PBDM+G : 그룹 개념을 이용한 목적 기반 데이터베이스 접근 제어 모델

PBDM＋G : Purpose-Based Database Access Control Model Using Group Concept

임지영(Ji-Young Lim)*, 김우철(Woo-Cheol Kim)**, 박상현(Sanghyun Park)**

요　약

온라인상에서 수집되어 사용되는 개인의 정보는 오용, 남용될 수 있다. 따라서 데이터 제공자가 허용하는 목적으로만 데이터를 사용할 수 있도록 하는 데이터 보안 기법이 필요하다. 기존의 LDHD 데이터베이스 보안 모델은 제공자의 프라이버시를 보호하기 위해 셀 단위로 제공 목적을 표시한다. 하지만 이 방식은 메타 데이터의 크기가 원본 데이터보다 커지며, 사용자-목적 별로 데이터 제공 여부를 조사하므로 새로운 데이터 제공자가 추가될 경우 변경되어야 하는 메타 데이터의 크기가 크다는 단점을 가지고 있다. 이러한 문제점을 해결하기 위해 본 논문에서는 먼저 제공자의 프라이버시를 보호하기 위한 데이터베이스 관리 시스템의 조건을 기술한 후, 이를 만족시키는 새로운 데이터베이스 보안 모델을 제안한다. 본 논문에서 제안하는 PBDM+G 모델은 중복된 데이터를 그룹으로 묶어 저장함으로써 메타 데이터의 크기를 크게 줄이며, 목적 단위로 데이터 제공 여부를 입력하게 함으로써 새로운 사용자가 추가되어도 데이터 제공 여부를 다시 입력해야 하는 번거로움을 줄이도록 한다. 실험을 통하여 PBDM+G 모델은 LDHD에 비해 약 2.6%에서 10%의 메타 데이터만을 사용하면서도 질의 처리 시간을 최대 23.6%까지 줄이는 것을 확인할 수 있었다.

주제어 : 데이터베이스 보안, 접근 제어, 프라이버시 보호

ABSTRACT

The personal information that is collected and used in on-line can be misused and abused. Therefore, data security techniques that restrict the usage of data only to the purpose of data provider are needed. The LDHD model, a well-known database security model, represents the purpose of data provision in the unit of "cell" in order to protect the privacy of data provider in detail. However, since the meta data is collected for every pair of users and purposes in this model, the size of the meta data is much larger than the original one and the introduction of a new user into the system causes the meta data to be changed significantly. To solve this problem, this paper first identifies the requirements of the database management systems supporting the privacy preservation and then suggests an effective and flexible database security model called PBDM＋G. The PBDM＋G model collects the meta data for every purpose rather than for every pair of users and purposes, and uses the concept of "grouping" to remove the duplicated meta data and thus reduce the size of meta data. The experimental result reveals that the PBDM＋G model consumes at most 10% of the space need for the LDHD model while reducing the query processing time up to 23.6%.

Key words : Database security, Access control, Privacy preservation

---

*　연세대학교 컴퓨터과학과, 교신저자
**　연세대학교 컴퓨터과학과
＋　논문접수 : 2006년 2월 26일, 심사완료 : 2006년 4월 13일

1. 서 론

인터넷 사용자가 급속히 증가하고 인터넷 뱅킹, 전자상거래를 비롯한 여러 가지 온라인 활동이 늘어남에 따라 많은 양의 개인 정보가 인터넷을 통해 전송 및 수집되고 있다. 이러한 개인 정보를 수집하는 기업은 웹사이트 이용자의 개인 정보를 파악하여 데이터베이스화하고 이를 분석함으로써 마케팅에 활용할 수 있다. 또한 이러한 개인 정보를 이용하여 웹사이트 콘텐츠를 발전시키고 제공하는 서비스나 상품의 수준을 높일 수 있다.

그러나 이러한 개인 정보의 수집 과정은 정보 제공자의 개별적 동의나 인지 없이 이루어지는 경우가 많다. 또한 합법적으로 수집된 정보의 경우에도 본래 사용 목적에서 벗어난 사용으로 개인의 프라이버시가 침해되기도 한다. 과도하게 수집된 개인 정보 때문에 생길 수 있는 프라이버시 침해에 대한 문제점과 프라이버시 보호의 필요성은 과거 여러 보고서 및 발표자료[1][2][3][4]를 통해서 주장되어 왔다. 실제로 2000년에 미국의 국가소비자연합(the National Consumer League)과 Harris International에 의해 수행된 조사[5]에 의하면, 전체 조사 응답자 중 56%가 자신의 프라이버시 보호에 대해 걱정하고 있는 것으로 나타났다. 또한 인터넷을 통해서 할 수 있는 일이 많아져 매우 편리하지만, 사용자들은 자신이 원치 않는 개인 정보 수집과 사용, 유통에 그들이 할 수 있는 것은 아무 것도 없다고 생각하는 것으로 밝혀졌다.

이러한 온라인상의 프라이버시 보호 문제 때문에 많은 기업들이 규제의 수단으로서 자사의 웹사이트에 프라이버시 보호 정책을 게재하고 있을 뿐 아니라 온라인 프라이버시 마크 제도를 도입하는 등 여러 가지 개인 정보 보호 방안을 강구하고 있다. 하지만 소비자들은 여전히 개인 정보의 사용과 처리, 유통에 대해 우려를 표명하고 있다. 이를 해결하기 위해 수많은 법ㆍ제도적 해결 방법이 제안되고 시행되었으나, 대부분의 방법은 이미 발생한 개인 정보의 오ㆍ남용 행위에 대한 사후 처리에 초점을 맞추고 있다. 그러나 프라이버시를 보다 근본적으로 보호하기 위해서는 개인의 데이터를 수집하는 단계에서 데이터의 사용 목적도 함께 수집하여 그 목적으로만 정보가 사용되도록 해야 한다.

기존의 데이터베이스 보안에서 주된 관심사는 이미 수집된 정보가 악의적인 목적을 가진 사용자에게 유출되지 않도록 하는 것이었다. 이를 위해 사용자들이 데이터에 접근하는 범위를 제한하는 방식을 사용하여 정보 유출을 방지하였다. 그 대표적인 방식들로 임의적 접근 제어(DAC : Discretionary Access Control), 강제적 접근 제어 (MAC : Mandatory Access Control), 역할 기반 접근 제어(RBAC : Role-Based Access Control)가 있다[6]. 이와 같은 기존의 데이터베이스 접근 제어 모델들은 이미 수집된 데이터의 보호에 초점을 맞추고 있기 때문에 데이터 제공자의 제공 목적에서 벗어난 데이터 사용을 차단할 수 없다는 문제점을 가지고 있다.

최근 히포크라테스 데이터베이스[7]에서는 프라이버시를 보호할 수 있는 새로운 데이터베이스 모델의 개념을 제시하였다. 기존의 데이터베이스 보안 모델과는 달리 데이터 수집

단계에서 프라이버시 보호에 필요한 메타데이터를 함께 수집하여 저장함으로써 프라이버시를 보호한다. 하지만 이 방법은 레코드 단위로 데이터를 제어하기 때문에 데이터 수집에 제한이 생길 수 있고 보안 정책이 변경되었을 때 정보 유출이 발생할 수 있다는 단점을 가지고 있다. 이러한 단점을 보완하기 위해서 LeFevre et al.는 [8] (이하 LDHD: Limiting Disclosure in Hippocratic Databases)에서 셀 단위로 데이터를 보호하는 방식을 제시하였다. 이 방식은 셀 단위로 메타데이터를 저장해야 하기 때문에 메타데이터의 크기가 매우 커지게 되어 대용량의 데이터베이스에는 적합하지 않은 단점을 가지고 있다. 또한 새로운 사용자가 시스템에 추가되는 경우에 기존에 수집된 모든 개인 정보에 대해서 새롭게 제공 여부를 확인해야 하는 번거로움이 있다.

