

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第5997379号
(P5997379)

(45) 発行日 平成28年9月28日(2016.9.28)

(24) 登録日 平成28年9月2日(2016.9.2)

(51) Int.Cl.

F 1

G06F 17/30 (2006.01)
G06F 21/62 (2013.01)G06F 17/30 120A
G06F 17/30 340B
G06F 17/30 412
G06F 21/62 354

請求項の数 15 (全 26 頁)

(21) 出願番号 特願2015-521987 (P2015-521987)
 (86) (22) 出願日 平成25年7月3日(2013.7.3)
 (65) 公表番号 特表2015-532737 (P2015-532737A)
 (43) 公表日 平成27年11月12日(2015.11.12)
 (86) 國際出願番号 PCT/EP2013/001958
 (87) 國際公開番号 WO2014/012627
 (87) 國際公開日 平成26年1月23日(2014.1.23)
 審査請求日 平成27年3月9日(2015.3.9)
 (31) 優先権主張番号 12290234.9
 (32) 優先日 平成24年7月16日(2012.7.16)
 (33) 優先権主張国 歐州特許庁(EP)

(73) 特許権者 391030332
 アルカテルルーセント
 フランス国、92100・プローニューピ
 ャンクール、ルート・ドゥ・ラ・レーヌ・
 148/152
 (74) 代理人 100094112
 弁理士 岡部 譲
 (74) 代理人 100106183
 弁理士 吉澤 弘司
 (74) 代理人 100170601
 弁理士 川崎 幸
 (74) 代理人 100187964
 弁理士 新井 剛

最終頁に続く

(54) 【発明の名称】ユーザ関心プロファイルのプライバシ保護されたクラスタ化のシステムおよび方法

(57) 【特許請求の範囲】

【請求項 1】

エンド・ユーザの関心プロファイルに基づいて少なくとも1つの関心プロファイル・セグメントを生成する、ユーザ・デバイスによるステップと、

前記エンド・ユーザのユーザ・デバイスに通信可能に結合されたセマンティック・メタデータ・データベースとの相互作用に基づいて、前記少なくとも1つの関心プロファイル・セグメントに対応する意味論的用語を取得するステップであって、前記意味論的用語は、前記セマンティック・メタデータ・データベース内で提供される1つまたは複数の意味論的用語の中から取得される、前記ユーザ・デバイスによるステップと、

前記意味論的用語に基づいて、前記少なくとも1つの関心プロファイル・セグメントのそれを少なくとも1つのセマンティック表現に変換する、前記ユーザ・デバイスによるステップと、

前記少なくとも1つのセマンティック表現に基づいて前記少なくとも1つの関心プロファイル・セグメントのそれぞれにクラスタ識別子を割り当てるステップであって、前記クラスタ識別子は、局所性鋭敏型ハッシュ技法を使用して生成される、前記ユーザ・デバイスによるステップと

を含む、エンド・ユーザの関心プロファイルのプライバシ保護されたクラスタ化の方法。

【請求項 2】

前記少なくとも1つの関心プロファイル・セグメント、および前記少なくとも1つの関

10

20

心プロファイル・セグメントのそれぞれに対応する前記クラスタ識別子をリモート・ノードに提供するステップをさらに含み、前記リモート・ノードは、前記クラスタ識別子に関連する関心グループの関心グループ・アグリゲータとして働くように構成され、前記クラスタ識別子は、匿名通信チャネルを介して提供される、請求項1に記載の方法。

【請求項3】

リモート・ノードとの複数の匿名通信チャネルを確立する、前記ユーザ・デバイスによるステップと、

前記関心プロファイルを1つまたは複数のプロファイル・スライスにスライス化するステップと、

前記クラスタ識別子、および前記1つまたは複数のプロファイル・スライスのそれぞれを前記複数の匿名通信チャネルを介して前記リモート・ノードに送信するステップであつて、前記1つまたは複数のプロファイル・スライスのそれぞれは、前記複数の匿名通信チャネルのうちの少なくとも1つを介して別々に送信される、前記ユーザ・デバイスによるステップと

をさらに含む、請求項1に記載の方法。

【請求項4】

前記エンド・ユーザの個人的に識別可能な情報を前記関心プロファイルから除去するステップをさらに含む、請求項3に記載の方法。

【請求項5】

複数の匿名通信チャネルを介して前記クラスタ識別子を送信する、前記ユーザ・デバイスによるステップと、

前記複数の匿名通信チャネルを介して他のユーザ・デバイスのクラスタ識別子を受信する、前記ユーザ・デバイスによるステップと、

送信する前記ステップおよび受信する前記ステップに基づいて、前記エンド・ユーザの前記クラスタ識別子に類似するクラスタ識別子を有する前記他のユーザ・デバイスを識別するステップと、

前記エンド・ユーザの前記ユーザ・デバイスと前記他のユーザ・デバイスとの間での前記複数の匿名通信チャネルのうちの少なくとも1つを介するデータの交換を可能にするステップと

をさらに含む、請求項1に記載の方法。

【請求項6】

割り当てる前記ステップは、

前記少なくとも1つの関心プロファイル・セグメントのそれぞれとハッシュ関数の集合との間の内積を取得するステップと、

前記少なくとも1つの関心プロファイル・セグメントに対応する前記クラスタ識別子としてハッシュ符号を取得するために前記内積の符号を連結するステップと

をさらに含む、請求項1乃至5のいずれか1項に記載の方法。

【請求項7】

変換する前記ステップは、前記少なくとも1つのセマンティック表現を取得するために前記意味論的用語に基づいて前記少なくとも1つの関心プロファイル・セグメントのそれぞれにタグ付けするステップをさらに含む、請求項1乃至6のいずれか1項に記載の方法。

。

【請求項8】

プロセッサと、

前記プロセッサに結合されたメモリとを含み、前記メモリは、

ユーザ・デバイスに結合されたセマンティック・メタデータ・データベースとの相互作用に基づいて、少なくとも1つの関心プロファイル・セグメントに対応する意味論的用語を取得することであつて、前記意味論的用語は、前記セマンティック・メタデータ・データベース内で提供される1つまたは複数の意味論的用語の中から取得される、取得することと、

10

20

30

40

50

前記少なくとも 1 つの関心プロファイル・セグメントのそれを少なくとも 1 つのセマンティック表現に変換することと

を行うように構成されたプロファイル変換モジュールと、

前記少なくとも 1 つのセマンティック表現に基づいて前記少なくとも 1 つの関心プロファイル・セグメントのそれぞれにクラスタ識別子を割り当てるよう構成されたクラスタ識別子モジュールであって、前記クラスタ識別子は、局所性鋭敏型ハッシュ技術を使用して生成される、クラスタ識別子モジュールと

を含む、ユーザ・デバイス。

【請求項 9】

前記ユーザ・デバイスに関連するエンド・ユーザのプロファイル・データに基づいて、前記エンド・ユーザの関心プロファイルを生成するように構成された関心プロファイル生成モジュールと、

前記関心プロファイルを前記少なくとも 1 つの関心プロファイル・セグメントにセグメント化するように構成されたプロファイル・セグメント化モジュールと

をさらに含む、請求項 8 に記載のユーザ・デバイス。

【請求項 10】

前記クラスタ識別子に関連する関心グループの関心グループ・アグリゲータとして働くように構成されたリモート・ノードとの匿名通信チャネルを確立することと、

前記少なくとも 1 つの関心プロファイル・セグメント、および前記少なくとも 1 つの関心プロファイル・セグメントのそれぞれに対応する前記クラスタ識別子を前記リモート・ノードに提供することであって、前記クラスタ識別子は、前記匿名通信チャネルを介して提供される、提供することと

を行うように構成されたデータ転送モジュールをさらに含む、請求項 8 または 9 に記載のユーザ・デバイス。

【請求項 11】

リモート・ノードとの複数の匿名通信チャネルを確立し、

前記関心プロファイルから前記エンド・ユーザの個人的に識別可能な情報を除去し、前記エンド・ユーザの前記クラスタ識別子および前記関心プロファイルを前記複数の匿名通信チャネルを介して前記リモート・ノードに送信する

ように構成されたデータ転送モジュールをさらに含む、請求項 8 または 9 に記載のユーザ・デバイス。

【請求項 12】

前記関心プロファイルを 1 つまたは複数のプロファイル・スライスにスライス化することと、

前記クラスタ識別子、および前記 1 つまたは複数のプロファイル・スライスのそれぞれを複数の匿名通信チャネルを介してリモート・ノードに送信することであって、前記 1 つまたは複数のプロファイル・スライスのそれぞれは、別々の匿名通信チャネルを介して送信される、送信することと

を行うように構成されたデータ転送モジュールをさらに含む、請求項 8 または 9 に記載のユーザ・デバイス。

【請求項 13】

複数の匿名通信チャネルを介して前記クラスタ識別子を送信し、

前記複数の匿名通信チャネルを介して他のユーザ・デバイスのクラスタ識別子を受信し、

前記送信および前記受信に基づいて前記エンド・ユーザの前記クラスタ識別子に類似するクラスタ識別子を有する前記他のユーザ・デバイスを識別し、

前記匿名通信チャネルを介する前記他のユーザ・デバイスとのデータの交換を可能にする

ように構成されたデータ転送モジュールをさらに含む、請求項 8 または 9 に記載のユーザ・デバイス。

10

20

30

40

50

【請求項 14】

前記ユーザ・デバイスのエンド・ユーザの関心グループの中央アグリゲータ・ノードの少なくとも部分的イメージを取得し、

複数の匿名通信チャネルを介して前記クラスタ識別子を他のユーザ・デバイスと交換し、

前記複数の匿名通信チャネルを介して受信された前記他のユーザ・デバイスのクラスタ識別子に基づいて、前記エンド・ユーザの前記クラスタ識別子に類似するクラスタ識別子を有する前記他のユーザ・デバイスを識別し、

前記クラスタ識別子に基づいて識別された前記他のユーザ・デバイスに、前記匿名通信チャネルを介して推奨サービスを提供する

ように構成されたデータ転送モジュールをさらに含む、請求項 8 または 9 に記載のユーザ・デバイス。

【請求項 15】

エンド・ユーザの関心プロファイルに基づいて少なくとも 1 つの関心プロファイル・セグメントを生成するステップと、

前記エンド・ユーザのユーザ・デバイスに結合されたセマンティック・メタデータ・データベースとの相互作用に基づいて、前記少なくとも 1 つの関心プロファイル・セグメントに対応する意味論的用語を取得するステップであって、前記意味論的用語は、前記セマンティック・メタデータ・データベース内で提供される 1 つまたは複数の意味論的用語の中から取得される、ステップと、

前記少なくとも 1 つの関心プロファイル・セグメントのそれぞれを少なくとも 1 つのセマンティック表現に変換するステップと、

前記少なくとも 1 つのセマンティック表現に基づいて前記少なくとも 1 つの関心プロファイル・セグメントのそれぞれにクラスタ識別子を割り当てるステップであって、前記クラスタ識別子は、局所性鋭敏型ハッシュ技法を使用して生成される、ステップと

を含む、エンド・ユーザの関心プロファイルのプライバシ保護されたクラスタ化の方法を実行するコンピュータ・プログラムを記憶したコンピュータ可読記憶媒体。

【発明の詳細な説明】

【技術分野】

【0001】

本主題は、パーソナライズされたサービスおよびデータ分析に関し、詳細には、それだけに限定されないが、パーソナライズされたサービスおよびデータ分析のためのユーザ関心プロファイル (user interest profile) をクラスタ化しながらエンド・ユーザのプライバシを保護することに関する。

