(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2011-0103038
(43) 공개일자 2011년09월20일

(51) Int. Cl.
H04L 12/26 (2006.01)
(21) 출원번호 10-2010-0022143
(22) 출원일자 2010년03월12일
심사청구일자 2010년03월12일

(71) 출원인
고려대학교 산학협력단
서울 성북구 안암동5가 1
(72) 발명자
박대희
서울특별시 강남구 도곡동 타워팰리스 A동 1601호
유재학
충청북도 옥천군 옥천읍 죽향리 옥향아파트 103동 803호
(뒷면에 계속)
(74) 대리인
특허법인엠에이피에스

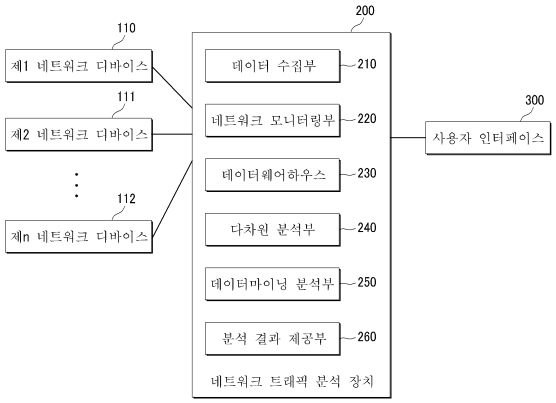
전체 청구항 수 : 총 14 항

(54) 네트워크 트래픽 분석 장치 및 방법

(57) 요 약

네트워크 트래픽 분석 장치는, 네트워크 디바이스로부터 수집된 네트워크 트래픽 데이터를 기설정된 기준에 따라 저장하고, 추상화 정도에 따라 복수의 계층으로 구성된 복수의 차원(dimension)에 기초하여 네트워크 트래픽 데이터에 대한 다차원 데이터 모델을 생성하고, 다차원 데이터 모델을 기선택된 차원 및 계층에 따라 OLAP(On-Line Analytical Processing) 연산하여 다차원적 분석을 수행하며, 다차원 데이터 모델에 대해 기선택된 차원 및 계층에 따른 군집 성장도 및 군집 상관도를 분석한다.

대 표 도 - 도1



(72) 발명자

이한성

경기도 용인시 기흥구 동백동 동백 택지개발지구 C2-1블럭 백현마을 2607동 2501호

임영희

경기도 수원시 영통구 매탄3동 우남퍼스트빌아파트 208동 702호

강봉수

경기도 용인시 기흥구 상하동 강남마을주공아파트 801동 103호

오승근

인천광역시 부평구 삼산동 삼산아파트 3동 501호

박승진

전라남도 목포시 옥암동 제일2차 하이빌마트 202동 105호

특허청구의 범위

청구항 1

네트워크 트래픽 분석 장치에 있어서,

네트워크 디바이스로부터 수집된 네트워크 트래픽 데이터가 기설정된 기준에 따라 저장된 데이터베이스;

추상화 정도에 따라 복수의 계층으로 구성된 복수의 차원(dimension)에 기초하여 상기 저장된 네트워크 트래픽 데이터에 대한 다차원 데이터 모델을 생성하고, 상기 다차원 데이터 모델을 기선택된 차원 및 계층에 따라 OLAP(On-Line Analytical Processing) 연산하여 다차원적 분석을 수행하는 다차원 분석부; 및

상기 다차원 데이터 모델에 대해 기선택된 차원 및 계층에 따른 군집 성장도 및 군집 상관도를 분석하는 데이터 마이닝 분석부를 포함하는 네트워크 트래픽 분석 장치.

청구항 2

제 1 항에 있어서,

상기 데이터마이닝 분석부는,

상기 다차원 데이터 모델에 대해 상기 기선택된 차원 및 계층에 따른 측도 값을 클러스터링하여 군집을 생성하고, 기설정된 적어도 두 개의 기간을 기준으로 상기 군집들에 대한 성장 지수 및 가중 성장 지수를 산출하여 군집 성장도를 분석하는 네트워크 트래픽 분석 장치.

청구항 3

제 1 항에 있어서,

상기 데이터마이닝 분석부는,

상기 다차원 데이터 모델에 대해 기선택된 제 1 차원 및 제 1 계층에 따른 측도 값을 클러스터링하여 군집들을 생성하고, 기선택된 제 2 차원 및 제 2 계층을 기준으로 상기 생성된 군집들에 대한 유사도 행렬을 산출하여 군집 상관도를 분석하는 네트워크 트래픽 분석 장치.

청구항 4

제 1 항에 있어서,

상기 다차원 분석부는,

상기 다차원 데이터 모델에 대해 상기 기선택된 차원 및 계층에 대응하는 측도 값을 요약 및 집계 중 적어도 하나로 처리하여 상기 다차원적 분석을 수행하는 네트워크 트래픽 분석 장치.

청구항 5

제 2 항 내지 제 4 항 중 어느 한 항에 있어서,

상기 측도 값은,

상기 복수의 차원 및 각 계층에 대응하는 대역폭 및 플로우 개수 중 적어도 하나인 네트워크 트래픽 분석 장치.

청구항 6

제 1 항 내지 제 4 항 중 어느 한 항에 있어서,

상기 복수의 차원은,

시간, 위치, 애플리케이션의 종류 및 프로토콜의 종류 중 적어도 하나의 차원을 포함하는 네트워크 트래픽 분석 장치.

청구항 7

제 1 항 내지 제 4 항 중 어느 한 항에 있어서,

상기 다차원 분석 및 데이터마이닝 분석의 기준이 되는 상기 차원 및 계층을 사용자가 선택할 수 있는 웹 페이지를 제공하는 사용자 인터페이스를 더 포함하는 네트워크 트래픽 분석 장치.

청구항 8

네트워크 트래픽 분석 방법에 있어서,

발생되는 네트워크 트래픽 데이터를 수집하는 단계;

상기 수집된 네트워크 트래픽 데이터들을 기설정된 기준에 따라 정렬한 이력데이터를 저장하는 단계;

상기 저장된 이력데이터를 기설정된 복수의 차원 및 차원 별 복수의 추상화 정도 계층에 따라 다차원 데이터 모델로 생성하는 단계;

상기 다차원 데이터 모델에 대해 기선택된 차원 및 추상화 정도 계층에 따른 OLAP(On-Line Analytical Processing) 연산을 수행하여 다차원적 분석을 수행하는 단계; 및

상기 다차원 데이터 모델에 대해 기선택된 차원 및 추상화 정도 계층에 따른 군집 성장도의 분석 및 군집 상관도의 분석 중 적어도 하나를 수행하여 데이터마이닝 분석을 수행하는 단계를 포함하는 네트워크 트래픽 분석 방법.

청구항 9

제 8 항에 있어서,

상기 데이터마이닝 분석을 수행하는 단계는,

상기 다차원 데이터 모델에 대해 기선택된 차원 및 추상화 정도 계층에 따른 측도 값을 클러스터링하여 군집을 생성하는 단계; 및

기설정된 적어도 두 개의 기간을 기준으로 상기 군집들에 대한 성장 지수 및 가중 성장 지수를 산출하여 상기 군집 성장도를 분석하는 단계를 포함하는 네트워크 트래픽 분석 방법.

청구항 10

제 8 항에 있어서,

상기 데이터마이닝 분석을 수행하는 단계는,

상기 다차원 데이터 모델에 대해 기선택된 제 1 차원 및 제 1 추상화 정도 계층에 따른 측도 값을 클러스터링하여 군집들을 생성하는 단계; 및

기선택된 제 2 차원 및 제 2 추상화 정도 계층을 기준으로 상기 생성된 군집들에 대한 유사도 행렬을 산출하여 상기 군집 상관도를 분석하는 단계를 포함하는 네트워크 트래픽 분석 방법.

청구항 11

제 8 항에 있어서,

상기 다차원적 분석을 수행하는 단계는,

상기 다차원 데이터 모델에 대해 상기 기선택된 차원 및 추상화 정도 계층에 대응하는 측도 값을 집계 및 요약 중 적어도 하나로 처리하여 상기 다차원적 분석을 수행하는 네트워크 트래픽 분석 방법.

청구항 12

제 9 항 내지 제 11 항 중 어느 한 항에 있어서,

상기 측도 값은,

상기 복수의 차원 및 각 추상화 정도 계층에 대응하는 대역폭 및 플로우 개수 중 적어도 하나인 네트워크 트래픽 분석 방법.

청구항 13

제 8 항 내지 제 11 항 중 어느 한 항에 있어서,

상기 복수의 차원은,

시간, 위치, 애플리케이션의 종류 및 프로토콜의 종류 중 적어도 하나의 차원을 포함하는 네트워크 트래픽 분석 방법.

청구항 14

제 8 항 내지 제 11 항 중 어느 한 항에 있어서,

상기 다차원적 분석을 수행하는 단계 이전에,

상기 다차원적 분석을 수행하는 단계 및 상기 데이터마이닝 분석을 수행하는단계의 기준이 되는 상기 차원 및 추상화 정도 계층에 대한 사용자의 선택 정보를 입력 받는 단계를 더 포함하는 네트워크 트래픽 분석 방법.

명 세 서

기 술 분 야

[0001] 본 발명은 네트워크 트래픽에 대한 다차원적 분석 및 데이터마이닝 분석을 수행하는 네트워크 트래픽 분석 장치 및 그 방법에 관한 것이다.

배 경 기 술

[0002] 인터넷 서비스의 급속한 발전에 의하여 인터넷 사용자들의 네트워크 트래픽 사용량이 폭발적으로 증가하였다. 이에 따라 네트워크 과부하를 최소화하고 안정적인 네트워크 서비스 지원 및 운용을 위한 네트워크 트래픽 분석 시스템이 요구되고 있다.

[0003] 이와 같은, 네트워크 트래픽 분석 시스템은 네트워크 내의 모든 통신 요소를 관찰, 제어하고 나아가 네트워크 변화에 응용함으로써 보다 안전하고 효율적으로 네트워크가 운영될 수 있도록 하는 모든 행위를 포함하는 시스템을 일컫는다.

[0004] 네트워크 트래픽 분석 시스템에 관한 연구 중 네트워크 사업자(NSP)나 인터넷 사업자(ISP)의 코어 네트워크 및

회사, 학교, 기관의 엔터프라이즈 네트워크의 트래픽 현황을 실시간으로 파악하기 위한 방법으로 네트워크 모니터링 시스템에 관한 연구가 활발히 진행 중이다.

[0005] 일반적으로, 네트워크 모니터링 시스템은 라우터나 스위치에 내재된 에이전트로부터 트래픽 통계정보를 주기적으로 수집하여 과거와 현재의 네트워크 트래픽 발생 현황을 다양한 형태로 보여준다. 이러한, 네트워크 모니터링 시스템에서는 트래픽 현황을 네트워크 토폴로지 정보와 함께 제공하는 Network Weather Map(NWM) 방법과 시간 추이에 따른 트래픽 변화를 그래프로 제공하는 방법 등이 있다.

