

UML 기반 OLAP 메타 데이터의 다이아그램 모델 설계

김 경 주* , 이 윤 배
조선대학교 대학원 전자계산학과
e-mail: *smile_99 @ naver.com

UML based Design of OLAP Meta Data Diagram Model

Kyung-ju Kim, Yun-bae Lee
Dept of Computer Science, Graduate School, Chosun University

요 약

데이터 웨어하우스(Data Warehouse : DW)는 데이터베이스에 저장되어 있는 데이터를 신속한 의사 결정 지원을 위해 최종 사용자가 여러 곳의 기업 내에 흩어져 있는 방대한 데이터를 손쉽게 빠르게 접근할 수 있도록 활용되고 있다. 현재 데이터 웨어하우스의 중요성이 부각되고 있는 가운데 온라인 분석 처리(On Line Analytical Processing : OLAP) 시스템이 데이터 웨어하우스 안에서 활용되고 발전되고 있다.

기존 연구에서는 서로 다른 OLAP 제품에서 공통으로 사용할 수 있는 모델을 적용하여 OLAP메타 데이터 교환 시스템을 설계해왔다. 그러나 본 논문에서는 서로 다른 OLAP 제품을 공통으로 사용할 수 있는 질의 언어 시스템 설계 전 단계인 논리적 설계를 UML snowflake 다이어그램을 이용하여 설계 하였다.

실험결과, XML 문서의 변환된 OLAP 메타 데이터를 이용하여 UML snowflake 다이어그램 설계를 통해 통합된 OLAP 제품의 XML 문서 구조가 논리적으로 설계되어 메타 데이터가 통합됨을 알 수가 있다.

1. 서 론

데이터 웨어하우스(Data Warehouse : DW)는 데이터베이스에 저장되어 있는 데이터를 신속한 의사 결정을 하기 위해 최종 사용자가 여러 곳의 기업 내에 흩어져 있는 방대한 데이터를 손쉽게 빠르게 접근할 수 있도록 활용되고 있다. 이와 같은 데이터 웨어하우스는 일관적이고, 주제 지향적이고, 통합적이며, 비 유동적인 데이터로써 이것을 바탕으로 사용자가 예측할 수 있고 결정할 수 있도록 의사 결정 정보를 제공한다[1].

현재 데이터 웨어하우스의 중요성이 대두되고 있는 가운데 온라인 분석 처리(On Line Analytical Processing : OLAP) 시스템은 데이터 웨어하우스

환경과 연관되어 데이터 웨어하우스에 대해 직접 수행할 수 있다. 그리고 의사 결정을 위한 데이터를 위해 최종 사용자들이 다차원적인 정보에 데이터 분석과 복잡한 질의 처리를 효율적으로 하도록 도와준다. 원래의 데이터에서 변경된 정보의 모든 가능한 것들이 다양성과 일관성 있게 빠르게 상호 작용하여 접근하도록 함으로써 데이터 웨어하우스의 활용으로 발전되고 있으며 데이터베이스의 분야에 중요한 비중을 차지하고 있다[3].

본 논문에서는 기존 연구에서 서로 다른 OLAP 제품에서 공통으로 사용할 수 있는 모델을 적용하여 OLAP메타 데이터 교환 시스템을 설계했지만 이 연구를 확장하여 서로 다른 OLAP 제품을 공통으로 사용할 수 있는 질의 언어 시스템 설계 전 단계의 논리적 설계를 OLAP 메타 데이터를 이용하여

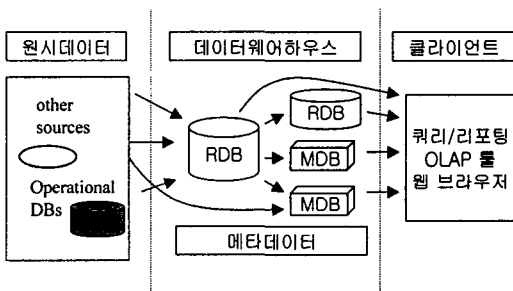
UML snowflake 다이어그램을 설계하였다.