【背景技術】

【0002】

ワールド・ワイド・ウェブを介して膨大な量のコンテンツを取得できるため、サービス・プロバイダによって提供されるコンテンツにアクセスするエンド・ユーザには、しばしば、サービス・プロバイダ、検索エンジン、ウェブ・パブリッシャ、および広告主によって、関連するコンテンツに簡単にアクセスするためのパーソナライズされたアシスタンスおよびデータ分析が提供される。通常、コンテンツベースの推奨、協調推奨、およびデータ分析などのさまざまな技法が、パーソナライズされたサービスをエンド・ユーザに提供するのに使用される。コンテンツベースの推奨では、エンド・ユーザは、過去にそのエンド・ユーザが使用し、もしくは好み、またはそのエンド・ユーザの関心または選択と一致する、コンテンツ、サービス、または製品に類似するコンテンツ、サービス、または製品を推奨される。協調推奨では、エンド・ユーザは、類似するまたは同一の関心または選択を有する他のエンド・ユーザが使用したもしくは好んだコンテンツ、サービス、または製品に類似するコンテンツ、サービス、または製品を推奨される。

【0003】

コンテンツベースの推奨の例では、映画レビュー・ウェブサイトは、エンド・ユーザが

10

20

30

40

50

特定のカテゴリの映画、たとえばアニメーション映画を定期的に見ることを監視することができる。したがって、アニメーション映画が見るために使用可能になるたびに、エンド・ユーザに、関連する支払を行うことによって映画をダウンロードするための、たとえば通知またはアラートなどの推奨を提供することができる。同様に、協調フィルタリングとも称する協調推奨では、サービス・プロバイダは、的が絞られた広告をエンド・ユーザに提供することができ、これらの広告は、エンド・ユーザに類似する関心およびプリファレンスを有する他のエンド・ユーザによって好まれた製品またはサービスに関する。たとえば、インターネット・プロトコル・テレビジョン（IPTV）サービス・プロバイダは、テレビ番組または映画がエンド・ユーザの関心と一致する関心を有する他のエンド・ユーザによって見られた場合に、そのテレビ番組または映画をエンド・ユーザに推奨することができる。

【0004】

さらに、ユーザは、たとえばデータ・マイニングのために、サービス・プロバイダがデータ分析アプリケーションを実行するために、彼らの個人データを提供する必要がある。サービス・プロバイダによるデータ分析とエンド・ユーザへの関連するコンテンツの提供とに関するこれらの技法のすべての中で、現在のユーザは、彼らの個人情報および潜在的に機密の情報に心配をますます感じている。これは主に、パーソナライズされたサービスから利益を得、データ分析を容易にするために、エンド・ユーザが、機密情報を明らかにしなければならないが、それと同時に、エンド・ユーザが情報のプライバシ保護について心配しているからである。たとえば、ソーシャル・ネットワーキング・サイトのエンド・ユーザは、ソーシャル・ネットワーキング・サイト上での更新に関する他のユーザおよび彼自身への匿名推奨を行うために、ソーシャル・ネットワーキング・サイトへの彼のアクセスの情報を使用することには異議を唱えない可能性があるが、他のエンド・ユーザ、サービス・プロバイダ、攻撃者、および悪意のある当事者などの他のエンティティが、エンド・ユーザが訪問するかーティングした特定のURLを知ることを望まない可能性がある。

【発明の概要】

【課題を解決するための手段】

【0005】

この課題を解決するための手段は、パーソナライズされたサービスを提供し、データ分析アプリケーションを実行するためのエンド・ユーザのクラスタ化しながらプライバシを保護することに関する概念を導入するために提供される。この課題を解決するための手段は、請求される主題の本質的特徴を識別することを意図されたものではなく、請求される主題の範囲を区切りまたは限定する際の使用を意図されたものでもない。

一実装形態では、ユーザ関心プロファイルのプライバシ保護されたクラスタ化の方法が説明される。この方法は、エンド・ユーザの関心プロファイルに基づいて少なくとも1つの関心プロファイル・セグメントを生成するステップを含む。この方法は、エンド・ユーザのユーザ・デバイスに結合されたセマンティック・メタデータ・データベースとの相互作用に基づいて、少なくとも1つの関心プロファイル・セグメントに対応する意味論的用語を取得するステップであって、意味論的用語は、セマンティック・メタデータ・データベース内で提供される1つまたは複数の意味論的用語の中から取得される、ステップをさらに含む。その後、少なくとも1つの関心プロファイル・セグメントのそれぞれが、少なくとも1つのセマンティック表現に変換される。この方法は、少なくとも1つのセマンティック表現に基づいて少なくとも1つの関心プロファイル・セグメントにクラスタ識別子を割り当てるステップであって、クラスタ識別子は、局所性鋭敏型ハッシュ（LSH）技法を使用して生成される、ステップをさらに含む。

【0006】

もう1つの実装形態では、ユーザ関心プロファイルのプライバシ保護されたクラスタ化のためのユーザ・デバイスが説明される。前記実装形態では、ユーザ・デバイスは、プロセッサと、プロセッサに結合されたメモリとを含む。メモリは、ユーザ・デバイスに結合

10

20

30

40

50

されたセマンティック・メタデータ・データベースとの相互作用に基づいて、少なくとも1つの関心プロファイル・セグメントに対応する意味論的用語を取得するように構成されたプロファイル変換モジュールを含む。一実施形態では、意味論的用語は、セマンティック・メタデータ・データベース内で提供される1つまたは複数の意味論的用語の中から取得される。プロファイル変換モジュールは、少なくとも1つの関心プロファイル・セグメントのそれぞれを少なくとも1つのセマンティック表現に変換するようにさらに構成される。メモリは、少なくとも1つのセマンティック表現に基づいて、局所性鋭敏型ハッシュ（L S H）を使用してクラスタ識別子を生成し、少なくとも1つのセマンティック表現に基づいて、少なくとも1つの関心プロファイル・セグメントにクラスタ識別子を割り当てるように構成されたクラスタ識別子モジュールをさらに含む。

10

【0007】

もう1つの実装形態では、方法を実行するコンピュータ可読プログラム・コードがその上で実施されたコンピュータ可読媒体が説明される。この方法は、エンド・ユーザの関心プロファイルに基づいて少なくとも1つの関心プロファイル・セグメントを生成するステップを含むことができる。この方法は、エンド・ユーザのユーザ・デバイスに結合されたセマンティック・メタデータ・データベースとの相互作用に基づいて、少なくとも1つの関心プロファイル・セグメントに対応する意味論的用語を取得するステップであって、意味論的用語は、セマンティック・メタデータ・データベース内で提供される1つまたは複数の意味論的用語の中から取得される、ステップをさらに含む。その後、少なくとも1つの関心プロファイル・セグメントのそれぞれは、少なくとも1つのセマンティック表現に変換される。この方法は、少なくとも1つのセマンティック表現に基づいて少なくとも1つの関心プロファイル・セグメントにクラスタ識別子を割り当てるステップであって、クラスタ識別子は、局所性鋭敏型ハッシュ（L S H）技法を使用して生成される、ステップをさらに含む。

20

【0008】

詳細な説明は、添付図面を参照して説明される。図面では、符号の左端の枠が、その符号がはじめて現れる図面を識別する。同一の符号が、複数の図面を通じて、同様の特徴および構成要素を参照するのに使用される。本主題の実施形態によるシステムおよび／または方法のいくつかの実施形態を、これから、例としてのみ、添付図面を参照して説明する。

30

【図面の簡単な説明】

【0009】

【図1】本主題の実施形態による、個人情報のプライバシ保護のためのユーザ・プロファイルのプライバシ保護されたクラスタ化の例示的なネットワーク環境実装形態を示す図である。

【図2】本主題の実施形態による、個人情報のプライバシ保護のためのユーザ・プロファイルのプライバシ保護されたクラスタ化を実施するユーザ・デバイスのコンポーネントを示す図である。

【図3（a）】本主題の実施形態による、ユーザ・プロファイルのプライバシ保護されたクラスタ化を実施するさまざまなアプリケーションの例示的なネットワーク環境実装形態を示す図である。

40

【図3（b）】本主題の実施形態による、ユーザ・プロファイルのプライバシ保護されたクラスタ化を実施するさまざまなアプリケーションの例示的なネットワーク環境実装形態を示す図である。

【図3（c）】本主題の実施形態による、ユーザ・プロファイルのプライバシ保護されたクラスタ化を実施するさまざまなアプリケーションの例示的なネットワーク環境実装形態を示す図である。

【図4】本主題の実施形態による、エンド・ユーザにプライバシ保護を提供するためのユーザ・プロファイルのプライバシ保護されたクラスタ化の方法を示す図である。

【発明を実施するための形態】

50

【0010】

本文書では、単語「例示的な」は、本明細書で「例、実例、または例示として働く」を意味するのに使用される。本明細書で「例示的な」として説明される本主題の任意の実施形態または実装形態は、必ずしも、他の実施形態より好ましいまたは有利と解釈されるべきものではない。

【0011】

当業者は、本明細書の任意のブロック図が本主題の原理を実施する例示的システムの概念図を表すことを了解するに違いない。同様に、任意のフロー・チャート、流れ図、状態遷移図、擬似コード、および類似物が、コンピュータ可読媒体内で実質的に表され、したがってコンピュータまたはプロセッサが明示的に図示されているかどうかに拘わらずそのようなコンピュータまたはプロセッサによって実行され得る、さまざまなプロセスを表すことを了解されたい。

10

【0012】

エンド・ユーザの個人情報に対するプライバシを提供するシステムおよび方法を説明する。このシステムおよび方法を、サービス・プロバイダによって提供されるサービスを役立てるかコンテンツを見るためにエンド・ユーザによって使用されるさまざまなユーザ・デバイス内で実施することができる。さらに、ユーザ・デバイスは、さまざまなネットワークを介して分散サーバおよび他のユーザ・デバイスと通信しているものとすることができます。説明される方法を実施できるユーザ・デバイスおよび分散サーバは、ラップトップ、デスクトップ・コンピュータ、ノートブック、携帯電話機、携帯情報端末、ワープステーション、メインフレーム・コンピュータ、セット・トップ・ボックス、メディア・プレイヤ、中央ディレクトリ・サーバ、データベース・サーバ、ファイル・サーバ、印刷サーバ、ウェブ・サーバ、アプリケーション・サーバ、および類似物などのコンピューティング・デバイスを含むが、これに限定されない。本明細書の説明は、推奨サービスを提供し、データ分析アプリケーションを実行するパーソナル・コンピュータおよび分散サーバを参照するが、当業者が理解するように、方法およびシステムを、エンド・ユーザのプライバシを保護するためにエンド・ユーザに接続性を提供する他のサーバおよびコンピューティング・システム内で実施することができる。

20

【0013】

推奨サービスおよびパーソナライズされたコンテンツが、さまざまなネットワークを介してサービス・プロバイダによって異なるエンド・ユーザに提供されるが、本明細書で説明される方法およびシステムは、アクセス独立であり、Global System for Mobile (GSM)、広帯域符号分割多元接続 (W-CDMA)、符号分割多元接続 (CDMA)、無線ローカル・エリア・ネットワーク (WLAN)、有線、ローカル・エリア・ネットワーク (LAN)、および接続性を提供する他のネットワークを含む複数のアクセス・タイプをサポートする。

30

【0014】

従来、多くのサービス・プロバイダは、エンド・ユーザの個人的な詳細、プリファレンス、および選択に基づいて、ビデオ、オーディオ、ニュース、広告などのコンテンツを提供するサービスなど、エンド・ユーザに提供されるサービスをパーソナライズすることを試みた。このために、サービス・プロバイダは、エンド・ユーザの年齢、性別、および位置などのエンド・ユーザの個人情報、エンド・ユーザの過去のアクション、またはエンド・ユーザの関心に類似する関心を有すると識別された他のエンド・ユーザが好んだサービス、コンテンツ、または製品を推奨し、データ分析アプリケーションを実行するのに、さまざまな技法を使用する。

40

【0015】

たとえば、エンド・ユーザ、たとえばユーザAが、特定の著者によって執筆された書籍を購入した場合に、サービス・プロバイダは、同一の著者によって執筆された他の書籍または同一のもしくは関連する主題に関する他の書籍などを購入するようにユーザAに提案することができる。類似するが別の従来の手法では、サービス・プロバイダは、エンド・

