

초기 데이터 웨어하우스 모델링을 지원하기 위한  
사례기반 추론의 응용

A Case-Based Reasoning Application to Support  
Initial Data Warehouse Modeling

이재식\*, 전용준\*\*

\*아주대학교 경영학부 경영학과 부교수

\*\*아주대학교 대학원 경영학과 박사과정

경기도 수원시 팔달구 원천동 산 5번지

Phone#: 0331-219-2719 Fax#: 0331-29-2190

E-Mail: leejsk@madang.ajou.ac.kr

Abstract

Since the primary purpose of Information Engineering focuses on transaction or operation processing, various information needs acquired in Information Strategy Planning phase are not properly utilized from the viewpoint of decision support systems development. In this research, we suggest a case-based reasoning application that supports initial Data Warehouse Modeling by expanding the activities in Information Strategy Planning.

1. 서론

최근에 들어 많은 기업들이 정보공학(Information Engineering)[Martin, 1989]을 시스템 개발 및 운영을 위한 방법론적 기반으로 사용하여 경쟁력 있는 정보 시스템을 구축하기 위한 프로젝트를 수행하고 있다. 하지만 정보공학이 일반적으로 고려하고 있는 개발 대상 시스템은 운영적(Operational) 관점이 강하며, 이제까지 기술적인 측면에서 전략적(Strategic) 또는 관리적인(Managerial) 목적에 부합되는 업무 정보의 제공을 위한 시스템을 구축하는 것이 어려웠다. 최근에 대두되고 있는 데이터 웨어하우스(Data Warehouse)의 개념은 이와같은 문제점을 해소하기 위한 하나의 방안으로 볼 수 있다. 기존의 정보공학에는 데이터 웨어하우스를 구축하기 위한 구체적인 방법이 제시되지 못하였으므로 정보공학과 데이터 웨어하우스를 연결하는 방법론적인 접근이 필요한 시점이다.

이 연구는 사례기반 추론(CBR: Case-

Based Reasoning) 기법을 활용하여 정보 전략 계획(ISP: Information Strategy Planning) 단계에서 전략적인 관점의 데이터 웨어하우스 설계가 가능하도록 하기 위하여 도출된 여러가지 전략 개념들로 부터 정보 요구 사항(IN: Information Needs)을 도출하고 도출된 정보 요구 사항을 만족 시켜줄 수 있는 초기 데이터 웨어하우스(Data Warehouse) 모델의 설계를 지원하는 시스템을 제시한다.

제 2 절에서는 이론적 배경으로 정보 전략 계획과 데이터 웨어하우스, 그리고 사례기반 추론의 기본 개념을 간략히 소개하고, 제 3 절에서는 데이터 웨어하우스 설계를 고려하기 위한 정보공학 수행경로의 확장을 제안한 후, 데이터 웨어하우스에 대한 초기 설계를 지원하기 위한 사례기반 추론 시스템 CBiDWM(Case-Based Support System for Initial Data Warehouse Modeling)의 기본 구조와 유용성등에 대하여 설명하며, 제 4 절에서는 결론 및 향후 연구방향을 제시한다.

2. 이론적 배경

ISP란 정보공학의 단계중 최초의 단계이며, 전사적 관점에서의 정보체계를 수립하는 과정이다[Martin, 1989]. 정보공학은 Finkelstein과 Martin[Finkestein and Martin, 1981]이 그 개념을 제시한 이후 1980년대를 거치면서 급속히 전파되었고, 오늘날에는 전사적 또는 부문별 기업 정보시스템 개발 방법의 표준의 하나로 자리 잡게 되었다. 정보공학에서의 단계 구분에 대해서