[0006] 이러한, 네트워크 트래픽 분석과 관련된 대부분의 연구들은 실시간적으로 네트워크의 트래픽 모니터링을 통하여 현재의 네트워크 상황을 간단한 통계적 분석결과와 함께 제공하는 서비스에 초점을 맞추고 있다.

[0007] 구체적으로, 종래의 네트워크 모니터링 시스템에서는 네트워크 관리자의 관리 범위에 포함된 모든 네트워크 장비 및 링크를 모니터링 한다. 이와 같이 함으로써, 트래픽 정보의 실시간 제공 및 저장을 제공하고, 트래픽 정보 수집 시 네트워크의 부하를 최소화하여 웹 상으로 관리자에게 네트워크 상황을 실시간으로 제공한다.

[0008] 그러나, 네트워크 관리자의 분석 대상이 늘어나게 되면 대용량의 트래픽 정보는 복잡한 구조로 표현되어 효율적인 분석 정보의 전달과 실시간 모니터링이 어려워지게 된다. 따라서, 보다 안정적인 네트워크 환경의 적극적인 지원 및 원활한 QoS의 제공을 위해서는 대용량의 네트워크 트래픽 데이터의 효과적인 저장 및 빠른 질의 응답 시간을 제공하고 네트워크 트래픽 데이터에 잠재된 패턴을 탐사할 수 있는 네트워크 분석 장치가 요구된다.

발명의 내용

해결하려는 과제

[0009] 본 발명은 상술한 문제점을 해결하기 위한 것으로써, 네트워크 트래픽에 대해 구축된 다차원 데이터 모델을 이용하여 각 차원의 추상화 정도에 따른 다차원적 분석을 수행하고, 상기 다차원적 분석 수행 결과를 이용하여 데이터마이닝 분석을 수행하는 네트워크 트래픽 분석 장치 및 방법을 제공하고자 한다.

과제의 해결 수단

[0010] 상술한 기술적 과제를 달성하기 위한 기술적 수단으로서, 본 발명의 일 측면에 따른 네트워크 트래픽 분석 장치는, 네트워크 디바이스로부터 수집된 네트워크 트래픽 데이터가 기설정된 기준에 따라 저장된 데이터베이스; 추상화 정도에 따라 복수의 계층으로 구성된 복수의 차원(dimension)에 기초하여 상기 저장된 네트워크 트래픽 데이터에 대한 다차원 데이터 모델을 생성하고, 상기 다차원 데이터 모델을 기선택된 차원 및 계층에 따라 OLAP(On-Line Analytical Processing) 연산하여 다차원적 분석을 수행하는 다차원 분석부; 및 상기 다차원 데이터 모델에 대해 기선택된 차원 및 계층에 따른 군집 성장도 및 군집 상관도를 분석하는 데이터마이닝 분석부를 포함한다.

[0011] 또한, 본 발명의 다른 측면에 따른 네트워크 트래픽 분석 방법은, 발생되는 네트워크 트래픽 데이터를 수집하는 단계; 상기 수집된 네트워크 트래픽 데이터들을 기설정된 기준에 따라 정렬한 이력데이터를 저장하는 단계; 상기 저장된 이력데이터를 기설정된 복수의 차원 및 차원 별 복수의 추상화 정도 계층에 따라 다차원 데이터 모델로 생성하는 단계; 상기 다차원 데이터 모델에 대해 기선택된 차원 및 추상화 정도 계층에 따른 OLAP(On-Line Analytical Processing) 연산을 수행하여 다차원적 분석을 수행하는 단계; 및 상기 다차원 데이터 모델에 대해 기선택된 차원 및 추상화 정도 계층에 따른 군집 성장도의 분석 및 군집 상관도의 분석 중 적어도 하나를 수행하여 데이터마이닝 분석을 수행하는 단계를 포함한다.

발명의 효과

[0012] 전술한 본 발명의 과제 해결 수단에 의하면, 네트워크 트래픽 데이터에 대한 다차원 데이터 모델을 이용하여 다차원적 분석 및 데이터마이닝 분석 등의 지적 구조 분석을 수행함으로써, 네트워크 과부하를 최소화하고 안정적인 네트워크 서비스를 제공할 수 있다.

[0013] 즉, 네트워크 관리자는 네트워크 트래픽에 내재된 잠재적 패턴을 발견함으로써 안정적인 네트워크 환경 지원을

위한 의사 결정과 QoS 정책 입안의 근거로써 상기 지적 구조 분석의 결과를 활용할 수 있다.

도면의 간단한 설명

[0014] 도 1은 본 발명의 실시예에 따른 네트워크 트래픽 분석 장치의 구성을 나타내는 블록도이다.

도 2는 본 발명의 실시예에 따른 다차원 분석부의 구성을 나타내는 도면이다.

도 3은 본 발명의 실시예에 따른 데이터마이닝 분석부의 구성을 나타내는 도면이다.

도 4는 본 발명의 실시예에 따른 데이터 큐브를 설명하기 위한 도면이다.

도 5는 본 발명의 실시예에 따른 데이터 큐브 모델의 큐보이드(cuboid)를 설명하기 위한 도면이다.

도 6은 본 발명의 실시예에 따른 스타 스키마 모델을 설명하기 위한 도면이다.

도 7은 본 발명의 실시예에 적용되는 데이터 큐브 모델에서 시간 차원의 계층을 나타내는 도면이다.

도 8은 본 발명의 실시예에 따른 데이터 큐브 모델에서 위치 차원의 계층을 나타내는 도면이다.

도 9는 본 발명의 실시예에 따른 데이터 큐브 모델에서 애플리케이션 차원의 계층을 나타내는 도면이다.

도 10은 본 발명의 실시예에 따른 데이터 큐브 모델에서 프로토콜 차원의 계층을 나타내는 도면이다.

도 11은 본 발명의 실시예에 따른 OLAP 연산 과정을 설명하기 위한 도면이다.

도 12는 본 발명의 실시예에 따른 네트워크 트래픽에 대한 다차원 분석 결과 제공 화면을 나타내는 도면이다.

도 13은 본 발명의 실시예에 따른 네트워크 트래픽에 대한 성장도 분석 결과를 나타내는 도면이다.

도 14는 본 발명의 실시예에 따른 네트워크 트래픽에 대한 상관도 분석 결과를 나타내는 도면이다.

도 15는 본 발명의 실시예에 따른 네트워크 트래픽 분석 방법을 설명하기 위한 순서도이다.

발명을 실시하기 위한 구체적인 내용

[0015] 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.

[0016] 명세서 전체에서,ㄹ 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.

[0017] 도 1은 본 발명의 실시예에 따른 네트워크 트래픽 분석 장치의 구성을 나타내는 블록도이다.

[0018] 그리고, 도 2는 본 발명의 실시예에 따른 다차원 분석부의 구성을 나타내는 도면이고, 도 3은 본 발명의 실시예에 따른 데이터마이닝 분석부의 구성을 나타내는 도면이다.

[0019] 도 1에서와 같이, 본 발명의 실시예에 따른 네트워크 트래픽 분석 장치(200)는 라우터(router), 스위치(switch) 등의 네트워크 디바이스(110, 111, 112)로부터 트래픽 데이터를 수집하여 다차원 분석 및 데이터마이닝 분석하고, 네트워크 트래픽에 대한 다차원 분석 및 데이터마이닝 분석의 결과를 사용자가 인지 가능하도록 제공하는 사용자 인터페이스(300)를 통해 제공한다.

[0020] 구체적으로, 네트워크 트래픽 분석 장치(200)는 데이터 수집부(210), 네트워크 모니터링부(220), 데이터웨어하우스(230), 다차원 분석부(240), 데이터마이닝 분석부(250) 및 분석 결과 제공부(260)를 포함한다.

[0021] 데이터 수집부(210)는 복수의 네트워크 디바이스(110, 111, 112)를 경유하는 네트워크 트래픽 데이터를 실시간, 주기적 또는 기설정된 특정 시기에 수집하여 저장한다. 이때, 데이터 수집부(210)는 네트워크 트래픽 데이터로써 SNMP(Simple Network Management Protocol) MIB(Management Information Base) 및 플로우를 수집한다.

[0022] 참고로, 네트워크 트래픽 데이터는 플로우(flow)의 형식으로 표현되며, 플로우는 일정 기간 간격으로 수집된 트래픽 데이터(즉, 패킷) 중 헤더 정보가 동일한 트래픽 데이터를 의미한다. 이때, 플로우는 트래픽 데이터를 수

집한 위치, 플로우 생성 시간, 패킷 헤더 정보, 플로우에 포함된 패킷 수 및 크기 정보를 포함한다.

[0023] 네트워크 모니터링부(220)는 데이터 수집부(210)를 통해 수집된 네트워크 트래픽 데이터를 이용하여 트래픽량 검출 및 트래픽의 애플리케이션 분류 등을 통해 비정상적인 트래픽 사용에 대한 모니터링을 수행한다. 이때, 네트워크 모니터링부(220)는 네트워크 모니터링 결과를 분석 결과 제공부(260)로 전송하여, 사용자(예를 들어, 네트워크 관리자 등)가 인지 가능한 형태로 제공할 수 있도록 한다.

[0024] 데이터웨어하우스(230)는 데이터 수집부(210)를 통해 수집되는 대용량의 네트워크 트래픽 데이터들을 의사 결정을 위한 공통의 형식으로 변환하여 저장 및 관리한다. 이때, 데이터웨어하우스(230)는 상기 대용량의 네트워크 트래픽 데이터들을 목적, 시간 등의 기준에 따라 정렬하여 저장 공간에 저장하여 통합 관리한다.

[0025] 다차원 분석부(240)는 데이터웨어하우스(230)에 저장된 대용량의 네트워크 트래픽 데이터들을 복수의 차원(dimension) 별 사실(fact, 팩트)로 정렬한 다차원 데이터 모델을 생성하고, 생성된 다차원 데이터 모델을 각 차원의 추상화 정도에 따라 분석한다. 참고로, 복수의 차원은 각각 추상화 정도에 따른 복수의 개념 계층으로 분류되는 구조를 가지며, 팩트는 각 차원 별 측도(measure) 값을 의미한다.

[0026] 구체적으로, 다차원 분석부(240)는 대용량의 네트워크 트래픽 데이터들에 대한 다차원 데이터 모델을 생성하고, 상기 복수의 차원 중 사용자가 선택하는 차원 및 개념 계층에 따라 OLAP(On-line Analytical Processing) 연산을 통해 적어도 하나의 차원에 대한 데이터 큐브(Data Cube)를 생성한다. 그리고, 다차원 분석부(240)는 생성된 데이터 큐브를 이용하여 사용자의 질의에 대한 응답으로써 다차원적 분석을 수행한다.

