



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2016년06월10일
(11) 등록번호 10-1628924
(24) 등록일자 2016년06월02일

(51) 국제특허분류(Int. Cl.)
H04L 9/30 (2006.01)
(21) 출원번호 10-2014-0174698
(22) 출원일자 2014년12월08일
심사청구일자 2014년12월08일
(56) 선행기술조사문헌
KR1020100069447 A
KR101174058 B1

(73) 특허권자
고려대학교 산학협력단
서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)
(72) 발명자
정익래
서울 광진구 긴고랑로4길 53, 203호 (중곡동, 호동아파트)
이상훈
서울 도봉구 시루봉로11길 25-8, 101동 1403호 (방학동, 극동아파트)
(74) 대리인
김등용, 김홍석

전체 청구항 수 : 총 10 항

심사관 : 전용해

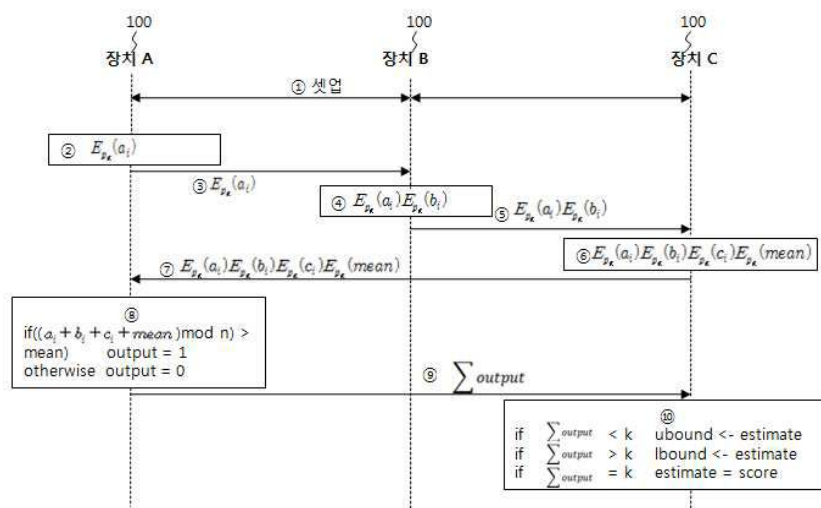
(54) 발명의 명칭 k번째 항목의 스코어 산출 방법 및 산출 시스템

(57) 요약

본 발명은 (a) 제1 장치의 공개키로 복수의 항목들 중 제1 항목의 암호화된 제1 속성 데이터를 수신하는 단계, (b) 제1 장치의 공개키로 제1 항목에 대응하는 제2 속성 데이터 및 추정값에 기초한 데이터를 각각 암호화하는 단계 및 (c) 암호화된 제1 속성 데이터, 암호화된 제2 속성 데이터 및 추정값에 기초한 데이터를 제1 장치로 전송하는 단계를 포함하며, 제1 속성 및 제2 속성은 제1 항목에서의 서로 다른 속성을 나타내는, k번째 항목의 스코어 산출 방법 및 산출 시스템에 관한 것이다.

본 발명을 이용함으로써 공모 공격에 안전하고 다자간에도 기존의 알려진 기법과 동일한 알고리즘 수행 횟수를 가지고서 k번째 항목을 찾을 수 있도록 하는 효과가 있다.

대표도 - 도2



(72) 발명자

홍준희

서울 중구 다산로 32, 27동 201호 (신당동, 남산타운아파트)

노건태

서울 강남구 개포로 516, 707동 1304호 (개포동, 개포주공아파트)

이 발명을 지원한 국가연구개발사업

과제고유번호 NIPA-2014-H0301-14-1004

부처명 미래창조과학부

연구관리전문기관 정보통신산업진흥원

연구사업명 ITRC

연구과제명 제어시스템 보안 연구

기 여 율 1/1

주관기관 고려대학교 산학협력단

연구기간 2014.01.01 ~ 2014.12.31

명세서

청구범위

청구항 1

(a) 암호화 장치는 공개키 제공장치의 공개키로 대용량 저장장치에 저장된 복수의 항목들 중 제1 항목의 속성값을 암호화한, 암호화된 제1 속성 데이터를 수신하는 단계;

(b) 공개키 제공장치의 공개키로 상기 제1 속성 데이터에 대응하는 속성값을 암호화한 제2 속성 데이터 및 스코어 추정값에 기초한 데이터를 각각 암호화하는 단계; 및

(c) 상기 암호화된 제1 속성 데이터, 암호화된 상기 제2 속성 데이터 및 스코어 추정값에 기초한 데이터를 상기 공개키 제공장치로 전송하는 단계를 포함하며,

상기 제1 항목의 속성값 및 상기 제1 속성 데이터에 대응하는 속성값은 상기 제1 항목에서의 서로 다른 속성을 나타내는 k번째 항목의 스코어 산출 방법.

청구항 2

제1항에 있어서,

상기 단계 (a) 이전에, 스코어 상한값, 스코어 하한값 및 스코어 추정값을 설정하는 단계; 및

스코어 추정값을 상기 공개키 제공장치로 전송하는 단계를 더 포함하며,

상기 스코어 추정값은 상기 스코어 상한값 및 스코어 하한값의 중간값에서 난수값을 적용하여 설정되는 k번째 항목의 스코어 산출 방법.

청구항 3

제1항에 있어서,

상기 단계 (a) 내지 (c)는 복수의 항목들 각각에 대해 스코어 산출 장치에 의해서 수행되고,

상기 방법은 상기 공개키 제공장치가, 상기 단계 (c)를 통해 수신된 데이터를 상기 공개키 제공장치의 개인키로 복호화하는 단계;

복호화된 데이터를 스코어 추정값과 비교하는 단계; 및

비교 결과를 상기 스코어 산출 장치로 전송하는 단계를 더 포함하는 k번째 항목의 스코어 산출 방법.

