



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2016년02월29일

(11) 등록번호 10-1595065

(24) 등록일자 2016년02월11일

(51) 국제특허분류(Int. Cl.)

G06F 21/00 (2006.01) G06F 21/60 (2013.01)

H04L 9/00 (2006.01)

(21) 출원번호 10-2014-0105931

(22) 출원일자 2014년08월14일

심사청구일자 2014년08월14일

(56) 선행기술조사문헌

KR1020140034222 A

(73) 특허권자

고려대학교 산학협력단

서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)

(72) 발명자

엄지은

서울 서대문구 가재울로 45, 102동 102호 (남가좌동, 현대아파트)

서민혜

서울 양천구 목동중앙남로 21-15, 101동 703호 (목동, 에이스아파트)

이동훈

서울 종로구 사직로8길 34, 1404호 (내수동, 경희궁의아침3단지)

(74) 대리인

김동용

전체 청구항 수 : 총 6 항

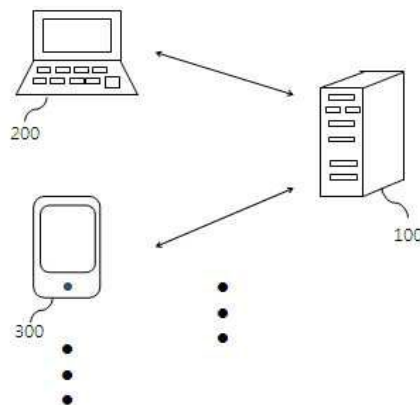
심사관 : 박진아

(54) 발명의 명칭 데이터 중복 제거 서버 및 데이터 중복 제거 방법

(57) 요약

중복 제거 방법이 개시된다. 상기 중복 제거 방법은, 서버가 단말기로부터 태그를 수신하는 단계, 상기 태그를 이용하여 상기 단말기가 업로드 하려는 메시지에 대한 암호문의 중복 여부를 결정하는 단계, 상기 암호문이 중복되지 않는 경우, 상기 단말기로 상기 메시지에 대한 업로드 요청을 송신하는 단계, 및 상기 단말기로부터 상기 암호문을 수신하는 단계를 포함한다.

대표도 - 도1



10

이 발명을 지원한 국가연구개발사업

과제고유번호 2013060756

부처명 미래창조과학부

연구관리전문기관 한국연구재단

연구사업명 중견연구자지원_도약연구(전략)

연구과제명 속성 기반 공개키 시스템에 관한 연구

기 여 율 1/1

주관기관 고려대학교 산학협력단

연구기간 2013.09.01 ~ 2014.08.31

명세서

청구범위

청구항 1

서버가 단말기로부터 태그와 질의값을 수신하는 단계;
 상기 태그를 이용하여 상기 단말기가 업로드 하려는 메시지에 대한 암호문의 중복 여부를 결정하는 단계;
 상기 암호문이 중복되지 않는 경우,
 상기 단말기로 상기 메시지에 대한 업로드 요청을 송신하는 단계;
 상기 단말기로부터 상기 암호문을 수신하는 단계;
 상기 암호문이 중복되는 경우,
 상기 질의값과 상기 암호문을 입력으로 하는 해시값인 응답값을 생성하는 단계; 및
 상기 응답값을 상기 단말기로 송신하는 단계를 포함하는 중복 제거 방법.

청구항 2

제1항에 있어서,
 상기 태그는 상기 메시지를 입력으로 하는 해시값인 암호화키를 입력으로 하는 해시값인,
 중복 제거 방법.

청구항 3

삭제

청구항 4

제1항에 있어서,
 상기 암호문이 중복되는 경우, 상기 단말기로부터 중복 제거 요청을 수신하는 단계를 더 포함하는 중복 제거 방법.

청구항 5

단말기가 메시지에 대한 해시값인 암호화키와 상기 암호화키에 대한 해시값인 태그를 생성하는 단계;
 상기 단말기가 질의값과 상기 태그를 서버로 송신하는 단계;
 상기 단말기가 상기 서버로부터 상기 메시지에 대한 암호문과 상기 질의값을 입력으로 하는 해시값인 응답값을 수신하는 단계;
 상기 단말기가 상기 암호문과 상기 질의값을 입력으로 하는 해시값과 상기 응답값을 비교하여, 상기 서버에 상기 메시지에 대한 암호문이 존재하는 여부를 검증하는 단계; 및
 상기 단말기가, 상기 메시지에 대한 암호문과 상기 질의값을 입력으로 하는 해시값과 상기 응답값이 일치하는 경우, 중복 제거 요청을 상기 서버로 송신하는 단계를 포함하는 중복 제거 방법.

청구항 6

제5항에 있어서,
 상기 메시지에 대한 암호문과 상기 질의값을 입력으로 하는 해시값과 상기 응답값이 일치하지 않는 경우, 중복 저장요청을 상기 서버로 송신하는 단계를 더 포함하는, 중복 제거 방법.

