에 의해 요청된 신규 봇넷에 대한 행위 트래픽으로서 C&C 통신 프로토콜의 컨텐츠 정보를 나타낸다.

[0040] C&C 기반 좀비리스트는 봇넷 매칭을 위하여 일정 대기시간동안 동일 C&C를 기준으로 유입되는 좀비 IP 리스트들

을 추가적으로 열거한다(ST130). 설정된 대기시간 이후 열거된 IP 리스트 개수가 임계값을 초과하면 봇넷 매칭을 수행한다(ST140). 한편, IP 리스트 개수가 임계값에 미치지 못할 경우에도, 관제시스템의 공유정보에 의하여 업데이트되는 C&C서버 블랙리스트와 매칭하여 봇넷의 트래픽이라 판단될 때에도 다음 단계로 봇넷 매칭을 수행한다(ST150). 먼저, “접속 Dst IP/Port”와 “C&C서버 IP/Port”를 매칭하는데, C&C서버 IP/Port의 블랙리스트는 관제시스템의 공유정보에 의하여 업데이트되며, C&C서버 도메인의 블랙리스트 업데이트는 필요하지 않다.

[0041] 또한, “요청 도메인에 대한 응답 IP”와 “DNS 싱크홀 IP”를 매칭하는데, DNS 싱크홀 IP 리스트는 관제시스템의 공유정보가 아니고 봇넷 탐지 시스템이 사전에 보유하고 있어야 하는 고정된 리스트이다. 구체적으로, 1) 관제시스템에 의해 최종 판단된 C&C 서버의 도메인을 추출하여, 2) 타 ISP들과 정보를 공유하고, 3) 모든 ISP가 해당 도메인에 대한 DNS 싱크홀을 적용한 후, 4) 해당 도메인을 요청하는 트래픽을 수집(센서)하면, 5) 해당 요청 도메인의 응답 IP가 싱크홀 IP일 때 좀비의 행위로 판단한다. ST130 내지 ST150의 모든 경우에 있어 결국 봇넷 트래픽 여부를 판별할 수 없다면 해당 트래픽은 무효 처리한다. 상기와 같이 트래픽 정보 수집 센서에 의해 수집되는 여러 트래픽들의 규정된 데이터 포맷 형태를 유지하면서 동일 C&C서버를 기준으로 좀비 IP 리스트를 업데이트한다.

[0042] 2. 봇넷 매칭 프로시저

[0043] 상기한 바와 같이, 봇넷 매칭 모듈은 IP/Port 기반 트래픽에 대하여 중앙집중 서버 정보와 기존 C&C서버 정보를 비교하고, 접속 클라이언트 IP 리스트와 기존 좀비리스트를 비교하여, 수집된 트래픽에 대하여 신종 봇넷 구성, 기존 봇넷 확장(A 플래그), C&C서버 재접속 및 에그 다운로드(B 플래그), C&C서버 이주 및 기타 주요 봇넷 행위(C 플래그)들로 분류하여 각 메시지 큐에 저장한다(ST160 내지 ST180).

[0044] 3. VDNS 체크 프로시저

[0045] VDNS 체크 모듈은 Domain 기반 트래픽에 대하여 수집/분류 관리 정책에 의하여 VDNS 여부를 체크하여 새로운 도메인 쿼리 정보를 메시지 큐에 저장한다(ST190).

[0046] 도 4는 본 발명에 따른 트래픽 분류 모듈의 각 구성모듈과 봇넷 구성 분석 모듈 및 봇넷 행위 분석 모듈 간의 동작 시퀀스를 나타낸 도이다.