2. 관련 연구

2.1 데이터 웨어하우스

데이터 웨어하우스는 의사 결정을 하기 위한 데이터의 집합으로서 외부의 DB에서 생긴 데이터를 주제별로 통합하여 여러 차원과 각도로 분석을 가능하게 하는 의사 결정 통합 시스템이다[7]. 데이터 웨어하우스의 구조는 원시 데이터 계층, 데이터 웨어하우스 계층, 클라이언트 계층으로 [그림 2.1]와 같이 나눌 수 있다.

[그림 2.1]은 데이터 웨어하우스의 구조를 나타낸 것으로 원시 데이터 계층은 외부에 저장되어 있는 데이터들이 변환하여 데이터 웨어하우스로 적재되는 계층이다. 그리고 데이터 웨어하우스 계층은 원시 데이터로부터 전송된 데이터들이 축적되어 사용자들이 필요한 데이터들을 사용하고 이 계층에 포함된 메타데이터는 데이터 웨어하우스 환경 내에서 중요한 데이터로써, 데이터에 대한 모든 정보가 들어있다. 클라이언트 계층에서는 이 데이터 웨어하우스에서 뽑아낸 데이터들을 분석하는 도구로 쿼리/리포팅, OLAP 툴, 웹 브라우저 등을 속할 수 있다.



[그림 2.1] 데이터 웨어하우스 구조

2.2 OLAP (On Line Analytical Processing)

2.2.1 OLAP 정의

온라인 분석 처리(On-Line Analytical Processing : OLAP) 시스템은 데이터 웨어하우스 환경과 연관되어 데이터 웨어하우스에 대해 직접 수행될 수 있으며, 데이터에 대해서 의사 결정을 위해 최종 사용자들이 다차원적인 정보에 데이터 분석과 복잡한 질의 처리를 효율적으로 하도록 도와준다. 원래의 데이터에서 변경된 정보의 모든 가능한 것들이 다양성과 일관성 있게 빠르게 상호 작용하여 접근하도록 함으로써 데이터 웨어하우스의 활용으로 발전되고 있으며 데이터베이스의 분야에 중요한 비중을 차지하고 있다[3].

2.2.2 데이터 큐브(Data Cube)

데이터 웨어하우스와 OLAP은 다차원 데이터모델을 기반으로 나타내는데, 다수의 차원 정보가 데이터 큐브를 통해 나타낸다. 다차원 데이터에는 여러 개의 차원(Dimension)정보와 하나의 사실(fact)정보로 이루어진다.

차원 정보는 데이터 레코드 상에 있는 이유의 대상이 되는 것이고, 사실 정보는 숫자적으로 표현되는 값이다. 예를 들어 자동차 회사의 자동차 판매 상황을 보면 차의 판매 시기, 자동차의 종류, 차가 팔리는 매장 등의 차원 정보와 차가 팔리는 판매액과 판매량 등의 사실 정보로 구성된다[4][6].

2.2.3 star schema

스타 스키마는 논리적인 모델로, 구현하기 전의 단계로서 RDBMS의 다차원 데이터를 스타 스키마를 이용하여 논리적으로 모델화 한 것이다. 스타 스키마는 중앙 중심에 필요한 하나의 사실 테이블로 구성되어 유일한 정규화된 테이블로 이루어진다. 이 테이블 주변에는 이 사실 테이블의 부가적인 정보들로 이루어진 여러 개의 차원 테이블이 있다.

2.2.4 Snowflake schema

snowflake schema는 스타 스키마의 확장된 개념이다. 이 스키마는 스타 스키마와 동일한 구조를 가지지만 차원 테이블에서 정규화가 되어 테이블들이 눈꽃 모양이 되도록 만든 스키마로 다차원 계층 구조를 가지고 있다.