청구항 7

DB(database);

단말기로부터 태그를 수신하고, 상기 태그를 이용하여 상기 단말기가 업로드 하려는 메시지에 대한 암호문이 상기 DB에 저장되어 있는지 여부를 결정하기 위한 중복 확인 모듈; 및

상기 암호문이 상기 DB에 저장되어 있지 않은 경우, 상기 메시지에 대한 업로드 요청을 상기 단말기로 송신하고, 상기 단말기로부터 상기 암호문을 수신하기 위한 중복 제거 모듈을 포함하고,

상기 암호문이 상기 DB에 저장되어 있는 경우,

상기 중복 확인 모듈은 상기 단말기로부터 수신된 질의값과 상기 암호문을 입력으로 하는 해시값인 응답값을 생성하고, 상기 응답값을 상기 단말기로 송신하고,

상기 중복 제거 모듈은 상기 단말기로부터 중복 제거 요청 또는 중복 저장 요청을 수신하는, 데이터 중복 제거 서버.

발명의 설명

기술 분야

[0001] 본 발명의 개념에 따른 실시 예는 데이터 중복 제거 서버에 관한 것으로, 특히 서버에 암호문이 존재하는지 여부를 확인함으로써 사용자의 데이터가 분실되는 것을 방지할 수 있는 데이터 중복 제거 서버 및 데이터 중복 제거 방법에 관한 것이다.

배경 기술

[0002] 데이터 중복제거(data deduplication)는 내용이 동일한 데이터를 하나만 저장하여 서버의 저장 공간을 효율적으로 관리하는 기술로써, 클라우드 스토리지(cloud storage)와 같이 동일한 데이터가 대량으로 삽입(또는 저장)되는 환경에서 유용하게 활용될 수 있다. 그러나 최근 신뢰할 수 없는 서버에 대한 보안 문제가 부각되면서 데이터를 암호화하여 저장하는 것이 필수적으로 요구되는데, 암호화는 데이터를 난수화하기 때문에 서버에서 해당 암호문이 동일한 데이터에 대한 정보인지를 판단하기 어려우며 암호화에 사용된 키를 공유하는 문제도 발생한다. 이를 해결하기 위해 암호화에 사용되는 키나 동일성 검사에 사용되는 태그를 데이터로부터 결정적(deterministic)으로 생성되도록 구성하는 기술들이 제안되고 있다.

[0003] 데이터 중복 제거 기법은 데이터 중복 제거를 위해 수행되는 프로토콜에 따라 사용자가 직접 데이터를 암호화하여 서버에 업로드하면 서버 측에서 동일성 여부를 판단한 뒤 하나의 데이터만 저장하는 서버측 중복제거(server-side deduplication)와 사용자가 서버에 동일한 데이터가 사전에 저장되어 있는지 확인한 뒤, 저장되어 있으면 해당 데이터에 접근할 수 있는 권한만 획득하고 데이터를 업로드하지 않는 사용자측 중복제거(client-side deduplication)로 나눌 수 있다. 전자의 경우 서버에 저장 공간에 대한 효율성만 제공하는 반면에 후자의 경우 저장 공간 효율성뿐만 아니라 사용자가 서버에 암호문을 업로드하는 전송량을 줄일 수 있는 장점이 있다.

선행기술문헌

비특허문헌

[0004] (비특허문헌 0001) Message-Locked Encryption and Secure Deduplication(M. Bellare, et al., EUROCRYPT 2013)

발명의 내용

해결하려는 과제

[0005] 본 발명이 이루고자 하는 기술적인 과제는 데이터를 전송하는 과정에서 서버에 암호문의 존재 여부를 확인함으

로써 사용자의 데이터가 분실되는 것을 방지할 수 있는 중복 제거 서버 및 중복 제거 방법을 제공하는 것이다.

과제의 해결 수단

[0006] 본 발명의 일 실시 예에 따른 데이터 중복 제거 방법은, 서버가 단말기로부터 태그를 수신하는 단계, 상기 태그를 이용하여 상기 단말기가 업로드 하려는 메시지에 대한 암호문의 중복 여부를 결정하는 단계, 상기 암호문이 중복되지 않는 경우, 상기 단말기로 상기 메시지에 대한 업로드 요청을 송신하는 단계, 및 상기 단말기로부터 상기 암호문을 수신하는 단계를 포함한다.

[0007] 또한, 본 발명의 다른 실시 예에 따른 데이터 중복 제거 방법은, 단말기가 메시지에 대한 해시값인 암호화키와 상기 암호화키에 대한 해시값인 태그를 생성하는 단계, 상기 단말기가 임의의 난수인 질의값과 상기 태그를 서버로 송신하는 단계, 상기 단말기가 상기 서버로부터 상기 질의값에 대응하는 응답값을 수신하는 단계, 상기 단말기가 상기 메시지에 대한 암호문과 상기 질의값을 입력으로 하는 해시값과 상기 응답값을 비교하여, 상기 서버에 상기 메시지에 대한 암호문이 존재하는 여부를 검증하는 단계, 및 상기 단말기가, 상기 메시지에 대한 암호문과 상기 질의값을 입력으로 하는 해시값과 상기 응답값이 일치하는 경우, 중복 제거 요청을 상기 서버로 송신하는 단계를 포함한다.

