http://dx.doi.org/10.3745/KIPSTC.2011.18C.4.199

원격 저장소 환경에서 다중 키워드를 이용한 효율적인 검색 가능한 대칭키 암호 시스템

이 선 호[†] · 이 임 영[††]

요 약

매우 가벼운 무게와 손안에 들어가는 작은 크기로 휴대성을 제공하는 휴대용 저장매체는 사용자들로부터 많은 호응을 받고 있다. 하지만 휴대용 저장매체의 휴대성으로 인하여 USB 메모리의 분실 및 도난이 잦아졌고 그로 인하여 저장매체 내부에 저장되어 있는 개인정보가 유출되는 사고가 발생하는 문제점이 발생되었다.

따라서 분실위험이 없고 네트워크를 통하여 언제든 자료를 저장하고 접근할 수 있는 원격 저장소 서비스가 등장하게 되었다. 정보통신산업이 발달함에 따라 여러 종류의 정보 기기를 통하여 언제 어디서든 빠른 네트워크에 접근할 수 있게 되었고, 이는 원격 저장소를 이용하는 사용자들을 더욱 증가하게 하였다. 하지만 여러 사용자의 주요 자료가 저장됨에 따라 비윤리적인 관리자 및 공격자로 인하여 서버에 저장된 여러 사용자의 주요자료가 동시에 유출될 수 있는 위험이 존재한다.

이를 해결하기 위해 서버에 저장되는 자료의 암호화가 필요해졌으며, 이와 동시에 암호화한 자료의 효율적인 검색 및 이용을 위하여 검색 가능 암호 시스템이 필요하다. 하지만, 기존의 대칭키 검색 가능 암호 시스템은 문서의 삽입/삭제 효율성 및 다중 키워드 검색 시 연산의 효율성이 떨어지는 문제점이 존재한다. 따라서 본 논문은 기존 대칭키 검색 가능 암호의 문제점을 해결할 수 있는 효율적인 대칭키 검색 가능 암호 시스템을 제안한다.

키워드 : 검색가능 암호, 다중키워드

Effective Searchable Symmetric Encryption System using Conjunctive Keyword on Remote Storage Environment

Sun-Ho Lee[†] · Im-Yeong Lee[††]

ABSTRACT

Removable Storage provides the excellent portability with light weight and small size which fits in one's hand, many users have recently turned attention to the high-capacity products. However, due to the easy of portability for Removable Storage, Removable Storage are frequently lost and stolen and then many problems have been occurred such as the leaking of private information to the public.

The advent of remote storage services where data is stored throughout the network, has allowed an increasing number of users to access data. The main data of many users is stored together on remote storage, but this has the problem of disclosure by an unethical administrator or attacker.

To solve this problem, the encryption of data stored on the server has become necessary, and a searchable encryption system is needed for efficient retrieval of encrypted data. However, the existing searchable encryption system has the problem of low efficiency of document insert/delete operations and multi-keyword search. In this paper, an efficient searchable encryption system is proposed.

Keywords : Serhcable Encryption, Conjunctive Keyword

1. 서 론

전 세계적으로 네트워크 산업에 관심이 높아지면서 미국, 유럽, 일본 등 세계 주요국에서 경쟁적으로 망 구축 사업이 진행되었다. 대한민국의 경우 1990년대부터 진행된 '초고속 정보통신 기반 구축사업'으로 세계 최고 수준의 네트워크를 구성하였으며, 사용자는 유무선 네트워크를 통해 언제 어디

※ 이 논문은 2010년도 정부(교육과학기술부)의 재원으로 한국연구재단의 기초연구사업 지원을 받아 수행된 것임(2010-0022607).
† 준 회 원 : 순천향대학교 컴퓨터소프트웨어공학과 박사과정
†† 종신회원 : 순천향대학교 컴퓨터소프트웨어공학과 교수
논문접수 : 2011년 1월 31일
수 정 일 : 1차 2011년 4월 25일
심사완료 : 2011년 4월 26일

서든 인터넷에 접속할 수 있게 되었다. 이와 같은 네트워크의 발달은 사용자들이 데이터를 저장하는 방식의 변화를 불러일으켰다. 사용자들은 자신의 데이터를 휴대하기 위해 주로 USB 메모리나 외장 하드와 같은 휴대용 저장매체를 이용하였다. 이와 같은 휴대용 저장매체는 분실 위험이 있으며, 소지하지 않았을 경우 저장된 데이터에 접근할 수 없는 불편함을 가지고 있다. 따라서 사용자들은 분실위험 없이 언제 어디서나 다양한 디바이스를 통해 자신의 자료를 저장 및 접근하길 원하게 되었으며, 이로 인하여 각종 원격 저장소 서비스가 등장하고 있다. 하지만, 원격 저장소는 해커나 비윤리적인 관리자로 말미암아 내부에 저장된 개인 정보 및 주요정보가 유출되는 사고가 발생할 위험이 존재한다. 이와 같은 서버의 비신뢰성 문제를 해결하기 위해 서버에 저장되는 자료를 오직 자신만이 알고 있는 키로 암호화 저장하고, 이를 서버에 정보유출 없이 검색하기 위한 검색가능암호 기술의 필요성이 대두 되고 있으며, 이메일 서버 환경 등을 고려한 검색가능암호가 활발히 연구되고 있다[1][2]. 본 논문은 기존방식과 달리 빈번한 데이터 삽입/삭제 및 다중 키워드를 통한 검색이 필요한 원격저장소 환경을 고려하였다. 먼저 데이터가 가지는 키워드들을 표현하는 색인이 빈번한 데이터 삽입/삭제로 자주 변경되어야 하는 측면을 고려하여 색인을 비트열로 표현, 비트 반전으로 색인을 수정하는 방법을 제안하였다. 또 다중 키워드 검색 시 키워드의 가지수 만큼 반복되는 검색 연산의 효율성을 제공하기 위해 여러 키워드에 대한 검색 쿼리를 통합하여 검색하는 응용방법을 제안한다. 본 논문의 구성은 다음과 같다. 2장에서는 기존의 검색 가능한 암호 시스템을 분석하고 요구 사항을 도출하며, 3장에서는 요구 사항을 만족하는 방식을 제안하며, 4장에서는 제안방식을 분석한다. 마지막으로 5장에서 결론을 맺는다.

