

데이터 웨어하우스의 개념적 설계를 위한 스타 스키마에서 ER 도형으로의 변환 기법

최은하¹⁾⁰ 김진호¹⁾ 옥수호²⁾

¹⁾강원대학교 컴퓨터과학과

²⁾고신대학교 컴퓨터과학부

eunha99@mail.kangwon.ac.kr, jhkim@kangwon.ac.kr, shok@kosin.ac.kr

Translation of Star Schema into Entity-Relationship Diagrams for Data Warehouse Conceptual Design

Eun-Ha Choi¹⁾⁰ Jin-Ho Kim¹⁾ Soo-Ho Ok²⁾

¹⁾Dept. of Computer Science, Kangwon National University

²⁾Dept. of Computer Science, Kosin University

요 약

데이터 웨어하우스는 기업의 의사 결정을 지원하기 위해 기업의 운영 데이터베이스로부터 추출한 데이터의 집합으로써 OLAP 분석에 이용된다. OLAP은 이들 데이터를 다양하게 분석할 수 있도록 다차원 데이터로 표현하고 이를 활용하여 복잡한 분석 질의 처리 및 다차원 데이터 분석에 활용한다. 이러한 OLAP의 다차원 데이터를 관계형 데이터베이스에서 표현하기 위해 스타 스키마가 널리 사용된다. 또한 다차원 데이터와 데이터 웨어하우스는 방대한 분량을 갖는 전체 기업의 데이터를 표현하고 있어 이를 설계하는 것이 매우 복잡하고 많은 노력이 소요된다. 따라서 이를 설계하기 위한 체계적인 설계 방법론이 필요하다. 데이터 웨어하우스의 원천 데이터가 되는 운영 데이터베이스는 현재 ER 도형을 이용하여 개념적인 방법으로 널리 설계되고 있다. 따라서 이 논문에서는 ER 도형으로 설계된 운영 데이터베이스로부터 데이터 웨어하우스를 설계하는 개념적인 방법론을 제시한다. 이에 따라 OLAP 분석을 위해 사용할 수 있는 다양한 유형의 스타 스키마에 대해 ER 도형으로 표현/변환하는 방법을 제시한다. 이를 통해 자신이 원하는 다차원 데이터를 얻기 위해 유지해야 할 데이터 웨어하우스를 ER 도형을 이용하여 개념적으로 편리하게 설계하는 방법/지침을 제공하며, 나아가 해당 유형의 스타 스키마가 갖는 의미를 개념적으로 쉽게 전달할 수 있도록 하였다.

1. 서 론

실시간 분석 처리(On-Line Analytical Processing : OLAP)는 기업의 의사 결정 지원을 위하여 다차원적인 데이터 분석 및 복잡 질의 처리에 효율적이며, 데이터 웨어하우스는 OLAP 분석에 사용할 데이터를 저장하고 있다[1][2]. OLAP에서는 데이터를 분석에 적합하도록 다차원 데이터 형태로 모델화하게 되는데 이를 큐브(cube)라 하며, 관계형 데이터 베이스에서 다차원 데이터 모델인 큐브를 표현하기 위해 등장한 데이터 구조가 스타 스키마이다. 스타 스키마는 하나의 사실 테이블과 여러 개의 차원 테이블로 구성되어 있고, 테이블간의 조인을 최소화하여 질의에 대한 응답시간을 줄일 수 있기 때문에 복잡한 질의에 적합한 데이터 구조이다[3]. 이를 데이터 웨어하우스는 기업의 운영 데이터베이스로부터 추출된다. 이러한 운영 데이터베이스는 현재 주로 ER 도형과 같은 개념적 모델링 기법을 사용하여 설계되어 있다.

