



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2013년01월24일
 (11) 등록번호 10-1226161
 (24) 등록일자 2013년01월18일

(51) 국제특허분류(Int. Cl.)

G06F 17/30 (2006.01)

(21) 출원번호 10-2012-0083208

(22) 출원일자 2012년07월30일

심사청구일자 2012년07월30일

(56) 선행기술조사문헌

JP06231181 A*

*는 심사관에 의하여 인용된 문헌

(73) 특허권자

한국과학기술정보연구원

대전광역시 유성구 대학로 245 (어은동)

(72) 발명자

김대홍

대전광역시 서구 청사로 253 꿈나무아파트
 203-302

최성필

대전광역시 유성구 가정로 65 대림두레아파트 10
 1동 1004호

(뒷면에 계속)

(74) 대리인

특허법인화우

전체 청구항 수 : 총 7 항

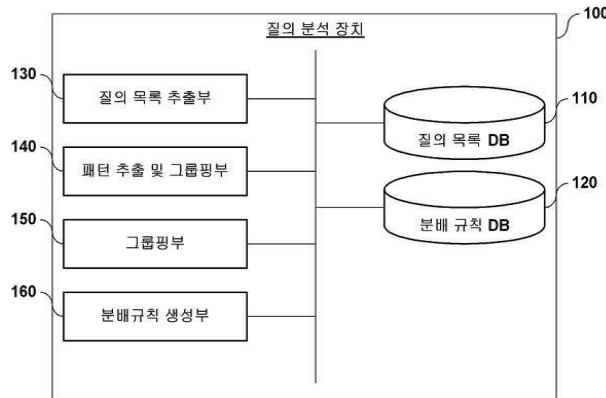
심사관 : 이명진

(54) 발명의 명칭 **분배규칙 생성을 위한 방법 및 그를 위한 질의 분석 장치**

(57) 요약

본 발명은 분배규칙 생성을 위한 방법 및 그를 위한 질의 분석 장치에 관한 것으로, 질의 분석 장치가 분배규칙을 생성하는 방법에 있어서, (a) 각 서비스의 질의 목록에 포함된 질의 셋들을 분석하여 패턴을 추출하고, 그 패턴을 근거로 질의 단위의 패턴 셋 구성 파일을 생성하는 단계, (b) 상기 질의 단위의 패턴 셋 구성 파일을 해당 서비스의 분배 규칙으로 생성하는 단계를 포함한다.

대표도 - 도1



(72) 발명자

정창후

대전 유성구 구암동 628-11 남산아트빌라 가동 40
2호

엄정호

전라북도 전주시 완산구 당산로 11 쌍용아파트 60
2동 1210호

정성재

인천광역시 남동구 소래역로 93 냇마을신영지웰아
파트 905동 2003호

정한민

대전광역시 유성구 어은동 99 한빛아파트 111동
1404호

특허청구의 범위

청구항 1

질의 분석 장치가 분배규칙을 생성하는 방법에 있어서,

(a) 각 서비스의 질의 목록에 포함된 질의 셋들을 분석하여 패턴을 추출하고, 질의 단위로 추출된 패턴들을 타입정보를 근거로 그룹핑하여 질의 단위의 타입패턴그룹을 생성한 후, 상기 타입패턴그룹들을 서비스 단위의 타입패턴그룹으로 생성하고, 상기 서비스 단위의 타입패턴그룹을 이용하여 질의 단위의 패턴 셋 구성 파일을 생성하는 단계; 및

(b) 상기 질의 단위의 패턴 셋 구성 파일을 해당 서비스의 분배 규칙으로 생성하는 단계;

를 포함하는 분배규칙 생성을 위한 방법.

