클라우드 환경에서 개인정보보호를 위한 분산 데이터 관리 기법

차정훈*, 강정호**, 박종혁*
서울과학기술대학교 컴퓨터공학과*
배화여자대학교 정보보호과**
e-mail:{ckwjdgns, jhpark1}@seoultech.ac.kr, kjh7548@naver.com

Distributed Information Management Scheme for Privacy in Cloud Environment

Jeonghun Cha*, Jungho Kang**, Jong Hyuk Park*
*Department of Computer Science and Engineering, Seoul National University of Science and Technology
**Department of Information Security, Baewha Women's University

요 약

최근 정보 기술의 발전으로 클라우드 컴퓨팅은 개개인에게 편의성을 제공하도록 기능하지만, 실생활에서 디지털 정보의 의존성을 높이게 되었다. 클라우드 컴퓨팅은 실시간으로 다양한 정보를 교환함으로써 다양한 어플리케이션 서비스를 제공한다. 특히, 사용자가 가지고 있는 정보들을 로컬 서버에 관리하기 어려운 문제를 해결하기 위해 아웃소싱 클라우드 스토리지 서비스를 이용하여 해결할 수 있다. 그러나, 사용자의 데이터를 외부 클라우드 서버에 업로드하여 저장하게 되면, 클라우드 서비스 제공자로 인한 프라이버시 문제가 발생할 수 있다. 최근, 클라우드 서버에서 발생할 수 있는 프라이버시 문제를 해결하기 위해서 사용자의 데이터를 암호화하여 클라우드 서비스 제공자로부터 사용자의 정보를 보호하는 연구가 진행되고 있다. 하지만 이 연구는 시간이 지남에 따라 암호화가 복호화될 수 있으며, 특히 클라우드 서버에서 Offline Bruteforce 공격이 발생할 수 있다. 본 논문에서는 클라우드 환경에서 사용자의 개인정보를 보호하기 위한 기존 연구의 한계점을 분석한다. 기존 연구 분석을 통해 개인정보보호를 위한 요구사항을 도출하고, 이를 기반으로 안전한 분산 데이터 관리 기법에 대해 고찰한다.

1. 서론

최근 정보 기술의 진화는 단순히 사물과 사물이 연결되는 IoT (Internet of Things) 기기들의 폭발적인 증가뿐만 아니라, 스마트 홈과 같이 실생활에 디지털화가 진행되고 있다. 스마트 시티의 이기종 시스템이나 대규모 데이터가 발생하는 환경은 사용자들이 점점 더 복잡하고 정보의 관리가 어렵기 때문에 클라우드 서비스가 종종 이러한 문제를 해결하기 위한 솔루션이 될 수 있다 [1]. 특히, 대규모 데이터 환경 및 관리에 어려움이 발생하기 때문에 사용자들은 아웃소싱 클라우드 스토리지 서비스를 이용한다. 사용자들은 IoT 기기의 증가에 따른 다양한 데이터와 도메인 외부의 실시간 대규모 데이터를 이용하기 위해 아웃소싱 스토리지를 사용할 수 있다. 그러므로, 사용자들은 IoT 기반의 대규모 데이터가 발생하는 환경인 스마트 시티에서 데이터를 효율적으로 관리하기 위해 클라우드 스토리지가 강제될 수 있다. 하지만, 스마트 시티 환경에서 개인 사용자들의 IoT 기기가 발생하는 데이터는 외부에 공개되면 안 되는 개인정보를 포함하고 있을 수 있다. 예를 들어, 가정용 스마트 홈 IP CCTV 데이터가 노출되면, 개인정보 침해와 더불어 심각한 보안 문제가 발생할 수 있다. 그러므로, CSP (Cloud Service Provider)에게 개인정보가 포함된 데이터에 대해 위탁하는 것은 해결해야 하는 보안 과제이다 [2].

CSP는 고객이 저장하는 개개인의 데이터를 관리하기 위해서 사본을 저장할 수 있으며, 고객이 데이터의 삭제를 요구하더라도 데이터의 원본 및 사본의 삭제를 검증할 방법이 부족하다 [1]. 이러한 문제를 해결하기 위해, 클라우드 스토리지에 데이터를 저장할 때, 암호화한 정보를 전송할 수 있다. 하지만, 암호화는 기본적으로 시간이 지남에 따라 취약해질 가능성이 있다. 즉, CSP가 데이터를 삭제하지 않고 보관한다면 언제든지 암호화된 정보에 대해 offline bruteforce와 같은 공격가능성이 존재하게 된다. 또한 사용자는 개인정보보호 문제뿐만 아니라 위탁한 데이터가 유실되어 원본 데이터를 복원하지 못하는 문제점도 발생할 수 있다.

