



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2011년11월30일
(11) 등록번호 10-1087761
(24) 등록일자 2011년11월22일

(51) Int. Cl.

H04L 12/26 (2006.01)

(21) 출원번호 10-2010-0054200

(22) 출원일자 2010년06월09일

심사청구일자 2010년06월09일

(56) 선행기술조사문헌

‘skype 트래픽 분류에 관한 연구’, 2009년 제3
1회 한국 정보처리학회 춘계학술 발표대회 논문집
제16권 제1호, pp.1112-1115, 2009.04.*

*는 심사관에 의하여 인용된 문헌

(73) 특허권자

고려대학교 산학협력단

서울 성북구 안암동5가 1

(72) 발명자

김명섭

충청남도 연기군 조치원읍 신흥리 푸르지오1차 아
파트 107동 401호

오영석

서울특별시 노원구 상계3동 불암대림아파트 202동
203호

(뒷면에 계속)

(74) 대리인

특허법인엠에이피에스

전체 청구항 수 : 총 6 항

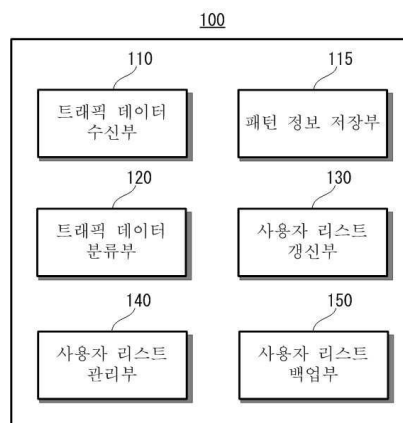
심사관 : 김대성

(54) 스카이프 트래픽 데이터의 분류가 가능한 트래픽 분류 장치 및 방법

(57) 요약

트래픽 분류 장치는 플로우 데이터 단위로 트래픽 데이터를 수신하는 트래픽 데이터 수신부, 수신된 트래픽 데이터의 노드에 관련된 IP(Internet Protocol) 및 포트번호 중 적어도 하나가 사용자 리스트에 포함되어 있는지 여부를 판단하는 사용자 리스트 비교부, 트래픽 데이터의 로그인 패턴이 미리 설정된 특정 어플리케이션의 로그인 패턴과 일치하는지 여부를 판단하는 로그인 패턴 비교부 및 사용자 리스트 비교부 및 로그인 패턴 비교부 중 적어도 하나의 판단에 따라 트래픽 데이터가 특정 어플리케이션의 트래픽 데이터로 분류되면, 트래픽 데이터의 노드에 관련된 IP 및 포트번호 중 적어도 하나를 사용자 리스트에 추가시키는 사용자 리스트 갱신부를 포함하되, 로그인 패턴 비교부는, 트래픽 데이터에 포함된 노드의 IP와 포트번호, 트래픽 데이터의 패킷 사이즈, 페이로드 시그니처(signature), 및 프로토콜 정보 중 적어도 하나를 이용하여 트래픽 데이터의 로그인 패턴이 특정 어플리케이션의 로그인 패턴과 일치하는지 여부를 판단한다.

대표도 - 도3a



(72) 발명자

이현신

충청남도 연기군 조치원읍 서창리 고려대학교세종
캠퍼스 과학기술대학 네트워크 관리연구실 210B

최미정

강원도 춘천시 효자2동 192-1번지 강원대학교 컴퓨
터학부

이상우

서울특별시 서초구 방배4동 878-24호

이 발명을 지원한 국가연구개발사업

과제고유번호 20090090455

부처명 한국과학재단

연구관리전문기관

연구사업명 일반연구자지원_기본연구

연구과제명 실시간 응용계층 트래픽 분류 시스템 설계 및 구현에 관한 연구

기여율

주관기관 고려대학교 산학협력단

연구기간 2009년 09월 01일 ~ 2010년 8월 31일

특허청구의 범위

청구항 1

트래픽 분류 장치에 있어서,

플로우 데이터 단위로 트래픽 데이터를 수신하는 트래픽 데이터 수신부,

상기 수신된 트래픽 데이터의 노드에 관련된 IP(Internet Protocol) 및 포트번호 중 적어도 하나가 사용자 리스트에 포함되어 있는지 여부를 판단하여 특정 어플리케이션의 트래픽 데이터인지를 판단하는 사용자 리스트 비교부,

상기 트래픽 데이터의 로그인 패턴이 미리 설정된 특정 어플리케이션의 로그인 패턴과 일치하는지 여부를 판단하여 상기 특정 어플리케이션의 트래픽 데이터인지를 판단하는 로그인 패턴 비교부,

상기 사용자 리스트 비교부 및 상기 로그인 패턴 비교부 중 적어도 하나의 판단에 따라 상기 트래픽 데이터가 상기 특정 어플리케이션의 트래픽 데이터로 분류되면, 상기 트래픽 데이터의 노드에 관련된 IP 및 포트번호 중 적어도 하나를 상기 사용자 리스트에 추가시키는 사용자 리스트 갱신부 및

상기 사용자 리스트에 추가된 노드에 대응되는 트래픽 데이터가 임계 시간 동안 발생하지 않은 경우, 상기 노드에 관련된 정보를 상기 사용자 리스트에서 삭제 처리하는 사용자 리스트 관리부를 포함하되,

상기 로그인 패턴 비교부는,

상기 트래픽 데이터에 포함된 노드의 IP와 포트번호, 상기 트래픽 데이터의 패킷 사이즈, 페이로드 시그니처(signature), 및 프로토콜 정보 중 적어도 하나를 이용하여 상기 트래픽 데이터의 로그인 패턴이 상기 특정 어플리케이션의 로그인 패턴과 일치하는지 여부를 판단하고,

상기 트래픽 데이터의 로그인 패턴이 미리 설정된 특정 어플리케이션의 로그인 패턴과 일치하는지 여부를 수신한 패킷을 검사하여 판단하되, 상기 검사를 수행한 상기 패킷의 개수가 미리 설정된 임계 개수를 초과한 경우에는 상기 트래픽 데이터가 상기 특정 어플리케이션의 트래픽 데이터가 아닌 것으로 판단하는 것을 특징으로 하는 트래픽 분류 장치.

청구항 2

제 1 항에 있어서,

상기 특정 어플리케이션은,

스카이프(Skype) 어플리케이션인 것을 특징으로 하는 트래픽 분류 장치.

청구항 3

삭제

청구항 4

제 1 항에 있어서,

상기 사용자 리스트 비교부를 통해 상기 사용자 리스트에 상기 트래픽 데이터의 노드에 관련된 IP 및 포트번호 중 적어도 하나가 포함되어 있지 않다고 판단된 경우, 상기 로그인 패턴 비교부를 통해 상기 트래픽 데이터의 로그인 패턴이 상기 미리 설정된 특정 어플리케이션의 로그인 패턴과 일치하는지 여부가 판단되도록 제어하는 분류 제어부를 더 포함하는 트래픽 분류 장치.

청구항 5

삭제

청구항 6

제 1 항에 있어서,

주기적으로 상기 사용자 리스트에 추가된 노드에 관련된 정보를 백업하는 사용자 리스트 백업부를 더 포함하는 트래픽 분류 장치.