청구항 2

제1항에 있어서,

상기 (a)단계는,

상기 질의 목록에 포함된 질의 셋들로부터 패턴을 질의 단위로 추출하는 단계;

상기 질의 단위로 추출된 패턴들을 타입정보를 근거로 질의 단위로 그룹핑하여 질의 단위의 타입패턴그룹을 생성하는 단계;

상기 질의 단위의 타입패턴그룹들에 대해 타입 정보를 근거로 서비스 단위의 타입패턴그룹을 생성하는 단계; 및

상기 서비스 단위의 타입패턴그룹을 이용하여 질의 단위의 패턴 셋 구성 파일을 생성하는 단계를 포함하는 것을 특징으로 하는 분배규칙 생성을 위한 방법.

청구항 3

각 서비스의 질의 목록에 포함된 질의 셋들로부터 패턴을 질의 단위로 추출하고, 그 패턴들을 타입정보를 근거로 질의 단위로 그룹핑하여 질의 단위의 타입패턴그룹을 생성하는 패턴 추출 및 그룹핑부;

상기 질의 단위의 타입패턴그룹들에 대해 타입 정보를 근거로 서비스 단위의 타입패턴그룹을 생성하는 그룹핑부; 및

상기 서비스 단위의 타입패턴그룹을 이용하여 질의 단위의 패턴 셋 구성 파일을 생성하고, 상기 질의 단위의 패턴 셋 구성 파일을 해당 서비스의 분배 규칙으로 생성하는 분배 규칙 생성부;

를 포함하는 분배규칙 생성을 위한 질의 분석 장치.

청구항 4

제3항에 있어서,

각 서비스별로 미리 정해진 질의 목록 및 질의 셋들이 저장된 질의 목록 데이터베이스를 더 포함하는 것을 특징으로 하는 분배규칙 생성을 위한 질의 분석 장치.

청구항 5

제3항에 있어서,

상기 분배 규칙 생성부에서 생성된 서비스별 분배 규칙이 저장된 분배규칙 데이터베이스를 더 포함하는 것을 특징으로 하는 분배규칙 생성을 위한 질의 분석 장치.

청구항 6

제3항에 있어서,

상기 그룹핑부는 각 서비스별로 생성된 질의 단위의 타입패턴그룹들에서 동일 타입의 타입패턴그룹은 병합하여 새로운 하나의 타입패턴 그룹으로 생성하는 것을 특징으로 하는 질의 분석 장치.

청구항 7

제3항에 있어서,

상기 분배 규칙 생성부는 각 질의별로 생성된 질의단위 타입패턴그룹을 상기 서비스 단위의 타입패턴그룹으로 대체하여 질의 단위의 패턴 셋 구성파일을 생성하는 것을 특징으로 하는 질의 분석 장치.

명세서

기술 분야

[0001] 본 발명은 분배규칙 생성을 위한 방법 및 그를 위한 질의 분석 장치에 관한 것으로, 보다 상세하게는 각 서비스의 질의 목록에 포함된 질의 셋들을 분석하여 패턴을 추출하고, 그 패턴을 근거로 질의 단위의 패턴 셋 구성 파일을 생성하여 해당 서비스의 분배 규칙으로 결정하는 분배규칙 생성을 위한 방법 및 그를 위한 질의 분석 장치에 관한 것이다.

배경 기술

[0002] 인터넷이 발전함에 따라 네티즌에 의해 하루에도 수없이 많은 데이터가 생성되어 유통되고 있으며, 최근 많은 기업, 특히 검색 엔진 회사 및 웹 포탈들 간에 이와 같은 엄청난 양의 데이터를 가능한 많이 수집 및 축적하여 수집된 데이터 중에서 가능한 빨리 의미있는 정보를 추출하는 것이 기업의 경쟁력이 되고 있다.

[0003] 때문에 많은 기업은 저비용으로 대규모 클러스터를 구축하여 대용량 분산 관리 및 작업 분산 병렬처리 기술에 대한 많은 연구를 하고 있다.

[0004] 즉, 기존 싱글 머신 시스템에서 처리하기 어려운 대용량 데이터의 가치가 부각되고, 이를 처리하기 위한 대안으로 분산 병렬 기반의 시스템이 다양한 분야에 도입/활용되고 있다.

