

	(19) 대한민국특허청(KR) (12) 공개특허공보(A)	(11) 공개번호 10-2009-0129124 (43) 공개일자 2009년12월16일
(51) Int. Cl. <i>G06F 17/21</i> (2006.01)	(21) 출원번호 10-2008-0055228 (22) 출원일자 2008년06월12일 심사청구일자 2008년06월12일	(71) 출원인 고려대학교 산학협력단 서울 성북구 안암동5가1 고려대학교 내 (72) 발명자 이상근 서울 동대문구 제기2동 153 안암골 벽산아파트 101동 203호 고혜경 서울 노원구 월계2동 롯데캐슬루나 121동 1004호 (74) 대리인 현종철

전체 청구항 수 : 총 14 항

(54) 브랜칭 노드 스트림을 이용한 XML 문서의 트위그 패턴질의 처리 방법, 그 장치 및 컴퓨터 판독 가능 저장 매체

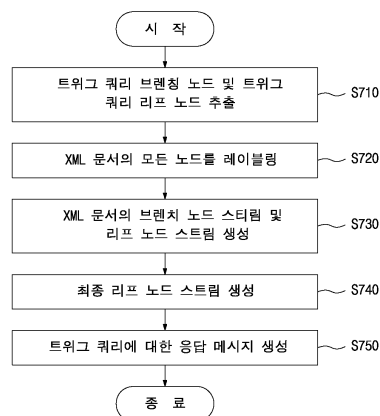
(57) 요약

브랜칭 노드 스트림을 이용한 XML 문서의 트위그 패턴 질의 처리 방법, 그 장치 및 컴퓨터 판독 가능 저장 매체가 개시된다.

본 발명에 따른 브랜칭 노드 스트림을 이용한 XML 문서의 트위그(Twig) 패턴 질의 처리 방법은, 확장된 듀웨이 레이블링(Extended Dewey Labeling)된 XML 문서의 트위그 쿼리의 질의 처리 방법에 있어서, 클라이언트로부터 상기 XML 문서에 대하여 특정의 트위그 쿼리(Twig query)를 수신하면, 상기 특정의 트위그 쿼리의 트위그 쿼리 브랜칭 노드(Branching node) 및 트위그 쿼리 리프 노드(leaf node)를 추출하는 단계; 상기 XML 문서를 파싱한 후, 상기 XML 문서에 상기 확장된 듀웨이 레이블링을 수행하여 상기 XML 문서의 모든 노드를 레이블링하는 단계; 상기 트위그 쿼리의 트위그 쿼리 브랜칭 노드 및 트위그 쿼리 리프 노드에 대응되는 상기 XML 문서의 브랜칭 노드 스트림 및 리프 노드 스트림을 생성하는 단계; 상기 리프 노드 스트림을 구성하는 리프 노드의 레이블링 값 중에서 상기 브랜칭 노드 스트림을 구성하는 브랜칭 노드의 레이블링 값을 포함하지 않는 리프 노드의 레이블링 값을 제거함으로써 최종 리프 노드 스트림을 생성하는 단계; 및 상기 브랜칭 노드 스트림 및 최종 리프 노드 스트림을 이용하여 상기 트위그 쿼리에 대한 응답 메시지를 생성하는 단계를 포함한다.

본 발명에 의하면, 클라이언트의 트위그 쿼리에 대하여 트위그 쿼리의 리프 노드에 매칭되어지는 XML 문서상의 노드의 탐색 시간을 줄임으로써 리프 노드 탐색 시간인 I/O 타임을 줄일 수 있으며, 트위그 쿼리에 대하여 탐색된 노드에 대하여 리프 노드에 레이블링된 값을 원래 노드 이름으로 변환하는 할 때, 트위그(Twig) 쿼리에 대응되는 브랜칭 노드와 중첩되는 레이블링된 값을 원래 노드 이름으로 변환하지 않아 중간 프로세싱 타임이 줄어들고, 전체적으로 쿼리에 대하여 응답하는 전체 프로세싱 타임이 줄어 신속한 질의 응답 처리를 수행할 수 있는 효과가 있다.

대표도 - 도7



(72) 발명자

박상현

서울 성북구 안암동2가 126-2번지 혜성빌딩 305호

류병걸

서울 광진구 노유2동 64-16호

특허청구의 범위

청구항 1

확장된 듀웨이 레이블링(Extended Dewey Labeling)된 XML 문서의 트위그 쿼리의 질의 처리 방법에 있어서,

클라이언트로부터 상기 XML 문서에 대하여 특정의 트위그 쿼리(Twig query)를 수신하면, 상기 특정의 트위그 쿼리의 트위그 쿼리 브랜칭 노드(Branching node) 및 트위그 쿼리 리프 노드(leaf node)를 추출하는 단계;

상기 XML 문서를 파싱한 후, 상기 XML 문서에 상기 확장된 듀웨이 레이블링을 수행하여 상기 XML 문서의 모든 노드를 레이블링하는 단계;

상기 트위그 쿼리의 트위그 쿼리 브랜칭 노드 및 트위그 쿼리 리프 노드에 대응되는 상기 XML 문서의 브랜칭 노드 스트림 및 리프 노드 스트림을 생성하는 단계;

상기 리프 노드 스트림을 구성하는 리프 노드의 레이블링 값 중에서 상기 브랜칭 노드 스트림을 구성하는 브랜칭 노드의 레이블링 값을 포함하지 않는 리프 노드의 레이블링 값을 제거함으로써 최종 리프 노드 스트림을 생성하는 단계; 및

상기 브랜칭 노드 스트림 및 최종 리프 노드 스트림을 이용하여 상기 트위그 쿼리에 대한 응답 메시지를 생성하는 단계를 포함하는 XML 문서의 트위그 쿼리의 질의 처리 방법.

청구항 2

제 1 항에 있어서,

상기 XML 문서의 브랜칭 노드 스트림 및 리프 노드 스트림을 생성하는 단계는

상기 트위그 쿼리 브랜칭 노드에 대응되며 상기 XML 문서의 브랜칭 노드의 레이블링 값으로 이루어진 브랜칭 노드 스트림을 생성하는 단계; 및

상기 트위그 쿼리 리프 노드에 대응되며 상기 XML 문서의 리프 노드의 레이블링 값으로 이루어진 리프 노드 스트림을 생성하는 단계를 포함하는 것을 특징으로 하는 XML 문서의 트위그 쿼리의 질의 처리 방법.

청구항 3

제 1 항에 있어서,

상기 최종 리프 노드 스트림을 생성하는 단계는

상기 브랜칭 노드 스트림과 상기 리프 노드 스트림을 상호 매칭하는 단계;

상기 리프 노드 스트림을 구성하는 각각의 리프 노드의 레이블링 값이 상기 브랜칭 노드 스트림 중 어느 하나의 브랜칭 노드의 레이블링 값을 포함하는지의 여부를 판단하는 단계; 및

상기 리프 노드의 레이블링 값이 어느 하나의 브랜칭 노드의 레이블링 값을 포함하지 않으면 상기 리프 노드의 레이블링 값을 상기 리프 노드 스트림에서 제거하는 단계를 포함하는 것을 특징으로 하는 XML 문서의 트위그 쿼리의 질의 처리 방법.

청구항 4

제 1 항에 있어서,

상기 트위그 쿼리에 대한 응답 메시지를 생성하는 단계는

상기 최종 리프 노드 스트림을 구성하는 복수 개의 리프 노드의 레이블링 값 중 상호 모두 중복되는 중첩 레이블링 엘리먼트를 제거하는 단계; 및

상기 중첩 레이블링 엘리먼트가 제거된 상기 리프 노드의 잔여 레이블링 엘리먼트를 XML 문서 상의 노드명으로 변환하는 단계를 포함하는 것을 특징으로 하는 XML 문서의 트위그 쿼리의 질의 처리 방법.

청구항 5

제 4 항에 있어서,

상기 응답 메시지는

상기 변환된 노드명을 이용하여 상기 클라이언트의 트위그 쿼리에 대응되는 구조로 형성된 XML 응답 메시지인 것을 특징으로 하는 XML 문서의 트위그 쿼리의 질의 처리 방법.

청구항 6

확장된 듀웨이 레이블링(Extended Dewey Labeling)된 XML 문서의 트위그 쿼리의 질의 처리 장치에 있어서,

클라이언트로부터 상기 확장된 듀웨이 레이블링된 XML 문서에 대하여 트위그 쿼리(Twig Query)를 수신하고, 상기 트위그 쿼리에 대한 응답메시지를 상기 클라이언트에 제공하는 송수신부;

상기 트위그 쿼리의 트위그 쿼리 브랜칭 노드(Branch node) 및 트위그 쿼리 리프 노드(leaf node)를 추출하는 트위그 쿼리 노드 추출부;

상기 XML 문서에 상기 확장된 듀웨이 레이블링을 수행하여 상기 XML 문서의 모든 노드를 레이블링하는 레이블링 부;

상기 XML 문서의 전체 노드 스트림과 상기 트위그 쿼리의 레이블링 값을 매칭하고, 상기 전체 노드 스트림에 대하여 상기 트위그 쿼리의 레이블링 값에 대응되는 상기 XML 문서의 브랜칭 노드 스트림 및 리프 노드 스트림을 생성하는 매칭 노드 스트림 생성부;

상기 리프 노드 스트림을 구성하는 리프 노드의 레이블링 값 중에서 상기 브랜칭 노드 스트림을 구성하는 브랜칭 노드의 레이블링 값을 포함하지 않는 리프 노드의 레이블링 값을 제거함으로써 최종 리프 노드 스트림을 생성하는 최종 리프 노드 스트림 생성부; 및

상기 브랜칭 노드 스트림 및 최종 리프 노드 스트림을 이용하여 상기 트위그 쿼리에 대한 응답 메시지를 생성하는 응답 메시지 생성부를 포함하는 XML 문서의 트위그 쿼리의 질의 처리 장치.

청구항 7

제 6 항에 있어서,

상기 XML 노드 스트림 생성부는

상기 트위그 쿼리 브랜칭 노드에 대응되며 상기 XML 문서의 브랜칭 노드의 레이블링 값으로 이루어진 브랜칭 노드 스트림을 생성하고,

상기 트위그 쿼리 리프 노드에 대응되며 상기 XML 문서의 리프 노드의 레이블링 값으로 이루어진 리프 노드 스트림을 생성하는 것을 특징으로 하는 XML 문서의 트위그 쿼리의 질의 처리 장치.

청구항 8

제 6 항에 있어서,

상기 최종 리프 노드 스트림 생성부는

상기 브랜칭 노드 스트림과 상기 리프 노드 스트림을 상호 매칭하는 매칭 모듈; 및

상기 리프 노드 스트림을 구성하는 각각의 리프 노드의 레이블링 값이 상기 브랜칭 노드 스트림 중 어느 하나의 브랜칭 노드의 레이블링 값을 포함하는지의 여부를 판단하고, 상기 리프 노드의 레이블링 값이 어느 하나의 브랜칭 노드의 레이블링 값을 포함하지 않으면 상기 리프 노드의 레이블링 값을 상기 리프 노드 스트림에서 제거하는 스트림 정제 모듈을 포함하는 것을 특징으로 하는 XML 문서의 트위그 쿼리의 질의 처리 장치.

청구항 9

제 6 항에 있어서,

상기 응답 메시지 생성부는

상기 최종 리프 노드 스트림을 구성하는 복수 개의 리프 노드의 레이블링 값 중 상호 모두 중첩되는 중첩 레이블

블링 엘리먼트를 제거하고, 상기 중첩 레이블링 엘리먼트가 제거된 상기 리프 노드의 잔여 레이블링 엘리먼트를 상기 XML 문서의 노드명으로 변환하는 노드명 변환 모듈을 포함하는 것을 특징으로 하는 XML 문서의 트위그 쿼리의 질의 처리 장치.

청구항 10

제 6 항에 있어서,

상기 응답 메시지는

상기 변환된 노드명을 이용하여 상기 클라이언트의 트위그 쿼리에 대응되는 구조로 형성된 XML 응답 메시지만 것을 특징으로 하는 XML 문서의 트위그 쿼리의 질의 처리 장치.