본 논문에서는 클라우드 환경에서 사용자의 개인정보를 보호하기 위한 기존 연구들의 한계점을 분석하고 요구사항을 도출한다. 이를 통해 사용자의 개인정보보호를 위한 분산 데이터 관리 기법에 대해 고찰한다. 기존 연구는 섹션 2에서 설명하고, 섹션 3에서는 분산 데이터 관리 기법에 대해 고찰한다. 마지막으로 섹션 4에서 결론을 짓는다.

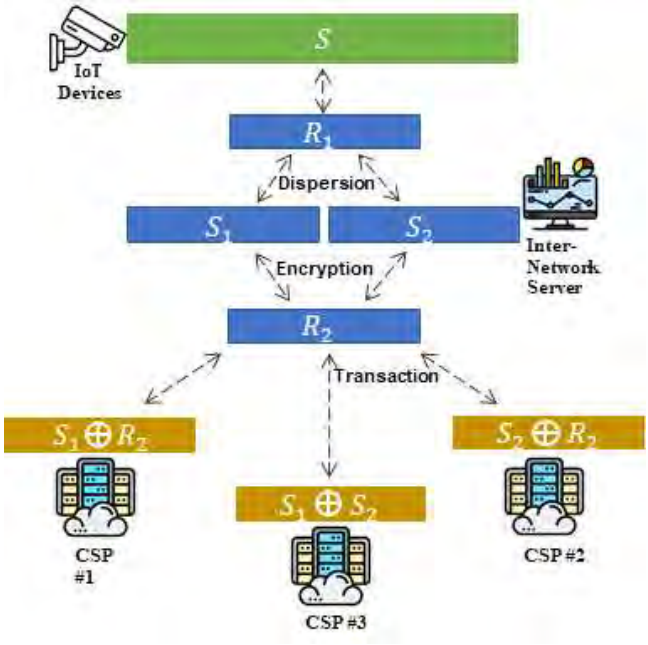
2. 관련 연구

최근 클라우드 컴퓨팅 환경에서 사용자의 개인정보를 보호하기 위해 사용자의 데이터를 암호화하여 클라우드 스토리지에 업로드 하는 연구가 진행되고 있다. [2]는 스토리지 효율성과, 퍼블릭 클라우드 스토리지에서 발생하는 프라이버시 문제를 해결하기 위해, 멀티미디어 데이터를 서로 다른 보호 수준으로 나누어 저장한다. 기밀성이 중요한 데이터 단편 조각은 프라이빗 조각으로써 신뢰할 수 있는 로컬에 저장하고, 기밀성이 낮은 퍼블릭 파편은 외부 퍼블릭 클라우드에 저장한다. 이 연구는 프라이빗 조각이 기밀성이 중요한 정보이기 때문에 반드시 신뢰할 수 있는 도메인에 저장해야 한다. 이 때, 단편화와 분산이라는 개념을 사용한다. 입력 데이터 D는 d1, d2로 나뉘고 데이터의 기밀성이 높은 d1은 암호화 후에 프라이빗 파편이 되고, d2는 퍼블릭 스토리지에 저장된다. 하지만 이 연구는 프라이빗 데이터가 무조건 신뢰할 수 있는 로컬에 저장해야하며, 퍼블릭 스토리지에 저장한 데이터의 무결성 검증을 고려하지 않았다. 결국, 데이터 저장을 안전하게 관리할 수 있으며, 스토리지 자원을 효율적으로 관리할 수 있었지만 프라이빗 데이터를 반드시 신뢰할 수 있는 트러스트 도메인에 저장해야 한다는 문제가 있다.

