데이터베이스 암호화 기술과 제품 동향

Database Encryption Technology and Current Product Trend

u-IT839의 정보보호 이슈 특집

이호균 (H.G. Lee) 보안게이트웨이연구팀 연구원
이승민 (S.M. Lee) 보안게이트웨이연구팀 선임연구원
남택용 (T.Y. Nam) 보안게이트웨이연구팀 책임연구원

목 차

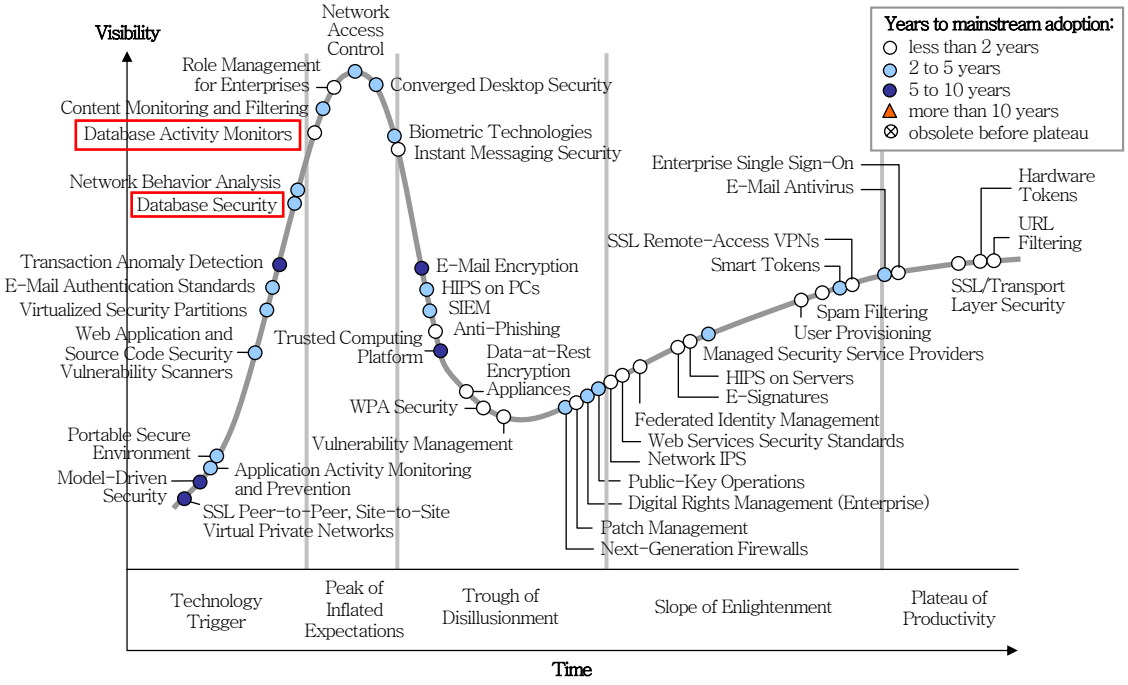
최근 개인정보보호법을 제정하려는 움직임으로 인해서 데이터베이스 보안 제품에 대한 관심이 국가기관뿐만 아니라 금융권, 포털사업자, ISP 등에서 급증하고 있다. 그 동안 써드 파티 업체들에 의해서 주도되어 오던 데이터베이스 보안 시장이 규모 면에서 크게 증가 추세에 있고, 이에 따라 오라클과 같은 DBMS 업체와 시만텍과 같은 메이저 보안 업체들도 속속 시장에 참여하고 있다. 본 고에서는 데이터베이스 보안 기술의 개요를 살펴보고 최근 기술 개발의 동향에 대해 정리하고자 한다. 현재까지, 시장에 출시된 제품들이 성능면에서 아직 시장의 요구를 만족시키지 못하고 있기 때문에 많은 기술 개발의 가능성이 남아 있는 상태이다. 또한 검색 가능한 암호 기술과 같은 학술적인 기술의 경우 기술적인 검증 문제 또한 남아 있기 때문에 이 분야에 대한 지속적인 연구 개발이 요구되는 바이다.

I. 서론

최근 사회가 급속하게 정보화 됨에 따라 개인 정보의 가치 또한 빠르게 상승하고 있다. IT 서비스가 점차 다양화되고 개인 맞춤형 서비스로 발전함에 따라, 수집된 개인 정보의 불법적인 접근 및 유출에 대한 우려가 증가하고 있다. 현재의 개인정보 이용 환경을 볼 때, 데이터 접근 시에 사용자의 질의 내용과 그에 대한 결과가 관리자 및 타인에게 그대로 노출되며, 이로 인해 발생하는 사용자의 프라이버시 침해가 반드시 해결되어야 할 문제로 지적되고 있다.

