

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6258246号
(P6258246)

(45) 発行日 平成30年1月10日(2018.1.10)

(24) 登録日 平成29年12月15日(2017.12.15)

(51) Int.Cl.

F 1

G 0 6 F 17/30 (2006.01)

G 0 6 F 17/30 2 1 0 D

請求項の数 9 (全 22 頁)

(21) 出願番号	特願2015-75618 (P2015-75618)	(73) 特許権者	500257300
(22) 出願日	平成27年4月2日(2015.4.2)		ヤフー株式会社
(65) 公開番号	特開2016-194885 (P2016-194885A)		東京都千代田区紀尾井町1番3号
(43) 公開日	平成28年11月17日(2016.11.17)	(74) 代理人	100107766
審査請求日	平成29年6月9日(2017.6.9)		弁理士 伊東 忠重
早期審査対象出願		(74) 代理人	100070150
			弁理士 伊東 忠彦
		(72) 発明者	真田 勉
			東京都港区赤坂九丁目7番1号 ヤフー株 式会社内
		審査官	齊藤 貴孝

最終頁に続く

(54) 【発明の名称】 解析装置、解析方法及びプログラム

(57) 【特許請求の範囲】

【請求項1】

複数のユーザ識別子のうち、同一ユーザが使用するユーザ識別子の組に関する情報を取得する取得手段と、

前記複数のユーザ識別子を、それぞれのユーザ識別子に対応付けられた情報に基づいてクラスタリングすることで得られた処理結果に基づいて、前記同一ユーザのユーザ識別子の組が、同じクラスタに分類される傾向を示す指標を算出する指標値算出手段と、

算出された前記指標を出力する出力手段と

を有することを特徴とする解析装置。

【請求項2】

算出された前記指標を含む評価関数を用いて、前記処理結果の評価値を算出する評価値算出手段を更に有することを特徴とする請求項1に記載の解析装置。

【請求項3】

算出された前記評価値を、解析結果として出力する解析結果出力手段を更に有することを特徴とする請求項2に記載の解析装置。

【請求項4】

算出された前記評価値に応じてクラスタリングするクラスタリング手段を更に有することを特徴とする請求項2に記載の解析装置。

【請求項5】

前記クラスタリング手段は、前記評価値が所定の条件を満たすまで、各クラスタに分類

10

20

されるためのクラスタリング条件を変更しながらクラスタリングを繰り返すことを特徴とする請求項 4 に記載の解析装置。

【請求項 6】

前記評価値が所定の条件を満たした場合の前記クラスタリング条件のもとでクラスタリングすることで得られた処理結果を、解析結果として出力する解析結果出力手段を更に有することを特徴とする請求項 5 に記載の解析装置。

【請求項 7】

前記評価値算出手段は、

前記処理結果が利用されたことによる効果を示す情報に基づいて、前記評価関数における前記指標の重み係数を変更することを特徴とする請求項 2 乃至 6 のいずれか 1 項に記載の解析装置。

10

【請求項 8】

解析装置における解析方法であって、

取得手段が、複数のユーザ識別子のうち、同一ユーザが使用するユーザ識別子の組に関する情報を取得する取得工程と、

指標値算出手段が、前記複数のユーザ識別子を、それぞれのユーザ識別子に対応付けられた情報に基づいてクラスタリングすることで得られた処理結果に基づいて、前記同一ユーザのユーザ識別子の組が、同じクラスタに分類される傾向を示す指標を算出する指標値算出工程と、

出力手段が、算出された前記指標を出力する出力工程と
を有することを特徴とする解析方法。

20

【請求項 9】

解析装置のコンピュータに、

複数のユーザ識別子のうち、同一ユーザが使用するユーザ識別子の組に関する情報を取得する取得工程と、

前記複数のユーザ識別子を、それぞれのユーザ識別子に対応付けられた情報に基づいてクラスタリングすることで得られた処理結果に基づいて、前記同一ユーザのユーザ識別子の組が、同じクラスタに分類される傾向を示す指標を算出する指標値算出工程と、

算出された前記指標を出力する出力工程と
を実行させるためのプログラム。

30

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、解析装置、解析方法及びプログラムに関する。

【背景技術】

【0002】

従来より、膨大なデータを複数のクラスタに分類するクラスタリング技術が知られており（例えば、下記特許文献 1、2 参照）、様々な分野への適用が進められている。

【0003】

例えば、当該クラスタリング技術を用いれば、インターネットを利用するユーザの購買履歴情報や閲覧履歴情報、検索履歴情報等の行動履歴に基づいて、ユーザを複数のクラスタに分類することができる。これにより、特定のクラスタに分類されたユーザだけに、所定の広告を配信したり、販売促進メールを送信するといった、効果的なマーケティング活動を実現することが期待できる。

40

【先行技術文献】

【特許文献】

【0004】

【特許文献 1】特開 2014-63343 号公報

【特許文献 2】特開 2002-278762 号公報

【発明の概要】

50

【発明が解決しようとする課題】

【0005】

ここで、より効果的なマーケティング活動を実現するためには、同じ嗜好性を有するユーザ同士が同じクラスタに分類されるようにクラスタリングが行われることが望ましい。

【0006】

しかしながら、これまでのクラスタリング技術は、行動履歴が似たユーザ同士が同じクラスタに分類されるようにクラスタリングが行われているにすぎず、嗜好性が同じユーザ同士が同じクラスタに分類されているか否かは必ずしも定かではない。

【0007】

本発明は上記課題に鑑みてなされたものであり、ユーザの嗜好性に基づくクラスタリングを行うための指標を提供することを目的とする。

10

【課題を解決するための手段】

【0008】

本発明の一実施形態に係る解析装置は、以下のような構成を有する。すなわち、複数のユーザ識別子のうち、同一ユーザが使用するユーザ識別子の組に関する情報を取得する取得手段と、

前記複数のユーザ識別子を、それぞれのユーザ識別子に対応付けられた情報に基づいてクラスタリングを実行することで得られた処理結果に基づいて、前記同一ユーザのユーザ識別子の組が、同じクラスタに分類される傾向を示す指標を算出する指標値算出手段と、

算出された前記指標を出力する出力手段とを有することを特徴とする。

20

【発明の効果】

【0009】

本発明の各実施形態によれば、ユーザの嗜好性に基づくクラスタリングを行うための指標を提供することが可能となる。

【図面の簡単な説明】

【0010】

【図1】 サービス提供システムの全体構成の一例を示す図である。

【図2】 解析装置のハードウェア構成を示す図である。

【図3】 ユーザ情報の一例を示す図である。

【図4】 クラスタ情報の一例を示す図である。

30

【図5】 評価情報の一例を示す図である。

【図6】 クラスタの数とパラメータとの関係を示す図である。

【図7】 係数とパラメータとの関係を示す図である。

【図8】 解析装置の機能構成の一例を示す図である。

【図9】 評価用情報の一例を示す図である。

【図10】 パラメータを説明するための図である。

【図11】 クラスタ情報の他の一例を示す図である。

【図12】 評価処理の流れを示すフローチャートである。

【図13】 サービス提供システムの全体構成の他の一例を示す図である。

【図14】 解析装置の機能構成の他の一例を示す図である。

40

【図15】 評価情報の他の一例を示す図である。

【図16】 係数最適化処理の流れを示すフローチャートである。

【図17】 サービス提供システムの全体構成の他の一例を示す図である。

【図18】 同一ユーザ推定装置及び解析装置の機能構成の一例を示す図である。

【発明を実施するための形態】

【0011】

以下、本発明の各実施形態の詳細について添付の図面を参照しながら説明する。なお、以下の各実施形態では、同じ嗜好性を有するユーザ同士が同じクラスに分類されるよう、同一人物で複数のユーザIDを使用するユーザに着目した。

【0012】

50

一般に、インターネットを利用するユーザは、同一の機器で複数のユーザ I Dを使い分けたり、複数の機器（パソコン、スマートフォン等）で別々のユーザ I Dを使用したりする。このため、同一ユーザ（人物）が複数のユーザ I Dを使用するケースが増えてきている。