2. 연구 배경

2.1 기존연구

검색 가능한 암호 시스템은 기존에 연구된 암호 시스템과 같이 대칭키 방식에서 공개키 방식으로 발전되었다. 대칭키 기반의 대표적 검색 가능한 암호 시스템은 최초의 실용적 검색가능 암호를 제안한 Song 등의 방식[6]으로부터 색인 크기의 효율성을 고려한 Goh의 방식[7], 링크드리스트로 색인을 생성하고 빠른 검색 속도를 제공하는 Curtmola 등의 방식[9]이 제안되었다. 또 공개키 기반의 대표적 검색 가능한 암호 시스템은 최초 Boneh등의 방식[5]를 시작으로 다중 키워드 검색을 고려한 Hwang 등의 방식[10], 필드에 제한 없이 유연한 다중 키워드 검색을 가능케 한 Wang등의 방식[8], 다중 사용자 환경을 고려한 Yang등의 방식[11], 다중 키워드 검색에 연산 효율성을 제공하는 Zhang등의 방식[4]이 대표적이다. 기존에 연구된 다양한 검색 가능한 암호 시스템은 각자 다양한 환경을 고려하여 연구되었다.

본 연구는 원격 저장소 환경을 고려하여 각 개인 사용자가 자신의 데이터를 서버에 저장하고 이를 빠르고 안전하게 검색하기 위한 방법을 제안하고자 한다. 따라서 본 절에선 이러한 환경에 적용 가능한 대표적 대칭키 검색 가능한 암호시스템에 대하여 살펴보도록 한다.

2.1.1 Practical Techniques for Searches on Encrypted Data

Practical Techniques for Searches on Encrypted Data는 2000년에 Song 등에 의해서 제안된 논문[6]으로 초기의 대칭키 기반 검색 가능한 암호로 평가되고 있다. Song 등은 신뢰할 수 없는 서버로부터 자료의 기밀성을 제공하면서 특정 키워드를 가지는 암호화데이터를 검색할 수 있는 방법을 제안하였다. 해당 방식의 내용은 다음과 같다.

• Encryption

해당 방식은 의사난수 순열과 의사난수 함수로 생성된 스트림 암호와 암호화 키워드를 XOR연산을 통해 재 암호화 한다(그림 1 Encrypt 참조).

Step 1. 해당 연구에서는 문서를 워드들 W_i로 나누며, 각 워드는 64비트 블록으로 표현한다.

Step 2. 각 워드의 블록을 사전 암호화하여 암호문 $X_i = E_{k_1}(W_i)$를 만든다. 여기서 사전 암호문 X_i를 2개 부분$X_i = \langle L_i, R_i \rangle$으로 나눈다. 여기서 L_i는 $n-m$ 비트, R_i는 m 비트이다.

Step 3. 키 k_i를 다음과 같이 $k_i = f_{k_2}(L_i)$생성한다. 의사난수생성기(Pseudorandom generator) G에 $seed$값을 넣어 $n-m$ 비트의 S_i를 생성하고, 의사난수함수(Pseudorandom function) F를 이용해 m 비트의 $F_{k_i}(S_i)$를 생성한다.

Step 4. 앞서 생성한 $\langle L_i, R_i \rangle$와 $\langle S_i, F_{k_i}(S_i) \rangle$를 XOR하여 암호문을 생성한다.

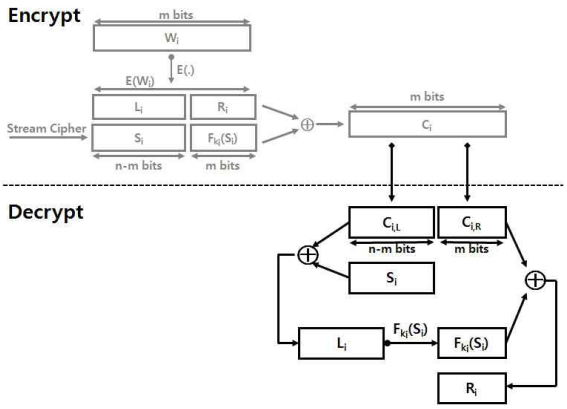
• Decryption

해당 방식은 암호화된 색인으로부터 다음과 같은 연산을 통하여 암호화된 키워드를 획득 및 비교한다(그림 1 Decrypt 참조).

Step 1. 암호문을 검색하기 위해 필요한 트랩 도어 $E_{k_1}(W_i) = (L_i \parallel R_i)$와 $k_i = f_{k_2}(L_i)$를 생성한다.

Step 2. 서버에 저장된 암호문 C_i에 대해서 $C_i \oplus E_{k_1}(W_i)$를 계산한다.