50

ユーザAの関心プロファイルに類似する関心プロファイルを有する他のエンド・ユーザによって好まれたコンテンツまたは製品をエンド・ユーザAに推奨することができる。パーソナライズされたサービスの提供とデータ分析とのために、さまざまな従来既知の方法が、類似する関心を有するエンド・ユーザを1つまたは複数の関心グループにクラスタ化するのに使用される。

【0016】

サービス・プロバイダによって実施されるクラスタ化の従来の方法は、エンド・ユーザの個人的なプリファレンス、選択などに関する情報の収集を必要とする。従来、サービス・プロバイダは、ログ・ファイル、アプリケーション履歴ファイル、またはエンド・ユーザのクライアント・デバイス上で保存されるかサービス・プロバイダに登録している間にユーザによって提供される他の個人的に識別可能な情報を分析することによるなど、さまざまな手段を介してエンド・ユーザに関する情報を監視し、収集する。別の従来の技法では、サービス・プロバイダは、エンド・ユーザに関する情報を収集するために、ハイパーテキスト転送プロトコル(h t t p) クッキーなどのテキスト・ファイルを保存することができる。たとえば、ウェブ・ポータルは、フォント・サイズ、ディスプレイ・ウィジェットの配置、その他など、エンド・ユーザのプリファレンスを格納するためにエンド・ユーザのウェブ・ブラウザのh t t p クッキーを含むユーザ・プリファレンス・データを保存することができる。さらに、ユーザ・プリファレンス・データは、エンド・ユーザのブラウジング詳細を格納し、これをウェブ・ポータルに送信することもできる。

【0017】

したがって、エンド・ユーザの個人的な選択に基づいて推奨サービス、パーソナライズされたコンテンツ、サービス、または製品をエンド・ユーザに提供することを試みて、サービス・プロバイダは、しばしば、エンド・ユーザのアクティビティに関する情報を監視し、収集する。ある種の状況で、サービス・プロバイダによって収集された情報に基づいてエンド・ユーザを識別することが可能になる。これは、エンド・ユーザの個人情報または機密情報を危険にさらすことをもたらし、エンド・ユーザを潜在的なプライバシ侵害にさらすか、エンド・ユーザを広告主またはスパマなどの標的にする可能性がある。さらに、極端な場合に、エンド・ユーザが、個人情報窃盗およびクレジット・カード詐欺などのさまざまな犯罪の被害者になる場合がある。

【0018】

さらに、今日、エンド・ユーザが、高速通信サポートを介して仮想ソーシャル・ネットワークを通じてお互いに接続される時に、エンド・ユーザは、S t u m b l e U p o n (登録商標)、G o o g l e N e w s (商標)、F o u r s q u a r e (登録商標)、F a c e b o o k (登録商標)、Y a h o o ! (登録商標)、およびt w i t t e r (商標)などの広いクラスのアプリケーションにプライベート情報および個人情報を常に共有している。明らかにされる機密の個人情報の性質は、ロケーション・ベースのサービス(L B S)の位置およびu n i f o r m r e s o u r c e l o c a t o r (U R L)リコメンダ・システムのブラウジング履歴など、アプリケーションのタイプに依存するが、オンライン・ソーシャル・ネットワーク(O S N)のすべてを包含するフレームワークへのこれらのアプリケーションの注入(O S Nは、推奨サービスを提供する「ワンストップショッピング(あらゆる商品を備えた総合店舗)」として働き始めている)は、ユーザの個人情報をこれまで以上に高いリスクにさらしている。

【0019】

最近、それでもデータ分析アプリケーションを実行し、推奨サービス、パーソナライズされたコンテンツ、パーソナライズされた検索/照会、ソーシャル・ネットワーク、データ・マイニング、および他の関連するサービスなどのパーソナライズされたサービスを提供しながらプライバシを保つ、異なる手法が提案された。

【0020】

そのような既存の手法の1つは、準同形暗号ベースの戦略を使用して関心プロファイルをクラスタ化することを含む。そのような手法では、ユーザの関心プロファイルは、ユー

10

20

30

40

50

ザにプライバシを提供するために、準同形暗号などの暗号技法を使用して暗号化される。そのような暗号技法は、分散計算セッティングを使用して関心プロファイルをクラスタ化するための加算および乗算などのプロトコル原始的操の実行を可能にする。分散計算セッティングの使用は、エンド・ユーザの静的集合のユーザ・デバイスを介するプロトコル原始的操の複数のフェーズの実行を可能にする。たとえば、K平均クラスタ化アルゴリズムを、準同形暗号の技法を使用するプライバシ保護方法で実施することができる。準同形暗号を使用する技法は、エンド・ユーザへのプライバシの提供を容易にするが、そのような技法は、通常、リソース集中的であり、実行が複雑である。さらに、エンド・ユーザの静的集合のユーザ・デバイスを介するプロトコル原始的操の複数のフェーズの実行は、エンド・ユーザの集合が動的に変化し、エンド・ユーザがオンライン・ポータルに常に参加し、離脱するオンライン・ポータルなどの動的環境でのそのような技法の適用を制限する。

【0021】

さらに、ある種の手法は、関心プロファイルをクラスタ化するように構成された集中化された媒介物がエンド・ユーザによって信頼されない場合の集中化された解決策を提供する。そのような状況では、各エンド・ユーザは、当初に、集中化された媒介物に通信される前に投影された関心プロファイルを取得するための暗号化など、ランダム・プロジェクション・ベースの手法を使用して関心プロファイルを前処理する。投影された関心プロファイルは、通常、集中化された媒介物を用いて取得可能ではないランダム・プロジェクション機能を使用して関心プロファイルを変換し、したがってエンド・ユーザのデータのプライバシを保証することによって取得される。しかし、ランダム・プロジェクション機能は、効率的で正しいクラスタ化を容易にするために、すべての類似する関心プロファイルが類似する投影された関心プロファイルに変換されることを保証するために、共通のクラスタの一部になるというエンド・ユーザの意図の間で共有される。しかし、エンド・ユーザの間でのランダム・プロジェクション機能の共有は、悪意のあるエンド・ユーザが、投影された関心プロファイルからオリジナルの関心プロファイルを取得するために、集中化された媒介物にランダム・プロジェクション機能を提供することができるので、ランダム・プロジェクション機能の漏洩を可能にする可能性がある。

【0022】

プライベート情報を保護するもう1つの技法は、エンド・ユーザの代わりに関心プロファイルを送信するように構成されたプロキシ・ノードを用いるゴシップオンビハーフ(gossip-on-behalf)手法である。そのような技法は、エンド・ユーザ・アイデンティティをプロキシ・ノードから隠蔽するために暗号化された2ホップ通信を使用する、媒介物ノードを介するプロキシ・ノードへのエンド・ユーザのユーザ・プロファイルの送信を含む。しかし、プロキシ・ノードとの完全な関心プロファイルの共有は、プロキシ・ノードがある時間期間にわたって関心プロファイルを分析することによってエンド・ユーザを識別することができるので、リンクアビリティ攻撃を容易にする可能性がある。

【0023】

本主題の実装形態によれば、エンド・ユーザの個人情報にプライバシを提供するためにエンド・ユーザのユーザ関心プロファイルをクラスタ化している間にプライバシを保証するシステムおよび方法が説明される。前に説明したように、このシステムおよび方法を、通信について定義されたさまざまな異なる標準規格に従ってネットワークと通信することができるさまざまな処理デバイスおよび通信デバイス内で実施することができる。さらに、本明細書で説明されるシステムおよび方法を、異なる手段を介して提供される有線ネットワークまたは無線ネットワークのいずれかを介して接続することができる。

【0024】

本明細書で説明されるシステムおよび方法は、一方では、アプリケーションの使用を制限せずに、各エンド・ユーザおよび彼らのプライベート情報を匿名性を提供することによって、エンド・ユーザのプライベート情報をに対する真の保護を提供し、他方では、限られた計算能力を有するモバイル・デバイスなどのシン・クライアントを含む任意のコンピュ

10

20

30

40

50

ーティング・デバイスを利用してパーソナライズされたサービスを受信するかプライベート・データの類似性に従うユーザ・クラスタ化を必要とする任意の他のアプリケーションまたはサービスを使用する能力をエンド・ユーザに提供する。一実施形態では、プライバシ保護されたネットワークは、お互いまたは1つまたは複数のアグリゲータ・ノードなどの分散サーバと通信するためにユーザ・デバイスに関連する複数のエンド・ユーザを含む。

【0025】

本主題の一実施形態によれば、パーソナライズされたコンテンツ・サービスおよびパーソナライズされた推奨サービスをエンド・ユーザに提供するために、類似する関心プロファイルを有するエンド・ユーザのクラスタが作成される。そのために、エンド・ユーザのそれぞれのユーザ・デバイスは、当初に、エンド・ユーザの個人的選択およびプリファレンスなどのプロファイル・データに基づいてエンド・ユーザの関心プロファイルを生成する。たとえば、ユーザ・デバイスは、エンド・ユーザによって訪問されたすべてのURL、エンド・ユーザによって検索されたか見られたすべてのビデオ、エンド・ユーザによって聞かれ、検索された音楽、およびエンド・ユーザによって探索された他の類似する関心分野に関するデータのリストを取得することができる。

【0026】

一実施形態では、関心プロファイルを、その後に、複数の関心プロファイル・セグメントにセグメント化することができ、各セグメントは、特定の関心分野に対応する。関心分野の例は、専門的関心、レジャー・プリファレンス、および宗教プリファレンスなどのコンテキスト的プリファレンスと、音楽、ダンス、eショッピング、およびソーシャル・ネットワーキングなどの関心の領域とを含むが、これに限定されない。関心プロファイル・セグメントを、その後に、さまざまな関心グループへのクラスタ化のために処理することができる。もう1つの実施形態では、関心プロファイル全体を、さまざまな関心グループへのクラスタ化のために処理することができる。

【0027】

関心プロファイル・セグメントは、その後に、セマンティック表現と称するコンパクトで意味論的に密な表現に変換される。関心プロファイル・セグメントのセマンティック表現は、限定ではなく一例として、タグ・ベースの表現、オントロジ概念ベースの表現、およびトピック・ベースの表現を含む。たとえば、ソーシャル・ネットワーキング・サイト「x y z . c o m」のURLに対応する関心プロファイルを、セマンティック表現「ソーシャル・ネットワーキング・サイト」に変換することができる。

【0028】

のために、ユーザ・デバイスは、匿名通信チャネルを介して、関心プロファイル・セグメントを対応するセマンティック表現に変換するのに使用できる意味論的用語のリストまたは辞書を有するセマンティック・メタデータ・データベースと相互作用することができる。一実施形態では、ユーザ・デバイスは、関心プロファイル・セグメントを対応するセマンティック表現に変換するのに関心プロファイル・セグメントのそれぞれに対応する意味論的用語を識別することができるセマンティック・メタデータ・データベースに関心プロファイル・セグメントを提供することができる。もう1つの実施形態では、ユーザ・デバイスは、セマンティック・メタデータ・データベースから意味論的用語のリストを取得し、その後、対応するセマンティック表現を取得するために関心プロファイル・セグメントと一致する意味論的用語を識別することができる。もう1つの実施形態では、ユーザ・デバイス自体が、エンド・ユーザによってブラウズされたウェブ・ページから抽出された意味論的用語を有するメタデータを含むことができる。前記実施形態では、ユーザ・デバイスは、意味論的用語の語彙をお互いと共有するように構成され、その結果、異なるユーザ・デバイスによって生成されたセマンティック表現が、同一の用語を使用して表現され、したがって、比較可能になる。