[0008] 또한, 본 발명의 일 실시 예에 따른 데이터 중복 제거 서버는 DB, 단말기로부터 태그를 수신하고, 상기 태그를 이용하여 상기 단말기가 업로드 하려는 메시지가 상기 DB에 저장되어 있는지 여부를 결정하기 위한 중복 확인 모듈, 및 상기 암호문이 상기 DB에 저장되어 있지 않은 경우, 상기 메시지에 대한 업로드 요청을 상기 단말기로 송신하고, 상기 단말기로부터 상기 암호문을 수신하기 위한 중복 제거 모듈을 포함한다.

발명의 효과

[0009] 본 발명의 실시 예에 따른 데이터 중복 제거 서버는 서버에 암호문의 존재 여부를 확인함으로써 중복 위장 공격(duplicate faking attack)으로 인해 사용자의 데이터가 분실되는 것을 방지할 수 있는 효과가 있다.

도면의 간단한 설명

[0010] 본 발명의 상세한 설명에서 인용되는 도면을 보다 충분히 이해하기 위하여 각 도면의 상세한 설명이 제공된다.

도 1은 본 발명의 일 실시 예에 따른 중복 제거 시스템을 도시한다.

도 2는 도 1에 도시된 중복 제거 서버의 기능 블록도이다.

도 3과 도 4는 도 1에 도시된 중복 제거 시스템의 동작 방법을 설명하기 위한 흐름도로써, 도 3은 최초 업로더를 대상으로 하는 동작 방법을 설명하기 위한 흐름도이고, 도 4는 이차 사용자를 대상으로 하는 동작 방법을 설명하기 위한 흐름도이다.

발명을 실시하기 위한 구체적인 내용

[0011] 본 명세서에 개시되어 있는 본 발명의 개념에 따른 실시 예들에 대해서 특정한 구조적 또는 기능적 설명은 단지 본 발명의 개념에 따른 실시 예들을 설명하기 위한 목적으로 예시된 것으로서, 본 발명의 개념에 따른 실시 예들은 다양한 형태로 실시될 수 있으며 본 명세서에 설명된 실시 예들에 한정되지 않는다.

[0012] 본 발명의 개념에 따른 실시 예들은 다양한 변경들을 가할 수 있고 여러 가지 형태들을 가질 수 있으므로 실시 예들을 도면에 예시하고 본 명세서에서 상세하게 설명하고자 한다. 그러나, 이는 본 발명의 개념에 따른 실시 예들을 특정한 개시 형태들에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물, 또는 대체물을 포함한다.

[0013] 제1 또는 제2 등의 용어는 다양한 구성 요소들을 설명하는데 사용될 수 있지만, 상기 구성 요소들은 상기 용어들에 의해 한정되어서는 안 된다. 상기 용어들은 하나의 구성 요소를 다른 구성 요소로부터 구별하는 목적으로만, 예컨대 본 발명의 개념에 따른 권리 범위로부터 벗어나지 않은 채, 제1 구성 요소는 제2 구성 요소로 명명될 수 있고 유사하게 제2 구성 요소는 제1 구성 요소로도 명명될 수 있다.

[0014] 어떤 구성 요소가 다른 구성 요소에 "연결되어" 있다거나 "접속되어" 있다고 언급된 때에는, 그 다른 구성 요소에 직접적으로 연결되어 있거나 또는 접속되어 있을 수도 있지만, 중간에 다른 구성 요소가 존재할 수도 있다고 이해되어야 할 것이다. 반면에, 어떤 구성 요소가 다른 구성 요소에 "직접 연결되어" 있다거나 "직접 접속되어" 있다고 언급된 때에는 중간에 다른 구성 요소가 존재하지 않는 것으로 이해되어야 할 것이다. 구성 요소들 간의

관계를 설명하는 다른 표현들, 즉 "~사이에"와 "바로 ~사이에" 또는 "~에 이웃하는"과 "~에 직접 이웃하는" 등도 마찬가지로 해석되어야 한다.

[0015] 본 명세서에서 사용한 용어는 단지 특정한 실시 예를 설명하기 위해 사용된 것으로서, 본 발명을 한정하려는 의도가 아니다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 명세서에서, "포함하다" 또는 "가지다" 등의 용어는 본 명세서에 기재된 특징, 숫자, 단계, 동작, 구성 요소, 부분품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성 요소, 부분품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.

[0016] 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 모든 용어들은 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가진다. 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 관련 기술의 문맥상 가지는 의미와 일치하는 의미를 갖는 것으로 해석되어야 하며, 본 명세서에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.

