(19)日本国特許庁(JP)　(12)公開特許公報(A)　(11)特許出願公開番号
特開2012-155409
(P2012-155409A)
(43)公開日　平成24年8月16日(2012.8.16)

(51)Int.Cl.		F I			テーマコード（参考）
G06F 17/30	(2006.01)	G06F	17/30	340A	5B075
		G06F	17/30	110C	
		G06F	17/30	120A	
		G06F	17/30	350C	

審査請求 未請求　請求項の数 6　ＯＬ　（全 19 頁）

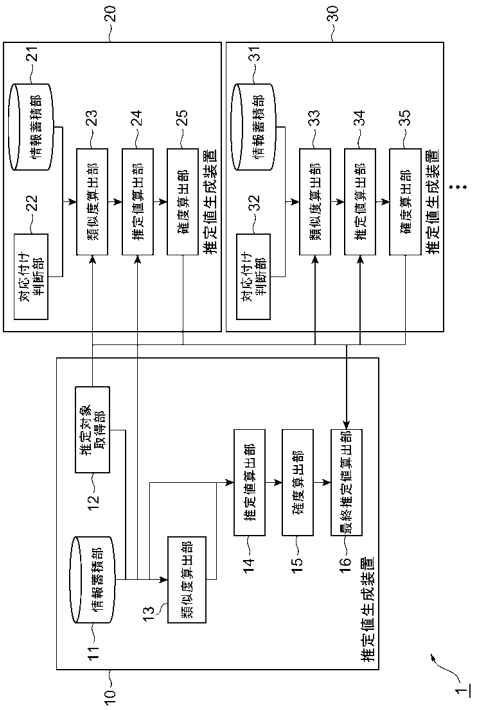
(21)出願番号　特願2011-12231(P2011-12231)
(22)出願日　平成23年1月24日(2011.1.24)

(71)出願人　392026693
株式会社エヌ・ティ・ティ・ドコモ
東京都千代田区永田町二丁目１１番１号
(74)代理人　100088155
弁理士　長谷川　芳樹
(74)代理人　100113435
弁理士　黒木　義樹
(74)代理人　100121980
弁理士　沖山　隆
(74)代理人　100128107
弁理士　深石　賢治
(72)発明者　寺田　雅之
東京都千代田区永田町二丁目１１番１号
株式会社エヌ・ティ・ティ・ドコモ内

最終頁に続く

(54)【発明の名称】評価値推定システム及び評価値推定方法

(57)【要約】
【課題】ユーザのアイテムの評価値に関する複数のデータベースから評価値を予測する場合に、ユーザのプライバシーを十分に保護すること。
【解決手段】この評価値推定システム１は、ユーザ毎のアイテム区分群に対応する評価値を表データDB1,DB2として格納する情報蓄積部１１，２１と、表データDB1,DB2中のユーザの対応付けを判断する対応付け判断部２２と、表データDB1,DB2に基づいて、対象ユーザと他のユーザとの間における評価値の類似度を計算する類似度算出部１３，２３と、対象ユーザの対象アイテム区分に対応する評価値に関する推定値s1,s2を算出する推定値算出部１４，２４と、推定値s1,s2の確度c1,c2を計算する確度算出部１５，２５と、推定値s1,s2及び確度c1,c2を基にして、対象ユーザの対象アイテム区分に対応する評価値に関する推定値sを算出する最終推定値算出部１６とを備える。
【選択図】図１



【特許請求の範囲】
【請求項１】
複数のユーザ毎の複数の評価区分毎の評価値を含む表データを分散管理し、前記表データに基づいて特定のユーザに関する特定の評価区分の評価値を推定する評価値推定システムであって、
複数のユーザ毎の第１の評価区分群に対応する評価値を第１の表データとして格納する第１の情報蓄積手段と、
複数のユーザ毎の第２の評価区分群に対応する評価値を第２の表データとして格納する第２の情報蓄積手段と、
前記第１の表データ中の前記複数のユーザと、前記第２の表データ中の前記複数のユーザとの対応付けを判断する対応付け判断手段と、
推定対象となる対象ユーザ及び前記第１の評価区分群に含まれる対象評価区分の指定を取得する推定対象取得手段と、
前記第１の表データに基づいて、前記対象ユーザと該対象ユーザ以外のユーザとの間における前記第１の評価区分群に対応する前記評価値の類似度を示す第１の類似度ベクトルを計算する第１の類似度算出手段と、
前記第２の表データに基づいて、前記対象ユーザと該対象ユーザ以外のユーザとの間における前記第２の評価区分群に対応する前記評価値の類似度を示す第２の類似度ベクトルを計算する第２の類似度算出手段と、
前記対象ユーザ以外の複数のユーザに対して計算された第１の類似度ベクトルと、前記第１の表データ中の該複数のユーザの前記対象評価区分に対応する評価値とに基づいて、前記対象ユーザの前記対象評価区分に対応する評価値に関する第１の推定値を算出する第１の推定値算出手段と、
前記第１の類似度ベクトルを算出する際に参照した前記第１の表データに応じて、前記第１の推定値に関する第１の確度を計算する第１の確度算出手段と、
前記対象ユーザ以外の複数のユーザに対して計算された第２の類似度ベクトルと、前記第１の表データ中の該複数のユーザの前記対象評価区分に対応する評価値と、前記対応付け判断手段による対応付けの判断とに基づいて、前記対象ユーザの前記対象評価区分に対応する評価値に関する第２の推定値を算出する第２の推定値算出手段と、
前記第２の類似度ベクトルを算出する際に参照した前記第２の表データに応じて、前記第２の推定値に関する第２の確度を計算する第２の確度算出手段と、
前記第１の推定値及び第１の確度と、前記第２の推定値及び第２の確度とを基にして、前記対象ユーザの前記対象評価区分に対応する評価値に関する第３の推定値を算出する第３の推定値算出手段と、
を備えることを特徴とする評価値推定システム。
【請求項２】
前記第１の情報蓄積手段と、前記推定対象取得手段と、前記第１の類似度算出手段と、前記第１の推定値算出手段と、前記第１の確度算出手段と、前記第３の推定値算出手段とを有する第１の推定値生成装置と、
前記第２の情報蓄積手段と、前記対応付け判断手段と、前記第２の類似度算出手段と、前記第２の推定値算出手段と、前記第２の確度算出手段とを有する第２の推定値生成装置と、
を備えることを特徴とする請求項１記載の評価値推定システム。
【請求項３】
処理対象の値ｘ，ｙに対して、下記式（１）；
Ｅ（ｘ）・Ｅ（ｙ）＝Ｅ（ｘ＋ｙ）　　…（１）
及び下記式（２）；
Ｄ（Ｅ（ｘ））＝ｘ　　…（２）
の関係を満たすような暗号化関数Ｅ及び復号関数Ｄを用いて、加法準同型性の性質を有する暗号化演算及び復号演算を行う暗号化手段及び復号手段と、

前記暗号化手段を用いて、前記第１の表データ中の前記複数のユーザの前記対象評価区分に対応する評価値から、暗号化ベクトルを生成する暗号化ベクトル生成手段と、をさらに備え、

前記第２の推定値算出手段は、前記暗号化ベクトルと前記第２の類似度ベクトルとを参照し、前記暗号化ベクトルの各要素を底とし、前記第２の類似度ベクトルの前記各要素に対応する値を指数とした累乗を計算し、前記累乗の積を計算することにより、暗号化された前記第２の推定値を算出し、

前記第３の推定値算出手段は、前記復号手段を用い、暗号化された前記第２の推定値を復号して前記第２の推定値を取得する、

ことを特徴とする請求項１記載の評価値推定システム。

【請求項４】

前記第１の推定値生成装置は、

処理対象の値ｘ，ｙに対して、下記式（１）；

Ｅ（ｘ）・Ｅ（ｙ）＝Ｅ（ｘ＋ｙ）　　…（１）

及び下記式（２）；

Ｄ（Ｅ（ｘ））＝ｘ　　…（２）

の関係を満たすような暗号化関数Ｅ及び復号関数Ｄを用いて、加法準同型性の性質を有する暗号化演算及び復号演算を行う暗号化手段及び復号手段と、

前記暗号化手段を用いて、前記第１の表データ中の前記複数のユーザの前記対象評価区分に対応する評価値から、暗号化ベクトルを生成する暗号化ベクトル生成手段と、をさらに有し、

