

유비쿼터스 데이터 웨어하우스: RFID와 다차원 온라인 분석의 통합

조대연*

*한동대학교 경영경제학부

Ubiquitous Data Warehouse: Integrating RFID with Multidimensional Online Analysis

Cho, Dai Yon

Handong Global University

E-mail : dyjoh@handong.edu

요 약

최근 RFID가 비즈니스의 여러 분야에서 추적시스템을 중심으로 도입되기 시작하고 있으며, 이러한 시스템들이 기업에 괄목할만한 효율성의 증가와 비용의 감소를 가져올 것으로 기대되고 있다. 한편, 기업의 대용량 정보저장고로 사용되어 온 데이터 웨어하우스는 생태적으로 과거의 정적인 데이터를 분석하도록 디자인 되었으며, 온라인 분석도구인 OLAP은 데이터 웨어하우스에 저장된 정적 데이터를 분석하는 도구로 사용됨으로 의사결정 지원의 역할을 하고 있다. 그렇다면 RFID를 통하여 실시간으로 수집되는 정보가 OLAP과 결합할 경우 그 정보는 추적과 같은 단순한 정보분석이 아니라 실시간 기반의 보다 중요한 의사결정을 위하여 사용될 수 있을 것이다. 본 연구에서는 데이터 웨어하우스의 정보소스를 확장하기 위한 노력의 일환으로 RFID와 데이터 웨어하우스의 통합모델을 제안한다. 그와 함께 RFID 기기를 통하여 수집된 실시간 기반의 동적인 정보를 분석할 수 있는 OLAP을 제시하고 있다. 그리고 이러한 개념을 프로토타입으로 구현함으로써 유비쿼터스 컴퓨팅 기술의 핵심을 이루고 있는 RFID가 데이터 웨어하우스에 정보소스를 제공할 수 있으며 온라인 분석도구와 결합될 경우 보다 강력한 의사결정 지원도구가 될 수 있음을 보여 준다.

1. 서론

최근 유비쿼터스 컴퓨팅의 핵심기술인 RFID가 기업의 공급사슬관리[1,2], 공항에서의 수하물 추적[3], 병원에서의 장비와 사람 추적[4], 그리고 학교와 놀이공원에서 미아를 추적하는 데 이르기까지 광범위하게 사용이 되고 있다[3]. RFID를 기반으로 하는 추적시스템이 프라이버시와 관련된 우려를 낳고 있기도 하지만[5], 기업에서는 RFID를 활용하여 재고 추적을 보다 용이하게 함으로서 매출 대비 재고비율을 낮추고 상당한 정도의 효율성 증진과 비용감소를 이루고 있다[6].

그러나 RFID 리더기로부터 입력되는 정보는 단순히 상품이나 수하물, 장비나 사람을 추적하는 것을 넘어서서 보다 유용한 목적을 위하여 사용될 수 있을 것이다. OLAP이 데이터 웨어하우스에 저장된 정보의 분석 기능을 제공하는 도구로써 중요한 의사결정지원 도구로 사용될 수 있다는 것이 여러 연구에서 증명된 만큼[7,8], 만일 OLAP과 결합된다면 RFID를 통해 입력된 실시간 기반의 동적인 정보는 의사결정을 위한 아주 중요한 정보소스가 될 수 있을 것이다.

데이터 웨어하우스는 다양한 정보소스로부터 수집된 정적인 과거의 정보들을 한 곳에 모아놓은 저장고라고 정의할 수 있다. 그동안 이러한 데이터 웨어하우스의 정보소스를 확장하기 위한 노력들이 이루어졌는데, 그 예로 Jensen과 Moller[9]는 데이터 웨어하우스의 정보소스를 XML을 기반으로 하는 정보로 확장하기 위한 연구를 하였다.

본 연구는 새로운 정보소스로 RFID를 데이터 웨어하우스에 통합한 프로토타입을 제시한다. 원래 과거의 정적 데이터의 분석도구로 디자인된 OLAP을 RFID 기술과 접목시킴으로써 본 연구에서 제안하는 데이터 웨어하

우스 시스템은 RFID 리더기로부터 입력된 실시간의 동적인 정보를 분석하는 기능을 제공한다.

2. 다차원 데이터베이스와 OLAP 디자인

다차원 데이터베이스는 다차원 OLAP 시스템이나 관계형 OLAP 시스템을 기반으로 구현된다[10]. 먼저, 다차원 OLAP 시스템은 특수한 다차원 구조를 지닌 저장장치에 데이터를 저장한다. 또한 대부분의 다차원 OLAP 시스템은 질의에 대한 답을 만들기 위해 데이터를 탐색할 때 고급 인덱스 기술과 해싱을 이용하고 있다. 한편 관계형 OLAP 시스템은 관계형 데이터베이스 기술을 사용하여 정보를 저장하며, 특별한 인덱스 구조를 사용하여 질의를 처리한다. 본 연구에서는 관계형 OLAP 시스템이 다차원 데이터베이스의 큐브를 다시 정의하거나 자주 업데이트할 수 있는 유연성을 지니고 있기 때문에 실험용 프로토타입 시스템을 구축하는데 보다 적합하다는 것을 고려하여, 다차원 OLAP 시스템보다는 관계형 OLAP 시스템을 기반으로 하여 다차원 데이터베이스를 구현하였다.

