클라우드 스토리지 구조를 고려한 데이터 저장 방법에 대한 연구

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

A Study on Data Storage Method for Cloud Storage Structure

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

요 약

USB flash drive와 같은 이동형 저장매체는 한 손에 쏙 들어오는 작은 크기와 가벼운 무게로 뛰어난 휴대성을 제공하고 있다. 많은 사용자들은 자신의 데이터를 저장하기 위해 고용량을 제공하는 이동형 저장매체에 관심을 보이고 있다. 하지만 이동형 저장매체는 휴대성으로 인한 도난 및 분실당할수 있다. 개인 정보가 유출되는 등의 많은 문제들이 발생하고 있다. 인터넷의 발달과 클라우드 컴퓨팅 붐을 통하여 이동형 저장매체의 문제점을 해결할 수 있는 클라우드 스토리지 서비스가 급증하고 있다. 하지만 이러한 클라우드 스토리지 서비스는 인터넷이 가지고 있는 취약성 및 서버와 서버의 관리자의 비신뢰성 등의 문제를 가지고 있으며, 이로 인한 몇몇 사고가 발생하였다. 이러한 문제를 해결하기 위해서 클라우드 스토리지 환경에서는 데이터를 암호화 저장하고 이를 복호화 과정 없이 검색할 수 있는 검색 가능한 암호 기술의 필요성이 대두되고 있다. 하지만 기존의 검색가능 암호 기술은 사용자가 저장하고자 하는 데이터를 직접 업로드하고, 해당 자료를 필요에 따라 공유 하고, 공유대상이 변화되는 클라우드 스토리지 환경에서 비효율성을 가지고 있어 실제 서비스에 적용하기 힘든 단점을 가지고 있다. 따라서 본 논문에서는 클라우드 컴퓨팅 환경을 고려하여 검색가능한 암호화 색인 생성 및 이를 재암호화를 통해 다른 사용자와 안전하게 공유할 수 있는 검색가능한 재 암호화 시스템을 제안한다.

1. 서론

가벼운 무게와 손안에 들어가는 작은 크기로 높은 휴대성을 제공하는 이동형 저장매체(예: USB flash drive, 외장하드 등)는 자신의 데이터를 휴대하려는 사용자들로부터 많은 호응을 받고 있다. 이러한 이동형 저장매체는 높은 휴대성으로 인해 분실 및 도난의 위험이 높으며, 이로 인하여 저장되어 있던 개인 정보가 유출되는 문제가 이슈화 되고 있다.

네트워크의 발달로 빠른 데이터 통신이 가능해짐에 따라 자신의 데이터를 원격 저장소에 저장하고 언제 어디서나 다양한 디바이스를 통해 자신의 데이터에 접근할 수 있는 클라우드 컴퓨팅 서비스가 등장하게 되었으며, 현재 MS의 sky drive, Apple의 iCloud 등 다양한 업체에서 경쟁적으로 클라우드 스토리지 서비스를 제공하고 있으며, 최근 경쟁적으로 고용량의 클라우드 스토리지를 무료로 제공하고 있다.

저장매체를 분실할 위험 없이 자신의 데이터를 네트워크를 통해 저장할 수 있게 됨에 따라, 사용자들은 데이터를 휴대하기 위해 이동형 저장매체가 아닌 클라우드 스토리지를 이용하게 되었다. 하지만 현재 사용되고 있는 인터넷은 개발 당시 패킷에 대한 암호화, 인증 등의 보안기능을 고려하지 않아 다양한 보안취약점을 가지고 있으며, 해커는 이를 이용하여 네트워크상에서 전송되는 패킷을 상대로 다양한 공격을 시도할 수 있다. 또한 사용자는 클라우드 컴퓨팅의 특성상 자신의 데이터가 어떤 서버에 저장되는지 알 수 없으며, 가격 경쟁력을 위해 여러 대의 저가 서버로 구성되는 클라우드 스토리지 서버에 대한 신뢰성 문제, 비윤리적인 관리자로 인한 데이터가 유출 등의 될 수 있는 등의 보안위협을 가지고 있다.

이러한 클라우드 스토리지 서비스의 신뢰성을 높이기 위해 검색 가능한 암호 기술을 도입하는 연구가 최근 진행되고 있다. 검색 가능 암호 시스템은 암호화된 자료를 복호화하지 않고도 키워드를 가지는 자료를 검색할 수 있도록 하는 암호 기반 기술을 말한다.

검색 가능 암호 시스템에서 암호화의 대상인 정보를 자료(document)라 부른다. 즉, 자료는 사용자가 숨기고 싶은 정보이다. 또한, 사용자가 자신이 원하는 자료를 검색하기 위해 서버에 제공하는 정보를 키워드(keyword)라고 부른다. 일반적으로 자료는 그 자료에 포함된 키워드들의 집합으로 정의된다. 검색 가능 암호 시스템은 다음과 같이 키 생성(key generation), 암호화(build index), 트랩도어 생성(trapdoor generation), 검색(search)의 4가지 단계로

이루어진다.

