서비스로의 데이터베이스에서 빈도수 기반의 추론공격 방지를 위한 인덱싱 기법

(An Indexing Method to Prevent Attacks based on Frequency in Database as a Service)

정 강 수 † 박 석 ††
(Kangsoo Jung) (Seog Park)

요 약 데이터의 소유권이 외부 업체로 이양되는 DaaS 모델은 신뢰할 수 없는 서비스 제공자에 의한 데이터 누출의 위험이 존재한다. 본 논문은 복수 개의 칼럼으로 이루어진 암호화 된 테이블의 인덱스를 통해 발생할 수 있는 추론 공격을 분석하여 b-anonymity라는 개념을 도입함으로써 이에 대한 해결책을 제시한다. 또한 데이터의 인덱싱에 R+-Tree 기법을 사용함으로써 인덱스를 사용할 때 발생할 수 있는 오탐률 (False-positive)을 최소화하여 데이터 처리의 효율성을 보장하였다.

키워드 : DaaS, 프라이버시, 익명화

Abstract DaaS model that surrogates their data has a problem of privacy leakage by service provider. In this paper, we analyze inference attack that can occur on encrypted data that consist of multiple column through index, and we suggest b-anonymity to protect data against inference attack. We use R+-tree technique to minimize false-positive that can happen when we use an index for efficiency of data processing.

Key words : DaaS, Privacy, Anonymization

1. 서 론

컴퓨터 사용 인구의 증대와 인터넷 인프라의 발전은 전산화 된 데이터의 폭발적인 증가를 가져왔다. 이에 따른 비용 증가 문제를 해결하기 위해 등장한 것이 DaaS (Database as a Service)모델이다[1].

DaaS 모델이 데이터의 급증에 따른 문제를 해결할 수 있는 대안으로 떠오르면서 데이터의 누출에 따른 프라이버시 위협이 부각되었다. DaaS 모델에서 서비스 사용자의 데이터 소유권은 서비스 제공자 측에 이양된다. 따라서 적절한 보호 장치가 없다면, 악의를 지닌 관리자에 의해 서비스 사용자의 데이터가 악용될 우려가 있다.

서비스 제공자에 의한 데이터의 누출을 막기 위해 가장 널리 쓰이는 기술은 암호화이다. 그러나 데이터에 암호화를 적용하면, 검색 성능 저하 및 기존의 데이터베이스에서 제공하는 기능을 사용하지 못 하는 문제가 발생한다. 이에 데이터 암호화에 따른 성능 저하 문제를 해결하고자 하는 시도가 활발히 이루어져왔다[2]. 암호화된 데이터 상에서 검색 성능을 향상시키기 위해 일반적으로 사용되는 방법은 인덱스를 생성하는 것이다[1,3,4].

본 논문은 암호화 된 데이터 상의 데이터 연산을 수행하기 위한 인덱스 작성 기법에 초점을 맞춘다. 특히 복수개의 칼럼으로 이루어진 테이블에 인덱스를 적용할 경우 발생할 수 있는 추론 공격을 분석하고 이에 대한 해결책을 제시한다. 또한 다차원 인덱스 작성에 효율적인 공간 트리를 사용하여 서로 상충하는 안전성과 효율성 간의 균형을 고려한 인덱스 작성 기법을 제안한다.

2. 관련 연구

2.1 버킷 기반 인덱스 작성 기법

Hacigumus는 2002년 SIGMOD에서 발표한 논문[1]에서 버킷에 기반한 인덱스를 사용해 질의를 처리하는 기법을 제안하였다. 기본적인 알고리즘은 다음과 같다. 그림 1과 같은 서비스 사용자의 데이터는 표준 암호 알고리즘에 의해 암호화된 후 Etuple이라는 칼럼에 저장된다. 각각의 데이터는 지정된 크기로 나뉘어진 파티션 함수를 기준으로 해당하는 버킷 ID와 매핑된다.