최근 인터넷상에서 노무현 대통령의 주민등록번호가 총 416회 사용되었고 280회의 성인인증을 받았으며, 한명숙 총리와 함께 여러 게임 사이트에도 가입되는 등 무분별하게 개인정보가 도용되어 심각한 사회적 문제로 대두되었다[1]. 또한 국내 모 이동통신사가 주최한 이벤트 응모에서 2,000여 명이 넘는 응모자들의 주민번호와 집주소 등 개인정보가 구글 검색 엔진에 뚫려 인터넷에 완전히 노출된 사건도 있었다. 이러한 문제를 근본적으로 해결하기 위한 방법으로는 DB 내에 데이터 자체를 암호화하여 저장하는 방식이 필수적이다. DB 암호화를 위한 방법으로 DBMS 전문 업체에서 제공하는 암호화와 제 3자에 의한 DB wrapper 형태의 암호화 솔루션이 있으나 현재까지 출시된 제품은 기술적 한계로 인한 성능저하 문제가 있어 실효성이 많이 떨어지는 상태이다.

현재 대부분의 DB 암호화 제품은 소프트웨어 방식을 적용하여 DB 내에 설치되며, 운용 시에 DB 서버의 성능을 상당히 떨어뜨리기 때문에 그 효용성이 부족하다. DB 암호화 기능을 적용할 때 암호화 기능을 적용하지 않은 상태와 비교해서 수 배 또는 수십 배 이상의 성능 저하를 가져온다. 성능 문제를 해결하기 위해 개발된 하드웨어 제품의 경우에도 초당 쿼리 수의 제한 때문에 ISP, 포털 및 게임 사이트 등에 적용할 만한 대용량 DB 암호화 솔루션이 없는 실정이다. 또한 국내외 DBMS 시장의 대다수를 점유하고 있는 오라클 DB의 경우, 8i 버전 이상에서는



<자료>: Gartner, 2006. 7.

(그림 1) 정보 보안 기술 Hype Cycle

암호화 모듈이 기본적으로 내부에 장착되어 있으나, 암호화 모듈을 이용하는 경우 성능저하가 매우 심하여 극히 제한적으로 적용되고 있다.

따라서, 고성능으로 DB 암호화를 수행하고 내부자 또는 외부에서의 개인 정보 불법 획득을 원천 차단할 수 있는 개인정보 보호 관련 기술 개발이 시급히 요구된다. 현재 세계 각국에서는 개인정보 유출을 방지하기 위하여 공공기관과 기업의 개인정보 보호를 강제하는 각종 법안을 시행중에 있다. 이와 같은 법안 시행 또는 법안 준비에 대응해서 국외 및 국내의 여러 공공기관과 기업에서 개인정보 DB 암호화 시스템을 도입하고 있다. 현재 미국의 은행 중 85%, 보험회사의 63%, 신용카드회사의 50%는 기밀/개인 정보 유출 방지 시스템을 적용하고 있다[2].

이와 같은 사회적, 법률적 배경을 바탕으로 데이터베이스 보안 시장은 빠르게 성장하고 있다. (그림 1)은 2006년 가트너에서 발표한 정보 보안 기술의 hype cycle이다[3]. 그림에서 보듯이 데이터베이스 보안 제품은 2~5년 이내에 시장 고점에 다다를 것이며 기술적인 면에서는 성장기에 접어들었다. 데이터베이스 활동 감시 제품은 2년 이내에 시장 고점이 예상되며 기술적으로는 성장기에 있음을 알 수 있다.

본 고에서는 현재 시장의 관심이 증가하고 있는 DB 보안 기술, 특히 암호화 관련 기술의 개요와 최신 동향에 대해서 중점적으로 다룬다. II장에서는 DB 암호화 기술의 개요에 대해서 살펴보고, III장에서는 최신 동향에 대해서 기술하며, 마지막으로 IV장에서 결론을 맺고자 한다.