본 논문에서는 LDHD와 동일하게 셀 단위의 데이터베이스 보안 레벨을 지원하면서도 LDHD의 문제점을 해결할 수 있는 새로운 메타데이터 관리 기법에 대해 논의한다. 첫 번째로, 사용자별로 데이터의 제공 여부를 수집하고 저장하는 LDHD 방식을 따르지 않고 대신 목적별로 데이터 제공 여부를 조사하는 방식을 제안한다. 이 방식에서는 새로운 목적이 시스템에 추가되는 경우에는 기존에 수집된 모든 개인 정보에 대해 제공 여부를 다시 확인해야 하지만, 새로운 사용자가 추가되는 경우에는 개인 정보의 제공 여부를 재확인할 필요가 없다. 일반적으로 새로운 목적이 추가되는 경우보다는 새로운 사용자가 추가되는 경우가 훨씬 더 많으므로, 제안된 방식을 사용함으로써 데이터 제공 여부를 다시 조사해야 하는 번거로움을 크게 줄일 수 있다. 두 번째로, 메타데이터의 크기를 줄이기 위해서 그룹의 개념을 도입하여 활용한다. 즉, 주어진 목적 집합에 대해 동일한 제공 여부 패턴을 가지는 제공자들을 하나의 그룹으로 묶은 후 각 데이터 제공자가 어떤 그룹에 속하는가를 저장하는 방식을 사용함으로써 메타데이터의 크기를 대폭 줄이도록 한다. Wisconsin 벤치마크 데이터 집합을 사용한 실험 결과를 통하여 본 논문에서 제안하는 방식이 LDHD에 비해 메타데이터의 크기와 질의 처리 시간 측면에서 각각 90%와 76.4%가 감소함을 확인할 수 있었다.

본 논문의 구성은 다음과 같다. 먼저 프라이버시를 위해 데이터베이스 관리 시스템이 갖추어야 할 조건을 2장에서 정의하고 그것과 관련된 기존의 데이터베이스 접근 제어 모델에 대해서 3장에서 기술한다. 4장에서는 프라이버시 보호를 위한 히포크라테스 데이터베이스와 LDHD의 단점을 지적하고 이러한 단점을 해결하기 위한 새로운 데이터베이스 모델을 5장에서 제시한다. 6장에서는 새로운 데이터베이스 모델을 실험을 통해서 기존의 모델들과 비교하고 7장에서 결론을 맺는다.

2. 프라이버시 보호를 위한 데이터베이스 관리 시스템의 조건

기존의 데이터베이스 보안 모델은 이미 수집된 정보의 보안에 초점을 두었다. 즉 저장된 정보를 데이터베이스 관리자가 임의로 구

분한 후 각 사용자별로 접근할 수 있는 권한을 부여하는 방식으로 사용자들의 정보 접근을 제한하였다. 하지만 프라이버시 보호를 위해서는 개인 정보 제공자가 자신의 정보의 접근 권한을 스스로 제어할 수 있어야 한다. 그렇게 하기 위해서는 데이터를 수집하는 단계에서부터 프라이버시 보호를 위한 설계가 필요하다. 즉, 프라이버시 보호를 위한 데이터베이스 관리 시스템은 기존 데이터베이스 보안 모델들의 방식과는 달리 다음과 같은 새로운 조건들을 고려해서 설계되어야 한다. 이 조건들을 데이터 흐름의 관점에 따라서 나누어 보면 다음과 같다.

2.1. 데이터 수집

일반적으로 데이터베이스의 데이터 처리 단계는 1) 데이터 수집, 2) 데이터 저장 및 관리, 3) 질의 처리의 3가지 과정으로 나누어진다. 이 중에서 데이터 흐름상 가장 먼저 고려할 단계는 데이터를 수집하는 단계이다. 기존의 데이터베이스 보안 모델에서는 깊게 고려하지 않아도 되는 단계이지만 프라이버시를 고려한 데이터베이스 보안 모델에서는 데이터 수집시 프라이버시 보호를 위한 사용 목적을 함께 수집해야 하므로 가장 중요한 단계라고 할 수 있다.

이 단계에서 가장 중요한 조건은 앞서 말한 바와 같이 데이터가 수집될 때 데이터의 사용 목적도 함께 수집되어야 한다는 것이다(조건 1). 이 조건은 프라이버시 보호를 위해서 꼭 지켜져야 한다. 또한 데이터 제공자가 사용 목적과 함께 제공하는 데이터는 데이터베이스에 모두 수집되어야 한다(조건 2). 그리고 좀 더 높은 수준의 프라이버시 보호를 위해서는 테이블 단위보다는 레코드 단위, 레코드 단위보다는 셀 단위와 같이 가능한 작은 정보 보호 단위가 제공되어야 한다(조건3). 마지막으로 데이터 사용 목적의 수집이 간편해야 한다(조건4). 예를 들어 사용 목적 수집 시 데이터 제공자가 개별적으로 사용 목적을 부여하는 방식보다는 시스템이 미리 정해둔 몇 가지 사용 목적 패턴 중 하나를 선택할 수 있다면 사용 목적 수집이 더 간편할 것이다.

2.2 데이터 저장 및 관리

현재 상용화된 대부분의 데이터베이스 관리 시스템은 여러 방식으로 수집된 데이터를 물리적 공간에 효율적으로 저장하고 관리하고 있다. 하지만 프라이버시 보호를 위한 데이터베이스의 경우 모든 데이터마다 사용 목적을 함께 저장해야 하기 때문에 대용량의 메타데이터를 유지해야 하는 특징을 가지고 있다. 따라서 기존의 데이터베이스 관리 시스템의 메타데이터 저장 및 관리 방식으로는 저장 공간의 제한을 받는다. 그러므로 대용량의 메타데이터의 특성을 잘 분석하여 적은 양의 저장 공간만을 사용하더라도 프라이버시를 보호할 수 있는 메타데이터 저장 및 관리 시스템이 제공되어야 한다(조건 5). 그 이외에 새로운 사용자가 추가되는 등 데이터베이스 사용 환경이 변화되었을 때에도 수정되어야 하는 메타데이터의 크기가 가급적 작아야 한다(조건 6).

<표 1> 프라이버시 보호를 위한 데이터베이스 관리 시스템의 조건

번호	조 건
1	데이터가 수집될 때 데이터 사용 목적도 같이 수집되어야 한다.
2	사용 목적과 함께 제공되는 데이터는 모두 수집되어야 한다.
3	셀 단위와 같이 가능한 작은 정보 보호 단위가 제공되어야 한다.
4	데이터 사용 목적의 수집이 간편해야 한다.
5	적은 양의 저장 공간만을 사용하더라도 프라이버시를 보호 할 수 있는 메타데이터 저장 및 관리 시스템이 제공되어야 한다.
6	데이터베이스 사용 환경이 변화되었을 때에도 수정되어야 하는 메타데이터의 크기가 가급적 작아야 한다.
7	메타데이터를 이용할 수 있도록 기존의 SQL문을 자동적으로 변형해 주는 질의 처리 방법이 제공되어야 한다.
8	2개 이상의 사용 목적을 이용하는 질의가 가능해야 한다.