【0029】

そのように取得されたセマンティック表現に、クラスタ識別子を割り当てて、エンド・

10

20

30

40

50

ユーザを関連付けることができる関心グループを識別することができる。クラスタ識別子を、局所性鋭敏型ハッシュ（L S H）の技法を使用して割り当てることができる。L S H技法は、セマンティック表現と、ユーザ・デバイスのそれぞれで生成されるランダム・ベクトルの共通シーケンスなどのランダム値によって定義されるハッシュ関数の集合とを使用して、セマンティック表現のそれぞれを対応するハッシュ符号すなわちクラスタ識別子に変換することを含む。ユーザ・デバイスでのクラスタ識別子の割当は、関心プロファイルがリモート・ノードと共有されないので、プライバシ保護を提供するのを容易にする。さらに、L S H技法の使用は、クラスタ化をハッシュ符号の照合によって簡単に実行できるので、ユーザ・プロファイルのクラスタ化での効率を高めるのを助ける。さらに、セマンティック・メタデータ・データベースの使用は、偽陰性および偽陽性の発生を減らすのを助け、したがって、クラスタ化の効率を高める。したがって、セマンティック表現を生成することによるプロファイル変換に関連するL S H技法によるクラスタ化の実行は、クラスタ化での誤りの減少と共に、プライバシ保護の提供を容易にする。

【0030】

そのように取得されたクラスタ識別子を、ユーザ・デバイスを1つまたは複数のクラスタすなわちクラスタ識別子に対応する関心グループにクラスタ化するのに使用することができる。さらに、クラスタ識別子は、1つまたは複数のリモート・ノード、たとえば、中央アグリゲータ、関心グループ・アグリゲータ、および別のユーザ・デバイスに、通信チャネル、たとえば匿名通信チャネルを介して提供される。

【0031】

一実装形態では、ユーザ・デバイスと通信するリモート・ノードを、ユーザ・デバイスをそれに使用できるアプリケーションまたは推奨ベースのサービスのタイプに依存して選択することができる。たとえば、協調フィルタリング・アプリケーションの場合に、リモート・ノードを、中央アグリゲータと対応することができ、各中央アグリゲータは、少なくとも1つのクラスタすなわち、映画、ダンス、eショッピング、ソーシャル・ネットワーキング・サイト、またはその組合せなどの関心グループに対応する。したがって、ユーザ・デバイスは、クラスタ識別子のそれぞれを、対応する関心グループ内の関心グループの関心分野の後に最高ランクのすなわちトップ・ランクの関心分野を中央アグリゲータが識別することを可能にするために、ユーザの関心プロファイルと一緒に対応する中央アグリゲータに提供することができる。たとえば、「音楽」に関する関心グループの責任を負う中央アグリゲータは、対応するクラスタ識別子を有するすべてのユーザ・デバイスの関心プロファイルを受信することができる。中央アグリゲータは、その後、たとえば大多数、たとえば90%のエンド・ユーザの関心プロファイルに存在することのゆえに、高くランクインされる他の関心分野または関心を識別するために、ユーザ・プロファイルを分析することができる。たとえば、中央アグリゲータは、最高ランクの関心として「ダンス」を識別することができ、したがって、関心分野「ダンス」の責任を負う中央アグリゲータからダンスに対応する推奨を取得することができる。したがって、中央アグリゲータは、クラスタ「音楽」に対応するクラスタ内に存在するすべてのユーザ・デバイスに、音楽ならびにダンスに関するコンテンツおよび推奨を提供することができる。したがって、中央アグリゲータは、ユーザの関心だけではなく、関心グループ内で高くランクインされている新しい関心または項目にも関するコンテンツおよび推奨を提供することができる。もう1つの実施形態では、中央アグリゲータを、関心グループの責任を負う中央アグリゲータとして働く各ユーザ・デバイスに置換することができる。

【0032】

同様に、コンテンツベースの推奨アプリケーションの場合に、リモート・ノードを、関心グループ・アグリゲータと対応することができ、各関心グループ・アグリゲータは、少なくとも1つのクラスタすなわち、映画、ダンス、eショッピング、ソーシャル・ネットワーキング・サイトなどの関心グループに対応する。ユーザ・デバイスは、その場合に、クラスタ識別子に対応する関心グループ・アグリゲータにクラスタ識別子およびセマンティック表現を送信するように構成され、これに基づいて、関心グループ・アグリゲータは、ク

10

20

30

40

50

ラスタ識別子に対応するコンテンツを提供することができる。

【0033】

さらに、ピアツーピア通信を含むアプリケーションでは、リモート・ノードを、他のユーザ・デバイスとすることができる。そのようなアプリケーションでは、ユーザ・デバイスは、匿名通信チャネルを使用してユーザ・デバイスの間でのチャットおよびデータの交換を可能にするために、類似するクラスタ識別子を有する他のユーザ・デバイスを識別するのにクラスタ識別子を使用することができる。匿名通信チャネルの使用は、クラスタ識別子および他のデータが匿名通信チャネルを形成するさまざまな媒介物ノードを使用して送信されるので、他のユーザまたはリモート・ノードが、クラスタ識別子を送信するエン

10

【0034】

したがって、本主題のシステムおよび方法は、ユーザ・デバイスでクラスタ識別子を関心プロファイルに割り当てるによって、エンド・ユーザにプライバシを提供するのを容易にする。LSH技法を使用してユーザ・デバイスでクラスタ識別子を関心プロファイルに割り当てるには、リモート・ノードがエンド・ユーザの関心プロファイルにアクセスすることなくリモート・ノードでの類似するユーザのクラスタ化を可能にし、したがって、エンド・ユーザにプライバシ保護を提供する。さらに、セマンティック表現の生成に関連してLSH技法を使用することは、セマンティック・メタデータ・データベースの使用が偽陰性および偽陽性の発生を減らすのを助け、したがってクラスタ化の効率を高めるので、プライバシが保護され、クラスタ化での誤りが減らされることを保証するのを助ける。したがって、ユーザ・デバイスでのクラスタ識別子を提供することは、サービス・プロバイダが関心プロファイルを分析する必要がなく、クラスタ識別子をコンテンツ・データと照合するのに任意のコンテンツ・マッチング技法を使用することができるので、コンテンツ・プロバイダがより高速で効率的にデータ分析を実行し、パーソナライズされたサービスを提供するのを容易にする。さらに、クラスタ識別子の使用は、サービス・プロバイダが関心プロファイルを処理したまたは分析する必要がないので、サービス・プロバイダの端でコストを下げ、リソースを節約する。

20

【0035】

この説明および図面が、単に、本主題の原理を示すことに留意されたい。したがって、当業者が、本明細書で明示的に説明されず、図示されないが、本主題の原理を実施し、その趣旨および範囲に含まれるさまざまな配置を考案できることを了解されたい。さらに、本明細書で列挙されるすべての例は、主に、当技術を促進するために本発明人によって貢献される本主題および概念の原理を理解する際に読者を助ける教育的目的のみのためであることが特に意図され、そのような具体的に列挙される例および条件への限定なしに解釈されなければならない。さらに、本主題の原理、態様、および実施形態ならびにその特定の例を列挙する本明細書のすべての言明は、その同等物を包含することが意図されている。

30

【0036】

また、当業者は、本明細書で使用される時に、単語 `during` (~中に)、`while` (~の間に)、および `when` (~時に) が、開始するアクションの際に瞬間的に行われるアクションを意味する正確な用語ではなく、初期アクションとその初期アクションによって開始される反応との間に、伝搬遅延など、短いが穩当な遅延があつてもよいことをも了解するであろう。さらに、単語「`connected` (接続された)」および「`coupled` (結合された)」は、説明の明瞭さのために終始使用され、直接接続または間接接続のいずれをも含むことができる。

40

【0037】

エンド・ユーザへのプライバシ保護のためにユーザ・プロファイルをクラスタ化するシステムおよび方法が実施される形を、図1～4に関して詳細に説明した。プライバシ保護を提供するためにユーザ・プロファイルをクラスタ化する説明されるシステムおよび方法の諸態様を、任意の個数の異なるコンピューティング・システム、伝送環境、およびノマ

50

たは構成で実施することができるが、諸実施形態は、次の例示的なシステムの文脈で説明される。

【0038】

図1に、本主題の実施形態による、データ分析とパーソナライズされたサービスへのアクセスとを容易にしながらエンド・ユーザのプライバシを保護するためのユーザ・プロファイルのプライバシ保護されたクラスタ化のネットワーク環境100実装形態を示す。ネットワーク環境100は、通信ネットワーク106を介してお互いおよびリモート・ノード104と通信する、以下では集合的にユーザ・デバイス102と称し、個別にユーザ・デバイス102と称する1つまたは複数のユーザ・デバイス102-1、102-2,...、および102-Nを含む。

10

【0039】

通信ネットワーク106は、無線ネットワークまたは有線ネットワークと無線ネットワークとの組合せとすることができます。通信ネットワーク106は、お互いに相互接続され、単一の大きいネットワーク（たとえば、インターネットまたはイントラネット）として機能する個々のネットワークの集合とすることができます。そのような個々のネットワークの例は、3rd Generation Partnership Project (3GPP)、Long Term Evolution (LTE)、Global System for Mobile Communication (GSM) ネットワーク、Universal Mobile Telecommunications System (UMTS) ネットワーク、Personal Communications Service (PCS) ネットワーク、時分割多元接続 (TDMA) ネットワーク、符号分割多元接続 (CDMA) ネットワーク、Next Generation Network (NGN)、公衆交換電話網 (PSTN)、およびサービス総合ディジタル網 (ISDN) を含むが、これに限定されない。通信ネットワーク106は、ローカル・エリア・ネットワーク (LAN)、広域ネットワーク (WAN)、インターネット、イントラネット、ピアツーピア・ネットワーク、および仮想プライベート・ネットワーク (VPN) を含む任意の公衆ネットワークまたは私有ネットワークとすることができ、ルータ、ブリッジ、サーバ、コンピューティング・デバイス、ストレージ・デバイス、その他などのさまざまなネットワーク・デバイスを含むことができる。

20

【0040】

ユーザ・デバイス102は、データ分析と、推奨サービス、パーソナライズされたコンテンツ、パーソナライズされた検索／照会、ソーシャル・ネットワーク、データ・マイニング、および他の関連するサービスなどのさまざまなパーソナライズされたサービスへのアクセスとを容易にする、お互いまたはリモート・ノード104および分散サーバなどの他のネットワーク・エンティティと通信するのにエンド・ユーザによって使用されるユーザ機器 (UE) と定義することができる。ユーザ・デバイス102の例は、限定なしに、ラップトップもしくは他のポータブル・コンピュータ、デスクトップ・コンピュータ、ノートブック、ネットワーク・コンピュータ、携帯電話機、ハンドヘルド・デバイス、携帯情報端末、ワークステーション、メインフレーム・コンピュータ、セット・トップ・ボックス、メディア・プレイヤ、中央ディレクトリ・サーバ、データベース・サーバ、ファイル・サーバ、印刷サーバ、ウェブ・サーバ、アプリケーション・サーバ、および類似物などのコンピューティング・デバイスを含むことができる。ユーザ・デバイス102のそれぞれは、ユーザ・デバイス102が結合される通信ネットワーク106によって定義される通信プロトコルに作用する。さらに、ユーザ・デバイス102は、以下で集合的に通信チャネル108と称する通信チャネル108-1、108-2、108-3、108-4、108-5...、108-nのうちの1つまたは複数を使用して通信ネットワーク106を介してお互いまたはリモート・ノード104などの他のネットワーク・エンティティと相互作用するように構成される。一実装形態では、通信チャネル108を、ユーザ・デバイス102の間およびユーザ・デバイス102とリモート・ノード104などの他のネットワーク・エンティティとの間の匿名通信をサポートするために、たとえばTORなどの

30

40

50

オニオン・ルーティング・フレームワークに基づいて構成された匿名通信チャネルとすることができる。

【0041】

リモート・ノード104を、ラップトップ・コンピュータ、デスクトップ・コンピュータ、ノートブック、ワークステーション、メインフレーム・コンピュータ、サーバ、および類似物など、さまざまなコンピューティング・デバイスとして実施することができる。リモート・ノード104は、エンティティとして図示されているが、リモート・ノード104を、各ノードをラップトップ・コンピュータ、デスクトップ・コンピュータ、ノートブック、ワークステーション、メインフレーム・コンピュータ、サーバ、および類似物などのコンピューティング・デバイスとして実施することができるネットワークを介して分散された複数の媒介物ノードを含む分散コンピューティング・システムとして実施することもできる。さらに、媒介物ノードを、通信およびデータの交換のために中間ネットワーク（この図には図示せず）を介して接続することができる。さらに、リモート・ノード104の例は、中央アグリゲータ、関心グループ・アグリゲータ、およびユーザ・デバイス102を含むが、これに限定されない。

