

	(19) 대한민국특허청(KR) (12) 공개특허공보(A)	(11) 공개번호 10-2011-0071651 (43) 공개일자 2011년06월29일
(51) Int. Cl. <i>H04N 7/173</i> (2011.01) <i>G06F 17/21</i> (2006.01) (21) 출원번호 10-2009-0128275 (22) 출원일자 2009년12월21일 심사청구일자 2009년12월21일	(71) 출원인 고려대학교 산학협력단 서울 성북구 안암동5가 1 (72) 발명자 정연돈 서울특별시 중구 황학동 롯데캐슬 베네치아 105-812 박준표 경기 의정부시 녹양동 341-8 박창섭 서울특별시 성북구 월곡동 23-1 (74) 대리인 조영현, 나승택	

전체 청구항 수 : 총 10 항

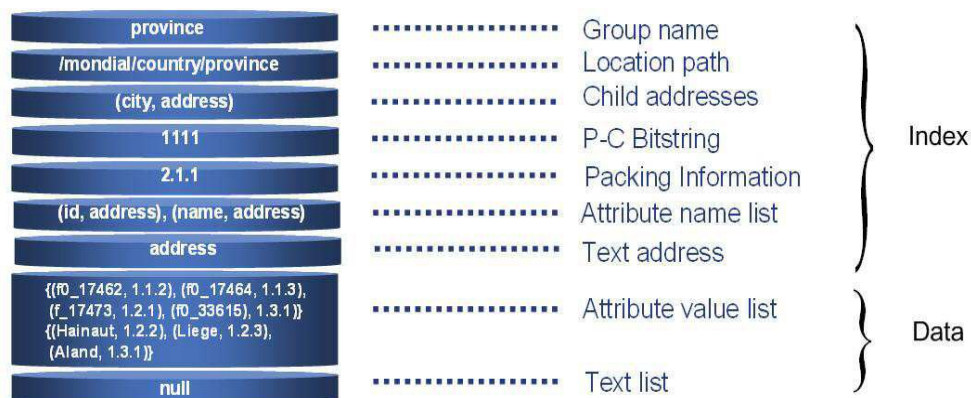
(54) 무선 방송 스트림에서 XML 질의 처리 방법

(57) 요약

본 발명은 XML 질의 처리 방법에 관한 것으로, 더욱 상세하게는 구조적인 요약과 속성 요약을 통해 방송 스트림의 크기를 최소화하고, P-C 비트열(P-C Bitstring)과 압축 정보(Packing Information)을 이용하여 노드간 부모-자식 관계를 유지함으로써 무선 방송 환경에서 분기 질의 처리를 지원하고, 속성과 텍스트에 대한 선택적인 수신을 통한 무선 방송 스트림에서 효율적인 XML 질의 처리 방법에 관한 것이다.

상기와 같은 본 발명은 분기 질의를 포함한 다양한 유형의 질의 처리에 있어 접근 시간, 튜닝시간이 획기적으로 단축되고, 이로 인해 에너지 효율이 탁월한 특징이 있다.

대 표 도 - 도3



이 발명을 지원한 국가연구개발사업

과제고유번호 20090083893

부처명 한국과학재단

연구관리전문기관

연구사업명 중견연구자지원_핵심연구

연구과제명 RIN형성과 분석 방법론 연구:잠재연구정보 및 연구진화방향의 효과적 추출

기여율

주관기관 고려대학교

연구기간 2009년 09월 01일 ~ 2010년 08월 31일

특허청구의 범위

청구항 1

구조적 요약 및 속성 요약을 통해 그룹 노드를 생성하고, 상기 그룹 노드에 부모-자식 관계를 표현하는 P-C 비트열(P-C Bitstring)과 압축 정보(Packing Information)를 포함하는 방송 스트림을 생성하는 단계와;

질의가 주어지면 질의를 분석하여 질의 트리를 생성하고 상기 방송 스트림을 선택적으로 수신하여 질의를 처리하는 단계를 포함하는 무선 방송 스트림에서 XML 질의 처리 방법.

청구항 2

제 1항에 있어서,

상기 속성 요약은

상기 방송 스트림 내의 중복된 태그 이름과 속성 이름을 생략하고 속성 값과 상기 속성 값을 가진 엘리먼트의 순서 정보만 저장하여 방송 스트림의 크기를 감소시키는 것을 특징으로 하는 무선 방송 스트림에서 XML 질의 처리 방법.

청구항 3

제 1항에 있어서,

상기 그룹 노드는

그룹 이름(Group name), 위치 경로(Location path), 자식노드 주소(Child addresses), P-C 비트열(P-C Bitstring), 압축 정보(Packing Information), 속성 이름 리스트(Attribute name list), 텍스트 주소(Text addresses)를 포함하는 인덱스와;

속성 값 리스트(Attribute value list)와 텍스트 리스트(Text list)를 포함하는 데이터를 포함하는 것을 특징으로 하는 무선 방송 스트림에서 XML 질의 처리 방법.

여기서, 상기 자식노드 주소는 상기 그룹 노드의 자식 그룹 노드들에 대한 주소 정보이며, 상기 P-C 비트열은 상기 그룹 노드의 부모 그룹 노드의 엘리먼트의 자식 엘리먼트로 존재하는지 여부를 0 또는 1의 연속된 비트열로 표현한 정보이며, 상기 압축 정보는 하나의 부모 엘리먼트에 소속된 자식 엘리먼트의 갯수를 나타낸 정보이며, 상기 속성 이름 리스트는 속성의 이름과 상기 속성의 값들이 연속적으로 저장된 스트림의 시작 주소 정보를 포함하고, 상기 텍스트 주소는 상기 그룹 노드에 속하는 엘리먼트의 텍스트들이 연속적으로 저장된 스트림의 시작 주소이며, 상기 속성 값 리스트는 상기 그룹 노드에 병합된 엘리먼트의 속성 값과 듀이 순서가 속성 이름별로 연속적으로 저장된 리스트이고, 상기 텍스트 리스트는 상기 그룹 노드에 병합된 엘리먼트의 텍스트들이 듀이 순서와 함께 그룹화되어 저장된 리스트이다.

청구항 4

제 3항에 있어서,

상기 질의를 처리하는 단계는

상기 질의가 분기 조건이 포함되지 않은 단순 경로 질의인 경우

질의 트리를 생성하고 상기 그룹 노드의 인덱스를 수신하여 상기 자식 주소에서 다음 위치의 그룹 노드를 찾는 방식으로 최상위 노드부터 하향식으로 질의 트리를 탐색하는 단계와;

상기 질의 트리의 마지막 노드까지 탐색한 후 질의 트리의 마지막 노드에 해당하는 그룹 노드의 정보를 질의의 결과로 반환하는 것을 특징으로 하는 무선 방송 스트림에서 XML 질의 처리 방법.

청구항 5

제 4항에 있어서,

상기 질의 트리를 탐색하는 단계는

사용자의 질의에 조건식이 포함된 경우

상기 속성 요약된 그룹 노드 중 속성 이름 리스트에 포함된 속성의 이름과 상기 속성의 값들이 연속적으로 저장된 스트림의 시작 주소 정보를 통해 주어진 조건식에 포함된 속성 값이나 텍스트만을 선택적으로 수신하여 상기 조건식을 만족하는 듀이 순서를 질의 트리에 저장하는 것을 특징으로 하는 무선 방송 스트림에서 XML 질의 처리 방법.

청구항 6

제 5항에 있어서,

상기 질의 트리에 조건식에 의해 저장된 듀이 순서가 있을 경우

상기 듀이 순서에 해당하는 정보만 질의의 결과로 반환하는 것을 특징으로 하는 무선 방송 스트림에서 XML 질의 처리 방법.

청구항 7

제 3항에 있어서,

상기 질의를 처리하는 단계는

상기 질의에 분기 조건이 포함된 경우

질의 트리를 생성하고 상기 질의 트리에 대한 경로를 탐색하는 단계와;

상기 분기 조건을 만족하는 엘리먼트를 선별하기 위해 P-C 비트열 일치 여부를 확인하는 단계와;

상기 선별된 P-C 비트열을 말단 그룹 노드에 적용하여 질의 결과를 도출하는 단계를 포함하는 것을 특징으로 하는 무선 방송 스트림에서 XML 질의 처리 방법.

청구항 8

제 7항에 있어서,

상기 경로를 탐색하는 단계는

찾고자 하는 데이터가 포함된 경로인 주 경로에 대한 탐색과 주어진 분기 조건을 포함하는 부 경로에 대한 탐색을 포함하고;

상기 주 경로와 부 경로에 해당하는 그룹 노드들을 수신하여 P-C 비트열 및 압축 정보를 질의 트리의 각 노드에 저장하는 것을 특징으로 하는 무선 방송 스트림에서 XML 질의 처리 방법.

청구항 9

제 8항에 있어서,

상기 P-C 비트열 일치 여부를 확인하는 단계는

상기 경로를 탐색하는 단계를 통해 질의 트리의 각 노드에 저장된 P-C 비트열 및 압축 정보를 이용하여 상기 분

기조건을 만족하는 엘리먼트를 선별하되;

상기 질의 트리에서 주 경로와 부 경로의 말단 노드부터 분기 노드까지 올라가면서 P-C 비트열을 압축 정보에 따라 압축하여 bit-wise AND 연산을 수행하고, 상기 질의트리의 분기 노드에서는 상기 주 경로와 부 경로에서 계산된 비트열을 bit-wise AND 연산하여 분기 조건을 만족하는 중간 결과 값을 산출하는 것을 특징으로 하는 무선 방송 스트림에서 XML 질의 처리 방법.

청구항 10

제 9항에 있어서,

상기 질의 결과를 도출하는 단계는

상기 분기 노드에서 계산된 중간 결과 값을 질의 트리에서 주 경로를 따라 내려가면서 원래의 비트열로 복원하고, 상기 주 경로의 말단 노드까지 복원된 비트열이 말단 그룹 노드의 데이터에 적용되어 비트열에서 1에 해당하는 엘리먼트의 내용만 질의 결과로 반환하는 것을 특징으로 하는 무선 방송 스트림에서 XML 질의 처리 방법.

명세서

발명의 상세한 설명

기술 분야

[0001] 본 발명은 XML 질의 처리 방법에 관한 것으로, 더욱 상세하게는 구조적인 요약과 속성 요약을 통해 방송 스트림의 크기를 최소화하고, P-C 비트열(P-C Bitstring)과 압축 정보(Packing Information)을 이용하여 노드간 부모-자식 관계를 유지함으로써 무선 방송 환경에서 분기 질의 처리를 지원하고, 속성과 텍스트에 대한 선택적인 수신을 통한 무선 방송 스트림에서 효율적인 XML 질의 처리 방법에 관한 것이다.

