韓國IT서비스學會誌
第13卷 第2號
2014年 6月, pp.163-172
http://dx.doi.org/10.9716/KITS.2014.13.2.163

OLAP 환경에서 개인정보보호를 위한 개인정보 분리 권한관리 모델

김형규* · 김민호** · 권정숙** · 최용락***

A Model of Authority Management for the Protection of Personal Information in OLAP

Hyoung-Gyu Kim* · Min-Ho Kim** · Jung-Sook Kwon** · Yong-Lak Choi***

■ Abstract ■

Personal information has been stolen continuously and it is also affected from development of the Internet. So the government requires that companies spend more effort for protecting customers' personal information. The OLAP server also should meet this requirement, but it is hard to satisfy for the authority management. The OLAP server must use personal information to extract required information from database. This thesis suggests a model of separating between general information and personal information, so this model can help to minimize the leakage of personal information. The model is implemented and tested as a prototype. This prototype can prove that the new model is better than the original one. This study presents that the authority management on the separation between personal information and general information helps protect the personal information of customers.

Keyword : OLAP, Personal Information, Authority Management

논문투고일 : 2014년 04월 24일 논문수정완료일 : 2014년 06월 08일 논문게재확정일 : 2014년 06월 13일
* 숭실대학교 SW특성화대학원 석사, 주저자
** 숭실대학교 SW특성화대학원 석사, 공동저자
*** 숭실대학교 SW특성화대학원 교수, 교신저자

1. 서 론

OLAP(On-Line Analytical Processing)는 최종 사용자가 축적된 대용량 데이터베이스를 이용하여 의사결정에 도움을 주는 처리기법이다. 이때 대용량 데이터베이스에 개인정보를 포함해야만, 기업이 요구하는 분석 자료를 도출할 수 있다. 그러나 기존 OLAP 시스템은 개인정보를 포함하여 데이터를 추출할 경우, 개인정보 유출 위험이 커지는 단점이 발생한다. 최근 카드사 및 통신사 등에서 개인정보가 외부로 유출되는 문제가 발생한 것을 보면 개인정보보호의 필요성은 더욱 요구된다. 따라서 본 논문에서 제시하는 OLAP 시스템은 개인정보와 일반정보를 분리하는 모델을 제시한다.

본 연구는 개인정보와 일반정보가 혼재된 데이터베이스를 이원화하는 모델로 제시하고, 개인정보 분리 전과 개인정보 분리 후 권한관리 모델을 비교 테스트한다. 이렇게 함으로써 개인정보에 접속할 수 있는 사용자를 최소화하고, 관리자의 유지 보수의 편리성을 증명한다.

본 논문의 구성은 다음과 같다. 제 2장에서는 관련 연구를 통해서 권한관리 모델의 필요성을 표현하고, 제 3장에서는 개인정보 분리 전후 모델을 비교하고, 제 4장에서는 프로토타입을 구현하며, 제 5장에서는 결론을 기술한다.

2. 관련 연구

본 장에서는 개인정보, 권한관리, OLAP와 OLTP (On-Line Transaction Processing)에 대해서 알아본다.

2.1 개인정보

2012년 2월 17일 개정된 개인정보보호법에 따르면 개인정보란 생존하는 개인에 관한 정보로서 성명 · 주민등록번호 등에 의하여 특정한 개인을 알아볼 수 있는 부호 · 문자 · 음성 · 음향 및 영상 등의 정보(해당 정보만으로는 특정 개인을 알아볼 수 없어도 다른 정보와 쉽게 결합하여 알아볼 수 있는 경우에는 그 정보를 포함한다)를 말한다[3]. 또한, 아이디, 생년월일, 전자우편, 전화번호, 주소 등 직접 수집하는 인적정보와 서비스 이용 과정에서 생성되어 다른 정보와 쉽게 결합하여 개인식별이 가능한 경우도 개인정보로 볼 수 있다. 이러한 개인정보가 유출될 경우 기업, 개인 모두에게 피해를 주게 된다. 따라서 기업은 개인정보 유출에 인한 기업의 손실비용을 추정하고 개인정보 침해사고 예방을 위해 합리적인 투자규모를 도출하여 개인정보 유출을 미리 방지해야 한다[2].

2.2 권한관리

개인정보 유출의 심각성을 알고 있는 기관과 기업은 암호화를 포함한 2차 이상의 정보보안을 적용하고 있다. 그러나 아직도 암호화가 전면적으로 운영되는 것은 아니다. 개인정보에 대한 암호화 개발 이전에 취할 수 있는 내부 사용자 권한관리에 대한 방법으로 개인정보를 보호할 수 있다[1].