[0047] 트래픽 정보 수집 센서 및 트래픽 수집 관리 모듈에 의해 수집된 트래픽은 봇넷 매칭 모듈에 의해 신규 봇넷/기존 봇넷으로 분류되어 메시지 큐에 저장 및 봇넷 구성 분석 모듈에 전달된다. 한편, 수집된 트래픽이 도메인 요청인 경우 VDNS 체크 모듈에 의해 VDNS 체크를 수행하고, 신규 VDNS 요청 도메인인 경우 역시 메시지 큐에 저장 및 봇넷 구성 분석 모듈에 전달된다. 봇넷 구성 분석 모듈은 신규 트래픽(신규 봇넷) 및 신규 VDNS 요청 도메인에 대하여 행위 정보를 요청하기 위하여 샘플링된 신규 좀비 IP 리스트를 전송한다.

[0048] 트래픽 정보 수집 센서와 트래픽 수집 관리 모듈은 계속해서 트래픽을 수집하며, 이 중 기존 봇넷에 해당하는 트래픽은 봇넷 매칭 모듈에 의해 분류되어 메시지 큐에 저장되고 봇넷 구성 분석 모듈 및 봇넷 행위 분석 모듈에 전달된다. 또한, 수집된 트래픽이 도메인 요청이고 VDNS 체크 모듈에 의해 신규 VDNS 요청 도메인으로 판단된 경우 메시지 큐에 저장 및 봇넷 구성 분석 모듈 및 봇넷 행위 분석 모듈에 전달된다. 한편, 트래픽 정보 수집 센서와 트래픽 수집 관리 모듈에 의해 수집된 트래픽이 봇넷 구성 분석 모듈에 의해 요청된 트래픽, 즉 봇넷 행위 정보인 경우 봇넷 매칭 모듈 및 VDNS 체크 모듈을 바이패스하여 메시지 큐에 저장되고 IRC Channel 정보가 봇넷 구성 분석 모듈로 전달된다.

[0049] 도 5는 본 발명에 따른 봇넷 구성 분석 모듈의 구성도이다. 또한, 도 6은 본 발명에 따른 봇넷 구성 분석 모듈의 각 구성모듈의 동작 시퀀스를 나타낸 도이다. 도 5 및 도 6을 참조하여 본 발명에 따른 봇넷 구성 분석 모듈의 각 구성모듈의 동작에 대하여 설명하면 이하와 같다.

[0050] 임시 구성 로그는 트래픽 수집 관리 모듈로부터 분류된 트래픽 정보(Domain, Dst_IP/Port)를 수신하여 저장한다. 한편, 미분류 서버 접속 IP 리스트 로그는 C&C 분석 및 탐지 모듈에서 정상 트래픽으로 분류된 Domain 및 IP/Port 등을 저장하고 그 결과를 분석 결과 로그에 전송한다.

[0051] C&C 분석 및 탐지 모듈은 임시 구성 로그로부터 분석에 필요한 트래픽 정보를 주기적으로 읽어 Domain 및 Dst_IP/Port 별 유사도를 분석한다. 먼저, Domain 유사도 분석은, 주기적으로 임시 구성 로그로부터 Domain 정보를 읽어 각 Domain 별로 요청한 소스 IP들을 매트릭스에 기록한 후 특정 시간이 지난 후 매트릭스를 분석하여 유사도를 측정하여 좀비 IP 리스트를 생성한다. 다음으로, Dst_IP/Port 유사도 분석은, 주기적으로 임시 구성 로그로부터 Dst_IP/Port 정보를 읽어 각 IP/Port 조합과 매칭되는 패킷을 전송한 소스 IP들을 매트릭스에 기록

한 후 특정 시간이 지난 후 매트릭스를 분석하여 유사도를 측정하여 좀비 IP 리스트를 생성한다. 한편, 접속 프로토콜(커맨드) 분석은 트래픽 정보 수집 센서로부터 샘플링된 트래픽 정보를 받아 각 봇넷들의 접속 프로토콜(커맨드)를 분석하며, 이 과정에서 봇넷 행위로 탐지된 트래픽은 C&C 추출 모듈로 전송한다.

