



**(19) 대한민국특허청(KR)**  
**(12) 등록특허공보(B1)**

(45) 공고일자 2016년05월13일  
 (11) 등록번호 10-1620881  
 (24) 등록일자 2016년05월09일

- (51) 국제특허분류(Int. Cl.)  
 H04L 29/06 (2006.01) H04L 29/08 (2006.01)  
 H04L 9/32 (2006.01)
- (52) CPC특허분류  
 H04L 63/0407 (2013.01)  
 H04L 63/18 (2013.01)
- (21) 출원번호 10-2015-7003929
- (22) 출원일자(국제) 2013년07월03일  
 심사청구일자 2015년02월13일
- (85) 번역문제출일자 2015년02월13일
- (65) 공개번호 10-2015-0036624
- (43) 공개일자 2015년04월07일
- (86) 국제출원번호 PCT/EP2013/001958
- (87) 국제공개번호 WO 2014/012627  
 국제공개일자 2014년01월23일
- (30) 우선권주장  
 12290234.9 2012년07월16일  
 유럽특허청(EPO)(EP)
- (56) 선행기술조사문헌  
 US20020091736 A1  
 US20080205774 A1  
 US6134597 B1

- (73) 특허권자  
**알까멜 루슨트**  
 프랑스 92100 불론뉴-비영꾸르 루뜨 들 라 렌느  
 148/152
- (72) 발명자  
**아가사라이언, 아르멘**  
 프랑스 91620 노자이 루트 드 빌퀴스페 상트르 드  
 빌라르세오스 알까멜-루슨트 벨 램 프랑스  
**부지드, 마크람**  
 프랑스 91620 노자이 루트 드 빌퀴스페 상트르 드  
 빌라르세오스 알까멜-루슨트 벨 램 프랑스  
 (뒷면에 계속)
- (74) 대리인  
**장훈**

전체 청구항 수 : 총 15 항

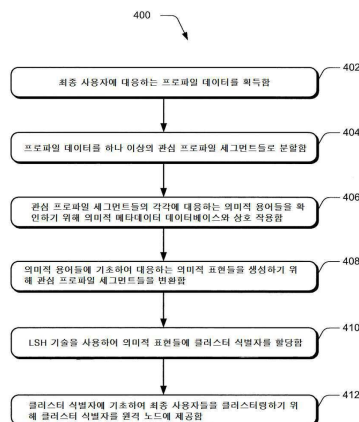
심사관 : 문형섭

**(54) 발명의 명칭 사용자 관심 프로파일들의 프라이버시 보호된 클러스터링을 위한 방법 및 장치**

**(57) 요약**

본 주제의 일 구현에 따라, 사용자 관심 프로파일들의 프라이버시 보호된 클러스터링을 위한 장치 및 방법들이 설명된다. 상기 방법은 최종 사용자의 관심 프로파일에 기초하여 적어도 하나의 관심 프로파일 세그먼트를 생성하는 단계(404)를 포함한다. 뿐만 아니라, 상기 적어도 하나의 관심 프로파일 세그먼트에 대응하는 의미 용어들(뒷면에 계속)

**대표도** - 도4



은 상기 최종 사용자의 사용자 디바이스에 결합된 의미 메타데이터 데이터베이스와의 상호작용에 기초하여 획득되고(406), 여기에서 상기 의미 용어들은 상기 의미 메타데이터 데이터베이스에서 제공된 하나 이상의 의미 용어들 중에서 획득된다. 상기 적어도 하나의 관심 프로파일 세그먼트의 각각은 그 뒤에 적어도 하나의 의미 표현으로 변환된다(408). 뿐만 아니라, 클러스터 식별자는 상기 적어도 하나의 의미 표현에 기초하여 상기 적어도 하나의 관심 프로파일 세그먼트에 할당되고, 상기 클러스터 식별자들은 지역에 맞는 해싱(LSH) 기술을 사용하여 생성된다(410).

(52) CPC특허분류

*H04L 67/306* (2013.01)

*H04L 9/3236* (2013.01)

(72) 발명자

**코타리, 모히트**

인도 구자라트 382475 아메다바드 한솔 니어 에어

포트 소사이어티 볼레쉬와르 코퍼레이티브 52

**난디, 아니메쉬**

인도 웨스트 벵갈 700074 콜카타 뭉뭉 모티즈힐 에

비뉴 26 4쓰 플로어 가우리 어파트먼트

## 명세서

### 청구범위

#### 청구항 1

최종 사용자들의 관심 프로파일들의 프라이버시 보호된 클러스터링(privacy protected clustering)을 위한 방법에 있어서:

최종 사용자의 관심 프로파일에 기초하여 적어도 하나의 관심 프로파일 세그먼트를 생성하는 단계;

상기 최종 사용자의 사용자 디바이스에 통신가능하게 결합된 의미 메타데이터 데이터베이스와의 상호작용에 기초하여 상기 적어도 하나의 관심 프로파일 세그먼트에 대응하는 의미 용어들을 획득하는 단계로서, 상기 의미 용어들은 상기 의미 메타데이터 데이터베이스에 제공된 하나 이상의 의미 용어들 중에서 획득되는, 상기 획득하는 단계;

상기 의미 용어들에 기초하여 상기 적어도 하나의 관심 프로파일 세그먼트의 각각을 적어도 하나의 의미 표현으로 변환하는 단계; 및

상기 적어도 하나의 의미 표현에 기초하여 상기 적어도 하나의 관심 프로파일 세그먼트의 각각에 클러스터 식별자를 할당하는 단계로서, 상기 클러스터 식별자는 지역에 맞는 해싱(locality sensitive hashing; LSH) 기술을 사용하여 생성되는, 상기 할당하는 단계를 포함하는, 프라이버시 보호된 클러스터링을 위한 방법.

#### 청구항 2

제 1 항에 있어서,

상기 방법은 상기 적어도 하나의 관심 프로파일 세그먼트 및 상기 적어도 하나의 관심 프로파일 세그먼트의 각각에 대응하는 상기 클러스터 식별자를 원격 노드에 제공하는 단계를 추가로 포함하고, 상기 원격 노드는 상기 클러스터 식별자와 연관된 관심 그룹에 대한 관심 그룹 집성기로서 동작하도록 구성되고, 상기 클러스터 식별자는 익명 통신 채널을 통해 제공되는, 프라이버시 보호된 클러스터링을 위한 방법.

#### 청구항 3

제 1 항에 있어서,

상기 방법은:

원격 노드에 대한 복수의 익명 통신 채널들을 확립하는 단계;

상기 관심 프로파일을 하나 이상의 프로파일 슬라이스들로 슬라이싱하는 단계; 및

상기 복수의 익명 통신 채널들을 통해 상기 클러스터 식별자 및 상기 하나 이상의 프로파일 슬라이스들 각각을 상기 원격 노드로 송신하는 단계로서, 상기 하나 이상의 프로파일 슬라이스들 각각은 별개의 익명 통신 채널들을 통해 송신되는, 상기 송신하는 단계를 추가로 포함하는, 프라이버시 보호된 클러스터링을 위한 방법.

#### 청구항 4

제 3 항에 있어서,

상기 방법은 상기 관심 프로파일로부터, 상기 최종 사용자의 개인적으로 식별가능한 정보를 제거하는 단계를 추가로 포함하는, 프라이버시 보호된 클러스터링을 위한 방법.

#### 청구항 5

제 1 항에 있어서,

상기 방법은:

복수의 익명 통신 채널들을 통해 상기 클러스터 식별자를 송신하는 단계;

상기 복수의 익명 통신 채널들을 통해 다른 사용자 디바이스들의 클러스터 식별자들을 수신하는 단계;

상기 송신 및 상기 수신에 기초하여, 상기 최종 사용자의 클러스터 식별자와 유사한 클러스터 식별자를 가진 상기 다른 사용자 디바이스들을 식별하는 단계; 및

상기 익명 통신 채널을 통해 상기 최종 사용자의 사용자 디바이스와 상기 다른 사용자 디바이스들 사이에서의 데이터의 교환을 가능하게 하는 단계를 추가로 포함하는, 프라이버시 보호된 클러스터링을 위한 방법.

#### 청구항 6

제 1 항 내지 제 5 항 중 어느 한 항에 있어서,

상기 할당하는 단계는:

상기 적어도 하나의 관심 프로파일 세그먼트의 각각과 해시 함수의 세트 사이에서 내적을 획득하는 단계; 및

상기 적어도 하나의 관심 프로파일 세그먼트에 대응하는 상기 클러스터 식별자로서 해시 코드를 획득하기 위해 상기 내적의 부호들을 연결시키는 단계를 추가로 포함하는, 프라이버시 보호된 클러스터링을 위한 방법.

#### 청구항 7

제 1 항 내지 제 5 항 중 어느 한 항에 있어서,

상기 변환하는 단계는 상기 적어도 하나의 의미 표현을 획득하기 위해 상기 의미 용어들에 기초하여 상기 적어도 하나의 관심 프로파일 세그먼트의 각각을 태깅하는 단계를 추가로 포함하는, 프라이버시 보호된 클러스터링을 위한 방법.

#### 청구항 8

사용자 디바이스(102)에 있어서:

프로세서(202); 및

상기 프로세서(202)에 결합된 메모리(206)를 포함하고,

상기 메모리(206)는,

프로파일 변환 모듈(110)로서:

상기 사용자 디바이스(102)에 결합된 의미 메타데이터 데이터베이스(114)와의 상호작용에 기초하여 적어도 하나의 관심 프로파일 세그먼트에 대응하는 의미 용어들을 획득하고, 상기 의미 용어들은 상기 의미 메타데이터 데이터베이스(114)에서 제공된 하나 이상의 의미 용어들 중에서 획득되고;

상기 적어도 하나의 관심 프로파일 세그먼트의 각각을 적어도 하나의 의미 표현으로 변환하도록 구성된, 상기 프로파일 변환 모듈(110); 및

상기 적어도 하나의 의미 표현에 기초하여 상기 적어도 하나의 관심 프로파일 세그먼트의 각각에 클러스터 식별자를 할당하도록 구성된 클러스터 식별자 모듈(112)로서, 상기 클러스터 식별자는 지역에 맞는 해싱(LSH) 기술을 사용하여 생성되는, 상기 클러스터 식별자 모듈(112)을 포함하는, 사용자 디바이스.

#### 청구항 9

제 8 항에 있어서,

최종 사용자의 프로파일 데이터에 기초하여 상기 사용자 디바이스(102)와 연관된 상기 최종 사용자의 관심 프로파일을 생성하도록 구성된 관심 프로파일 생성 모듈(212); 및

상기 관심 프로파일을 상기 적어도 하나의 관심 프로파일 세그먼트로 분할하도록 구성된 프로파일 분할 모듈(214)을 추가로 포함하는, 사용자 디바이스.

#### 청구항 10

제 8 항 또는 제 9 항에 있어서,

데이터 전송 모듈(216)을 추가로 포함하고,

상기 데이터 전송 모듈(216)은:

상기 클러스터 식별자와 연관된 관심 그룹에 대한 관심 그룹 집성기로서 동작하도록 구성된 원격 노드(104)에 대한 익명 통신 채널(108)을 확립하고;

상기 적어도 하나의 관심 프로파일 세그먼트 및 상기 적어도 하나의 관심 프로파일 세그먼트들 각각에 대응하는 상기 클러스터 식별자를 상기 원격 노드(104)에 제공하고, 상기 클러스터 식별자는 상기 익명 통신 채널(108)을 통해 제공되도록 구성되는, 사용자 디바이스.

**청구항 11**

제 9 항에 있어서,

데이터 전송 모듈(216)을 추가로 포함하고,

상기 데이터 전송 모듈(216)은:

원격 노드(104)에 대한 복수의 익명 통신 채널들(108)을 확립하고;

상기 관심 프로파일로부터 상기 최종 사용자의 개인적으로 식별가능한 정보를 제거하고;

상기 복수의 익명 통신 채널들(108)을 통해 상기 최종 사용자의 클러스터 식별자 및 상기 관심 프로파일을 상기 원격 노드(104)로 송신하도록 구성되는, 사용자 디바이스.

**청구항 12**

제 8 항 또는 제 9 항에 있어서,

데이터 전송 모듈(216)을 추가로 포함하고,

상기 데이터 전송 모듈(216)은:

상기 관심 프로파일을 하나 이상의 프로파일 슬라이스들로 슬라이싱하고;

복수의 익명 통신 채널들(108)을 통해 상기 클러스터 식별자 및 상기 하나 이상의 프로파일 슬라이스들 각각을 원격 노드(104)로 송신하고, 상기 하나 이상의 프로파일 슬라이스들 각각은 별개의 익명 통신 채널들(108)을 통해 송신되도록 구성되는, 사용자 디바이스.

**청구항 13**

제 9 항에 있어서,

데이터 전송 모듈(216)을 추가로 포함하고,

상기 데이터 전송 모듈(216)은:

복수의 익명 통신 채널들(108)을 통해 상기 클러스터 식별자를 송신하고;

상기 복수의 익명 통신 채널들(108)을 통해 다른 사용자 디바이스들의 클러스터 식별자들을 수신하고;

상기 송신 및 상기 수신에 기초하여 상기 최종 사용자의 클러스터 식별자와 유사한 클러스터 식별자를 가진 상기 다른 사용자 디바이스들을 식별하고;

상기 익명 통신 채널(108)을 통해 상기 다른 사용자 디바이스들과의 데이터의 교환을 가능하게 하도록 구성되는, 사용자 디바이스.

**청구항 14**

제 8 항 또는 제 9 항에 있어서,

데이터 전송 모듈(216)을 추가로 포함하고,

상기 데이터 전송 모듈(216)은:

상기 사용자 디바이스(102)의 최종 사용자의 관심 그룹들의 중앙 집성기 노드의 적어도 부분적 이미지를 획득하고;

복수의 익명 통신 채널들(108)을 통해 다른 사용자 디바이스들과 상기 클러스터 식별자를 교환하고;

상기 복수의 익명 통신 채널들(108)을 통해 수신된 상기 다른 사용자 디바이스들의 클러스터 식별자들에 기초하여 상기 최종 사용자의 클러스터 식별자와 유사한 상기 클러스터 식별자를 가진 상기 다른 사용자 디바이스들을 식별하고;

상기 클러스터 식별자들에 기초하여 상기 익명 통신 채널(108)을 통해 추천 서비스들을 상기 다른 사용자 디바이스들에 제공하도록 구성되는, 사용자 디바이스.