【0013】

一方で、ユーザ I Dが異なっているとしても、同一ユーザ（人物）であれば、嗜好性は同じであり、使用しているユーザ I Dには依存しないはずである。そこで、以下の各実施形態に係る解析装置では、同一ユーザ（人物）が使用する複数のユーザ I Dが、同じクラスタに分類されるようにクラスタリングを行う。

【0014】

具体的には、同一ユーザ（人物）の複数のユーザ I Dが、同じクラスタに分類される傾向を示す指標を提供し、当該指標を含む評価関数の計算結果（評価値）が所定の条件を満たすように、クラスタリングを行う。これにより、同一ユーザ（人物）の複数のユーザ I Dが、同じクラスタに分類されることとなり、嗜好性が同じユーザ（人物）のユーザ I Dが、同じクラスタに分類される確率を上げることができる。その結果、より効果的なマーケティング活動を実現するクラスタリングを行うことが可能となる。

【0015】

加えて、このようなクラスタリングを行うことで得られたクラスタリング条件（各クラスタに分類するための条件）を用いれば、同一人物であろうと推定された結果の妥当性を評価することも可能となる。つまり、複数のユーザ I Dを使用しているユーザがおそらく同一人物であろうと推定された結果の妥当性を、その複数のユーザ I Dが同じクラスタに分類されるか否かをクラスタリング条件を用いて判定することで、評価することが可能となる。

【0016】

以下、各実施形態の詳細について説明する。なお、各実施形態に係る明細書及び図面の記載に際して、実質的に同一の機能構成を有する構成要素については、同一の符号を付することにより重複した説明を省く。

【0017】

[第1の実施形態]

<1. サービス提供システムの全体構成>

はじめに、本実施形態に係る解析装置を有するサービス提供システムの全体構成について説明する。図1は、サービス提供システムの全体構成の一例を示す図である。図1に示すように、サービス提供システム100は、サービス提供サーバ110と、解析装置120と、ユーザ端末131、132、・・・とを有する。サービス提供システム100において、サービス提供サーバ110と解析装置120とは、ネットワーク140を介して通信可能に接続されている。また、サービス提供サーバ110とユーザ端末131、132、・・・とは、ネットワーク140を介して通信可能に接続されている。

【0018】

サービス提供サーバ110には、サービス提供プログラム及びマーケティングプログラムがインストールされており、当該プログラムが実行されることで、サービス提供サーバ110は、サービス提供部111、マーケティング部112として機能する。

【0019】

サービス提供部111は、ユーザ端末131、132、・・・にインターネット上の各種サービスを提供する。サービス提供部111では、各種サービスの提供に際して、ユーザ端末131、132、・・・より、ユーザの属性情報を取得し、ユーザ I Dと対応付けてユーザ情報としてユーザ情報格納部113に格納する。サービス提供部111により取得されるユーザの属性情報には、例えば、ユーザの氏名、生年月日、性別、住所等が含まれる。

【0020】

また、サービス提供部111は、各種サービスの提供において、ユーザが購入した商品

10

20

30

40

50

の購買履歴やユーザが閲覧したWebサイトの閲覧履歴、ユーザが検索した検索履歴等を取得し、ユーザIDと対応付けて、ユーザ情報としてユーザ情報格納部113に格納する。

【0021】

マーケティング部112は、解析装置120に対して解析依頼を送信する。具体的には、ユーザ情報格納部113に格納されたユーザ情報を解析し、ユーザの嗜好性に基づくクラスタリングを行うよう、解析依頼を送信する。なお、マーケティング部112により解析装置120に対して送信される解析依頼には、ユーザ情報が含まれるものとする。

【0022】

また、マーケティング部112は、解析依頼に応じて解析装置120より送信される解析結果を取得する。解析装置120より送信される解析結果には、ユーザ情報に含まれる各ユーザのユーザIDが複数のクラスタに分類されたクラスタ情報等が含まれる。マーケティング部112では、取得したクラスタ情報を利用して、所定のクラスタに分類されたユーザに対して、所定の広告を提供したり、所定の商品の販売促進メールを送信したりする。

10

【0023】

一方、解析装置120には、評価プログラムがインストールされており、当該プログラムが実行されることで、解析装置120は、評価部121として機能する。

【0024】

評価部121は、解析依頼に含まれるユーザ情報についてクラスタリングを行い、クラスタ情報をクラスタ情報格納部122に格納する。評価部121では、ユーザ情報に含まれる同一ユーザ（人物）の複数のユーザIDが、同じクラスタに分類されるようにクラスタリングを行う。具体的には、同一ユーザ（人物）の複数のユーザIDが同じクラスタに分類される傾向を示す指標が含まれる評価関数を評価情報格納部123より読み出し、当該評価関数の値（評価値）が所定の条件を満たすようにクラスタリングを行う。

20

【0025】

また、評価部121は、クラスタ情報格納部122に格納したクラスタ情報等を、解析結果としてサービス提供サーバ110に送信する。

【0026】

ユーザ端末131、132、・・・は、スマートフォンやパソコン等の情報端末であり、Webブラウザ等を介して、サービス提供サーバ110より各種サービスの提供を受ける。ユーザ端末131、132、・・・のユーザは、各種サービスの提供を受けるにあたり、自身の属性情報をサービス提供サーバ110に送信し、登録する。これにより、ユーザ端末131、132、・・・のユーザは、サービス提供サーバ110から送信されるサービスを受信したり、各種広告、販売促進メール等を受信したりすることができる。

30

【0027】

＜2. 解析装置のハードウェア構成＞

次に、解析装置120のハードウェア構成について説明する。図2は、解析装置120のハードウェア構成を示す図である。

【0028】

図2に示すように、解析装置120は、CPU（Central Processing Unit）201、ROM（Read Only Memory）202、RAM（Random Access Memory）203、補助記憶部204を備える。また、解析装置120は、入力部205、表示部206、通信部207を備える。なお、解析装置120の各部は、バス208を介して相互に接続されている。

40

【0029】

CPU201は、補助記憶部204に格納された各種プログラム（評価プログラム等）を実行するコンピュータである。

【0030】

ROM202は不揮発性メモリである。ROM202は、補助記憶部204に格納され

50

た各種プログラムをCPU 201が実行するために必要な、各種プログラム、データ等を格納する主記憶部として機能する。具体的には、BIOS (Basic Input/Output System) やEFI (Extensible Firmware Interface) 等のブートプログラムなどを格納する主記憶部として機能する。

【0031】

RAM 203は、DRAM (Dynamic Random Access Memory) やSRAM (Static Random Access Memory) 等の主記憶部である。RAM 203は、補助記憶部204に格納された各種プログラムがCPU 201によって実行される際に展開される、作業領域として機能する。

【0032】

10

補助記憶部204は、CPU 201により実行される各種プログラム及び各種プログラムがCPU 201により実行される際に利用される各種情報を格納する。補助記憶部204の一領域は、各種情報を格納する格納部として機能し、当該格納部には、例えば、クラスタ情報格納部122、評価情報格納部123が含まれる。

【0033】

入力部205は、解析装置120の操作者が各種指示を入力する際に用いる。また、表示部206は、解析装置120の内部状態を表示する。通信部207は、ネットワーク140を介して、サービス提供サーバ110と通信を行う。

【0034】

なお、サービス提供サーバ110及びユーザ端末131、132、・・・のハードウェア構成も、解析装置120のハードウェア構成と概ね同じであるため、ここでは説明を省略する。

20

【0035】

< 3. ユーザ情報の説明 >

次に、サービス提供サーバ110のユーザ情報格納部113に格納されるユーザ情報について説明する。図3は、ユーザ情報300の一例を示す図である。図3に示すように、ユーザ情報300には情報の項目として、"ユーザID"、"氏名"、"生年月日"、"性別"、"住所"、"購買履歴"、"閲覧履歴"、"検索履歴"等が含まれる。

【0036】

"ユーザID"には、サービス提供部111により提供される各種サービスを受けるユーザ（またはユーザ端末）を識別するユーザ識別子（ユーザID）が記録される。なお、ここでいうユーザIDには、ユーザがサービス提供サーバ110の管理者（プロバイダ）との間でプロバイダ契約を結ぶことで付与されるユーザIDのほか、Webブラウザを利用する際に発行されるクッキー等が含まれる。あるいは、ユーザが使用するユーザ端末131、132、・・・のシリアル番号等のユーザ端末ID、メールアドレス、会員番号、クレジットカード番号、IPアドレス、電話番号等が含まれる。