관계형 OLAP 시스템은 통상적으로 데이터를 사실테이블(fact table)과 차원(dimension)으로 저장하는 방법인 스키마(schema), 즉 스타(star) 스키마나 스노우플레이크(snowflake) 스키마 중 하나를 사용하여 구현된다. 이 두 가지의 스키마는 각각의 장단점을 지니고 있다. 스타 스키마는 데이터 중복의 문제와 차원간의 계층적 관계를 속성 속에 암묵적으로 표현함으로써 잘 드러나지 않는 문제를 지니고 있는 반면, 스노우플레이크 스키마는 각 차원의 level에 하나의 테이블을 위치시켜 계층을 명백하게 표현한다[11]. 데이터 중복의 문제가 있기는 하지만, 스타 스키마가 질의를 처리하는데 있어서 보다 탁월한 기능을 가지고 있다는 것과, 동적인 데이터를 다루기 때

문에 질의를 보다 정확하고 신속하게 처리하는 것이 중요하다는 것을 고려하여 본 연구에서는 스타 스키마를 채택하였다.

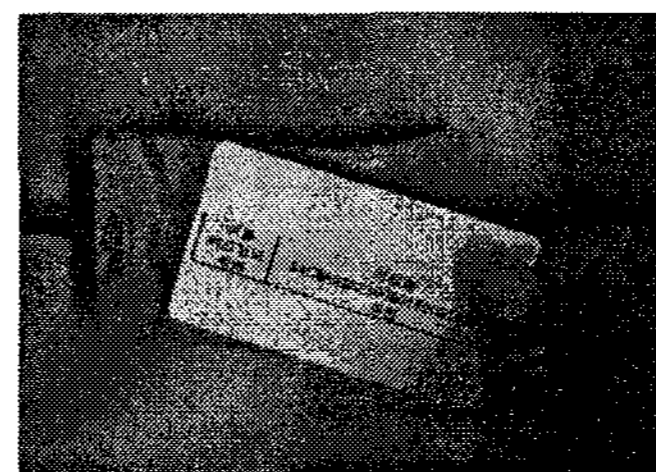
차원은 다차원 데이터베이스의 핵심에 해당한다. 차원의 목적은 사실을 표현하는데 있어 가능한 한 많은 컨텍스트를 제공하는데 있다. 일반적으로 대부분의 다차원 데이터베이스에서는 정보의 가치를 증진시키기 위하여 통제가 가능한 어느 정도의 데이터 중복을 허용한다. Body et al.[12]은 여기에서 한 걸음 더 나아가 차원의 구조를 정적으로 만드는 제약조건을 완화하여 차원이 동적으로 변화할 수 있도록 하였다. 그들의 연구는 시간이 지남에 따라 차원이 진화할 수 있도록 하는 방법을 제안하고 있다. 본 연구에서는 RFID와 OLAP의 통합에 주안점을 두기 위하여 차원의 구조는 변화하지 않는다는 것을 전제로 한다. 그리고 차원의 구조는 OLAP의 분석기능을 극대화하기 위하여 컨트롤 가능한 범위 내에서 데이터의 중복을 허용함으로써 하나의 차원이 여러 개의 레벨을 갖도록 하였다.

전형적인 OLAP의 구조는 다계층(multi-tier)의 플랫폼을 가진다. 데이터를 레거시 데이터베이스로부터 추출하고 정리한 후 데이터 웨어하우스 모델과 부합하는 포맷으로 변환하여 데이터 웨어하우스에 로드시키는 것이 첫 번째 계층이다. 이러한 과정을 통하여 데이터 웨어하우스는 동질적인 과거의 데이터를 갖게 된다. OLAP의 두 번째 계층은 데이터 웨어하우스로부터 특정 주제를 주안점으로 하여 그와 관련된 데이터들이 추출된 데이터 마트이다. OLAP 서버는 세 번째 계층을 형성하는데 이 서버는 미리 계산되고 최적화된 하이퍼큐브의 사실값들을 담고 있다. 본 연구에서는 단계층 OLAP 구조가 사용되었다. 프로토타입 시스템의 데이터 웨어하우

스는 한 가지 주제를 중심으로 선택해야 할 만큼 크지 않기 때문에 본 연구에서 두 번째 계층인 데이터 마트는 생략되었다. RFID 기기를 통해 입력되는 실시간 기반의 정보 분석에 본 연구의 최우선적인 관심이 있기 때문에 세 번째 계층인 OLAP 서버도 구현되지 않았다.