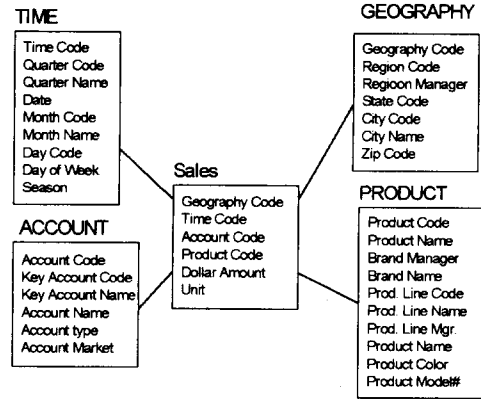
는 Navigator, IEF Composer, TCSM(The Client Server Methodology)등 상용화된 방법론의 공급자들에 따라 여러가지 명칭들과 서로 다른 작업 목록들이 제시되고 있으나, 전체적인 체계는 유사하다. 정보 전략 계획 단계에서는 컴퓨터 시스템의 관점 보다는 사용자의 관점에서 기업의 전략을 확인하고 이를 기반으로 기업의 정보 전략을 수립하는 것이 주된 작업이다. 이 과정에서 도출되는 것이 정보 요구사항[Texas Instruments Inc., 1995]이다. 정보 요구사항은 장차 개발될 시스템이 제공해야 할 전사적으로 중요한 정보의 목록이다. 이 단계에서 도출된 정보에 대한 요구는 단순한 운영 수준의 데이터만으로는 충족될 수 없으며, 필연적으로 전사적으로 수집된 기초 데이터에 대한 요약, 집계, 가공등의 작업이 수반되어야 한다.

한편 데이터 웨어하우스는 최근 산업계의 주목을 받고 있는 개념으로, '경영 의사 결정을 지원하기 위한 주제 지향적이고, 통합적이며, 시간의 흐름을 담고 있고, 휘발성을 가지지 않은 데이터의 집합'라고 정의할 수 있다[Inmon and Hackathorn, 1994].

데이터 웨어하우스 개발을 위한 데이터베이스 구조 모델링 방법으로는 크게 Star schema와 Snow Flake Schema 등의 방법이 사용되고 있다[Inmon and Hackathorn, 1994; Kimball, 1996; Stanford Technology Group Inc., 1996]. <그림 1>은 Stanford Technology Group의 자료[Stanford Technology Group Inc., 1996]를 수정하여 작성한 Star Schema의 예이다. <그림 1>에서 Sales는 Fact Table이며, Time, Geography, Product, Account는 Sales를 분석하기 위한 Dimension Table이다. 즉, 이 모델에서 Sales는 4가지 차원들의 조합을 분석의 차원으로 사용할 수 있다. Snow Flake Schema는 스타 스키마를 사용한 차원 테이블들이 상호간에 가지는 관계를 포함할 수 있도록 스타 스키마를 확장 시킨 것이다[Stanford Technology Group Inc., 1996].

한편 사례기반 추론은 최근에 대두되고 있는 전문가 시스템의 한 유형으로 현재 직면한 문제와 유사한 과거의 문제를 해결한 경험으로부터 현재의 문제에 대한 해를 구하는 방식을 사용한다[Kolodner, 1993]. 사례기반 추론은 지식 획득이 용이하여 개발에 소요되는 시간 및 비용이 적다는 장점을 가지고 있다[Turban, 1992].

<그림 1> Star Schema의 예



### 3. 정보 전략 계획 수행경로의 확장파 사례 기반 추론을 통한 지원

전통적인 정보 공학 방법론의 일반적인 ISP 수행 경로는 <그림 2>의 왼쪽에 보이고 있는 바와 같이 먼저 기업의 전략 개념을 도출한 후, 정보 요구 사항을 도출한다. 이를 기반으로 업무 영역 분석(BAA: Business Area Analysis)을 수행한 후 주로 운영적 시스템의 구축으로 이어지게 된다.

이와 같은 수행 경로는 데이터 웨어하우스의 개념이 등장하기 이전에 제안된 것이며, 주로 일상적인 거래 처리 시스템(Transaction Processing Systems)을 구축하기 위한 것이다. 그러므로 의사 결정 분석에 대한 요구 사항에는 대응하지 못하고 있다. 따라서 전통적 수행경로는 수정되어야 할 필요성이 있다.

