의료용 데이터 웨어하우스를 위한 메타데이터 기반의 보안 프로토타입 시스템 구현 [1)]

김종호*·김태훈**·송해용*·홍수희*·박진두*·민성우***·이희석**

Implementing Metadata-based Security Prototype System for Medical Data Warehouse

Jongho Kim*, Taehun Kim**, Haeyong Song*, Suehee Hong*, Jindoo Park*,
Sungwoo Min***, Heeseok Lee**

요 약

본 연구는 통합병원정보시스템 (Integrated Hospital Information System) 에서 의료용 데이터 웨어하우스 (Medical Data Warehouse) 부분의 보안 프로토타입 시스템을 메타데이터 기반으로 설계하고 구현하는 데 주안점을 두었다. 특히, 의료용 데이터 웨어하우스 중에서도 임상 데이터 웨어하우스 (Clinical Data Warehouse) 에 초점을 두었으며 이에 대한 프로토타입은 ㅈ 병원에 적용되어서 개발되었다.

Key words: Data Warehouse, Metadata, Security, Prototype System

1. 서론

데이터 웨어하우스 (Data Warehouse) 는 상당한 양의 분류된 데이터를 가지고 있으며, 그 데이터는 가치있고 중요한 기업의 자원으로 여겨지고 있다. 의사 결정자 (Decision-maker) 는 의사 결정을 할 때, 이러한 데이터를 기초로 사용한다. 이러한 데이터는 믿을 수 있어야 하며, 위협받지 말아야 한다. 예를들면, 의료정보시스템은 환자 정보를 원하는 의사에게 즉각적으로 그 정보를 제공해야 하지만, 허가되지 않은 접근에 대해서는 그 정보를 제공하지 말아야한다. 의사결정 과정에서 정보시스템의 활용은 상당한 이점을 가지고 있지만, 그로 인한 정보시스템의 보안이 새로운 문제로 대두된다 [3].

데이터 웨어하우징 시스템 (Data Warehousing System) 을 위한 보안 요구사항은 다른 분산 컴퓨팅 시스템의 보안 요구사항과 크게 다르지 않으나, 대부분의 데이터 웨어하우스 개발은 개발과정 단계에서 보안에 대한 고려가 거의 이루어지지 않고 있다 [5, 6, 11]. 데이터 웨어하우스가 운영중일 때, 보안 요구사항을 첨가하는 것은 매우 힘든 일이기 때문에, 설계단계에서 보안 요구사항이 반영되어져야 한다 [5].

근본적으로 데이터 웨어하우스는 공개적이고 이용하기 쉬운 시스템이다. 데이터 웨어하우스의 목적은 일반적으로 사용자에게 하여금 많은 양의 데이터를 손쉽게 사용하여서 의사결정 과정에 도움이 되려고 하는 것이기 때문에, 어떤 보안상의 제한은 제약조건으로 대두되고 데이터 웨어하우스의 목적에 위배되는 것처럼 보여진다. 그러나 데이터

1) 본 연구는 산업자원부 산업기반기술개발사업과제 (관리번호: A00-981-3302-09-1-2) 로 수행되었음.
* ㈜ 비트컴퓨터
** 한국과학기술원 테크노경영대학원
*** 서울중앙병원

웨어하우스가 업무주도적인 (Mission Critical) 운영 데이터를 관리하는 것이 아니기 때문에 보안 위협의 성질은 데이터를 손상시키는 것이 아니라, 기업의 비밀과 전략을 누설하는 것이다 [4]. 그러나 데이터 웨어하우스가 데이터 탐색 등에 초점을 두고 있다면, 다른 운영 시스템들보다 접근 통제 권한을 관리하기가 쉽다 [8].

본 연구의 목적은 데이터 웨어하우징 시스템에서 필요한 보안 요구사항을 구조화하는 데 있다. 이러한 보안 데이터도 일종의 메타데이터이기 때문에 [5, 8], 통합병원정보시스템 (Integrated Hospital Information System) 에서 의료용 데이터 웨어하우스 (Medical Data Warehouse) 부분의 보안 프로토타입 시스템을 메타데이터 기반으로 설계하고 구현하는 데 주안점을 두었다. 특히, 의료용 데이터 웨어하우스 중에서도 임상 데이터 웨어하우스 (Clinical Data Warehouse) 에 초점을 두었으며 이에 대한 프로토타입은 ㅈ 병원에 적용되어서 개발되었다.