청구항 11

컴퓨터 판독가능 저장매체에 있어서,

상기 컴퓨터 판독가능 저장매체에 상기 확장된 듀웨이 레이블링(Extended Dewey Labeling)된 XML 문서의 트위그 질의 처리를 컴퓨터에서 수행할 수 있도록 하는 프로그램이 기록되어 있고, 상기 프로그램은,

클라이언트로부터 상기 확장된 듀웨이 레이블링된 XML 문서에 대하여 특정의 트위그 쿼리(Twig query)를 수신하면, 상기 트위그 쿼리의 트위그 쿼리 브랜칭 노드(Branching node) 및 트위그 쿼리 리프 노드(leaf node)를 추출하도록 하는 코드;

상기 XML 문서를 파싱한 후, 상기 XML 문서에 상기 확장된 듀웨이 레이블링을 수행하여 상기 XML 문서의 모든 노드를 레이블링하도록 하는 코드;

상기 트위그 쿼리의 레이블링 값에 대응되는 상기 XML 문서의 브랜칭 노드 스트림 및 리프 노드 스트림을 생성하도록 하는 매칭 노드 스트림 생성 코드;

상기 리프 노드 스트림을 구성하는 리프 노드의 레이블링 값 중에서 상기 브랜칭 노드 스트림을 구성하는 브랜칭 노드의 레이블링 값을 포함하지 않는 리프 노드의 레이블링 값을 제거함으로써 최종 리프 노드 스트림을 생성하도록 하는 코드; 및

상기 브랜칭 노드 스트림 및 최종 리프 노드 스트림을 이용하여 상기 트위그 쿼리에 대한 응답 메시지를 생성하도록 하는 코드를 포함하는 컴퓨터 판독 가능 저장 매체.

청구항 12

제 11 항에 있어서,

상기 매칭 노드 스트림 생성 코드는

상기 트위그 쿼리 브랜칭 노드에 대응되며 상기 XML 문서의 브랜칭 노드의 레이블링 값으로 이루어진 브랜칭 노드 스트림을 생성하고, 상기 트위그 쿼리 리프 노드에 대응되며 상기 XML 문서의 리프 노드의 레이블링 값으로 이루어진 리프 노드 스트림을 생성하는 코드를 포함하는 것을 특징으로 하는 컴퓨터 판독 가능 저장 매체.

청구항 13

제 11 항에 있어서,

상기 최종 리프 노드 스트림을 생성하도록 하는 코드는

상기 브랜칭 노드 스트림과 상기 리프 노드 스트림을 상호 매칭하고, 상기 리프 노드 스트림을 구성하는 각각의 리프 노드의 레이블링 값이 상기 브랜칭 노드 스트림 중 어느 하나의 브랜칭 노드의 레이블링 값을 포함하는지의 여부를 판단하여, 상기 리프 노드의 레이블링 값이 어느 하나의 브랜칭 노드의 레이블링 값을 포함하지 않으면 상기 리프 노드의 레이블링 값을 상기 리프 노드 스트림에서 제거하는 코드를 포함하는 것을 특징으로 하는 컴퓨터 판독 가능 저장 매체.

청구항 14

제 11 항에 있어서,

상기 트위그 쿼리에 대한 응답 메시지를 생성하도록 하는 코드는

상기 최종 리프 노드 스트림을 구성하는 복수 개의 리프 노드의 레이블링 값 중 상호 모두 중복되는 중첩 레이블링 엘리먼트를 제거하고, 상기 중첩 레이블링 엘리먼트가 제거된 상기 리프 노드의 잔여 레이블링 엘리먼트를 XML 문서 상의 노드명으로 변환하는 코드를 포함하는 것을 특징으로 하는 컴퓨터 판독 가능 저장 매체.

명세서

발명의 상세한 설명

기술 분야

- <1> 본 발명은 XML 문서의 질의 처리에 관한 것으로서, 특히 XML 문서를 처리하여 데이터 베이스화함에 있어, 데이터 베이스에 저장된 XML 문서에 대하여 XML 트위그 패턴(twig pattern)을 가지는 질의처리에 대하여, XML 문서의 파싱(parsing)시 제공되는 정보를 가공함으로써 XML 트위그 패턴(twig pattern)을 가지는 질의처리시 질의 처리 시간을 효율적으로 감소시킬 수 있는 브랜칭 노드 스트림을 이용한 XML 문서의 트위그 패턴 질의 처리 방법, 그 장치 및 컴퓨터 판독 가능 저장 매체를 제공하는 것이다.

배경 기술

- <2> 확장성 마크업 언어(eXtensible Markup Language:XML)은 하이퍼텍스트 마크업 언어(HyperText Markup Language:HTML)의 단점을 보완하기 위하여 채택된 마크업 언어이다.
- <3> XML은 문서 구조, 문서 내용 및 출력 형식이 각각 분리되어 전체 문서를 정의하므로, 문서 구조의 재활용 및 출력형식의 유연성, 문서구조에 대한 검색 기능 등의 문서구조화에 따른 다양한 특성을 제공한다는 점에서 XML 문서의 확산이 더욱 가속화되고 있다.
- <4> 현재 웹 상에서 데이터 표현을 위한 XML의 사용과 관심이 늘어나면서, XML 문서의 효율적인 질의 처리에 대한 관심 또한 증가하고 있다.
- <5> 따라서, XML 문서의 효율적인 질의 처리를 위한 많은 질의 처리 방법이 제안되었다.
- <6> 기존의 효율적인 질의 처리 관련 연구는 단일 경로는 가지는 단순 패스 질의(Simple Path Query) 처리가 주를 이루고 있었으나, XML 문서 처리시 가장 처리 시간을 많이 요구하는 두 개 이상의 경로를 가지는 트위그 질의 처리(Twig query processing)가 최근 들어 가장 중요한 문제로 인식되고 있다.
- <7> 이러한 트위그 질의(Twig query)를 효율적으로 처리하기 위하여 지역 인코딩 레이블링(Region Encoding Labeling), 듀웨이 레이블링(Dewey Labeling), 확장된 듀웨이 레이블링(Extended Dewey Labeling)등의 레이블링 기법이 제안되었으며, 제안된 레이블링 기법을 활용하여, TJFast, TwigStack, Twig²Stag등 효율적으로 질의를 처리하기 위한 트위그 패턴 매칭 알고리즘(Twig pattern matching algorithm)이 제안되었다.
- <8> 상기 TJFast 알고리즘은 확장된 듀웨이 레이블링을 기반으로 하고, TwigStack, Twig²Stag등의 알고리즘은 효율적인 질의 처리를 위해 지역 인코딩(region encoding) 레이블링 기반의 레이블을 이용한다.
- <9> 상기 TJFast, TwigStack, Twig²Stag등의 알고리즘은 효율적인 질의 처리를 위해 모두 지역 인코딩 레이블링 기반의 레이블을 이용한다.
- <10> 도 1은 XML 문서의 일반적인 예를 도시한 것이다.
- <11> 도 1을 참조하면, XML 문서는 계층적 트리 형태로 구성되며, 이러한 구성으로 인하여 문서 구조에 대한 검색이 가능하다.
- <12> XML 질의 언어는 W3C에서 표준으로 제안된 XML 질의 언어인 XML Query Language(XQL), 질의 언어 Query(A Query Language for XML)를 그 예로 들 수 있는데, 이들은 질의 표현에 있어서 XML 문서의 엘리먼트나 텍스트들의 경로를 나타내는 Xpath(XML Path Language)라는 별도의 스펙을 사용하고 있다.
- <13> 도 2는 XML 문서에 대한 클라이언트의 특정 쿼리(query)의 구조를 도시한 것이다.
- <14> 도 2를 참조하면 클라이언트의 특정 쿼리는 노드 간의 관계를 나타내는 구조 표현식인 '/' 또는 '//'를 사용함

으로써 XML 문서를 구성하는 노드 간의 관계를 지정한다.

- <15> 여기서, 구조 표현식 '/'는 부모 노드와 자식 노드 간의 관계를 나타내고, '//'는 조상 노드와 자손 노드와의 관계를 나타낸다.
- <16> 이를 기반으로 도 2를 상술하면, 클라이언트의 쿼리는 노드 A와 노드 D에 있어서 중간 노드가 어떤 노드인지 상관없이 노드 A에 대하여 자식 노드를 제외한 자손 노드 D를 가지며, 자식 노드가 B이고, 자손 노드가 C인 노드에 대한 쿼리를 요청하는 것임을 확인할 수 있다.
- <17> 여기서, 노드 A는 도 2에서 볼 수 있는 바와 같이 분기되는 노드를 가지기 때문에 브랜칭 노드(Branching node)라 하고, 노드 C 및 노드 D는 말단에 위치하기 때문에 리프 노드(Leaf node)라고 한다.
- <18> 도 3은 일반적인 지역 인코딩(Region encoding) 기반의 레이블링 기법을 도시한 것이다.
- <19> 지역 인코딩 기반의 레이블링 기법은 도 3에 도시된 바와 같이 부모 노드와 자식 노드 간의 관계에 있어서, 자식 노드의 레이블링은 레이블링된 부모 노드의 레이블링 값 사이에 존재하여야 한다.
- <20> 이는, XML 문서의 구조 탐색시, 문서 구조 탐색 순서를 레이블링된 값으로 대체할 수 있고, 아울러, 부모 노드와 자식 노드 간의 관계를 쉽게 알 수 있다.
- <21> 그러나, 이러한 지역 인코딩 기반의 레이블링 기법은 클라이언트의 질의 요청시 트위그 쿼리에 기술된 노드를 모두 탐색하여야 한다는 문제점을 가지고 있다.
- <22> 하기에 상술되는 TJFast의 알고리즘은 효율적인 질의 처리를 위하여 Dewey Labeling 방법을 확장한 확장된 듀웨이 레이블링(Extended Dewey Labeling)의 레이블링 기법을 적용한다.
- <23> 도 4는 XML 문서 탐색에 있어서의 TJ Fast기법의 일례를 도시한 것이다.
- <24> TJ Fast 기법은 확장된 듀웨이 레이블링(Extended Dewey Labeling)을 이용하여 문서 구조를 레이블링 하고, 트위그 패턴 매칭 알고리즘을 이용하여 클라이언트의 질의에 대한 응답을 수행할 수 있다.
- <25> 확장된 듀웨이 레이블링 기법은 리프 노드의 레이블링 정보를 알면 XML 문서에서 리프 노드의 상위 계층의 노드 정보를 모두 알 수 있다는 점에서 출발한다.
- <26> 도 4를 참조하면, 상위 계층에서 하위 계층으로 갈수록 레이블링 수가 1씩 증가되며, 동일 계층에서 동일 부모 노드를 가지는 자식노드는 부모 노드의 레이블링을 상속받고, 자신의 고유 레이블링을 가지게 된다.
- <27> 이에 의하면, 최상위 계층인 루트 노드(Root node)는 인덱스 되지 않으며, 루트 노드의 하위 계층인 제 1 계층은 (0)으로, 상기 제 1 계층의 하위에 속하는 제 2 계층인 노드 f_1 은 (0, 0)이고, 제 3계층에 속하며 노드 f_1 의 자식 노드 중 왼쪽 노드 a_1 은 (0, 0, 0)으로 레이블링 되고, 그 옆의 오른쪽 노드 a_2 는 (0, 0, 1)로 레이블링된다.
- <28> 한편, 이와 같은 방식으로 노드 d_1 은 (0, 0, 0, 0)으로, b_1 은 (0, 0, 0, 1)로 레이블링되고, 노드 d_2 는 노드 a_2 의 자식 노드이므로 부모 노드 a_2 의 레이블링 값에 레이블링 엘리먼트를 추가하여 (0, 0, 1, 0)으로 레이블링되고, 노드 b_2 는 동일한 방식으로 (0, 0, 1, 1)로 레이블링되며, 노드 c_2 는 (0, 0, 1, 1, 0)으로 레이블링된다.
- <29> 이와 같은 방식으로 XML 문서를 구성하는 모든 노드가 레이블링될 수 있다. 이와 같은 레이블링 기법은 말단 노드의 레이블링 정보에 따라 XML 문서의 상위 구조를 파악할 수 있고, 레이블링 정보만으로 상위 노드의 레벨 값 및 노드 네임(name)을 알 수 있다.
- <30> 여기서, 상기 도 2와 같은 클라이언트의 트위그 쿼리(Twig Query)가 전송되면, 쿼리에 대하여 TJ Fast 기법에 따라 문서를 탐색하게 된다.
- <31> TJ Fast 기법은 클라이언트의 쿼리의 노드 구조와 동일한 구조를 가지는 XML 문서의 결과값을 추출하여 클라이언트에게 전송하기 위하여 XML 전체 문서 중에 쿼리의 리프 노드에 대응되는 XML 문서의 노드를 모두 탐색하

다.