2.3 OLAP와 OLTP 비교

OLAP는 1990년대 중반부터 DW(Data Warehouse)와 결합하여 기업의 의사결정을 지원하는 시스템이 되었다[7]. 관계형 데이터베이스에 공헌한 Codd는 본인이 만든 관계형 모델은 OLTP에 적합하지만, 온라인 분석처리에는 부적합하다고 판단하여 OLAP를 만들었다[6]. OLAP에서 사용하는 DW는 OLTP를 통해서 생성된 데이터를 대용량 데이터베이스에 저장하여 만들고, 필요한 업무 단위별로 데이터베이스를 만들면 DM(Data Mart)이 된다.

기업에서 개인정보를 위해서 OLTP에는 많은 투자를 하지만 OLAP는 개인정보보호가 잘 적용되지 않는 상태이다. 이에 OLTP와 OLAP의 특징을 비교해 보고자 한다. OLTP와 OLAP를 비교하면 <표 1>과 같다. <표 1>에서 보는 바와 같이 OLAP

역시 개인정보를 포함하기 때문에 개인정보보호를 해야 하는 필요성이 있다.

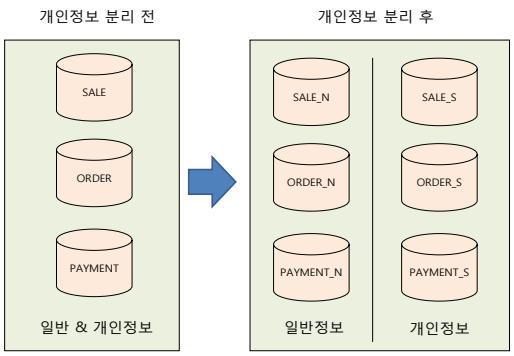
〈표 1〉 OLAP와 OLTP 비교

	OLAP	OLTP
정의	분석처리	거래처리
목적	비즈니스 분석	비즈니스 운영
저장내용	요약, 종합자료	갱신된 현재 값
사용법	정형, 비정형	단순 반복
자료량	많다	적다
개인정보 포함 유무	개인정보 포함	개인정보 포함

OLAP를 사용하면 최종 사용자가 DBMS의 사용법을 몰라도 다차원 분석을 통해 기업에 필요한 데이터를 분석하여 의사결정에 도움을 줄 수 있다[5, 8]. 이처럼 지금까지 대부분의 연구는 OLAP 서버를 활용하여 의사결정을 지원하는데 연구가 이루어졌다[9, 10]. 그러나 개인정보의 유출 및 오남용 사건이 계속하여 증가하고 있으며[1, 4], 특히 최근 대한민국 국민 대부분의 정보를 가지고 있는 카드사나 통신사가 개인정보 침해사고를 당했다. 이는 기업의 손실일 뿐만 아니라, 개인 차원에서도 개인정보가 악용되고 있고 이러한 피해 사례는 증가하고 있다. 따라서 정부도 개인정보보호를 요구하고 있지만, OLAP 서버에서 개인정보보호는 아직 미흡한 상태이다.

3. 개인정보 분리 전후 모델 비교

대부분 기업은 운영상의 편리성을 위해서 일반정보와 개인정보를 포함한 데이터마트를 만들고, 이렇게 구축된 데이터를 OLAP 서버에서 개인별로 권한을 설정하여 사용한다. 이는 사용자에게는 편리성을 제공하지만, 개인정보가 유출될 가능성을 높이게 된다. 개인정보 유출될 가능성을 줄이기 위해 개인정보 분리 모델을 제시하고자 한다. [그림 1]의 개인정보 분리 전 모델에서는 일반정보와 개인정보가 혼재된 데이터마트를 생성하고 사용자 접근 권한을 설정했다. 반면에 개인정보 분리 후에는 일반정보만 포함하는 데이터마트와 개인정보를 포함하는 데이터마트를 생성했다. 이러한 분리방법은 데이터의 중복을 허용하지만, 일반 사용자는 개인정보에 접근하는 방법을 원천적으로 차단할 수 있다.