● 용 어 해 설 ●

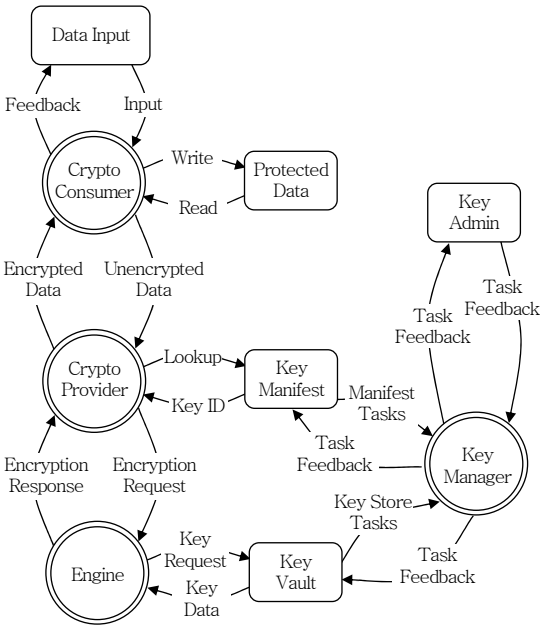
데이터베이스 보안(Database Security): 데이터베이스 보안이란 정보 보호 관리에 있어서 최종적이고 핵심적 관리 대상인 데이터베이스를 비인가된 변경, 파괴, 정보 누출을 발생시키는 사건으로부터 보호하기 위한 방법을 말한다. 데이터베이스 보안 방법은 크게 데이터베이스의 입출력 경로를 감시하는 접근제어/감사 제품과 데이터베이스 내부의 데이터 자체를 암호화하는 암호화 제품으로 분류된다.

Ⅱ. 데이터베이스 암호화 기술 개요

DB 보안을 위한 암호 시스템 구조는 총 7개의 논리 구성 요소로 구성된다(4개의 데이터 저장 과정과 3개의 처리 과정)[4]. 이 구성은 (그림 2)에 정리되어 있다. 7개의 구성요소는 다음과 같다.

- 암호화 엔진(Cryptographic engine): 암호화 기능을 수행한다.
- 키 저장소(Key vault): 안전한 키 저장소이다.
- 키 목록(Key manifest): 앨리어스, 패밀리, 상태, 엔진 등을 포함하는 키의 세부정보 트랙이다.
- 키 관리자(Key manager): 키 저장소와 키 목록 안에서 키를 관리한다.
- 보호 데이터(Protected data): 암호화를 통해서 보호되는 데이터들이다.
- 암호 소비자(Cryptographic consumer): 암복호화가 필요한 데이터를 관리하고 처리하는 주체이다.
- 암호 제공자(Cryptographic provider): 암호 엔진과 암호 소비자 사이를 연결하는 주체이다.

(그림 2)의 암호 인프라는 다음과 같이 동작한다.



(그림 2) 암호 인프라스트럭처

암호 소비자가 암호화가 필요한 데이터를 암호 제공자에 넘기면서, 해당 작업에 어떤 키 패밀리가 사용되는지를 확인한다. 암호 제공자는 키 목록을 사용해서 현재 키 패밀리 안에서 어떤 키가 사용되는지를 확인하고, 어떤 키 저장소와 엔진이 그 키에 할당되는지를 확인한다. 암호 제공자는 또한 초기 벡터(IV)와 같은 부수적인 데이터들을 수집하고 생성한다. 이 모든 정보들은 암호 엔진으로 넘어가고, 암호 엔진은 키 저장소에서 실제 키를 검색하고, 암호 작업을 수행한다. 암호화된 데이터를 암호 제공자로 반환하고, 암호 제공자는 키 ID와 그 외 필요한 정보들로 구성된 영수증(receipt)을 준비한다. 영수증은 암호화된 데이터와 함께 암호 소비자로 반환된다.

암호 소비자는 암호화된 데이터와 영수증을 DB에 기록한다. 나중에 암호 소비자가 데이터를 복호할 때, 암호 소비자는 암호화된 데이터와 영수증을 암호 제공자에 넘기기만 하면 된다. 암호 제공자는 영수증으로부터 필요한 정보를 추출하고, 키 목록에서 키가 유효한 상태로 있는지를 확인한다. 만일 모든 조건이 만족된다면 암호 제공자는 엔진으로 복호 요청을 넘기고 엔진은 복호 작업을 수행한다. 그리고 그 결과를 다시 암호 제공자를 통해서 암호 소비자로 넘겨준다.

암호화 기술은 많은 비밀을 보호해야 하는 문제를 몇 개의 비밀만 보호하면 되는 문제로 작게 만들어 준다.

이때 남게 되는 몇 개의 비밀이 암호화 키(cryptography keys)이다. 이런 암호화 키는 매우 중요하기 때문에 이 장의 나머지는 암호화 키의 속성에 대해서 할당하도록 한다. 키의 길이는 키를 사용하는 알고리듬이 제공하는 보안성에 결정적인 영향을 미치는 중요한 속성이다. 임의의 알고리듬이 주어졌을 때는 더 긴 키가 짧은 키에 비해서 더 강한 보안성을 제공하지만, 서로 다른 알고리듬이 사용될 때는 키 길이에 따라서 제공되는 보안성의 성능에 대해서 일반적인 기술을 할 수 없다. 예를 들어 알고리듬 A에서 사용되는 128비트 키가 항상 알고리듬 B에서 사용되는 128비트 키보다 보안성이 더 높다 또는 낮다고 얘기할 수 없다.