2. 데이터 웨어하우스 보안 모형

일반적으로 보안 모형은 보안 주체 (Security Subject), 보안 객체 (Security Object), 접근 유형 (Access Type) 및 허가 (Authorization) 의 네가지 요소 (Element) 들로 구성되어 있다 [2, 5, 6, 9, 10]; 보안 주체는 보안 객체에 접근하길 원하는 사용자 (Users), 프로그램, 프로세스 등을 나타내는 것으로 능동적 요소 (Active Element) 이다. 반대로, 보안 객체는 보호 대상인 수동적 요소 (Passive Element) 로써, 파일이나 테이블 등이다. 접근 유형은 보안 객체에 대해 행하여지는 접근 종류로써, 테이블에 대한 선택, 갱신, 삭제, 읽기, 쓰기 등을 나타낸다. 데이터 웨어하우스 환경에서는 갱신, 쓰기 같은 경우가 드물게 일어나기 때문에 대부분의 접근 유형은 읽기이다. 허가는 임의의 접근유형에 적용된 보안 객체에 대해서 보안 주체에게 허가된 권한을 의미하는 것으로, 임의의 사용자가 시스템 내의 임의의 테이블에 접근할 수 있도록 부여된 권한 등을 나타낸다.

데이터 웨어하우스와 같은 데이터베이스의 보안 모형은 임의적 보안모형 (Discretionary Security Model), 강제적 보안모형 (Mandatory Security Model), 수정된 강제적 보안모형 (Adapted Mandatory Security Model), 개인지식 접근모형 (Personal Knowledge Approach Model), 클락 & 윌슨 모형 (Clark and Wilson Model) 으로 구분된다 [6]; 임의적 보안모형은 보안 객체에 대한 접근을 보안 주체의 사용자 개인 또는 그룹의 식별자를 기반으로 제한하는 모형이다. 또한 어떤 종류의 접근 권한을 갖는 사용자는 다른 사용자에게 권한을 줄 수도 있다. 이때 사용되는 접근 규칙 (Access Rule) 은 <o, s, t, p> 로 표현되며, 함수 f 는 허가 f(o, s, t, p) 가 유효한지 안한지를 판단하는 것으로 정의되어진다 [6]:

$f: O \times S \times T \times P \rightarrow$ {True, False}

어떤 <o, s, t, p> 에 대해서, f(o, s, t, p) 값이 유효하면, 보안주체 s 는 술어 p 에 정의된 영역 내에서 보안 객체 o 에 대하여 허가 t 를 가진다.

이때, 술어 (Predicate) p 는 고려중인 보안 객체의 속성값이거나 비교값이다.

강제적 보안모형은 어떤 주체가 어떤 객체에 접근하려 할 때, 객체와 주체가 레이블 (Label) 정보에 기초한 보안 등급 (Security Level) 에 할당되어서, 높은 보안을 요하는 정보가 낮은 보안수준의 주체에게 노출되지 않도록 접근을 제한하는 접근통제 모형이다. 수정된 강제적 보안모형은 보안 주체의 사용자에게 역할 (Role) 을 부여해서 만든, 역할에 기반한 강제적 모형이다.

개인지식 모형은 데이터베이스나 정보시스템 내에 저장된 개인 정보에 대하여 접근을 제한함으로써 개인의 프라이버시를 보호하는 데 초점을 두고있다.

<표 1> 기존연구와 본 연구의 비교표

연구자 / 범주	Essmayr et al. [2]	Katic et al. [5]	Kirkgöze et al. [6]	본 연구
데이터베이스 보안모형	임의적 보안모형	임의적 보안모형	수정된 강제적 보안모형	임의적 보안모형
보안모형 메타데이터 구조	UML	파일 구조	파일 구조	ERD
보안모형 초점 및 보안주체	기업전역의 정보시스템 사용자	데이터 웨어하우징 시스템 사용자	데이터 웨어하우징 시스템 역할	데이터 웨어하우징 시스템 사용자와 그룹
시스템 구현	None	1. Security Manager 2. Security Query Management Layer	None	1. Metadata Component 2. Application Component

클락 & 윌슨 모형은 시스템의 사용자가 오직 어떤 트랜잭션만을 실행하도록 제한을 가하는 것이며, 각 트랜잭션은 할당된 데이터 객체에만 운영되어지는 것이다.