Step 3. 결과가 $S_i \parallel F_{K_1}(S)$이면 해당 문서가 원하는 워드를 가지는 검색 결과가 되며, 이를 사용자에게 전달한다. 이렇게 함으로써 암호문 C를 해독하지 않은 상태에서도 평문의 특정 정보를 포함하고 있는 문서들에 대한 검색을 수행할 수 있다.

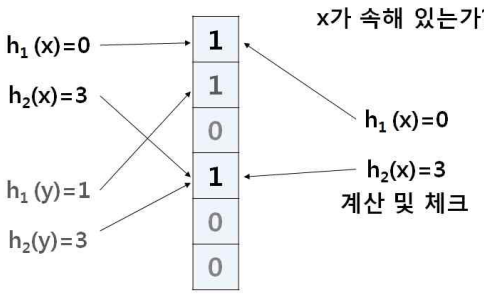


(그림 1) Song et al. 방식의 키워드 암/복호화 과정

해당 방식은 명확한 안전성이 정의되지는 않았지만 초기의 검색 가능 암호 시스템으로써 의미가 있다고 평가된다.

2.1.2 Secure Indexes

Secure Indexes는 2003년 Goh가 제안한 논문[7]이다. 처음으로 검색 가능 암호시스템에 대한 명확한 안전성 모델을 정의하였으며, 제안된 검색 가능 암호 시스템이 안전함을 증명한 것으로 평가되고 있다. 해당 논문은 Bloom filter를 사용하여 문서가 가지는 키워드를 표현하고 있다. Bloom filter는 (그림 2)와 같이 여러 값들을 다양한 해시 함수에 넣어 나온 값을 비트 열로 표시하고, 차후 특정 값이 해당되는지 포함 여부를 알아 볼 수 있는 데이터 표현 방법이다. 해당방식은 색인 생성 단계와 색인으로부터 키워드로 문서를 검색하는 단계로 구성되어 있다.



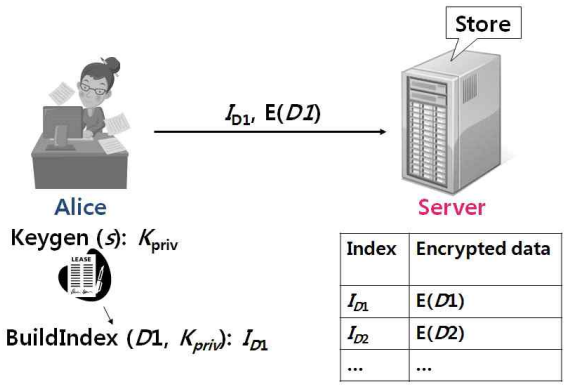
(그림 2) Bloom filter 예시

• 색인 생성

Secure Index는 검색할 문서와 문서가 가지는 키워드들로 Bloom Filter를 이용하여 사전에 색인을 생성하며, 그 과정은 암호화키를 생성하는 Keygen단계와 색인을 생성하는 BuildIndex단계로 구성된다(그림 3 참조).

Keygen(s): 먼저 보안 파라미터 s가 주어지며, 의사난수 함수 $f: \{0,1\}^n \times \{0,1\}^s \rightarrow \{0,1\}^s$를 선택한다. 이를 통하여 마스터키인 $K_{priv} = (k_1,\ldots,k_r) \leftarrow \{0,1\}^{sr}$를 생성한다.

BuildIndex(D, K_{priv}): 주어진 문서 $D: D_{id} \in \{0,1\}^n$와 워드들$(W_0, \ldots, W_t) \in \{0,1\}^{nt}$ 마스터키 $K_{priv} = (k_1, \ldots, k_r) \in \{0,1\}^{sr}$로 Bloom Filter에 각 문서에 해당하는 워드들을 표현한다. 문서 D_{id}에 해당하는 각 워드 $W_i (1 \le i \le t)$를 의사난수 함수를 통하여 $x_j = f(W_i, k_j)$ $(1 \le j \le r)$를 생성하고 이를 통하여 코드워드 $y_j = f(D_{id}, x_j)$ $(1 \le j \le r)$를 생성하여 블룸필터BF에 y_j에 해당하는 비트를 1로 치환하는 방식으로 색인을 생성한다.



(그림 3) Secure index의 색인 생성 개념도

• 문서 검색

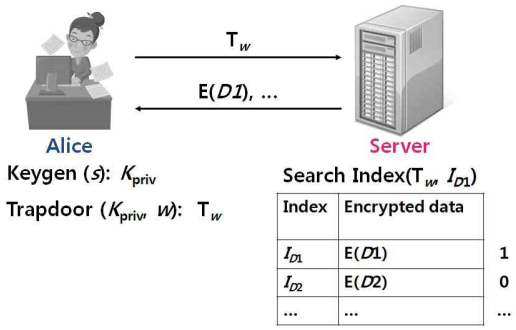
Secure Index는 검색하고자 하는 키워드로 트랩도어를 생성하여 키워드가 문서에 포함되는지 탐색하며, 그 과정은 문서 검색을 위한 키를 생성하는 Keygen단계와 문서 검색을 위한 트랩도어를 생성하는 Trapdoor단계, 트랩도어를 이용하여 실제로 문서를 찾는 SearchIndex단계로 구성된다(그림 4 참조).

Keygen(s): 먼저 보안 파라미터 s가 주어지며, 의사난수 함수 $f: \{0,1\}^n \times \{0,1\}^s \rightarrow \{0,1\}^s$를 선택한다. 이를 통하여 마스터키인 $K_{priv} = (k_1,\ldots,k_r) \leftarrow \{0,1\}^{sr}$를 생성한다.