[0027] 이때, 다차원 분석부(240)는 생성된 적어도 하나의 차원에 대한 데이터 큐브 및 다차원적 분석 결과를 데이터마이닝 분석부(250) 및 분석 결과 제공부(260)에 전달한다.

[0028] 데이터마이닝 분석부(250)는 다차원 분석부(240)를 통해 생성된 다차원 데이터 모델에 대해 설정된 기준에 따라 데이터마이닝을 수행하여 네트워크 트래픽 데이터에 대한 군집 성장도 및 군집 상관도를 분석한다.

[0029] 구체적으로, 데이터마이닝 분석부(250)는 사용자가 설정한 적어도 두 개의 기간을 기준으로 다차원 데이터 모델에 대해 사용자가 설정한 차원 및 개념 계층의 측도들을 클러스터링(clustering)한다. 그리고, 데이터마이닝 분석부(250)는 클러스터링 결과로써 생성되는 상기 두 개의 기간 별 군집(cluster)들에 대해서 성장 지수(GI: Growth Index) 및 가중 성장 지수(WGI: Weighted GI)를 산출하여 해당 차원에 대한 기간별 군집 성장도를 분석한다.

[0030] 또한, 데이터마이닝 분석부(250)는 사용자가 설정한 제 1 차원을 기준으로 다차원 데이터 모델에 대해 사용자가 설정한 제 2 차원 및 상기 제 2 차원에 대한 개념 계층의 측도들을 클러스터링한다. 그리고, 데이터마이닝 분석부(250)는 클러스터링 결과로써 생성되는 복수의 군집들에 대해서 상기 제 2 차원 및 상기 제 2 차원에 대한 개념 계층에 따른 유사도 함수를 산출하고, 산출된 유사도 함수를 이용하여 군집 상관도를 분석한다.

[0031] 이때, 데이터마이닝 분석부(250)는 데이터마이닝 분석 결과를 분석 결과 제공부(260)로 전달한다.

[0032] 분석 결과 제공부(260)는 다차원적 분석 결과 및 데이터마이닝 분석 결과를 사용자가 인지 가능한 정보로 구성하여 사용자 인터페이스(300)로 전달하고, 사용자 인터페이스(300)를 통해 사용자가 선택하는 차원 및 개념 계층의 정보를 수신하여 다차원 분석부(240) 및 데이터마이닝 분석부(250)에 전달한다.

[0033] 예를 들어, 본 발명의 실시예에 따른 분석 결과 제공부(260)는 다차원적 분석 결과 및 데이터마이닝 분석 결과를 그래프, 문자 또는 음성 데이터 등의 인지 가능한 정보로 변환한 웹 페이지 형식의 분석 결과 페이지를 생성하여 사용자 인터페이스(300)로 전달할 수 있다. 이때, 분석 결과 제공부(260)는 사용자가 다차원 데이터 모델에 대한 차원 및 개념 계층 정보를 선택하거나 직접 입력할 수 있는 정보 입력 기능을 포함하는 상기 웹 페이지 형식의 분석 결과 페이지를 생성할 수 있다.

[0034] 한편, 본 발명의 실시예에서는 네트워크 트래픽 분석 장치(200)가 사용자 인터페이스(300)를 통해 사용자가 선택한 차원 및 개념 계층의 정보를 수신하는 것을 나타내었다. 그런데, 본 발명의 다른 실시예에서는 네트워크 트래픽 분석 장치(200) 내에 사용자 인터페이스부(미도시)가 더 포함되고, 다차원 분석부(240) 및 데이터마이닝 분석부(250)가 사용자 인터페이스부(미도시)를 통해 상기 사용자가 선택한 차원 및 개념 계층의 정보를 직접 수신하는 것도 가능하다. 이때, 네트워크 트래픽 장치(200) 자체에 입력 수단(미도시) 및 출력 수단(미도시)이 더 설치되어 사용자가 차원 및 개념 계층의 정보를 네트워크 트래픽 장치(200)에 직접 입력하는 것이 가능하며, 네트워크 트래픽 장치(200)에서 출력되는 다차원적 분석 결과 및 데이터마이닝 분석 결과를 확인하는 것도 가능

하다.

[0035] 이하, 도 2 및 도 4 내지 도 10을 참조하여 본 발명의 실시예에 적용되는 다차원 데이터 모델 및 다차원 분석부(240)의 다차원적 분석 방법에 대해서 상세히 설명하도록 한다.

[0036] 먼저, 도 2에서와 같이, 본 발명의 실시예에 따른 다차원 분석부(240)는 다차원 분석 제어 모듈(241), 데이터 큐브 생성 모듈(242) 및 OLAP 연산 모듈(243)을 포함한다.

[0037] 다차원 분석 제어 모듈(241)은 데이터웨어하우스(230)에 저장된 대용량의 네트워크 트래픽 데이터들을 획득하고, 상기 네트워크 트래픽 데이터들을 데이터 큐브 생성 모듈(242)로 전송하여 네트워크 트래픽에 대한 다차원 데이터 모델을 생성하도록 한다.

[0038] 이때, 다차원 분석 제어 모듈(241)은 네트워크 트래픽에 대한 사용자의 질의에 응답하기 위해 사용자가 선택한 적어도 하나의 차원 및 개념 계층의 정보를 수신하여 데이터 큐브 생성 모듈(242) 및 OLAP 연산 모듈(243)로 전달한다.

[0039] 데이터 큐브 생성 모듈(242)은 기설정된 복수의 차원 및 각 차원 별 복수의 개념 계층을 기준으로 상기 대용량의 네트워크 트래픽 데이터들을 정렬하여 다차원 데이터 큐브를 생성한다.

[0040] 이때, 데이터 큐브 생성 모듈(242)는 자체적으로 다차원 데이터 모델의 생성 기준이 되는 복수의 차원 및 개념 계층의 정보가 기저장되어 있거나, 외부로부터 상기 복수의 차원 및 개념 계층의 정보를 입력 받을 수 있다.

[0041] 참고로, 다차원 데이터 큐브는 복수의 차원 및 각 하나의 팩트를 기준으로 생성된다. 본 발명의 실시예에서, 차원이란 네트워크 트래픽에 대한 질의에 응답하기 위한 대상이 되는 측면을 의미하며, 팩트란 숫자적으로 표현되는 측도 값을 의미한다. 이때, 본 발명의 실시예에서는 차원으로써 시간(time), 위치(location), 애플리케이션의 종류 (application) 및 프로토콜의 종류(protocol)가 설정되며, 팩트로써 대역폭 및 플로우 개수 중 적어도 하나가 설정되는 것을 나타내었다.

[0042] 아래에서는, 도 4 내지 도 6을 참조하여 데이터 큐브 생성 모듈(242)이 네트워크 트래픽 데이터에 대한 데이터 큐브를 생성하는 방법에 대해서 설명하도록 한다.

[0043] 도 4는 본 발명의 실시예에 따른 데이터 큐브를 설명하기 위한 도면이다.

[0044] 그리고, 도 5는 본 발명의 실시예에 따른 데이터 큐브 모델의 큐보이드(cuboid)를 설명하기 위한 도면이고, 도 6은 본 발명의 실시예에 따른 스타 스키마 모델을 설명하기 위한 도면이다.

[0045] 본 발명의 실시예에 따른 데이터 큐브 생성 모듈(242)은 대용량의 네트워크 트래픽 데이터에 대하여 4개 차원의 집합으로 큐보이드 격자를 구축하며, 각 큐보이드는 각 차원의 개념 계층 별로 대역폭 및 플로우 개수 중 적어도 하나에 대한 측도 값을 제공한다.

[0046] 예를 들어, 도 4에서는 2 차원의 데이터 큐브를 나타내었다.

[0047] 구체적으로, 도 4에서는 데이터 큐브 생성 모듈(242)이 네트워크 트래픽 데이터들에 대해 애플리케이션 차원 및 시간 차원과 두 차원의 팩트를 기준으로 데이터 큐브를 생성하는 것을 나타내었다.

[0048] 이때, 도 4에서는 시간 차원에 대한 복수의 개념 계층 중 월(month) 계층을 나타내었고, 애플리케이션 차원에 대한 복수의 개념 계층 중 세부 프로그램(specific types) 계층을 나타내었다. 그리고, 팩트는 해당 차원에 따른 대역폭을 나타내었다.

[0049] 또한, 도 5에서는 데이터 큐브 생성 모듈(242)이 4 차원의 데이터 큐브를 생성하기 위해 형성하는 큐보이드의 격자(lattice)를 나타내었다.

[0050] 구체적으로, 도 5에서는 데이터 큐브 생성 모듈(242)이 시간 차원, 위치 차원, 애플리케이션 종류 차원 및 프로토콜 종류 차원에 대한 큐보이드의 격자를 형성하는 것을 나타내었다.

[0051] 이때, 도 5에서는 가장 낮은 수준의 요약을 의미하는 4 차원의 기본 큐보이드(base cuboid)로부터 가장 높은 수준의 요약을 의미하는 0 차원의 정점 큐보이드(apex cuboid)를 나타내었다.

[0052] 참고로, 데이터 큐브 생성 모듈(242)은 큐보이드 격자들을 이용하여 다차원 데이터 모델을 생성할 시에 각 차원에 대한 팩트 테이블을 키(key)로써 연결하는 스타 스키마(star schema) 기법을 이용할 수 있다.

[0053] 예를 들어, 도 6에서는 본 발명의 실시예에 적용되는 스타 스키마에서 4 개의 차원 테이블과 각 차원 테이블 별

로 포함된 각 차원의 계층들을 나타내고, 각 차원에 대한 팩트 테이블(fact table)이 키로써 연결되는 것을 나타내었다.

[0054] 한편, 도 6에서 나타낸 바와 같이 본 발명의 실시예에 따른 다차원 데이터 큐브 생성 시 상기 4 개의 차원은 각각 복수의 추상화 정도 즉, 개념 계층으로 분류된다.

[0055] 아래에서는, 도 7 내지 도 10을 참조하여 본 발명의 실시예에 적용되는 각 차원들의 개념 계층에 대해서 설명하도록 한다.

[0056] 도 7은 본 발명의 실시예에 따른 데이터 큐브 모델에서 시간 차원의 계층을 나타내는 도면이고, 도 8은 본 발명의 실시예에 따른 데이터 큐브 모델에서 위치 차원의 계층을 나타내는 도면이다. 또한, 도 9는 본 발명의 실시예에 따른 데이터 큐브 모델에서 애플리케이션 차원의 계층을 나타내는 도면이며, 도 10은 본 발명의 실시예에 따른 데이터 큐브 모델에서 프로토콜 차원의 계층을 나타내는 도면이다.

[0057] 먼저, 도 7에서와 같이 본 발명의 실시예에 따른 시간 차원은 추상화 정도에 따라 시각(hour), 날짜(day), 주(week), 월(month), 분기(quarter), 년(year)으로 구분되는 복수의 계층으로 추상화된다.