청구항 4

제3항에 있어서,

상기 스코어 산출 장치가, 상기 비교 결과의 수신에 따라 스코어 상한값, 스코어 하한값 및 스코어 추정값을 변경하는 단계; 및 변경된 스코어 추정값을 상기 공개키 제공장치를 포함하는 복수의 장치들로 전송하는 단계를 더 포함하며,

상기 변경된 스코어 추정값은 변경된 스코어 상한값 및 스코어 하한값의 중간값에서 난수값을 적용하여 설정되고, 후속하는 비교 결과 산출 라운드에서 이용될 상기 난수값은 먼저 수행되는 상기 단계 (a) 내지 (c)를 위해 설정된 난수 범위보다 좁은 난수 범위에서 선택되는 k번째 항목의 스코어 산출 방법.

청구항 5

제1항에 있어서,

상기 단계 (c) 이후에,

상기 공개키 제공장치를 포함하는 복수의 장치들 중 다른 하나인 공개키 제공장치의 공개키로 대용량 저장장치

에 저장된 복수의 항목들 중 제1 항목의 속성값을 암호화한, 암호화된 제1 속성 데이터를 수신하는 단계;

상기 다른 하나의 공개키 제공장치의 공개키로 상기 제1 속성 데이터에 대응하는 속성값을 암호화한 제2 속성 데이터 및 스코어 추정값에 기초한 데이터를 각각 암호화하는 단계; 및

상기 암호화된 제1 속성 데이터, 암호화된 상기 제2 속성 데이터 및 스코어 추정값에 기초한 데이터를 상기 다른 하나의 공개키 제공장치로 전송하는 단계;를 더 포함하는 k번째 항목의 스코어 산출 방법.

청구항 6

제1항에 있어서,

상기 단계 (a) 및 (b)에서의 암호화는 준동형(homomorphic) 성질을 갖는 Paillier의 공개키 암호 기법을 이용하여 이루어지는,

k번째 항목의 스코어 산출 방법.

청구항 7

공개키 제공장치의 공개키로 대용량 저장장치에 저장된 복수의 항목들 중 제1 항목의 속성값을 암호화한, 암호화된 제1 속성 데이터를 수신하고, 공개키 제공장치의 공개키로 상기 제1 항목에 대응하는 속성값을 암호화한, 제2 속성 데이터 및 스코어 추정값에 기초한 데이터를 각각 암호화하며, 상기 암호화된 제1 속성 데이터, 암호화된 상기 제2 속성 데이터 및 스코어 추정값에 기초한 데이터를 상기 공개키 제공장치로 전송하는 스코어 산출 장치를 포함하며,

상기 제1 항목의 속성값 및 상기 제1 속성 데이터에 대응하는 속성값은 상기 제1 항목에서의 서로 다른 속성을 나타내는 k번째 항목의 스코어 산출 시스템.

청구항 8

제7항에 있어서,

상기 스코어 산출 장치는 상기 제1 속성 데이터의 수신 이전에, 스코어 상한값, 스코어 하한값 및 스코어 추정값을 설정하고 스코어 추정값을 상기 공개키 제공장치로 전송하며,

상기 스코어 추정값은 상기 스코어 상한값 및 스코어 하한값의 중간값에서 난수값을 적용하여 설정되는 k번째 항목의 스코어 산출 시스템.

청구항 9

제7항에 있어서,

상기 스코어 산출 장치에 연결된 공개키 제공장치를 더 포함하며,

상기 공개키 제공장치는 상기 스코어 산출 장치로부터 수신된 데이터를 상기 공개키 제공장치의 개인키로 복호화하고 복호화된 데이터를 추정값과 비교하고 비교 결과를 상기 스코어 산출 장치로 전송하는 k번째 항목의 스코어 산출 시스템.

청구항 10

제7항에 있어서,

상기 스코어 산출 장치에 연결된 공개키 제공장치 및 다른 하나의 공개키 제공장치를 더 포함하며,

상기 스코어 산출 장치는, 상기 암호화된 제1 속성 데이터, 암호화된 상기 제2 속성 데이터 및 스코어 추정값에 기초한 데이터를 상기 공개키 제공장치로의 전송 이후에, 상기 다른 하나의 공개키 제공장치의 공개키로 대용량 저장장치에 저장된 복수의 항목들 중 제1 항목의 속성값을 암호화한, 암호화된 제1 속성 데이터를 수신하고, 다른 하나의 공개키 제공장치의 공개키로 상기 제1 속성 데이터에 대응하는 속성값을 암호화한, 암호화된 제2 속성 데이터 및 스코어 추정값에 기초한 데이터를 각각 암호화하고, 상기 암호화된 제1 속성 데이터, 암호화된 상기 제2 속성 데이터 및 스코어 추정값에 기초한 데이터를 상기 다른 하나의 공개키 제공장치로 전송하는 k번째 항목의 스코어 산출 시스템.

발명의 설명

기술 분야

- [0001] 본 발명은 k번째 항목의 스코어 산출 방법 및 산출 시스템에 관한 것으로서, 구체적으로는 다자간 환경하에서 서로의 데이터 프라이버시를 보호하면서 상위 k번째 항목을 찾을 수 있도록 하는, k번째 항목의 스코어 산출 방법 및 산출 시스템에 관한 것이다.

