



(19) 대한민국특허청(KR)

(12) 공개특허공보(A)

(11) 공개번호 10-2014-0093317

(43) 공개일자 2014년07월28일

(51) 국제특허분류(Int. Cl.)

H04L 12/851 (2013.01)

(21) 출원번호 10-2013-0003965

(22) 출원일자 2013년01월14일

심사청구일자 2013년01월14일

(71) 출원인

주식회사 시큐아이

서울특별시 강남구 테헤란로83길 15, 1호 ~9층 (삼성동, 제이엠빌딩)

고려대학교 산학협력단

서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)

(72) 발명자

김명섭

충청남도 연기군 조치원읍 신흥리 푸르지오1차 111-902

윤성호

경상북도 포항시 북구 우현동 금호어울림아파트 103-702

조원용

서울 강남구 테헤란로83길 15, 5층 (삼성동, JM빌딩)

(74) 대리인

특허법인다나

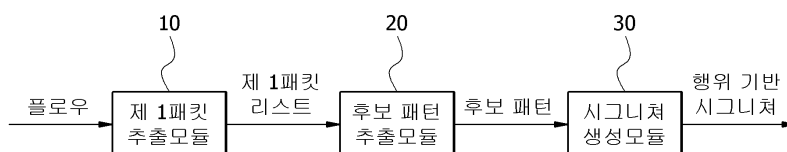
전체 청구항 수 : 총 12 항

(54) 발명의 명칭 행위기반 시그니처 생성 시스템 및 방법

(57) 요약

행위 기반 시그니처 생성 시스템 및 방법이 개시된다. 상기 행위 기반 시그니처 생성 시스템은 입력받은 플로우들의 첫 번째 패킷을 추출하여 일련의 순서로 조합한 제1패킷 리스트를 생성하는 제1 패킷 추출 모듈; 상기 제1 패킷 리스트에 포함된 패킷을 이용하여 조합 가능한 후보 패턴을 추출하는 후보 패턴 추출 모듈 및 상기 후보 패턴 중에서 기 설정된 임계값을 초과하는 패턴을 추출하여 시그니처를 생성하는 시그니처 생성 모듈을 포함한다.

대표도 - 도1



특허청구의 범위

청구항 1

입력받은 플로우들의 첫 번째 패킷을 추출하여 일련의 순서로 조합한 제1패킷 리스트를 생성하는 제1 패킷 추출 모듈;

상기 제1패킷 리스트에 포함된 패킷을 이용하여 조합 가능한 후보 패턴을 추출하는 후보 패턴 추출 모듈 및

상기 후보 패턴 중에서 기 설정된 임계값을 초과하는 패턴을 추출하여 시그니처를 생성하는 시그니처 생성 모듈을 포함하는 행위 기반 시그니처 생성 시스템.

청구항 2

제1항에 있어서,

상기 제1패킷 추출 모듈은 입력받은 플로우 중 페이로드가 존재하는 플로우의 첫 번째 요청 패킷을 추출하여 상기 제1패킷 리스트를 생성하는 행위 기반 시그니처 생성 시스템.

청구항 3

제1항에 있어서,

상기 제1패킷 추출 모듈은 상기 첫 번째 패킷을 발생 시간 순으로 정렬하여 상기 제1패킷 리스트를 생성하는 행위 기반 시그니처 생성 시스템.

청구항 4

제1항에 있어서,

상기 후보 패턴 추출 모듈은 상기 제1패킷 리스트에서 일정 인터벌내에 존재하는 패킷들을 이용하여 상기 후보 패턴을 추출하는 행위 기반 시그니처 생성 시스템.

청구항 5

제4항에 있어서,

상기 후보 패턴 추출 모듈은 기 설정된 행위 기반 시그니처 모델을 기반으로 상기 제1패킷 리스트에서 일정 인터벌내에 존재하는 패킷에서 트래픽의 특징을 추출하여 구성된 엔트리를 복수개 조합하여 상기 후보 패턴을 추출하는 행위 기반 시그니처 생성 시스템.

청구항 6

제5항에 있어서,

상기 행위 기반 시그니처 모델은 타입, 인터벌 및 엔트리 구성요소를 포함하는 행위 기반 시그니처 생성 시스템.

청구항 7

제1항에 있어서,

상기 시그니처 생성 모듈은 최대 패턴 크기, 최대 인터벌, 최소 발생 호스트 갯수 및 최소 발생 횟수 중 하나 이상을 이용하여 상기 임계값을 설정하는 행위 기반 시그니처 생성 시스템.