前記第２の推定値生成装置の前記第２の推定値算出手段は、前記暗号化ベクトルと前記第２の類似度ベクトルとを参照し、前記暗号化ベクトルの各要素を底とし、前記第２の類似度ベクトルの前記各要素に対応する値を指数とした累乗を計算し、前記累乗の積を計算することにより、暗号化された前記第２の推定値を算出し、

前記第１の推定値生成装置の前記第３の推定値算出手段は、前記復号手段を用い、暗号化された前記第２の推定値を復号して前記第２の推定値を取得する、

ことを特徴とする請求項２記載の評価値推定システム。

【請求項５】

複数のユーザ毎の複数の評価区分毎の評価値を含む表データを分散管理し、前記表データに基づいて特定のユーザに関する特定の評価区分の評価値を推定する評価値推定方法であって、

第１の情報蓄積手段が、複数のユーザ毎の第１の評価区分群に対応する評価値を第１の表データとして格納する第１の情報蓄積ステップと、

第２の情報蓄積手段が、複数のユーザ毎の第２の評価区分群に対応する評価値を第２の表データとして格納する第２の情報蓄積ステップと、

対応付け判断手段が、前記第１の表データ中の前記複数のユーザと、前記第２の表データ中の前記複数のユーザとの対応付けを判断する対応付け判断ステップと、

推定対象取得手段が、推定対象となる対象ユーザ及び前記第１の評価区分群に含まれる対象評価区分の指定を取得する推定対象取得ステップと、

第１の類似度算出手段が、前記第１の表データに基づいて、前記対象ユーザと該対象ユーザ以外のユーザとの間における前記第１の評価区分群に対応する前記評価値の類似度を示す第１の類似度ベクトルを計算する第１の類似度算出ステップと、

第２の類似度算出手段が、前記第２の表データに基づいて、前記対象ユーザと該対象ユーザ以外のユーザとの間における前記第２の評価区分群に対応する前記評価値の類似度を示す第２の類似度ベクトルを計算する第２の類似度算出ステップと、

第１の推定値算出手段が、前記対象ユーザ以外の複数のユーザに対して計算された第１の類似度ベクトルと、前記第１の表データ中の該複数のユーザの前記対象評価区分に対応する評価値とに基づいて、前記対象ユーザの前記対象評価区分に対応する評価値に関する第１の推定値を算出する第１の推定値算出ステップと、

第１の確度算出手段が、前記第１の類似度ベクトルを算出する際に参照した前記第１の表データに応じて、前記第１の推定値に関する第１の確度を計算する第１の確度算出ステップと、

第２の推定値算出手段が、前記対象ユーザ以外の複数のユーザに対して計算された第２の類似度ベクトルと、前記第１の表データ中の該複数のユーザの前記対象評価区分に対応する評価値と、前記対応付け判断手段による対応付けの判断とに基づいて、前記対象ユーザの前記対象評価区分に対応する評価値に関する第２の推定値を算出する第２の推定値算出ステップと、

第２の確度算出手段が、前記第２の類似度ベクトルを算出する際に参照した前記第２の表データに応じて、前記第２の推定値に関する第２の確度を計算する第２の確度算出ステップと、

第３の推定値算出手段が、前記第１の推定値及び第１の確度と、前記第２の推定値及び第２の確度とを基にして、前記対象ユーザの前記対象評価区分に対応する評価値に関する第３の推定値を算出する第３の推定値算出ステップと、

を備えることを特徴とする評価値推定方法。

【請求項６】

暗号化手段及び復号手段が、それぞれ、処理対象の値ｘ，ｙに対して、下記式（１）；

Ｅ（ｘ）・Ｅ（ｙ）＝Ｅ（ｘ＋ｙ）　　…（１）

及び下記式（２）；

Ｄ（Ｅ（ｘ））＝ｘ　　…（２）

の関係を満たすような暗号化関数Ｅ及び復号関数Ｄを用いて、加法準同型性の性質を有する暗号化演算及び復号演算を行う暗号化ステップ及び復号ステップと、

暗号化ベクトル生成手段が、前記暗号化手段を用いて、前記第１の表データ中の前記複数のユーザの前記対象評価区分に対応する評価値から、暗号化ベクトルを生成する暗号化ベクトル生成ステップと、をさらに備え、

前記第２の推定値算出ステップでは、前記暗号化ベクトルと前記第２の類似度ベクトルとを参照し、前記暗号化ベクトルの各要素を底とし、前記第２の類似度ベクトルの前記各要素に対応する値を指数とした累乗を計算し、前記累乗の積を計算することにより、暗号化された前記第２の推定値を算出し、

前記第３の推定値算出ステップでは、前記復号手段を用い、暗号化された前記第２の推定値を復号して前記第２の推定値を取得する、

ことを特徴とする請求項５記載の評価値推定方法。

【発明の詳細な説明】

【技術分野】

【０００１】

本発明は、評価値推定システム及び評価値推定方法に関するものである。

【背景技術】

【０００２】

従来から、複数のユーザによって複数の商品等のアイテムが評価付けされた情報を管理するデータベースを用いて、他のユーザの評価に関する値を基に特定のユーザの評価されていないアイテムの評価値を予測する「協調フィルタリング」という手法が知られている。

【０００３】

この手法においては、データベースは表データで管理されており、例えば、その表データの各行はそれぞれ各ユーザに対応し、各列はそれぞれ各商品区分に対応している。また、この表データの各セルには、対応する行のユーザによる対応する列の商品区分に対する評価付けの値（評価値）が格納されている。なお、全てのユーザが全ての商品に対して評価をしているとは限らないため、表データ中の一部のセルには評価値が格納されていない場合が起こり得る。この協調フィルタリングでは、このような評価されていない特定ユーザの特定アイテムに関する評価値を、他のユーザを含めた複数のユーザの様々なアイテム

に関する評価値を基に推測するというものである。

【発明の概要】

【発明が解決しようとする課題】

【０００４】

ここで、複数のデータ管理組織がそれぞれ異なる商品群についての評価値を管理することが通常行われる。このような場合に、協調フィルタリングによってユーザの評価値を推測しようとすると、単純な手法としては各組織の管理しているデータベースを一つに統合し、その統合したデータベースから評価値の予測値を計算することが考えられる。しかしながら、そのような手法では、複数のユーザのプライバシーが保護されないという懸念が生じる。これは、各組織の管理している各ユーザの各商品群に関する評価値は、重要な秘密情報であり、プライバシー保護の対象になりうるからである。

【０００５】

そこで、本発明は、かかる課題に鑑みて為されたものであり、ユーザのアイテムの評価値に関する複数のデータベースから評価値を予測する場合に、ユーザのプライバシーを十分に保護することが可能な評価値推定システム及び評価値推定方法を提供することを目的とする。

