



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2011-0115467
(43) 공개일자 2011년10월21일

(51) Int. Cl.

H04L 12/26 (2006.01) H04L 12/56 (2006.01)

(21) 출원번호 10-2010-0034980

(22) 출원일자 2010년04월15일

심사청구일자 2010년04월15일

(71) 출원인

고려대학교 산학협력단

서울 성북구 안암동5가 1

(72) 발명자

김명섭

충청남도 연기군 조치원읍 신흥리 푸르지오1차
107동 401호

윤성호

경상북도 포항시 북구 용흥1동 현대타워 104동
1205호

박진완

충청남도 연기군 조치원읍 서창리 고려대학교세종
캠퍼스 과학기술대학 210-B

(74) 대리인

특허법인엠에이피에스

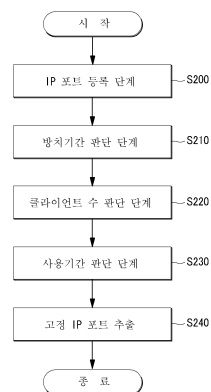
전체 청구항 수 : 총 13 항

(54) 고정 IP 포트 추출장치 및 그 방법

(57) 요약

고정 IP(Internet Protocol) 포트를 추출하는 장치에 있어서, IP 네트워크 상에서 분석 대상 어플리케이션의 트래픽 정보를 수집하는 트래픽 수집부 및 상기 수집된 트래픽 정보로부터 IP 포트를 사용하여 통신한 클라이언트 호스트의 개수 및 상기 IP 포트를 사용하여 상기 클라이언트 호스트와 통신한 기간에 기초하여 고정 IP 포트를 추출하는 고정 IP 포트 추출부를 포함하는 고정 IP 포트 추출장치이다.

대표도 - 도2



이 발명을 지원한 국가연구개발사업

과제고유번호 D00387

부처명 학술진흥재단

연구관리전문기관

연구사업명 신진교수 연구지원 사업

연구과제명 인터넷 트래픽 응용레벨 분석 방법에 관한 연구

기여율

주관기관 고려대학교 산학협력단

연구기간 2009년 08월 01일 ~ 2010년 07월 31일

특허청구의 범위

청구항 1

고정 IP(Internet Protocol) 포트를 추출하는 장치에 있어서,

IP 네트워크 상에서 분석 대상 어플리케이션의 트래픽 정보를 수집하는 트래픽 수집부 및

상기 수집된 트래픽 정보로부터 IP 포트를 사용하여 통신한 클라이언트 호스트의 개수 및 상기 IP 포트를 사용하여 상기 클라이언트 호스트와 통신한 기간에 기초하여 고정 IP 포트를 추출하는 고정 IP 포트 추출부

를 포함하는 고정 IP 포트 추출장치.

청구항 2

제 1 항에 있어서,

상기 고정 IP 포트 추출부는 상기 IP 포트를 사용하여 통신한 상기 클라이언트 호스트의 개수가 클라이언트 임계값 이상이면 상기 IP 포트를 사용하여 상기 클라이언트 호스트와 통신한 기간이 사용기간 임계값 이상인 IP 포트를 추출하는 것인 고정 IP 포트 추출장치.

청구항 3

제 1 항에 있어서,

상기 고정 IP 포트는 트래픽 분석을 위한 시그니처로 이용되는 것인 고정 IP 포트 추출장치.

청구항 4

제 1 항에 있어서,

상기 트래픽 수집부가 수집하는 상기 트래픽 정보는 네트워크 소켓 정보 및 상기 소켓을 생성한 프로세스 정보를 포함하는 고정 IP 포트 추출장치.

청구항 5

제 1 항에 있어서,

상기 고정 IP 포트 추출부는 온라인 상에서 동작하는 제 1 고정 IP 포트 추출부 및 오프라인 상에서 동작하는 제 2 고정 IP 포트 추출부를 포함하고, 상기 제 2 고정 IP 포트 추출부는 상기 제 1 고정 IP 포트 추출부가 온라인 상에서 실시간으로 동작하지 않는 경우에 동작하는 고정 IP 포트 추출장치.

청구항 6

고정 IP 포트 정보를 추출하는 방법에 있어서,

클라이언트 호스트와 통신하는 IP 포트를 등록하는 단계,

상기 등록된 IP 포트와 통신하는 클라이언트 호스트의 개수가 클라이언트 임계값 이상인지 판단하는 단계,

상기 등록된 IP 포트를 사용하여 상기 클라이언트 호스트와 통신하는 기간이 사용기간 임계값 이상인지 판단하는 단계 및

상기 클라이언트 호스트의 개수가 상기 클라이언트 임계값 이상이면서 상기 클라이언트 호스트와 통신하는 기간이 상기 사용기간 임계값 이상인 IP 포트를 추출하는 단계를 포함하는 고정 IP 포트 추출방법.

청구항 7

제 6 항에 있어서,
상기 IP 포트를 등록하는 단계는 상기 IP 포트가 기등록된 것인지 판단하는 단계 및
상기 IP 포트가 기존의 어플리케이션과 다른 어플리케이션에 의해 발생된 것인지 검사하는 단계를 포함하는 고정 IP 포트 추출방법.

청구항 8

제 6 항에 있어서,
상기 등록된 IP 포트 중 등록된 시점으로부터 방치기간 임계값 내에 고정 IP 포트로 추출되는지 검사하는 단계를 더 포함하는 고정 IP 포트 추출방법.

청구항 9

제 6 항 내지 제 8 항 중 어느 한 항에 있어서,
상기 클라이언트 임계값은 2인 고정 IP 포트 추출방법.

청구항 10

제 6 항 내지 제 8 항 중 어느 한 항에 있어서,
상기 사용기간 임계값은 사용기간 임계값에 따른 IP 포트 개수의 변화를 확인하는 실험을 통해 결정하되,
상기 사용기간 임계값은 임시 IP 포트가 더 이상 추출되지 않는 때의 값인 고정 IP 포트 추출방법.

청구항 11

트래픽을 분석하는 장치에 있어서,
분석 대상 어플리케이션의 트래픽 정보를 입력 받아 해당 어플리케이션의 고정 IP 포트를 추출하는 고정 IP 포트 추출장치,
상기 추출된 고정 IP 포트에 기초하여 분석 대상 네트워크에 대해 트래픽을 분석하는 분석장치 및
상기 트래픽 분석 결과를 검증하는 검증장치를 포함하되,
상기 고정 IP 포트 추출장치는 상기 트래픽 정보로부터 IP 포트를 사용하여 통신한 클라이언트 호스트의 개수 및 상기 IP 포트를 사용하여 상기 클라이언트 호스트와 통신한 기간에 기초하여 고정 IP 포트를 추출하는 트래픽 분석 장치.

청구항 12

제 11 항에 있어서,

상기 분석장치는 분석 대상 네트워크 링크에서 패킷을 수집하는 패킷 캡처부,

상기 패킷 캡처부에서 수집한 패킷들을 이용하여 플로우를 만드는 플로우 생성부 및

상기 플로우 생성부에서 생성된 플로우 헤더와 상기 고정 IP 포트 추출장치에서 추출된 고정 IP 포트 헤더를 비교하는 헤더 비교부를 포함하는 트래픽 분석 장치.

청구항 13

제 11 항에 있어서,

상기 검증장치는 네트워크에서 분석 대상 어플리케이션의 트래픽 정보를 수집하는 하는 검증용 트래픽 수집부,

상기 검증용 트래픽 수집부에서 수집된 트래픽 정보와 상기 플로우 생성부에서 생성된 플로우 데이터를 이용하여 정답지 데이터를 만드는 정답지 데이터 생성부 및

상기 정답지 데이터 생성부로부터 만들어진 정답지 데이터와 상기 분석장치의 트래픽 분석 결과를 비교하여 트래픽 분석 결과를 검증하는 매칭부를 포함하는 트래픽 분석 장치.

명세서

기술 분야

[0001] 본 발명은 트래픽 분석에 사용할 수 있는 고정 IP 포트 추출장치 및 그 방법에 관한 것이다.

배경 기술

[0002] 최근 인터넷 사용자의 증가와 고속 네트워크의 보급으로 인해 네트워크 트래픽이 급증하였다. 이것은 단순히 WWW, FTP, e-mail 과 같은 전통적인 인터넷 서비스뿐만 아니라 멀티미디어 스트리밍, P2P(peer to peer) 파일 공유, 게임과 같은 다양한 멀티미디어 서비스 대중화에 원인이 있다. 따라서 인터넷 트래픽이 급증함에 따라 효과적인 네트워크 관리를 위한 트래픽 모니터링 및 분석의 중요성이 커지고 있다.

[0003] 네트워크 트래픽 분석은 네트워크 링크로부터 패킷을 수집하고 분석 기준에 맞게 각 패킷을 분류하는 일련의 과정이다. 이러한 분석 결과는 네트워크 관리와 제어의 관점에서 효과적으로 사용된다. 한편, 트래픽 분석은 여러 방법에 의해 수행된다.

