(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2011-0028968
(43) 공개일자 2011년03월22일

(51) Int. Cl.
*G06F 11/07* (2006.01) *G06F 15/16* (2006.01)
*G06F 21/24* (2006.01)
(21) 출원번호 10-2009-0086625
(22) 출원일자 2009년09월14일
심사청구일자 2009년09월14일

(71) 출원인
**고려대학교 산학협력단**
서울 성북구 안암동5가 1
(72) 발명자
**이동훈**
서울특별시 종로구 내수동 72 경희궁의아침3단지 아파트 1404호
**정익래**
서울특별시 노원구 공릉동 81번지 태강아파트 1014동 1303호
*(뒷면에 계속)*
(74) 대리인
**특허법인충현**

전체 청구항 수 : 총 17 항

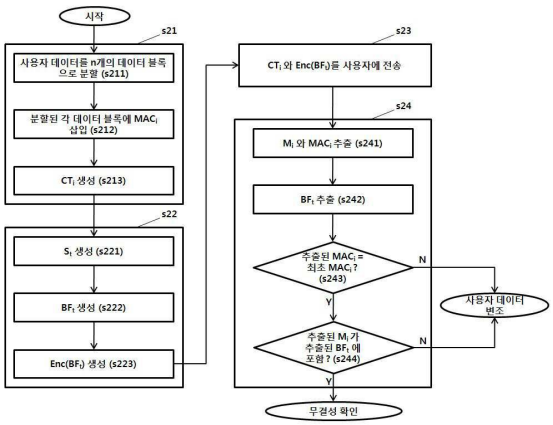
# (54) 원격 컴퓨팅 환경에서 사용자 데이터의 무결성 검증 방법 및 그 시스템

## (57) 요 약

본 발명은 원격 컴퓨팅 환경에서 사용자 데이터의 무결성 검증 방법에 관한 것이다.

본 명세서에서 개시하는 원격 컴퓨팅 환경에서 사용자 데이터의 무결성 검증 방법은 (a)사용자 데이터의 암호문과 상기 사용자 데이터에 대한 블룸 필터(Bloom filter)를 원격 컴퓨팅 서비스를 제공하는 서버의 스토리지(storage)에 저장시키는 단계; (b)사용자로부터 자신의 사용자 데이터에 대한 다운로드 요청이 있는 경우, 상기 요청된 사용자 데이터에 상응하는 암호문의 일부와 블룸 필터의 일부를 상기 스토리지로부터 호출하여 상기 사용자에게 전송하는 단계; 및 (c)상기 전송받은 일부 암호문이 상기 전송받은 일부 블룸 필터에 속하는지 판단하여 상기 다운로드 요청한 사용자 데이터의 무결성(integrity)을 검증하는 단계를 포함하여 본 방법 발명의 과제를 해결한다.

*대 표 도* - 도2



(72) 발명자

**천지영**

서울특별시 성북구 하월곡4동 두산위브 아파트 118동 104호

**노건태**

서울특별시 강남구 개포3동 주공아파트 707동 1304호

이 발명을 지원한 국가연구개발사업

과제고유번호 ITAA1100090100480001000100100

부처명 정보통신연구진흥원

연구관리전문기관

연구사업명 지식경제기술혁신사업

연구과제명 Car-헬스케어 보안 기술개발

기여율

주관기관 고려대학교

연구기간 2009년 03월 01일 ~ 2010년 02월 28일

## *특허청구의 범위*

**청구항 1**

(a)사용자 데이터의 암호문과 상기 사용자 데이터에 대한 블룸 필터(Bloom filter)를 원격 컴퓨팅 서비스를 제공하는 서버의 스토리지(storage)에 저장시키는 단계;

(b)사용자로부터 자신의 사용자 데이터에 대한 다운로드 요청이 있는 경우, 상기 요청된 사용자 데이터에 상응하는 암호문의 일부와 블룸 필터의 일부를 상기 스토리지로부터 호출하여 상기 사용자에게 전송하는 단계; 및

(c)상기 전송받은 일부 암호문이 상기 전송받은 일부 블룸 필터에 속하는지 판단하여 상기 다운로드 요청한 사용자 데이터의 무결성(integrity)을 검증하는 단계를 포함하는 원격 컴퓨팅 환경에서 사용자 데이터의 무결성 검증 방법.

**청구항 2**

제 1 항에 있어서, 상기 사용자 데이터의 암호문 생성은

상기 사용자 데이터를 n(n은 2 이상의 자연수)개의 데이터 블록($M_i$, i=1, 2, …, n)으로 분할하는 단계;

상기 분할된 각 데이터 블록 $M_i$에 메시지 인증 코드(Message Authentication Code: MAC) $MAC_i$을 생성하여 상기 각 데이터 블록 $M_i$에 각각 삽입하는 단계; 및

의사 난수 함수(pseudo random function) f(k, i)의 결과값과 상기 $M_i$와 $MAC_i$가 결합된 데이터를 이용하여 상기 사용자 데이터에 대한 암호문($CT_i$, i=1, 2, …, n)을 생성하는 단계를 포함하여 이루어지는 것을 특징으로 하는 원격 컴퓨팅 환경에서 사용자 데이터의 무결성 검증 방법.

여기서, 상기 k는 사용자의 비밀키이다.

