

사용자 요구사항 정의를 위한 OLAP View의 제안 및 활용

최 승 교[†] · 박 종 모^{**}

요 약

OLAP 시스템은 사용자가 다양한 각도에서 정보에 접근하여 대화식으로 정보를 분석할 수 있는 의사결정 지원 시스템이다. 그러나 기존 OLAP 시스템의 구축방법은 사용자의 잦은 요구사항의 변경에 따라 시간과 비용이 증가하는 비효율성을 가진다. 본 연구에서는 OLAP 시스템의 기존 구축과정이 갖는 문제점을 분석하여, 사용자의 요구사항에 효율적으로 대응할 수 있도록 스타 스키마에 기반을 둔 프로토타이핑 OLAP View를 제안한다. 제안 기법은 S 쇼핑몰의 구현 사례를 통해 사용자의 요구사항을 정확히 파악하여 OLAP의 구축 전략 수립이 가능하고, 데이터에 대한 이해와 정확한 파악이 가능하게 됨을 보인다. OLAP View를 사용함으로써 OLAP 시스템 구축에 필요한 시간과 비용을 줄일 수 있다.

Proposal And Application of An OLAP View For the Definition of the Users Requirements

Sung-Kyo Choi[†] · Jong-Mo Park^{**}

ABSTRACT

An OLAP system is the decision support tool with which a user can analyze the information interactively in the various aspects. However, the traditional existing construction of an OLAP system has the inefficiency problem of increasing the time and cost caused by the frequent changes of the users requirements. In this paper, we analyse existing construction procedure of OLAP systems and propose a prototyping OLAP View based on the star schema which can help and guide the designer in supporting efficiently of users requirements. Through an implementation of S shopping mall using the OLAP View, we show that our proposal is able to support OLAP construction strategy and provide accurate understanding of data resources. With an OLAP View, we can reduce the required time and cost of implementing OLAP system.

키워드 : OLAP(OnLine-Analysis Process), 데이터웨어하우스(Datawarehouse), OLAP View, 스타 스키마(Star Schema)

1. 서 론

지식 정보화 사회에서는 축적된 방대한 정보 중에서 필요한 정보만을 획득하여 이를 잘 활용하는 기업은 경쟁우위를 확보할 수 있으나, 그렇지 못한 기업은 경쟁우위를 상실하여 도태될 수 있다. 따라서, 기업은 효율적이고 효과적인 정보만을 얻도록 도와주는 다양한 의사결정 지원 시스템을 필요로 한다. OLTP(OnLine Transaction Process) 시스템은 업무 프로세스를 중심으로 트랜잭션(입력/조회/삭제/수정 등)을 처리하기 위해 사용되는 시스템이다. 그러나 OLTP 시스템은 저장되어 있는 데이터에 대해 평면적이고 정규화된 정보를 제공하기 때문에 기업의 의사결정 환경에는 유용하지 못하다. 즉, 질의의 수나 복잡성에 비례하여 운영시스템에 부하를 부여하고, 다양한 관점에서 비정규적으로 원하는 정보를 습득하기가 어렵다는 문제점을 가진다.

이러한 OLTP 시스템의 문제점을 극복하기 위해 OLAP(OnLine-Analysis Process) 시스템이 등장하였다. OLAP 시스템은 다양한 관점의 정보를 획득하기 위해 OLTP에서의 업무처리 중심이 아닌 주제영역 중심(예를 들면 고객, 제품, 판매 등)으로 분석 수행을 가능하게 해 준다. 즉, OLAP 시스템은 사용자가 다양한 각도에서 정보에 접근하여 대화식으로 정보를 분석할 수 있는 의사결정 지원 시스템이다.

OLTP 환경에서 데이터베이스는 데이터의 중복으로부터 생기는 삽입, 삭제, 수정의 문제점을 없애기 위하여 정규화라는 방법을 사용한다. 그러나 OLAP 환경에서는 데이터의 수정이나 삭제는 거의 일어나지 않고, 분석 데이터를 수집하기 위해 대용량의 데이터가 한번에 삽입되는 경우가 많으므로 데이터의 검색이 중요한 작업이다. 따라서 정규화보다는 중복이 일어나더라도 복잡한 검색 질의를 효율적으로 처리할 수 있도록 데이터를 모델링하고 구축하는 것이 중요하다. 즉, 관계형 데이터베이스의 연산 중 가장 처리비용이 비싼 조인 연산과 집계 연산을 효율적으로 처리할 수 있도록 데이터베이스를 모델링하고 구축해야 한다. 이러한 OLAP환경

* 본 논문은 삼척대학교 2002학년도 장기연수 연구비 지원에 의해 연구되었음.

[†] 정 회 원 : 삼척대학교 컴퓨터공학과 교수

^{**} 준 회 원 : 단국대학교 대학원 전산통계학과

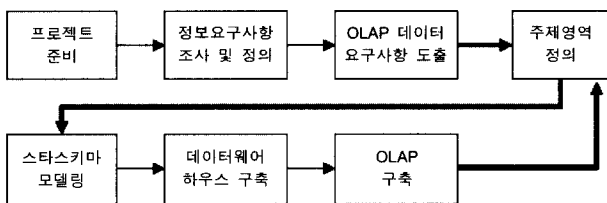
논문접수 : 2004년 3월 29일, 심사완료 : 2004년 7월 1일

에서 이용할 수 있도록 데이터웨어하우스가 제시되었다.

