(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2016년10월12일
(11) 등록번호 10-1665369
(24) 등록일자 2016년10월06일

(51) 국제특허분류(Int. Cl.)
H04L 29/06 (2006.01)
(52) CPC특허분류
H04L 63/1425 (2013.01)
H04L 2463/144 (2013.01)
(21) 출원번호 10-2015-0081869
(22) 출원일자 2015년06월10일
심사청구일자 2015년06월10일
(56) 선행기술조사문헌
JP2005151289 A
KR1020120068612 A

(73) 특허권자
고려대학교 산학협력단
서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)
(72) 발명자
권종훈
서울특별시 성북구 안암로 145, 우정정보통신관 307B (안암동5가, 고려대학교)
이희조
서울특별시 마포구 새터산14길 56, 2층 (성산동)
(뒷면에 계속)
(74) 대리인
특허법인엠에이피에스

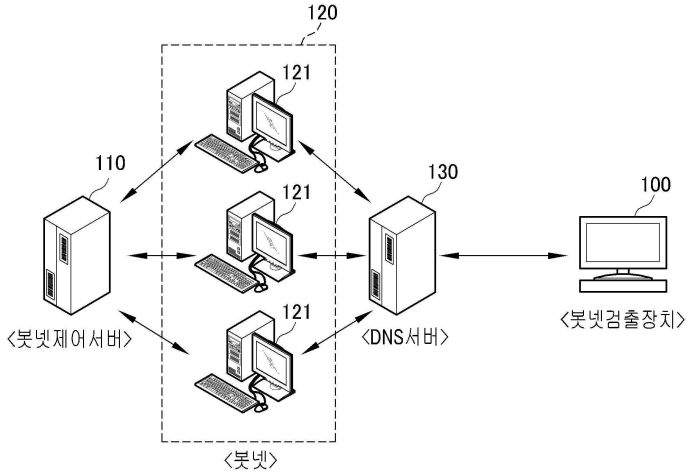
전체 청구항 수 : 총 14 항 심사관 : 문형섭

(54) 발명의 명칭 **DNS 트래픽을 통하여 봇넷을 검출하는 방법, 장치 및 컴퓨터 판독가능 기록매체**

(57) 요 약

본 발명의 일 실시예에 따르면, 봇넷검출장치에 의해 수행되는 DNS 트래픽을 통하여 봇넷을 검출하는 방법은, (a) DNS(Domain Name System) 서버로 수신되는 패킷을 분석하여, 상기 DNS 서버로 질의한 각 호스트 별로 질의 요청주기를 측정하는 단계; (b) 상기 질의요청주기를 기초로 주파수 신호를 생성하는 단계; 및 (c) 상기 주파수 신호에 주기적 질의 성분이 포함되어 있는 경우, 상기 주기적 질의 성분의 전력값이 미리 설정된 임계값을 넘는다면, 상기 주기적 질의 성분을 포함하는 주파수 신호에 대응하는 호스트를 봇의심객체로 판단하는 단계를 포함한다.

대 표 도 - 도1



(72) 발명자

김정식

충청북도 청주시 상당구 영운천로129번길 31, 3층 (용정동)

이제현

서울특별시 성북구 종암로28길 72-5, 5층 501호 (종암동)

공지예외적용 : 있음

명 세 서

청구범위

청구항 1

봇넷검출장치에 의해 수행되는 DNS 트래픽을 통하여 봇넷을 검출하는 방법에 있어서,

(a) DNS(Domain Name System) 서버로 수신되는 패킷을 분석하여, 상기 DNS 서버로 질의한 각 호스트 별로 질의요청주기를 측정하는 단계;

(b) 상기 질의요청주기를 기초로 주파수 신호를 생성하는 단계;

(c) 상기 주파수 신호에 주기적 질의 성분이 포함되어 있는 경우, 상기 주기적 질의 성분의 전력값이 미리 설정된 임계값을 넘는다면, 상기 주기적 질의 성분을 포함하는 주파수 신호에 대응하는 호스트를 봇의심객체로 판단하는 단계; 및

(d) 상기 (c) 단계를 통하여 복수의 봇의심객체를 검출한 경우, 상기 복수의 봇의심객체의 주파수 신호 간의 전력값의 크기 차이를 산출하고, 상기 전력값의 크기 차이가 미리 설정된 값 이하인 경우, 동일한 그룹으로 분류하여, 봇넷을 검출하는 단계;를 포함하며,

상기 봇넷은 악성코드에 감염된 다수의 봇이 네트워크로 연결되어 있는 집합체를 나타내는 것인, DNS 트래픽을 통하여 봇넷을 검출하는 방법.

청구항 2

제 1 항에 있어서,

상기 주기적 질의 성분은,

상기 주파수 신호에 포함된 타 대역 주파수의 전력값보다 미리 설정된 크기 이상으로 높은 전력값을 갖는 피크인, DNS 트래픽을 통하여 봇넷을 검출하는 방법.

청구항 3

제 1 항에 있어서,

상기 (a) 단계는,

서로 상이한 시간대 별로 상기 각 호스트마다 질의요청주기를 측정하는 단계를 포함하는, DNS 트래픽을 통하여 봇넷을 검출하는 방법.

청구항 4

제 1 항에 있어서,

상기 (a) 단계는,

각 호스트 별로 일정시간동안의 단위시간당 질의요청횟수를 측정하는 단계를 포함하는, DNS 트래픽을 통하여 봇넷을 검출하는 방법.

청구항 5

제 4 항에 있어서,

상기 (b) 단계는,

상기 일정시간동안의 단위시간당 질의요청횟수를 기초로 시간 별 질의요청횟수에 관한 질의요청신호를 생성하는 단계; 및

상기 질의요청신호를 상기 주파수 신호로 변환하는 단계를 포함하는, DNS 트래픽을 통하여 봇넷을 검출하는 방

법.

청구항 6

제 1 항에 있어서,

상기 (c) 단계는,

상기 주파수 신호에 포함된 적어도 하나의 주기적 질의 성분 중 전력값이 가장 높은 주기적 질의 성분과 상기 임계값과의 비교를 통하여 상기 봇의심객체를 판단하는 단계를 포함하는, DNS 트래픽을 통하여 봇넷을 검출하는 방법.

청구항 7

삭제

청구항 8

제 1 항에 있어서,

상기 (d) 단계는,

상기 분류된 각 그룹에 포함된 봇의심객체의 도메인 주소와 상기 봇넷검출장치에 미리 저장된 봇넷도메인주소를 비교하여 봇넷을 검출하는 단계;를 포함하는, DNS 트래픽을 통하여 봇넷을 검출하는 방법.

청구항 9

DNS(Domain Name System) 트래픽을 통한 봇넷검출장치에 있어서, 상기 봇넷검출장치는,

DNS 트래픽을 통하여 봇넷을 검출하기 위한 프로그램이 저장된 메모리; 및

상기 프로그램을 실행하는 프로세서;를 포함하며,

상기 프로세서는, 상기 프로그램의 실행에 따라,

DNS 서버로 수신되는 패킷을 분석하여, 상기 DNS 서버로 질의한 각 호스트 별로 질의요청주기를 측정하고,

상기 질의요청주기를 기초로 주파수 신호를 생성하고,

상기 주파수 신호에 주기적 질의 성분이 포함되어 있는 경우, 상기 주기적 질의 성분의 전력값이 미리 설정된 임계값을 넘는다면, 상기 주기적 질의 성분을 포함하는 주파수 신호에 대응하는 호스트를 봇의심객체로 판단하며,

복수의 봇의심객체를 검출한 경우, 상기 복수의 봇의심객체의 주파수 신호 간의 전력값의 크기 차이를 산출하고, 상기 전력값의 크기 차이가 미리 설정된 값 이하인 경우, 동일한 그룹으로 분류하여, 봇넷을 검출하며,

상기 봇넷은 악성코드에 감염된 다수의 봇이 네트워크로 연결되어 있는 집합체를 나타내는 것인, DNS 트래픽을 통한 봇넷검출장치.