클라우드 스토리지를 이용하는 데이터 소유자는 CSP가 데이터를 변조하거나 손실에 대한 데이터 무결성을 검증하기 위해, 블록체인 기술이 적용될 수 있다. [3]에서는 블록체인의 낮은 자원 효율성, 복잡성, 제한된 확장성, 높은 오버헤드 및 대기시간 때문에 IoT 환경에 적용하는 것이 부적합하다고 언급했다. 이 논문은 기존의 IoT 데이터 통신에서 부족했던 보안성과 프라이버시 문제를 해결하는 효율적인 블록체인 기술을 제안했다. 데이터 소유자가 데이터와 상호작용 하는 클라이언트의 데이터 액세스를 효율적으로 한다. 하지만, 리소스 서버에 저장하는 스토리지 효율성이나, 퍼블릭 스토리지에서 발생할 수 있는 프라이버시 문제점에 대해 다루지 않았다. [4] 은 블록체인을 이용한 P2P 클라우드 스토리지 접근법을 제안했다. 이 연구는 사용자의 파일을 32MB로 나누고 P2P 네트워크에 개인키를 이용해 암호화된 블록을 업로드 한다. P2P 블록체인 노드들은 암호화된 데이터를 분산화 및 복제하여 데이터 센터에 저장한다. 이 연구는 블록체인을 이용하여 신뢰할 수 있는 CSP가 없어도, P2P 통신을 통해 데이터의 무결성을 보장하는데 초점을 맞췄다. 하지만, P2P 블록체인 트랜잭션을 감소하기 위해 블록에 데이터를 최소화 하였지만, 실시간 대규모 데이터 통신 환경의 속도 측면을 고려하지 않았다. 추가적으로 데이터 유실에 대응하기 위해 동일한 크기의 복제본을 만들기 때문에 스토리지 효율성을 고려하지 않았다.

3. 개인정보보호를 위한 분산 데이터 관리 기법

기존 연구 분석을 기반으로 안전한 정보 관리 기법을 설계하기 위해서 프라이버시, 효율성, 분산 저장의 조건을 만족시켜야 한다. 그림 1은 요구사항을 충족시키기 위한 분산 정보 관리 기법이다.



(그림 1) 개인정보보호를 위한 분산 정보 관리 기법

기존 연구를 기반으로 도출한 요구사항을 통해 제안한 분산 정보 관리 기법은 분산, 암호화 두 가지 방식으로 작동한다. 먼저, 원본 데이터 S는 분산화 과정을 통해 S1, S2로 데이터를 나누어 원본 데이터를 유추하기 어렵게 한다. 그 이후 비밀 조각인 S1, S2는 로컬 서버에서 생성한 개인키를 기반으로 암호화하고 서로 다른 CSP에게 업로드한다. 제안한 개인정보보호를 위한 분산 정보 관리 기법은 CSP에서 발생할 수 있는 프라이버시 문제를 분산화하여 예방할 수 있으며, 분산화를 통해 데이터의 크기가 늘어나지 않아 효율적으로 저장할 수 있다.

4. 결론

최근 정보 기술 발전에 따라, 클라우드 스토리지 서비스는 사용자의 데이터를 효율적으로 관리하기 위한 필수 솔루션이다. 하지만, 클라우드 스토리지 서비스 제공자에 의해 개인정보가 포함된 민감 데이터가 유출될 수 있으며, 데이터의 손실에 대한 무결성을 검증할 방안이 부족하다. 본 논문에서는 개인정보보호를 위한 기존 연구들을 분석하여 요구사항들을 도출하였고, 이를 기반으로 안전한 분산 데이터 관리 기법에 대해 제안했다. 제안한 분산 데이터 관리 기법은 클라우드 서비스에서 발생할 수 있는 데이터 손실을 방지하기 위해 블록체인을 사용한 무결성 검증과 원본 민감 데이터를 안전하게 보호하기 위한 비밀 공유 방안에 대해 고찰했다. 향후 연구에서는 제안한 분산 데이터 관리 기법을 효율적이고 스마트 시티와 같은 대규모 이기종 시스템에서 적용할 수 있는 분산 공유 알고리즘을 연구할 계획이다.

Acknowledgement

This study was supported by the Advanced Research Project funded by the SeoulTech(Seoul National University of Science and Technology).

참고문헌

[1] Sookhak, Mehdi, et al. "Security and privacy of sma rt cities: a survey, research issues and challenges." IEE E Communications Surveys & Tutorials 21.2 (2018): 171 8-1743.

[2] Qiu, Han, et al. "A user-centric data protection met hod for cloud storage based on invertible DWT." IEEE Transactions on Cloud Computing (2019).

[3] Mohanty, Sachi Nandan, et al. "An efficient Lightwe ight integrated Blockchain (ELIB) model for IoT securit y and privacy." Future Generation Computer Systems 1 02 (2020): 1027-1037.

[4] Li, Jiaxing, Jigang Wu, and Long Chen. "Block-secu re: Blockchain based scheme for secure P2P cloud stora ge." Information Sciences 465 (2018): 219-231.