- <32> 왜냐하면, TJ Fast 기법은 전술한 Extended Dewey Labeling 기법에 의해 문서가 레이블링 되어 있으므로 쿼리의 리프 노드에 대응되는 XML 문서의 노드의 레이블링 값을 알면, XML 문서의 상위 구조를 파악할 수 있기 때문이다.
- <33> 한편, 도 4에 도시된 테이블과 같이, 쿼리의 리프 노드 D에 대응되는 XML 문서 중 노드 D로부터 구성한 리프 노드 테이블인 T_D , 쿼리의 노드 C에 대응되는 XML 문서 중 노드 C로부터 구성한 리프 노드 테이블인 T_C 에 대하여 레이블링 값을 노드 명으로 변환하는 과정이 수행된다.
- <34> 즉, XML 문서 중 노드 D로부터 구성한 리프 노드 테이블인 T_D 는 (0, 0, 0, 0), (0, 0, 1, 0)로 구성되며, 노드 C로부터 구성한 리프 노드 테이블인 T_C 는 (0, 0, 0, 1, 0), (0, 0, 1, 1, 0)으로 구성된다.
- <35> 여기서, 리프 노드 테이블인 T_D 는 중간 값이 레이블링 값으로 구성되어 있으므로 최종적으로 노드 명으로 변환하는 과정이 수행된다.
- <36> 따라서, 도 4를 참조할 때, 리프 노드 D로부터 구성한 리프 노드 테이블인 T_D 에서 (0, 0, 0, 0)이면 (e1, f1, a1, d1)의 노드 명으로 변환되고, (0, 0, 1, 0)이면 (e1, f1, a2, d2)의 노드 명으로 변환된다.
- <37> 그리고, 노드 C로부터 구성한 리프 노드 테이블인 T_C 에서 (0, 0, 0, 1, 0)이면 (e1, f1, a1, b1, c1)이고, (0, 0, 1, 1, 0)이면 (e1, f1, a2, b2, c2)가 된다.
- <38> 이와 같이, TJ Fast 레이블링 기법은 쿼리의 리프 노드에 대응되는 XML 문서 상의 노드의 레이블링 정보를 알아야 XML 문서에서 노드의 상위 구조를 파악할 수 있기 때문에 상기 XML 문서 중 쿼리의 리프 노드에 대응되는 노드를 모두 파악하여야 하여야 하고, 레이블링 값으로 출력되는 중간 결과값에 대하여 레이블링 값을 구성하는 모든 엘리먼트 값을 노드 명으로 변환하여야 한다.
- <39> 도 5는 종래의 지역 인코딩 기반으로 레이블링된 XML 문서에서의 트위그 스택(TwigStack) 탐색 기법의 일 실시예를 도시한 것이다.
- <40> 도 5를 참조하면, 도 3의 지역 인코딩 기법으로 레이블링된 XML 문서를 도시하고 있다.
- <41> 트위그 스택에서는 트위그 패턴 매칭(Twig Pattern Matching)을 위하여 스택(Stack) 구조를 이용한다.
- <42> 이는 도 5에서 볼 수 있는 바와 같이, 트위그 패턴 매칭 알고리즘에서는 쿼리 안에 있는 노드에 맞는 노드들이 삽입된다.
- <43> 여기서는 몇 가지 문제점이 존재하는데 이를 살펴보면 다음과 같다.
- <44> 첫 번째는 XML 문서에 있는 모든 노드가 쿼리 노드와 동일한 경우 스택의 크기가 XML 문서의 크기와 동일해지는 문제점이 발생한다.
- <45> 두 번째는 부모 노드와 자식 노드(Parent-Children:PC)간의 관계에서, 트위그 스택 알고리즘은 좋지 못한 성능을 보여준다.
- <46> 예를 들어, A/[C]/B와 쿼리가 주어졌을 때, 트위그 스택 알고리즘은 A/B의 관계와 같은 조상노드-자손노드(Ancessor-Descendant:AD)의 관계에서는 최적의 처리를 보여주나, 이와 더불어 A/C와 같은 부모 노드와 자식 노드간의 관계에서는 질의처리 속도가 매우 느려진다.
- <47> 지역 인코딩 레이블링 기반의 다른 트위그 패턴 매칭 알고리즘으로는 트위그 스퀘어 스택(Twig²Stack)이 있다.
- <48> 도 6은 종래의 지역 인코딩 기반으로 레이블링된 XML 문서에서의 트위그 스퀘어 스택(Twig²Stack) 탐색 기법의 일 실시예를 도시한 것이다.
- <49> 이 트위그 스퀘어 스택 탐색 기법은 전술한 트위그스택의 부모 노드와 자식 노드의 관계 처리시 처리 속도가

느려지는 처리 속도 문제를 효율적으로 개선한 기법이다.

- <50> 트위그 스퀘어 스택 탐색 기법은 도 6에 도시된 바와 같이, 변형된 스택 구조를 사용한다.
- <51> 이 스택 구조는 계층적 스택(Hierarchical Stack:HS)이라고 불려진다. 이 계층적 스택을 이용하여 트위그 스퀘어 스택은 클라이언트의 트위그 쿼리의 관계를 표현한다.
- <52> 이 방법을 사용함으로써 트위그 스퀘어 스택은 트위그 스택에서 비효율적이었던 조상노드-자손노드의 관계에 대한 처리를 좀 더 효율적으로 개선할 수 있고, 트위그 쿼리의 처리 시간을 좀 더 빠르게 할 수 있다.
- <53> 그러나, 트위그 스퀘어 스택 역시 트위그 스택과 마찬가지로 최악의 경우 문서의 모든 노드를 스택 안에 저장하고 있어야 되는 문제점이 발생한다.
- <54> 이와는 별도로, 트위그 쿼리 안의 노드와 XML 문서 상의 노드를 매칭시킬 때 최종 결과와 상관없는 중간 노드들이 스택안에 들어가 쿼리 프로세싱의 불필요한 처리 과정을 생성하게 된다.
- <55> 또한, 트위그 스퀘어 스택은 노드 간의 관계(relationship)를 가지고 있으며, 만일 노드 간의 관계가 복잡한 경우에는 쿼리 프로세싱 시간이 길어지는 문제점을 안고 있다.
- <56> 이와 같이, 트위그 스택과 트위그 스퀘어 스택 알고리즘은 지역 인코딩 레이블링 방법을 사용하였기 때문에 지역 인코딩 레이블링의 문제점을 그대로 가지고 있다.
- <57> 지역 인코딩 레이블링 은 상위 노드와 하위 노드와의 관계를 파악할 수 있는 정보 이외에 다른 정보를 가지고 있지 않다.
- <58> 따라서, 스택 안에는 XML 문서 노드의 이름을 알기 위하여 XML 문서에서 기술된 내용을 모두 가지고 있어야 한다. 이와 같은 경우에는 저장 공간의 크기가 커지게 되는 문제점을 안게 된다.
- <59> 따라서, 클라이언트의 트위그 쿼리에 대하여 트위그 스택과 트위그 스퀘어 스택 알고리즘은 지역 인코딩 레이블링 기법을 적용하므로, 상위 노드와 하위 노드와의 관계를 나타내는 정보외에 XML 문서에서 기술된 정보를 가지고 있어야 하기 때문에 저장 공간이 커지게 되는 문제점이 있고, 더불어 트위그 스택 알고리즘은 브랜칭되는 부분을 기준으로 노드 구조를 분할하여 탐색하므로 탐색 시간이 길어져 전체 프로세싱 타임이 오래 소요되는 문제점이 있고, TJ Fast 알고리즘은 쿼리의 리프 노드에 대응되는 XML 문서상의 노드를 모두 탐색하여야 하므로 리프 노드 탐색 시간이 오래 소요되며, 탐색된 노드의 레이블링 값을 모두 노드명으로 변환하여야 하므로 프로세싱 타임이 길어지는 문제점이 있다.

발명의 내용

해결 하고자하는 과제

- <60> 따라서, 본 발명이 해결하고자 하는 첫 번째 과제는 데이터베이스에 저장된 XML 문서에 대하여 XML 트위그 패턴을 가지는 질의를 처리함에 있어 XML 데이터의 과잉시 제공되는 정보를 새롭게 가공하여 질의 응답 시간을 효율적으로 줄일 수 있는 브랜칭 노드 스트림을 이용한 XML 문서의 트위그 패턴 질의 처리 방법을 제공하는 것이다.
- <61> 본 발명이 해결하고자 하는 두 번째 과제는 상기 브랜칭 노드 스트림을 이용한 XML 문서의 트위그 패턴 질의 처리 방법을 이용한 브랜칭 노드 스트림을 이용한 XML 문서의 트위그 패턴 질의 처리 시스템을 제공하는 것이다.
- <62> 본 발명이 해결하고자 하는 세 번째 과제는 상기 브랜칭 노드 스트림을 이용한 XML 문서의 트위그 패턴 질의 처리 방법을 컴퓨터에서 실행할 수 있도록 프로그램으로 기록된 컴퓨터 판독 가능 저장매체를 제공하는 것이다.

과제 해결수단

- <63> 상기 첫 번째 과제를 해결하기 위하여 본 발명은,
- <64> 확장된 듀웨이 레이블링(Extended Dewey Labeling)된 XML 문서의 트위그 쿼리의 질의 처리 방법에 있어서, 클라이언트로부터 상기 XML 문서에 대하여 특정의 트위그 쿼리(Twig query)를 수신하면, 상기 특정의 트위그 쿼리의 트위그 쿼리 브랜칭 노드(Branching node) 및 트위그 쿼리 리프 노드(leaf node)를 추출하는 단계; 상기 XML 문

서를 파싱한 후, 상기 XML 문서에 상기 확장된 듀웨이 레이블링을 수행하여 상기 XML 문서의 모든 노드를 레이블링하는 단계; 상기 트위그 쿼리의 트위그 쿼리 브랜칭 노드 및 트위그 쿼리 리프 노드에 대응되는 상기 XML 문서의 브랜칭 노드 스트림 및 리프 노드 스트림을 생성하는 단계; 상기 리프 노드 스트림을 구성하는 리프 노드의 레이블링 값 중에서 상기 브랜칭 노드 스트림을 구성하는 브랜칭 노드의 레이블링 값을 포함하지 않는 리프 노드의 레이블링 값을 제거함으로써 최종 리프 노드 스트림을 생성하는 단계; 및 상기 브랜칭 노드 스트림 및 최종 리프 노드 스트림을 이용하여 상기 트위그 쿼리에 대한 응답 메시지를 생성하는 단계를 포함하는 XML 문서의 트위그 쿼리의 질의 처리 방법을 제공한다.

<65> 한편, 상기 XML 문서의 브랜칭 노드 스트림 및 리프 노드 스트림을 생성하는 단계는 상기 트위그 쿼리 브랜칭 노드에 대응되며 상기 XML 문서의 브랜칭 노드의 레이블링 값으로 이루어진 브랜칭 노드 스트림을 생성하는 단계; 및 상기 트위그 쿼리 리프 노드에 대응되며 상기 XML 문서의 리프 노드의 레이블링 값으로 이루어진 리프 노드 스트림을 생성하는 단계를 포함하는 것을 특징으로 한다.

<66> 그리고, 상기 최종 리프 노드 스트림을 생성하는 단계는 상기 브랜칭 노드 스트림과 상기 리프 노드 스트림을 상호 매칭하는 단계; 상기 리프 노드 스트림을 구성하는 각각의 리프 노드의 레이블링 값이 상기 브랜칭 노드 스트림 중 어느 하나의 브랜칭 노드의 레이블링 값을 포함하는지의 여부를 판단하는 단계; 및 상기 리프 노드의 레이블링 값이 어느 하나의 브랜칭 노드의 레이블링 값을 포함하지 않으면 상기 리프 노드의 레이블링 값을 상기 리프 노드 스트림에서 제거하는 단계를 포함한다.

<67> 아울러, 상기 트위그 쿼리에 대한 응답 메시지를 생성하는 단계는 상기 최종 리프 노드 스트림을 구성하는 복수 개의 리프 노드의 레이블링 값 중 상호 모두 중복되는 중복 레이블링 엘리먼트를 제거하는 단계; 및 상기 중복 레이블링 엘리먼트가 제거된 상기 리프 노드의 잔여 레이블링 엘리먼트를 XML 문서 상의 노드명으로 변환하는 단계를 포함하는 것을 특징으로 한다.

<68> 또한, 상기 응답 메시지는 상기 변환된 노드명을 이용하여 상기 클라이언트의 트위그 쿼리에 대응되는 구조로 형성된 XML 응답 메시지인 것을 특징으로 한다.

<69> 상기 두 번째 과제를 해결하기 위하여 본 발명은,

<70> 확장된 듀웨이 레이블링(Extended Dewey Labeling)된 XML 문서의 트위그 쿼리의 질의 처리 장치에 있어서, 클라이언트로부터 상기 확장된 듀웨이 레이블링된 XML 문서에 대하여 트위그 쿼리(Twig Query)를 수신하고, 상기 트위그 쿼리에 대한 응답메시지를 상기 클라이언트에 제공하는 송수신부; 상기 트위그 쿼리의 트위그 쿼리 브랜칭 노드(Branch node) 및 트위그 쿼리 리프 노드(leaf node)를 추출하는 트위그 쿼리 노드 추출부; 상기 XML 문서에 상기 확장된 듀웨이 레이블링을 수행하여 상기 XML 문서의 모든 노드를 레이블링하는 레이블링부; 상기 XML 문서의 전체 노드 스트림과 상기 트위그 쿼리의 레이블링 값을 매칭하고, 상기 전체 노드 스트림에 대하여 상기 트위그 쿼리의 레이블링 값에 대응되는 상기 XML 문서의 브랜칭 노드 스트림 및 리프 노드 스트림을 생성하는 매칭 노드 스트림 생성부; 상기 리프 노드 스트림을 구성하는 리프 노드의 레이블링 값 중에서 상기 브랜칭 노드 스트림을 구성하는 브랜칭 노드의 레이블링 값을 포함하지 않는 리프 노드의 레이블링 값을 제거함으로써 최종 리프 노드 스트림을 생성하는 최종 리프 노드 스트림 생성부; 및 상기 브랜칭 노드 스트림 및 최종 리프 노드 스트림을 이용하여 상기 트위그 쿼리에 대한 응답 메시지를 생성하는 응답 메시지 생성부를 포함하는 XML 문서의 트위그 쿼리의 질의 처리 장치를 제공한다.