[0004] 트래픽 분석 방법들은 크게 4가지로 구분할 수 있는데, 이들은 포트 기반 분석, 시그니처 기반 분석, 머신 러닝 기반의 분석 및 트래픽 상관관계 기반 분석이다.

[0005] 시그니처 기반의 트래픽 분석 방법은 특정 어플리케이션 트래픽이 다른 어플리케이션들과 구분하여 사용하는 공통 요소를 찾아내 그것을 그 어플리케이션의 시그니처라 명명하고, 동일한 시그니처를 가졌는지 아닌지를 판단하는 방법으로 트래픽을 분석하는 방법이다. 이 방법은 시그니처가 확인된 어플리케이션에 대해서는 정확한 분석이 가능하다는 장점이 있으나, 모든 어플리케이션 별로 시그니처를 수작업으로 찾아야 하고, 시그니처를 확인하기 힘든 어플리케이션들이 존재하며, 또 확인된 시그니처가 어플리케이션의 변화에 적절히 대처하지 못한다는 단점이 있었다.

발명의 내용

해결하려는 과제

[0006] 본 발명은 전술한 종래기술의 문제점을 해결하기 위한 것으로, 한 어플리케이션을 대표하는 고정 IP 포트를 자동 추출 시스템을 통해 추출하고, 이를 기반으로 트래픽을 분석하는 장치를 제공하는 것을 목적으로 한다.

과제의 해결 수단

[0007] 상술한 기술적 과제를 달성하기 위한 기술적 수단으로서, 본 발명의 제 1 측면은 고정 IP(Internet Protocol) 포트를 추출하는 장치에 있어서, IP 네트워크 상에서 분석 대상 어플리케이션의 트래픽 정보를 수집하는 트래픽 수집부 및 상기 수집된 트래픽 정보로부터 IP 포트를 사용하여 통신한 클라이언트 호스트의 개수 및 상기 IP 포

트를 사용하여 상기 클라이언트 호스트와 통신한 기간에 기초하여 고정 IP 포트를 추출하는 고정 IP 포트 추출부를 포함하는 고정 IP 포트 추출장치를 제공한다.

[0008] 또한, 본 발명의 제 2 측면은 고정 IP 포트 정보를 추출하는 방법에 있어서, 클라이언트 호스트와 통신하는 IP 포트를 등록하는 단계, 상기 등록된 IP 포트와 통신하는 클라이언트 호스트의 개수가 클라이언트 임계값 이상인지 판단하는 단계, 상기 등록된 IP 포트를 사용하여 상기 클라이언트 호스트와 통신하는 기간이 사용기간 임계값 이상인지 판단하는 단계 및 상기 클라이언트 호스트의 개수가 상기 클라이언트 임계값 이상이면서 상기 클라이언트 호스트와 통신하는 기간이 상기 사용기간 임계값 이상인 IP 포트를 추출하는 단계를 포함하는 고정 IP 포트 추출방법을 제공한다.

[0009] 또한, 본 발명의 제 3 측면은 트래픽을 분석하는 장치에 있어서, 분석 대상 어플리케이션의 트래픽 정보를 입력받아 해당 어플리케이션의 고정 IP 포트를 추출하는 고정 IP 포트 추출장치, 상기 추출된 고정 IP 포트에 기초하여 분석 대상 네트워크에 대해 트래픽을 분석하는 분석장치 및 상기 트래픽 분석 결과를 검증하는 검증장치를 포함하되, 상기 고정 IP 포트 추출장치는 상기 트래픽 정보로부터 IP 포트를 사용하여 통신한 클라이언트 호스트의 개수 및 상기 IP 포트를 사용하여 상기 클라이언트 호스트와 통신한 기간에 기초하여 고정 IP 포트를 추출하는 트래픽 분석 장치를 제공한다.

발명의 효과

[0010] 진술한 본 발명의 과제 해결 수단에 의하면, 분석 대상 어플리케이션의 트래픽 정보를 입력받아 자동으로 상기 어플리케이션의 고정 IP 포트를 추출한다. 특히, 고정 IP 포트를 추출하는 방법에 있어서, IP 포트와 통신하는 클라이언트 호스트의 개수 및 상기 IP 포트를 사용하여 상기 클라이언트 호스트와 통신한 기간을 조건으로 함으로써 고정 IP 포트 추출의 정확도를 높일 수 있다.

도면의 간단한 설명

[0011] 도 1은 본 발명의 일실시예에 따른 고정 IP 포트 추출장치를 도시한 블록도.
 도 2는 본 발명의 일실시예에 따른 고정 IP 포트 추출방법을 도시한 동작 흐름도.
 도 3은 본 발명의 일실시예에 따른 IP 포트 등록 단계를 도시한 동작 흐름도.
 도 4는 본 발명의 일실시예에 따라 IP 포트의 상태 전이를 설명하기 위한 상태 흐름도.
 도 5는 본 발명의 일실시예에 따른 고정 IP 포트 추출 알고리즘을 설명하기 위한 도면.
 도 6a는 본 발명의 일실시예에 따라 클라이언트 임계값에 따른 분석률 및 정확도 변화를 나타낸 도면.
 도 6b 본 발명의 일실시예에 따라 사용기간 임계값에 따른 IP 포트 개수 변화를 나타낸 도면.
 도 6c는 본 발명의 일실시예에 따라 방치기간 임계값에 따른 IP 포트 개수 변화를 나타낸 도면.
 도 7은 본 발명의 일실시예에 따른 고정 IP 포트를 이용한 트래픽 분석 장치를 도시한 블록도.

발명을 실시하기 위한 구체적인 내용

[0012] 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.

[0013] 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.

[0014] 도 1은 본 발명의 일실시예에 따른 고정 IP 포트 추출장치를 도시한 블록도이다.

[0015] 도 1을 참조하면 고정 IP 포트 추출장치(100)는 트래픽 수집부(110), 트래픽 전송부(120) 및 고정 IP 포트 추출부(130)를 포함한다.

- [0016] 트래픽 수집부(110)는 네트워크 상에서 트래픽 정보를 수집한다. 수집하는 트래픽 정보는 분석 대상 어플리케이션에 관한 트래픽 정보로, 네트워크 소켓 정보, 해당 소켓을 생성 시킨 프로세스의 이름 및 프로세스가 설치된 경로 정보를 포함할 수 있다. 네트워크 소켓 정보는 SrcIP(Source IP), DstIP(Destination IP), SRCPort(Source Port), DstPort(Destination Port) 또는 프로토콜을 포함할 수 있다.
- [0017] 트래픽 수집부(110)는 실제 트래픽이 발생하는 종단 호스트에 설치될 수 있으며, 실시간으로 발생하는 트래픽 정보를 모을 수 있는 모든 형태의 하드웨어, 소프트웨어 또는 하드웨어와 소프트웨어의 조합일 수 있다. 또한 트래픽 수집부(110)는 동일한 구성을 갖는 다수의 장치로 구성될 수 있다.
- [0018] 트래픽 전송부(120)는 트래픽 수집부(110)에서 수집된 정보를 고정 IP 포트 추출부(130)로 전송한다. 트래픽 전송부(120)는 유무선에 의한 정보전달 매체일 수 있으며, 바람직하게는 온라인 서버일 수 있다. 필요에 따라 트래픽 전송부(120)가 트래픽 수집부(110) 내에 포함되어, 트래픽 수집부(110)가 트래픽 정보를 수집하는 동시에 전송하도록 구성할 수도 있다.
- [0019] 고정 IP 포트 추출부(130)는 트래픽 전송부(120)로부터 전송된 트래픽 정보를 이용하여 고정 IP 포트 정보를 추출한다.
- [0020] IP 포트는 고정 IP 포트, 동적 IP 포트 및 임시 IP 포트에 분류될 수 있다. 고정 IP 포트는 한 어플리케이션이 단독으로 고정적으로 사용하는 IP 포트이다. 그 예로 로그인 서버의 IP 포트가 있다. 동적 IP 포트는 한 어플리케이션에서 사용된다는 점에서 고정 IP 포트와 같지만, IP 포트가 계속 바뀐다는 점에서 차이가 있다. 임시 IP 포트는 IP 포트가 여러 어플리케이션과 공유한다는 점에서 고정 IP 포트 및 동적 IP 포트와 다르다.
- [0021] IP 포트에서 고정 IP 포트를 추출하기 위해, 고정 IP 포트 추출부(130)는 IP 포트를 사용하여 통신한 클라이언트 호스트의 개수 및 상기 IP 포트를 사용하여 클라이언트 호스트와 통신한 기간 정보를 이용할 수 있다. 고정 IP 포트 추출부(130)가 클라이언트 호스트의 개수 및 사용기간 정보를 이용하여 고정 IP 포트를 추출하는 방법에 대해서는 도 2 및 도 3에서 상세히 설명한다.
- [0022] 고정 IP 포트 추출부(130)는 제 1 고정 IP 포트 추출부(132)뿐만 아니라 제 2 고정 IP 포트 추출부(134)를 포함할 수 있다. 제 1 고정 IP 포트 추출부(132) 및 제 2 고정 IP 포트 추출부(134)는 같은 구성일 수 있으며, 제 2 고정 IP 포트 추출부(134)와 같은 구성의 제 3 고정 IP 포트 추출부, 제 4 고정 IP 포트 추출부 등을 더 포함할 수 있다.
- [0023] 일실시예로 제 1 고정 IP 포트 추출부(132) 및 제 2 고정 IP 포트 추출부(134)는 트래픽 전송부(120)로부터 전송 받은 서로 다른 트래픽 정보를 병렬로 입력 받아 각각 처리할 수 있다. 이로써 고정 IP 포트 추출 속도를 향상시킬 수 있다.
- [0024] 변형된 일실시예로 제 1 고정 IP 포트 추출부(132)는 실시간으로 처리하고, 제 2 고정 IP 포트 추출부(134)는 비실시간으로 처리할 수 있다. 제 1 고정 IP 포트 추출부(132)는 온라인 전송하는 트래픽 전송부(120)로부터 트래픽 정보를 입력 받아 실시간으로 고정 IP 포트를 추출한다. 한편, 제 2 고정 IP 포트 추출부(134)는 트래픽 전송부(120)로부터 온라인 또는 오프라인 전송 받아 비실시간으로 처리할 수 있다. 제 2 고정 IP 포트 추출부(134)는 장기간 많은 데이터를 유지하지 못하는 실시간 추출의 단점을 보완하기 위한 것으로, 모아진 데이터를 정기적 또는 비정기적으로 처리할 수 있다. 이 경우 제 1 고정 IP 포트 추출부(132) 및 제 2 고정 IP 포트 추출부(134)는 같은 트래픽 정보를 전송 받아 동시에 가동시킬 수 있고, 제 1 고정 IP 포트 추출부(132)에 의해 실시간 처리가 어려운 경우에만 선택적으로 제 2 고정 IP 포트 추출부(134)를 가동시킬 수도 있다. 트래픽 정보가 제 2 고정 IP 포트 추출부(134)로 오프라인 전송되는 경우, 모든 형태의 데이터 저장 매체를 이용할 수 있다.
- [0025] 고정 IP 포트 추출부(130)는 전송된 정보에 대해 특정한 조건에 따라 분류 또는 추출할 수 있는 모든 형태의 하드웨어, 소프트웨어 또는 하드웨어와 소프트웨어의 조합일 수 있다.
- [0026] 고정 IP 포트 추출장치(100)는 분석 대상 어플리케이션에 대해 독립적으로 실행될 수 있다. 서로 다른 어플리케이션은 서로 다른 고정 IP 포트를 사용한다. 따라서 각 어플리케이션 별로 고정 IP 포트를 추출하는 동작이 독립적으로 실행될 수 있다.
- [0027] 도 2는 본 발명의 일실시예에 따른 고정 IP 포트 추출방법을 도시한 동작 흐름도이다.
- [0028] 먼저 IP 포트를 등록한다(S200). 어플리케이션 서버 IP 포트 중 클라이언트 호스트와 통신하는 것을 감지한 후, 기등록된 IP 포트에 해당되지 않는다면 신규로 등록한다. IP 포트 등록 단계에 대해서는 도 3에서 상세히