3. 프로토타입 시스템

먼저, 프로토타입 시스템이 작동하는 RFID 환경은 어떻게 구성되어 있는지 보기로 하자. RFID 태그를 회원제로 운영되고 있는 놀이공원의 회원카드에 삽입하여 스마트카드의 형태로 만들어, 놀이공원 입구의 티켓판매소에 있는 RFID 리더기가 이 회원카드를 읽어 들인다 (그림 1, 좌측). 놀이기구의 담당자도 RFID 태그가 있는 스마트카드 형태의 신분증을 소지하고 있는데 RFID 리더기가 이 카드를 읽으면 특정 놀이기구의 담당자 정보



가 시스템에 등록된다 (그림 1, 우측).

그림 1. RFID 리더기가 고객 회원카드에 삽입되어 있는 RFID 태그를 읽는 모습 (좌측). PDA에 연결된 RFID 리더기가 놀이기구 담당직원의 RFID 태그를 읽고 있는 모습 (우측).

RFID 리더기가 카드를 읽는 순간 RFID 태그의 정보는 고객과 직원 데이터베이스가 있는 컴퓨터로 전송되고 시스템은 데이터베이스의 정보를 이용하여 고객이나 직원의 신분을 확

인하게 된다. RFID 태그의 정보를 이용하여 데이터베이스에서 추출한 고객의 정보가 그림 2의 우측상단과 같이 매표소의 컴퓨터 터미널에 나타난다. 회원으로 등록된 고객이 놀이기구의 티켓을 구매할 때 회원카드의 정보에 의하여 고객의 상세정보가 매표소의 컴퓨터 모니터에 나타나게 되고, 매표소의 직원은 단지 터미널에 보여지는 놀이기구 중 고객이 원하는 것을 클릭하여 선택하게 된다 (그림 2, 우측하단). 고객의 티켓구매에 관한 정보는 그 즉시 무선전송으로 컴퓨터의 데이터 웨어하우스에 저장된다.