청구항 10

제 9항에 있어서,

상기 주기적 질의 성분은,

상기 주파수 신호에 포함된 타 대역 주파수의 전력값보다 미리 설정된 크기 이상으로 높은 전력값을 갖는 피크인, DNS 트래픽을 통한 봇넷검출장치.

청구항 11

제 9 항에 있어서,

상기 프로세서는,

서로 상이한 시간대 별로 상기 각 호스트마다 질의요청주기를 측정하는, DNS 트래픽을 통한 봇넷검출장치.

청구항 12

제 9 항에 있어서,

상기 프로세서는,

각 호스트 별로 일정시간동안의 단위시간당 질의요청횟수를 측정하는, DNS 트래픽을 통한 봇넷검출장치.

청구항 13

제12 항에 있어서,

상기 프로세서는,

상기 일정시간동안의 단위시간당 질의요청횟수를 기초로 시간 별 질의요청횟수에 관한 질의요청신호를 생성하고,

상기 질의요청신호를 상기 주파수 신호로 변환하는, DNS 트래픽을 통한 봇넷검출장치.

청구항 14

제 9 항에 있어서,

상기 프로세서는,

상기 주파수 신호에 포함된 적어도 하나의 주기적 질의 성분 중 전력값이 가장 높은 주기적 질의 성분과 상기 임계값과의 비교를 통하여 상기 봇의심객체를 판단하는, DNS 트래픽을 통한 봇넷검출장치.

청구항 15

삭제

청구항 16

제9 항에 있어서,

상기 프로세서는,

상기 분류된 각 그룹에 포함된 봇의심객체의 도메인 주소와 상기 봇넷검출장치에 미리 저장된 봇넷도메인주소를 비교하여 봇넷을 검출하는, DNS 트래픽을 통한 봇넷검출장치.

발명의 설명

기 술 분 야

[0001] 본 발명은 DNS 트래픽을 통하여 봇넷을 검출하는 방법, 장치 및 컴퓨터 판독가능 기록매체에 관한 것으로서, 보다 상세하게는, 점차 지능화되는 악성코드에 능동적으로 대응하기 위하여 DNS 트래픽을 기초로 효율적으로 봇넷을 검출하기 위한 DNS 트래픽을 통하여 봇넷을 검출하는 방법, 장치 및 컴퓨터 판독가능 기록매체 에 관한 것이다.

배 경 기 술

[0002] 컴퓨터 시스템과 인터넷 사용자의 증가로 컴퓨터 네트워크 시스템은 양적, 질적으로 폭발적인 발전을 보이고 있다. 특히 사회 시스템 전반에 걸쳐 컴퓨터를 활용한 자동화, 지능화가 이루어지고 있고, 모든 정보가 네트워크를 통해 전달되면서 인터넷 환경의 중요성은 점차 높아지고 있다.

[0003] 이러한 현상에 대응하여, 컴퓨터 시스템을 감염시키는 악성코드를 통해 부당한 이득을 추구하는 공격 역시 급속도로 증가하고 있다. 악성코드가 네트워크로 연결되어 공격의 파급력을 극대화한 악성코드 집단을 봇넷(Botnet)이라고 하며, 최근의 악성코드는 분산 서비스 공격, 개인정보 탈취, 대용량 스팸메일 발송, 피싱을 수행하는 등 공격 형태 또한 다양한 양상을 지니고 있다.

[0004] 특히 지능화된 악성코드들은 기존의 탐지 기술을 회피하기 위해 코드 난독화, 통신 암호화 및 행위 패턴 변조 등 다양한 회피 기법들을 적용하고 있어, 악성코드 탐지 및 대응이 점차 어려워지고 있다. 또한, 악성코드 감염 트래픽을 발생시키는 사용자 단말과 공격을 수행하는 봇(bot) 간의 구분이 용이하지 않아 봇넷 검출이 어렵다는 문제점이 있다. 또한 급속도로 증가하는 네트워크 트래픽 양이 매우 방대해짐에 따라 악성코드의 악성행위를 검출하여 차단하는 종래의 기술들은 악성코드를 탐지하는 데에 많은 어려움을 겪고 있다.

선행기술문헌

[0005] 한국공개특허공보 제10-2011-0070182호 (공개일 : 2011.06.24)

발명의 내용

해결하려는 과제

[0006] 본 발명은 점차 지능화되는 악성코드에 대하여 능동적이고 본질적으로 대처하기 위하여, 봇넷이 DNS 서버로 주기적으로 질의를 수행하는 점에 착안함에 따라, 주기적인 DNS 트래픽을 발생시키는 호스트를 감지하여 봇넷을 검출하는 방법 및 장치를 제공하는 데에 목적이 있다.

과제의 해결 수단

[0007] 상술한 기술적 과제를 달성하기 위한 기술적 수단으로서, 본 발명의 제 1 측면에 따르면, 봇넷검출장치에 의해 수행되는 DNS 트래픽을 통하여 봇넷을 검출하는 방법은, (a) DNS(Domain Name System) 서버로 수신되는 패킷을 분석하여, 상기 DNS 서버로 질의한 각 호스트 별로 질의요청주기를 측정하는 단계; (b) 상기 질의요청주기를 기초로 주파수 신호를 생성하는 단계; 및 (c) 상기 주파수 신호에 주기적 질의 성분이 포함되어 있는 경우, 상기 주기적 질의 성분의 전력값이 미리 설정된 임계값을 넘는다면, 상기 주기적 질의 성분을 포함하는 주파수 신호에 대응하는 호스트를 봇의심객체로 판단하는 단계를 포함한다.

[0008] 또한, 본 발명의 제 2 측면에 따르면, DNS(Domain Name System) 트래픽을 통한 봇넷검출장치는, DNS 트래픽을 통하여 봇넷을 검출하기 위한 프로그램이 저장된 메모리; 및 상기 프로그램을 실행하는 프로세서;를 포함하며, 상기 프로세서는, 상기 프로그램의 실행에 따라, DNS 서버로 수신되는 패킷을 분석하여, 상기 DNS 서버로 질의한 각 호스트 별로 질의요청주기를 측정하고, 상기 질의요청주기를 기초로 주파수 신호를 생성하고, 상기 주파수 신호에 주기적 질의 성분이 포함되어 있는 경우, 상기 주기적 질의 성분의 전력값이 미리 설정된 임계값을 넘는다면, 상기 주기적 질의 성분을 포함하는 주파수 신호에 대응하는 호스트를 봇의심객체로 판단한다.