이러한 형태를 가지는 기존의 검색가능 암호를 클라우드 스토리지에 적용하려면 먼저 클라우드 스토리지의 서비스 특성과 구조를 알아야 한다. 클라우드 스토리지 서비스를 이용하는 사용자는 데이터를 업로드 후 이를 안전하게 저장하고, 공유하고자 하는 사용자와 안전한 데이터 공유가 필요하다. 또한 데이터를 특정 크기의 블록으로 나누어 여러 서버에 분산 저장하는 클라우드 스토리지의 특성을 고려해야 한다. 따라서 본 연구에서는 클라우드 스토리지 서비스 환경을 고려하여 데이터를 안전하게 저장하고 이를 공유하는 방법을 연구하였다.

2. 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.
- 통신량: 클라이언트와 서버간의 에너지 효율 및 네트워크 자원의 효율성을 위하여 통신량이 적어야 한다.
- 연산 효율성: 색인의 생성 및 검색을 수행하기 위한 연산의 효율성이 제공되어야 한다. 또한 색인을 공유하기 위한 연산의 효율성이 제공되어야 한다.

3. 제안방식

앞서 정의된 요구사항을 만족하기 위해 본 논문에서 다음과 같은 방식을 제안한다.

4.1 시스템 계수

먼저 제안방식은 다음과 같은 시스템 계수를 사용한다.

- p: 소수
- G: p를 법으로 하는 덧셈군
- g: G의 생성자
- e: 겹선현 사상, $G\times G\rightarrow G_T$
- sk: 개인키
- pk: 공개키
- w: 키워드
- m: 평문 데이터
- $H_1()$: 해시함수, $\{0,1\}^*\rightarrow G$
- $H_2()$: 해시함수, $\{0,1\}^*\rightarrow G$
- $H_3()$: 해시함수, $G_T\rightarrow\{0,1\}^*$
- T_*: 키워드 *을 검색하는 트랩도어
- $rk_{a\rightarrow b}$: a의 암호문을 b의 암호문으로 변경하는 재 암호화 키

4.2 KeyGen(1^k)

클라우드 스토리지 이용자들은 신뢰기관으로부터 키쌍을 안전하게 전송받는다.

$x\in Z_q$ 선택

$sk=x$ 설정

$pk=g^x$ 설정

4.2 Enc(sk, pk, w, m)

사용자는 키워드 검색 및 사용자간 공유를 위한 재 암호화가 가능한 암호문 E를 다음과 같이 생성한다.

$r\in Z_q$

$A=pk^r$

$B=e(g,g)^{sk\cdot r}$

$C=e(g,H_2(sk))^r\cdot m$

$D=H_3(e(g,H_1(w))^r)$

$E=(A,B,C,D)$암호문으로 출력

4.3 ReKeyGen(sk_a, pk_b)

데이터의 소유주가 자신의 데이터를 다른 사용자에게 공유하고자 할대 재 암호화를 위한 키를 생성한다. 사용자 a가 사용자b에게 데이터를 공유하고자 할 경우 a의 개인키와 b의 공개키로 재 암호화키를 다음과 같이 생성한다.

$$rk_{a\rightarrow b}=pk_b^{-sk_a}=g^{sk_b/sk_a}\bmod p$$

4.4 ReEnc($rk_{a\rightarrow b}$, E_a)$\rightarrow E_b$

클라우드 스토리지 서비스 서버는 사용자로부터 입력된 재 암호화키와 재 암호화 하려는 목표 암호문 그리고 공개키를 가지고 재 암호화를 다음과 같이 수행한다.

$B'=e(A,pk_b^{-sk_a})$

$=e(g^{sk_a\cdot r},g^{sk_b/sk_a})$

$=e(g,g)^{sk_2\cdot r}$

$E_b=(A,B',C,D)$

4.5 TrapdoorGen(sk, w)

데이터를 검색할 사용자는 검색하고자 하는 키워드와 자신의 개인키로 트랩도어를 생성한다.

$T_w=H_1(w)^{1/sk}$

4.6 Test(C, T_w)→'yes'or'no'

사용자는 데이터가 자신이 찾고자하는 키워드를 가지고 있는지 확인하기 위하여, 자신의 공개키와 트랩도어, 암호문을 입력 받아 다음과 같이 테스트를 수행한다.

$D = ?H_3(e(A, T_w))$
$= H_3(e(pk^r, H_1(w)^{1/sk}))$
$= H_3(e(g, H_1(w))^r)$

4.8 Dec(sk, E)→m

데이터 소유자는 자신의 비밀키 sk 그리고 복호화 하고자 하는 암호문을 다음과 같이 복호화 한다.