설명한다.

- [0029] 다음으로 등록된 IP 포트의 방치기간을 판단한다(S210). 등록된 IP 포트가 방치기간 임계값 기간 내에 하기의 클라이언트 임계값($T_{clienthost}$) 및 사용기간 임계값($T_{duration}$) 조건을 만족하는지 판단한다. 이 단계는 IP 포트가 하기의 클라이언트 임계값 및 사용기간 임계값 조건을 만족하지 못한 채 무한히 한 상태에 남아있는 것을 방지하기 위함이다. 하드웨어 측면에서 보면, 등록된 IP 포트가 무한히 한 상태에 남아있다는 것은 IP 포트 정보가 메모리 상에 무한히 방치되는 것으로 이는 시스템 오버헤드를 초래할 수 있기 때문이다. 방치기간 판단단계에 대해서는 도 4에서 상세히 설명한다.
- [0030] 다음으로 클라이언트 수를 판단한다(S220). 등록된 IP 포트를 사용하여 통신하는 클라이언트 호스트의 개수가 클라이언트 임계값($T_{clienthost}$) 이상인지 판단한다. 상기 단계 이전에, 등록된 IP 포트를 사용하여 통신하는 클라이언트 호스트의 개수를 산출하는 단계를 더 포함할 수 있다. 클라이언트 호스트의 개수를 산출하는 방법으로, 산출 시점을 기준으로 통신하고 있는 클라이언트 호스트의 개수를 산출할 수 있다. 또는 산출 시점에는 통신하지 않지만 IP 포트와 통신했었던 클라이언트 호스트의 개수까지 합산해서 산출하거나, 그밖에 클라이언트 호스트의 개수를 산출하는 다른 방법을 이용할 수도 있다.
- [0031] 많은 클라이언트 호스트를 사용하여 통신하는 IP 포트일수록 고정 IP 포트일 확률이 높다. 따라서 클라이언트 임계값 이상의 클라이언트 호스트 개수를 가지는 IP 포트만이 고정 IP 포트로 판단된다. 클라이언트 임계값이 증가하면 입력 데이터 양이 풍부한 고정 IP 포트만 추출되므로, 이를 이용한 트래픽 분석의 정확도(Accuracy)는 증가할 것이다. 하지만, 임계값에 도달하지 못하는 고정 IP 포트들이 추출되지 못해 분석률(Completeness)은 감소할 것이다.
- [0032] 다음으로 IP 포트의 사용기간을 판단한다(S230). 사용기간이란 IP 포트의 최초 등록 시각으로부터 마지막 사용 시각까지의 기간을 지칭하고, 상기 단계에서는 사용기간이 사용기간 임계값($T_{duration}$) 이상인지 판단한다.
- [0033] 상기 단계 이전에, 사용기간을 산출하는 단계를 더 포함할 수 있다. 사용기간은 일반적으로 IP 포트가 최초 등록된 시점부터 마지막 사용 시점까지 기간을 지칭한다. 그러나 클라이언트 호스트 별로 각각 사용 기간을 평균하여 산출하는 등 그밖에 등록된 IP 포트를 사용하는 시간 또는 빈도를 산출하는 다른 방법을 이용할 수도 있다.
- [0034] 만약 클라이언트 임계값만을 조건으로 하여 고정 IP 포트를 결정한다면 P2P 어플리케이션을 사용하는 임시 IP 포트를 고정 IP 포트로 결정하는 실수를 저지할 수 있다. 따라서 본 발명에서는 이러한 임시 IP 포트를 제거하기 위해 사용기간 조건을 이용한다. P2P 어플리케이션에서 발생하는 트래픽은 짧은 사용기간의 특징을 가지므로, 사용기간 임계값을 이용하여 정확한 고정 IP 포트를 추출할 수 있다.
- [0035] 다음으로 고정 IP 포트를 추출한다(S240). 상기 방치기간 판단 단계(S210), 클라이언트 수 판단 단계(S220) 및 사용기간 판단 단계(S230)를 통해 각각의 임계값 조건을 만족시키는 IP 포트를 고정 IP 포트로 추출한다. 추출된 고정 IP 포트는 IP 주소, 포트 번호 및 전송 프로토콜 정보를 포함할 수 있다. 또한 고정 IP 포트 추출장치(100)가 분석 대상 어플리케이션에 대해 독립적으로 실행될 수 있으므로, 상기 IP 포트를 발생시킨 어플리케이션에 대한 정보도 포함할 수 있다.
- [0036] 고정 IP 포트로 추출된 IP 포트들은 트래픽 분석 또는 검증을 위해 사용될 수 있다. 추출된 고정 IP 포트가 트래픽 분석에 사용되는 경우 상기 고정 IP 포트는 트래픽 분석을 위한 시그니처로 이용될 수 있다. 시그니처 기반의 트래픽 분석 방법은 특정 어플리케이션 트래픽이 다른 어플리케이션들과 구분하여 사용하는 공통 요소를 찾아내 그것을 그 어플리케이션의 시그니처라 명명하고, 동일한 시그니처를 가졌는지 아닌지를 판단하는 방법으로 트래픽을 분석하는 방법이다.
- [0037] IP 포트가 방치기간 임계값, 클라이언트 임계값 및 사용기간 임계값 조건에 의해 고정 IP 포트로 판단되는 상태 흐름에 대해서는 도 4를 통해 상세히 설명한다.
- [0038] 도 3은 본 발명의 일실시예에 따른 IP 포트 등록 단계를 도시한 동작 흐름도이다.
- [0039] 먼저 IP 포트를 감지한다(S300). 어플리케이션 서버 IP 포트 중 클라이언트 호스트와 통신하는 IP 포트를 감지한다. 이 단계에서는 IP 포트 발생 자체에 대한 감지뿐 아니라 IP 포트를 발생시킨 어플리케이션에 대한 정보에 대해서도 감지하는 것을 포함할 수 있다. 예를 들어, 클라이언트 호스트가 어플리케이션A의 로그인 서버 포트번호 5001를 사용하여 어플리케이션 서버에 접속하면, 통신하는 IP 포트 자체 포트번호 5001뿐 아니라 상기

IP 포트를 발생시키는 어플리케이션A에 대한 정보도 감지된다. IP 포트에는 고정 IP 포트뿐 아니라 동적 IP 포트 및 임시 IP 포트도 포함된다. 또한 최초 발생하는 IP 포트뿐 아니라 기존에 이미 발생한 IP 포트일 수 있다.