<71> 한편, 상기 XML 노드 스트림 생성부는 상기 트위그 쿼리 브랜칭 노드에 대응되며 상기 XML 문서의 브랜칭 노드의 레이블링 값으로 이루어진 브랜칭 노드 스트림을 생성하고, 상기 트위그 쿼리 리프 노드에 대응되며 상기 XML 문서의 리프 노드의 레이블링 값으로 이루어진 리프 노드 스트림을 생성하는 것을 특징으로 한다.

<72> 그리고, 상기 최종 리프 노드 스트림 생성부는 상기 브랜칭 노드 스트림과 상기 리프 노드 스트림을 상호 매칭하는 매칭 모듈; 및 상기 리프 노드 스트림을 구성하는 각각의 리프 노드의 레이블링 값이 상기 브랜칭 노드 스트림 중 어느 하나의 브랜칭 노드의 레이블링 값을 포함하는지의 여부를 판단하고, 상기 리프 노드의 레이블링 값이 어느 하나의 브랜칭 노드의 레이블링 값을 포함하지 않으면 상기 리프 노드의 레이블링 값을 상기 리프 노드 스트림에서 제거하는 스트림 정제 모듈을 포함하는 것을 특징으로 한다.

<73> 아울러, 상기 최종 리프 노드 스트림을 구성하는 복수 개의 리프 노드의 레이블링 값 중 상호 모두 중복되는 중복 레이블링 엘리먼트를 제거하고, 상기 중복 레이블링 엘리먼트가 제거된 상기 리프 노드의 잔여 레이블링 엘리먼트를 상기 XML 문서의 노드명으로 변환하는 노드명 변환 모듈을 포함하는 것을 특징으로 한다.

<74> 한편, 상기 응답 메시지는 상기 변환된 노드명을 이용하여 상기 클라이언트의 트위그 쿼리에 대응되는 구조로

형성된 XML 응답 메시지인 것을 특징으로 한다.

- <75> 상기 세 번째 과제를 해결하기 위하여 본 발명은,
- <76> 컴퓨터 판독가능 저장매체에 있어서, 상기 컴퓨터 판독가능 저장매체에 상기 확장된 듀웨이 레이블링(Extended Dewey Labeling)된 XML 문서의 트위그 질의 처리를 컴퓨터에서 수행할 수 있도록 하는 프로그램이 기록되어 있고, 상기 프로그램은, 클라이언트로부터 상기 확장된 듀웨이 레이블링된 XML 문서에 대하여 특정의 트위그 쿼리(Twig query)를 수신하면, 상기 트위그 쿼리의 트위그 쿼리 브랜칭 노드(Branching node) 및 트위그 쿼리 리프 노드(leaf node)를 추출하도록 하는 코드; 상기 XML 문서를 파싱한 후, 상기 XML 문서에 상기 확장된 듀웨이 레이블링을 수행하여 상기 XML 문서의 모든 노드를 레이블링하도록 하는 코드; 상기 트위그 쿼리의 레이블링 값에 대응되는 상기 XML 문서의 브랜칭 노드 스트림 및 리프 노드 스트림을 생성하도록 하는 매칭 노드 스트림 생성 코드; 상기 리프 노드 스트림을 구성하는 리프 노드의 레이블링 값 중에서 상기 브랜칭 노드 스트림을 구성하는 브랜칭 노드의 레이블링 값을 포함하지 않는 리프 노드의 레이블링 값을 제거함으로써 최종 리프 노드 스트림을 생성하도록 하는 코드; 및 상기 브랜칭 노드 스트림 및 최종 리프 노드 스트림을 이용하여 상기 트위그 쿼리에 대한 응답 메시지를 생성하도록 하는 코드를 포함하는 컴퓨터 판독 가능 저장 매체를 제공한다.
- <77> 한편, 상기 매칭 노드 스트림 생성 코드는 상기 트위그 쿼리 브랜칭 노드에 대응되며 상기 XML 문서의 브랜칭 노드의 레이블링 값으로 이루어진 브랜칭 노드 스트림을 생성하고, 상기 트위그 쿼리 리프 노드에 대응되며 상기 XML 문서의 리프 노드의 레이블링 값으로 이루어진 리프 노드 스트림을 생성하는 코드를 포함할 수 있다.
- <78> 또한, 상기 최종 리프 노드 스트림을 생성하도록 하는 코드는 상기 브랜칭 노드 스트림과 상기 리프 노드 스트림을 상호 매칭하고, 상기 리프 노드 스트림을 구성하는 각각의 리프 노드의 레이블링 값이 상기 브랜칭 노드 스트림 중 어느 하나의 브랜칭 노드의 레이블링 값을 포함하는지의 여부를 판단하여, 상기 리프 노드의 레이블링 값이 어느 하나의 브랜칭 노드의 레이블링 값을 포함하지 않으면 상기 리프 노드의 레이블링 값을 상기 리프 노드 스트림에서 제거하는 코드를 포함할 수 있다.
- <79> 그리고, 상기 트위그 쿼리에 대한 응답 메시지를 생성하도록 하는 코드는 상기 최종 리프 노드 스트림을 구성하는 복수 개의 리프 노드의 레이블링 값 중 상호 모두 중복되는 중복 레이블링 엘리먼트를 제거하고, 상기 중복 레이블링 엘리먼트가 제거된 상기 리프 노드의 잔여 레이블링 엘리먼트를 XML 문서 상의 노드명으로 변환하는 코드를 포함하는 것을 특징으로 한다.

효 과

- <80> 본 발명에 의하면, 클라이언트의 트위그 쿼리에 대하여 트위그 쿼리의 리프 노드에 매칭되어지는 XML 문서상의 노드의 탐색 시간을 줄임으로써 리프 노드 탐색 시간인 I/O 타임을 줄일 수 있으며, 트위그 쿼리에 대하여 탐색된 노드에 대하여 리프 노드에 레이블링된 값을 원래 노드 이름으로 변환하는 할 때, 트위그 쿼리에 대응되는 브랜칭 노드와 중첩되는 레이블링된 값을 원래 노드 이름으로 변환하지 않아 중간 프로세싱 타임이 줄어들고, 전체적으로 쿼리에 대하여 응답하는 전체 프로세싱 타임이 줄어 신속한 질의 응답 처리를 수행할 수 있는 효과가 있다.

발명의 실시를 위한 구체적인 내용

- <81> XML 쿼리 프로세싱 분야에서 가장 핵심적인 연구는 불필요한 중간 연산 값을 제거하는 방법과 쿼리의 응답시간을 최소화하는 방법이다.
- <82> 여기서 기술하는 불필요한 중간 연산 값은 최종 결과에는 포함되지 않는 것이지만, 프로세싱 시에 처리하여야 하는 값을 말한다. 쿼리의 응답시간은 클라이언트가 쿼리를 주었을 때, 프로세싱을 하여 클라이언트가 원하는 최종 값을 전송하기까지의 시간을 의미한다.
- <83> 본 발명은 확장된 듀웨이 레이블링((Extended Dewey Labeling) 기법을 기반으로 하여 탐색하여야 하는 리프 노드의 수를 줄일 수 있고, 브랜칭 노드를 우선적으로 탐색하기 때문에 종래에 제안된 트위그 쿼리의 질의 처리 기법보다 중간 연산과정을 줄이고, 확장된 듀웨이 레이블링을 이용할 경우에 레이블링 값을 XML 문서의 노드명으로 변환하는 과정을 줄임으로써 쿼리 프로세싱 시간을 전체적으로 효율적으로 감소시킬 수 있다.
- <84> 도 7은 본 발명의 일 실시예에 따른 XML 문서의 트위그 쿼리의 질의 처리 방법의 흐름도이다.
- <85> 도 7을 참조하면, 우선, 확장된 듀웨이 레이블링(Extended Dewey Labeling)된 XML 문서의 트위그 쿼리의 질의 처리 방법에 있어서, 클라이언트로부터 XML 문서에 대하여 특정의 트위그 쿼리(Twig query)를 수신하면, 상기

트위그 쿼리의 트위그 쿼리 브랜칭 노드(Branching node) 및 트위그 쿼리 리프 노드(leaf node)로 이루어진 상기 트위그 쿼리의 노드들을 추출한다(S710).

<86> 본 발명은 우선 2가지의 과정을 전제로 한다.

<87> 첫 번째는 본 발명의 XML 문서에서의 각 노드는 확장된 듀웨이 레이블링(Extended Dewey Labeling) 기법에 따라 레이블링된다고 가정한다.

<88> 그리고, 두 번째는 본 발명에 적용되는 클라이언트의 쿼리는 트위그 쿼리로 가정한다. 물론 트위그 쿼리가 아니라 단일 경로를 가지는 단순 패스 쿼리(Simple Path Query)에도 본 발명이 적용되지만, 본 발명에서의 탁월한 질의 처리 성능은 트위그 쿼리에서 그 우수한 성능을 확인할 수 있다.

<89> 한편, 본 발명은 우선, 클라이언트로부터 특정 경로에 대한 트위그 쿼리를 수신한다.

<90> 그러면, 트위그 쿼리에 대하여 트위그 쿼리 자신의 트위그 쿼리 브랜칭 노드와 트위그 쿼리 리프 노드를 우선 탐색한다.

<91> 이 과정(S710)을 통하여, 트위그 쿼리에 대하여 브랜칭 노드와 리프 노드를 구별할 수 있다. 본 발명에서는 트위그 쿼리에 존재하는 브랜칭 노드를 트위그 쿼리 브랜칭 노드라 명명하고, 트위그 쿼리에 존재하는 리프 노드를 트위그 쿼리 리프 노드라 명명하여 용어의 혼동을 피하기로 한다.

<92> 이와 같이, 트위그 쿼리 브랜칭 노드와 트위그 쿼리 리프 노드를 추출하고, XML 문서에서 상기 트위그 쿼리의 트위그 쿼리 브랜칭 노드와 트위그 쿼리 리프 노드에 대응되는 XML 문서의 노드를 레이블링된 값으로 추출하게 된다.

<93> 그 다음, 상기 XML 문서를 파싱한 후, 상기 XML 문서에 상기 확장된 듀웨이 레이블링을 수행하여 상기 XML 문서의 모든 노드를 레이블링한다(S720).

<94> 이는 XML 문서의 모든 노드를 레이블링 값으로 추출해내는 과정이다. 즉, 본 발명에서는 XML 문서를 파싱하면서, 계층적 문서 구조의 형식상 확장된 듀웨이 레이블링되고, XML 문서의 각각의 노드는 레이블링 값으로 표현된다.

<95> 그 다음, 상기 트위그 쿼리의 트위그 쿼리 브랜칭 노드 및 트위그 쿼리 리프 노드에 대응되는 상기 XML 문서의 브랜칭 노드 스트림 및 리프 노드 스트림을 생성한다(S730).

<96> 본 발명은 트위그 쿼리와 XML 문서를 매칭하여 트위그 쿼리에 대응되는 노드 스트림을 생성한다.

<97> 본 발명에서는 트위그 쿼리에 대응되며, XML 문서의 브랜칭 노드의 레이블링 값을 모두 포함하는 값을 브랜칭 노드 스트림이라 정의하고, 트위그 쿼리에 대응되며 XML 문서의 리프 노드의 레이블링 값을 모두 포함하는 값을 리프 노드 스트림이라 정의한다.

<98> 상세하게는, XML 문서의 브랜칭 노드 스트림 및 리프 노드 스트림을 생성하는 과정은, 트위그 쿼리 브랜칭 노드에 대응되며 상기 XML 문서의 브랜칭 노드의 레이블링 값으로 이루어진 브랜칭 노드 스트림을 생성하고, 그 다음 상기 트위그 쿼리 리프 노드에 대응되며 상기 XML 문서의 리프 노드의 레이블링 값으로 이루어진 리프 노드 스트림을 생성한다.