[0052] C&C 추출 모듈은 C&C 분석 및 탐지 모듈에서 탐지된 봇넷 트래픽을 전송받아 추출된 C&C를 저장하고, 분석을 마친 트래픽은 다시 좀비리스트 추출 모듈로 전송한다.

[0053] 좀비리스트 추출 모듈은 C&C 추출 모듈에서 봇넷 트래픽을 전송받아 봇넷으로 탐지된 IRC Channel에 접근하는 좀비리스트를 추출하여 분석결과 로그에 저장한다.

[0054] 구성 이벤트 트리거는 각 모듈에서 분석된 결과를 종합하여 로그 관리자에게 전송하고, 분석 결과에서 추후 정책에 사용될 트리거 메시지를 생성하여 이벤트 트리거에 전송한다.

[0055] 도 7은 트래픽 정보 수집 센서와 봇넷 탐지 시스템 및 관제 시스템 간의 송수신 데이터 흐름도이다.

[0056] 도 7에 도시한 바와 같이, 트래픽 정보 수집 센서로부터 전달되는 수집 트래픽 데이터는 중앙집중형 트래픽 데이터로 Class O-1(Domain 기반 트래픽)과 Class O-2(IP/Port 기반 트래픽)을 포함하고, 행위트래픽 데이터로 Class B-1 내지 6을 포함한다. 또한, 봇넷 탐지 시스템, 구체적으로 봇넷 구성 분석 모듈로부터 전달되는 행위 트래픽 수집 요청 데이터는 수집 대상 좀비리스트를 전달하는 Class R로 표현된다. 또한, 봇넷 탐지 시스템으로부터 관제 시스템으로 전달되는 탐지 결과 데이터는, 봇넷의 초기 구성단계 탐지 결과인 Class O-1, 봇넷의 초기 구성단계로 의심될만한 비정상 트래픽 탐지 결과인 Class O-2, 그리고 봇넷의 행위 분류 결과인 Class B-1 내지 8을 포함한다.

[0057] 도 8은 트래픽 수집 관리 모듈에 의해 수집된 트래픽 정보의 전송 데이터 포맷을 나타낸 도이다. 트래픽 수집 관리 모듈에 의하여 감지된 중앙집중형 트래픽의 기본적인 가공 데이터 포맷은 하나의 C&C 서버 정보를 기준으로 이에 접속하는 좀비들의 IP를 열거하는 형식으로 생성한다.

[0058] 먼저, Domain 기반 트래픽의 경우, 헤더에는 Time(트래픽 발생시간)이 기록되고, 중앙집중 서버(봇넷의 C&C 서버) 필드에는 C&C Domain(C&C 서버의 DNS 쿼리 도메인명)과 C&C IP(이때의 응답 IP)가 기록되며, 서버 접속 호스트(봇넷의 좀비리스트) 필드에는 Count(발견된 총 Src 개체수), Time Window(처음 좀비 발생부터 마지막 좀비 발생까지의 시간 구간) 및 좀비 IP 리스트(접속한 총 좀비들의 IP 리스트)가 기록된다. 이때, 중앙집중 서버 필드의 값은 C&C 서버와의 직접 통신 트래픽이 아닌 DNS서버와의 통신 트래픽으로부터 수집하며, 수집 대상 포트번호는 53번 포트이다.

[0059] 다음으로, IP/Port 기반 트래픽의 경우, 헤더에는 역시 Time(트래픽 발생시간)이 기록되고, 중앙집중 서버(봇넷의 C&C 서버) 필드에는 C&C IP(C&C 서버의 IP)와 C&C Port(접속 포트번호)가 기록되며, 서버 접속 호스트(봇넷의 좀비리스트) 필드에는 Count(발견된 총 Src 개체수), Time Window(처음 좀비 발생부터 마지막 좀비 발생까지의 시간 구간) 및 좀비 IP 리스트(접속한 총 좀비들의 IP 리스트)가 기록된다. 이때, 중앙집중 서버 필드의 값은 C&C 서버와의 직접 통신 트래픽으로부터 수집되며, 수집 대상 포트번호는 모든 포트가 된다.