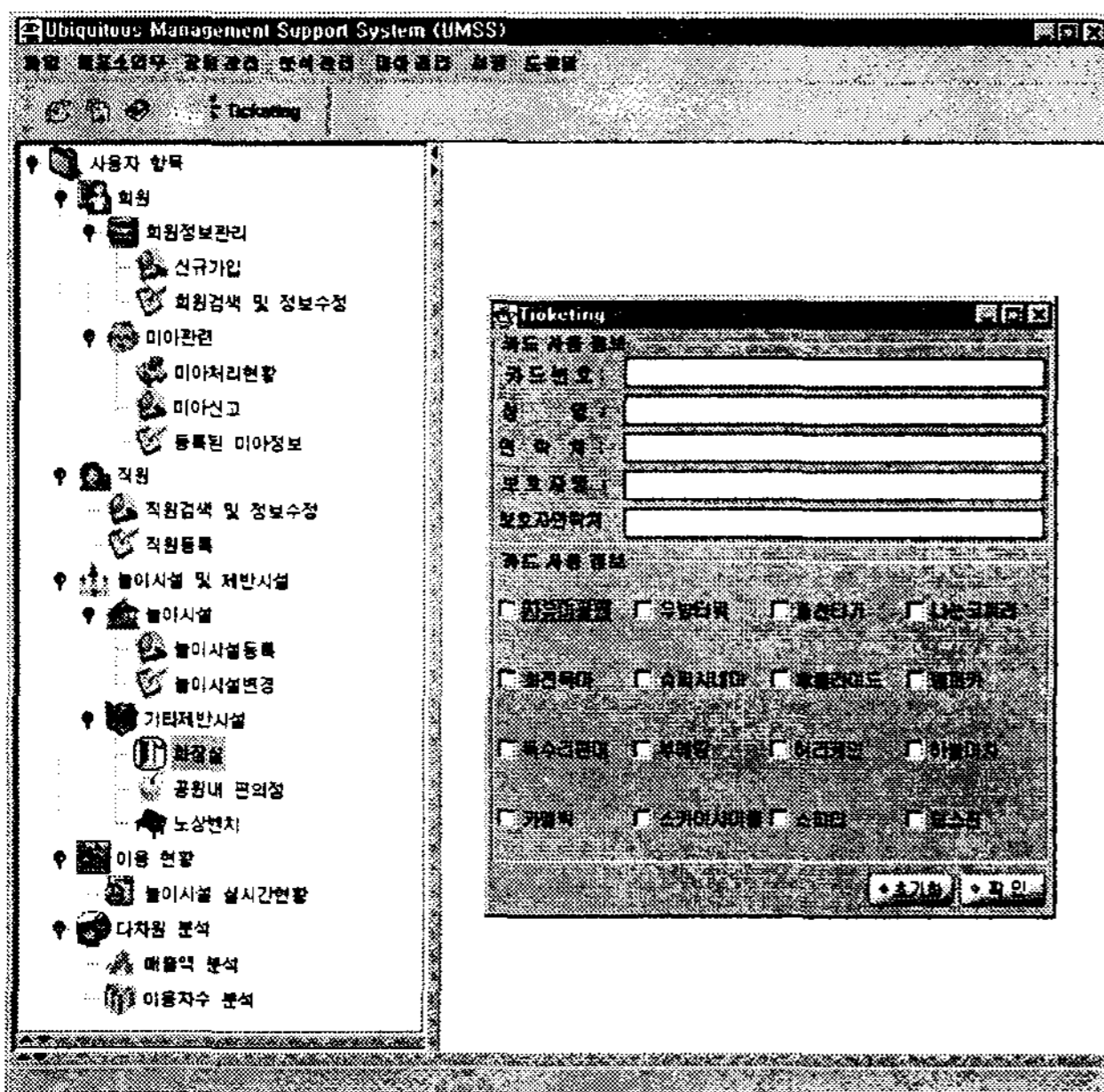


그림 2. 고객정보를 보여주는 화면(우측상단)과 놀이동산의 놀이기구 리스트를 보여주는 매표소의 컴퓨터 화면 (우측하단).

프로토타입 시스템은 실시간 분석 기능을 제공한다. RFID 리더기가 누가 어느 놀이기구를 이용하고 있는가 하는 정보를 실시간으로 획득하기 때문에 어느 놀이기구를 이용하려고 대기하고 있는 인원이 몇 명인가와 같은 정보가 수집되고 분석되어서 실시간으로

보여질 수 있다. 고객이 각 놀이기구의 입구에서 놀이기구 관리직원의 PDA에 연결되어 있는 RFID 리더기에 자신의 회원카드를 읽힐 때 해당 고객의 RFID 태그 정보가 데이터 웨어하우스에 전송된다. 모든 놀이기구로부터 획득된 이러한 정보는 실시간으로 제공이 된다. 그림 3은 놀이기구를 탑승하고 있는 사람의 수와 놀이기구를 타기 위해 대기하고 있는 사람의 수를 보여주는 매표소의 컴퓨터 화면이다. 모든 놀이기구의 담당직원들은 PDA를 이용하여 언제든지 이와 같은 실시간 정보를 확인할 수 있다.