[0005] 그러나, 다수의 노드에 데이터를 저장하고 처리하는 분산 병렬 시스템에서 하나의 태스크를 처리하는 과정에서 각 노드 간 소요되는 네트워크 IO와 다수의 Join 연산으로 인한 부하로 인해 전체적인 시스템의 처리 속도 저하가 불가피하여 빠른 속도로 대용량 데이터를 처리하는데 본질적인 문제가 있었다.

발명의 내용

해결하려는 과제

[0006] 본 발명은 상기한 문제점을 해결하기 위하여 안출된 것으로, 분산 병렬 시스템의 각 노드 간 네트워크 IO 타입 및 Join 연산을 최소화하여 전체적인 시스템의 응답속도를 단축시킬 수 있는 분배규칙을 생성하는 분배규칙 생성을 위한 방법 및 그를 위한 질의 분석 장치를 제공하는데 그 목적이 있다.

[0007] 본 발명의 또 다른 목적은 질의 목록을 분석하여 연관 데이터들을 독립노드에 저장하기 위한 분배 규칙을 얻을

수 있는 분배규칙 생성을 위한 방법 및 그를 위한 질의 분석 장치를 제공하는데 있다.

과제의 해결 수단

- [0008] 상기 목적들을 달성하기 위하여 본 발명의 일 측면에 따르면, 질의 분석 장치가 분배규칙을 생성하는 방법에 있어서, (a) 각 서비스의 질의 목록에 포함된 질의 셋들을 분석하여 패턴을 추출하고, 그 패턴을 근거로 질의 단위의 패턴 셋 구성 파일을 생성하는 단계, (b) 상기 질의 단위의 패턴 셋 구성 파일을 해당 서비스의 분배 규칙으로 생성하는 단계를 포함하는 분배규칙 생성을 위한 방법이 제공된다.
- [0009] 상기 (a)단계는, 상기 질의 목록에 포함된 질의 셋들로부터 패턴을 질의 단위로 추출하는 단계, 상기 질의 단위로 추출된 패턴들을 타입정보를 근거로 질의 단위로 그룹핑하여 질의 단위의 패턴타입그룹을 생성하는 단계, 상기 질의 단위의 패턴타입그룹들에 대해 타입 정보를 근거로 서비스 단위의 타입패턴그룹을 생성하는 단계 및 상기 서비스 단위의 타입패턴그룹을 이용하여 질의 단위의 패턴 셋 구성 파일을 생성하는 단계를 포함할 수 있다.
- [0010] 본 발명의 다른 측면에 따르면, 각 서비스의 질의 목록에 포함된 질의 셋들로부터 패턴을 질의 단위로 추출하고, 그 패턴들을 타입정보를 근거로 질의 단위로 그룹핑하여 질의 단위의 패턴타입그룹을 생성하는 패턴 추출 및 그룹핑부, 상기 질의 단위의 패턴타입그룹들에 대해 타입 정보를 근거로 서비스 단위의 패턴타입그룹을 생성하는 그룹핑부, 상기 서비스 단위의 패턴타입그룹을 이용하여 질의 단위의 패턴 셋 구성 파일을 생성하고, 상기 질의 단위의 패턴 셋 구성 파일을 해당 서비스의 분배 규칙으로 생성하는 분배 규칙 생성부를 포함하는 분배규칙 생성을 위한 질의 분석 장치가 제공된다.
- [0011] 상기 질의 분석 장치는 각 서비스별로 미리 정해진 질의 목록 및 질의 셋들이 저장된 질의 목록 데이터베이스를 더 포함할 수 있다.
- [0012] 또한, 상기 질의 분석 장치는 상기 분배 규칙 생성부에서 생성된 서비스별 분배 규칙이 저장된 분배규칙 데이터베이스를 더 포함할 수 있다.
- [0013] 상기 그룹핑부는 각 서비스별로 생성된 질의 단위의 타입패턴그룹들에서 동일 타입의 타입패턴그룹은 병합하여 새로운 하나의 타입패턴 그룹으로 생성할 수 있다.
- [0014] 상기 분배 규칙 생성부는 각 질의별로 생성된 질의단위 타입패턴그룹을 상기 서비스 단위의 타입패턴그룹으로 대체하여 질의 단위의 패턴 셋 구성파일을 생성할 수 있다.