1. 키 분리

키 분리는 암호 키가 오직 한 가지 목적을 위해서만 사용되도록 요구하는 보안 개념이다. 키 분리의 기본 목적과 이익은 다음과 같다.

- 키에 접근할 필요가 있는 개체 수를 최소화한다.
- 키 교환 동안 다루어야 하는 데이터의 양을 관리 가능한 수준으로 유지한다.
- 주어진 특정 키로 암호화된 데이터 수를 줄임으로써 암호화가 깨졌을 때, 공격자에게 더 적은 정보가 노출되도록 한다.
- 만일 특정 키가 위태롭게 됐을 때, 입을 수 있는 손상을 제한한다.
- 서로 다른 종류의 데이터에게 다른 수준의 보안성을 제공할 수 있도록 한다.

암호 키는 보안성 유지를 위해 주기적으로 교체되어야 한다. 키를 교체하는 동안, 과거 키로 암호화된 모든 데이터들은 복호화된 후, 새로운 키로 다시 암호화해야 한다. 키 분리는 다양한 교체 스케줄링과 병렬 교환이 가능하게 함으로써 키 교체 작업을 관리 가능한 수준으로 유지시켜 준다.

암호화된 데이터를 해킹하는 방법 중의 하나인 “알려진 암호문 공격” 방법은 동일한 키로 암호화된 대량의 데이터들을 분석함으로써 성공할 수 있다. 키 분리는 각각의 키로 작은 분량의 데이터만을 암호화하는 것을 의미하므로 “알려진 암호문 공격” 방법을 거의 효과 없게 만들 수 있다.

여기서 암호키가 한 가지 목적 이상으로 쓰여서는 안된다는 문구가 애매할 수 있다. DB 암호화 시스템에서는 한 가지 목적이란 한 개의 데이터베이스로 생각할 수 있다. 즉 임의의 키는 한 개의 DB 이상에 걸쳐서 사용되어서는 안된다. 또한 구현 시에 키 분리는 키 패밀리와 키 범위(scope)를 통해서 실현된다.

2. 키 패밀리

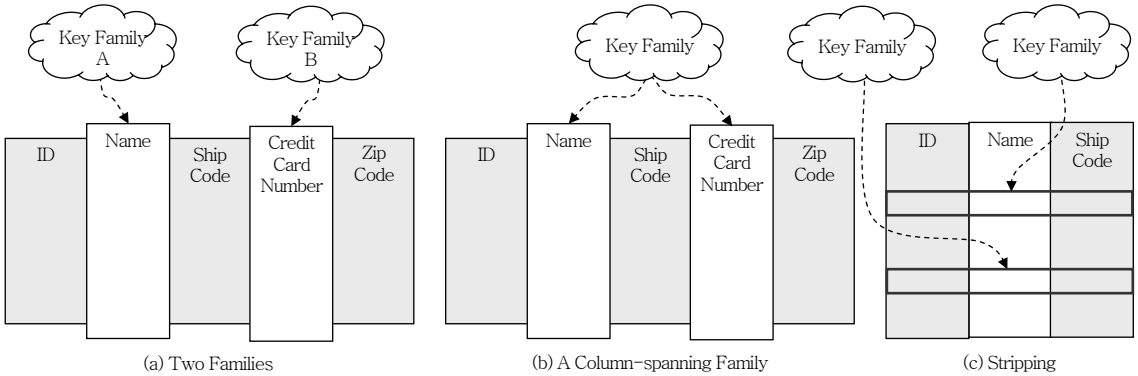
키 패밀리는 같은 집합의 데이터의 운용에 사용되는 키의 그룹이다. 예를 들어 하나의 키 패밀리가 신용 카드 번호를 위해 사용되면, 다른 또 하나의 키 패밀리가 의료 기록을 위해 사용되는 식이다. 키들은 그들의 패밀리 이름으로 라벨이 붙는다. 패밀리 안에서 키들은 그들이 어떻게 사용될지를 결정하는 특정한 역할을 갖고 있다. 예를 들어 한 패밀리 안에서 특정 순간에 오직 하나의 키만이 암호화를 위해서 사용될 수 있다. 복호화를 위해서는 여러 개의 키가 사용될 수 있다. 키 패밀리는 종종 더 이상 암복호화에 사용되지 않는 오래된 키를 보관할 수도 있다. 시간에 따른 역할의 변화는 키 라이프 사이클에 의해서 결정된다.