기존의 데이터 웨어하우스 보안 연구모형 [2, 5, 6] 은 임의적 모형, 수정된 강제적 모형에 초점을 두고 있다 (참조 <표 1>).

Essmayr et al. [2] 은 기업전역의 정보시스템의 보안 모형을 임의적 보안 모형에 기반을 두고 UML (Unified Modeling Language) 로 표현하였다. 보안 모형에 근거한 프로토타입 시스템은 향후연구로 제시하고 있다.

Katic et al. [5] 은 메타데이터를 구조 메타데이터 (Structural Metadata) 와 접근 메타데이터 (Access Metadata) 의 두 가지 관점으로 구분하고, 파일 형태의 구조를 제시하고 있으며, Security Manager 와 Security Query Management Layer 를 인터넷상에서 구현하였다. 보안주체를 사용자와 그룹 두개로 사용하나 사용자는 그룹에 무조건 할당되어야 하기 때문에 실제 주체는 사용자로 되어 있다. Kirkgöze et al. [6] 은 보안 객체를 역할만으로 표현하는, 수정된 강제모형에 기반한 파일구조를 제시하고 있다. 이 또한 프로토타입 시스템 구현을 향후연구 과제로 제시하고 있다.

본 연구는 사용자와 그룹의 보안 주체 하에서 임의적 보안모형에 기반하고 있다. 그러한 보안모형은 ERD (Entity Relationship Diagram) 으로 표현되었으며, 보안 프로토타입 시스템은 의료용 데이터 웨어하우징 시스템의 메타데이터 컴포넌트 (Metadata Component) 내의 일부와 어플리케이션 컴포넌트 (Application Component) 내의 일부로 개발되어져 있다.

3. 보안 프로토타입 시스템

의료용 데이터 웨어하우스 (Medical Data Warehouse) 는 사용자에 따라, 진료와 진료지원을 기반으로 하는 임상 데이터 웨어하우스 (Clinical Data Warehouse) 와 원무관리와 일반관리를 기반으로 하는 비임상 데이터 웨어하우스 (Non-clinical Data Warehouse) 의 두 부분으로 나누어진다. 비임상 데이터 웨어하우스는 경영층의 의사결정을 위한 비즈니스 데이터 웨어하우스 (Business Data Warehouse) 와 유사한 것이다 [1]; 임상 데이터 웨어하우스는 대규모의 환자에 대한 정보를 모아 놓

은 것으로 환자에 대한 의료 서비스와 의료 연구를 주목적으로 한다. 이러한 임상 데이터 웨어하우스와 비임상 데이터 웨어하우스는 ㅈ 병원을 대상으로 구현되었으며, 데이터 웨어하우징 시스템의 7 컴포넌트 [1] 중에서 운영 데이터 저장소 (Operational Data Store) 와 데이터 마트 (Data Mart) 컴포넌트는 구현되어 있지 않다. 기본적으로 적용된 데이터 웨어하우스 아키텍쳐는 임상 데이터 웨어하우스와 비임상 데이터 웨어하우스를 축으로 이루어져 있다.

구현된 시스템을 통해서, 사용되는 데이터 웨어하우스의 데이터가 사용자 또는 그룹에게 전부인 것처럼 보이게 하였다 (Transparency). 다른 사용자 또는 그룹은 데이터 웨어하우스의 다른 보안 객체를 볼 수 있는 것이다. 이렇게 함으로 인해서 사용자 또는 그룹은 보안상의 제한을 받고 있다는 생각을 안하게 되고 금지된 데이터를 찾으려는 시도는 하지 않을 것이다.

<그림 1>은 전체 메타데이터 스키마 중에서 보안 메타데이터 스키마 일부를 PLATINUM ERwin 3.5.2 로 나타낸 것이다.