발명의 효과

- [0015] 본 발명에 따르면, 분산 병렬 시스템의 각 노드 간 네트워크 IO 타임 및 Join 연산을 최소화하여 전체적인 시스템의 응답속도를 단축시킬 수 있는 분배규칙을 생성할 수 있다.
- [0016] 또한, 질의 목록을 분석하여 연관 데이터들을 독립노드에 저장하기 위한 분배 규칙을 얻을 수 있다.

도면의 간단한 설명

- [0017] 도 1은 본 발명에 따른 분배규칙 생성을 위한 질의 분석 장치의 구성을 개략적으로 나타낸 블럭도.
- 도 2는 본 발명에 따른 질의 분석 장치가 서비스별 질의를 분석하여 분배 규칙을 생성하는 방법을 나타낸 흐름도.
- 도 3은 본 발명에 따른 질의 분석 장치가 분배 규칙을 생성하는 방법을 나타낸 흐름도.
- 도 4는 본 발명에 따른 질의 분석 장치가 분배 규칙을 생성하는 방법을 설명하기 위한 예시도.

발명을 실시하기 위한 구체적인 내용

- [0018] 본 발명의 전술한 목적과 기술적 구성 및 그에 따른 작용 효과에 관한 자세한 사항은 본 발명의 명세서에 첨부된 도면에 의거한 이하 상세한 설명에 의해 보다 명확하게 이해될 것이다.