[0060] 따라서, 트래픽 수집 관리 모듈은 기본 수집 대상 트래픽의 데이터 포맷으로, 요청 Domain 기반 분류 리스트(DNS 트래픽으로부터 추출되는 정보)로는 DNS서버 요청 도메인명(C&C Domain), DNS서버 응답 IP 주소(C&C IP), 열거된 총 좀비수(Count), 처음 좀비 발생부터 마지막 좀비 발생까지의 시간 구간(Time Window) 및 DNS서버에 동일한 도메인을 요청한 좀비들의 IP 리스트(좀비 IP 리스트)를 수집하고, Dst IP/Port 기반 분류 리스트(서버 접속 트래픽으로부터 추출되는 정보)로는 중앙집중 트래픽이 유발된 C&C서버의 IP 주소(C&C IP), 중앙집중 트래픽이 유발된 C&C서버의 접속 Port 번호(C&C Port), 열거된 총 좀비수(Count), 처음 좀비 발생부터 마지막 좀비 발생까지의 시간 구간(Time Window) 및 동일한 IP 주소와 Port 번호로 접속한 좀비 IP 리스트(좀비 IP 리스트)를 수집하게 된다.

[0061] 도 9는 봇넷 탐지 시스템에서 탐지된 봇넷 구성 정보의 전송 데이터 포맷을 나타낸 도이다. 본 발명에 따른 봇넷 탐지 시스템은 새로이 탐지된 봇넷에 대하여 도 9와 같은 포맷으로 가공하여 저장 또는 관제시스템에 전달한다. 도 9에서 탐지는 일정 수 이상의 좀비들에 의하여 발견된 트래픽을 나타내고, 비정상은 일정 수 이하의 클라이언트들에 의하여 발견된 중앙집중형 트래픽을 나타낸다.

[0062] 먼저, 탐지된 봇넷 구성 정보(Class O-1)는 헤더에 local ID(탐지시스템 인식용 지역적인 봇넷 ID 번호)와 Time(최초 구성 트래픽 발생 시간)을 포함하고, C&C 서버 필드에 Type(봇넷 구성 프로토콜(IRC)), Domain(DNS 쿼리 발생시 C&C서버 도메인명), IP(C&C서버 IP), Port(C&C서버 포트번호) 및 Locator(C&C 서버 내 접속 위치

표시자(Channel 이름))를 포함하고, 좀비리스트 필드에 Count(접속한 총 좀비 개체수)와 Zombie IP List(접속한 모든 좀비들의 IP 리스트)를 포함하고, 분석결과 필드에 Similarity(봇넷의 구성 탐지를 위한 척도로써 이용되는 유사도 분석값)을 포함한다.

[0063] 또한, 비정상 봇넷 구성 정보(Class 0-2)는 기본적으로 Class 0-1 데이터 포맷과 동일하며, 다만 봇넷 ID 번호는 불필요하므로 local ID 필드가 제외된다. 이는 비정상 봇넷 구성 정보의 경우 행위 트래픽 정보와 동기화 될 필요가 없기 때문이다.

[0064] 본 발명의 바람직한 실시예가 특정 용어들을 사용하여 기술되어 왔지만, 그러한 기술은 오로지 설명을 하기 위한 것이며, 다음의 청구범위의 기술적 사상 및 범위로부터 이탈되지 않고서 여러가지 변경 및 변화가 가해질 수 있는 것으로 이해되어져야 한다.

도면의 간단한 설명

[0065] 도 1은 본 발명에 따른 봇넷 탐지 시스템의 구성도이다.

[0066] 도 2는 본 발명에 따른 트래픽 분류 모듈의 구성도이다.