Trapdoor(K_{priv}, W): 주어진 마스터키 $K_{priv} = (k_1, \ldots, k_r) \in \{0,1\}^{sr}$와 워드 W를 입력받아 검색을 위해 필요한 트랩도어를 생성한다. 입력받은 마스터키와 워드로 W를 위한 트랩도어 $T_W = (f(W, k_1), \ldots, f(W, k_r)) \in \{0,1\}^{sr}$를 생한다.

SearchIndex(T_W, $I_{D_{id}}$): 주어진 트랩도어 $T_W = (x_1, \ldots, x_r) \in \{0,1\}^{sr}$와 색인$I_{D_{id}} = (D_{id}, BF)$로 문서가 해당키워드를 가지는지 판별하며 상세 내용은 다음과 같다. 서버는 사용자로부터 주어진 트랩도어를 통해 검색할 문서 D_{id}와 x_j $(1 \le j \le r)$로 코드워드 $y_j = f(D_{id}, x_j)$

$(1 \le j \le r)$를 생성하여 블룸필터BF에 y_j에 해당하는 비트가 1인지 확인한다. 만약 y_j에 해당하는 비트가 전부 1이면 해당 문서에 검색하고자 하는 키워드가 존재하는 것으로 판별한다.

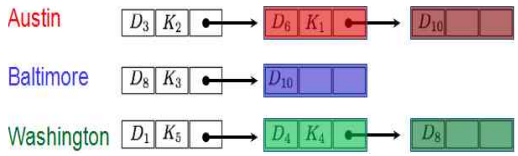


(그림 4) Secure index의 문서 검색 개념도

해당 방식은 사용자가 기억하고 있어야 할 키 $(k_1, k_2, \ldots, k_r)$가 길어 비효율 적이며, Bloom filter를 사용하였기 때문에 긍정 오류가 발생될 확률을 가지고 있다. 긍정오류의 경우 Bloom filter의 크기와 사용되는 함수의 가지수를 늘려 해결 가능하다[3].

2.1.3 Searchable Symmetric Encryption

기존의 대칭키 방식 중 가장 빠른 검색 속도를 제공하는 Curtmola 등이 제안한 Searchable Symmetric Encryption (SSE)의 경우 (그림 5)와 같이 링크드리스트로 각 키워드를 가지는 문서들을 구현하였다[9]. 해당방식의 색인 생성 단계 및 검색 단계의 내용은 다음과 같다.



(그림 5) SSE 색인 구조

• 색인 생성

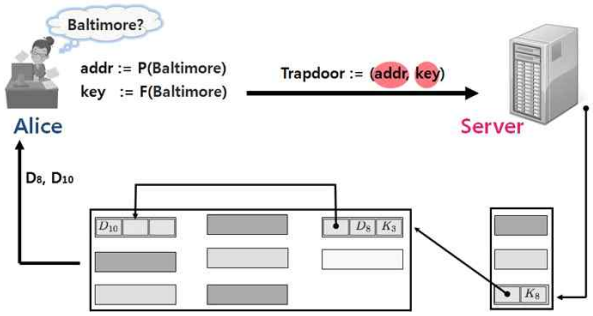
SSE는 링크드리스트로 각 키워드를 가지는 문서들을 구현하였다. 각 링크드리스트는 키워드별로 존재하고 있으며, 키워드를 가지는 문서들의 목록을 표현하고 있다.

각 노드는 문서의 주소 값, 다음 노드의 복호화를 위한 킷값, 다음 노드의 주소값과 같은 부수적인 데이터를 가지고 있다. 특히, 링크드리스트의 최초의 노드 위치는 키워드의 의사난수 순열을 통해 지정되며, 최초노드의 암호키는 키워드의 의사난수 함수값으로 지정된다.

• 문서 검색

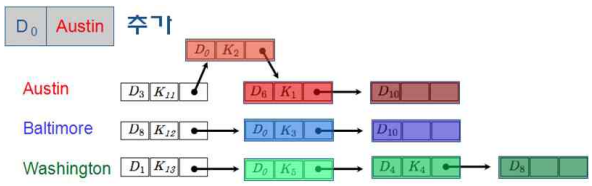
Searchable Symmetric Encryption의 문서 검색과정은 기존 검색가능 암호 시스템과 마찬가지로 트랩도어를 이용한다.

본 방식에서 사용하는 트랩도어는 검색하고자하는 키워드의 의사난수 순열값 $P(keyword)$와 의사난수 함수값 $F_k(keyword)$으로 구성된다. 여기서 $P(keyword)$는 최초노드의 주소값, $F_k(keyword)$는 최초노드의 암호키가 된다. 이를 통하여 해당 키워드의 최초 노드 위치를 찾고 최초노드를 복호화 한다. 최초노드에 있는 다음노드의 주소값과 킷값을 통하여 다음노드들을 반복하여 복호화하고, 다음노드 위치가 Null이 나오게 되면 지금까지 탐색한 문서의들을 사용자에게 반환한다(그림 6 참조).



(그림 6) SSE의 검색 과정

해당 방식은 노드별로 문서의 주소 값, 다음 노드의 복호화를 위한 킷값, 다음 노드의 주소값과 같은 부수적인 데이터가 추가되어 저장공간의 효율성이 떨어진다. 또한, SSE 방식에서 문서를 추가할 때 (그림 7)과 같이 최초 노드의 복호화 및 링크를 변경해야 하는 복잡한 구조로 되어 있다. 문서의 삽입뿐만 아니라 문서를 삭제할 때 문서에 해당하는 각 키워드의 링크드리스트에서 삭제할 문서의 노드를 찾아 삭제해야 하며, 이를 위해서 복호화 연산이 지속적으로 발생되어 문서의 삽입/삭제 시 서버의 과부하를 유발하는 문제점이 있다.