데이터 웨어하우스 설계자는 여러 사이트로부터 정보를 수집하고 사용자의 요구사항을 분석하여 OLAP에 적합한 데이터 웨어하우스를 설계 한다[4][5][6]. 이 과정에서 데이터 웨어하우스 설계자들에게 주요 이슈로 대두되고 있는 것은 ER 도형으로 설계되어 있는 데이터 소스들을 어떻게 효율적으로 분석에 적합한 다차원 데이터 모델로 변환하며, 이러한 다차원 데이터를 위한 데이터 웨어하우스를 구축하는가에 관한 것이다. 이를 위한 데이터 웨어하우스의 개념적 및 논리적 모델링 방법이 계속적으로 연구되어 왔다.

이 논문에서는 개념적인 방법으로 데이터 웨어하우스를 편리하게 설계하는 방법을 제시하기 위해 다양한 유형의 스타 스키마를 ER 도형으로 표현하는 방법을 제시한다. 이를 통해 자신이 목표로 하는 다차원 데이터(즉 스타 스키마)를 얻기 위해 유지해야 할 데이터 웨어하우스내의 테이블들을 ER 도형을 이용해 개념적으로 설계할 수 있도록 하였다. 이 논문에서는 여러 차원 테이블과 하나의 사실 테이블로 이루어진 단순한 스타 스키마, 여러개의 사실 테이블을 갖는 다중 사실 테이블을 갖는 스타 스키마, 사실 테이블내에 별도의 식별자를 갖는 멀티 스타 스키마, 차원과 사실간에 1:N이 아닌 M:N 관계를 갖는 스타 스키마 등등 실제계에 존재할 수 있는 다양한 형태의 스타 스키마에 대해 ER 도형으로 개념적 모델링하는 방법을 제시하였다. 이렇게 함으로써 원하는 다차원 데이터에 대한 데이터 웨어하우스를 개념적으로 설계하여 설계에 소요되는 시간과 비용을 절약할 수 있다. 뿐만 아니라 다양한 유형의 스타 스키마를 쉽게 이해할 수 있는 ER 도형으로 표현하여 스타 스키마의 의미를 명확하고 쉽게 이해할 수 있다는 장점이 있다.

이 논문의 구성은 다음과 같다. 2장에서는 관련 연구와 이 논문의 연구 동기에 대해 서술하고, 3장에서는 OLAP을 위한 데이터 모델링 요소와 스타 스키마의 유형에 대해 논한다. 4장에서는 각 유형의 스타 스키마를 ER 도형으로 변환하는 방법과 과정을 설명한 후, 5장에서 결론을 맺는다.

2. 관련 연구

2.1 데이터 웨어하우스와 OLAP

데이터 웨어하우스는 의사 결정에 필요한 정보를 사용자에게 효율적으로 제공하기 위해 구축된 거대한 데이터 저장소이다. 데이터 웨어하우스에 근간을 두

1) 이 논문은 첨단정보기술연구센터(AITrc)를 통하여 한국과학재단의 지원을 받았음.

고 있는 OLAP은 기업의 의사 결정 지원을 위하여 다차원적인 데이터 분석 및 복잡 질의 처리를 제공하는데 효율적이다. 데이터 웨어하우스 설계자는 여러 사이트로부터 통합된 데이터 소스와 사용자의 요구사항을 분석하여 데이터 웨어하우스를 설계하게 된다. 이 과정에서 ER 도형으로 설계되어 있는 데이터 소스들을 분석에 적합한 다차원 데이터 모델 즉 스타 스키마로 변환하게 되는데, 좀 더 효율적인 변환 방법을 찾는 것이 그들의 주요 관심 대상이 되고 있다. 이에 따라 데이터 웨어하우스의 스타 스키마 설계를 위한 개념적 및 논리적 모델링 기법에 대한 연구가 계속해서 진행되고 있다.

2.2 데이터 웨어하우스 설계

OLAP을 위한 데이터 웨어하우스 설계에 대한 연구는 지금까지 여러 관점에서 연구되어 왔다. M. Golfarelli와 S. Rizzi[4]는 데이터 웨어하우스 설계를 위한 방법론을 다음과 같은 단계로 설명하고 있다. 먼저, 운영 데이터베이스로부터 수집된 정보를 분석하고 사용자의 요구 사항을 명세하는 단계를 거친 후 개념적 설계를 통해 사실과 차원 구조를 생성하게 된다. 다음으로 개념적 설계 결과에 대한 작업량과 유효성을 판단한 후 테이블 구조로 실제화하기 위한 논리적 설계 단계를 거치고, 마지막으로 물리적 설계를 수행한다.