2.3 질의 처리

질의 처리 과정은 수집된 데이터 중에서 사용자가 원하는 데이터를 제공하는 과정이다. 기존의 데이터베이스 관리 시스템에서도 SQL문을 이용하여 이 기능이 지원되고 있다. 하지만 프라이버시 보호를 위해 메타데이터를 이용해야 하므로 새로운 SQL문의 형태를 지원하거나 기존의 SQL문을 자동적으로 변형해 주는 질의 처리 방법이 제공되어야 한다(조건 7). 또한 데이터베이스의 사용자는 여러 가지 목적을 복합적으로 이용해 질의를 수행할 수도 있으므로, 이를 지원하기 위해서 2개 이상의 목적이 질의에 사용 가능해야 한다(조건 8).

위의 내용을 종합해서 표로 나타내면 다음과 같다.

3. 전통적인 데이터베이스 보안 모델

보안을 필요로 하는 데이터에 대한 접근 제어 정책은 권한의 소유자에 의해 객체에 대한 접근 권한이 정해지는 임의적 접근 제어(DAC : Discretionary Access Control), 정보의 보안 수준과 사용자의 보안 등급에 따라 접근을 제어하는 강제적 접근 제어 (MAC : Mandatory Access Control), 그리고 역할과 해당 역할의 권한을 정의하는 역할 기반 접근 제어(RBAC : Role-Based Access Control)로 나눌 수 있다.

이러한 정책을 실현하기 위한 접근 제어 메커니즘으로는 권한 주체와 권한 객체 사이의 권한 관계를 매트릭스로 표현하는 접근 제어 매트릭스(Access Control Matrix)[9], 권한주체에 대해 권한이 부여된 객체들을 리스트로 표현하는 접근 제어 리스트(Access Control

List)[10], 그리고 권한 객체에 대해 접근이 허용된 주체들을 리스트로 표현하는 자격 리스트(Capability List)[11]등이 있다.

3.1 임의적 접근 제어

임의적 접근 제어는 시스템이 인증하는 식별자에 근거하여 사용자나 사용자 그룹의 객체에 대한 접근을 제한하는 방법이다[12]. 이 정책은 사용자가 특정 객체에 대한 접근 권리를 갖는 한 자율적으로 다른 사용자에게 자신의 권리를 넘겨줄 수 있기 때문에 자율적 정책이다. 대표적인 모델은 선취-허가 모델(Take-Grant Model)[13]과 우드 모델 (Wood Model) [14]이 있다.

3.2 강제적 접근 제어

강제적 접근 제어는 객체의 보안 등급과 사용자의 인가 등급이 일치하는 경우에만 데이터에 대한 사용자의 접근을 허용하는 방법이다[12]. 이에 해당하는 대표적인 보안 모델은 벨-라파듈라 모델(Bell-Lapadula Model)[15]과 비바 모델(Biba Model)[16]이 있다.

3.3 역할 기반 접근 제어

역할 기반 접근 제어는 사용자가 가진 역할에 기반을 두고 사용자의 데이터에 대한 접근을 제어하는 기법이다. 이는 사용자가 특정 데이터에 임의로 접근할 수 없도록 하며, 접근 권한을 역할(Role)에 부여하고 사용자는 적절한 역할에 소속됨으로써 역할의 수행에 필요한 최소한의 데이터에만 접근할 수 있도록 한다. 사용자는 그들의 업무적 권한과 책임에 따라 특정 역할의 구성원이 되며 접근 구조의 변경 없이도 쉽게 역할을 변경할 수 있다.

4. 프라이버시를 고려한 기존의 보안 모델

본 장에서는 프라이버시를 고려한 기존의 보안 모델에 대해서 설명한다. 간편한 설명을 위해서 데이터베이스 환경을 다음과 같이 제한한다. 데이터베이스에는 하나의 데이터 테이블만이 존재한다. 또한 테이블을 구성하는 각 레코드는 오직 한 데이터 제공자로부터 제공된 데이터만을 기록한다. 이와 같은 제한적 환경은 실제 데이터베이스 환경으로 쉽게 확장이 가능하다.

4.1 히포크라테스 데이터베이스

히포크라테스 데이터베이스는 기존의 관계형 데이터베이스 관리 시스템 상에서 프라이버시-정책과 프라이버시-권한 등과 같은 메타데이터를 이용하여 프라이버시 보호 기능을 제공하는 데이터베이스이다. 프라이버시-정책은 데이터 수집을 위한 데이터베이스 관리자의 정책을 나타내는 데이터로서 데이터 테이블의 각 속성별로 사용 목적을 기술한 데이터이다. 프라이버시-권한은 데이터의 접근을 제어하기 위해 사용하는 접근 제어 정책으로써 데이터베이스 관리자가 데이터 접근자

에게 부여하는 권한을 기술한 데이터이다. 이는 데이터 테이블의 각 속성별로 허용된 데이터 접근자와 사용 목적을 명시함으로써 데이터 접근자가 허용된 목적에 따라서만 데이터 테이블의 속성을 접근하도록 한다.

히포크라테스 데이터베이스에서는 데이터를 수집할 때 데이터 제공자의 데이터 제공 여부를 나타내는 프라이버시-선호와 데이터베이스 시스템의 프라이버시-정책이 모두 일치하는 경우에만 데이터를 수집한다. 따라서 데이터 수집에 한계가 있다. 이로 인해 2장에서 기술한 조건 2를 만족시키지 못한다. 또한 기본적으로 레코드 단위의 접근 제어 방식이기 때문에 셀 단위 접근 제어보다 세밀하지 않다는 단점을 지니고 있다. 이는 2장에서 기술한 조건 3에 부합되지 못한다.

4.2 LDHD

히포크라테스 데이터베이스의 단점을 보완하기 위해서 LeFev et al.은 셀 단위의 접근 제어 방식을 제공하는 새로운 형태의 데이터베이스 보안 모델인 LDHD(Limiting Disclosure in Hippocratic Databases)를 제안하였다. LDHD의 데이터 수집 과정을 살펴보면 다음과 같다. 먼저 데이터 제공자는 새로운 데이터를 제공할 때 데이터베이스 관리 시스템이 허용하는 모든 (사용자—목적)에 따라 제공자의 선호(Preference)를 기입한다. 즉 데이터베이스 관리 시스템이 허용하는 사용자가 A와 B이고 각각 P_1과 P_2의 접근 목적을 가지고 있다면 (A-P_1)과 (B-P_2)에 따라서 데이터 제공 여부를 기록해야 한다. 이렇게 제공된 선호는 데이터를 제공하는 경우에는 “1"로 표시하고 그렇지 않은 경우에는 “0"으로 표시해 별도의 메타데이터 테이블에 저장한다. 이때 (사용자—목적)에 따른 선호 테이블을 선호(사용자—목적) 테이블(PreferenceTableBy(Recipient—Purpose))이라고 한다.

프라이버시를 보호하기 위해서는 사용자 질의가 선호(사용자—목적) 테이블을 이용하도록 수정해야 한다. 이를 위해 LDHD에서는 선호(사용자—목적) 테이블에서 데이터의 제공을 허용하면 사용자에게 실제 데이터를 제공하고 그렇지 않으면 “NULL"을 반환하는 case문을 이용한 질의 수정 알고리즘을 사용한다.

하지만 LDHD 방식을 실제 데이터베이스 환경에 적용할 경우 몇 가지 문제점이 발생한다. 먼저 데이터 제공시 데이터베이스를 이용하는 사용자 수만큼 제공자의 선호를 기입해야 한다는 단점을 지니고 있다. 예를 들어 어떤 제공자가 데이터베이스에 자신의 ‘이름’ 데이터를 제공한다고 하자. 제공자가 이 데이터의 사용을 허용하는 선호도를 표시하기 위해서는 사용자 A가 P1이라는 목적으로 이용함을“허용"하고, 사용자 B가 P2라는 목적으로 이용함을 “불허용" 한다는 제공자 자신의 선호를 사용자별로 각각 기입해야 한다. 이러한 문제점은 2장에서 기술한 조건 4를 충족시키지 못한다.