<99> 이는 XML 문서 상에서 트위그 쿼리의 노드와 동일한 노드가 존재하므로, 클라이언트의 트위그 쿼리의 브랜칭 노드와 동일한 XML 문서의 브랜칭 노드의 레이블링 값으로 이루어진 브랜칭 노드 스트림을 생성하고, 트위그 쿼리의 리프 노드에 대응되는 XML 문서의 노드의 레이블링 값으로 이루어진 리프 노드 스트림을 생성하는 것이다.

<100> 그 다음, 상기 리프 노드 스트림을 구성하는 리프 노드의 레이블링 값 중에서 상기 브랜칭 노드 스트림을 구성하는 브랜칭 노드의 레이블링 값을 포함하지 않는 리프 노드의 레이블링 값을 제거함으로써 최종 리프 노드 스트림을 생성한다(S740).

<101> 즉, 확장된 듀웨이 레이블링은 리프 노드의 레이블링 값을 알 수 있으면 상위 노드의 구조 정보를 파악할 수 있기 때문에 쿼리의 리프 노드에 대응되는 XML 문서의 리프 노드를 탐색하는 과정이 중요하다.

<102> 그러나, 모든 리프 노드를 탐색할 경우에는 쿼리의 브랜칭 노드와 관련이 없는 리프 노드까지 탐색될 수 있으므로, 본 발명에서는 탐색 되는 리프 노드의 갯수를 최대한 감소시키기 위하여 쿼리의 리프 노드에 대응되는 XML 문서의 리프 노드 뿐만 아니라, 쿼리의 브랜칭 노드에 대응되는 XML 문서의 브랜칭 노드까지 탐색한다.

- <103> 상세하게는, 최종 리프 노드 스트림을 생성하기 위하여, 브랜칭 노드 스트림과 상기 리프 노드 스트림을 상호 매칭하고, 리프 노드 스트림을 구성하는 각각의 리프 노드의 레이블링 값이 상기 브랜칭 노드 스트림 중 어느 하나의 브랜칭 노드의 레이블링 값을 포함하는지의 여부를 판단하여, 리프 노드의 레이블링 값이 어느 하나의 브랜칭 노드의 레이블링 값을 포함하지 않으면 상기 리프 노드의 레이블링 값을 상기 리프 노드 스트림에서 제거함으로써 최종 리프 노드 스트림을 생성한다.
- <104> 본 발명의 특징은 클라이언트의 트위그 쿼리에 대하여 모든 노드를 탐색하는 것이 아니라, 우선적으로 트위그 쿼리의 최상위에 위치한 트위그 노드를 기반으로 XML 문서의 브랜칭 노드를 탐색하는 것이 가장 탐색 속도가 빠르다는 것에 그 기반을 두고 있다.
- <105> 왜냐하면, 모든 노드를 탐색하는 것보다는 출발점인 브랜칭 노드를 먼저 탐색하고, 그 하위의 노드를 탐색하는 것이, 모든 노드를 탐색하는 것보다 불필요한 탐색 경로를 없앨 수 있기 때문이다.
- <106> 이와 같이, XML 문서의 브랜칭 노드를 우선 탐색하여 탐색된 브랜칭 노드를 기반으로 브랜칭 노드 스트림을 구성하고, 상기 리프 노드 스트림을 구성하는 리프 노드의 레이블링 값 중에서 상기 브랜칭 노드 스트림에 속한 브랜칭 노드의 레이블링 값을 포함하지 않는 리프 노드의 레이블링 값을 제거함으로써 불필요한 리프 노드를 제거할 수 있다.
- <107> 이 과정은 전술한 바와 같이, 본 발명은 모든 리프 노드를 탐색하지 않기 위하여 브랜칭 노드와 조상 노드-자손 노드의 관계에 있지 않은 리프 노드의 레이블링 값을 제거하는 과정을 포함하게 된다.
- <108> 이를 위하여 확장된 듀웨이 레이블링 기법에서는 리프 노드는 상위 조상 노드의 구조를 포함하는 형식으로 레이블링되기 때문에 상위에 브랜칭 노드가 있을 경우, 브랜칭 노드의 레이블링 값을 포함하는 형식으로 레이블링되어 있다.
- <109> 따라서, 이 과정에서는 리프 노드 스트림을 구성하는 리프 노드 각각의 레이블링 값 중에서 브랜칭 노드 스트림을 구성하는 브랜칭 노드의 레이블링 값을 포함하지 않는 리프 노드가 있을 경우, 이 리프 노드의 레이블링 값을 제거하여 최종 리프 노드 스트림을 생성함으로써 클라이언트의 쿼리에 대하여 XML 문서의 탐색시 탐색하여야 할 리프 노드의 수를 줄일 수 있다.
- <110> 즉, 본 발명은 클라이언트의 트위그 쿼리에 대하여 XML 문서상에 매칭되는 트위그 노드를 우선적으로 탐색함으로써 트위그 노드가 아닌 노드를 탐색대상에서 우선적으로 제외시키고, 그 다음, 트위그 노드의 하위 노드에 해당되는 리프 노드만을 리프 노드 스트림에 포함시킴으로써, 상기 리프 노드 스트림 중 불필요한 리프 노드를 삭제하여 최종 리프 노드 스트림을 생성하므로 쿼리에 매칭되는 노드를 빨리 탐색할 수 있게 된다.
- <111> 마지막으로, 상기 브랜칭 노드 스트림 및 최종 리프 노드 스트림을 이용하여 상기 트위그 쿼리에 대한 응답 메시지를 생성한다(S750).
- <112> 즉, 브랜칭 노드 스트림에 포함된 노드의 레이블링 값과 최종 리프 노드 스트림에 포함된 레이블링 값을 이용하여, 트위그 쿼리에 대한 응답 메시지를 생성하게 된다.
- <113> 이는 본 발명의 모든 노드는 확장된 듀웨이 레이블링 기법으로 레이블링 되어 있으므로, 각 노드의 레이블링 값만으로도 상위 노드의 구조를 파악할 수 있고, 이를 기반으로 클라이언트의 질의에 맞는 응답 메시지를 생성할 수 있는 것이다.
- <114> 이와 같이 클라이언트의 질의에 맞는 응답 메시지를 생성하기 위하여, 최종 리프 노드 스트림을 구성하는 복수개의 리프 노드의 레이블링 값 중 상호 모두 중복되는 중첩 레이블링 엘리먼트를 제거하고, 중첩 레이블링 엘리먼트가 제거된 상기 리프 노드의 잔여 레이블링 엘리먼트를 XML 문서 상의 노드명으로 변환한다.
- <115> 여기서, 응답 메시지는 상기 변환된 노드명을 이용하여 상기 클라이언트의 트위그 쿼리에 대응되는 구조로 형성된 XML 응답 메시지일 수 있다.
- <116> 본 발명은 확장된 듀웨이 레이블링 기법에 따라 모든 노드가 레이블링 되는데, 전술한 바와 같이, 각 레이블링 값은 상위 노드의 구조를 모두 포함하는 값으로 레이블링된다.
- <117> 이 때문에, 레이블링 값을 노드 명으로 변환하여 노드 구조를 형성하는 과정에서 노드명으로 변환하는 과정을 가급적 줄임으로써 클라이언트로의 응답메시지를 빨리 생성할 수 있게 한다.
- <118> 즉, 확장된 듀웨이 레이블링 기법에 따라 레이블링 된 브랜칭 노드와 리프 노드에 있어서, 리프 노드의 레이블

링 값을 모두 노드 명으로 변환하면, 레이블링 엘리먼트 중 반드시 중첩되는 노드 명이 발생한다.

- <119> 따라서, 본 발명에서는 가급적 노드명으로 변환하는 과정을 줄여 클라이언트로의 응답시간인 전체 프로세싱 타임을 줄일 수 있으므로 보다 빠른 응답시간을 제공한다.
- <120> 도 8은 본 발명의 일 실시예에 따른 XML 문서의 트위그 쿼리의 질의 처리 장치의 블록도이다.
- <121> 도 8을 참조하면, 본 발명의 일 실시예에 따른 XML 문서의 트위그 쿼리의 질의 처리 장치(800)는 송수신부(810), 트위그 쿼리 노드 추출부(820), 레이블링부(830), 매칭 노드 스트림 생성부(840), 최종 리프 노드 스트림 생성부(850) 및 응답 메시지 생성부(860)를 포함할 수 있다.
- <122> 우선, 송수신부(810)는 클라이언트(890)로부터 XML 문서에 대하여 특정의 트위그 쿼리(Twig query)를 수신하고, 클라이언트의 트위그 쿼리에 대한 응답 메시지를 상기 클라이언트(890)에 제공한다.
- <123> 그리고, 트위그 쿼리 노드 추출부(820)는 상기 트위그 쿼리의 트위그 쿼리 브랜칭 노드(Branch node) 및 트위그 쿼리 리프 노드(leaf node)를 추출한다.
- <124> 즉, 트위그 쿼리 노드 추출부(820)는, 클라이언트의 트위그 쿼리에 대하여 트위그 쿼리 브랜칭 노드와 트위그 쿼리 리프 노드를 우선적으로 탐색하는 역할을 수행한다. 이에 의해 트위그 쿼리의 브랜칭 노드와 리프 노드를 구별할 수 있다.
- <125> 본 발명에서는 트위그 쿼리에 존재하는 브랜칭 노드를 트위그 쿼리 브랜칭 노드라 명명하고, 트위그 쿼리에 존재하는 리프 노드를 트위그 쿼리 리프 노드라 명하여 용어의 혼동을 피하기로 한다.
- <126> 한편, 트위그 쿼리 노드 추출부(820)는 트위그 쿼리에 대하여 트위그 쿼리 브랜칭 노드와 트위그 쿼리 리프 노드를 추출하고, 추출된 트위그 쿼리 브랜칭 노드와 트위그 쿼리 리프 노드를 매칭 노드 스트림 생성부(840)에 전송하여, 트위그 쿼리에 대응되는 XML 문서의 브랜칭 노드의 레이블링 값과, 트위그 쿼리에 대응되는 XML 문서의 리프 노드의 레이블링 값을 매칭 노드 스트림 생성부(840)에서 추출할 수 있도록 한다.
- <127> 그리고, 레이블링부(830)는 XML 문서를 파싱하고, XML 문서에 존재하는 모든 노드에 대하여 레이블링 값을 가지도록 하는 역할을 수행할 수 있도록 한다.
- <128> 본 발명에서, XML 문서는 계층적 문서 구조의 형식을 가지고 있으므로, XML 문서에 해당하는 각각의 노드를 확장된 듀웨이 레이블링 기법에 따라 레이블링할 수 있다.
- <129> 한편, 매칭 노드 스트림 생성부(840)는 트위그 쿼리의 트위그 쿼리 브랜칭 노드 및 트위그 쿼리 리프 노드에 대응되는 XML 문서의 브랜칭 노드 스트림 및 리프 노드 스트림을 생성한다.
- <130> 본 발명에서는 트위그 쿼리에 대응되는 XML 문서의 브랜칭 노드의 레이블링 값을 모두 포함하는 값을 브랜칭 노드 스트림이라 하고, 쿼리에 대응되는 XML 문서의 리프 노드의 레이블링 값을 모두 포함하는 값을 리프 노드 스트림이라 칭한다.
- <131> 상세하게는 상기 매칭 노드 스트림 생성부(840)는 트위그 쿼리 브랜칭 노드에 대응되며 상기 XML 문서의 브랜칭 노드의 레이블링 값으로 이루어진 브랜칭 노드 스트림을 생성하고, 그 다음, 상기 트위그 쿼리 리프 노드에 대응되며, 상기 XML 문서의 리프 노드의 레이블링 값으로 이루어진 리프 노드 스트림을 생성한다.
- <132> 이는 XML 문서 상에서 트위그 쿼리의 노드와 동일한 역할을 하는 노드가 존재하므로, 클라이언트의 트위그 쿼리의 브랜칭 노드와 동일한 XML 문서의 브랜칭 노드의 레이블링 값으로 이루어진 브랜칭 노드 스트림을 생성하고, 트위그 쿼리의 리프 노드에 대응되는 XML 문서의 레이블링 값으로 이루어진 리프 노드 스트림을 생성하는 것이다.
- <133> 한편, 최종 리프 노드 스트림 생성부(850)는 리프 노드 스트림을 구성하는 리프 노드의 레이블링 값 중에서 브랜칭 노드 스트림을 구성하는 브랜칭 노드의 레이블링 값을 포함하지 않는 리프 노드의 레이블링 값을 추출함으로써, 최종 리프 노드를 생성한다.
- <134> 즉, 확장된 듀웨이 레이블링은 리프 노드의 레이블링 값을 알 수 있으면, 상위 노드의 구조정보를 파악할 수 있기 때문에, 트위그 쿼리의 리프 노드에 대응되는 XML 문서의 리프 노드를 탐색하는 것이 아주 중요하다.
- <135> 그러나, 모든 리프 노드를 탐색할 경우에는 쿼리의 브랜칭 노드와 관련이 없는 리프 노드까지 탐색될 수 있으므로, 본 발명에서는 탐색되는 리프 노드의 갯수를 최대한 줄이기 위하여 쿼리의 리프 노드에 대응되는 XML 문서

의 리프 노드 뿐만 아니라, 쿼리의 브랜칭 노드에 대응되는 XML문서의 브랜칭 노드까지 탐색한다.