$C/e(A, H_2(sk))^{1/sk}$
$= m$

4. 제안방식 분석

제안방식은 아래와 같은 요구사항을 만족한다.

- 기밀성: 제안 방식은 페어링을 이용하여 악의적인 제3자가 클라이언트와 서버 간의 통신을 도청한다고 해도 통신 내용을 유추하기 어렵다.
- 검색 속도: 한 번의 페어링연산과 해시 연산만으로 문서에 키워드가 포함되는지 확인할 수 있어 빠른 검색 속도를 제공한다.
- 통신량: 키워드 검색 및 재 암호화를 위해 한 라운드의 통신과정만이 필요여 통신량의 효율성을 제공한다.
- 연산 효율성: 경량화된 페어링 연산을 기반으로 색인을 생성 및 검색하며, 재 암호화 과정을 수행하여 연산의 효율성을 제공한다.

5. 결론

클라우드 스토리지 서비스의 등장으로 많은 사용자들이 해당 서비스를 통해 데이터를 저장 및 접근할 수 있게 되었다. 이러한 저장소에 저장되는 데이터의 안전성을 보장 받고자 최근 클라우드 스토리지에 검색 가능한 암호 기술을 적용하고자 하는 연구가 시작되고 있다. 하지만 기존의 연구된 대부분의 검색 가능한 암호의 경우 주로 이메일 환경을 고려하고 있어 데이터를 공유할 대상을 사전에 정하고, 공유대상을 추가하는데 비효율적인 문제를 가지고 있다. 클라우드 스토리지 환경에서는 사용자 자신이 사용할 데이터를 직접 올리고, 이를 필요에 따라 원하는 사용자와 안전하게 공유하는 유형으로 사용되어 기존방식은 클라우드 스토리지 환경에 적합하지 않다. 따라서 우리는 이러한 클라우드 스토리지 환경을 고려하여 보안 요구사항을 설정하였으며, Rroxy Re-encryption과 검색가능한 재 암호화 기능을 동시에 제공하는 방식을 제안하였다. 제안 방식은 기존 연구에 대비하여 연산량적 효율성을 제공한다. 클라우드 스토리지에서의 데이터를 유연하고 쉽게 검색하기 위해서는 다수의 키워드를 이용한 검색이 중요한 이슈가 될 것으로 사료된다. 따라서 차후에는 가변길이의 다중 키워드로 구성되어 있는 색인을 암호화 하고, 또 이를 유연하게 검색할 수 있는 재 암호화 시스템에 대한 연구가 필요하다.

참고문헌

[1] D. Boneh, G.D. Crescenzo, R. Ostrovsky, and G. Persiano, "Public Key Encryption with Keyword Search," In Advances in Cryptology-Eurocrypt 2004, LNCS 3027, pp. 506-522, 2004.

[2] D. Boneh and X.Boyen, "Efficient Selective-id Secure Identity Based Encryption without Random Oracles," In Advances in Cryptology -EUROCRYPT 2004, LNCS 3027, Springer - Verlag, 2004.

[3] E. Goh and T. Matsuo. "Proposal for P1363.3 Proxy Re-encryption," http://grouper.ieee.org/groups/1363/IBC/submissions/NTTDataProposal-for-P1363.3-2006-08-14.pdf.

[4] M. Green and G. Ateniese, "Identity-Based Proxy Re-encryption," In Applied Cryptography and Network Security'07, LNCS 4521, Springer - Verlag,2007.

[5] P. Golle, J. Staddon, and B. Waters, "Secure conjunctive keyword search over encrypted data," In proceedings of ACNS 2004, LNCS 3089, pp. 31-45, 2004.

[6] T. Matsuo. "Proxy Re-encryption Systems for Identity-Based Encryption," In First International Conference on Pairing-Based Cryptography -Pairing 2007,LNCS 4575, Springer - Verlag,2007.

[7] W. Ogata and K. Kurosawa, "Oblivious Keyword Search," Journal of Complexity, Vol. 20, No. 2-3, pp. 356-371, Apr.-June, 2004.

[8] X. Wang and X. Yang, "Analysis on the Security of an Identity Based Proxy Re-encryption," In Third International Conference on Applied Cryptography and Network Security(ACNS 2005), LNCS 3531, pp. 442-455, 2005.

[9] Y.C. Chang and M. Mitzenmacher, "Privacy preserving keyword searches on remote encrypted data," In Third International Conference on Applied Cryptography and Network Security(ACNS 2005), LNCS 3531, pp. 442-455, 2005.

[10] Y.H. Hwang and P.J. Lee, "Public Key Encryption with Conjunctive Keyword Search and Its Extension to a Multi-user System," In proceedings of Pairing 2007, LNCS 4575, pp. 2-22, 2007.

[11] 이현숙, 박종환, 이동훈, "다중 수신자 환경에서 키워드 검색 가능한 공개키 암호시스템", 정보보호학회논문지, 19(2), pp. 31-38, 2009.