- [0040] 다음으로 감지된 IP 포트가 기등록된 IP 포트인지 판단한다(S310). 등록된 IP 포트에 대한 정보를 확인할 수 있으므로, 새로 감지되는 IP 포트가 이미 등록된 IP 포트인지 판단할 수 있다. 이미 등록된 IP 포트라면 다음 단계로 넘어간다. 한편, 최초 발생한 IP 포트라면 IP 포트에 등록한다(S320).
- [0041] 다음으로 감지된 IP 포트의 발생이 기존의 어플리케이션에 의한 것인지 판단한다(S330). IP 포트 감지 단계에서 IP 포트를 발생시키는 어플리케이션에 대한 정보도 함께 감지될 수 있으므로, 새로 감지되는 IP 포트가 기존의 어플리케이션에 의해 발생한 것인지 판단할 수 있다. 만약 기존의 어플리케이션에 의한 것이라면 종료한다. 기존의 어플리케이션에 의해 최초 발생한 때 등록되었으므로 다시 등록할 필요가 없기 때문이다.
- [0042] 이 경우 등록된 IP 포트에 대한 새로운 정보를 업데이트하는 단계를 더 포함할 수 있다. 새로운 정보는 상기 IP 포트를 사용하여 통신한 클라이언트 수 및 마지막 사용 시각을 포함할 수 있다. 구체적으로, 새로운 클라이언트 호스트가 상기 IP 포트를 사용하여 통신하는 경우 기존의 IP 포트 정보에 클라이언트 수를 추가할 수 있다. 또는 클라이언트 호스트에 의해 IP 포트가 사용된 마지막 시각 정보도 수정할 수 있다. 업데이트된 IP 포트 정보는 고정 IP 포트 추출을 위한 단계(S220, S230)에서 사용될 수 있다.
- [0043] 한편, IP 포트의 발생이 새로운 어플리케이션에 의해 발생한 것이라면 등록된 IP 포트를 삭제한다(S350). 새로운 어플리케이션이 이미 등록된 IP 포트와 동일한 IP 포트를 이용하여 통신한다면, 등록된 IP 포트는 고정 IP 포트가 아니기 때문이다. 만약 시간 차이를 두고 서로 다른 어플리케이션이 하나의 IP 포트를 사용한다면 등록된 IP 포트는 임시 IP 포트일 수 있다. 또는 동시에 서로 다른 어플리케이션이 하나의 IP 포트를 사용한다면 IP 포트 충돌일 수 있다. 어느 경우에 의하든 등록된 IP 포트는 고정 IP 포트가 아니므로 삭제한다.
- [0044] 도 4는 본 발명의 일실시예에 따라 IP 포트의 상태 전이를 설명하기 위한 상태 흐름도이다.
- [0045] 모든 IP 포트들은 미등록(S400), 등록(S410), 후보자(S420), 고정(S430) 및 지연(S440) 상태 중 어느 하나에 속한다.
- [0046] 최초 모든 IP 포트는 미등록(S400) 상태에 속한다. 아직 감지되지 않아 등록(S410) 이전인 상태이거나, 서로 다른 어플리케이션에 의해 IP 포트가 발생하여 IP 포트가 삭제된 경우 미등록(S400) 상태에 속한다. 하드웨어 측면에서 보면, 미등록(S400) 상태의 IP 포트 정보는 고정 IP 포트 추출장치 메모리 상에 존재하지 않는다.
- [0047] 클라이언트 호스트와 통신하는 IP 포트가 감지되면 미등록(S400) 상태의 IP 포트는 등록(S410) 상태가 된다. 트래픽을 분석하여 하나의 클라이언트와 통신하는 특정 어플리케이션 서버 IP 포트를 등록(S410) 상태로 전이시킬 수 있다. 선택적으로, 미등록(S400) 상태의 IP 포트가 등록(S410) 상태로 전이되는 조건을 부가할 수 있다. 예를 들면, 앞서 도 3에서 설명한 기등록된 IP 포트인지 여부 또는 기존 어플리케이션에 의해 발생한 것인지 여부를 판단하는 단계가 그것이다.
- [0048] 등록(S410) 상태의 IP 포트가 클라이언트 임계값 이상의 클라이언트와 통신한다면 후보자(S420) 상태가 된다. 많은 클라이언트 호스트와 통신하는 IP 포트일수록 고정 IP 포트일 확률이 높다. 따라서 등록(S410) 상태의 IP 포트 중 클라이언트 임계값 이상의 클라이언트 호스트의 개수를 가지는 IP 포트만이 후보자(S420) 상태가 된다.
- [0049] 클라이언트 임계값이 증가하면 입력 데이터 양이 풍부한 고정 IP 포트만 추출되므로, 이를 이용한 트래픽 분석의 정확도(Accuracy)는 증가할 것이다. 하지만, 임계값에 도달하지 못하는 고정 IP 포트들이 추출되지 못해 분석률(Completeness)은 감소할 것이다. 따라서 실험을 통해 최적의 클라이언트 임계값을 정하는 것이 필요하다. 이것과 관련해서는 도 6a에서 상세히 설명한다.
- [0050] 후보자(S420) 상태의 IP 포트가 사용기간 임계값 이상 통신한다면 고정(S430) 상태가 된다. IP 포트의 사용기간이 사용기간 임계값 이상인지 판단하는데, 사용기간이란 일반적으로 IP 포트의 최초 등록 시각으로부터 마지막 사용 시각까지의 기간을 지칭한다. 그러나 클라이언트 호스트 별로 각각 사용 기간을 평균하여 산출하는 등 그밖에 등록된 IP 포트를 사용하는 시간 또는 빈도를 산출할 수 있는 다른 기간을 지칭할 수도 있다.
- [0051] 만약 클라이언트 임계값만을 이용하여 고정 IP 포트를 결정한다면 P2P 어플리케이션을 사용하는 임시 IP 포트를 고정 IP 포트에 결정하는 실수를 저지를 수 있다. 따라서 본 발명에서는 이러한 임시 IP 포트를 제거하기 위해 사용기간 조건을 이용한다. P2P 어플리케이션에서 발생하는 트래픽은 짧은 사용기간의 특징을 가지므로, 사용기간 임계값을 이용하여 정확한 고정 IP 포트를 추출할 수 있다. 사용기간 임계값을 크게 잡으면 잡을수록 임

시 IP 포트 제거는 달성할 수 있어서 정확도는 증가할 것이나, 임계값에 도달하지 못하는 고정 IP 포트들이 추출되지 못해 분석률이 감소하고 시스템 오버헤드도 증가할 것이다. 따라서 실험을 통해 최적의 클라이언트 임계값을 정한 것이 필요하다. 이것과 관련해서는 도 6b에서 상세히 설명한다.