[0058] 그리고, 도 8에서와 같이 본 발명의 실시예에 따른 위치(location) 차원은 네트워크 구성원의 IP 주소인 호스트(host), 부서(department), 층(floor), 건물 별 스위치 또는 라우터를 의미하는 건물(building) 등의 복수의 계층으로 추상화된다.

[0059] 그리고, 도 9에서와 같이 본 발명의 실시예에 따른 애플리케이션(Application) 차원은 최하위 개념인 각각의 애플리케이션의 종류(each application), 각 애플리케이션들의 세부 타입(specific types) 및 최상위 개념인 P2P 및 non P2P로 구분되는 복수의 계층으로 추상화된다.

[0060] 또한, 도 10에서와 같이 본 발명의 실시예에 따른 프로토콜(Protocol) 차원은 최하위 개념인 애플리케이션 계층(application layer), 전송 계층(transport layer) 및 최상위 개념인 네트워크 계층(network layer)으로 구분되는 복수의 계층으로 추상화된다.

[0061] 다시 도 2로 돌아와서, OLAP 연산 모듈(243)은 생성된 다차원 데이터 큐브를 이용하여 네트워크 트래픽 데이터에 대한 OLAP 연산을 통해 다차원적 분석을 수행한다.

[0062] 이때, OLAP 연산 모듈(243)은 사용자의 분석 목적(즉, 질의)에 따라 사용자가 선택한 차원의 정보와, 해당 차원 테이블에서의 추상화 정도(즉, 개념 계층)의 정보를 수신한다. 그리고, OLAP 연산 모듈(243)은 수신한 차원 및 개념 계층에 따라 OLAP 연산을 수행하고, OLAP 연산 수행에 따라 생성된 데이터 큐브에 해당하는 팩트 테이블의 측도(measure)를 분석한다.

[0063] 구체적으로, 본 발명의 실시예에 따른 OLAP 연산 모듈(243)은 도 11에서 나타낸 바와 같이 다차원 데이터 큐브에 대한 OLAP 연산을 수행한다.

[0064] 도 11은 본 발명의 실시예에 따른 OLAP 연산 과정을 설명하기 위한 도면이다.

[0065] 도 11에서는 4 개 차원에 대해 생성된 다차원 데이터 큐브들의 예를 나타내었으며, OLAP 연산 모듈(243)은 롤-업(roll-up), 드릴-다운(drill-down), 슬라이스(slice)와 다이스(dice) 등의 다양한 OLAP 연산을 수행하여 해당하는 데이터 큐브를 생성할 수 있다.

[0066] 이때, 롤-업 연산은 선택한 차원에 대한 개념 계층을 따라 올라가거나 차원 축소에 의하여 데이터 큐브에 대한 집계를 수행하는 연산이다. 그리고, 드릴-다운 연산은 롤-업의 반대 연산으로서, 선택한 차원의 정보를 더욱 상세히 제공하는 연산을 의미한다. 또한, 슬라이스 연산은 주어진 데이터 큐브에서 한 차원을 선택하여 부분적인 큐브를 만드는 연산을 의미하며, 다이스 연산은 2 개 또는 그 이상의 차원을 선택하여 부분적인 큐브를 만들어주는 연산이다.

[0067] 예를 들어, 도 11에서 나타낸 다차원 데이터 큐브 중 시간 차원의 개념 계층은 분기(quarters)이고, 위치 차원의 개념 계층은 부서(department)이고, 프로토콜 차원의 개념 계층은 UDP 계층이고, 애플리케이션 차원의 개념 계층은 각각의 애플리케이션의 종류(each application)인 4 차원 데이터 큐브에 대해서 사용자가 롤-업 연산을 선택한 경우에 대해서 설명하도록 한다.

[0068] 구체적으로, 4차원 데이터 큐브에 대해서 시간 차원, 위치 차원 및 애플리케이션 차원에 대한 개념 계층의 롤-업 연산이 수행되면, 시간 차원의 개념 계층이 년(year)이고, 위치 차원의 개념 계층이 건물(building)이고, 애

플리케이션 차원의 개념 계층이 세부 타입(specific types)으로 롤-업된 데이터 큐브가 생성된다.

[0069] 이와 같은 방식으로, OLAP 연산 모듈(243)은 다차원 데이터 큐브들에 대해 다양한 OLAP 연산을 수행하여 차원이 추가 및 제외되거나 차원 별 개념 계층이 상위 계층 또는 하위 계층으로 변경된 데이터 큐브들을 생성할 수 있다.

[0070] 그리고, OLAP 연산 모듈(243)은 사용자가 선택한 차원 및 개념 계층에 따라 생성된 데이터 큐브로부터 해당 차원 및 개념 계층에 대한 측도 값(즉, 대역폭 및 플로우 개수 중 적어도 하나의 값)을 요약 및/또는 집계 처리하는 다차원적 분석을 수행한다. 또한, OLAP 연산 모듈(243)은 다차원적 분석 결과를 분석 결과 제공부(260)로 전송한다.

[0071] 그러면, 분석 결과 제공부(260)는 다차원적 분석 결과를 그래프 및 표 등으로 출력하는 웹 페이지를 생성하여 사용자 인터페이스(300)으로 제공할 수 있다.

[0072] 이때, 분석 결과 제공부(260)는 다차원 분석부(240)로부터 네트워크 트래픽에 대한 다차원적 분석 결과를 수신하여, 사용자가 선택한 차원 및 추상화 정도(즉, 개념 계층)에 따른 측도 값의 요약 및 집계 결과를 그래프 및 표로 출력할 수 있다.

[0073] 예를 들어, 분석 결과 제공부(260)는 도 12에서와 같이 사용자가 선택한 적어도 하나의 차원 및 차원 별 계념 계층에 따라 각 차원의 측도 값 즉, 대역폭을 표로 출력하는 분석 결과 웹 페이지를 생성할 수 있다.

[0074] 도 12는 본 발명의 실시예에 따른 네트워크 트래픽에 대한 다차원 분석 결과 제공 화면을 나타내는 도면이다.

[0075] 구체적으로, 도 12에서는, 사용자가 시간 차원의 개념 계층으로써 월(month, 참고로 도 12에서는 “7월”을 선택한 것을 나타냄)을 선택하고, 애플리케이션 차원의 개념 계층으로써 각각의 애플리케이션 종류를 선택하고, 위치 차원의 개념 계층으로써 부서(department) 중 컴퓨터 정보 학과(컴정과)를 선택하였을 때의 분석 결과 웹 페이지를 나타내었다. 이처럼, 분석 결과 제공부(260)는 다차원 분석부(240)로부터 수신한 다차원적 분석 결과에 따라, 해당 네트워크 트래픽에 대한 애플리케이션 종류 별 사용 대역폭을 그래프로써 출력하는 분석 결과 웹 페이지를 생성한다.

[0076] 또한, 분석 결과 제공부(260)는 도 12에서와 같이 사용자가 네트워크 트래픽에 대한 각 차원 별 추상화 정도를 선택할 수 있는 선택 버튼(roll up, drill down)과 사용자가 차원 별 추상화 정도를 확인 및 선택할 수 있는 입력 아이콘들을 포함하는 제어 메뉴판(control panel)이 구성된 웹 페이지를 생성할 수 있다.

[0077] 이하, 도 3을 참조하여 본 발명의 실시예에 따른 데이터마이닝 분석부(250)의 네트워크 트래픽에 대한 데이터마이닝 분석 방법에 대해서 상세히 설명하도록 한다.

[0078] 먼저, 도 3에서와 같이 본 발명의 실시예에 따른 데이터마이닝 분석부(250)는 데이터마이닝 분석 제어 모듈(251), 성장도 분석 모듈(252) 및 상관도 분석 모듈(253)을 포함한다.

[0079] 데이터마이닝 분석 제어 모듈(251)은 다차원 분석부(240)로부터 생성된 다차원 데이터 모델을 이용하여, 사용자가 선택한 차원 및 개념 계층에 따른 군집 성장도 및 군집 상관도를 분석하도록 제어한다.

[0080] 이때, 데이터마이닝 분석 제어 모듈(241)은 사용자가 선택한 차원 및 개념 계층의 정보를 획득하고, 획득한 차원 및 개념 계층의 정보 및 다차원 데이터 모델을 성장도 분석 모듈(252) 및 상관도 분석 모듈(253)으로 전달한다.

[0081] 성장도 분석 모듈(252)은 다차원 데이터 모델을 이용하여 상기 선택된 차원 및 개념 계층에 따른 측도 값에 대해 클러스터링을 수행하여 두 개의 군집을 생성한다. 그리고, 성장도 분석 모듈(252)은 기설정된 적어도 두 개의 기간에 해당하는 군집을 생성한다. 참고로, 상기 기설정된 적어도 두 개의 기간은 다차원 데이터 모델에 대한 성장도를 비교하는 기준이 되는 기간으로써, 사용자가 선택한 기간의 정보를 입력 받거나 또는 기설정된 기준 기간이 저장되어 있을 수 있다.

[0082] 성장도 분석 모듈(252)은 상기 생성된 군집들의 측도 값에 대해 성장 지수(GI) 및 가중 성장 지수(WGI)를 산출한다.

[0083] 구체적으로, 성장도 분석 모듈(252)은 하기 수학식 1을 이용하여 군집 성장 지수(GI)를 산출한다.

[0084] [수학식 1]

[0085] $$GI=\frac{\text{제2기간의차원별측도값}-\text{제1기간의차원별측도값}}{\text{제2기간의차원별측도값}+\text{제1기간의차원별측도값}}$$

[0086] 이때, 수학식 1에서 제1기간은 제2기간보다 이전 기간을 의미하며, 측도 값은 기선택된 차원 및 개념 계층에 대응하는 대역폭 또는 플로우 개수를 의미한다.

[0087] 예를 들어, 본 발명의 실시예에서 사용자가 데이터마이닝 분석을 위해 선택한 차원이 애플리케이션 차원이고, 선택한 측도 값의 종류가 대역폭일 때 성장 지수(GI)는 하기 수학식 2와 같이 나타낼 수 있다.

[0088] [수학식 2]

[0089] $$AGI(ApplicationGI)=\frac{\text{제2기간의애플리케이션에대한대역폭}-\text{제1기간의애플리케이션에대한대역폭}}{\text{제2기간의애플리케이션에대한대역폭}+\text{제1기간의애플리케이션에대한대역폭}}$$

[0090] 이때, AGI는 제1기간과 제2기간에 대해서 애플리케이션 차원의 상대적인 성장도를 나타내는 지표이다.

[0091] 그런 다음, 성장도 분석 모듈(252)은 산출된 성장 지수(GI)를 이용하여 해당 차원의 절대적인 증가 정도인 가중 성장 지수(WGI)를 산출한다.