청구항 7

트래픽 분류 장치를 이용한 트래픽 분류 방법에 있어서,

(a) 플로우 데이터 단위로 트래픽 데이터를 수신하는 수신 단계,

(b) 상기 수신된 트래픽 데이터의 노드에 관련된 IP(Internet Protocol) 및 포트번호 중 적어도 하나가 사용자 리스트에 포함되어 있는지 여부를 판단하는 단계,

(c) 상기 트래픽 데이터의 노드에 관련된 IP 및 포트번호 중 적어도 하나가 상기 사용자 리스트에 포함되어 있지 않은 경우, 상기 트래픽 데이터의 로그인 패턴과 미리 설정된 특정 어플리케이션의 로그인 패턴과의 일치 여부를 판단하는 단계,

(d) 상기 (b) 단계 또는 상기 (c) 단계의 판단에 따라, 상기 트래픽 데이터가 상기 특정 어플리케이션의 트래픽 데이터로 분류된 경우, 상기 트래픽 데이터의 노드에 관련된 IP 및 포트번호 중 적어도 하나를 상기 사용자 리스트에 추가시키는 단계 및

(e) 상기 사용자 리스트에 추가된 노드에 대응되는 트래픽 데이터가 임계 시간 동안 발생하지 않은 경우, 상기 노드에 관련된 정보를 상기 사용자 리스트에서 삭제 처리하는 단계를 포함하되,

상기 (c) 단계는,

상기 트래픽 데이터에 포함된 노드의 IP와 포트번호, 상기 트래픽 데이터의 패킷 사이즈, 페이로드 시그니처(signature), 및 프로토콜 정보 중 적어도 하나를 이용하여 상기 트래픽 데이터의 로그인 패턴이 미리 설정된 스카이프 어플리케이션의 로그인 패턴과 일치하는지 여부를 판단하는 단계 및

상기 트래픽 데이터의 로그인 패턴이 미리 설정된 스카이프 어플리케이션의 로그인 패턴과 일치하는지 여부를 수신한 패킷을 검사하여 판단하되, 상기 검사를 수행한 상기 패킷의 개수가 미리 설정된 임계 개수를 초과한 경우에는 상기 트래픽 데이터가 상기 스카이프 어플리케이션의 트래픽 데이터가 아닌 것으로 판단하는 단계를 포함하는 트래픽 분류 방법.

청구항 8

제 7 항에 있어서,

상기 특정 어플리케이션은,

스카이프(Skype) 어플리케이션인 것을 특징으로 하는 트래픽 분류 방법.

청구항 9

삭제

청구항 10

삭제

명세서

기술 분야

[0001] 본 발명은 트래픽 분류 장치 및 방법에 관한 것으로, 보다 상세하게는 응용 프로그램 별로 트래픽을 분류하는 트래픽 분류 장치 및 방법에 관한 것이다.

배경 기술

[0002] 최근 인터넷 사용자의 증가와 고속 네트워크 망의 보급으로 네트워크 트래픽이 급증하였다. 그리고 단순히 WWW, FTP, SMTP, DNS와 같은 전통적인 인터넷 서비스뿐만 아니라 멀티미디어, P2P(Peer-to-Peer), 게임, 스마트폰 응용 프로그램 등의 다양한 서비스의 증가로 인한 트래픽도 증가하고 있다. 따라서, 이에 따른 효과적인 네트워크 관리를 위한 트래픽 모니터링 및 분석의 중요성이 커지고 있다.

[0003] 이때, 효과적인 네트워크 관리를 위해 선행되어야 할 것은 해당 트래픽이 어떤 응용 프로그램 또는 어떤 서비스에서 발생되었는가를 판별하는 것이다. 그러나, 응용 프로그램 별 트래픽 데이터 분류를 위해 시그니처(signature) 기반, 패킷 사이즈와 포트 기반, 기계 학습(machine learning) 방법들이 제안되고 있지만, p2p 방식의 메신저 프로그램에 대한 적용 결과는 응용 프로그램에 대한 분석을 및 정확성이 떨어져 이에 대한 연구가 더 필요한 실정이다.

[0004] 예를 들어, 스카이프(skype) 어플리케이션(이하, “스카이프” 라고도 함)은 p2p 방식의 메신저로서 사용자간 채팅, 음성통화, 화상통화, 전화 교환망을 통한 일반 전화, 파일전송 등의 기능을 제공한다.

[0005] 그러나, 스카이프의 트래픽 데이터들은 기본적으로 암호화가 되어있고, 인스톨(Install)시에 동적 포트를 할당하며, 일반적인 프로토콜을 사용하지 않아 네트워크 관리자의 입장에서 스카이프 응용프로그램의 트래픽 데이터 분류 기준을 정하기 어려운 문제점이 있다.

[0006] 따라서, 스카이프 등과 같은 어플리케이션의 트래픽 특성을 분석해 해당 트래픽을 보다 정확하게 탐지할 필요성이 제기된다.

발명의 내용

해결하려는 과제

[0007] 본 발명의 일 실시예는 특정 어플리케이션의 로그인 패턴 특성 및 사용자 리스트에 포함된 정보 중 적어도 하나를 이용하여, 트래픽 데이터가 특정 어플리케이션의 트래픽 데이터인지 여부를 판단하여 트래픽 데이터를 특정 어플리케이션의 트래픽 데이터로 분류하는 트래픽 분류 장치 및 방법을 제공하는 데에 그 목적이 있다.

[0008] 또한, 본 발명의 일 실시예는 트래픽 데이터의 노드에 관련된 IP(Internet Protocol) 및 포트번호 중 적어도 하나가 사용자 리스트에 포함되어 있는지 여부를 먼저 판단하고, 트래픽 데이터의 노드에 관련된 IP 및 포트번호 중 적어도 하나가 사용자 리스트에 포함되어 있지 않은 경우, 트래픽 데이터의 로그인 패턴과 미리 설정된 특정 어플리케이션의 로그인 패턴과의 일치 여부를 판단하도록 하여, 어느 하나의 판단 결과에 따라 해당 트래픽 데이터를 특정 어플리케이션의 트래픽 데이터로 분류함으로써, 보다 신속한 트래픽 데이터에 대한 분류가 이루어지도록 하는 트래픽 분류 장치 및 방법을 제공하는 데에 그 목적이 있다.

[0009] 또한, 본 발명의 일 실시예는 특정 어플리케이션의 트래픽 데이터를 효과적으로 분류하여 네트워크 관리의 효율성을 높이는 트래픽 분류 장치 및 방법을 제공하는 데에 그 목적이 있다.

[0010] 또한, 본 발명의 일 실시예는 특정 어플리케이션의 로그인 패턴 특성을 갖는 노드의 IP 및 포트 번호를 포함하는 사용자 리스트를 이용하여 보다 효과적으로 특정 어플리케이션의 트래픽 데이터를 분류하는 트래픽 분류 장치 및 방법을 제공하는 데에 그 목적이 있다.

과제의 해결 수단

[0011] 상술한 기술적 과제를 달성하기 위한 기술적 수단으로서, 본 발명의 제 1 측면에 따른 트래픽 분류 장치는 플로우 데이터 단위로 트래픽 데이터를 수신하는 트래픽 데이터 수신부, 수신된 트래픽 데이터의 노드에 관련된 IP(Internet Protocol) 및 포트번호 중 적어도 하나가 사용자 리스트에 포함되어 있는지 여부를 판단하여 특정 어플리케이션의 트래픽 데이터인지를 판단하는 사용자 리스트 비교부, 트래픽 데이터의 로그인 패턴이 미리 설

정된 특정 어플리케이션의 로그인 패턴과 일치하는지 여부를 판단하여 특정 어플리케이션의 트래픽 데이터인지를 판단하는 로그인 패턴 비교부 및 사용자 리스트 비교부 및 로그인 패턴 비교부 중 적어도 하나의 판단에 따라 트래픽 데이터가 특정 어플리케이션의 트래픽 데이터로 분류되면, 트래픽 데이터의 노드에 관련된 IP 및 포트번호 중 적어도 하나를 사용자 리스트에 추가시키는 사용자 리스트 갱신부를 포함하되, 로그인 패턴 비교부는, 트래픽 데이터에 포함된 노드의 IP와 포트번호, 트래픽 데이터의 패킷 사이즈, 페이로드 시그니처(signature), 및 프로토콜 정보 중 적어도 하나를 이용하여 트래픽 데이터의 로그인 패턴이 특정 어플리케이션의 로그인 패턴과 일치하는지 여부를 판단한다.