- [0052] 고정(S430) 상태의 IP 포트들이 본 발명에서 제안하는 고정 IP 포트이다. 고정 IP 포트는 트래픽 분석 및 검증을 위한 정보로 사용될 수 있다.
- [0053] 등록(S410), 후보자(S420) 및 고정(S430) 상태의 IP 포트들은 충돌 발생 여부가 판단된다. 충돌은 동시에 서로 다른 어플리케이션에 의해 하나의 IP 포트가 공유되는 것을 지칭한다. 충돌을 일으킨 IP 포트는 고정 IP 포트가 아니기 때문에 지연(S440) 상태를 거쳐 미등록(S400) 상태로 전이된다. 이 단계에서 등록된 IP 포트를 삭제시키는 단계를 더 포함할 수 있다. 선택적으로 고정(S430) 상태의 IP 포트에 대해서는 충돌 발생 여부를 판단하지 않을 수 있다.
- [0054] 충돌이 발생한 IP 포트는 지연(S440) 상태가 된다. 지연(S440) 상태는 IP 포트가 충돌 발생 후 미등록(S400) 상태가 되기 전 중간 단계로, 일정 시간이 지연된다. 만약 충돌이 발생한 IP 포트가 지연(S440) 상태를 거치지 않고 바로 미등록(S400) 상태로 전이된다면 다시 곧바로 등록(S410) 상태가 되어 미등록(S400) 상태와 등록(S410) 상태를 무한히 반복하여 전이될 수 있다. 지연(S440) 상태를 설정함으로써 불필요한 전이를 막을 수 있으며, 시스템 오버헤드 발생도 감소시킬 수 있다. 다만 다른 알고리즘의 부가 또는 필요에 따라 지연(S440) 상태를 제외할 수도 있다.
- [0055] 미등록(S400)을 제외한 등록(S410), 후보자(S420), 고정(S430) 및 지연(S440) 상태에는 각각 타임아웃(time out)이 존재한다. 타임아웃은 임계값을 이용하여 일정시간 상태의 변화가 없는 IP 포트를 미등록(S400) 상태로 전이시키는 것이다.
- [0056] 등록(S410), 후보자(S420) 및 고정(S430) 상태의 모든 IP 포트들은 클라이언트 임계값과 사용기간 임계값을 넘지 못하면 계속 같은 상태로 남아있게 된다. 하드웨어 측면에서 보면 같은 상태로 남아있는 IP 포트 정보들이 계속 메모리상에 존재하게 되는 것이다. 이 상태에서 만약 새로 등록되는 IP 포트만 점점 늘어난다면 물리적인 메모리의 제한으로 인해 메모리가 포화될 것이다. 따라서 등록 요건을 갖춘 모든 IP 포트 정보들을 저장하지 못하게 되는 문제가 생기고, 무한히 많은 데이터들은 고정 IP 포트 추출 시 높은 오버헤드를 발생하여 실시간 분석을 불가능하게 할 수 있다. 이러한 문제들을 해결하기 위한 방안으로 방치기간 임계값을 설정한다.
- [0057] 방치기간 임계값은 IP 포트의 각 상태마다 다르게 정의될 수 있다. 등록(S410) 상태에서는 등록 타임아웃, 후보자(S420) 상태에서는 후보자 타임아웃, 고정(S430) 상태에서는 고정 타임아웃, 지연(S440) 상태에서는 지연 타임아웃이 그것이다. 각 타임아웃 값은 서로 다를 수 있고, 의도에 따라 특정 상태의 타임아웃은 포함하지 않을 수 있다. 미등록(S400) 상태를 제외한 모든 상태에서 각각 타임아웃이 되면 미등록(S400) 상태로 전이된다. 이 경우 등록된 IP 포트를 삭제하는 단계를 더 포함할 수 있다.
- [0058] 도 5는 본 발명의 일실시예에 따른 고정 IP 포트 추출 알고리즘을 설명하기 위한 도면이다.
- [0059] 도 5를 참조하면 코드는 크게 두 부분으로 나뉜다. 새로운 IP 포트의 등록에 관한 알고리즘과, 등록되어 있는 IP 포트 중 고정 IP 포트를 추출하기 위한 알고리즘이 그것이다.
- [0060] 코드에서 IPT는 IP 포트 정보가 기록되는 영역을 의미하고, 도 4와 비교하여 Discard 상태는 미등록(S400) 상태, New 상태는 등록(S410) 상태, Candidate 상태는 후보자(S420) 상태, Fixed 상태는 고정(S430) 상태, Time_Wait는 지연(S440) 상태를 뜻한다.
- [0061] 새로운 IP 포트의 등록에 대해서는 3번째 줄부터 7번째 줄까지 기록되어 있다. 먼저 새로운 IP 포트가 IPT에 등록되어 있는지 판단하여 새로 등록 시키거나, 이미 등록되어 있다면 새로운 정보를 업데이트 한다. 업데이트 되는 정보는 충돌 여부, 클라이언트의 수 또는 마지막 사용 시각일 수 있다. 충돌은 동시에 서로 다른 어플리케이션에 의해 하나의 IP 포트가 공유되는 것으로, 충돌 여부는 이미 등록된 IP 포트의 발생이 기존의 어플리케이션에 의한 것인지를 통해 판단할 수 있다. 만약 동일한 어플리케이션에 의해 발생한 것이라면 충돌이 아닌 것으로 판단된다. 이 경우, 마지막 사용 시각 정보를 업데이트할 수 있다.
- [0062] 고정 IP 포트를 추출에 대해서는 8번째 줄부터 23번째 줄까지 기록되어 있다. IPT에 등록되어 있는 IP 포트들을 앞서 설명한 도 4에 따라 미등록(S400), 등록(S410), 후보자(S420), 고정(S430) 또는 지연(S440) 상태로 전이시킨다.
- [0063] 일실시예로 등록(S410) 상태의 경우, 상기의 충돌 여부 판단에 의해 충돌이 발생한 경우 지연(S440) 상태로 전

이시킨다. 충돌이 발생하지 않았다면 클라이언트 수를 체크하여 임계값 이상인 경우 후보자(S420) 상태로 전이시킨다. 만약 임계값 미만인 경우 방치기간을 체크하여 임계값 이상인 경우 미등록(S400) 상태로 전이시킨다. 후보자(S420) 상태, 고정(S430) 상태 및 지연(S440) 상태에 대해서도 상기와 유사한 판단을 거친 후 전이될 수 있으며, 이는 도 4에서 설명하였다.

[0064] 도 6a는 본 발명의 일실시예에 따라 클라이언트 임계값에 따른 분석률 및 정확도 변화를 나타낸 도면이다.

[0065] x축은 클라이언트 임계값을 나타내고, y축은 분석률 및 정확도를 나타낸다.

[0066] 도 6a에서 확인하듯 클라이언트 임계값은 고정 IP 포트 추출에 있어 정확도 및 분석률에 영향을 준다. 분석률은 전체 트래픽 중에 트래픽 분석 장치(400)가 분석한 트래픽 비율을 나타낸다. 정확도는 트래픽 분석 장치(400)가 분석한 트래픽 중 정확하게 분석한 트래픽 비율을 나타낸다. 얼마나 정확하고 많은 수의 고정 IP 포트를 추출하는가가 고정 IP 포트를 기반으로 하는 트래픽 분석의 분석률 및 정확도를 좌우한다.

[0067] 클라이언트 임계값이 높을수록, 즉 많은 수의 클라이언트 호스트와 통신하는 고정 IP 포트일수록 더 높은 정확도를 보인다. 하지만 클라이언트 임계값이 높아질수록 추출되는 고정 IP 포트의 개수가 줄어들어 분석률은 떨어진다. 따라서 실험을 통해 적절한 임계값을 설정할 필요가 있다. 실험 결과를 도시한 도 6a를 참조하면, 적절한 클라이언트 임계값은 2이다. 클라이언트 임계값이 1인 경우 정확도가 보장되지 않고, 클라이언트 임계값이 3 이상인 경우 분석률이 보장되지 않기 때문이다.

[0068] 도 6b 본 발명의 일실시예에 따라 사용기간 임계값에 따른 IP 포트 개수 변화를 나타낸 도면이다.

[0069] x축은 사용기간 임계값을 나타내고, y축은 IP 포트 개수를 나타낸다.

[0070] 앞서 설명했듯 사용기간 임계값은 고정 IP 포트 추출에 있어 불필요한 임시 IP 포트를 제거하는 데에 목적이 있다. 임시 IP 포트는 짧은 사용기간을 특징으로 하기 때문에 적절한 사용기간 임계값을 설정하여 고정 IP 포트와 구별할 수 있다. 적절한 임계값을 설정하기 위해 동일한 데이터를 대상으로 다양한 임계값을 적용시켜 총 IP 포트의 개수, 임시 IP 포트 및 임시 IP 포트를 포함한 고정 IP 포트의 개수를 보여준다.

[0071] 도 6b를 참조하면, 적절한 사용기간 임계값은 9시간이다. 9시간보다 작게 설정한다면 임시 IP 포트를 다 제거할 수 없고, 9시간보다 더 크게 설정한다면 고정 IP 포트 추출 시스템의 오버헤드를 증가시킬 수 있기 때문이다. 그러나, 사용기간 임계값은 분석 대상 네트워크 및 분석 시스템 환경에 맞게 재조정될 수 있다. 다만, 사용기간 임계값이 임시 IP 포트가 더 이상 추출되지 않는 때의 값이라는 것은 어느 환경에서나 동일할 수 있다.

[0072] 도 6c는 본 발명의 일실시예에 따라 방치기간 임계값에 따른 IP 포트 개수 변화를 나타낸 도면이다.

[0073] x축은 방치기간 임계값을 나타내고, y축은 IP 포트 개수를 나타낸다.

[0074] 앞서 설명했듯 방치기간 임계값은 추출 알고리즘이 실행될 때 IP 포트가 클라이언트 임계값 또는 사용기간 임계값 조건을 만족하지 못한 상태에서 영원히 잔류되는 것을 방지한다. 방치기간 임계값을 너무 작게 하면 추출되는 고정 IP 포트 수에 손실이 생기고, 너무 크게 할 경우 방치기간 임계값을 설정한 실익이 없게 된다.

[0075] 도 6c를 참조하면, 방치기간 임계값을 48시간으로 정하였을 때 고정 IP 포트의 손실 없이 전체 IP 포트의 유지량을 줄일 수 있었다. 방치기간 임계값은 분석 대상 네트워크 또는 분석 시스템 환경에 맞게 재조정될 수 있다. 다만, 고정 IP 포트가 손실 없이 유지되기 시작하는 시점이 방치기간 임계값이 된다는 것은 어느 환경에서나 동일할 수 있다.

[0076] 도 7은 본 발명의 일실시예에 따른 고정 IP 포트를 이용한 트래픽 분석 장치를 도시한 블록도이다.

