

Drill-across 연산자를 이용한 Snowflake schema 개념 설계

김경주 · 오근탁 · 이윤배
조선대학교 대학원 전자계산학과

Design of Snowflake schema concept using Drill-across Operator

Kyung-ju Kim · Gun-tak Oh · Yun-bae Lee
Dept. of Computer Science, Graduate School, Chosun University
E-mail : diskerror @ hanmail.net

요 약

데이터 웨어하우스는 주제 지향적이고, 통합적이며, 비 유동적인 데이터로써 의사결정에 필요한 정보들을 추출하여 OLAP(On-Line Analytical Processing) 분석에 이용된다. 현재 OLAP 연산자와 스타 스키마 사이의 다차원 모델링에 관한 연구는 많이 진행되어 왔다.

본 논문에서는 drill-across 연산자를 통해 객체 지향 개념적 관계와 스타 스키마보다 한층 더 확장된 스노우 플레이크(snowflake) 스키마를 이용하여 관계를 설계하고자 한다. 설계를 통해 star schema에 적용되지 않았던 객체 지향 관계가 개선되었음을 알 수 있었다.

ABSTRACT

Data warehouse is subject-oriented, integrated, non-volatile data, and it used for OLAP(On-Line Analytical Processing) the extraction of information from making decision processing. In the present, lots of study have been devoted to multidimensional modeling between OLAP operator and star schema.

In this paper, the design of using the snowflake schema for object-oriented conceptual relation is more extended than using drill-across operator. The object-oriented relation schema which was not applicable has been improved through the design.

키워드

snowflake schema, drill-across operator, object-oriented conceptual, OLAP

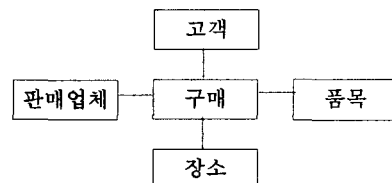
1. 서 론

데이터 웨어하우스는 일관적이고, 주제 지향적이고, 통합적이며, 비 유동적인 데이터로써 이것을 바탕으로 사용자가 예측할 수 있고 결정할 수 있도록 의사 결정 정보를 제공한다[1].

온라인 분석 처리(OLAP : On-Line Analytical Processing)는 데이터에 대해서 의사 결정을 위해 최종 사용자들이 다차원적인 데이터 분석과 복잡한 질의 처리를 효율적으로 하도록 도와준다. 그리고 원래의 데이터에서 변경된 정보의 모든 가능한 것들이 다양성과 일관성 있게 빠르게 상호 작용하여 접근하도록 한다. 또한 OLAP는 최종 사용자가 통합된 기업 데이터들과 향해(navigation) 활동들의 동적인 다차원을 분석하는데도 이용 할 수 있다[2].

최종 사용자는 롤업 (roll-up), 드릴다운(drill-

down),스크린(screening),스코핑(scoping), 슬라이싱(slicing),드릴어크로스(drill-across) 등을 통해 OLAP 연산자를 수행할 수 있다[3]. [그림 1]은 구매의 star schema에서 하나의 사실(Fact)과 차원(Dimension)이 주어진 star schema에서 분석할 수 있음을 보여주고 있다. 다차원 분석은 사실적인 데이터와 차원적인 데이터를 근거로 분석하고 있다.



[그림 1] 구매의 star schema

다차원 분석 중 각기 다른 스타 스키마 사이에서 "drill-across"를 이용하여 의미적 관련성을 나타내는 방법이 있다 [5][8]. 즉, 두개의 사실과 공유된 차원의 관계를 나타내는 것으로 단지 star 스키마와 관련성만을 나타낸다. 이 개념들을 발전시켜 drill-across 사이에서 객체지향 개념적 관계와 스타 스키마와의 의미적 관계를 나타낼 수 있다[4].

본 논문에서는 drill-across연산자를 통해 객체지향 관계와 스타 스키마보다 한층 더 확장된 스노우플레이크(snowflake) 스키마를 이용하여 의미적 관계를 설계한다.

II. 관련연구

2.1 OLAP 정의와 연산자 기능

다차원 데이터는 여러 개의 차원 데이터와 사실 값으로 구성된다. 예를 들면, 회사의 차 판매 상황을 보면 차의 판매 시기, 차의 종류, 차가 팔리는 매장 등의 차원 정보와 차가 팔리는 판매액과 판매량 등의 사실의 측정 값으로 구성 될 수 있다.