**청구항 3**

제 2 항에 있어서, 상기 블룸 필터의 생성은

상기 $M_1$, $M_2$, …, $M_n$을 복수개(j개)씩 그룹핑하여 각 그룹 블록($S_t$)을 생성하는 단계; 및

상기 $S_t$의 블룸 필터($BF_t$)를 생성하는 단계를 포함하여 이루어지는 것을 특징으로 하는 원격 컴퓨팅 환경에서 사용자 데이터의 무결성 검증 방법.

**청구항 4**

제 3 항에 있어서, 상기 $S_t$의 블룸 필터($BF_t$)의 생성은

상기 $S_t$의 키워드($KW_t$)를 삽입하여 이루어지는 것을 특징으로 하는 사용자 데이터의 무결성 검증 방법.

**청구항 5**

제 4 항에 있어서,

상기 생성된 $BF_t$는 의사 난수 함수 g(k, t)의 결과값과 상기 $BF_t$를 이용하여 암호화한 암호문($Enc(BF_t)$)으로 변형 생성되고,

상기 $Enc(BF_t)$가 상기 스토리지에 저장되는 것을 특징으로 하는 원격 컴퓨팅 환경에서 사용자 데이터의 무결성 검증 방법.

**청구항 6**

제 5 항에 있어서, 상기 (b)단계는

상기 일부 암호문으로 상기 $CT_i$를, 상기 일부 블룸 필터로 상기 $Enc(BF_t)$를 호출하여 상기 사용자에게 전송하는

것으로 구현되는 것을 특징으로 하는 원격 컴퓨팅 환경에서 사용자 데이터의 무결성 검증 방법.

**청구항 7**

제 6 항에 있어서, 상기 (c)단계는

(c1)상기 전송받은 $CT_i$와 f(k, i)로부터 상기 $M_i$와 $MAC_i$를 추출하는 단계;

(c2)상기 전송받은 Enc($BF_t$)와 g(k, t)로부터 상기 $BF_t$를 추출하는 단계;

(c3)상기 추출된 $MAC_i$가 상기 최초 생성된 $MAC_i$와 일치하는지 판단하는 단계; 및

(c4)상기 추출된 $M_i$가 상기 추출된 $BF_t$에 속하는지 판단하는 단계를 포함하는 것을 특징으로 하는 원격 컴퓨팅 환경에서 사용자 데이터의 무결성 검증 방법.

**청구항 8**

제 1 항 내지 제 7 항 중 어느 한 항에 있어서,

상기 원격 컴퓨팅은 클라우드 컴퓨팅(cloud computing)인 것을 특징으로 하는 원격 컴퓨팅 환경에서 사용자 데이터의 무결성 검증 방법.

**청구항 9**

제 1 항 내지 제 8 항 중 어느 한 항의 방법을 컴퓨터에서 실행하기 위한 프로그램을 기록하는 컴퓨터 판독 가능한 기록 매체.

**청구항 10**

사용자 데이터의 암호문을 생성하는 암호문 생성부;

상기 사용자 데이터에 대한 블룸 필터(Bloom filter)를 생성하는 블룸 필터 생성부;

상기 사용자 데이터의 암호문과 상기 생성된 블룸 필터를 저장하는 원격 컴퓨팅 서비스를 제공하는 서버의 스토리지(storage);

사용자로부터 자신의 사용자 데이터에 대한 다운로드 요청이 있는 경우, 상기 요청된 사용자 데이터에 상응하는 암호문의 일부와 블룸 필터의 일부를 상기 스토리지로부터 호출하여 상기 사용자에게 전송하는 전송부; 및

상기 전송받은 일부 암호문이 상기 전송받은 일부 블룸 필터에 속하는지 판단하여 상기 다운로드 요청한 사용자 데이터의 무결성(integrity)을 검증하는 무결성 검증부를 포함하는 원격 컴퓨팅 환경에서 사용자 데이터의 무결성 검증 시스템.

**청구항 11**

제 10 항에 있어서, 상기 암호문 생성부는

상기 사용자 데이터를 n(n은 2 이상의 자연수)개의 데이터 블록($M_i$, i=1, 2, …, n)으로 분할하고, 상기 분할된 각 데이터 블록 $M_i$에 메시지 인증 코드(Message Authentication Code: MAC) $MAC_i$을 생성하여 상기 각 데이터 블록 $M_i$에 각각 삽입하며,

의사 난수 함수(pseudo random function) f(k, i)의 결과값과 상기 $M_i$와 $MAC_i$가 결합된 데이터를 이용하여 상기 사용자 데이터에 대한 암호문($CT_i$, i=1, 2, …, n)을 생성하는 것을 특징으로 하는 원격 컴퓨팅 환경에서 사용자 데이터의 무결성 검증 시스템.

여기서, 상기 k는 사용자의 비밀키이다.

**청구항 12**

제 11 항에 있어서, 상기 블룸 필터 생성부는

상기 $M_1$, $M_2$, …, $M_n$을 복수개(j개)씩 그룹핑하여 각 그룹 블록($S_t$)을 생성하고, 상기 $S_t$에 대한 블룸 필터($BF_t$)를 생성하는 것을 특징으로 하는 원격 컴퓨팅 환경에서 사용자 데이터의 무결성 검증 시스템.

**청구항 13**

제 12 항에 있어서, 상기 블룸 필터 생성부는

상기 $S_t$의 키워드($KW_t$)를 삽입하여 상기 $BF_t$를 생성하는 것을 특징으로 하는 사용자 데이터의 무결성 검증 시스템.