[0077] 도 7을 참조하면, 트래픽 분석 장치(400)는 고정 IP 포트 추출장치(100), 분석장치(200) 및 검증장치(300)를 포함한다. 화살표는 데이터의 흐름을 나타낸다. 한편, 본 실시예에서 고정 IP 포트 추출장치(100)는 이전의 실시예와 그 구성 및 작동 과정이 동일하므로 이에 대한 중복되는 설명은 생략하기로 한다.

[0078] 분석장치(200)는 분석 대상 네트워크 링크에서 패킷을 수집하는 패킷 캡처부(210), 패킷 캡처부(210)에서 수집한 패킷들을 이용하여 플로우를 만드는 플로우 생성부(220) 및 플로우 생성부(220)에서 생성된 플로우 헤더와 상기 고정 IP 포트 추출장치(100)에서 추출된 고정 IP 포트 헤더를 비교하는 헤더 비교부(230)를 포함한다.

[0079] 패킷 캡처부(210)는 네트워크 모니터링 지점을 지나는 패킷을 캡처한다. 캡처된 패킷이 단편화되지 않은 패킷 같은 일반 패킷일 경우 이 일반 패킷의 헤더를 분석해 트래픽 분석에 필요한 패킷 헤더 정보를 추출하여 플로우

생성부(220)에 인가한다. 만약 캡처된 패킷이 단편화된 패킷일 경우 상기 패킷을 재조립하는 장치를 더 포함할 수 있다.

- [0080] 플로우 생성부(220)는 패킷 캡처부(210)로부터 제공되는 패킷 헤더 정보 내의 특성을 결정짓는 SrcIP(Source IP), DstIP(Destination IP), SRCPort(Source Port), DstPort(Destination Port) 및 프로토콜을 검사한다. 입력되는 패킷이 기존에 존재하던 플로우의 패킷이면 해당 플로우에 추가되는 패킷 수 및 플로우 길이를 갱신시킨다. 반면에 기존에 존재하지 않던 패킷이 입력된다면 새로운 플로우를 생성시킨다.
- [0081] 헤더 비교부(230)는 플로우 생성부(220)에서 생성된 헤더와 고정 IP 포트 추출장치(100)에서 추출한 고정 IP 포트의 헤더를 비교함으로써 트래픽을 분석한다. 단순한 헤더 비교이므로 분석 속도가 빨라 실시간으로 처리될 수 있어 바람직하다.
- [0082] 검증장치(300)는 네트워크에서 분석 대상 어플리케이션의 트래픽 정보를 수집하는 검증용 트래픽 수집부(310), 검증용 트래픽 수집부(310)에서 수집된 트래픽 정보와 플로우 생성부(220)에서 생성된 플로우 데이터를 이용하여 정답지 데이터(ground truth)를 만드는 정답지 데이터 생성부(320) 및 정답지 데이터 생성부(320)로부터 만들어진 정답지 데이터와 분석장치(200)의 트래픽 분석 결과를 비교하여 트래픽 분석 결과를 검증하는 매칭부(330)를 포함한다.
- [0083] 검증용 트래픽 수집부(310)는 네트워크 상에서 트래픽 정보를 수집한다. 수집하는 트래픽 정보는 분석 대상 어플리케이션에 관한 트래픽 정보로서, 네트워크 소켓 정보, 해당 소켓을 생성 시킨 프로세스의 이름 및 프로세스가 설치된 경로를 포함할 수 있다.
- [0084] 검증용 트래픽 수집부(310)는 고정 IP 포트 추출장치(100)의 트래픽 수집부(110)와 같은 구성 및 기능을 가질 수 있다. 따라서 검증용 트래픽 수집부(310)는 실제 트래픽이 발생하는 종단 호스트에 설치될 수 있으며, 실시간으로 발생하는 트래픽 정보를 수집할 수 있는 모든 형태의 하드웨어, 소프트웨어 또는 하드웨어와 소프트웨어의 조합일 수 있다. 또한 검증용 트래픽 수집부(310)는 동일한 구성을 갖는 다수의 장치로 구성될 수 있다.
- [0085] 정답지 데이터 생성부(320)는 검증용 트래픽 수집부(310)에서 수집된 트래픽 정보와 플로우 생성부(220)에서 생성된 플로우 데이터를 이용하여 정답지 데이터를 생성한다. 정답지 데이터는 트래픽 분석 결과를 검증하기 위한 기준이 된다.
- [0086] 매칭부(330)는 분석장치(200)의 트래픽 분석 결과와 정답지 데이터를 비교하여 트래픽 분석 결과를 검증한다. 매칭부(330)가 검증하는 평가 요소는 분석률 및 정확도를 포함할 수 있다. 분석률은 전체 트래픽 중에 트래픽 분석 장치(400)가 분석한 트래픽 비율을 나타낸다. 정확도는 트래픽 분석 장치(400)가 분석한 트래픽 중 정확하게 분석한 트래픽 비율을 나타낸다.
- [0087] 한편, 본 발명에 적용되는 네트워크는 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.
- [0088] 참고로, 본 발명의 실시예에 따른 도 1 및 도 7에 도시된 구성 요소들은 소프트웨어, FPGA(Field Programmable Gate Array) 또는 ASIC(Application Specific Integrated Circuit)와 같은 하드웨어 구성 요소를 의미하며, 소정의 역할들을 수행한다.
- [0089] 그렇지만 '구성 요소들'은 소프트웨어 또는 하드웨어에 한정되는 의미는 아니며, 각 구성 요소는 어드레싱할 수 있는 저장 매체에 있도록 구성될 수도 있고 하나 또는 그 이상의 프로세서들을 재생시키도록 구성될 수도 있다.
- [0090] 따라서, 일 예로서 구성 요소는 소프트웨어 구성 요소들, 객체지향 소프트웨어 구성 요소들, 클래스 구성 요소들 및 태스크 구성 요소들과 같은 구성 요소들과, 프로세스들, 함수들, 속성들, 프로시저들, 서브루틴들, 프로그램 코드의 세그먼트들, 드라이버들, 펌웨어, 마이크로 코드, 회로, 데이터, 데이터베이스, 데이터 구조들, 테이블들, 어레이들 및 변수들을 포함한다.
- [0091] 구성 요소들과 해당 구성 요소들 안에서 제공되는 기능은 더 작은 수의 구성 요소들로 결합되거나 추가적인 구성 요소들로 더 분리될 수 있다.
- [0092] 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 관독가능 매체는 컴퓨터 저장 매체 및 통신 매체를 모두 포함할 수 있다. 컴퓨터 저장 매체는 컴퓨터 관독가능 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 관독가능 명령어, 데이터 구조, 프로그램 모듈, 또는 반송파와 같은 변조된 데이터 신호의 기타 데이터, 또는 기타 전송 메커니즘을 포함하며, 임의의 정보 전달 매체를 포함한다.

[0093] 본 발명의 방법 및 시스템은 특정 실시예와 관련하여 설명되었지만, 그것들의 구성 요소 또는 동작의 일부 또는 전부는 범용 하드웨어 아키텍처를 갖는 컴퓨터 시스템을 사용하여 구현될 수 있다.

[0094] 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.

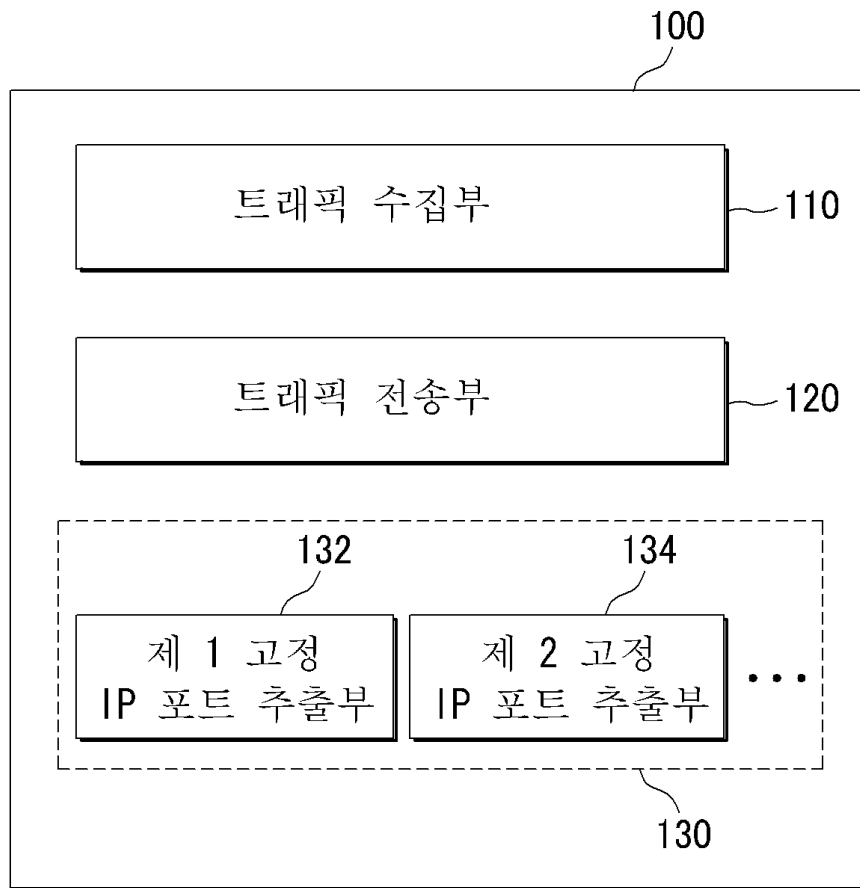
[0095] 본 발명의 범위는 상기 상세한 설명보다는 후술하는 특허청구범위에 의하여 나타내어지며, 특허청구범위의 의미 및 범위 그리고 그 균등 개념으로부터 도출되는 모든 변경 또는 변형된 형태가 본 발명의 범위에 포함되는 것으로 해석되어야 한다.