(그림 7) SSE 문서 추가 방법

2.2 요구 사항

원격 저장 환경에서 검색 가능 암호 시스템은 다음과 같은 요구 사항을 만족해야 한다.

- 기밀성: 원격 저장소 서버와 클라이언트 단말기 간의 통신내용 및 저장된 색인을 통하여 서버나 제3의 공격자가 검색하는 키워드 및 키워드에 해당하는 문서를 유추할 수 없어야 한다.
- 검색 속도: 제한적인 시스템 자원을 가지는 클라이언트에서도 원격저장소에 저장된 검색 하고자 하는 키워드를 포

함하는 문서의 검색이 빠르게 제공되어야 한다.

- 서버 저장공간의 효율성: 검색 가능 암호 시스템에서 빠른 검색을 위하여, 생성하는 색인의 용량이 크지 않아야 한다.
- 통신량의 효율성: 클라이언트와 서버 간의 네트워크 자원의 효율성을 위하여 통신량이 적어야 한다.
- 문서 삽입/삭제의 효율성: 대용량의 문서의 빈번한 삽입/삭제 환경에서 서버의 연산 효율성이 제공되어야 한다.
- 문서 검색의 효율성: 한 번의 검색만으로 단일 키워드 검색이 아닌 유연한 다중 키워드 검색을 지원하는 효율성이 제공되어야 한다.

3. 제안방식

제안방식은 기존의 대칭키 검색 가능 암호 시스템인 SSE 암호 방식의 단점인 문서 삽입/삭제의 효율성 및 다중 키워드 검색 효율성을 제공하기 위하여 제안된 방식으로, 색인을 의사 난수 순열 및 의사 난수 함수로 작성된 $n \times m$ 비트 표현(색인은 DB에 저장하며, 색인의 무결성은 DBMS를 통해 제공된다고 가정한다.)되며, 문서를 검색할 수 있도록 색인을 생성하는 색인 생성 단계, 색인이 생성된 문서를 검색하기 위한 트랩도어 생성 및 테스트를 수행하는 문서 검색 단계와 문서 삽입/삭제의 효율성을 보여주기 위해 추가적인 문서 삽입단계 및 문서 삭제 단계의 과정으로 구성된다.

3.1 시스템 계수

다음은 제안방식에서 사용되는 시스템 계수이다.

k : 암호화 키
m : 전체 문서의 개수
n : 전체 키워드의 개수
j : 특정 문서가 가지는 키워드의 개수 ($j \leq n$)
q : 특정 키워드를 가지는 문서의 개수 ($q \leq m$)
d_i : i번째 문서
w_i : i번째 키워드
$w_{i,j}$: i번째 문서의 j번째 키워드
D : 문서들의 집합
W : 키워드들의 집합
BF_{d_i} : i번째 문서의 블룸필터
$E_*[\,]$: *로 암호화
$D_*[\,]$: *로 복호화
$F_*[\,]$: *의 키로 의사 난수 함수
$P[\,]$: 의사 난수 순열
$H[\,]$: 안전한 일방향 함수
IT_i : 색인 테이블의 i번째 레코드
DT_i : 자료 테이블의 i번째 레코드
T_* : *키워드를 가지는 문서를 검색하기 위한 트랩도어

3.2 제안방식

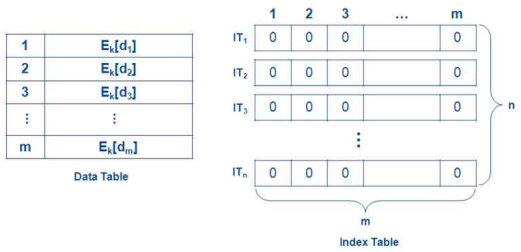
3.2.1 색인 생성 단계

문서를 추가할 때 암호화된 문서가 가지는 키워드를 표현하기 위한 색인을 생성하게 된다.

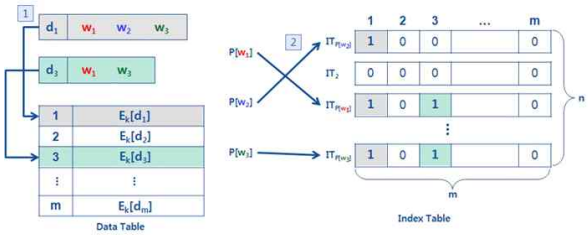
Step 1. 색인을 생성하고자 하는 전체 문서의 개수 m과 전체 문서가 가지는 키워드들의 개수 n으로 (그림 8)과 같은 자료 테이블(Data Table)과 색인 테이블(Index Table)을 작성한다. 각 키워드는 고유한 색인 테이블을 가지며 색인 테이블은 m비트로 구성된다. 키워드를 가지는 문서 번호에 해당하는 색인 비트를 1로 표시하여 해당 키워드를 가지는 문서들을 표현한다.

Step 2. 색인을 생성할 문서들을 자료 테이블에 암호화하여 저장 한 뒤, 해당 문서들의 키워드 목록으로 색인 테이블을 생성한다. 각 키워드는 $P[w_i]$연산을 통하여 m비트의 테이블 $IT_{P[w_i]}$를 할당받는다. 해당 키워드 w_i를 가지고 있는 문서에 해당하는 비트를 1로 표시한다(그림 9 참조).