청구항 8

제1항에 있어서,

상기 제1패킷 추출 모듈은 복수개의 호스트로부터 플로우를 입력받는 행위 기반 시그니처 생성 시스템.

청구항 9

제1패킷 추출 모듈이 입력받은 플로우들의 첫 번째 패킷을 추출하여 일련의 순서로 조합한 제1패킷 리스트를 생성하는 단계;

후보 패턴 추출 모듈이 상기 제1패킷 리스트에 포함된 패킷을 이용하여 조합 가능한 후보 패턴을 추출하는 단계 및

시그니처 생성 모듈이 상기 후보 패턴 중에서 기 설정된 임계값을 초과하는 패턴을 추출하여 시그니처를 생성하는 단계를 포함하는 행위 기반 시그니처 생성 방법.

청구항 10

제9항에 있어서, 상기 제1패킷 리스트를 생성하는 단계는

입력받은 플로우를 메모리에 로드하는 단계;

상기 메모리에 로드된 플로우가 리스트 생성 대상 플로우인지 판단하는 단계 및

리스트 생성 대상 플로우의 첫 번째 패킷을 추출하여 일련의 순서로 조합하는 단계를 포함하는 행위 기반 시그니처 생성 방법.

청구항 11

제10항에 있어서,

상기 리스트 생성 대상 플로우인지 판단하는 단계는 입력받은 플로우의 요청 패킷 포함 여부를 기준으로 하는 행위 기반 시그니처 생성 방법.

청구항 12

제9항에 있어서, 상기 후보 패턴을 추출하는 단계는

상기 제1패킷 리스트 상에서 인터벌을 설정하는 단계,

기 설정된 행위 기반 시그니처 모델을 기반으로 상기 인터벌내에 존재하는 패킷에서 트래픽의 특징을 추출하여 엔트리를 구성하는 단계 및

상기 엔트리를 복수개 조합하여 상기 후보 패턴을 추출하는 단계를 포함하는 행위 기반 시그니처 생성 방법.

명세서

기술분야

[0001] 본 발명은 행위기반 시그니처 생성 시스템 및 방법에 관한 것으로, 더욱 상세하게는 응용 트래픽 분석을 위한 시그니처를 생성하기 위한 행위기반 시그니처 생성 시스템 및 방법에 관한 것이다.

배경 기술

- [0002] 오늘날의 네트워크에서는 다양한 응용의 등장으로 인해 트래픽이 복잡 다양해지고 있다. 이러한 상황 속에서 네트워크의 효과적인 운용과 관리를 위해 네트워크 트래픽 분석은 필수적인 요소가 되었다. 네트워크의 의존도가 앞으로 계속 증대됨에 따라 트래픽 분석의 중요성은 계속 커질 전망이다.
- [0003] 엔터프라이즈 네트워크나 캠퍼스 네트워크에서는 네트워크 자원의 효율적인 운용 및 관리를 위해서 다양한 정책들을 마련하고 있다. 예를 들어, 학교나 공공 기관에서 P2P 트래픽, 게임 트래픽 등 네트워크 자원을 많이 차지하거나 업무와 관련이 없는 트래픽들을 제한하는 정책들이 있다. 이러한 정책들을 펼치기 위해서는 빠르고 정확한 응용계층 트래픽 분류가 필수적인 요소이다.
- [0004] 응용계층 트래픽 분류란 네트워크 패킷을 수집하여 해당 패킷을 발생시킨 응용계층의 정체를 알아내고 이를 기준으로 트래픽을 분류하는 것을 말한다.
- [0005] 실시간의 정확한 응용계층 트래픽 분류의 중요성에 의해 현재까지 많은 연구들이 진행되어 왔다. 기존의 분류 방법들로는 포트 기반 분석, 페이로드 시그니처 기반 분석, 머신러닝 기반 분석, 플로우 상관관계 기반 분석 등 다양한 방법론들이 존재한다.
- [0006] 페이로드 시그니처 기반 분석은 가장 보편적으로 사용되고 있는 분석 방법으로, 특정 응용을 사용할 때 패킷의 페이로드에서 공통적으로 관찰되는 스트링을 페이로드 시그니처로 생성하고 이를 이용하여 트래픽을 분석하는 방법이다.
- [0007] 그러나 페이로드 시그니처 기반 분석 방법은 시스템 오버헤드 문제를 발생시키고, 암호화 된 트래픽 분석에는 적합하지 못하다는 문제가 있다.
- [0008] 또한, 새로운 응용 프로그램의 등장, 응용 프로그램의 업데이트 등 응용의 변화에 따라 트래픽이 변하고 있기 때문에 오늘날의 응용 별 트래픽 분류는 더욱 더 힘들어지고 있다. 따라서, 보다 정확한 응용 별 트래픽 분류를 위하여 다양한 분류 방법론에 대한 지속적인 연구가 필요하다.