**청구항 14**

제 13 항에 있어서, 상기 블룸 필터 생성부는

상기 생성된 $BF_t$를 의사 난수 함수 g(k, t)의 결과값과 상기 $BF_t$를 이용하여 암호화한 암호문($Enc(BF_t)$)으로 변형 생성하고, 상기 $Enc(BF_t)$를 상기 스토리지에 저장하는 것을 특징으로 하는 원격 컴퓨팅 환경에서 사용자 데이터의 무결성 검증 시스템.

**청구항 15**

제 14 항에 있어서, 상기 전송부는

상기 일부 암호문으로 상기 $CT_i$를, 상기 일부 블룸 필터로 상기 $Enc(BF_t)$를 호출하여 상기 사용자에게 전송하는 것을 특징으로 하는 원격 컴퓨팅 환경에서 사용자 데이터의 무결성 검증 시스템.

**청구항 16**

제 15 항에 있어서, 상기 무결성 검증부는

상기 전송부로부터 전송받은 $CT_i$와 f(k, i)로부터 상기 $M_i$와 $MAC_i$를 추출하고, 상기 전송부로부터 전송받은 $Enc(BF_t)$와 g(k, t)로부터 상기 $BF_t$를 추출하며,

상기 추출한 $MAC_i$가 상기 최초 생성된 $MAC_i$와 일치하는지와 상기 추출한 $M_i$가 상기 추출한 $BF_t$에 속하는지 판단하여 상기 사용자 데이터의 무결성을 검증하는 것을 특징으로 하는 원격 컴퓨팅 환경에서 사용자 데이터의 무결성 검증 시스템.

**청구항 17**

제 10 항 내지 제 16 항 중 어느 한 항에 있어서,

상기 원격 컴퓨팅은 클라우드 컴퓨팅(cloud computing)인 것을 특징으로 하는 원격 컴퓨팅 환경에서 사용자 데이터의 무결성 검증 시스템.

# *명 세 서*

## *발명의 상세한 설명*

### *기 술 분 야*

[0001] 본 발명은 원격 컴퓨팅 환경에서 사용자 데이터의 무결성 검증 방법 및 그 시스템에 관한 것이다.

### *배 경 기 술*

[0002] 유비쿼터스(ubiquitous) 환경이 일반화되어 감에 따라 원격 컴퓨팅(remote computing)도 이에 따라 확장 일로의 추세에 있다. 원격 컴퓨팅 중 특히 클라우드 컴퓨팅(Cloud Computing)이 미래의 컴퓨팅 환경을 주도할 새로운 패러다임이 될 것으로 예상되는데, 특히 개인별 또는 기업 등의 집단별 IT 자원들을 저비용으로 효율적으로 관리하기 위한 필요성의 증가에 힘입어 클라우드 컴퓨팅은 더욱 활성화될 것으로 예상된다.

[0003] 클라우드 컴퓨팅은 서로 다른 물리적 공간에 있는 IT 자원들(예를 들어 소프트웨어, 스토리지, 서버, 네트워크 등)을 가상화(virtualization)시켜 통합 서비스로 제공하는 컴퓨팅으로서, 인터넷을 통해 사용자에게 IT 자원을 제공하는 컴퓨팅으로 정의할 수 있으며, 이는 사용자가 인터넷을 통해 IT 자원의 이용 요청을 하여 자신이 필요한 만큼 빌려서 사용하고 사용한 만큼 비용을 지불하는 컴퓨팅 방식이다. 즉, 클라우드 컴퓨팅은 사용자 주문형(On-Demand) IT 자원 제공 서비스라고 볼 수 있다.

[0004] 클라우드 컴퓨팅이 미래의 컴퓨팅 환경에서의 지배적인 위치를 확보하기 위해 선결되어야 할 사항은 바로 사용자 데이터에 대한 무결성(integrity)이 보장되어야 하는 것이다. 클라우드 컴퓨팅은 사용자가 자신의 단말에 IT 자원을 소유함이 없이 일부 또는 모두를 아웃소싱(outsourcing)하는 형태이며, 모든 데이터가 클라우드 컴퓨팅 서비스를 제공하는 서버에 집중되어 있기 때문에 필연적으로 보안 문제가 제기될 수밖에 없다. 더불어 최근 해킹이나 기타 원인으로 인한 개인 정보의 유출이 심각한 사회 문제화 되고 있는 것이 현재의 상황인바, 이러한 상황에서 차세대 컴퓨팅 환경을 지배할 것으로 예상될 클라우드 컴퓨팅에서의 사용자 데이터에 대한 무결성의 보장은 그 무엇보다도 중요하다.

## 발명의 내용

### 해결 하고자하는 과제

[0005] 본 발명은 상기와 같은 배경을 인식하여 창안된 것으로, 본 발명이 해결하고자 하는 과제는 원격 컴퓨팅 환경 특히, 클라우드 컴퓨팅 환경에서 사용자 데이터에 대한 무결성을 확보할 수 있는 방안을 제공하는 것이다.