[0092] 구체적으로, 성장도 분석 모듈(252)은 하기 수학식 3을 이용하여 군집 가중 성장 지수(WGI)를 산출한다.

[0093] [수학식 3]

[0094] $$WGI=|\text{제2기간의차원별측도값}-\text{제1기간의차원별측도값}|*GI$$

[0095] 예를 들어, 본 발명의 실시예에서 사용자가 애플리케이션 차원을 선택하고, 측도 값의 종류로써 대역폭으로 선택하였을 때, 가중 성장 지수(WGI)는 하기 수학식 4와 같이 나타낼 수 있다.

[0096] [수학식 4]

[0097] $$WAGI=|\text{제2기간의애플리케이션에대한대역폭}-\text{제1기간의애플리케이션에대한대역폭}|*AGI$$

[0098] 한편, 성장도 분석 모듈(252)은 사용자가 선택한 차원 및 개념 계층에 따른 가중 성장 지수(WGI)를 산출한 결과를 분석 결과 제공부(260)로 전송한다.

[0099] 그러면, 분석 결과 제공부(260)는 산출된 가중 성장 지수(WGI)를 이용한 데이터마이닝 분석 결과로써 해당 차원 및 개념 계층에 따른 기간별 군집 성장도 그래프를 출력한다.

[0100] 예를 들어, 본 발명의 실시예에 따른 분석 결과 제공부(260)는 도 13에서와 같이 사용자가 선택한 애플리케이션 차원에 대한 기간별 측도 값 즉, 대역폭 및 플로우 개수를 그래프로 생성하여 출력할 수 있다.

[0101] 도 13은 본 발명의 실시예에 따른 네트워크 트래픽에 대한 성장도 분석 결과를 나타내는 도면이다.

[0102] 이때, 도 13의 (a) 및 (b)에서는, 사용자가 선택한 애플리케이션 차원의 개념 계층이 P2P 및 non-P2P 계층인 것을 나타내었다. 이때, 도 13의 (a)의 해당 차원 및 개념 계층에 따른 측도 값의 종류는 대역폭이고, 도 13의 (b)의 해당 차원 및 개념 계층에 따른 측도 값의 종류는 플로우 개수인 것을 나타내었다.

[0103] 그리고, 도 13의 (c) 및 (d)에서는, 사용자가 선택한 애플리케이션 차원의 개념 계층이 세부 타입인 것을 나타내었다. 이때, 도 13의 (c)의 차원 및 개념 계층에 따른 측도 값의 종류는 대역폭이고, 도 13 (d)의 해당 차원 및 개념 계층에 따른 측도 값의 종류는 플로우 개수인 것을 나타내었다.

[0104] 도 13의 (a) 내지 (d)를 통해서, 기간별 애플리케이션 차원의 성장도는 제2기간에 P2P 애플리케이션의 사용이 크게 증가된 것을 알 수 있으며, P2P 애플리케이션 중 세부 타입이 file-sharing인 애플리케이션의 사용이 크게 증가된 것을 알 수 있다.

[0105] 다시 도 3으로 돌아가서, 상관도 분석 모듈(253)은 다차원 데이터 모델을 이용하여 상기 선택된 차원 및 개념 계층에 따른 측도 값에 대해 클러스터링을 수행하여 군집을 생성한다. 그리고, 상관도 분석 모듈(253)은 생성된 군집들의 측도 값에 대해 유사도 함수를 산출한다.

[0106] 구체적으로, 상관도 분석 모듈(253)은 하기 수학식 5를 이용하여 군집 상관도를 분석하기 위한 유사도 행렬을

생성한다.

[0107] [수학식 5]

$$\cos(d_x, d_y) = \frac{\sum_{i=1}^{n}(w(a_i, d_x) \times w(a_i, d_y))}{\sqrt{\sum_{i=1}^{n} w(a_i, d_x)^2 \times \sum_{i=1}^{n} w(a_i, d_y)^2}}$$

[0108]

[0109] 이때, 수학식 5의 코사인 유사도 함수에서 ai는 유사도 산출 시 비교 기준이 되는 차원의 개념 계층의 측도 값이고, dx 및 dy는 유사도 산출의 비교 대상이 되는 차원의 개념 계층을 의미한다.

[0110] 그런 다음, 상관도 분석 모듈(253)은 생성된 유사도 행렬들을 분석 결과 제공부(260)로 전달하여, 사용자가 선택한 차원 및 개념 계층에 따른 상관도 분석 결과가 출력되도록 한다.

[0111] 그러면, 분석 결과 제공부(260)는 산출된 유사도 행렬을 이용한 데이터마이닝 분석 결과로써, 해당 차원 및 개념 계층에 따른 군집별 상관도 분석 결과를 출력한다.

[0112] 예를 들어, 본 발명의 실시예에 따른 분석 결과 제공부(260)는 도 14에서와 같이 사용자가 상관도 비교 결과를 나타내는 기준이 되는 제 1 차원으로써 애플리케이션 차원을 선택하고, 유사도를 비교할 제 2 차원으로써 위치 차원을 선택한 경우 패스파인더 네트워크(Pathfinder Network) 모델을 생성하여 출력할 수 있다. 참고로, 패스파인더 네트워크 모델에서는 군집 개체의 수가 많이 포함된 군집일수록 원의 지름이 크게 표현되며, 각 원의 링크들의 굵기에 따라 가중치가 표현된다.

[0113] 도 14는 본 발명의 실시예에 따른 네트워크 트래픽에 대한 상관도 분석 결과를 나타내는 도면이다.

[0114] 이때, 도 14에서는 본 발명의 실시예에 따른 상관도 분석 모듈(253)이 대학 내 네트워크 트래픽 데이터에 대한 상관도 분석을 수행할 경우, 상기 제 2 차원인 위치 차원의 개념 계층인 부서로써 학과를 기준으로 유사도를 비교할 수 있다.

[0115] 즉, 도 14에서는 상관도 분석 모듈(253)이 애플리케이션 차원의 측도 값에 대한 학과 별 유사도 행렬을 생성하고, 상기 생성된 유사도 행렬을 이용하여 분석한 학과별 상관도를 나타내었다. 이때, 도 14에서는 상기 수학식 5에서 ai로써 애플리케이션 차원에 대한 측도 값(즉, 대역폭)을 대입하고, dx 및 dy로써 각각 유사도를 비교할 학과를 대입하여 유사도 행렬을 산출하였다.

[0116] 구체적으로, 도 14에서 나타낸 패스파인더 네트워크의 각 노드의 크기는 해당 학과에서 사용한 전체 애플리케이션의 대역폭 양이고, 각 노드를 연결하는 링크의 굵기는 각 노드 간의 상관 정도 즉, 학과 별로 사용한 애플리케이션의 유사도를 나타낸다. 이때, 도 14에서는 사용한 대역폭 양과 사용한 애플리케이션의 종류의 유사도에 기초하여, 경제학과, 전자정보공학부, 시스템실 그리고 정보통계학과가 매우 밀접한 관계인 것 즉, 상관도가 높은 것을 알 수 있다. 또한, 컴퓨터 공학과의 경우 사용한 전체 애플리케이션의 대역폭의 양이 매우 많고, 다른 학과와 상관도가 크지 않은 것으로 즉, 기설정된 기준 값 이하인 것으로 분석되어 별도의 큰 군집으로 구분된 것을 알 수 있다.

[0117] 이하, 도 15를 참조하여 본 발명의 실시예에 따른 네트워크 트래픽 분석 방법에 대해서 설명하도록 한다.

[0118] 도 15는 본 발명의 실시예에 따른 네트워크 트래픽 분석 방법을 설명하기 위한 순서도이다.

[0119] 먼저, 네트워크 디바이스들로부터 네트워크 트래픽에 대한 정보 데이터들을 수집한다(S151).

[0120] 이때, 네트워크 트래픽에 대한 정보 데이터는, 해당 네트워크 디바이스들을 경유하는 연속되는 패킷들을 일정 기간 별로 동일 헤더 정보를 갖는 패킷들로 정렬한 플로우 형식의 네트워크 트래픽 데이터와 네트워크 트래픽 데이터 별 대역폭 정보 데이터를 포함한다.

[0121] 그런 다음, 수집되는 대용량의 네트워크 트래픽 데이터들을 목적 또는 시간 등을 기준으로 정렬하여 네트워크 트래픽에 대한 이력데이터를 적재한다(S152).

[0122] 이때, 본 발명의 실시예에서는 상기 이력데이터들이 데이터웨어하우스에 적재될 수 있다.

[0123] 그런 다음, 상기 이력데이터에 대한 다차원 데이터 큐브 모델을 생성한다(S153).

[0124] 구체적으로, 상기 이력데이터들을 기설정된 차원 및 차원 별 개념 계층을 기준으로 재정렬하여 다차원 데이터

큐브를 생성한다. 이때, 본 발명의 실시예에서는 시간, 위치, 애플리케이션의 종류 및 프로토콜의 종류를 기준으로 복수의 차원이 분류되고, 각 차원 별로 최하위 개념부터 최상위 개념까지 복수의 개념 계층으로 구분되는 다차원 데이터 큐브가 생성된다.

[0125] 그런 다음, 생성된 다차원 데이터 큐브 모델에 대해서 선택된 차원 및 계층에 따른 OLAP 연산을 수행하여 다차원적 분석을 수행한다(S154).

[0126] 이때, OLAP 연산은 다차원 데이터 큐브 모델에서 네트워크 트래픽 데이터를 정렬하는 기준이 되는 차원을 추가 또는 제거하는 연산 및 각 차원 별로 개념 계층의 추상화 정도를 하향 또는 상향 시키는 연산 등을 포함할 수 있다.

[0127] 또한, 본 발명의 실시예에서는 상기 다차원 데이터 큐브 모델에 대해 사용자가 선택한 차원 및 개념 계층에 기초하여 OLAP 연산을 수행한 결과로써 1 차내지 4차의 데이터 큐브를 생성하여 다차원적 분석을 수행할 수 있다. 즉, 사용자가 선택한 차원의 개수에 대응하는 데이터 큐브가 생성되며, 생성된 데이터 큐브를 이용하여 해당 차원 및 개념 계층에 따른 측도 값 즉, 대역폭 또는 플로우 개수를 집계 및/또는 요약하여 사용자가 인지할 수 있는 정보로써 출력할 수 있다.

[0128] 다음으로, 상기 다차원 데이터 큐브 모델을 이용하여, 사용자가 선택한 차원 및 개념 계층에 따른 군집 성장도를 분석한다(S155).

[0129] 구체적으로, 본 발명의 실시예에 따른 네트워크 트래픽 분석 방법에서는, 상기 다차원 데이터 큐브 모델에 대해 상기 선택한 차원 및 개념 계층에 따른 측도 값을 클러스터링하여 군집을 생성하고, 기설정된 적어도 두 개의 기간을 기준으로 상기 생성된 군집의 성장 지수 및 가중 성장 지수를 산출하여 기간별 군집 성장도를 분석한다.