데이터 웨어하우스 설계 단계 중 개념적 설계에 관한 연구가 [5]에서 이루어졌다. 이 논문에서는 데이터 웨어하우스 설계에 관한 새로운 개념적 데이터 모델인 starER을 소개하였다. 이것은 ER 모델을 바탕으로 스타 구조를 표현할 수 있는 새로운 모델링 구조로 확장한 것이다. 이러한 새로운 모델링 구조는 그 모델이 복잡하며, 다양한 유형의 스타 스키마를 starER을 통해 어떻게 표현하는지에 대한 연구는 부족하다.

L. Cabibbo 와 R. Tortone[3]은 운영 데이터베이스의 ER 도형으로부터 OLAP에서 필요로 하는 다차원 데이터 모델을 얻기 위한 방법을 제시하고 있다. 이 방법에서는 ER 도형을 다차원 데이터로 변환하기 용이하도록 ER 도형을 변형한 후 이로부터 다차원 그래프를 추출하고, 이 다차원 그래프로부터 다차원 데이터 모델로 변환하는 방법을 제시하였다. 이 방법은 ER 스키마를 재구성하는 과정이 복잡하고 설계자의 직관에 많이 의존한다는 문제점이 있다.

ER 도형에 의해 데이터 웨어하우스를 개념적으로 설계하는 방법을 제시하기 위해 스타 스키마를 ER 도형으로 변환하는 시도는 M. Krippeerdorf 와 Il-Yeol Song[6]에 의해 시도된 바 있다. 이 논문에서는 스타 스키마의 유형을 표준 유형(standard case)과 비표준 유형(non-standard case)으로 구분하여 각각 ER 도형으로 변환하기 위한 규칙을 유도하고, 그 규칙에 따라 스타 스키마를 ER 도형으로 변환하는 과정을 보여주고 있다. 이 논문에서는 이들의 연구를 확장하여 좀 더 복잡한 유형의 스타 스키마를 ER로 변환하는 과정을 포함하였으며, [6]에서 표현하지 않았던 차원의 계층 구조를 추가로 표현하도록 하였다.

3. 스타 스키마와 그 유형들

본 장에서는 OLAP을 위한 다차원 데이터 모델링에서 사용되는 스타 스키마 및 주요 이슈 사항에 대해 논한다.

3.1 스타 스키마와 ER 도형

스타 스키마는 데이터 웨어하우스와 OLAP의 출현으로 인해 요구되는 새로운 데이터 모델로서, 분석을 위한 다차원 데이터를 표현하기 위해 사실 테이블과 차

원 테이블로 구성된다. 테이블간의 조인을 최소화 함으로써 질의에 대한 응답속도를 향상시킬 수 있다는 장점을 가지고 있기 때문에 OLAP의 복잡 질의에 적합하며, 데이터 웨어하우스 테이블 구조의 논리적 설계에 사용되어 진다. ER 도형은 사용자의 요구에 대한 분석사항을 엔티티(entity), 애트리뷰트(attribute), 관계(release)로 나타낸 개체 관계 도형을 말한다. 현재 대부분의 DBMS 개념적 모델링 기법으로 ER 도형이 사용되고 있으며 데이터 웨어하우스를 위한 개념적 모델링에서도 ER 도형이 이용된다.