두 번째 단점은 저장하는 메타데이터의 크기가 크다는 것이다. 첫 번째 단점에서 지적했듯이 LDHD 방식에서는 데이터 테이블의 한 셀이 여러 개의 선호를 가질 수 있다. 이

<표 2> 프라이버시 보호를 위한 데이터베이스 관리 시스템의 조건과 비교표

번호	히포그라테스 데이터베이스	LDHD
1	사용 목적 수집	사용 목적 수집
2	일부 미수집	모두 수집
3	레코드 단위	셀 단위
4	간편	불편
5	작은 메타데이터	많은 메타데이터
6	적음	많음
7	없음	있음
8	없음	없음

경우 메타데이터의 선호(사용자—목적) 테이블은 여러 개의 선호를 모두 각각의 테이블에 저장하게 된다. 따라서 메타데이터의 크기는 원본 데이터의 크기보다 커지게 된다. 이러한 문제점은 2장에서 기술한 조건 5를 충족시키지 못한다.

세 번째 단점은 새로운 데이터 사용자가 추가되는 경우 변경되어야 하는 메타데이터의 크기가 크다는 것이다. 예를 들어 어떤 데이터베이스 시스템에 P_1이라는 목적을 가진 사용자 A와 P_2라는 목적을 가진 사용자 B가 있다고 하자. 이때 P_2의 목적을 가진 새로운 사용자 C가 생긴다면 LDHD에서는 선호(P_2, C) 테이블을 채우기 위해서 기존의 데이터 제공자들로부터 새롭게 선호를 조사하여야 한다. 이러한 문제점은 2장에서 기술한 조건 6을 잘 충족시키지 못한다.

2장에서 기술한 프라이버시 보호를 위한 데이터베이스 관리 시스템의 조건에 맞추어 4.1절의 히포크라테스 데이터베이스와 4.2절의 LDHD를 비교해보면 <표 2>와 같은 결과를 얻을 수 있다.

5. 프라이버시 선호 그룹을 이용한 목적 기반의 데이터베이스 모델

5.1 목적 단위로 프라이버시 선호를 수집하는 데이터베이스 모델 (PBDM)

앞서 살펴본 것처럼 LDHD의 경우 (사용자-목적)별로 셀의 제공 여부를 수집하고 저장한다. 따라서 데이터 제공자는 프라이버시 보호를 위해 각 사용자마다 사용가능한 데이터를 모두 표시해야 하므로 사용자 수가 많은 경우 많은 입력이 필요하다. 따라서 메타데이터의 크기가 커질 수밖에 없다. 하지만 일반적인 데이터 제공자의 관점에서 보자면 자신의 데이터를 사용하는 사용자의 신원보다는 데이터의 사용 목적이 중요한 관심사 일 것이다. 따라서 본 논문에서는 데이터 제공자의 데이터 공개 여부를 조사할 때 (사용자—목적) 단위가 아닌 (목적)단위로 데이터 공개 여부를 조사하여 불필요한 데이터 중복 입력을 데

MetadataTable

SID	A_1C	A_2C	A_3C
1	O	X	O
2	X	O	O
3	X	O	O
4	O	X	O

DataTable

SID	A_1	A_2	A_3
1	AA	11	!
2	BB	22	@
3	CC	33	#
4	DD	44	$

(a) 그룹 개념 적용전

MetadataTable

SID	GID
1	1
2	2
3	2
4	1

GID	A_1C	A_2C	A_3C
1	O	X	O
2	X	O	O

(b) 그룹 개념 적용후

<그림 1> 그룹 개념 적용 전과 후의 메타데이터

이터 수집시 방지한다. 이렇게 (목적)단위로 데이터 제공자의 선호를 조사함으로써 데이터 제공자의 입력의 수가 줄어들고 프라이버시 보호를 위한 메타데이터의 양이 줄어든다.

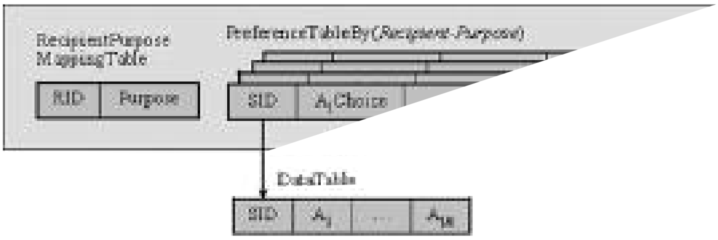
5.2 프라이버시 선호 그룹을 이용한 데이터베이스 모델 (PBDM+G)

PBDM은 목적 단위로 프라이버시 선호를 조사함으로써 LDHD보다 메타데이터의 크기가 줄어든다. 하지만 각 데이터별로 제공목적을 각각 기입할 경우 메타데이터의 크기는 원본 데이터보다 커지게 된다. 이 경우 메타데이터에는 많은 중복된 데이터들이 존재하게 된다. 예를 들어 3개의 속성으로 구성된 데이터 테이블이 있을 때 테이블의 레코드를 구성하는 데이터 제공자는 제공 목적에 따라서 각 속성의 접근 허용 여부를 결정한다. 이 경우 데이터 제공자의 선호의 종류는 최대 8(= 23)가지의 경우의 수가 생기게 된다. 이 때 데이터 제공자의 수가 8명 이상이 되면 선호를 조사한 선호(P1)테이블에서는 중복된 레코드들이 존재하게 된다. 이 경우 중복된 레코드를 제거함으로써 메타데이터의 크기를 줄일 수 있다.

이렇게 중복된 레코드들을 효율적으로 관리하기 위해서 메타데이터를 구성하는 테이블들에 대한 정규화 (Normalization)과정이 필요하다. 이를 위해 공통의 선호를 갖는 데이터 제공자(레코드)를 하나의 그룹으로 묶어서 프라이버시 선호 그룹을 생성한다. 즉 모든 레코드의 프라이버시 선호를 각각 저장하지 않고 프라이버시 선호 그룹의 선호만을 저장한 후 각 레코드가 어떤 프라이버시 선호 그룹에 속하는지 저장한다. 이렇게 함으로써 선호(목적)테이블의 중복된 데이터를 제거할 수 있다. 이를 그림으로 나타내면 아래 <그림 1>과 같다.

5.3 데이터베이스 스키마의 비교

본 논문에서 제안한 데이터베이스 모델의 메타데이터 스키마를 설명하기 전에 앞서 설명한 다양한 모델들의 스키마를 설명하고자 한다. 먼저 프라이버시를 고려하지 않는 일반적인 데이터베이스 환경은 그림 2의 데이터 테이블만을 가진다. 즉 SID라는 키 속성과 |A|개의 일반속성으로 구성된 데이터 테이블



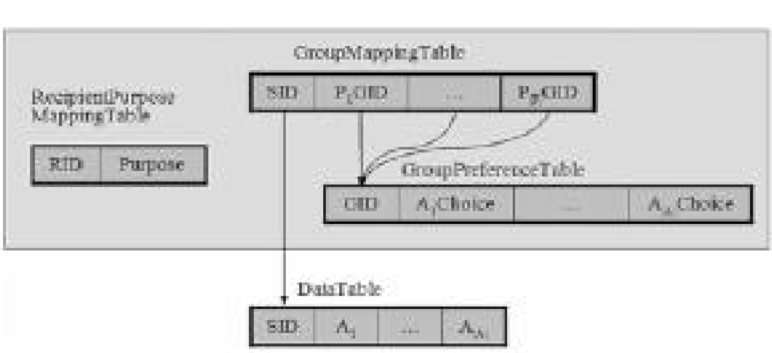
<그림 2> LDHD의 메타데이터의 스키마

스키마를 가질 뿐 메타데이터를 가지고 있지 않다.