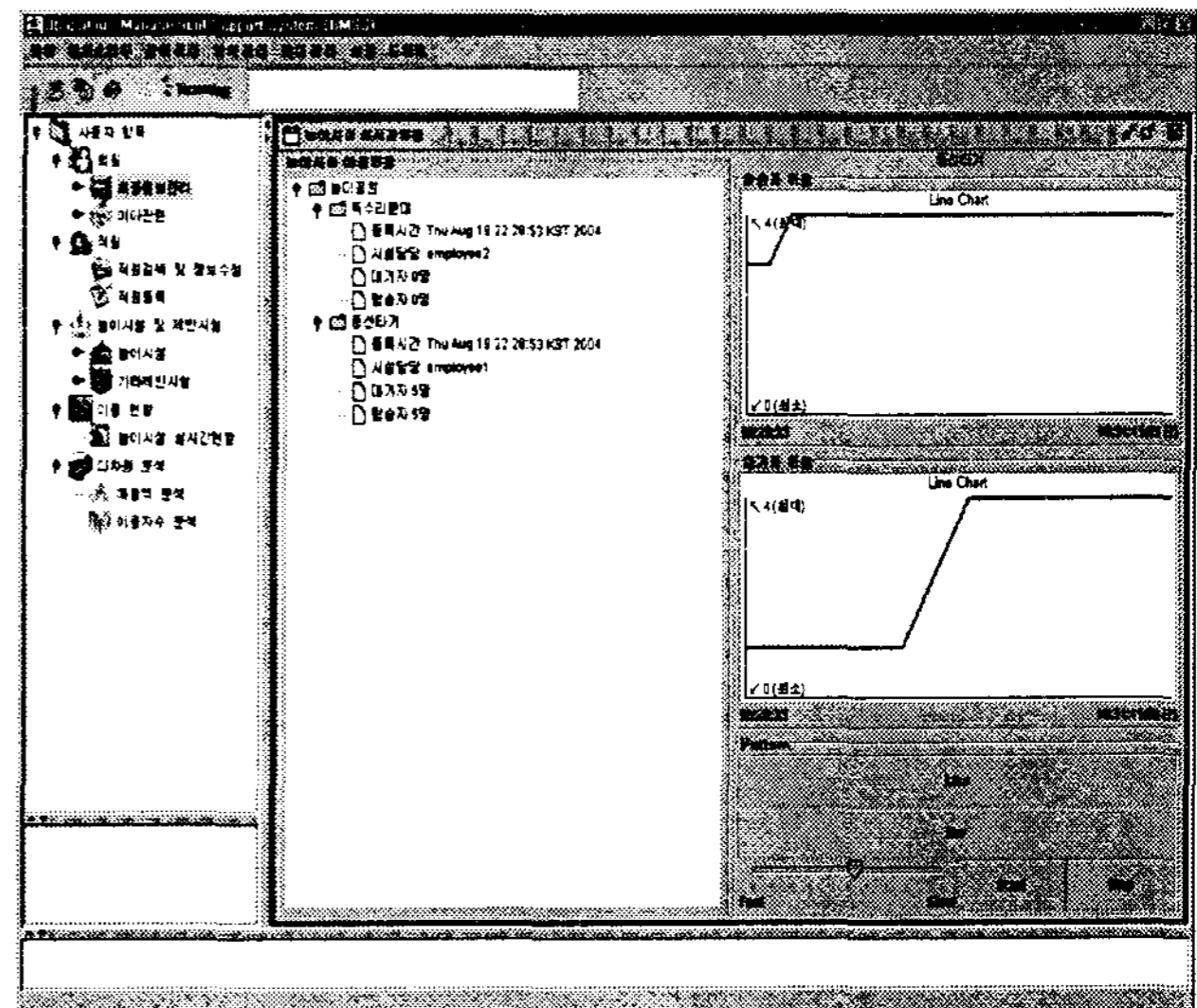


그림 3. 놀이기구 탑승인원의 수와 탑승대기인원의 수를 실시간으로 보여주는 화면

그림 4는 특정 놀이기구에 대기하는 사람이 아무도 없어서 이용 가능하다는 것을 보여주는 PDA 화면이다.

프로토타입 시스템의 다차원 분석기능을 살펴보면, 프로토타입 시스템의 fact는 판매 데이터이다. 다차원 데이터베이스의 디자인에는 지역, 연령, 성별, 놀이기구, 그리고 시간이 차원으로 사용되었고 따라서 fact granularity는 지역 대 연령 대 놀이기구 대 시간으로 표시된다. 각 차원에는

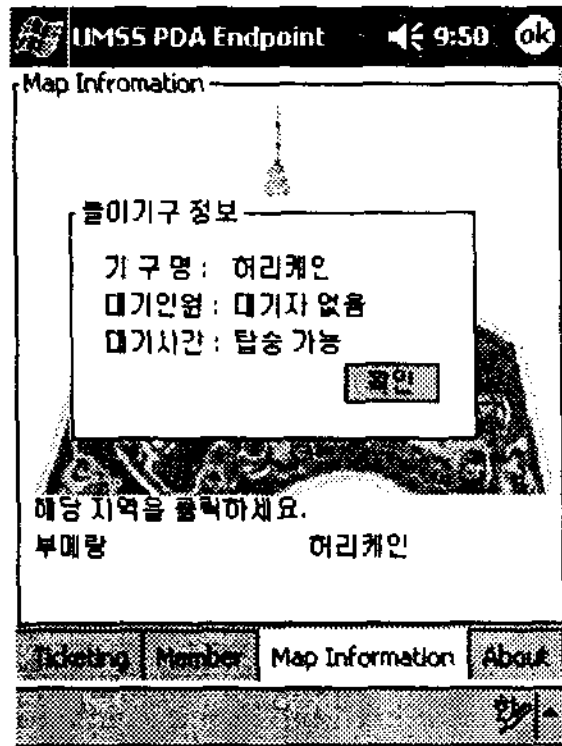


그림 4. 탑승대기자 수를 보여주는 PDA 화면

하나 이상의 level을 두었다. 예를 들어 주소 차원은 행정구역 구조를 따라 광역(도, 광역시), 중간(시, 구), 소 행정단위(동, 리)를 나타내는 세 개의 level을 두었고, 한 개의 차원에 속한 모든 level들은 하나의 관계형 테이블로 구현하였다. 광역 주소지가 한 단계 낮은 level을 따라 작은 지역으로 나누어 지는 이러한 구조를 사용하여 OLAP은 드릴다운(drill-down)이나 롤업(roll-up)과 같은 질의를 실행할 수 있게 된다. 그림 5에서는 OLAP을 이용하여 매출 데이터를 분석하는 화면을 보여준다.