**청구항 15**

최종 사용자들의 관심 프로파일들의 프라이버시 보호된 클러스터링의 방법을 실행하기 위한 컴퓨터 프로그램이 구현된 컴퓨터-판독가능한 매체에 있어서:

상기 방법은:

최종 사용자의 관심 프로파일에 기초하여 적어도 하나의 관심 프로파일 세그먼트를 생성하는 단계;

상기 최종 사용자의 사용자 디바이스에 결합된 의미 메타데이터 데이터베이스와의 상호작용에 기초하여 상기 적어도 하나의 관심 프로파일 세그먼트에 대응하는 의미 용어들을 획득하는 단계로서, 상기 의미 용어들은 상기 의미 메타데이터 데이터베이스에서 제공된 하나 이상의 의미 용어들 중에서 획득되는, 상기 획득하는 단계;

상기 적어도 하나의 관심 프로파일 세그먼트의 각각을 적어도 하나의 의미 표현으로 변환하는 단계; 및

상기 적어도 하나의 의미 표현에 기초하여 상기 적어도 하나의 관심 프로파일 세그먼트의 각각에 클러스터 식별자를 할당하는 단계로서, 상기 클러스터 식별자들은 지역에 맞는 해싱(LSH) 기술을 사용하여 생성되는, 상기 할당하는 단계를 포함하는, 컴퓨터-판독가능한 매체.

**발명의 설명**

**기술 분야**

[0001] 본 주제는 개인화된 서비스들 및 데이터 분석들에 관한 것이며, 특히 전적으로는 아니지만 개인화된 서비스들 및 데이터 분석들을 위한 사용자 관심 프로파일들을 클러스터링(cluster ing)하는 동안 최종 사용자들의 프라이버시 보호에 관한 것이다.

**배경 기술**

[0002] 월드 와이드 웹을 통해 이용가능한 엄청나게 많은 콘텐츠 덕분에, 서비스 제공자들에 의해 제공된 콘텐츠에 액세스하는 최종 사용자들은 종종 관련 콘텐츠에 쉽게 액세스하기 위해 서비스 제공자들, 탐색 엔진들, 웹 퍼블리셔들, 및 광고 회사들에 의한 개인화된 보조 및 데이터 분석들을 제공받는다. 종래에, 콘텐츠 기반 추천, 협업 추천, 및 데이터 분석들과 같은, 다양한 기술들이 최종 사용자들에게 개인화된 서비스들을 제공하기 위해 사용된다. 콘텐츠 기반 추천에서, 최종 사용자들은 콘텐츠, 과거에 최종 사용자들에 의해 사용되거나 선호된 콘텐츠 서비스들 또는 제품들과 유사한 또는 최종 사용자의 관심 또는 선택과 일치하는 콘텐츠, 서비스들 또는 제품들을 추천받는다. 협업 추천에서, 최종 사용자는 유사하거나 동일한 관심 또는 선택들을 가진 다른 최종 사용자들에 의해 사용되거나 선호된 콘텐츠, 서비스들 또는 제품들과 유사한 콘텐츠, 서비스들 또는 제품들을 추천받는다.

[0003] 콘텐츠 기반 추천의 예에서, 영화 리뷰 웹사이트는 특정한 카테고리의 영화들, 예를 들면, 애니메이션 영화들을 정기적으로 시청하기 위해 최종 사용자를 모니터링할 수 있다. 그에 따라, 애니메이션 영화가 시청을 위해 이용 가능할 때마다, 최종 사용자는 예를 들면, 관련 결체들을 함으로써 영화를 다운로드하도록, 통보 또는 통지와 같은 추천을 제공받을 수 있다. 유사하게, 또한 협업 필터링으로서 알려진, 협업 추천에서, 서비스 제공자들은 타겟팅된 광고들을 최종 사용자에게 제공할 수 있으며 이들 광고들은 최종 사용자와 유사한 관심들 및 선호들을 갖는 다른 최종 사용자들에 의해 선호되어 온 제품 또는 서비스들과 관련된다. 예를 들면, 인터넷 프로토콜 텔레비전(IPTV) 서비스 제공자는, 텔레비전 쇼들 또는 영화들이 그 관심들이 최종 사용자들의 관심들과 일치하는

다른 최종 사용자들에 의해 시청되어 왔다면, 최종 사용자에게 텔레비전 쇼들 또는 영화들을 추천할 수 있다.

[0004]

뿐만 아니라, 사용자들은 데이터 분석 애플리케이션들을 실행하기 위해, 예를 들면, 데이터 마이닝 목적들을 위해 서비스 제공자들에게 그것들의 인적 데이터를 제공하도록 요구한다. 데이터 분석 및 서비스 제공자들에게 의해 최종 사용자들에게 관련 콘텐츠를 제공하기 위한 모든 이들 기술들의 가운데에서, 오늘날의 사용자들은 그들의 개인적 및 잠재적으로 민감한 정보를 점차 우려하게 된다. 이것은 주로 개인화된 서비스들로부터 이득을 얻으며 데이터 분석들을 용이하게 하기 위해서, 최종 사용자들이 민감 정보를 누설해야 하지만, 동시에, 그들은 정보의 프라이버시 보호를 우려하기 때문이다. 예를 들면, 소셜 네트워킹 사이트의 최종 사용자는 이와 같이 소셜 네트워킹 사이트들 상에서의 업데이트들에 관해 스스로에 및 다른 최종 사용자들에게 익명 추천들을 하기 위해 그의 액세스하는 소셜 네트워킹 사이트들의 정보의 사용에 반대하지 않을 수 있지만, 최종 사용자는 다른 최종 사용자들, 서비스 제공자들, 공격자들 및 악의적인 당사자들과 같은, 다른 엔티티들이 최종 사용자가 방문하거나 평가한 특정한 URL들을 아는 것을 원하지 않을 수 있다.

**발명의 내용**

**해결하려는 과제**

[0005]

본 요약은 개인화된 서비스들을 제공하며 데이터 분석 애플리케이션들을 실행하기 위해 최종 사용자들을 클러스터링하는 동안 프라이버시 보호에 관련된 개념들을 도입하기 위해 제공된다. 본 요약은 청구된 주제의 필수적인 특징들을 식별하기 위해 의도되지 않으며 그것은 청구된 주제의 범위를 결정하거나 제한할 때 사용하기 위해 의도되지도 않는다.

**과제의 해결 수단**

[0006]

하나의 구현에서, 사용자 관심 프로파일들의 프라이버시 보호된 클러스터링을 위한 방법이 설명된다. 상기 방법은 최종 사용자의 관심 프로파일에 기초하여 적어도 하나의 관심 프로파일 세그먼트를 생성하는 단계를 포함한다. 상기 방법은 상기 최종 사용자의 사용자 디바이스에 결합된 의미 메타데이터 데이터베이스와의 상호작용에 기초하여 상기 적어도 하나의 관심 프로파일 세그먼트에 대응하는 의미 용어들을 획득하는 단계를 추가로 포함하고, 상기 의미 용어들은 상기 의미 메타데이터 데이터베이스에서 제공된 하나 이상의 의미 용어들 중에서 획득된다. 상기 적어도 하나의 관심 프로파일 세그먼트의 각각은 그 뒤에 적어도 하나의 의미 표현으로 변환된다. 상기 방법은 상기 적어도 하나의 의미 표현에 기초하여 상기 적어도 하나의 관심 프로파일에 클러스터 식별자를 할당하는 단계를 추가로 포함하고, 상기 클러스터 식별자들은 지역에 맞는 해싱(locality sensitive hashing; LSH) 기술을 사용하여 생성된다.

[0007]

또 다른 구현에서, 사용자 관심 프로파일들의 프라이버시 보호된 클러스터링을 위한 사용자 디바이스가 설명된다. 상기 구현에서, 상기 사용자 디바이스는 프로세서 및 상기 프로세서에 결합된 메모리를 포함한다. 상기 메모리는 상기 사용자 디바이스에 결합된 의미 메타데이터 데이터베이스와의 상호작용에 기초하여 적어도 하나의 관심 프로파일 세그먼트에 대응하는 의미 용어들을 획득하도록 구성된 프로파일 변환 모듈을 포함한다. 하나의 실시예에서, 상기 의미 용어들은 상기 의미 메타데이터 데이터베이스에서 제공된 하나 이상의 의미 용어들 중에서 획득된다. 상기 프로파일 변환 모듈은 상기 적어도 하나의 관심 프로파일 세그먼트의 각각을 적어도 하나의 의미 표현으로 변환하도록 추가로 구성된다. 상기 메모리는 상기 적어도 하나의 의미 표현에 기초하여, 지역에 맞는 해싱(LSH)을 사용하여 클러스터 식별자를 생성하며 상기 적어도 하나의 의미 표현에 기초하여 상기 클러스터 식별자를 적어도 하나의 관심 프로파일 세그먼트에 할당하도록 구성된 클러스터 식별자 모듈을 추가로 포함한다.

[0008]

또 다른 구현에서, 방법을 실행하기 위한 컴퓨터 판독가능한 프로그램 코드가 구현된 컴퓨터-판독가능한 매체가 설명된다. 상기 방법은 최종 사용자의 관심 프로파일에 기초하여 적어도 하나의 관심 프로파일 세그먼트를 생성하는 단계를 포함할 수 있다. 상기 방법은 상기 최종 사용자의 사용자 디바이스에 결합된 의미 메타데이터 데이터베이스와의 상호작용에 기초하여 상기 적어도 하나의 관심 프로파일 세그먼트에 대응하는 의미 용어들을 획득하는 단계를 추가로 포함하며 상기 의미 용어들은 상기 의미 메타데이터 데이터베이스에서 제공된 하나 이상의 의미 용어들 중에서 획득된다. 상기 적어도 하나의 관심 프로파일 세그먼트의 각각은 그 뒤에 적어도 하나의 의미 표현으로 변환된다. 상기 방법은 상기 적어도 하나의 의미 표현에 기초하여 상기 적어도 하나의 관심 프로파일 세그먼트에 클러스터 식별자를 할당하는 단계를 추가로 포함하며 상기 클러스터 식별자들은 지역에 맞는 해싱(LSH) 기술을 사용하여 생성된다.

[0009] 상세한 설명은 첨부한 도면들을 참조하여 설명된다. 도면들에서, 참조 번호의 가장 왼쪽 숫자(들)는 참조 부호가 먼저 나타나는 도면을 식별한다. 동일한 숫자들은 유사한 특징들 및 구성요소들을 참조하기 위해 도면들 전체에 걸쳐 사용된다. 본 주제의 실시예들에 따른 시스템 및/또는 방법들의 몇몇 실시예들은 이제 단지 예로서, 및 첨부한 도면들을 참조하여 설명된다.

**도면의 간단한 설명**

[0010] 도 1은 본 주제의 일 실시예에 따른, 개인 정보의 프라이버시 보호를 위해 사용자 프로파일들의 프라이버시 보호된 클러스터링의 일 대표적인 네트워크 환경 구현을 도시한 도면.

도 2는 본 주제의 일 실시예에 따른, 개인 정보의 프라이버시 보호를 위해 사용자 프로파일들의 프라이버시 보호된 클러스터링을 구현하는 사용자 디바이스의 구성요소들을 도시한 도면.

도 3a, 도 3b, 및 도 3c은 본 주제의 일 실시예에 따른, 사용자 프로파일들의 프라이버시 보호된 클러스터링을 구현하는 다양한 애플리케이션들의 대표적인 네트워크 환경 구현들을 도시하는 도면들.

도 4는 본 주제의 일 실시예에 따른, 최종 사용자들에게 프라이버시 보호를 제공하기 위해 사용자 프로파일들의 프라이버시 보호된 클러스터링의 방법을 도시한 도면.

**발명을 실시하기 위한 구체적인 내용**

[0011] 본 문서에서, 단어("대표적인")는 여기에서 "예, 인스턴스, 또는 예시로서 작용하는"을 의미하기 위해 사용된다. "대표적인"과 같은 여기에 설명된 본 주제의 임의의 일 실시예 또는 구현은 반드시 다른 실시예들에 비해 선호되거나 유리한 것으로서 해석되지는 않는다.

[0012] 여기에서의 임의의 블록 다이어그램들은 본 주제의 원리들을 구현하는 예시적인 시스템들의 개념적인 뷰들을 표현한다는 것이 당업자들에 의해 이해되어야 한다. 유사하게, 컴퓨터 또는 프로세서가 명시적으로 도시되는지의 여부에 관계없이, 임의의 플로우 차트들, 흐름도들, 상태 전이도들, 의사 코드 등이 컴퓨터 판독가능한 매체에서 실질적으로 표현될 수 있으며, 따라서 이러한 컴퓨터 또는 프로세서에 의해 실행될 수 있는 다양한 프로세스들을 나타낸다는 것이 이해될 것이다.

[0013] 최종 사용자들의 개인 정보에 프라이버시를 제공하기 위한 시스템들 및 방법들이 설명된다. 시스템들 및 방법들은 서비스 제공자에 의해 제공된 콘텐츠를 시청하거나 서비스들을 이용하기 위해 최종 사용자들에 의해 사용된 다양한 사용자 디바이스들에 구현될 수 있다. 뿐만 아니라, 사용자 디바이스들은 다양한 네트워크들을 통해 분산된 서버들 및 다른 사용자 디바이스들과 통신하고 있을 수 있다. 설명된 방법(들)을 구현할 수 있는 사용자 디바이스들 및 분산된 서버들은 이에 제한되지 않지만, 랩탑, 데스크탑 컴퓨터, 노트북, 모바일 전화, 개인 휴대용 정보 단말기, 워크스테이션, 메인프레임 컴퓨터, 셋 탑 박스, 미디어 플레이어, 중앙 디렉토리 서버들, 데이터베이스 서버, 파일 서버, 프린트 서버, 웹 서버, 애플리케이션 서버 등과 같은, 컴퓨팅 디바이스들을 포함한다. 여기에서의 설명은 추천 서비스들을 제공하며 데이터 분석 애플리케이션들을 실행하는 개인 컴퓨터들 및 분산 서버들을 참조하지만, 방법들 및 시스템들은 당업자에 의해 이해될 바와 같이, 그들의 프라이버시를 보호하기 위해 최종 사용자들에게 연결을 제공하는 다른 서버 및 컴퓨팅 시스템들에서 구현될 수 있다.

[0014] 추천 서비스들 및 개인화된 콘텐츠는 다양한 네트워크들을 통해 서비스 제공자들에 의해 상이한 최종 사용자들에게 제공되지만, 여기에 설명된 방법들 및 시스템들은 액세스 독립적이며, 모바일을 위한 전역적 시스템(GSM), 광대역 코드 분할 다중 액세스(W-CDMA), 코드 분할 다중 액세스(CDMA), 무선 근거리 네트워크(WLAN), 와이어라인, 근거리 네트워크(LAN), 및 연결을 제공하는 다른 네트워크들을 포함하는 다수의 액세스 유형들을 지원한다.

[0015] 종래에, 많은 서비스 제공자들은 최종 사용자들의 개인 세부사항들, 선호들, 및 선택들에 기초하여, 비디오들, 오디오, 뉴스, 광고들과 같은, 콘텐츠를 제공하는 서비스들과 같이, 최종 사용자들에게 제공된 서비스들을 개인화하려고 시도한다. 이를 위해, 서비스 제공자들은 최종 사용자의 나이, 성별, 및 위치와 같은 개인 정보와 같은 인자들에 기초하여; 최종 사용자의 과거 동작들, 또는 최종 사용자의 것과 유사한 관심들을 갖는 것으로 식별되어 온 다른 최종 사용자들에 의한 과거 선호들에 기초하여 서비스들, 콘텐츠들, 또는 제품들을 추천하며 데이터 분석 애플리케이션들을 실행하기 위해 다양한 기술들을 사용한다.