3.2 스타 스키마 유형

스타 스키마는 나타내고자 하는 다차원 데이터의 여러 가지 유형에 따라 다양한 종류의 형태로 구분될 수 있다. 먼저 하나의 사실 테이블과 다수의 차원 테이블로 구성된 간단한 스타 스키마가 있고, 정규화된 차원을 포함하는 스타 스키마가 있다. 또한 공통된 차원에 대해 집계의 정도가 다르거나 요구되는 측정값이 다른 경우 두 개 이상의 사실 테이블을 가진 스타 스키마가 있고, 차원 테이블의 애트리뷰트들간에 재귀 관계가 있는 스타 스키마가 있다. 그리고, 사실 테이블의 튜플에 유일성을 부여하기 위해 키를 추가한 경우이거나, 차원의 특정 애트리뷰트만이 필요할 때 이를 키로 추가한 경우의 멀티 스타 스키마가 있다.

3.3 차원의 계층구조

스타 스키마의 차원에서, 측정값을 기술하는 상세 정도를 수준(level)이라고 하며, 차원을 구성하는 구성원들은 그에 해당하는 수준에 속함으로써 차원 내에서 수직적인 상하 관계를 형성한다. 이러한 하나의 전체적인 차원의 구조를 계층구조(Hierarchy)라고 하며, 측정값을 어떤 하나의 차원을 기준으로 분석하고자 할 때 대개 계층구조상의 가장 상위 수준에 해당하는 데이터부터 분석하고 점점 세부 수준으로 들어가며 데이터를 분석하게 된다. 이와 같은 계층구조는 차원 테이블에서 애트리뷰트로 나타나게 되며, 본 논문에서는 이러한 차원의 계층구조에 대해 ER 도형으로 변환하는 방법을 제시하고자 한다.

3.4 사실과 차원 사이의 다 대 다 관계

일반적으로 스타 스키마에서 사실 테이블과 차원 테이블은 다 대 일(N:1) 관계로 형성된다. 그러나 실생활에서는 사실과 차원 사이의 다 대 다(M:N) 관계가 종종 나타나게 된다. 이러한 M:N 관계를 N:1 관계로 스타 스키마에서 표현하기가 어렵다. 이러한 사실과 차원의 M:N 관계를 비정규화(denormalization) 기법을 이용하여 N:1 관계의 스타 스키마로 표현하는 방법들이 [An Analysis of Many-to-Many Relationship Between Fact and Dimension Tables in Dimensional Modeling]에서 제시되었다. 이 논문에서는 이러한 사실과 차원의 M:N 관계를 비정규화하는 다양한 방법들을 ER 도형으로 표현하는 방법을 제시한다.

4. 스타 스키마 변환

이 장에서는 앞에서 설명한 여러 유형의 스타 스키마를 ER 도형으로 변환하는 방법을 제시한다. 스타 스키마의 유형, 차원의 계층구조, 그리고 사실과 차원 사이의 M:N 관계 해결에 따른 ER 도형으로의 변환 방법을 설명한다.

4.1 스타 스키마 유형에 따른 ER 도형으로 변환 방법

본 절에서는 스타 스키마 유형에 따라 ER 도형으로 변환하는 방법을 제시한다. 스타 스키마의 사실 테이블과 차원 테이블은 각각 엔티티로 표현되고 N:1 관계로 연결된다. 또한 정규화된 차원 테이블도 관계를 통해 연결된다. 그림 1은 간단한 스타 스키마의 ER 도형으로의 변환을 보여주고, 그림 2는 정규화된 차원을 포함한 스타 스키마의 변환을 보여주고 있다.