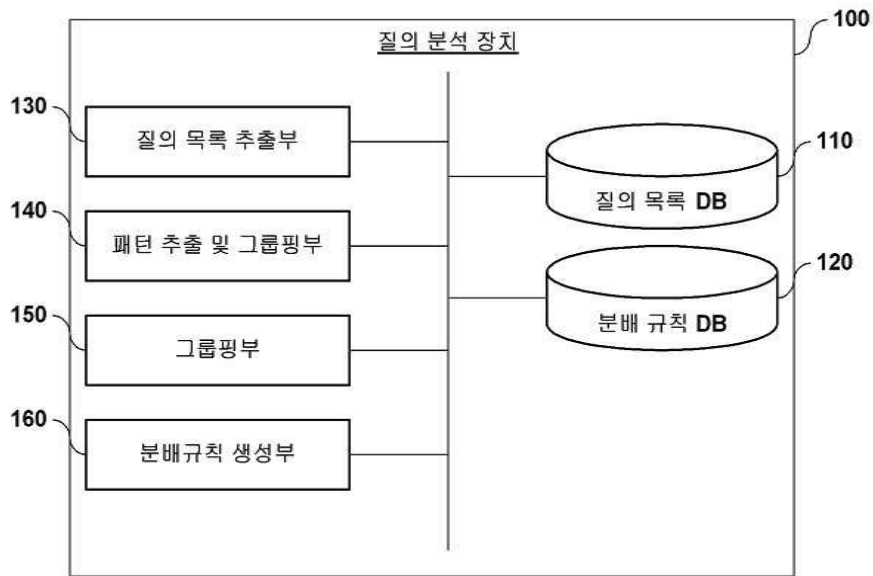
- [0019] 이하에서는 분배규칙을 생성하기 위한 장치를 질의 분석 장치로 표현하였으나, 질의 분석 서버, 단말 등으로 다양하게 표현될 수 있다.
- [0020] 도 1은 본 발명에 따른 분배규칙 생성을 위한 질의 분석 장치의 구성을 개략적으로 나타낸 블럭도이다.
- [0021] 도 1을 참조하면, 분배규칙 생성을 위한 질의 분석 장치(100)는 질의 목록 데이터베이스(110), 분배 규칙 데이터베이스(120), 질의 목록 추출부(130), 패턴 추출 및 그룹핑부(140), 그룹핑부(150), 분배 규칙 생성부(160)를 포함한다.
- [0022] 상기 질의 목록 데이터베이스(110)에는 각 서비스별로 미리 정해진 질의 목록 및 질의 셋들이 저장되어 있다. 여기서, 서비스는 논문 분석 서비스, 기술 분석 서비스 등과 같은 서비스를 의미하고, 각 서비스에 사용되는 질의는 그 유형이 제한되어 있다.
- [0023] 예를 들어, 논문 분석 서비스의 경우, 질의는 "2010년 생물학 분야의 논문 목록을 모두 얻기 위한 질의, 한국의 A라는 기관의 생물학 분야 논문 목록과 그의 저자 목록을 얻기 위한 질의, 한국의 연구투자 예산을 얻는 질의" 등과 같이 그 유형이 정해져 있다.
- [0024] 상기 분배규칙 데이터베이스(120)에는 분배 규칙 생성부(160)에서 생성된 서비스별 분배 규칙이 저장되어 있다. 상기 분배 규칙은 분산 병렬 기반 저장소를 이용하는 경우, 연관 데이터들을 독립 노드에 저장하기 위한 규칙일 수 있다.
- [0025] 상기 질의 목록 추출부(130)는 분배규칙 생성 명령이 입력된 경우, 상기 질의 목록 데이터베이스(110)로부터 해당 서비스의 질의 목록 및 질의 셋을 추출하여 패턴 추출 및 그룹핑부(140)로 전달한다.
- [0026] 상기 패턴 추출 및 그룹핑부(140)는 각 서비스의 질의 목록에 포함된 질의 셋들로부터 패턴을 질의 단위로 추출하고, 그 패턴들을 타입정보를 근거로 질의 단위로 그룹핑하여 질의 단위의 타입패턴(typepattern)그룹을 생성한다. 여기서 패턴은 각 질의 셋에 포함된 미리 정해진 형태로서, 예를 들면, '?논문 a ex:Paper', '?논문 ex:Subject ?주제', '?주제 a ex:Sub', '?주제 rdf:lable "생물학 "', '?논문 ex:Year "2010".'일 수 있다. 상기 타입은 패턴에서 첫번째 단어를 의미하는 것으로서, '?논문 ex:Subject의 경우, 논문, '?주제 a ex:Sub'의 경우 주제가 타입일 수 있다.