[0017] 이하, 본 명세서에 첨부된 도면들을 참조하여 본 발명의 실시 예들을 상세히 설명한다.

[0018] 도 1은 본 발명의 일 실시 예에 따른 중복 제거 시스템을 도시한다.

[0019] 도 1을 참조하면, 중복 제거 시스템(10)은 중복 제거 서버(100), 제1 단말기(200), 및 제2 단말기(300)를 포함한다. 도 1에는 두 개의 단말기(200, 300)만이 도시되어 있으나, 중복 제거 시스템(10)에 포함되는 단말기는 두 개 이상으로 구현될 수 있다.

[0020] 중복 제거 서버(100)는 제1 사용자의 단말기, 즉 제1 단말기(200)와 제2 사용자의 단말기, 즉 제2 단말기(300) 각각으로부터 데이터를 수신하고, 수신된 데이터를 내부의 DB(database)에 저장한다. 이 때, 중복 제거 서버(100)는 저장하려는 데이터가 상기 DB에 이미 저장되어 있는 경우, 데이터 중복 제거(data deduplication)를 수행할 수 있다. 즉, 중복 제거 서버(100)는 상기 DB에 저장된 데이터와 동일한 데이터를 수신하지 않음으로써, 상기 DB의 저장 공간을 효율적으로 활용할 수 있다.

[0021] 제1 단말기(200)와 제2 단말기(300)는 PC(personal computer), 태블릿(tablet) PC, 노트북(notebook), 넷-북(net-book), e-리더(e-reader), PDA(personal digital assistant), PMP(portable multimedia player), MP3 플레이어, 또는 MP4 플레이어와 같은 데이터 처리 장치로 구현될 수 있다. 실시 예에 따라, 제1 단말기(200)와 제2 단말기(300)는 모바일폰(mobile phone), 또는 스마트폰(smart phone)과 같은 핸드헬드 장치(handheld device)로 구현될 수 있다.

[0022] 본 발명은 메시지 고정 암호(Message-Locked Encryption) 기법을 이용하여 데이터 중복 제거를 수행한다. 메시지 고정 암호 기법의 하나로써, Bellare 등이 제안한 HCE(Hash and Convergent Encryption) 기법은 PGen 알고리즘(시스템 파라미터 생성 알고리즘), KGen 알고리즘(키 생성 알고리즘), TGen 알고리즘(태그 생성 알고리즘), Enc 알고리즘(암호화 알고리즘), 및 Dec 알고리즘(복호화 알고리즘)으로 구성된다.

[0023] HCE 기법에 포함된 각각의 알고리즘에 대해 설명하면, PGen 알고리즘은 다음과 같이 정의될 수 있다.

[0024] $PGen(1^\lambda) \rightarrow P$

[0025] 즉, PGen 알고리즘은 보안상수(λ)를 입력받아 시스템 파라미터(P)를 생성하는 알고리즘이다. 시스템 파라미터(P)는 암호학적 해시함수(H)와 암호화에 사용되는 결정적 대칭키 암호(Deterministic Symmetric Encryption; SE)에 대한 정의를 포함한다.

[0026] KGen 알고리즘은 다음과 같이 정의될 수 있다.

[0027] $KGen(P, M) \rightarrow K$

[0028] 즉, KGen 알고리즘은 시스템 파라미터(P)와 메시지(M)을 입력받아 키(K)를 생성하는 알고리즘이다. 메시지(M)에 대한 해시값을 생성하여 H(M)을 키(K)로 출력한다.

[0029] TGen 알고리즘은 다음과 같이 정의될 수 있다.