배경 기술

[0002] 무선 통신 기술의 발달에 따라 휴대용 단말기(랩탑, PDA, 휴대폰 등)를 통해 데이터를 주고받는 무선 모바일 컴퓨팅 기술이 확산되고 있다. 무선 모바일 환경에서 무선 방송을 통한 데이터의 전달은 대역폭의 활용의 이점과 에너지 효율성, 확장성을 제공하는 효과적인 방법이다.

[0003] 무선 방송 환경에서 서버는 방송 채널을 통해 계속해서 데이터를 방송하고, 이동 사용자들은 방송 채널로부터 필요한 데이터를 수신한다. 서버 입장에서는 브로드캐스트로 데이터를 송출하므로 한정된 네트워크 대역폭을 효율적으로 활용할 수 있으며, 이동 사용자의 수가 증가하더라도 추가적인 비용이 발생하지 않는다는 장점이 있다. 또한, 이동 사용자의 입장에서는 에너지 소모가 많은 "송신" 작업없이 "수신"만으로 데이터에 접근할 수 있는 이점이 있다.

[0004] 무선 방송 환경에서는 이동 사용자들은 한정된 배터리 자원의 휴대용 단말기를 사용하므로 에너지 절약을 위해 데이터의 선택적인 수신에 요구된다. 선택적인 수신을 통해 이동 사용자는 동작모드(active mode)에서 데이터를 수신하고 필요한 데이터가 방송되지 않을 때는 대기 모드(doze mode)에서 필요한 데이터의 방송시점까지 대기함으로써 에너지를 효율적으로 이용할 수 있다. 또한, 이동 사용자에게 빠른 응답을 제공하기 위해 전체 질의 처리 시간이 최소화되어야 한다.

[0005] 일반적으로 상기 전자를 에너지 효율성, 후자를 접근시간 효율성이라 일컫는다. 방송 기법을 통한 데이터 전달에 있어서 이동 사용자의 에너지 효율성과 접근시간 효율성을 측정하기 위해 각각 튜닝시간(tuning time)과 접근시간(access time)이 성능 측정의 척도로 활용된다.

[0006] 상기 접근시간은 이동 사용자의 질의가 시작한 시점부터 모든 필요한 데이터를 수신받은 시점까지의 소요시간을 의미하고, 상기 튜닝시간은 이동 사용자가 방송 채널에서 실제로 데이터를 읽은 시간들의 합을 의미한다.

[0007] 일반적으로 무선 방송 환경에서 XML(eXtensible Markup Language)가 정보의 표현 및 교환의 표준으로서 널리

활용되고 있으며, 이에 따라 XML 문서에서 효율적인 질의처리가 중요한 이슈로 떠오르고 있다.

[0008] XML 문서는 엘리먼트들이 계층적으로 내포되어 있는 준-구조적인 언어로서 루트(Root)로부터 시작하여 순서가 있고 레이블이 있는 트리로 표현될 수 있다. XML 문서에서 엘리먼트(element), 속성(attribute), 텍스트(text)는 각각 트리에서 노드(node)로 표현되며, 부모-자식 관계는 각 노드간의 간선으로 표현될 수 있다.

[0009] 도 1a는 종래의 XML 문서를 도시한 것이다. 상기 도 1a와 같은 XML 문서는 도 1b와 같이 트리 형태로 표현될 수 있다.

[0010] 도 1b에서 실선인 타원으로 표현된 노드는 XML 문서의 엘리먼트를 의미하고 각 노드 사이에 간선은 부모-자식 관계를 의미한다. 타원 안에 포함된 숫자는 각 노드의 듀이 순서를 나타낸다. XML 문서에서 듀이 순서는 문서의 루트에서부터 그 노드까지의 경로를 표현하는 백터로, 각 노드는 유일한 듀이 순서를 가진다. 속성과 텍스트는 설명의 편의를 위해 생략하였다.

[0011] 한편 XML 문서에 대한 질의 표현은 XPath나 XQuery와 같은 경로 표현식 질의 언어를 사용한다. 예를 들어, 도 1a의 XML 문서에서 "체코(Czech Republic)에 소재한 도시를 검색하라"는 질의는 XPath 경로 표현식 질의를 사용하여 다음과 같이 표현할 수 있다.

[0012]

/mondial/country[@name="Czech Republic"]/province/city

[0013] XPath 질의 처리는 위치 경로(location path)의 탐색으로 이루어지고, 상기 위치 경로의 탐색은 질의를 구성하는 각 위치 단계들의 처리로 이루어지고, 상기 위치 단계의 처리는 질의에서 각 구성요소(term)에 해당하는 축(axis)이나, 노드 또는 조건식을 검사하는 것을 의미한다.

[0014] 예를 들어, 상기 질의는 4개의 위치 단계(즉, modial,/country[@name="Czech Republic"],/province,/city)로 구성되어 있다. 따라서, 첫번째 위치 단계 처리는 루트인 "modial" 노드를 탐색하는 것이고, 두번째 위치 단계 처리는 "modial" 노드의 자식 노드들 중에서 "name" 속성 값이 "Czech Republic"인 "country" 노드를 탐색하는 것이고, 세번째 위치 단계 처리는 상기 두번째 위치 단계를 만족하는 "country" 노드의 자식들 중에서 "province" 노드를 탐색하는 것이고, 마지막 위치 단계는 질의의 말단 노드인 "city"노드를 탐색하는 것이다.

[0015] 이하, 분기 질의에 대해 간단히 살펴보면, 분기 질의(Twig Query)란 트리 구조의 XML 데이터에서 여러 엘리먼트에 대한 조건식을 동시에 만족하는 패턴을 찾는 질의 표현이다.

[0016] 상기 분기 질의는 엘리먼트 태그들, 속성 또는 텍스트에 대한 조건식을 포함하는 둘 이상의 부모-자식 관계, 조상-후손 관계로 구성된다. 예를 들어, 도 1a의 예제 문서에서 국가의 이름이 명시된 주(province)에 소재한 도시를 검색하는 분기 질의는 다음과 같이 표현할 수 있다.

[0017]

//country[name]/province/city

[0018] 마찬가지로, 국가의 이름이 명시된 주(province) 중에서 "Liege"에 소재한 도시를 검색하는 질의는 다음과 같이 표현된다.

[0019]

//country[name]/province[@name="Liege"]/city

[0020] 도 2는 상기와 같은 분기 질의를 트리 형태로 표현한 것이고, 도 2에서 점선으로 표시된 간선과 노드는 속성을 의미한다.

[0021] 분기 질의는 복잡한 내용을 포함할 수 있는 이점으로 인해 널리 사용되고 있는 질의 유형으로 XML 질의 처리에 있어 핵심적인 부분이다.

[0022] 지금까지 XML 분기 질의 처리 및 무선 XML 방송 기법에 대해 많은 연구가 이루어져 왔으나, 지금까지 제안되었던 XML 분기 질의처리 기법들은 무선 모바일 컴퓨팅 환경에서 중요한 척도인 에너지 및 접근시간 효율성을 고려하지 않았으며, 종래의 무선 XML 방송에 관한 연구들은 분기 조건이 없는 단순 경로 질의만을 고려하였기 때문에 분기 질의 처리를 할 수 없거나 비효율적인 문제가 있었다.

[0023] 한편, XML 질의 처리에 있어서 경로가 같은 엘리먼트들을 묶어 색인 정보의 크기를 효과적으로 줄이는 구조적인 요약과 관련 여러 기법들이 활용되고 있다. 그러나, 상기와 같은 구조적인 요약 방법은 무선 방송 환경에서

접근시간을 줄이는 데에는 커다란 효과가 있지만 엘리먼트의 순서를 보존하지 않기 때문에 조건식이 포함된 질의를 처리할 수 없으며 엘리먼트 간의 부모-자식 관계가 사라지기 때문에 조건식이 포함된 질의를 처리할 수 없으며 엘리먼트 간의 부모-자식 관계가 사라지기 때문에 분기 질의를 처리할 수 없는 문제가 있다.

발명의 내용

해결 하고자하는 과제

[0024] 상기와 같은 문제점을 해결하기 위해 안출된 것으로서 본 발명의 목적은 무선 방송 환경에서 에너지 및 접근시간 효율적인 XML 분기 질의 처리를 위해 구조적인 요약과 속성 요약을 통해 방송스트림의 크기를 최소화하고, 그룹 노드에 P-C 비트열과 압축 정보를 삽입하여 부모-자식 관계를 유지함으로써 분기 질의 처리를 지원하고, 속성과 텍스트에 대한 선택적인 색인 기법을 제공함으로써 이동 사용자의 접근 시간 및 튜닝 시간을 최소화할 수 있는 무선 방송 스트림에서 XML 질의 처리 방법을 제공하는 데 있다.

과제 해결수단

[0025] 본 발명에 따른 무선 방송 스트림에서 XML 질의 처리 방법은 하기와 같은 특징으로 요약될 수 있다.

[0026] ○ 구조적인 요약과 속성 요약

[0027] 중복된 태그 이름과 속성 이름을 생략함으로써 방송스트림의 크기를 최소화한다.

[0028] ○ 속성 값과 텍스트에 대한 선택적인 수신

[0029] 질의에 주어진 조건식에 포함된 속성 값이나 텍스트만 선택적으로 수신함으로써 이동 사용자의 에너지 효율성을 높인다. 특히, 속성 값과 텍스트를 듀이 순서와 함께 저장함으로써 조건식이 포함된 질의 처리가 불가능했던 구조적인 요약 방법의 한계를 극복한다.

[0030] ○ 분기 질의 처리를 위한 부모-자식 관계 유지 기법

[0031] 본 발명에서는 P-C 비트열(P-C Bitstring)과 압축정보(Packing Information)을 통해 부모-자식 관계를 유지함으로써 분기 질의 처리를 지원한다.

[0032] ○ 질의 트리를 이용한 직관적인 질의 처리

[0033] 주어진 질의에 대한 질의 트리를 생성하여 질의 트리에서 질의 처리 단계에 따른 스트림 수신과 직관적인 질의 처리 방법을 제공한다.

[0034] 상기와 같은 본 발명의 과제 해결 수단을 통한 XML 질의 처리 방법은 구조적 요약 및 속성 요약을 통해 그룹 노드를 생성하고, 상기 그룹 노드에 부모-자식 관계를 표현하는 P-C 비트열(P-C Bitstring)과 압축 정보(Packing Information)를 포함하는 방송 스트림을 생성하는 단계와 질의가 주어지면 질의를 분석하여 질의 트리를 생성하고 상기 방송 스트림을 선택적으로 수신하여 질의를 처리하는 단계를 포함하는 것을 특징으로 한다.