발명의 내용

해결하려는 과제

- [0009] 본 발명이 이루고자 하는 기술적 과제는 응용 실행시 발생하는 응용 트래픽을 분석하기 위하여 복수개의 플로우를 이용하여 시그니처를 생성하고, 또한 복수개의 호스트로부터 발생하는 플로우를 이용하여 시그니처를 생성하기 때문에 정확한 응용 트래픽 분석을 위한 시그니처를 생성할 있는 행위기반 시그니처 생성 시스템 및 방법을 제공하는 데 있다.

과제의 해결 수단

- [0010] 본 발명의 일 양태에 따르면 입력받은 플로우들의 첫 번째 패킷을 추출하여 일련의 순서로 조합한 제1패킷 리스트를 생성하는 제1 패킷 추출 모듈; 상기 제1패킷 리스트에 포함된 패킷을 이용하여 조합 가능한 후보 패턴을 추출하는 후보 패턴 추출 모듈 및 상기 후보 패턴 중에서 기 설정된 임계값을 초과하는 패턴을 추출하여 시그니처를 생성하는 시그니처 생성 모듈을 포함하는 행위 기반 시그니처 생성 시스템을 제공한다.
- [0011] 상기 제1패킷 추출 모듈은 입력받은 플로우 중 페이로드가 존재하는 플로우의 첫 번째 요청 패킷을 추출하여 상기 제1패킷 리스트를 생성할 수 있다.
- [0012] 상기 제1패킷 추출 모듈은 상기 첫 번째 패킷을 발생 시간 순으로 정렬하여 상기 제1패킷 리스트를 생성할 수 있다.

- [0013] 상기 후보 패턴 추출 모듈은 상기 제1패킷 리스트에서 일정 인터벌내에 존재하는 패킷들을 이용하여 상기 후보 패턴을 추출할 수 있다.
- [0014] 상기 시그니처 생성 모듈은 최대 패턴 크기, 최대 인터벌, 최소 발생 호스트 갯수 및 최소 발생 횟수 중 하나 이상을 이용하여 상기 임계값을 설정할 수 있다.
- [0015] 상기 제1패킷 추출 모듈은 복수개의 호스트로부터 플로우를 입력받을 수 있다.
- [0016] 본 발명의 다른 양태에 따르면 제1패킷 추출 모듈이 입력받은 플로우들의 첫 번째 패킷을 추출하여 일련의 순서로 조합한 제1패킷 리스트를 생성하는 단계; 후보 패턴 추출 모듈이 상기 제1패킷 리스트에 포함된 패킷을 이용하여 조합 가능한 후보 패턴을 추출하는 단계 및 시그니처 생성 모듈이 상기 후보 패턴 중에서 기 설정된 임계값을 초과하는 패턴을 추출하여 시그니처를 생성하는 단계를 포함하는 행위 기반 시그니처 생성 방법을 제공한다.
- [0017] 상기 제1패킷 리스트를 생성하는 단계는 입력받은 플로우를 메모리에 로드하는 단계; 상기 메모리에 로드된 플로우가 리스트 생성 대상 플로우인지 판단하는 단계 및 리스트 생성 대상 플로우의 첫 번째 패킷을 추출하여 일련의 순서로 조합하는 단계를 포함하여 구성될 수 있다.
- [0018] 상기 리스트 생성 대상 플로우인지 판단하는 단계는 입력받은 플로우의 요청 패킷 포함 여부를 기준으로 할 수 있다.
- [0019] 상기 후보 패턴을 추출하는 단계는 상기 제1패킷 리스트 상에서 인터벌을 설정하는 단계 및 상기 인터벌내에 존재하는 패킷들을 조합하는 단계를 포함하여 구성될 수 있다.

발명의 효과

- [0020] 본 발명에 따른 행위기반 시그니처 생성 시스템 및 방법은 응용 실행시 발생하는 응용 트래픽을 분석하기 위하여 복수개의 플로우를 이용하여 시그니처를 생성하고, 또한 복수개의 호스트로부터 발생되는 플로우를 이용하여 시그니처를 생성하기 때문에 정확한 응용 트래픽 분석을 위한 시그니처를 생성할 수 있다.