발명의 효과

[0009] 본 발명은 DNS 질의 패턴을 이용하여 봇넷을 검출하므로 종래기술보다 더욱 정확하고 효율적으로 봇넷을 검출할 수 있다. DNS 질의 패턴은 봇넷이 임의로 변경할 수 없는 근본적인 행위 특징이다. 따라서, 더욱 지능화된 악성코드가 나타난다고 하더라도 규칙적인 DNS 질의 패턴을 가지게 되므로, 이를 이용하여 봇넷을 검출하는 본 발명의 일 실시예는 종래의 봇넷 탐지 기법에 비하여 더욱 근본적인 악성코드 탐지 및 대응의 효과를 가질 수 있다. 또한, DNS 질의 주기성을 분석함으로써, 악성코드에 감염된 컴퓨터의 사용자가 발생시키는 트래픽 여부와 상관없이 안정적인 봇넷 행위 분석이 가능하다. 또한, 시간대 별로 DNS 질의 주기성을 분석하기 때문에, 봇넷의 활동시간이 변경되더라도 어느 한 시간 대에서 주기적 질의 특성이 검출되기만 한다면 봇넷으로 검출될 수 있으므로, 효과적인 봇넷 검출이 가능하다.

도면의 간단한 설명

[0010] 도 1은 본 발명의 일 실시예에 따르는 DNS 트래픽을 통하여 봇넷을 검출하는 장치 및 주변 객체에 관한 구조도이다.

도 2는 본 발명의 일 실시예에 따르는 봇넷그룹 검출장치의 구성에 대한 블록도이다.

도 3a는 본 발명의 일 실시예에 따라 생성되는 질의요청신호의 파형에 대한 그래프이다.

도 3b는 본 발명의 일 실시예에 따라 질의요청신호를 변환하여 생성되는 주파수 신호의 파형에 대한 그래프이다.

도 4a는 본 발명의 일 실시예에 따르는 주기적 질의요청신호 및 이를 변환한 주파수 신호를 나타내는 그래프이

다.

도 4b는 본 발명의 일 실시예에 따르는 비주기적 질의요청신호 및 이를 변환한 주파수 신호를 나타내는 그래프이다.

도 5는 본 발명의 일 실시예에 따라, 서로 상이한 시간대 별로 DNS 트래픽을 감지하고 주기성을 검사하는 과정을 개념적으로 나타낸 도식도이다.

도 6은 본 발명의 일 실시예에 따라, DNS 트래픽을 통하여 봇넷을 검출하는 방법을 설명하기 위한 순서도이다.

발명을 실시하기 위한 구체적인 내용

[0011] 아래에서는 첨부한 도면을 참조하여 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있도록 본 발명의 실시예를 상세히 설명한다. 그러나 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며 여기에서 설명하는 실시예에 한정되지 않는다. 그리고 도면에서 본 발명을 명확하게 설명하기 위해서 설명과 관계없는 부분은 생략하였으며, 명세서 전체를 통하여 유사한 부분에 대해서는 유사한 도면 부호를 붙였다.

[0012] 명세서 전체에서, 어떤 부분이 다른 부분과 "연결"되어 있다고 할 때, 이는 "직접적으로 연결"되어 있는 경우뿐 아니라, 그 중간에 다른 소자를 사이에 두고 "전기적으로 연결"되어 있는 경우도 포함한다. 또한 어떤 부분이 어떤 구성요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있는 것을 의미한다.

[0013] 본 명세서에 있어서 '부(部)'란, 하드웨어에 의해 실현되는 유닛(unit), 소프트웨어에 의해 실현되는 유닛, 양방을 이용하여 실현되는 유닛을 포함한다. 또한, 1 개의 유닛이 2 개 이상의 하드웨어를 이용하여 실현되어도 되고, 2 개 이상의 유닛이 1 개의 하드웨어에 의해 실현되어도 된다. 한편, '~부'는 소프트웨어 또는 하드웨어에 한정되는 의미는 아니며, '~부'는 어드레싱 할 수 있는 저장 매체에 있도록 구성될 수도 있고 하나 또는 그 이상의 프로세서들을 재생시키도록 구성될 수도 있다. 따라서, 일 예로서 '~부'는 소프트웨어 구성요소들, 객체지향 소프트웨어 구성요소들, 클래스 구성요소들 및 태스크 구성요소들과 같은 구성요소들과, 프로세스들, 함수들, 속성들, 프로시저들, 서브루틴들, 프로그램 코드의 세그먼트들, 드라이버들, 펌웨어, 마이크로코드, 회로, 데이터, 데이터베이스, 데이터 구조들, 테이블들, 어레이들 및 변수들을 포함한다. 구성요소들과 '~부'들 안에서 제공되는 기능은 더 작은 수의 구성요소들 및 '~부'들로 결합되거나 추가적인 구성요소들과 '~부'들로 더 분리될 수 있다. 뿐만 아니라, 구성요소들 및 '~부'들은 디바이스 또는 보안 멀티미디어카드 내의 하나 또는 그 이상의 CPU들을 재생시키도록 구현될 수도 있다.

[0014] 이하에서 언급되는 "사용자 단말"은 네트워크를 통해 서버나 타 단말에 접속할 수 있는 컴퓨터나 휴대용 단말기로 구현될 수 있다. 여기서, 컴퓨터는 예를 들어, 웹 브라우저(WEB Browser)가 탑재된 노트북, 데스크톱(desktop), 랩톱(laptop) 등을 포함하고, 휴대용 단말기는 예를 들어, 휴대성과 이동성이 보장되는 무선 통신 장치로서, PCS(Personal Communication System), GSM(Global System for Mobile communications), PDC(Personal Digital Cellular), PHS(Personal Handyphone System), PDA(Personal Digital Assistant), IMT(International Mobile Telecommunication)-2000, CDMA(Code Division Multiple Access)-2000, W-CDMA(W-Code Division Multiple Access), Wibro(Wireless Broadband Internet) 단말 등과 같은 모든 종류의 핸드헬드(Handheld) 기반의 무선 통신 장치를 포함할 수 있다. 또한, "네트워크"는 근거리 통신망(Local Area Network; LAN), 광역 통신망(Wide Area Network; WAN) 또는 부가가치 통신망(Value Added Network; VAN) 등과 같은 유선 네트워크나 이동 통신망(mobile radio communication network) 또는 위성 통신망 등과 같은 모든 종류의 무선 네트워크로 구현될 수 있다.

[0015] 또한, "봇(Bot)"은 로봇(robot)의 줄임말로서 악의적 의도를 가진 소프트웨어에 감염된 개인용 컴퓨터(Personal Computer, PC)를 의미한다. 그리고 "봇넷(Botnet)"이란 악성코드에 감염된 다수의 컴퓨터들 또는 복수의 봇이 네트워크로 연결되어 있는 집합을 의미한다. 봇넷은 봇 마스터(Bot Master)에 의해 원격 조종되어 디도스(DDoS) 공격, 개인정보 수집, 피싱, 악성코드배포, 스팸메일 발송 등 다양한 악성행위에 이용되고 있다. 이러한 봇넷은 봇넷이 사용하는 프로토콜에 따라 분류될 수 있다. "봇넷 그룹"이란 적어도 하나의 봇으로 구성된 그룹을 의미한다.

[0016] 이하, 도면을 참고하여, 본 발명의 일 실시예에 대하여 구체적으로 설명하도록 한다.