버킷 ID는 인덱스 값으로서 인덱스 칼럼에 저장되어 그림 2와 같이 Etuple과 함께 서비스 제공자 측에 보관되어 데이터 연산에 사용된다.

서비스 사용자가 원하는 질의에 대한 값을 얻기 위해

· 본 연구는 한국과학재단 세계수준의 연구중심대학(WCU)육성사업(R33-2008-000-10110-0)지원으로 수행되었음
· 이 논문은 제36회 추계학술발표회에서 '서비스로의 데이터베이스에서 빈도수 기반의 추론 공격 방지를 위한 인덱싱 기법'의 제목으로 발표된 논문을 확장한 것임

† 학생회원 : 서강대학교 컴퓨터공학과
azure84@naver.com
†† 종신회원 : 서강대학교 컴퓨터공학과 교수
spark@sogang.ac.kr

논문접수 : 2009년 12월 24일
심사완료 : 2010년 7월 6일

정보과학회논문지: 컴퓨팅의 실제 및 레터 제16권 제8호(2010.8)

eid	ename	salary	addr	Did
23	Tom	70k	Maple	40
860	Mary	60k	Main	80
320	John	50k	River	50
875	Jerry	55k	Hoppewell	110

그림 1 서비스 사용자의 데이터[1]

etuple	eid^s	ename^s	salary^s	addr^s	did^s
1100110011110010...	2	19	81	18	2
1000000000011101...	4	31	59	41	4
1111101000010001...	7	7	7	22	2
1010101010111110...	4	71	49	22	4

그림 2 인덱스 칼럼[1]

서는 서비스 사용자 측에 보관된 파티션 함수를 사용하여 쿼리를 변형한다. 서비스 제공자 측은 변형된 쿼리에 대해 보관된 인덱스를 이용하여 데이터 연산을 수행한다. 이 기법은 숫자와 문자에 대하여 모두 적용 가능하며 일치 검색과 범위 연산을 모두 수행할 수 있다는 장점 을 지닌다. 그러나 일정 범위에 같은 ID를 부여하는 버 킷 기반 인덱스의 특성상 실제로 원하는 검색 결과가 아님에도 원하는 결과값으로 간주되어 결과값에 포함되는(False-positive) 경우가 발생할 수 있다.

2.2 배포된 데이터에 대한 프라이버시 노출 방지 기술

통계 기관에선 연구나 통계 분석 등의 목적을 위해 수집된 데이터를 배포한다. 이와 같은 데이터에는 개인의 민 감한 정보가 포함되어 있다. 따라서 이름이나 주민등록 번호와 같이 명시적인 개인 신원 정보를 암호화하거나 삭제한 후 배포하는 탈 식별화 방법이 사용되었다. 그러나 탈 식별화 방법은 배포된 데이터와 공통 속성을 포 함하는 외부 데이터와 연결, 취합함으로써 개인 정보를추론할 수 있는 추론 공격의 위험을 내포한다.

이와 같은 추론 공격을 방지하기 위해 k-anonymity [4]를 비롯한 익명화 기법이 연구되었다. 익명화 기법은 일부 데이터를 변경하여 비구별성을 얻는 연구 분야로 추론 공격을 방지하여 프라이버시를 보호하는 기법으로 사용되고 있다.

k-anonymity는 배포된 데이터에 대해 k개의 동일한 튜플을 유지함으로써 정확한 추론의 확률을 1/k로 만드는 기법이다. 그러나 이 기법은 익명화의 과정에서 정보의 손실을 가져오므로 연구에 쓰일 데이터의 활용도를 감소시키는 단점이 있다. 따라서 정보 손실률을 최소화하면서 데이터가 추론될 확률을 낮추는 기법과 새로운 추론 공격에 의한 데이터 누출을 방지하는 기법에 관해 많은 연구들이 이루어졌다.