[그림 1] 개인정보 분리 전후 모델 비교

4. 프로토타입 구현

4.1 프로토타입 구현 방법 및 대상 시스템 선정

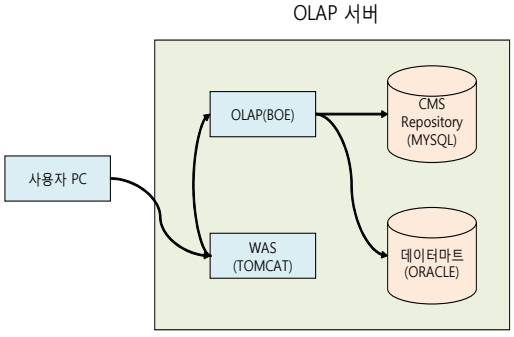
개인정보 분리 모델을 OLAP 서버에 적용하는 프로토타입을 구현하고자 한다. 프로토타입은 일반정보와 개인정보가 혼재되어있는 모델과 개인정보가 분리된 모델을 비교한다. 현재 가장 많이 사용되는 OLAP 서버는 SAP BO(Business Objects), MSTR(MicroSTRategy), SAS(Statistical Analysis System) 등이 있다. MSTR, SAS는 권한 관리가 어려운 단점이 있다. BO는 국내외 기업에 많이 배포되어 있고, 안정성을 인정받았으며, 유니버스 기능을 이용하여 데이터의 권한관리를 쉽게 할 수 있다. 또한, 리포트와 사용자의 권한관리도 수월하다.

이 테스트는 개인정보 분리 전과 분리 후 모델을 비교하는 프로토타입이므로 그룹과 권한 설정이 수월한 BO를 활용하여 프로토타입을 진행했다. 특히

BO는 일반정보만 포함하는 데이터베이스를 중복으로 생성하지 않고, 일반정보와 개인정보가 같이 있는 데이터베이스에서 일반정보만 불러와 일반정보만 포함하는 유니버스를 만들고, 개인정보를 포함하는 유니버스를 각각 만들 수 있다. 연구에 사용된 OLAP 서버는 국내 대표적인 온라인 쇼핑몰의 비즈니스 일부분을 기준으로 했다.

4.2 프로토타입 시스템 환경 및 구성

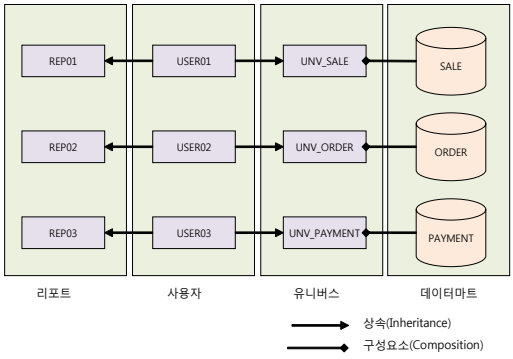
[그림 2]는 프로토타입에 사용된 시스템의 물리 구성도를 나타낸다. OLAP 서버는 CPU Intel®Core TM i7-3632QM @ 2.20GHz, 메모리 4GB, Windows 7으로 구축했다. 이 서버에는 OLAP(BOE : Business Objects Enterprise), WAS(TOMCAT), CMS Reposity(MYSQL), 데이터마트(ORACLE)를 설치했다. BOE는 Business Objects Enterprise XI Professional sp3 Fix Pack 3.4를 사용했다. WAS 서버는 TOMCAT 5.5.20을 사용했다. CMS Reposity는 MYSQL 5를 사용했다. 데이터마트는 ORACLE 10g를 사용했다. 사용자 PC는 CPU Intel(R) Pentium(R) Dual T2390@ 1.86GHz, 메모리 2GB, Windows 7을 사용했다. 네트워크는 100Mbps LAN으로 구성했다.