[0017] 도 1을 참고하면, 봇넷(120)은 적어도 하나의 봇(121)으로 구성되며, 봇넷제어서버(110)의 명령에 의해 제어된다. 또한, 일반적으로 봇넷(120)은 하나의 도메인 네임을 통하여 공격대상 타겟에 접근할 수 있다. 봇넷제어서

버(110)는 봇마스터(미도시)에 의해 제어되며, 봇마스터에 의해 설계된 프로그램대로 봇넷(120)을 제어할 수 있다. 한편, 봇넷(120)은 봇넷제어서버(100)나 공격대상 타겟의 IP 주소를 찾기 위하여 주기적으로 DNS(Domain Name System) 서버로 DNS 질의를 전송하는 특성을 갖는다. 이에 따라, DNS 서버(130)는 DNS 질의에 따라 문자열로 구성된 도메인 네임을 인터넷 상의 컴퓨터에 할당된 IP 주소로 변환하여 IP 주소를 봇넷(120)으로 제공한다.

[0018] 한편, 네트워크에 연결된 일반 사용자들이 사용하는 컴퓨터는 봇넷(120)과 달리 DNS 서버(130)로 비주기적인 질의를 전송한다. 일반 사용자들은 컴퓨터에 접속하여 인터넷을 이용하는 때에만 컴퓨터를 통하여 질의를 전송하므로, 미리 설정된 프로그램대로 질의를 전송하는 봇넷(120)과 달리 비주기적인 질의 특성을 가질 수 밖에 없다.

[0019] 이러한 점에 착안하여 본 발명의 일 실시예에 따르는 봇넷검출장치(100)는 DNS 서버(130)에 대한 DNS 트래픽을 수집 및 분석하여, 주기적 질의 특성을 갖는 트래픽을 추출하고, 해당 트래픽을 발생시킨 호스트들을 봇넷(120)으로서 검출한다.

[0020] 이하, 도 2를 참조하여, 본 발명의 일 실시예에 따르는 봇넷검출장치(100)의 구성에 대하여 구체적으로 설명한다.

[0021] 봇넷검출장치(100)는 서버 또는 PC(Personal Computer)의 형태로 구현될 수 있다. 봇넷검출장치(100)는DNS 트래픽을 통하여 봇넷(120)을 검출하는 프로그램(또는 애플리케이션)이 저장된 메모리와 위 프로그램을 실행하는 프로세서를 포함하여 구성될 수 있다. 여기서 프로세서는 메모리에 저장된 프로그램의 실행에 따라 다양한 기능을 수행할 수 있는데, 각 기능에 따라 프로세서에 포함되는 세부 구성요소들을 DNS 트래픽 수집부(210), 질의주기성 분석부(220), 봇의심객체 검출부(230), 유사그룹 분류부(240), 봇넷 검출부(250)로 나타낼 수 있다.

[0022] DNS 트래픽 수집부(210)는 네트워크를 통하여 DNS 서버(130)와 연결되어, DNS 서버(130)에 대하여 송신 또는 수신되는 패킷에 대한 데이터를 수집한다. 구체적으로, DNS 트래픽 수집부(210)는 DNS 서버(130)로 질의를 요청한 각 호스트의 IP 주소, 도메인 네임, 타임 스탬프 등을 수집할 수 있다. 또한, DNS 트래픽 수집부(210)는 일정시간 동안의 단위시간당 질의요청횟수에 관한 정보인 질의요청주기를 각 호스트 별로 수집할 수 있다. 여기서, 봇넷(120)에 포함된 호스트는 단위시간당 일정한 횟수로 질의를 요청할 것이며, 일반 사용자 단말의 경우 단위시간당 불규칙적인 횟수로 질의를 요청할 것이다.

[0023] 질의주기성 분석부(220)는 각 호스트 별로 질의주기성을 분석하기 위하여 주파수 신호를 생성한다. 구체적으로, 질의주기성 분석부(220)는 일정시간 동안의 단위시간당 질의요청횟수를 기초로 도 3a와 같이 타임 도메인 상의 질의요청신호를 생성한다. 그리고 질의주기성 분석부(220)는 질의요청신호를 도 3b와 같은 주파수 신호로 변환할 수 있다. 이때, 질의주기성 분석부(220)는 PSD(Power Spectral Density) 기법을 이용하여 질의요청신호를 주파수 신호로 변환할 수 있다. PSD 는 수학식 1을 통하여 정의될 수 있다.

[0024] [수학식 1]

$$P_{xx}(\omega) = \frac{(\Delta t)^2}{T}\left|\sum_{n=1}^{N} f_n e^{-i\omega n}\right|^2$$

[0025]

[0026] T는 DNS 질의가 측정된 총 시간을 의미하며, N은 총 시간을 세그먼트한 개수를 의미하고, Δt는 T/N이고, f(n)은 질의요청신호의 각 세그먼트에서의 데이터값(즉, 질의요청횟수)을 의미하며, Pxx(ω)는 질의요청신호를 주파수 대역으로 변환한 후의 주파수 신호를 의미한다. 즉, 질의주기성 분석부(220)는 질의요청신호를 시간 기준으로 소정의 개수만큼 세그먼트한 후, 푸리에 변환을 통하여 주파수 신호를 획득할 수 있다.

[0027] 이 경우, 도 3a에 나타나는 주기적 질의는 도 3b에서 가장 높은 피크인 주기적 질의 성분으로 나타나게 되며, 도 3a의 비주기적 질의는 도 3b에서 주기적 질의 성분 보다 전력값이 낮은 성분으로 나타날 수 있다. 도 3a 및 도 3b를 통하여 확인할 수 있듯이, 주기적 질의 성분은 주파수 신호에 포함된 타 대역 주파수의 전력값보다 미리 설정된 크기 이상으로 높은 전력값을 갖는 피크(peak)를 의미하며, 비주기적 질의 성분은 주기적 질의 성분을 제외한 성분을 의미한다. 타임 도메인 상에서 주기적 특성을 갖는 질의는 일정한 빈도(주파수)로 발현되기 때문에, 주파수 도메인 상에서 비주기적 질의에 비하여 높은 전력값을 갖는 피크로 나타나게 된다.

[0028] 구체적으로, 도 4a를 참조하면, 일정한 빈도로 질의한 호스트의 질의요청신호를 주파수 신호로 변환할 경우, 일부 개수의 주기적 질의 성분을 포함하는 주파수 신호가 나타날 수 있다. 반면, 도 4b를 참조하면, 불규칙적인

빈도로 질의한 호스트의 질의요청신호를 주파수 신호로 변환할 경우, 주파수 신호에서 다수의 피크가 검출되기는 하나, 주기적 질의 성분은 검출되지 않음을 확인할 수 있다. 이처럼, PSD 기법을 이용할 경우, 각 호스트 별로 주기적 질의 특성을 갖는지, 주기적 질의 특성이 얼마나 자주 발현되는지 여부 등을 간단하게 파악할 수 있다.

[0029] 봇의심객체 검출부(230)는 주파수 신호에 주기적 질의 성분이 있는 경우, 해당 주기적 질의 성분이 얼마나 중요한 신호인지 판단하여, 대응하는 호스트를 봇의심객체로 검출할 수 있다. 도 4a 및 도 4b를 통하여 확인하였듯이, 질의요청신호가 주기적 패턴을 갖는 것과 무관하게 주파수 신호에서는 항상 피크가 검출될 수 있다. 따라서, 봇의심객체 검출부(230)는 이러한 피크의 중요도를 판단함으로써, 주기적 질의 성분을 포함하는 호스트를 구별할 수 있다. 그리고, 주기적 질의 성분을 포함하는 호스트는 봇의심객체로 간주될 수 있다.