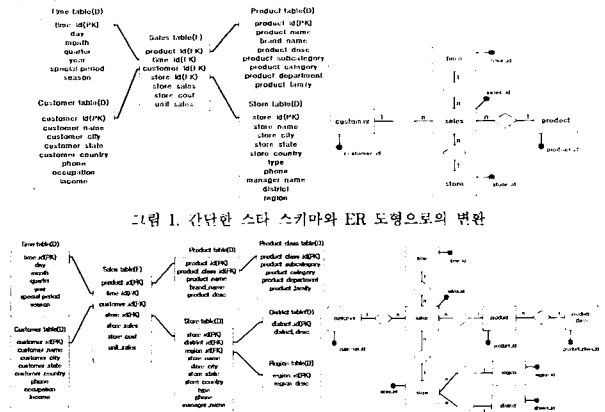


그림 1. 간단한 스타 스키마와 ER 도형으로의 변환

그림 2는 정규화된 차원을 포함한 스타 스키마와 ER 도형으로 변환한 것이다. 그림 3은 두 개의 사실 테이블, 판매(Sales) 테이블과 반품>Returns) 테이블을 가진 스타 스키마와 그에 따른 ER 도형으로의 변환을 보여주고 있고, 그림 4는 재귀관계를 가지는 차원에 대한 ER 도형으로의 변환을 보여주고 있다.

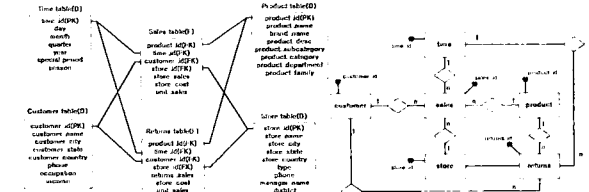


그림 3. 두 개의 사실 테이블을 가진 스타 스키마와 ER 도형으로의 변환

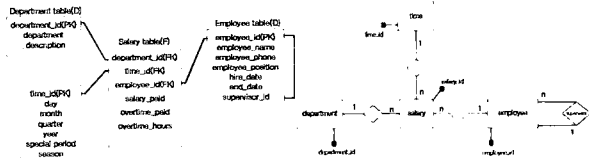
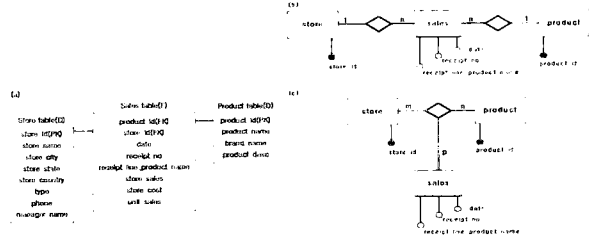


그림 4. 재귀 관계를 가지는 차원에 대한 ER 도형으로의 변환

멀티 스타 스키마는 사실 테이블의 기본키와 외래키가 동일하지 않은 경우의 스타 스키마를 말한다. 그림 5는 사실 테이블(Sales)이 두 개의 차원 테이블(Product, Store)의 기본키를 외래키로 가지고 있지만, 이 두 개의 외래키 만으로는 사실 테이블의 튜플에 유일성을 부여할 수 없기 때문에, 즉 식별될 수 없기 때문에 부가적으로 키(date, receipt.no, receipt.line.product.name)를 추가한 경우의 스타 스키마를 보여주고 있다. 사실 테이블은 약(weak) 엔티티로 표현되며, 여기서는 (b), (c) 두 가지 ER 도형 표현을 모두 허용한다.



(a) 멀티 스타 스키마 (b)(c) ER 도형 변환
그림 5. 멀티 스타 스키마와 ER 도형으로의 변환

4.2 차원의 계층구조에 대한 ER 도형으로의 변환 방법

본 절에서는 3.3절에서 논의했던 스타 스키마에서의 차원의 계층구조를 ER 도형의 모델링 구조를 그대로 사용하여 표현하는 방법에 대해 기술한다. 먼저 스타 스키마에서 차원의 계층구조는 ER 도형에서 복합 애트리뷰트(composite attribute)로 표현하고, 계층 수준을 알 수 있도록 각 애트리뷰트 이름에 (숫자)를 기록한다. 여기서 (숫자)가 클수록 높은 계층 수준을 나타낸다. 계층구조의 유형을 간단한 형태(customer), 이중 계층구조(time), 정규화된 차원에서의 계층구조(store, product)의 3가지 유형으로 나누어 설명한다. 그림 6은 이러한 계층구조에 대한 ER 도형으로의 변환을 보여준다.