2.3 기존 연구와의 관계 및 기존 연구의 문제점

본 연구는 암호화된 데이터베이스 상에서의 데이터 연산을 위한 기법으로 버킷에 기반한 인덱스에 초점을 맞춘다. 이는 버킷을 사용한 인덱스 작성 기법이 일치, 범위 연산에 대한 쿼리를 지원할 수 있으며 상대적으로 간단한 방법으로 인덱스를 생성하기에 실제 시스템에서 구현하기 유리한 이점을 지니기 때문이다. 본 연구는 배포된 데이터에 대한 익명화 기법을 DaaS 모델에서의 인덱스에 적용함으로써 버킷 기반 인덱스가 지닌 추론 공격에 대한 문제점을 보완한 인덱스 작성 기법을 제안한다.

3. DaaS 모델에서의 프라이버시

3.1 DaaS 모델

DaaS 모델은 기업이나 조직이 별도의 데이터 센터를 직접 구축하는 대신 데이터 처리를 전문적으로 수행하는 외부 업체를 이용하는 것이다. DaaS 모델의 기본 구조는 그림 3과 같다.

본 연구에서는 DaaS 모델의 요구 사항을 두 가지로 파악한다.

첫 번째는 인덱스를 이용하여 암호화된 데이터에 대한 연산을 수행할 때 False-positive가 발생할 수 있다는 문제점이다. 따라서 실제 데이터를 잘 반영한 인덱스를 생성함으로써 False-positive의 발생을 최소화하는 것이 필요하다.

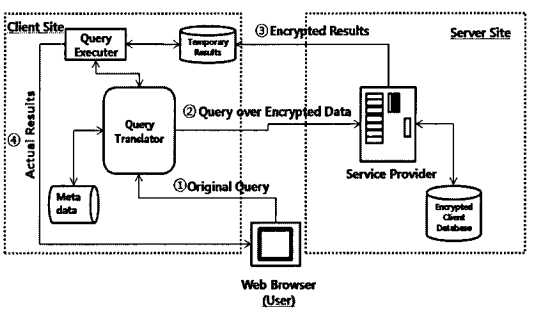

그림 3 DaaS 모델 기본 구조[1]

두 번째는 인덱스를 통한 추론 공격의 위험성이다. 인덱스는 실제 데이터를 반영하여 작성되므로 실제 데이터의 추론 가능성을 최소화한 인덱스 생성이 필요하다. 이 두 가지 요구사항은 서로 상충된다. 따라서 이에 대한 균형을 고려하여 가장 적절한 수준의 데이터 안전성과 효율성을제공하는 인덱스를 생성하는 것이 필요하다.

3.2 버킷 기반 인덱스에서의 데이터 누출

버킷 기반 인덱스는 그 특성상 버킷의 아이디와 실제 데이터 간의 상관관계를 통해 데이터의 누출이 가능하

다. [3]의 연구에서는 인덱스를 통한 데이터 누출을 고려한 인덱스 작성 기법을 제안하여 추론에 의한 공격을 방지하려 했다. 이 외에도 여러 연구들에서 인덱스를 통한 데이터 누출에 대한 해결책을 제시하였으나 이들 연구들은 하나의 칼럼만이 존재하는 일차원 데이터(Single-dimension)에 대한 추론 공격만을 고려하였기에 복수 개의 칼럼이 존재하는 다차원 데이터(Multi-dimension) 환경에서 발생할 수 있는 데이터 누출이 발생할 수 있는 한계를 지닌다. 또한 버킷의 인덱스의 빈도수를 통한 추론 공격에 의해 버킷 자체의 정보가 노출될 수 있는 위험이 존재한다.

3.3 추론에 의한 데이터 누출과 그 보호 기법

3.3.1 다차원 데이터 환경에서의 추론 공격

일차원 데이터 환경에서 각각의 칼럼들이 k-anonymity를 지킬 경우 추론 공격에 대해 정확한 값이 추론될 확률은 각 튜플 당 1/k 이하로 떨어진다. 그러나 하나의 칼럼이 아닌 2개 이상의 칼럼을 결합하여 추론 공격을 시도할 경우 각각의 칼럼에 대해 k-anonymity가 유지되더라도 튜플의 정확한 값이 추론될 수 있다. 이에 대해 예를 들어 설명한다.