【課題を解決するための手段】

【０００６】

上記課題を解決するため、本発明の評価値推定システムは、複数のユーザ毎の複数の評価区分毎の評価値を含む表データを分散管理し、表データに基づいて特定のユーザに関する特定の評価区分の評価値を推定する評価値推定システムであって、複数のユーザ毎の第１の評価区分群に対応する評価値を第１の表データとして格納する第１の情報蓄積手段と、複数のユーザ毎の第２の評価区分群に対応する評価値を第２の表データとして格納する第２の情報蓄積手段と、第１の表データ中の複数のユーザと、第２の表データ中の複数のユーザとの対応付けを判断する対応付け判断手段と、推定対象となる対象ユーザ及び第１の評価区分群に含まれる対象評価区分の指定を取得する推定対象取得手段と、第１の表データに基づいて、対象ユーザと該対象ユーザ以外のユーザとの間における第１の評価区分群に対応する評価値の類似度を示す第１の類似度ベクトルを計算する第１の類似度算出手段と、第２の表データに基づいて、対象ユーザと該対象ユーザ以外のユーザとの間における第２の評価区分群に対応する評価値の類似度を示す第２の類似度ベクトルを計算する第２の類似度算出手段と、対象ユーザ以外の複数のユーザに対して計算された第１の類似度ベクトルと、第１の表データ中の該複数のユーザの対象評価区分に対応する評価値とに基づいて、対象ユーザの対象評価区分に対応する評価値に関する第１の推定値を算出する第１の推定値算出手段と、第１の類似度ベクトルを算出する際に参照した第１の表データに応じて、第１の推定値に関する第１の確度を計算する第１の確度算出手段と、対象ユーザ以外の複数のユーザに対して計算された第２の類似度ベクトルと、第１の表データ中の該複数のユーザの対象評価区分に対応する評価値と、対応付け判断手段による対応付けの判断とに基づいて、対象ユーザの対象評価区分に対応する評価値に関する第２の推定値を算出する第２の推定値算出手段と、第２の類似度ベクトルを算出する際に参照した第２の表データに応じて、第２の推定値に関する第２の確度を計算する第２の確度算出手段と、第１の推定値及び第１の確度と、第２の推定値及び第２の確度とを基にして、対象ユーザの対象評価区分に対応する評価値に関する第３の推定値を算出する第３の推定値算出手段と、を備える。

【０００７】

或いは、本発明の評価値推定方法は、複数のユーザ毎の複数の評価区分毎の評価値を含む表データを分散管理し、表データに基づいて特定のユーザに関する特定の評価区分の評価値を推定する評価値推定方法であって、第１の情報蓄積手段が、複数のユーザ毎の第１の評価区分群に対応する評価値を第１の表データとして格納する第１の情報蓄積ステップと、第２の情報蓄積手段が、複数のユーザ毎の第２の評価区分群に対応する評価値を第２の表データとして格納する第２の情報蓄積ステップと、対応付け判断手段が、第１の表デ

ータ中の複数のユーザと、第２の表データ中の複数のユーザとの対応付けを判断する対応付け判断ステップと、推定対象取得手段が、推定対象となる対象ユーザ及び第１の評価区分群に含まれる対象評価区分の指定を取得する推定対象取得ステップと、第１の類似度算出手段が、第１の表データに基づいて、対象ユーザと該対象ユーザ以外のユーザとの間における第１の評価区分群に対応する評価値の類似度を示す第１の類似度ベクトルを計算する第１の類似度算出ステップと、第２の類似度算出手段が、第２の表データに基づいて、対象ユーザと該対象ユーザ以外のユーザとの間における第２の評価区分群に対応する評価値の類似度を示す第２の類似度ベクトルを計算する第２の類似度算出ステップと、第１の推定値算出手段が、対象ユーザ以外の複数のユーザに対して計算された第１の類似度ベクトルと、第１の表データ中の該複数のユーザの対象評価区分に対応する評価値とに基づいて、対象ユーザの対象評価区分に対応する評価値に関する第１の推定値を算出する第１の推定値算出ステップと、第１の確度算出手段が、第１の類似度ベクトルを算出する際に参照した第１の表データに応じて、第１の推定値に関する第１の確度を計算する第１の確度算出ステップと、第２の推定値算出手段が、対象ユーザ以外の複数のユーザに対して計算された第２の類似度ベクトルと、第１の表データ中の該複数のユーザの対象評価区分に対応する評価値と、対応付け判断手段による対応付けの判断とに基づいて、対象ユーザの対象評価区分に対応する評価値に関する第２の推定値を算出する第２の推定値算出ステップと、第２の確度算出手段が、第２の類似度ベクトルを算出する際に参照した第２の表データに応じて、第２の推定値に関する第２の確度を計算する第２の確度算出ステップと、第３の推定値算出手段が、第１の推定値及び第１の確度と、第２の推定値及び第２の確度とを基にして、対象ユーザの対象評価区分に対応する評価値に関する第３の推定値を算出する第３の推定値算出ステップと、を備える。

【０００８】

このような評価値推定システム及び評価値推定方法によれば、異なる評価区分群に対するユーザによる評価値に関する２つの表データが別々に格納されており、評価値の推定対象の対象ユーザ及び対象評価区分が指定されると、第１の類似度算出手段によって、第１の表データを基に対象ユーザとそれ以外のユーザの評価値の類似度を示す第１の類似度ベクトルが計算され、第２の類似度算出手段によって、第２の表データを基に対象ユーザとそれ以外のユーザの評価値の類似度を示す第２の類似度ベクトルが計算され、第１の推定値算出手段及び第１の確度算出手段によって、第１の類似度ベクトルと第１の表データに含まれる対象評価区分の評価値とを基に第１の推定値及びその推定値の確度が計算され、第２の推定値算出手段及び第２の確度算出手段によって、第２の類似度ベクトルと第１の表データに含まれる対象評価区分の評価値とを基に、第１の表データと第２の表データのユーザの対応付けの判断結果を用いて第２の推定値及びその推定値の確度が計算され、第１の推定度及びその確度と、第２の推定値及びその確度とから、最終的な評価値の推定値である第３の推定値が算出される。これにより、第１の表データを管理するデータベース側から第２の表データを管理するデータベース側に対して、対象評価区分に関する評価値の通知が必要な以外には、その他の評価区分に関する評価値の通知の必要がないので、対象評価区分以外の評価区分の評価値に関するユーザのプライバシーが十分に保護される。

【０００９】

本発明の評価値推定システムは、第１の情報蓄積手段と、推定対象取得手段と、第１の類似度算出手段と、第１の推定値算出手段と、第１の確度算出手段と、第３の推定値算出手段とを有する第１の推定値生成装置と、第２の情報蓄積手段と、対応付け判断手段と、第２の類似度算出手段と、第２の推定値算出手段と、第２の確度算出手段とを有する第２の推定値生成装置と、を備えていてもよい。この場合、第１の表データを管理する第１の推定値生成装置から第２の表データを管理する第２の推定値生成装置に対して、対象評価区分に関する評価値の通知が必要な以外には、その他の評価区分に関する評価値の通知の必要がないので、対象評価区分以外の評価区分の評価値に関するユーザのプライバシーが十分に保護される。

【００１０】

また、処理対象の値ｘ，ｙに対して、下記式（１）；
Ｅ（ｘ）・Ｅ（ｙ）＝Ｅ（ｘ＋ｙ）　…（１）
及び下記式（２）；
Ｄ（Ｅ（ｘ））＝ｘ　…（２）
の関係を満たすような暗号化関数Ｅ及び復号関数Ｄを用いて、加法準同型性の性質を有する暗号化演算及び復号演算を行う暗号化手段及び復号手段と、暗号化手段を用いて、第１の表データ中の複数のユーザの対象評価区分に対応する評価値から、暗号化ベクトルを生成する暗号化ベクトル生成手段と、をさらに備え、第２の推定値算出手段は、暗号化ベクトルと第２の類似度ベクトルとを参照し、暗号化ベクトルの各要素を底とし、第２の類似度ベクトルの各要素に対応する値を指数とした累乗を計算し、累乗の積を計算することにより、暗号化された第２の推定値を算出し、第３の推定値算出手段は、復号手段を用い、暗号化された第２の推定値を復号して第２の推定値を取得する、ことも好ましい。