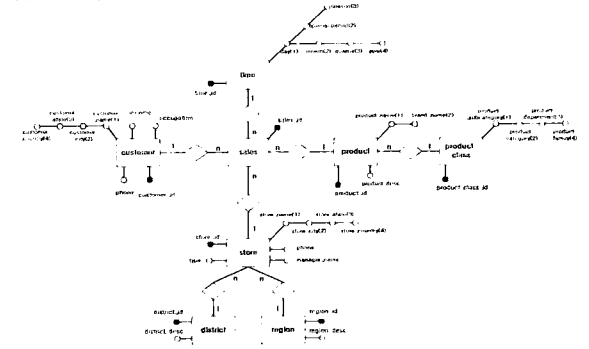


그림 6. 스타 스키마의 계층구조에 대한 ER 도형으로의 변환

4.3 사실과 차원 사이의 M:N 관계를 ER 도형으로 변환 방법

본 절에서는 사실과 차원 사이의 다 대 다 관계를 다 대 일 관계로 변형하여 스타 스키마 형태를 유지하는 방법과 그에 따른 ER 도형으로의 변환 방법에 대해 논의한다. 사실과 차원 사이의 다 대 다 관계를 다 대 일 관계로 비정규화하는 방법들이 [An Analysis of Many-to-Many Relationship Between Fact and

Dimension Tables in Dimensional Modeling]에서 연구되었다. 이 논문에서 이 연구를 토대로 M:N 관계의 스키마를 ER 도형으로 변환하는 방법을 제시한다.

4.3.1 사실과 차원 사이의 M:N 관계의 예

아래 릴레이션은 수업료를 분석하기 위해 사용되는 릴레이션들이다.

- 학기(학기_id, year, 학기)
- 학생(학생_id, 학생_이름, 학생_학년, 학생_전과번호)
- 과목(과목_id, 과목_이름, 과목_수강료)
- 수업료(학기_id, 학생_id, 과목_id, 수업료)

수업료는 스타 스키마에서 사실이 되고 나머지는 차원이 된다. 수업료(사실)와 과목(차원) 사이의 관계를 살펴보면, 한 학생의 수업료에는 여러 과목이 포함되어 있고, 하나의 과목은 여러 학생의 수업료에 포함된다. 또한 한 학생에 대한 하나의 학기의 수업료에 여러 과목이 포함되며, 몇몇 과목에 대하여 재수강을 할 경우 하나의 과목이 여러 학기 수업료에 포함된다. 즉 수업료(사실)와 과목(차원) 사이 다-대-다 관계가 발생하게 된다. 이를 스타 스키마 형태로 표현한 것이 그림 7에 보여지고 있다.

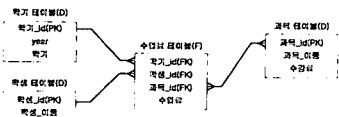


그림 7. 사실과 차원 사이의 다-대-다 관계 있는 스타 스키마

4.3.2 사실과 차원 사이의 M:N 관계의 해결

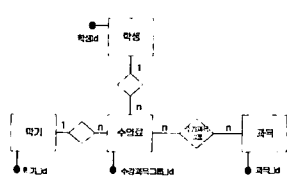
[An Analysis of Many-to-Many Relationship Between Fact and Dimension Tables in Dimensional Modeling]에서는 M:N 관계를 비정규화하는 6가지 방법을 제시하였지만, 이 중에서 스타 스키마의 개념내에서 표현될 수 있는 아래의 4가지 경우에 대해 ER 도형으로 변환하는 방법을 설명한다.

① Bridge 테이블을 이용하는 방법

사실과 차원 사이에 테이블을 하나 추가하는 방법이다. 이 테이블을 Bridge 테이블이라 하는데 본 예에서는 수강과목그룹 테이블이 이에 해당한다. 그림 8의 (a)는 Bridge 테이블을 나타낸 것이고, (b)는 Bridge 테이블이 있는 스타 스키마를 ER 도형으로 변환한 것인데, 여기서 Bridge 테이블은 관계(Relationship)로 표현한 것을 볼 수 있다.