[0130] 그런 다음, 상기 다차원 데이터 큐브 모델을 이용하여, 사용자가 선택한 차원 및 개념 계층에 따른 군집 상관도를 분석한다(S156).

[0131] 구체적으로, 본 발명의 실시예에 따른 네트워크 트래픽 분석 방법에서는, 상기 다차원 데이터 큐브 모델에 대해 상기 선택한 차원 및 개념 계층에 따른 측도 값을 클러스터링하여 군집을 생성한다. 그리고, 사용자가 선택한 유사도 산출 시 비교 기준이 되는 차원의 개념 계층의 측도 값과 유사도 산출의 비교 대상이 되는 차원의 개념 계층을 이용하여 각 군집 간 유사도 행렬을 산출한다. 그런 후, 상기 유사도 행렬을 이용하여 상기 다차원 데이터 큐브 모델에 대한 패스파인더 네트워크 모델을 생성하여 군집 상관도를 분석한다.

[0132] 다음으로, 네트워크 트래픽 데이터에 대한 다차원적 분석 결과, 성장도 분석 결과 및 상관도 분석 결과 중 적어도 하나의 결과를 사용자(예를 들어, 네트워크 관리자)가 인지할 수 있도록 생성하여 제공한다(S157).

[0133] 이때, 본 발명의 실시예에 따른 네트워크 트래픽 분석 방법에서는, 다차원적 분석 결과, 성장도 분석 결과 및 상관도 분석 결과를 사용자가 효율적으로 인지할 수 있는 그래프 및 표 등의 문자 정보로 출력하고, 다차원적 분석 및 데이터마이닝 분석의 목적인 차원 및 개념 계층을 사용자가 편리하게 선택할 수 있는 입력 아이콘 또는 버튼 형식의 입력 수단을 출력하는 웹 페이지를 생성하여 제공할 수 있다.

[0134] 본 발명의 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 판독가능 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.

[0135] 본 발명의 장치 및 방법은 특정 실시예와 관련하여 설명되었지만, 그것들의 구성 요소 또는 동작의 일부 또는 전부는 범용 하드웨어 아키텍처를 갖는 컴퓨터 시스템을 사용하여 구현될 수 있다.

[0136] 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, 마찬가

지로 분산된 것으로 설명되어 있는 구성 요소들도 결합된 형태로 실시될 수 있다.

[0137] 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

[0138]

110: 네트워크 디바이스 210: 데이터 수집부

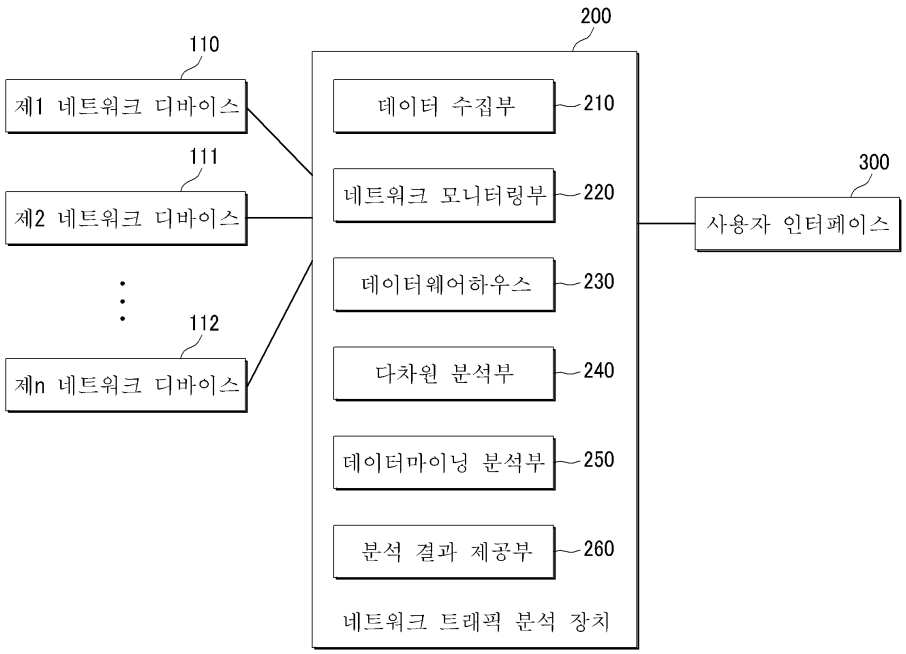
220: 네트워크 모니터링부 230: 데이터웨어하우스

240: 다차원 분석부 250: 데이터마이닝 분석부

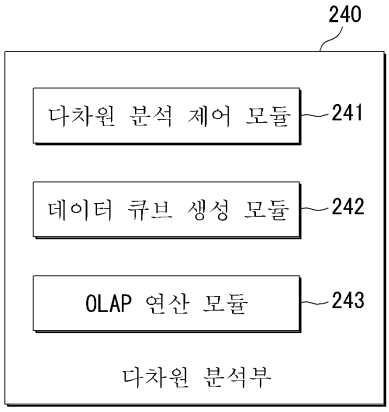
260: 분석 결과 제공부 300: 사용자 인터페이스

도면

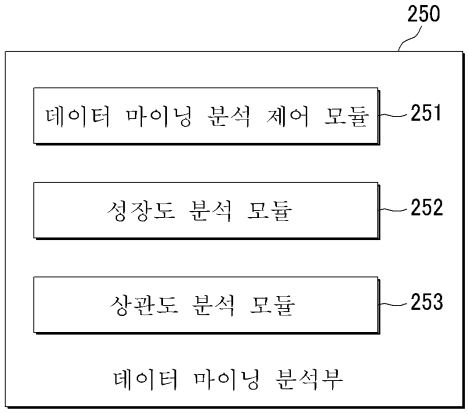
도면1



도면2



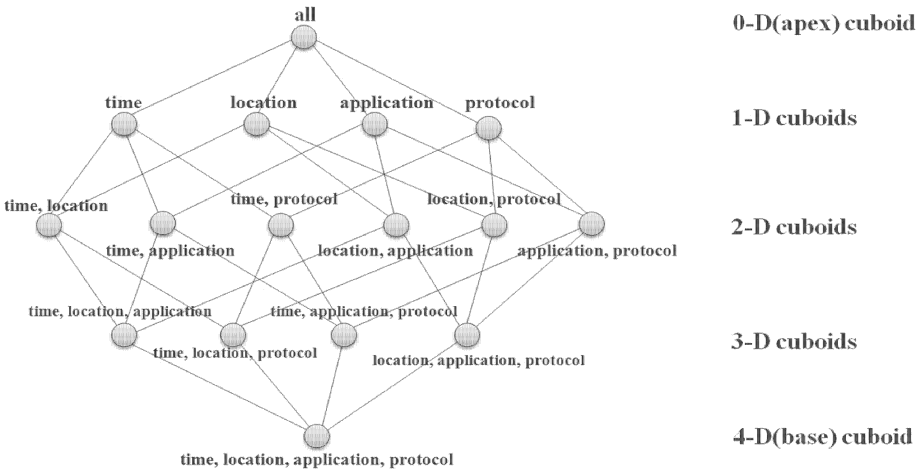
도면3



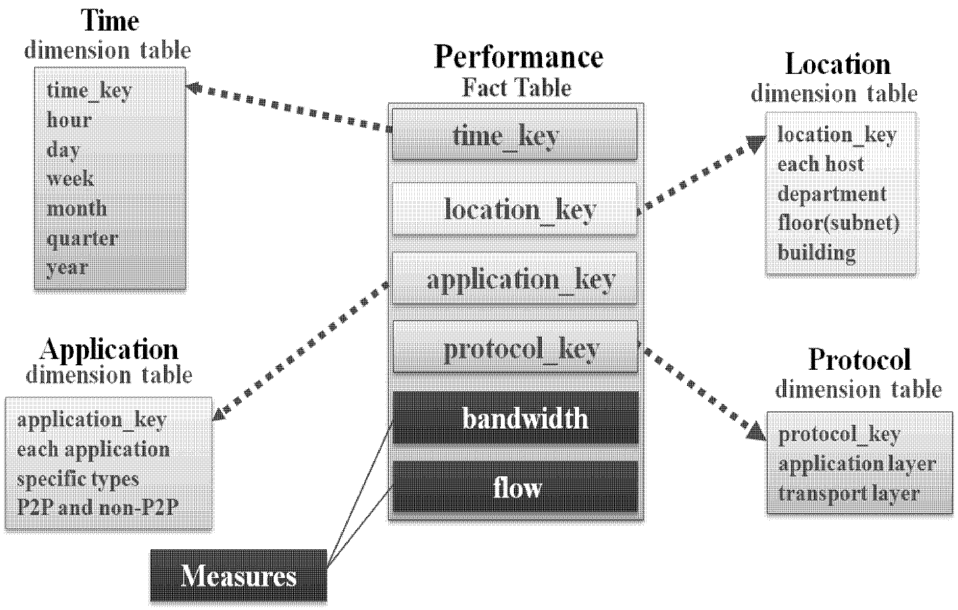
도면4

Time(month)	Application(specific types)					
	file-sharing	messenger	streaming(tv)	traditional client/server	streaming	game
3 월	46282.99	3074.95	1712.65	432560.53	344.21	282.06
4 월	280105.14	3814.57	1088.80	557369.23	3110.88	496.90
5 월	118734.83	2101.70	680.40	168254.24	422.36	718.31
6 월	102802.26	2290.39	4254.11	199962.13	144.10	761.00
7 월	767787.83	4456.49	412.98	102991.15	244.91	37.30

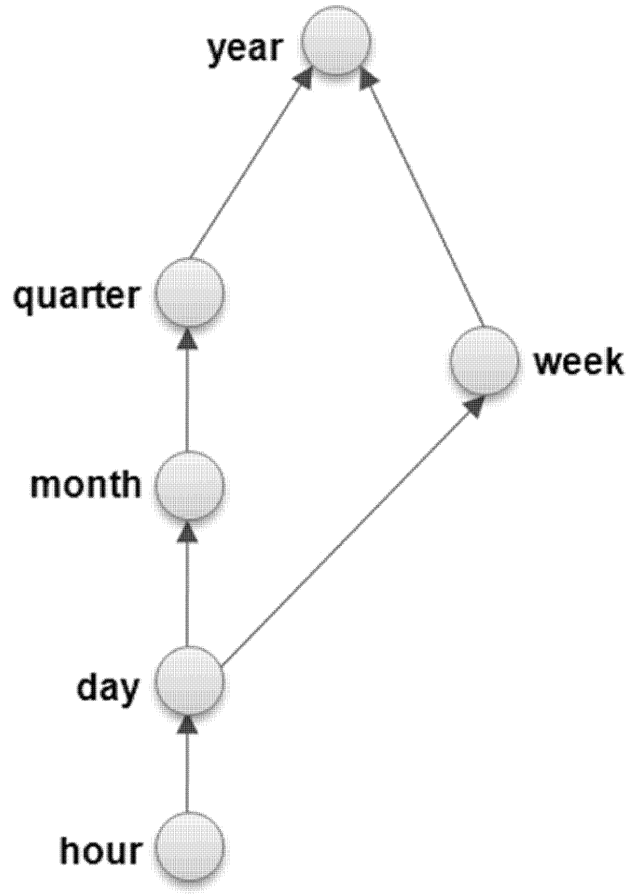
도면5


all
time
location
application
protocol
time, location
time, protocol
location, protocol
time, application
location, application
application, protocol
time, location, application
time, application, protocol
time, location, protocol
location, application, protocol
time, location, application, protocol
0-D(apex) cuboid
1-D cuboids
2-D cuboids
3-D cuboids
4-D(base) cuboid