【００１１】
また、第１の推定値生成装置は、処理対象の値ｘ，ｙに対して、下記式（１）；
Ｅ（ｘ）・Ｅ（ｙ）＝Ｅ（ｘ＋ｙ）　…（１）
及び下記式（２）；
Ｄ（Ｅ（ｘ））＝ｘ　…（２）
の関係を満たすような暗号化関数Ｅ及び復号関数Ｄを用いて、加法準同型性の性質を有する暗号化演算及び復号演算を行う暗号化手段及び復号手段と、暗号化手段を用いて、第１の表データ中の複数のユーザの対象評価区分に対応する評価値から、暗号化ベクトルを生成する暗号化ベクトル生成手段と、をさらに有し、第２の推定値生成装置の第２の推定値算出手段は、暗号化ベクトルと第２の類似度ベクトルとを参照し、暗号化ベクトルの各要素を底とし、第２の類似度ベクトルの各要素に対応する値を指数とした累乗を計算し、累乗の積を計算することにより、暗号化された第２の推定値を算出し、第１の推定値生成装置の第３の推定値算出手段は、復号手段を用い、暗号化された第２の推定値を復号して第２の推定値を取得する、ことも好ましい。

【００１２】
また、暗号化手段及び復号手段が、それぞれ、処理対象の値ｘ，ｙに対して、下記式（１）；
Ｅ（ｘ）・Ｅ（ｙ）＝Ｅ（ｘ＋ｙ）　…（１）
及び下記式（２）；
Ｄ（Ｅ（ｘ））＝ｘ　…（２）
の関係を満たすような暗号化関数Ｅ及び復号関数Ｄを用いて、加法準同型性の性質を有する暗号化演算及び復号演算を行う暗号化ステップ及び復号ステップと、暗号化ベクトル生成手段が、暗号化手段を用いて、第１の表データ中の複数のユーザの対象評価区分に対応する評価値から、暗号化ベクトルを生成する暗号化ベクトル生成ステップと、をさらに備え、第２の推定値算出ステップでは、暗号化ベクトルと第２の類似度ベクトルとを参照し、暗号化ベクトルの各要素を底とし、第２の類似度ベクトルの各要素に対応する値を指数とした累乗を計算し、累乗の積を計算することにより、暗号化された第２の推定値を算出し、第３の推定値算出ステップでは、復号手段を用い、暗号化された第２の推定値を復号して第２の推定値を取得する、ことも好ましい。

【００１３】
この場合、加法準同型性の性質を有する暗号化演算及び復号演算を行う暗号化手段及び復号手段と、第１の表データ中の対象評価区分の評価値を暗号化した暗号化ベクトルを生成する暗号化ベクトル生成手段が備えられ、第２の推定値算出手段によって、暗号化ベクトルを用いて暗号化された第２の推定値が算出され、第３の推定値算出手段によって、暗号化された第２の推定値が復号される。これによって、２つのデータベース間で評価値が通知されずに、対象ユーザの対象評価区分に関する評価値が予測されるので、複数の評価区分に関する評価値に関するユーザのプライバシーがより一層保護されるとともに、複数のデータベースを用いても正確な予測値が得られる。

【発明の効果】
【００１４】
本発明によれば、ユーザのアイテムの評価値に関する複数のデータベースから評価値を予測する場合に、ユーザのプライバシーを十分に保護することができる。
【図面の簡単な説明】
【００１５】
【図１】本発明の好適な一実施形態にかかる評価値推定システム１の概略構成図である。
【図２】図１の情報蓄積部１１に格納されている表データDB1のデータ構成の一例を示す図である。
【図３】図１の情報蓄積部２１に格納されている表データDB2のデータ構成の一例を示す図である。
【図４】図１の対応付け判断部２２が管理するテーブルデータの構成を示す図である。
【図５】図２の表データDB1を対象に計算された類似度ベクトルsim1の計算結果を示す図である。
【図６】図３の表データDB2を対象に計算された類似度ベクトルsim2の計算結果を示す図である。
【図７】図１の評価値推定システム１の動作を示すフローチャートである。
【図８】本発明の変形例である評価値推定システム１０１の概略構成図である。
【図９】図８の評価値推定システム１０１の動作を示すフローチャートである。
【発明を実施するための形態】
【００１６】
以下、図面とともに本発明による評価値推定システム及び評価値推定方法の好適な実施形態について詳細に説明する。なお、図面の説明においては同一要素には同一符号を付し、重複する説明を省略する。
【００１７】
図１は、本発明の好適な一実施形態にかかる評価値推定システム１の概略構成図である。同図に示す評価値推定システム１は、ＬＡＮ、ＷＡＮ等のネットワークを介して相互にデータの送受信が可能なように接続された２つ以上のデータベース管理装置である推定値生成装置１０、推定値生成装置２０、推定値生成装置３０、…によって構成されている。この評価値推定システム１は、複数のユーザによる複数の商品等のアイテムに関する評価値を、複数のデータベースによって分散管理し、これらの複数のデータベースを参照することによって特定ユーザの特定アイテムに関する評価値を予測するコンピュータシステムである。以下、評価値推定システム１を構成する各装置の機能構成について詳細に説明する。ここで、推定値生成装置１０を除く推定値生成装置２０、推定値生成装置３０、…の各装置の機能構成は同一であるので、重複する説明を省略するとともに、以下の説明においては、評価値推定システム１は、推定値生成装置１０、推定値生成装置２０、及び推定値生成装置３０の３つのデータベース管理装置によって構成されているものとする。もちろん、評価値推定システム１は、４つ以上のデータベース管理装置を含んで構成されていてもよい。
【００１８】
推定値生成装置１０は、情報蓄積部（第１の情報蓄積手段）１１、推定対象取得部（推定対象取得手段）１２、類似度算出部（第１の類似度算出手段）１３、推定値算出部（第１の推定値算出手段）１４、確度算出部（第１の確度算出手段）１５、及び最終推定値算出部（第３の推定値算出手段）１６を備えている。この推定値生成装置１０は、予測対象のアイテムを含む複数のアイテムに関する予測値を管理するデータベース管理装置である。
【００１９】
情報蓄積部１１には、予め、複数のユーザ毎の複数のアイテム区分（評価区分）毎に対応する評価値が表データDB1として格納されている。このアイテム区分は、商品等のアイテムを評価する単位を示す区分情報であり、アイテムの種類を示す識別情報である。図２

には、情報蓄積部１１に格納されている表データDB1のデータ構成の一例を示している。同図に示すように、情報蓄積部１１には、複数のユーザを識別するユーザ識別情報“user_1”、“user_2”、“user_3”、“user_4”毎に、アイテム区分群“Item_1”、“Item_2”、“Item_3”に対応付けて、評価値を示す数値データが格納されている。詳細には、情報蓄積部１１に格納された表データDB1において、各行が各ユーザに対応し、各列が各アイテム区分に対応している。例えば、表データDB1の３行２列のセルDB1[3,2]は、ユーザ“user_3”によるアイテム区分“Item_2”に対する評価値“５”を示している（以下、表データDB1におけるm行n列のセルを、“DB1[m,n]”と表記する）。これらの表データDB1のセルには、既にユーザによって対応するアイテム区分の評価が反映されている場合には数値データが格納され、ユーザによって対応するアイテム区分の評価が反映されていない場合には数値データは格納されず、“スペース”や“NULL値”等によって空白セルが設定されている。このような空白セルに対応する評価値が、評価値推定システム１における推定対象となりうる。

【００２０】

推定対象取得部１２は、評価値を予測する対象となる対象ユーザと対象アイテム区分の指定を取得する。詳細には、推定対象取得部１２は、情報蓄積部１１に格納された表データDB1の各行が示す複数のユーザの中から選択された対象ユーザを示す対象ユーザ識別情報と、その表データDB1の各列が示す複数のアイテム区分の中から選択された対象アイテム区分を示す対象区分識別情報を取得する。例えば、対象ユーザ及び対象アイテム区分として、表データDB1中の４行１列のセルDB1[4,1]（図２中の斜線で示すセル）に対応するユーザ及びアイテム区分が選択された場合には、推定対象取得部１２は、対象ユーザ識別情報“user_4”及び対象区分識別情報“Item_1”を取得する。なお、これらの対象ユーザと対象アイテム区分の指定は、外部からの入力を受け付けることによって取得してもよいし、外部からネットワークを介して受信してもよい。ここで、推定対象取得部１２によって取得される対象アイテム区分に対応する各ユーザの評価値は、推定値生成装置１０の情報蓄積部１１にのみ格納されており、後述する推定値生成装置２０，３０の情報蓄積部２１，３１には格納されていないものとする。