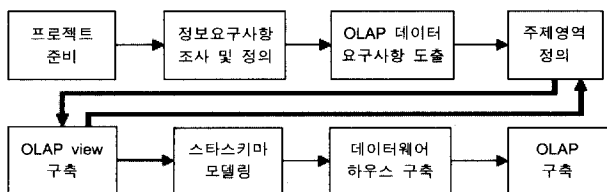
데이터웨어하우스는 OLTP(OnLine Transaction Process) 시스템으로부터 생성된 데이터에서 다양한 분석 정보를 추출하여 OLAP 영역에 사용하도록 하는 데이터의 저장고이다[2,3]. 즉, 데이터웨어하우스는 사용자들의 다양한 분석을 편리하게 제공하기 위하여 분석의 기준이 되는 차원들과 분석할 내용에 해당하는 측정값들로 구성된 다차원 형태로 데이터를 표현한다. 따라서 데이터웨어하우스를 이용함으로써 OLAP 질의에 빠른 응답을 처리할 수 있게 된다.

데이터웨어하우스의 설계를 위해 스타 스키마를 이용하는 데, 스타 스키마는 사실 테이블을 중심으로 차원 테이블들이 마치 별 모양과 같은 형태를 취한다[4,5]. 스타 스키마는 다차원 데이터를 관계 데이터베이스의 테이블을 이용하여 저장하면서, 비교적 단순한 형태로 표현하기 때문에 사용자들이 쉽게 이용할 수 있다는 장점이 있다.

전통적인 OLAP 시스템을 구축하기 위해서는 (그림 1)의 과정을 거치므로, OLAP 구축 후에도 자주 변하는 사용자의 요구분석에 효율적이지 못하였다. OLAP 시스템의 구축 과정에서는 스타스키마를 위한 모델링과 데이터웨어하우스 구축에 많은 시간이 소요되고, 사용자의 잦은 요구사항의 변경에 따라 전체 모델링과 구현과정을 반복 수행해야 한다. (그림 1)의 기존 구축과정에서 굵은 화살표의 과정들이 반복 수행되는 비효율성과 이로 인한 과도한 시간과 자원의 소요라는 오버헤드 문제가 발생한다. S 쇼핑몰과 같이 방대한 OLAP 구축이 요구되는 곳에서는 잘못된 분석으로 인해 OLAP 구축이 끝나고 처음부터 다시 시스템을 구축해야 하는 경우가 발생한다.



(그림 1) 기존의 OLAP 구축 과정



(그림 2) 제안된 OLAP 구축 과정

이러한 문제점을 최소화하기 위해 본 논문에서는 (그림 2)과 같은 프로토타이핑 OLAP View가 추가된 구축과정을 제안한다. 즉, 구축될 OLAP의 모습을 미리 확인함으로써, 데이터웨어하우스 구축과 OLAP 구축 후 반복 수행되는 과

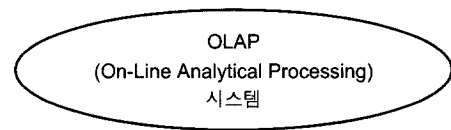
정을 줄이고자 한다. 왜냐하면 OLAP View의 구축일정이 반복 수행되는 일정보다 상대적으로 작기 때문이다.

본 논문에서는 S쇼핑몰의 실제 구축과정을 실례로 보이고, 또한 구축과정에서 발생한 문제점을 분석하여, 제안된 프로토타이핑 OLAP View 기법의 효율성을 제시한다.

본 논문의 구성을 다음과 같다. 2장에서는 OLAP의 개념과 데이터웨어하우스의 개념을 소개하고, 관련 연구를 제시한다. 3장에서는 사용자 요구사항 정의의 문제점을 분석하여, 사용자 요구사항 정의과정 중에 발생하는 문제점을 제시하며, 이에 대한 해결책으로 OLAP View의 활용을 제안한다. 4장에서는 제안된 OLAP View를 활용한 시스템 구축의 예를 보인다. 5장에서는 제안 기법의 특성을 분석한다. 즉, 제안 기법을 통해 프로젝트 전략의 정립이 가능하고, 데이터의 정확한 파악이 가능함을 보인다. 6장의 결론으로 본 논문을 마무리한다.

2. 관련 연구

OLAP은 OLTP에 상대되는 개념으로 공유되는 다차원 정보에 대한 신속한 분석을 위해 제시되었다[1]. 즉, OLAP은 최종 사용자가 다차원 정보에 직접 접근하여 대화식으로 정보를 분석하고 의사결정에 활용하는 과정이다. 이러한 OLAP은 데이터웨어하우스에 체계적으로 쌓여 있는 데이터 속에 담겨 있는 필요한 정보만을 효율적으로 분석하고자 할 때 효율적임이 제시되었다[10]. 데이터웨어하우스에 기반을 둔 OLAP은 기업의 의사결정 지원을 위하여 다차원적인 데이터 분석 및 복잡한 질의의 처리를 제공하는데 효율적이다. 이러한 OLAP 시스템은 [1]에서 제시된 FASMI(Fast Analysis of Shared Multidimensional Information)특성을 가지며 (그림 3)과 같이 요약할 수 있다.



| 특징 | 설명 |
|------------------|--------------------|
| Fast | 사용자 질의에 신속히 반응 |
| Analysis | 제약성이 없고 사용자 정의가 가능 |
| Shared | 동일 데이터에 다중 사용자가 접근 |
| Multidimensional | 다차원 분석 가능 |
| Information | 얼은 정보 및 유추된 정보 |

(그림 3) OLAP 시스템의 특성

또한, OLTP 시스템에 대한 OLAP 시스템의 특성은 <표 1>과 같이 비교할 수 있다.