하지만 프라이버시를 보호하기 위한 모델들은 메타데이터가 필요하다. LDHD 모델의 경우 (사용자ㅡ목적)에 따라 조사된 데이터 제공자의 선호를 저장하기 위해서 선호(사용자ㅡ목적) 테이블을 사용한다. 메타데이터에 이 테이블은 (사용자ㅡ목적)의 개수만큼 저장된다. 또한 각 사용자(Recipient)가 데이터에 접근할 수 있는 목적을 나타내는 사용자 목적 연결 테이블 (RecipientPurpose MappingTable)이 필요하다. 이를 그림으로 표현해 보면 <그림 2>와 같다.

5.1절에서 제안하였던 PBDM의 사용자 목적 연결 테이블은 LDHD와 같은 구조를 갖는다. 하지만 LDHD 방식의 선호(사용자-목적) 테이블을 이용하는 대신 선호(목적)테이블을 이용한다.

5.2절에서 설명한 PBDM+G 데이터베이스 모델의 경우 새로운 형태의 스키마를 이용한다. 앞서 설명한 모델들과 가장 큰 차이는 그룹 선호 테이블(GroupPreferenceTable)을 이용한다는 점이다. 그룹 선호 테이블은 각 레코드의 선호를 기록하는 것이 아니라 프라이버시 선호 그룹의 선호를 기록한다. 기록된 프라이버시 선호 그룹을 이용해 각 레코드의 그룹을 찾는다. 예를 들어 SID가 5인 레코드가 P_2라는 목적에 대해서 {1, 0, 1, 0, 0}라는 선



<그림 3> PBDM+G의 데이터베이스 스키마

호를 갖는다면 그룹 선호 테이블에서 그러한 선호를 나타내는 그룹을 찾아 그룹 아이디인 GID를 확인한다. 다음 그룹 연결 테이블(GroupMappingTable)의 P2GID라는 속성에 그 GID를 기록한다.

5.4 메타데이터의 크기 비교

앞에서 제안한 데이터베이스 모델들의 대략적인 메타데이터 크기를 계산하기 위해 다음과 같이 변수들을 정의한다. 먼저 프라이버시 보호가 필요한 데이터 테이블은 |A|개의 속성으로 구성된 |S|개의 레코드로 구성되어 있다. 데이터베이스에서 사용된 목적의 개수는 |P|이고 사용자 수를 |R|이라고 한다. 이 변수들을 사용하면 메타데이터를 구성하는 셀들의 개수를 계산할 수 있다. 이 셀들의 개수를 각 모델별 메타데이터의 크기를 비교하기 위한 측정 단위로 이용한다. 이 때 다른 테이블들에 비해 상대적으로 크기가 작고 각 모델마다 공통적으로 가지고 있는 사용자 목적 연결테이블은 메타데이터 크기 비교대상에서 제외한다.

먼저 LDHD의 메타데이터는 (사용자－목적)의 개수만큼 선호(사용자－목적) 테이블이 필요하다. 선호(사용자－목적) 테이블은 데이터 테이블과 동일하게 (|A|+1)×S개의 셀을 갖는다. 이때 모든 사용자가 1개의 목적을 가지는 경우 (사용자－목적)의 개수는 가장 작은|R|개이다. 하지만 사용자가 데이터베이스에서 사용하는 모든 목적으로 접근 가능하다면 (사용자－목적)의 개수는 최대 |P|×|R|개까지 가능하다. 따라서 전체 메타데이터의 크기는 최소 |R|×(|A|+1)×|S|부터 최대 |P|×|R|×(|A|+1)×|S|까지의 셀로 구성된다.

본 논문에서 제시한 PBDM+G의 경우에는 새로운 변수인 프라이버시 선호 그룹의 개수 |G|가 필요하다. 먼저 그룹 연결 테이블은 |S|×(|P|+1)의 셀이 필요하고 그룹 선호 테이블은 |G|×(|A|+1)개의 셀이 필요하다. 따라서 메타데이터의 크기는 |S|×(|P|+1)+|G|×(|A|+1)개의 셀을 갖는다.

위의 수식으로 정의된 메타데이터의 크기를 각 모델별로 비교하기 위해 예를 들어보자. 실제 온라인 상점의 데이터베이스에 프라이버시 보호가 필요한 테이블이 한 개가 있고 그 테이블은 8개의 속성과 1,000,000개의 레코드로 구성되어 있다. 이 데이터베이스를 이용하는 사용자는 16명이고 목적의 개수는 8개이다. 이때 LDHD의 메타데이터는 최소 144,000,000개부터 최대 1,152,000,000개의 셀이 필요하다. 이에 비해 PBDM은 72,000,000개의 셀이 필요하다. PBDM+G의 경우 그룹의 개수에 따라 테이블의 크기가 달라진다. 그룹의 개수는 테이블이 8개의 속성으로 이루어져 있기 때문에 최대 256(28)개이다. 따라서 최대 9,002,304개의 셀이 필요하다. 이를 비교해 보면 LDHD에서 이용하는 최소의 셀 개수에 비해 PBDM+G는 단지 7%의 개수만큼의 셀만을 이용하는 것을 알 수 있다.

5.5 질의 수정 알고리즘

LDHD 모델에서는 case문을 이용해 프라이버시 선호를 보호할 수 있도록 질의를 수정한다. 즉 선호(사용자－목적) 테이블에서 데이

```
select  case when exists ( select  Ai
                           from    PreferenceTableBy(Pr)-1 as PT
                           where   DT.SID = PT.SID and PT.AiChoice = 1 )
              then Ai else null end,
        case when exists ( select  Aj
                           from PreferenceTableBy(Pr)-1 as PT
                           where   DT.SID = PT.SID and PT.AjChoice = 1 )
              then Aj else null end
from DataTable as DT
```

<그림 4> LDHD의 방식으로 수정된 질의

```
select  case when PT1.AiChoice = 0 then null else DT.Ai end,
        case when PT2.AjChoice = 0 then null else DT.Aj end
from    DataTable as DT left outer join GroupMappingTable as GMT
                            on DT.SID = GMT.SID
                    left outer join GroupPreferenceTable as PT1
                            on GMT.PrGID = PT1.GID and PT1.AiChoice = 1
                    left outer join GroupPreferenceTable as PT2
                            on GMT.PrGID = PT2.GID and PT2.AjChoice = 1
```

<그림 5> 프라이버시 선호 그룹을 이용한 데이터베이스 모델 방식으로 수정된 질의

<표 3> 프라이버시를 고려한 데이터베이스 관리 시스템의 조건과 비교표

번호	PBDM	PBDM+G
1	사용 목적 수집	사용 목적 수집
2	모두 수집	모두 수집
3	셀 단위	셀 단위
4	간편	아주 간편
5	많은 메타데이터	적은 메타데이터
6	적음	적음
7	있음	있음
8	없음	없음

터의 제공을 허용하면 사용자에게 실제 데이터를 제공하고 그렇지 않으면 “NULL"을 반환하는 질의 수정 알고리즘을 사용한다. 즉 사용자 R이 Pr이라는 목적으로 데이터를 사용할 때 사용자가 작성한 ‘select Ai, Aj from DataTable’ 질의는 <그림 4>와 같이 수정된다.