[0035] 여기서, 상기 속성 요약은 상기 방송 스트림 내의 중복된 태그 이름과 속성 이름을 생략하고 속성 값과 상기 속성 값을 가진 엘리먼트의 순서 정보만 저장하여 방송 스트림의 크기를 감소시키는 것을 특징으로 한다.

[0036] 그리고 상기 그룹 노드는 그룹 이름(Group name), 위치 경로(Location path), 자식노드 주소(Child addresses), P-C 비트열(P-C Bitstring), 압축 정보(Packing Information), 속성 이름 리스트(Attribute name list), 텍스트 주소(Text addresses)를 포함하는 인덱스와 속성 값 리스트(Attribute value list)와 텍스트 리스트(Text list)를 포함하는 데이터를 포함하는 것을 특징으로 한다.

[0037] 여기서, 상기 자식노드 주소는 상기 그룹 노드의 자식 그룹 노드들에 대한 주소 정보이며, 상기 P-C 비트열은 상기 그룹 노드의 부모 그룹 노드의 엘리먼트의 자식 엘리먼트로 존재하는지 여부를 0 또는 1의 연속된 비트열로 표현한 정보이며, 상기 압축 정보는 하나의 부모 엘리먼트에 소속된 자식 엘리먼트의 갯수를 나타낸 정보이며, 상기 속성 이름 리스트는 속성의 이름과 상기 속성의 값들이 연속적으로 저장된 스트림의 시작 주소 정보를 포함하고, 상기 텍스트 주소는 상기 그룹 노드에 속하는 엘리먼트의 텍스트들이 연속적으로 저장된 스트림의 시

작 주소이며, 상기 속성 값 리스트는 상기 그룹 노드에 병합된 엘리먼트의 속성 값과 듀이 순서가 속성 이름별로 연속적으로 저장된 리스트이고, 상기 텍스트 리스트는 상기 그룹 노드에 병합된 엘리먼트의 텍스트들이 듀이 순서와 함께 그룹화되어 저장된 리스트이다.

[0038] 또한, 상기 질의를 처리하는 단계는 상기 질의가 분기 조건이 포함되지 않은 단순 경로 질의 인 경우 질의 트리를 생성하고 상기 그룹 노드의 인덱스를 수신하여 상기 자식 주소에서 다음 위치의 그룹 노드를 찾는 방식으로 최상위 노드부터 하향식으로 질의 트리를 탐색하는 단계와 상기 질의 트리의 마지막 노드까지 탐색한 후 질의 트리의 마지막 노드에 해당하는 그룹 노드의 정보를 질의의 결과로 반환하는 것을 특징으로 한다.

[0039] 그리고 상기 질의 트리를 탐색하는 단계는 사용자의 질의에 조건식이 포함된 경우 상기 속성 요약된 그룹 노드 중 속성 이름 리스트에 포함된 속성의 이름과 상기 속성의 값들이 연속적으로 저장된 스트림의 시작 주소 정보를 통해 주어진 조건식에 포함된 속성 값이나 텍스트만을 선택적으로 수신하여 상기 조건식을 만족하는 듀이 순서를 질의 트리에 저장하는 것을 특징으로 한다.

[0040] 여기서, 상기 질의 트리에 조건식에 의해 저장된 듀이 순서가 있을 경우 상기 듀이 순서에 해당하는 정보만 질의의 결과로 반환하는 것을 특징으로 한다.

[0041] 또한, 상기 질의를 처리하는 단계는 상기 질의에 분기 조건이 포함된 경우 질의 트리를 생성하고 상기 질의 트리에 대한 경로를 탐색하는 단계와 상기 분기 조건을 만족하는 엘리먼트를 선별하기 위해 P-C 비트열 일치 여부를 확인하는 단계와 상기 선별된 P-C 비트열을 말단 그룹 노드에 적용하여 질의 결과를 도출하는 단계를 포함하는 것을 특징으로 한다.

[0042] 여기서, 상기 경로를 탐색하는 단계는 찾고자 하는 데이터가 포함된 경로인 주 경로에 대한 탐색과 주어진 분기 조건을 포함하는 부 경로에 대한 탐색을 포함하고, 상기 주 경로와 부 경로에 해당하는 그룹 노드들을 수신하여 P-C 비트열 및 압축 정보를 질의 트리의 각 노드에 저장하는 것을 특징으로 한다.

[0043] 그리고 상기 P-C 비트열 일치 여부를 확인하는 단계는 상기 경로를 탐색하는 단계를 통해 질의 트리의 각 노드에 저장된 P-C 비트열 및 압축 정보를 이용하여 상기 분기조건을 만족하는 엘리먼트를 선별하되, 상기 질의 트리에서 주 경로와 부 경로의 말단 노드부터 분기 노드까지 올라가면서 P-C 비트열을 압축 정보에 따라 압축하여 bit-wise AND 연산을 수행하고, 상기 질의트리의 분기 노드에서는 상기 주 경로와 부 경로에서 계산된 비트열을 bit-wise AND 연산하여 분기 조건을 만족하는 중간 결과 값을 산출하는 것을 특징으로 한다.

[0044] 그리고 상기 질의 결과를 도출하는 단계는 상기 분기 노드에서 계산된 중간 결과 값을 질의 트리에서 주 경로를 따라 내려가면서 원래의 비트열로 복원하고, 상기 주 경로의 말단 노드까지 복원된 비트열이 말단 그룹 노드의 데이터에 적용되어 비트열에서 1에 해당하는 엘리먼트의 내용만 질의 결과로 반환하는 것을 특징으로 한다.

[0045]

효 과

[0046] 본 발명은 무선 방송 환경에서 중요한 척도인 접근 시간과 튜닝 시간을 감소시키고 효율적으로 XML 분기 질의를 방법에 관한 것으로, 구조적인 요약과 속성 요약 방법을 통해 방송스트림의 크기를 감소 및 접근시간을 단축시키고, 속성 값과 텍스트에 대한 선택적인 수신을 통해 튜닝 시간을 효과적으로 단축시키는 탁월한 효과가 발생한다.

[0047] 또한, 본 발명은 속성 값과 텍스트를 듀이 순서와 함께 저장함으로써 무선 방송 환경에서 구조적인 요약을 이용한 질의 처리 방법에서 조건식을 포함한 질의를 처리할 수 없는 한계를 극복하였으며, P-C 비트열과 압축 정보를 사용하여 부모-자식 관계를 유지함으로써 효율적인 분기 질의 처리를 지원하는 탁월한 효과가 발생한다.

[0048] 또한, 본 발명은 주어진 질의에 부합되는 질의 트리를 생성하고 질의 처리 단계에 따라 방송스트림을 선택적으로 수신함으로써 이동 사용자에게 직관적이고 효율적인 분기 질의 처리 방법을 지원한다. 즉, 이동 사용자는 질의 트리에서 각 지의 처리의 위치 단계에 해당하는 스트림의 그룹 노드를 수신하여 질의 트리에 분기 질의 처리를 위한 색인 정보인 P-C 비트열과 압축 정보를 저장하고, bit-wise 연산을 수행함으로써 분기 조건을 만족하는 엘리먼트의 정보를 선별할 수 있으며, 이동 사용자는 XML 스트림에서 색인 정보들을 포함하고 있는 인덱스를 우선 수신한 후 질의 트리에서 속성이나 텍스트에 대한 조건식이 주어질 때만 해당하는 정보를 선택적으로 수신함으로써 에너지 효율성을 확보할 수 있는 탁월한 효과가 발생한다.

발명의 실시를 위한 구체적인 내용

- [0049] 이하, 본 발명의 구체적인 실시예에 대하여 도면을 참조하여 상세하게 설명하기로 한다.
- [0050] 본 발명은 XML 질의 처리 방법에 관한 것으로, 무선 방송스트림을 생성하여 전송하는 서버단과 상기 방송스트림을 수신하는 이동 사용자의 클라이언트단으로 구성되어 XML 질의 처리가 수행된다.
- [0051] 즉, 본 발명에 따른 XML 질의 처리 방법은 서버단의 방송스트림 생성 장치 또는 프로그램에 의해 생성된 방송스트림을 이동 사용자의 클라이언트단(단말)이 수신하여 클라이언트단의 프로세서가 상기 수신된 방송스트림에 대한 질의 처리를 수행한다.
- [0052] 상기와 같이 본 발명에 따른 방송스트림의 생성은 서버단의 방송스트림 생성 장치 또는 생성 프로그램이 수행하고, 질의 처리는 클라이언트단의 프로세서가 수행하는 것을 전제로 설명하기로 한다.
- [0053] 이하, 본 발명의 구성상의 특징들에 대해 상세하게 살펴본 후, 구체적인 XML 질의 처리 방법에 대해 살펴보기로 한다.
- [0054] <무선 방송을 위한 XML 방송스트림 구조>
- [0055] 본 발명은 XML 방송 스트림에서 구조적인 요약을 사용하여 경로가 같은 엘리먼트들이 병합된 그룹 노드의 구조를 가지고, 속성 요약과 속성에 따른 선택적인 수신 방법을 특징으로 하고 있으므로 차례대로 살펴보기로 한다.
- [0056] -그룹 노드 구조
- [0057] 본 발명은 경로가 같은 엘리먼트들을 하나의 노드로 병합하여 하나의 그룹 노드를 생성함으로써 XML 문서의 방송스트림을 그룹 노드의 방송스트림으로 생성한다.
- [0058] 본 발명에 따른 그룹 노드는 다음과 같이 나타낼 수 있다.
- [0059] $GN_p = [I_p, D_p]$, 여기서 GN_p 는 원본 XML 문서에서 경로가 P인 모든 엘리먼트(e_p)를 병합한 그룹 노드이며, I_p 와 D_p 는 각각 상기 그룹 노드의 인덱스(Index)와 데이터(Data)를 의미한다.
- [0060] 상기 I_p 는 GN_p 에 병합된 e_p 의 태그 이름(tn_e), 위치 경로 IP_e , 자식 주소(Child addresses), P-C 비트열(P-C Bitstring), 압축 정보(Packing Information), 속성 이름 리스트(Attribute name list) 및 텍스트 주소(Text address)를 포함하여 구성될 수 있다.
- [0061] 여기서, 상기 자식 주소는 자식 그룹 노드들에 대한 주소 정보이며, 상기 P-C 비트열과 압축 정보는 그룹 노드에 병합된 엘리먼트들의 부모-자식 관계를 표현하는 정보이고, 상기 속성 이름 리스트는 속성의 이름과 상기 속성의 값들이 연속적으로 저장된 스트림의 시작 주소 정보를 포함하고, 텍스트 주소는 그룹 노드에 속하는 엘리먼트의 텍스트들이 연속적으로 저장된 스트림의 시작 주소이다.
- [0062] 상기 DP 는 속성 값 리스트(Attribute value list)와 텍스트 리스트(Text list)를 포함하여 구성될 수 있다.
- [0063] 여기서, 상기 속성 값 리스트는 그룹 노드에 병합된 엘리먼트의 속성 값과 듀이 순서가 속성 이름별로 연속적으로 저장된 리스트이며, 상기 텍스트 리스트는 그룹 노드에 병합된 엘리먼트의 텍스트들이 듀이 순서와 함께 그룹화되어 저장된 리스트이다.
- [0064] 도 3은 본 발명의 바람직한 실시예에 따른 그룹 노드의 구조를 개략적으로 도시한 것으로, 도 1a의 XML 문서 예제에서 경로가 "/mondial/country/province"인 그룹 노드의 구조를 도시한 것이다.
- [0065] 도 3을 참조하면, 그룹 노드의 인덱스(Index, I_p)는 이동 사용자가 XML 스트림을 선택적으로 수신하기 위한 색인 정보의 모음으로 방송 채널에서 그룹 노드를 수신할 때 필수적으로 수신하는 부분이다.
- [0066] 상기 인덱스는 그룹 노드의 태그 이름(그룹 이름)인 "province", 위치 경로인 "/modial/country/province.", 자식 주소인 "city" 그룹 노드에 대한 주소, P-C 비트열인 비트열 1111, 압축 정보인 문자열(string) 2.1.1, 속성 이름 리스트인 "province" 엘리먼트들의 속성인 "id" 속성과 "name" 속성 값의 스트림 시작 주소, 그리고 텍스트 주소인 "province" 엘리먼트들의 텍스트들의 스트림 시작 주소로 구성된다.
- [0067] 여기서, 각 인덱스 정보에 포함된 주소는 XML 데이터 스트림에서 해당 데이터가 방송되는 시간 정보를 의미한다.