10

【0042】

一実装形態では、ユーザ・デバイス102と通信するリモート・ノード104を、ユーザ・デバイス102をそれに使用できるアプリケーションまたは推奨ベースのサービスのタイプに依存して選択することができる。たとえば、協調フィルタリング・アプリケーションの場合に、リモート・ノード104を、中央アグリゲータとすることでき、同様に、コンテンツベースの推奨アプリケーションの場合に、リモート・ノード104を、関心グループ・アグリゲータとして実施することができる。

20

【0043】

一実施形態では、エンド・ユーザの関心プロファイルは、エンド・ユーザのアクティビティに基づいて生成され、ユーザ・デバイス102によってローカルに保存される。たとえば、エンド・ユーザの関心プロファイルを、エンド・ユーザに対応するプロファイル・データに基づいて生成することができる。プロファイル・データは、たとえば、エンド・ユーザによって訪問されたウェブサイト、エンド・ユーザによって再生されたかダウンロードされた曲またはビデオ、エンド・ユーザによって使用された製品もしくはエンド・ユーザによって役立てられたか再検討されたサービスなどを示すことができる。生成された関心プロファイルに基づいて、クライアント・デバイスは、エンド・ユーザを1つまたは複数の事前定義の関心グループに分類する。関心グループは、類似する関心および選択を共有するエンド・ユーザのグループと理解することができる。

30

【0044】

そのために、ユーザ・デバイス102は、プロファイル変換モジュール110とクラスタ識別子モジュール112とを含む。プロファイル変換モジュール110は、関心プロファイルを、セマンティック表現と称するコンパクトで意味論的に密な表現に変換するよう構成される。セマンティック表現は、限定ではなく例として、タグ・ベースの表現、オントロジ概念ベースの表現、およびトピック・ベースの表現を含む。一実施形態では、完全な関心プロファイルを、プロファイル変換モジュール110によってセマンティック表現に変換することができる。もう1つの実施形態では、以下で関心プロファイル・セグメントと称する関心プロファイルのセグメントを、セマンティック表現に変換することができる。本明細書の説明は、関心プロファイル・セグメントを参照するが、当業者によって理解されるように、この方法およびシステムを、多少の変更は伴うが、完全な関心プロファイルを変換するために実施することができる。

40

【0045】

関心プロファイルを変換するために、プロファイル変換モジュール110は、関心プロファイル・セグメントのそれぞれに関するメタデータを取得するために通信チャネル108-1を介してセマンティック・メタデータ・データベース114などのデータベースと相互作用することができる。セマンティック・メタデータ・データベース114は、関心

50

プロファイル・セグメントを対応するセマンティック表現に変換するのに使用することができる意味論的用語のリストを有する辞書と理解することができる。意味論的用語は、その広義の意味または定義に基づいてエンド・ユーザのさまざまな関心分野を識別する用語と理解することができ、したがって、各関心分野の概念に基づいてさまざまな関心プロファイルの効率的で誤りのない照合に使用することができる。関心分野の例は、専門的関心、レジャー・プリファレンス、および宗教プリファレンスなどのコンテキスト的プリファレンスと、音楽、ダンス、eショッピング、およびソーシャル・ネットワーキングなどの関心の領域とを含むが、これに限定されない。たとえば、ソーシャル・ネットワーキング・サイト「x y z . c o m」の2つの異なるURLに対応する関心プロファイルを、URLが異なるバージョンであり、したがって構文的に異なるにも関わらず、意味論的用語「ソーシャル・ネットワーキング・サイト」と照合することができる。セマンティック・メタデータ・データベース114は、ユーザ・デバイス102の外部であるものとして図示されているが、セマンティック・メタデータ・データベース114を、他の実施形態ではユーザ・デバイス102の内部とすることを理解されたい。さらに、横の場合に、セマンティック・メタデータ・データベースを、ユーザ・デバイスの間で語彙および用語を交換することによって、ユーザ・デバイスにまたがって一致させることができる。

【0046】

プロファイル変換モジュール110によってそのように生成されたセマンティック表現を、セマンティック表現のそれぞれにクラスタ識別子を割り当てるためにクラスタ識別子モジュール112に提供することができる。クラスタ識別子は、クラスタすなわちエンド・ユーザを関連付けることができるクラスタ識別子に対応する関心グループを効率的に識別するのに使用できる関心グループ識別符号と理解することができる。クラスタ識別子モジュール112を、クラスタ識別子を割り当てるためにLSHの技法を使用するように構成することができる。LSH技法を、クラスタ識別子モジュール112によって、セマンティック表現のそれぞれを対応するハッシュ符号すなわちクラスタ識別子に変換するのに使用することができる。そのために、クラスタ識別子モジュール112は、ユーザ・デバイス102のそれぞれで生成されるランダム・ベクトルの共通シーケンスなどのランダム値によって定義されるハッシュ関数の集合を利用する。

【0047】

一実装形態では、ハッシュ関数は、クラスタ識別がすべてのユーザ・デバイス102によって均一に行われることを保証するためにすべてのユーザ・デバイス102によってアクセス可能なランダム・シード・ジェネレータ（この図には図示せず）などの中央エンティティからクラスタ識別子モジュール112によって取得される。もう1つの実装形態では、クラスタ識別子モジュール112を、たとえば中央エンティティから取得されたシード生成関数に基づいてハッシュ関数を生成するように構成することができる。したがって、ユーザ・デバイス102でクラスタ識別子を割り当てるることは、関心プロファイルがリモート・ノードと共有されないので、プライバシ保護を提供するのを容易にする。したがって、セマンティック表現を生成することによるプロファイル変換に関連するLSH技法によるクラスタ化の実行は、クラスタ化での誤りの減少と共に、プライバシ保護の提供を容易にする。

【0048】

このように取得されたクラスタ識別子を、ユーザ・デバイス102すなわちエンド・ユーザのユーザ・プロファイルを1つまたは複数のクラスタにクラスタ化するのに使用することができる。さらに、クラスタ識別子モジュール112は、推奨サービスの提供またはピアツーピア通信に関するクラスタ化など、1つまたは複数のサービスに関して通信チャネル108を介して1つまたは複数のリモート・ノード104にクラスタ識別子を提供する。ユーザ・デバイス102が利用することを望む可能性があるアプリケーションまたは推奨ベースのサービスのタイプに基づいて、ユーザ・デバイス102と通信するリモート・ノード104は、中央アグリゲータおよび関心グループ・アグリゲータの中から変化することができる。代替案では、ユーザ・デバイス102は、お互いと通信することができ

10

20

30

40

50

、各ユーザ・デバイス102は、他のユーザ・デバイス102のリモート・ノードとして働く。

【0049】

たとえば、協調フィルタリング・アプリケーションの場合に、リモート・ノードを、中央アグリゲータとすることができる、各中央アグリゲータは、少なくとも1つのクラスタすなわち、映画、ダンス、eショッピング、ソーシャル・ネットワーキング・サイト、およびその組合せなどの関心グループに対応する。その場合のリモート・ノード104は、グループに対応する関心分野だけではなく、クラスタのメンバであるエンド・ユーザの他のトップ・レーティングのすなわち高くランキングされた関心分野にも関する推奨をクラスタの各メンバに提供するように構成される。もう1つの実施形態では、中央アグリゲータとして働くリモート・ノード104を、取り除くことができ、各ユーザ・デバイス102は、関心グループの責任を負う中央アグリゲータとして働く。

10

【0050】

同様に、コンテンツベースの推奨アプリケーションの場合に、リモート・ノード104を、1つまたは複数の関心グループ・アグリゲータとして実施することができ、各関心グループ・アグリゲータは、映画、ダンス、eショッピング、ソーシャル・ネットワーキング・サイトなどの少なくとも1つのクラスタに対応する。さらに、ピアツーピア通信を含むアプリケーションでは、リモート・ノード104を、もう1つの他のユーザ・デバイスとすることができます。そのようなアプリケーションでは、ユーザ・デバイス102は、類似するクラスタ識別子を有する他のユーザ・デバイス102を識別するのにクラスタ識別子を使用することができる。

20

【0051】

したがって、本主題のシステムおよび方法は、ユーザ・デバイスで関心プロファイルにクラスタ識別子を割り当てることによって、エンド・ユーザにプライバシを提供するのを容易にする。

【0052】

図2に、本主題の実施形態による、ユーザ・プロファイルのプライバシ保護されたクラスタ化を実施するユーザ・デバイス102のコンポーネントを示す。前記実施形態では、ユーザ・デバイス102は、1つまたは複数のプロセッサ202、I/Oインターフェース204、およびプロセッサ202に結合されたメモリ206を含む。プロセッサ202を、1つまたは複数のマイクロプロセッサ、マイクロコンピュータ、マイクロコントローラ、デジタル信号プロセッサ、中央処理装置、状態機械、論理回路網、および／または動作命令に基づいて信号を操作する任意のデバイスとして実施することができる。プロセッサ202を、单一の処理ユニットまたは、すべてが複数のコンピューティング・ユニットを含むこともできる複数のユニットとすることができる。他の能力の中でも、プロセッサ202は、メモリ206に格納されたコンピュータ可読命令を取り出し、実行するように構成される。

30

【0053】

「プロセッサ」としてラベルを付けられたすべての機能ブロックを含む図面に示されたさまざまな要素の機能を、専用ハードウェアならびに適当なソフトウェアに関連してソフトウェアを実行することのできるハードウェアの使用を介して提供することができる。プロセッサによって提供される時に、单一の専用プロセッサ、单一の共有されるプロセッサ、またはそのうちのいくつかを共有できる複数の個別のプロセッサによって機能を提供することができる。さらに、用語「プロセッサ」の明示的使用は、ソフトウェアを実行できるハードウェアを排他的に指すと解釈されなければならない、限定なしに、暗黙のうちにデジタル信号プロセッサ(DSP)ハードウェア、ネットワーク・プロセッサ、特定用途向け集積回路(ASIC)、フィールド・プログラマブル・ゲート・アレイ(FPGA)、ソフトウェアを格納する読み取り専用メモリ(ROM)、ランダム・アクセス・メモリ(RAM)、および不揮発性ストレージを含むことができる。通常のおよび／またはカスタムの他のハードウェアを含めることもできる。

40

50

【0054】

I/Oインターフェース204は、さまざまなソフトウェア・インターフェースおよびハードウェア・インターフェース、たとえば、キーボード、マウス、外部メモリ、プリンタ、その他などの周辺デバイスのインターフェースを含むことができる。さらに、I/Oインターフェース204は、ユーザ・デバイス102が通信ネットワーク106を介して通信することを可能にすることができる、ウェブ・サーバおよび外部データベースなどの他のコンピューティング・デバイスにユーザ・デバイス102を接続するための1つまたは複数のポートを含むことができる。I/Oインターフェース204は、LAN、ケーブル、その他などの有線ネットワークおよびWLAN、セルラ、衛星、その他などの無線ネットワークを含むネットワークなど、さまざまなプロトコルおよびネットワーク内の複数の通信を容易にすることができます。10

【0055】

メモリ206は、たとえばスタティック・ランダム・アクセス・メモリ(SRAM)およびダイナミック・ランダム・アクセス・メモリ(DRAM)などの揮発性メモリならびに/または読み取り専用メモリ(ROM)、消去可能プログラム可能ROM、フラッシュ・メモリ、ハード・ディスク、光ディスク、および磁気テープなどの不揮発性メモリを含む、当技術分野で既知の任意のコンピュータ可読媒体を含むことができる。メモリ206は、さらに、モジュール208およびデータ210を含むことができる。モジュール208は、特定のタスクを実行するか特定の抽象データ型を実施する、ルーチン、プログラム、オブジェクト、コンポーネント、データ構造などを含む。データ210は、とりわけ、モジュール208のうちの1つまたは複数によって処理され、受け取られ、生成されるデータを格納するリポジトリとして働く。20