### 과제 해결수단

[0006] 상기와 같은 과제를 해결하기 위해 본 명세서에서 개시하는 원격 컴퓨팅 환경에서 사용자 데이터의 무결성 검증 방법은

[0007] (a)사용자 데이터의 암호문과 상기 사용자 데이터에 대한 블룸 필터(Bloom filter)를 원격 컴퓨팅 서비스를 제공하는 서버의 스토리지(storage)에 저장시키는 단계; (b)사용자로부터 자신의 사용자 데이터에 대한 다운로드 요청이 있는 경우, 상기 요청된 사용자 데이터에 상응하는 암호문의 일부와 블룸 필터의 일부를 상기 스토리지로부터 호출하여 상기 사용자에게 전송하는 단계; 및 (c)상기 전송받은 일부 암호문이 상기 전송받은 일부 블룸 필터에 속하는지 판단하여 상기 다운로드 요청한 사용자 데이터의 무결성(integrity)을 검증하는 단계를 포함하여 본 방법 발명의 과제를 해결한다.

[0008] 상기와 같은 과제를 해결하기 위해 본 명세서에서 개시하는 원격 컴퓨팅 환경에서 사용자 데이터의 무결성 검증 시스템은

[0009] 사용자 데이터의 암호문을 생성하는 암호문 생성부; 상기 사용자 데이터에 대한 블룸 필터를 생성하는 블룸 필터 생성부; 상기 사용자 데이터의 암호문과 상기 생성된 블룸 필터를 저장하는 원격 컴퓨팅 서비스를 제공하는 서버의 스토리지(storage); 사용자로부터 자신의 사용자 데이터에 대한 다운로드 요청이 있는 경우, 상기 요청된 사용자 데이터에 상응하는 암호문의 일부와 블룸 필터의 일부를 상기 스토리지로부터 호출하여 상기 사용자에게 전송하는 전송부; 및 상기 전송받은 일부 암호문이 상기 전송받은 일부 블룸 필터에 속하는지 판단하여 상기 다운로드 요청한 사용자 데이터의 무결성(integrity)을 검증하는 무결성 검증부를 포함하여 본 시스템 발명의 과제를 해결한다.

### 효 과

[0010] 본 발명에 의한 무결성 검증 방안에 의하면 POR 기법과는 달리 사용자 데이터의 무결성 검증에 특히 검증 횟수에 있어 아무런 제한이 없다. 아울러 사용자(원본) 데이터의 암호화을 위한 데이터의 최소한의 삽입만으로 무결성 검증이 가능하므로 서버의 데이터 저장 부하를 획기적으로 감소시킬 수 있다.

### 발명의 실시를 위한 구체적인 내용

[0011] 본 발명에 관한 구체적인 내용의 설명에 앞서 이해의 편의를 위해 본 발명이 해결하고자 하는 과제의 해결 방안의 개요를 우선 제시한다.

[0012] 클라우드 컴퓨팅 환경에서 사용자 데이터는 사용자가 보유한 단말(로컬 스토리지(local storage))에 저장되는 것이 아니라 클라우드 스토리지 즉, 클라우드 컴퓨팅 서비스를 제공하는 서버에 저장된다. 이러한 이유로 클라

우드 컴퓨팅 환경에서 사용자의 데이터는 제3자에 의한 불법적인 변조 가능성이 상존하고 매우 높을 수밖에 없다.

[0013] 제3자에 의한 불법적인 변조 가능성을 차단하기 위한 각종 방안들이 제안되거나 실시 중인데, 현재까지의 방안들은, 사용자의 원본 데이터를 안전하게 보호하기 위해 암호화하였을 경우, 사용자가 원하는 데이터를 효율적으로 검색하는 방안에 대해서만 초점이 맞춰져왔다.

[0014] 그러나 사용자가 검색을 통해 원하는 데이터를 다운로드하기 이전에, 자신이 저장한 데이터가 변조되지 않았음을 검증하는 즉, 사용자 데이터에 대한 무결성 검증이 필요하다. 무결성 검증은 서버의 관리 소홀 또는 제3자의 의도된(불법적) 변조 등으로 사용자의 원본 데이터가 변조되었을 경우 이를 정기적으로 감사(audit)하기 위해 필요하다. 아울러 무결성 검증은 특히 사용자가 원하는 데이터가 대량인 경우 아웃소싱 이전에 변조 유무를 미리 검사하여 변조되었을 경우 대량의 데이터를 아웃소싱하는 수고를 덜어줄 수 있다.

[0015] 이러한 필요성을 인식하여 POR(Proof Of Retrievability) 기법이라는 데이터 무결성 검증 기법이 제안된 바가 있다. 이 기법은 사용자의 원본 데이터를 암호화한 암호문 사이에 랜덤한 값인 센티널(Sentinel)을 생성하여 서버에 저장하고, 원본 데이터의 변조 여부를 검증시에 서버에게 센티널의 위치를 알려준 뒤 서버로부터 센티널을 받아와서 사용자 자신이 가지고 있는 센티널과 같은지를 비교하여 데이터 무결성을 검증하는 기법이다.

[0016] 이 기법은 사용자가 원하는 데이터를 아웃소싱 받지 않고도 데이터 변조 유무를 알 수 있으나, 상기 암호문을 구성하기 위해 원본 데이터에 추가적인 데이터(센티널, 에러 정정 코드(ECC) 등)의 삽입으로 인해 데이터 용량이 지나치게 증가하며 아울러 검증의 횟수가 센티널의 개수에 종속되는 결점을 갖는다. 아울러 센티널이 삽입되지 아니한 부분에 대해서는 본 기법에 의한 검증 자체가 이루어질 수 없기 때문에 제3자에 의한 불법적인 변조가 아주 쉽게 이루어질 수 있다.