[0030] 이때, 봇의심객체 검출부(230)는 SPT(Significant Peak Testing) 기법을 이용하여 주기적 질의 성분의 중요도를 판단할 수 있다. 구체적으로, 봇의심객체 검출부(230)는 임계값을 미리 설정한다. 그리고, 아래 수학식2에 따라 주파수 신호에 포함된 주기적 질의 성분 중 전력값이 가장 높은 주기적 질의 성분에 대한 데이터 값(g*x)을 구한다.

[0031] [수학식 2]

$$g_x^* = \frac{max(P_{xx}[k])}{N^{-1}\sum_{k=1}^{N/2} P_{xx}[k]}$$

[0032]

[0033] 위 데이터값(g*x)은 각 주파수 대역의 전력값의 평균 대비 최대 전력값을 갖는 주기적 질의 성분의 전력값의 비율을 의미한다.

[0034] 이어서, 위 데이터값이 미리 설정된 임계값을 넘는 경우, 해당 주기적 질의 성분이 중요한 주기적 질의 성분인 것으로 간주하여, 대응하는 호스트를 봇의심객체로 판단할 수 있다. 예를 들어, 임계값이 23.52로 설정되는 경우, 도 4a의 주파수 신호의 전력값이 35.34로 측정되고, 도 4b의 주파수 신호의 전력값이 5.39로 측정된다면, 도 4a의 주파수 신호에 대응하는 호스트를 봇의심객체로 판별할 수 있다.

[0035] 유사그룹 분류부(240)는 봇의심객체 검출부(230)를 통하여 복수의 봇의심객체가 검출된 경우, 유사한 특성의 주기적 질의 성분을 갖는 봇의심객체를 동일한 그룹으로 분류한다. 일반 사용자 단말도 운영체제의 주기적 업데이트 등과 같은 기능에 의해 주기적 질의 특성을 나타낼 수 있다. 다만, 봇넷은 일반 사용자 단말의 주기적 질의 특성과 다른 주기패턴을 가질 수 있으므로, 유사 그룹 분류를 통하여, 봇의심객체들을 하나의 그룹으로 묶음으로써 봇넷을 검출할 수 있다.

[0036] 여기서, 유사그룹 분류부(240)는 pDist(Power Distance) 기법을 이용하여 유사 그룹을 분류할 수 있다. 구체적으로, 유사그룹 분류부(240)는 임의의 두 개의 봇의심객체에 관한 주파수 신호로부터 각각 주기적 질의 성분의 주파수값(x값)과 전력값(y값)을 파악한다. 그리고 유사그룹 분류부(240)는 양 주기적 질의 성분의 주파수값 대비 전력값의 차이를 계산하여 양 봇의심객체 간의 유사도를 판단한다. pDist(Power Distance) 기법에 의한 유사도의 수식은 수학식 3과 같다.

[0037] [수학식 3]

$$pDist = \|f_a - f_b\|$$

[0038]

[0039] 수학식 3은 fa는 a라는 봇의심객체의 주기적 질의 성분의 피크지점과 fb는 b라는 봇의심객체의 주기적 질의 성분의 피크지점 간의 거리차이를 의미한다. 계산된 거리차이가 미리 설정된 값 이하인 경우, 동일한 그룹으로 분류한다. 즉, 두 개의 봇의심객체 간 주기적 질의 성분의 주파수값과 전력값의 크기가 유사할수록, 해당하는 두 개의 봇의심객체는 동일한 그룹으로 분류될 가능성이 높다.

[0040] 이때, 한 그룹으로 분류된 각 봇의심객체들을 하나의 봇넷(120)으로 간주할 수도 있으나, 더욱 정밀한 봇넷(120) 검출을 위해 다음의 과정을 수행할 수 있다.

[0041] 봇넷 검출부(250)는 각 그룹에 포함된 봇의심객체들의 도메인 주소와 봇넷검출장치(100)에 미리 저장된 봇넷도메인주소 리스트 및 비봇넷 도메인주소 리스트를 비교하여 봇넷(120)을 검출할 수 있다. 구체적으로, 각 그룹에

포함된 봇의심객체의 도메인 주소가 봇넷도메인주소 리스트에 포함되어 있는 경우, 해당 봇의심객체를 봇(121)으로 판단하며, 동일한 도메인 주소를 갖는 복수의 봇(121)들을 하나의 봇넷(120)으로 간주할 수 있다. 반면, 봇의심객체의 도메인 주소가 비봇넷 도메인주소 리스트에 포함된 경우, 일반 사용자 단말이나 합법적인 웹사이트 등과 같은 정상 그룹으로 판단할 수 있다. 또한, 어느 한 그룹에 포함된 봇의심객체들의 도메인 주소가 봇넷도메인주소 리스트 및 비봇넷 도메인주소 리스트 중 어디에도 포함되지 않은 경우, 봇넷의심그룹으로 판단할 수 있다.

[0042] 추가적으로, 본 발명의 일 실시예의 프로세서는 검출된 봇넷(120)의 도메인 주소를 접속차단 리스트에 등록하여, 접속차단 리스트에 등록된 도메인 주소와 관련된 모든 패킷들을 사용자 단말에 송수신되지 않도록 차단할 수 있다. 예를 들어, 사용자 단말에 대한 송수신 패킷이 경유하는 로컬 서버나 프록시 서버 등에서 접속차단 리스트에 포함된 도메인 주소와 대응하는 패킷이 발견된 경우, 해당 패킷의 송수신을 차단할 수 있다.

[0043] 또한, 추가적으로, 본 발명의 일 실시예의 프로세서는 서로 상이한 시간대 별로 DNS 트래픽 수집부(210) 내지 봇넷 검출부(250)가 수행하는 과정을 반복 수행할 수 있다. 구체적으로, 도 5를 참조하면, 프로세서는 DNS 서버(130)로 송수신되는 패킷들을 항시 수집하되, 시간을 여러 개의 세그먼트들로 분할하여, 각 세그먼트마다 수집된 패킷에 대하여 질의주기성을 분석하고 봇넷(120)을 검출하는 작업을 수행할 수 있다. 종래의 봇넷 검출 또는 악성코드 검출 기술들은 특정 시간대에서만 분석을 수행하기 때문에, 봇넷 제어시간이 변경되거나 봇넷(120)을 구성하는 봇(121)들의 전원이 상이한 시간 대에 켜지는 등 특정 시간 대의 행동 규칙을 벗어나는 경우가 발생될 경우, 봇넷(120)을 검출하기 어려웠다. 이를 개선하기 위하여, 본 발명의 일 실시예에 따르는 프로세서는 각 시간 대 별로 분석을 수행하므로, 어느 시간 대이든지 한번이라도 주기적 질의 특성을 검출하는 경우, 주기적 질의 특성이 검출된 해당 호스트를 봇(121)으로 검출할 수 있다. 구체적으로, DNS 트래픽 수집부(210)가 각 호스트 별로 질의요청주기를 측정하는 과정을 상이한 시간 대 별로 반복 수행함으로써, 상술한 기능을 달성할 수 있다.

[0044] 이하, 도 6을 참조하여, 본 발명의 일 실시예에 따르는 DNS 트래픽을 통하여 봇넷(120)을 검출하는 방법에 대하여 구체적으로 설명한다. 본 발명의 일 실시예에 따르는 방법은 상술한 봇넷검출장치(100)에 의하여 수행되는 것으로서, 이하에서 생략되는 내용이라 하더라도 봇넷검출장치(100)에 의해 수행되는 구성의 방법적인 특징들을 모두 포함한다.