암호화가 필요한 각 컬럼에 대해서 각각 하나씩, 고유의 키 패밀리를 할당하는 것이 최선의 해결책이다. (그림 3a)는 두 개의 컬럼에 대해서 두 개의 패밀리가 할당된 경우를 보이고 있다. 그러나 컬럼과 키 패밀리의 일대일 관계는 행(row) 전체를 읽을 때는 각 키들이 모두 정렬되어서 복호화에 적용되어야 한다. 이는 성능 면에서 눈에 띄는 저하를 가져올 수 있다. 임의의 키 패밀리가 모든 컬럼을 처리하도록 하는 것이 성능 상의 오버헤드를 다소간 줄일 수 있다. (그림 3b)는 하나의 키 패밀리가 두 컬럼에 할당되는 상황을 보이고 있다. 또 다른 키 패밀리 할당 방법은 컬럼 한 개에 여러 개의 키 패밀리를 할당하는 방법이다. 이를 통해서 하나의 행이 한 키 패밀리로 암호화될 수 있다. 이를 스트리핑이라 한다. (그림 3c)는 스트리핑의 예를 보이고 있다.

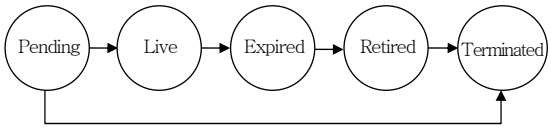
3. 키 라이프 사이클

키는 영속적이지 않다. 하나의 키가 많이 사용될수록 시스템의 보안성은 점점 더 약해진다. 따라서 키교체가 필요하다. 키를 교체할 때, 과거의 키로 암호화된 모든 데이터는 복호화된 후 새로운 키로 다시 암호화되어야 한다. (그림 4)는 키의 라이프 사이클을 보이고 있다. <표 1>은 라이프 사이클의 각 단계에서 허용되는 암호화 동작을 정리하고 있다.

- Pending: 활성화 날짜가 되면 키가 활성 상태가 된다. 패밀리 안에 여러 개의 pending 상태의 키가 있을 수 있다.
- Live: 현재 암복호화를 위해서 사용되는 키이다. 패밀리 안에 오직 한 개의 live 상태의 키만 있을 수 있다.
- Expired: 과거에 live 상태였으나 더 새로운 live 키에 의해서 밀려난 키로써 과거의 키로 암호화된 데이터를 복호화 할 때 사용된다. 패밀리 안에 여러 개의 expired 상태의 키가 있을 수 있다.
- Retired: 더 이상 사용되지 않는 키들로 패밀리 안에 여러 개의 retired 키가 있을 수 있다.
- Terminated: 키 저장소에서 삭제된 키들로 패밀리 안에 여러 개의 terminated 상태의 키가 있을 수 있다.



(그림 3) 키 패밀리 할당 방법



(그림 4) 암호화 키의 라이프 사이클

〈표 1〉 라이프 사이클 각 단계별로 가능한 암호기능

키 상태		
상태	암호화	복호화
Pending	No	No
Live	Yes	Yes
Expired	No	Yes
Retired	No	No
Terminated	No	No

4. Key Fatigue

임의의 키로 암호화되는 데이터의 양은 제한되어야 한다. 암호화 엔진이 같은 키를 오랫동안 이용할수록 정보가 유출될 확률이 점점 높아지기 때문이다.

5. Key Aliases와 Key Manifest

키 저장소와 암호화 엔진이 진짜 키에 접근할 수 있는 유일한 개체들이다. 그 이외의 모든 개체들은 키 앨리어스를 통해서 키를 언급하도록 한다. 각 키 앨리어스들은 시스템 전체에 걸쳐서 유일한 식별자이다. (키 ID와 키 저장소 ID로 구성된다.)

키 목록(manifest)은 키 저장소와 키 ID 간의 연결 트랙 정보를 유지한다. 또한 키 패밀리와 활성화 날짜, 키 상태 정보들을 관리한다. 모든 암호화 엔진이 모든 키 저장소에 대한 접근 권한을 갖고 있는 것이 아니기 때문에 키 목록은 어떤 엔진에 어떤 키가 할당되었는지에 대한 정보를 저장하고 있어야 한다.

Ⅲ. 데이터베이스 암호화 제품 동향

데이터베이스 보안 제품들은 크게 암호화 제품과 접근제어/감사 제품으로 구분될 수 있다. 본 절에서 제품 군에 따라서 특허 동향을 정리한다.