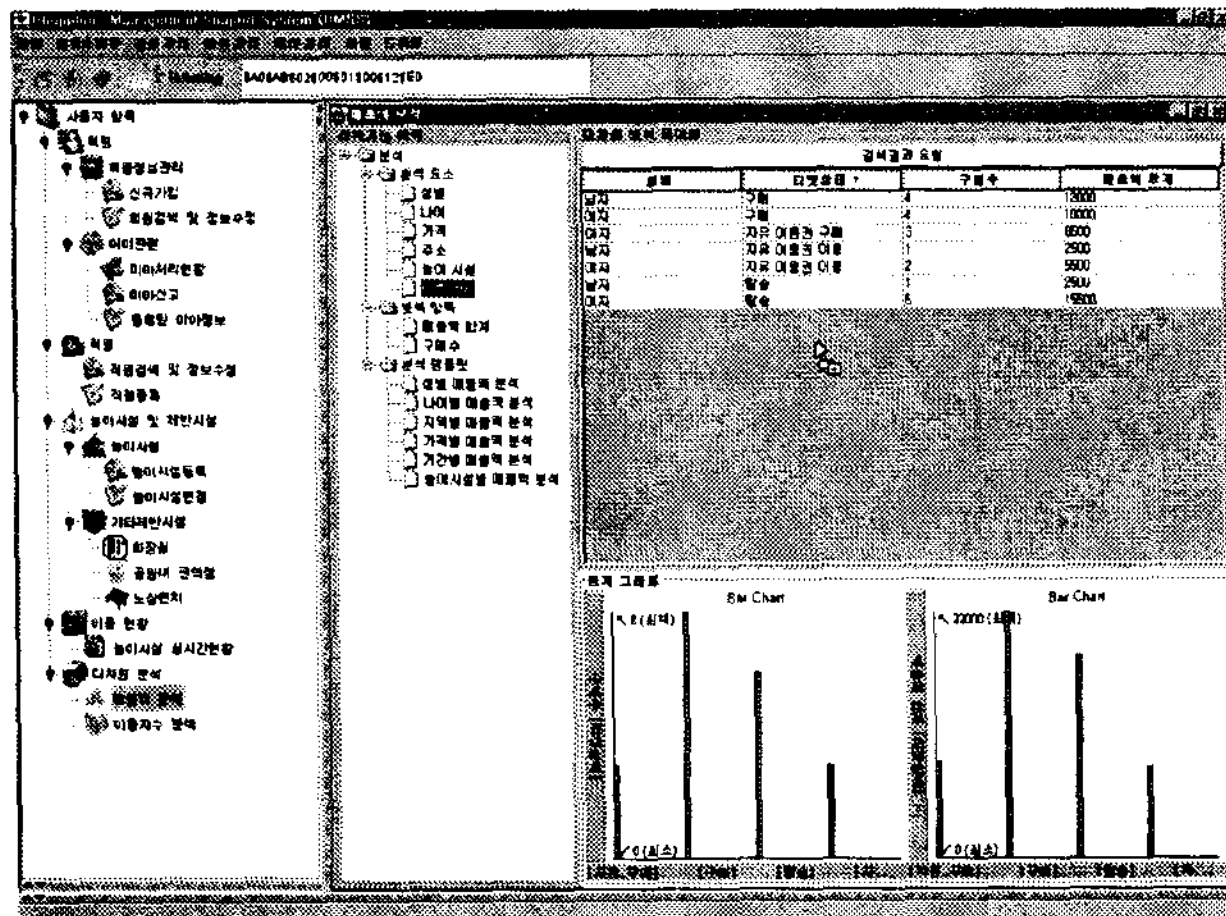


그림 5. OLAP 분석의 결과를 보여주는 화면

프로토타입 시스템은 동일한 데이터 웨어하우스 시스템에 추적기능과 분석기능이 동시에 구현되어 작동될 수 있는지를 살펴보기 위해 추적 기능도 구현하였다. 놀이공원에서 미아가 발생하였

을 때 추적기능을 이용하여 미아를 찾아주는 기능을 구현하였는데, 이 기능은 Swartz(2004)가 발표한 학교와 놀이공원에서 어린이들을 추적하는 시스템과 매우 흡사하기 때문에 본 논문에서는 자세한 설명을 생략한다.

4. 시스템 아키텍처

그림 6은 프로토타입 시스템의 아키텍처를 보여주고 있다. RFID가 데이터 웨어하우스와 통합될 수 있는지를 보다 잘 살펴볼 수 있도록 하기 위해 프로토타입 시스템의 아키텍처는 가능한 한 간단히 구현함으로써 그 초점을 RFID 리더기와 PDA, 데스크탑 컴퓨터 그리고 데이터 웨어하우스의 연결에 두고자 하였다. 프로토타입 시스템은 JAVA, Visual Basic과 Visual C 언어를 이용하여 구현되었다. 매표소의 판매관리 프로그램은 JAVA로 구현되었으며, PDA 어플리케이션은 Visual Basic으로 구현되었고 RFID 리더기는 Visual C를 사용하였다. 데이터 웨어하우스의 구축에는 MS SQL을 사용하였으며 JAVA를 이용하여 OLAP을 구현하였다. 프로토타입 시스템은 윈도우 XP 환경에서 구동되었으며 IIS와 HTTP가 데이터 웨어하우스, 판매관리 프로그램, RFID 리더기를 장착한 PDA 간의 커뮤니케이션을 위하여 사용되었다.