배경 기술

- [0002] 데이터 마이닝(data mining)은 대용량 데이터로부터 의미가 있는 정보를 추출하는 기술을 의미한다. 추출한 정보는 데이터 간의 연관성을 밝혀내거나, 전체 데이터의 내용을 함축해 주기도 한다. 이러한 특징으로 인해 실생활의 다양한 분야에서 활용된다.
- [0003] 데이터 마이닝 알고리즘 중 k번째 항목을 찾는 것은 데이터 마이너(miner)들의 전체 숫자 데이터에서 상위 k번째 항목을 추정하여 찾아내는 것이다. 이러한 알고리즘은 주로 수직 분할된 데이터에 주로 적용이 된다. 수직 분할된 데이터는 데이터 마이너들 각각이 갖고 있는 데이터가 동일한 데이터 집합에 관한 것이고, 다만 그 집합에서의 서로 다른 속성을 마이너들이 가지고 있다.
- [0004] 예를 들어, 한 학년의 국어 과목 담당자와 수학 과목 담당자가 각각 자신이 담당하는 과목의 데이터는 가지고 있으나, 다른 과목의 데이터는 가지고 있지 않기 때문에 각각 국어 점수와 수학 점수를 수직 분할된 데이터라고 할 수 있다. 수직 분할된 데이터에서 k번째 항목을 찾는다는 것은 예를 들었던 국어, 수학 점수의 전체 합산 데이터에서 상위 k번째 점수를 갖는 항목을 찾아내는 것 또는 k번째 점수를 찾아내는 것을 의미한다. 이러한 환경에서 데이터 마이닝을 적용할 때 프라이버시 침해 문제가 나타날 수 있다.
- [0005] 예를 들어 기업에서 직원들을 대상으로 인사고과를 수행할 때, 해당 직원에 대한 평가 점수가 그대로 드러난다면 개인 프라이버시 문제로 발전할 수 있게 된다. 이런 문제를 해결하기 위한 데이터 마이닝 기법이 프라이버시 보호 데이터 마이닝이다. 프라이버시 보호 데이터 마이닝은 기존의 데이터 마이닝에서 사용하던 원본 데이터의 분포를 변화시키거나, 암호학 기술을 사용하여 데이터의 민감한 정보를 보호하면서 원본 데이터의 마이닝 결과와 동일하도록 하는 것이다.
- [0006] 논문(J. Vaidya, and C. W. Clifton, K-th Element Score over Vertically Partitioned Data, Knowledge and Data Engineering, vol. 21, No. 2, pp. 253-258, 2009.)에서 처음으로 Yao의 비교법을 사용하여 프라이버시 보장 k번째 항목 찾기 기법을 제안하였다. 하지만 해당 기법의 경우 수행 가능한 마이너가 둘뿐이라는 단점이 존재했다.
- [0007] 이에 따라, 여러 사람이 동시에 k번째 항목을 찾는 데이터 마이닝을 수행하면 알고리즘을 여러 번 수행해야 하는 단점이 존재하고 공모 공격에 취약한 단점이 존재한다.
- [0008] 따라서, 공모 공격에 안전하고 다자간에도 기존의 알려진 기법과 동일한 알고리즘 수행 횟수를 가지고서 k번째 항목을 찾을 수 있도록 하는 k번째 항목의 스코어 산출 방법 및 산출 시스템이 필요하다.

발명의 내용

해결하려는 과제

- [0009] 본 발명은 상술한 문제점을 해결하기 위해서 안출한 것으로서, 공모 공격에 안전하고 다자간에도 기존의 알려진 기법과 동일한 알고리즘 수행 횟수를 가지고서 k번째 항목을 찾을 수 있도록 하는, k번째 항목의 스코어 산출 방법 및 산출 시스템을 제공하는 데 그 목적이 있다.

과제의 해결 수단

- [0010] 상기와 같은 목적을 달성하기 위한 k번째 항목의 스코어 산출 방법은 (a) 제1 장치의 공개키로 복수의 항목들 중 제1 항목의 암호화된 제1 속성 데이터를 수신하는 단계, (b) 제1 장치의 공개키로 제1 항목에 대응하는 제2 속성 데이터 및 추정값에 기초한 데이터를 각각 암호화하는 단계 및 (c) 암호화된 제1 속성 데이터, 암호화된 제2 속성 데이터 및 추정값에 기초한 데이터를 제1 장치로 전송하는 단계를 포함하며, 제1 속성 및 제2 속성은 제1 항목에서의 서로 다른 속성을 나타낸다.

[0011] 상기와 같은 목적을 달성하기 위한 k번째 항목의 스코어 산출 시스템은 제1 장치의 공개키로 복수의 항목들 중 제1 항목의 암호화된 제1 속성 데이터를 수신하고, 제1 장치의 공개키로 제1 항목에 대응하는 제2 속성 데이터 및 추정값에 기초한 데이터를 각각 암호화하며, 암호화된 제1 속성 데이터, 암호화된 제2 속성 데이터 및 추정값에 기초한 데이터를 제1 장치로 전송하는 스코어 산출 장치를 포함하며, 제1 속성 및 제2 속성은 상기 제1 항목에서의 서로 다른 속성을 나타낸다.

발명의 효과

[0012] 상기와 같은 본 발명에 따른 k번째 항목의 스코어 산출 방법 및 산출 시스템은 공모 공격에 안전하고 다자간에도 기존의 알려진 기법과 동일한 알고리즘 수행 횟수를 가지고서 k번째 항목을 찾을 수 있도록 하는 효과가 있다.

도면의 간단한 설명

[0013] 도 1은 k번째 항목의 스코어 산출 시스템의 예시적인 시스템도를 도시한 도면이다.
 도 2는 k번째 항목의 산출을 위해 협업하는 장치간의 데이터 흐름을 도시한 도면이다.
 도 3은 k번째 항목의 산출을 위해 협업하는 장치간 후속하는 데이터 흐름을 도시한 도면이다.

발명을 실시하기 위한 구체적인 내용

[0014] 상술한 목적, 특징 및 장점은 첨부된 도면을 참조하여 상세하게 후술 되어 있는 상세한 설명을 통하여 더욱 명확해 질 것이며, 그에 따라 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 본 발명의 기술적 사상을 용이하게 실시할 수 있을 것이다. 또한, 본 발명을 설명함에 있어서 본 발명과 관련된 공지 기술에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에 그 상세한 설명을 생략하기로 한다. 이하, 첨부된 도면을 참조하여 본 발명에 따른 바람직한 실시 예를 상세히 설명하기로 한다.

[0015] 도 1은 k번째 항목의 스코어 산출 시스템의 예시적인 시스템도를 도시한 도면이다.