【００２１】

類似度算出部１３は、情報蓄積部１１に格納された表データDB1を参照しながら、対象ユーザとそれ以外のユーザとの間におけるアイテム区分群に亘った評価値の類似度を計算する。すなわち、対象ユーザ“user_i”（iは整数）と、それ以外のユーザ“user_k”（kはi以外の整数）との類似度は、表データDB1のk行目のセルによって構成されるベクトルDB1[k,*]と、表データDB1のi行目のセルによって構成されるベクトルDB1[i,*]との二乗誤差和の逆数を、下記式（３）；

$$sim1_\{k,i\} = 1 / \{(DB1[i,1] - DB1[k,1])^2 + (DB1[i,2] - DB1[k,2])^2 + \cdots + (DB1[i,n] - DB1[k,n])^2\} \quad \cdots (3)$$

によって計算した後に、全てのユーザ“user_k”に対応する二乗誤差和の逆数の総和が“1”になるように正規化された類似度sim1_{k,i}として計算される。この類似度sim1_{k,i}は、表データDB1の行に対応する対象ユーザ以外の全てのユーザ“user_k”に関して計算されることにより、類似度ベクトルsim1_{k,i}（kはi以外の整数）として得られる。

【００２２】

推定値算出部１４は、表データDB1に含まれる対象ユーザ以外のユーザ“user_k”（kはi以外の整数）及び対象アイテム区分“Item_j”(jは整数)に対応する評価値と、類似度ベクトルsim1_{k,i}とを基にして、対象ユーザ及び対象アイテム区分に対応する評価値の推定値（第１の推定値）を計算する。詳細には、推定値算出部１４は、ユーザ“user_k”及び対象アイテム区分“Item_j”に対応する評価値によって構成されるベクトルDB1[*,j]と、類似度ベクトルsim1_{k,i}との内積を、下記式（４）；

$$s1 = DB1[1,j]*sim1_\{1,i\} + DB1[2,j]*sim1_\{2,i\} + DB1[m,j]*sim1_\{m,i\} \quad \cdots (4)$$

を用いて計算し、この内積を対象ユーザ及び対象アイテム区分に対応する評価値の推定値

s1とする。

【００２３】

確度算出部１５は、類似度算出部１３が類似度ベクトルsim1_{k,i}を算出する際に参照した表データDB1を基にして、評価値の推定値s1に関する確度（第１の確度）を計算する。すなわち、確度算出部１５は、類似度ベクトルsim1_{k,i}を計算する際に表データDB1から参照されたベクトルDB1[k,*]、ベクトルDB1[i,*]に対応するアイテム区分の数、言い換えれば、表データDB1の列数に比例する値を確度c1として計算する。また、確度算出部１５は、表データDB1に含まれる評価値の数、言い換えれば、表データDB1の空白セル以外のセルの数に比例する数を確度c1として計算してもよい。

【００２４】

最終推定値算出部１６は、推定値算出部１４及び確度算出部１５によって計算された評価値に関する推定値s1及び確度c1と、後述する推定値生成装置２０，３０によって別の表データDB2,DB3を基に計算されて受信された推定値s2,s3及びそれぞれの確度c2,c3とを基にして、対象ユーザの対象アイテム区分に対応する評価値に関する最終的な推定値s（第３の推定値）を算出する。詳細には、最終推定値算出部１６は、確度c1,c2,c3を重みとした推定値s1,s2,s3の重み付け平均を、下記式（５）；

$$s = (c1*s1 + c2*s2 + c3*s3) / (c1 + c2 + c3) \quad \cdots (5)$$

を用いて計算し、その値を推定値sとして求める。このようにして算出された推定値sは、３つの表データDB1,DB2,DB3から求められた対象ユーザ”user_i”及び対象アイテム区分“Item_j”に対する評価値の最終的な予測値となる。この予測値は、その値が高ければユーザ“user_i”はアイテム区分“Item_j”に属するアイテムを高く評価付けすると予測されるので、このユーザに対して当該アイテムを推奨すべきかどうかの判断基準となる情報として利用可能となる。一方で、予測値が低ければ、ユーザ“user_i”はアイテム区分“Item_j”に属するアイテムを低く評価付けすると予測されることになる。

【００２５】

また、推定値生成装置２０は、情報蓄積部（第２の情報蓄積手段）２１、対応付け判断部（対応付け判断手段）２２、類似度算出部（第２の類似度算出手段）２３、推定値算出部（第２の推定値算出手段）２４、及び確度算出部（第２の確度算出手段）２５を備えている。この推定値生成装置２０は、予測対象のアイテム以外の複数のアイテムに関する予測値を管理するデータベース管理装置である。

【００２６】

情報蓄積部２１には、表データDB1に含まれるアイテム区分とは異なるアイテム区分毎に、表データDB1の行に対応する複数のユーザに関する評価値が、表データDB2として格納されている。図３には、情報蓄積部２１に格納されている表データDB2のデータ構成の一例を示している。同図に示すように、情報蓄積部２１には、複数のユーザを識別するユーザ識別情報“user_1”、“user_2”、“user_3”、“user_4”毎に、アイテム区分群“Item_4”、“Item_5”に対応付けて、評価値を示す数値データが格納されている。例えば、表データDB2の２行２列のセルDB2[2,2]は、ユーザ“user_2”によるアイテム区分“Item_5”に対する評価値“４”を示している（以下、表データDB2におけるm行n列のセルを、“DB2[m,n]”と表記する）。これらの表データDB2に含まれるアイテム区分は、表データDB1に含まれるアイテム区分と完全に異なっていてもよく、一部のアイテム区分が重複していてもよい。ただし、評価値推定システム１における評価値の推定対象となりうるアイテム区分は、表データDB1にのみ存在しており、表データDB2には存在していないものとする。

【００２７】

対応付け判断部２２は、表データDB1の各行が示す各ユーザと、表データDB2の各行が示す各ユーザとの対応付けを判断する。例えば、対応付け判断部２２は、図４に示すようなテーブルデータを管理し、このテーブルデータによってユーザ識別情報“user_m”によって特定されるユーザがそれぞれの表データDB1,DB2の何行目に対応するかが管理されている。そして、対応付け判断部２２は、このようなテーブルデータを参照することにより、各ユーザ“user_m”に関する２つの表データDB1,DB2における行間の対応付けを判断する

。

【００２８】

類似度算出部２３は、推定対象取得部１２から対象ユーザ識別情報“user_i”及び対象区分識別情報“Item_j”を取得し、情報蓄積部２１に格納された表データDB2を参照しながら、対象ユーザとそれ以外のユーザとの間におけるアイテム区分群に亘った評価値の類似度を計算する。すなわち、類似度算出部２３は、類似度算出部１３と同様な計算方法を用いて、対象ユーザ“user_i”（iは整数）と、それ以外のユーザ“user_k”（kはi以外の整数）との類似度を示す類似度ベクトルsim2_{k,i}（kはi以外の整数）を計算する。このとき、ユーザ“user_i”，“user_k”に対応する表データDB2の行は、対応付け判断部２２による判断結果を基に特定される。

【００２９】

推定値算出部２４は、表データDB1に含まれる対象ユーザ以外のユーザ“user_k”（kはi以外の整数）及び対象アイテム区分“Item_j”(jは整数)に対応する評価値と、類似度ベクトルsim2_{k,i}とを基にして、対象ユーザ及び対象アイテム区分に対応する評価値の推定値（第２の推定値）s2を計算する。詳細には、推定値算出部２４は、ユーザ“user_k”及び対象アイテム区分“Item_j”に対応する評価値によって構成されるベクトルDB1[*,j]を情報蓄積部１１から取得し、類似度ベクトルsim2_{k,i}との内積を、下記式（６）；

s2 = DB1[1,j]*sim2_{1,i} +DB1[2,j]*sim2_{2,i} + DB1[m,j]*sim2_{m,i}　　…（６）

を用いて計算し、この内積を対象ユーザ及び対象アイテム区分に対応する評価値の推定値s2とする。ここで、類似度ベクトルsim2_{k,i}の各要素とベクトルDB1[*,j]の各要素との間における各ユーザに関する対応付けは、対応付け判断部２２による判断結果を利用して判断される。