[0017] 본 발명은 POR 기법의 이러한 문제를 해결하기 위해 창안된 것으로, 사용자의 데이터를 암호화하기 위한 추가적인 데이터의 최소한의 삽입만으로도 무결성 검증이 가능하며 데이터 무결성 검증에 제한 조건을 두지 않는 그리고 사용자 데이터의 모든 부분에 대해 무결성 검증이 가능한 무결성 검증 방안을 구현한 것이다.

[0018] 이하, 본 발명이 해결하고자 하는 과제의 해결 방안을 명확하게 하기 위한 발명의 구성을 본 발명의 바람직한 실시예에 근거하여 첨부 도면을 참조하여 상세히 설명하되, 도면의 구성요소들에 참조번호를 부여함에 있어서 동일 구성요소에 대해서는 비록 다른 도면상에 있더라도 동일 참조번호를 부여하였으며 당해 도면에 대한 설명시 필요한 경우 다른 도면의 구성요소를 인용할 수 있음을 미리 밝혀둔다. 아울러 본 발명과 관련된 공지 기능 혹은 구성에 대한 구체적인 설명 그리고 그 이외의 제반 사항이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우, 그 상세한 설명을 생략한다.

[0019] 도 1a는 본 시스템 발명의 바람직한 일 구성을, 도 2는 본 방법 발명의 바람직한 일 흐름을 제시한 도면이다.

[0020] 암호문 생성부(11)는 사용자 데이터(원본 데이터)의 암호문을 생성하는(s21) 기능을 수행하는 부분으로, 암호문 생성의 일례는 다음과 같다.

[0021] 암호문 생성부(11)는 우선 사용자 데이터를 n(n은 2 이상의 자연수)개의 데이터 블록으로 분할한다(s211). 분할의 가장 큰 이유는 특히 사용자 데이터가 대량인 경우를 대비하기 위함인데, 대량의 원본 데이터는 아웃소싱 측면뿐만이 아니라 추후에 언급할 무결성 검증 그 자체만으로도 상당한 부하가 소요되기 때문이다.

[0022] 따라서 본 발명은 분할된 각 데이터 블록($M_i$, i=1, 2, …, n) 중 임의의 데이터 블록에 대해서만 검증을 수행하여 검증에 소요되는 부하를 최소화시킬 수 있다. 다만 임의의 데이터 블록 몇 개에 대해 검증을 수행할지는 일률적인 기준이 없으며, 원격 컴퓨팅이 구축된 제반 환경에 의해 결정된다. 아울러 분할은 임의의 데이터 블록에 대해 검증을 수행하게 하는 발판을 제공하여 사용자 데이터의 모든 부분에 대한 검증을 가능하게 한다.

[0023] 암호문 생성부(11)는 다음으로 분할된 각 데이터 블록($M_i$, i=1, 2, …, n)에 메시지 인증 코드(Message Authentication Code: MAC) $MAC_i$을 생성하여 각 $M_i$에 삽입한다(s212). MAC는 적은 비트수의 데이터로서, 사용자 데이터(원본 데이터)의 변조 여부를 확인하기 위해 사용되는 코드이다. MAC에 관한 일반적인 사항은 기 공지된 문헌 등에 상세히 기술되어 있으므로, 본 명세서에서는 이에 대한 자세한 설명은 생략한다.

[0024] 암호문 생성부(11)는 다음으로 의사 난수 함수(pseudo random function) f(k, i)의 결과값과 상기 $M_i$와 $MAC_i$가 결합된 데이터를 이용하여 상기 $M_i$에 대한 암호문($CT_i$)을 생성한다(s213). $CT_i$는 다음의 식에 의해 생성되는 것

을 그 예로 들 수 있다.

[0025] $CT_i = (M_i \,||\, MAC_i) \oplus f(k, i)$.

[0026] 여기서 k는 사용자의 비밀키, ||는 이어붙이기(concatenation), $\oplus$는 배타적 논리합(exclusive OR)을 나타낸다. 암호문 생성부(11)는 생성된 암호문($CT_1$, $CT_2$, …, $CT_i$)을 원격 컴퓨팅 서비스를 제공하는 서버의 스토리지(storage)에 저장시킨다(s22).

[0027] 블룸 필터 생성부(12)는 사용자 데이터($M_1$, $M_2$, …, $M_n$)에 대한 블룸 필터를 생성하여 원격 컴퓨팅 서비스를 제공하는 서버의 스토리지(storage)에 저장시킨다(s22).

[0028] 블룸 필터는 어떤 원소(element)가 어떤 집합(set)의 구성원(member)에 속하는지 테스트하는데 이용되는 공간 효율적인 확률적 데이터 구조로서, 많은 양의 데이터를 줄여서 공간 효율적으로 데이터를 빠르게 검색을 할 수 있는 장점을 제공한다.

[0029] 블룸 필터 생성부(12)는 우선 $M_i$을 복수(j)개씩 그룹핑(grouping)하여 각 그룹 블록($S_t$)을 생성한다(s221). 그룹핑의 이유는 하나의 메시지 $M_i$마다 블룸 필터를 생성할 경우 저장 공간이 많이 소요될 우려가 있기 때문에 이를 미연에 방지하기 위함이다. 그룹핑의 예를 들면 j=3인 경우 $S_1=[M_1, M_2, M_3]$, $S_2=[M_4, M_5, M_6]$, …, $S_t=[M_{3t-2}, M_{3t-1}, M_{3t}]$, …, $S_{n/3}=[M_{n-2}, M_{n-1}, M_n]$이다. 한편 $S_t$로부터 키워드($KW_t$)를 임의로 결정할 수 있다. 키워드($KW_t$)는 사용자 데이터에 포함된 가시적인 임의의 데이터로서, 후술할 무결성 검증 과정에서 키워드($KW_t$)가 추출되는지의 여부로 사용자가 서버에 다운로드 요청한 자신의 데이터가 스토리지에 안전하게 존재하는지의 여부를 판단할 수 있도록 하는데 이용된다.