[0030] $TGen(P, K) \rightarrow T$

- [0031] 즉, TGen 알고리즘은 시스템 파라미터(P)와 키(K)를 입력받아 태그(T)를 생성하는 알고리즘이다. 키(K)에 대한 해시값을 생성하여 H(K)를 태그(T)로 출력한다.
- [0032] Enc 알고리즘은 다음과 같이 정의될 수 있다.
- [0033] $Enc(P,K,M) \rightarrow C$
- [0034] 즉, Enc 알고리즘은 시스템 파라미터(P), 키(K), 및 메시지(M)을 입력받아 메시지(M)를 암호화하는 알고리즘이다. 초기에 정의된 대칭키 암호를 사용하여 생성된 암호문을 출력한다.
- [0035] Dec 알고리즘은 다음과 같이 정의될 수 있다.
- [0036] $Dec(P,K,C) \rightarrow M$
- [0037] 즉, Dec 알고리즘은 시스템 파라미터(P), 키(K), 및 암호문(C)을 입력받아 복호화하는 알고리즘이다.
- [0038] 상기와 같은 메시지 고정 암호 기법은 메시지로부터 추출된 키를 사용하여 결정적 대칭키 암호화를 수행하기 때문에 동일한 메시지는 언제나 동일한 암호문을 생성한다.
- [0039] 메시지 고정 암호 기법에 대한 보다 자세한 내용은 "Message-Locked Encryption and Secure Deduplication(M. Bellare, et al., EUROCRYPT 2013)"가 참조될 수 있다.
- [0040] 또한, 본 발명에서 이용되는 해시함수, 특히 충돌 저항(collision resistance) 해시 함수는 $H:\{0,1\}^* \rightarrow \{0,1\}^n$ (SHA256의 경우 $n=256$)로 정의될 수 있다.
- [0041] 도 2는 도 1에 도시된 중복 제거 서버의 기능 블록도이다.
- [0042] 도 1과 도 2를 참조하면, 중복 제거 서버(10)는 중복 확인 모듈(110), 중복 제거 모듈(130), 및 DB(150)를 포함한다.
- [0043] 중복 확인 모듈(110)은 제1 단말기(200) 또는 제2 단말기(300)로부터 태그(Tag)와 질의값(R)을 수신한다. 중복 확인 모듈(110)은 수신된 태그(Tag)를 이용하여 제1 단말기(200) 또는 제2 단말기(300)가 업로드 하려는 데이터의 중복 여부, 즉 상기 데이터가 DB(150)에 저장되어 있는지 여부를 확인한다.
- [0044] 중복 확인 모듈(110)은 소정의 룩업 테이블(lookup table)을 이용하여 데이터의 중복 여부를 확인할 수 있다. 즉, 중복 확인 모듈(110)은 수신된 태그(Tag)가 상기 룩업 테이블에 포함되어 있는 경우 상기 업로드하려는 데이터와 동일한 데이터가 DB(150)에 저장되어 있는 것으로 결정하고, 수신된 태그(Tag)가 상기 룩업 테이블에 포함되어 있지 않은 경우 상기 업로드하려는 데이터와 동일한 데이터가 DB(150)에 저장되어 있지 않은 것으로 결정할 수 있다.
- [0045] 또한, 중복 확인 모듈(110)은 질의값(R)에 대한 응답으로 응답값(H(CT,R))을 제1 단말기(200) 또는 제2 단말기(300)로 송신함으로써, 중복 제거 서버(100)에 암호문이 존재함을 증명할 수 있다.
- [0046] 중복 제거 모듈(130)은 제1 단말기(200) 또는 제2 단말기(300)로 데이터의 업로드 요청(Req1)을 송신하고, 제1 단말기(200) 또는 제2 단말기(300)로부터 암호문(CT)을 수신할 수 있다. 즉, 중복 제거 모듈(130)은 업로드하려는 데이터가 중복 저장되어 있지 않다고 판단되는 경우, 데이터의 업로드를 요청하는 업로드 요청(Req1)을 제1 단말기(200) 또는 제2 단말기(300)로 송신할 수 있다. 또한, 업로드 요청(Req1)을 수신한 제1 단말기(200) 또는 제2 단말기(300)는 메시지에 대한 암호문(CT)을 중복 제거 모듈(130)로 송신할 수 있다. 암호문(CT)을 수신한 중복 제거 모듈(130)은 수신된 암호문(CT)을 DB(150)에 저장할 수 있다. 이때, 중복 제거 모듈(130)은 룩업 테이블을 업데이트하여, 이후에 암호문(CT)의 중복 저장 여부를 판단 때 이용할 수 있다.
- [0047] 실시 예에 따라, 중복 제거 모듈(130)은 제1 단말기(200) 또는 제2 단말기(300)로부터 중복 제거 요청(Req2) 또는 중복 저장 요청(Req3)을 수신할 수 있다.
- [0048] 구체적으로, 중복 제거 모듈(130)이 중복 제거 요청(Req2)을 수신한 경우는, 업로드하려는 데이터가 DB(150)에 이미 저장되어 있는 경우이다. 이 때, 중복 제거 모듈(130)은 업로드하려는 데이터를 수신하지 않음으로써, 저장 공간을 효율적으로 활용할 수 있다.
- [0049] 중복 제거 모듈(130)이 중복 저장 요청(Req3)을 수신한 경우는, 업로드하려는 데이터가 DB(150)에 저장되어 있지 않은 경우이다. 이 때, 중복 제거 모듈(130)을 제1 단말기(200) 또는 제2 단말기(300)로부터 메시지에 대한

암호문(CT)을 수신하고, 수신된 암호문(CT)을 DB(150)에 저장한다. 또한, 암호문(CT)은 중복 제거 모듈(130)이 제1 단말기(200) 또는 제2 단말기(300)로 송신한 업로드 요청(Req1)에 대한 응답으로 수신될 수도 있다.

[0050] 실시 예에 따라, 중복 제거 모듈(130)은 업로드하려는 데이터가 DB(150)에 이미 저장되어 있는 경우라도, 제1 단말기(200) 또는 제2 단말기(300)의 중복 저장 요청(Req3)에 응답하여, 메시지에 대한 암호문(CT)을 수신하고, 수신된 암호문(CT)을 DB(150)에 저장할 수도 있다.

