

전자상거래를 위한 확장된 디지털 카탈로그 및 질의 모델 제안

정지혜 이상구 우치수

서울대학교 전산과학과

stop@europa.snu.ac.kr sglee@europa.snu.ac.kr wuchisu@selab.snu.ac.kr

Extended Electronic Catalog for dynamic and flexible Electronic Commerce

Jihye Jung Sang-goo Lee Chisu Wu

Dept. of Computer Science Seoul National Univ.

World Wide Web은 하이퍼미디어라는 뛰어난 사용자 인터페이스 기능을 제공함으로써 인터넷을 대중화 시켰고, 전자상거래라는 인터넷의 상업화도 가능하게 하였다. 이와 같은 전자상거래에서 필수적인 기술 중의 하나는 사용자가 원하는 상품의 카탈로그를 쉽고 빠르게 찾는 것이다. 본 논문의 목적은 전자 카탈로그를 정의하고 질의하는 모델을 제안하여 전자 카탈로그 시스템을 보다 쉽게 구축하고 유지하며 사용자의 요구사항을 만족하는 상품에 대한 카탈로그를 보다 쉽게 검색할 수 있는 방법을 제공하는 것이다. 본 모델의 주된 아이디어는 상품에 대한 정보를 표현하는 기존의 카탈로그와 계층적 검색을 위해 존재하는 분류체계를 통합하여 하나의 객체로 정의하고 그 객체에 대한 질의 언어를 정의함으로써 기존의 카탈로그에 대한 개념을 확장하여 전자 카탈로그 시스템 전반에 관한 검색을 용이하도록 하는 것이다. 확장된 카탈로그는 하나의 객체로 정의되기 때문에 질의에 의해 전체를 필터링해서 일부분만을 보여주거나, 사용자가 원하는 형태로 재구성하는 것이 가능하게 된다. 이를 위해 본 모델에서는 상품에 대한 정보를 그래프 형태로 정의하고 그들을 계층적으로 분류하는 분류 체계에 대해서 설명한다. 그리고 확장된 카탈로그를 각 상품과 카테고리들 노드로 한 그래프 형태로 정의하고 그에 대한 질의어를 제안한다.

1. 서론

상품을 사기 위해서 사람들은 직접 매장을 방문하거나, 홈쇼핑이나 전화주문 등 다양한 방법을 이용한다. 그와 같은 방법 중 최근 가장 주목을 받는 것이 전자상거래이다. 전자상거래에서는 각 쇼핑몰이 인터넷상에 판매할 상품이나 서비스에 대한 정보를 올리고, 소비자들은 그와 같은 정보를 바탕으로 원하는 물건을 선택해서 전자적인 거래를 통해 물품을 구매한다. 상품의 구매에 있어서 첫 번째 단계는 상품에 대한 정보를 수집하는 것이므로 보다 쉽고 빠르게 상품에 대한 정보에 접근할 수 있도록 해 주는 기술은 전자상거래에서 필수적이며 중요한 기술이라고 할 수 있다. 본 논문의 목적은 확장된 전자 카탈로그를 정의하고 질의하는 모델을 제안하여 전자 카탈로그 시스템을 보다 쉽게 구축하고 유지하며 사용자의 요구사항을 만족하는 상품에 대한 카탈로그를 보다 쉽게 검색할 수 있는 방법을 제공하는 것이다.

전자 카탈로그는 한 상품을 설명하기 위해 하나 이상의 페이지를 이용하여 그 정보를 표현한 웹 문서들의 집합이다. 이와 같은 전자 카탈로그에 대한 연구는 대부분 상품의 효과적인 표현과 빠른 검색에만 초점을 맞추고 있었다. 하지만 소비자는

상품에 대한 정보만이 아니라 분류체계와 같은 정보도 필요로 한다. 예를 들어 사용자가 자신의 요구사항을 명확히 모를 때는 직접적으로 상품에 대한 검색을 하기보다는 분류정보를 이용하여 계층적인 검색을 통해 최종적으로 자신이 어떤 상품을 원하는지 탐색해 갈 수 있다. 각 분류들의 이름이나 그것들이 포함하고 있는 상품들은 사용자들에게 탐색의 실마리를 제공할 수 있고, 그와 같은 정보를 바탕으로 소비자는 효과적으로 자신의 마음에 드는 상품을 선택할 수 있는 것이다.