1. 데이터베이스 암호화 관련 특허

데이터베이스 암호화 제품은 크게 DBMS 자체에 포함된 암호화 기능과 써드 파티에서 제공하는 암호화 기능으로 구분된다. 써드 파티 제품은 다시 제품 구현 형태에 따라서 DBMS 시스템에 소프트웨어 에이전트 형식으로 장착되는 방식과 별도의 하드웨어 어플라이언스 형태의 제품으로 구분된다. 소프트웨어 방식으로 유명한 국외 업체로는 Protegrity와 nCipher가 있다[5]. 특히 Protegrity는 데이터베이스 보안과 관련된 논문과 특허를 매우 활발하게 발표하고 있다. <표 2>의 1, 2, 4, 5번이 소프트웨어 방식과 관련된 특허이다. 특히, 특허 1은 소프트웨어 방식의 암호화 제품이 키 관리를 위해서 하드웨어 방식의 HSM 장비와 혼용되는 형태를 고안하고 있는데 Protegrity는 이 형태를 이용하여 수위의 시장 매출을 보이고 있다. 하드웨어 어플라이언스 형태의 제품으로 국외 제품 평가에서 1위를 차지하고 있는 Ingrian 또한 여러 개의 관련 특허를 보유하고 있다(<표 2>의 3번 특허).

〈표 2〉 데이터 암호화 제품 관련 특허

번호	국가	출원일자	출원인	특허 제목
1	US	2000.11.16.	Protegrity Corporation	Combined Hardware and Software Based Encryption of Databases
2	US	2003.11.27.	nCipher Corporation Ltd.	Biometric Key Generation for Secure Storage
3	WO	2005.5.18.	Ingrian Networks, Inc.	Encrypted Table Indexes and Searching Encrypted Tables
4	WO	2000.11.27.	Protegrity Corporation	Method for Reencryption of a Database
5	EP	2004.12.2.	Protegrity Corporation	Database System with Second Preprocessor and Method for Accessing a Database

2. 데이터베이스 접근제어/감사 관련 특허

데이터베이스 접근제어/감사 제품 관련 특허는 <표 3>과 같다. <표 3>은 국내외 데이터베이스 접근제어/감사 제품을 출시한 회사들의 대표 특허를 정리하고 있다. 시장 조사 전문 기관인 Forrester에서는 DB 접근 제어/감사의 대표적인 전문 업체로 Application Security, IPLocks, Quest Software 등을 선정하였다[6]. 이들 회사의 제품 및 특허는 데이터베이스의 접속 선로를 통과하는 트래픽을 감시하고 사용자 질의 및 명령어들을 분석함으로써 보안 사고를 예방, 탐지하는 방법과 관련이 있다.

접근제어/감사 형태의 제품 및 특허 최신 동향은 실시간 감시 및 침입 차단 기능의 도입을 들 수 있다. 감사 기능은 처리해야 할 데이터가 방대하고 분석 모듈의 고지능화가 필요하기 때문에 실시간 감사를 제공하는 업체가 많지 않았지만, 외국 제품을 시작으로 실시간 감사 기능이 지원되는 추세이다. 그리고 기존 네트워크 침입 탐지 시스템과 유사한 개념으로 데이터베이스 시스템 자체에 대한 크래킹 시도를 차단할 수 있는 DB 침입 탐지 기능을 통합적으로 제공하기 시작하였다.

3. 검색 가능한 암호 기술

현재까지 시장에 출시된 DB 암호화 제품에 적용된 검색방법에는 매우 원시적인 접근방법이 사용되고 있는 반면, 학술적인 접근방법으로는 검색 가능한 암호 기술 방안이 있다. 현재 DB 암호화 제품의 최고 기술 이슈는 속도 저하가 크지 않으면서 컬럼 암호화 기능을 제공하는 것이지만, 인덱스 컬럼 암호화 시의 속도 저하 문제와 문자열 필드의 부분 검색 문제 등이 해결되지 않고 있다. 인덱스 컬럼 암호화 시의 속도 저하 문제는 정렬 순서가 유지되는 암호화 기술(order preserving encryption)을 통해서 해결이 가능할 것이고, 문자열 필드의 부분 검색 문제는 검색 가능한 암호화 기술(searchable encryption)로 해결이 가능할 것으로 보인다. <표 4>는 지금까지 발표된 검색 가능한 암호 기술 관련 논문을 정리하고 있다.