도면의 간단한 설명

- [0021] 도1 은 본 발명의 일실시예에 따른 행위 기반 시그니처 생성 시스템의 블록도,
- 도2 는 본 발명의 일실시예에 따른 행위 기반 시그니처 생성 개념도,
- 도3 은 본 발명의 일실시예에 따른 제1패킷 추출 모듈의 동작 개념도,
- 도4 는 본 발명의 일실시예에 따른 후보 패턴 추출 모듈의 동작 개념도,
- 도5 는 본 발명의 일실시예에 따른 시그니처 생성 모듈의 동작 개념도 및
- 도6 은 본 발명의 일실시예에 따른 행위기반 시그니처 생성 방법의 순서도이다.

발명을 실시하기 위한 구체적인 내용

- [0022] 본 발명은 다양한 변경을 가할 수 있고 여러 가지 실시예를 가질 수 있는 바, 특정 실시예들을 도면에 예시하고 설명하고자 한다. 그러나, 이는 본 발명을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다.
- [0023] 제2, 제1 등과 같이 서수를 포함하는 용어는 다양한 구성요소들을 설명하는데 사용될 수 있지만, 상기 구성요소들은 상기 용어들에 의해 한정되지는 않는다. 상기 용어들은 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 사용된다. 예를 들어, 본 발명의 권리 범위를 벗어나지 않으면서 제2 구성요소는 제1 구성요소로 명명될 수 있고, 유사하게 제1 구성요소도 제2 구성요소로 명명될 수 있다. 및/또는 이라는 용어는 복수의 관련된 기재된 항목들의 조합 또는 복수의 관련된 기재된 항목들 중의 어느 항목을 포함한다.
- [0024] 어떤 구성요소가 다른 구성요소에 "연결되어" 있다거나 "접속되어" 있다고 언급된 때에는, 그 다른 구성요소에 직접적으로 연결되어 있거나 또는 접속되어 있을 수도 있지만, 중간에 다른 구성요소가 존재할 수도 있다고 이

해되어야 할 것이다. 반면에, 어떤 구성요소가 다른 구성요소에 "직접 연결되어" 있다거나 "직접 접속되어" 있다고 언급된 때에는, 중간에 다른 구성요소가 존재하지 않는 것으로 이해되어야 할 것이다.

- [0025] 본 출원에서 사용한 용어는 단지 특정한 실시예를 설명하기 위해 사용된 것으로, 본 발명을 한정하려는 의도가 아니다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 출원에서, "포함하다" 또는 "가지다" 등의 용어는 명세서상에 기재된 특징, 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.
- [0026] 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 모든 용어들은 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가지고 있다. 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 관련 기술의 문맥 상 가지는 의미와 일치하는 의미를 가지는 것으로 해석되어야 하며, 본 출원에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.
- [0027] 이하, 첨부된 도면을 참조하여 실시예를 상세히 설명하되, 도면 부호에 관계없이 동일하거나 대응하는 구성 요소는 동일한 참조 번호를 부여하고 이에 대한 중복되는 설명은 생략하기로 한다.