이 스키마에서 보안주체는 사용자와 그룹을 이용하였다. 한 사용자는 여러 그룹에 속할 수 있으며 (예를 들면, 정형외과 레지던트는 정형외과 그룹에도 속할 수 있고, 레지던트 그룹에도 속할 수 있다), 어느 그룹에도 속할지 않을 수도 있다 (이 경우에는 Public Group에 할당된다). 보안 주체가 사용자로 로그온하면, 사용자로써 가지는 접근 유형 (MD_SUBJECTUSER) 와 그룹으로써 가지는 접근 유형 (MD_SUBJECTGRP) 둘 다를 허가/체크해서 어플리케이션 컴포넌트 화면에 보여준다. 그래서 한 화면에서 모든 업무를 수행하게끔 메타데이터가 도와주게 된다.

보안 객체 (MD_OBJECT) 에는 테이블, 분석오브젝트들, 차트, 케이스, 질의, 리포트 등이 속하며 <그림 1>에는 나타나 있지 않다. 객체에는 새로운 분석이라든지 형태들이 계속해서 포함될 수 있으며, 각 객체에서 파생되거나 생성되는 것들이 파일의 형태 (MD_OBJECTFILE) 로 저장되어진다.

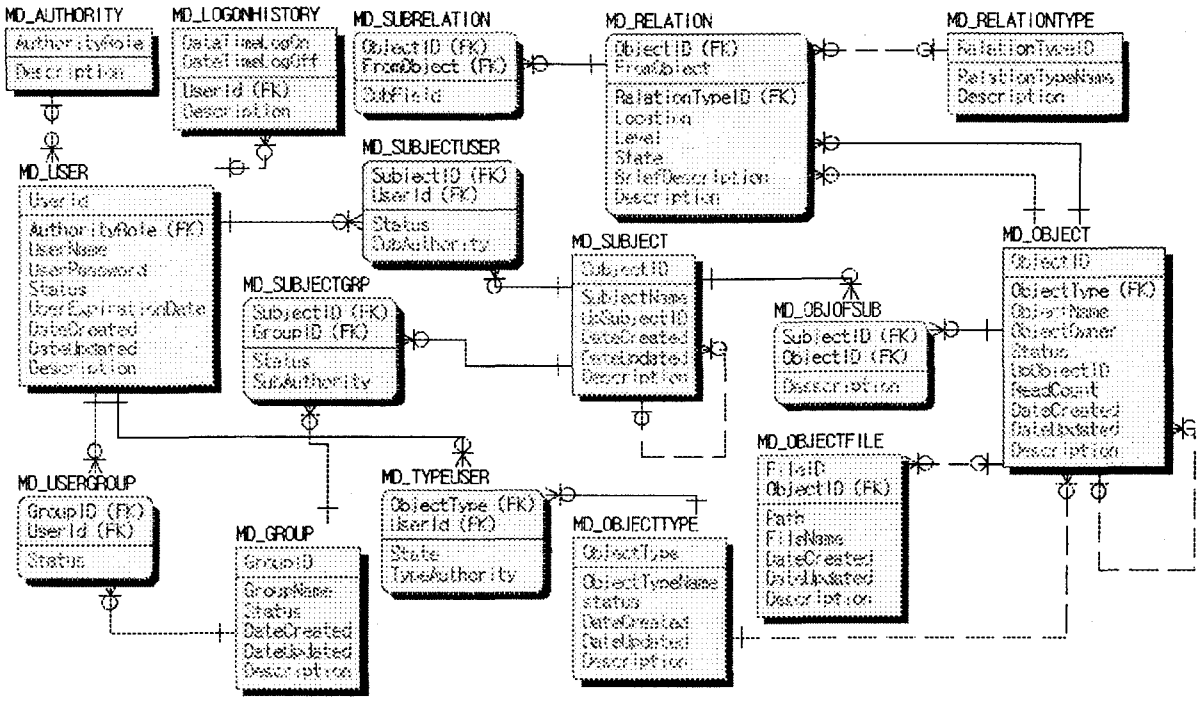
<그림 1> 에서, MD_SUBJECT 는 보안 주체가 가질 수 있는 보안 객체를 트리 구조로 나타내기 위한 것과 보안 주체 스스로 생성한 보안 객체를 등록시키는 것 등을 나타내는 것이다. 접근 유형은 사용자에 한정된 것이기 때문에 MD_TYPEUSER 로 나타내었다.