[0045] 먼저, 봇넷검출장치(100)는 DNS 트래픽을 수집하여 각 호스트 별로 DNS 서버(130)에 대한 단위시간당 질의요청 횟수를 측정한다(S600). 이때, 봇넷검출장치(100)는 1일을 전체 시간으로 가정하였을 경우, 1일을 여러 개의 시간대로 분할하여, 각 시간대마다 단위시간당 질의요청횟수를 측정할 수 있다.

[0046] 봇넷검출장치(100)는 단위시간당 질의요청횟수를 이용하여 타임 도메인 상의 질의요청신호를 생성한다(S610).

[0047] 이어서, 봇넷검출장치(100)는 PSD 기법을 이용하여 질의요청신호를 주파수 신호로 변환한다(S620). 그에 따라, 임의의 호스트가 DNS 서버(130)에 대한 주기적 질의패턴을 가질 경우, 도 4a와 같이 특정 주파수 대역에서 주기적 질의 성분(즉, 피크)이 나타나는 주파수 신호가 생성될 수 있다. 반면, 다른 호스트가 DNS 서버(130)에 대한 비주기적 질의패턴을 가질 경우, 도 4b와 같이 주기적 질의 성분이 나타나지 않는 주파수 신호가 생성될 수 있다.

[0048] 그리고 봇넷검출장치(100)는 SPT 기법을 이용하여 주파수 신호에 포함된 주기적 질의 성분의 중요도에 따라 봇의심객체 여부를 판단한다(S630). 구체적으로, 전체 주파수 대역의 전력값 평균에 대비한 주기적 질의 성분의 전력값이 미리 설정된 임계값 이상인 경우, 중요도가 높은 주기적 질의 성분으로 간주되어, 해당 주기적 질의 성분을 갖는 호스트를 봇의심객체로 판단할 수 있다. 일반 사용자의 PC의 경우에도 일시적으로 주기적 질의패턴을 나타낼 수도 있으나, 중요도 판단 과정을 통해 제거될 수 있다.

[0049] 이어서, 봇넷검출장치(100)는 각 주파수 신호 간의 유사도를 분석하여 봇넷(120)의심객체를 유사그룹으로 분류함에 따라 봇넷(120)을 검출할 수 있다(S640). 구체적으로, 봇넷검출장치(100)는 pDist(Power Distance) 기법을 이용하여 각 주파수 신호 간의 전력값 크기 차이를 계산하고, 전력값 크기 차이가 미리 설정된 크기 이하인 경우 동일한 그룹으로 분류함으로써 유사 그룹들을 분류할 수 있다. 이러한 과정들을 통하여 유사한 주기적 질의 패턴을 갖는 봇(121)들은 하나의 그룹으로 분류되어 봇넷 그룹을 형성할 수 있다. 또한, 일반 사용자 단말의 경우에도 운영체제 업데이트 등의 과정으로 주기적 질의 패턴을 가질 수 있으나 유사그룹 분류 과정을 통하여 봇넷 그룹과 다른 그룹으로 분류될 수 있다. 그리고, 봇넷검출장치(100)는 미리 저장된 봇넷도메인주소 리스트 및 비봇넷도메인주소 리스트와 각 그룹의 봇의심객체들의 도메인 주소를 비교하여 각 그룹이 봇넷 그룹, 정상

그룹, 봇넷의심그룹 그룹인지 판단할 수 있다.

[0050] DNS 질의 패턴은 봇넷(120)이 임의로 변경할 수 없는 근본적인 행위 특징이다. 따라서, 더욱 지능화된 악성코드가 나타난다고 하더라도 규칙적인 DNS 질의 패턴을 가지게 되므로, 이를 이용하여 봇넷을 검출하는 본 발명의 일 실시예는 종래의 봇넷 탐지 기법에 비하여 더욱 근본적인 악성코드 탐지 및 대응의 효과를 가질 수 있다. 또한, DNS 질의 주기성을 분석함으로써, 악성코드에 감염된 사용자 컴퓨터가 발생시키는 트래픽 여부와 상관없이 안정적인 봇넷 행위 분석이 가능하다. 또한, 시간대 별로 DNS 질의 주기성을 분석하기 때문에, 봇넷의 활동시간이 변경되더라도 어느 한 시간 대에서 주기적 질의 특성이 검출되기만 한다면 봇넷으로 검출될 수 있으므로, 효과적인 봇넷 검출이 가능하다.

[0051] 본 발명의 일 실시예는 컴퓨터에 의해 실행되는 프로그램 모듈과 같은 컴퓨터에 의해 실행가능한 명령어를 포함하는 기록 매체의 형태로도 구현될 수 있다. 컴퓨터 판독 가능 매체는 컴퓨터에 의해 액세스될 수 있는 임의의 가용 매체일 수 있고, 휘발성 및 비휘발성 매체, 분리형 및 비분리형 매체를 모두 포함한다. 또한, 컴퓨터 판독 가능 매체는 컴퓨터 저장 매체 및 통신 매체를 모두 포함할 수 있다. 컴퓨터 저장 매체는 컴퓨터 판독가능 명령어, 데이터 구조, 프로그램 모듈 또는 기타 데이터와 같은 정보의 저장을 위한 임의의 방법 또는 기술로 구현된 휘발성 및 비휘발성, 분리형 및 비분리형 매체를 모두 포함한다. 통신 매체는 전형적으로 컴퓨터 판독가능 명령어, 데이터 구조, 프로그램 모듈, 또는 반송파와 같은 변조된 데이터 신호의 기타 데이터, 또는 기타 전송 메커니즘을 포함하며, 임의의 정보 전달 매체를 포함한다.

[0052] 본 발명의 방법 및 시스템은 특정 실시예와 관련하여 설명되었지만, 그것들의 구성 요소 또는 동작의 일부 또는 전부는 범용 하드웨어 아키텍쳐를 갖는 컴퓨터 시스템을 사용하여 구현될 수 있다.

[0053] 전술한 본 발명의 설명은 예시를 위한 것이며, 본 발명이 속하는 기술분야의 통상의 지식을 가진 자는 본 발명의 기술적 사상이나 필수적인 특징을 변경하지 않고서 다른 구체적인 형태로 쉽게 변형이 가능하다는 것을 이해할 수 있을 것이다. 그러므로 이상에서 기술한 실시예들은 모든 면에서 예시적인 것이며 한정적이 아닌 것으로 이해해야만 한다. 예를 들어, 단일형으로 설명되어 있는 각 구성 요소는 분산되어 실시될 수도 있으며, 마찬가지로 분산된 것으로 설명되어 있는 구성 요소들도 결합된 형태로 실시될 수 있다.

[0054] 본 발명의 범위는 상기 상세한 설명보다는 후술하는 특허청구범위에 의하여 나타내어지며, 특허청구범위의 의미 및 범위 그리고 그 균등 개념으로부터 도출되는 모든 변경 또는 변형된 형태가 본 발명의 범위에 포함되는 것으로 해석되어야 한다.

