제36회 한국정보처리학회 추계학술발표대회 논문집 제18권 제2호 (2011. 11)

분산된 원격 저장소 환경을 고려한 검색 가능 암호 시스템+)

이선호, 이임영
순천향대학교 컴퓨터소프트웨어학과
e-mail:[sunho431, imylee]@sch.ac.kr

Searchable Encryption System: Considering Distributed Remote Storage Environment

Sun-Ho Lee, Im-Yeong Lee
Dept of Computer Software Engineering, Soonchunhyang University

요 약

많은 사용자들이 자신의 데이터를 휴대하기 위해 이동형 저장매체를 이용하고 있다. 하지만 이와 같은 이동형 저장매체는 분실 위험이 있으며, 휴대하지 않았을 경우 데이터에 접근하지 못한다는 불편함을 가지고 있다. 통신기술의 발달로 언제어디서나 다양한 기기를 통해 네트워크 접속이 가능하게 되었고, 많은 사람들은 원격 저장소 서비스를 통해 언제 어디서든 자신의 데이터를 저장 및 접근할 수 있게 되었다.

하지만 최근 이러한 원격 저장소 서비스를 제공하는 서버의 신뢰성 문재가 제기 되고 있다. 이를 해결하기 위해 데이터의 암호화, 암호화된 데이터를 안전하게 검색할 수 있도록 하는 검색 가능한 암호 기술이 연구되고 있다. 하지만, 기존의 검색 가능한 암호 기술은 분산된 서버 구조를 가지는 원격 저장소의 특성을 고려하지 않고 있어 바로 적용하기 어렵다는 문제점을 가지고 있다. 따라서 본 논문은 원격 저장소 구조에 특화되어 있으며 빠른 암, 복호화 속도를 제공하는 검색가능한 암호 기술을 제안한다.

1. 서론

우리나라는 정보화 사업으로 인하여 세계최고 수준의 네트워크를 구성하였다. 사용자들은 이와 같이 빠른 네트워크를 통하여 자신의 주요 자료의 백업 및 언제 어디서든 자료에 접근할 수 있는 접근성을 보장 받기 위해 웹하드와 같은 원격 데이터 저장 서비스를 사용하게 되었다. 하지만 이와 같은 서비스를 제공하는 서버가 해커나 관리자에 인하여 개인 정보 및 주요정보가 노출되는 사건이 빈번하게 발생되었다. 이로 인하여 사용자는 신뢰할 수 없는 서버에 데이터 저장을 하는 부담해소하기 위해 데이터를 암호화 저장할 필요가 생겼으며 이와 함께 암호화된 데이터를 안전하게 검색할 수 있는 검색 가능한 암호 기술의 필요성이 대두 되었다. 하지만 기존의 검색 가능한 암호 시스템의 경우 이메일 환경을 기반으로 설계되어 필드에 제한적인 검색을 지원하고, 키워드 가지 수의 제한을 가지고 있다.

따라서 본 논문은 분산된 서버구조를 가지는 원격 저장소 환경을 고려하여 빠른 데이터 입/출력, 검색속도를 제공하는 검색 가능한 암호 시스템을 제안한다.

Ⅱ. 요구사항

검색가능 암호 시스템은 아래와 같은 요구사항을 만족해야한다.

- 기밀성: 원격 데이터 서버와 클라이언트 단말기 간의 통신 데이터는 정당한 개체만이 확인할 수 있어야 한다.
- 검색 속도: 제한적 시스템 자원을 가지는 클라이언트에서도 웹하드 시스템에 저장된 문서에서 검색하고자 하는 워드를 포함하는 문서를 빠르게 검색할 수 있어야 한다.
- 서버 저장공간: 검색가능 암호 시스템에서 생성하는 인덱스의 용량이 크지 않아야 한다.
- 통신량: 클라이언트와 서버간의 에너지 효율 및 네트워크 자원의 효율성을 위하여 통신량이 적어야 한다.
- 문서 검색 효율성: 한번의 검색만으로 단일 키워드 검색이 아닌 유연한 결합 키워드 검색을 지원하는 효율성이 제공되어야 한다.

Ⅲ. 제안 방식

본 논문에서는 블룸필터를 사용하여 다중 키워드 검색을 지원하는 대칭키 기반 검색가능 암호시스템을 제안한다.

- 시스템 계수

k : 암호화 키
m : 전체 문서의 개수
n : 전체 키워드의 개수
j : 특정 문서가 가지는 키워드의 개수 ($j \leq n$)

+) 이 논문은 2010년도 정부(교육과학기술부)의 재원으로 한국연구재단의 기초연구사업 지원을 받아 수행된 것임 (2010-0022607)

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

▪ 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 자료 테이블(Data Table)과 색인 테이블(Index Table)을 작성한다. 각 키워드는 고유한 색인 테이블을 가지며 색인 테이블은 m 비트로 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.

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.

▪ 문서 검색 단계
클라이언트가 서버에 저장된 키워드 w_i를 가지는 문서들을 검색하기 위해서 기존의 검색 가능 암호 시스템과 같이 트랩도어를 이용한다.

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.

Ⅳ. 제안 방식 분석

- 기밀성: 제안 방식은 페어링암호을 이용하여 악의적인 제3자가 클라이언트와 서버 간의 통신을 도청한다고 해도 통신 내용을 유추하기 어렵다.
- 검색 속도: 결합 키워드 검색 시 키워드의 개수만큼 검색 량이 증가하지 않아 빠른 검색속도를 제공한다.
- 저장 공간 효율성: 블룸필터를 사용하여 암축된 고정크기의 저장 공간을 차지하여 저장 공간의 효율성을 제공한다.
- 통신량 효율성: 단 한 번의 통신으로 다중 키워드 검색을 수행하는 통신량 효율성을 제공한다.
- 문서 검색 효율성: 필드를 사용하지 않는 다중키워드 검색을 지원하여 검색의 효율성을 제공한다.

5. 결론

본 논문은 기존의 공개키 검색 가능 암호 시스템의 문제점인 결합키워드 검색의 비효율성을 해결하기 위해 블룸필터를 이용하여 한 번의 테스트 과정만으로 여러 키워드에 해당 유무를 확인할 수 있는 효율성을 제공하였다.

앞으로 다양한 단말기를 통한 일반사용자의 원격 서버의 데이터 저장이 증가할 것으로 보인다. 따라서 그룹 사용자 환경등 다양한 환경 및 다양한 추가 기능을 제공하는 검색가능 암호 시스템의 지속적인 연구가 필요한 것으로 본다.

참고문헌

[1] B.Bloom, "Space/time trade-offs in hashcoding with allowable errors." Communications of the ACM, 13(7), pp.422 - 426, 1970.
[2] B. Zhang, and F. Zhang, "An efficient public key encryption with conjunctive-subset keywords search",

Journal of Network and Computer Applications, 34(1), 2011.

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

[4] 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.

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

[6] M. Green, and G. Ateniese "Identity-Based Proxy Re-encryption," ACNS2007, 2007.

[7] 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.

[8] 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.

[9] 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.

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