[0051] 도 2에 도시된 중복 제거 서버(100)의 구성들 각각은 기능 및 논리적으로 분리될 수 있음으로 나타내는 것이며, 반드시 각각의 구성이 별도의 물리적 장치로 구분되거나 별도의 코드로 작성됨을 의미하는 것이 아님을 본 발명의 기술분야의 평균적 전문가가 용이하게 추론할 수 있을 것이다.

[0052] 또한, 본 명세서에서 모듈이라 함은, 본 발명의 기술적 사상을 수행하기 위한 하드웨어 및 상기 하드웨어를 구동하기 위한 소프트웨어의 기능적, 구조적 결합을 의미할 수 있다. 예컨대, 상기 모듈은 소정의 코드와 상기 소정의 코드가 수행되기 위한 하드웨어 리소스의 논리적인 단위를 의미할 수 있으며, 반드시 물리적으로 연결된 코드를 의미하거나, 한 종류의 하드웨어를 의미하는 것이 아니다.

[0053] 또한, 본 명세서에서 DB라 함은, 각각의 DB에 대응되는 정보를 저장하는 소프트웨어 및 하드웨어의 기능적 구조적 결합을 의미할 수 있다. DB는 적어도 하나의 테이블로 구현될 수도 있으며, 상기 DB에 저장된 정보를 검색, 저장, 및 관리하기 위한 별도의 DBMS(Database Management System)를 더 포함할 수도 있다. 또한, 링크드 리스트(linked-list), 트리(Tree), 관계형 DB의 형태 등 다양한 방식으로 구현될 수 있으며, 상기 DB에 대응되는 정보를 저장할 수 있는 모든 데이터 저장매체 및 데이터 구조를 포함한다.

[0054] 도 3과 도 4는 도 1에 도시된 중복 제거 시스템의 동작 방법을 설명하기 위한 흐름도로써, 도 3은 최초 업로더를 대상으로 하는 동작 방법을 설명하기 위한 흐름도이고, 도 4는 이차 사용자를 대상으로 하는 동작 방법을 설명하기 위한 흐름도이다.

[0055] 도 1 내지 도 4를 참조하면, 제1 단말기(200)는 메시지(M1)에 대한 해시값으로 암호화키($K1=H(M1)$)를 생성하고, 메시지(M1)와 암호화키(K1)에 대한 해시값으로 태그($Tag1=H(M1, K1)$)를 생성한다(S110). 본 발명에서 이용되는 태그와 암호화키는 데이터로부터 결정적(deterministic)이다. 즉, 암호화키와 태그는 데이터에 따라 결정되므로 동일한 데이터를 소유하는 자는 누구라도 동일한 암호화키와 태그를 생성할 수 있다.

[0056] 또한, 제1 단말기(200)는 질의(challenge)값으로 사용할 난수(R1)를 선택하고, 태그(Tag1)와 난수(R1)를 중복 제거 서버(100)로 전송한다(S130). 상기 질의값은 중복 제거 서버(100)가 메시지(M1) 또는 메시지(M1)에 대한 암호문(CT1)을 보유하고 있는지 여부를 확인하기 위한 용도로 활용된다.

[0057] 중복 제거 서버(100)는 태그(Tag1)와 난수(R1)를 수신하고, 수신된 태그(Tag1)에 기초하여 중복 제거 여부를 결정한다(S150). 즉, 수신된 태그(Tag1)의 값과 동일한 값이 존재하는 경우, 제1 단말기(100)의 사용자, 즉 제1 사용자는 최초 업로더(first uploader)가 되고, 수신된 태그(Tag1)와 동일한 값이 존재하지 않는 경우, 상기 제1 사용자는 이차 사용자(secondary user)가 된다. 이와 같이, 태그(Tag1)를 통하여 메시지의 중복 여부가 판단될 수 있다. 또한, 상기 비교는 룩업 테이블을 통하여 수행될 수 있다.

[0058] 상기 제1 사용자가 최초 업로더로 결정된 경우, 중복 제거 서버(100)는 제1 단말기(200)로 데이터의 업로드 요청(Req1)을 송신한다(S170).

[0059] 업로드 요청(Req1)을 수신한 제1 단말기(200)는 암호화키(K1)를 이용하여 메시지(M1)를 암호화하여 암호문($CT1=SE(K1, M1)$)을 생성할 수 있다. 또한, 제1 단말기(200)는 생성된 암호문(CT1)을 중복 제거 서버(100)로 송신한다(S190).

[0060] 암호문(CT1)을 수신한 중복 제거 서버(100)는 수신된 암호문(CT1)과 태그(Tag1)를 DB에 저장한다.

[0061] 이와 같은 과정을 통하여 최초 업로더의 데이터는 중복 제거 서버(100)에 저장된다. 이하에서는 제2 단말기(300)를 이차 사용자로 가정하여 기술하기로 한다.