다. 그룹 노드의 이름 및 경로는 상기 그룹 노드를 식별하는데 사용되며, 상기 자식 주소는 방송스트림에서 "city" 그룹 노드의 선택적인 수신을 위해 사용된다.

- [0068] 그리고 상기 속성 이름 리스트와 텍스트 주소는 데이터 중에서 조건식에 포함된 속성 값이나 텍스트만 선택적으로 수신하는데 사용된다.
- [0069] P-C 비트열은 상기 그룹 노드의 엘리먼트가 부모 그룹 노드의 엘리먼트의 자식 엘리먼트로 존재하는지 여부를 0 또는 1의 연속된 비트열로 표현한 정보이며, 압축 정보는 하나의 부모 엘리먼트에 소속된 자식 엘리먼트의 갯수를 나타낸 정보이다.
- [0070] 또한, 데이터(DP)는 "id" 속성과 "name" 속성의 값들이 듀이 순서와 함께 연속적으로 저장된 속성 값 리스트를 포함하고 있다. 여기서, "province" 엘리먼트들은 텍스트를 가지고 있지 않기 때문에 텍스트 리스트에는 아무 것도 저장되어 있지 않다.(null)
- [0071] 그룹 노드는 같은 경로의 여러 엘리먼트들을 한 개의 노드로 병합하되 각 엘리먼트의 속성 값과 텍스트를 그 엘리먼트의 듀이 순서와 함께 저장함으로써 원본 XML 문서에서의 엘리먼트 순서 정보를 유지하여 조건식이 포함된 질의를 처리할 수 있다.
- [0072] - 속성 요약
- [0073] 본 발명은 스트림의 크기를 최소화하고 이동 사용자의 튜닝 시간을 단축시키기 위해 그룹 노드의 속성들을 요약하고, 요약된 속성 및 텍스트들을 선택적으로 수신한다.
- [0074] XML에서 속성은 속성 이름과 속성 값의 쌍으로 이루어져 있으며, 하나의 엘리먼트가 여러 개의 속성을 가질 수 있다. 이때, 경로와 태그 이름이 같은 엘리먼트들은 일반적으로 같은 속성 이름을 사용하는 구조적 특성이 있다.
- [0075] 예를 들어, 도 1a의 XML 문서에서 "province" 엘리먼트들은 "id"와 "name"이라는 공통적인 속성을 지닌다. 따라서 같은 속성 이름이 여러 개 반복될 경우, 반복되는 속성 이름을 생략하고 속성값과 이 속성 값을 가진 엘리먼트의 순서 정보만 저장함으로써 원래 XML 문서에서의 엘리먼트 순서를 유지하면서 스트림의 크기를 최소화하고 이동 사용자의 튜닝시간을 단축시킬 수 있다.
- [0076] 본 발명에 따른 속성 요약 방법은 상기에서 설명한 바와 같이 중복되는 속성 이름은 생략하고, 속성 값과 상기 값을 가진 엘리먼트의 듀이 순서를 저장하여 속성에 대한 조건식이 포함된 질의를 효율적으로 처리할 수 있는 색인 방법이다.
- [0077] 도 4는 도 1a의 예제 XML 문서에서 "province" 노드의 속성들을 속성 요약 방법을 이용해 간략화한 것을 도시한 것이다.
- [0078] 도 4를 통해 알 수 있듯이, 본 발명에 따른 속성 요약 방법을 통해 중복되는 속성 이름들을 모두 생략함으로써 스트림의 크기를 효과적으로 줄일 수 있다.
- [0079] 한편, 속성 이름은 상기 속성의 값이 연속적으로 저장되어 있는 스트림에서의 시작 주소와 함께 속성 이름 리스트에 저장되며, 같은 이름의 속성 값들은 각각의 듀이 순서와 함께 그룹 노드의 속성 값 리스트에 연속적으로 저장된다.
- [0080] XML 문서에서 인덱스 정보에 비해 문장(string) 형태인 텍스트 및 속성이 전체 스트림의 대부분을 차지하기 때문에 구조적인 요약을 통해 같은 경로를 가진 엘리먼트들의 태그를 생략하고 속성 이름과 값을 분리하여 주어진 조건식에 속성 값만 수신함으로써 접근 시간과 튜닝 시간을 단축할 수 있다.
- [0081] 도 5는 본 발명의 바람직한 실시예에 따른 조건식이 포함된 질의에 대한 속성 요약 방법을 이용하여 선택적으로 수신하는 예를 도시한 것이다.
- [0082] 상기 도 5는 조건식이 "/mondial/country/province[@id= "f0_17473"]"인 경우에 대한 것이다. 도 5를 참조하면, 속성 이름 리스트는 "id"와 "name" 속성 값이 연속적으로 저장되어 있는 부분의 시작 주소를 포함하고 있으므로, 이동 사용자의 클라이언트 단의 프로세서는 "province" 엘리먼트들의 속성 전체를 탐색하지 않고 "id"란 속성 이름에 해당하는 속성 값들만 선택적으로 수신 및 검색함으로써 조건식이 포함된 질의를 효율적으로 처리할 수 있다.
- [0083] <분기질의 처리를 위한 P-C 비트열(P-C Bitstring)과 압축 정보(Packing Information)>

- [0084] 종래의 구조적인 요약을 이용한 방법에서는 부모-자식 관계의 정보가 유지되지 않기 때문에 분기 질의 처리시에 정확한 결과를 반환할 수 없는 문제점이 있다.
- [0085] 본 발명에서는 상기와 같이 부모-자식 관계의 정보의 부재로 인해 잘못된 결과 집합을 반환하는 문제를 해결하기 위해 P-C비트열과 압축 정보를 이용하였다.
- [0086] 도 6은 국가명이 명시된 국가에 소재한 주(province)를 검색하는 질의(//country[name]/province)를 원래 XML 문서의 트리(a)와 구조적인 요약을 사용한 트리(b)에서 처리한 결과를 도시한 것이다.
- [0087] 도 6의 (a)와 같이 원래 XML 문서에서는 분기 조건을 만족하는 2번 "country" 노드의 자식인 4번과 7번 "province" 노드만이 정확한 결과이다.
- [0088] 그러나 구조적인 요약을 사용한 방법에서는 엘리먼트의 부모-자식 관계가 유지되지 않으므로 병합된 "province"엘리먼트들 가운데 어느 엘리먼트가 2번 "country" 엘리먼트의 자식인지 식별할 수 없기 때문에 분기 질의 처리를 할 수 없다.
- [0089] 본 발명은 부모-자식 관계를 P-C 비트열과 압축정보를 통해 나타냄으로써 분기 질의 처리시 정확한 결과를 반환한다. P-C 비트열은 해당 그룹 노드의 엘리먼트가 부모 그룹 노드에 속하는 엘리먼트의 자식 엘리먼트중에서 존재하는지를 여부를 비트(bit)로 나타낸 것이다.
- [0090] 즉, 상기 해당 그룹 노드에 속하는 엘리먼트가 부모 그룹 노드에 속하는 엘리먼트의 자식들 중에 존재한다면 1로 나타내고, 존재하지 않는다면 0으로 표현한다. 비트를 사용해 부모-자식 관계를 표현함으로써 다음과 같은 장점을 얻을 수 있다.
- [0091] 크기가 작은 비트를 사용해 부모-자식관계를 표현함으로써 색인 정보의 경량화하고 스트림의 크기를 줄일 수 있으며, Bit-Wise 연산을 통해 직관적이고 빠른 연산이 가능하다.
- [0092] 한편, XML에서는 하나의 엘리먼트가 여러 개의 같은 이름을 가진 엘리먼트를 자식으로 가질 수 있기 때문에 부모-자식 관계를 P-C 비트열로 표현할 때, 각 부모 엘리먼트에 대응되는 자식 엘리먼트들을 정확히 구분할 수 없다.
- [0093] 본 발명은 압축정보를 사용해 각 부모 엘리먼트에 대응되는 자식 엘리먼트들을 구분한다. 압축정보는 그룹 노드에 속하는 엘리먼트 중에서 부모가 같은 엘리먼트들의 갯수를 나타낸 정보로 압축정보 내의 각 숫자는 부모 그룹 노드에 속하는 각 엘리먼트에 순서대로 대응된다.
- [0094] 도 7은 본 발명의 바람직한 실시예에 따른 예제 XML 문서로부터 구조적인 요약을 통해 생성된 그룹 노드의 P-C 비트열과 압축정보를 도시한 것이다.
- [0095] 예를 들어, 3번 "name" 그룹 노드의 P-C 비트열(100)은 부모 그룹 노드인 "country"그룹 노드에 속하는 엘리먼트중에 첫 번째 "country"엘리먼트만이 자식으로 "name"엘리먼트를 가지고 있음을 의미한다.
- [0096] 압축정보(1.1.1)은 P-C 비트열의 비트들이 각각 1개의 부모 엘리먼트에 대응된다는 것을 의미한다. 마찬가지로, 4번 "province"그룹 노드의 P-C 비트열(1111)은 부모 그룹 노드인 "country"그룹 노드에 속하는 엘리먼트들의 자식 엘리먼트들 중에 총 4개의 "province"엘리먼트가 존재함을 의미한다.
- [0097] 압축정보(2.1.1)은 첫 번째와 두 번째 "province"엘리먼트의 부모가 같으며, 세 번째와 네 번째 비트에 해당하는 부모 엘리먼트는 다르다는 것을 의미한다.
- [0098] <XML 방송 스트림에서의 질의 처리 방법>
- [0099] 이하, 방송 스트림을 이용한 질의 처리 방법에 대해 살펴보기로 한다. XML 방송 스트림에서 질의 처리는 분기 조건이 없는 일반적인 경로 표현식 질의와 분기 조건이 포함된 분기 질의로 구분할 수 있다. 질의가 주어지면 이동 사용자는 질의를 분석하여 질의 트리(query tree)를 생성한 후 하향식으로 질의 트리를 탐색한다. 따라서, 질의 트리 생성에 대해 간단하게 설명한 후 질의 처리 방법에 대해 상세하게 설명하기로 한다.
- [0100] 질의가 주어지면 이동 사용자가 소지한 단말의 프로세서는 질의를 분석하여 효율적인 질의 처리를 위한 질의 트리를 생성한 후 각 단계에 따라 스트림을 선택적으로 수신하여 질의 처리를 수행한다.
- [0101] 질의 트리의 각 노드는 위치 경로(location path)에서 위치 단계(location step)에 해당되며 사용자는 질의 트리를 하향식으로 탐색하면서 질의 트리의 각 노드에 해당하는 그룹 노드를 수신한다.