[6]은 데이터웨어하우스의 설계 방법론을 다음과 같은 단계로 제시하였다. 먼저, 운영 데이터베이스로부터 수집된 정

보를 분석하고 사용자의 요구 사항을 명세하는 단계를 거친 후 개념적 설계를 통해 사실과 차원 구조를 생성한다. 다음으로 개념적 설계 결과에 대한 작업량과 유효성을 판단한 후 테이블 구조로 실체화하기 위한 논리적 설계 단계를 거치고, 마지막으로 물리적 설계를 수행한다.

〈표 1〉 OLTP 및 OLAP의 특성비교

| 구분 | OLTP | OLAP |
|----|----------------------|-----------------|
| 기반 | 작업 흐름 기반(업무 프로세스 중심) | 분석 수행 기반(주제 중심) |
| 목적 | 트랜잭션 처리(입력/조회/삭제/수정) | 보고서, 분석, 계획 |
| 특징 | 평면적, 정규화 | 다차원, 계층구조 |
| 구조 | 상세 데이터, 중복성 배제 | 요약 정보, 중복성 수용 |
| 관점 | '무엇'에 초점 | '왜'에 초점 |

데이터웨어하우스에서 다양한 분석정보를 위해 다차원 데이터를 효율적으로 표현할 수 있는 여러 개념적 모델들이 제안되었다[6, 7]. 즉, 데이터웨어하우스를 설계하기 위해 체계적인 단계를 제시하고, 데이터웨어하우스의 다차원 데이터를 개념적으로 표현하기 위한 DFM(차원 사실 모델)을 제안하였다. [5]에서는 운영 데이터베이스의 개념적 설계에 널리 활용되는 ER 도형을 확장하여 다차원데이터를 개념적으로 모델링 하기 위한 starER이라는 모델을 제시하였다. 이러한 모델링을 기초로 [7]에서는 실제 프로젝트에 이용하기 위한 체계적인 방법론을 제시하였다. 제시된 방법론을 이용한 다양한 구축사례도 발표되었다[8, 9].

3. 사용자의 요구사항 정의의 문제점

3.1 사용자 요구사항 정의과정

(그림 1)에서 제시된 OLAP 시스템의 기존 구축과정을 분석하면 다음과 같다. 1) 사용자의 요구사항을 분석하여 정의하고, 2) 이 요구사항에 따라 스타 스키마 모델링을 수행하며, 3) 그 모델에 따라 데이터웨어하우스를 구축하고, 4) 데이터웨어하우스로부터 사용자의 편리한 질의를 위해 OLAP 시스템을 구축한다. 위의 과정들 중에서 문제가 되는 것은 OLAP시스템 구축을 위한 첫 단계인 사용자의 요구사항 정의이다. 요구사항 정의의 목적은 사용자로부터 업무분석에 필요한 의미 있는 정보 요구 조사서를 도출하기 위해서이다. 요구사항 정의에 필요한 정보 요구 조사서와 주제 영역 정의서를 작성하기 위해서는 다음과 같은 세 단계가 필요하다.

첫 번째 단계는 사용자로부터 데이터베이스 레이아웃(Layout), 현업에서 사용하고 있는 각종 분석 보고서와 각종 코드집 등을 요청하여 수집한다. 제안내용과 수집한 자료에 근거하여 주제영역을 도출해내고, 이를 기반으로 정보 요구 조사서를 작성한다. 이를 위해, 정보 요구 조사의 목적과 기본 개념의 설명이 필요하다. 즉, 주제영역/분석관점/분석항목 등에 대한 실 예를 제시하여 구체적으로 설명한다. 정보 요구

조사서가 작성되면 부서별로 제시된 정보요구 사항을 구체화하여 현실성을 검토하고, 필요한 경우 아이디어를 도출하기 위해 현업 각 부서를 상대로 인터뷰를 실시한다. 인터뷰 할 내용은 업무 팀의 주된 역할, 업무팀원 구성, 업무 진행 방식(수작업 또는 전산시스템) 등이며, 현업 보고서(분석보고서, 코드집 등)를 분석하여 정보요구 사항의 필요 배경을 확인한다.

두 번째 단계는 정보요구를 정의하고, 업무범위를 확정한다. 전사적 관점의 분석정보(분석영역/분석관점/분석항목)를 최종 정리하고, 현업 담당자로부터 확인을 받는다. 이는 시스템 구축 기간이나 범위 내에 적용 가능한 정보요구를 고정한 작업이다. 취합된 정보요구를 토대로 양식에 맞게 주제영역을 정의한 후, 회사 관련자(전담팀, 현업팀)들이 함께 공유하여 최종확인 하는 순으로 진행된다.