[0012] 여기서, 특정 어플리케이션은 스카이프(Skype) 어플리케이션일 수 있다.

[0013] 또한, 상기 로그인 패턴 비교부는 트래픽 데이터의 로그인 패턴이 스카이프 어플리케이션의 로그인 패턴과 일치하는지 여부를 수신한 패킷을 검사하여 판단하되, 검사를 수행한 패킷의 개수가 미리 설정된 임계 개수를 초과한 경우에는 트래픽 데이터가 스카이프 어플리케이션의 트래픽 데이터가 아닌 것으로 판단할 수 있다.

[0014] 또한, 상기 사용자 리스트 비교부를 통해 상기 사용자 리스트에 상기 트래픽 데이터의 노드에 관련된 IP 및 포트번호 중 적어도 하나가 포함되어 있지 않다고 판단된 경우, 상기 로그인 패턴 비교부를 통해 상기 트래픽 데이터의 로그인 패턴이 상기 미리 설정된 특정 어플리케이션의 로그인 패턴과 일치하는지 여부가 판단되도록 제어하는 분류 제어부를 더 포함할 수 있다.

[0015] 또한, 상기 사용자 리스트에 추가된 노드에 대응되는 트래픽 데이터가 임계 시간 동안 발생하지 않은 경우, 상기 노드에 관련된 정보를 상기 사용자 리스트에서 삭제 처리하는 사용자 리스트 관리부를 더 포함할 수 있다.

[0016] 그리고, 주기적으로 상기 사용자 리스트에 추가된 노드에 관련된 정보를 백업하는 사용자 리스트 백업부를 더 포함할 수 있다.

[0017] 본 발명의 제 2 측면에 따른 트래픽 분류 방법은 (a) 플로우 데이터 단위로 트래픽 데이터를 수신하는 단계, (b) 수신된 트래픽 데이터의 노드에 관련된 IP(Internet Protocol) 및 포트번호 중 적어도 하나가 사용자 리스트에 포함되어 있는지 여부를 판단하는 단계, (c) 트래픽 데이터의 노드에 관련된 IP 및 포트번호 중 적어도 하나가 사용자 리스트에 포함되어 있지 않은 경우, 트래픽 데이터의 로그인 패턴과 미리 설정된 스카이프 어플리케이션의 로그인 패턴과의 일치 여부를 판단하는 단계, (d) 상기 (b) 단계 또는 상기 (c) 단계의 판단에 따라, 트래픽 데이터가 스카이프 어플리케이션의 트래픽 데이터로 분류된 경우, 트래픽 데이터의 노드에 관련된 IP 및 포트번호 중 적어도 하나를 사용자 리스트에 추가시키는 단계를 포함한다.

[0018] 여기서, 특정 어플리케이션은 스카이프(Skype) 어플리케이션일 수 있다.

[0019] 또한, (c) 단계는 상기 트래픽 데이터에 포함된 노드의 IP와 포트번호, 상기 트래픽 데이터의 패킷 사이즈, 페이로드 시그니처(signature), 및 프로토콜 정보 중 적어도 하나를 이용하여 상기 트래픽 데이터의 로그인 패턴이 상기 미리 설정된 스카이프 어플리케이션의 로그인 패턴과 일치하는지 여부를 판단할 수 있다.

[0020] 그리고, (c) 단계는 트래픽 데이터의 로그인 패턴이 스카이프 어플리케이션의 로그인 패턴과 일치하는지 여부를 수신한 패킷을 검사하여 판단하되, 검사를 수행한 패킷의 개수가 미리 설정된 임계 개수를 초과한 경우에는 트래픽 데이터가 스카이프 어플리케이션의 트래픽 데이터가 아닌 것으로 판단할 수 있다.

발명의 효과

[0021] 전술한 본 발명의 과제 해결 수단 중 어느 하나에 의하면, 특정 어플리케이션의 로그인 패턴 특성 및 사용자 리스트에 포함된 정보 중 적어도 하나를 이용하여, 트래픽 데이터가 특정 어플리케이션의 트래픽 데이터인지 여부를 판단하여 트래픽 데이터를 특정 어플리케이션의 트래픽 데이터로 분류할 수 있다.

[0022] 또한, 전술한 본 발명의 과제 해결 수단 중 어느 하나에 의하면, 트래픽 데이터의 노드에 관련된 IP(Internet Protocol) 및 포트번호 중 적어도 하나가 사용자 리스트에 포함되어 있는지 여부를 먼저 판단하고, 트래픽 데이터의 노드에 관련된 IP 및 포트번호 중 적어도 하나가 사용자 리스트에 포함되어 있지 않은 경우, 트래픽 데이터의 로그인 패턴과 미리 설정된 특정 어플리케이션의 로그인 패턴과의 일치 여부를 판단하도록 하여, 어느 하나의 판단 결과에 따라 해당 트래픽 데이터를 특정 어플리케이션의 트래픽 데이터로 분류함으로써, 보다 신속한 트래픽 데이터에 대한 분류가 이루어질 수 있다.

[0023] 또한, 전술한 본 발명의 과제 해결 수단 중 어느 하나에 의하면, 스카이프 어플리케이션의 트래픽 데이터를 효과적으로 분류하여 네트워크 관리의 효율성을 높일 수 있다.

[0024] 또한, 전술한 본 발명의 과제 해결 수단 중 어느 하나에 의하면, 스카이프 어플리케이션의 로그인 패턴 특성을 갖는 노드의 IP 및 포트 번호를 포함하는 사용자 리스트를 이용하여 보다 효과적으로 스카이프 어플리케이션의 트래픽 데이터를 분류할 수 있다.

도면의 간단한 설명

[0025] 도 1은 본 발명의 일 실시예에 따른 트래픽 수집 및 검증 시스템을 도시한다.

도 2는 스카이프 어플리케이션의 로그인 과정에 대한 개념도이다.

도 3a는 본 발명의 일 실시예에 따른 트래픽 분류 장치의 블록도이다.

도 3b는 상기 트래픽 분류 장치에 포함된 트래픽 데이터 분류부의 블록도이다.

도 4는 본 발명의 일 실시예에 따른 IP 및 포트번호를 이용하여 트래픽 데이터를 분류하는 방법을 도시한다.

도 5는 본 발명의 일 실시예에 따른 로그인 패턴을 이용하여 트래픽 데이터를 분류하는 방법을 도시한다.

도 6은 본 발명의 실시예에 따른 스카이프 트래픽 분류 방법의 순서도이다.

발명을 실시하기 위한 구체적인 내용

[0026] 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.

[0027] 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.

[0028] 이하에서는 특정 어플리케이션으로서 스카이프 어플리케이션(이하, "스카이프" 라고도 함)을 예로 들어 설명하나 반드시 이에 한정되지는 않고, 스카이프 외에 특정 어플리케이션에도 본 발명의 실시예들이 적용될 수 있음은 물론이다.

[0029] 도 1은 본 발명의 일 실시예에 따른 트래픽 수집 및 검증 시스템을 도시한다.

[0030] 외부 인터넷 망과 연결되는 라우터에 설치된 실시간 트래픽 모니터링 시스템인 KU-MON은, 트래픽 데이터의 수집을 위해 실시간으로 패킷을 수집한다.

[0031] 여기서, KU-MON(박상훈, 박진완, 김명섭, "플로우 기반 실시간 트래픽 수집 및 분석 시스템", 정보처리학회 추계학술대회, 목포대학교, 전주, Nov. 9-10, 1047, pp. 1061.)은 실시간 트래픽 모니터링 시스템으로서, 외부 인터넷 망과 연결되어 있는 라우터와 두 대의 코어스위치 사이에 있는 2개의 링크로부터 트래픽 데이터를 수집할 수 있다.