- [0028] 본 발명에서의 시그니처란 응용 프로그램 별로 사용하여 다른 응용들과 구분되고 전체 응용프로그램의 트래픽으로부터 해당 응용프로그램을 분류할 수 있는 고유한 패턴으로 정의될 수 있다.
- [0029] 본 발명에서 트래픽은 호스트와 호스트 사이의 모든 트래픽을 대상으로 하고, 단일 플로우를 동일한 소스 IP(source IP), 목적지 IP(destination IP), 소스 포트(source port), 목적지 포트(destination port) 및 전송 계층 프로토콜(protocol)을 갖는 패킷의 집합으로, 양방향의 플로우로 정의될 수 있다. 플로우에서 소스 IP(source IP), 목적지 IP(destination IP), 소스 포트(source port), 목적지 포트(destination port) 및 전송 계층 프로토콜(protocol)은 5-튜플(tuple)이라 한다.
- [0030] 특정 네트워크에서 트래픽을 분류하기 위해서는 네트워크에 연결된 호스트 사이의 트래픽을 수집해야 한다. 예를 들어, 어떤 네트워크가 라우터를 통해 인터넷과 연결되어 있다면 하나의 라우터에서 미러링을 통해 인터넷과 통신하는 모든 트래픽을 수집할 수 있다.
- [0031] 도1 은 본 발명의 일실시예에 따른 행위 기반 시그니처 생성 시스템의 블록도, 도2 는 본 발명의 일실시예에 따른 행위 기반 시그니처 생성 개념도이다..
- [0032] 도1 을 참조하면, 본 발명의 일실시예에 따른 행위 기반 시그니처 생성 시스템은 입력받은 플로우들의 첫 번째 패킷을 추출하여 일련의 순서로 조합한 제1패킷 리스트를 생성하는 제1 패킷 추출 모듈(10), 상기 제1패킷 리스트에 포함된 패킷을 이용하여 조합 가능한 후보 패턴을 추출하는 후보 패턴 추출 모듈(20) 및 상기 후보 패턴 중에서 기 설정된 임계값을 초과하는 패턴을 추출하여 시그니처를 생성하는 시그니처 생성 모듈(30)을 포함하여 구성될 수 있다.
- [0033] 먼저, 제1패킷 추출 모듈(10)은 입력받은 플로우들의 첫 번째 패킷을 추출하여 일련의 순서로 조합한 제1패킷 리스트를 생성할 수 있다.
- [0034] 제1패킷 추출 모듈(10)은 미러링 포트 등을 통하여 수집되는 플로우를 입력으로 할 수 있다. 제1패킷 추출 모듈(10)은 입력되는 플로우 중 페이로드(payload)가 존재하는 플로우의 첫 번째 요청 패킷을 추출하여 일련의 순서로 조합한 제1패킷 리스트를 생성할 수 있다. 제1패킷 추출 모듈(10)은 리스트 생성 대상 플로우에 한하여 플로우에 포함된 제1패킷을 추출할 수 있다. 제1패킷 추출 모듈(10)은 추출한 제1패킷을 일련의 순서로 조합하여 제1패킷 리스트를 생성할 수 있으며, 예를 들면, 패킷의 발생 시간을 기준으로 추출한 제1패킷을 정렬하여 제1패킷 리스트를 생성할 수 있다.
- [0035] 제1패킷 추출 모듈(10)은 플로우의 입력이 중단되는 시점까지 제1패킷 리스트를 생성하는 과정을 반복 수행하게 된다.
- [0036] 도3 은 본 발명의 일실시예에 따른 제1패킷 추출 모듈의 동작 개념도이다.
- [0037] 도3 을 참조하면, 입력받은 4개의 플로우가 모두 리스트 생성 대상 플로우인 경우에 제1패킷 추출 모듈(10)은 각 플로우의 첫 번째 패킷을 조합하여 P₁-P₂-P₃-P₄로 조합되는 제1패킷 리스트를 생성할 수 있다. 4개의 플로우