30

【0037】

"氏名"、"生年月日"、"性別"、"住所"には、ユーザIDにより特定されるユーザの氏名、生年月日、性別、住所が記録される。ただし、ユーザの属性を示すこれらの属性情報は、ユーザIDが付与される際に、ユーザによる入力が必要なケースと不要なケースとがある。このため、ユーザ情報300には、"ユーザID"に対応する"氏名"、"生年月日"、"性別"、"住所"が空欄となっているユーザIDも存在する。

40

【0038】

なお、本実施形態では、"氏名"と"生年月日"が記録されているユーザIDの中で、氏名と生年月日が同一のユーザIDを、同一ユーザ（人物）が使用するユーザIDであると判断するものとする。つまり、属性情報のうち、"氏名"と"生年月日"は、本実施形態において、同一ユーザ（人物）のユーザIDか否かの判断に用いる情報である。

【0039】

"購買履歴"には、ユーザIDにより特定されるユーザ（人物）が、商品等を購入した場合の購買履歴が記録される。"閲覧履歴"には、ユーザIDにより特定されるユーザ（人物

50

）がWebサイトを閲覧した際の履歴等が記録される。"検索履歴"には、ユーザIDにより特定されるユーザ（人物）がネット検索した際の履歴等が記録される。

【0040】

<4. クラスタ情報>

次に、解析装置120のクラスタ情報格納部122に格納されるクラスタ情報について説明する。図4は、クラスタ情報400の一例を示す図である。図4に示すように、クラスタ情報400には情報の項目として、"クラスタ番号"、"クラスタリング条件"、"クラスタリング結果"が含まれる。

【0041】

"クラスタ番号"には、評価部121によりクラスタリングが行われることで生成されたクラスタを識別するクラスタ番号が記録される。

10

【0042】

"クラスタリング条件"には、各クラスタに分類されるためのクラスタリング条件が記録される。クラスタリング条件は、複数の項目に分けて規定される。図4の例では、クラスタリング条件として、"購買履歴"、"閲覧履歴"、"検索履歴"等が規定されている。

【0043】

"クラスタリング結果"には、各クラスタに分類されたユーザIDが記録される。具体的には、各クラスタのクラスタリング条件に合致するユーザ情報を有するユーザIDが記録される。図4の例では、クラスタ番号="クラスタ1"の場合、クラスタリング条件として検索履歴="スポーツA"が規定されているため、少なくとも検索履歴="スポーツA"をユーザ情報に含むユーザIDが、"クラスタ1"に分類される。具体的には、ユーザID="Y0001"、"Y0005"等が"クラスタ1"に分類される。

20

【0044】

<5. 評価情報>

次に、解析装置120の評価情報格納部123に格納される評価情報について説明する。評価情報は、評価部121がクラスタリングを行う際に用いる評価関数を規定したものである。本実施形態において用いる評価関数は、以下の式により定義される。

（評価関数）＝（第1の評価関数）＋（係数）×（第2の評価関数）

図5は、評価情報500の一例を示す図である。図5に示すように、評価情報500には情報の項目として、"第1の評価関数"、"第2の評価関数"、"係数"、"指標値"、"評価値"が含まれる。

30

【0045】

"第1の評価関数"に記録される"S"は、従来のクラスタリングにおける評価関数である。例えば、K平均法（k-means clustering）によるクラスタリングの場合、ユーザIDと各ユーザIDが分類されたクラスタの中心との距離を評価する関数である。また、ワード法によるクラスタリングの場合、クラスタ間の距離を評価する関数である。なお、評価関数は、値が「大きいほど良い」とする場合と値が「小さいほど良い」とする場合とがあるが、符号を変更すれば数学的には同等であるため、ここでは値が「大きいほど良い」とする場合に限ることとする。

【0046】

40

"第2の評価関数"に記録される" R/Q "は、同一ユーザ（人物）の複数のユーザIDが、同じクラスタに分類される傾向を示す指標である。なお、" R/Q "の詳細は後述する。"係数"に記録される" α "は、評価関数全体における第2の評価関数（ R/Q ）の影響度を示す値である。

【0047】

"指標値"には、クラスタリングを実行した際に、"第2の評価関数"に基づいて算出された値である指標値が記録される。"評価値"には、クラスタリングを実行した際に、"評価関数"に基づいて算出された値である評価値が記録される。

【0048】

このように、本実施形態において、評価部121は、クラスタリングを行う際に、従来

50

の評価関数（第1の評価関数）の値に加え、同一ユーザ（人物）の複数のユーザIDが同じクラスタに分類される傾向を示す指標が含まれた評価値を用いる。当該評価値は、同一ユーザ（人物）の複数のユーザIDが同じクラスタに分類されるほど、値が大きくなる。つまり、クラスタリングを行った際に、従来の評価関数を用いた場合と比較して（第1の評価関数のみを用いた場合と比較して）、同一ユーザ（人物）のユーザIDが、同じクラスタに分類される傾向がより強いクラスタ情報が得られることとなる。

【0049】

<6. 第2の評価関数の説明>

次に、第2の評価関数（ R/Q ）について詳細に説明する。第2の評価関数（ R/Q ）は以下のように導き出される。

10

【0050】

はじめに、クラスタリングにより、N個のクラスタが生成され、それぞれのクラスタに分類されたユーザIDの数が、 m_1 、 m_2 、 $\dots$ 、 m_N であったとする（つまり、ユーザIDの総数 $M = m_1 + m_2 + \dots + m_N$ であったとする）。

【0051】

この場合、ランダムに選択したユーザIDが、i番目のクラスタに分類されている確率 p_i は、

【0052】

【数1】

20

$$p_i = m_i / M$$

となる。また、ランダムに選択した2つのユーザID（ ID_1 と ID_2 ）が、ともにi番目のクラスタに分類される確率 q_i は、

【0053】

【数2】

30

$$q_i = p_i \times p_i = (m_i / M)^2$$

となる。したがって、ランダムに選択した2つのユーザID（ ID_1 と ID_2 ）が、同じクラスタに分類される確率Qは、

【0054】

【数3】

40

$$Q = \sum_i q_i = \sum_i \left(\frac{m_i}{M} \right)^2$$

となる。

【0055】

50

ここで、同一ユーザ（人物）のユーザ I D の組が全部で L 組あり、そのうち同一のクラスタに分類されているユーザ I D の組が K 組であったとすると、同一のクラスタに分類されている組の割合 R は、

【0056】

【数4】

$$R = K/L$$

10

となる。

【0057】

本実施形態では、確率 Q と割合 R の比、すなわち "R/Q" を第 2 の評価関数としている。このようにして算出される第 2 の評価関数 (R/Q) は、同一ユーザ（人物）の複数のユーザ I D が、同一のクラスタに分類されている傾向が強いほど値（指標値）が大きくなる。逆に、同一ユーザ（人物）の複数のユーザ I D が、同一のクラスタに分類されている傾向が弱いほど値（指標値）が小さくなる。

【0058】

なお、第 2 の評価関数として、R ではなく Q に対する比 R/Q を採用したのは以下の理由による。図 6 は、確率 Q、割合 R、第 2 の評価関数 (R/Q) の、クラスタの数 N に対する関係を示した図である。同一ユーザ（人物）の 2 つのユーザ I D が同じクラスタに分類される確率 R は、例えば、クラスタの数 N = 2 で、かつ、各クラスタに分類されるユーザ I D の数がそれぞれ同じであれば、出鱈目に生成したクラスタであっても、0.5 となる（図 6 参照）。一方、適切に生成したクラスタであっても、クラスタの数 N が大きくなると、確率 R は通常 0.5 よりはるかに小さくなる。