부호의 설명

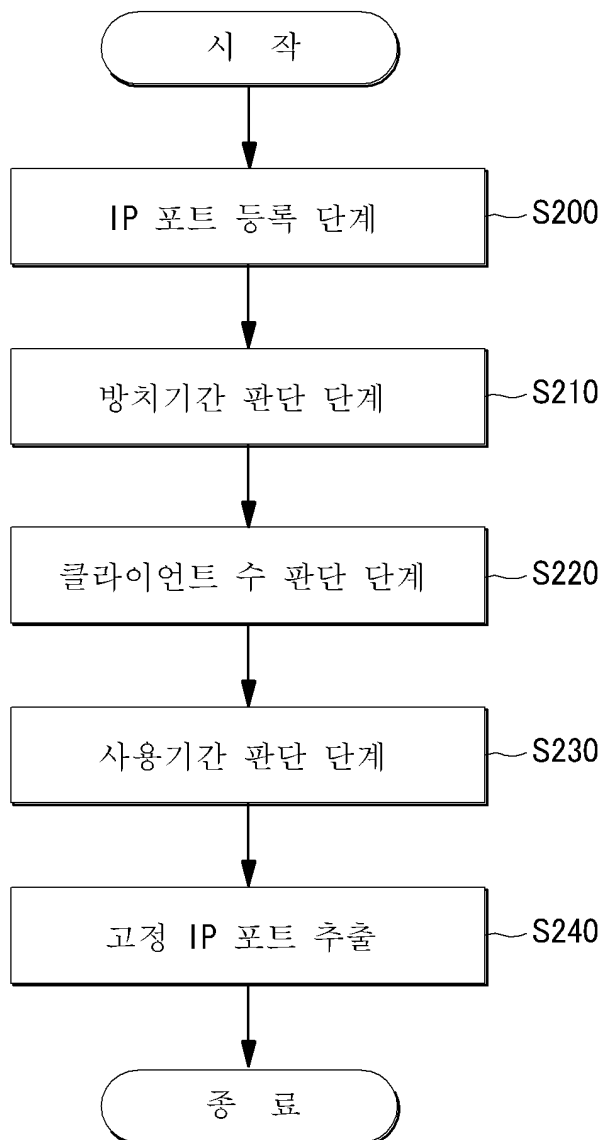
[0096]	100: 고정 IP 포트 추출장치	110: 트래픽 수집부
	120: 트래픽 전송부	130: 고정 IP 포트 추출부
	132: 제 1 고정 IP 포트 추출부	134: 제 2 고정 IP 포트 추출부
	200: 분석장치	210: 패킷 캡처부
	220: 플로우 생성부	230: 헤더 비교부
	300: 검증장치	310: 검증용 트래픽 수집부
	320: 정답지 데이터 생성부	330: 매칭부
	400: 트래픽 분석 장치	