[0030] 블룸 필터 생성부(12)는 다음으로 $S_t$와 $KW_t$를 이용하여 블룸 필터($BF_t$)를 생성하게 된다(s222). 즉, 블룸 필터는 $S_t$ 단위로 이루어지게 된다.

[0031] 도 1b는 블룸 필터를 생성하는 일반적인 방안을 설명하기 위해 제시한 도면이다.

[0032] 우선 블룸 필터를 저장할 m-비트(m-bit) 저장 공간을 설정한다(도 1b에서 A 부분). 공간 설정시 모든 비트는 0으로 초기화한다. 다음으로 각 데이터 블록($M_i$, i=1, 2, …, n)에 대해 0에서 m-1까지의 값을 생성하는 k개의 해시 함수(hash function) $h_1$, $h_2$, …, $h_k$를 적용한 값으로부터 그 값에 해당하는 m-비트 저장 공간의 해당 비트의 값을 1로 설정한다. 즉 $M_i$마다 각각 k개의 해시 함수($h_1$, $h_2$, …, $h_k$)를 사용해서 m 비트 저장 공간에 값을 저장하여(값이 겹칠 수도 있으며, 이 경우 역시 1로 설정함) 블룸 필터를 생성한다.

[0033] 본 발명에서는 각 데이터 블록($M_i$, i=1, 2, …, n) 단위로 블룸 필터를 생성하는 것이 아니다. 상기 생성된 $S_t$와 $KW_t$를 하나의 메시지 $SYN_t=[S_t, KW_t]$로 구현하고, $SYN_t$에 대해 해시 함수를 적용하여 $BF_t$를 생성한다.

[0034] 한편 블룸 필터 생성부(12)는 $BF_t$ 그 자체를 서버의 스토리지에 저장시키는 것이 아니라, 암호화시켜 저장시키는데(s223) $BF_t$에 대한 암호화는 다음의 식에 의해 이루어질 수 있음을 그 예로 들 수 있다.

[0035] $Enc(BF_t) = BF_t \oplus g(k, t)$.

[0036] 여기서 $Enc(BF_t)$가 $BF_t$의 암호문이며, g(k, t)는 상기 f(k, i)와 마찬가지로 의사 난수 함수이다. $Enc(BF_1)$, $Enc(BF_2)$, …, $Enc(BF_t)$, …, $Enc(BF_{n/j})$가 스토리지에 저장된다.

[0037] 전송부(13)는 사용자로부터 스토리지에 저장된 사용자 데이터($M_i$, i=1, 2, …, n)의 다운로드 요청이 있는 경우, 이에 상응하는 암호문($CT_1$, $CT_2$, …, $CT_i$, …, $CT_n$) 중 하나인 $CT_i$와 $Enc(BF_t)$를 스토리지로부터 호출하여 사용자에게 전송한다(s23). 즉, 사용자에게로의 전송은 아웃소싱이 요청된 사용자 데이터의 암호문 모두에 대해 행하는 것이 아니라 요청된 데이터의 암호문 중 일부만이 전송된다. 이는 위에서 언급한 바와 같이 무결성 검증을 위한 부하를 최소화시키기 위해서이다.

[0038] 무결성 검증부(14)는 전송부(13)로부터 전송받은 $CT_i$와 f(k, i)로부터 $M_i$와 $MAC_i$를 추출하고(s241), 전송부(13)로부터 전송받은 $Enc(BF_t)$와 g(k, t)로부터 $BF_t$를 추출하고(s242), 추출한 $MAC_i$가 암호문 생성부(11)에 의해 최초 생성된 $MAC_i$와 일치하는지 판단하며(s243), 추출한 $M_i$가 상기 추출된 $BF_t$에 속하는지 판단하여(s244) 상기 아웃소싱 요청한 사용자 데이터의 무결성을 검증하게 된다(s24). s241 또는 s242에서 각각 $M_i$와 $MAC_i$, $BF_t$가 추출되지 아니하거나 s243 내지 s244 어느 한 단계라도 만족하지 아니하면 사용자가 다운로드 요청한 사용자 데이터의 무결성은 훼손된 것이다.

[0039] 한편 상기 추출한 $BF_t$에는 $KW_t$가 포함되어 있기 때문에, $BF_t$가 추출되면 $KW_t$가 자동으로 추출된다. 따라서 사용자는 추출한 $M_i$가 상기 추출된 $BF_t$에 속하는지 판단하는 과정(s24)에서 $KW_t$를 확인할 수 있으므로 자신이 다운로드 요청한 사용자 데이터가 스토리지에 안전하게 존재하는지의 여부를 쉽게 손쉽게 확인할 수 있다.

[0040] 무결성 검증 과정을 식을 일례로 들어 요약해보면 다음과 같다.

[0041] $CT_i \oplus f(k, I) = M_i \,||\, MAC_i$.

[0042] $Enc(BF_t) \oplus g(k, t) = BF_t$.

[0043] $MAC_i$ = 암호문 생성부에 의해 생성된 $MAC_i$?