[0032] 그리고, 수집된 패킷은 플로우 데이터 단위로 그룹화되는데, 여기서 플로우 데이터(이하, "플로우" 라고도 함)는 패킷 헤더의 5-튜플(tuple)(소스 IP, 목적지 IP, 소스 포트, 목적지 포트, 프로토콜) 정보를 공유하는 양방향 패킷들의 집합으로 정의된다.

[0033] 도 1에 도시된 바와 같이, 상술된 KU-MON 기반의 트래픽 수집 및 검증 시스템에 있어서, 라우터에 연결된 TCS(Traffic Capture System)(104)는 미러링을 통해 패킷들을 수집하여 실시간으로 플로우 데이터를 생성한다.

[0034] 그리고, 생성된 플로우 데이터는 TAS(Traffic Analysis Server)(102)와 TMS(Traffic Measurement Server)(200)로 각각 전송된다. 여기서, TCS(104)와 TAS(102)는 "Real-time SIS" (Skype Identification System)(101)으로 정의될 수 있으며, "Real-time SIS" (101)는 후술될 본 발명의 실시 예에 따른 도 3의 구성요소들을 포함할 수 있다.

[0035] 그리고, TMA(300) 검증 시스템은 네트워크 내의 중단 노드에 설치된 TMA(300)를 이용하여 TMA 기반 데이터인 스카이프 트래픽 정답지(Ground Truth Data)를 생성하고 이를 바탕으로 분류결과를 검증할 수 있다.

- [0036] 즉, 중단 노드에 설치된 TMA(300)는 해당 노드의 현재 활성화된 소켓 정보를 토대로 TMA 정보(소스 IP, 목적지 IP, 소스 포트, 목적지 포트, 프로토콜)를 추출하여 TMS(200)로 전송한다.
- [0037] 그리고 TMS(200)는 TMA 정보를 이용하여 TCS(104)로부터 전달받은 플로우 데이터와 비교하여 스카이프에 대한 정답지 플로우를 생성한다. 그리고 TVS(Traffic Verification System)(400)는 TAS(102)에서의 트래픽 탐지 결과와 정답지 플로우의 비교를 통해 트래픽 탐지 결과의 정확도를 측정하게 된다.
- [0038] 이하 도 2에서는 스카이프 어플리케이션의 로그인 과정에 대해서 설명하기로 한다.
- [0039] 도 2는 스카이프 어플리케이션의 로그인 과정에 대한 개념도이다.
- [0040] 도 2에 도시된 바와 같이, 스카이프의 로그인 과정에서 스카이프의 로그인 패턴 특성을 알 수 있다.
- [0041] 먼저 스카이프의 로그인 과정에 있어서, SC(Skype Client)가 자신에 저장되어 있던 SN(Super Node)의 리스트를 기반으로 하여 SN에게 UDP 패킷을 전송한다(S11). 다음으로, SC가 SN으로부터 응답 패킷을 수신한다(S21). 다음으로, SC가 응답 패킷을 전달받은 SN과 TCP(Transmission Control Protocol) 연결을 수행한다(S31).
- [0042] 여기서, 복수의 SN으로부터 응답이 온 경우 응답 패킷을 보낸 1번째 또는 2 번째 SN이 선택되어 TCP 연결이 이루어진다. 그리고 TCP 연결을 수행한 SN으로부터 뉴(new) SN 리스트를 수신한다고 추정될 수 있다. 또한, SC는 리스트를 수신하고 UDP(User Datagram Protocol) 패킷을 뉴 SN에게 전송한다. 이때, 응답을 전송한 뉴 SN에 한하여 최종 확정된 새로운 리스트를 저장한다고 추정될 수 있다. 그리고, 예컨대 dstport(Destination Port, 목적지 포트번호) 33033인 노드와의 TCP 연결이 이루어질 수 있다.
- [0043] 다음으로, SC가 예컨대 dstport 12350인 노드로 UDP 패킷을 전송한다(S41). 이때, SC의 스카이프 포트번호가 인지될 수 있고, 후술될 스카이프 트래픽 분류 알고리즘에 의해 SC의 IP(Internet Protocol)와 포트(Port)가 기록될 수 있다.
- [0044] 그리고, SC와 버디(Buddy)(스카이프 친구등록) 관계인 SC들에게 UDP 패킷이 전송되며, 프로필 변경, 사진 변경 등 갱신 정보가 있을 때 버디 관계의 SC와 TCP 연결이 이루어져 해당 데이터에 대한 전송이 수행될 수 있다.
- [0045] 다음으로, SC가 DNS 서버로 쿼리를 전송하여 스카이프 웹 서버의 IP를 수신한다(S51). 여기서, dstport 12350인 노드와 TCP 연결을 맺는 경우와 맺지 않는 경우의 두 가지 경우가 발생할 수 있다.
- [0046] 이때, 단계(S41)에서 SC가 이전에 로그인하였던 노드에서 다시 로그인 하였을 경우, 단계(S51)에서 dstport 12350인 노드와 TCP 연결을 맺지 않는다. 그리고, 단계(S41)에서 SC가 이전에 다른 노드에서 로그인을 하였고 이번 로그인이 그전의 노드와 다를 경우, dstport 12350인 노드와 TCP 연결을 맺게 된다. 따라서 단계(S41)의 버디 SC에게 UDP 패킷을 전송하는 과정은 한 과정 뒤로 미루어지게 된다.
- [0047] 다음으로, SC가 DNS서버로부터 수신한 IP를 갖는 웹 서버와 TCP 연결을 맺고 버전을 체크한다(S61). 그리고, SC가 dstport 33033인 노드와 접속을 해제하며, 단계(S51)에서 dstport 12350인 노드와 TCP 연결을 이루었을 경우 해당 접속을 해제한다(S71). 이후, SC가 버디 SC와 통신, 그 외 SC와 통신, 및 로그아웃할 경우 단계(S31) 과정에서 SN과 맺었던 TCP 연결을 해제하게 된다.
- [0048] 한편, 본 발명의 실시 예에서는 SC 및 SN의 스카이프 포트번호 추출을 위해, 단계(S41)에서 목적지 포트번호 12350인 노드로 트래픽 데이터를 전송할 때 발생하는 UDP 패킷의 사이즈와 페이로드 내용을 분석할 수 있다. 또한, 로그인 단계가 끝난 단계(S71) 이후에서도 구축된 스카이프 노드 리스트와의 상관관계를 통하여 다른 노드의 IP 및 포트 또한 추출이 가능하다.
- [0049] 도 3a는 본 발명의 일 실시예에 따른 트래픽 분류 장치의 블록도이다. 그리고, 도 3b는 트래픽 분류 장치에 포함된 트래픽 데이터 분류부의 블록도이다.
- [0050] 도 3a에 도시된 바와 같이, 트래픽 분류 장치(100)는 트래픽 데이터 수신부(110), 패턴 정보 저장부(115), 트래픽 데이터 분류부(120), 사용자 리스트 갱신부(130), 사용자 리스트 관리부(140), 사용자 리스트 백업부(150)를 포함하며, 상술된 “Real-time SIS” (101)에 구성될 수 있다.
- [0051] 이하, 각 구성요소에 대해서 보다 구체적으로 설명하기로 한다.
- [0052] 트래픽 데이터 수신부(110)는 플로우 데이터 단위로 트래픽 데이터를 수신한다. 여기서, 플로우 데이터는 일반적으로 사용되는 패킷 헤더의 5-튜플(tuple) 정보(소스 IP, 목적지 IP, 소스 포트, 목적지 포트, 프로토콜)가 동일한 단 방향(Uni-flow) 패킷들의 집합으로 정의될 수 있다.