세 번째 단계는 확정된 업무 범위를 통해 주제 영역 정의서를 작성한다. 먼저 주제영역을 정의하고, 도출된 정보요구를 주제영역별로 분류한다. 또한, 정보요구를 만족하도록 분석관점과 분석항목을 정의하고, 관련성이 높은 분석항목별로 분류한다.

3.2 사용자 요구사항 정의의 문제점

앞 절에서 설명된 세 번째 단계의 주제 영역 정의서 작성에 다음과 같이 사용자의 요구사항에 대한 두 가지 측면의 문제점이 발생한다.

첫째, OLAP에 대한 이해의 부재로 인한 문제점이다. OLAP 시스템은 의사 결정지원 시스템을 위한 다차원 분석시스템을 구축하는 것으로, 기존의 단순 보고서와는 다르다. OLAP라는 지원 시스템을 통해 원하는 시점에 원하는 데이터를 가공하여 즉시 의사결정을 할 수 있도록 지원하는 것이 최종목적이다. 기존의 가공되지 않은 단순한 형태의 보고서 개념을 가지고 있는 고객은 시스템 구축에 비용만 낭비하고 OLAP를 제대로 활용하지 못하게 된다.

둘째, 고객이 가지고 있는 데이터에 대한 파악 부족의 문제점이다. 오랜 기간 축적된 데이터이기 때문에, 신규코드의 추가 또는 기존코드의 변경을 파악하지 못하여, 전체 시스템을 다시 설계하여 구축작업을 진행해야 하는 경우도 발생한다. 또한 기존의 구축된 시스템과의 연계에 따른 매핑 문제가 발생한다. 매핑되지 않는 코드로 인해 실제 존재하는 데이터의 코드가 없는 경우가 발생할 수 있다. OLAP에서는 데이터웨어하우스의 집계 코드와 기존에 구축된 시스템의 코드와의 일치성이 보장되어야 하기 때문에 이러한 매핑 문제가 발생하게 된다.

3.3 문제점 해결을 위한 제안

앞 절에서 제시된 사용자의 요구사항 정의에 대한 두 가지 문제점을 해결하기 위해 본 연구에서는 다음과 같은 특

성을 갖는 프로토타이핑 OLAP View을 이용한 구축과정을 제안한다.

기존의 구축 과정에서는 (그림 1)에서 표시된 것 같이 데이터웨어하우스 구축을 통해 만들어지는 차원 테이블과 사실 테이블을 이용하여 OLAP 분석 보고서를 생성하였다. 그러나, 제안한 OLAP View는 이와 동일한 형태를 보이나 주제 영역 정의서를 만든 뒤에 View의 형태로 생성되어지는 OLAP View이다.

(그림 2)에서 표시된 것과 같이 주제 영역 정의서를 작성한 뒤에 프로토타이핑 형태로 생성하게 된다. OLAP View는 Microsoft Analysis services를 이용하여 스타 스키마 형태로 테이블을 직접 선정하는 방식으로 구현되는데, 차원 테이블과 사실 테이블을 복합적으로 포함된 형태로 구성된다. 즉, 구축이 완료된 후에 발생하는 문제점을 최소화하기 위해, 구축 전에 미리 OLAP View의 형태로 파일럿 시스템을 생성한다. 생성 과정은 먼저 주제 영역 정의서에서 제시된 주제영역을 참고하여 운영시스템의 테이블을 선정한다. 선정된 테이블에서 직접 차원과 측정치를 선정하고 이를 기반으로 OLAP 형태로 구성하게 된다. 이와 같은 OLAP 형태는 마지막 단계인 OLAP 구축 후 나온 결과물과 같으나, 요구사항 분석시 구축하기 때문에 OLAP View라고 정의하였다. 실제 OLAP View의 생성과정은 4장의 (그림 8)~(그림 10)에서 제시된다.

OLAP View를 활용함으로써 다음과 같은 문제해결과 활용이 가능하다.

첫 번째로 OLAP에 대한 이해 부재에 대한 문제점의 대안으로 OLAP View를 이용하여 파일럿 형태의 시스템을 미리 구축하고 OLAP에서 제공하는 분석의 형태가 기존의 단순 보고서와 어떻게 다른지 확인할 수 있고, 다차원 분석의 관점에서 시스템 구축의 방향을 설정할 수 있다.

두 번째로 데이터에 대한 파악 부족에 대한 문제점의 해결을 위해 주제 영역 정의서의 대상이 되는 업무분야 중에서 OLAP의 형태로 분석하고자 하는 테이블을 선정한다. 선정된 테이블에 대해 프로토타이핑 OLAP View를 구축하여 분석할 필요가 있는 테이블인지 또는 다른 테이블과의 통합성에 문제가 있는지 확인한다. 이 과정을 통해 최종 OLAP 분석에 필요한 적합한 테이블을 미리 선정할 수 있다. 기존 시스템과의 연계는 매핑에 필요한 코드 테이블을 선정하여 차원을 미리 생성하고 이 차원을 중심으로 사실 테이블의 모습을 확인하여 매핑 문제가 발생하는 코드를 사전에 확인할 수 있다.

4. 제안된 OLAP View를 활용한 구축

본 연구에서 활용한 구축 환경은 (그림 4)와 같이 데이터웨어하우스 서버와 OLAP서버로 구성된다.