20

【0059】

このため、第 2 の評価関数として確率 R を採用したとすると、出鱈目に生成したクラスタ数が小さいクラスタの方が、適切に生成したクラスタ数の大きいクラスタよりも評価関数の値が大きくなってしまう。一方、第 2 の評価関数として R/Q を採用した場合、確率 Q と割合 R とが共に、クラスタの数 N と、それぞれのクラスタに含まれるユーザ I D の数とに依存する。このため、その比、すなわち評価関数の値は、クラスタの数 N とそれぞれのクラスタに含まれるユーザ I D の数が同一の出鱈目に作成したクラスタに対する相対評価となる。そこで、本実施形態では、第 2 の評価関数として R/Q を採用する。

30

【0060】

また、評価関数として、第 2 の評価関数だけでなく第 1 の評価関数を含めたのは以下の理由による。図 7 は、係数 α と、確率 R、ユーザ I D の総数 M に対するクラスタの数 N の 2 倍の比 ($= N \times 2 / M$) との関係を示した図である。図 7 に示すように、係数 α の値を大きくしていくと、第 1 の評価関数の影響度が小さくなり、第 2 の評価関数の値が支配的となる。

40

【0061】

この場合、同一ユーザ（人物）が使う 2 つのユーザ I D のみで、1 つのクラスタが形成されるとき、R の値は 1 となり、評価関数 R/Q は最大値となる。つまり、第 2 の評価関数の値は大きくなるものの、同一ユーザ（人物）の数だけクラスタが生成されることとなる。しかしながら、通常この数（生成されるクラスタの数）は膨大な値であり、実用的でない結果となる。このようなことから、本実施形態では、評価関数に、第 1 の評価関数と第 2 の評価関数の両方を含める構成とした。

【0062】

< 7. 解析装置の機能構成 >

次に、解析装置 120 の機能構成について説明する。図 8 は、解析装置 120 の機能構

50

成の一例を示す図である。図8に示すように、解析装置120の評価部121は、解析依頼取得部801、評価用情報抽出部802、同一ユーザ特定部803、クラスタリング部804、パラメータ算出部805、指標値算出部806、評価値算出部807を有する。

【0063】

解析依頼取得部801は、サービス提供サーバ110より送信された解析依頼を取得する。

【0064】

評価用情報抽出部802は、解析依頼取得部801において取得された解析依頼に含まれるユーザ情報300のうち、ユーザの属性情報（氏名、生年月日、性別、住所）であって、氏名、生年月日が記録されているユーザ情報を、評価用情報として抽出する。

10

【0065】

同一ユーザ特定部803は、評価用情報抽出部802により抽出された評価用情報のうち、氏名と生年月日とが同一のユーザIDを、同一ユーザ（人物）が使うユーザIDと特定する。同一ユーザ特定部803は、特定した同一ユーザ（人物）のユーザIDの組と、特定した同一ユーザ（人物）のユーザIDの組の数とをパラメータ算出部805に通知する。

【0066】

図9は、評価用情報900の一例を示す図である。図9に示すように、評価用情報900は、ユーザ情報300のうち、氏名、生年月日が記録されているユーザ情報を抽出したものである。図9の例では、ユーザID="Y0001"と"Y0007"のユーザが、氏名と生年月日が同じであるため同一ユーザ（人物）であると特定されている。また、ユーザID="Y0002"と"Y0011"のユーザも同一ユーザ（人物）と特定されている。

20

【0067】

同一ユーザ特定部803では、ユーザID="Y0001"と"Y0007"の組と、ユーザID="Y0002"と"Y0011"の組とを、パラメータ算出部805に通知する。また、同一ユーザ特定部803は、同一ユーザ（人物）のユーザIDの組の数Lをパラメータ算出部805に通知する。

【0068】

図8の説明に戻る。クラスタリング部804は、解析依頼に含まれるユーザ情報300をクラスタリングし、クラスタ情報400を取得する。また、クラスタリング部804は、取得したクラスタ情報400をクラスタ情報格納部122に格納する。

30

【0069】

パラメータ算出部805は、第2の評価関数の算出に用いる各パラメータを算出する。図10は、第2の評価関数の算出に用いる各パラメータを説明するための図である。図10の例によれば、パラメータ算出部805は、クラスタ情報400の"クラスタ番号"に基づいて、クラスタの数Nを導出する。また、パラメータ算出部805は、"クラスタリング結果"に基づいて、各クラスタに分類されたユーザIDの数 m_i ($1 \leq i \leq N$)を算出する。更に、パラメータ算出部805は、各クラスタに分類されたユーザIDの数 m_i ($1 \leq i \leq N$)に基づいて、ユーザIDの総数Mを算出する。

【0070】

また、パラメータ算出部805は、同一ユーザ（人物）のユーザIDの組の数Lについて、同一ユーザ特定部803より通知を受ける。更に、パラメータ算出部805は、クラスタ情報400の"クラスタリング結果"に基づいて、同一ユーザ（人物）のユーザIDの組のうち、同一クラスタに分類されているユーザIDの組の数Kを算出する。

40

【0071】

再び図8の説明に戻る。指標値算出部806は、パラメータ算出部805において算出された、第2の評価関数の算出に用いる各パラメータに基づいて、第2の評価関数を算出し、指標値を得る。具体的には、パラメータ算出部805において算出された m_i 及びMに基づいて、確率Qを算出する。また、パラメータ算出部805において算出された、ユーザIDの組の数LとKに基づいて、割合Rを算出する。更に、算出した確率Qと割合R

50

とに基づいて第2の評価関数（R/Q）を算出し、指標値を得る。

【0072】

評価値算出部807は、評価情報格納部123に格納されている評価関数を用いて、クラスタ情報400についての評価値を算出する。また、評価値算出部807は、算出した評価値をクラスタリング部804にフィードバックする。これにより、クラスタリング部804では、クラスタリング条件を変更して、再び、クラスタリングを行う。

【0073】

図11は、クラスタ情報の他の一例を示す図である。クラスタ情報1100は、クラスタ情報400について算出した評価値に基づいて、クラスタリング部804が再びクラスタリングを行うことで得られたクラスタ情報の一例である。

10

【0074】

図11に示すように、クラスタ情報1100の"クラスタリング結果"によれば、1回目のクラスタリングでは、同一のクラスタに分類されなかった、同一ユーザ（人物）の2つのユーザID（"Y0001"、"Y0007"）が同一のクラスタに分類されている。

【0075】

このように、評価関数により算出された評価値に基づいてクラスタリングを繰り返し実行することで、同一ユーザ（人物）の複数ユーザIDが、同一のクラスタに分類される確率が上がっていく。

【0076】

再び、図8の説明に戻る。評価値算出部807は、クラスタリングが繰り返し実行されることで評価値が所定の条件を満たすと判断した場合に、クラスタリング部804に通知する。これにより、クラスタリング部804では、クラスタリングを終了する。また、評価値算出部807は、評価値が所定の条件を満たすと判断した際の指標値及び評価値を評価情報格納部123に格納する。

20

【0077】

解析結果出力部808は、クラスタ情報格納部122に格納されているクラスタ情報及び評価情報格納部123に格納されている指標値及び評価値を、解析結果としてサービス提供サーバ110に送信する。

【0078】

＜8. 評価処理の流れ＞

30

次に、評価部121による評価処理の流れについて説明する。図12は、評価部121による評価処理の流れを示すフローチャートである。ステップS1201において、解析依頼取得部801は、サービス提供サーバ110より、解析依頼を受信したか否かを判定する。

【0079】

ステップS1201において、解析依頼を受信していないと判定した場合には、受信したと判定するまで待機する。一方、ステップS1201において、解析依頼を受信したと判定した場合には、ステップS1202に進む。

【0080】

ステップS1202において、評価用情報抽出部802は、解析依頼に含まれるユーザ情報より、評価用情報900を抽出する。ステップS1203において、同一ユーザ特定部803は、ステップS1202において抽出された評価用情報900に基づいて、同一ユーザ（人物）のユーザIDの組を特定する。更に、同一ユーザ特定部803は、同一ユーザ（人物）のユーザIDの組及び組の数をパラメータ算出部805に通知する。

40

【0081】

ステップS1204において、クラスタリング部804は、ステップS1201において抽出されたユーザ情報をクラスタリングし、クラスタ情報を取得する。更に、クラスタリング部804は、取得したクラスタ情報を、クラスタ情報格納部122に格納する。