- [0102] 도 8은 여러 질의에 대한 다양한 질의 트리의 예를 도시한 것이다.
- [0103] 질의 트리에서 말단 노드(leaf node)를 제외한 중간 노드(internal node)는 속성이나 텍스트에 대한 조건식이 주어지지 않을 경우, 다음 위치 단계에 해당하는 노드 탐색을 위해 그룹 노드의 인덱스만을 수신한다. 중간 노드에 조건식이 주어질 경우 이동 사용자는 인덱스에 포함된 색인 정보를 이용하여 그룹 노드의 데이터 중에서 조건식에 해당하는 속성 값 혹은 텍스트 데이터만 선택적으로 수신한다.
- [0104] -단순 경로 질의 처리
- [0105] 먼저, 분기 조건이 포함되지 않은 단순 경로 질의 처리에 대해 설명한다.
- [0106] 질의가 주어지면 이동 사용자가 소지한 단말의 프로세서는 질의 트리를 생성하여 최상위 노드부터 하향식으로 질의 트리를 탐색한다.
- [0107] 질의 트리의 탐색은 그룹 노드의 인덱스를 수신하여 자식 주소(Child Addresses)에서 다음 위치 단계의 그룹 노드를 찾는 것으로 이루어진다.
- [0108] 이때, 질의 트리의 위치 단계에 속성이나 텍스트에 대한 조건식이 포함되어 있다면 인덱스의 속성 이름 리스트(Attribute name list)나 텍스트 주소(Text address)를 이용하여 그룹 노드의 데이터 중에서 조건식에 해당하는 속성 값이나 텍스트만 선택적으로 수신하여 조건식을 만족하는 듀이 순서를 질의 트리에 저장한다.
- [0109] 질의 트리의 마지막 노드까지 탐색한 후에는 질의 트리의 마지막 노드에 해당하는 그룹 노드의 정보를 질의의 결과로 반환한다. 질의 트리에 조건식에 의해 저장된 듀이 순서가 있을 경우에는 이 듀이 순서에 해당하는 정보만 질의의 결과로 반환한다.
- [0110] 도 9는 본 발명의 바람직한 실시예에 따른 속성에 대한 조건식을 포함한 단순 경로 질의 (/mondial/country/province[@name="Aland"/city/text()])에 대한 질의 트리 탐색 과정(a)과 XML 데이터 스트림의 선택적인 수신 과정(b)을 도시한 것이다.
- [0111] 질의가 주어지면 이동 사용자는 도 9의 (a)와 같은 질의 트리를 생성하고 하향식으로 질의 트리를 탐색한다.
- [0112] 먼저, 그룹 노드 트리의 루트인 "mondial"그룹 노드를 수신하여 자식 주소들 중에서 질의 트리의 다음 단계에 해당하는 "country"그룹 노드의 주소를 탐색하여 수신한다.
- [0113] 같은 방식으로 "country"그룹 노드를 수신한 후에는 질의 트리의 다음 단계인 "province"그룹 노드의 주소를 탐색한다. 상기 "province"그룹 노드를 수신한 후에는 이 위치 단계에 주어진 조건식을 처리하기 위해 인덱스의 속성 이름 리스트에서 "name"속성의 값들이 저장된 시작 주소를 탐색하여 수신한다.
- [0114] "province"그룹 노드의 속성 값 리스트(Attribute value list)에서 속성 값이 "Aland"엘리먼트의 듀이 순서(1.3.1)를 탐색하여 질의 트리의 "province"노드에 저장한다.
- [0115] 마지막으로 질의 트리의 말단 노드인 "city"그룹 노드를 수신하여 텍스트 주소로부터 텍스트가 저장된 부분만 선택적으로 수신한다. 이 질의의 결과로 텍스트 데이터 중에서 질의 트리에 저장된 듀이 순서의 후순(1.3.1.1)에 해당하는 텍스트 데이터(Mariehamn)만 결과로 반환된다.
- [0116] - 분기 질의 처리
- [0117] 이어서, 분기 조건이 주어졌을 때 질의 처리 방법에 대해 설명한다.
- [0118] 분기 조건이 주어졌을 때의 질의 처리는 질의 트리에 대한 경로 탐색 과정(path traversal phase), 분기 조건을 만족하는 엘리먼트를 선별하기 위한 P-C 비트열 일치 여부 확인 과정(matching phase), 그리고 선별된 P-C 비트열을 말단 그룹 노드에 적용하여 결과를 도출하는 과정(retrieval phase)을 포함하여 이루어진다.
- [0119] 상기 질의 트리에 대한 경로 탐색 과정(path traversal phase)은 주 경로에 대한 탐색(main path traversal)과 부 경로에 대한 탐색(sub path traversal)으로 구성되며, 여기서 주 경로란 찾고자 하는 데이터가 포함된 경로이고, 부 경로란 주어진 분기 조건을 포함하는 경로이다.
- [0120] 예를 들어, 인구 정보가 있는 도시의 이름을 검색하는 질의(/mondial/country/city[population]/name)에서 주 경로는 찾고자 하는 데이터인 도시의 이름을 포함하는 경로(/mondial/country/city/name)이며, 부 경로는 분기 조건으로 주어진 인구 정보를 포함하는 경로(/mondial/country/city/population)이다. 부 경로는 질의에 주어진 분기 조건에 따라 여러 개가 존재할 수 있다.

- [0121] 상기 경로에 대한 탐색 과정(path traversal phase)에서는 XML 스트림에서 주 경로와 부 경로에 해당하는 그룹 노드들을 수신하여 P-C 비트열과 압축정보를 질의 트리의 각 노드에 저장한다.
- [0122] 상기 P-C 비트열 일치 여부 확인 과정(matching phase)에서는 상기 경로 탐색 과정(path traversal phase)을 통해 질의 트리의 각 노드에 저장된 P-C 비트열과 압축정보를 이용해 분기 조건을 만족하는 엘리먼트를 선별한다.
- [0123] 이 과정에서는 질의 트리에서 주 경로와 부 경로의 말단 노드부터 분기 노드까지 올라가면서 P-C 비트열을 압축정보에 따라 압축(packing)하여 bit-wise AND 연산을 수행한다.
- [0124] 질의 트리의 분기 노드에서는 주 경로와 부 경로에서 계산된 비트열을 bit-wise AND 연산하여 분기 조건을 만족하는 중간 결과 값을 산출한다.
- [0125] 마지막으로 상기 P-C 비트열을 말단 그룹 노드에 적용하여 결과를 도출하는 과정(retrieval phase)에서는 분기 노드에서 계산된 중간 결과 값을 질의 트리에서 주 경로를 따라 내려가면서 원래의 비트열로 복원(unpacking)한다.
- [0126] 주 경로의 말단 노드까지 복원된 비트열은 말단 그룹 노드의 Data에 적용되어 비트열에서 1에 해당하는 엘리먼트의 내용만 결과로 반환된다.
- [0127] 도 10은 본 발명의 바람직한 실시예에 따른 분기 질의(/mondial/ country[name]/province/city)를 처리하는 예를 도시한 것이다.
- [0128] 이동 사용자가 소지한 단말의 프로세서는 주어진 질의로부터 질의 트리를 생성하여 주 경로(/mondial/country/province/city)와 부 경로(/mondial/country/name)를 탐색한다.
- [0129] 질의 트리의 루트인 "mondial"그룹 노드를 수신하여 자식주소들(Child addresses)로부터부터 질의 트리의 다음 단계인 "country"그룹 노드를 수신한다.
- [0130] 질의 트리에서 분기가 이루어지는 "country"그룹 노드를 수신한 후에는 이 그룹 노드의 자식 주소들(Child addresses)을 이용하여 주 경로의 그룹 노드와 부 경로의 그룹 노드 중에서 먼저 방송되는 순서대로 그룹 노드를 수신한다.
- [0131] 먼저 "name"그룹 노드를 수신하여 P-C 비트열과 압축정보를 저장하고, 다음으로 "province"그룹 노드를 수신하여 P-C 비트열과 압축정보를 질의 트리에 저장한다.
- [0132] 마지막으로 "province"그룹 노드의 자식 주소들(Child addresses)로부터 주 경로의 말단 노드인 "city"그룹 노드를 수신한다.
- [0133] 질의 트리에 대한 경로 탐색 과정(path traversal phase)이 끝난 후에는 주 경로와 부 경로의 말단 노드부터 분기 노드까지 P-C 비트열 일치 여부 확인 과정(matching phase)을 수행한다.
- [0134] 먼저 주 경로의 말단 노드인 "city"그룹 노드의 P-C 비트열(11111)과 부모 노드인 "province"그룹 노드의 P-C 비트열(1111)을 bit-wise AND 연산한다.
- [0135] 이때 "city"그룹 노드의 P-C 비트열은 압축정보(2.1.2.1)에 의해 1111로 압축된다. 즉, 원래 P-C 비트열에서 첫 번째와 두 번째, 네 번째와 다섯 번째 비트들은 부모가 같기 때문에 OR-ing하여 한 개의 비트로 압축된다.
- [0136] 마찬가지로 "province"그룹 노드까지 계산된 결과(1111)와 부모 그룹 노드인 "country"그룹 노드의 P-C 비트열(111)을 bit-wise AND 연산하여 주 경로의 중간 결과(111)를 산출한다.
- [0137] 같은 방식으로 부 경로의 말단 노드인 "name"그룹 노드부터 분기 노드인 "country"그룹 노드까지 P-C 비트열을 bit-wise AND 연산하여 부 경로의 중간 결과(100)를 산출한다.
- [0138] 그리고 분기 노드에서 주 경로와 부 경로의 중간 결과를 bit-wise AND 연산하여 분기 조건을 만족하는 비트열(100)을 산출한다.
- [0139] 상기 분기 노드에서 계산된 비트열은 주 경로를 따라 내려가면서 각 노드에 저장된 P-C 비트열과 압축정보를 이용해 원래의 비트열로 복원된다. "province"노드의 압축정보는 (2.1.1)이므로 중간 결과(100)에서 첫 번째 1이 압축되기 전의 비트열인 11로 복원되어 중간 결과는 1100이 된다.
- [0140] 다음으로 "city"노드의 압축정보는 (2.1.2.1)이므로 1100에서 첫 번째와 세 번째 비트는 각각 11과 00으로 복