(그림 4) 구축환경

본 연구에 사용된 시스템은 컴팩 ML570R01서버로서 Dual CPU Pentium III 850MHz, 주메모리 1GB, 보조 저장용량 100GB를 가지며, 데이터웨어하우스 서버와 OLAP서버를 동일 시스템 상에 설치하였다.

본 연구에서는 [8]에서 제시한 S 쇼핑몰 구축을 대상으로 제안된 OLAP View를 활용하여 사용자 요구사항 정의에 효과적임을 보이도록 한다. 제시된 업무 분야로는 S 쇼핑몰에서의 주문 업무인 주문을 선택하였다.

주문을 위하여 고객은 쇼핑몰 웹사이트에 회원으로 가입한 후에 온라인 고객으로 로그인하고, 상품의 카테고리를 검색하면서 원하는 상품을 선택한다. 고객이 선택한 상품을 장바구니에 담고, 쇼핑을 마친 후에 장바구니에 담긴 상품에 대해 결제를 한 후 주문을 종료한다. 쇼핑몰에서는 주문에 대한 즉시 취소가 가능하며, 배송 받은 후에 주문한 상품과 동일하지 않거나 상품에 하자가 있을 경우 주문에 대해 취소가 가능하다. 제시된 업무에 대한 개략적인 주제 영역 정의서는 (그림 5)와 같다.

- 일별/요일별/시간대별 주문건수, 주문금액, 취소건수, 취소금액
- 월별 가장 많이 주문된 상위 10개의 상품
- 월별 주문이 전혀 없었던 상품분류
- 결제방법별 주문비율, 주문취소비율, 판매비율
- 회원 연령대별 주문비율, 주문취소비율, 판매비율, 클레임비율
- 회원 거주지역별 주문비율, 주문취소비율, 판매비율, 클레임비율

(그림 5) 주제 영역 정의서

쇼핑몰의 주문에 관한 주제영역을 작성하기 위해 여러 관점들이 필요한데, 일자의 관점에 대한 차원은 (그림 6)과 같이 설계가 가능하다.

(그림 6) 일자의 관점에 대한 차원 설계

일자(Date) 차원을 중심으로 한 사실 테이블은 일자와 시간의 키값을 중심으로 한 스타스키마 형태의 모델 구성이 가능하다. 쇼핑몰의 주문업무에 대한 스타스키마 모델은 (그림 7)과 같이 표현될 수 있다. (그림 7)의 스타스키마 테이블은 (그림 1)에서 설명된 기존 구축과정 중에서 스타스키마 모델링 및 데이터웨어하우스 구축에 사용된다.

기존 구축과정과의 비교를 위해 본 논문에서는 제시된 주문 사실 테이블에 대해 OLAP View를 구축하는 과정을 다음과 같이 세 단계로 제시한다.

(그림 8) 일자 차원 테이블 생성

(그림 7) 주문에 대한 사실 테이블 설계

첫 단계로 OLTP 테이블 중에서 주문 테이블을 찾는다. 주문 테이블에는 주문에 사용하는 각종 수치 값과 코드 값이 들어있다.

두 번째 단계로 주문 테이블에서 최근 한 달 기간에 해당하는 데이터를 추출한다. 통상적으로는 3년 이상의 데이터를 분석해야 하나, 제안된 OLAP View에서는 구축될 형태를 미리 확인하는 것이 목적이기 때문에 전체 데이터를 대상으로 하지 않고 최근의 데이터를 대상으로 한다.

세 번째 단계로 Microsoft Analysis Service에서 주문 테이블을 중심으로 선택된 한 달 치의 데이터를 대상으로 직접 OLAP의 형태를 구성하여 주문 사실 테이블을 생성한다. 이와 같은 구성이 가능한 이유는 Microsoft Analysis service가 사실 테이블을 만드는 기능을 제공하기 때문이다. 기존 구축과정에서는 차원 테이블을 설계 및 구축, 사실 테이블을 설계 및 구축한 후에 이를 이용하여 OLAP 형태로 구성되었다. 그러나 제시된 OLAP View에서는 주문 테이블에서 직접 수치 값을 선택하여 차원을 만들고 테이블 자체를 사실 테이블로서 선택하여 구성한다.

이 과정은 (그림 8) 및 (그림 9), (그림 10)과 같이 진행된다. (그림 8)은 (그림 6)의 일자 차원 설계에 대한 실제 구축 과정이다.

(그림 9) 주문 사실 테이블 생성

다. 즉, 보다 많은 상품을 보다 많은 사람에게 판매하여 매출증대를 이루고자 한다. 그러나 온라인 쇼핑물 시장이 점점 포화상태에 이르러, 지속적인 매출향상은 어려운 실정이다. 이의 대안이 쇼핑물 운영의 손실을 줄이는 것이다. 예를 들어, 불량고객을 선별 관리하거나 제품 반송율을 줄여 이익을 얻을 수 있다.

본 연구의 대상인 S 쇼핑물은 손실을 줄이기 위해서 기존의 평면적인 분석으로는 원하는 결과를 얻기 힘들어 다차원 분석을 위한 OLAP을 도입하게 되었다. 그러나 OLAP을 도입할 당시에는 OLAP에 대한 이해의 부족, 어떻게 다차원 분석을 해야 하며, 분석 결과를 어떻게 이용해야 손실을 줄일 수 있는지가 불확실 하였다. 이와 같이, 실제 OLAP 시스템을 구축하기 전에 어떤 데이터를 다차원으로 분석해야 하는지 파악하기 힘들다. 또한 주제 영역 정의서에서 정의된 업무 분야가 실제 구축된 시스템에 정확하게 반영되지 않는 경우가 발생한다.