[0016] 도 1에 따르면 스코어 산출 시스템은 복수의 장치(100)를 포함하고 바람직하게는 3개 이상의 장치(100)들을 포함한다. 각 장치(100)들은 네트워크를 통해 서로 연결되고 네트워크를 통해 k번째 항목의 스코어 산출을 위한 데이터를 송수신할 수 있다.

[0017] 도 1의 각 장치(100)들을 살펴보면, 복수의 장치(100)들 각각은 k번째(k는 1 이상) 항목(element)의 스코어 산출을 위해서 서로 협업하도록 구성된다. 각 장치(100)들은 부여된 기능에 따라 다른 용어로 지칭될 수 있다. 암호화에 이용될 공개키를 제공하는 장치(100)는 '공개키 제공 장치'로 지칭될 수 있고 제공된 공개키로 장치(100)가 가지고 있는 항목의 속성을 암호화하는 장치(100)를 '암호화 장치'로 지칭될 수 있다. 또한, k번째 항목의 스코어를 산출하는 장치(100)를 '스코어 산출 장치'로 지칭될 수 있다.

[0018] 각 장치(100)들은 스코어 산출 장치(100)와 협업하여 부여된 기능에 따라 공개키를 제공하거나 공개키에 따라 암호화하여 다른 장치(100)나 스코어 산출 장치(100)로 암호화된 데이터를 제공 가능하도록 구성된다. k번째 항목의 스코어 계산은 여러 라운드(Round)에 걸쳐서 이루어질 수 있고 각 라운드에서의 각 장치(100)에 부여된 기능은 변경가능하다. 따라서, 첫 번째 라운드에서의 공개키 제공 장치(100)는 두 번째 라운드에서 암호화 장치(100) 또는 스코어 산출 장치(100)로서 기능할 수 있다.

[0019] 장치(100) 각각은 입력 인터페이스, 출력 인터페이스, 메모리, 대용량 저장매체, 통신 인터페이스 및 프로세서를 포함한다. 각 장치(100)의 구성요소를 간단히 살펴보면, 입력 인터페이스는 사용자 입력을 수신하도록 구성된다. 예를 들어, 입력 인터페이스는 마우스, 키보드, 마이크 등을 구비하여 사용자 입력을 수신한다. 입력 인터페이스를 통해 k 번째 항목의 계산 요청을 장치(100)가 수신할 수 있다. 출력 인터페이스는 사용자 입력에 따른 처리 결과 또는 프로그램에 의한 처리 결과를 출력할 수 있도록 구성된다. 예를 들어, 출력 인터페이스는 디스플레이, 스피커, LED 등을 구비하여 처리 결과를 출력한다.

[0020] 장치(100)의 메모리는 휘발성 메모리를 포함하고 비휘발성 메모리를 더 포함할 수 있다. 메모리는 각종 데이터와 프로그램의 코드를 임시로 저장하거나 영구히 저장할 수 있다. 대용량 저장매체는 각종 데이터와 프로그램을 저장한다. 대용량 저장매체는 적어도 본 발명에 따른 k번째 항목의 스코어 산출을 위한 산출 프로그램과 k번째

항목의 스코어 산출에 이용될 속성값들을 저장한다. 산출 프로그램은 다른 장치(100)와 협업하여 k번째 항목의 스코어를 산출하고 그 산출 결과를 출력한다.

[0021] 대용량 저장매체에 저장되는 속성값은 항목의 식별자와 함께 각 항목별로 저장된다. 속성값은 예를 들어 특정 과목의 점수값이거나 특정 관리 대상자의 특정 분야의 고과 점수 등일 수 있다. 한 장치(100)에 저장되는 속성값에 대응하는 속성(property)(예를 들어 국어 점수)은 다른 장치(100)에 저장되는 속성값의 속성(예를 들어 영어 점수)과 서로 다르고 동일한 항목에 대한 수직 분할 데이터를 구성한다. 장치(100)에서 관리하는 데이터 집합은 복수 개의 항목으로 구성된다. 데이터 집합은 각각의 항목에 대해 적어도 항목의 식별자 및 대응하는 항목의 속성값을 식별할 수 있도록 구성된다. 데이터 집합은 리스트의 형태로 구성될 수도 있다. 항목의 식별자는 예를 들어 학생들을 식별할 수 있거나 인사고과 대상자 등을 식별할 수 있다. 설명의 간단한 이해를 위해 항목의 식별자는 이하 숫자(i)로 표시되고 관리 대상자의 인원만큼의 인덱스(인원수가 2000명인 경우 1에서 2000의 인덱스)를 가지는 것으로 가정한다.

[0022] 장치(100)의 통신 인터페이스는 네트워크에 연결되어 네트워크 패킷을 송수신한다. 통신 인터페이스는 예를 들어 무선랜 또는 유선랜을 통해 다른 장치(100)에 서로 연결되고 데이터의 네트워크 패킷을 서로 송수신할 수 있다.

[0023] 장치(100)의 프로세서는 입력 인터페이스를 통한 사용자 입력 또는 통신 인터페이스를 통한 요청에 따라 대용량 저장매체에 저장된 프로그램 및 데이터를 로딩하고 프로그램을 수행하여 각 블록들을 제어한다. 예를 들어 프로세서는 산출 프로그램을 휘발성 메모리 등에 로딩하여 산출 프로그램의 프로그램 코드에 따라 블록들을 제어하여 k 번째 항목의 스코어를 계산할 수 있다. k는 데이터 집합의 장치(100)들에 산재된 각 항목별 속성값들의 합산된 정렬 집합(예를 들어 내림 차순으로 정렬된 집합)에서 k번째로 큰(또는 작은) 합산값을 찾기 위해서 이용된다. k는 입력 인터페이스를 통한 사용자 입력에 따라 지정될 수 있다. 프로세서는 소위 CPU이거나 MPU 등일 수 있다. 프로세서에 이루어지는 제어 흐름은 도 2 및 도 3을 통해서 상세히 살펴보도록 한다.