III. OLAP 메타 데이터의 UML Diagram 모델 설계

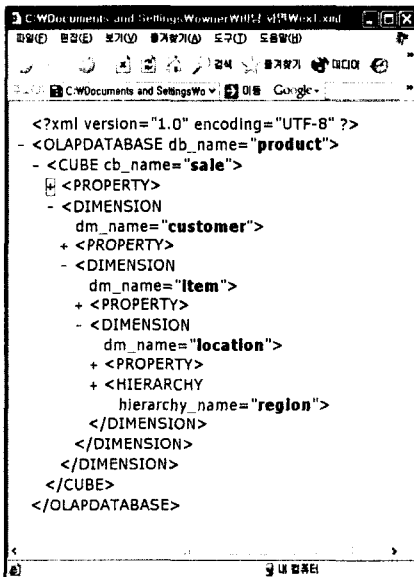
서로 다른 OLAP 제품의 메타 데이터를 통합하여 공통으로 사용할 수 있는 질의 언어 시스템 설계 전 단계의 논리적 설계로, 통합된 OLAP 메타 데이터를 이용하여 UML snowflake 다이어그램을 설계한다.

3.1 메타 데이터의 기본 요소

전체적인 OLAP 메타 데이터의 모델의 기본적인 요소인 Database, Cube, Dimension, Measure, Hierarchy, Level, Member, Attribute 으로 구성된다.

3.2 메타 데이터에서의 XML 문서 변환

OLAP 메타 데이터의 기본 요소를 가지고 서로 다른 OLAP 제품의 메타데이터를 통합하여 OLAP 메타데이터를 이용하여 [그림 3.1]와 같이 통합된 XML 문서를 변환한다.



[그림 3.1] 변환된 XML 문서

3.2.1 UML Snowflake Diagram 모델 생성

XML에서 추출되어진 메타 데이터를 이용하여 공통 언어의 논리적으로 설계하기 위해 UML Snowflake 다이어그램을 이용하여 설계하였다.

UML Snowflake Diagram의 생성은 다음과 같다.

UML Snowflake Diagram의 클래스는 사실 클래스로 구성된 루트와 서브클래스로 구성된 차원 클래스로 이루어진다. [표3.1]는 Snowflake Diagram의 요소를 이룬다.

Snowflake Diagram = {(class_name, attributeList, PointTo)}

[표3.1] snowflake diagram 의 기본 요소

class_name은 클래스에서 식별할 수 있는 이름이고, attributeList는 클래스안에서 관련된 속성의 집합이다. 그리고 PointTo는 diagram 사이에서의 클래스간 연결 할 수 있는 요소이다. [표3.2]는 AttributeList 와 PointTo의 구성요소 나타낸 것이다.

pointTo = {(class_name, link_type, source_cardinality, target_cardinality | source_cardinality ∈ {0, *, 1, *}, target_cardinality = 1}
 attributeList = {(at_name, at_type, modifier, AD, calc_att, link_path_type, 집계함수 | at_type ∈ {DataType ? modifier ∈ {?, 1}} ? type ? type ∈ {속성, descriptive} ? 집계함수 ∈ {SUM, COUNT, MIN, MAX, AVG, NULL}}

[표3.2] attributeList 와 PointTo의 구성요소

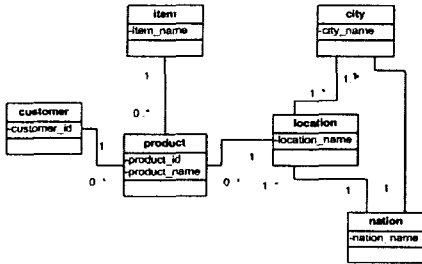
이에 따른 요소들은 XML 문서 구조에 있는 데이터를 스노우 플레이크 다이어그램을 표현하기 위해 앞에서 정의한 요소로 차원 들의 AttributeList 요소와 PointTo의 요소 [표3.3]와 같이 나타낼 수 있다.