그러나 제안된 OLAP View를 활용하면 주제 영역 정의서에 대상이 되는 분석해야할 데이터를 미리 확인할 수 있고, 손실을 줄이기 위해 어떤 데이터를 분석해야 하는지 테스트를 실행할 수 있게 된다. 이를 기반으로 실제 시스템을 구축하면 다차원 분석 시스템 구축에 필요한 시간과 비용을 줄일 수 있다. 또한 고객의 요구사항을 명시한 주제 영역 정의서에 있는 각 주제영역에 대해 파일럿 시스템 구축이 가능하다. 파일럿 시스템을 통해 주제 영역 정의서에 있는 고객의 요구사항을 확인하여 상세하고 정확한 요구사항을 전체 시스템 구축과정에 반영시킬 수 있다. 이는 구축과정 중에 잘못 반영될 수 있는 요구사항의 오류를 줄여서 전체 구축 일정을 줄일 수 있게 된다.

5.2 데이터의 정확한 파악

S 쇼핑물에 대한 실제 분석시스템의 구축과정에서 3.2절에서 제시한 바와 같이 코드의 불일치, 전체 테이블 파악 미비, 불확실한 매핑 등의 기존에 관리되었던 데이터에 대한 문제점이 발생될 수 있다.

데이터에 대한 문제점은 테이블에 대한 정의가 정확하거나 분명하지 않은 것이다. 즉, 기존의 반복적 시스템 구축으로 인해 가장 최신의 ER(Entity Relation) 다이어그램이 존재하지 않게 된다. 결국, 과거 ER 다이어그램을 기초로 분석 데이터를 생성하였는데 데이터 형태의 불일치로 인해 분석 데이터를 다시 생성해야 하는 일이 발생한다. 예를 들어, 숫자가 입력되어야할 필드에 문자가 입력되는 경우가 있다. 또는 코드값이 불일치하는 경우도 있고, 코드값에 널 값이 들어 있는 경우도 있다. 또한 테이블에 대한 참조가 정확하지 않아 외래 키(Foreign Key)의 불일치가 발생하는 경우가 빈번하다. 주문 테이블에는 남아있지만 고객 테이블에는 없는 고객이 있는 경우도 발생하였다. 이로 인해 정확한 데이

(그림 10) OLAP View

(그림 9)는 주문에 대한 사실 테이블을 생성하는 과정이다. 생성과정은 다음과 같다.

(그림 10)은 (그림 9)에서 제시한 주문 사실 테이블에 대한 OLAP View를 보여준다.

기존의 구축과정에서는 분석 및 모델 설계가 끝나고, 코드에 해당하는 테이블을 선정한 이후에 각 코드에 대한 차원 테이블을 생성하였다. 이로 인해 3장에서 제시한 데이터의 불일치, 잘못된 매핑과 같은 문제점이 발생한다. 이러한 문제점이 초기에 발견되면 다행이지만 구축이 끝나고 테스트 시점에 발견 될 경우 전체 시스템을 다시 구축해야 한다.

그러나 제안된 OLAP View를 활용하면 (그림 9)와 같이 주문에 해당하는 테이블을 선정한 이후에 직접 차원 테이블을 생성한다. 이 과정은 기존 시스템의 구축단계에서 OLAP을 생성하는 것과 달리 분석 단계에서 진행되므로, 3장에서 발생한 문제점을 최소화함으로써 재구축의 요인을 줄여 시간과 비용을 최소화 할 수 있다. 즉, 구축이 완료된 후에 발생하는 문제점을 최소화하기 위해, 구축 전에 미리 OLAP View의 형태로 파일럿 시스템을 생성한다. 생성된 차원 테이블은 OLAP View의 형태로 구성되며, 주제 영역 정의서에서 제시된 주제영역에 대해 OLAP View를 활용하여 의사 결정 시스템을 지원할 수 있게 된다.

5. 제안 기법의 특성 분석

4장에서 제시된 S 쇼핑물의 주문업무 구축사례를 통해 사용자 요구사항 분석 및 정의를 위한 프로토타이핑 OLAP View의 활용 특성과 장점을 다음과 같이 제시할 수 있다.