【0056】

モジュール208は、さらに、関心プロファイル生成モジュール212、プロファイル・セグメント化モジュール214、プロファイル変換モジュール110、クラスタ識別子モジュール112、データ転送モジュール216、および他のモジュール218を含む。他のモジュール218は、ユーザ・デバイス102のアプリケーションおよび機能を補足するプログラムまたはコーディングされた命令、たとえば、オペレーティング・システム内のプログラムを含むことができる。

【0057】

データ210は、プロファイル・データ220、変換されたプロファイル・データ222、クラスタ識別子データ224、および他のデータ226を含む。他のデータ226は、他のモジュール218内の1つまたは複数のモジュールの実行の結果として生成されたデータを含むことができる。30

【0058】

前に説明したように、ユーザ・デバイス102は、ユーザ・プロファイルをクラスタ化するためにクラスタ識別子を計算することによって、エンド・ユーザにプライバシ保護を提供するように構成される。一実施形態では、関心プロファイル生成モジュール212は、プロファイル・データ220に基づいてユーザ・デバイス102のエンド・ユーザの関心プロファイルを生成するように構成される。プロファイル・データ220は、たとえば、エンド・ユーザによって訪問されたウェブサイト、エンド・ユーザによって再生されたかダウンロードされた曲またはビデオ、エンド・ユーザによって使用された製品もしくはエンド・ユーザによって役立てられたか再検討されたサービスなどを示すことができる。たとえば、関心プロファイル生成モジュール212は、エンド・ユーザによって検索されたか見られたすべてのビデオ、エンド・ユーザによって訪問されたすべてのURL、エンド・ユーザによって聞かれ、検索された音楽、およびエンド・ユーザによって探索された他の関心分野に関する類似するデータの詳細を取得することができる。そのように生成された関心プロファイルを、プロファイル・データ220に格納することができる。プロファイル変換モジュール110は、その後、クラスタ識別子を割り当てるために、関心プロファイルをセマンティック表現に変換することができる。前に議論したように、プロファ4050

イル変換モジュール 110 を、完全な関心プロファイルまたは関心プロファイル・セグメントのいずれかをセマンティック表現に変換するように構成することができる。後者の場合に、関心プロファイルは、プロファイル変換モジュール 110 によって変換される前に、プロファイル・セグメント化モジュール 214 によってアクセスされ、セグメント化される。

【0059】

一実施形態では、プロファイル変換モジュール 110 は、関心プロファイルを複数の関心プロファイル・セグメントにセグメント化するように構成され、各関心プロファイル・セグメントは、特定の関心分野に対応する。一実装形態では、プロファイル変換モジュール 110 は、関心プロファイルをセグメント化するためにエンド・ユーザの関心プロファイルを分析することができる。プロファイル変換モジュール 110 は、レジャー・プリファレンスおよび専門的関心などのコンテキスト的プリファレンスまたは音楽、ダンス、e ショッピング、およびソーシャル・ネットワーキングなどの関心の明示的な領域のいずれかに基づいて、関心プロファイルを関心プロファイル・セグメントにセグメント化することができる。このように取得された関心プロファイル・セグメントを、プロファイル変換モジュール 110 に提供することができる。

【0060】

プロファイル変換モジュール 110 は、ユーザ・デバイス 102 の内部または外部とすることができるセマンティック・メタデータ・データベース 114 内で提供されるメタデータに基づいて、関心プロファイル・セグメントまたは完全な関心プロファイルを対応するセマンティック表現に変換する。一実施形態では、プロファイル変換モジュール 110 は、関心プロファイル・セグメントをセマンティック・メタデータ・データベース 114 に提供することができる。セマンティック・メタデータ・データベース 114 は、その場合に、関心プロファイル・セグメントのそれぞれに対応する意味論的用語を識別し、これを、関心プロファイル・セグメントを対応するセマンティック表現に変換するためにプロファイル変換モジュール 110 に提供することができる。もう 1 つの実施形態では、プロファイル変換モジュール 110 は、セマンティック・メタデータ・データベース 114 から意味論的用語のリストを取得することができる。プロファイル変換モジュール 110 を、その場合に、対応するセマンティック表現を取得するために、意味論定期用語のリストから、関心プロファイル・セグメントと一致する意味論的用語を識別するように構成することができる。

【0061】

もう 1 つの実施形態では、プロファイル変換モジュール 110 は、ユーザ・デバイス 102 のエンド・ユーザによって訪問されたウェブ・ページから抽出されたタグなどのメタデータを使用して、関心プロファイル・セグメントまたは完全な関心プロファイルにタグ付けすることによって、セマンティック表現を生成することができる。そのようなローカルに抽出されたタグを、内部セマンティック・メタデータとしてユーザ・デバイス 102 の他のデータ 226 に格納することができる。前記実施形態では、ユーザ・デバイスは、意味論的用語の語彙をお互いと共有するように構成され、その結果、異なるユーザ・デバイスによって生成されたセマンティック表現が、同一の用語を使用して表現され、したがって、比較可能になる。セマンティック・メタデータ・データベースの使用は、偽陰性および偽陽性の発生を減らすのを助け、したがって、クラスタ化の効率を高める。このように取得されたセマンティック表現を、変換されたプロファイル・データ 222 内に格納し、さらなるクラスタ化のためにセマンティック表現のそれぞれにクラスタ識別子を割り当てるためにクラスタ識別子モジュール 112 によってアクセスすることができる。

【0062】

クラスタ識別子モジュール 112 は、図 1 で説明したように、エンド・ユーザのプロファイルのクラスタ化を可能にするためにセマンティック表現のそれぞれにクラスタ識別子を割り当てるように構成される。クラスタ識別子モジュール 112 は、セマンティック表現にクラスタ識別子を割り当てるために LSH 技法を使用する。当業者によって理解され

10

20

30

40

50

るよう、LSH技法では、2つのプロファイルのハッシュ衝突確率が、次式を使用して表されるように2つのプロファイルの間の類似性に比例するので、2つの類似するオブジェクトが、高い確率で同一の値にハッシュ化される。

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

ここで、 x および y は、2つのプロファイルであり、 $h()$ は、LSH関数であり、 $sim()$ は、2つのプロファイルの間の類似性尺度である。さらに、LSH技法の機能性を達成するために、類似するハッシュ関数が、LSH技法のためにユーザ・デバイス 102 によって使用されることが保証される。

【 0 0 6 3 】

クラスタ識別子モジュール 112 は、クラスタ識別子すなわち類似する関心を有するエンド・ユーザすなわち類似する関心プロファイルを有するエンド・ユーザの関心グループのラベルまたはグループ id としてハッシュ関数によって生成された値を使用するように構成される。そのために、クラスタ識別子モジュール 112 は、当初に、ランダム・シード・ジェネレータなどの中央エンティティから、ハッシュ関数とも称するランダム・ベクトルの集合を取得する。クラスタ識別子モジュール 112 は、その後、たとえば LSH 技法のコサイン類似性方法を使用する。前記方法では、クラスタ識別子モジュール 112 は、セマンティック表現とランダム・ベクトルのそれぞれとの間の内積を取得し、セマンティック表現のクラスタ識別子としてハッシュ符号を取得するために内積の符号を連結する。そのように取得されたクラスタ識別子を、クラスタ識別子データ 224 に格納することができる。本明細書での説明は、LSH 技法のコサイン類似性方法を参照するが、当業者によって理解されるように、この方法およびシステムを、LSH 技法の他の実装形態を使用して実施することができる。セマンティック表現を生成することによるプロファイル変換に関連する LSH 技法によるクラスタ化の実行は、クラスタ化での誤りの減少と共に、プライバシ保護の提供を容易にする。さらに、LSH 技法を使用してクラスタ化を実行する前にセマンティック表現を生成することは、リモート・ノード 104 がクラスタ・メンバに対する線形検索を実行できないことを保証するのを助ける。

【 0 0 6 4 】

データ転送モジュール 216 は、その後、たとえばサービス・プロバイダから推奨またはコンテンツを受信するために、通信チャネル 108 を介してリモート・ノード 104 にクラスタ識別子を送信する。一実装形態では、リモート・ノード 104 にクラスタ識別子を転送する方法は、クラスタ化が実行されるアプリケーションに依存して変化することができる。たとえば、協調フィルタリング・アプリケーションの場合に、データ転送モジュール 216 を、リモート・ノードがグループ推奨を提供するためにユーザ・プロファイルを分析することを可能にするために、クラスタ識別子だけではなく、エンド・ユーザの関心プロファイルをもリモート・ノード 104 に送信するように構成することができる。データ転送モジュール 216 を、その場合に、関心プロファイルをより小さいプロファイル・スライスにスライスし、通信チャネル 108 などの異なる匿名通信チャネルを介して独立にリモート・ノード 104 により小さいプロファイル・スライスのそれぞれを送信するように構成することができる。プロファイル・スライスは、プロファイル・セグメントとは異なって、リモート・ノードがプロファイル・スライスを分析することを可能にするために、セマンティック表現に変換されることなく送信される。匿名通信チャネルを介して異なるスライスとしてユーザ・プロファイルを送信することは、リモート・ノード 104 が関心プロファイルのスライスに基づいてエンド・ユーザを簡単に識別することができない可能性があるので、プライバシ保護を保証する。協調フィルタリング・アプリケーションのもう 1 つの実施形態では、リモート・ノード 104 を置換することができ、各ユーザ・デバイス (102) を、関心グループの責任を負う中央アグリゲータとして働くように構成することができる。

〔 0 0 6 5 〕

さらに、コンテンツベースの推薦アプリケーションの場合に、データ転送モジュール 2 16 を、サービス・プロバイダによって提供される推薦サービスまたは分析サービスを利用

用するためにリモート・ノード104にクラスタ識別子およびセマンティック表現を送信するように構成することができる。ピアツーピア通信の場合と同様に、データ転送モジュール216を、ユーザ・デバイス102の間でのチャットおよびデータの交換を可能にするために、類似するクラスタ識別子を有する他のユーザ・デバイス102を識別するように構成することができる。他のユーザ・デバイス102を識別する際に、データ転送モジュール216を、他のユーザ・デバイスの間で匿名メッセージング・チャネルおよび匿名チャット・チャネルまたは匿名検索プロードキャスト・チャネルを送信確立するように構成することができる。アプリケーションのそれぞれでのデータ転送モジュール216の働きを、図3(a)、3(b)、および3(c)でアプリケーションを説明しながらより詳細に説明する。

10

【0066】

図3(a)、3(b)、および3(c)に、それぞれ、本主題の実施形態による、ユーザ・プロファイルのプライバシ保護されたクラスタ化を実施するさまざまなアプリケーションの例示的なネットワーク環境実装形態300、302、および304を示す。

【0067】

図3(a)に、クラスタのメンバにグループ推奨を提供する協調フィルタリング・アプリケーションのユーザ・プロファイルのプライバシ保護されたクラスタ化を実施するネットワーク環境300を示す。協調フィルタリング・アプリケーションの前記実施形態では、リモート・ノード104は、映画、ダンス、eショッピング、ソーシャル・ネットワーキング・サイトなどの少なくとも1つのクラスタに対応する中央アグリゲータとして実施される。したがって、ユーザ・デバイス102は、対応する関心グループ内の関心グループの関心分野の後に最高ランキングのすなわちトップ・ランクの関心分野を中央アグリゲータが識別することを可能にするために、エンド・ユーザの関心プロファイルと一緒に対応する中央アグリゲータにクラスタ識別子のそれぞれを提供することができる。