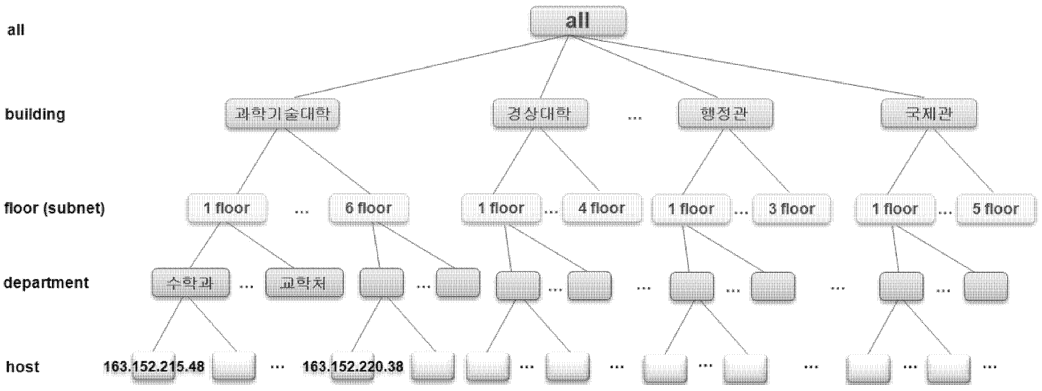
도면6


Time
dimension table
time_key
hour
day
week
month
quarter
year
Performance
Fact Table
time_key
location_key
application_key
protocol_key
bandwidth
flow
Location
dimension table
location_key
each host
department
floor(subnet)
building
Application
dimension table
application_key
each application
specific types
P2P and non-P2P
Protocol
dimension table
protocol_key
application layer
transport layer
Measures

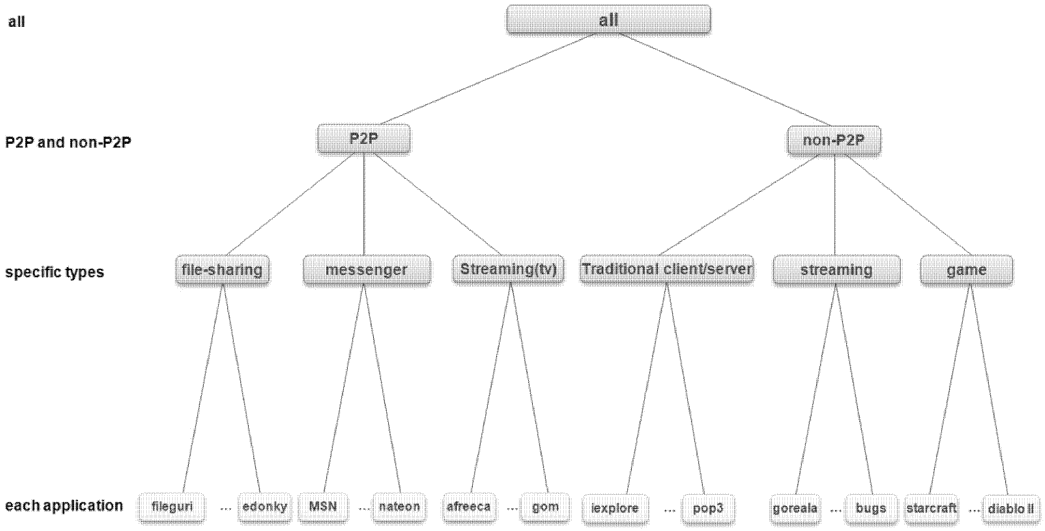
도면7


year
quarter
week
month
day
hour

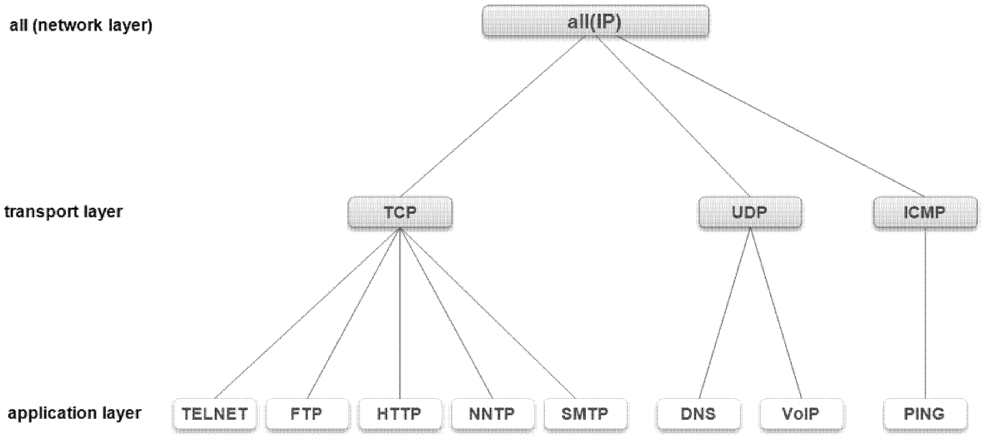
도면8


all
building
floor (subnet)
department
host
all
과학기술대학
경상대학
행정관
국제관
1 floor
6 floor
4 floor
3 floor
5 floor
수학과
교학처
163.152.215.48
163.152.220.38

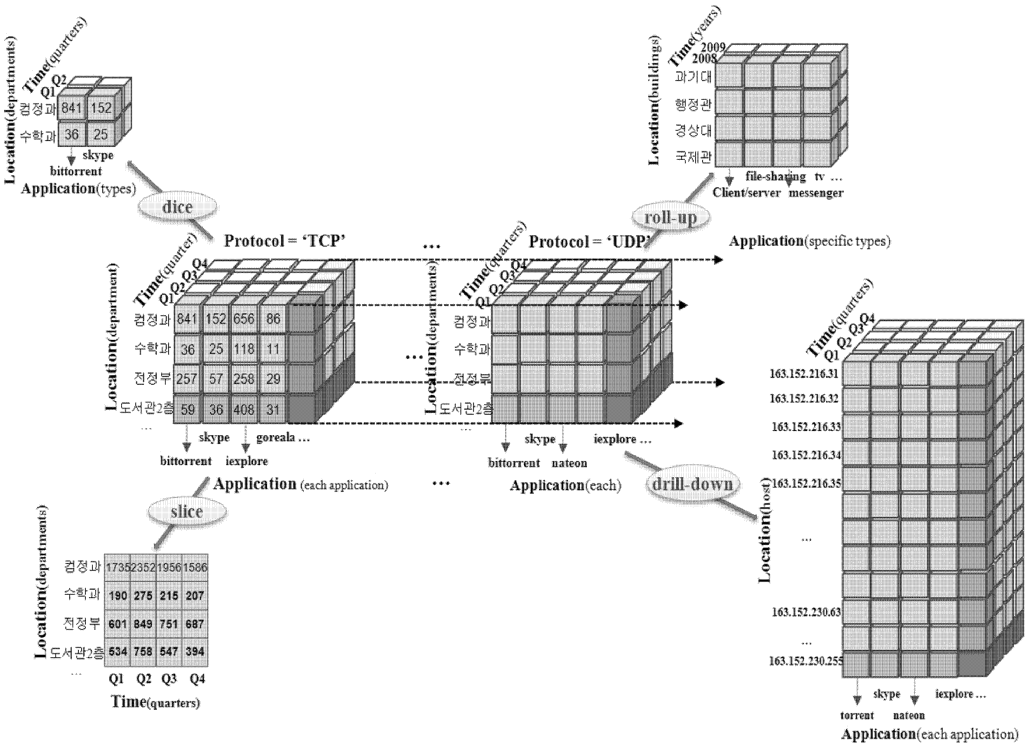
도면9


all
all
P2P and non-P2P
P2P
non-P2P
specific types
file-sharing
messenger
Streaming(tv)
Traditional client/server
streaming
game
each application
fileguri
edonky
MSN
nateon
afreeca
gom
iexplore
pop3
goreala
bugs
starcraft
diablo II

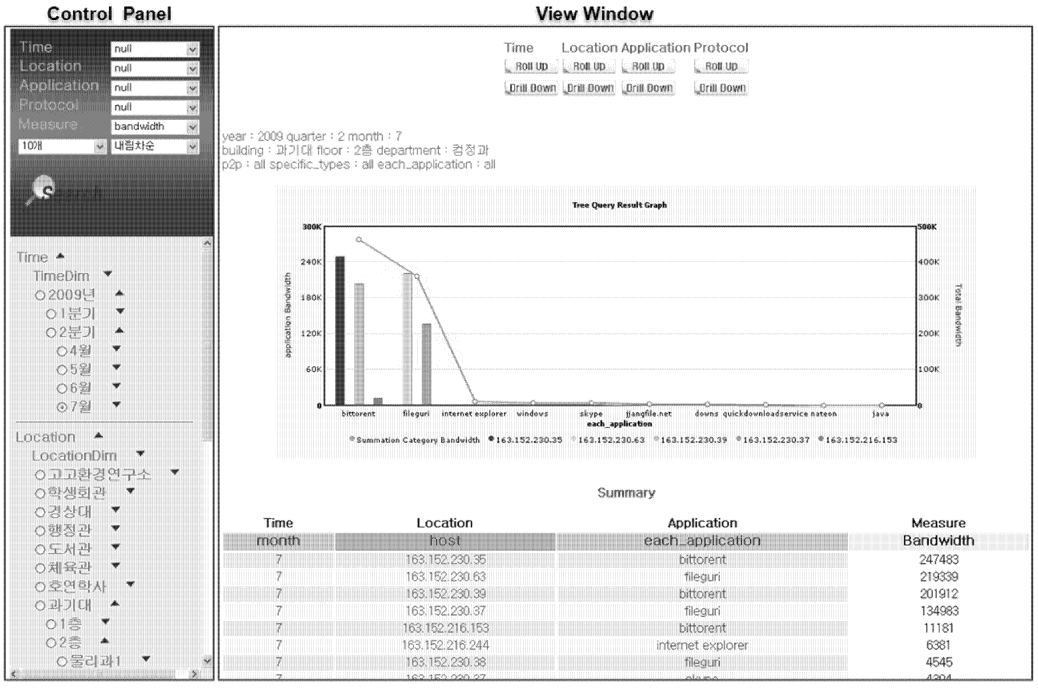
도면10


all (network layer)
all(IP)
transport layer
TCP
UDP
ICMP
application layer
TELNET
FTP
HTTP
NNTP
SMTP
DNS
VoIP
PING