5.1 OLAP 구축 전략의 정립

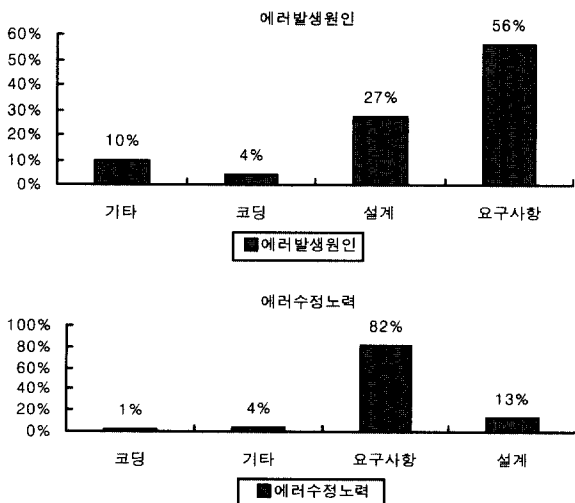
쇼핑물은 매출을 향상시켜 이익을 증대하는 목적을 가진

터의 양을 분석하지 못하고, 데이터의 정확도가 떨어지게 된다. 이러한 불일치는 구축 후 테스트 단계가 되어야 파악되는 경우가 있어 처음부터 다시 데이터의 적합성을 맞추어야 하고, 이로 인해 매우 많은 노력과 시간이 소요된다. 또한 기존의 구축과정에서 상품분류의 변경과 같은 코드체계의 변경은 전체 OLAP을 구성한 후에 수정해야 하므로 노력과 시간의 낭비가 심하다.

이와 같은 오류를 줄이기 위해 프로토타이핑 OLAP View를 통해 구축될 모습을 미리 확인하여 앞에서 제시된 불일치를 제거할 수 있다. 또한, 데이터에 대한 정확한 파악을 통해 오류의 발생을 미리 예측하고, 실제 프로젝트 수행 시 오류에 대한 교정이 가능하게 된다. 그리고 프로토타이핑 OLAP View를 구축한 이후에 사용자의 확인을 거쳐 필요 없는 데이터에 대해 필터링(filtering)과 클렌징(cleansing)이 가능하다. 마지막으로 제안된 OLAP View를 통해서 상 품 분류 체계를 수정한 뒤에 변경에 의해 파생되는 문제점을 확인하면서 전체 OLAP을 구성하지 않고도 반복적인 구축 및 확인이 가능하다.

5.3 OLAP View 구축 전략의 평가

요구사항 결함에 따른 전체적인 시스템 구축의 오버헤드를 분석하면 다음과 같다. James Martin에 의하면 (그림 11)과 같이 결함의 발생원인 중에서 요구사항 결함이 전체의 56%나 되고 결함 수정노력의 82%가 요구사항 결함을 수정하는데 사용되며[11], Bohm이나 Dion에 의하면 전체 프로젝트 비용의 40%~50%가 이러한 결함 수정을 위한 재작업에 투입되고 있다. 따라서 전체 프로젝트 비용의 30%가 재작업에 사용되고, 재작업 비용의 70%가 요구사항에 기인한다고 가정하면, 요구사항 에러를 10%만 감소시키면 요구사항 관리에 투입되는 비용의 153% 또한 20%를 감소시키면 407%의 투자대비 효과를 기대할 수 있다[12].



(그림 11) 시스템구축의 오버헤드

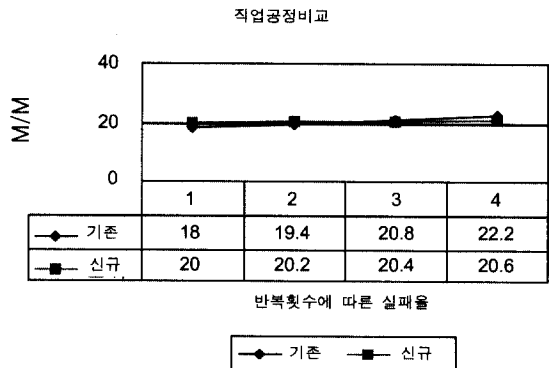
이와 같은 연구에 따라서 실제 S 소평몰 구축 시 투입되었던 구축 기간을 기반으로, 본 연구의 OLAP View를 도입했을 경우의 구축 기간을 <표 2>에서 비교 분석한다.

<표 2> 작업공정 비교 (M/M)

| 작업공정 | | 기존 | 기존 30% | 제안 | 제안 30% |
|--------------------|------------------|-----|--------|-----|--------|
| 데이터웨어 하우스/ OLAP 계획 | 프로젝트 준비 | 1 | 1 | 1 | 1 |
| | 정보요구사항 조사 및 정의 | 1.5 | 1.5 | 1.5 | 1.5 |
| | OLAP 데이터 요구사항 도출 | 1.5 | 1.5 | 1.5 | 1.5 |
| 데이터웨어 하우스/ OLAP 구축 | 주제영역정의 | 2 | 2.6 | 2 | 2 |
| | OLAP VIEW 구축 | | | 2 | 2.6 |
| | 스타스키마모델링 | 3 | 3.9 | 3 | 3 |
| | 데이터웨어하우스 구축 | 5 | 6.5 | 5 | 5 |
| OLAP 구축 | | 4 | 5.2 | 4 | 4 |
| 합 계 | | 18 | 22.2 | 20 | 20.6 |

M / M: Man/Month(시스템 구축시 기간과 투입비용의 기준으로 <표 2>의 수치값)
 기존: 실제 S 소평몰 구축시 투입된 M/M
 기존30%: 기존 방법론을 적용시 실패비용을 30%로 가정
 제안: OLAP View 구축 방법이 추가된 경우
 제안30%: OLAP View 구축 방법 적용시 실패비용을 30%로 가정

<표 2>를 분석하면 다음과 같은 결과를 얻을 수 있다. 기존의 방법으로 작업하였을 경우 실패 비용(재작업 및 반복 비용)을 30%로 보았을 때 총 M/M는 22.2 M/M이 된다. 그러나 제안된 OLAP View를 이용한 구축 전략의 경우에는 OLAP View 구축의 실패비용만 추가되기 때문에 전체적인 M/M는 20.6이 된다. 이와 같이 실패 비용적인 측면에서 반복 실행에 대해 상대적인 비교가 가능하다. (그림 12)에서 보듯이 실패에 따른 반복횟수가 증가할수록 실패비용이 증가하게 되고, 3회 이상 반복 시 OLAP를 이용한 신규방법이 기존의 방법보다 실패비용이 더 낮음을 알 수 있다. 즉, 실패로 인한 반복작업이 증가할수록 OLAP View를 이용한 구축 전략이 더 효과적임을 알 수 있다.