연산자들 중 Slice와 Dlice는 사용자가 어떤 부분을 볼 것인지에 대해서 정의하는 것이다. 드릴다운(Drill-down)은 데이터를 보다 더 구체적으로 상세시켜 단계적으로 접근하여 분석하는 기법이다. 드릴업(Drill-up)은 드릴 다운과 반대 방향으로 사용자가 정보를 분석하는 것이다. 드릴어크로스(Drill-across)는 서로 연관되어 있는 다른 큐브의 데이터에 접근하는 것이다. 드릴스루는 OLAP시스템에 존재하는 상세 데이터에 접근하는 것이다[7].

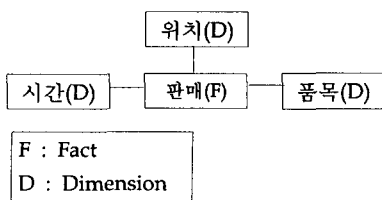
2.2 스키마 유형

데이터 웨어 하우스의 개념적 모델링을 표현하기 위해 스키마는 다음과 같이 세 가지 종류로 분류할 수 있다.

2.2.1 스타 스키마(star schema)

스타 스키마는 논리적인 모델을 구현하기 전에 관계형 데이터베이스의 다차원 데이터를 스타 스키마를 이용하여 논리적으로 모델화한다. 스타 스키마는 중앙에 하나의 사실 테이블로 이루어지고 이 사실 테이블은 유일한 정규화 된 테이블이다. 이 테이블 주위에는 여러 개의 차원 테이블이 있다. 이 차원 테이블은 사실 데이터에 대한 사용자 관점을 나타내어 서술화된 정보를 저장하며 비정규적인 데이터 구조를 나타낸다.

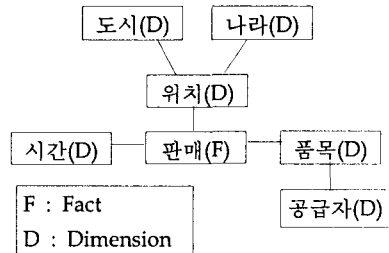
[그림 2]는 판매의 스타 스키마를 나타낸 것이다.



[그림 2] 판매의 스타스키마

2.2.2 snowflake schema

snowflake schema는 스타 스키마의 확장된 개념이라고 할 수 있다. 이 스키마는 스타 스키마와 동일한 구조를 가지지만 차원 테이블에서 정규화가 되어 있는 구조라 할 수 있다.

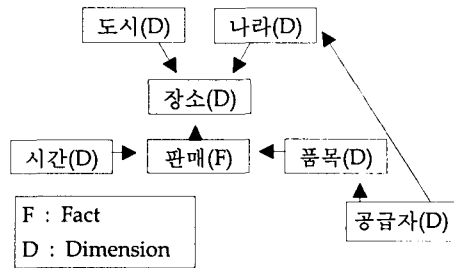


[그림 3] 판매의 snowflake 스키마

[그림 3]은 판매 사실 테이블에 포함되어 있는 기본키와 이를 참조하는 품목, 위치, 시간 등 외래키들의 차원 테이블을 나타낸 것이다. 이 테이블에 나타난 기본키를 이용하여 위치나 품목 차원 테이블들이 확대되어 이 기본키를 참조하는 외래키를 가진다. 이 스키마는 다차원 계층 구조를 가지고 있다.

2.2.3 3NF schema

3NF schema는 snowflake schema의 정보를 포함한다. [그림4]에서는 ER이나 UML모델처럼 의미적 데이터 모델부터 전개되는 것을 보여 주고 있다.



[그림 4] 판매의 3NF schema

2.4 객체 지향 관계

객체 지향의 모델링에 쓰이는 UML(unified modeling language)은 일반화(Generation), 연관성(Association), 흐름(flow), 전개(derivation) 4가지의 관계로 나눌 수 있다[6].

첫째, 일반화는 어떤 의미를 명확히 나타낸다.

둘째, 연관성은 객체(object)와 객체(object)사이에서 어떤 의미를 지정해 놓으면 이것들에 의해 관계를 연관 지어 나타낸다.

셋째, 흐름 관계는 모델 사이에 두 요소를 흐름으로 나타낸다.

넷째 전개 관계는 객체(object)와 객체 사이의 전개 관계를 나타낸다.

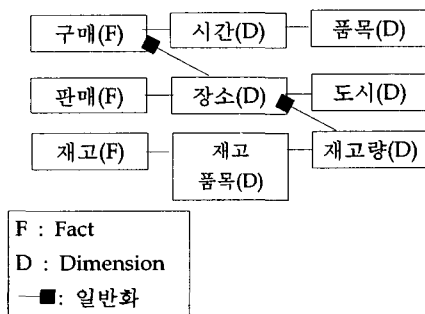
관계를 나타낸다. [표1]과 같이 이들 구조의 객체관계를 볼 수 있다[4].