최근 전자상거래를 제공하는 쇼핑몰들이 증가함에 따라 각 상품에 대한 카탈로그를 생성하는 비용을 줄이기 위해 전자 카탈로그 저장소를 만들어서 생산자는 상품에 대한 정보를 제공하고 각 쇼핑몰은 전체 카탈로그 중 자신의 요구에 부응하는 일부분을 다운로드 받아서 자신의 쇼핑몰에 맞게 변형시켜 사용하고자 하는 시도가 있다.[3] 이와 같은 공용 카탈로그 저장소가 제공되는 경우는 여러 다양한 분야의 수많은 카탈로그를 관리하고 검색하기 쉽도록 효율적인 분류 체계가 제공되고, 각 쇼핑몰에서는 그 중 일부분만을 사용하게 된다. 이와 같은 경우 상품 카탈로그만 검색해서 사용하는 것 보다 전체 구조를 일부분이나 주어진 조건에 의해 필터링한 뷰를 가져와서 사용자가 원하는 형태로 재구성하는 것이 필수적인 기능으로 제공되어야 한다. 이를 위해 본 논문에서는 기존의 카탈로그와 계층적 검색을 위해 존재하는 분류 체계를 통합하여 하나의

* 본 연구는 정보통신부가 주관하는 산·학·연 공동기술개발사업인 1999년도 "전자상거래를 위한 상품카탈로그의 표준 데이터베이스 구축 및 원격 입출력 기술개발 연구"에 의한 결과임 (과제번호:A2-98-106-00)

객체로 정의하고 그 객체에 대한 질의 언어를 정의함으로써 기존의 카탈로그에 대한 개념을 확장한다. 그리고 확장된 카탈로그가 보다 동적이고, 유연하며 유지가 용이하다.

본 논문의 구성은 다음과 같다. 2장에서는 상품 정보를 정의할 수 있는 전자 카탈로그 모델을 정의하고, 3장에서는 계층적 검색을 위한 분류 체계와 확장된 카탈로그 모델에 대하여 설명한다. 그리고 4장에서는 확장된 카탈로그를 검색하고 조작하기 위한 질의어의 종류와 예들을 살펴보고, 끝으로 5장에서 결론을 기술한다.

2. 전자 카탈로그 모델

전자 카탈로그는 상품을 효과적으로 설명하는 속성이나 그림, 설명문 등을 이용하여 상품을 표현하는 웹 문서의 집합이다. 따라서 하나의 상품을 그 상품을 설명하는 속성의 집합으로 정의할 수 있다. 이러한 방법은 지금까지의 상용 제품들 - Internet Commerce Server, LiveCommerce, Commerce21 - 이나 기존 연구에서 전자 카탈로그를 바라보는 입장이었다. 하지만 상품을 속성의 집합으로만 정의하면 컴퓨터와 같은 다른 제품들을 결합하여 만들어진 상품을 설명하기가 힘들다. 예를 들어 개인용 컴퓨터에 관한 카탈로그는 메인보드, CPU, 메모리, 하드 디스크 등 여러 부속에 대한 정보를 포함하고 있어야 하는데, 각각의 부속들도 그들을 설명하는 속성을 가지기 때문에 단순한 속성의 집합만으로는 그 특성을 효과적으로 설명하기 힘들다. 이와 같은 문제점을 해결하기 위해 본 논문에서는 상품을 semi-structured 모델과 유사한 형식의 그래프 모델로 정의한다.