- [0053] 패턴 정보 저장부(115)는 스카이프의 로그인 패턴 정보를 저장한다. 이러한 미리 설정된 로그인 패턴 정보는 해당 트래픽 데이터가 스카이프의 트래픽 데이터인지 여부를 판단할 때 사용될 수 있다.
- [0054] 트래픽 데이터 분류부(120)는 미리 설정된 스카이프 어플리케이션의 로그인 패턴 및 사용자 리스트에 포함된 정보 중 적어도 하나를 이용하여, 수신된 트래픽 데이터가 스카이프의 트래픽 데이터인지 여부를 판단한다.
- [0055] 이를 위해 도 3b에 도시된 바와 같이, 트래픽 데이터 분류부(120)는 사용자 리스트 비교부(122), 로그인 패턴 비교부(124) 및 분류 제어부(126)를 포함한다.
- [0056] 사용자 리스트 비교부(122)는 수신된 트래픽 데이터의 노드에 관련된 정보가 사용자 리스트에 포함되어 있는지 여부를 판단하여, 수신된 트래픽 데이터를 스카이프의 트래픽 데이터로 분류할 수 있다. 예컨대, 수신된 트래픽 데이터의 노드에 관련된 정보로서 IP 및 포트번호 중 적어도 하나가 이용될 수 있으며, 해당 IP 및 포트번호가 사용자 리스트에 포함되어 있는지 여부에 따라 수신된 트래픽 데이터를 스카이프의 트래픽 데이터로 분류할 수 있다.
- [0057] 로그인 패턴 비교부(124)는 사용자 리스트를 사용하지 않고도 수신된 트래픽 데이터의 로그인 패턴이 미리 설정된 스카이프의 로그인 패턴과 일치하는지 여부를 판단하여, 수신된 트래픽 데이터를 스카이프의 트래픽 데이터로 분류할 수 있다.
- [0058] 보다 구체적으로, 로그인 패턴 비교부(124)는 트래픽 데이터에 포함된 노드의 IP와 포트번호, 트래픽 데이터의 패킷 사이즈, 페이로드 시그니처(signature), 및 프로토콜 정보 중 적어도 하나를 이용하여 해당 트래픽 데이터의 로그인 패턴이 미리 설정된 스카이프의 로그인 패턴과 일치하는지 여부를 판단할 수 있다.
- [0059] 분류 제어부(126)는 사용자 리스트 비교부(122)를 통해 사용자 리스트에 수신된 트래픽 데이터의 노드에 관련된 정보가 포함되어 있지 않다고 판단된 경우, 로그인 패턴 비교부(124)를 통해 수신된 트래픽 데이터의 로그인 패턴이 미리 설정된 스카이프의 로그인 패턴과 일치하는지 여부가 판단되도록 제어한다. 여기서, 스카이프의 로그인 패턴 특성에 대해서는 상술된 도 2를 참조하기 바란다.
- [0060] 즉, 분류 제어부(126)는 사용자 리스트를 이용하여 수신된 트래픽 데이터에 대한 1차 분류가 이루어지도록 하고, 1차 분류를 통해 수신된 트래픽 데이터가 스카이프의 트래픽 데이터가 아닌 경우, 미리 설정된 스카이프의 로그인 패턴 정보를 이용하여 수신된 트래픽 데이터에 대한 2차 분류가 수행되도록 하여, 수신된 트래픽 데이터가 스카이프의 트래픽 데이터인지 여부가 판단되도록 한다.
- [0061] 이는, 수신된 트래픽 데이터가 스카이프의 트래픽 데이터인지 여부를 판단함에 있어서, IP와 포트번호로 구성된 사용자 리스트를 통해 판단하는 것이 로그인 패턴 분석을 통해 판단하는 것에 비해 속도 및 성능 면에 있어서 효율적이기 때문이다.
- [0062] 이때, 분류 제어부(126)는 트래픽 데이터의 로그인 패턴이 미리 설정된 스카이프 어플리케이션의 로그인 패턴과 일치하는지 여부를 수신한 패킷을 검사하여 판단하되, 검사를 수행한 패킷의 개수가 미리 설정된 임계 개수를 초과한 경우에는 해당 트래픽 데이터가 스카이프 어플리케이션의 트래픽 데이터가 아닌 것으로 판단할 수 있다.
- [0063] 즉, 미리 설정된 스카이프의 로그인 패턴을 통해 판단할 경우, 로그인 패턴 비교부(124)를 통해 트래픽 데이터의 임계 개수의 패킷까지 검사가 수행되도록 하여, 트래픽 데이터를 스카이프의 트래픽 데이터로 분류할 수 있다. 만약, 임계 개수에 도달하지 못한 경우, 해당 개수를 채울 때까지 상기 판단이 보류되도록 할 수 있다. 또한, 검사를 수행한 패킷의 개수가 임계 개수를 초과한 경우에는 더 이상 검사를 수행하지 않고 해당 트래픽 데이터를 스카이프의 트래픽 데이터가 아니라고 판단할 수 있다.
- [0064] 사용자 리스트 갱신부(130)는 사용자 리스트 비교부(122) 및 로그인 패턴 비교부(124) 중 적어도 하나의 판단에 따라 트래픽 데이터가 스카이프의 트래픽 데이터로 분류되면, 해당 트래픽 데이터의 노드에 관련된 IP 및 포트번호 중 적어도 하나를 상기 사용자 리스트에 추가시킨다.
- [0065] 또한, 사용자 리스트 갱신부(130)는 해당 트래픽 데이터가 발생한 마지막 시간을 사용자 리스트에 저장할 수도 있다.
- [0066] 사용자 리스트 관리부(140)는 사용자 리스트에 등록된 노드에 관련된 정보를 관리한다. 예를 들어, 사용자 리스트 관리부(140)는 사용자 리스트에 포함된 노드에 대응되는 트래픽 데이터가 임계 시간 동안 발생하지 않은 경우, 해당 노드에서 더 이상 스카이프가 실행되고 있지 않다고 판단하고, 해당 노드에 관련된 정보를 사용자 리스트에서 삭제 처리할 수 있다. 이를 통해, 트래픽 분류 장치(100)의 부하를 감소시키고, 항상 최신의 노드

정보를 유지 및 관리할 수 있게 된다.

[0067] 사용자 리스트 백업부(150)는 주기적으로 사용자 리스트에 등록된 노드에 관련된 정보를 백업한다. 따라서, 트래픽을 분류하는 시스템의 작동이 중단되었을 경우, 백업된 정보를 통해 해당 정보들이 계속 유지 및 관리될 수 있다.

[0068] 본 발명의 실시 예에서는 TCP 및 UDP 프로토콜의 양방향 플로우 형태로 가공된 패킷을 수신하여 사용자 리스트가 구축될 수 있으며, 이에 따라 트래픽 데이터가 보다 정확하게 분류될 수 있다.

[0069] 한편, 후술될 도 4는 사용자 리스트를 기반으로 해당 트래픽 데이터를 스카이프의 트래픽 데이터로 분류하는 것을 나타낸다.

[0070] 그리고 후술될 도 5는 사용자 리스트를 사용하지 않고도 미리 설정된 스카이프의 로그인 패턴에 의해 해당 트래픽 데이터가 스카이프의 트래픽 데이터인지 여부를 판단하는 것을 나타낸다. 이하 도 4 및 도 5를 통해 이와 관련된 내용을 구체적으로 설명하기로 한다.

[0071] 도 4는 본 발명의 일 실시예에 따른 IP 및 포트번호를 이용하여 트래픽 데이터를 분류하는 방법을 도시한다. 도 5는 본 발명의 일 실시예에 따른 로그인 패턴을 이용하여 트래픽 데이터를 분류하는 방법을 도시한다.