[0067] 도 3은 본 발명에 따른 트래픽 분류 모듈의 트래픽 분류 과정을 도시한 흐름도이다.

[0068] 도 4는 본 발명에 따른 트래픽 분류 모듈의 각 구성모듈과 봇넷 구성 분석 모듈 및 봇넷 행위 분석 모듈 간의 동작 시퀀스를 나타낸 도이다.

[0069] 도 5는 본 발명에 따른 봇넷 구성 분석 모듈의 구성도이다.

[0070] 도 6은 본 발명에 따른 봇넷 구성 분석 모듈의 각 구성모듈의 동작 시퀀스를 나타낸 도이다.

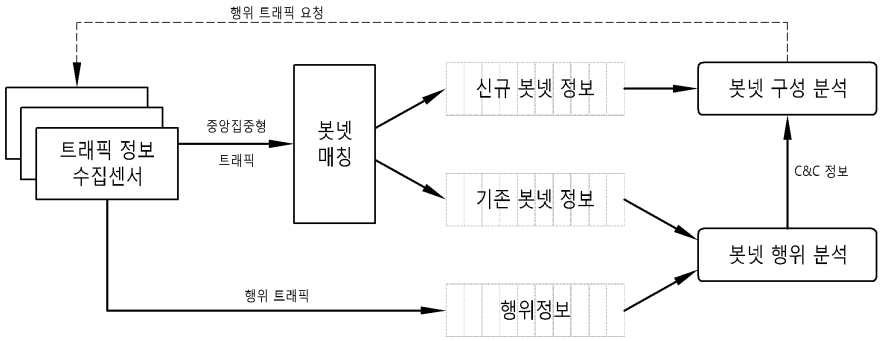
[0071] 도 7은 트래픽 정보 수집 센서와 봇넷 탐지 시스템 및 관제 시스템 간의 송수신 데이터 흐름도이다.

[0072] 도 8은 트래픽 수집 관리 모듈에 의해 수집된 트래픽 정보의 전송 데이터 포맷을 나타낸 도이다.

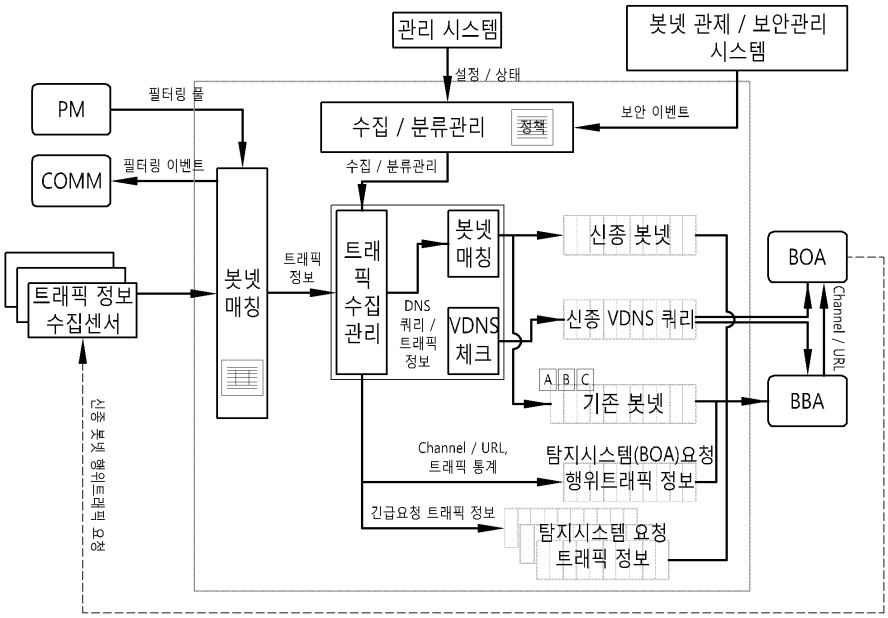
[0073] 도 9는 봇넷 탐지 시스템에서 탐지된 봇넷 구성 정보의 전송 데이터 포맷을 나타낸 도이다.