[그림 2] 프로토타입에 사용된 시스템 환경

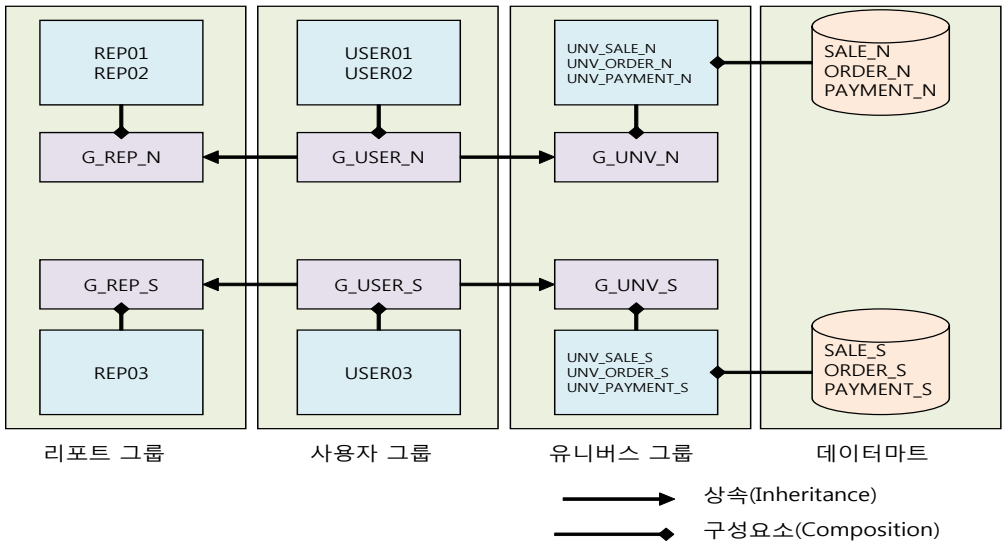
4.3 개인정보 분리 모델 적용 및 결과 분석

[그림 3]은 개인정보 분리 전 모델을 나타낸다. BOE는 크게 4가지(리포트, 사용자, 유니버스, 데이터마트)로 나누어진다. 여기서는 개인정보 분리 전 모델을 표현하기 위해서 데이터마트 3개(SALE, ORDER, PAYMENT)를 생성했다. 생성된 SALE로 UNV_SALE 유니버스를 만들면, USER01 사용자는 UNV_SALE을 사용하여 REP01 리포트를 만들 수 있다. 생성된 ORDER로 UNV_ORDER 유니버스를 만들면, USER02 사용자는 UNV_ORDER를 사용하여 REP02 리포트를 만들 수 있다. 생성된 PAYMENT로 UNV_PAYMENT 유니버스를 만들면, USER03 사용자는 UNV_PAYMENT를 사용하여 REP03 리포트를 만들 수 있다. 즉 각각의 사용자는 데이터마트 3개(SALE, ORDER, PAYMENT)가 포함하고 있는 개인정보를 이용하여 리포트를 생성할 수 있다.



[그림 3] 개인정보 분리 전 모델

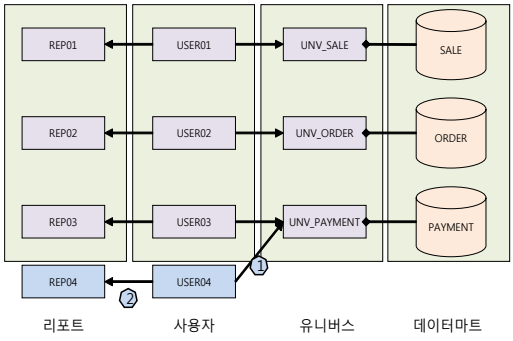
[그림 4]는 개인정보 분리 후 모델을 나타낸다. 여기서는 개인정보 분리 후 모델을 표현하기 위해서 일반정보만 포함하는 데이터마트(SALE_N, ORDER_N, PAYMENT_N)과 개인정보를 포함하는 데이터마트(SALE_S, ORDER_S, PAYMENT_S)를 생성했다. 만들어진 SALE_N은 UNV_SALE_N 유니버스 명으로 만들고, ORDER_N은 UNV_ORDER_N 유니버스 명으로 만들고, PAYMENT_N은 UNV_PAYMENT_N 유니버스 명으로 만든다. 이 유니버스들은 G_UNV_N 유니버스 그룹에 속하게 한다. G_USER_N 사용자 그룹을 만들어 USER01, USER02



[그림 4] 개인정보 분리 후 모델

사용자들을 속하게 하면, 이 사용자들은 G_USER_N 사용자 그룹의 권한을 상속받는다. G_REP_N 리포트 그룹을 만들어 REP01, REP02를 속하게 하면, G_REP_N의 권한을 가지고 있는 사용자는 REP01, REP02를 사용할 수 있다. 결국, 이 그룹의 권한을 가지고 있는 USER01, USER02 사용자는 SALE_N, ORDER_N, PAYMENT_N가 포함하는 일반정보만 사용할 수 있다.

반면에, SALE_S는 UNV_SALE_S 유니버스 명으로 만들고, ORDER_S는 UNV_ORDER_S 유니버스 명으로 만들고, PAYMENT_S은 UNV_PAYMENT_S 유니버스 명으로 만든다. 이 유니버스들은 G_UNV_S 유니버스 그룹에 속하게 한다. G_USER_S 사용자 그룹을 만들어 USER03 사용자를 속하게 하면, 이 사용자들은 G_USER_S 사용자 그룹의 권한을 상속받는다. G_REP_S 리포트 그룹을 만들어 REP03을 속하게 하면, G_REP_S의 사용권한을 가지고 있는 사용자는 REP03을 사용할 수 있다. 결국, 이 그룹의 권한을 가지고 있는 USER03 사용자는 SALE_S, ORDER_S, PAYMENT_S가 포함하는 개인정보를 사용할 수 있다. 따라서 개인정보를 분리하면 개인정보 유출을 최소화할 수 있다.