【0082】

ステップS1205において、パラメータ算出部805は、ステップS1202におい

50

て抽出されたクラスタ情報と、ステップS 1 2 0 3において通知された、同一ユーザ（人物）のユーザ I Dの組及び組の数とに基づいて、パラメータを算出する。

【0083】

ステップS 1 2 0 6において、指標値算出部806は、ステップS 1 2 0 5において算出されたパラメータに基づいて第2の評価関数の値（指標値）を算出する。ステップS 1 2 0 7において、評価値算出部807は、ステップS 1 2 0 2において抽出されたクラスタ情報に基づいて第1の評価関数の値を算出する。更に、評価値算出部807は、ステップS 1 2 0 6において算出された第2の評価関数の値（指標値）に係数 α をかけあわせてうえて第1の評価関数の値に加算することで、評価値を算出する。

【0084】

10

ステップS 1 2 0 8において、評価値算出部807は、評価値が所定の条件を満たすかを判定する。ステップS 1 2 0 8において、所定の条件を満たさないと判定された場合には、ステップS 1 2 0 4に戻り、クラスタリング条件を変更したうえでステップS 1 2 0 4からステップS 1 2 0 7の処理を実行する。

【0085】

一方、ステップS 1 2 0 8において、所定の条件を満たすと判定された場合には、ステップS 1 2 0 9に進む。ステップS 1 2 0 9において、評価値算出部807は、算出された指標値及び評価値を評価情報格納部123に格納する。

【0086】

ステップS 1 2 1 0において、解析結果出力部808は、クラスタ情報格納部122に格納されているクラスタ情報と、評価情報格納部123に格納されている指標値及び評価値とを、解析結果として出力する。

20

【0087】

<9. まとめ>

以上の説明から明らかなように、本実施形態に係る解析装置では、嗜好性が同じユーザとして、氏名と生年月日が同一のユーザ（人物）が使用する複数のユーザ I Dに着目した。そして、同一ユーザ（人物）が使用する複数のユーザ I Dが同じクラスタに分類される傾向を示す指標を算出する構成とした。

【0088】

これにより、ユーザの嗜好性に基づくクラスタリングを行うための指標を提供することが可能となる。

30

【0089】

また、本実施形態に係る解析装置では、当該指標を含む評価関数を用いて、クラスタリングの処理結果を評価し、当該評価関数の値（評価値）が所定の条件を満たすまで、クラスタリング条件を変更しながらクラスタリングを繰り返す構成とした。

【0090】

これにより、同一ユーザ（人物）が使用するユーザ I Dが同じクラスタに分類されることとなり、結果として、嗜好性が同じユーザ（人物）のユーザ I Dが同じクラスタに分類される確率が上がる。つまり、マーケティング活動の観点から、適切なクラスタリングを実現することが可能となる。

40

【0091】

[第2の実施形態]

上記第1の実施形態では、評価情報に格納される係数 α を固定値として説明した。しかしながら、係数 α は固定値である必要はなく、例えば、サービス提供サーバ110のマーケティング部112がクラスタ情報を利用したことによって得られた効果を反映させるようにしてもよい。以下、第2の実施形態について説明する。なお、説明は第1の実施形態との相違点を中心に行う。

【0092】

<1. サービス提供システムの全体構成>

はじめに、本実施形態に係る解析装置120を有するサービス提供システム1300の

50

全体構成について説明する。図13は、サービス提供システム1300の全体構成の他の一例を示す図である。図1に示すサービス提供システム100との相違点は、サービス提供サーバ110が、コンバージョン情報格納部1301を有する点、及び、解析装置120の評価部1302の機能が、評価部121の機能と異なる点である。

【0093】

コンバージョン情報格納部1301は、マーケティング部112により取得されたクリック回数を、コンバージョン情報として格納する。具体的には、マーケティング部112が、特定のクラスタに分類されたユーザIDにより特定されるユーザ（人物）に対して広告を配信した場合に、マーケティング部112では、当該配信に応じて、当該ユーザ（人物）が当該広告をクリックした回数をカウントする。コンバージョン情報格納部1301は、マーケティング部112によりカウントしたクリック回数を、コンバージョン情報として格納する。

10

【0094】

なお、マーケティング部112においてカウントするクリック数は、係数 α が変更されるごとにリセットされるものとする。また、マーケティング部112では、ユーザ情報についての解析依頼を送信する際に、コンバージョン情報格納部1301に格納されたコンバージョン情報を解析依頼に含めて送信するものとする。

【0095】

評価部1302は、解析依頼に含まれるコンバージョン情報と、当該コンバージョン情報が取得された際に格納されていた係数 α との関係に基づいて、係数 α を変更する。そして、変更後の係数 α を用いて、評価処理（図12）を行う。

20

【0096】

< 2. 解析装置の機能構成 >

次に、本実施形態に係る解析装置120の機能構成について説明する。図14は、解析装置120の他の機能構成の一例を示す図である。図8に示した解析装置120の機能構成との相違点は、解析依頼取得部1401の機能、評価値算出部1402の機能、及び評価情報格納部1403に格納される評価情報の内容である。以下、当該相違点について説明する。

【0097】

解析依頼取得部1401は、サービス提供サーバ110より送信された解析依頼を取得する。また、解析依頼取得部1401は、解析依頼に含まれるコンバージョン情報を抽出し、係数 α と対応付けて、評価情報格納部1403に格納する。

30

【0098】

評価値算出部1402は、評価情報格納部1403に格納されたコンバージョン情報と係数 α との関係を解析することで、最適な係数 α を決定する。評価値算出部1402では、評価関数を用いて評価値を算出するにあたり、最適な係数 α を用いる。

【0099】

評価情報格納部1403は、第1の評価関数、第2の評価関数、係数 α 、指標値、評価値に加え、コンバージョン情報を格納する。コンバージョン情報は、それぞれのコンバージョン情報が取得された際に格納されていた係数 α と対応付けて格納される。

40

【0100】

< 3. 評価情報の説明 >

次に、解析装置120の評価情報格納部1403に格納される評価情報1500について図15を用いて説明する。図5に示す評価情報500との相違点は、評価情報1500が情報の項目として、"コンバージョン情報"を含む点である。

【0101】

図15に示すように、"コンバージョン情報"には、それぞれの係数を用いてクラスタリングが行われることで得られたクラスタ情報に基づいて、マーケティング部112が、広告を配信することで取得したクリック回数が記録される。

【0102】

50

図15の例では、係数="α1"の場合、コンバージョン情報は"10000"であり、係数="α2"の場合、コンバージョン情報は、"15000"である。また、係数="α3"の場合、コンバージョン情報は、"12000"である。

【0103】

このように、係数とコンバージョン情報とは相関があることから、係数とコンバージョン情報との関係を解析することで、最適な係数を導出することができる。

【0104】

<4. 係数最適化処理の流れ>

次に、評価値算出部1402による係数最適化処理の流れについて説明する。図16は、係数最適化処理の流れを示すフローチャートである。

10

【0105】

ステップS1601において、解析依頼取得部1401は、サービス提供サーバ110より解析依頼を受信したか否かを判定する。ステップS1601において、解析依頼を受信していないと判定した場合には、解析依頼を受信するまで待機する。

【0106】

一方、ステップS1601において解析依頼を受信したと判定した場合には、ステップS1602に進む。ステップS1602において、解析依頼取得部1401は、解析依頼に含まれるコンバージョン情報を抽出する。更に、解析依頼取得部1401は、抽出したコンバージョン情報を、係数と対応付けて評価情報格納部1403に格納する。

20

【0107】

ステップS1603において、評価値算出部1402は、評価情報格納部1403に格納されている、係数とコンバージョン情報とを読み出し、係数とコンバージョン情報との関係を解析する。

【0108】

ステップS1604において、評価値算出部1402は、解析結果に基づいて係数の値を最適化する。更に、評価値算出部1402は、最適化した係数を評価情報1500に新たに格納する。

【0109】

<5. まとめ>

以上の説明から明らかなように、本実施形態に係る解析装置では、評価関数全体における第2の評価関数(R/Q)の影響度を示す係数を可変とし、コンバージョン情報に基づいて変更する構成とした。