[0024] 장치(100)는 예를 들어 개인용 컴퓨터, 노트북, 태블릿 PC, 휴대폰, 스마트폰, 서버 등일 수 있다. 복수의 장치(100)들은 모두 동일한 유형의 장치(100)일 필요는 없다. 예를 들어, 하나의 장치(100)는 개인용 컴퓨터이고 다른 하나의 장치(100)는 노트북이고 또 다른 하나의 장치(100)는 스마트폰이거나 서버 등일 수 있다.

[0025] 장치(100)들은 서로 네트워크를 통해서 연결되는 데, 네트워크는 근거리 네트워크이거나 광대역 네트워크일 수 있다. 근거리 네트워크는 예를 들어 유선랜 및/또는 무선랜으로 구성된다. 광대역 네트워크는 하나 이상의 근거리 네트워크를 포함한다. 네트워크를 통해 전송 계층(Transport Layer) 이상의 네트워크 패킷을 장치(100)간에 서로 송수신할 수 있다.

[0026] 도 2는 k번째 항목의 산출을 위해 협업하는 장치(100)간의 데이터 흐름을 도시한 도면이다. 도 2의 데이터 흐름은 3개의 장치(100)가 협업하는 것의 예를 도시한 도면이다. 도 2의 데이터 흐름에서 각 장치(100)들은 바람직하게는 각 장치(100)의 프로세서 및 통신 인터페이스를 이용하여 다른 장치(100)들과 데이터를 송수신할 수 있도록 구성된다.

[0027] 도 2의 데이터 흐름은 바람직하게는 특정 한 장치(100)에서의 입력 인터페이스를 통한 사용자 입력으로 시작된다. k를 나타내는 사용자 입력에 따라 장치(100)들은 각 장치(100)들의 기능을 부여하고 셋업 데이터를 공유하기 위한 셋업 과정을 진행(① 참조)한다. 기능의 부여는 임의의 장치(100)에 의해서 수행된다. 예를 들어 사용자 입력을 수신한 장치(100)가 장치(100)들의 기능을 부여하거나 지정된 특정 장치(100)가 장치(100)들의 기능을 부여할 수 있다. k번째 항목의 스코어는 예를 들어 데이터 집합에서 항목들의 속성값들의 합산값 중 k번째로 큰 항목의 합산값을 나타낸다.

[0028] 셋업 과정 동안에, 특정 하나의 장치(100)는 공개키 제공 장치(100)로 설정(도 2의 장치 A)되고, 다른 하나 이상의 장치(100)는 암호화 장치(100)로 설정(도 2의 장치 B)되고 또 다른 하나의 장치(100)는 스코어 산출 장치(100)로 설정(도 2의 장치 C)된다.

[0029] 셋업 과정에서 각각의 장치(100)는 공개키 및 개인키를 생성한다. 특히 공개키 제공 장치(100)는 이 장치(100)의 공개키를 다른 장치(100)들로 전송한다. 예를 들어 k 번째 항목의 스코어가 가질 수 있는 최대값을 $|F|$ (각 장치(100)가 가지는 속성이 국어 점수, 영어 점수 또는 수학 점수일 때, $|F|$ 는 300점)라고 가정할 때 공개키(n)는 $2 \cdot |F|$ 보다 큰 값으로 설정된다. 공개키는 각 장치(100)별로 설정되고 소수 p, q의 곱으로 만들어진다. 공개키는 k 번째 항목 스코어 산출 과정에서 각 속성값의 노출을 막기 위해서 이용되는 추정값으로 인해 최소

2*|F|보다 크도록 설정된다.