도면

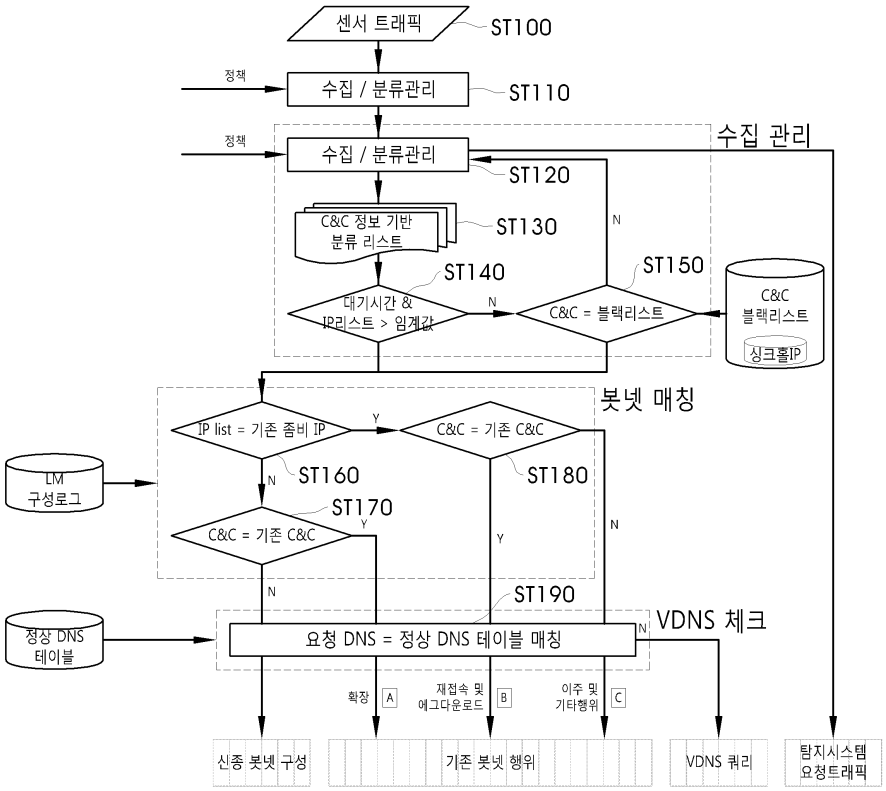
도면1



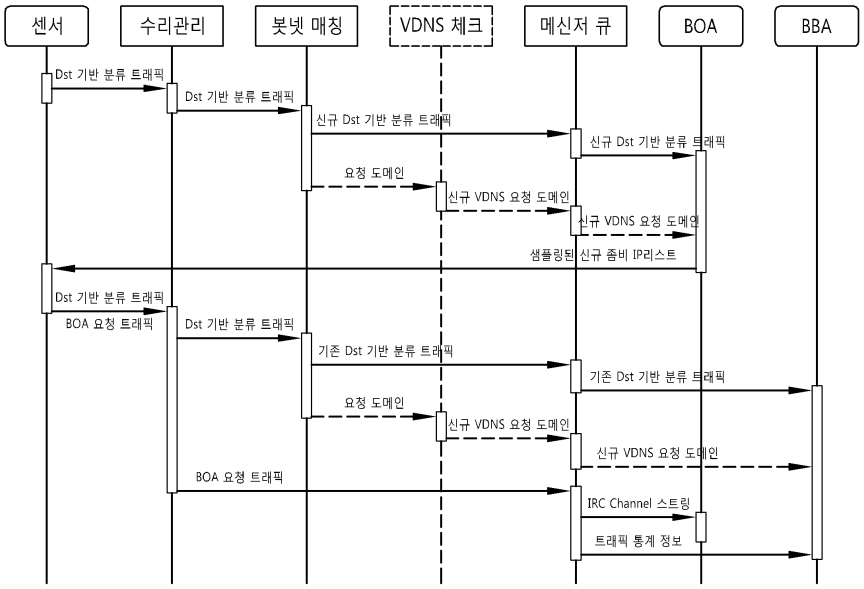
도면2


관리 시스템
봇넷 관제 / 보안관리 시스템
설정 / 상태
PM
필터링 룰
수집 / 분류관리
정책
보안 이벤트
COMM
필터링 이벤트
수집 / 분류관리
트래픽 정보 수집센서
봇넷 매칭
트래픽 정보
트래픽 수집 관리
봇넷 매칭
VDNS 체크
DNS 쿼리 / 트래픽 정보
신종 봇넷
신종 VDNS 쿼리
기존 봇넷
A B C
BOA
BBA
Channel / URL
신종 봇넷 행위트래픽 요청
Channel / URL, 트래픽 통계
탐지시스템(BOA)요청 행위트래픽 정보
긴급요청 트래픽 정보
탐지시스템 요청 트래픽 정보