는 각각 순차적으로 제1패킷 추출 모듈(10)에 입력되었으며, 따라서 제1패킷 추출 모듈(10)은 입력시간을 기준으로 제1패킷 리스트를 생성하였음을 확인할 수 있다.

- [0038] 후보 패턴 추출 모듈(20)은 제1패킷 리스트에 포함된 패킷을 이용하여 조합 가능한 모든 후보 패턴을 추출할 수 있다. 후보 패턴 추출 모듈(20)은 제1패킷 리스트상에서 일정 인터벌을 설정하고, 상기 인터벌내에 포함되는 패킷을 대상으로 후보 패턴을 추출할 수 있다. 하나의 후보 패턴은 복수개의 엔트리로 구성되며, 각각의 엔트리는 행위 기반 시그니처 모델로 정의된다. 여기서 인터벌이란 제1패킷 리스트상에서 임의의 갯수의 엔트리가 포함되는 일정 구간을 의미할 수 있다.
- [0039] 후보 패턴 추출 모듈(20)은 후보 패턴을 추출함에 있어서, 기 설정된 행위 기반 시그니처 모델을 기반으로 제 1 패킷 리스트에 포함된 패킷에서 트래픽의 특징을 추출하여 엔트리를 구성한다. 각 패킷에서 추출된 엔트리의 조합이 후보 패턴으로 추출될 수 있다.
- [0040] 행위 기반 시그니처 모델은 타입, 인터벌 및 엔트리 구성요소 (목적지 IP, 목적지 port, 페이로드, 전송 계층 프로토콜)로 구성된다.
- [0041] 타입은 Seq 타입과 Set 타입으로 나뉘 수 있으며, Seq 타입은 응용 실행 시 발생하는 플로우들의 첫 질의 패킷이 행위 기반 시그니처의 엔트리와 동일한 순서로 적용됨을 의미하며, Set타입은 적용 순서가 임의로 변경될 수 있음을 의미한다.
- [0042] 인터벌은 첫 엔트리와 마지막 엔트리가 매칭되는 일정한 시간 간격(ms)을 의미한다. 즉, 트래픽 발생 시간을 기준으로 해당 패턴이 적용되는 기간을 의미한다.
- [0043] 행위 기반 시그니처 모델의 각각의 엔트리들은 목적지 IP, 목적지 port, 페이로드, 전송 계층 프로토콜로 구성된다.
- [0044] 목적지 IP는 해당 엔트리가 전송되는 목적지 IP주소를 의미하며 CIDR표기법을 이용하여 표시할 수 있다.
- [0045] 목적지 port는 해당 엔트리가 전송되는 목적지 port를 의미한다.
- [0046] 전송 계층 프로토콜은 해당 엔트리가 전송될 때 사용되는 전송 계층 프로토콜(TCP, UDP)를 의미한다.
- [0047] 페이로드는 해당 엔트리가 전송시키는 콘텐츠이다. 페이로드 전체를 엔트리 구성요소로 사용하지 않고, 페이로드의 첫 몇 byte만을 엔트리의 구성요소로 사용한다. HTTP를 사용하는 트래픽의 경우 트래픽 첫 부분에 위치하는 Mthod(GET, POST, PUT 등)를 구별하기 위해 페이로드의 첫 10byte를 사용하고 Non-HTTP인 경우, 페이로드의 첫 2byte만을 사용한다
- [0048] 도4 는 본 발명의 일실시예에 따른 후보 패턴 추출 모듈의 동작 개념도이다. 도4 를 참조하면, 후보 패턴 추출 모듈(20)은 P1 내지 P4가 포함되는 4개의 패킷에서 행위 기반 시그니처 모델을 기반으로 각각의 엔트리(E1~E4)를 추출하고 이를 조합한 후보 패턴을 추출하였다. 후보 패턴 추출 모듈(20)은 E1 내지 E4의 조합 가능한 모든 후보 패턴을 추출하였으며, 그 결과는 상위 7개를 제외하고는 생략하였다.
- [0049] 시그니처 생성 모듈(30)은 후보 패턴 중 기 설정된 임계값을 초과하는 패턴을 추출하여 시그니처를 생성할 수 있다. 시그니처 생성 모듈(30)은 최대 패턴 크기, 최대 인터벌, 최소 발생 호스트 갯수 및 최소 발생 횟수 중 하나 이상을 이용하여 임계값을 설정할 수 있다.
- [0050] 최소 발생 호스트 갯수는 제1패킷 추출 모듈에 입력되는 플로우들이 발생된 특정 호스트의 갯수가 일정 갯수를 초과하는지 여부를 판단하기 위한 기준이다. 본 발명의 일실시예에 따라 생성되는 행위 기반 시그니처는 최소 2개 호스트에서 동일한 패턴으로 발생하는 후보 패턴 중에서 추출되어야 하므로 최소 발생 호스트 갯수는 이를 판단하기 위한 기준이다.
- [0051] 최대 패턴 크기는 행위 기반 시그니처가 가질 수 있는 최대 엔트리의 개수를 의미한다. 각각의 행위 기반 시그니처는 복수개의 엔트리로 구성되어 있으므로, 이를 제한함이 필요하다. 즉, 최대 패턴 크기를 설정함으로써 행위 기반 시그니처를 생성하기 위하여 사용되는 플로우의 갯수를 한정할 수 있다.
- [0052] 임계값은 하나의 행위 기반 시그니처가 생성될 수 있는 최대 시간 간격을 의미한다. 행위 기반 시그니처는 특정 시간내에 발생하는 복수개의 트래픽으로부터 추출되는 것이므로 일정한 시간 간격인 임계값을 설정하고, 임계값내에 발생하는 트래픽을 대상으로 하기 위한 것이다. 본 발명의 일실시예의경우 시그니처 생성 모듈(30)은 임계값을 5000ms로 설정할 수 있다.