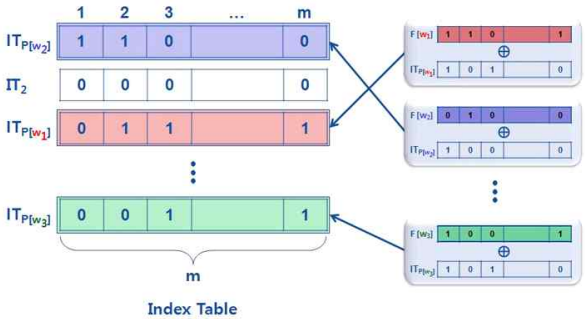
Step 3. 색인 테이블의 내용이 평문으로 노출되지 않도록 하기 위하여 마스킹을 수행하게 되는데 $F_k[w_i]$를 통해 생성된 m비트 배열과 $IT_{P[w_i]}$를 XOR 연산하여 암호화 색인을 생성한다(그림 10 참조).



(그림 8) 색인 테이블 및 자료 테이블 생성 및 초기화



(그림 9) 각 테이블에 문서 정보 추가



(그림 10) 색인 테이블에 마스킹 수행

3.2.2 문서 검색 단계

클라이언트가 서버에 저장된 키워드 w_i를 가지는 문서들을 검색하기 위해서 기존의 검색 가능 암호 시스템과 같이 트랩도어를 이용한다(그림 11 참조).

Step 1. 의사 난수 생성기 P와 의사 난수 함수 F로 검색하고자 하는 키워드 w_i의 트랩도어를 생성한다.

$T_{w_i} = [\, P[w_i] \,\|\, F_k[w_i] \,]$

Step 2. 트랩도어를 받은 서버는 $P[w_i]$를 참조하여 키워드에 대한 색인 테이블을 추출한다.

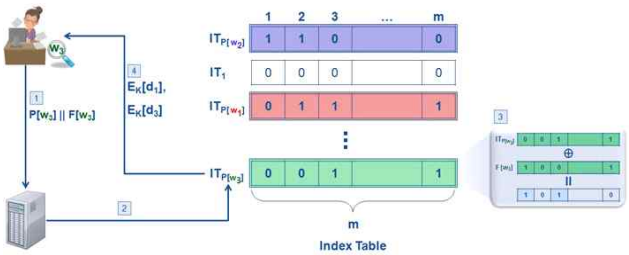
$IT_{P[w_i]}$

Step 3. 마스크 되어있는 색인 테이블과 마스크 비트를 XOR 연산하여 마스크를 해지 한 뒤 1로 표시되어진 비트들을 확인한다.

$IT_{P[w_i]} \oplus F_k[w_i]$

Step 4. 색인에 1로 표시되어진 비트의 해당하는 문서를 추출하여 클라이언트에게 전송한다.

Step 5. 클라이언트는 전송된 암호화된 문서들을 키 k로 복호화 한다.



(그림 11) 문서 검색 과정

3.3 제안방식의 응용

3.3.1 문서 삽입

본 방식은 문서 삽입/삭제를 신속하게 처리하기 위해서 색인 테이블의 암/복호화 과정 없이 문서가 가지고 있는 키워드의 색인 테이블에 추가하고자 하는 문서 번호 비트를 반전시켜 빠른 문서 삽입 환경을 제공한다(그림 12 참조).

Step 1. 새로이 삽입하고자 하는 문서d_i를 자료 테이블에 키k로 암호화하여 저장 한다.

Step 2. 추가된 문서d_i의 키워드$[w_1, w_2, \ldots, w_j]$들을 의사 난수 순열에 넣어 나온 결과로 해당 키워드의 색인 테이블 위치를 추출한다.

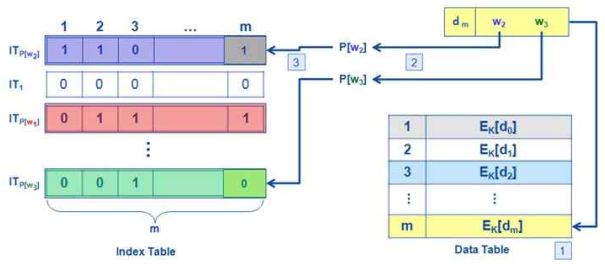
$d_i = [w_1, w_2, \ldots, w_j]$

$P[w_1] \sim P[w_j]$

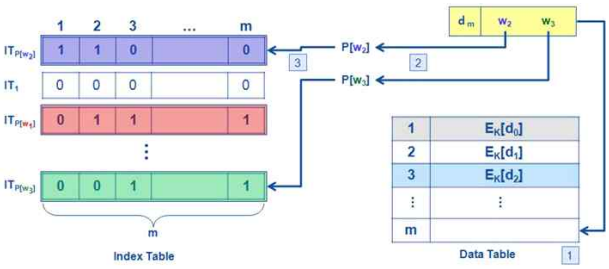
Step 3. 추가된 문서가 가지는 키워드의 색인 테이블 중 문서의 번호에 해당하는 비트 값을 반전시킨다.

$IT_{P[w_j]}[i] = IT_{P[w_j]}[i] \oplus 1$

3.3.2 문서 삭제



(그림 12) 문서 삽입 과정



(그림 13) 문서 삭제 과정

문서의 삭제 역시 삽입과 동일한 순서로 이루어진다(그림 13 참조).

Step 1. 삭제하고자 하는 문서를 자료 테이블에서 삭제한다.

Step 2. 삭제한 문서의 키워드들을 의사난수 순열에 넣어 해당 키워드의 색인 테이블 위치를 추출한다.

$d_i = [w_1, w_2, \ldots, w_j]$

$IT_{P[w_1]} \sim IT_{P[w_j]}$

Step 3. 키워드의 색인 테이블 중 추가한 문서에 해당하는 비트 값을 반전시킨다.