5.3절에서 설명한 것과 같이 본 논문에서 제안하는 PBDM+G은 프라이버시 선호 그룹을 이용하므로 메타데이터를 구성하는 데이터베이스 스키마가 LDHD와 다르다. 즉 사용자 R이 Pr이라는 목적으로 데이터를 사용할 때 사용자가 작성한 ‘select Ai, Aj from DataTable’ 질의는 그림 5와 같이 수정된다. 먼저 사용자가 검색하려는 데이터 Ai, Aj의 제공 여부는 그룹 선호 테이블에 표시되어 있다. 이 테이블에서 데이터 Ai, Aj가 제공되는 GID를 찾아 이 GID를 갖는 SID를 그룹 연결 테이블에서 확인하고, 이 SID에 해당하는 데이터만을 데이터 테이블에서 찾아 제공한다. PBDM+G는 데이터 제공 여부를 나타내는 선호(사용자－목적) 테이블 대신 그룹 선호테이블과 그룹 연결 테이블을 이용하므로 질의 처리시 이용하는 테이블의 종류가 늘어난다. 따라서 정규화 되어 있는 테이블들의 조인 비용을 줄여 효율적인 질의 처리를 하기 위해서 left outer join을 이용한다. 이렇게 조인된 테이블을 바탕으로 case문을 이용하는 질의 수정 알고리즘을 사용한다.

앞장에서 소개한 프라이버시를 고려한 데이터베이스 관리 시스템의 조건들과 PBDM과 PBDM+G를 비교하면 <표 3>과 같다.

6. 성능 평가

본 장에서는 실험에 의한 성능 평가를 통하여 제안한 모델을 이전 모델들과 비교한다. 6.1절에서는 실험환경 및 방법을 설명하고 6.2절에서는 실험 결과를 보여주고 분석한다.

6.1 실험 환경 및 방법

본 실험에서는 제안한 모델의 성능 평가를 위해 1) 프라이버시를 고려하지 않은 질의(Unmodified), 2) LDHD 모델에서 사용한 질의(LDHD), 3) 그룹 개념을 적용하지 않은 PBDM 모델에서 사용하는 질의(PBDM), 4) 그룹 개념 적용 후의 PBDM 모델에서 사용하는 질의(PBDM+G)의 질의 처리 시간과 메타데이터 크기를 비교한다.

LDHD 논문에서는 case문을 이용하는 질의 수정 방법과 outer join을 이용하는 질의 수정 방법이 제안되었다. 하지만 LDHD의 실험 결과에 따르면 case문을 이용하는 질의가 outer join을 이용하는 질의보다 더 좋은 성능을 보여주는 것으로 나타났다. 따라서 본 논문에서는 case문을 이용한 질의 수정 방법만을 비교 대상으로 사용한다. PBDM 질의는 LDHD 질의와 동일한 스키마를 가지므로 질의 처리 시간은 따로 측정하지 않고 LDHD 질의와 동일한 질의 처리 시간이 걸리는 것으로 간주한다. 하지만 PBDM과 LDHD의 메타데이터 스키마는 다르기 때문에 메타데이터 크기는 각자 비교한다.

실험 결과의 측정값으로 사용하는 질의 처리 시간은 각 모델 별로 10회를 측정하고 상

<표 4> 벤치마크 데이터셋과 choice column

Column	Description
Unique2 (int)	Primary Key, Sequential order
Unique1 (int)	Candidate key, random order
Onepercent (int)	Values 0-99, random order
Tenpercent (int)	Values 0-9, random order
Twentypercent (int)	Values 0-4, random order
Fiftypercent (int)	Values 0-1, random order
stringu1 (32-byte str)	Unique character string
stringu2 (32-byte str)	Unique character string
Choice0 (int)	Values 0-1 (5% = 1), indexed
Choice1 (int)	Values 0-1 (20% = 1), indexed
Choice2 (int)	Values 0-1 (50% = 1), indexed
Choice3 (int)	Values 0-1 (80% = 1), indexed
Choice4 (int)	Values 0-1 (100% = 1), indexed

하위 2개씩을 제외한 나머지 6개의 평균을 이용한다. 메타데이터 크기는 메타데이터 테이블의 크기와 각 메타데이터 테이블의 인덱스 크기의 합으로 측정한다.

LDHD의 경우와 동일하게 <표 4>의 Wisconsin Benchmark[17] 데이터셋을 사용하여 실험을 수행한다.

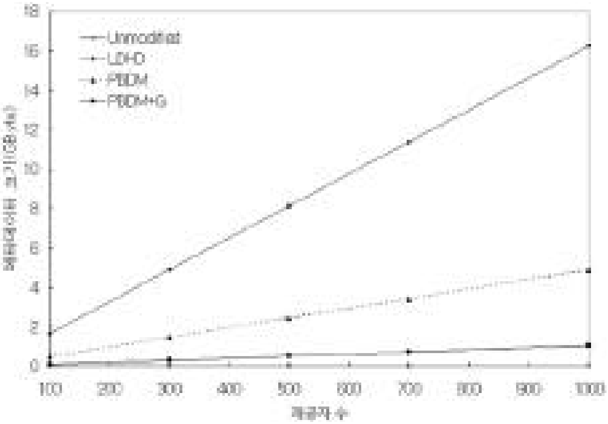
LDHD 논문의 실험과 비교하기 위해 본 논문에서 사용된 데이터베이스에서도 인덱스를 생성하여 시간을 측정한다. 모든 모델들의 시간 측정에 공통적으로 사용되는 데이터 테이블의 경우 주 키(primary key)에 해당하는 SID 컬럼에 대해서 인덱스를 생성한다. LDHD 모델에서 사용하는 선호(사용자－목적) 테이블의 경우 LDHD 논문의 실험과 같은 방식으로 전체 컬럼에 인덱스를 생성한다. PBDM+G 모델 실험에서 필요한 그룹 선호 테이블은 primary key인 GID컬럼에 인덱스를 생성한다. 또한 그룹 연결 테이블의 경우 전체 컬럼에 인덱스를 생성한다.

실험은 2.60GHz CPU와 512MB 메모리를 갖는 Pentium IV PC를 사용한다. 운영체제는 Microsoft Server 2003을 운영체제로 사용하고 DBMS는 Microsoft SQL Server 2005를 사용한다.

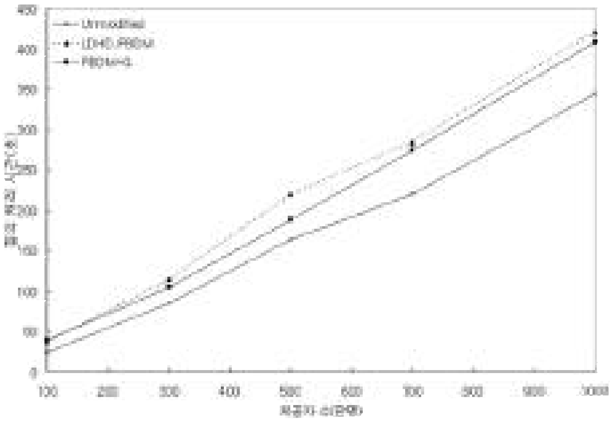
6.2 실험결과 및 분석

실험 1 : 데이터 제공자 수 변화에 따른 질의 처리 시간 및 메타데이터 크기 비교

실험 1에서는 데이터 제공자의 수를 100만명, 300만명, 500만명, 700만명, 1000만명으로 증가시켜 가며 각 모델들의 질의 처리 시간 및 메타데이터 크기를 측정한다. 이 때 데이