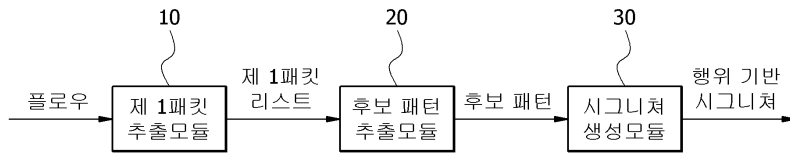
- [0053] 최소 발생 횟수는 후보 패턴 중 해당 패턴이 반복하여 발생한 횟수를 의미한다. 이는 복수개의 시그니처 중에서 자주 사용되는 시그니처를 판단하기 위한 것이다.
- [0054] 도5 는 본 발명의 일실시예에 따른 시그니처 생성 모듈의 동작 개념도이다. 도5 를 참조하면, 시그니처 생성 모듈(30)은 기 설정된 임계값을 후보 패턴에 적용하여 2개의 시그니처를 생성하였음을 확인할 수 있다.
- [0055] 도6 은 본 발명의 일실시예에 따른 행위기반 시그니처 생성 방법의 순서도이다.
- [0056] 도6 을 참조하면 본 발명의 일실시예에 따른 행위기반 시그니처 생성 방법은 입력 받은 플로우들의 첫 번째 패킷을 추출하여 일련의 순서로 조합한 제1패킷 리스트를 생성하는 단계(S601), 상기 제1패킷 리스트에 포함된 패킷을 이용하여 조합 가능한 후보 패턴을 추출하는 단계(S602) 및 상기 후보 패턴 중에서 기 설정된 임계값을 초과하는 패턴을 추출하여 시그니처를 생성하는 단계(S603)를 포함하여 구성될 수 있다.
- [0057] 먼저, 제1패킷 추출 모듈은 입력되는 플로우에 요청 패킷이 포함되었는지 여부를 기준으로 해당 플로우가 리스트 생성 대상 플로우인지 판단한다. 제1패킷 추출 모듈은 리스트 생성 대상 플로우에 한하여 플로우에 포함된 제1패킷을 추출한다. 제1패킷 추출 모듈은 추출한 제1패킷을 일련의 순서로 조합하여 제1패킷 리스트를 생성하며, 예를 들면, 입력시간을 기준으로 추출한 제1패킷을 정렬하여 제1패킷 리스트를 생성한다(S601).
- [0058] 후보 패턴 추출 모듈은 제1패킷 리스트 상에서 인터벌을 설정하고, 기 설정된 행위 기반 시스니처 모델을 기반으로 인터벌내에 존재하는 패킷에서 트래픽의 특징을 추출하여 엔트리를 구성한다. 후보 패턴 추출 모듈은 구성된 엔트리를 조합하여 후보 패턴을 추출한다(S602).
- [0059] 시그니처 생성 모듈은 후보 패턴 중 기 설정된 임계값을 초과하는 패턴을 추출하여 시그니처를 생성한다. 시그니처 생성 모듈은 최대 패턴 크기, 최대 인터벌, 최소 발생 호스트 갯수 및 최소 발생 횟수 중 하나 이상을 이용하여 임계값을 설정할 수 있다(S603).
- [0060] 본 실시예에서 사용되는 '~부'라는 용어는 소프트웨어 또는 FPGA(field-programmable gate array) 또는 ASIC과 같은 하드웨어 구성요소를 의미하며, '~부'는 어떤 역할들을 수행한다. 그렇지만 '~부'는 소프트웨어 또는 하드웨어에 한정되는 의미는 아니다. '~부'는 어드레싱할 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.
- [0061] 상기에서는 본 발명의 바람직한 실시예를 참조하여 설명하였지만, 해당 기술 분야의 숙련된 당업자는 하기의 특허 청구의 범위에 기재된 본 발명의 사상 및 영역으로부터 벗어나지 않는 범위 내에서 본 발명을 다양하게 수정 및 변경시킬 수 있음을 이해할 수 있을 것이다.

부호의 설명

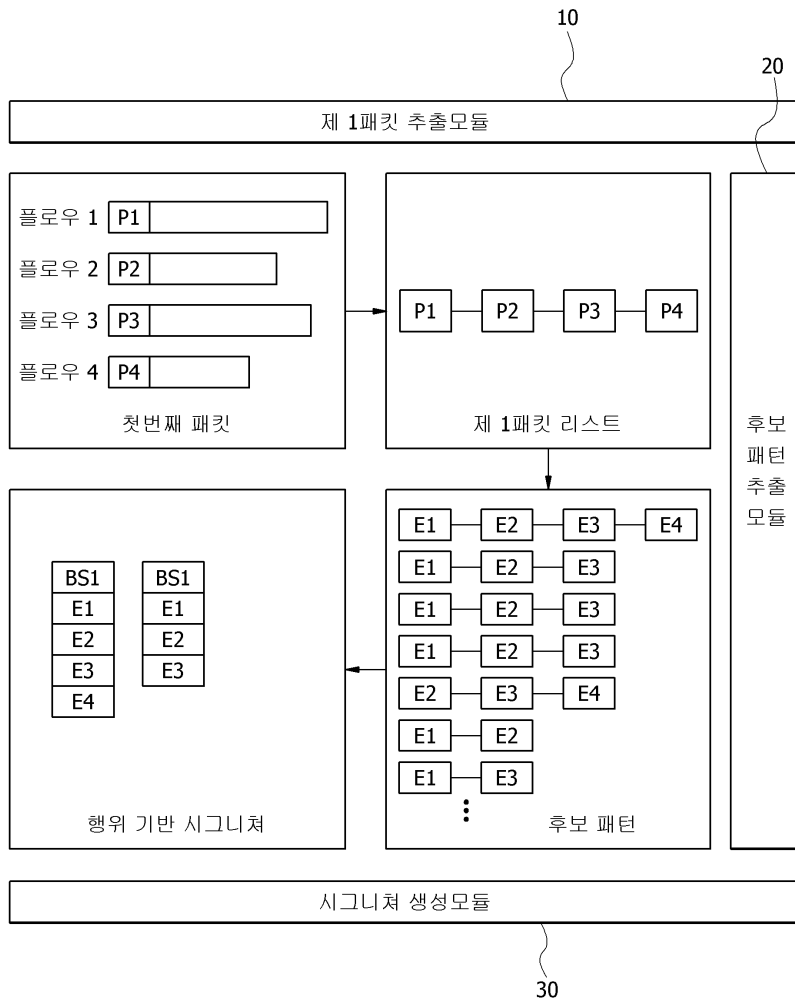
- [0062] 10: 제1패킷 추출 모듈
20: 후보 패턴 추출 모듈
30: 시그니처 생성 모듈