[0072] 먼저, 도 4 및 도 5에 표시된 각 심볼들은 아래 표 1과 같이 정의될 수 있다.

표 1

심볼	설명
src	트래픽 데이터에 포함된 소스 노드 {IP, 포트} 정보
dst	트래픽 데이터에 포함된 목적지 노드 {IP, 포트} 정보
L	사용자 {IP, 포트} 리스트
P[]	패킷 페이로드의 옥텟(octet) 배열 (16진수)
P _n	트래픽 데이터 내의 페이로드가 있는 n번째 패킷의 크기
a → L	a를 L에 추가함

[0073]

[0074] 그리고, 도 4 및 도 5에서 트래픽 데이터가 스카이프의 트래픽 데이터인지 여부를 검사할 때 수행되는 검사 조건(①~⑥), 프로토콜 및 패킷의 개수는 이하 표 2와 같이 정의될 수 있다.

표 2

패킷 개수		-	1	2	3	4	5
프로토콜	조건						
TCP/UDP	{IP, 포트}	①					
TCP	{IP}					②	
	{포트}		③				
UDP	{포트}		④	⑤			
	-		⑥				

[0075]

[0076] 상기 표 2에 있어서, 예컨대 ①번은 패킷의 페이로드(payload)를 검사하지 않고, 프로토콜이 TCP/UDP인 트래픽 데이터의 IP, 및 포트번호로 해당 트래픽 데이터가 스카이프의 로그인 패턴과 일치하는지 여부를 판단하는 것을 나타낸다. 그리고, ⑤번의 경우는 프로토콜이 UDP인 트래픽 데이터의 포트번호와 2번째 패킷의 페이로드를 이

용하여 해당 트래픽 데이터가 스카이프의 로그인 패턴과 일치하는지 여부를 판단하는 것을 나타낸다.

- [0077] 여기서, 상술된 표 2에 나타난 바와 같이, 예컨대 트래픽 데이터에서 최대 5개의 패킷만을 참조하기 때문에 플로우 데이터 단위로 수신된 트래픽 데이터가 스카이프의 로그인 패턴 특성을 갖는지 여부에 대한 빠른 실시간 검사가 가능하다.
- [0078] 이러한 조건들(①~⑥)을 적용시켜, 해당 트래픽 데이터를 스카이프의 트래픽 데이터로 분류하는 방법에 대해서 보다 구체적으로 설명하기로 한다. 이해를 돕기 위해, 표 1, 표 2, 도 4 및 도 5를 함께 참조하기 바란다. 그리고, 도 4 및 도 5의 ①~⑥은 상술된 표 2의 ①~⑥의 조건들을 대응시켜 표시한 것이다.
- [0079] 도 4에 도시된 바와 같이, 패킷의 페이로드를 검사하지 않는 ①번 조건의 경우, 트래픽 데이터의 IP 및 포트번호에 매칭되는 사용자 리스트의 정보가 존재하는지 여부를 비교함으로써, 수신된 트래픽 데이터가 스카이프의 트래픽 데이터인지 여부가 판단될 수 있다.
- [0080] 구체적으로, 도 4에 도시된 바와 같은 트래픽 데이터에 포함된 소스(src)의 IP 및 포트번호에 매칭되는 정보가 사용자 리스트에 존재할 경우(src=L), 트래픽 데이터에 포함된 패킷 목적지(dst)의 IP 및 포트번호가 사용자 리스트에 추가될 수 있다(dst->L).
- [0081] 여기서, 해당 포트번호가 스카이프의 포트번호라면 그와 연관되어 트래픽을 발생하는 다른 노드의 포트번호도 스카이프의 포트번호일 수 있음을 나타낸다.
- [0082] 또한, 트래픽 데이터에 포함된 목적지(dst)의 IP 및 포트번호에 매칭되는 정보가 사용자 리스트에 존재할 경우(dst=L), 트래픽 데이터에 포함된 패킷 목적지(src)의 IP 및 포트번호가 사용자 리스트에 추가될 수 있다(src->L).
- [0083] 또한, 트래픽 데이터에 포함된 목적지(dst)의 IP가 사용자 리스트에 존재하고(dst=L), 목적지(dst)의 포트가 80443인 경우 트래픽 데이터에 포함된 패킷 목적지(src)의 IP 및 포트번호가 사용자 리스트에 추가될 수 있다(src->L).
- [0084] 그리고, 프로토콜이 TCP인 트래픽 데이터의 IP와 3~5번째 패킷의 사이즈를 검사하는 ②번 조건의 경우, 트래픽 데이터의 소스 IP에 매칭되는 사용자 리스트의 정보가 존재하는지 여부를 비교하고(src(IP)=L), 트래픽 데이터의 3 내지 5번째 패킷의 사이즈(|P3|, |P4|, |P5|)를 이용하여 해당 트래픽 데이터가 스카이프의 트래픽인지를 판단한다.
- [0085] 한편, 도 5에 도시된 바와 같이, 상기 ③, ④, ⑤, 및 ⑥번 조건의 경우, 상술된 사용자 리스트를 참조하지 않고도, 패킷의 사이즈와 포트번호, 페이로드 시그니처(signature) 등을 통해 임계 개수의 패킷(예컨대 최대 2개의 패킷)까지 검사하여 수신된 트래픽 데이터가 스카이프의 로그인 패턴과 일치하는지 여부가 판단될 수 있다.
- [0086] 구체적으로, 프로토콜이 TCP인 트래픽 데이터의 포트번호 및 1번째 패킷의 사이즈를 검사하는 ③번 조건의 경우, 도 5에 도시된 바와 같은 로그인 패턴과 일치하여 해당 트래픽 데이터의 목적지 포트번호가 33033이고 패킷 페이로드의 사이즈가 5 바이트, 페이로드 시그니처가 16 03 01 00 00 일 경우, 이는 스카이프 로그인 과정에서 나타나는 패턴으로 판단된다. 그리고 목적지 노드가 SN이므로, 이는 스카이프 네트워크의 또 다른 사용자를 나타내므로, 사용자 리스트에 추가된다(dst->L).
- [0087] 그리고, 프로토콜이 UDP인 트래픽 데이터의 포트번호 및 2번째 패킷의 사이즈를 검사하는 ⑤번 조건의 경우도 마찬가지로 도 5에 도시된 바와 같은 패턴을 나타내면, 스카이프 로그인 과정에서 나타나는 패턴으로 판단되고, 이 또한 스카이프 네트워크의 다른 사용자인기 때문에 소스와 목적지의 IP 및 포트번호가 사용자 리스트에 추가될 수 있다(src->L, dst->L).
- [0088] 그리고, 프로토콜이 UDP인 트래픽 데이터의 포트번호 및 1번째 패킷의 사이즈를 검사하는 ④번과 프로토콜이 UDP인 트래픽 데이터의 포트번호 및 1~2번째 패킷의 사이즈를 검사하는 ⑥번 조건의 경우, 해당 스카이프 어플리케이션 사용시 주기적으로 다른 새로운 노드와 계속적으로 UDP 패킷을 주고 받고, 패킷의 사이즈와 페이로드 시그니처가 존재하여 해당 스카이프 어플리케이션이 주기적으로 새로운 SN 노드들을 찾는 패킷으로 추정될 수 있다.
- [0089] 상술된 바와 같이, 트래픽 데이터가 스카이프의 트래픽 데이터로 판단된 경우, 사용자 리스트에 대한 갱신이 이루어지게 된다. 여기서, 도 4 및 도 5에서 설명된 바와 같이 사용자 리스트에 해당 노드에 관련된 정보의 추가가 이루어질 수 있다.