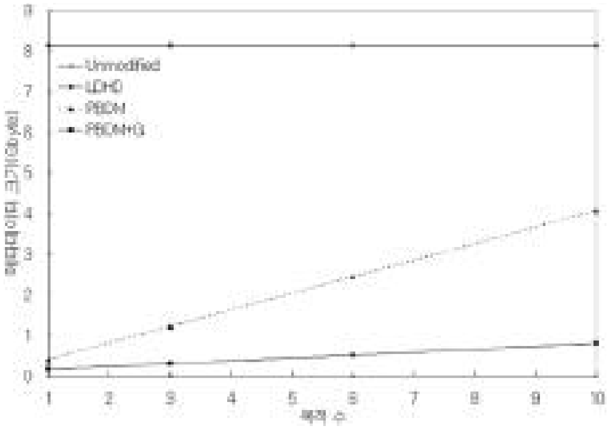
<그림 6> 데이터 제공자 수 변화에 따른 메타데이터 크기 비교



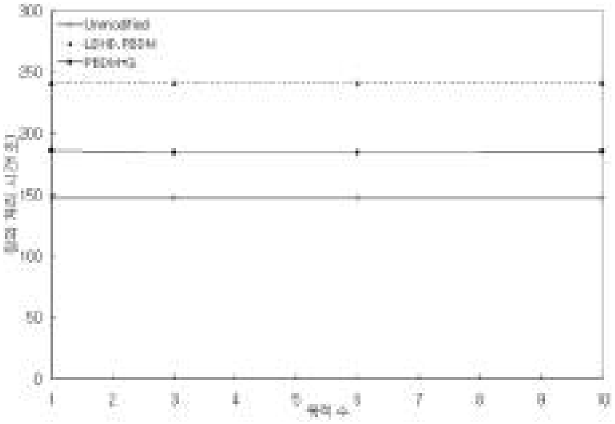
<그림 7> 데이터 제공자 수 변화에 따른 질의 처리 시간 비교

터 제공자가 어떤 목적에 대해서 자신의 데이터를 공개하는 비율인 선택률은 50%(예를 들어 선택률을 100%로 한다면 자신의 정보를 모두 제공하는 것을 의미한다.), 목적 수는 6개로 정하여 사용한다.

<그림 6>을 보면 데이터 제공자의 수가 증가할수록 LDHD와 PBDM+G가 많은 차이를 보이기는 하지만 두 모델 모두 메타데이터의 크기가 선형적으로 증가함을 알 수 있다. <그림 7> 역시 데이터 제공자의 수가 증가함에 따라서 질의 처리 시간이 선형적으로 증가함을 보여준다. 하지만 메타데이터의 크기와는 다르게 LDHD와 PBDM+G 두 모델간의 질의 처리 시간은 거의 차이가 나지 않는다. 그 이유는 PBDM+G의 경우는 전체 메타데이터의 크기가 줄어들지만 정규화에 따라 질의



<그림 8> 목적 수 변화에 따른 메타데이터 크기 비교



<그림 9> 목적 수 변화에 따른 질의 처리 시간 비교

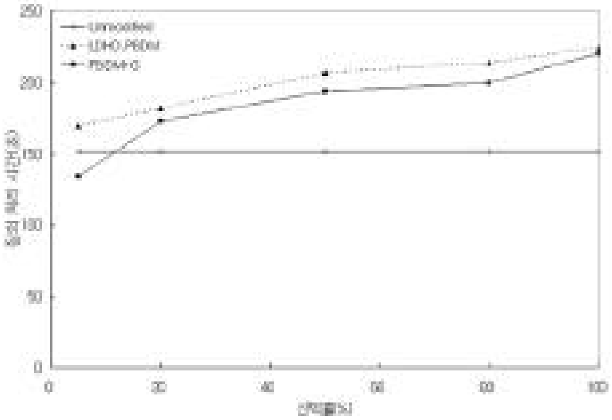
처리 과정에서 한번의 조인 연산이 추가되어 질의 처리 시간이 증가되기 때문이다.

실험에 따르면 LDHD에 비해 메타데이터의 크기는 PBDM의 경우 약 30%로, PBDM+G의 경우 약 6.5%로 감소한다.

실험 2 : 목적 수 변화에 따른 질의 처리 시간 및 메타데이터 크기 비교

실험 2는 목적 수를 1개, 3개, 6개, 10개로 변경시키며 각 모델별 질의 처리 시간 및 메타데이터 크기를 측정한다. 이 때 데이터 제공자의 수는 500만명, 선택률은 50%로 정하여 사용한다.

<그림 8>을 보면 LDHD의 경우는 메타데이터의 크기가 일정한 반면 PBDM과 PBDM+G는 일정하게 증가함을 알 수 있다.



<그림 10> 선택률 변화에 따른 질의 처리 시간 비교

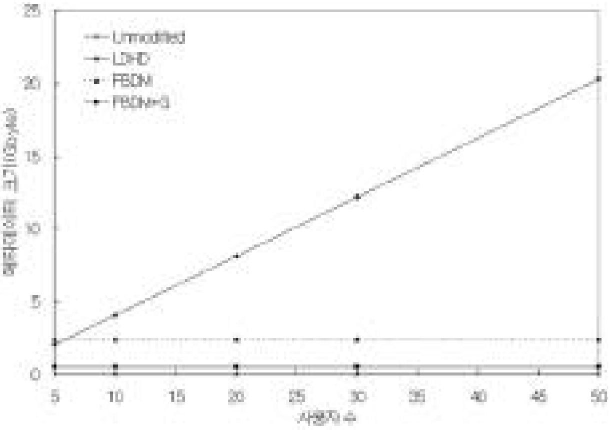
LDHD의 경우 한명의 사용자가 하나의 목적만을 가지고 데이터를 사용하므로 목적 개수에 상관없이 일정한 크기의 메타데이터를 갖는다. 하지만 PBDM과 PBDM+G의 경우 목적 수가 증가할수록 선호(목적)테이블의 수와 그룹 연결 테이블의 크기가 각각 커지기 때문에 메타데이터의 크기가 증가한다. 그러나 PBDM과 PBDM+G가 가장 큰 메타데이터 크기를 갖는 목적수가 10개인 경우에도 LDHD에 비해서 메타데이터 크기가 각각 50%와 10%로 감소함을 알 수 있다.

하지만 메타데이터의 크기와는 다르게 목적 수가 증가하여도 질의 처리 시간은 일정하다. 그 이유는 PBDM의 경우 질의에 사용된 목적의 수는 질의 처리에 사용되는 메타 데이터의 크기와 무관하기 때문이다. 그리고 PBDM+G의 경우 질의 처리에 사용되는 그룹 연결 테이블의 크기가 비록 커지기는 하지만 인덱스의 사용으로 인해 질의 처리 시간은 거의 증가하지 않는다.

실험 3 : 선택률 변화에 따른 질의 처리 시간 비교

실험 3에서는 선택률을 5%, 20%, 50%, 80%, 100%로 변화시키며 각 모델의 질의 처리 시간을 측정한다. 이때 선택률의 변화는 메타데이터의 크기에 영향을 미치지 않으므로 메타데이터의 크기 측정 실험은 하지 않는다. 실험에서 데이터 제공자 수는 500만명, 목적 수는 6개로 정하여 사용한다.

<그림 10>을 보면 unmodified의 경우는 데이터 제공자의 전체 데이터를 공개하므로 모든 경우 선택률이 100%가 되어서 항상 일정한 질의 처리 시간을 갖는다. 하지만 LDHD과 PBDM, PBDM+G은 선택률이 증가할수록 공개되는 데이터가 더 많아지므로 선택률이 증가함에 따라 질의 처리 시간도 증가한다. 실험 결과에 따르면 PBDM+G의 질의 처리 시간은 LDHD방식에 비해서 최소 1.9%에서 최대 21%까지 빠름을 알 수 있다.