[0016] 예를 들면, 최종 사용자, 즉 사용자(A)가 특정한 저자에 의해 쓰여진 책을 구매하였다면, 서비스 제공자는 동일한 저자에 의해 쓰여진 다른 책들 또는 동일하거나 관련된 주제들에 대한 다른 책들 등을 구매하도록 사용자(A)에게 제안할 수 있다. 유사하지만 또 다른 종래의 접근법에서, 서비스 제공자는 최종 사용자(A)의 것과 유사

한 관심 프로파일을 가진 다른 최종 사용자들에 의해 선호되어 온 이들 콘텐츠들, 또는 제품들을 최종 사용자(A)에게 추천할 수 있다. 개인화된 서비스들을 제공하기 위해 및 데이터 분석을 위해, 다양한 종래에 알려진 방법들이 하나 이상의 관심 그룹들로 유사한 관심들을 가진 최종 사용자들을 클러스터링하기 위해 사용된다.

[0017] 서비스 제공자에 의해 구현된 클러스터링의 종래의 방법들은 최종 사용자의 개인 선호들, 선택들 등에 관련된 정보의 수집을 요구한다. 종래에, 서비스 제공자들은 로그 파일들, 애플리케이션 이력 파일들 또는 최종 사용자의 클라이언트 디바이스 상에 저장되거나 서비스 제공자에 등록하면서 사용자에게 의해 제공된 다른 개인적으로 식별가능한 정보를 분석함으로써와 같은, 다양한 수단들을 통해 최종 사용자들과 관련되는 정보를 모니터링 및 수집한다. 또 다른 종래의 기술에서, 서비스 제공자는 최종 사용자와 관련되는 정보를 수집하기 위해 하이퍼 텍스트 전송 프로토콜(http) 쿠키와 같은, 텍스트 파일을 저장할 수 있다. 예를 들면, 웹 포털은 폰트 크기, 디스플레이 위젯들의 배열 등과 같은, 최종 사용자의 선호들을 저장하기 위해 최종 사용자의 웹 브라우저의 http 쿠키를 포함하는 사용자 선호 데이터를 저장할 수 있다. 뿐만 아니라, 사용자 선호 데이터는 또한 최종 사용자의 브라우징 세부사항들을 저장하며 이를 웹 포털로 전송할 수 있다.

[0018] 따라서, 최종 사용자의 개인 선택에 기초하여 최종 사용자에게 추천 서비스들, 개인화된 콘텐츠, 서비스들 또는 제품들을 제공하려는 시도로, 서비스 제공자는 종종 최종 사용자의 활동들에 관련된 정보를 모니터링 및 수집한다. 특정한 상황들에서, 서비스 제공자에 의해 수집된 정보에 기초하여 최종 사용자를 식별하는 것이 가능하게 된다. 이것은 최종 사용자의 개인적 또는 비밀 정보를 손상시키는 것을 야기하며 최종 사용자들을 잠재적인 프라이버시 침해들에 노출시키거나 그를 광고주들 또는 스파머들 등의 타겟으로 만들 수 있다. 뿐만 아니라, 극단적인 경우들에서, 최종 사용자는 아이덴티티 절도 및 신용 카드 사기들과 같은, 다양한 범죄들의 희생자가 될 수 있다.

[0019] 뿐만 아니라, 오늘날 최종 사용자들이 고속 통신 지원을 경유하여 가상 소셜 네트워크들을 통해 서로에 연결될 때, 최종 사용자들은 스템블어폰(StumbleUpon)<sup>®</sup>, 구글뉴스(GoogleNews)<sup>™</sup>, 포스퀘어(Foursquare)<sup>®</sup>, 페이스북(Facebook)<sup>®</sup>, 야후!(Yahoo!)<sup>®</sup>, 및 트위터(twitter)<sup>™</sup>와 같은, 광범위한 클래스의 애플리케이션들에 사설 및 개인 정보를 끊임없이 공유하고 있다. 누설된 민감한 개인 정보의 특징은 위치 기반 서비스들(LBS)에 대한 위치 및 고유 리소스 로케이터(URL) 추천기 시스템들에 대한 브라우징 이력과 같은, 애플리케이션의 유형에 의존하지만, 온라인 소셜 네트워크들(OSN들)의 모두를 아우르는 프레임워크로의 이들 애플리케이션들의 주입은 이전보다 더 높은 위험들에 사용자들의 개인 정보를 두며, 여기에서 OSN들은 추천 서비스들을 제공하기 위한 '원-스톱-숍(one-stop-shop)'으로서 작용하기 시작한다.

[0020] 최근에, 추천 서비스들, 개인화된 콘텐츠, 개인화된 탐색/질의, 소셜 네트워크들, 데이터 마이닝, 및 다른 관련된 서비스들과 같은, 여전히 데이터 분석 애플리케이션들을 실행하며 개인화된 서비스들을 제공하면서, 프라이버시를 보존하기 위한 상이한 접근법들이 제안되어 왔다.

[0021] 이러한 기존의 접근법 중 하나는 준동 암호화 기반 전략들을 사용하여 관심 프로파일들을 클러스터링하는 것을 수반한다. 이러한 접근법에서, 사용자들의 관심 프로파일들은 사용자들에게 프라이버시를 제공하기 위해, 준동 암호화와 같은 암호 기술들을 사용하여 암호화된다. 이러한 암호 기술들은 분산 계산 설정을 사용하여 관심 프로파일들을 클러스터링하기 위한 덧셈 및 곱셈과 같은, 프로토콜 프리미티브 연산들의 실행을 가능하게 한다. 분산 계산 설정을 사용하는 것은 정적 세트의 최종 사용자들의 사용자 디바이스들을 통해 다수의 단계들의 프로토콜 프리미티브 연산들을 실행하는 것을 허용한다. 예를 들면, KMeans 클러스터링 알고리즘은 준동 암호화의 기술을 사용한 프라이버시 보존 방법에서 구현될 수 있다. 준동 암호화를 사용한 기술들은 최종 사용자들에게 프라이버시를 제공하는 것을 용이하게 하지만, 이러한 기술들은 보통 리소스 집중적이며 실행하기에 복잡하다. 뿐만 아니라, 정적 세트의 최종 사용자들의 사용자 디바이스들을 통해 다수의 단계들의 프로토콜 프리미티브 연산들을 실행하는 것은 최종 사용자들의 세트가 온라인 포털들에 끊임없이 가입하고 떠나는 최종 사용자들에 따라 동적으로 변화하는 온라인 포털들과 같은, 동적 환경들에서 이러한 기술들의 적용을 제한한다.

[0022] 뿐만 아니라, 특정한 접근법들은 관심 프로파일들을 클러스터링하도록 구성된 집중화된 중재자들이 최종 사용자들에 의해 신뢰되지 않는 집중화된 해결책을 제공한다. 이러한 상황들에서, 각각의 최종 사용자는 처음에 집중화된 중재자에게 전달되기 전에 투영된 관심 프로파일을 획득하기 위해 암호화와 같은, 랜덤 투영 기반 접근법을 사용하여 관심 프로파일을 사전-프로세싱한다. 투영된 관심 프로파일은 통상적으로 집중화된 중재자들을 갖고 이용가능하지 않은 랜덤 투영 함수들을 사용하여 관심 프로파일들을 변환함으로써 획득되고, 따라서 최종 사용자의 데이터의 프라이버시를 보장한다. 랜덤 투영 함수들은 그러나, 모든 유사한 관심 프로파일들이 효율적이며 정확한 클러스터링을 용이하게 하기 위해 유사한 투영된 관심 프로파일로 변환됨을 보장하기 위해 공통 클러

스터의 일부이도록 의도하는 최종 사용자들 중에서 공유된다. 그러나, 최종 사용자들 중에서 랜덤 투영 함수들을 공유하는 것은 악의적인 최종 사용자들이 투영된 관심 프로파일들로부터 원래 관심 프로파일들을 획득하기 위해 집중화된 중재자들에게 랜덤 투영 함수를 제공할 수 있기 때문에 랜덤 투영 함수의 누설을 가능하게 할 수 있다.

[0023] 사실 정보를 보호하기 위한 또 다른 기술은 최종 사용자를 대신하여 관심 프로파일을 송신하도록 구성된 프록시 노드를 가진 원조에 대한 가십(gossip-on-behalf) 접근법의 사용이다. 이러한 기술은 프록시 노드로부터 최종 사용자들 아이덴티티를 은닉하기 위해 암호화된 2-홉 통신을 사용하여, 중재자 노드를 통해, 최종 사용자들의 사용자 프로파일을 프록시 노드들로 송신하는 것을 수반한다. 그러나, 프록시 노드와 완전한 관심 프로파일을 공유하는 것은 프록시 노드들이 시간 기간에 걸쳐 관심 프로파일들을 분석함으로써 최종 사용자들을 식별할 수 있기 때문에 연계성 공격들을 용이하게 할 수 있다.

[0024] 본 주제의 일 구현에 따르면, 최종 사용자들의 개인 정보에 프라이버시를 제공하기 위해 최종 사용자들의 사용자 관심 프로파일들을 클러스터링하면서 프라이버시를 보장하기 위한 시스템들 및 방법들이 설명된다. 이전에 설명된 바와 같이, 시스템들 및 방법들은 통신을 위해 정의된 다양한 상이한 표준들에 따라 네트워크와 통신할 수 있는 다양한 프로세싱 및 통신 디바이스들에 구현될 수 있다. 뿐만 아니라, 여기에 설명된 시스템들 및 방법들은 상이한 수단들을 통해 제공된 유선 네트워크들 또는 무선 네트워크들을 통해 연결될 수 있다.

[0025] 한편으로, 여기에 설명된 바와 같은 시스템들 및 방법들은 애플리케이션들의 사용을 제한하지 않고 그들의 사실 정보 및 각각의 최종 사용자에게 익명성을 제공함으로써 최종 사용자들의 사실 정보에 대한 진정한 보호를 제공하고, 다른 한편으로, 또한 개인화된 서비스들을 수신하기 위해 또는 그들의 사실 데이터의 유사성에 따라 사용자 클러스터링을 요구하는 임의의 다른 애플리케이션들 또는 서비스들을 사용하기 위해 제한된 계산 능력을 가진 모바일 디바이스와 같은, 썬(thin) 클라이언트를 포함하는 임의의 컴퓨팅 디바이스를 이용하는 능력을 최종 사용자에게 제공한다. 하나의 실시예에서, 프라이버시 보호된 네트워크는 하나 이상의 집성기(aggregator) 노드들과 같은, 분산 서버들 또는 서로와 통신하기 위한 사용자 디바이스들과 연관된 복수의 최종 사용자들을 포함한다.

[0026] 본 주제의 일 실시예에 따르면, 최종 사용자에게 개인화된 콘텐츠 및 추천 서비스들을 제공하기 위해, 유사한 관심 프로파일들을 가진 최종 사용자들의 클러스터들이 생성된다. 이를 위해, 최종 사용자들 각각의 사용자 디바이스들은 처음에 최종 사용자의 개인 선택들 및 선호들과 같은, 프로파일 데이터에 기초하여 최종 사용자의 관심 프로파일을 생성한다. 예를 들면, 사용자 디바이스는 최종 사용자에게 의해 방문된 모든 URL들의 리스트, 최종 사용자를 위해 탐색되거나 그에 의해 보여진 모든 비디오들, 최종 사용자에게 청취되거나 그에 의해 탐색된 음악, 및 최종 사용자들에 의해 탐색된 다른 유사한 관심 필드들에 관련된 데이터를 획득할 수 있다.

[0027] 하나의 실시예에서, 관심 프로파일들은 그 다음에 각각의 세그먼트가 특정한 관심 필드에 대응하는 다수의 관심 프로파일 세그먼트들로 분할될 수 있다. 관심 필드들의 예들은, 이에 제한되지 않지만, 전문적인 관심들, 여가 관심들, 및 종교적 선호들과 같은 맥락과 관련된 선호들 및 음악, 댄스, e-쇼핑, 및 소셜 네트워킹과 같은 관심 영역들을 포함한다. 관심 프로파일 세그먼트들은 그 다음에 다양한 관심 그룹들로 클러스터링하기 위해 프로세싱될 수 있다. 또 다른 실시예에서, 전체로서 관심 프로파일은 다양한 관심 그룹들로 클러스터링하기 위해 프로세싱될 수 있다.

[0028] 관심 프로파일 세그먼트들은 그 다음에 의미 표현들로서 불리우는 작으며 의미적으로 난해한 표현들로 변환된다. 관심 프로파일 세그먼트들의 의미 표현들은, 제한으로서가 아닌 예로서, 태그 기반 표현들, 온톨로지 개념 기반 표현들, 및 토픽 기반 표현들을 포함한다. 예를 들면, 소셜 네트워킹 사이트("xyz.com")의 URL에 대응하는 관심 프로파일은 의미 표현("소셜 네트워킹 사이트")으로 변환될 수 있다.

[0029] 상기 목적을 위해, 사용자 디바이스는, 익명 통신 채널을 통해, 관심 프로파일 세그먼트들을 대응하는 의미 표현들로 변환하기 위해 사용될 수 있는 의미 용어들의 리스트 또는 사전을 가진 의미 메타데이터 데이터베이스와 상호작용할 수 있다. 하나의 실시예에서, 사용자 디바이스는 관심 프로파일 세그먼트들을 대응하는 의미 표현들로 변환하기 위해 관심 프로파일 세그먼트들 각각에 대응하는 의미 용어를 식별할 수 있는 관심 프로파일 세그먼트들을 의미 메타데이터 데이터베이스에 제공할 수 있다. 또 다른 실시예에서, 사용자 디바이스는 의미 메타데이터 데이터베이스로부터 의미 용어들의 리스트를 획득하며 그 후 대응하는 의미 표현들을 획득하기 위해 관심 프로파일 세그먼트와 일치하는 의미 용어를 식별할 수 있다. 또 다른 실시예에서, 사용자 디바이스 자체는 최종 사용자에게 의해 브라우징된 웹 페이지들로부터 추출된 의미 용어들을 가진 메타데이터를 포함할 수 있다. 상기 실시예에서, 사용자 디바이스들은 상이한 사용자 디바이스들에 의해 생성된 의미 표현들이 동일한 용어들

을 사용하여 표현되며 그러므로 비교가능하도록 서로와 의미 용어들의 어휘를 공유하도록 구성된다.