- [0090] 한편, 트래픽 데이터 분류부(120)는 수신된 트래픽 데이터의 로그인 패턴이 미리 설정된 패턴과 일치하는지 여부 및 수신된 트래픽 데이터의 노드가 사용자 리스트에 포함되어 있는지 여부 중 적어도 하나의 판단 결과에 따라 수신된 트래픽 데이터가 스카이프의 트래픽 데이터가 아니라고 판단된 경우, 해당 트래픽 데이터를 다른 어플리케이션의 트래픽 데이터로 분류하게 된다.
- [0091] 여기서 트래픽 데이터 분류부(120)는, 트래픽 데이터의 임계 개수의 패킷(예컨대 프로토콜이 TCP인 트래픽 데이터의 경우 최대 5개, 프로토콜이 UDP인 트래픽 데이터의 경우 최대 2개)을 검사하여 해당 트래픽 데이터의 스카이프 트래픽 데이터인지 여부를 결정할 수 있다.
- [0092] 이때, 정해진 트래픽 데이터의 패킷 개수에 대해서 검사가 이루어지지 않아도, 트래픽 데이터의 수신이 종료되면 스카이프의 트래픽 데이터가 아니라고 판단할 수 있다.
- [0093] 이와 같이, 갱신된 사용자 리스트를 이용하여 트래픽 데이터를 분류할 경우, 해당 트래픽 데이터의 헤드 정보(IP, 포트)만으로도 편리하게 스카이프 트래픽 인지 여부가 결정될 수 있다.
- [0094] 도 6은 본 발명의 실시예에 따른 스카이프 트래픽 분류 방법의 순서도이다.
- [0095] 도 6에 도시된 바와 같이, 먼저, 트래픽 데이터 수신부(110)가 플로우 데이터 단위로 스카이프 어플리케이션의 트래픽 데이터를 수신한다(S601).
- [0096] 다음으로, 미리 설정된 스카이프 어플리케이션의 로그인 패턴 및 사용자 리스트 중 적어도 하나를 이용하여 수신된 트래픽 데이터가 스카이프의 트래픽 데이터인지 여부를 판단한다(S611). 여기서, 스카이프 어플리케이션의 트래픽 데이터의 노드에 관련된 IP 및 포트번호 중 적어도 하나를 포함하는 사용자 리스트를 이용하여 수신된 트래픽 데이터가 스카이프 어플리케이션의 트래픽 데이터인지 여부를 먼저 판단할 수 있다.
- [0097] 그리고, 판단한 결과, 트래픽 데이터의 노드에 관련된 정보가 사용자 리스트에 포함되어 있지 않은 경우, 미리 설정된 스카이프 어플리케이션의 로그인 패턴과 수신된 트래픽 데이터의 로그인 패턴과의 일치 여부를 판단할 수 있다.
- [0098] 그리고, 설정된 임계 개수의 패킷까지 검사하여 해당 트래픽 데이터가 스카이프의 트래픽 데이터인지 여부를 판단할 수 있다. 여기서, 임계 개수에 도달하지 못한 경우에는 판단을 보류하고, 임계 개수를 초과한 경우에는 트래픽 데이터가 스카이프가 아닌 것으로 판단할 수 있다.
- [0099] 다음으로, 판단 결과에 따라 해당 트래픽 데이터를 스카이프 어플리케이션의 트래픽 데이터로 분류한다(S621).
- [0100] 다음으로, 스카이프 어플리케이션의 트래픽 데이터로 분류된 트래픽 데이터에 포함된 노드에 관련된 정보를 사용자 리스트에 추가시킨다(S631). 여기서, 해당 트래픽 데이터의 노드에 관련된 IP 및 포트번호 중 적어도 하나가 사용자 리스트에 추가될 수 있다.
- [0101] 한편, 상술된 스카이프의 로그인 패턴 특성은, 도 2에서 설명된 바와 같이, SC에 저장되어 있던 SN의 리스트를 기반으로 하여 SC가 SN에게 UDP 패킷을 전송하고, SC가 UDP 패킷에 대해 응답 패킷을 전달해온 SN과 TCP 연결을 수행하고, SC가 제 1 목적지 포트번호(예컨대 dstport 12350)를 갖는 노드로 UDP 패킷을 전송하는 과정에서의 패턴 특징을 포함할 수 있다.
- [0102] 또한, 상술된 스카이프의 로그인 패턴 특성은, SC가 DNS 서버로 쿼리를 전송하여 스카이프 웹 서버의 IP를 수신하고, SC가 스카이프 웹 서버와 TCP 연결을 수행하고, SC가 제 2 목적지 포트의 노드와 접속을 해제하며 제 1 목적지 포트번호(예컨대 dstport 33033)의 노드와 TCP 연결을 이루었을 경우 해당 접속을 해제하는 과정에서의 패턴 특징을 포함할 수 있다.
- [0103] 그리고, 상술된 내용에서는 스카이프의 트래픽 데이터를 탐지하고 분류하는 것을 예로 들어 설명하였지만, 이외에도 특정 어플리케이션의 트래픽 데이터를 탐지하고 분류할 수도 있음은 물론이다.
- [0104] 한편 이하의 표 3에 나타난 바와 같이, 본 발명의 실시예에 따른 사용자 포트 추출 방법은 별도의 트레이닝 시간이 필요 없으며 분류결과 또한 높은 정확도를 가지고 있다.
- [0105] 또한, 특정 방법론(Signature-based, Port-based)만을 사용하지 않고 스카이프 트래픽의 패턴을 분석하여 동적 포트를 찾아내어 스카이프 노드의 IP와 포트의 리스트를 생성하고, 특정 방법론(Signature-based, Port-based)을 혼합하여 트래픽들을 분류할 수 있었다.

[0106] 그리고, Precision(100%), Recall(100%) 이루는 만족할만한 결과를 내었다. 참고로, FN(False Negative)은 스카이프의 트래픽 데이터이며, FP(False Positive)는 다른 어플리케이션의 트래픽이나 스카이프로 분류한 트래픽이다.

표 3

	수치
TP (# of flow)	11828
FP (# of flow)	0
FN (# of flow)	0
Precision (%)	100
Recall (%)	100

[0107]

[0108] 한편, 상술된 패턴 정보 저장부(115)는 캐쉬, ROM(Read Only Memory), PROM(Programmable ROM), EPROM(Erasable Programmable ROM), EEPROM(Electrically Erasable Programmable ROM) 및 플래쉬 메모리(Flash memory)와 같은 비휘발성 메모리 소자 또는 RAM(Random Access Memory)과 같은 휘발성 메모리 소자 또는 하드디스크 드라이브(HDD, Hard Disk Drive), CD-ROM과 같은 저장 매체 중 적어도 하나로 구현될 수 있으나 이에 한정되지는 않는다.

[0109] 그리고, 도 3a 및 도 3b에서 도시된 각각의 구성요소는 일종의 '모듈'로 구성될 수 있다. 상기 '모듈'은 소프트웨어 또는 Field Programmable Gate Array(FPGA) 또는 주문형 반도체(ASIC, Application Specific Integrated Circuit)와 같은 하드웨어 구성요소를 의미하며, 모듈은 어떤 역할들을 수행한다. 그렇지만 모듈은 소프트웨어 또는 하드웨어에 한정되는 의미는 아니다. 모듈은 어드레싱할 수 있는 저장 매체에 있도록 구성될 수도 있고 하나 또는 그 이상의 프로세서들을 실행시키도록 구성될 수도 있다. 구성요소들과 모듈들에서 제공되는 기능은 더 작은 수의 구성요소들 및 모듈들로 결합되거나 추가적인 구성요소들과 모듈들로 더 분리될 수 있다.

[0110] 한편 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.

[0111] 그리고 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.