$IT_{P[w_j]}[i] = IT_{P[w_j]}[i] \oplus 1$

3.3.3 다중 키워드 검색

다중 키워드 검색은 다음과 같은 순서로 이루어진다.

Step 1. 클라이언트는 검색하고자 하는 각 키워드에 대한 트랩도어를 연접해 서버에 전송한다.

Step 2. 트랩도어를 받은 서버는 첨부된 키워드들의 색인 테이블을 추출한다.

Step 3. 키워드들에 해당하는 색인 테이블과 마스크 비트를 XOR 연산하여 마스크를 해지 한 뒤 각 결과들에 AND 연산을 취한 뒤, 1로 표시되어진 비트들을 확인한다.

Step 4. 색인에 1로 표시되어진 비트의 해당하는 문서를 추출하여 클라이언트에게 전송한다.

Step 5. 클라이언트는 전송된 암호화된 문서들을 키 k로 복호화 한다.

4. 제안방식 분석

본 제안 방식은 대용량 문서의 빈번한 삽입/삭제가 일어

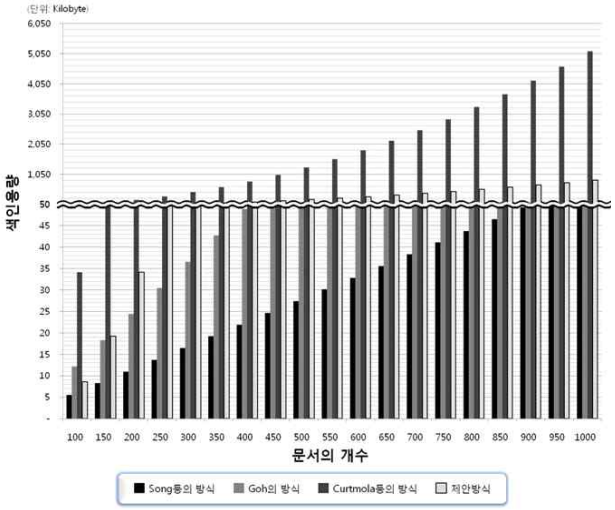
나는 환경에서 효율적인 검색 가능 암호 시스템을 제공한다. 본 방식은 아래와 같은 요구 사항을 만족하며 기존에 제안된 방식과의 비교표는 <표 1>과 같다.

• 기밀성: 제안방식은 문서가 암호화 저장되며, 문서가 가지는 키워드를 표현하는 색인 또한 의사난수 함수로부터 생성된 암호 스트림과 XOR 연산되어 이를 통하여 키워드를 유추하기 어렵다. 또한 키워드의 색인 위치를 키로부터 도출된 의사난수 순열로만 알 수 있어, 서버는 사용자가 어떠한 키워드를 검색하는지 유추할 수 없으며, 제안방식은 의사난수 순열 및 의사난수 함수를 이용하는 기존 방식들과 동일한 안전성 모델을 가진다.

• 검색 속도: 제안방식은 검색할 키워드에 해당하는 고유 색인 테이블을 의사난수 순열 연산 한번만으로 찾아낼 수 있으며, 마스크 비트와의 XOR 연산을 통해 마스크 제거 후 1로 표현된 비트의 위치로 검색 키워드에 해당하는 문서를 빠르게 찾을 수 있다. 제안방식은 빠른 검색속도를 제공하는 Curtmola 등의 방식과 달리 링크 노드를 찾고 복호화 하는 과정이 없어 더욱 빠른 검색속도를 제공한다.

• 서버 저장공간의 효율성: 제안방식은 m*n비트를 사용하며, Curtmola 등의 방식보다 소폭 상승된 저장공간의 효율성을 제공한다. 기존 방식과 제안방식의 효율성 비교는 (그림 14)와 같다. 해당 그래프를 얻기 위해 본 연구에서는 문서 하나당 키워드의 개수는 10개, 문서의 증가에 따른 키워드의 중복확률은 30%로 가정하였으며, Goh등의 방식에서 사용하는 문서별 블룸필터의 크기는 1024비트, Curtmola등의 방식의 node의 크기는 20비트로 가정하였다.

• 통신량의 효율성: 제안방식은 기존의 검색 가능 암호 시스템과 같이 한 번의 트랩도어 전송으로 결과 값을 받아



(그림 14) 문서 개수에 따른 색인용량 비교

통신의 효율성을 제공한다.

• 문서 삽입/삭제의 효율성: 제안방식은 문서의 삽입/삭제 시 해당 문서가 가지는 키워드 색인에서 문서의 번호에 해당하는 비트를 암/복호화 없이 단순 반전시킴으로써 문서의 삽입/삭제가 간단하게 수행되어, Curtmola등의 방식[9]과 달리 매우 빠른 문서 삽입/삭제 효율성을 제공한다.

• 문서 검색의 효율성: Song등의 방식[6], Curtmola등의 방식[9]은 다중 키워드 검색을 위하여 검색할 키워드의 개수만큼의 검색을 수행한다. 하지만 Goh의 방식[7]과 제안방식은 한 번에 각 키워드에 대한 쿼리를 합쳐 전송하고, 이를 통해 1라운드의 통신만으로 다중 키워드 검색을 지원하는 문서 검색의 효율성을 제공한다.