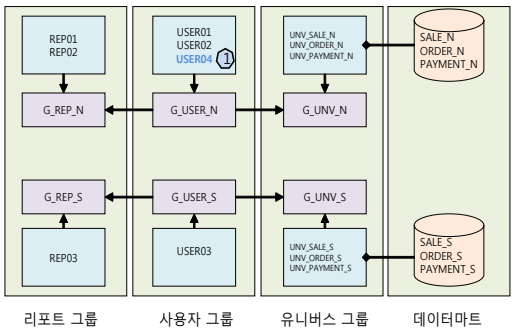


[그림 5] 개인정보 분리 전 모델에서 사용자 추가할 경우

[그림 5]는 개인정보 분리 전 모델에서 사용자 추가 시 관리자가 처리해야 할 단계를 나타낸다. 이미 생성된 유니버스 중 UNV_PAYMENT를 사용할 경우 USER04 사용자 생성 후, UNV_PAYMENT에 권한 할당하고, REP04 권한 할당하는 2단계 작업을 실행해야 한다.

[그림 6]은 개인정보 분리 후 모델에서 사용자 추가 시 관리자가 처리해야 할 단계를 나타낸다. 이미 생성된 유니버스를 사용할 경우 USER04 사용자 생성 후, G_USER_N 사용자 그룹에 속하게 하는 1단계만 실행하면 된다. 개인정보 분리 전 모

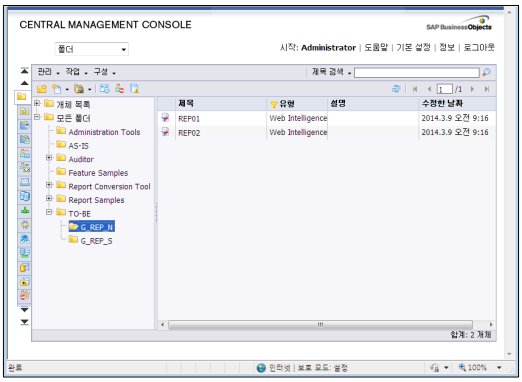
델에서 사용자 100명을 추가할 경우 200번의 권한 관리를 해야 하지만, 개인정보 분리 후 모델에서 사용자 100명을 추가할 경우 100번의 권한관리만 하면 된다. 결국, 관리자의 유지보수 작업이 0.5배로 줄어드는 효과를 볼 수 있다.



[그림 6] 개인정보 분리 후 모델에서 사용자 추가할 경우

4.4 개인정보 분리 모델로 개발된 프로토타입

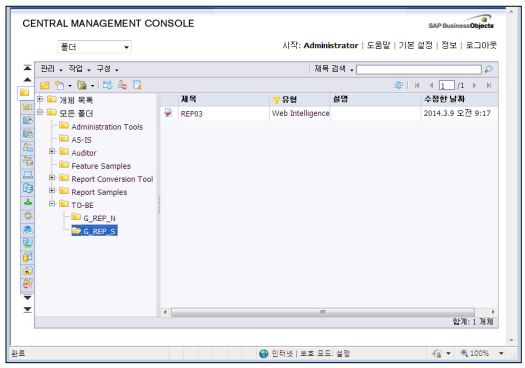
본 절에서는 [그림 4]의 개인정보 분리 후 모델을 기준으로 BOE를 가지고 프로토타입을 개발한 내용이다. 프로토타입의 시스템 환경 및 구성은 제 4.2절의 내용과 같다.



[그림 7] 일반정보 사용 가능한 리포트 그룹

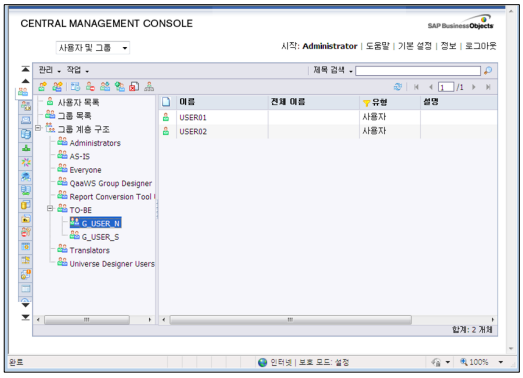
[그림 7]은 BOE의 CMC(Central Management Console)를 이용하여 일반정보만 사용 가능한 리포트 그룹(G_REP_N) 관리 기능을 개발한 화면이다. G_REP_N 그룹은 REP01, REP02의 CRUD(Create, Read, Update, Delete) 권한을 가진다.