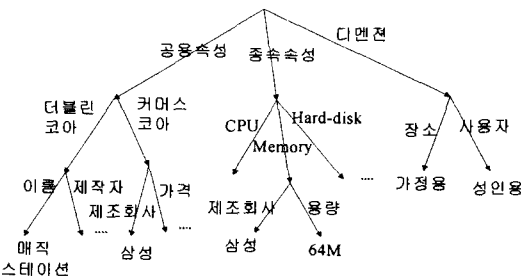


그림 1 개인용 컴퓨터를 설명하는 카탈로그 그래프의 예

위의 그림1은 개인용 컴퓨터를 설명하기 위한 카탈로그 그래프의 예이다. 그래프의 간선은 각 속성을 나타내고, 말단 노드는 속성의 값을 나타낸다.

이와 같이 그래프 모델로 정의되는 상품 정보에 대해 질의를 효과적으로 수행하기 위해서는 각 상품별로 속성을 정의할 필요가 있다. 각 상품 별 속성들을 효율적으로 정의하기 위해 우리는 모든 상품에 공통적인 공통 속성을 정의하고 유사한 상품들을 묶어 하나의 상품군 - TV, 냉장고, 전자레인지 등 - 으로 지정하여 각 상품군 별로 종속 속성을 정의하였다. 그리고 상품을 여러 다른 여각도에서 비교하는 것을 용이하게 하

기 위해서는 각 상품별로 시간, 장소, 사용자등의 디멘전 정보도 정의하였다. 이와 같은 디멘전 정보는 특히 사용자들이 자신이 원하는 상품에 대한 요구사항이 명확하지 않을 때 유용하게 사용될 수 있으므로 상품을 분류하는데 있어서 필수적인 요소이고 가장 많이 검색되는 정보이다. 이에 대한 자세한 사항은 [3] 논문에서 제안한 바 있다.

3. 확장된 전자 카탈로그 모델

소비자들이 원하는 상품의 카탈로그를 검색하는 방법은 일반적으로 키워드 검색, 속성별 검색, 계층적 검색의 세 종류가 있고, 계층적 검색을 제공하기 위해서는 공통적인 속성을 가지는 상품들을 여러 카테고리 분류하고 이들 카테고리를 계층적으로 구성한 분류 체계가 필요하다. 이와 같은 분류체계는 대부분의 쇼핑몰에서 트리 형태로 제공되고 있고, 각 상품들은 자신이 어느 카테고리에 속하는 지의 정보를 가지고 있거나, 각 카테고리가 해당 상품에 대한 링크를 모두 가지고 있는 것이 대부분이다.

하지만 쇼핑몰에서 검색을 하는 소비자들은 구매 패턴이나 관심의 깊이가 상이하기 때문에, 하나의 분류체계만으로 모든 사용자를 만족시킬 수 없으므로 여러 개의 다른 분류체계를 제공해 줄 수 있어야 한다.

예를 들어 운동복 상의를 사고자 하는 사람의 경우,

- 가) 브랜드를 먼저 선택하고, 상품의 종류를 선택한 후, 사이즈를 고르는 경우가 있고,
나이키 → 운동복 상의 → 중간 사이즈
- 나) 사이즈를 먼저 선택하고, 상품의 종류를 선택한 후, 브랜드를 고르는 경우가 있다.
중간 사이즈 → 운동복 상의 → 나이키

이와 같이 최종 카테고리에 도달하기 위한 경로를 여러 개 두기 위해서 분류체계는 트리 구조가 아닌 그래프 구조가 되어야 한다.

위의 예와 같이 상품의 계층적 검색을 위한 카테고리는 사이즈나 브랜드, 상품군의 종류, 사용자 층, 사용되는 장소등의 여러 다양한 속성에 의해 정의될 수 있다. 예를 들어 나이키에서 판매하는 테니스화는 상품군 = "운동화" and 브랜드 = "나이키" and 목적 = "테니스" 와 같이 정의될 수 있다. 이와 같은 규칙을 각 카테고리 별로 부여하면 각 상품이 소속 카테고리에 대한 정보를 가질 필요도 없고, 각 카테고리들이 모든 소속 상품들에 대한 링크를 가질 필요도 없게된다. 따라서 분류 체계의 변경도 보다 쉬워지고 유연해진다.