- <136> 상세하게는 상기 최종 리프 노드 생성부(850)는 최종 리프 노드 스트림을 생성하기 위하여, 브랜칭 노드 스트림과 상기 리프 노드 스트림을 매칭하고, 리프 노드 스트림을 구성하는 각각의 리프 노드의 레이블링 값이 상기 브랜칭 노드 스트림 중 어느 하나의 브랜칭 노드의 레이블링 값을 포함하는지의 여부를 판단하여, 리프 노드의 레이블링 값이 어느 하나의 브랜칭 노드의 레이블링 값을 포함하지 않으면 상기 리프 노드의 레이블링 값을 상기 리프 노드 스트림에서 제거함으로써 최종 리프 노드 스트림을 생성한다.
- <137> 여기서, 브랜칭 노드 스트림과 상기 리프 노드 스트림의 매칭은 매칭 모듈(미도시)에 의해 수행될 수 있고, 리프 노드 스트림을 구성하는 각각의 리프 노드의 레이블링 값이 상기 브랜칭 노드 스트림 중 어느 하나의 브랜칭 노드의 레이블링 값을 포함하는지의 여부를 판단하여, 리프 노드의 레이블링 값이 어느 하나의 브랜칭 노드의 레이블링 값을 포함하지 않으면 상기 리프 노드의 레이블링 값을 상기 리프 노드 스트림에서 제거하는 스트림 정제 모듈(미도시)을 포함할 수 있다.
- <138> 즉, 본 발명의 특징 중의 하나는 전술한 바와 같이 클라이언트의 트위그 쿼리에 대하여 모든 노드를 탐색하는 것이 아니라 우선적으로 트위그 쿼리의 최상위에 위치한 트위그 노드를 기반으로 XML 문서의 브랜치 노들 탐색하는 것이 가장 탐색 속도가 빠르다는 사실에 그 기반을 두고 있다.
- <139> 왜냐하면, 모든 노드를 탐색하는 것보다는 쿼리의 출발점인 브랜칭 노드를 먼저 탐색하고, 그 하위의 노드를 탐색하는 것이 모든 노드를 탐색하는 것보다 불필요한 탐색 경로를 없앨 수 있기 때문이다.
- <140> 이와 같이, XML 문서의 브랜칭 노드들 우선 탐색하고, 탐색된 브랜칭 노드를 기반으로 브랜칭 노드 스트림을 생성하고, 상기 리프 노드 스트림을 구성하는 리프 노드 레이블링 값 중에서 상기 브랜칭 노드 스트림에 속한 브랜칭 노드의 레이블링 값을 포함하지 않는 리프 노드의 레이블링 값을 제거함으로써 불필요한 리프 노드를 제거할 수 있다. 즉, 본 발명은 모든 노드를 탐색하지 않기 위하여 브랜칭 노드와 조상 노드-자손 노드의 레이블링 값을 제거함으로써 불필요한 리프 노드 를 제거할 수 있다.
- <141> 상세히 설명하면, 확장된 듀웨이 레이블링 기법에서는 리프 노드는 상위 조상 노드의 구조를 포함하는 형식으로 레이블링되기 때문에 상위에 브랜칭 노드가 있을 경우, 브랜칭 노드의 레이블링 값을 포함하는 형식으로 레이블링 되어 있다.
- <142> 따라서, 최종 리프 노드 스트림 생성부(850)는 리프 노드 스트림을 구성하는 리프 노드 각각의 레이블링 값 중에서 브랜칭 노드 스트림을 구성하는 브랜칭 노드의 레이블링 값을 포함하지 않는 리프 노드가 있을 경우, 이 리프 노드의 레이블링 값을 제거하여 최종 리프 노드 스트림을 생성함으로써, 클라이언트의 쿼리에 대하여 XML 문서의 탐색시, 탐색하여야 할 리프 노드의 수를 줄일 수 있다.
- <143> 즉, 본 발명은 클라이언트의 트위그 쿼리에 대하여 XML 문서상에 매칭되는 트위그 노드를 우선적으로 탐색함으로써 트위그 노드가 아닌 노드를 탐색대상에서 우선적으로 제외시키고, 그 다음 트위그 노드의 하위 노드에 해당되는 리프 노드만을 리프 노드 스트림에 포함시킴으로써 상기 리프 노드 스트림 중 불필요한 리프 노드 만을 삭제하여 최종 리프 노드 스트림을 생성하므로 쿼리에 매칭되는 노드를 빨리 탐색할 수 있다.
- <144> 한편, 응답 메시지 생성부(860)는 상기 브랜칭 노드 스트림 및 최종 리프 노드 스트림을 이용하여 상기 트위그 쿼리에 대한 응답 메시지를 생성한다.
- <145> 즉, 브랜칭 노드 스트림에 포함된 노드의 레이블링 값과 최종 리프 노드 스트림에 포함된 레이블링 값을 이용하여 트위그 쿼리에 대한 응답 메시지를 생성한다.
- <146> 이는 본 발명의 모든 노드는 확장된 듀웨이 레이블링 기법으로 레이블링 되어 있으므로, 각 노드의 레이블링 값 만으로도 상위 노드의 구조를 파악할 수 있고, 이를 기반으로 클라이언트의 질의에 맞는 응답 메시지를 생성할 수 있다.
- <147> 이와 같이, 클라이언트의 질의에 맞는 응답 메시지를 생성하기 위하여 노드명 변환 모듈(미도시)을 포함할 수 있고, 상기 노드명 모듈은 최종 리프 노드 스트림을 구성하는 복수 개의 리프 노드의 레이블링 값 중 상호 모두 중복되는 중첩 레이블링 엘리먼트를 제거하고, 중첩 레이블링 엘리먼트가 제거된 상기 리프 노드의 잔여 레이블링 엘리먼트를 XML 문서상의 노드명으로 변환한다.
- <148> 여기서, 응답 메시지는 상기 변환된 노드명을 이용하여 상기 클라이언트의 트위그 쿼리에 대응되는 구조로 형성된 XML 응답 메시지일 수 있다.

- <149> 본 발명은 확장된 듀웨이 레이블링 기법에 따라 모든 노드가 레이블링되는데, 전술한 바와 같이, 각 레이블링 값은 상위 노드의 구조를 모두 포함하는 값으로 레이블링된다.
- <150> 이 때문에 레이블링 값을 노드명으로 변환하여 노드 구조를 형성하는 과정에서, 노드명으로 변환되는 변환량을 줄일 수 있기 때문에 클라이언트로의 응답메시지를 빨리 생성할 수 있다.
- <151> 즉, 확장된 듀웨이 레이블링 기법에 따라 레이블링된 브랜칭 노드와 리프 노드에 있어서, 리프 노드의 레이블링 값을 모두 노드명으로 변환하면, 레이블링 엘리먼트 중 반드시 중첩되는 노드명이 발생한다.
- <152> 따라서, 본 발명에서는 가급적 노드명으로 변환하는 과정을 줄여 클라이언트로의 응답시간인 전체 프로세싱 타임을 줄일 수 있으므로 보다 빠른 응답시간을 제공한다.
- <153> 도 9는 도 8의 매칭 노드 스트림 생성부(840)에서 생성하는 매칭 노드 스트림을 도시한 것이다.
- <154> 상기 도 8의 매칭 노드 스트림 생성부(840)는 상기 도 2와 같은 클라이언트의 쿼리에 대하여 클라이언트의 트위그 쿼리와 매칭되는 상기 XML 문서의 브랜칭 노드 스트림 및 리프 노드 스트림을 생성하는 것을 도시하고 있다.
- <155> 클라이언트의 트위그 쿼리(910)에 대하여 브랜칭 노드 A에 대응되는 XML 문서의 노드는 a1 및 a2로, 각각의 레이블링 값은 (0, 0, 0)과 (0, 0, 1)이다.
- <156> 그리고, 트위그 쿼리(910)에 대하여 리프 노드 D에 대응되는 XML 문서의 노드는 d1 및 d2이고, 이의 레이블링 값은 (0, 0, 0, 0)과 (0, 0, 1, 0)이다.
- <157> 한편, 트위그 쿼리(910)에 대하여 리프 노드 C에 대응되는 XML 문서의 노드는 c1 및 c2이고, 이의 레이블링 값은 (0, 0, 0, 1, 0), (0, 0, 1, 1, 0)이다.
- <158> 이를 기반으로 트위그 쿼리 브랜칭 노드 A에 대한 매칭 노드 스트림 생성부에서 생성하는 XML 문서의 브랜칭 노드 스트림 BNS_A 는 0.0.0, 0.0.1이고, 트위그 쿼리 리프 노드 D에 대한 매칭 노드 스트림 생성부에서 생성하는 XML 문서의 리프 노드 스트림 LNS_D 는 0.0.0.0, 0.0.1.0 이 된다.
- <159> 아울러, 트위그 쿼리 리프 노드 C에 대한 매칭 노드 스트림 생성부에서 생성하는 XML 문서의 리프 노드 스트림 LNS_C 는 0.0.0.1.0, 0.0.1.1.0 이 된다.
- <160> 도 10은 도 8의 응답 메시지 생성부(860)에서 수행되는 노드명 변환 과정을 도시한 것이다.
- <161> 응답 메시지 생성부(860)는 전술한 바와 같이, 브랜칭 노드 스트림 및 최종 리프 노드 스트림을 이용하여 상기 트위그 쿼리에 대한 응답 메시지를 생성한다.
- <162> 도 10에서 브랜칭 노드 스트림 BNS_A 와 트위그 쿼리 리프 노드 D에 대하여 생성된 최종 리프 노드 스트림 LNS_D 와 트위그 쿼리 리프 노드 C에 대하여 생성된 최종 리프 노드 스트림 LNS_C 가 도 10에 도시된 바와 같을 때, 이와 같은 노드 스트림을 기반으로 노드명으로 변환 연산한 노드명 테이블을 도시하고 있다.
- <163> 즉, 본 발명에서는 최종 리프 노드 스트림을 구성하는 복수 개의 리프 노드의 레이블링 값 중 상호 모두 중복되는 중첩 레이블링 엘리먼트를 제거하여 최적화된 리프 노드 스트림(Leaf Node Stream:LNS)을 생성하고, 중첩 레이블링 엘리먼트가 제거된 상기 리프 노드의 잔여 레이블링 엘리먼트를 포함하는 상기 최적화된 리프 노드 스트림을 XML 문서상의 노드명으로 변환한다.
- <164> 도 10을 참조하면, 리프 노드 C와 리프 노드 D모두가 중첩되는 중첩 레이블링 엘리먼트는 앞 부분의 2개의 0이므로, 이를 제거하여 최적화된 리프 노드 스트림을 생성하고 있음을 알 수 있다.
- <165> 이와 같이 최적화된 리프 노드 스트림 OLS_C 및 OLS_D 을 기반으로 XML 문서상의 노드명으로 변환한다.
- <166> 도 11a 내지 도 11c는 각 XML 문서에 있어서 종래의 트위그 질의 처리 기법과 본 발명에 따른 XML 문서의 트위그 쿼리의 질의 처리 방법의 성능을 비교한 그래프이다.