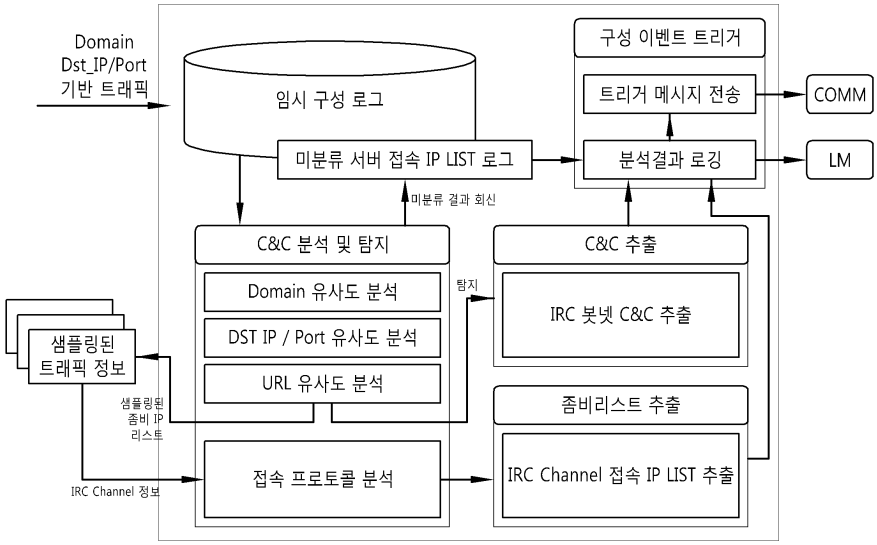
도면3


센서 트래픽
ST100
정책
수집 / 분류관리
ST110
정책
수집 / 분류관리
수집 관리
ST120
C&C 정보 기반 분류 리스트
ST130
N
ST140
ST150
대기시간 & IP리스트 > 임계값
N
C&C = 블랙리스트
C&C 블랙리스트
싱크홀IP
봇넷 매칭
IP list = 기존 좀비 IP
Y
C&C = 기존 C&C
LM 구성로그
N
ST160
ST180
ST170
C&C = 기존 C&C
Y
Y
N
N
ST190
VDNS 체크
정상 DNS 테이블
요청 DNS = 정상 DNS 테이블 매칭
N
확장
A
재접속 및 에그다운로드
B
이주 및 기타행위
C
신종 봇넷 구성
기존 봇넷 행위
VDNS 쿼리
탐지시스템 요청트래픽