원된다. 따라서 질의 단말 그룹노드인 "city"에 대한 최종 비트열은 111000이므로 분기 질의를 만족하는 엘리먼트는 "city"그룹 노드에 병합된 엘리먼트들 가운데 첫 번째부터 세 번째 엘리먼트들이다.

- [0141] 따라서, 이 질의에 대한 결과로 5번, 6번, 8번 "city"엘리먼트들만 결과로 반환한다.
- [0142] 상기와 같은 본 발명에 따른 질의 처리 방법의 성능을 측정하기 위해 실제 데이터 집합에 대해 단순 경로 질의와 분기 질의를 사용하여 평가하였다.
- [0143] 성능 비교는 한 개의 노드로 스트리밍하여 방송하는 방법(이하, S-node)과 경로가 같은 엘리먼트들의 데이터를 하나의 노드로 간략화하여 방송하는 방법(이하, Path-summary)에 대해 접근 시간 및 튜닝 시간을 평가하였다.
- [0144] 단, 상기 Path Summary는 조건식이 포함된 질의나 분기 질의를 지원하지 않기 때문에 단순 경로 질의에 대해서만 실험하였다.
- [0145] 본 실험에서는 실제 데이터 집합인 *mondial* XML 문서를 사용하여 무선 방송을 위한 XML 스트림을 생성하고, 질의 처리에 소요되는 접근 시간 및 튜닝 시간을 측정하였다.
- [0146] 성능 측정을 위해 표준 언어 질의인 XPath 질의를 사용하여 조건식이 포함되지 않은 단순 경로 질의, 조건식이 포함된 단순 경로 질의, 조건식이 포함되지 않은 분기 질의, 그리고 조건식이 포함된 분기 질의의 4가지의 질의 유형에 대해 평가하였다.
- [0147] 하기의 <표 1>은 실험에 사용된 데이터 집합과 질의들을 기술한 것이다.
- [0148] S1과 S2는 단순 경로 질의를 측정하기 위한 질의들이다. 조건식이 포함된 단순 경로 질의는 SP1과 SP2를 사용하여 측정하였다. T1과 T2는 조건식이 포함되지 않은 분기 질의들이며, 조건식을 포함한 분기 질의는 TP1과 TP2를 사용하여 측정하였다.

표 1

Query #	Test Query (XPath expression)	Description
S1	/mondial/organization	Simple path query
S2	/mondial/country/name	
SP1	/mondial/country/name[text()='CzechRepublic']	Simple path query with Predicate
SP2	/mondial/country/city/located_at[@type="lake"]	
T1	//country[name]/city/name	Twig query
T2	/mondial/country/city[population]/name	
TP1	/mondial/country[religions]/city[@id="f0_1751"]/name	Twig query with Predicate
TP2	/mondial/country/province[@name="Praha"]/city[population]/name	

- [0149]
- [0150] 도 11은 원본 XML 문서, Path Summary, SL, SD, SP, 그리고 본 발명에 따라 생성된 XML 스트림의 크기를 비교한 것이다. 여기서 SL, SD, 그리고 SP 방법은 S-node에서 제안한 공지된 방법들이다.
- [0151] 실험 결과와 같이 Path Summary, SL, SD, SP, 그리고 본 발명에 따라 생성된 XML 스트림은 원본 XML 문서의 크기를 각각 52.9%, 38.6%, 33.5%, 33.5%, 그리고 42.4% 감소시켰다.
- [0152] 실험에 사용된 데이터 집합은 크기가 작으며 속성을 포함하고 있지 않은 엘리먼트의 수가 많다. 따라서 본 발명에 따라 생성된 XML 스트림에서는 P-C비트열(P-C Bitstring), 압축정보(Packing Information) 등의 추가된 색인 정보로 인한 스트림의 증가에 비해 속성 요약에 의한 스트림 감소 효율이 낮기 때문에 제안 기법에 의해 생성된 XML 스트림의 크기는 원본 XML 문서나 S-node에 의해 생성된 XML 스트림보다 작았지만 Path Summary에 의해 생성된 XML 스트림보다는 크게 나타났다. 그러나, 실제 방송스트림 환경과 같이 속성을 포함하고 있는 엘리먼트의 수가 많을 경우 Path Summary에 의해 생성된 XML 스트림보다 작아질 것이 확실하다.
- [0153] 도 12는 접근 시간을 측정한 결과를 도시한 것이다.
- [0154] 이 실험에서 Path Summary 방법은 단순 질의 처리만 지원하기 때문에 SP1~TP2에 대한 성능 측정은 수행하지 않

았다. 본 발명은 다른 방법들에 비해 평균적으로 우수한 성능을 보였으며 SL, SD, SP, 그리고 Path Summary 방법에 비해 각각 73.4%, 66.9%, 63.3%, 30.8%의 접근 시간이 단축되었다.

- [0155] 특히 S2~TP1 질의에 대한 질의 측정 결과에서는 제안 방법의 성능이 S-node에 비해 크게 향상되었는데, 이는 S-node는 XML 문서의 나타난 엘리먼트 순서대로 스트림을 구성하기 때문에 스트림 전체에 흩어져 있는 노드들을 탐색하는 한편 제안방법은 구조적인 요약을 사용하여 하나의 그룹 노드로 병합하였기 때문에 질의 처리가 일찍 종료 되기 때문이다.
- [0156] 그러나 S1 질의에 대한 접근 시간 측정에서는 질의의 결과 값이 실험에서 사용된 XML 문서의 후반부에 위치하기 때문에 스트림의 크기가 가장 작은 Path Summary 방법이 가장 우수한 성능을 보였으며, 나머지 방법들의 접근 시간은 스트림의 크기와 비례하였다.
- [0157] 도 13은 튜닝 시간을 측정한 결과를 도시한 것이다.
- [0158] 접근 시간 측정과 마찬가지로 Path Summary 방법의 튜닝 시간은 S1과 S2 질의에 대해서만 측정하였다. 실험 결과에서 알 수 있듯이 본 발명은 다른 방법들에 비해 접근 시간 보다 튜닝 시간 성능 측정에서 우수한 성능을 보였다. 본 발명은 SL, SD, SP, Path Summary 방법에 비해 각각 95.5%, 90.4%, 75.4%, 40.7%의 접근 시간을 단축시켰다.
- [0159] 이는 다음과 같은 이유로 설명된다.
- [0160] 1) 본 발명에서는 Index와 Data를 분리하여 Index 부분만 수신한 후, 조건식에 주어진 속성 값들 혹은 텍스트를 선택적으로 수신하기 때문에 질의에 조건식이 포함되어 있거나 질의의 말단 노드에 텍스트가 존재하지 않을 경우 튜닝 시간이 효과적으로 감소된다.
- [0161] 2) 특히 분기 질의의 경우, S-node는 분기 조건을 만족하는 결과를 선별하기 위해 단순 경로 질의에 비해 많은 노드를 수신하기 때문에 튜닝 시간이 크게 증가 하였으나, 제안 방법에서는 그룹 노드의 Index에 포함된 P-C Bitstring과 Packing Information을 이용하여 분기 조건을 만족하는 엘리먼트를 선별하기 때문에 분기 질의 처리에 뛰어난 성능을 보였다.
- [0162] 이상에서 설명한 본 발명의 상세한 설명에서는 본 발명의 바람직한 실시예를 참조하여 설명하였지만, 본 발명의 보호범위는 상기 실시예에 한정되는 것이 아니며, 해당 기술분야의 통상의 지식을 갖는 자라면 본 발명의 사상 및 기술영역으로부터 벗어나지 않는 범위 내에서 본 발명을 다양하게 수정 및 변경시킬 수 있음을 이해할 수 있을 것이다.

도면의 간단한 설명

- [0163] 도 1a는 종래의 XML 문서를 도시한 것이고, 도 1b는 도 1a의 XML 문서를 트리 형태로 표현한 것을 도시한 것이다.
- [0164] 도 2는 상기와 같은 분기 질의를 트리 형태로 표현한 것이다.
- [0165] 도 3은 본 발명의 바람직한 실시예에 따른 그룹 노드의 구조를 개략적으로 도시한 것으로, 도 1a의 XML 문서 예제에서 경로가 "/mondial/country/province"인 그룹 노드의 구조를 도시한 것이다.
- [0166] 도 4는 도 1a의 예제 XML 문서에서 "province" 노드의 속성들을 속성 요약 방법을 이용해 간략화한 것을 도시한 것이다.
- [0167] 도 5는 본 발명의 바람직한 실시예에 따른 조건식이 포함된 질의에 대한 속성 요약 방법을 이용하여 선택적으로 수신하는 예를 도시한 것이다.
- [0168] 도 6은 국가명이 명시된 국가에 소재한 주(province)를 검색하는 질의(//country[name]/province)를 원래 XML 문서의 트리(a)와 구조적인 요약을 사용한 트리(b)에서 처리한 결과를 도시한 것이다.
- [0169] 도 7은 본 발명의 바람직한 실시예에 따른 예제 XML 문서로부터 구조적인 요약을 통해 생성된 그룹 노드의 P-C 비트열과 압축정보를 도시한 것이다.
- [0170] 도 8은 여러 질의에 대한 다양한 질의 트리의 예를 도시한 것이다.