DB 보안시스템을 위한 암호 알고리듬 고려 시에 문제 정의에 따라서 여러 가지 후보 알고리듬을 고려할 수 있다. 첫번째로 DB를 신뢰할 수 있느냐 없느냐에 따라서 접근 방법이 달라진다. DB를 신뢰할 수 있는 경우에는 단순히 외부로 자료가 유출되는

〈표 3〉 데이터베이스 접근제어/감사 제품 관련 특허

번호	국가	출원일자	출원인	특허 제목
1	US	2003.2.24.	IPLocks, Inc.	Method and Apparatus for Monitoring a Database System
2	US	2001.2.12.	Quest Sotfware, Inc.	System and Method for Reconciling Transactions between a Replication System
3	KO	2004.4.3.	피앤피시큐어	데이터베이스 감시 및 보안 방법 및 장치
4	KO	2004.1.9.	바넷정보기술	3-Tier 구조 기반의 데이터베이스 접근 통제 시스템 및 방법

〈표 4〉 검색 가능한 암호 기술 관련 논문 리스트

연도	저자	논문 제목
2000년	Dawn Xiaodong Song, David Wagner	Practical Techniques for Searches on Encrypted Data
2005년	Abdalla, Bellare, Catalano	Searchable Encryption Revisited: Consistency Properties, Relation to Anonymous IBE and Extensions
2006년	Mihir Bellare, Alexandra Boldyreva	Efficiently-Searchable and Deterministic Asymmetric Encryption
2006년	Reza Curtmola, Juan Garay	Searchable Symmetric Encryption: Improved Definitions and Efficient Constructions

경우에 대비해서 자료를 암호화해서 저장하는 간단한 방식만 고려하면 되지만, DB를 신뢰할 수 없는 경우 DB에 질의와 데이터를 노출하지 않기 위해 여러 가지 방안을 고려해야 한다.

<표 4>의 논문 중 Dawn Xiaodong Song의 논문은 검색 가능한 암호화 기법을 다룬 대표적인 논문 중의 하나이다. Song의 논문은 DB를 신뢰할 수 없는 경우에 사용할 수 있는 암호화 문제를 풀고 있지만 검색 가능 암호화 기법에서 쓸 수 있는 좋은 아이디어를 포함하고 있기 때문에 많이 인용되고 있다[7]. 논문에서 제안하고 있는 방안을 간단히 정리하면 다음과 같다. 우선 문제를 정의한다.

- Alice가 자신의 문서를 신뢰할 수 없는 서버 Bob에 저장하는 경우 Alice는 문서를 암호화하여 저장하길 원한다.
- 문서는 단어들(words)로 나눌 수 있으며, 본 논문에서 각 단어는 동일한 길이를 갖는 것으로 가정한다.
- Alice는 low-bandwidth network connection을 가지기 때문에 찾으려는 단어를 포함한 문서만을 추출하기를 원한다.

이때 Alice는 자신의 질의 내용을 Bob에게 들키지 않으면서 자신이 원하는 단어를 포함한 문서를 Bob의 DB에서 찾고 싶다. 이 문제를 해결하기 위해 암호문에 특정 연산 수행을 통하여 DB에 저장된 문서(암호문)가 특정 단어 W를 포함하는지 Bob이 결정할 수 있도록 하는 기법의 설계가 필요하다. 단 이때 Bob은 어떠한 정보도 알지 못해야 한다.

저자는 해결 방안을 설명하기 위해 네 가지 단계를 거치고 있다(basic 스키마, controlled searching 스키마, support for hidden search 스키마, final 스키마).

첫번째, 가장 기본적인 검색 방법은 모든 단어(W_i)를 암호화 키(K_i)를 이용해서 암호화해서 DB에 저장한 뒤 Bob에게 찾고자 하는 단어와 단어의 위치, 그리고 해당 위치의 키를 보내거나, 단어와 암호화 키 모두를 전달하는 방법이다. 이 방법은 찾고자 하는 단어의 위치에 해당하는 키를 정확하게 보내거나 키 모두를 보내야 한다는 점에서 실용성이 없다.

두번째 단계는 모든 키를 다 보내야 하는 첫번째 단계의 단점을 보완하기 위해서 암호화 단계에서 W_i마다 고유한 키를 갖도록 하는 방법이다. 사용자는 W_i와 해당 단어에 고유한 키만 Bob에게 보내면 되기 때문에 위치를 정확히 알아야 하거나, 키 모두를 보내야 하는 첫번째 문제는 해결이 된다. 하지만 Bob에게 찾고자 하는 질의 단어 W_i가 공개되는 문제는 여전히 남는다.

세번째 단계는 질의 단어가 공개되는 문제를 해결하고 있다. 우선 암호화 단계에서 Pseudo Random Permutation 함수 E를 이용해서 W_i를 암호화한다. 이 암호화 결과를 다시 K_i를 이용해서 암호화한다. Alice에서 Bob으로 질의를 던질 때는 W_i를 보내는 것이 아니라 함수 E에 의한 1차 암호화 결과를 보내므로 Bob에게 질의 단어를 누출시키지 않는 장점이 있다. 하지만 Alice는 Bob이 검색 결과를 반환해 주더라도 이 반환 결과를 다시 평문으로 복호화할 수 없는 단점이 있다.