30

【0110】

これにより、マーケティング部112がクラスタ情報を利用したことによって得られた効果を、評価値に反映させることが可能となる。

【0111】

[第3の実施形態]

上記第1及び第2の実施形態では、氏名と生年月日が同一のユーザは同一人物であるとみなし、同一ユーザ(人物)の複数ユーザIDが同じクラスタに分類されるようにクラスタリングを行うこととした。ところが、氏名または生年月日が不明のユーザ同士でも、何らかの手法により同一人物であろうと推定されることがある。第3の実施形態に係る解析装置では、同一人物によって利用されていると推定された複数ユーザIDが、第1または第2の実施形態により作成されたクラスタリング条件により、同じクラスタに分類されているか否かを解析し、同一ユーザ推定の妥当性を評価する。以下、第3の実施形態について説明する。

40

【0112】

<1. サービス提供サーバの全体構成>

はじめに、本実施形態に係る解析装置を有するサービス提供システムの全体構成について説明する。図17は、サービス提供システム1700の全体構成の他の一例を示す図である。図1に示すサービス提供システム100との相違点は、同一ユーザ推定装置172

50

0を有する点である。また、サービス提供サーバ110が同一ユーザ推定情報格納部1701を有する点、解析装置1730の評価部1731の機能が解析装置120の評価部121の機能とは異なる点である。

【0113】

同一ユーザ推定装置1720には、同一ユーザ推定解析プログラムがインストールされており、当該プログラムが実行されることで、同一ユーザ推定装置1720は、同一ユーザ推定解析部1721として機能する。

【0114】

同一ユーザ推定解析部1721は、サービス提供サーバ110より送信される解析依頼に含まれるユーザ情報を用いて、同一ユーザ推定を行い、その推定結果（同一ユーザ推定結果）を取得する。また、同一ユーザ推定解析部1721は、取得した同一ユーザ推定結果を解析装置1730に送信する。更に、同一ユーザ推定解析部1721は、同一ユーザ推定結果の送信に応じて解析装置1730から受信した妥当性結果を、同一ユーザ推定結果と共に解析結果としてサービス提供サーバ110に送信する。なお、サービス提供サーバ110に送信された解析結果は、同一ユーザ推定情報格納部1701に格納される。

10

【0115】

解析装置1730には、評価プログラムがインストールされており、当該プログラムが実行されることで、解析装置1730は、評価部1731として機能する。また、解析装置1730は、第1または第2の実施形態により作成されたクラスタリング条件を含むクラスタ情報が格納されたクラスタ情報格納部122を有する。

20

【0116】

評価部1731は、同一ユーザ推定装置1720より送信された同一ユーザ推定結果について、妥当性を評価する。

【0117】

具体的には、評価部1731は、同一ユーザ推定装置1720から送信された同一ユーザ推定結果について、同じクラスタに分類されるか否かを、クラスタリング条件に基づき解析し、同一ユーザ推定解析部1721による同一ユーザ推定の妥当性を評価する。

【0118】

また、評価部1731は、同一ユーザ推定の妥当性を評価した結果（妥当性結果）を同一ユーザ推定装置1720に送信する。

30

【0119】

<2. 同一ユーザ推定装置及び解析装置の機能構成>

次に、同一ユーザ推定装置1720及び解析装置1730の機能構成について説明する。図18は、同一ユーザ推定装置1720及び解析装置1730の機能構成を示す図である。

【0120】

図18に示すように、同一ユーザ推定解析部1721は、解析依頼取得部1801、同一ユーザ推定部1802、解析結果出力部1803を有する。

【0121】

解析依頼取得部1801は、サービス提供サーバ110より送信された解析依頼を取得する。

40

【0122】

同一ユーザ推定部1802は、解析依頼に含まれるユーザ情報から同一ユーザであると推定される複数ユーザIDの情報を同一ユーザ推定結果として取得する。また、同一ユーザ推定部1802は、取得した同一ユーザ推定結果を解析装置1730に送信する。

【0123】

解析結果出力部1803は、解析装置1730より送信された妥当性結果を取得し、同一ユーザ推定部1802により取得された同一ユーザ推定結果とともに、解析結果としてサービス提供サーバ110に送信する。

【0124】

50

また、図 18 に示すように、評価部 1731 は、同一ユーザ推定結果取得部 1811、分類部 1812、妥当性評価部 1813 を有する。

【0125】

同一ユーザ推定結果取得部 1811 は、同一ユーザ推定装置 1720 より送信された同一ユーザ推定結果を取得する。分類部 1812 は、クラスタ情報格納部 122 に格納されたクラスタ情報よりクラスタリング条件を読み出し、同一ユーザ推定結果に含まれる各ユーザ ID を各クラスタに分類する。

【0126】

妥当性評価部 1813 は、同一ユーザ推定結果として、同一ユーザ（人物）であると推定された各ユーザ ID が、分類部 1812 により、同一のクラスタに分類されているか否かを解析することで、同一ユーザ推定の妥当性を評価する。また、妥当性評価部 1813 は、妥当性結果を、解析結果出力部 1803 に送信する。

10

【0127】

<3. まとめ>

以上の説明から明らかなように、本実施形態に係る解析装置では、同一ユーザ推定装置において得られた同一ユーザ推定結果の妥当性を、クラスタリング条件を用いて評価する。同一ユーザ推定結果の妥当性が高いほど同じクラスタに分類される可能性が高いので、本実施形態に係る解析装置を用いれば、同一ユーザ推定の妥当性を、定量的に評価することが可能となる。

【0128】

20

[第4の実施形態]

上記第1乃至第3の実施形態では、同一ユーザ（人物）の複数のユーザ ID を、同一ユーザ特定部がユーザの属性情報に基づいて特定する構成としたが、本発明はこれに限定されない。例えば、同一ユーザ（人物）の複数のユーザ ID の組及び組の数が既知の評価用情報を用いるようにしてもよい。

【0129】

なお、上記実施形態に挙げた構成等に、その他の要素との組み合わせなど、ここで示した構成に本発明が限定されるものではない。これらの点に関しては、本発明の趣旨を逸脱しない範囲で変更することが可能であり、その応用形態に応じて適切に定めることができる。

30

【符号の説明】

【0130】

100	: サービス提供システム
110	: サービス提供サーバ
111	: サービス提供部
112	: マーケティング部
120	: 解析装置
121	: 評価部
131、132	: ユーザ端末
300	: ユーザ情報
400	: クラスタ情報
500	: 評価情報
801	: 解析依頼取得部
802	: 評価用情報抽出部
803	: 同一ユーザ特定部
804	: クラスタリング部
805	: パラメータ算出部
806	: 指標値算出部
807	: 評価値算出部
808	: 解析結果出力部

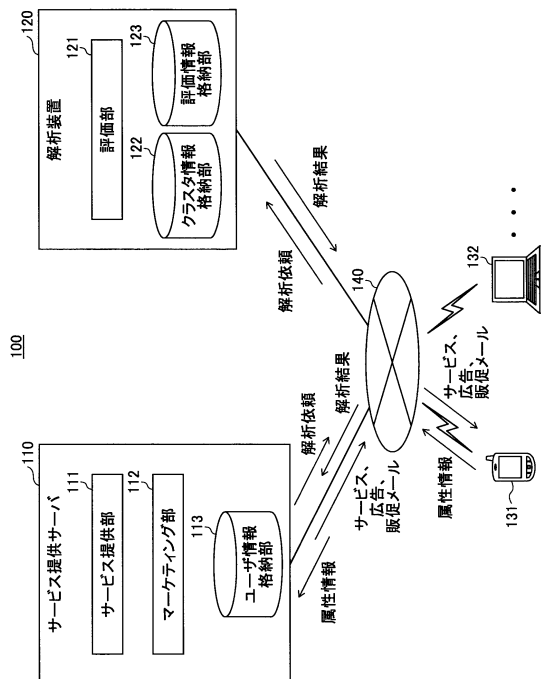
40

50

9 0 0 : 評価用情報
1 1 0 0 : クラスタ情報
1 6 0 0 : コンバージョン情報
1 7 2 0 : 同一ユーザ推定装置
1 7 2 1 : 同一ユーザ推定解析部