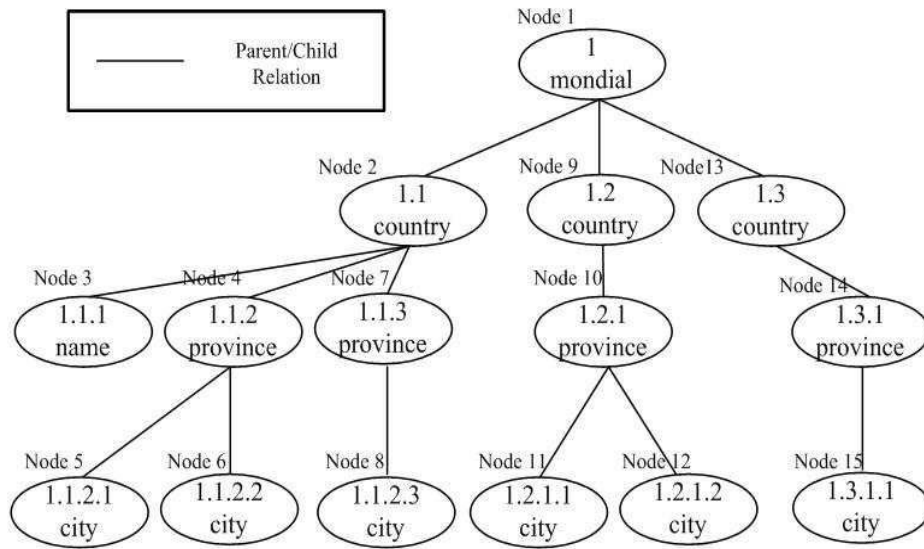
- [0171] 도 9는 본 발명의 바람직한 실시예에 따른 속성에 대한 조건식을 포함한 단순 경로 질의 (/mondial/country/province[@name="Aland"/city/text()]) 에 대한 질의 트리 탐색 과정(a)과 XML 데이터 스트림의 선택적인 수신 과정(b)을 도시한 것이다.
- [0172] 도 10은 본 발명의 바람직한 실시예에 따른 분기 질의(/mondial/ country[name]/province/city)를 처리하는 예를 도시한 것이다.
- [0173] 도 11은 원본 XML 문서, Path Summary, SL, SD, SP, 그리고 본 발명에 따라 생성된 XML 스트림의 크기를 비교한 것이다.
- [0174] 도 12는 접근 시간을 측정한 결과를 도시한 것이다.
- [0175] 도 13은 튜닝 시간을 측정한 결과를 도시한 것이다.

도면

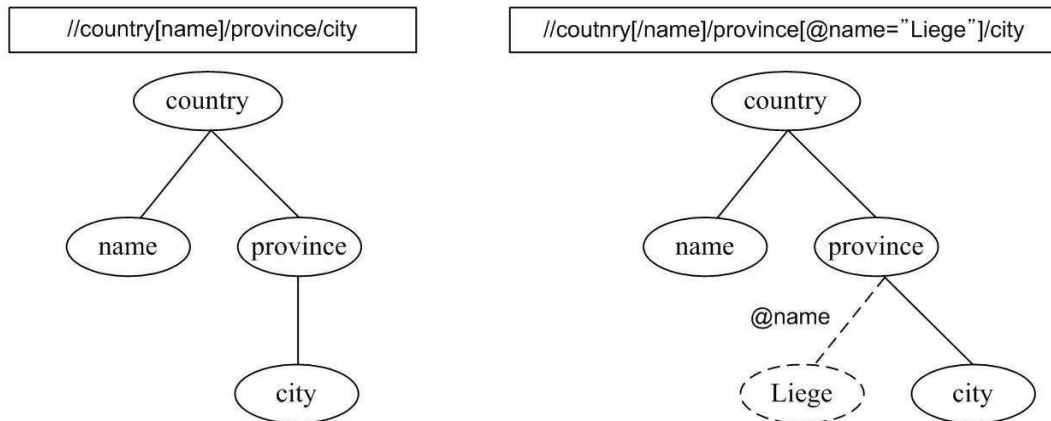
도면1a

```
<mondial>
  <country id="f0_162" name="Belgium" capital="f0_1477" population="10170241">
    <name>Belgium</name>
    <province id="f0_17462" name="Hainaut">
      <city id="f0_2335" country="f0_162" province="f0_17462">Charleroi</city>
      <city id="f0_2345" country="f0_162" province="f0_17462">Mons</city>
    </province>
    <province id="f0_17464" name="Liege">
      <city id="f0_2323" country="f0_162" province="f0_17464">Liege</city>
    </province>
  </country>
  <country id="f0_184" name="Czech Republic" capital="f0_1493" population="10321120">
    <province id="f0_17473" name="Jihomoravsky">
      <city id="f0_2371" country="f0_184" province="f0_17473">Brno</city>
      <city id="f0_2394" country="f0_184" province="f0_17473">Zlin</city>
    </province>
  </country>
  <country id="f0_208" name="Finland" capital="f0_1507" population="5105230">
    <province id="f0_33615" name="Aland" country="f0_208">
      <city id="f0_35399" country="f0_208" province="f0_33615">Marichamn</city>
    </province>
  </country>
</mondial>
```

도면1b



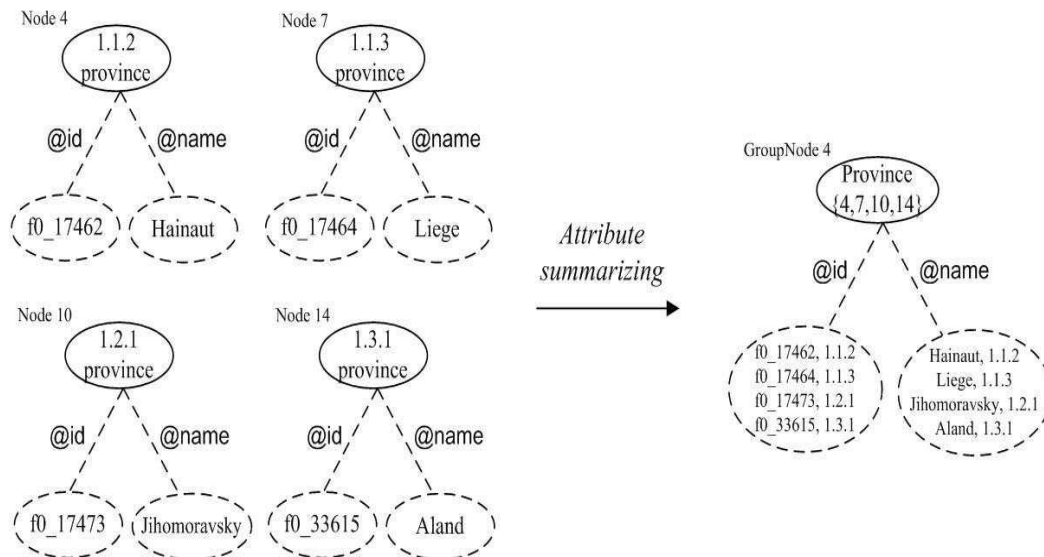
도면2



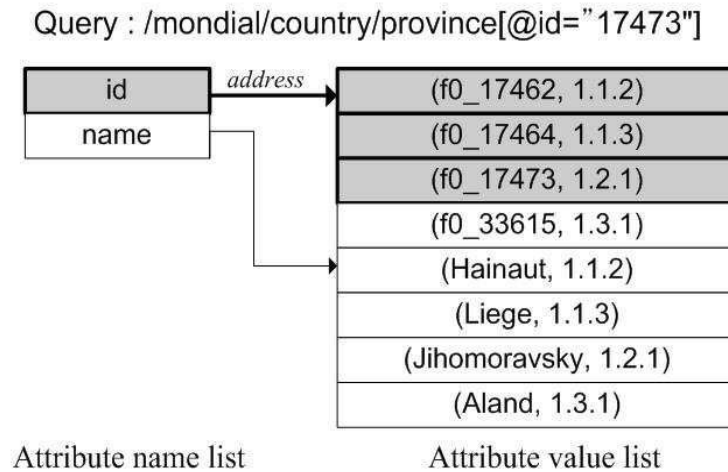
도면3

province		Group name	} Index
/mondial/country/province		Location path	
(city, address)		Child addresses	
1111		P-C Bitstring	
2.1.1		Packing Information	
(id, address), (name, address)		Attribute name list	
address		Text address	} Data
{{(f0_17462, 1.1.2), (f0_17464, 1.1.3), (f_17473, 1.2.1), (f0_33615), 1.3.1}}		Attribute value list	
{{(Hainaut, 1.2.2), (Liege, 1.2.3), (Aland, 1.3.1)}}		Text list	
null			