【００３０】

確度算出部２５は、類似度算出部２３が類似度ベクトルsim2_{k,i}を算出する際に参照した表データDB2を基にして、評価値の推定値s2に関する確度（第２の確度）を計算する。すなわち、確度算出部２５は、類似度ベクトルsim2_{k,i}を計算する際に表データDB2から参照されたベクトルDB2[k,*]、ベクトルDB2[i,*]に対応するアイテム区分の数、言い換えれば、表データDB2の列数に比例する値を確度c2として計算する。また、確度算出部１５は、表データDB2に含まれる評価値の数、言い換えれば、表データDB2の空白セル以外のセルの数に比例する数を確度c2として計算してもよい。そして、確度算出部２５は、計算した確度c2を、推定値算出部２４によって計算された推定値s2と共に、推定値生成装置１０の最終推定値算出部１６に送信する。

【００３１】

以下、図５～７を参照して、評価値推定システム１の動作について説明するとともに、併せて評価値推定システム１における評価値推定方法について詳述する。図７は、評価値推定システム１による評価値の推定処理時の動作を示すフローチャートであり、図５及び図６は、評価値推定システム１による評価値の推定処理時に算出される類似度ベクトルの一例を示す図である。なお、ここでは、図２及び図３に示したデータ構成の表データDB1,DB2が参照され、対応付け判断部２２では図４に示したテーブルデータを用いて２つの表データDB1,DB2の対応関係が判断されるものとする。

【００３２】

まず、推定値生成装置１０の推定対象取得部１２により、評価値を予測する対象となる対象ユーザと対象アイテム区分の指定が取得される（ステップＳ１０１）。例えば、推定対象取得部１２により、対象ユーザ識別情報“user_4”及び対象区分識別情報“Item_1”が取得される。次に、類似度算出部１３により、表データDB1に基づいて、対象ユーザ“user_4”と、それ以外のユーザ“user_1”、“user_2”、“user_3”との間の類似度sim1_{1,4}、sim1_{2,4}、sim1_{3,4}が計算される（ステップＳ１０２）。一例として、二乗誤差和の逆数を正規化した値を類似度とする計算例を示すと、まず、下記の計算式を用いて正規化前の類似度d1_{1,4}、d1_{2,4}、d1_{3,4}が得られる。

$d1_\{1,4\} = 1/(2-1)^2 = 1$
$d1_\{2,4\} = 1/(2-4)^2 = 1/4$
$d1_\{3,4\} = 1/\{(2-5)^2 + (2-2)^2\}= 1/9$
その後、これらの類似度d1_{1,4}、d1_{2,4}、d1_{3,4}の総和が“１”となるように正規化される。具体的には、総和が下記式；
$d1_\{1,4\}+d1_\{2,4\}+d1_\{3,4\} = 49/36$
によって計算されるので、正規化された類似度sim1_{1,4}、sim1_{2,4}、sim1_{3,4}が以下のように算出される（図５）。
$sim1_\{1,4\} = 1/(49/36)$　0.73
$sim1_\{2,4\} = (1/4)/(49/36)$　0.18
$sim1_\{3,4\} = (1/9)/(49/36)$　0.08

【００３３】
その一方で、推定値生成装置２０の類似度算出部２３により、表データDB2に基づいて、対象ユーザ“user_4”と、それ以外のユーザ“user_1”、“user_2”、“user_3”との間の類似度sim2_{1,4}、sim2_{2,4}、sim2_{3,4}が計算される（ステップＳ１０３）。まず、下記の計算式を用いて正規化前の類似度d2_{1,4}、d2_{2,4}、d2_{3,4}が得られる。
$d2_\{1,4\} = 1/\{(4-3)^2 + (3-2)^2\}= 1/2$
$d2_\{2,4\} = 1/\{(4-1)^2 + (3-4)^2\}= 1/10$
$d2_\{3,4\} = 1/(4-3)^2 = 1$
この場合、これらの類似度d2_{1,4}、d2_{2,4}、d2_{3,4}の総和が下記式；
$d2_\{1,4\}+d2_\{2,4\}+d2_\{3,4\} = 8/5$
によって計算されるので、正規化された類似度sim2_{1,4}、sim2_{2,4}、sim2_{3,4}が以下のように算出される（図６）。
$sim2_\{1,4\} = (1/2)/(8/5)$　0.31
$sim2_\{2,4\} = (1/10)/(8/5)$　0.06
$sim2_\{3,4\} = 1/(8/5)$　0.63

【００３４】
その後、推定値生成装置１０の推定値算出部１４により、対象ユーザ以外のユーザ“user_1”、“user_2”、“user_3”による対象アイテム区分“Item_1”対する評価値ベクトル{DB1[1,1],DB1[2,1],DB1[3,1]}と、類似度ベクトル[sim1_{1,4},sim1_{2,4},sim1_{3,4}]との内積を算出することにより、対象ユーザによる対象アイテム区分に対する評価値に関する推定値s1が下記式により計算される（ステップＳ１０４）。
$s1 = 5*0.73 + 3*0.18 + 4*0.08 = 4.51$

【００３５】
さらに、推定値生成装置１０の確度算出部１５によって、類似度ベクトル[sim1_{1,4},sim1_{2,4},sim1_{3,4}]を計算するときに参照された表データDB1を基に、推定値s1の確度c1が計算される（ステップＳ１０５）。例えば、類似度ベクトルを計算するときに用いたアイテム区分の数を基に確度を求める場合には、２つのアイテム区分“Item_2”、“Item_3”を用いて計算したので、確度c1=“２”と求められる。

【００３６】
それと同時に、推定値生成装置２０の推定値算出部２４により、対象ユーザ以外のユーザ“user_1”、“user_2”、“user_3”による対象アイテム区分“Item_1”対する評価値ベクトル{DB1[1,1],DB1[2,1],DB1[3,1]}と、類似度ベクトル[sim2_{1,4},sim2_{2,4},sim2_{3,4}]との内積を算出することにより、対象ユーザによる対象アイテム区分に対する評価値に関する推定値s2が下記式により計算される（ステップＳ１０６）。
$s2 = 5*0.31 + 3*0.06 + 4*0.63 = 4.23$

【００３７】
さらに、推定値生成装置２０の確度算出部２５によって、類似度ベクトル[sim2_{1,4},sim2_{2,4},sim2_{3,4}]を計算するときに参照された表データDB2を基に、推定値s2の確度c2が計算される（ステップＳ１０７）。例えば、類似度ベクトルを計算するときに用い

たアイテム区分の数を基に確度を求める場合には、２つのアイテム区分“Item_4”、“Item_5”を用いて計算したので、確度c2=“２”と求められる。

【００３８】

その後、推定値生成装置１０の最終推定値算出部１６によって、推定値s1、s2とそれぞれの推定値に対応する確度c1、c2とを基に、最終的な推定値sが算出される（ステップＳ１０８）。例えば、確度c1、c2を重みとした推定値s1、s2の重み付け平均を計算する場合には、推定値sは次のように計算される。

s = (2 * 4.51 + 2 * 4.23) / (2+2) = 4.37

このようにして計算された推定値sは、表データDB1、DB2の両方に基づく対象ユーザによる対象アイテム区分に対応する評価値に関する予測値となる。例えば、評価値の全体の平均値が“３”である場合には、ユーザ“user_4”はアイテム区分“Item_1”に属する商品を高く評価すると推測されることから、ユーザ“user_4”に対してアイテム区分“Item_1”に属する商品を推奨すべきであると判断することができる。

【００３９】

そして、算出された推定値sは、推定値生成装置１０からネットワークを介して外部の端末装置等に出力される。また、この推定値sを基にした商品推奨情報等が生成された後に、その商品推奨情報がユーザ“user_4”の端末装置等の外部装置に出力されてもよい（ステップＳ１０９）。