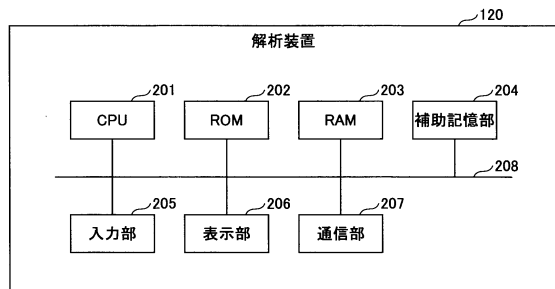
【図 1】

サービス提供システムの全体構成の一例を示す図



【図 2】

解析装置のハードウェア構成を示す図



【図 3】

ユーザ情報の一例を示す図

ユーザID	氏名	生年月日	性別	住所	購買履歴	閲覧履歴	検索履歴	...
Y0001	山田 太郎	1987年10月22日	男	...	車A	パソコンA	スポーツA	...
Y0002	田中 花子	1986年05月16日	女	...	服A	化粧品A	国内A	...
Y0003	田中 二郎	1979年02月16日	男	...	服B	書籍A	経済A	...
Y0004	山本 三郎	1991年04月18日	男	...	パソコンB	書籍A	芸能A	...
Y0005	高橋 一子	1961年07月06日	女	...	食品A	書籍B	スポーツA	...
Y0006	...	...	...	...	アクセサリA	服C	国内C	...
Y0007	山田 太郎	1987年10月22日	男	...	パソコンC	パソコンA	スポーツB	...
Y0008	高橋 花子	1974年09月01日	女	...	服A	書籍C	経済B	...
Y0009	...	...	...	...	パソコンD	車B	スポーツC	...
Y0010	高橋 花子	1988年12月19日	女	...	アクセサリA	服B	経済A	...
Y0011	田中 花子	1986年05月16日	女	...	服A	書籍D	食品C	...
Y0012	...	...	...	...	服D	化粧品B	経済C	...
Y0013	佐藤 二子	1983年11月26日	女	...	アクセサリA	服C	国内C	...
Y0014	...	...	...	...	食品C	車C	芸能B	...
...	...	...	...	...	...	...	...	...

【図 4】

クラスタ情報の一例を示す図

クラスタ番号	クラスティング条件				クラスティング結果
	購買履歴	閲覧履歴	検索履歴	...	
クラスタ1	—	—	スポーツA	...	Y0001、Y0005、...
クラスタ2	—	書籍A	—	...	Y0003、Y0004、...
クラスタ3	服A	—	—	...	Y0002、Y0011、...
クラスタ4	アクセサリA	—	—	...	Y0006、Y0010、...
...	...	...	...	...	...
クラスタN	車A	—	—	...	Y0009、...

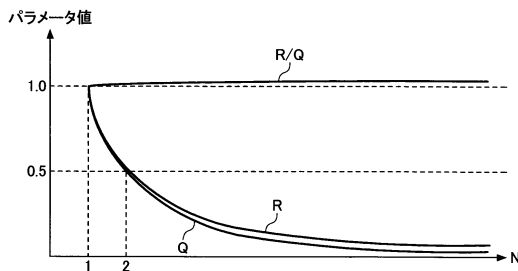
【図 5】

評価情報の一例を示す図

第1の評価関数	第2の評価関数	係数	指標値	評価値
S	R/Q	α	...	...

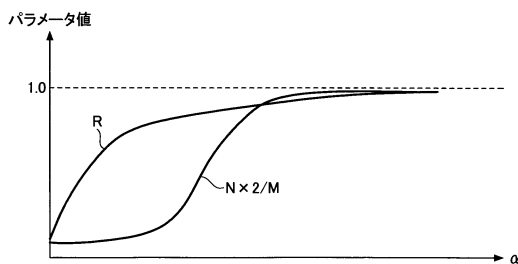
【図 6】

クラスタの数とパラメータとの関係を示す図



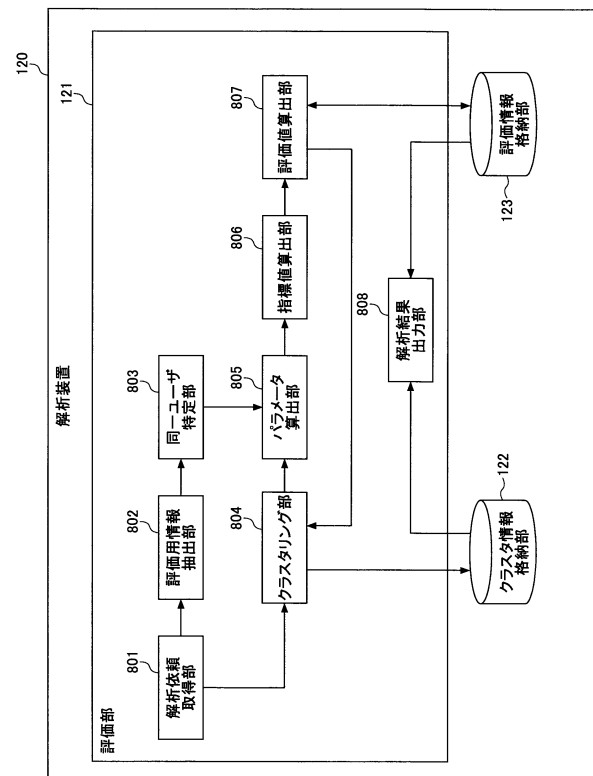
【図 7】

係数とパラメータとの関係を示す図



【図 8】

解析装置の機能構成の一例を示す図



【図 9】

評価用情報の一例を示す図

ユーザID	氏名	生年月日	性別	住所	クラスタリング条件				検索履歴	閲覧履歴	検閲履歴	...
					購買履歴	閲覧履歴	検索履歴	検閲履歴				
Y0001	山田 太郎	1987年10月22日	男	...	車A	...	...	...	スポーツA	...	...	...
Y0002	田中 花子	1986年05月16日	女	...	服A	...	...	...	化粧品A	...	...	...
Y0003	田中 二郎	1979年02月16日	男	...	服B	...	...	...	書籍A	...	...	...
Y0004	山本 三郎	1991年04月18日	男	...	パソコンB	...	...	...	書籍A	...	...	...
Y0005	高橋 一子	1981年07月06日	女	...	食品A	...	...	...	書籍B	...	...	...
Y0007	山田 太郎	1987年10月22日	男	...	パソコンC	...	...	...	パソコンA	...	...	...
Y0008	高橋 花子	1974年09月01日	女	...	服A	...	...	...	書籍C	...	...	...
Y0010	高橋 花子	1988年12月19日	女	...	アクセサリA	...	...	...	服B	...	...	...
Y0011	田中 花子	1986年05月16日	女	...	服A	...	...	...	書籍D	...	...	...
Y0013	佐藤 二子	1983年11月26日	女	...	アクセサリA	...	...	...	服C	...	...	...

【図 10】

パラメータを説明するための図

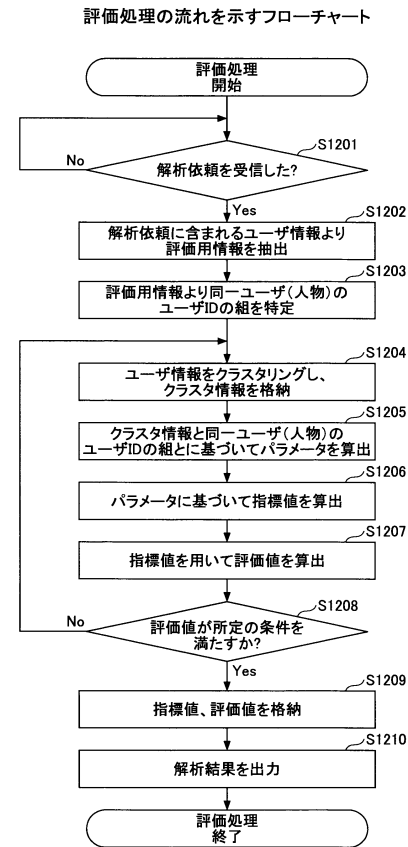
クラスタ番号	クラスタリング条件				クラスタリング結果
	購買履歴	閲覧履歴	検索履歴	検閲履歴	
クラスタ1	—	—	スポーツA	...	Y0001、Y0005、...
クラスタ2	—	書籍A	—	...	Y0003、Y0004、...
クラスタ3	服A	—	—	...	Y0002、Y0011、...
クラスタ4	アクセサリA	—	—	...	Y0006、Y0010、...
...	...	...	...	...	...
クラスタN	車A	—	—	...	Y0009、...