- [0030] 셋업 과정에서 지정된 스코어 산출 장치(100)는 상한값(upper bound, 'ubound'라고도 지칭됨.), 하한값(lower bound, 'lbound'라고도 지칭됨.) 및 추정값(estimate)을 설정하고 설정된 추정값, 나아가 상한값 및 하한값을 다른 장치(100)들로 전송한다. 적어도 스코어 산출 장치(100)는 추정값을 공개키 제공 장치(100)로 전송한다.
- [0031] 스코어 산출 장치(100)는 스코어 산출을 위해 최초로 상한값을 최대값(|F|)으로 설정하고 하한값을 속성들의 합의 최소값(예를 들어 0)으로 설정한다. 스코어 산출 장치(100)는 추정값을 설정하여 적어도 공개키 제공 장치(100)로 전송하는 데, 스코어 산출 장치(100)는 지정된 난수 범위(Q^α)에서 난수값을 선택한다. 초기의 난수 범위는 데이터 집합의 항목들의 개수에 기초해서 결정되며 k번째 스코어 산출을 위해 반복이 예상되는 최대 라운드 개수에 기초해서 설정된다. 초기의 난수 범위는 항목들의 개수와 최대 라운드 개수에 따라 예를 들어 +/- 10으로 설정되고, 스코어 산출 장치(100)는 초기의 난수 범위에서 임의의 난수값을 선택하고 난수값을 상한값과 하한값의 중간값에 적용(합산)하여 추정값을 설정한다.
- [0032] 셋업 과정의 완료 후에, 공개키 제공 장치(100)는 대용량 저장매체에 포함된 데이터 집합의 각 항목(i)의 속성값을 자신의 공개키로 암호화(②)의 $E_{Pa}(a_i)$, 공개키는 Pa , 속성값은 a_i 임.)하고 셋업과정에서 설정된 후속하는 암호화 장치(100)로 암호화된 속성 데이터를 전송(③ 참조)한다.
- [0033] 암호화 장치(100)는 암호화된 속성 데이터를 수신하고 공개키 제공 장치(100)의 공개키로 암호화된 속성 데이터에 대응하고 대용량 저장매체에 저장된 항목(i)의 속성값(b_i)을 암호화하고 암호화된 데이터($E_{Pa}(b_i)$)와 수신된 암호화 데이터를 곱하여 전송할 속성 데이터를 생성(④ 참조)한다. 생성된 속성 데이터는 물론 공개키로 암호화되어 있고 적어도 특정 항목에 대한 암호화 장치(100)에서의 속성값과 공개키 제공 장치(100)에서의 속성값을 포함한다. 각각의 속성값은 서로 다른 속성에 대한 속성값을 나타낸다.
- [0034] 이후 암호화 장치(100)는 ④를 통해 생성된 암호화된 속성 데이터를 스코어 산출 장치(100)로 전송하고 스코어 산출 장치(100)는 이 속성 데이터를 수신(⑤ 참조)한다.
- [0035] 암호화된 속성 데이터의 수신에 따라, 스코어 산출 장치(100)는 공개키 제공 장치(100)의 공개키로 암호화된 속성 데이터에 대응하고 대용량 저장매체에 저장된 항목(i)의 속성값(c_i)을 암호화($E_{Pa}(c_i)$)한다. 또한 스코어 산출 장치(100)는 셋업 과정을 통해 설정된 추정값에 기초한 데이터를 암호화($E_{Pa}(-m)$)한다. 이후 스코어 산출 장치(100)는 수신되고 암호화된 속성 데이터, 스코어 산출 장치(100)에 저장된 속성값의 암호화된 속성 데이터 및 암호화된 추정값의 데이터를 곱하여 공개키 제공 장치(100)로 전송할 데이터를 생성(⑥ 참조)한다. 추정값에 기초한 데이터('mean' 값이라고도 지칭됨)는 "- 추정값 mod n"으로 설정될 수 있고 n은 스코어 산출 장치(100)의 공개키로 사용되는 값이다. 이 mean 값은 공개키 제공 장치(100)에서 이용되고 추정값보다 큰 값의 개수만을 파악하기 위한 용도로 이용되며 추정값과의 비교를 위해서 이용된다. 스코어 산출 장치(100)에 저장된 속성값의 속성과 수신된 속성 데이터에 포함된 속성들은 동일한 항목(i)에 대한 서로 다른 속성을 나타낸다.
- [0036] 이후 스코어 산출 장치(100)는 생성된 데이터를 공개키 제공 장치(100)로 전송(⑦ 참조)한다.
- [0037] ① 내지 ⑦의 과정은 데이터 집합의 모든 항목($1 \leq i < n$, n은 항목들의 인덱스 최대값)들에 대해서 각각 수행되고 이에 따라 공개키 제공 장치(100)는 각각의 항목(i)별 암호화된 데이터를 수신하도록 구성된다.
- [0038] 공개키 제공 장치(100)는 각각의 항목별 암호화된 데이터를 자신의 개인키로 복호화한다.
- [0039] 여기서, 본 발명에 이용되는 암호 기법에 대해서 살펴보면, 본 발명은 준동형(homographic) 성질을 가지는 암호 기법을 이용하고 바람직하게는 Paillier의 공개키 암호 기법을 이용하여 암호화 및 복호화가 이루어진다.
- [0040] Paillier의 공개키 암호 기법은 지수 연산의 어려움에 기초하고 RSA와 달리 준동형 성질을 이용하여 암호화된 상태에서 연산이 가능하다. 즉, Paillier의 공개키 암호 기법은 곱해져 있는 각각의 암호화된 데이터를 복호화하여 복호화된 데이터의 덧셈 결과값을 알 수 있도록 한다.
- [0041] 복호화에 따라 그리고 준동형 성질에 따라 공개키 제공 장치(100)는 각 항목의 복호화된 데이터

($a_i + b_i + c_i - m$)를 추정값과 비교하고 각 항목의 비교 결과(output)를 합산(⑧ 참조, Q_{output})한다. 합산된 비교 결과는 스코어 산출 장치(100)로 전송(⑨ 참조)된다.