네번째 단계는 마지막 단계로 세번째 단계에서 풀지 못했던 검색 결과의 복호화 문제를 해결한다. 세번째 단계에서 복호화가 불가능했던 이유는 Bob이 1차 암호화 결과 $E(W_i)$의 마지막 m bit 부분을 알 수 없었기 때문이므로 이 부분을 알 수 있도록 함수 E를 이용하는 1차 암호화 단계에서 $E(W_i)$를 두 부분(L_i와 R_i)으로 분리함으로써 마지막 m bit를 알 수 있도록 한다. 이를 통해서 검색 결과의 복호화가 가능하다.

네 가지 접근 방법 모두 단어 W가 고정 길이라는 가정을 하고 있다는 단점이 있고 이를 해결하기 위해서 패딩 방안과 길이 필드를 추가하는 방안 등이 있을 수 있다.

정렬 순서가 유지되는 암호화 기술과 검색 가능한 암호 기술 개발은 아직까지는 학술적인 논의 단계에 있고 제품 기술로 적용하기에는 해결해야 할 문제가 많이 남아 있는 상태이다. 하지만 현재 데이터베이스 암호화 제품들이 직면하고 있는 많은 난제

들을 해결해 줄 수 있다는 점에서 앞으로 지속적인 연구 개발이 있어야 할 것이다.

Ⅳ. 결론

지금까지 최근 각광 받고 있는 데이터베이스 보안 기술의 개요와 최신 제품 기술 동향에 대해서 살펴보았다. 데이터베이스 보안 분야는 개인정보보호법의 제정 움직임으로 인해서 급속한 시장 확대가 예상되는 분야이다. 또한 기존 제품들이 아직 성능 개선 문제 등의 시장의 요구를 만족시키지 못하고 있기 때문에 많은 연구 개발의 가능성이 남아 있는 분야이기도 하다. 암호화 기능을 적용하는 경우 기존의 데이터베이스 운용 속도 저하 문제뿐만 아니라 검색 가능한 암호 기술 등에 대한 학술적 검증이 남아 있기 때문에 향후 국내외 학계, 연구계, 산업계에서 많은 연구와 제품 개발이 요구되는 바이다.

● 용 어 해 설 ●

DES: DES (Data Encryption Standard)는 블록 암호의 일종으로 메시지를 보내려는 사람과 받는 사람이 56비트의 동일한 열쇠를 가지고 있어야 한다. 보내려는 사람은 열쇠를 이용하여 데이터를 암호화하여 안전하지 않은 채널로 전송하면, 받는 사람은 암호화된 데이터를 받아 보낸 사람의 열쇠와 동일한 열쇠를 이용하여 데이터를 복호화한다. DES의 역사는 1970년대까지 거슬러 올라간다. 1972년에 미국 NBS (National Bureau of Standards, 오늘날의 NIST)는 암호 기술의 필요성을 절감하고 미국 정부 규모의 표준적인 암호 알고리듬을 개발하기로 했다. 이에 1974년 8월 27일, IBM에서 알고리듬을 제안했고, 이것을 수정하여 DES로 정했다.

약 어 정 리

AES	Advanced Encryption Standard
CBC	Cipher Block Chaining mode
DES	Data Encryption Standard
ECB	Electronic Code Book mode
FIPS	Federal Information Processing Standard
GLBA	Gramm-Leach-Bliley Act
HIPAA	Health Insurance Portability and Accountability Act
HSM	Hardware Security Module
IV	Initialization Vector
SHAs	Secure Hash Algorithms

참 고 문 헌

[1] 매일경제신문, "盧대통령 주민번호 인터넷 떠돈다," 2006. 6. 27.

[2] Rich Mogull and Ray Wagner, "Data Security Enters the Spotlight," Gartner, 24 Oct. 2005.

[3] Rich Mogull, Ray Wagner, John Girard, and Vic Wheatman, "Hype Cycle for Data Security, 2005," Gartner, 12 July 2005.

[4] Kevin Kenan, "Cryptography in the Database–The Last Line of Defense," Addison-Wesley, Oct. 2005.

[5] Noel Yuhanna, "The Forrester Wave: Database Encryption Solutions, Q3 2005," Forrester, 8 Aug. 2005.

[6] Rich Mogull, "Database Activity Monitoring Is a Viable Stopgap to Database Encryption for the Payment Card Industry Data Security Standard (and Beyond)," Gartner, 3 July 2006.

[7] Dawn Xiaoding Song, D. Wagner, and A. Perrig, "Practical Techniques for Searches on Encrypted Data," *Proc. IEEE Symp. on Security and Privacy*, 2000.