프로토타입 시스템은 MS Visual Basic 6.0 을 이용해서 개발되었으며, 메타데이터 데이터베이스는 의료용 데이터 웨어하우스와 마찬가지로 Sun Ultra Enterprise 450 하드웨어의 Sybase IQ 12 DBMS 안에 저장되어있다.

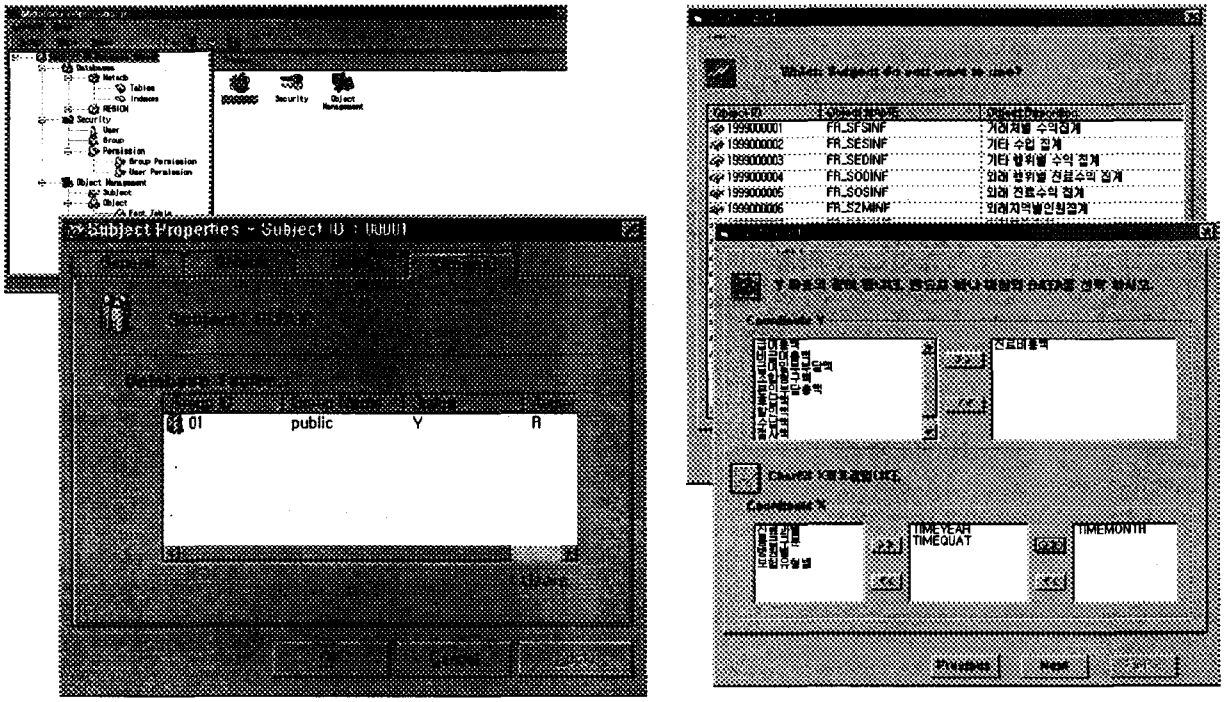
메타데이터 컴포넌트 (<그림 2>의 (a) 참조) 는 보안과 관련해서, 보안 주체들을 등록하고, 보안 주체를 보안 객체와 연결한다. 보안 주체에 알맞은 보안 객체를 선택함으로써, 선택된 정보를 메타데이터 데이터베이스에 저장하며, 이 정보를 어플리케이션 컴포넌트에 호출되어서, 어플리케이션 화면이 구성되어진다.

어플리케이션 컴포넌트 (<그림 2>의 (b) 참조) 에서는 메타데이터를 체크해서 화면에 보안 주체가 사용할 수 있는 사용 객체들을 보여주고, 질의, 리포트, 차트 등을 등록시키고 삭제할 수 있는 접근 유형 기능을 제공하다. 보안 주체가 사용할 수 있는 객체를 이용할 때, 질의는 미리 정의된 (Canned) 질의 또는 임시 (Ad hoc) 질의로 구분되는 데, 이 때 트리 기능이나 마법사 (Wizard) 기능 등을 통해서 질의가 어플리케이션 컴포넌트에 의해서 수행되며, 결과를 보안 주체가 메타데이터 안으로 저장하거나 삭제할 수 있다.