【００４０】

以上説明した評価値推定システム１及び評価値推定方法によれば、異なるアイテム区分群に対するユーザによる評価値に関する２以上の表データが別々に格納されており、評価値の推定対象の対象ユーザ“user_i”及び対象アイテム区分“Item_j”が指定されると、類似度算出部１３によって、表データDB1を基に対象ユーザとそれ以外のユーザの評価値の類似度を示す類似度ベクトルsim1_{k,i}が計算され、類似度算出部２３によって、表データDB2を基に対象ユーザとそれ以外のユーザの評価値の類似度を示す第２の類似度ベクトルsim2_{k,i}が計算される。また、推定値算出部１４及び確度算出部１５によって、類似度ベクトルsim1_{k,i}と表データDB1に含まれる対象アイテム区分の評価値DB1[*,j]とを基に推定値s1及びその推定値s1の確度c1が計算され、推定値算出部２４及び確度算出部２５によって、類似度ベクトルsim2_{k,i}と表データDB1に含まれる対象アイテム区分の評価値DB1[*,j]とを基に、表データDB1、DB2間のユーザに関する対応付けの判断結果を用いて、推定値s2及びその推定値s2の確度c2が計算される。さらに、推定値s1、s2及びそれらに対応する確度c1、c2から、最終的な評価値の推定値sが算出される。これにより、表データDB1を管理する推定値生成装置１０から表データDB2を管理する推定値生成装置２０に対して、対象アイテム区分に関する評価値の通知が必要な以外には、その他のアイテム区分に関する評価値の通知の必要がない。従って、複数のデータベース管理装置で分散管理されている評価値のうちで互いに通知される範囲を最低限にすることができるので、アイテムに対するユーザによる評価情報に関するプライバシーが十分に保護される。

【００４１】

なお、本発明は、前述した実施形態に限定されるものではない。例えば、図８には、本発明の変形例である評価値推定システム１０１の構成を示している。同図に示す評価値推定システム１０１の評価値推定システム１との相違点は、推定値生成装置１１０が、暗号化部（暗号化手段）１１７、復号部（復号手段）１１８、及び暗号化ベクトル生成部（暗号化ベクトル生成手段）１１９をさらに備える点と、推定値生成装置１２０，１３０の推定値算出部１２４，１３４の機能が異なる点である。

【００４２】

暗号化部１１７及び復号部１１８は、それぞれ、暗号化関数E及び復号関数Dを用いて加法準同型性の性質を有する暗号化演算及び復号演算を行う。すなわち、この暗号化関数Eは、処理対象の任意値x,yに対して、下記式（１）；

Ｅ（ｘ）・Ｅ（ｙ）＝Ｅ（ｘ＋ｙ）　　…（１）

の関係を満たすような関数であり、任意値x,yの和を暗号化した値と、それぞれの任意値x

,yを暗号化した値の積とが等しくなるような性質を持つ。また、暗号化関数E及び復号関数Dは、下記式（２）；

Ｄ（Ｅ（ｘ））＝ｘ　　…（２）

の関係を満たすような関数であり、任意値xの暗号化値を復号すると基の値xに戻されるような性質も有する。このような暗号方式としては、Paillier暗号を用いることができる。

【００４３】

暗号化ベクトル生成部１１９は、暗号化部１１７によって対象ユーザ以外のユーザ“user_k”及び対象アイテム区分“Item_j”に対応する評価値によって構成されるベクトルDB1[*,j]の構成要素を暗号化させることにより、暗号化ベクトルE_j=[E(DB1[1,j]), E(DB1[2,j]),…, E(DB1[m,j])]を生成する。この場合、暗号化ベクトル生成部１１９は、暗号化対象のベクトルDB1[*,j]の構成要素が整数でない場合には、予め定められた方法、例えば、全ての構成要素を定数倍する等の方法で整数化する。そして、暗号化ベクトル生成部１１９は、生成した暗号化ベクトルE_jを、推定値生成装置１２０，１３０の推定値算出部１２４，１３４に引き渡す。

【００４４】

これに対し、推定値生成装置１２０の推定値算出部１２４は、暗号化ベクトルE_jと、類似度ベクトルsim2_{k,i}とを参照し、暗号化ベクトルE_jの各構成要素を底とし、類似度ベクトルsim2_{k,i}の各構成要素に対応する値を指数とした累乗を計算し、その累乗の積を算出することにより、暗号化された推定値s2を取得する。具体的には、推定値算出部１２４は、暗号化ベクトル生成部１１９と同様にして、類似度ベクトルsim2_{k,i}の構成要素を整数化し、整数化されたsim2_{k,i}の各構成要素をv1,…,vmとし、整数化されたベクトルDB1[*,j]の各構成要素をu1,…umとしたときに、下記式（７）；

$E(s2) = E(u1)^{v1} * E(u2)^{v2} * \cdots * E(um)^{vm}$　　…（７）

によって暗号化された推定値E(s2)を算出する。

【００４５】

推定値生成装置１１０の復号部１１８は、推定値生成装置１２０の推定値算出部１２４によって生成された暗号化後の推定値E(s2)を取得すると、復号関数Dを用いて推定値E(s2)を復号することにより、推定値s2 = D(E(s2))を取得することができる。そして、復号部１１８は、取得した推定値s2を最終推定値算出部１６に引き渡す。

【００４６】

図９には、評価値推定システム１０１による評価値の推定処理時の動作手順を示している。なお、同図に示すステップＳ２０１～Ｓ２０５，Ｓ２０８，Ｓ２１０～Ｓ２１１の処理は、図７に示したステップＳ１０１～Ｓ１０５、Ｓ１０７～Ｓ１０９の処理と同様である。

【００４７】

まず、図２及び図３に示したデータ構成の表データDB1,DB2を対象に評価値推定処理が開始されると、ステップＳ２０１～Ｓ２０５において、２つの表データDB1,DB2を基にした類似度ベクトルsim1_{k,i}, sim2_{k,i}と、表データDB1を基にした推定値s1及びその確度c1とが算出される。次に、暗号化ベクトル生成部１１９によって、ベクトルDB1[*,j]の構成要素u1,…,umを暗号化されることにより、暗号化ベクトルE_j=[E(u1),E(u2),…,E(um)]が生成される（ステップＳ２０６）。例えば、図２に示したようなデータ構成の表データDB1を対象にした場合には、暗号化ベクトルE_j=[E(5), E(3), E(4)]が生成される。

【００４８】

その後、推定値算出部１２４によって、暗号化ベクトルE_jと類似度ベクトルsim2_{k,i}とを用いて、上記式（７）により暗号化後の推定値E(s2)が生成される（ステップＳ２０７）。例えば、図３に示したようなデータ構成の表データDB2を対象にした場合には、次の計算式によって暗号化後の推定値E(s2)が計算される。

$E(s2) = E(5)^{(0.31*100)} * E(3)^{(0.06*100)} * E(4)^{(0.63*100)}$

$= E(5)^{31} * E(3)^{6} * E(4)^{63}$

ここでは、類似度ベクトルsim2_{k,i}の構成要素は実数値をとるため、それぞれを100倍

することによって整数化される。なお、上記の暗号化後の推定値E(s2)は、暗号化関数Eが持つ加法準同型性により、以下の値と等価となる。

$E(s2) = E(5)^{31} * E(3)^{6} * E(4)^{63}$

$= E(5*31) * E(3*6) * E(4*63)$

$= E(5*31 + 3*6 + 4*63)$

【００４９】

その後、ステップＳ２０８において、推定値s2に対する確度c2が算出される（ステップＳ２０８）。次に、推定値生成装置１１０の最終推定値算出部１６によって、復号部１１８を用いて暗号化後の推定値E(s2)が復号されることによって、推定値s2が取得される（ステップＳ２０９）。具体的には、下記の計算式によって、推定値s2が取得される。