III. 객체 지향 관계와 Snowflake schema의 의미적 관계

본 논문에서는 drill-across 연산자를 통해 객체 지향 의미에 스타 스키마보다 한층 더 확장된 스노우 플레이크(snowflake)스키마를 이용하여 의미적 관계를 설계한다. 이것은 스타 스키마의 사실-차원 구조의 스타스키마에서 나타나지 않았던 일반화 관계를 drill-across를 통하여 의미적 구조를 나타낸다.

3.1 Snowflake schema의 사실-차원-차원관계

Snowflake schema는 사실-차원-차원 테이블 구조로 되어 있다. drill-across의 연산자를 이용하여 한쪽의 Snowflake schema의 사실 부분과 Snowflake schema의 차원 부분 그리고 또 다른 Snowflake schema의 차원 부분을 이용하여 객체 지향 관계인 일반화 관계의 의미적 구조 나타낼 수 있는데 [그림5]와 같이 설계할 수 있다. 특히 재고량 차원 테이블에서 장소 차원 테이블을 일반화하여 이것에 의해 구매 사실 테이블을 일반화 관계로 나타낼 수 있다. 사용자가 어떤 물건을 구매 하고자 할 때 만약 소비자가 구매하고자 하는 아이템이 없을 경우 다른 장소에 있는 재고량을 검색하여 사용자가 원하는 것을 구매할 수 있도록 하여 일반화할 수 있다.



[그림5] 사실-차원-차원사이의 일반화 관계

IV. star schema와 Snowflake schema를 적용한 관계

4.1 star schema를 적용한 객체지향관계

스타 스키마의 구조에서는 차원(Dimension)-차원과의 관계, 사실(Fact) - 사실(Fact), 사실(Fact)-차원(Dimension)관계의 구조로 이루어져 객체지향

[표1] 객체 지향 관계

relationship	F-F	D-D	F-D
일반화	가능	가능	불가능
연관화	가능	가능	가능

4.2 snowflake schema를 적용한 객체지향관계

기존의 스타 스키마를 적용하여 객체지향관계를 나타내어 본 결과[표1] 차원-차원, 사실-사실 사이에서 drill-across를 이용하여 일반화, 연관화의 의미적 관계는 잘 나타내었지만 차원-사실 사이에서 연관성의 의미적 관계가 성립되었지만 일반화 관계는 성립되지 않았다.

그러나 Snowflake schema를 적용하여 일반 관계를 [그림5]와 같이 설계를 부여한 결과 사실-차원-차원의 일반화 관계를 통하여 사실-차원의 관계가 성립된 것을 알 수 있다.

V. 결 론

데이터 웨어 하우스는 주제 지향적이고, 통합적이고, 비유동적인 데이터로서 의사결정에 필요한 정보들을 추출하여 OLAP 분석에 이용된다. OLAP 연산자와 스타 스키마 사이의 다차원 모델링에 관한 연구는 현재 많이 진행되어 왔다.

본 논문에서는 drill-across 연산자를 통해 객체 지향 개념적 관계와 스타 스키마보다 한층 더 확장된 스노우 플레이크(snowflake) 스키마를 이용하여 의미적 관계를 설계하였다. 이 방법은 기존의 스타 스키마에서 사실-차원 구조에서 적용하여 나타내지 못했던 일반화 관계를 나타낼 수 있는 특징을 가지고 있음을 보였다.

향후 연구 과제로는 스노우 플레이크 스키마 (snowflake schema)의 의미적 관계 설계를 통해 구현으로 나타낼 수 있도록 해야 한다.

참고문헌

- [1] W.H.INmon.Building the Data Warehouse. John Wiley & Sons. 1996
- [2] OLAP Council. OLAP and OLAP Server Definitions. Available at the URL <http://www.olapcouncil.org/research/glossaryly.htm> 1997.
- [3] E. Franconi, F. Baader, U. Satter, and P. Vassiliadis. Fundamentals of Data Warehouse, chapter Multidimensional Data Model a-

- nd Aggregation. Springer, 2000. M. Jarke, M. Lenzerini, Y. Vassilios and P. Vassiliadis editors
- [4] A. Abello, J. Samos, and F. Saltor. On Relationships Offering New Drill-across Possibilities 2002
- [5] A. Abello, J. Samos, and F. Saltor. Implementing Operations to Navigate Semantic star schemas
- [6] OMG. Unifide Modeling Language Specification, September 2001. Version 1.4 Available at <http://www.org/cgi-bin/doc/formal/01-09-67>
- [7] 조재희, 박성진. □□OLAP 테크노로지□□ 시그마 컨설팅 그룹, 1999
- [8] J. Akoka, I. Comyn-Wattiau, and N. Prat. Dimension Hierarchies Design from UML Generlization and Aggregation. Springer, 2001