(그림 12) 작업공정비교

6. 결론

지식 정보화 사회에서의 경영환경은 과거보다 더욱 경쟁

적이고 역동적이어서, 기업의 사활은 정보시스템의 상황 적응 능력과 정보를 분석/활용하는 능력에 의존하게 된다. 따라서 정보를 의사결정에 반영시킬 수 있는 효과적인 의사결정시스템을 필요로 한다. 다차원 분석을 이용한 OLAP시스템은 사용자가 데이터를 여러 관점에서 비교 분석하여 필요한 데이터를 업무 분야별로 구조화함으로써 분석항목의 범위를 명확히 하고, 경영자에게 전략적인 방향을 제시해 줄 수 있다. 그러나 기존에는 OLAP의 설계 및 구축이 진행되는 과정 중에 사용자의 잦은 요구사항의 변경으로 모델링과 구현이 반복되어 전체 시간과 비용이 증가하는 비효율성을 가진다. 본 논문에서는 사용자의 정확한 요구사항을 도출하기 위해 스타스키마에 기반을 둔 프로토타이핑 형태의 OLAP View를 제안하고 구현하였다. 제안된 OLAP View의 구축을 통해 사용자의 요구사항을 보다 더 정확히 파악할 수 있게 되고, OLAP 구축 전략 수립이 가능하며, 데이터에 대한 정확한 이해와 파악이 가능함을 보였다. 결과적으로 OLAP View를 사용함으로써 분석 시스템 구축에 필요한 시간과 비용을 줄일 수 있게 된다.

본 연구를 개선하기 위한 향후 과제는 OLAP에 필요한 데이터웨어하우스를 효율적으로 구축하기 위한 OLAP 메타데이터에 대한 연구이다.

참 고 문 헌

[1] Pendse, N. and Creeth, R., "The OLAP Report : Succeeding with on line analytical processing," Vol.1, Business Intelligence, 1995.
 [2] Immon, W. H., Building the Data Warehouse, 2nd Ed. NewYork : John Wiley and Sons Inc., 1996.
 [3] Kimball, R., Reeves. L., Ross, M. and Thornthwaite, W., "The Data Warehouse LifeCycle Toolkit," New York : John Wiley and Sons Inc., 1998.
 [4] Poe, V., "Building a Data Warehouse for Decision Support," Upper Saddle River, Prentice Hall, 1996.
 [5] Tryfoan, N., Busborg, F. and Chritiansen, J. G. B., "starER : A conceptual Model for Data Warehouse Design," Proc. of DOLAP, 1999.

[6] M. Golfarelli and S. Rizzi, "A Methodological Framework for Data Warehouse Design," Proc. of DOLAP, pp.3-9, 1998.
 [7] Kimball, R., "A Dimensional Manifesto," DBMS, August, 1997.
 [8] MS Enterprise 2001, "삼성물 분석CRM 사례", 위세아이텍, October, 2001.
 [9] Microsoft 구축사례, "SQL 서버 2000 기반의 DW/OLAP", 삼양사, <http://www.microsoft.com/korea/sql/evaluation/case-studies/2000/samyang.asp>, December, 2001.
 [10] 권오주, "OLAP Solutions + α ", 대림출판사, 2001.
 [11] Gray Mogyorodi, "What is Requiriement based Testing?," Technology Builder Inc, <http://www.tbi.com/products/asq/>.
 [12] Dean Leffingwell, "Calculating Your Return on Investment from More Effective Requirements Management," Rational Software Corporaton, <http://www.rational.com/products/reqpro/whitepapers.jsp>, 1997.

최 승 교

e-mail : skchoi@samcheok.ac.kr
 1982년 단국대학교 전기공학과(학사)
 1992년 단국대학교 대학원 전산통계학과 (석사)
 2001년 단국대학교 대학원 전산통계학과 (박사)
 1994년~현재 삼척대학교 컴퓨터공학과 부교수
 관심분야 : 컴퓨터 구조론, 성능평가, 시뮬레이션, 컴퓨터통신

박 종 모

e-mail : tinytree@nate.com
 1995년 단국대학교 전자계산학과(학사)
 1997년 단국대학교 대학원 전산통계학과 (석사)
 1997년~2000년 (주)한국정보시스템 선임 연구원
 2000년~2003년 (주)위세아이텍 책임 연구원
 2003년~현재 단국대학교 대학원 전산통계학과 박사과정
 2003년~현재 (주)이지엠텍 경영지원실 전산개발팀
 관심분야 : 전사적자원관리(ERP), 고객관계관리(CRM), 컴퓨터 통신