【図 11】

クラスタ情報の他の一例を示す図

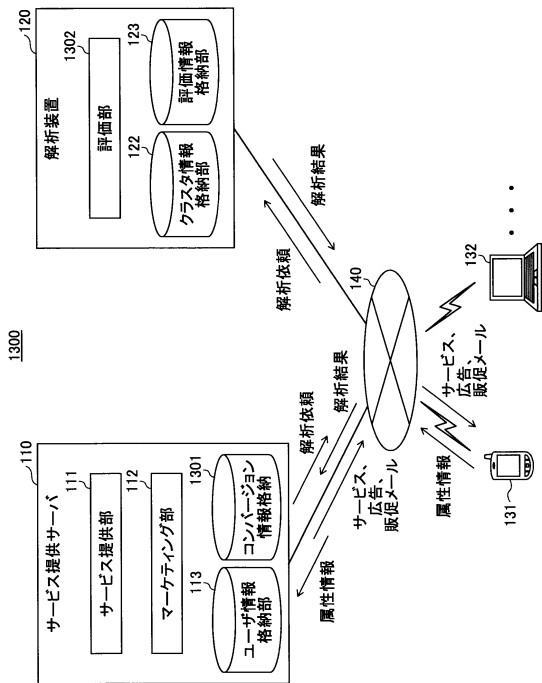
クラスタ番号	クラスタリング条件				クラスタリング結果
	購買履歴	閲覧履歴	検索履歴	検閲履歴	
クラスタ1	—	パソコンB	—	...	Y0001、Y0007、...
クラスタ2	服A	—	—	...	Y0003、Y0004、...
クラスタ3	—	服A	—	...	Y0002、Y0011、...
クラスタ4	アクセサリA	—	—	...	Y0006、Y0010、...
...	...	...	...	...	...
クラスタN	車A	—	—	...	Y0009、...

【図 12】



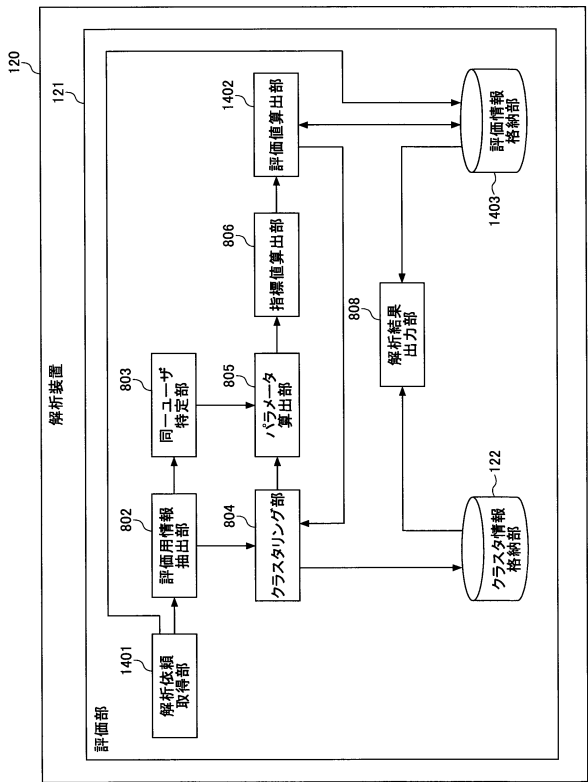
【図 1 3】

サービス提供システムの全体構成の他の一例を示す図



【図 1 4】

解析装置の機能構成の他の一例を示す図



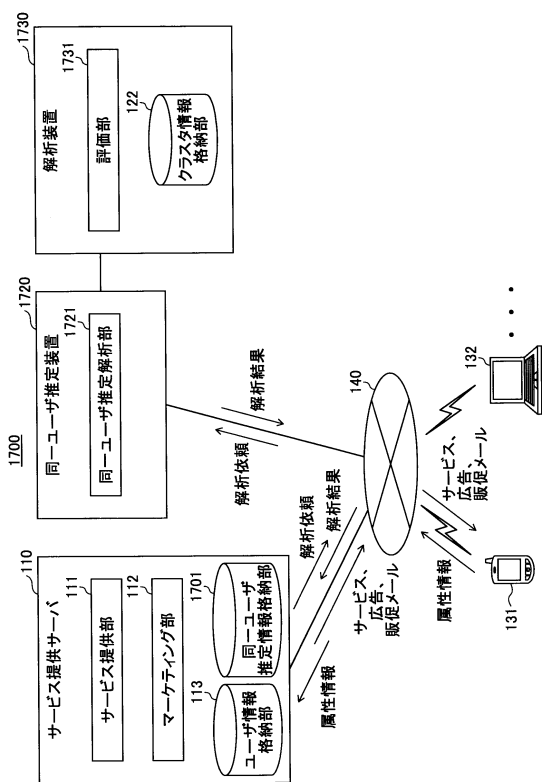
【図 1 5】

評価情報の他の一例を示す図

第1の評価関数	第2の評価関数	係数	指標値	評価値	コンバージョン情報
S	R/Q	$\alpha 1$	...	...	10000
		$\alpha 2$	...	...	15000
		$\alpha 3$	...	...	12000
		$\alpha 4$	...	...	

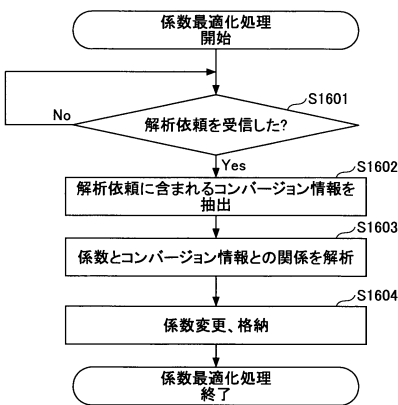
【図 1 7】

サービス提供システムの全体構成の他の一例を示す図



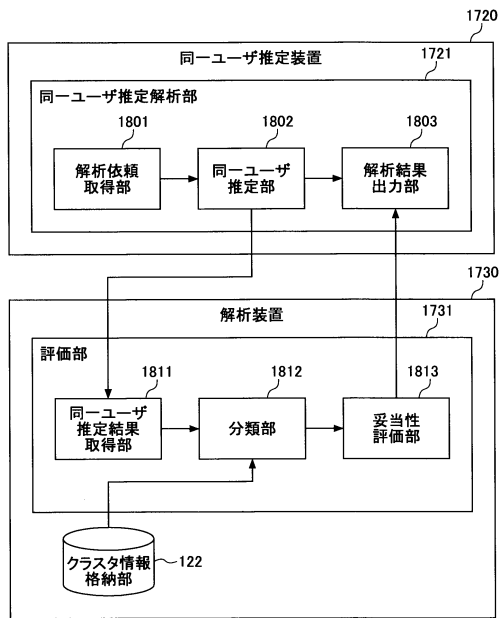
【図 1 6】

係数最適化処理の流れを示すフローチャート



【図 18】

同一ユーザ推定装置及び解析装置の機能構成の一例を示す図



フロントページの続き

- (56)参考文献 特開2009-193465 (JP, A)
特開2014-115715 (JP, A)
特表2012-518237 (JP, A)
特開2004-355404 (JP, A)
陳 伯翰、外2名、クラスタ検証による話者ダイアライゼーションの高精度化, 日本音響学会
2012年 秋季研究発表会講演論文集CD-ROM [CD-ROM], 日本, 日本音響学会
, 2012年10月 3日, p. 41-44

- (58)調査した分野(Int.Cl., DB名)
G06F 17/30