- [0027] 따라서, 상기 패턴 추출 및 그룹핑부(140)는 '?논문 a ex:Paper', '?논문 ex:Subject ?주제','?논문 ex:Year "2010"'을 논문 기준으로 하나의 타입패턴그룹을 생성하고, '?주제 a ex:Sub', '?주제 rdf:lable "생물학 "'을 주제 기준으로 하나의 타입패턴 그룹을 생성한다.
- [0028] 상기 그룹핑부(150)는 상기 패턴 추출 및 그룹핑부(140)에서 생성된 질의 단위의 타입패턴그룹들에 대해 타입 정보를 근거로 서비스 단위로 타입패턴그룹을 생성한다. 즉, 상기 그룹핑부(150)는 각 서비스별로 생성된 질의 단위의 타입패턴그룹들에서 동일 타입의 타입패턴그룹은 병합하여 새로운 하나의 타입패턴 그룹으로 생성한다.
- [0029] 예를 들어, 3개의 질의 셋으로 구성된 논문 분석 서비스의 경우, 제1질의 셋에는 타입패턴그룹1, 2가 포함, 제2질의 셋에는 타입패턴그룹3 내지 5가 포함, 제3질의 셋에는 타입패턴그룹6 및 7이 포함되어 있다고 가정하여 설명하기로 한다.
- [0030] 이 경우, 제1질의 셋의 타입패턴그룹1과 제2질의 셋의 타입패턴그룹3의 타입이 동일한 경우, 상기 그룹핑부(150)는 타입패턴그룹1과 타입패턴그룹3을 하나로 병합하여 타입패턴그룹8을 생성한다. 이때, 타입패턴그룹1과 타입패턴그룹3은 동일하므로 하나는 삭제한다.
- [0031] 상기 분배 규칙 생성부(160)는 상기 그룹핑부(150)에서 생성된 서비스 단위의 타입패턴그룹을 이용하여 질의 단위의 패턴 셋 구성 파일을 생성하고, 상기 질의 단위의 패턴 셋 구성 파일을 해당 서비스의 분배 규칙으로 상기 분배규칙 데이터베이스(120)에 저장한다. 즉, 상기 분배 규칙 생성부(160)는 각 질의별로 생성된 질의단위 타입패턴그룹을 상기 서비스 단위의 타입패턴그룹으로 표시하여 질의 단위의 패턴 셋 구성파일을 생성한다.
- [0032] 상기와 같이 구성된 질의 분석 장치(100)는 분산 병렬 기반 저장소를 이용하는 특정 애플리케이션이나 서비스 등에 사용되는 질의 목록 파일을 입력받고, 그 입력된 질의에서 나타난 패턴을 분석하여 데이터 저장기에서 사용될 수 있는 분배 규칙을 생성한다. 즉, 질의 목록을 분석하여 연관 데이터들을 독립노드에 저장하기 위한 분배 규칙을 얻는다.
- [0033] 분산 병렬 기반 저장소를 이용하는 특정 어플리케이션이나 서비스 등에 사용되는 질의는 일반적으로 그 유형(패턴)이 제한되어 있으므로, 이러한 유형을 분석하여 하나의 질의 패턴에서 사용되는 데이터는 하나의 데이터 노