〈표 1〉 제안방식 분석

	Song등의 방식[6]	Goh의 방식[7]	Curtmola등의 방식[9]	제안방식
기밀성	○	○	○	○
검색속도	△	△	○	○
검색 시 비교횟수	$j*m$	c	1	1
저장공간 효율성	×	○	×	○
색인 용량(비트)	m^*키워드크기$^*j$	m^*블룸필터의 크기	n^*q^*노드크기	m^*n
통신 효율성	○	○	○	○
통신 횟수	2	2	2	2
문서 삽입/삭제 연산 효율성	○	○	×	○
문서 삽입 시 암·복호화 연산횟수	2	0	3	0
문서 삭제 시 암·복호화 연산횟수	2	0	$2\sim q$+1	0
다중 키워드 검색 효율성	×	○	×	○
다중 키워드 검색 시 통신 및 검색 라운드 수	키워드 개수	1	키워드개수	1

○: 좋음, 제공됨, △: 보통, 부분 제공됨 ×:나쁨, 제공되지 않음

5. 결 론

네트워크의 환경 변화에 따라 데이터를 저장하는 방식이 오프라인 매체를 이용하는 방식에서 원격 저장소를 이용하는 방식으로 변화하고 있다. 이러한 변화는 언제 어디서나 빠르게 접속 가능한 유/무선 네트워크의 발전과 함께 더욱 가속화 될 전망이다. 따라서 이러한 원격 저장소에 적용 가능한 검색가능 암호 시스템의 개발이 시급한 실정이다.

기존의 대칭키 검색 가능 암호 시스템의 경우 문서의 삽입/삭제의 비효율성, 문서 검색을 위해 요구되는 연산량에 비효율성을 가지고 있으며. 다중 키워드 검색을 위해 여러 차례 통신이 발생되는 문제점을 가지고 있다. 이는 빠른 검색 및 빈번한 문서의 삽입/삭제가 발생되는 원격저장소 환경에 적용하기 어렵다. 따라서 제안 방식에서는 이를 해결하기 위해 문서가 가지는 키워드 정보를 색인하는 것이 아닌, 키워드를 가지는 문서들의 정보를 색인으로 표기하여 의사난수 순열 연산 한번만으로 키워드를 가지는 문서들의 목록을 확인 가능하도록 하였으며, 문서의 삽입/삭제 시 색인에서 문서에 해당하는 비트를 반전시키는 방식으로 빠르게 색인을 갱신하는 연산의 효율성을 제공하도록 하였다. 또 한 라운드의 통신만으로 다중 키워드 검색이 가능한 응용 방안을 제안하였다. 제안 방식은 기존 대칭키 기반 검색 가능 암호 보다 빠른 검색 및 문서 삽입/삭제 속도를 제공하지만 저장소 효율성 부분에서는 Curtmola 등의 방식보다 소폭 상승된 효율성을 제공할 뿐 뛰어난 용량의 효율성을 제공하지는 못하고 있다. 따라서 향후에는 빠른 검색 속도 및 저장공간의 효율성을 동시에 제공 가능한 방법에 대한 연구가 진행되어야 할 것으로 사료된다.

참 고 문 헌

[1] 김선영, 서재우, 이필중, "검색 가능 암호 기술의 연구 동향," 정보보호학회지, 19(2), pp.63-73, 2009.

[2] 조남수, 홍도원, "검색 가능 암호 시스템 기술 동향," 전자통신동향분석 23(4), pp.1-9, 2008.

[3] B.Bloom, "Space/time trade-offs in hashcoding with allowable errors." Communications of the ACM, 13(7), pp.422 - 426, 1970.

[4] B. Zhang, and F. Zhang, "An efficient public key encryption with conjunctive-subset keywords search", Journal of Network and Computer Applications, 34(1), 2011.

[5] D. Boneh, G. Di. Crescenzo, and R Ostrovsky, "Public key encryption with keyword search," Eurocrypt'04,

[6] D. Song, D. Wagner, and A. Perrig, "Practical Techniques for Searching on Encrypted Data," Proceedings of IEEE Symposium on Security and Privacy, pp.44-55, 2000.

[7] E. J. Goh, "Secure Indexes," Technical Report, 2003/216, IACR ePrint Cryptography Archive, 2003.

[8] P. Wang, H. Wang and J. Pieprzyk, "Keyword Field-Free Conjunctive Keyword Searches on Encrypted Data and Extension for Dynamic Groups", Cryptology and Network Security, pp.178-195, 2008.

[9] R. Curtmola, J. Garay, S. Kamara, and R. Ostrovsky, "Searchable Symmetric Encryption: Improved Definitions and Efficient Constructions," Proceedings of the 13th ACM conference on computer and communication security-ACM-CCS, pp.79-88, 2006.

[10] Y. H. Hwang, and P. J. Lee, "Public key encryption with conjunctive keyword search and its extension to a multi-user system", International Conference on Pairing-Based Cryptography, Pairing'07, 2007.

[11] Y. Yang, F. Bao, X. Ding, and R. H. Deng, "International Journal of Applied Cryptography", 1(4), 2009.

이 선 호

e-mail : sunho431@sch.ac.kr

2009년 순천향대학교 정보기술 공학부(학사)

2011년 순천향대학교 컴퓨터학과(석사)

2011년년~현 재 순천향대학교 컴퓨터소프트웨어공학과 박사과정

관심분야 : 보안USB, 검색가능한 암호

이 임 영

e-mail : imylee@sch.ac.kr

1981년 홍익대학교 전자공학과

1986년 오사카대학 통신공학전공(석사)

1989년 오사카대학 통신공학전공(박사)

1989년~1994년 한국전자통신연구원 선임연구원

1994년~현 재 순천향대학교 컴퓨터소프트웨어공학과 교수

관심분야 : 암호이론, 정보이론, 컴퓨터 보안