그림 4는 2개의 칼럼으로 구성된 테이블과 이에 대한 인덱스이다. 그리고 공격자는 그림 5와 같이 실제 데이터의 빈도수 정보를 알고 있다 가정한다.

각각의 칼럼은 2-anonymity을 적용하여 최소 2개 이상의 같은 인덱스 값을 지니는 튜플이 존재하므로 튜플의 정확한 값에 대한 추론은 불가능하다.

나이	질병
23	독감
25	결핵
28	고혈압
31	당뇨
33	당뇨
36	고혈압
37	당뇨
36	고혈압
42	에이즈
45	에이즈
47	암
49	암

나이	질병
1	A
1	A
1	B
2	B
2	B
2	B
2	B
2	B
3	C
3	C
3	C
3	C

그림 4 2개의 칼럼을 지니는 데이터

(나이, 질병)
{(23,독감), (25,결핵), (28,고혈압), (31,당뇨), (33,당뇨), (36,고혈압), (37,고혈압), (38,고혈압),(42,에이즈), (45,에이즈), (47,암), (49,암)}

그림 5 나이와 질병에 대한 실제 데이터의 빈도수 정보

나이	질병
1	A
1	A
1	B
2	B
2	B
2	B
2	B
2	B
3	C
3	C
3	C
3	C

나이	질병
1/2	1/2
1/2	1/2
1	1
1/5	1/5
1/5	1/5
1/5	1/5
1/5	1/5
1/5	1/5
1/4	1/4
1/4	1/4
1/4	1/4
1/4	1/4

그림 6 Multi-dimension의 추론 확률

그러나 나이 칼럼과 질병 칼럼 모두를 고려하여 추론 공격이 이루어질 경우 유일하게 구별되는 튜플이 발생한다. 그림 6의 좌측은 두 개의 칼럼을 결합한 인덱스 테이블이다. 두 개의 칼럼을 동시에 고려한 인덱스 값의 빈도수는 {1,A}=2, {2,B}=5, {3,C}=3, {1,B}=1이다. 두 개의 칼럼을 결합하여 인덱스의 빈도수를 계산했을 때 {1,B}에 해당하는 튜플의 빈도수가 1이므로 2-anonymity가 위반된다. 이를 실제 데이터의 빈도수와 비교한다면 해당 튜플의 정확한 값이 추론될 수 있다. 그림 6의 우측은 해당 튜플의 값이 추론될 수 있는 확률을 나타낸다.

3.3.3 다차원 데이터 환경에서 추론 공격에 대한 보호

다차원 데이터 환경에서 추론 공격에 의한 데이터 누출을 막기 위해서는 다차원 데이터 환경에 적합하게 k-anonymity를 적용한 인덱스 작성 기법이 필요하다. 본 연구에서는 인덱스 작성을 위해 기존의 다차원 인덱싱 기법 중 공간 트리 기법을 적용하였다. 이에 대한 설명은 4장에서 이루어진다.

3.4 빈도수를 이용한 추론에 의한 데이터 누출 방지

3.4.1 빈도수 정보를 통한 추론에 의한 버킷 정보 누출

3.3.1절에서 두 개의 칼럼의 결합으로 유일하게 튜플이 식별되는 추론에 의한 데이터 누출의 예를 보였다. 이에 대한 해결책은 다차원 데이터에서도 k-anonymity를 유지하도록 강제하는 것이다. 본 절에서는 k-anonymity를 적용하여도 빈도수에 의한 추론으로 버킷의 범위 정보가 노출될 수 있는 위험성을 보인다. 그림 6에서 {2,B}에 해당하는 인덱스 값의 빈도수는 5이다. 만약 이 데이터에 대한 익명화 기준값 k가 5 이하라면 이 데이터는 충분히 익명화되어 데이터 누출이 방지되었다고 이야기할 수 있다. 그러나 전체 데이터에서 5에 해당하는 빈도수는 유일하며 따라서 빈도수에 관련된 연관 데이터와의 비교를 통해 해당 버킷의 정보가 누출될 수 있다.