따라서 분류 체계는 각 카테고리를 노드로 하고 카테고리 사이의 관계를 간선으로 하는 DAG로 표현이 된다. 카테고리 사이의 관계에는 테니스화와 운동화 사이에 존재하는 상하관계와 운동화와 운동 기구 사이에 존재하는 연관관계가 있다. 이와 같이 DAG로 표현되고 각 카테고리가 규칙을 가지는 분류 체계에 대한 상세한 모델은 [4]에서 제안하였다. 그림 2는 제안된 모델에 의해 정의된 분류 체계의 한 예이다.

일반적으로 이와 같은 분류 체계와 상품을 설명하는 전자 카탈로그는 따로 관리가 되었고 분류 체계는 상품을 검색하기 위한 틀로써만 존재하였다. 하지만 기존의 전자 카탈로그와 계층적 검색을 위해 존재하는 분류 체계를 통합하여 하나의 객체로 정의하게 되면 질의에 의해 전체 구조를 필터링해서 일

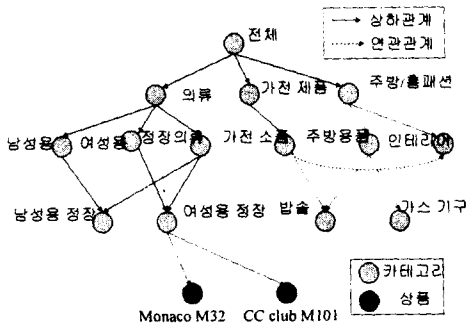


그림 2 분류 체계의 예

부분만을 보여주거나, 사용자가 원하는 형태로 재구성하는 것이 가능하게 된다. 그리고 다른 곳에서 만들어진 카탈로그와 분류체계를 자신의 분류체계에 통합하는 데도 유용하게 사용될 수 있다. 또한 각 카테고리를 설명하는 카탈로그 - 카테고리 위한 웹 문서들도 본 논문에서는 전자 카탈로그로 본다. -에서 소속 상품의 정보나 하위 카테고리의 정보를 추출하여 사용하는 것도 가능하게 된다. 따라서 본 논문에서는 전자 카탈로그의 범위를 하나의 상품을 설명하는 웹 문서의 집합에서 전체 상품 카탈로그와 그들을 분류하는 분류체계를 통합한 전체 구조로 확장하고 이들을 하나의 객체로 모델링하여 보다 능적이고, 유연하며 유지가 용이하도록 한다. 확장된 카탈로그는 앞서 설명한 분류 체계의 모습을 따르며 각 노드는 카테고리나 상품을 설명하는 카탈로그를 나타내고 각 간선은 그들 사이의 링크를 보여준다.

4. 확장된 전자 카탈로그를 위한 질의 언어

위와 같이 정의된 확장된 전자 카탈로그에 대한 첫 번째 질의는 사용자의 요구 사항에 부합하는 상품 카탈로그나 카테고리 카탈로그를 검색하는 것이다. 2장에서 설명한 바와 같이 상품 카탈로그나 카테고리 카탈로그의 내용은 그래프 형태로 표현이 되어있다. 그래프 형식은 다른 제품을 결합하여 만들어진 컴퓨터와 같은 제품을 표현하는 데는 효과적이지만 질의하기가 어려운 면이 있다. 그래서 본 논문에서는 사용자들이 속성별 검색과 유사한 방식으로 카탈로그를 검색할 수 있도록 브랜드(c1), 상품군(c1), 제조회사(c1)과 같은 외부함수를 정의할 수 있도록 한다. 아래는 "삼성"에서 만든 "개인용 컴퓨터"에 관한 모든 상품 카탈로그를 얻는 질의이다.

```
Select 삼성컴퓨터(p)
From Catalogs AllCatalogs(E1) as p
Condition 제조회사(p) = "삼성" AND
상품군(p) = "개인용 컴퓨터"
```

위의 질의에서 E1은 어떤 확장된 카탈로그 객체를 나타내고, AllCatalog(E1)은 E1에 존재하는 모든 상품 카탈로그와 카테고리 카탈로그를 나타내는 함수이다.