[0044] 추출된 $M_i \in$ 추출된 $BF_t$?

[0045] 이처럼 본 발명은 모든 데이터 블록 $M_1, M_2, \cdots, M_n$에 대해 암호문($CT_1, CT_2, \cdots, CT_i, \cdots, CT_n$)을 생성하고 블룸 필터를 적용하기 때문에, 사용자 데이터의 모든 부분에 대한 무결성 검증을 임의의 데이터 블록을 선택하여 수행할 수 있다. 즉, POR 기법에 의하면 센티널이 삽입된 데이터 블록에 대해서만 무결성 검증이 가능하고 센티널의 삽입 위치를 서버에 알려주어야 하며 아울러 무결성 검증은 센티널의 삽입 횟수에 제한되므로 무결성 검증이 지속가능하지 않으나, 본 발명은 모든 데이터 블록 $M_1, M_2, \cdots, M_n$에 대해 암호문($CT_1, CT_2, \cdots, CT_i, \cdots, CT_n$)을 생성하고 블룸 필터를 적용하기 때문에, 사용자 데이터의 모든 부분에 대한 무결성 검증이 언제든지 지속적으로 가능하다.

[0046] 본 방법발명은 또한 컴퓨터로 읽을 수 있는 기록매체에 컴퓨터가 읽을 수 있는 코드로서 구현하는 것이 가능하다. 컴퓨터가 읽을 수 있는 기록매체는 컴퓨터 시스템에 의하여 읽혀질 수 있는 데이터가 저장되는 모든 종류의 기록 장치를 포함한다.

[0047] 컴퓨터가 읽을 수 있는 기록매체의 예로는 ROM, RAM, CD-ROM, 자기 테이프, 플로피디스크, 광 데이터 저장장치 등이 있으며, 또한 캐리어 웨이브(예를 들어 유무선 네트워크를 통한 전송)의 형태로 구현되는 것도 포함한다. 또한 컴퓨터가 읽을 수 있는 기록매체는 네트워크로 연결된 컴퓨터 시스템에 분산되어 분산방식으로 컴퓨터가 읽을 수 있는 코드가 저장되고 실행될 수 있다.

[0048] 이제까지 본 발명에 대하여 그 바람직한 실시예를 중심으로 살펴보았다. 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자는 본 발명이 본 발명의 본질적인 특성에서 벗어나지 않는 범위에서 변형된 형태로 구현될 수 있음을 이해할 수 있을 것이다.

[0049] 그러므로 개시된 실시예들은 한정적인 관점이 아니라 설명적인 관점에서 고려되어야 한다. 본 발명의 범위는 전술한 설명이 아니라 특허청구범위에 나타나 있으며, 그와 균등한 범위 내에 있는 모든 차이점은 본 발명에 포함된 것으로 해석되어야 할 것이다.

## *도면의 간단한 설명*

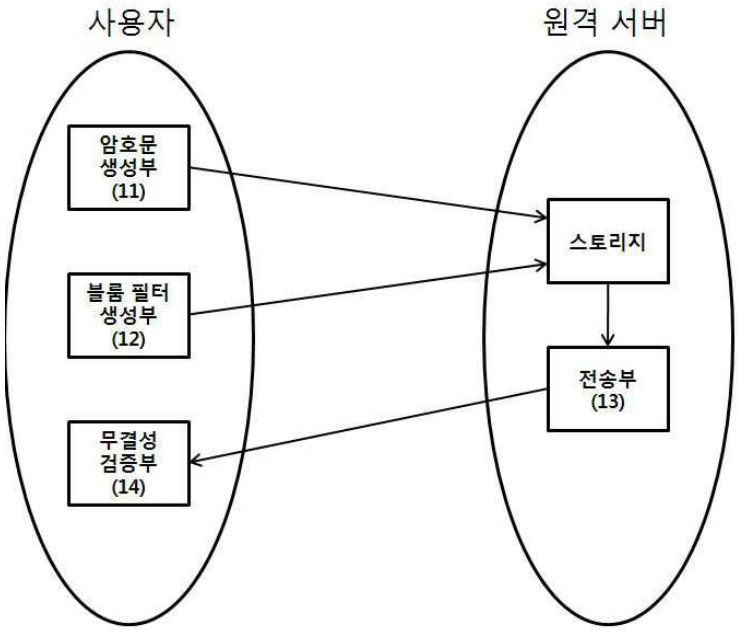
[0050] 도 1a는 본 시스템 발명의 바람직한 일 구성을 제시한 도면이다.

[0051] 도 1b는 블룸 필터를 생성하는 일반적인 방안을 설명하기 위해 제시한 도면이다.