[0030] 이렇게 하여 획득된 의미 표현들은 최종 사용자가 연관될 수 있는 관심 그룹들을 식별하기 위해 클러스터 식별자를 할당받을 수 있다. 클러스터 식별자들은 지역에 맞는 해싱(LSH)의 기술을 사용하여 할당받을 수 있다. LSH 기술은 사용자 디바이스의 각각에서 발생된 랜덤 벡터들의 공통 시퀀스와 같은, 랜덤 값들에 의해 정의된 해시 함수들의 세트 및 의미 표현들을 사용하여, 의미 표현들 각각을 대응하는 해시 코드들, 즉 클러스터 식별자들로 변환하는 것을 수반한다. 사용자 디바이스에서 클러스터 식별자들을 할당하는 것은 관심 프로파일들이 원격 노드들과 공유되지 않기 때문에 프라이버시 보호를 제공할 때 용이하게 한다. 뿐만 아니라, LSH 기술을 사용하는 것은 클러스터링이 이제 해시 코드들을 매칭시킴으로써 쉽게 수행될 수 있기 때문에 사용자 프로파일들의 클러스터링에서 효율성을 증가시키는 것을 돕는다. 더욱이, 의미 메타데이터 데이터베이스를 사용하는 것은 위음성(false negative)들 및 위양성(false positive)들의 발생을 감소시키며 따라서 클러스터링의 효율성을 증가시키는 것을 돕는다. 따라서, 의미 표현을 생성함으로써 프로파일 변환과 함께 LSH 기술에 의해 클러스터링을 수행하는 것은 클러스터링시 감소된 에러들을 갖고 프라이버시 보호를 제공할 때 용이하게 한다.

[0031] 이렇게 하여 획득된 클러스터 식별자들은 사용자 디바이스를 하나 이상의 클러스터들, 즉 클러스터 식별자들에 대응하는 관심 그룹들로 클러스터링하기 위해 사용될 수 있다. 뿐만 아니라, 클러스터 식별자들은 통신 채널, 예를 들면 익명 통신 채널을 통해 하나 이상의 원격 노드들, 예를 들면, 중앙 집성기, 관심 그룹 집성기, 및 또 다른 사용자 디바이스에 제공된다.

[0032] 하나의 구현에서, 사용자 디바이스와 통신하는 원격 노드는 사용자 디바이스가 사용될 수 있는 추천 기반 서비스의 유형 또는 애플리케이션에 의존하여 선택될 수 있다. 예를 들면, 협업 필터링 애플리케이션의 경우에, 원격 노드들은 영화들, 댄스들, e-쇼핑들, 소셜 네트워킹 사이트들, 또는 그것의 조합과 같은, 적어도 하나의 클러스터, 즉 관심 그룹에 대응하는 각각의 중앙 집성기를 가진 중앙 집성기들일 수 있다. 사용자 디바이스는 따라서 중앙 집성기로 하여금 최고 랭킹된, 즉 그것들의 대응하는 관심 그룹들에서의 관심 그룹의 관심 필드 후 최상위 랭킹된 관심 필드를 식별할 수 있게 하기 위해 사용자의 관심 프로파일들과 함께 대응하는 중앙 집성기에 클러스터 식별자들 각각을 제공할 수 있다. 예를 들면, '음악'에 대한 관심 그룹을 책임지는 중앙 집성기는 대응하는 클러스터 식별자를 가진 사용자 디바이스들 모두의 관심 프로파일들을 수신할 수 있다. 중앙 집성기는 그 다음에 예를 들면, 대다수, 즉 90% 최종 사용자들의 관심 프로파일들에 존재하기 때문에 다른 관심 필드들 또는 높게 랭킹된 관심들을 식별하기 위해 사용자 프로파일들을 분석할 수 있다. 예를 들면, 중앙 집성기는 최고 랭킹된 관심으로서 '댄스'를 식별할 수 있으며, 따라서 관심 필드('댄스')를 책임지는 중앙 집성기로부터 댄스에 대응하는 추천들을 획득할 수 있다. 중앙 집성기는 따라서 클러스터('음악')에 대응하는 클러스터에 존재하는 사용자 디바이스들 모두에 댄스 뿐만 아니라, 음악과 관련된 콘텐츠들 및 추천들을 제공할 수 있다. 중앙 집성기는 따라서 사용자의 관심 뿐만 아니라, 또한 관심 그룹에서 높게 랭킹된 새로운 관심 또는 아이템과 관련된 콘텐츠들 및 추천들을 제공할 수 있다. 또 다른 실시예에서, 중앙 집성기는 관심 그룹을 책임지는 중앙 집성기로서 동작하는 각각의 사용자 디바이스로 교체될 수 있다.

[0033] 유사하게, 콘텐츠-기반 추천 애플리케이션들의 경우에, 원격 노드들은 각각의 관심 그룹 집성기가 영화들, 댄스들, e-쇼핑들, 소셜 네트워킹 사이트들과 같은, 적어도 하나의 클러스터, 즉 관심 그룹에 대응하는, 관심 그룹 집성기들일 수 있다. 이러한 경우에 사용자 디바이스들은 어떤 관심 그룹 집성기가 클러스터 식별자에 대응하는 콘텐츠를 제공할 수 있는지에 기초하여, 클러스터 식별자에 대응하는 관심 그룹 집성기에 클러스터 식별자 및 의미 표현을 송신하도록 구성된다.

[0034] 뿐만 아니라, 피어-투-피어 통신들을 수반한 애플리케이션들에서, 원격 노드들은 다른 사용자 디바이스들일 수 있다. 이러한 애플리케이션들에서, 사용자 디바이스들은 익명 통신 채널들을 사용하여 사용자 디바이스들 사이에서 데이터의 교환 및 채팅을 가능하게 하도록 유사한 클러스터 식별자들을 가진 다른 사용자 디바이스들을 식별하기 위해 클러스터 식별자들을 사용할 수 있다. 익명 통신 채널들을 사용하는 것은 다른 사용자들 또는 원격 노드들이, 클러스터 식별자들 및 다른 데이터가 익명 통신 채널들을 형성하는 다양한 중재자 노드들을 사용하여 전송되기 때문에 클러스터 식별자들을 송신하는 최종 사용자를 식별할 수 없다는 것을 보장하는 것을 돕는다.

[0035] 따라서 본 주제의 시스템들 및 방법들은 사용자 디바이스들에서 관심 프로파일들에 클러스터 식별자들을 할당함으로써 최종 사용자들에게 프라이버시를 제공할 때 용이하게 한다. LSH 기술을 사용하여 사용자 디바이스들에서 관심 프로파일들에 클러스터 식별자들을 할당하는 것은 원격 노드들이 최종 사용자들의 관심 프로파일들에 액세스하지 않고 원격 노드들에서 유사한 사용자들의 클러스터링을 가능하게 하고, 따라서 최종 사용자들에게 프라이버시 보호를 제공한다. 뿐만 아니라, 의미 표현들과 함께 LSH 기술을 사용하는 것은 의미 메타데이터 데이터

베이스를 사용하는 것이 위음성들 및 위양성들의 발생들을 감소시키는데 도움을 주기 때문에 프라이버시가 보호되며 클러스터링시 에러들이 감소됨을 보장하도록 도우며, 따라서 클러스터링시 효율성을 증가시킨다. 따라서, 사용자 디바이스들에서 클러스터 식별자들을 제공하는 것은 서비스들이 이제 관심 프로파일들을 분석하도록 요구하지 않기 때문에 데이터 분석을 수행하며 보다 빠르고 효율적으로 개인화된 서비스들을 제공하는데 콘텐츠 제공자들을 용이하게 하고 콘텐츠 데이터와 클러스터 식별자를 매칭시키기 위해 임의의 콘텐츠 매칭 기술들을 사용할 수 있다. 부가적으로, 클러스터 식별자들을 사용하는 것은 서비스 제공자들이 이제 관심 프로파일들을 프로세싱하거나 분석하도록 요구하지 않기 때문에 서비스 제공자의 단부에서 비용을 감소시키며 리소스들을 절약한다.

[0036] 설명 및 도면들은 단지 본 주제의 원리들을 도시한다는 것이 주의되어야 한다. 따라서 당업자들은 여기에 명시적으로 설명되거나 도시되지 않지만, 본 주제의 원리들을 구현하며 그것의 사상 및 범위 내에 포함되는 다양한 배열들을 고안할 수 있다는 이해될 것이다. 더욱이, 여기에 나열된 모든 예들은 원칙적으로 이 기술을 발전시키기 위해 발명자(들)에 의해 기여된 개념들 및 본 주제의 원리들을 이해하도록 관독자를 돕기 위해 단지 교육적인 목적들을 위한 것으로 명확하게 의도되고, 이러한 구체적으로 나열된 예들 및 조건들에 대한 제한이 없는 것으로 해석될 것이다. 게다가, 본 주제의 원리들, 양태들, 및 실시예들을 여기에 나열한 모든 문장들, 뿐만 아니라 그것의 특정 예들은 그것의 등가물들을 포함하도록 의도된다.

[0037] 여기에 사용된 바와 같은 단어들(동안, ~하는 동안, 및 할 때)은 개시 동작시 즉시 동작이 발생함을 의미하지만 초기 동작 및 상기 초기 동작에 의해 개시되는 반응 사이에, 전파 지연과 같은, 몇몇 작지만 적정한 지연이 있을 수 있는 정확한 용어들은 아니라는 것이 당업자들에 의해 또한 이해될 것이다. 부가적으로, 단어("연결된" 및 "결합된")는 설명의 명료함을 위해 전체에 걸쳐 사용되며 직접 연결 또는 간접 연결을 포함할 수 있다.

[0038] 최종 사용자들에게 프라이버시 보호를 위한 사용자 프로파일들을 클러스터링하는 시스템들 및 방법들이 구현되는 방식이 도 1 내지 도 4에 대하여 상세히 설명되었다. 프라이버시 보호를 제공하기 위해 사용자 프로파일들을 클러스터링하기 위한 설명된 시스템들 및 방법들의 양태들이 임의의 수의 상이한 컴퓨팅 시스템들, 송신 환경들, 및/또는 구성들에서 구현될 수 있지만, 실시예들이 다음의 대표적인 시스템(들)의 맥락에서 설명된다.

[0039] 도 1은 본 주제의 일 실시예에 따른, 데이터 분석을 용이하게 하며 개인화된 서비스들에 액세스하면서 최종 사용자들의 프라이버시 보호를 위해 사용자 프로파일들의 프라이버시 보호된 클러스터링의 네트워크 환경(100) 구현을 도시한다. 네트워크 환경(100)은 통신 네트워크(106)를 통해 서로 및 원격 노드(104)와 통신하는, 총괄하여 사용자 디바이스들(102)로 불리우며, 이후 개별적으로 사용자 디바이스(102)로 불리우는, 하나 이상의 사용자 디바이스들(102-1, 102-2, ..., 102-N)을 포함한다.

[0040] 통신 네트워크(106)는 무선 네트워크, 또는 유선 및 무선 네트워크의 조합일 수 있다. 통신 네트워크(106)는 서로와 상호 연결되며 단일 대형 네트워크(예로서, 인터넷 또는 인트라넷)로서 기능하는, 개개의 네트워크들의 컬렉션일 수 있다. 이러한 개개의 네트워크들의 예들은, 이에 제한되지 않지만, 제 3 세대 파트너십 프로젝트(3GPP), 롱 텀 에볼루션(LTE), 모바일 통신을 위한 전역적 시스템(GSM) 네트워크, 범용 모바일 원격통신 시스템(UMTS) 네트워크, 개인 통신 서비스(PCS) 네트워크, 시간 분할 다중 액세스(TDMA) 네트워크, 코드 분할 다중 액세스(CDMA) 네트워크, 차세대 네트워크(NGN), 공중 스위칭 전화 네트워크(PSTN), 및 통합 서비스 디지털 네트워크(ISDN)를 포함한다. 통신 네트워크들(106)은 근거리 네트워크(LAN), 광역 네트워크(WAN), 인터넷, 인트라넷, 피어 투 피어 네트워크 및 가상 사설 네트워크(VPN)를 포함하는, 임의의 공중 또는 사설 네트워크일 수 있으며 라우터들, 브리지들, 서버들, 컴퓨팅 디바이스들, 저장 디바이스들 등과 같은, 다양한 네트워크 디바이스들을 포함할 수 있다.

[0041] 사용자 디바이스들(102)은 추천 서비스들; 개인화된 콘텐츠, 개인화된 탐색/질의, 소셜 네트워크들, 데이터 마이닝, 및 다른 관련된 서비스들과 같은, 데이터 분석을 용이하게 하며 다양한 개인화된 서비스들에 액세스하기 위한 원격 노드(104) 및 분산 서버들과 같은 다른 네트워크 엔티티들과 또는 서로와 통신하기 위해 최종 사용자들에 의해 사용된 사용자 장비들(UE들)로서 정의될 수 있다. 사용자 디바이스들(102)의 예들은 제한 없이, 랩탑 또는 다른 휴대용 컴퓨터들, 데스크탑 컴퓨터, 노트북, 네트워크 컴퓨터들과 같은 컴퓨팅 디바이스들; 모바일 전화; 핸드-헬드 디바이스들; 개인 휴대용 정보 단말기; 워크스테이션; 메인프레임 컴퓨터; 셋 탑 박스 미디어 플레이어; 중앙 디렉토리 서버들; 데이터베이스 서버; 파일 서버; 프린트 서버; 웹 서버; 애플리케이션 서버 등을 포함할 수 있다. 사용자 디바이스들(102) 각각은 사용자 디바이스(102)가 결합되는 통신 네트워크(106)에 의해 정의된 바와 같이 통신 프로토콜 상에서 작동한다. 뿐만 아니라, 사용자 디바이스들(102)은 이후 총괄하여 통신 채널들(108)로서 불리우는, 통신 채널들(108-1, 108-2, 108-3, 108-4, 108-5, ..., 108-n) 중 하나 이상을

사용하여 통신 네트워크(106)를 통해 원격 노드(104)와 같은, 다른 네트워크 엔티티들 또는 서로와 상호작용하도록 구성된다. 하나의 구현에서, 통신 채널들(108)은 예를 들면, 사용자 디바이스들(102) 사이에서 및 사용자 디바이스들(102)와 원격 노드(104)와 같은 다른 네트워크 엔티티들 사이에서 익명 통신을 지원하기 위해 TOR과 같은, 어니언 라우팅 프레임워크에 기초하여 구성된 익명 통신 채널들일 수 있다.

[0042] 원격 노드(104)는 랩탑 컴퓨터, 데스크탑 컴퓨터, 노트북, 워크스테이션, 메인프레임 컴퓨터, 서버 등과 같은, 다양한 컴퓨팅 디바이스들로서 구현될 수 있다. 원격 노드(104)가 엔티티로서 도시되지만, 원격 노드(104)는 또한 각각의 노드가 랩탑 컴퓨터, 데스크탑 컴퓨터, 노트북, 워크스테이션, 메인프레임 컴퓨터, 서버 등과 같은, 컴퓨팅 디바이스들로서 구현될 수 있는 네트워크를 통해 분포된 다수의 중재자 노드들을 포함하는 분산 컴퓨팅 시스템으로서 구현될 수 있다. 뿐만 아니라, 중재자 노드들은 데이터의 통신들 및 교환의 목적을 위해 중간 네트워크(도면에 도시되지 않음)를 통해 연결될 수 있다. 뿐만 아니라, 원격 노드(104)의 예들은, 이에 제한되지 않지만, 중앙 집성기, 관심 그룹 집성기, 및 사용자 디바이스(102)를 포함한다.