수강과목그룹 테이블

수강과목 그룹_id	과목_id
group_1	A
group_1	B
group_1	C
group_1	D
group_2	A
group_2	B
...	...



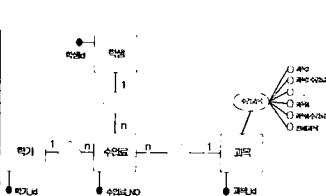
(a) Bridge 테이블 (b) ER 도형으로 변환

② 차원 테이블을 변형하는 방법

기존의 과목 테이블(차원)을 그림 9의 (a)와 같이 변형하는 방법이다. 한 학생이 한 학기에 수강할 수 있는 최대 과목수가 4과목이라고 가정할 경우, 중복이 없는 모든 가능한 과목의 조합을 전부 테이블에 나타내는 것이다. 이렇게 되면 여러 수업료가 하나의 과목_id와 관계되기 때문에 사실과 차원 사이에 다-대-일 관계를 유지할 수 있다. 그림 9의 (a)는 변형된 과목 테이블이고 (b)는 스타 스키마를 ER 도형으로 변환한 것인데 수강과목이라는 복합 애트리뷰트로 나타낸 것을 볼 수 있다.

과목 테이블

과목_id	과목_이름	과목_수강료	과목_수강인원	연계과목
001	영어	1000	사학	국어, 영어, 수학, 과학
002	국어	1000	민립	국어, 영어
003	수학	2000	국사	수학, 과학, 사회, 국사
004	과학	2500	민립	국어, 사회, 국사
...



(a) 과목 테이블 (b) ER 도형 변환

③ 사실 테이블에 키를 추가하는 방법

수업료 테이블(사실)에 수강과목그룹_id를 추가하는 방법이다. 이렇게 하면 수강과목그룹_id와 하나의 과목_id가 복합되어 하나의 기본키가 된다. 따라서 과목 테이블(차원)과 다-대-일 관계가 이루어 질 수 있다. 그림 10의 (a)는 수강과목그룹_id가 추가된 스타 스키마를 나타내고 있고, (b)는 ER 도형으로 변환한 것인데 추가된 수강과목그룹_id와 과목_id가 결합하여 하나의 기본키를 이루므로, 이것을 외부/혼합 식별자(External/mixed identifier)로 나타낸 것을 볼 수 있다.

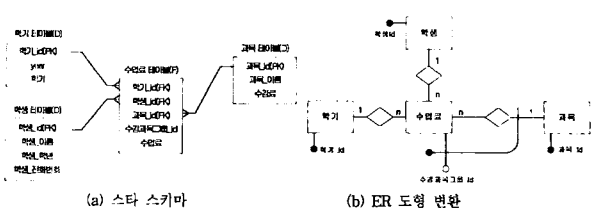


그림 10. 사실 테이블에 키를 추가하는 방법

④ 사실 테이블을 2개로 나누는 방법

사실 테이블을 분석 관점에 따라 2개의 테이블로 분리하는 방법이다. 여기서 수업료에만 관심이 있는 경우의 수업료 테이블과 수업료에는 관심 없고 어떤 과목을 수강하는지에 관심이 있는 경우의 수강내역 테이블로 나눌 수 있다. 또한 두 개의 사실 테이블은 수업료_id를 통해 서로 연결되어 있기 때문에 서로를 참조할 수 있다. 이렇게 두 개의 사실 테이블로 나누면 기존에 문제가 되었던 사실과 차원 사이의 다-대-다 관계는 없어지게 된다. 그림 11의 (a)는 두 개의 사실 테이블을 나타내는 스타 스키마이고 (b)는 ER 도형으로 변환한 것이다.