또한 확장된 카탈로그 객체를 일부분이나 어떤 조건에 의해 필터링을 한 뷰를 얻기 위해서 아래와 같은 형식으로 질의를 할 수 있다.

Select Subgraph E2

From Graph E1

Start 101

이 질의는 의류의 노드 번호가 "101"일 때 의류로부터 시작하는 서브 그래프를 얻어서 E2로 지정하는 것이다.

Select Subgraph E3

From Graph E1

Consistent 장소 = "가정용"

위의 질의는 확장된 카탈로그의 뷰를 얻는 질의로써 가정에 쓰이는 것들만을 얻고자 할 때 사용하게 되고, 그 결과는 그림 2에서 "가전제품", "주방/홈패션", "가전소품", "주방용품", "발수", 그리고 "가스기구" 만으로 이루어진 그래프와 같다. 필터링을 할 때는 사용자에 의해 주어진 조건이 각 카테고리의 규칙에 위배되지 않는 지와 각 상품의 속성에 부합하는지를 바탕으로 이루어진다. 따라서 각 카테고리를 규칙을 이용하여 표현하지 않으면 이러한 질의는 수행될 수 없다.

마지막으로 카탈로그 내용의 일부를 추출하여 새로이 재구성하기 위한 질의는 다음과 같이 사용할 수 있다.

Select \$1, \$m

From 삼성컴퓨터(p)

Where

```
<공용속성> <더블런코어> <이름> </> Element_As $1 </></>
<종속속성> <memory> <용량> $c </></> Element_As $m </>
IN p, $c > 64M
```

Construct <result> \$1 \$m</result> AS Unique MemoryPage

위의 질의는 상품 정보가 그래프 형태로 표현되므로, 유사한 구조 위에서 정의된 XML-QL[2]을 적용한 것이다. 상품이나 카테고리 카탈로그에서 정보를 추출하여 재구성하는 질의는 기존에 존재하는 확장된 카탈로그에 외부에서 가져온 확장된 카탈로그를 통합하는 데 유용하게 사용될 수 있다.

5. 결론

본 논문에서는 상품 정보를 공용속성, 종속속성, 디멘션을 이용하여 그래프 형태로 정의하고, 계층적 검색을 위한 분류 체계를 정의한 후 그 두 가지를 통합한 확장된 카탈로그를 정의하였다. 그리고 확장된 카탈로그에 대하여 네가지 종류의 질의를 정의하였다.

이와 같은 확장된 카탈로그 모델을 사용하면 기존의 ICS나 LiveCommerce와 같은 시스템보다 유연하고 동적이며 통합이 쉬운 카탈로그 시스템을 구축할 수 있으며, WebOQL이나 Strudel, ARANEUS에서 제공하는 웹 사이트 유지 보수의 기능[1]도 지원할 수 있다.

참고문헌

[1] A. O. Mendelzon, G. A. Mihaila, T. Milo, Querying the World Wide Web, Int. Conf. On Parallel and Distributed Information Systems, 1996
 [2] XML-QL: A Query Language for XML, Submission to the W3C 19-August-1998, available via http://www.w3.org/TR/NOTE-xml-ql/
 [3] Sang-goo Lee, Chisu Wu, Kapsoo Kim, Dongkyu Kim, Woosung Shin, Digital Catalog Library: A Shared Repository of Online Catalogs for Electronic Commerce, WECWIS 98, 1998
 [4] 정지해, 이상규, 우치수, 전자상거래에서의 체계적인 상품 카탈로그 구축을 위한 분류 체계 모델 및 구현, 99 한국 데이터베이스 학술대회, 1999