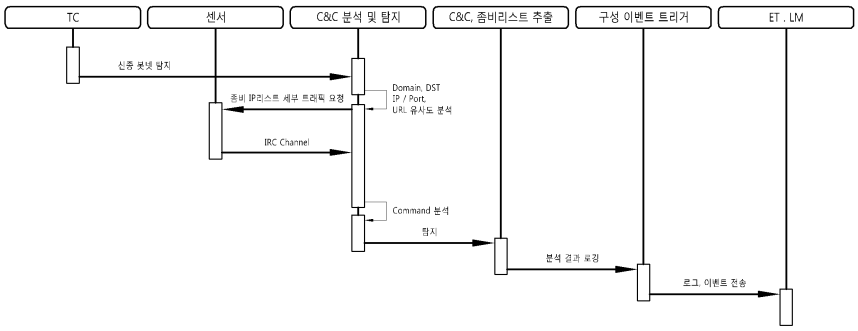
도면4


센서
수리관리
봇넷 매칭
VDNS 체크
메신저 큐
BOA
BBA
Dst 기반 분류 트래픽
Dst 기반 분류 트래픽
신규 Dst 기반 분류 트래픽
신규 Dst 기반 분류 트래픽
요청 도메인
신규 VDNS 요청 도메인
신규 VDNS 요청 도메인
샘플링된 신규 좀비 IP리스트
Dst 기반 분류 트래픽
BOA 요청 트래픽
Dst 기반 분류 트래픽
기존 Dst 기반 분류 트래픽
기존 Dst 기반 분류 트래픽
요청 도메인
신규 VDNS 요청 도메인
신규 VDNS 요청 도메인
BOA 요청 트래픽
IRC Channel 스트링
트래픽 통계 정보

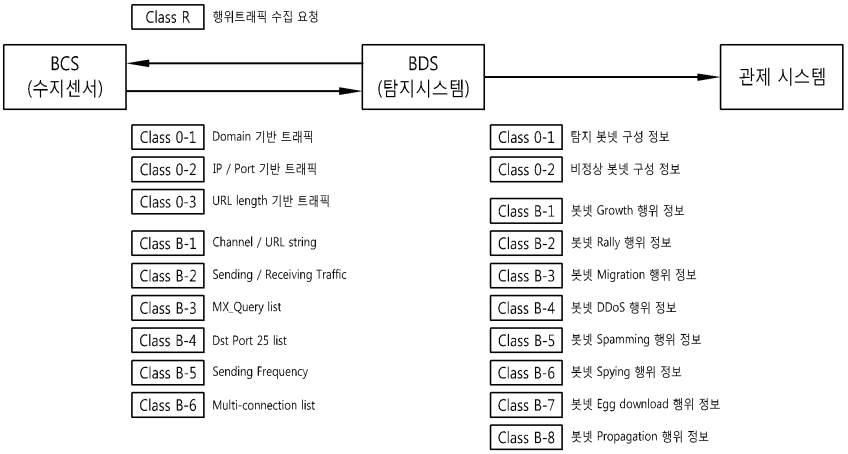
도면5


Domain
Dst_IP/Port
기반 트래픽
임시 구성 로그
미분류 서버 접속 IP LIST 로그
구성 이벤트 트리거
트리거 메시지 전송
분석결과 로깅
COMM
LM
미분류 결과 회신
C&C 분석 및 탐지
Domain 유사도 분석
DST IP / Port 유사도 분석
URL 유사도 분석
탐지
C&C 추출
IRC 봇넷 C&C 추출
좀비리스트 추출
IRC Channel 접속 IP LIST 추출
샘플링된
트래픽 정보
샘플링된
좀비 IP
리스트
IRC Channel 정보
접속 프로토콜 분석

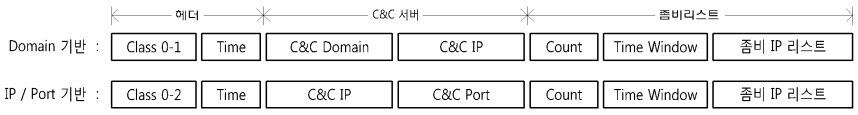
도면6



도면7



도면8



도면9