이러한 문제점을 해결하기 위하여 본 연구에서 제안하는 새로운 수행경로는 <그림 2>의 오른쪽에 제시되는 흐름을 가진다. 이 경로에서는 전략 개념이 확정되면 정보 요구 사항을 도출하고, 이들을 충족시킬 수 있는 요약된 수준의 개념적인 초기 데이터 웨어하우스 모델을 개발하여 제시한다. ISP 프로젝트에 참가한 구성원들과 중역들의 검증이 이루어지고 나면, 이 모델은 데이터 웨어하우스 초기 모델로 확정되며, 운영적 시스템의 개발을 위한 BAA가 아닌 별도의 데이터 웨어하우스 구축 프로젝트에 사용된다. 이러한 수행 경로는 운영적 시스템과 의사 결정의 지원을 위한 시스템 모두에 대하여 구축 과정상의 복잡성을 감소 시키는 효과를 가져오게 되며, 상위 단계에서 정보 시스템 구축을 위해

주어진 전체 자원중에서 데이터 웨어하우스 구축을 위한 자원을 할당하는데 중요한 정보를 제공하게 된다. 그러나 앞에서 제시한 경로를 따르는 경우 하나의 추가적인 산출물이 되는 초기 데이터 웨어하우스 모델의 개발은 통상 3개월 이상이 소요되는 ISP의 수행에 또 다른 부담을 제공할 수 있다. 이에 본 연구에서는 초기 데이터 웨어하우스 모델링 지원 시스템(CBiDWM)을 제안한다.

사례기반 추론으로 모델링을 지원한다는 것은 모델링을 수행하는 과정에서 일반화된 패턴(Generic Pattern)을 사용하는 접근 방식[Hay, 1996]과 유사하다.

<그림 3>은 CBiDWM의 구조와 입출력을 보여준다. CBiDWM은 크게 정보 요구 사항의 설계를 제공하는 부분(IN Designer)과 iDWM(Initial Data Warehouse Model)을 설계하는 부분(iDWM Designer), 그리고 새로운 사례를 입수하는 학습 기관(Learning Facility)으로 구성된다. <그림 3>에서 iDWM은 Star Schema로 표현되는 초기 데이터 웨어하우스 모델이다. <그림 4>는 [이재식과 전용준, 1995]의 제안에 근거하여 CBiDWM의 대상 문제를 개념적인 실체-관계(Entity-Relationship) 모델로 표현한 것이다. <그림 5>는 CBiDWM의 사례베이스에 저장되는 사례의 예이다. CBiDWM의 사례는 크게 기업의 면모에 관한 속성과 ISP 프로젝트 개요에 대한 속성, 도출된

전략 개념과 결론이 되는 IN에 대한 속성과 iDWM 등의 속성들로 구성된다. CBiDWM에 사용되는 유사성 판단의 근거가 되는 지식(Similarity Rule)들과 과거 사례에 부여되는 iDWM은 각각 정보공학 전문가와 데이터 웨어하우스 모델링 전문가에 대한 면담으로부터 획득된다.

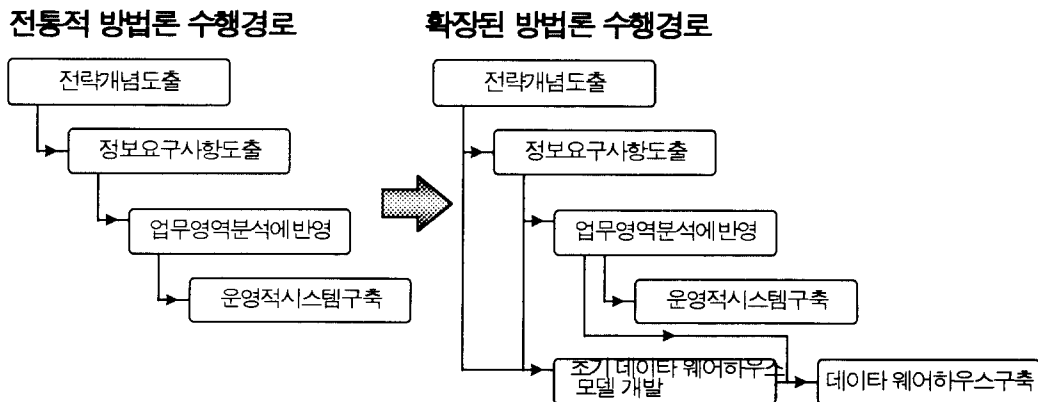
이와 같은 CBiDWM의 활용은 모델링 과정의 어려움을 감소시키고, 모델의 품질을 향상시키며, 모델링에 소요되는 시간을 단축시키는 효과를 제공할 수 있을 것으로 기대된다.