도면

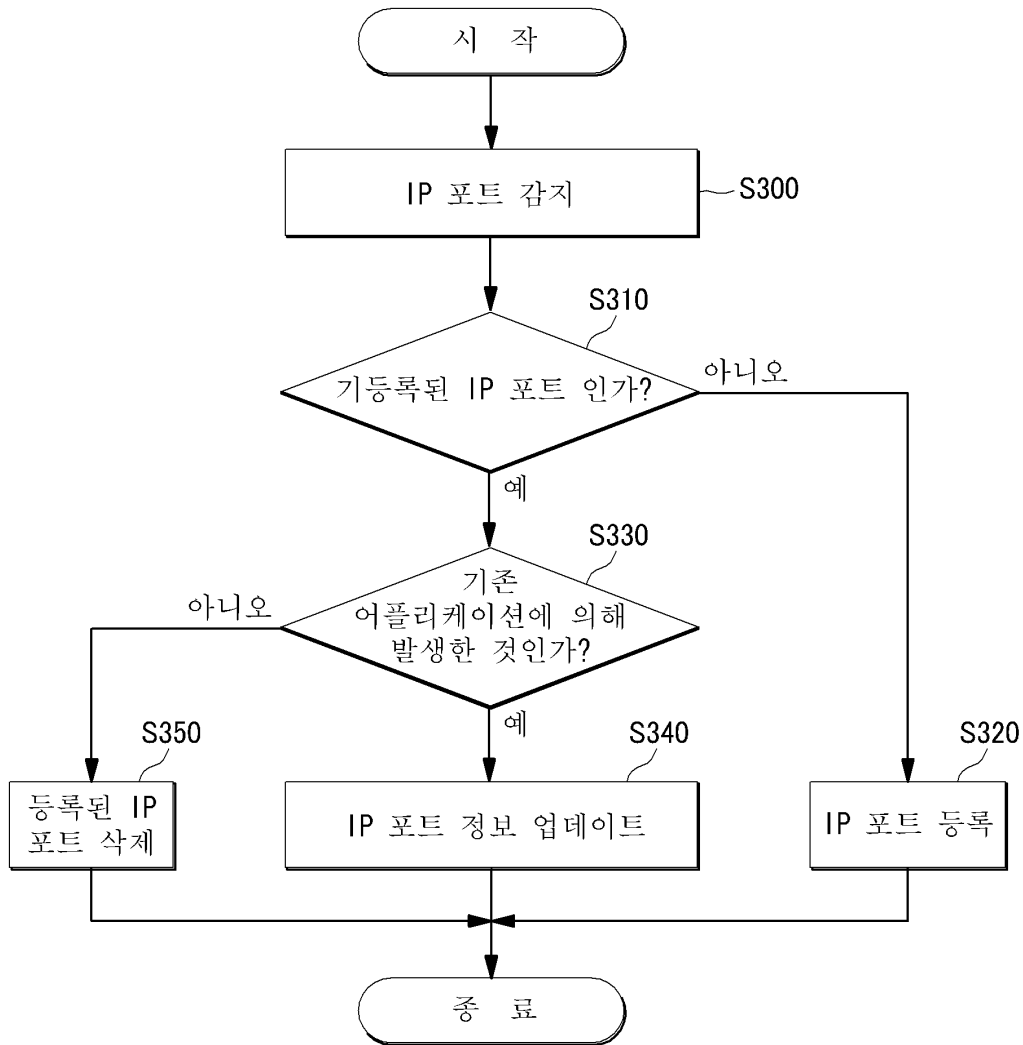
도면1



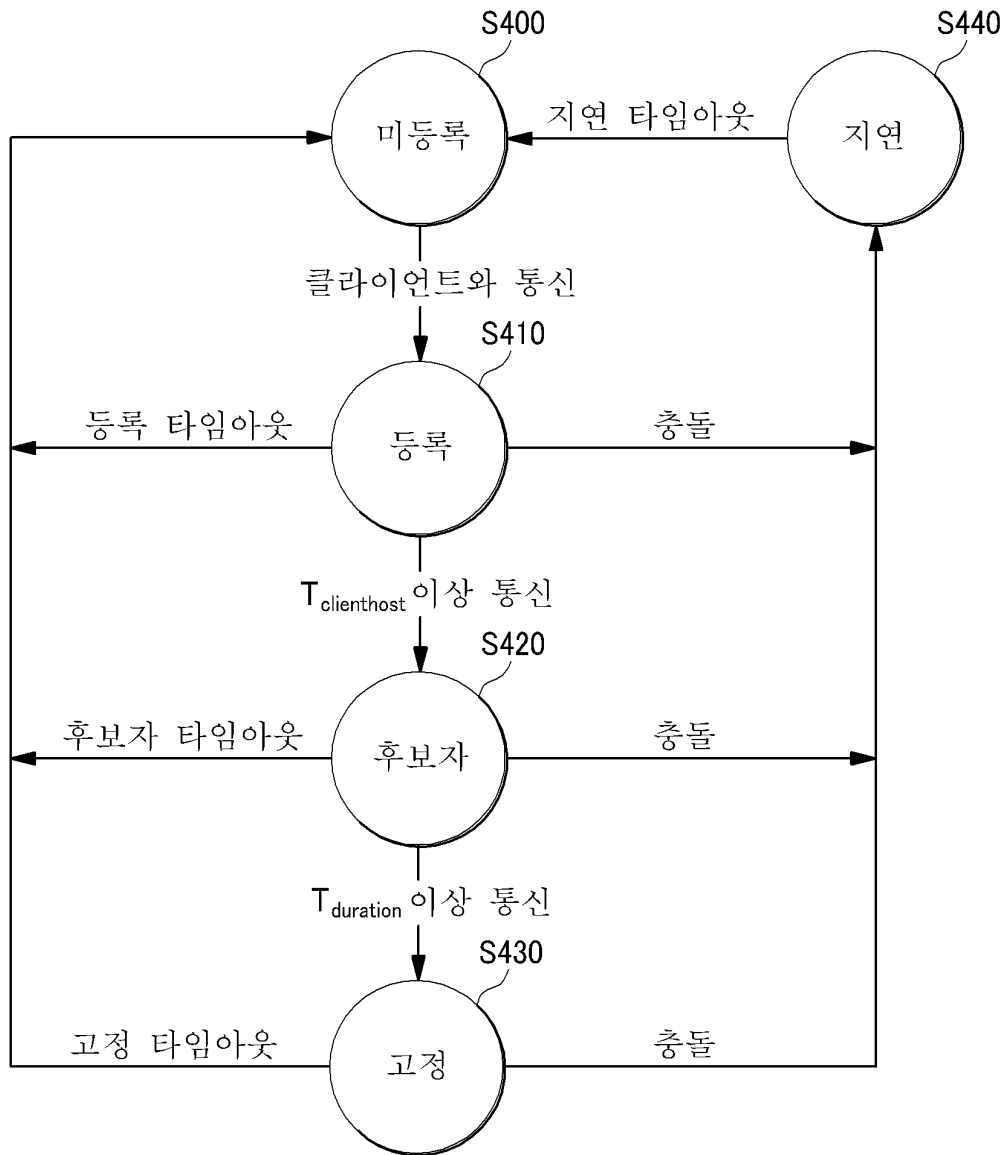
도면2



도면3



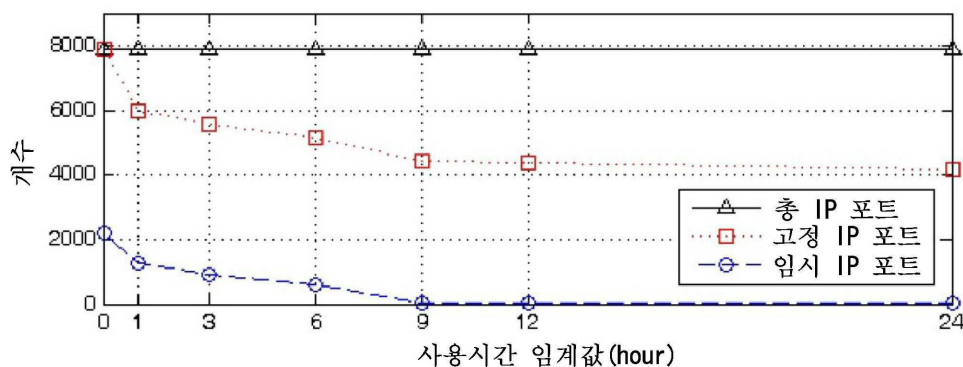
도면4



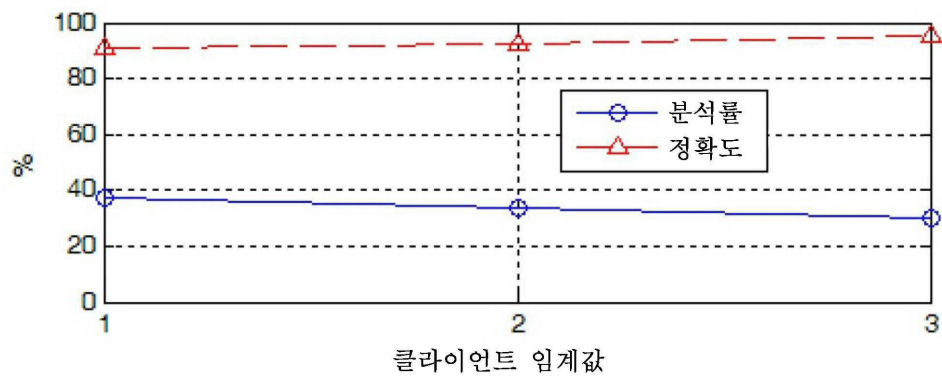
도면5

1:	procedure Fixed IP-port Extraction for an Application A
2:	IPT $\leftarrow$ IP-port Table
3:	for each input flow record do
4:	search the corresponding IP-port record in IPT
5:	if found in IPT then update the status of the record;
6:	else add a new record with the state as New ;
7:	end for
8:	for each flow record in IPT do
9:	Check the state of the record
10:	if state == New then
11:	if conflict then move to Time_Wait ;
12:	else if over $T_{clienthost}$ then move to Candidate ;
13:	else if over $T_{timelimit}$ then move to Discard ;
14:	else if State == Candidate then
15:	if conflict then move to Time_Wait ;
16:	else if over $T_{duration}$ then move to Fixed ;
17:	else if over $T_{timelimit}$ then move to Discard ;
18:	else if State == Fixed then
19:	if conflict then move to Time_Wait ;
20:	else if over $T_{timelimit}$ then move to Discard ;
21:	else if State == Time_Wait then
22:	if over $T_{timelimit}$ then move to Discard ;
23:	end for
24:	end procedure

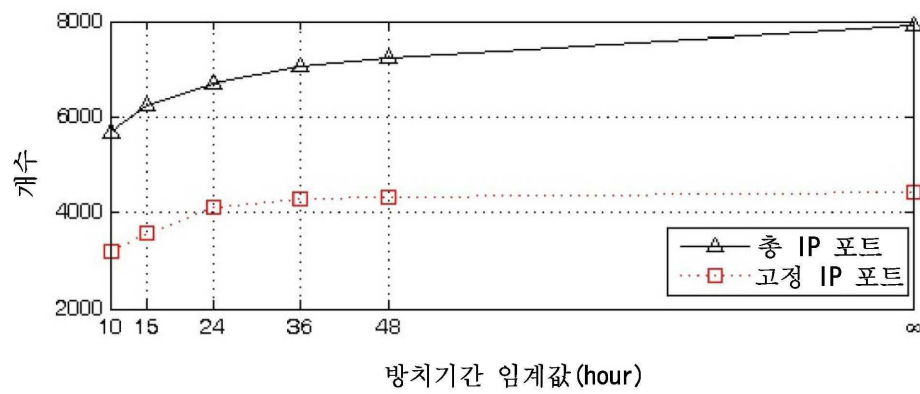
도면6a



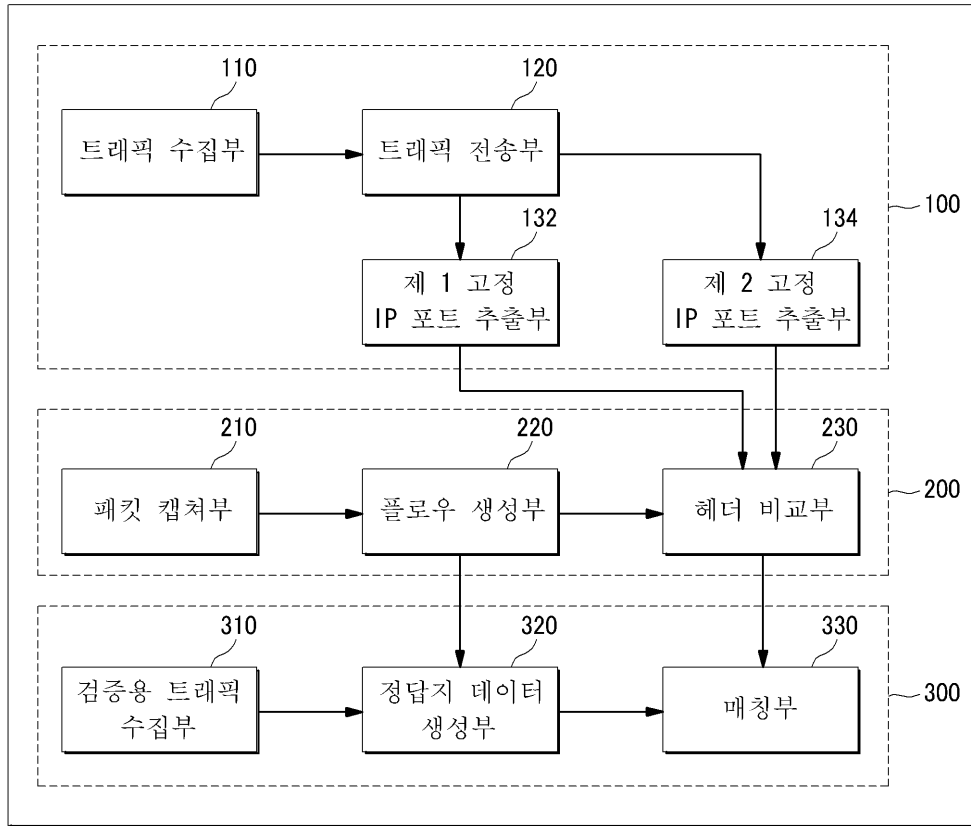
도면6b



도면6c



도면7



400