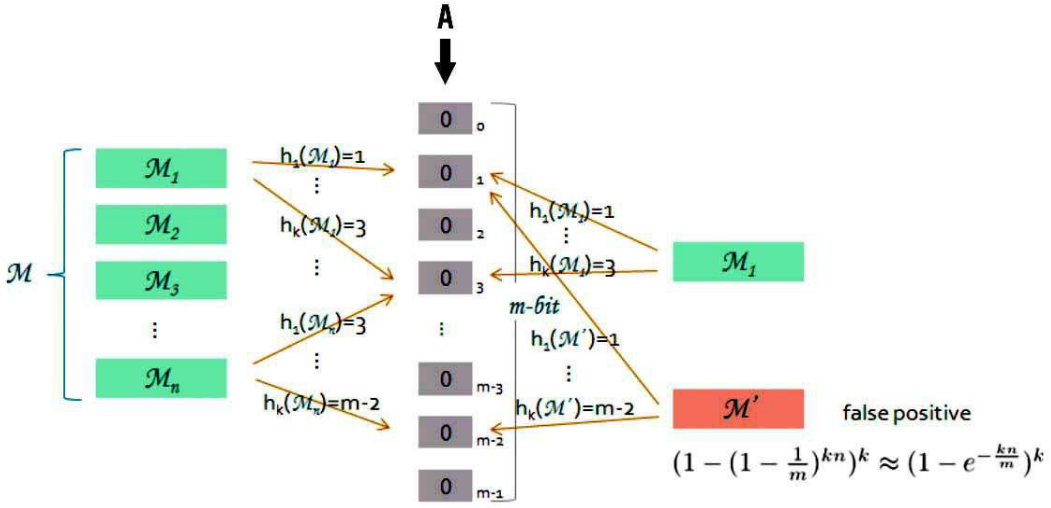
[0052] 도 2는 본 방법 발명의 바람직한 일 흐름을 제시한 도면이다.

## 도면

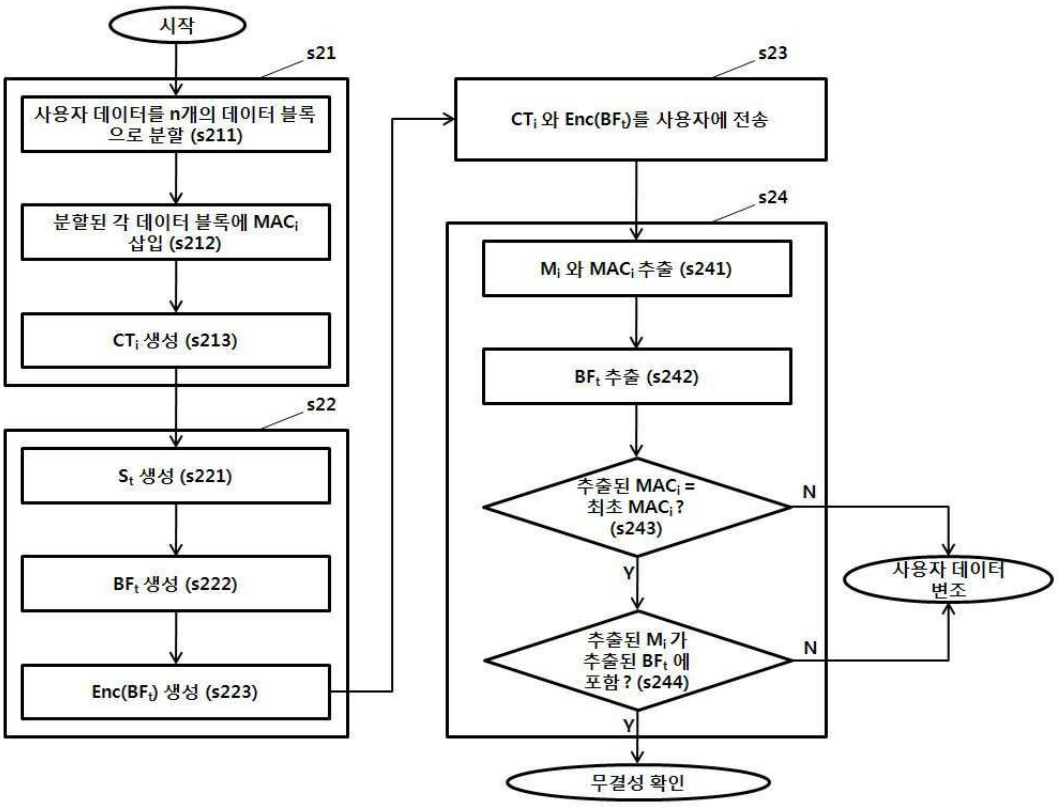
### 도면1a


사용자
원격 서버
암호문 생성부 (11)
블룸 필터 생성부 (12)
무결성 검증부 (14)
스토리지
전송부 (13)

### 도면1b


A
M
$M_1$
$M_2$
$M_3$
$M_n$
$h_1(M_1)=1$
$h_k(M_2)=3$
$h_1(M_3)=3$
$h_k(M_n)=m-2$
0
m-bit
$h_1(M_2)=1$
$h_k(M_1)=3$
$h_1(M')=1$
$h_k(M')=m-2$
$M'$
false positive
$(1-(1-\frac{1}{m})^{kn})^k \approx (1-e^{-\frac{kn}{m}})^k$

도면2


시작
s21
사용자 데이터를 n개의 데이터 블록으로 분할 (s211)
분할된 각 데이터 블록에 $MAC_i$ 삽입 (s212)
$CT_i$ 생성 (s213)
s22
$S_t$ 생성 (s221)
$BF_t$ 생성 (s222)
$Enc(BF_t)$ 생성 (s223)
s23
$CT_i$ 와 $Enc(BF_t)$를 사용자에 전송
s24
$M_i$ 와 $MAC_i$ 추출 (s241)
$BF_t$ 추출 (s242)
추출된 $MAC_i$ = 최초 $MAC_i$ ? (s243)
N
Y
추출된 $M_i$ 가 추출된 $BF_t$ 에 포함 ? (s244)
N
Y
사용자 데이터 변조
무결성 확인