<그림 11> 사용자 수 변화에 따른 메타데이터 크기 비교

실험 4 : 사용자 수 변화에 따른 메타데이터 비교

실험 4는 사용자 수 변화에 따른 각 모델의 메타데이터 크기를 비교한다. 사용자 수가 변화함에 따라 메타데이터의 크기는 변하지만 질의에 사용되는 테이블은 변하지 않기 때문에 질의 처리 시간을 제외한 메타데이터 크기만을 비교한다. 그림 11은 사용자 수가 5명, 10명, 20명, 30명, 50명으로 증가함에 따른 각 모델의 메타데이터 크기를 보인다. 이 때 데이터제공자 수는 500만명, 선택률은 50%, 목적 수는 6개로 한다.

데이터 사용자가 많아질수록 LDHD, PBDM, PBDM+G 모두 사용자 선호 연결 테이블의 크기가 증가하지만 그 크기는 다른 메타데이터보다는 아주 작은 크기이다. 따라서 그림 11의 PBDM과 PBDM+G는 데이터 사용자 수가 증가하여도 거의 일정한 크기의 메타데이터를 갖는다. 하지만 LDHD 모델은 데이터 사용자의 목적에 따라 각각 선호(사용자–목적) 테이블이 필요하기 때문에 사용자가 증가함에 따라 메타데이터 크기가 크게 증가하게 된다.

PBDM, PBDM+G 모두 데이터 사용자 수가 50명일 때 가장 큰 메타데이터를 가지는데, 이 경우에도 LDHD에 비해서는 PBDM의 경우 메타데이터 크기가 12%, PBDM+G의 경우 2.6%로 감소한다.

7. 결　론

본 논문에서는 프라이버시 보호가 필요한 데이터를 위하여 목적 기반의 효율적인 데이터베이스 보안 모델을 제시하였다.

제안한 모델은 다른 모델들에 비해서 다음과 같은 장점을 가지고 있다. 1) 먼저 LDHD에서 (사용자–목적)단위로 프라이버시 선호를 수집하는 대신 (목적)단위로 수집함으로

써 불필요한 정보 수집을 줄일 수 있다. 2) 수집된 프라이버시 정보를 프라이버시 선호 그룹을 이용한 정규화 과정을 통해서 기존의 LDHD모델의 데이터베이스 스키마에 비해 적은 양의 메타데이터를 이용한다. 3) 프라이버시 선호를 조사할 때 데이터 제공자는 미리 조사되어 있는 프라이버시 선호 그룹 정보를 이용해서 제시된 그룹 중에 한 그룹을 선택함으로써 손쉽게 프라이버시 선호를 수집할 수 있다.

실험에 따르면 본 논문에서 제안한 PBDM+G 방식은 LDHD방식에 비해서 질의 처리 시간은 최대 23.6%까지 빨라지고, 메타데이터 크기는 LDHD 방식에 비해 약 2.6%에서 10%만을 가지는 효율적인 구조임을 알 수 있었다.

본 논문의 주요 공헌은 1) 프라이버시 보호를 위한 데이터베이스 관리 시스템의 조건을 정의하고 2) 적은 메타데이터를 사용하여 프라이버시를 효율적으로 보호할 수 있는 새로운 형식의 메타데이터 스키마를 제안한 것이다. 또한 3) 제안한 메타데이터 테이블의 스키마에 맞는 효율적인 질의 처리 방법을 제시하였다.

참 고 문 헌

[1] Office of the Information and Privacy Commissioner, Ontario, Data Mining : Staking a Claim on Your Privacy, 1998.

[2] The Economist. The End of Privacy, May 1999.

[3] European Union. Directive on Privacy Protection, October 1998.

[4] Time. The Death of Privacy, August 1997.

[5] Online Americans More Concerned about Privacy than Health Care, Crime, and Taxes, New Survey Reveals , http : //www.nclnet.org/pressessentials.htm

[6] R. S. Sandhu, E. J. Coyner, H. L. Feinstein, and C. E. Youman, Role-Based Access Control Models, IEEE Computer, Vol. 29. No. 2, pp. 38-47, 1996.

[7] R. Agrawal, J. Kiernan, R. Srikant, and Y. Xu, Hippocratic Databases, In Proc. International Conference on Very Large Data Bases, 2002.

[8] K. LeFevre, R. Agrawal, V. Ercegovac, R. Ramakrishnan, Y. Xu, and D. DeWitt, " Limiting Disclosure in Hippocratic Databases", VLDB 2004, pp 108-119.

[9] M. H. Harrison, W. L. Ruzzo, and J. D. Ullman. Protection in operating systems. Communications of the ACM, 19(8) : 461-471, 1976.

[10] C. J. McCollum, J. R. Messing, and L. Notargiacomo. Beyond the pale of MAC and DAC-Defining new forms of access control. In Proc. of the IEEE Symposium on Security and Privacy, pages 190-200, Oakland, CA, 1990.

[11] P.P. Griffiths and B. W. Wade. An authorization mechanism for a relational database system. ACM Transactions on

Database Systems, 1(3) : 242-255, 1976.
[12] S. Castano, M. G. Fugini, G. Martella, and P. Samarati, Database Security, Addison-Wesley, 1995.
[13] GA. K. Jones, R. J. Lipton, and L. Snyder. " A linear time algorithm for deciding security", In Proc. FOCS, pp 33-41. IEEE, 1976.
[14] C. Wood, R. C. Summers, and E. B. Feranadez, " Authorization in multilevel database models", Information Systems, Vol. 4, No. 2, 1979.
[15] D. E. Bell and L. J. La Padula, Secure Computer Systems : mathematical foundations and model, Technical report M74-244, MITRE Corp., 1974.
[16] K. J. Biba, Integrity considerations for secure computer systems, Technical report 76-372, MITRE Corp., 1977.
[17] D. DeWitt. The Wisconsin benchmark : Past, present, and future. In J. Gray, editor, The benchmark Handbook. Morgan Kaufmann, 1993.

임 지 영

2002. 2 : 전북대학교 컴퓨터과학과(학사)
2005. 3 ~ 현재 : 연세대학교 컴퓨터과학과 석사과정 재학중
관심 분야 : 데이터베이스 보안
e-mail : jylim@cs.yonsei.ac.kr

김 우 철

2003. 2 : 연세대학교 컴퓨터과학과(학사)
2006. 2 : 연세대학교 컴퓨터과학과(석사)
2006. 3 ~ 현재 : 연세대학교 컴퓨터과학과 박사과정 재학중
관심분야 : 데이터베이스 보안, 바이오인포메틱스, LBS, 멀티미디어 데이터베이스 등
e-mail : twelvepp@cs.yonsei.ac.kr

박 상 현

1989. 2 : 서울대학교 컴퓨터공학과(학사)
1991. 2 : 서울대학교 컴퓨터공학과(석사)
2001. 2 : UCLA대학교 전산학과(박사)
1991. 3 ~ 1996. 8 : 대우통신 연구원
2001. 2 ~ 2002. 6 : IBM T. J Watson Research Center Post-Doctoral Fellow.
2002. 8 ~ 2003. 8 : 포항공과대학교 컴퓨터공학과 조교수
2003. 9 ~ 현재 : 연세대학교 컴퓨터과학과 조교수
관심분야 : 데이터베이스 보안, 데이터마이닝, 바이오인포매틱스, XML
e-mail : sanghyun@cs.yonsei.ac.kr