보안 주체는 임시 질의 등을 통해서 나온 결과를 다른 보안 주체들에게 공개 질의 (Public Query) 형태로 제공할 수도 있으며, 보안 문제 때문에 특정 보안 주체에게만 제공할 수도 있다. 그렇지 않을 경우에는 혼자만 쓸 수 있게 할 수도 있다.



<그림 1> 전체 메타데이터 스키마 중에서 보안 메타데이터 스키마 일부



(a) (b)

<그림 2> 메타데이터 컴포넌트 및 어플리케이션 컴포넌트 화면 일부

4. 결론

본 연구에서, 우리는 데이터 웨어하우징 시스템을 위한 보안 시스템을 메타데이터 접근으로 관리할 수 있는 방법을 제시하였다.

본 연구는 보안에 대한 이론적 기반과 정성적 차원에서 병원정보시스템의 근본적 이해가 부족한 점을 제외하고도 다음과 같은 한계를 가진다: 첫째, 데이터 웨어하우징 시스템에만 적용되었고, 둘째, 커스터마이징 (Customizing) 문제, 셋째, 성능의 저하에 또한 한계점이 대두되며, 넷째, 감사추적 (Audit Trail) 요구사항, 네트워크 요구사항 등과 같은 기타 보안 요구사항을 추가하지 못한 점이다. 본 연구의 보안 메타데이터 스키마는 유사한 데이터 웨어하우징 시스템 프로젝트에 적용가능하지만, 다른 영역의 고유한 요구사항을 반영해서 커스터마이징을 해야한다. 특히, 조직구성 [7] 은 그 영역에 맞게끔 재구성되어야 한다. 메타데이터를 사용함으로 인해, 어플리케이션 컴포넌트 등은 계속해서 메타데이터를 체크해야하기 때문에 성능의 저하를 야기시킬 수도 있다.

메타데이터는 기업 전역의 보안 정보를 한 곳에서 관리할 수 있는 기초를 제공할 수 있다 [2]. 향후 연구로써, 데이터 웨어하우징 시스템과 기존 시스템을 포함한, 이질적이고 분산된 정보시스템을 위한 통합 보안 메타데이터 리파지토리 (Repository) 설계에 초점을 둘 것이다.

참 고 문 헌

[1] 김종호, 김태훈, 민성우, 이희석, "의료용 데이터 웨어하우스 아키텍쳐 설계 및 구현," *'99 한국데이터베이스학회/한국지능정보시스템학회 춘계공동학술대회*, pp. 393-402, 1999.

[2] W. Essmayr, E. Kapsammer, and A. M. Tjoa, "Metadata for enterprise-wide security administration," *3rd IEEE Metadata Conference*, 1998.

[3] T. Finne, "What are the information security risks in decision support systems and data warehousing?," *Computer & Security*, vol. 16, no. 3, pp. 197-204, 1997.

[4] H. S. Gill and P. C. Rao, *Computing Guide to Data Warehousing*, Que Corporation, 1996.

[5] N. Katic, G. Quirchmayr, J. Schiefer, M. Stolba, and A. M. Tjoa, "A prototype model for data warehouse security based on metadata," *Proc. of IEEE 9th International Workshop on Database and Expert Systems Applications*, pp. 300-308, 1998.

[6] R. Kirkgöze, N. Katic, M. Stolba, and A. M. Tjoa, "A security concept for OLAP," *Proc. of IEEE 8th International Workshop on Database and Expert Systems Applications*, pp. 619-626, 1997.

[7] T. Moriarty, "Structuring organizations," *Intelligent Enterprise*, vol. 2, no. 7, pp. 60-63, 1999.

[8] G. M. Nussbaum, "The best little data warehouse," *Journal of Healthcare Information Management*, vol. 12, no. 4, pp. 79-93, 1998.

[9] T. C. Rindfleisch, "Privacy, information technology, and health care," *Communications of the ACM*, vol. 40, no. 8, pp. 93-100, 1997.

[10] E. Smith, J. H. P. Eloff, "Security in health-care information systems- current trends," *International Journal of Medical Informatics*, pp. 39-54, 1999.

[11] S. Warigon, "Data warehouse control and security," *Association of College and University Auditors LEDGER*, vol. 41, no. 2, pp.3-7, 1997.