[0043] 하나의 구현에서, 사용자 디바이스(102)와 통신하는 원격 노드(104)는 사용자 디바이스(102)가 사용될 수 있는 추천 기반 서비스의 유형 또는 애플리케이션에 의존하여 선택될 수 있다. 예를 들면, 협업 필터링 애플리케이션의 경우에, 원격 노드(104)는 중앙 집성기일 수 있으며, 유사하게 콘텐츠-기반 추천 애플리케이션들의 경우에, 원격 노드(104)는 관심 그룹 집성기들로서 구현될 수 있다.

[0044] 하나의 실시예에서, 최종 사용자들의 관심 프로파일들은 최종 사용자들의 활동들에 기초하여 생성되며 사용자 디바이스(102)에 의해 국부적으로 저장된다. 예를 들면, 최종 사용자들의 관심 프로파일들은 최종 사용자들에 대응하는 프로파일 데이터에 기초하여 생성될 수 있다. 프로파일 데이터는 예를 들면, 최종 사용자들에 의해 방문된 웹사이트들, 최종 사용자들에 의해 플레이되거나 다운로드된 노래들 또는 비디오들, 최종 사용자들에 의해 사용된 제품들 또는 그에 의해 이용되거나 검토된 서비스들 등을 표시할 수 있다. 생성된 관심 프로파일에 기초하여, 클라이언트 디바이스는 하나 이상의 미리-정의된 관심 그룹들에서 최종 사용자를 분류한다. 관심 그룹들은 유사한 관심들 및 선택들을 공유한 최종 사용자들의 그룹들로서 이해될 수 있다.

[0045] 상기 목적을 위해, 사용자 디바이스(102)는 프로파일 변환 모듈(110) 및 클러스터 식별자 모듈(112)을 포함할 수 있다. 프로파일 변환 모듈(110)은 관심 프로파일을, 의미 표현들로서 불리우는, 작고 의미적으로 난해한 표현들로 변환하도록 구성된다. 의미 표현들은 제한이 아닌 예로서, 태그 기반 표현들, 온톨로지 개념 기반 표현들, 및 토픽 기반 표현들을 포함한다. 하나의 실시예에서, 완전한 관심 프로파일은 프로파일 변환 모듈(110)에 의해 의미 표현들로 변환될 수 있다. 또 다른 실시예에서, 이후 관심 프로파일 세그먼트들로서 불리우는, 관심 프로파일의 세그먼트들은 의미 표현들로 변환될 수 있다. 여기에서의 설명은 관심 프로파일 세그먼트들을 참조하지만, 방법들 및 시스템들은 당업자에 의해 이해될 바와 같이, 몇 개의 변화들이 있을지라도, 완전한 관심 프로파일을 변환하기 위해 구현될 수 있다.

[0046] 관심 프로파일들을 변환하기 위해, 프로파일 변환 모듈(110)은, 관심 프로파일 세그먼트들 각각에 관련된 메타 데이터를 획득하기 위해 통신 채널(108-1)을 통해 의미 메타데이터 데이터베이스(114)와 같은, 데이터베이스와 상호작용할 수 있다. 의미 메타데이터 데이터베이스(114)는 관심 프로파일 세그먼트들을 대응하는 의미 표현들로 변환하기 위해 사용될 수 있는 의미 용어들의 리스트를 가진 사전으로서 이해될 수 있다. 의미 용어들은 그것들의 광범위한 의미 또는 정의에 기초하여 최종 사용자의 다양한 관심 필드들을 식별하는 용어들로서 이해될 수 있으며 그러므로 각각의 관심 필드의 개념에 기초하여 다양한 관심 프로파일들의 효율적이며 애러 없는 매칭을 위해 사용될 수 있다. 관심 필드들의 예들은 이에 제한되지 않지만, 전문적인 관심들, 여가 관심들, 및 종교적 관심들과 같은 맥락과 관련된 선호들 및 음악, 댄스, e-쇼핑, 및 소셜 네트워킹과 같은 관심 영역들을 포함한다. 예를 들면, 소셜 네트워킹 사이트("xyz.com")의 두 개의 상이한 URL들에 대응하는 관심 프로파일들은 버전들이 상이하며 따라서 의미적으로 상이한 URL들에도 불구하고 의미 용어("소셜 네트워킹 사이트")에 매칭될 수 있다. 의미 메타데이터 데이터베이스(114)가 사용자 디바이스(102)의 외부에 있는 것으로 도시되지만, 의미 메타데이터 데이터베이스(114)는 다른 실시예들에서 사용자 디바이스(102)의 내부에 있을 수 있다는 것이 이해될 것이다. 뿐만 아니라, 측면 경우들에서, 의미 메타데이터 데이터베이스들은 사용자 디바이스들 사이에서 어휘 및 용어들을 교환함으로써 사용자 디바이스들에 걸쳐 조화가 이루어질 수 있다.

[0047] 프로파일 변환 모듈(110)에 의해 이렇게 하여 생성된 의미 표현들은 의미 표현들 각각에 클러스터 식별자를 할당하기 위해 클러스터 식별자 모듈(112)에 제공될 수 있다. 클러스터 식별자들은 클러스터들, 즉 최종 사용자가 연관될 수 있는 클러스터 식별자들에 대응하는 관심 그룹들을 효율적으로 식별하기 위해 사용될 수 있는 관심 그룹 아이덴티티 코드들로서 이해될 수 있다. 클러스터 식별자 모듈(112)은 클러스터 식별자들을 할당하기 위한

LSH의 기술을 사용하도록 구성될 수 있다. LSH 기술은 의미 표현들 각각을 대응하는 해시 코드들, 즉 클러스터 식별자들로 변환하기 위해 클러스터 식별자 모듈(112)에 의해 사용될 수 있다. 상기 목적을 위해, 클러스터 식별자 모듈(112)은 사용자 디바이스들(102)의 각각에서 생성된 랜덤 벡터들의 공통 시퀀스와 같은, 랜덤 값들에 의해 정의된 해시 함수들의 세트를 이용한다.

[0048] 하나의 구현에서, 해시 함수들은 클러스터 식별이 모든 사용자 디바이스들(102)에 의해 균일하게 행해짐을 보장하기 위해 모든 사용자 디바이스(102)에 의해 액세스가능한 랜덤 시드 발생기(도면에 도시되지 않음)와 같은, 중앙 엔티티로부터 클러스터 식별자 모듈(112)에 의해 획득된다. 또 다른 구현에서, 클러스터 식별자 모듈(112)은 예를 들면, 중앙 엔티티로부터 획득된 시드 발생 함수들에 기초하여 해시 함수들을 생성하도록 구성될 수 있다. 사용자 디바이스(102)에서 클러스터 식별자들을 할당하는 것은 따라서 관심 프로파일이 원격 노드들과 공유되지 않기 때문에 프라이버시 보호를 제공할 때 용이하게 한다. 의미 표현들을 생성함으로써 프로파일 변환과 함께 LSH 기술에 의해 클러스터링을 수행하는 것은 따라서 클러스터링시 감소된 에러들과 함께 프라이버시 보호를 제공하는 것을 용이하게 한다.

[0049] 그렇게 하여 획득된 클러스터 식별자들은 사용자 디바이스(102)의 사용자 프로파일들, 즉 최종 사용자들을 하나 이상의 클러스터들로 클러스터링하기 위해 사용될 수 있다. 뿐만 아니라, 클러스터 식별자 모듈(112)은 추천 서비스들을 제공하기 위한 또는 피어-투-피어 통신을 위한 클러스터링과 같은, 하나 이상의 서비스들을 위해 통신 채널(108)을 통해 하나 이상의 원격 노드들(104)에 클러스터 식별자들을 제공한다. 사용자 디바이스(102)가 이용하기를 원할 수 있는 추천 기반 서비스의 유형 또는 애플리케이션에 기초하여, 사용자 디바이스(102)와 통신하는 원격 노드(104)가 중앙 집성기 및 관심 그룹 집성기 중에서 변할 수 있다. 대안적으로, 사용자 디바이스들(102)은 서로와 통신할 수 있으며 각각의 사용자 디바이스(102)는 다른 사용자 디바이스들(102)에 대한 원격 노드로서 동작한다.

[0050] 예를 들면, 협업 필터링 애플리케이션의 경우에, 원격 노드들은 적어도 하나는 중앙 집성기들일 수 있으며 각각의 중앙 집성기는 적어도 하나의 클러스터, 즉 영화들, 댄스들, e-쇼핑, 소셜 네트워킹 사이트들, 및 그것의 조합과 같은, 관심 그룹에 대응한다. 이러한 경우에 원격 노드(104)는 그룹에 대응하는 관심 필드에 관련된 뿐만 아니라 다른 최상위 랭킹된, 즉 클러스터의 멤버인 최종 사용자들의 매우 높게 랭킹된 관심 필드들에 관련된 추천들을 클러스터의 각각의 멤버에 제공하도록 구성된다. 또 다른 실시예에서, 중앙 집성기로서 동작하는 원격 노드(104)는 관심 그룹을 책임지는 중앙 집성기로서 동작하는 각각의 사용자 디바이스(102)를 갖고 제거될 수 있다.

[0051] 유사하게, 콘텐츠-기반 추천 애플리케이션들의 경우에, 원격 노드(104)는 하나 이상의 관심 그룹 집성기들로서 구현될 수 있으며 각각의 관심 그룹 집성기는 영화들, 댄스들, e-쇼핑, 소셜 네트워킹 사이트들과 같은, 적어도 하나의 클러스터에 대응한다. 뿐만 아니라, 피어-투-피어 통신들을 수반한 애플리케이션들에서, 원격 노드(104)는 또 다른 다른 사용자 디바이스일 수 있다. 이러한 애플리케이션들에서, 사용자 디바이스들(102)은 유사한 클러스터 식별자들을 가진 다른 사용자 디바이스들(102)을 식별하기 위해 클러스터 식별자들을 사용할 수 있다.

[0052] 본 주제의 시스템 및 방법들은 따라서 사용자 디바이스들에서 관심 프로파일들에 클러스터 식별자들을 할당함으로써 최종 사용자들에게 프라이버시를 제공하는 것을 용이하게 한다.

[0053] 도 2는 본 주제의 일 실시예에 따른, 사용자 프로파일들의 프라이버시 보호된 클러스터링을 구현하는 사용자 디바이스(102)의 구성요소들을 도시한다. 상기 실시예에서, 사용자 디바이스(102)는 하나 이상의 프로세서(들)(202), I/O 인터페이스(들)(204), 및 상기 프로세서(들)(202)에 결합된 메모리(206)를 포함한다. 상기 프로세서(들)(202)는 하나 이상의 마이크로프로세서들, 마이크로컴퓨터들, 마이크로제어기들, 디지털 신호 프로세서들, 중앙 프로세싱 유닛들, 상태 기계들, 논리 회로들, 및/또는 동작 지시들에 기초하여 신호들을 조작하는 임의의 디바이스들로서 구현될 수 있다. 상기 프로세서(들)(202)는 단일 프로세싱 유닛 또는 다수의 유닛들일 수 있으며, 그 모두는 또한 다수의 컴퓨팅 유닛들을 포함할 수 있다. 다른 능력들 중에서, 프로세서(들)(202)는 메모리(206)에 저장된 컴퓨터-관독가능한 지시들을 인출하며 실행하도록 구성된다.

[0054] "프로세서(들)"로서 라벨링된 임의의 기능 블록들을 포함하는, 도면에 도시된 다양한 요소들의 기능들이 적절한 소프트웨어에 관련하여 소프트웨어를 실행할 수 있는 하드웨어 뿐만 아니라, 전용 하드웨어의 사용을 통해 제공될 수 있다. 프로세서에 의해 제공될 때, 기능들은 단일 전용 프로세서에 의해, 단일 공유 프로세서에 의해, 또는 그 몇몇이 공유될 수 있는 복수의 개개의 프로세서들에 의해 제공될 수 있다. 게다가, 용어("프로세서")의 명시적인 사용은 전적으로 소프트웨어를 실행할 수 있는 하드웨어를 나타내는 것으로 해석되어서는 안되고, 제한 없이, 디지털 신호 프로세서(DSP) 하드웨어, 네트워크 프로세서, 주문형 반도체(ASIC), 필드 프로그래밍가능

한 게이트 어레이(FPGA), 소프트웨어를 저장하기 위한 판독 전용 메모리(ROM), 랜덤 액세스 메모리(RAM), 및 비휘발성 저장 장치를 암시적으로 포함할 수 있다. 종래의 및/또는 맞춤형, 다른 하드웨어가 또한 포함될 수 있다.