따라서 본 연구에서는 이에 대한 보호 기법을 제안한다. 본 연구에서 제안하는 기법은 가상 튜플(counterfeit tuple)을 사용하여 실제 데이터의 빈도수와 서비스 제공

자 측에 저장되는 인덱스 값의 빈도수를 불일치 시켜 실제 데이터의 빈도수와 인덱스 값의 빈도수의 비교를 통한 버킷의 범위 정보의 노출을 방지하는 것이다.

정의 1. 가상 튜플(counterfeit tuple)

실제 데이터 연산에 사용되지는 않으나 서버 측에 존재함으로써 빈도수에 의한 추론을 방지하는 역할을 하는 튜플이다. 가상 튜플의 정보는 클라이언트 측에 저장되며 이를 통해 가상 튜플의 삽입과 삭제를 조정할 수 있다. ■

3.4.1절에서 제시한 예에서 문제시되는 부분은 테이블에서 버킷 안에 존재하는 튜플의 개수가 5개인 버킷이 유일하게 존재하여 버킷의 정보가 식별된다는 것이었다. 따라서 빈도수 4에 해당하는 버킷과 빈도수 1에 해당하는 버킷에 가상 튜플을 1개씩 삽입함으로써 서비스 제공자 측에 저장된 버킷들 중 튜플의 개수가 5인 버킷을 2개로, 2인 버킷 역시 2개로 만들어 실제 데이터의 빈도수에 의한 버킷의 범위 정보 노출을 방지하는 것이다.

즉 빈도수에 의해 유일하게 구별되는 버킷을 구별하지 못 하게 하기 위해, k-anonymity 기법을 버킷에 적용시켜 버킷 자체의 비구별성을 추구하는 것이 본 연구에서 제안하는 기법이다. 이를 b(bucket)-anonymity라 정의한다. b-anonymity의 정의는 아래와 같다.

정의 2. b-anonymity

b-anonymity는 버킷 기반 인덱스 작성 기법에서 어떤 하나의 버킷이 가진 튜플의 개수가 최소 b-1개 이상의 다른 버킷이 지닌 튜플의 개수와 일치해야 함을 의미한다. 이를 통해 튜플의 빈도수에 구별될 수 있는 버킷의 식별 가능성을 1/b로 만든다. ■

4. 트리 구조를 통한 인덱스 작성 기법

4.1 다차원 데이터에 대한 인덱스 작성

다차원 데이터는 일차원 데이터와 다른 요구사항을 지니기에 기존의 인덱스 작성 기법과 다른 기법이 필요하다. 이를 위해 데이터베이스 커뮤니티는 다차원 인덱스 작성 기법에 대한 연구를 진행하여왔다[5]. 본 연구는 다차원 인덱스 작성 기법 중 공간 트리를 사용한 인덱스 작성 기법을 DaaS 환경에서 효율적인 인덱스 연산과 데이터 누출 방지를 위한 인덱스 작성을 위해 적용한다. 사용하는 공간 트리는 R-트리의 변형인 R^+-트리를 사용한다.

4.2 R^+-트리를 사용한 인덱스 작성 기법

R^+-트리를 사용하여 인덱스를 작성하는 알고리즘은 다음과 같다.