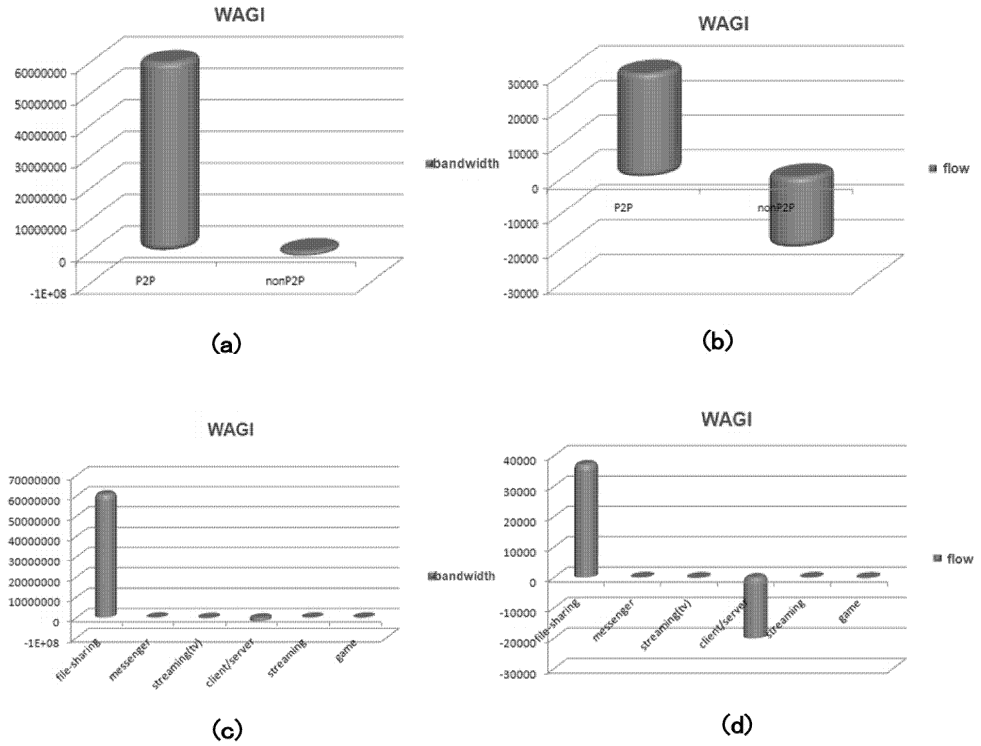
도면11


Location(departments)
Time(quarters)
Q1
Q2
컴정과 841 152
수학과 36 25
skype
bittorrent
Application(types)
dice
Protocol = 'TCP'
Protocol = 'UDP'
Application(specific types)
roll-up
Location(buildings)
Time(years)
2009
2008
과기대
행정관
경상대
국제관
file-sharing tv …
Client/server messenger
Time(quarter)
Location(department)
Q1 Q2 Q3 Q4
컴정과 841 152 656 86
수학과 36 25 118 11
전정부 257 57 258 29
도서관2층 59 36 408 31
skype goreala …
bittorrent iexplore
skype iexplore …
bittorrent nateon
Application (each application)
Application(each)
drill-down
slice
Location(departments)
컴정과 1735 2352 1956 1586
수학과 190 275 215 207
전정부 601 849 751 687
도서관2층 534 758 547 394
Q1 Q2 Q3 Q4
Time(quarters)
Location(host)
163.152.216.31
163.152.216.32
163.152.216.33
163.152.216.34
163.152.216.35
163.152.230.63
163.152.230.255
skype iexplore …
torrent nateon
Application(each application)

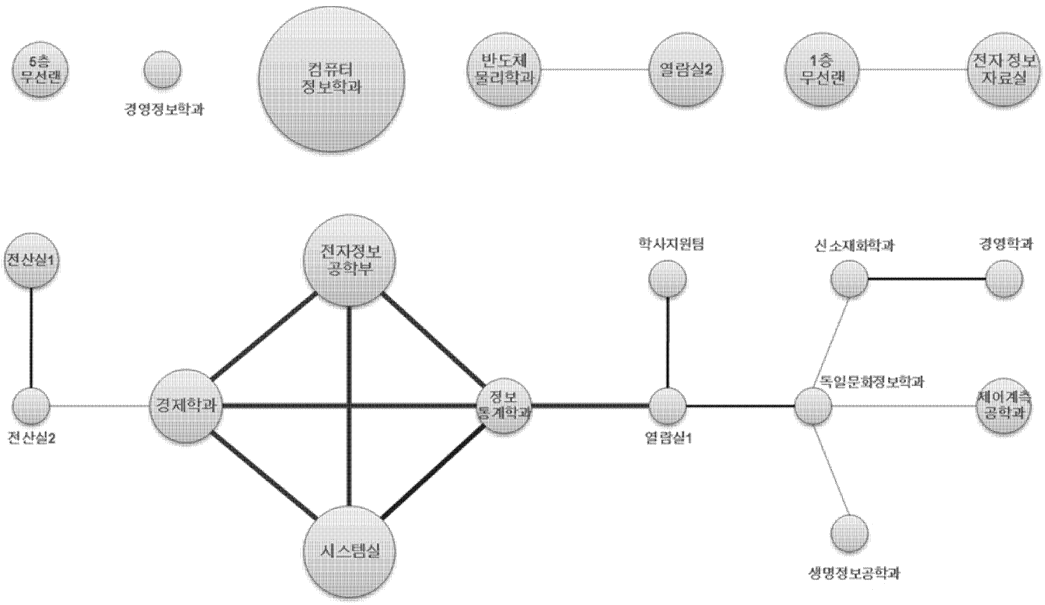
도면12


Control Panel
View Window
Time null
Location null
Application null
Protocol null
Measure bandwidth
Time Location Application Protocol
Roll Up
Drill Down
year : 2009 quarter : 2 month : 7
building : 과기대 floor : 2층 department : 컴정과
p2p : all specific_types : all each_application : all
Tree Query Result Graph
application Bandwidth
Total Bandwidth
bittorrent
fileguri
internet explorer
windows
skype
jiangfile.net
downs
quickdownloadservice
nateon
java
each_application
Summation Category Bandwidth
163.152.230.35
163.152.230.63
163.152.230.39
163.152.230.37
163.152.216.153
Summary
Time
Location
Application
Measure
month
host
each_application
Bandwidth
7 163.152.230.35 bittorent 247483
7 163.152.230.63 fileguri 219339
7 163.152.230.39 bittorent 201912
7 163.152.230.37 fileguri 134983
7 163.152.216.153 bittorent 11181
7 163.152.216.244 internet explorer 6381
7 163.152.230.38 fileguri 4545

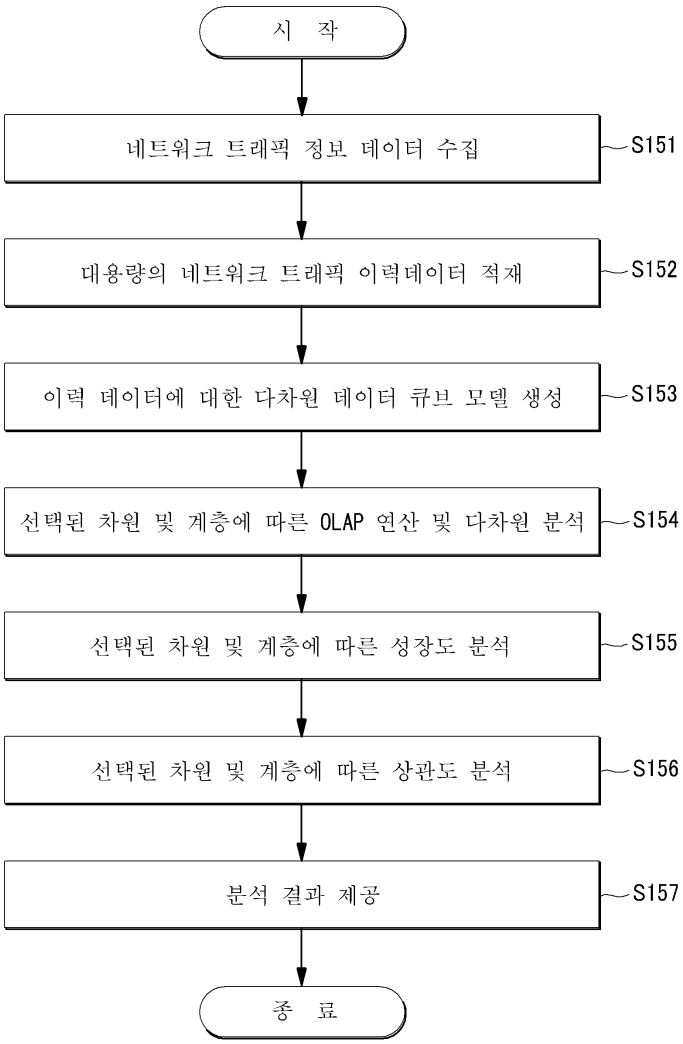
도면13


WAGI
60000000
50000000
40000000
30000000
20000000
10000000
0
-1E+08
P2P
nonP2P
(a)
WAGI
bandwidth
30000
20000
10000
0
-10000
-20000
-30000
P2P
nonP2P
flow
(b)
WAGI
70000000
60000000
50000000
40000000
30000000
20000000
10000000
0
-1E+08
file-sharing
messenger
streaming(tv)
client/server
streaming
game
(c)
WAGI
bandwidth
40000
30000
20000
10000
0
-10000
-20000
-30000
file-sharing
messenger
streaming(tv)
client/server
streaming
game
flow
(d)

도면14


5층 무선랜
경영정보학과
컴퓨터 정보학과
반도체 물리학과
열람실2
1층 무선랜
전자정보 자료실
전산실1
전산실2
경제학과
전자정보 공학부
정보 통계학과
시스템실
학사지원팀
열람실1
신소재화학과
경영학과
독일문화정보학과
제어계측 공학과
생명정보공학과

도면15


시 작
네트워크 트래픽 정보 데이터 수집
S151
대용량의 네트워크 트래픽 이력데이터 적재
S152
이력 데이터에 대한 다차원 데이터 큐브 모델 생성
S153
선택된 차원 및 계층에 따른 OLAP 연산 및 다차원 분석
S154
선택된 차원 및 계층에 따른 성장도 분석
S155
선택된 차원 및 계층에 따른 상관도 분석
S156
분석 결과 제공
S157
종 료