- <167> 도 11a는 DBLP라는 XML 문서에 대하여 종래의 트위그 쿼리의 질의 처리 방법 중 TJ Fast(1110), 트위그 스택 기법(1120)과 본 발명에 따른 XML 문서의 트위그 쿼리의 질의 처리 방법(1130)의 성능을 비교한 그래프이다.
- <168> 여기서, Q1은 트위그 쿼리가 //dblp/inproceedings[title]/author일 경우, 이에 대한 전체 응답 처리 시간(Execution time)을 비교한 것이고, Q2는 트위그 쿼리가 //dblp/article[author][.//title]/year일 경우, 이에 대한 전체 응답 처리 시간을 비교한 것이며, Q3는 트위그 쿼리가 //inproceedings[author][.//title]/booktitle일 경우의 전체 응답 처리 시간을 비교한 것이다.
- <169> 도 11a에서 트위그 쿼리 Q1, Q2, Q3 모두에 대하여 종래의 트위그 쿼리의 질의 처리 방법 중 TJ Fast(1110), 트위그 스택 기법(1120)보다 XML 문서의 트위그 쿼리의 질의 처리 방법(1130)이 25% 내지 50% 정도의 성능 향상을 보이고 있음을 알 수 있다.
- <170> 도 11b는 TreeBank라는 XML 문서에 대하여 종래의 트위그 쿼리의 질의 처리 방법 중 TJ Fast(1110), 트위그 스택 기법(1120)과 본 발명에 따른 XML 문서의 트위그 쿼리의 질의 처리 방법(1130)의 성능을 비교한 그래프이다.
- <171> 여기서, Q1은 트위그 쿼리가 //S/VP/PP[.//NP/VBN]/N일 경우, 이에 대한 전체 응답 처리 시간(Execution time)을 비교한 것이고, Q2는 트위그 쿼리가 //S/VP/PP[IN]/NP/VBN일 경우, 이에 대한 전체 응답 처리 시간을 비교한 것이며, Q3는 트위그 쿼리가 //VP[DT]/PRP_DOLLAR_일 경우의 전체 응답 처리 시간을 비교한 것이다.
- <172> 도 11b에서 트위그 쿼리 Q1, Q2, Q3 모두에 대하여 종래의 트위그 쿼리의 질의 처리 방법 중 TJ Fast(1110), 트위그 스택 기법(1120)보다 XML 문서의 트위그 쿼리의 질의 처리 방법(1130)이 20% 내지 80% 정도의 성능 향상을 보이고 있음을 알 수 있다.
- <173> 도 11c는 XMark라는 XML 문서에 대하여 종래의 트위그 쿼리의 질의 처리 방법 중 TJ Fast(1110), 트위그 스택 기법(1120)과 본 발명에 따른 XML 문서의 트위그 쿼리의 질의 처리 방법(1130)의 성능을 비교한 그래프이다.
- <174> 여기서, Q1은 트위그 쿼리가 //site/open_auctions[.//bidder/personref]/reserve일 경우, 이에 대한 전체 응답 처리 시간(Execution time)을 비교한 것이고, Q2는 트위그 쿼리가 //people/person[.//address/zipcode]/profile/education일 경우, 이에 대한 전체 응답 처리 시간을 비교한 것이며, Q3는 트위그 쿼리가 //item[location]/description/keyword일 경우의 전체 응답 처리 시간을 비교한 것이다.
- <175> 도 11c에서 트위그 쿼리 Q1, Q2, Q3 모두에 대하여 종래의 트위그 쿼리의 질의 처리 방법 중 TJ Fast(1110), 트위그 스택 기법(1120)보다 XML 문서의 트위그 쿼리의 질의 처리 방법(1130)이 10% 내지 70% 정도의 성능 향상을 보이고 있음을 알 수 있다.
- <176> 상기 도 11a 내지 도 11c에서 볼 수 있는 바와 같이, 본 발명은 클라이언트의 트위그 쿼리에 대하여 트위그 쿼리의 리프 노드에 매칭되어지는 XML 문서상의 노드의 탐색 시간을 줄임으로써 리프 노드 탐색 시간인 I/O 타임을 줄일 수 있으며, 트위그 쿼리에 대하여 탐색된 노드에 대하여 리프 노드에 레이블링된 값을 원래 노드 이름으로 변환하는 할 때, 트위그 쿼리에 대응되는 브랜칭 노드와 중첩되는 레이블링된 값을 원래 노드 이름으로 변환하지 않아 중간 프로세싱 타임이 줄어들고, 전체적으로 쿼리에 대하여 응답하는 전체 프로세싱 타임이 줄어 신속한 질의 응답 처리를 수행할 수 있음을 다시 한번 확인 할 수 있다.
- <177> 본 발명에 따른 XML 문서의 트위그 쿼리의 질의 처리 방법은 또한, 상기 XML 문서의 트위그 쿼리의 질의 처리를 컴퓨터로 수행할 수 있는 컴퓨터로 판독 가능한 저장매체에 프로그램 코드로서 구현하는 것이 가능하다.
- <178> 상기 컴퓨터 판독가능 저장매체는 컴퓨터 시스템에 의하여 읽혀질 수 있는 데이터가 기록되는 모든 종류의 저장장치를 포함한다.
- <179> 컴퓨터 판독가능 저장매체의 예로는 ROM, RAM, CD/DVD-ROM, 블루레이, 자기 테이프, 플로피디스크, 광데이터 저장장치 등이 있으며, 또한 캐리어 웨이브(예를 들어 인터넷을 통한 전송)의 형태로 구현되는 것도 포함한다. 또한 컴퓨터 판독가능 저장매체는 네트워크로 연결된 컴퓨터 시스템에 분산되어 분산방식으로 컴퓨터가 읽을 수 있는 코드가 저장되고 실행될 수 있다.
- <180> 이제까지 본 발명에 대하여 그 바람직한 실시예를 중심으로 살펴보았다. 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자라면 본 발명이 본 발명의 본질적인 특성에서 벗어나지 않는 범위에서 변형된 형태로 구현될 수 있음을 이해할 수 있을 것이다.
- <181> 그러므로 개시된 실시예들은 한정적인 관점이 아니라 설명적인 관점에서 고려되어야 한다. 본 발명의 범위는 전

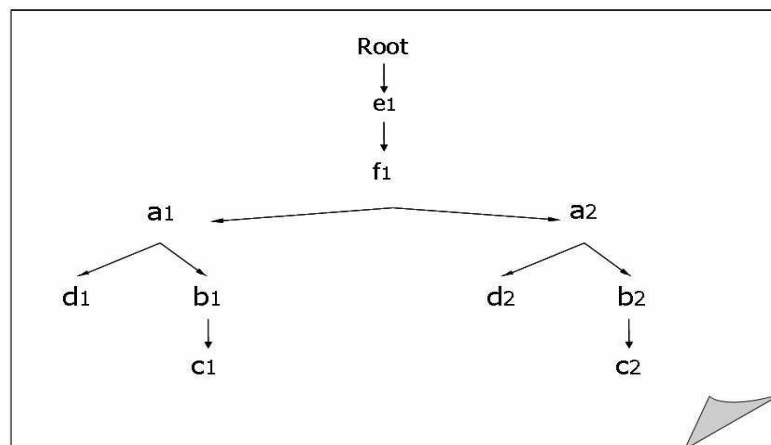
술한 설명이 아니라 특허청구범위에 나타나 있으며, 그와 균등한 범위 내에 있는 모든 차이점은 본 발명에 포함 된 것으로 해석되어야 할 것이다.

도면의 간단한 설명

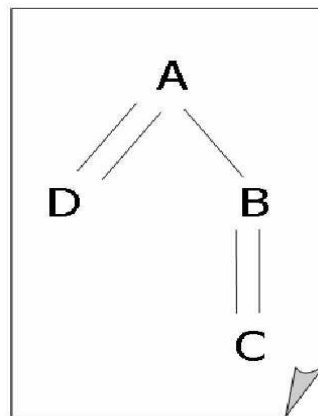
- <182> 도 1은 XML 문서의 일반적인 예를 도시한 것이다.
- <183> 도 2는 XML 문서에 대한 클라이언트의 특정 쿼리(query)의 구조를 도시한 것이다.
- <184> 도 3은 일반적인 지역 인코딩(Region encoding) 기반의 레이블링 기법을 도시한 것이다.
- <185> 도 4는 XML 문서 탐색에 있어서의 TJ Fast기법의 일례를 도시한 것이다.
- <186> 도 5는 종래의 지역 인코딩 기반으로 레이블링된 XML 문서에서의 트위그 스택(TwigStack) 탐색 기법의 일 실시 예를 도시한 것이다.
- <187> 도 6은 종래의 지역 인코딩 기반으로 레이블링된 XML 문서에서의 트위그 스퀘어 스택(Twig²Stack) 탐색 기법의 일 실시예를 도시한 것이다.
- <188> 도 7은 본 발명의 일 실시예에 따른 XML 문서의 트위그 쿼리의 질의 처리 방법의 흐름도이다.
- <189> 도 8은 본 발명의 일 실시예에 따른 XML 문서의 트위그 쿼리의 질의 처리 장치의 블록도이다.
- <190> 도 9는 도 8의 매칭 노드 스트림 생성부(840)에서 생성하는 매칭 노드 스트림을 도시한 것이다.
- <191> 도 10은 도 8의 응답 메시지 생성부(860)에서 수행되는 노드명 변환 과정을 도시한 것이다.
- <192> 도 11a 내지 도 11c는 각 XML 문서에 있어서 종래의 트위그 질의 처리 기법과 본 발명에 따른 XML 문서의 트위그 쿼리의 질의 처리 방법의 성능을 비교한 그래프이다.