인덱스의 작성은 기존의 R^+-트리에서의 인덱스 작성 기법을 그대로 사용한다. 버킷은 R^+-트리에서의 MBR로 간주되고 각 버킷에 포함된 최소 튜플의 개수는 k-anonymity를 유지하기 위해 k개 이상이어야 하며, 최대 튜플의 개수는 효율성을 위해, max{2k-1, 3k-5}개의 튜플로 제한한다. k-anonymity를 만족시키는 인덱스를 작성한 후엔 b-anonymity에 대한 적합성 검사를 수행한다. 위반 사항이 발생한 버킷에 대해서는 데이터 사용자 측에 저장되어 있는 빈도수에 대한 정보를 바탕으로 가상 튜플을 삽입하여 b-anonymity를 유지하게 한다. 가상 튜플을 삽입한 후에는 다시 b-anonymity에 대한 검사를 수행하여 가상 튜플의 삽입으로 인해 b-anonymity 제약 조건을 위반하였는지에 대한 검사를 수행한다. 위반 사항이 발생하지 않았다면 표준 암호화 알고리즘을 사용해 데이터를 암호화한 후 s인덱스 값과 함께 서버 측에 이를 저장한다.

5. 실 험

5.1 실험 환경

본 실험을 위해 사용한 데이터는 UC Irvine machine learning repository의 성인 데이터를 사용하였다. 실험을 위해 데이터의 속성 중 Age와 Zip code의 2개의 속성을 사용하여 인덱스 작성 기법에 대한 실험을 실시하였다. 데이터 중 "?"의 값을 갖거나 값이 비어있는 튜플을 제외하고 남은 튜플 중 10,000개를 대상으로 실험을 사용하였다. 실험에 사용한 기기의 사양은 다음과 같다.

- Operation System : Microsoft Window XP Professional Version 2002
- CPU : Intel Pentium 3.20GHz
- RAM : 1.00GB
- Programming Language : Java (jdk 1.6.0)

5.2 실험 결과 및 분석

그림 7의 결과는 버킷의 범위를 일정하게 정한 equi-width와, 버킷 내의 튜플의 개수를 일정하게 유지하는 equi-depth을 사용한 인덱스 생성 기법을 제안 기법과 비교하여 얻어진, k-anonymity를 위반한 튜플 개수에 대한 결과이다. 실험 결과 equi-width와 equi-depth 모두 일차원 데이터에 대해 k-anonymity를 검사하였을 때는 위반 사항이 없음에도 불구하고 다차원 데이터에 적용했을 때는 이와 같은 위반 사항이 발견되었다. 이는 기존의 기법이 다차원 데이터 환경에 대한 고려를 하지 않았기 때문이다. 다차원 데이터 환경을 고려한 본 제안 기법에서는 한 건의 위반 사항도 발견되지 않았다.

그림 8의 결과는 본 연구에서 지적한 버킷의 식별가능성에 대한 실험 결과이다. avcs는 Average cell size의 약자로 버킷의 평균 크기를 의미한다. 버킷의 평균 크기는 전체 튜플의 개수/생성된 버킷의 수로 구해지며 버킷의 평균 크기가 클수록 추론의 위험은 적어진다. 실험 결과 avcs가 100, b=4의 경우 equi-width의 경우

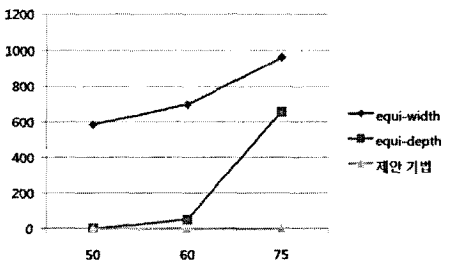

그림 7 k-anonymity 위반 튜플 개수(k=75)

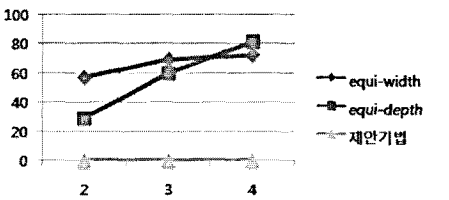

그림 8 b-anonymity 위반 버킷 개수(avcs=100,50,25)

약 72%, equi-depth의 경우 81%의 b-anonymity를 위반한 버킷이 발견되었다. 반면 제안 기법은 b-anonymity의 위반 사항이 발생할 경우 가상 튜플을 삽입해 이에 대한 위반을 방지하였으므로 한 건의 위반 사항도 발견되지 않았다.