- [0055] I/O 인터페이스(들)(204)는 다양한 소프트웨어 및 하드웨어 인터페이스들, 예를 들면, 키보드, 마우스, 외부 메모리, 프린터 등과 같은, 주변 디바이스(들)를 위한 인터페이스를 포함할 수 있다. 뿐만 아니라, I/O 인터페이스(들)(204)는 사용자 디바이스(102)로 하여금 통신 네트워크(106)를 통해 통신할 수 있게 할 수 있으며, 웹 서버들 및 외부 데이터베이스들과 같은, 다른 컴퓨팅 디바이스들과 사용자 디바이스(102)를 연결하기 위한 하나 이상의 포트들을 포함할 수 있다. I/O 인터페이스(들)(204)는 유선 네트워크들, 예로서 LAN, 케이블 등, 및 무선 네트워크들, 예로서 WLAN, 셀룰러, 위성 등을 포함하는, 네트워크와 같은 매우 다양한 네트워크들 및 프로토콜들 내에서 다수의 통신들을 용이하게 할 수 있다.
- [0056] 메모리(206)는 예를 들면, 정적 랜덤 액세스 메모리(SRAM) 및 동적 랜덤 액세스 메모리(DRAM)와 같은 휘발성 메모리, 및/또는 판독 전용 메모리(ROM), 삭제 가능한 프로그래밍가능한 ROM, 플래시 메모리들, 하드 디스크들, 광 디스크들, 및 자기 테이프들과 같은 비-휘발성 메모리를 포함하는, 이 기술분야에 알려진 임의의 컴퓨터-판독가능한 매체를 포함할 수 있다. 메모리(206)는 모듈들(208) 및 데이터(210)를 추가로 포함한다. 상기 모듈들(208)은 루틴들, 프로그램들, 오브젝트들, 구성요소들, 데이터 구조들 등을 포함하고, 이것은 특정한 태스크들을 수행하거나 특정한 추상 데이터 유형들을 구현한다. 데이터(210)는 다른 것들 중에서, 모듈들(208) 중 하나 이상에 의해 프로세싱되고, 수신되고, 생성된 데이터를 저장하기 위한 저장소로서 작용한다.
- [0057] 상기 모듈들(208)은 관심 프로파일 생성 모듈(212), 프로파일 분할 모듈(214), 프로파일 변환 모듈(110), 클러스터 식별자 모듈(112), 데이터 전송 모듈(216), 및 다른 모듈들(218)을 추가로 포함한다. 다른 모듈(들)(218)은 사용자 디바이스(102)의 애플리케이션들 및 기능들을 보완하는 프로그램들 또는 코딩된 지시들, 예를 들면 운영 시스템에서의 프로그램들을 포함할 수 있다.
- [0058] 데이터(210)는 프로파일 데이터(220), 변환된 프로파일 데이터(222), 클러스터 식별자 데이터(224), 및 다른 데이터(226)를 포함한다. 다른 데이터(226)는 다른 모듈(들)(218)에 하나 이상의 모듈들의 실행의 결과로서 생성된 데이터를 포함할 수 있다.
- [0059] 이전에 설명된 바와 같이, 사용자 디바이스(102)는 사용자 프로파일들을 클러스터링하기 위한 클러스터 식별자들을 계산함으로써 최종 사용자에게 프라이버시 보호를 제공하도록 구성된다. 하나의 실시예에서, 관심 프로파일 생성 모듈(212)은 프로파일 데이터(220)에 기초하여 사용자 디바이스(102)의 최종 사용자의 관심 프로파일을 생성하도록 구성된다. 프로파일 데이터(220)는 예를 들면, 최종 사용자들에 의해 방문된 웹사이트들, 최종 사용자들에 의해 플레이되거나 다운로드된 노래들 또는 비디오들, 최종 사용자들에 의해 사용된 제품들 또는 그에 의해 이용되거나 검토된 서비스들 등을 표시할 수 있다. 예를 들면, 관심 프로파일 생성 모듈(212)은 최종 사용자들 위해 탐색되거나 그에 의해 시청된 모든 비디오들, 최종 사용자에 의해 방문된 모든 URL들, 최종 사용자에게 청구되며 그에 의해 탐색된 음악, 및 최종 사용자들에 의해 탐색된 다른 관심 필드들에 관련된 유사한 데이터의 세부사항들을 획득할 수 있다. 따라서, 생성된 관심 프로파일들은 프로파일 데이터(220)에 저장될 수 있다. 프로파일 변환 모듈(110)은 그 다음에 관심 프로파일을 클러스터 식별자들을 할당하기 위한 의미 표현들로 변환할 수 있다. 이전에 논의된 바와 같이, 프로파일 변환 모듈(110)은 완전한 관심 프로파일 또는 관심 프로파일 세그먼트들을 의미 표현들로 변환하도록 구성될 수 있다. 후자의 경우들에서, 관심 프로파일들은 프로파일 변환 모듈(110)에 의해 변환되기 전에 프로파일 분할 모듈(214)에 의해 액세스되고 분할된다.
- [0060] 하나의 실시예에서, 프로파일 변환 모듈(110)은 관심 프로파일을 다수의 관심 프로파일 세그먼트들로 분할하도록 구성되고, 각각의 관심 프로파일 세그먼트들은 특정한 관심 필드에 대응한다. 하나의 구현에서, 프로파일 변환 모듈(110)은 관심 프로파일을 분할하기 위해 최종 사용자의 관심 프로파일을 분석할 수 있다. 프로파일 변환 모듈(110)은 여가 선호도 및 전문적 관심들과 같은 맥락과 관련된 선호도 또는 음악, 댄스, e-쇼핑, 및 소셜 네트워킹과 같은 명시적인 관심 영역들에 기초하여 관심 프로파일들을 관심 프로파일 세그먼트들로 분할할 수 있다. 따라서, 획득된 관심 프로파일 세그먼트들은 프로파일 변환 모듈(110)에 제공될 수 있다.
- [0061] 관심 변환 모듈(110)은 사용자 디바이스(102)의 내부에 있거나 외부에 있을 수 있는 의미 메타데이터 데이터베이스(114)에 제공된 메타데이터에 기초하여 관심 프로파일 세그먼트들 또는 완전한 관심 프로파일들을 대응하는 의미 표현들로 변환한다. 하나의 실시예에서, 프로파일 변환 모듈(110)은 관심 프로파일 세그먼트들을 의미 메타데이터 데이터베이스(114)에 제공할 수 있다. 이러한 경우에 의미 메타데이터 데이터베이스(114)는 관심 프로파일 세그먼트들 각각에 대응하는 의미 용어들을 식별하며 이를 관심 프로파일 세그먼트들을 대응하는 의미 표

현들로 변환하기 위한 프로파일 변환 모듈(110)에 제공한다. 또 다른 실시예에서, 프로파일 변환 모듈(110)은 의미 메타데이터 데이터베이스(114)로부터 의미 용어들의 리스트를 획득할 수 있다. 이러한 경우에 프로파일 변환 모듈(110)은 의미 용어들의 리스트로부터, 대응하는 의미 표현들을 획득하기 위해 관심 프로파일 세그먼트와 일치하는 의미 용어를 식별하도록 구성될 수 있다.

[0062] 또 다른 실시예에서, 프로파일 변환 모듈(110)은 사용자 디바이스(102) 상에서 최종 사용자에게 의해 방문된 웹 페이지들로부터 추출된 태그들과 같은, 메타데이터를 사용하여 완전한 관심 프로파일들 또는 관심 프로파일 세그먼트들을 태깅함으로써 의미 표현들을 생성할 수 있다. 이러한 국부적으로 추출된 태그들은 사용자 디바이스(102)의 다른 데이터(226)에서 내부 의미 메타데이터로서 저장될 수 있다. 상기 실시예에서, 사용자 디바이스들은 상이한 사용자 디바이스들에 의해 생성된 의미 표현들이 동일한 용어들을 사용하여 표현되며 그러므로 비교 가능하도록 서로와 의미 용어들의 어휘를 공유하도록 구성된다. 의미 메타데이터 데이터베이스를 사용하는 것은 위음성들 및 위양성들의 발생들을 감소시키며, 따라서 클러스터링의 효율성을 증가시키도록 돕는다. 그에 따라 획득된 의미 표현들은 변환된 프로파일 데이터(222)에 저장되며 추가 클러스터링을 위한 의미 표현들 각각에 클러스터 식별자들을 할당하기 위해 클러스터 식별자 모듈(112)에 의해 액세스될 수 있다.

[0063] 도 1에 설명된 바와 같이, 클러스터 식별자 모듈(112)은 최종 사용자의 프로파일의 클러스터링을 가능하게 하기 위해 의미 표현들 각각에 클러스터 식별자들을 할당하도록 구성된다. 클러스터 식별자 모듈(112)은 의미 표현들에 클러스터 식별자들을 할당하기 위해 LSH 기술을 사용한다. 당업자에 의해 이해될 바와 같이, LSH 기술에서, 두 개의 프로파일들의 해시 충돌 확률과 동일한 높은 확률을 가진 값에 대한 두 개의 유사한 오브젝트들 해시는 다음의 수학적식을 사용하여 표현된 바와 같이, 두 개의 프로파일들 사이의 유사성에 비례한다:

[0064] 
$$\Pr[h(x) = h(y)] \approx \text{sim}(x,y) \dots \dots \dots (1)$$

[0065] 여기에서  $x$  및  $y$ 는 두 개의 프로파일들이고,  $h()$ 는 LSH 함수이며,  $\text{sim}()$ 은 프로파일들 사이의 유사성 측정치이다. 뿐만 아니라, LSH 기술의 기능을 달성하기 위해, 유사한 해시 함수들이 LSH 기술을 위해 사용자 디바이스(102)에 의해 사용된다는 것이 보장된다.

[0066] 클러스터 식별자 모듈(112)은 클러스터 식별자, 즉 유사한 관심들을 가진 최종 사용자들, 즉 유사한 관심 프로파일들을 가진 최종 사용자들의, 관심 그룹의 라벨 또는 그룹 id로서 해시 함수들에 의해 생성된 값을 사용하도록 구성된다. 상기 목적을 위해, 클러스터 식별자 모듈(112)은 처음에 랜덤 시드 발생기와 같은, 중앙 엔티티로부터, 또한 해시 함수들로서 불리우는, 랜덤 벡터들의 세트를 획득한다. 클러스터 식별자 모듈(112)은 그 다음에, 예를 들면, LSH 기술의 코사인 유사성 방법을 사용한다. 상기 방법에서, 클러스터 식별자 모듈(112)은 의미 표현들과 랜덤 벡터들 사이에서 내적(dot product)을 획득하고 의미 표현에 대한 클러스터 식별자로서 해시 코드를 획득하기 위해 상기 내적들의 부호를 연결한다. 따라서, 획득된 클러스터 식별자들은 클러스터 식별자 데이터(224)에 저장될 수 있다. 여기에서의 설명은 LSH 기술의 코사인 유사성 방법을 참조하지만, 방법들 및 시스템들은 당업자에 의해 이해될 바와 같이, LSH 기술의 다른 구현들을 사용하여 구현될 수 있다. 의미 표현들을 생성함으로써 프로파일 변환과 함께 LSH 기술에 의해 클러스터링을 수행하는 것은 클러스터링시 감소된 에러들을 가진 프라이버시 보호를 제공하는 것을 용이하게 한다. 뿐만 아니라, LSH 기술을 사용하여 클러스터링을 수행하기 전에 의미 표현들을 생성하는 것은 원격 노드(104)가 클러스터 멤버들에 걸쳐 선형 탐색을 수행할 수 없음을 보장하도록 돕는다.

[0067] 데이터 전송 모듈(216)은 그 다음에 예를 들면, 서비스 제공자로부터 추천 또는 콘텐츠들을 수신하기 위해, 통신 채널들(108)을 통해 원격 노드(104)에 클러스터 식별자들을 송신한다. 하나의 구현에서, 원격 노드(104)에 클러스터 식별자들을 전송하는 방법은 클러스터링이 수행되는 애플리케이션에 의존하여 변할 수 있다. 예를 들면, 협업 필터링 애플리케이션의 경우에, 데이터 전송 모듈(216)은 원격 노드로 하여금 그룹 추천들을 제공하기 위한 사용자 프로파일을 분석할 수 있게 하기 위해 클러스터 식별자들 뿐만 아니라, 최종 사용자의 관심 프로파일을 또한 원격 노드(104)로 송신하도록 구성될 수 있다. 이러한 경우에 데이터 전송 모듈(216)은 관심 프로파일들을 보다 작은 프로파일 슬라이스들로 슬라이싱하며 통신 채널들(108)과 같은, 상이한 익명 통신 채널을 통해 보다 작은 프로파일 슬라이스들 각각을 원격 노드(104)에 독립적으로 송신하도록 구성될 수 있다. 프로파일 세그먼트들과 달리, 프로파일 슬라이스들은 원격 노드로 하여금 프로파일 슬라이스들을 분석할 수 있게 하기 위해 의미 표현들로 변환되지 않고 송신된다. 익명 통신 채널들을 통해 상이한 슬라이스들로서 사용자 프로파일들을 송신하는 것은 원격 노드(104)가 관심 프로파일의 슬라이스들에 기초하여 최종 사용자들을 쉽게 식별할 수 없기 때문에 프라이버시 보호를 보장한다. 협업 필터링 애플리케이션의 또 다른 실시예에서, 원격 노드(104)는 교체될 수 있으며 각각의 사용자 디바이스(102)는 관심 그룹을 책임지는 중앙 집성기로서 동작하도록 구성될 수

있다.

- [0068] 뿐만 아니라, 콘텐츠-기반 추천 애플리케이션들의 경우에, 데이터 전송 모듈(216)은 서비스 제공자들에 의해 제공된 추천 또는 분석 서비스들을 이용하기 위해 클러스터 식별자 및 의미 표현들을 원격 노드(104)로 송신하도록 구성될 수 있다. 유사하게, 피어-투-피어 통신들의 경우에, 데이터 전송 모듈(216)은 사용자 디바이스들(102) 사이에서 데이터의 교환 및 채팅을 가능하게 하기 위해 유사한 클러스터 식별자들을 가진 다른 사용자 디바이스들(102)을 식별하도록 구성될 수 있다. 다른 사용자 디바이스들(102)을 식별할 때, 데이터 전송 모듈(216)은 다른 디바이스들 사이에 익명 메시징 채널들 및 익명 채팅 또는 탐색 브로드캐스트 채널들을 확립하도록 구성될 수 있다. 애플리케이션들 각각에서 데이터 전송 모듈(216)의 작동이 도 3a, 도 3b, 및 도 3c에서의 애플리케이션들을 설명하면서 보다 상세히 설명될 것이다.
- [0069] 도 3a, 도 3b, 및 도 3c은 본 주제의 일 실시예에 따른, 다양한 애플리케이션들을 위해 사용자 프로파일들의 프라이버시 보호된 클러스터링의 대표적인 네트워크 환경 구현들(300, 302, 및 304)을 각각 도시한다.
- [0070] 도 3a은 클러스터의 멤버들에 그룹 추천들을 제공하는 협업 필터링 애플리케이션을 위해 사용자 프로파일들의 프라이버시 보호된 클러스터링을 구현하는 네트워크 환경(300)을 도시한다. 협업 필터링 애플리케이션의 상기 실시예에서, 원격 노드(104)는 영화들, 댄스들, e-쇼핑, 소셜 네트워킹 사이트들과 같은, 적어도 하나의 클러스터에 대응하는 중앙 집성기로서 구현된다. 사용자 디바이스(102)는 따라서 중앙 집성기로 하여금 최고 랭킹된, 즉 그들의 대응하는 관심 그룹들에서의 관심 그룹의 관심 필드 후 최상위 랭킹된 관심 필드를 식별할 수 있게 하기 위해 최종 사용자의 관심 프로파일들과 함께 대응하는 중앙 집성기에 클러스터 식별자들 각각을 제공할 수 있다.
- [0071] 하나의 실시예에서, 데이터 전송 모듈(216)은 원격 노드(104)에 의해 수신된 다양한 프로파일 세그먼트들 사이의 불연계성을 보장하기 위해 다양한 통신 채널들(108)을 통해 전달될 수 있는 보다 작은 독립적 슬라이스들로 관심 프로파일을 슬라이싱하도록 구성된다. 불연계성을 제공하는 것은 원격 노드(104)가 상이한 프로파일 세그먼트들이 동일한 사용자 또는 상이한 사용자에 속하는지의 여부를 식별할 수 없을 수 있음을 보장한다. 데이터 전송 모듈(216)은 최종 사용자의 관심 프로파일의 어떤 슬라이스도 홀로 완전한 관심 프로파일을 구성하며 최종 사용자의 아이덴티티를 추론하기 위해 사용될 수 있는 충분한 프로파일 정보를 포함하지 않는다는 것을 보장한다. 또 다른 실시예에서, 전체 관심 프로파일은 단일 슬라이스로서 원격 노드(104)로 전송될 수 있다.
- [0072] 뿐만 아니라, 관심 프로파일 슬라이스들 각각 및 클러스터 식별자는 익명성을 보장하는 메커니즘들을 이용한 네트워크, 예를 들면, 어니언 라우팅을 구현하는 네트워크를 통해 익명 통신 채널들을 경유하여 데이터 전송 모듈(216)에 의해 전송된다. 하나의 구현에서, 어니언-라우팅 경로가 확립되고, 여기에서 데이터 전송 모듈(216)은 어니언-라우팅 경로의 출구 노드의 공개-키를 갖고 최종 사용자와 관련된 클러스터 식별자 및 프로파일 정보의 슬라이스들을 암호화한다. 최종 사용자와 관련된 클러스터 식별자 및 프로파일 정보의 다양한 슬라이스들이 출구 노드에 도달하기 전에 하나 이상의 중간 노드들을 통해 송신된다. 출구-노드는 정보를 복호화하며 이를 원격 노드(104)로 송신한다.
- [0073] 뿐만 아니라, 사용자 프라이버시를 보장하기 위해, 데이터 전송 모듈(216)은 원격 노드(104)에 최종 사용자들의 아이덴티티를 드러낼 수 있는 URL들과 같은, 개인적으로 식별가능한 정보를 가진 모든 데이터를, 관심 프로파일로부터 제거하도록 구성될 수 있다.
- [0074] 최종 사용자의 관심 프로파일 및 클러스터 식별자들을 수신할 때, 원격 노드(104)는 원격 노드(104)에 대응하는 클러스터에서 클러스터링된 최종 사용자들의 최고 또는 하나 이상의 높게 랭킹된 관심을 결정하기 위해 관심 프로파일을 분석한다. 예를 들면, 관심 그룹('소셜 네트워킹')을 책임지는 원격 노드(104)는 클러스터('소셜 네트워킹')에 대응하는 클러스터 식별자를 가진 사용자 디바이스들 모두의 관심 프로파일들을 수신할 수 있다. 원격 노드(104)는 그 다음에 높게 랭킹된 다른 관심들을 식별하기 위해 사용자 디바이스(102)의 사용자 프로파일들을 분석할 수 있다. 예를 들면, 최고 랭킹된 관심으로서 '뉴스'를 식별할 때, 원격 노드(104)는 관심 필드('뉴스')를 책임지는 중앙 집성기로부터 '뉴스'에 대응하는 추천들을 획득할 수 있다. 원격 노드(104)는 따라서 클러스터('소셜 네트워킹')에 대응하는 클러스터에 존재하는 사용자 디바이스들(102) 모두에 뉴스 뿐만 아니라, 소셜 네트워킹에 관련된 콘텐츠들 및 추천들을 제공할 수 있다.
- [0075] 협업 필터링 애플리케이션의 또 다른 실시예에서, 중앙 집성기로서 동작하는 원격 노드(104)는 관심 그룹을 책임지는 중앙 집성기로서 동작하는 각각의 사용자 디바이스(102)를 갖고 제거될 수 있다. 상기 목적을 위해, 데이터 전송 모듈(216)은 최종 사용자의 관심 그룹들의 중앙 집성기 노드의 적어도 부분적인 이미지를 유지하도록