도면

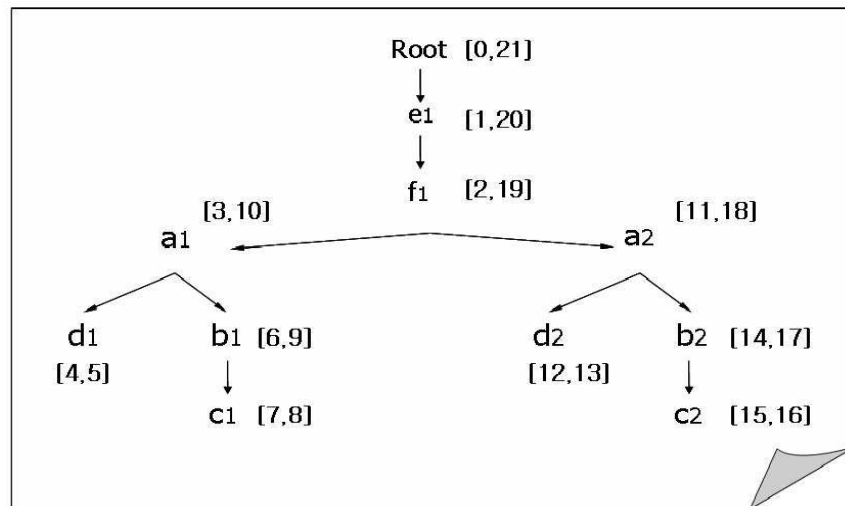
도면1



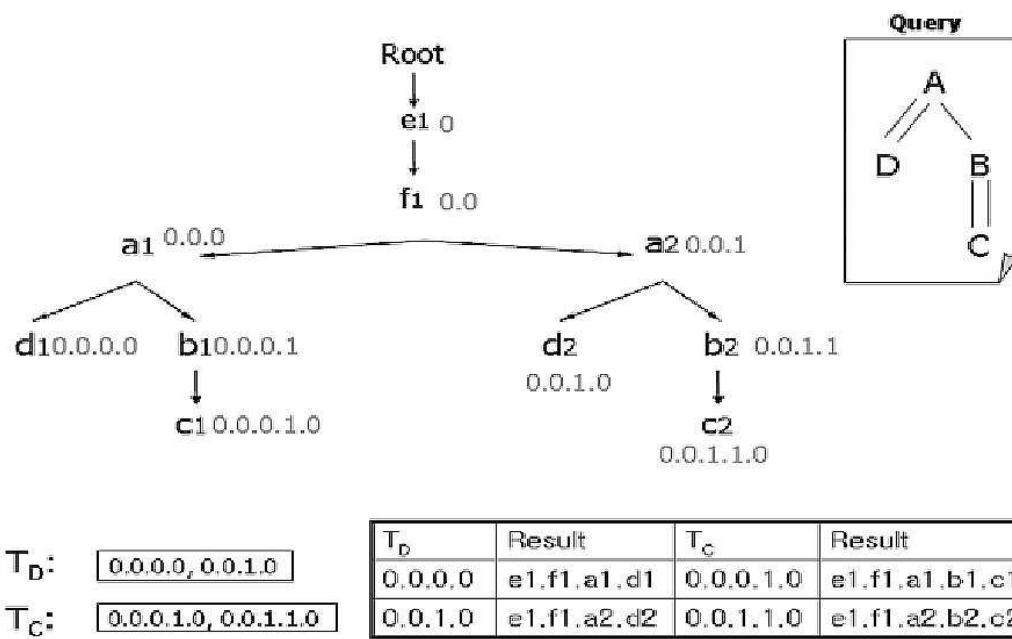
도면2



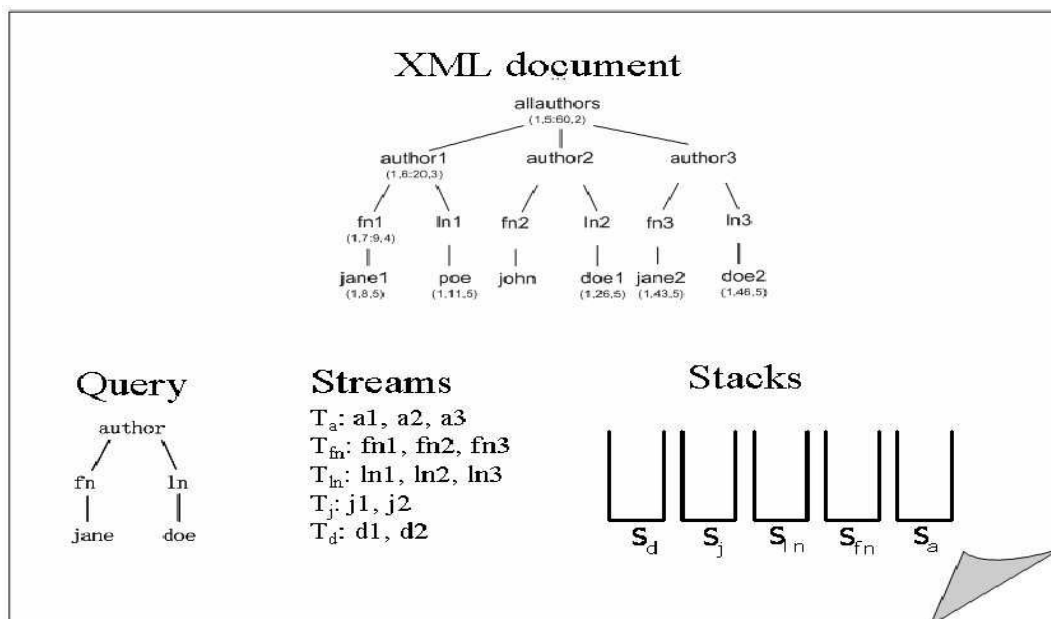
도면3



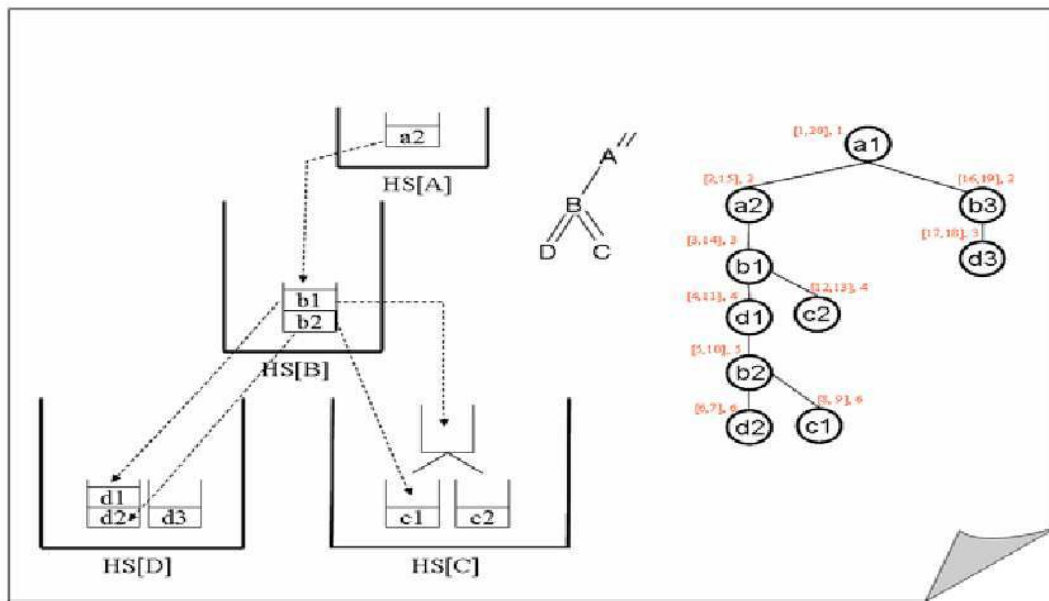
도면4



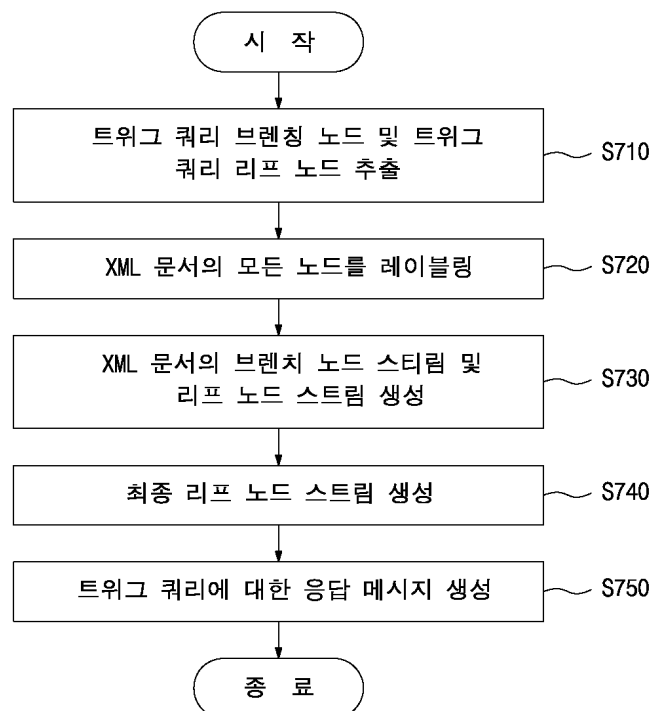
도면5



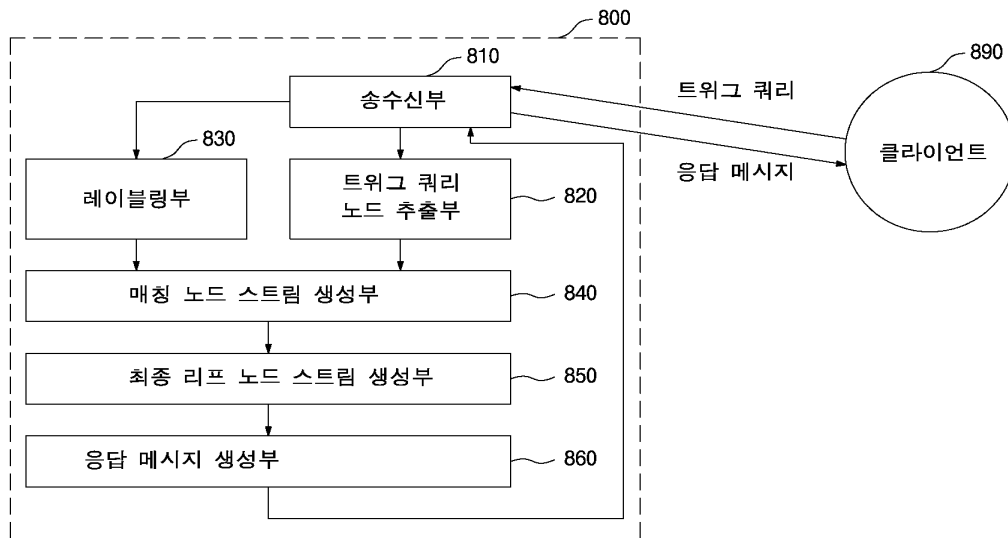
도면6



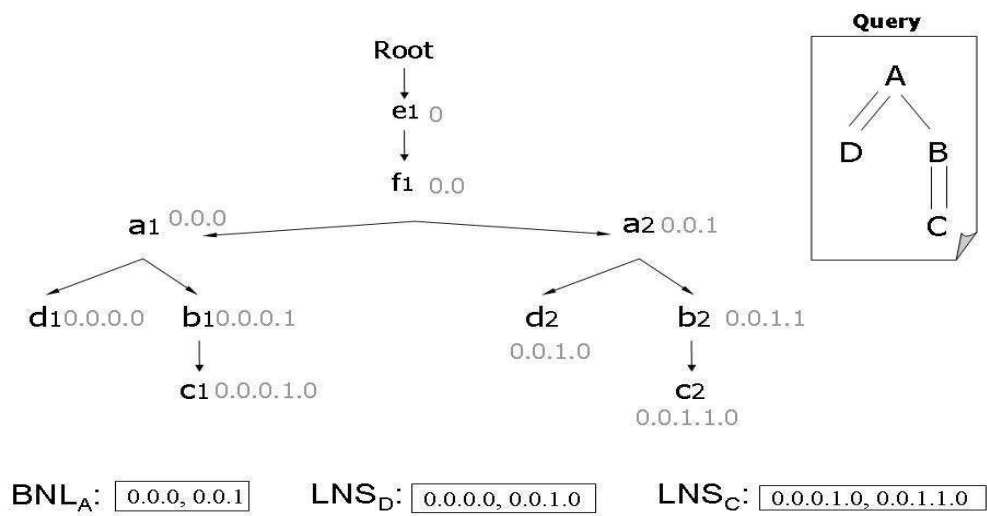
도면7



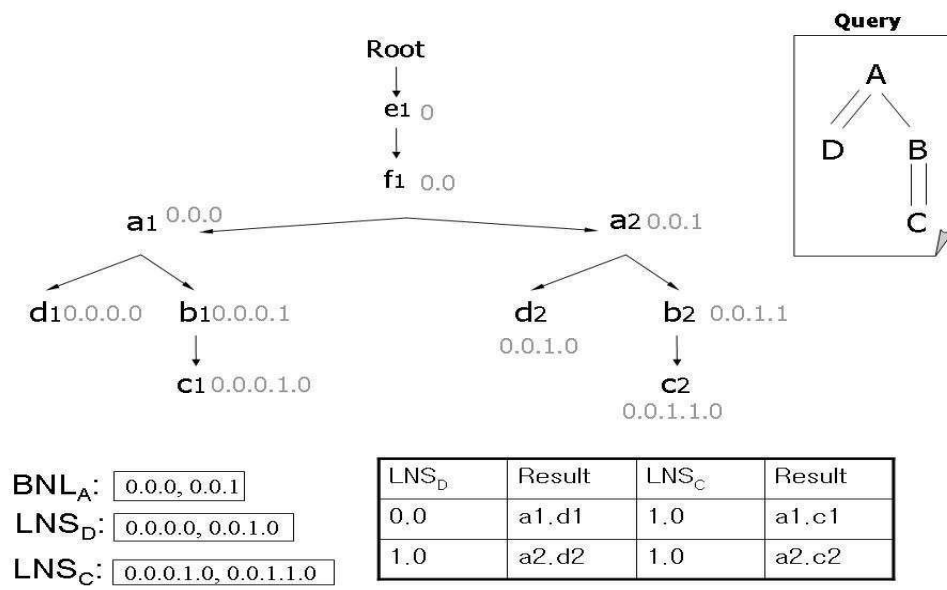
도면8



도면9



도면10



도면11a

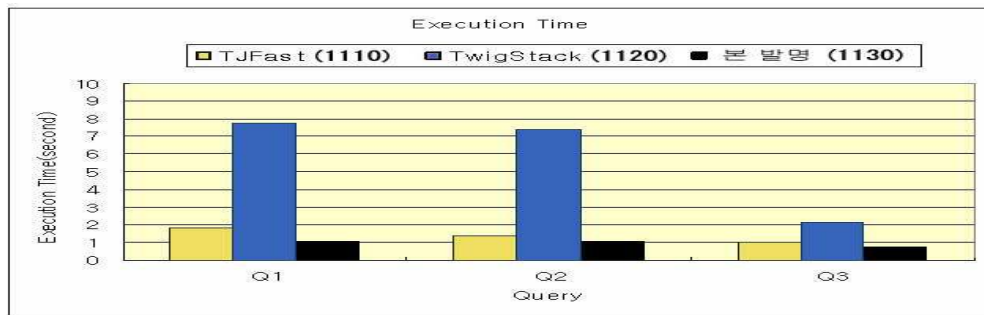


Q1 : //dblp/inproceedings[title]/author

Q2 : //dblp/article[author] [../title]//year

Q3 : //inproceedings[author] [../title]//booktitle

도면11b

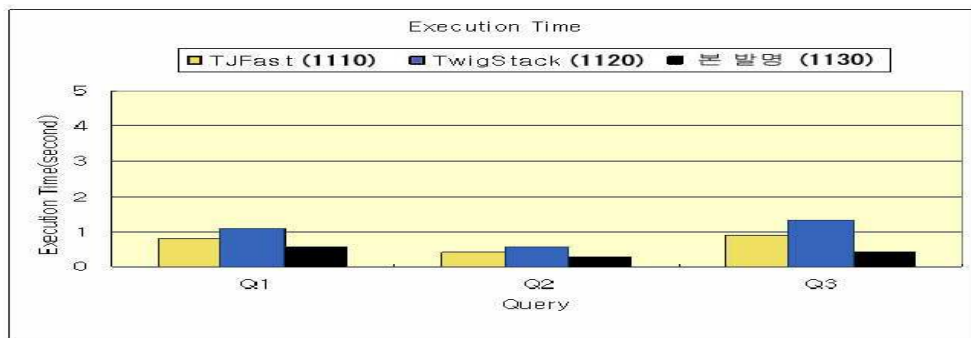


Q1 : //S/NP//PP[./NP/VBN]/IN

Q2 : //S/NP/PP[IN]/NP/VBN

Q3 : //NP[DT]//PRP_DOLLAR_

도면11c



Q1 : /site/open_auctions[./bidder/personref]//reserve

Q2 : //people/person[./address/zipcode]/profile/education

Q3 : //item[location]/description//keyword