$s2 = D\{E(s2)\} / 100$

$= D(E(5*31 + 3*6 + 4*63)) / 100 = 423 /100 = 4.23$

ここでは、ステップＳ２０７において類似度ベクトルsim2_{k,i}の整数化のために“100”を乗じているため、復号された値が“100”で除される。その後、ステップＳ２１０～Ｓ２１１において、推定値生成装置１１０により、最終的な評価値の推定値sの計算及びその推定値sの出力処理が行われる。

【００５０】

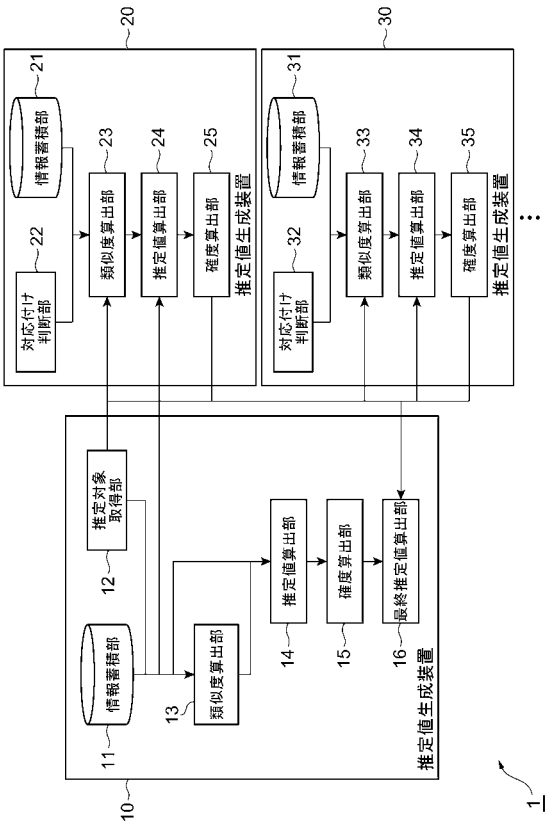
このような評価値推定システム１０１及び評価値推定方法によれば、複数のデータベースDB1,DB2,DB3間で対象アイテム区分に関する評価値が平文のまま通知されずに、対象ユーザの対象アイテム区分に関する評価値が予測されるので、複数のアイテム区分に関する評価値に関するユーザのプライバシーがより一層保護される。それに加えて、複数のデータベースを対象にした場合でも、加法準同型性の性質を有する暗号化演算及び復号演算を用いることで正確な予測値が得られる。

【符号の説明】

【００５１】

１，１０１…評価値推定システム、１０，１１０…推定値生成装置、１１…情報蓄積部（第１の情報蓄積手段）、１２…推定対象取得部（推定対象取得手段）、１３…類似度算出部（第１の類似度算出手段）、１４…推定値算出部（第１の推定値算出手段）、１５…確度算出部（第１の確度算出手段）、１６…最終推定値算出部（第３の推定値算出手段）、２０，１２０…推定値生成装置、２１…情報蓄積部（第２の情報蓄積手段）、２２…対応付け判断部（対応付け判断手段）、２３…類似度算出部（第２の類似度算出手段）、２４，１２４…推定値算出部（第２の推定値算出手段）、２５…確度算出部（第２の確度算出手段）、１１７…暗号化部（暗号化手段）、１１８…復号部（復号手段）、１１９…暗号化ベクトル生成部（暗号化ベクトル生成手段）。

【図１】



【図２】

	Item_1	Item_2	Item_3
user_1	5	1	
user_2	3		4
user_3	4	5	2
user_4		2	2

【図３】

	Item_4	Item_5
user_1	3	2
user_2	1	4
user_3	3	
user_4	4	3

【図４】

	DB1	DB2	…
user_1	1行目	1行目	…
user_2	2行目	2行目	…
user_3	3行目	3行目	…
user_4	4行目	4行目	…

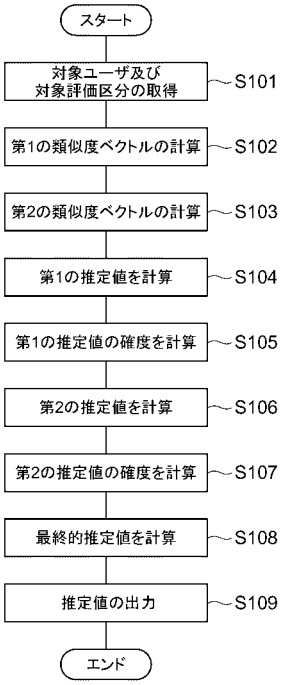
【図５】

	Item_1	Item_2	Item_3	sim1
user_1	5	1		0.73
user_2	3		4	0.18
user_3	4	5	2	0.08
user_4		2	2	

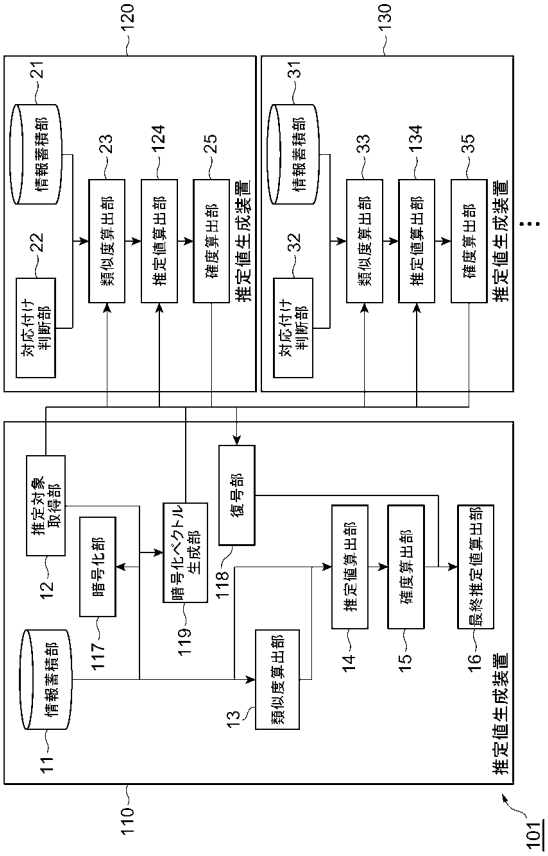
【図６】

	Item_4	Item_5	sim2
user_1	3	2	0.31
user_2	1	4	0.06
user_3	3		0.63
user_4	4	3	

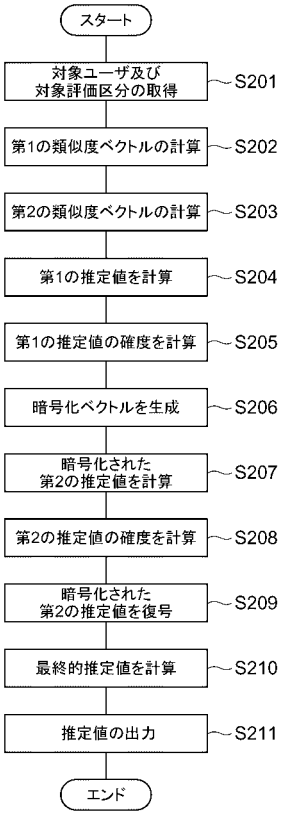
【図７】



【図８】



【図９】


スタート
対象ユーザ及び
対象評価区分の取得
S201
第1の類似度ベクトルの計算
S202
第2の類似度ベクトルの計算
S203
第1の推定値を計算
S204
第1の推定値の確度を計算
S205
暗号化ベクトルを生成
S206
暗号化された
第2の推定値を計算
S207
第2の推定値の確度を計算
S208
暗号化された
第2の推定値を復号
S209
最終的推定値を計算
S210
推定値の出力
S211
エンド

フロントページの続き

(72)発明者　石井　一彦
東京都千代田区永田町二丁目１１番１号　株式会社エヌ・ティ・ティ・ドコモ内
(72)発明者　関野　公彦
東京都千代田区永田町二丁目１１番１号　株式会社エヌ・ティ・ティ・ドコモ内
Ｆターム(参考) 5B075 KK02　KK53　KK54　NS10　QM05