[그림 8]은 CMC를 이용하여 개인정보 사용 가능한 리포트 그룹(G_REP_S)관리 기능을 개발한 화면이다. G_REP_S 그룹은 REP03의 CRUD 권한을 가진다.



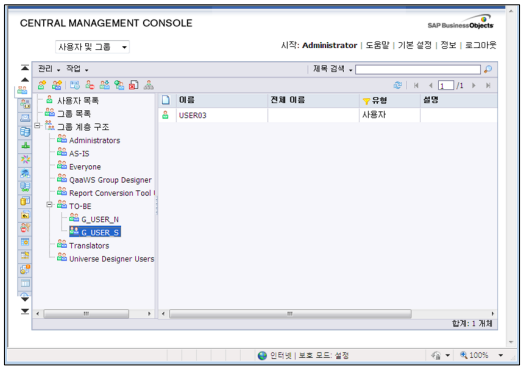
[그림 8] 개인정보 사용 가능한 리포트 그룹

[그림 9]는 CMC를 이용하여 일반정보 사용 가능한 사용자 그룹(G_USER_N) 관리 기능을 개발한 내용이다. USER01, USER02는 G_USER_N 그룹에 소속되어 G_USER_N 그룹의 권한을 상속받게 된다.



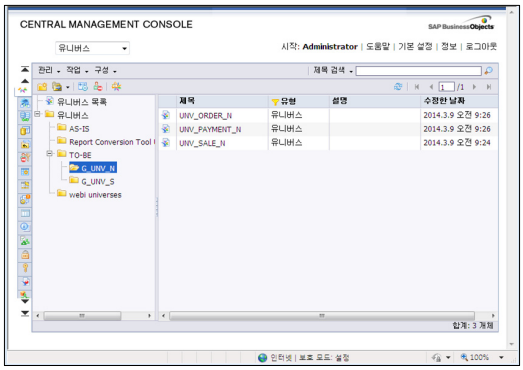
[그림 9] 일반정보 사용 가능한 사용자 그룹

[그림 10]은 CMC를 이용하여 개인정보 사용 가능한 사용자 그룹(G_USER_S) 관리 기능을 개발한 내용이다. USER03은 G_USER_S 그룹에 소속되어 G_USER_S 그룹의 권한을 상속받게 된다.



[그림 10] 개인정보 사용 가능한 사용자 그룹

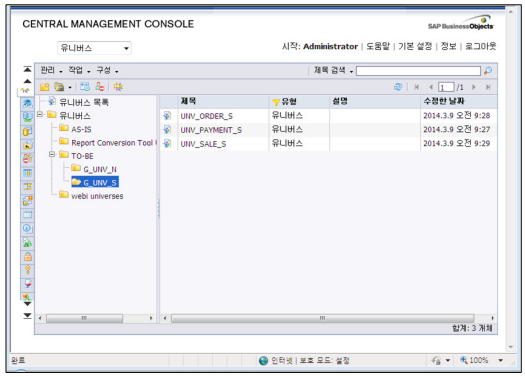
[그림 11]은 CMC를 이용하여 일반정보 사용 가능한 유니버스 그룹(G_UNV_N) 관리 기능을 개발한 내용이다. UNV_ORDER_N, UNV_PAYMENT_N, UNV_SALE_N는 G_UNV_N 그룹에 소속되어 G_UNV_N 그룹의 권한을 상속받게 된다.



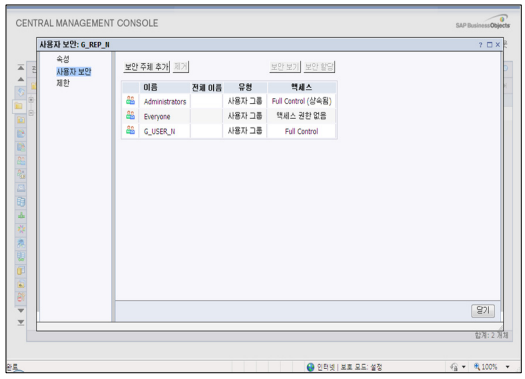
[그림 11] 일반정보 사용 가능한 유니버스 그룹

[그림 12]는 CMC를 이용하여 개인정보 사용 가능한 유니버스 그룹(G_UNV_S) 관리 기능을 개발한 내용이다. UNV_ORDER_S, UNV_PAYMENT_S, UNV_SALE_S는 G_UNV_S 그룹에 소속되어 G_UNV_S 그룹의 권한을 상속받게 된다.