드에 저장하기 위해 분배 규칙을 생성할 수 있다.

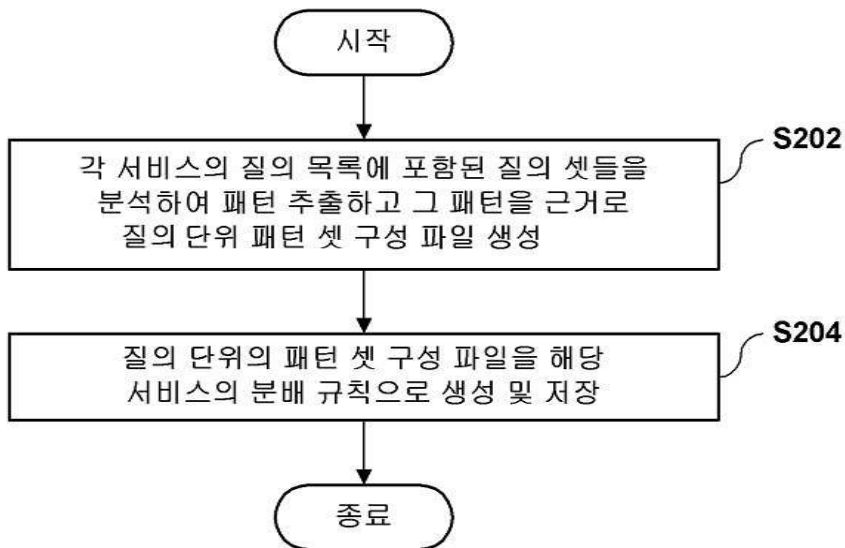
- [0034] 도 2는 본 발명에 따른 질의 분석 장치가 서비스별 질의를 분석하여 분배 규칙을 생성하는 방법을 나타낸 흐름도이다.
- [0035] 도 2를 참조하면, 질의 분석 장치는 각 서비스의 질의 목록에 포함된 질의 셋들을 분석하여 패턴을 추출하고, 그 패턴을 근거로 질의 단위의 패턴 셋 구성 파일을 생성한다(S202). 각 서비스에 사용되는 질의는 그 패턴이 제한되어 있으므로, 상기 질의 분석 장치는 각 서비스별로 입력되는 질의 셋을 분석하여 패턴을 추출하고, 그 추출된 패턴을 근거로 패턴 셋 구성파일을 생성할 수 있다.
- [0036] 상기 질의 분석 장치가 질의단위 패턴 셋 구성 파일을 생성하는 방법에 대한 상세한 설명은 도 3을 참조하기로 한다.
- [0037] 상기 S202가 수행되면, 상기 질의 분석 장치는 상기 질의 단위의 패턴 셋 구성 파일을 해당 서비스의 분배 규칙으로 생성하여 저장한다(S204).
- [0038] 도 3은 본 발명에 따른 질의 분석 장치가 분배 규칙을 생성하는 방법을 나타낸 흐름도, 도 4는 본 발명에 따른 질의 분석 장치가 분배 규칙을 생성하는 방법을 설명하기 위한 예시도이다.
- [0039] 도 3을 참조하면, 질의 분석 장치는 특정 서비스의 질의 목록이 입력되면, 상기 질의 목록에 포함된 질의 셋들로부터 패턴을 질의 단위로 추출한다(S302).
- [0040] 그런 다음 상기 질의 분석 장치는 상기 질의 단위로 추출된 패턴들을 타입정보를 근거로 질의 단위로 그룹핑하여 질의 단위의 타입패턴그룹들을 생성한다(S304).
- [0041] 상기 S304의 수행 후, 상기 질의 분석 장치는 상기 질의 단위의 타입패턴그룹들에 대해 타입 정보를 근거로 서비스 단위의 타입패턴그룹을 생성한다(S306).
- [0042] 즉, 상기 질의 분석 장치는 각 서비스에 포함된 질의들의 질의 단위 타입패턴그룹들중에서 동일한 타입을 갖는 타입패턴그룹을 병합하여 하나의 타입패턴 그룹을 생성한다.
- [0043] 상기 S306의 수행 후, 상기 질의 분석 장치는 상기 서비스 단위의 타입패턴그룹을 이용하여 질의 단위의 패턴 셋 구성 파일을 생성한다(S308). 즉, 상기 질의 분석 장치는 각 질의별로 생성된 질의단위 타입패턴그룹을 상기 서비스 단위의 타입패턴그룹으로 표시하여 질의 단위의 패턴 셋 구성파일을 생성한다.
- [0044] 상기 생성된 질의 단위의 패턴 셋 구성 파일이 해당 서비스의 분배 규칙일 수 있다.
- [0045] 상기 질의 분석 장치가 질의 목록을 분석하여 분배 규칙을 생성하는 방법에 대해 도 4를 참조하여 설명하기로 한다.
- [0046] 도 4를 참조하면, 논문 분석 서비스에 사용되는 "2010년에 생물학 분야의 논문 목록을 모두 얻기 위한 제1질의", "한국의 KISTI라는 기관의 생물학 분야 논문 목록과 그의 저자목록을 얻기 위한 제2질의", "한국의 연구투자 예산을 얻는 질의가 논문 분석 서비스에 대해 정해진 제3질의"의 3개의 질의가 입력되는 경우이다.
- [0047] 상기 질의 중에서 제1 질의의 경우,
- [0048] "prefix ex:<http://example.com>
- [0049] Select ?논문 where {
- [0050] ?논문 a ex:Paper.
- [0051] ?논문 ex:Subject ?주제.
- [0052] ?주제 a ex:Sub.
- [0053] ?주제 rdf:lable "생물학 "
- [0054] ?논문 ex:Year "2010".