구성된다. 상기 실시예에서, 당업자들에게 알려질 바와 같이, 유사한 관심 프로파일들을 가진 최종 사용자들은 종래의 가십 기반 랜덤 피어링 프로토콜에 기초하여 식별될 수 있다. 데이터 전송 모듈(216)은 따라서 유사한 클러스터 식별자, 즉 유사한 관심 그룹들을 가진 최종 사용자들을 식별하기 위해 사용자 디바이스들(102) 모두의 클러스터 식별자들을 송신 및 수신하도록 구성될 수 있다.

[0076] 도 3b는 콘텐츠-기반 추천 애플리케이션들을 위해 사용자 프로파일들의 프라이버시 보호된 클러스터링을 구현하는 네트워크 환경(302)을 도시한다. 상기 실시예에서, 원격 노드(104)는 각각의 관심 그룹 집성기가 영화들, 덴스들, e-쇼핑, 소셜 네트워킹 사이트들과 같은, 적어도 하나의 클러스터에 대응하는, 하나 이상의 관심 그룹 집성기들로서 구현된다.

[0077] 이러한 경우에, 데이터 전송 모듈(216)은 통신 채널(108)과 같은, 익명 통신 채널을 통해 원격 노드(104)에 클러스터 식별자 및 의미 표현들을 송신하도록 구성된다. 상기 목적을 위해, 데이터 전송 모듈(216)은 도 3a에 설명된 바와 같은 어니언-라우팅 경로의 출구 노드의 공개-키를 갖고 최종 사용자와 관련된 클러스터 식별자를 암호화하는 데이터 전송 모듈(216)을 가진 어니언-라우팅 경로를 확립한다. 이러한 경우에 원격 노드(104)는 잘 알려진 콘텐츠 매칭 기술들 중 임의의 것을 사용하여 관심 필드들의 리스트와 클러스터 식별자를 매칭시킴으로써 클러스터 식별자에 의해 표현된 관심 필드에 관련된 추천들 또는 콘텐츠들을 선택하도록 구성된다. 예를 들면, 최종 사용자, 즉 사용자(A)가 관심 필드('책들')를 가진 클러스터에 속한다면, 서비스 제공자는 사용자(A)에게 최근 톱 셀러들의 리스트를 제공할 수 있다.

[0078] 도 3c는 피어-투-피어 통신들에 관련된 애플리케이션들을 위해 사용자 프로파일의 프라이버시-보호된 클러스터링을 구현하는 네트워크 환경(304)을 도시한다. 상기 실시예에서, 원격 노드(104)는 사용자 디바이스들(102)이 프라이버시 보호된 환경에서 서로와 상호작용하기 위해 클러스터링을 사용하기 때문에 제거된다. 이러한 애플리케이션들에서, 사용자 디바이스들(102)은 유사한 클러스터 식별자들을 가진 다른 사용자 디바이스들(102)을 식별하기 위해 클러스터 식별자들을 사용할 수 있다. 이러한 경우에, 데이터 전송 모듈(216)은 통신 채널(108)과 같은, 익명 통신 채널을 통해 클러스터 식별자를 다른 사용자 디바이스들(102)로 송신하도록 구성된다. 데이터 전송 모듈(216)은 클러스터 식별자를 송신하며 유사한 클러스터 식별자들을 가진 사용자 디바이스(102)를 식별하기 위해 가십 기반 랜덤 피어링 프로토콜을 사용한다. 공통 클러스터 식별자들을 가진 사용자 디바이스들(102)은 그 다음에 그것들의 아이덴티티 또는 인적 정보를 누설하지 않고 사용자 디바이스들 사이에서의 데이터의 교환 및 채팅을 가능하게 하기 위해, 예를 들면, TOR과 같은 어니언 라우팅 프레임워크를 사용하여 익명으로서 서로에 연결할 수 있다.

[0079] 도 4는 본 주제의 일 실시예에 따른, 최종 사용자들에게 프라이버시 보호를 제공하기 위해 사용자 프로파일들의 프라이버시 보호된 클러스터링의 방법을 도시한다. 방법이 설명되는 순서는 제한으로서 해석되도록 의도되지 않으며, 임의의 수의 설명된 방법 블록들은 방법(400) 또는 임의의 대안적인 방법들을 구현하기 위해 임의의 순서로 조합될 수 있다. 부가적으로, 개개의 블록들은 여기에 설명된 주제의 사상 및 범위로부터 벗어나지 않고 방법들로부터 삭제될 수 있다. 더욱이, 방법들은 임의의 적절한 하드웨어, 소프트웨어, 펌웨어, 또는 그것의 조합으로 구현될 수 있다.

[0080] 방법(들)은 컴퓨터 실행가능한 지시들의 일반적인 맥락에서 설명될 수 있다. 일반적으로, 컴퓨터 실행가능한 지시들은 특정한 기능들을 수행하거나 특정한 추상 데이터 유형들을 구현하는 루틴들, 프로그램들, 오브젝트들, 구성요소들, 데이터 구조들, 절차들, 모듈들, 기능들 등을 포함할 수 있다. 방법은 또한 기능들이 통신 네트워크를 통해 연계되는 원격 프로세싱 디바이스들에 의해 수행되는 분산 컴퓨팅 환경에서 실시될 수 있다. 분산 컴퓨팅 환경에서, 컴퓨터 실행가능한 지시들은 메모리 저장 디바이스들을 포함하는, 로컬 및 원격 컴퓨터 저장 매체들 양쪽 모두에 위치될 수 있다.

[0081] 당업자는 방법들의 단계들이 프로그램된 컴퓨터들에 의해 수행될 수 있다는 것을 쉽게 인식할 것이다. 여기에서, 몇몇 실시예들은 또한 기계 또는 컴퓨터 판독가능하며 지시들의 기계-실행가능한 또는 컴퓨터-실행가능한 프로그램들을 인코딩하는 프로그램 저장 디바이스들, 예를 들면, 디지털 데이터 저장 매체들을 커버하도록 의도되고, 여기에서 상기 지시들은 설명된 방법의 단계들 중 일부 또는 모두를 수행한다. 프로그램 저장 디바이스들은 예를 들면, 디지털 메모리들, 자기 디스크들 및 자기 테이프들과 같은 자기 저장 매체들, 하드 드라이브들, 또는 광학적으로 판독가능한 디지털 데이터 저장 매체들일 수 있다. 실시예들은 또한 대표적인 방법들의 상기 단계들을 수행하도록 구성된 통신 디바이스들 및 통신 네트워크 양쪽 모두를 커버하도록 의도된다.

[0082] 블록(402)에서, 최종 사용자에게 대응하는 프로파일 데이터는 사용자 디바이스, 예를 들면, 사용자 디바이스(102)에 의해 획득된다. 하나의 구현에서, 프로파일 데이터는 최종 사용자들에 의해 방문된 웹사이트들, 최종 사용

자들에 의해 플레이되거나 다운로드된 노래들 또는 비디오들, 최종 사용자들에 의해 사용된 제품들 또는 그에 의해 이용되거나 검토된 서비스들 등에 기초하여 획득될 수 있다. 관심 프로파일 생성 모듈, 예를 들면, 관심 프로파일 생성 모듈(212)은 최종 사용자를 위해 탐색되거나 그에 의해 시청된 모든 비디오들, 최종 사용자에 의해 방문된 모든 URL들, 최종 사용자에 의해 이루어진 모든 질의들, 및 최종 사용자들에 의해 탐색된 다른 관심 필드들에 관련된 유사한 데이터의 세부사항들을 획득하도록 구성될 수 있다. 이렇게 하여 생성된 관심 프로파일들은 프로파일 데이터로서 저장될 수 있다.

[0083] 블록(404)에서, 프로파일 데이터는 하나 이상의 관심 프로파일 세그먼트들로 분할된다. 하나의 구현에서, 프로파일 분할 모듈(214)과 같은, 프로파일 분할 모듈은 관심 프로파일 세그먼트들로서 불리우는, 다수의 세그먼트들로 관심 프로파일을 분할하도록 구성된다. 관심 프로파일 세그먼트들은 관심 프로파일들의 독립 세그먼트들로서 이해될 수 있으며 각각의 세그먼트는 특정한 관심 필드에 대응한다. 관심 필드들의 예들은, 이에 제한되지 않지만, 전문적 관심들, 여가 관심들, 및 종교적 관심들과 같은 맥락과 관련된 선호들, 및 음악, 댄스, e-쇼핑, 및 소셜 네트워킹과 같은 관심의 영역들을 포함한다. 관심 프로파일 세그먼트들 각각은 그 다음에 클러스터링을 위해 클러스터 식별자들을 획득하기 위해 독립적으로 프로세싱된다. 뿐만 아니라, 또 다른 실시예에서, 전체로서 관심 프로파일은 클러스터링을 위한 단일 관심 프로파일 세그먼트로서 제공될 수 있다.

[0084] 블록(406)에서, 관심 프로파일 세그먼트들 각각에 대응하는 의미 용어들은 의미 메타데이터 데이터베이스와의 상호작용에 기초하여 확인된다. 하나의 구현에서, 프로파일 변환 모듈(110)과 같은, 프로파일 변환 모듈은 익명 통신 채널을 통해 관심 프로파일 세그먼트들을 의미 메타데이터 데이터베이스에 제공하도록 구성될 수 있다. 관심 프로파일 세그먼트들을 수신할 때, 의미 메타데이터 데이터베이스는 관심 프로파일 세그먼트들 각각에 대응하는 의미 용어들을, 의미 용어들의 리스트로부터 식별한다. 대안적으로, 프로파일 변환 모듈(110)은 사용자 디바이스(102) 상에서 최종 사용자에 의해 방문된 웹 페이지들로부터 추출된 태그들과 같은, 데이터를 사용하여 실장된 사용자 디바이스(102)의 내부 의미 메타데이터와 상호작용할 수 있다. 상기 실시예에서, 사용자 디바이스들(102)은 사용자 디바이스들(102) 모두에 의해 사용된 균일성을 생성된 의미 용어들에 걸쳐 유지하기 위해 서로와 의미 용어들의 어휘를 공유하도록 구성된다.

[0085] 블록(408)에서, 관심 프로파일 세그먼트들은 의미 용어들에 기초하여 대응하는 의미 표현들을 생성하기 위해 변환된다. 예를 들면, 프로파일 변환 모듈(110)과 같은, 프로파일 변환 모듈은 예를 들면, 의미 메타데이터 데이터베이스로부터 획득된 대응하는 의미 표현들을 갖고 관심 프로파일 세그먼트들 각각을 태깅함으로써, 관심 프로파일 세그먼트들을 변환하도록 구성된다. 관심 프로파일 세그먼트들은 대안적으로 온톨로지 개념 기반 또는 토픽 기반 의미 표현들로 변환될 수 있다. 그렇게 하여 획득된 의미 표현들은 변환된 프로파일 데이터(222)에 저장될 수 있다.

[0086] 블록(410)에서, 클러스터 식별자들은 LSH 기술을 사용하여 의미 표현들에 할당된다. 하나의 구현에서, 랜덤 벡터들의 공통 시퀀스와 같은, 랜덤 값들의 세트에 의해 정의된 해시 함수들은 클러스터 식별자들을 할당하기 위해 사용된다. 예를 들면, 클러스터 식별자 모듈(112)과 같은 클러스터 식별 모듈은 균일한 클러스터 식별을 보장하기 위해 사용자 디바이스들에 의해 액세스가능한 랜덤 시드 발생기와 같은, 중앙 엔티티로부터 해시 함수들을 획득할 수 있다. 의미 표현들의 공통 세트를 사용한 프로파일 변환과 함께 LSH 기술에 의해 클러스터링을 수행하는 것은 따라서 클러스터링 감소된 에러들을 갖고 프라이버시 보호를 제공하는 것을 용이하게 한다.

[0087] 블록(412)에서, 관심 프로파일 세그먼트들 각각에 대응하는 클러스터 식별자들은 클러스터 식별자들에 기초하여 사용자 관심 프로파일들을 클러스터링하기 위해 원격 노드에 제공된다. 하나의 구현에서, 원격 노드는 중앙 엔티티일 수 있으며 클러스터링이 수행되는 애플리케이션의 유형에 의존하여 중앙 집성기 노드로부터 관심 집성기 노드로 변할 수 있다. 예를 들면, 원격 노드는 콘텐츠-기반 추천 애플리케이션들의 경우에 관심 그룹 집성기로서 구현될 수 있다. 이러한 경우에 사용자 디바이스들은 익명 통신 채널을 통해, 클러스터 식별자들 및 의미 표현들 각각을 클러스터 식별자들에 대응하는 관심 그룹 집성기로 송신하도록 구성된다.