[0062] 제2 단말기(300)는 메시지(M2)에 대한 해시값으로 암호화키($K2=H(M2)$)를 생성하고, 메시지(M2)와 암호화키(K2)에 대한 해시값으로 태그($Tag2=H(M2, K2)$)를 생성한다(S310). 또한, 제2 단말기(300)는 질의값으로 사용할 난수(R2)를 선택한 뒤, 태그(Tag2)와 질의값(R2)를 중복 제거 서버(100)로 전송한다(S330).

[0063] 태그(Tag2)와 질의값(R2)을 수신한 중복 제거 서버(100)는 수신된 태그(Tag2)를 이용하여 제2 단말기(300)가 업로드하려는 데이터(M2)가 DB(150)에 이미 저장되어 있는지 여부를 결정한다(S350).

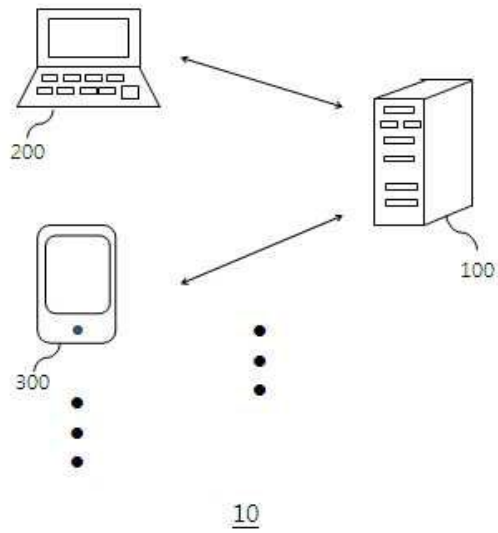
- [0064] 태그(Tag2)와 일치하는 태그가 존재하는 경우, 제2 단말기(300)의 사용자, 즉 제2 사용자는 최초 업로더가 된다. 이 경우, 상술한 바와 같이, 제1 단말기(100)와의 송수신 과정을 반복하게 된다. 태그(Tag2)와 일치하는 태그가 존재하지 않는 경우, 상기 제2 사용자는 이차 사용자가 되므로, 중복 제거 수행 여부가 결정되어야 한다.
- [0065] 우선, 중복 제거 서버(100)는 메시지(M2)의 존재 여부를 증명하기 위하여, 제2 단말기(300)로부터 수신한 질의 값(R2)과 메시지(M2)에 대한 암호문(CT2)에 대한 해시값을(H(CT,R2))을 생성한다. 중복 제거 서버(100)는 생성된 해시값을 응답값으로 하여 제2 단말기(300)로 송신한다(S370).
- [0066] 제2 단말기(300)는 메시지(M2)가 중복 제거 서버(100)에 존재하는지 여부, 즉 메시지의 중복 여부를 검증한다(S380). 구체적으로, 제2 단말기(300)는 암호화키(K2)로 대칭키 암호화하여 메시지(M2)에 대한 암호문(CT2=SE(K2,M2))을 생성한다. 제2 단말기(300)는 암호문(CT2)과 질의값(R2)을 해시함수의 입력값으로 하여 H(CT2,R2)를 생성한다. 제2 단말기(300)는 생성된 해시값(H(CT2,R2))과 중복 제거 서버(100)로부터 수신된 응답값(H(CT2,R2))을 비교함으로써, 메시지(M2)가 중복 제거 서버(100)에 존재하는지 여부를 검증할 수 있다. 즉, 제2 단말기(300)는 생성된 해시값(H(CT2,R2))과 수신된 응답값(H(CT2,R2))의 값이 동일한 경우, 메시지(M2)가 중복 제거 서버(100)에 존재하는 것으로 판단하고, 생성된 해시값(H(CT2,R2))과 수신된 응답값(H(CT,R2))이 동일하지 않는 경우, 메시지(M2)가 중복 제거 서버(100)에 존재하지 않는 것으로 판단할 수 있다.
- [0067] 제2 단말기(300)는 생성된 해시값(H(CT2,R2))과 수신된 응답값(H(CT,R2))이 동일한 경우, 중복 제거 요청(Req 2)을 중복 제거 서버(100)로 전송하고, 생성된 해시값(H(CT2,R2))과 수신된 응답값(H(CT,R2))이 동일하지 않은 경우, 중복 저장 요청(Req3)을 중복 제거 서버(100)로 전송한다(S390). 또한, 중복 제거 서버(100)는 제2 단말기(300)로부터 메시지(M2)에 대한 암호문(CT2)을 수신하고, 수신된 암호문(CT2)을 DB(150)에 저장한다. 이 때, 제2 단말기(300)는 중복 제거 서버(100)가 제2 단말기(300)로 송신한 업로드 요청에 대한 응답으로 암호문(CT 2)을 송신할 수도 있다.
- [0068] 상술한 바와 같이, 중복 제거 시스템(10)은 사용자측 중복 제거 기술에서 악의적인 최초 업로더의 중복 위장 공격으로부터 이차 사용자의 데이터를 보호할 수 있다.
- [0069] 본 발명은 도면에 도시된 실시 예를 참고로 설명되었으나 이는 예시적인 것에 불과하며, 본 기술 분야의 통상의 지식을 가진 자라면 이로부터 다양한 변형 및 균등한 타 실시 예가 가능하다는 점을 이해할 것이다. 따라서, 본 발명의 진정한 기술적 보호 범위는 첨부된 등록청구범위의 기술적 사상에 의해 정해져야 할 것이다.