부호의 설명

[0055] 100: 봇넷검출장치

210: DNS 트래픽 수집부

220 : 질의주기성 분석부

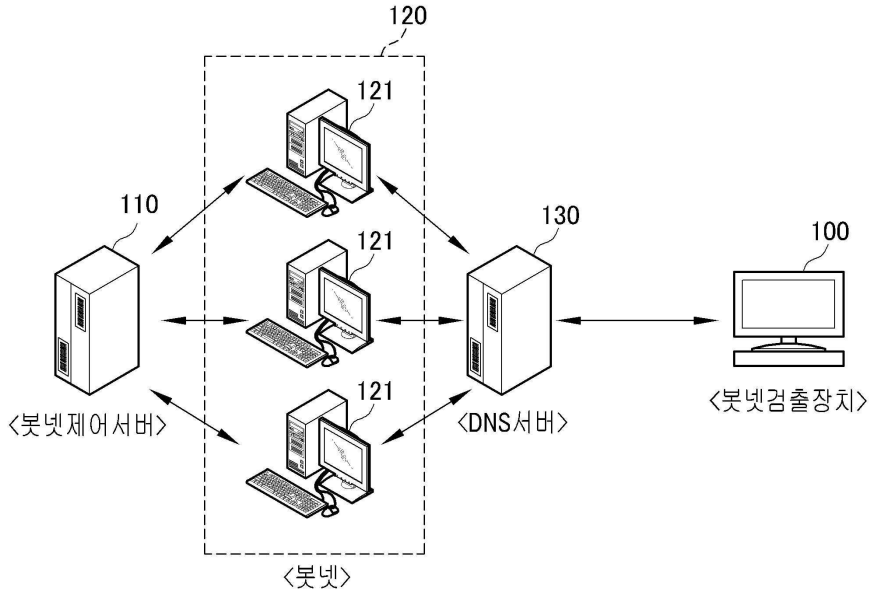
230 : 봇의심객체 검출부

240 : 유사그룹 분류부

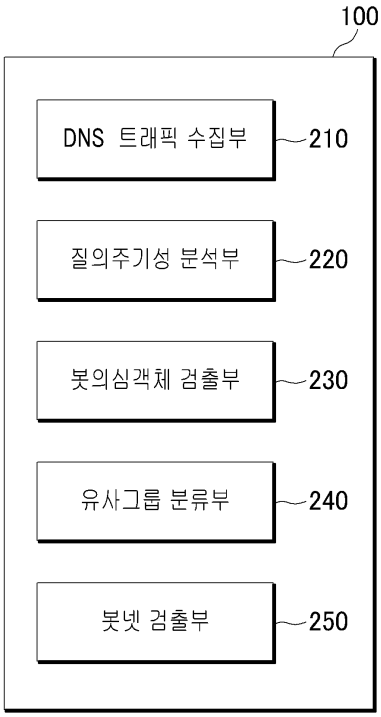
250 : 봇넷 검출부

도면

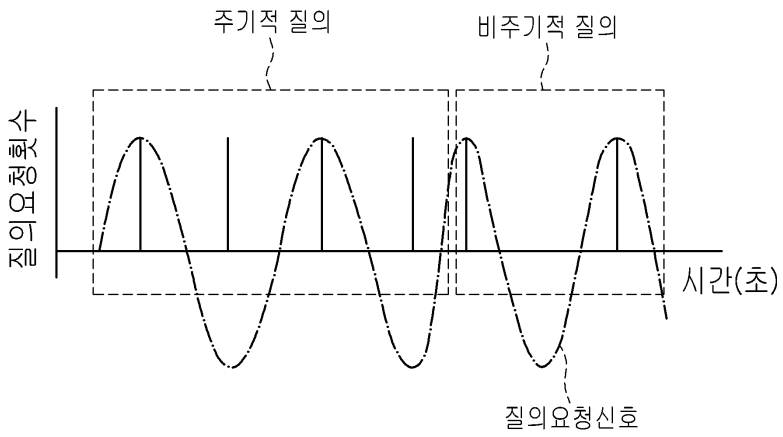
도면1


120
121
110
130
100
121
121
<봇넷제어서버>
<DNS서버>
<봇넷검출장치>
<봇넷>

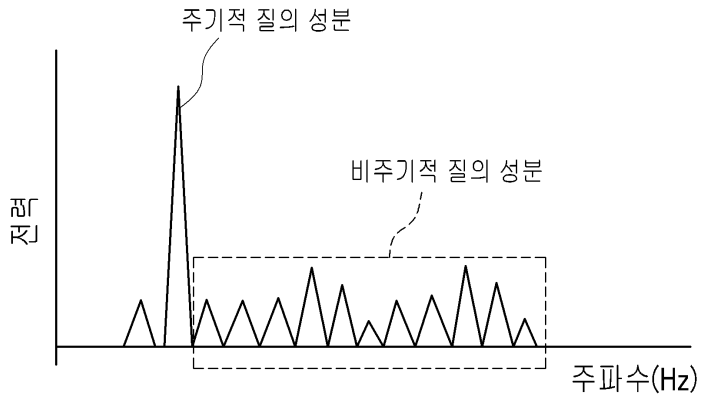
도면2


100
DNS 트래픽 수집부
210
질의주기성 분석부
220
봇의심객체 검출부
230
유사그룹 분류부
240
봇넷 검출부
250

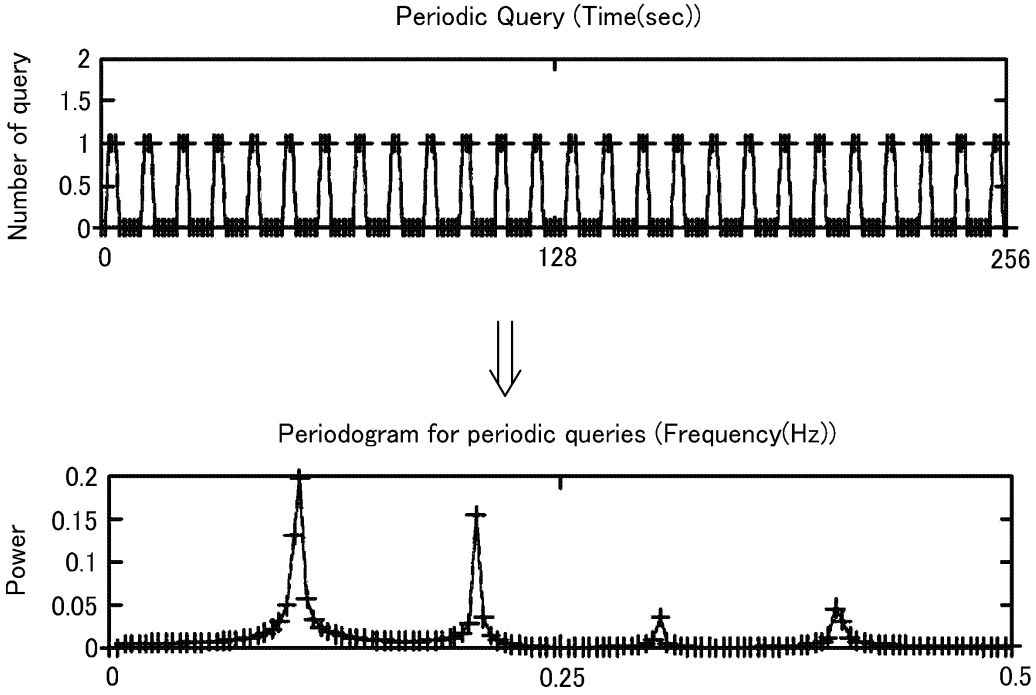
도면3a



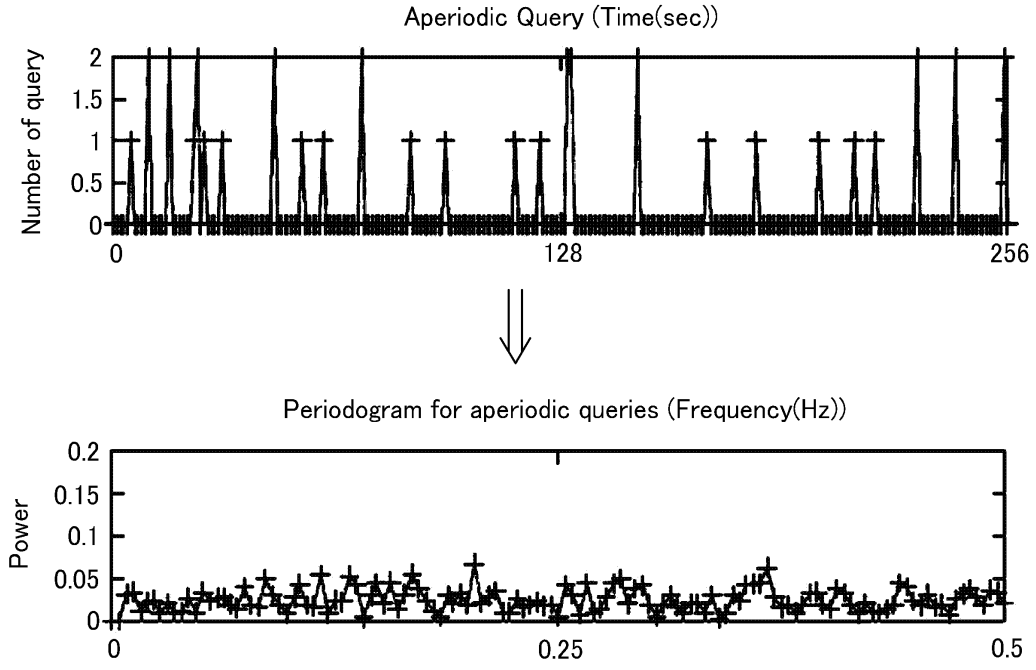
도면3b



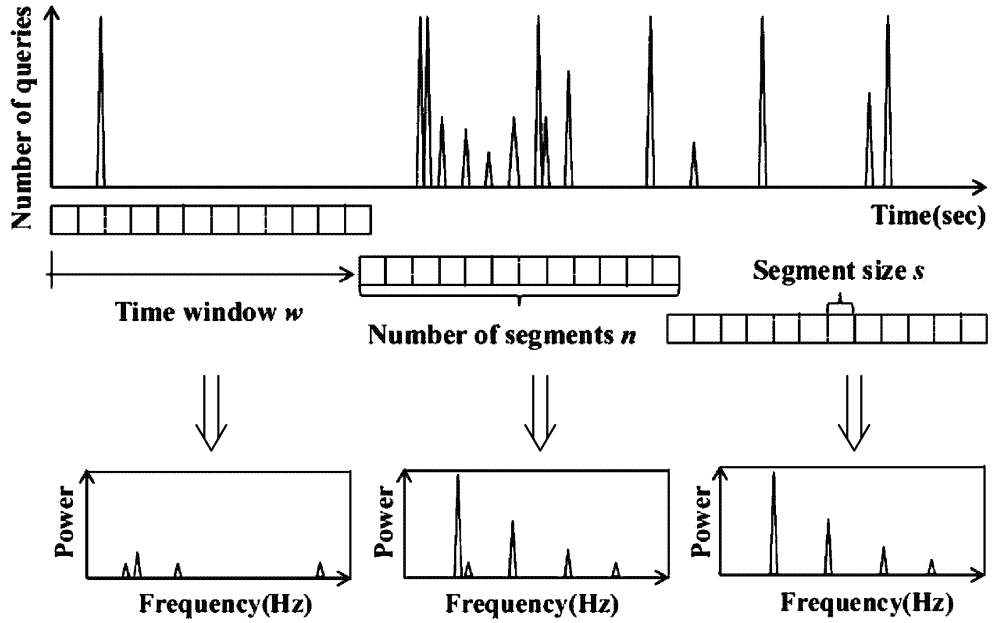
도면4a


Periodic Query (Time(sec))
Number of query
2
1.5
1
0.5
0
0
128
256