20

【0068】

一実施形態では、データ転送モジュール216は、リモート・ノード104によって受信されるさまざまなプロファイル・セグメントの間のアンリンクアビリティ(unlinkability)を保証するために、関心プロファイルをさまざまな通信チャネル108を介して通信され得るより小さい独立のスライスにスライス化するように構成される。アンリンクアビリティの提供は、リモート・ノード104が、異なるプロファイル・セグメントが同一のユーザまたは異なるユーザに属するかどうかを識別することができないことを保証する。データ転送モジュール216は、エンド・ユーザの関心プロファイルのスライス自体が、完全な関心プロファイルを構成し、エンド・ユーザのアイデンティティを推論するのに使用できる十分なプロファイル情報を含まないことを保証する。もう1つの実施形態では、関心プロファイル全体を、単一のスライスとしてリモート・ノード104に送信することができる。

30

【0069】

さらに、関心プロファイル・スライスおよびクラスタ識別子のそれぞれは、匿名性を保証する機構を使用するネットワーク、たとえばオニオン・ルーティングを実施するネットワーク上の匿名通信チャネルを介してデータ転送モジュール216によって送信される。一実装形態では、オニオンルーティング・パスが確立され、ここで、データ転送モジュール216は、オニオンルーティング・パスの出口ノードの公開鍵を用いて、エンド・ユーザに関するプロファイル情報のスライスおよびクラスタ識別子を暗号化する。エンド・ユーザに関するプロファイル情報のさまざまなスライスおよびクラスタ識別子は、出口ノードに達する前に、1つまたは複数の中間ノードを介して送信される。出口ノードは、情報を暗号化解除し、これをリモート・ノード104に送信する。

40

【0070】

さらに、ユーザ・プライバシを保証するために、データ転送モジュール216を、エンド・ユーザのアイデンティティをリモート・ノード104に明らかにする可能性があるURLなど、個人的に識別可能な情報を有するすべてのデータを関心プロファイルから除去

50

するように構成することができる。

【0071】

エンド・ユーザの関心プロファイルおよびクラスタ識別子を受信した時に、リモート・ノード104は、リモート・ノード104に対応するクラスタ内にクラスタ化されたエンド・ユーザの最高のまたは1つもしくは複数の高くランキングされた関心を判定するために関心プロファイルを分析する。たとえば、関心グループ「ソーシャル・ネットワーキング」の責任を負うリモート・ノード104は、クラスタ「ソーシャル・ネットワーキング」に対応するクラスタ識別子を有するすべてのユーザ・デバイスの関心プロファイルを受信することができる。リモート・ノード104は、その後、高くランキングされた他の関心を識別するためにユーザ・デバイス102のユーザ・プロファイルを分析することができる。たとえば「ニュース」を最高ランキングの関心として識別する時に、リモート・ノード104は、関心分野「ニュース」の責任を負う中央アグリゲータから「ニュース」に対応する推奨を取得することができる。したがって、リモート・ノード104は、ソーシャル・ネットワーキングならびにニュースに関するコンテンツおよび推奨を、クラスタ「ソーシャル・ネットワーキング」に対応するクラスタ内に存在するすべてのユーザ・デバイス102に提供することができる。10

【0072】

協調フィルタリング・アプリケーションのもう1つの実施形態では、中央アグリゲータとして働くリモート・ノード104を除去することができ、各ユーザ・デバイス102は、関心グループの責任を負う中央アグリゲータとして働く。そのために、データ転送モジュール216は、エンド・ユーザの関心グループの中央アグリゲータ・ノードの少なくとも部分的イメージを維持するように構成される。前記実施形態では、類似する関心プロファイルを有するエンド・ユーザを、当業者に既知のように、従来のゴシップ・ベースのランダム・ピアリング・プロトコルに基づいて識別することができる。したがって、データ転送モジュール216を、類似するクラスタ識別子すなわち類似する関心グループを有するエンド・ユーザを識別するために、すべてのユーザ・デバイス102のクラスタ識別子を送信し、受信するように構成することができる。20

【0073】

図3(b)に、コンテンツベースの推奨アプリケーションに関するユーザ・プロファイルのプライバシ保護されたクラスタ化を実施するネットワーク環境302を示す。前記実施形態では、リモート・ノード104は、1つまたは複数の関心グループ・アグリゲータとして実施され、各関心グループ・アグリゲータは、映画、ダンス、eショッピング、ソーシャル・ネットワーキング・サイトなど、少なくとも1つのクラスタに対応する。30

【0074】

データ転送モジュール216は、その場合に、通信チャネル108などの匿名通信チャネルを介してリモート・ノード104にクラスタ識別子およびセマンティック表現を送信するように構成される。そのために、データ転送モジュール216は、オニオングルーティング・パスを確立し、データ転送モジュール216は、図3(a)で説明したように、エンド・ユーザに関するクラスタ識別子をオニオングルーティング・パスの出口ノードの公開鍵を用いて暗号化する。リモート・ノード104は、その場合に、周知のコンテンツ照合技法のいずれかを使用してクラスタ識別子を関心分野のリストと照合することによって、クラスタ識別子によって表される関心分野に関する推奨またはコンテンツを選択するように構成される。たとえば、エンド・ユーザ、たとえばユーザAが、関心分野「書籍」を有するクラスタに属する場合に、サービス・プロバイダは、最新のベスト・セラーのリストをユーザAに提供することができる。40

【0075】

図3(c)に、ピアツーピア通信に関するアプリケーションについてユーザ・プロファイルのプライバシ保護されたクラスタ化を実施するネットワーク環境304を示す。前記実施形態では、ユーザ・デバイス102が、プライバシ保護された環境内でお互いと相互作用するためにクラスタ化を使用するので、リモート・ノード104は除去される。その50

のようなアプリケーションでは、ユーザ・デバイス102は、類似するクラスタ識別子を有する他のユーザ・デバイス102を識別するのにクラスタ識別子を使用することができる。データ転送モジュール216は、その場合に、他のユーザ・デバイス102に通信チャネル108などの匿名通信チャネルを介してクラスタ識別子を送信するように構成される。データ転送モジュール216は、クラスタ識別子を送信し、類似するクラスタ識別子を有するユーザ・デバイス102を識別するのに、ゴシップ・ベースのランダム・ピアリング・プロトコルを使用する。共通のクラスタ識別子を有するユーザ・デバイス102は、その後、そのアイデンティティまたは個人情報を明らかにせずにユーザ・デバイスの間でのチャットおよびデータの交換を可能にするために、たとえばTORなどのオニオン・ルーティング・フレームワークを使用して、お互いに匿名で接続することができる。

10

【0076】

図4に、本主題の実施形態による、エンド・ユーザにプライバシ保護を提供するためのユーザ・プロファイルのプライバシ保護されたクラスタ化の方法を示す。この方法が説明される順序は、限定として解釈されることを意図されたものではなく、任意の個数の説明される方法ブロックを任意の順序で組み合わせて、方法400または任意の代替方法を実施することができる。さらに、本明細書で説明される主題の趣旨および範囲から逸脱せずに、個々のブロックをこの方法から削除することができる。さらに、この方法を、任意の適切なハードウェア、ソフトウェア、ファームウェア、またはその組合せで実施することができる。

【0077】

20

方法を、コンピュータ実行可能命令の全般的な文脈で説明することができる。一般に、コンピュータ実行可能命令は、特定の機能を実行するか特定の抽象データ型を実施するルーチン、プログラム、オブジェクト、コンポーネント、データ構造、手続き、モジュール、関数などを含むことができる。この方法を、機能が通信ネットワークを介してリンクされたリモート処理デバイスによって実行される分散コンピューティング環境で実践することもできる。分散コンピューティング環境では、コンピュータ実行可能命令を、メモリ・ストレージ・デバイスを含む、ローカルとリモートとの両方のコンピュータ記憶媒体内に配置することができる。

【0078】

30

当業者は、この方法のステップを、プログラムされたコンピュータによって実行できることをたやすく了解する。本明細書では、いくつかの実施形態は、プログラム記憶デバイス、たとえば、機械可読またはコンピュータ可読であり、命令の機械実行可能またはコンピュータ実行可能なプログラムを符号化するデジタル・データ記憶媒体を包含することも意図され、前記命令は、説明される方法のステップの一部またはすべてを実行する。プログラム記憶デバイスは、たとえば、デジタル・メモリ、磁気ディスクまたは磁気テープなどの磁気記憶媒体、ハード・ドライブ、または光学的に可読のデジタル・データ記憶媒体とすることができる。実施形態は、例示的な方法の前記ステップを実行するように構成された通信ネットワークと通信デバイスとの両方を包含することも意図されている。

【0079】

40

ブロック402では、エンド・ユーザに対応するプロファイル・データが、ユーザ・デバイス、たとえばユーザ・デバイス102によって取得される。一実装形態では、プロファイル・データを、エンド・ユーザによって訪問されたウェブサイト、エンド・ユーザによって再生されたかダウンロードされた曲またはビデオ、エンド・ユーザによって使用された製品もしくはエンド・ユーザによって役立てられたか再検討されたサービスなどに基づいて取得することができる。関心プロファイル生成モジュール、たとえば関心プロファイル生成モジュール212を、エンド・ユーザによって検索されたか見られたすべてのビデオ、エンド・ユーザによって訪問されたすべてのURL、エンド・ユーザによって行われたすべての照会、およびエンド・ユーザによって探索された他の関心分野に関する類似するデータの詳細を取得するように構成することができる。そのように生成された関心プロファイルを、プロファイル・データとして格納することができる。

50

【0080】

ブロック404では、プロファイル・データが、1つまたは複数の関心プロファイル・セグメントにセグメント化される。一実装形態では、プロファイル・セグメント化モジュール214などのプロファイル・セグメント化モジュールが、関心プロファイルを、関心プロファイル・セグメントと称される複数のセグメントにセグメント化するように構成される。関心プロファイル・セグメントは、各セグメントが特定の関心分野に対応する、関心プロファイルの独立のセグメントと理解することができる。関心分野の例は、専門的関心、レジャー・プリファレンス、および宗教プリファレンスなどのコンテキスト的プリファレンスと、音楽、ダンス、eショッピング、およびソーシャル・ネットワーキングなどの関心の領域とを含むが、これに限定されない。関心プロファイル・セグメントのそれぞれは、その後、クラスタ化のためにクラスタ識別子を取得するために独立に処理される。さらに、もう1つの実施形態では、関心プロファイル全体を、クラスタ化のために単一の関心プロファイル・セグメントとして提供することができる。10

【0081】

ブロック406では、関心プロファイル・セグメントのそれぞれに対応する意味論的用語が、セマンティック・メタデータ・データベースとの相互作用に基づいて突き止められる。一実装形態では、プロファイル変換モジュール110などのプロファイル変換モジュールを、匿名通信チャネルを介してセマンティック・メタデータ・データベースに関心プロファイル・セグメントを提供するように構成することができる。関心プロファイル・セグメントの受信時に、セマンティック・メタデータ・データベースは、意味論的用語のリストから、関心プロファイル・セグメントのそれぞれに対応する意味論的用語を識別する。代替案では、プロファイル変換モジュール110が、ユーザ・デバイス102のエンド・ユーザによって訪問されたウェブ・ページから抽出されたタグなどのデータを使用して投入されたユーザ・デバイス102の内部セマンティック・メタデータと相互作用することができる。前記実施形態では、ユーザ・デバイス102は、すべてのユーザ・デバイス102によって生成使用された意味論的用語にまたがって均一性を維持するために、お互いに意味論的用語の語彙を共有するように構成される。20

【0082】

ブロック408では、関心プロファイル・セグメントが、意味論的用語に基づいて対応するセマンティック表現を生成するために変換される。たとえば、プロファイル変換モジュール110などのプロファイル変換モジュールは、たとえば関心プロファイル・セグメントのそれぞれに、セマンティック・メタデータ・データベースから取得された対応するセマンティック表現を用いてタグ付けすることによって、関心プロファイル・セグメントを変換するように構成される。関心プロファイル・セグメントを、その代わりに、オントロジ概念ベースのセマンティック表現またはトピック・ベースのセマンティック表現に変換することができる。このように取得されたセマンティック表現を、変換されたプロファイル・データ222に保存することができる。30