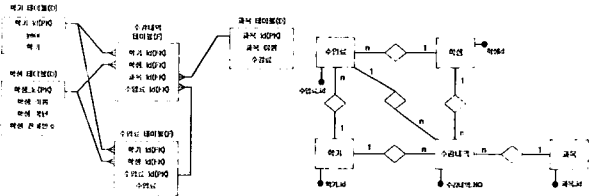


그림 11. 사실 테이블을 2개로 나누는 방법

5. 결 론

데이터 웨어하우스는 기업의 의사 결정을 위해 OLAP에서 사용할 정보를 유지하고 있는 데이터의 집합체이다. 이 데이터 웨어하우스를 설계하기 위해서는 기업의 활동으로 발생하는 운영 데이터베이스로부터 다차원 분석에 사용할 수 있도록 체계적으로 쌓을 데이터를 추출하고 구축해야 한다. 데이터 웨어하우스는 대형 조직의 방대한 데이터로부터 구축해야 하므로 정보 수집 분석 및 사용자 요구사항 명세 과정을 거쳐 개념적, 논리적, 물리적 설계 단계로 수행된다. 이 논문에서는 ER 도형을 이용하여 데이터 웨어하우스를 개념적으로 설계하는 방법에 대해 연구하였다. OLAP에서는 데이터 웨어하우스에 축적된 데이터를 스타 스키마 등의 다차원 모델로 표현한다. 이 논문에서는 OLAP에서 사용할 다양한 유형의 스타 스키마에 대해 ER 도형으로 표현하는 방법을 제시하여 원하는 OLAP 분석을 지원하기 위한 데이터 웨어하우스를 개념적으로 설계하는 방법을 제안하였다.

이 방법에서는 기존의 ER 모델을 그대로 사용하여 편리하고 빠르게 데이터 웨어하우스를 설계할 수 있도록 하였다. 이를 위해 스타 스키마로부터 ER 도형을 변환하는 규칙을 제시하였으며, 스타 스키마의 유형을 5가지로 구분하여 각각의 유형에 대해 ER 도형으로 표현하는 방법을 기술하였다. 또한 각 차원의 분류 계층 구조를 표현하는 방법뿐만 아니라 차원과 사실간에 M:N의 관계를 스타 스키마로 형태로 표현하는 방법들을 ER 도형으로 표현하는 방법을 제시하였다. 이렇게 함으로써 OLAP 분석에 사용할 데이터에 대한 개념적인 설계 방법의 기초로 활용하며, 스타 스키마에 대한 의미를 편리한 ER 도형으로 명확하게 전달할 수 있다. 뿐만 아니라 ER 도형으로 설계된 소스 데이터베이스나 데이터 웨어하우스로부터 스타 스키마를 자동적으로 추출할 수 있는 알고리즘을 개발하는데 기초로 활용할 수 있다. 이러한 ER로부터 스타 스키마를 자동적으로 추출하는 알고리즘은 향후에 연구될 예정이다.

6. 참 고 문 헌

- [1] S. Chaudhuri and U. Dayal, "An Overview of Data Warehousing and OLAP Technology", In Proc. of ACM SIGMOD Conf., pp. 65-74, 1997.
- [2] 조세환, 박성진, "OLAP 테크놀로지", 시그마 컨설팅 그룹, 1999.
- [3] L. Cabibbo and R. Torlone, "A Logical Approach to Multidimensional Database", In Proc. of EDBT, pp. 183-197, 1998.
- [4] M. Golfarelli and S. Rizzi, "A Methodological Framework for Data Warehouse Design", In Proc. of DOLAP, pp. 3-9, 1998.
- [5] N. Tryfona, F. Busborg and J. Christiansen, "starER: A Conceptual Model for Data Warehouse Design", In Proc. of DOLAP, pp. 3-8, 1999.
- [6] M. Kripendorf and Il-Yeol Song, "The Translation of Star Schema into Entity-Relationship Diagrams", In Proc. of DEXA, pp. 390-395, 1997.
- [7] Il-Yeol Song, W. Rowen, C. Medsker and E. Ewen, "An Analysis of Many-to-Many Relationships Between Fact and Dimension Tables in Dimensional Modeling", In Proc. of DMDW, pp. 6.1-6.13, 2001.