Periodogram for periodic queries (Frequency(Hz))
Power
0.2
0.15
0.1
0.05
0
0
0.25
0.5

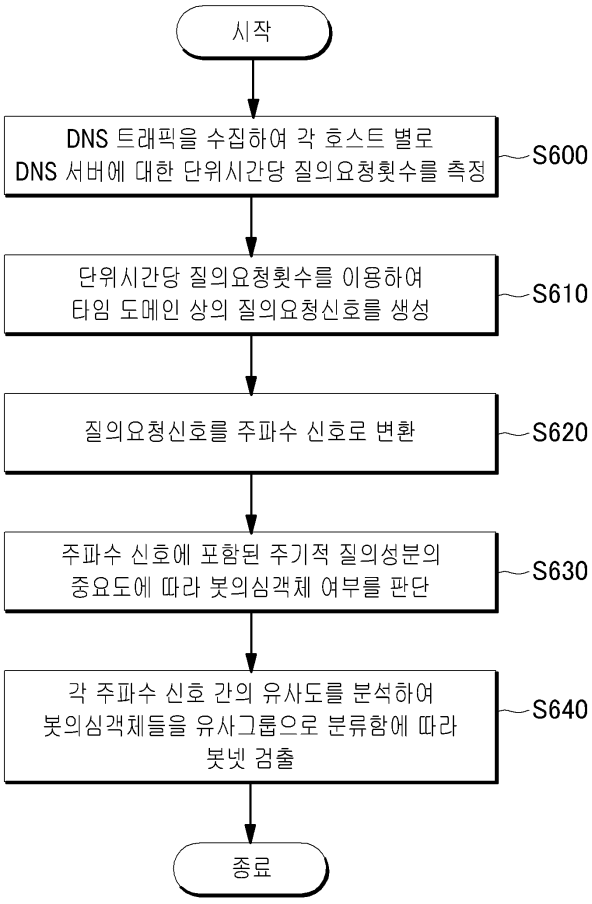
도면4b


Aperiodic Query (Time(sec))
Number of query
2
1.5
1
0.5
0
0
128
256
Periodogram for aperiodic queries (Frequency(Hz))
Power
0.2
0.15
0.1
0.05
0
0
0.25
0.5

도면5


Number of queries
Time(sec)
Time window w
Number of segments n
Segment size s
Power
Frequency(Hz)
Power
Frequency(Hz)
Power
Frequency(Hz)

도면6


시작
DNS 트래픽을 수집하여 각 호스트 별로
DNS 서버에 대한 단위시간당 질의요청횟수를 측정
S600
단위시간당 질의요청횟수를 이용하여
타임 도메인 상의 질의요청신호를 생성
S610
질의요청신호를 주파수 신호로 변환
S620
주파수 신호에 포함된 주기적 질의성분의
중요도에 따라 봇의심객체 여부를 판단
S630
각 주파수 신호 간의 유사도를 분석하여
봇의심객체들을 유사그룹으로 분류함에 따라
봇넷 검출
S640
종료