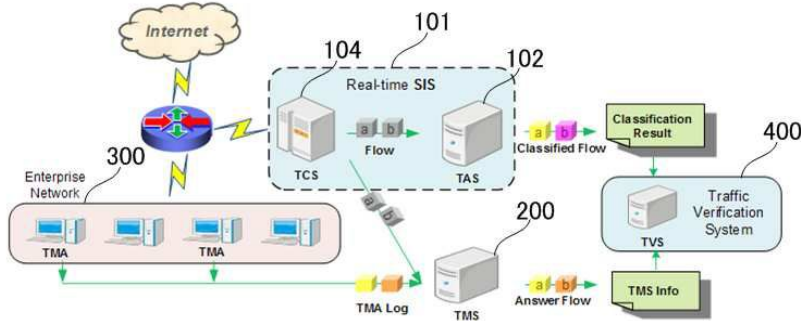
부호의 설명

- [0112]
- 110: 트래픽 데이터 수신부
 - 115: 패턴 정보 저장부
 - 120: 트래픽 데이터 분류부
 - 122: 사용자 리스트 비교부
 - 124: 로그인 패턴 비교부
 - 126: 분류 제어부
 - 130: 사용자 리스트 갱신부
 - 140: 사용자 리스트 관리부

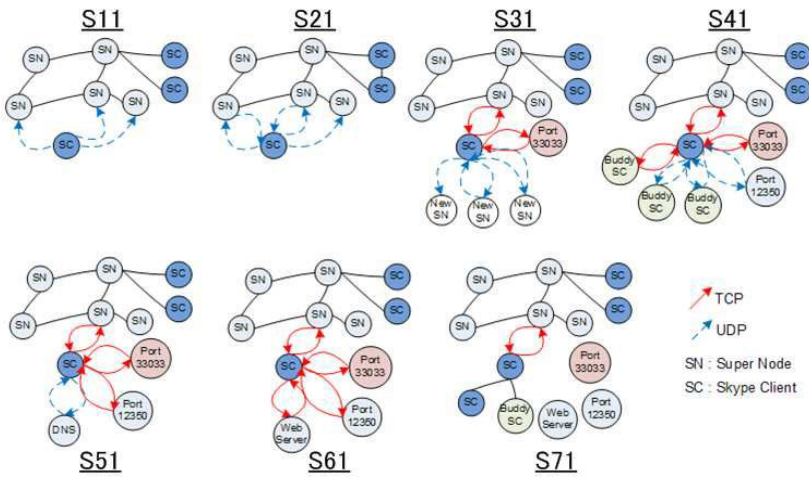
150: 사용자 리스트 백업부

도면

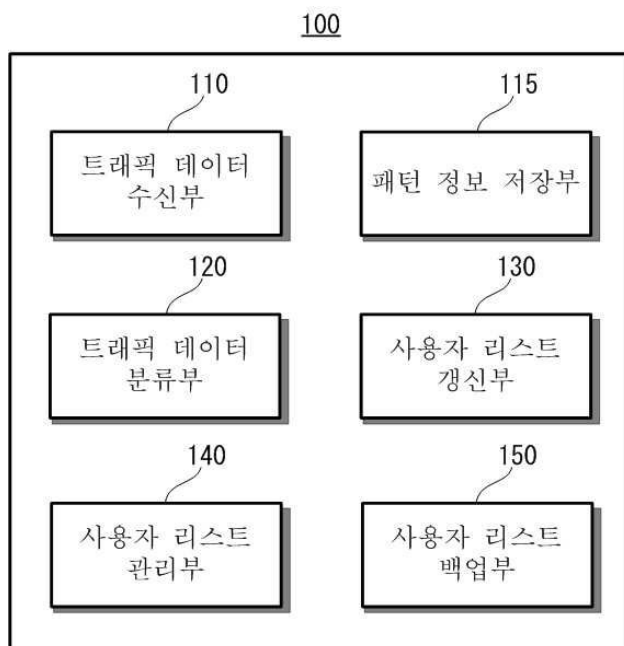
도면1



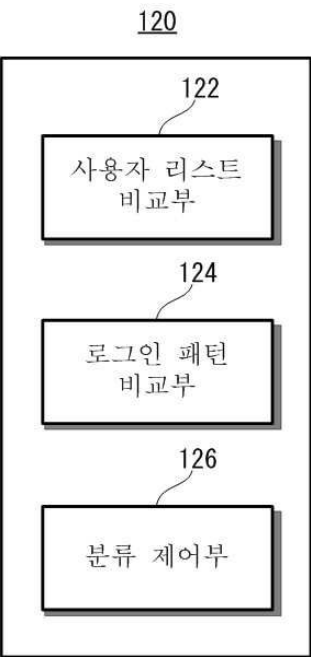
도면2



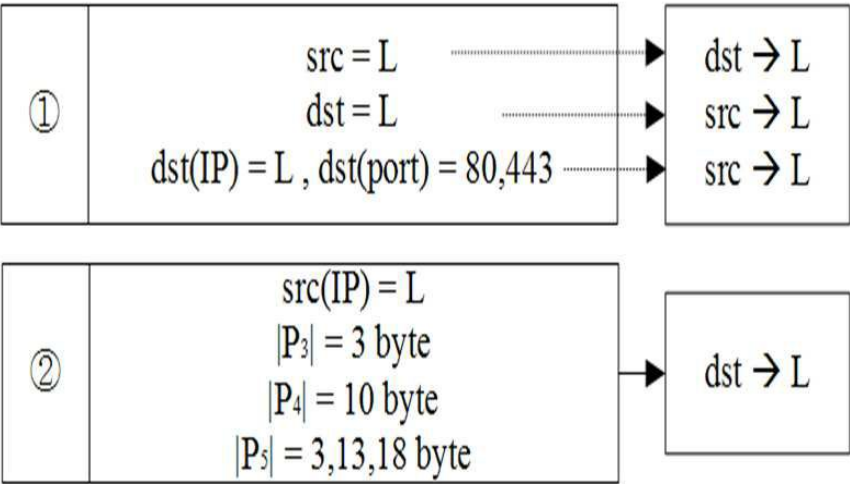
도면3a



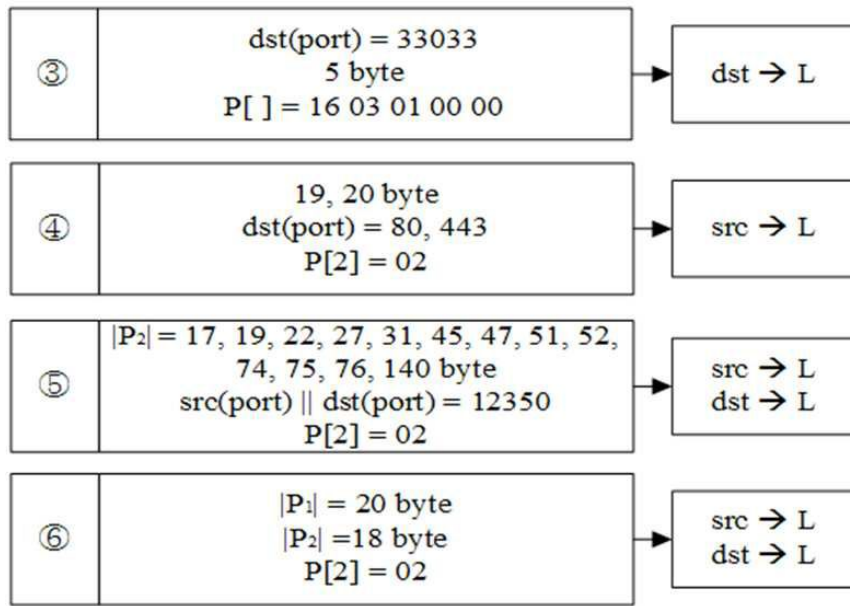
도면3b



도면4



도면5



도면6