[0088] 유사하게, 협업 필터링 애플리케이션의 경우에, 원격 노드들은 그룹에 대응하는 관심 필드에 관련될 뿐만 아니라, 또한 클러스터의 최종 사용자들의 다른 높게 랭킹된 관심 필드들에 관련된 추천들을 클러스터의 각각의 멤버에 제공하도록 구성된 중앙 집성기들로서 구현될 수 있다. 이러한 경우에 사용자 디바이스들은 익명 통신 채널을 통해, 클러스터 식별자들에 대응하는 관심 그룹 집성기로서 사용자 디바이스들을 가능하게 하기 위해 최종 사용자의 관심 프로파일들의 슬라이스들과 함께 클러스터 식별자들 각각을 송신하도록 구성된다. 또 다른 실시예에서, 사용자 디바이스들은 사용자 디바이스들 각각이 관심 그룹을 책임지는 중앙 집성기 또는 다른 사용자 디바이스들에 대한 원격 노드로서 동작하도록 다른 사용자 디바이스들과 상호작용할 수 있다.

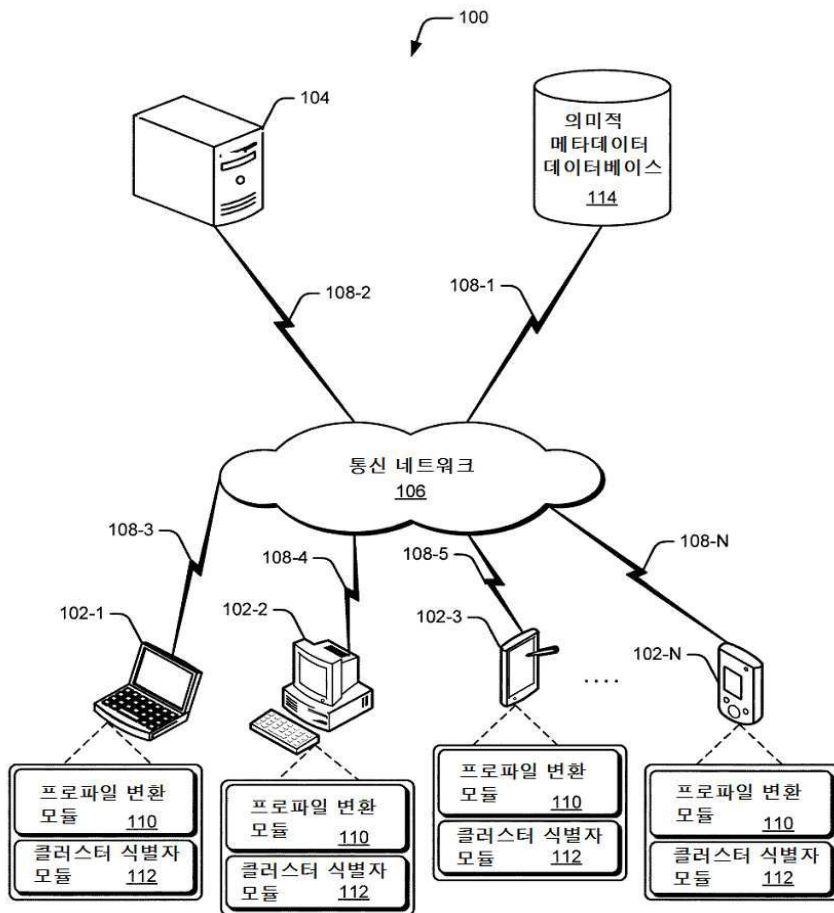
[0089] 통신 네트워크에서 사용자 관심 프로파일들의 프라이버시 보호된 클러스터링을 위한 방법들 및 시스템들에 대한 실시예들이 구조적 특징들 및/또는 방법들에 특징적인 언어로 설명되었지만, 본 발명은 반드시 설명된 특정 특징들 또는 방법들에 제한되는 것은 아니라는 것이 이해될 것이다. 오히려, 특정 특징들 및 방법들은 사용자 관심 프로파일들의 프라이버시 보호된 클러스터링을 제공하기 위한 대표적인 실시예들로서 개시된다.

**부호의 설명**

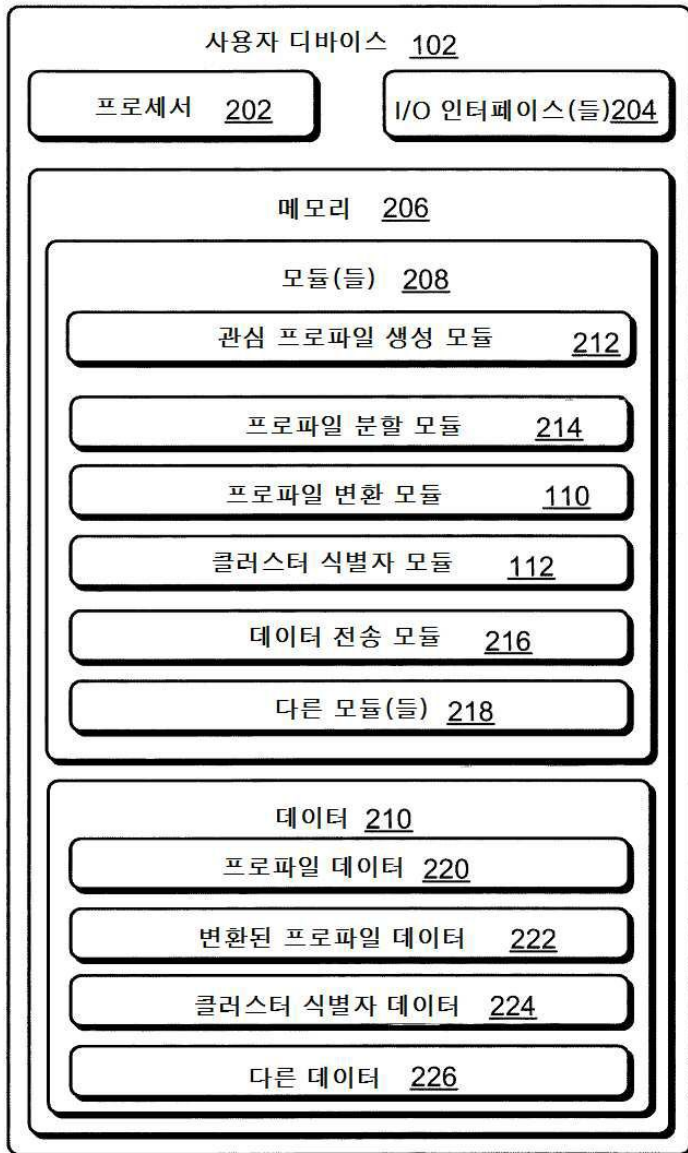
- |        |                      |                 |
|--------|----------------------|-----------------|
| [0090] | 100: 네트워크 환경         | 102: 사용자 디바이스   |
|        | 104: 원격 노드           | 106: 통신 네트워크    |
|        | 108: 통신 채널           | 110: 프로파일 변환 모듈 |
|        | 112: 클러스터 식별자 모듈     |                 |
|        | 114: 의미 메타데이터 데이터베이스 | 202: 프로세서       |
|        | 204: I/O 인터페이스       | 206: 메모리        |
|        | 208: 모듈              |                 |
|        | 212: 관심 프로파일 생성 모듈   | 214: 프로파일 분할 모듈 |
|        | 216: 데이터 전송 모듈       | 218: 다른 모듈      |

**도면**

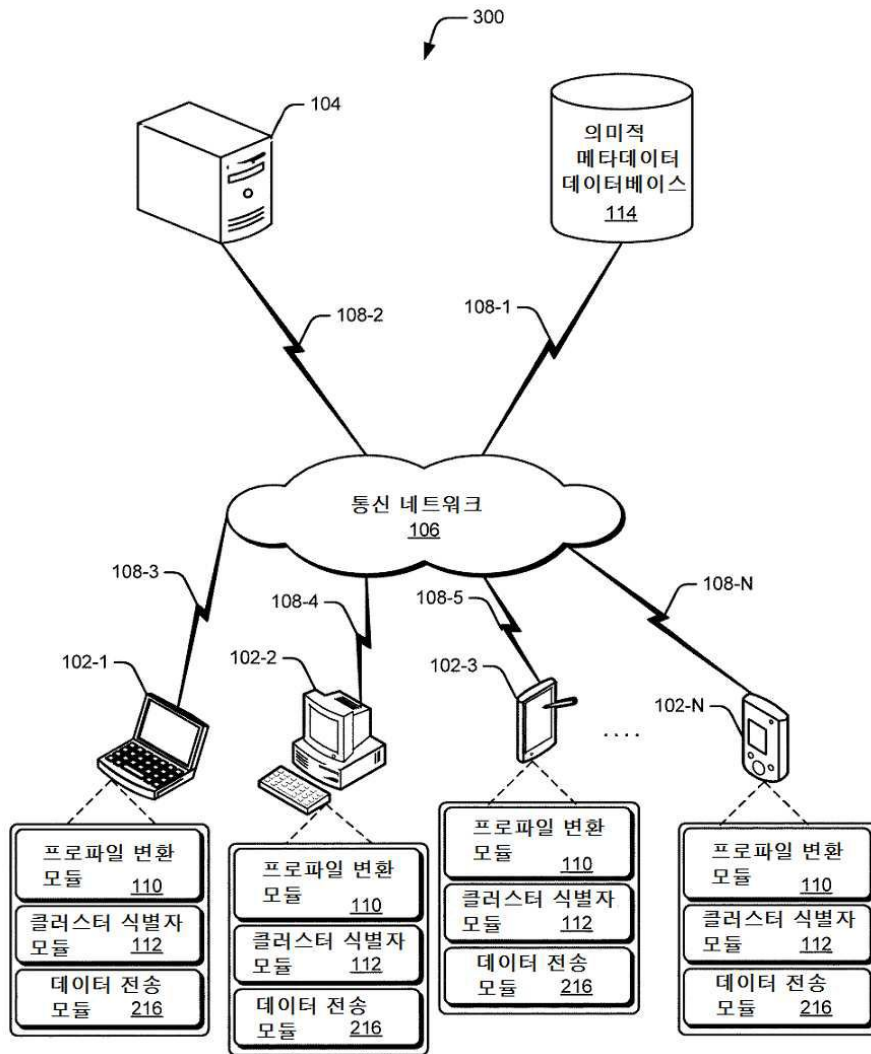
**도면1**



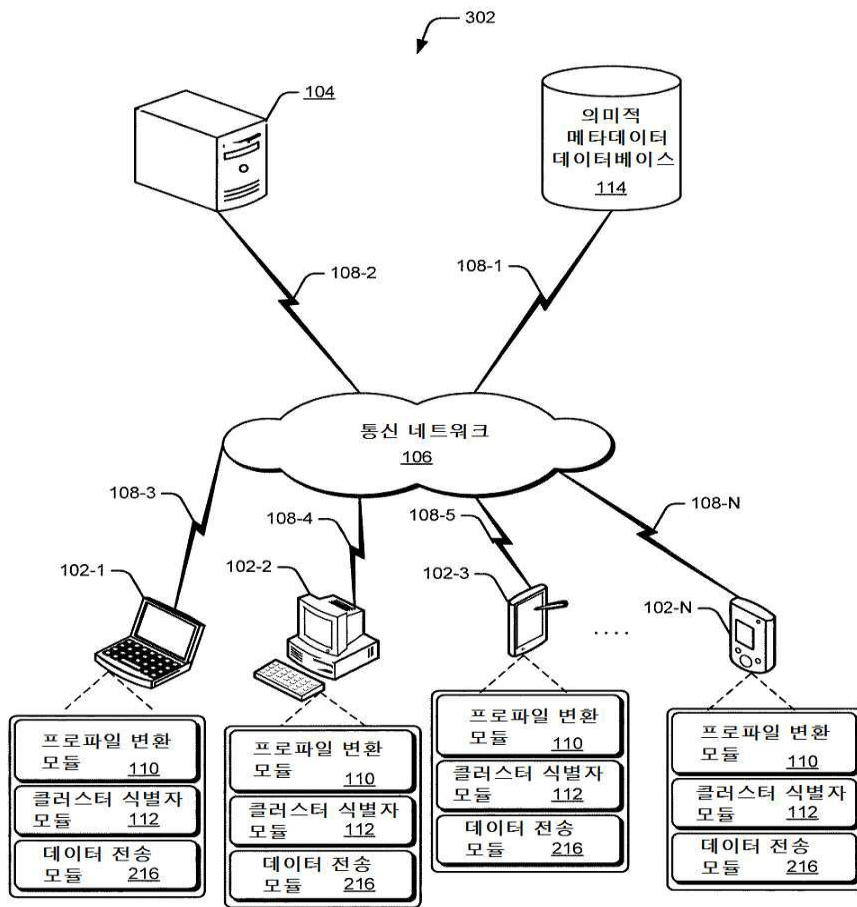
도면2



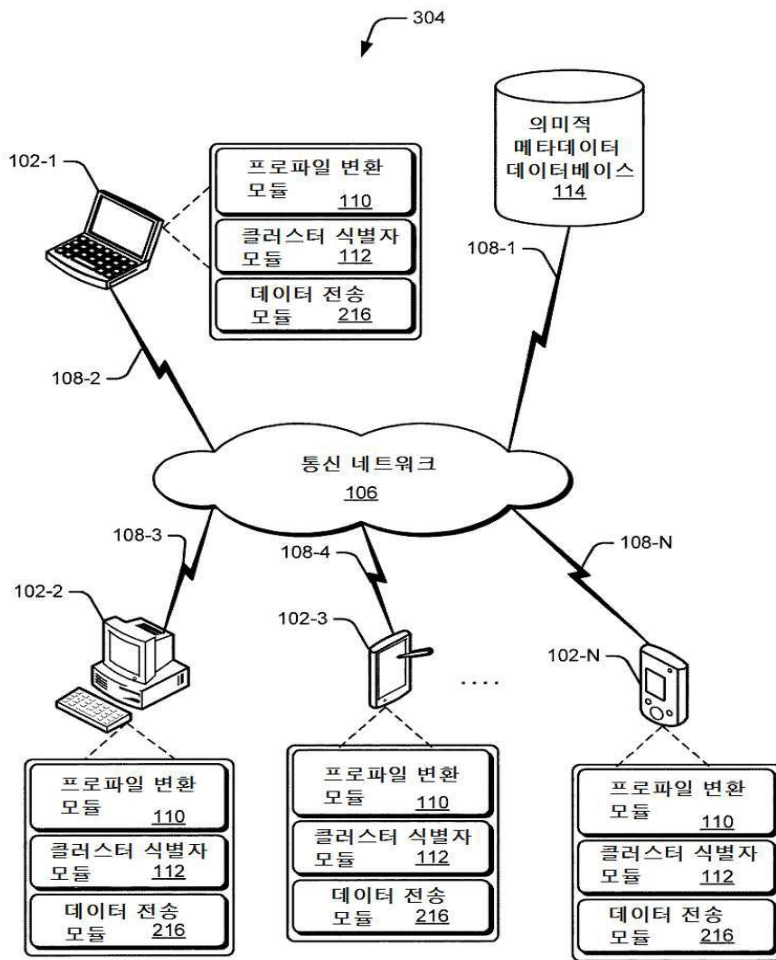
도면3a



도면3b



도면3c



도면4