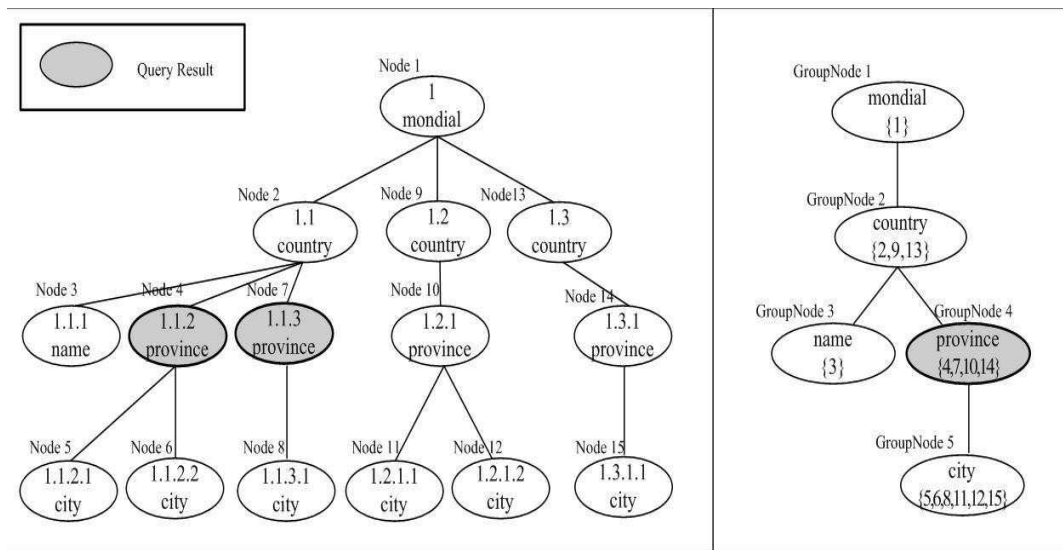
도면4



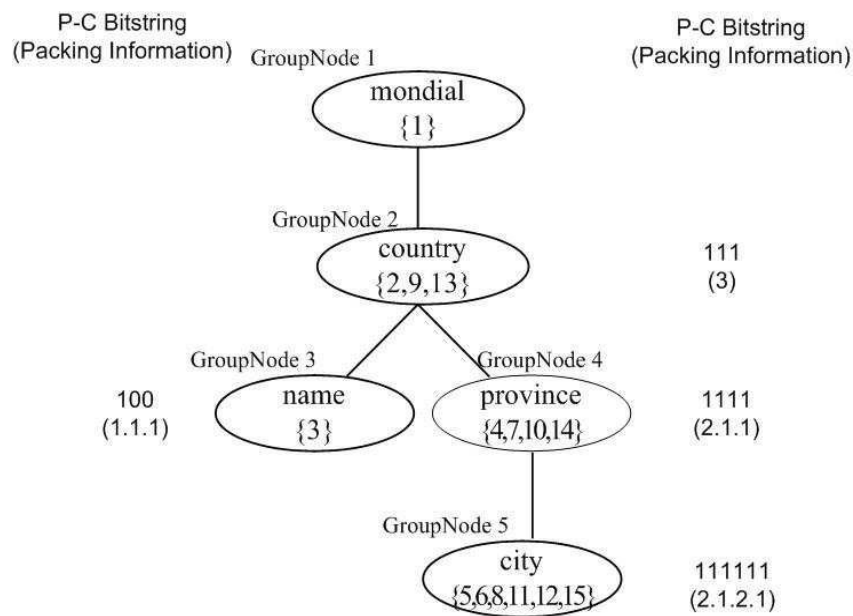
도면5



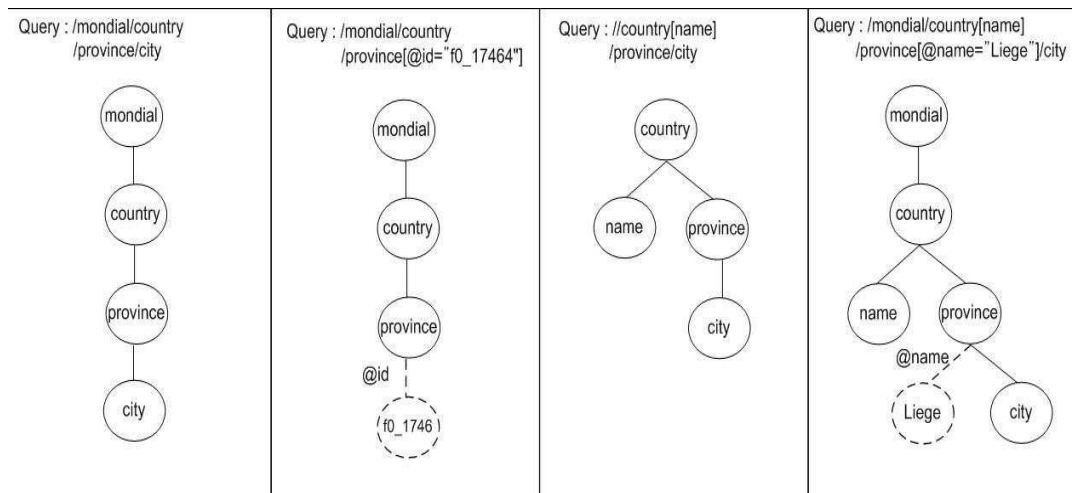
도면6



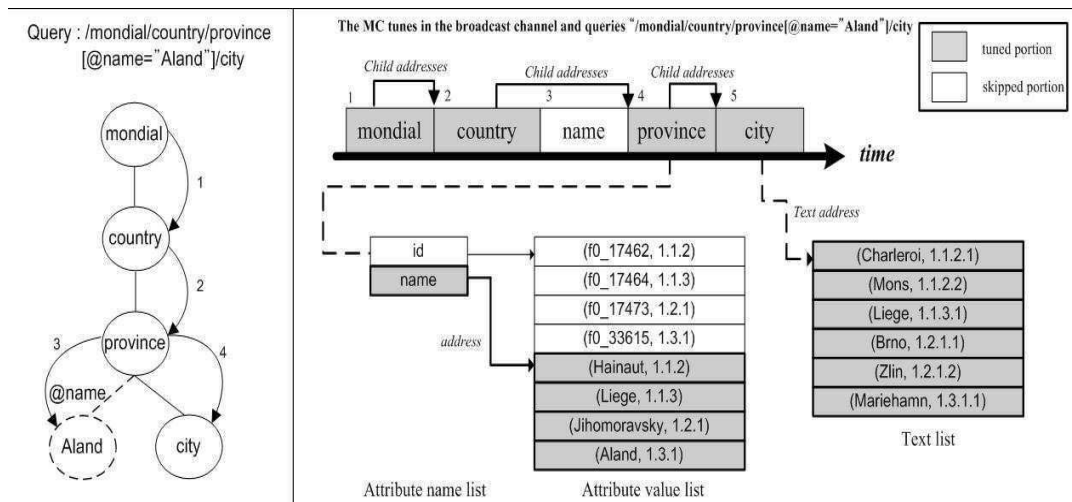
도면7



도면8



도면9



도면10

Traversal Phase

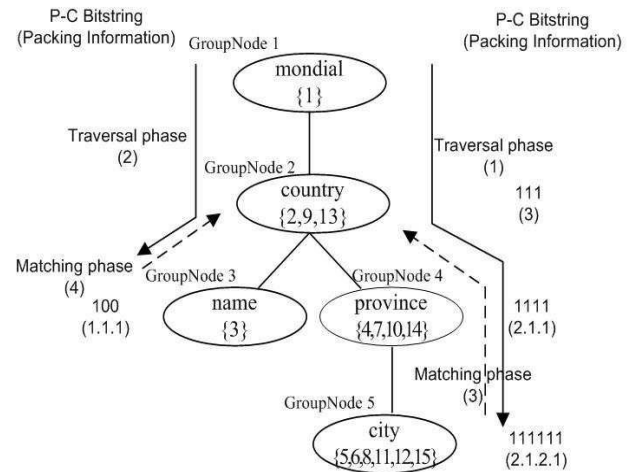
- (1) main path traversal
: mondial - country - province - city
(2) sub path traversal
: name

Matching Phase

- (3) main path matching
: city -> province
111111 & 1111 => 1111 & 1111=1111 ... ①
//packing 11 1 11 1 into 1111 by (2.1.2.1)
: province -> country
① & 111 = 1111 & 111 => 111 & 111=111 ... ②
//packing 11 1 1 into 111 by (2.1.1)

- (4) sub path matching
name -> country
100 & 111 = 100 ... ③

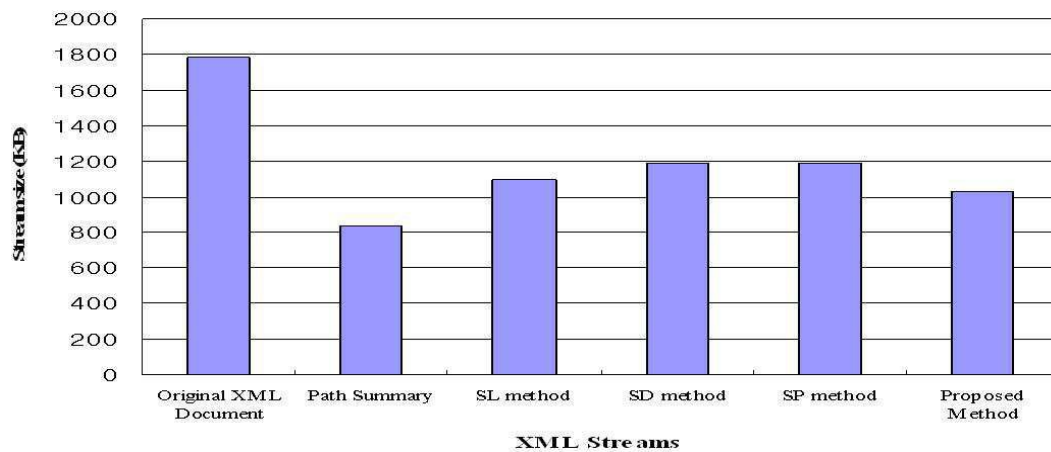
At branch node (country)
Main path & sub path = ② & ③ = 111 & 100 = 100



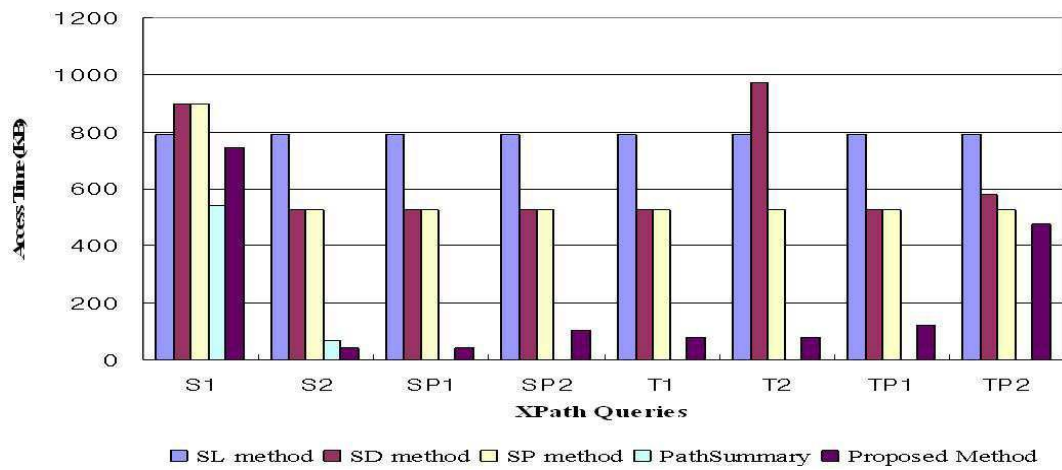
Retrieval Phase

country -> province : 100 is unpacked to 11 0 0 by (2.1.1)
province -> city : 1100 is unpacked to 11 1 00 0 by (2.1.2.1)

도면11



도면12



도면13