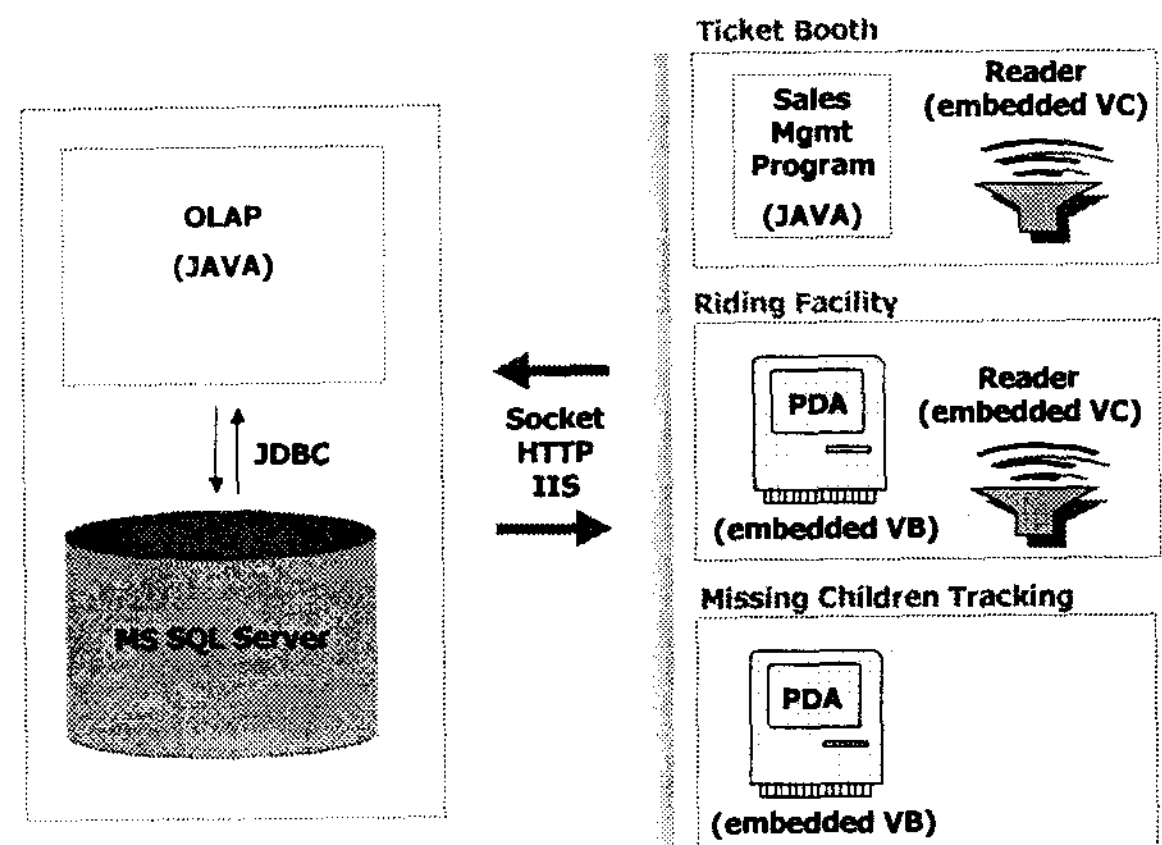


그림 6. 시스템 아키텍처

프로토타입 시스템은 국내의 한 놀이공원을 상정하여 구현되었는데, 주요 아이디어는 RFID 기술을 이용하여 놀이기구 탑승자나 대기자에 관한 실시간 정보를 제공함으로써 고객이 티켓을 구매할 때 대기자가 적은 놀이기구의 티켓을 구매하도록 의사결정을 돕는 것이다. 이러한 기능은 실험적 환경 하에서는 아무런 문제없이 구현되고 구동되었으나, 실제 현장에 구축하는 데는 다음과 같은 몇 가지 애로사항이 있었다. 첫째, Silwa [13]에서 언급된 것과 같이 라디오 주파수를 사용할 때 아직도 해결되어야 할 문제들이 있는 점이다. 둘째, RFID의 표준이 아직 확립되어 있지 않아서 다른 장비회사가 만든 장비간에는 호환성이 없다는 점이다 [14]. 따라서 어떤 회사의 장비를 선택하느냐 하는 문제는 매우 사용자에게 있어서 매우 결정하기 힘든 문제이다. 셋째, 대량으로 소요되는 RFID 리더기나 태그의 생산단가가 아직도 높아서 수요자들이 아직 RFID 시스템의 도입을 망설이고 있다는 점이다.