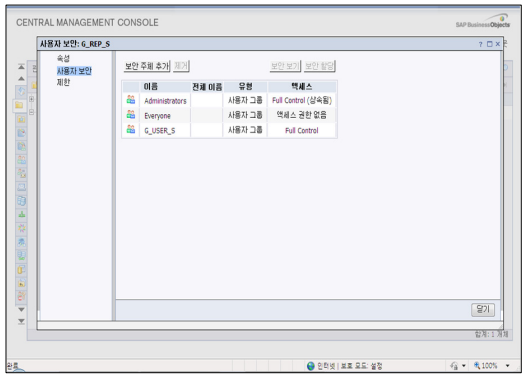
[그림 13]은 CMC를 이용하여 일반정보 사용자 그룹(G_USER_N)이 일반정보 리포트 그룹(G_REP_N)의 모든 권한을 상속받도록 Full Control을 설정한 화면이다. 따라서 일반정보 사용자 그룹은 일반정보 리포트 그룹의 사용 권한을 상속받게 된다.



[그림 12] 개인정보 사용 가능한 유니버스 그룹



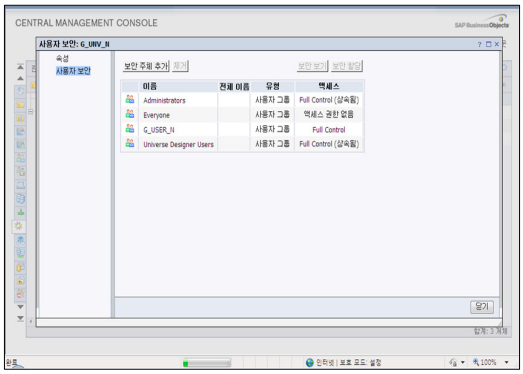
[그림 13] 일반정보 사용자 그룹에 일반정보 리포트 그룹 권한 설정



[그림 14] 개인정보 사용자 그룹에 개인정보 리포트 그룹 권한 설정

[그림 14]는 CMC를 이용하여 개인사용자 그룹(G_USER_S)이 개인정보 리포트 그룹(G_REP_S)의 모든 권한을 상속받도록 Full Control을 설정한

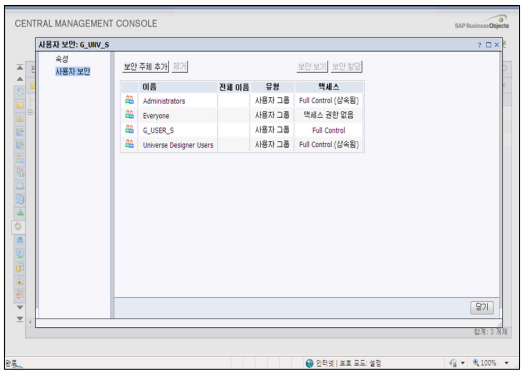
화면이다. 따라서 개인정보 사용자 그룹은 개인정보 리포트 그룹의 모든 사용 권한을 상속받게 된다.



[그림 15] 일반정보 사용자 그룹에 일반정보 유니버스 그룹 권한 설정

[그림 15]는 CMC를 이용하여 일반정보 사용자 그룹(G_USER_N)이 일반정보 유니버스 그룹(G_UNV_N)의 모든 권한을 상속받도록 Full Control을 설정한 화면이다. 따라서 일반정보 사용자 그룹은 일반정보 유니버스 그룹의 사용 권한을 상속받게 된다.

[그림 16]은 CMC를 이용하여 개인사용자 그룹(G_USER_S)이 개인정보 유니버스 그룹(G_UNV_S)의 모든 권한을 상속받도록 Full Control을 설정한 화면이다. 따라서 개인정보 사용자 그룹은 개인정보 유니버스 그룹의 모든 사용 권한을 상속받게 된다.



[그림 16] 개인정보 사용자 그룹에 개인정보 유니버스 그룹 권한 설정

5. 결 론

최근 기업 및 정부기관들의 개인정보 관리 소홀 때문에 개인정보가 유출되었고. 이에 정부는 기업 및 정부기관들이 반드시 지켜야 하는 개인정보보호법을 제정했다. 이러한 중요성을 인식하여 대부분의 홈페이지나 OLTP 환경에서 개인정보보호를 강화하고 있다. 그러나 시간과 자원의 부족으로 아직도 대부분의 기업 및 정부기관들은 적절한 개인정보보호를 하지 못하고 있다. 특히 OLAP 환경에서 일반정보와 개인정보가 혼재되어있는 데이터베이스를 이용하고 있다.