부호의 설명

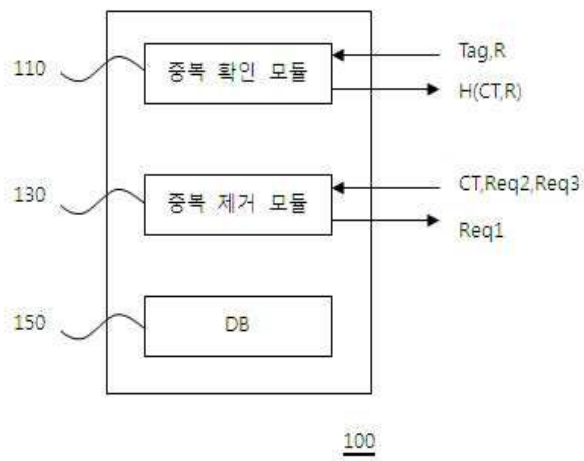
- [0070] 10 : 중복 제거 시스템
- 100 : 중복 제거 서버
- 110 : 중복 확인 모듈
- 130 : 중복 제거 모듈
- 150 : DB
- 200 : 제1 단말기
- 300 : 제2 단말기

도면

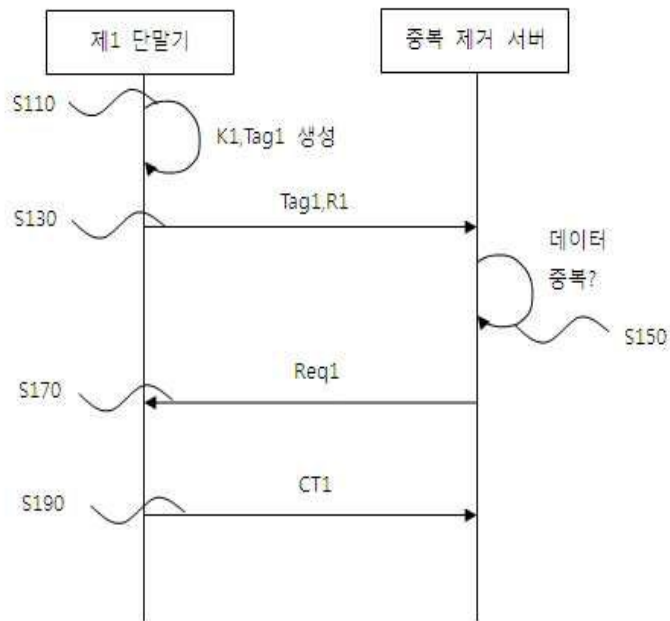
도면1



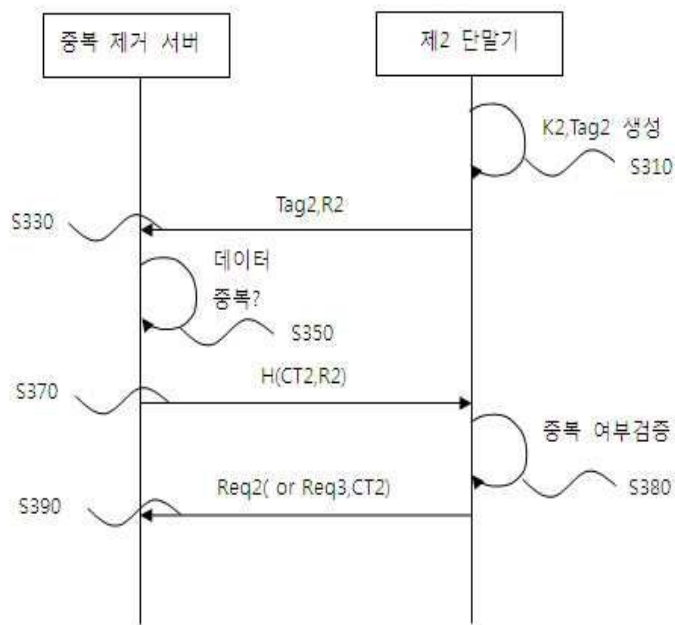
도면2



도면3



도면4



【심사관 직권보정사항】

【직권보정 1】

【보정항목】 청구범위

【보정세부항목】 제6항

【변경전】

중복저장요청을 상시 서버로 송신하는

【변경후】

중복저장요청을 상기 서버로 송신하는