또한 총 4가지 기법에서의 False-positive에 대한 실험을 수행하였다. equi-width와 equi-depth 외에 b-anonymity를 적용했을 때와 적용하지 않았을 때에 대한 실험도 수행하였다. b-anonymity를 적용하지 않은 것을 제안 기법, b-anonymity를 적용한 것을 제안 기법2라 명명하였다. 사용한 계산식은 그림 9와 같다.

TFP는 Total False Positive의 약자로 전체 튜플에서 발생하는 False-positive에 대한 측정값을 의미한다.

그림 10은 각 기법에 대한 False-positive이다. 단위는 100,000을 기준으로 하였다. 제안 기법은 실험 결과에서 보여지듯 11% 가량 향상된 성능을 보였다. 이는 다차원 데이터에 대한 인덱스 작성에서 공간 트리의 사용이 기존의 인덱스 작성 기법보다 더 효율적임을 보여준다. 또한 b-anonymity가 적용된 제안 기법2 역시 제안 기법과 거의 유사한 성능을 보였다. 이는 제안 기법이 성능을 저하시키지 않으면서 보다 철저히 데이터 누출을 방지할 수 있음을 의미한다.

$$TFP = \sum_{\forall \epsilon Q} (|R^{S}_{T(q)}| - |R_q|) \approx N_b \times F_b$$

(R=질의 결과값, S=서버 측에 저장된 값, T(q)=인덱스 질의, q=원 질의, N=버킷의 크기, F=튜플의 개수, b=버킷ID)

그림 9 False-positive 측정식

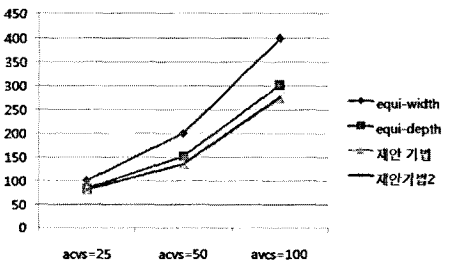

그림 10 False-positive

6. 결 론

본 연구는 DaaS 모델을 대상으로 성능과 프라이버시 간의 상충을 고려한 인덱스 작성 기법에 대한 연구를 수행하였다. 다차원 데이터 환경에서 발생할 수 있는 데이터 누출에 대해 고려하였으며 이에 대한 보호 기법을 제시하였다. 나아가 DaaS 모델에서 인덱스가 가진 특성을 분석하고 이를 바탕으로 가상 튜플을 사용하여 빈도수를 통한 데이터 누출 방지가 가능한 기법을 제안하였다. 또한 다차원 데이터의 인덱스 작성을 위하여 공간 트리를 적용하여 기존에 사용되었던 인덱스 작성 기법에 비해 더 효율적인 인덱스 작성 기법을 제안하였다.

참 고 문 헌

[1] Hakan Hacigumus, Bala Iyer, Chen Li, Sharad Mehrotra, "Executing SQL over Encrypted Data in the Database Service Provider Model," *Proc. of the ACM SIGMOD 2002*, pp.216-227, June. 2002.

[2] Eu-Jin Goh, "Secure Indexes," In submission, 2004.

[3] Bijit Hore, Sharad Mehrotra, Gene Tsudik, "A Privacy-Preserving Index for Range Queries," *Proc. of the 30th VLDB Conference*, pp.720-731, Aug. 2004.

[4] Latanya Sweeney, "K-anonymity: A model for protecting privacy," *Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol.10, no.5, pp.557-570, Oct. 2002.

[5] Volker Gaede, Oliver Gunther, "Multidimensional AccessMethods," Technical Report, Humboldt-University of Berlin, ACM Computing Surveys, vol.30, pp.170-231, Jun. 1998.