5. 결론

유비쿼터스 컴퓨팅의 핵심기술 중 하나인 RFID는 향후 몇 년 동안 비즈니스 영역에서 활발하게 사용될 것이다. RFID 테크놀로지가 동적이며 실시간 기반의 데이터를 획득하는데 적합한 기술이기 때문에 이러한 기술이 의사결정에 있어서 성공적으로 사용될 수 있으려면 동적이며 실시간 기반의 데이터를 분석할 수 있는 도구가 필요하다.

본 연구는 동적이며 실시간 기반의 데이터를 획득하기 위하여 고안된 RFID와, 과거의 정적인 데이터를 수집하고 저장하고 분석하는데 주로 사용되어 왔던 데이터 웨어하우스와 OLAP을 통합하는 프로토타입 시스템을 제안한다. 이 프로토타입 시스템은 RFID가 OLAP과 다차원 데이터베이스가 적절하게 디자인된다면 데이터 웨어하우스의 중요한 데이터 소스가 될 수 있다는 것을 보여주고 있다.

반면, 본 연구에서 제안된 프로토타입 시스템이 갖는 한계는 다음과 같다. 프로토타입 시스템의 다차원 데이터베이스와 OLAP은 실험적 연구 목적을 위하여 빈번히 발생하는 업데이트의 용이성과 시스템의 유연성을 위주로 디자인 되었다. 따라서 데이터베이스의 다차원 구조는 변하지 않는 것을 전제로 하였다. 따라서 차원이 변화하지 않는다는 것과 OLAP 구조가 하나의 계층으로 구성된다는 전제를 완화하여 보다 안정적인 유비쿼터스 데이터 웨어하우스를 디자인하는 것이 향후의 연구 과제가 될 수 있을 것이다.

프로토타입 시스템이 실험 환경 하에서는 아무런 문제없이 구동이 되었지만, 아직 라디오 주파수를 사용할 때 드러나는 문제점들을 해결하기 위한 기술이 개발되고 있고 RFID 관련 표준화 작업이 진행 중에 있으며 RFID 리더기와 태그의 생산단가가 아직 높다는 것을 고려하면 실제로 이러한 시스템이 현장에서 사용될 수 있기까지는 시간이 다소 필요할 것이다.

[참고문헌]

- [1] Angeles, R. "RFID Technologies: Supply-Chain Applications and Implementation Issues," *Information Systems Management* (22:1), 2005, pp. 51-65.
- [2] Srivastava, B. "Radio Frequency ID Technology: The Next Revolution in SCM," *Business Horizon* (47:6), 2004, p. 60-68.
- [3] Swartz, N. "Tagging Toothpaste and Toddlers," *Information Management Journal* (38:5), 2004, p. 22.
- [4] Wireless News. "Radianse Unveils Single-Use Active-RFID Tag for Hospitals," *Wireless News. Coventry*, 2005, Feb. 13, p. 1.
- [5] Fusaro, R. A. "None of Our Business?" *Harvard*

Business Review (82:12), 2004, p. 33-46.

[6] Lahart, J. "Ahead of the Tape," *Wall Street Journal*, 2005, Feb. 9, p. C1.

[7] Chaudhuri, S. and Dayal U. "An Overview of Data Warehousing and OLAP Technology," *SIGMOD Record* (26:1), 1997, pp. 23-29.

[8] Shim, J. P., Warkentin, M., Courtney, J. F., Power, D. J., Sharda, R. and Carlsson, C. "Past Present and Future of Decision Support Technology," *Decision Support Systems* (33), 2002, pp. 111-126.

[9] Jensen, M. R., Moller, T. H. and Pederson, T. B. "Specifying OLAP Cubes on XML Data," *Journal of Intelligent Information Systems* (17:2-3), 2001, pp. 255-280.

[10] Pedersen, T. B. and Jensen, C. S. "Multidimensional Database Technology," *IEEE Computer*, 2001, December, pp. 40-46.

[11] Kimball, R. *The Data Warehouse Toolkit: Practical Techniques for Building Dimensional Data Warehouses*, John Wiley & Sons, New York, 1996.

[12] Body, M., Miquel, M., Bedard, Y. and Tchounikine, A. "Handling Evolutions in Multidimensional Structures," *ICDE*, 2003, pp. 214-217.

[13] Sliwa, C. "Retailers Drag Feet on RFID Initiatives," *Computerworld* (39:4), 2005, pp. 1-2.

[14] Wailgum, T. "Tag, You're Late," *CIO*, Nov. 15, 2004.