- [0042] 구체적으로 살펴보면, 각 항목의 수신된 데이터에 대한 복호화로 각 장치(100)의 속성값과 mean(-추정값) 값들의 합산값을 추출하고 추출된 합산값을 n(공개키 제공 장치(100)의 공개키)으로 모듈러 연산하고 모듈러 연산된 데이터를 추정값과 비교한다. 만일 합산값이 추정값보다 큰 경우 비교 결과(출력값)를 1로 그렇지 않은 경우 0으로 설정한다. 각 항목의 비교 결과를 합산하여 합산된 비교 결과를 스코어 산출 장치(100)로 전송한다. 합산된 비교 결과는 적어도 추정값 보다 큰 항목의 개수가 몇 개인지를 알 수 있도록 한다.
- [0043] 합산된 비교 결과의 수신에 따라, 스코어 산출 장치(100)는 합산된 비교 결과와 k(셋업 과정에서 설정됨)와 비교하고 k번째 항목의 스코어를 산출하거나 이후 k번째 항목의 스코어의 산출을 위한 준비를 수행(⑩ 참조)하고 후속하는 라운드(이하 '비교 결과 산출 라운드' 또는 줄여서 '라운드'라고도 함)(도 3 참조)을 진행한다.
- [0044] 구체적으로, 스코어 산출 장치(100)는 수신된 합산된 비교 결과의 결과값(Q_{output})을 k와 비교한다. 그 비교 결과에 따라 결과값이 k 보다 작으면 상한값을 mean 값(또는 추정값)으로 변경하고 상한값과 하한값의 차이가 1이면 mean 값(또는 추정값)을 k번째 항목의 스코어로 결정한다. 만일 비교 결과의 결과값이 k보다 크면 하한값을 mean 값(또는 추정값)으로 변경한다. 만일 비교 결과 결과값이 k이면 mean 값(또는 추정값)을 k번째 항목의 스코어로 결정한다. 비교 결과의 결과값이 k보다 크거나 작은 경우(즉 k번째 항목의 스코어가 결정되지 못한 경우) 추정값이 또한 변경된다.
- [0045] 변경된 추정값은 후속하는 비교 결과 산출 라운드에서 이용되고, 변경되는 추정값은 난수값을 적용하여 설정된다. 스코어 산출 장치(100)는 후속하는 비교 결과 산출 라운드를 위해 난수 범위를 먼저 설정한다. 이 난수 범위는 앞선 비교 결과 산출 라운드에서 이용된 난수 범위보다 좁은 범위이다. 예를 들어 앞선 라운드에서 +/-10의 범위였다면 후속하는 라운드에서는 1/2로 줄은 범위(+/-5)에서 난수 범위가 결정된다. 이와 같이 각 비교 결과 산출 라운드에서 이용되는 난수 범위는 이전 라운드의 난수 범위보다 좁아지고 스코어 값의 범위(상한값과 하한값의 범위)보다 항상 좁도록 구성되어 스코어 산출을 용이하게 한다.
- [0046] 스코어 산출 장치(100)는 새로이 설정된 난수 범위 내에서 임의의 난수값을 선택하고 변경된 상한값과 하한값 사이의 중간값에서 선택된 난수값을 적용하여 후속 라운드를 위한 추정값을 변경한다. 변경된 상한값, 하한값 및 추정값과 나아가 난수 범위는 다른 장치(100)들로 전송될 수 있다.
- [0047] 만일 k번째 항목의 스코어가 산출된 경우, 스코어 산출 장치(100)는 산출된 k번째 항목의 스코어를 다른 장치(100)들로 전송하고 더 이상의 비교 결과 산출 라운드를 진행하지 않고 도 2의 흐름이 종료된다. 여기서 전송되는 스코어는 적어도 k번째 항목과 k+1 번째 항목을 구분가능하도록 한다.
- [0048] k번째 항목의 수신에 따라 장치(100)들(특히 k번째 항목의 검색 요청한 장치(100))은 출력 인터페이스를 통해 해당 스코어를 출력할 수 있다.
- [0049] 도 3은 k번째 항목의 산출을 위해 협업하는 장치(100)간 후속하는 데이터 흐름을 도시한 도면이다. 도 3의 데이터 흐름은 도 2의 데이터 흐름의 완료 후에 후속하는 비교 결과 산출 라운드로서 진행된다. 도 3의 데이터 흐름은 각 비교 결과 산출 라운드에서 k번째 항목의 스코어를 결정하지 못한 경우에 반복적으로 수행되고 각 반복에서 이용되는 상한값, 하한값 및 추정값들은 변경된다.
- [0050] 도 3의 데이터 흐름은 도 2와 대부분 동일하므로, 여기서는 도 2와 차이가 나는 내용을 중심으로 설명하도록 한다. 본 설명에서 생략된 부분은 도 2의 설명을 통해 대체된다.
- [0051] 후속 라운드의 수행을 위해 장치(100)들은 서로 협업을 통해 각 장치(100)들의 기능을 부여하고 셋업 데이터를 공유하기 위해 셋업 과정을 진행(① 참조)한다.
- [0052] 협업하는 각 장치(100)들은 기능이 부여된다. 예를 들어 하나의 장치(100)는 공개키 제공 장치(100)로 설정되고 다른 하나의 장치(100)는 암호화 장치(100)로 설정되고 또 다른 하나의 장치(100)는 스코어 산출 장치(100)로 설정된다. 적어도 이전 라운드에서의 공개키 제공 장치(100)와 현재 라운드에서의 공개키 제공 장치(100)는 다르게 설정될 수 있다. 예를 들어, 이전 라운드에서의 공개키 제공 장치(100)는 암호화 장치(100)로 설정되고 이전 라운드에서의 한 암호화 장치(100)는 공개키 제공 장치(100)로 설정된다. 현재 설정된 공개키 제공 장치

(100)는 다른 장치(100)들로 공개키를 전송한다. 각 장치(100)의 공개키 및 개인키는 최초의 라운드에서 생성된 공개키 및 개인키일 수 있다.

[0053] 각 라운드 별로 공개키를 제공하는 공개키 제공 장치(100)를 서로 다른 장치(100)로 설정하는 데, 이는 동일한 장치(100)에 권한이 집중되는 것을 방지하기 위한 것이다. 특히 동일한 장치(100)의 공개키가 각 라운드에서 지속적으로 사용되게 되면 외부 공격에 취약해지게 된다. 또한 서로 다른 장치(100)를 공개키 제공 장치(100)로 설정함으로써 각 라운드별 각 장치(100)의 연산량을 분산할 수 있도록 하고 나아가 각 라운드에서 사용되는 공개키가 달라 외부 공격으로부터 안전하게 설계 가능하다.

[0054] 셋업 과정의 설정 이후에, 공개키 제공 장치(100)는 저장되어 있는 항목의 속성값을 자신의 공개키로 암호화하여 암호화된 속성 데이터를 다른 암호화 장치(100)로 전송(② 참조)한다.

[0055] 암호화된 속성 데이터를 수신한 암호화 장치(100)는 자신의 항목 속성값을 수신된 공개키로 암호화하고 수신되고 암호화된 속성 데이터를 자신의 속성값으로 암호화한 속성 데이터에 곱하여 곱해진 속성 데이터를 스코어 산출 장치(100)로 전송(③ 참조)한다.

[0056] 스코어 산출 장치(100)는 암호화된 속성 데이터를 수신하고 자신의 항목 속성값과 추정값에 기초한 데이터를 각각 암호화하고 각각의 데이터를 곱하여 곱해진 데이터를 공개키 제공 장치(100)로 전송(④ 참조)한다. 추정값에 기초한 데이터(mean 값)은 " $-$ 추정값 mod n "으로 설정될 수 있고 n 은 스코어 산출 장치(100)의 공개키로 사용되는 값이다. 각 장치(100)에서 암호화되는 속성값은 동일한 항목(i)에 대한 서로 다른 속성에 대한 값을 나타낸다.