AttributeList _{customer} = {(id, NUMERIC, 1, URLcomponent: class/customer.id, NULL, NULL)} {(name, TEXT, 1, URLcomponent: class/customer.id, NULL, NULL)}
AttributeList _{Item} = {(id, NUMERIC, 1, URLcomponent: class/Item.id, NULL, NULL)} {(name, TEXT, 1, URLcomponent: class/Item.id, NULL, NULL)}
AttributeList _{location} = {(name, NUMERIC, 1, URLcomponent: class/Item.id, NULL, NULL)}
PointTo = { (customer, ASSOCIATION, 0, *), (Item, ASSOCIATION, 0, *), (location, ASSOCIATION, 0, *), }

[표3.3] 차원 AttributeList 와 PointTo

이런 요소들을 사용하여 [그림3.1]에 나타난 변환된 XML 문서의 데이터 소스 product에서 생성되어

진 3개의 차원 customer, location, item이 구성되며 UML snowflake Diagram을 [그림3.2]와 같이 나타낼 수 있다.

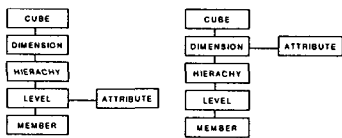


[그림 3.2] UML snowflake Diagram

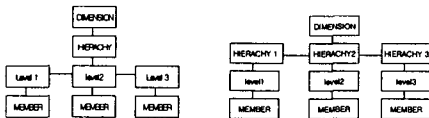
IV. 기존 연구와의 비교

4.1 서로 다른 OLAP 제품의 모델 적용과 OLAP 메타 데이터가 적용된 UML Diagram 모델 설계

서로 다른 OLAP 제품에서 공통으로 사용할 수 있도록 [그림4.1]과 [그림4.2]와 같이 속성과 차원 모델을 적용하였다. 이를 통해 OLAP메타 데이터 교환 시스템을 설계를 통해 통합된 OLAP제품의 메타 데이터의 XML문서를 형성하여 교환 할 수 있었다.



[그림4.1] 속성 모델



[그림 4.2] 차원 모델

그러나 통합된 OLAP 메타 데이터의 XML문서를 통해 서로 다른 OLAP 제품의 공통 언어를 생성할 수 있으며, 질의 언어 시스템 설계 하기 위한 논리적 설계로서, 이를 통해 OLAP 메타 데이터를 [그림

3.2]와 같이 논리적 모델을 설계함으로써 메타 데이터가 통합됨을 알 수가 있다.

V. 결론

본 논문에서는 서로 다른 OLAP 제품을 공통으로 사용할 수 있는 질의 언어 시스템 설계 전 단계인 논리적 설계로, OLAP 메타 데이터를 이용하여 논리적인 UML snowflake 다이어그램을 설계하였다. 이 설계를 통해 통합된 OLAP 제품의 XML의 문서 구조가 논리적으로 설계되어 메타 데이터가 통합됨을 알 수가 있었다.

향후 연구 과제로는 이 논리적 설계를 확장하여 OLAP 메타데이터를 기반으로 서로 다른 OLAP제품의 공통 질의 언어를 표현하기 위한 질의 기반 언어 시스템을 설계할 수 있도록 해야 한다.

참고 문헌

- [1] W.H.INmon. Building the Data Warehouse. John Wiley & Sons. 1996
- [2] Query-based Data Warehousing Tool, Rami Rifaieh, 2002
- [3] E.Thomsen, OLAP Solution, John Wiley & Sons, 1997
- [4] 데이터 웨어하우스 성능 관리를 위한 DBMax의 확장, 김은주, 용환승, 이상원, 정보처리학회, 2002
- [5] 웹 기반의 OLAP 메타 데이터 교환 시스템의 설계 및 구현, 이인기, 이민수, 용환승 정보처리학회, 2002
- [6] Pedersen extending OLAP Querying to external Object Database, In processing of the ninth International conference on information and knowledge Management 2000
- [7] Tapio Niemi, constructing an OLAP Cube from Distributed XML data, Proceedings of the 5th ACM international workshop on Data Warehousing and OLAP, 2002