OLAP에서 이용하는 데이터는 의사 결정의 분석을 위해 OLTP의 데이터를 데이터웨어하우스나 데이터마트에 담아서 데이터를 이용하기 때문에 당연히 개인정보를 포함하고 있고, 오히려 아이디, 생년월일, 전자우편, 전화번호, 주소 등 직접 수집된 인적정보와 서비스 분석 과정에서 생성된 다른 정보와 결합하여 개인식별이 가능하여 개인정보를 추출하는데 더욱 편리하다. 이는 OLAP에서 개인정보가 유출될 가능성을 높이는 문제점이 있다. 따라서 본 논문에서 개인정보가 분리된 모델을 제시했다. 이 모델을 적용하면 OLAP에서 개인정보를 원천적으로 분리하여 개인정보 유출을 사전에 차단하고, 관리자의 유지 보수 편리성까지 도모할 수 있다.

참 고 문 헌

[1] 서우석, 전문석, “효율적인 계층별 관리자 및 내부 사용자 권한관리 및 접근제어에 관한 연구”, 「한국전자통신학회학술대회지」, 제5권, 제2호(2011), pp.181-185.

[2] 유진호, 지상호, 임종인, “개인정보 유노출 사고로 인한 기업의 손실비용 추정”, 「정보보호학회논문지」, 제19권, 제4호(2009), pp.63-75.

[3] 정보통신망 이용촉진 및 정보보호 등에 관한 법률, 국가법령정보센터, 개정 2012. 2. 17, http://www.law.go.kr/lsInfoP.do?lsiSeq=123210&

efYd=20120818#0000.

[4] 조성규, 전문석, “개인정보보호를 위한 개인정보 유출 모니터링 시스템의 설계”, 「정보보호학회논문지」, 제22권, 제1호(2012), pp.99-106.

[5] 조재희, “OLAP 테크놀로지의 이해”, 「정보과학회지」, 제21권, 제10호(2003), pp.23-30.

[6] Codd, E. F., S. B. Codd, and C. T. Salley, *Providing OLAP to User-Analysts : An IT Mandate*, White Paper, Codd and Date Inc., 1993.

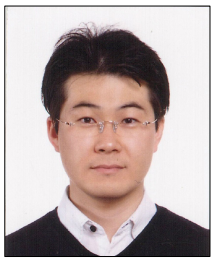
[7] Chaudhuri, S. and U. Dayal, *An Overview of Data Warehousing and OLAP Technology*, ACM, 1997.

[8] Sarawagi, S., R. Agrawal, and N. Megiddo, *Discovery-Driven Exploration of OLAP Data Cubes*, Springer, 1998.

[9] Vassiliadis, P. and T. Sellis, “A Survey of Logical Models for OLAP Databases”, *ACM Sigmod Record*, Vol.28, No.4(1999), pp.64-69.

[10] Zaiane, O. R., M. Xin, and J. Han, *Discovering Web Access Patterns and Trends by Applying OLAP and Data Mining Technology on Web Logs*, IEEE, 1998.

◈ 저 자 소 개 ◈



김 형 규 (corea2u@naver.com)

현재 숭실대학교 소프트웨어특성화대학원 석사 과정에 있다. 1999년 10월부터 벤처 기업에서 IT 일을 시작했으며, 2012년 12월까지 핸디소프트에서 외산 솔루션 컨설턴트로 근무를 했다. 자격증에는 CISA, OCP-DBA, 정보처리기사를 보유하고 있다. 주요 관심분야는 빅데이터, 데이터모델링, OLAP, 정보분석, 정보보호 등이다.

김 민 호 (kimito094@gmail.com)

현재 숭실대학교 소프트웨어특성화대학원 석사 과정에 있다. 경희대학교에서 서어서문 학사 학위를 취득하였고, 주요 관심분야는 빅데이터, 데이터베이스 모델링, 정보보안 및 개인정보 보호 등이다.

권 정 숙 (pippie@daum.net)

현재 숭실대학교 소프트웨어특성화대학원 석사 과정에 있다. 숙명여자대학교에서 전산학 학사 학위를 취득하였고, 주요 관심분야는 선진금융시스템, 품질보증, 정보보호, 빅데이터 등이다.

최 용 락 (ylchoi58@ssu.ac.kr)

숭실대학교 대학원에서 공학박사 학위 취득하여, 세종대 평생대학원 교수와, 숭실대학교 정보과학대학원 소프트웨어공학과 교수를 거쳐, 현재는 숭실대학교 소프트웨어특성화대학원 교수로 재직 중이며, 주요 관심분야는 데이터모델링, 소프트웨어공학, 정보전략기획 등이다.