[0057] 공개키 제공 장치(100)는 단계 ① 내지 ④에 따라 암호화된 데이터를 각 항목별로 수신하고 수신된 데이터를 개인키로 복호화하고 복호화된 데이터와 mean 값을 비교하고 1 또는 0으로 출력되는 비교 결과를 합산한다. 합산된 비교 결과는 스코어 산출 장치(100)로 전송(⑤ 참조)한다.

[0058] 스코어 산출 장치(100)는 합산된 비교 결과를 k 와 비교하고 그 비교 결과에 따라 상한값, 하한값 및/또는 추정값을 변경하고 후속하는 라운드를 진행할 수 있다. 후속하는 라운드는 도 3과 동일하다. 또는 스코어 산출 장치(100)는 k 번째 항목을 찾은 경우에 해당 k 번째 항목의 스코어를 다른 장치(100)들로 전송한다. 수신된 스코어는 장치(100)들의 출력 인터페이스를 통해 출력될 수 있다. 여기서 전송되는 스코어는 적어도 k 번째 항목과 $k+1$ 번째 항목을 구분가능하도록 한다.

[0059] 도 2 및 도 3의 흐름을 통해서, 각 장치(100)간의 데이터(속성값)의 노출 없이 k 번째 항목의 산출을 위한 연산이 가능하다. 예를 들어 한 반의 석차를 계산하고자 하는 경우 각 과목의 담당자들이 서로 데이터(속성값)의 공유 없이도 석차 계산이 가능하다. 인사과과를 통한 승진 대상자 심사의 경우에도 실적, 근무태도 등의 정보를 각 담당자간의 공유 없이도 종합가능하고 이에 따라 승진 여부를 판단할 수 있다.

[0060] 도 2 및 도 3에서, 임의의 난수값을 추가하여 둘 이상의 장치(100)(사용자)가 다른 장치(100)의 속성값을 알아내기 위한 공모 공격에 효과적으로 대처할 수 있다. 예를 들어 둘 이상의 장치(100)가 공모하는 경우라도 임의의 난수값에 의해 다른 장치(100)(특히 스코어 산출 장치(100))의 속성값을 알 수 없어 효과적으로 공모 공격에 대처할 수 있다.

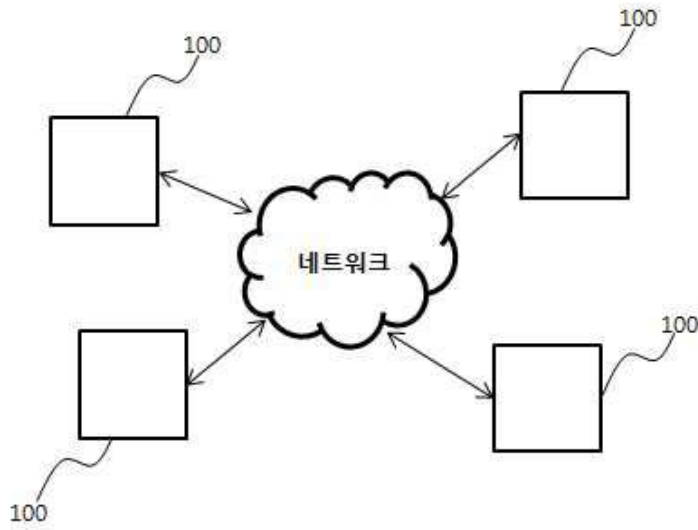
[0061] 이상에서 설명한 본 발명은, 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에게 있어 본 발명의 기술적 사상을 벗어나지 않는 범위 내에서 여러 가지 치환, 변형 및 변경이 가능하므로 전술한 실시 예 및 첨부된 도면에 의해 한정되는 것이 아니다.

부호의 설명

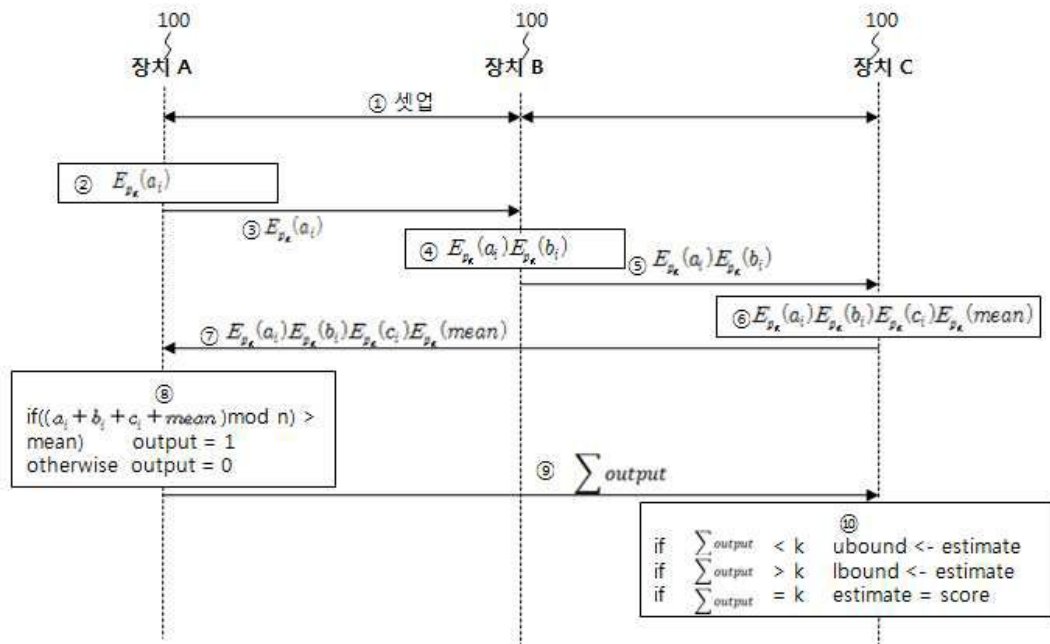
[0062] 100 : 장치

도면

도면1



도면2



도면3