도면

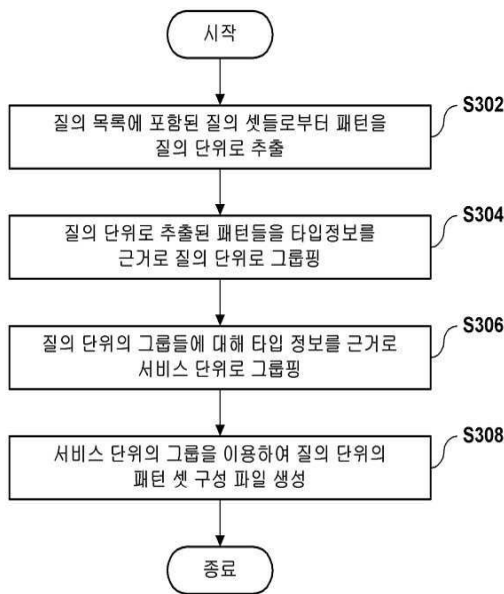
도면1



도면2



도면3



도면4

Sparql Query	1. 2010년에 발행한 분야의 논문 목록을 모두 알기 위한 질의	2. 한국의 KISTI라는 기관의 생물학 분야 논문 목록과 그의 저자 목록 알기 위한 질의	3. 한국의 연구투자 예산을 알기 위한 질의
1. 질의 질의 셋	<pre> prefix ex-<http://example.com> Select ?논문 where { ?논문 a ex:Paper ?논문 ex:Subject ?주제 ?주제 a ex:Sub ?주제 rdfs:label "생물학" ?논문 ex:Year "2010" } </pre>	<pre> prefix ex-<http://example.com> Select ?저자 ?논문 where { ?논문 a ex:Paper ?논문 ex:Subject ?주제 ?주제 a ex:Sub ?주제 rdfs:label "생물학" ?논문 ex:writen ?저자 ?저자 a ex:Author ?저자 ex:Affiliation ?기관 ?기관 a ex:Organization ?기관 rdfs:label "KISTI" ?기관 ex:Nationality ?국가 ?국가 a ex:Nation ?국가 rdfs:label "KOREA" } </pre>	<pre> prefix ex-<http://example.com> Select ?Nation ?Budget where { ?Nation a ex:Nation ?Nation rdfs:label "KOREA" ?Nation ex:R_Budget ?예산 ?예산 a ex:Budget ?예산 ex:Amount ?Budget } </pre>
2. 패턴 추출 및 그룹화	패턴 개수 : 5개 - ID#1 타입정보 : 2개 - TypePattern#1 ?논문 a ex:Paper ?논문 ex:Subject ?주제 ?논문 ex:Year [문자] TypePattern#2 ?주제 a ex:Sub ?주제 rdfs:label [문자]	패턴 개수 : 12개 - ID#2 타입정보 : 5개 TypePattern#3 ?논문 a ex:Paper ?논문 ex:Subject ?주제 ?논문 ex:writen ?저자 TypePattern#4 ?주제 a ex:Sub ?주제 rdfs:label "생물학" TypePattern#5 ?저자 a ex:Author ?저자 ex:Affiliation ?기관 TypePattern#6 ?기관 a ex:Organization ?기관 rdfs:label "KISTI" ?기관 ex:Nationality ?국가 TypePattern#7 ?국가 a ex:Nation ?국가 rdfs:label "KOREA"	패턴개수: 5개 - ID#3 타입정보 : 2개 TypePattern#8 ?Nation a ex:Nation ?Nation rdfs:label "KOREA" ?Nation ex:R_Budget ?예산 TypePattern#9 ?예산 a ex:Budget ?예산 ex:Amount ?Budget
3. 패턴 병합	<ul style="list-style-type: none"> 패턴의 서브셋 포함 여부 판단 동일 type 의 typepattern 은 병합 (TypePattern#12) #3은 하나로 묶을 => TypePattern#10 생성(#1#3은 삭제) TypePattern#2와 #4를 합치므로 하나로 삭제 이 과정을 거치면 ID#1 은 TypePattern#10 과 TypePattern #2를 갖음) 		
4. 패턴 생성	<ul style="list-style-type: none"> 질의 단위의 패턴 셋 구성 파일로 저장 ex) ID #1 => Typepattern#1,2 ID #2 => ID#1 + ID#3 + Typepattern#5,6) ID #3 => Typepattern#3,9		