#### 4. 결론

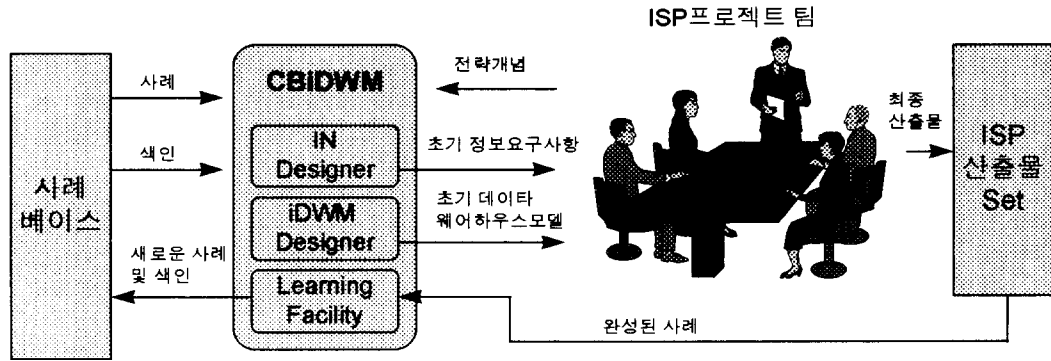
이 연구에서는 ISP 수립 과정에서 도출된 전략 개념들로 부터 IN과 초기 데이터 웨어하우스 모델을 생성 시킨다는 정보공학 방법론의 확장 방안과 사례기반 추론 기법을 활용하여 이를 지원하는 시스템의 설계를 제시하였다.

본 연구의 의의는 새로운 기술과 기존의 방법론을 접목시키기 위한 방안을 제시하였다는 점에서 찾을 수 있을 것이다. 아직까지 본 연구는 완성된 형태는 아니므로 본 연구의 기본적인 제안을 보강하는 후속 연구와 제안된 시스템의 구현 및 실제 프로젝트에의 적용을 통한 유용성 검증 작업이 수반 되어야 할 것이다.

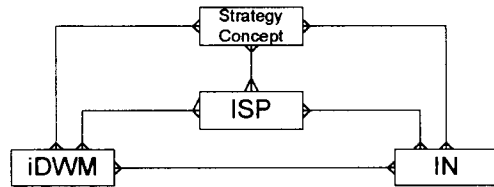
<그림 2> 데이터 웨어하우스 설계를 위한 전통적인 정보공학 방법론의 확장



<그림 3> CBiDWM 의 구조



<그림 4> 사례기반 추론에 사용되는 개념적인 데이터 모델



<그림 5> 초기 데이터 웨어하우스 모델링에 사용될 사례의 예

Company Name: HHH Inc.  
 Industry: Film Distribution  
 Number of Employee: 2000  
 Strategic Concept: Increase Customer Satisfaction  
 Index: Customer(0.9), CS(1.0), Customer Satisfaction(1.0)  
 Concept Type: Goal  
 Information Needs: Customer rating details,  
 Customer complaint volumes by reason,  
 Average customer rating by store,  
 Average customer rating by film category  
 iDWM: [CUSTOMER<= Rating, Film Category, Store],  
 [CUSTOMER COMPLAINT<=Reason]

[2] Finkelstein, C. and J. Martin, Information Engineering, Savant Institute, 1981.  
 [3] Hay, D.C., Data Model Patterns-Conventions of Thought, Dorset House, New York NY, 1996.  
 [4] Inmon, W. H., and R. D. Hackathorn, Using the Data Warehouse, John Wiley and Sons, New York, 1994.  
 [5] Kimball, R., The Data Warehouse Toolkit - Practical Techniques for Building Dimensional Data Warehouses, John Wiley & Sons, 1996.  
 [6] Kolodner, J., Case-Based Reasoning, Morgan Kaufmann, San Mateo CA, 1993.  
 [7] Martin, J., Informaion Engineering Book II, 1989.  
 [8] Stanford Technology Group Inc., "Designing the Data Warehouse on Relational Databases," White Paper, 1996.  
 [9] Texas Instruments Inc., Information Strategy Planning Workshop Student Guide, Texas Instruments Inc., 1995.  
 [10] Turban, E., Expert Systems and Applied Artificial Intelligence, Macmillan, 1992.

참고문헌

[1] 이재식과 전용준, "사례기반 추론을 위한 데이터 중심의 문제 구조화", '95 한국전문가 시스템 학회 춘계 학술대회 논문집, 1995, pp. 103-113.