【0083】

ブロック410では、クラスタ識別子を、LSH技法を使用してセマンティック表現に割り当てる。一実装形態では、ランダム・ベクトルの共通シーケンスなどのランダム値の集合によって定義されるハッシュ関数が、クラスタ識別子の割当に使用される。たとえば、クラスタ識別子モジュール112などのクラスタ識別モジュールが、均一なクラスタ識別を保証するためにユーザ・デバイスによってアクセス可能なランダム・シード・ジェネレータなどの中央エンティティからハッシュ関数を取得することができる。したがって、セマンティック表現の共通の集合を使用するプロファイル変換に関連するLSH技法によるクラスタ化の実行は、クラスタ化での誤りの減少と共に、プライバシ保護の提供を容易にする。40

【0084】

ブロック412では、関心プロファイル・セグメントのそれぞれに対応するクラスタ識別子が、クラスタ識別子に基づくユーザ関心プロファイルのクラスタ化のためにリモート50

・ノードに提供される。一実装形態では、リモート・ノードを、中央エンティティとすることができ、リモート・ノードは、クラスタ化がそれに関して実行されつつあるアプリケーションのタイプに応じて、中央アグリゲータ・ノードから関心アグリゲータ・ノードに変化することができる。たとえば、コンテンツベースの推奨アプリケーションの場合に、リモート・ノードを関心グループ・アグリゲータとして実施することができる。ユーザ・デバイスは、その場合に、匿名通信チャネルを介して、クラスタ識別子およびセマンティック表現のそれぞれを、クラスタ識別子に対応する関心グループ・アグリゲータに送信するように構成される。

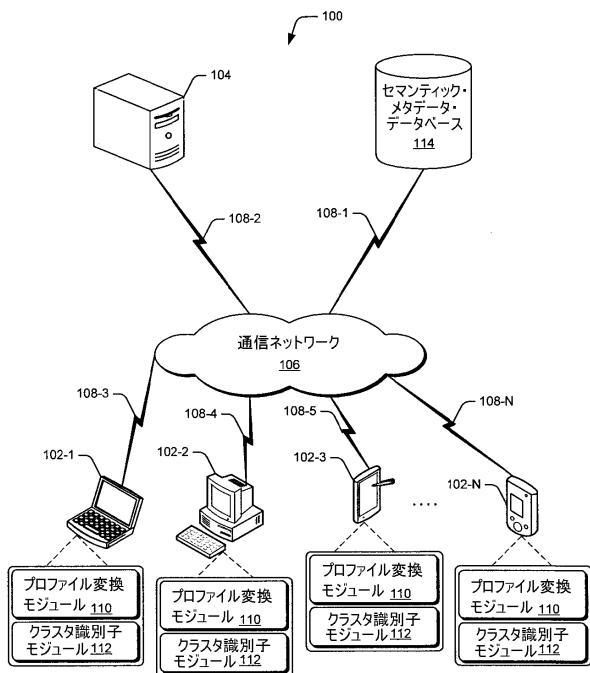
【0085】

同様に、協調フィルタリング・アプリケーションの場合に、リモート・ノードを、グループに対応する関心分野だけではなく、クラスタのエンド・ユーザの他の高くランクインされた関心分野にも関する推奨をクラスタの各メンバに提供するように構成された中央アグリゲータとして実施することができる。ユーザ・デバイスは、その場合に、クラスタ識別子に対応する関心グループ・アグリゲータに対してユーザ・デバイスを使用可能にするためにエンド・ユーザの関心プロファイルのスライスと一緒にクラスタ識別子のそれぞれを匿名通信チャネルを介して送信するように構成される。もう1つの実施形態では、ユーザ・デバイスは、ユーザ・デバイスのそれぞれが他のユーザ・デバイスのリモート・ノードまたは関心グループの責任を負う中央アグリゲータのいずれかとして働くように、他のユーザ・デバイスと相互作用することができる。

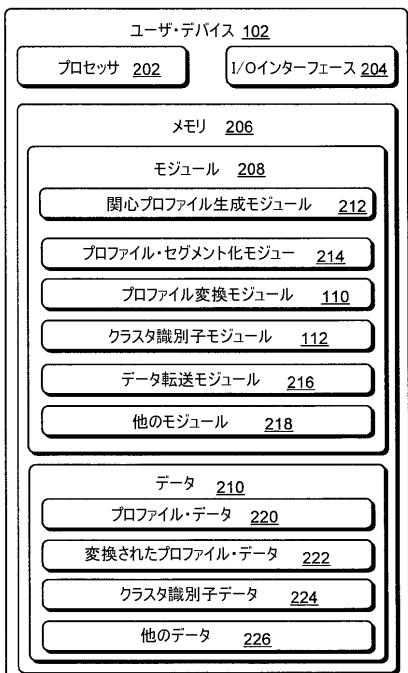
【0086】

通信ネットワークでのユーザ関心プロファイルのプライバシ保護されたクラスタ化の方法およびシステムの実施形態を、構造的特徴および／または方法に固有の言葉で説明したが、本発明が、必ずしも説明された特定の特徴または方法に限定されないことを理解されたい。そうではなく、特定の特徴および方法は、ユーザ関心プロファイルのプライバシ保護されたクラスタ化を提供する例示的実施形態として開示されるものである。

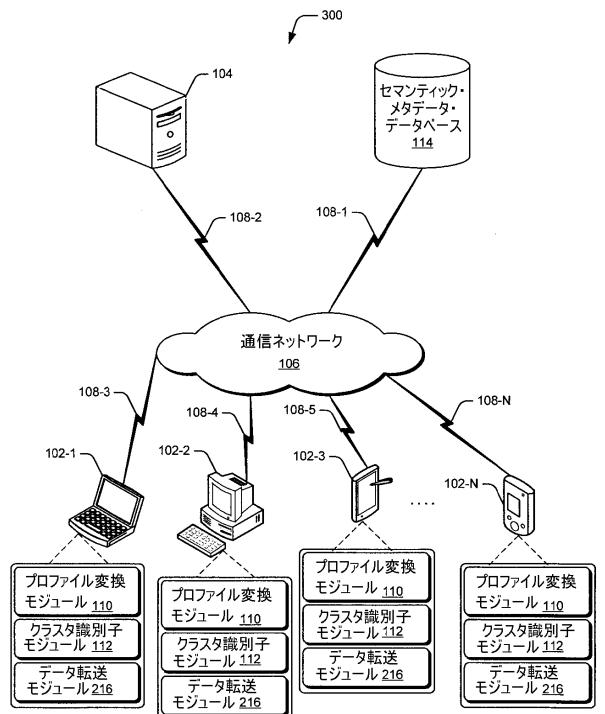
【図1】



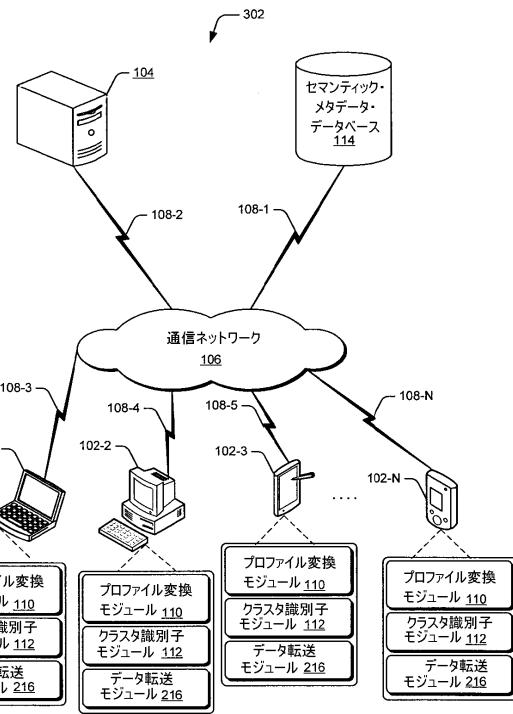
【図2】



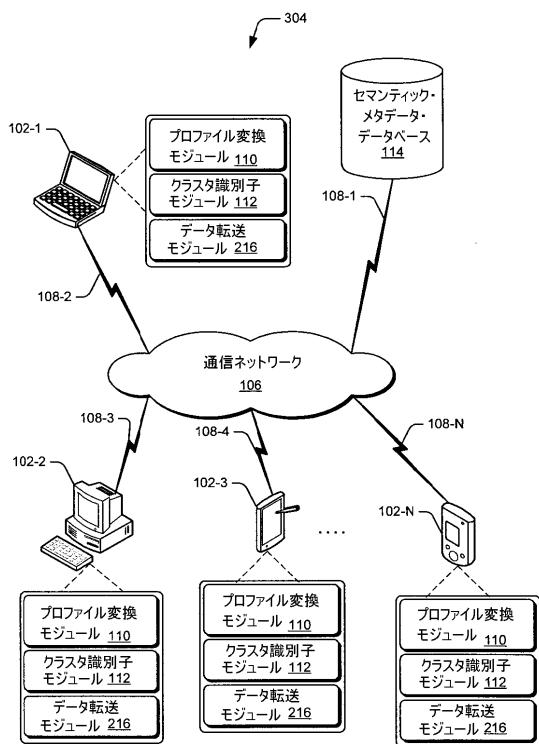
【図3(a)】



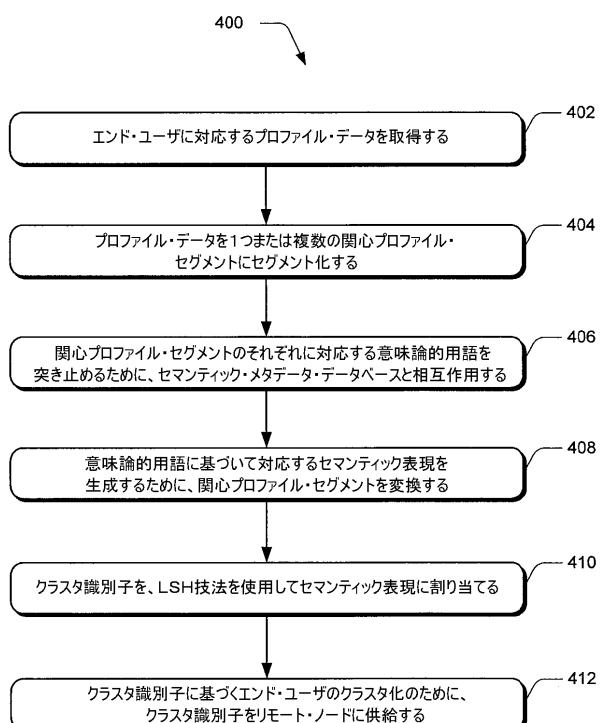
【図3(b)】



【図3(c)】



【図4】



フロントページの続き

- (72)発明者 アガハサリヤン, アーメン
フランス エフ-91620 ノゼ, ルート ドゥ ヴィルジュ, セントレ ドゥ ヴィラルソー
, アルカテル-ルーセント ベル ラボ フランス
- (72)発明者 ブジッド, マクラム
フランス エフ-91620 ノゼ, ルート ドゥ ヴィルジュ, セントレ ドゥ ヴィラルソー
, アルカテル-ルーセント ベル ラボ フランス
- (72)発明者 コサリ, モヒト
インド 382475 グジャラート, アフマダーバード, ハンソル, ニア エアポート, エスオ
ーシー., ボレシュワー シーオー-オーピー 52
- (72)発明者 ナンディ, アニメシュ
インド 560045 バンガロール, マンヤタ エンパシー ビジネス パーク, ウタ
リング ロード, カサバ タルク, ナガワラ ヴィレッジ, アルカテル-ルーセント インディア
リミテッド

審査官 早川 学

(56)参考文献 特表2009-505290 (JP, A)
特開2010-067175 (JP, A)

(58)調査した分野(Int.Cl., DB名)

G 06 F 17/30
G 06 F 21/62