도면

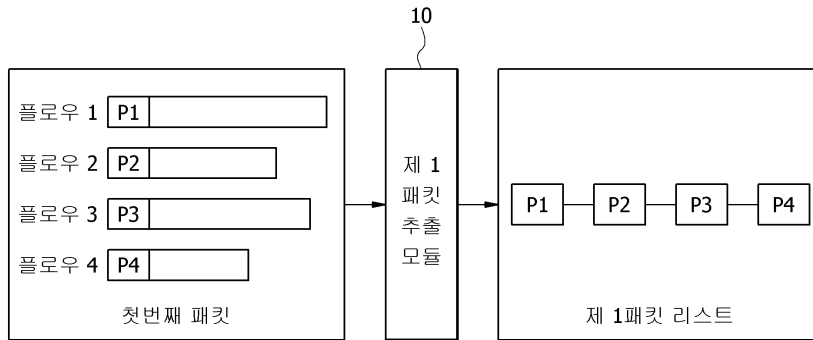
도면1



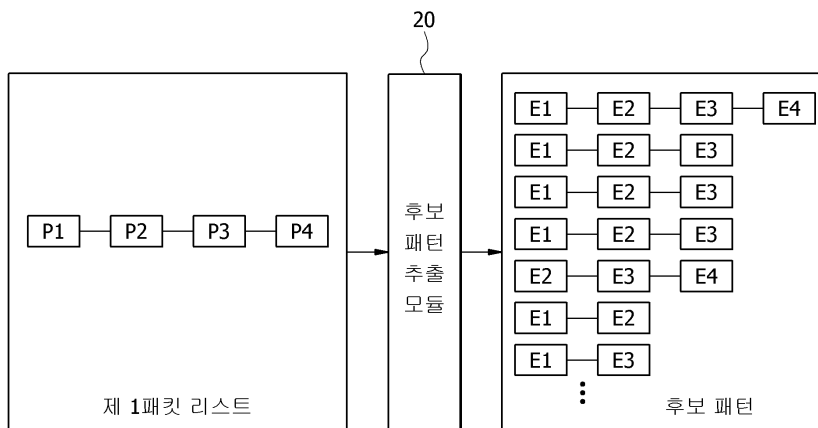
도면2



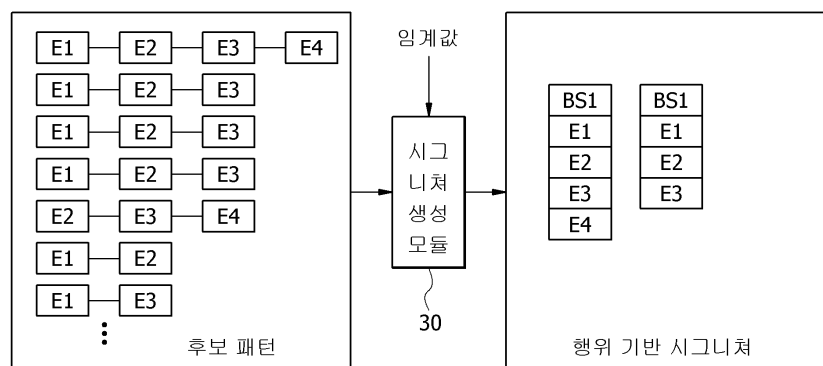
도면3



도면4



도면5



도면6

