(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2012-118878
(P2012-118878A)
(43)公開日　平成24年6月21日(2012.6.21)

(51)Int.Cl.		F I			テーマコード（参考）
G06F 17/30	(2006.01)	G06F	17/30	340A	5B017
G06F 13/00	(2006.01)	G06F	13/00	540P	5B075
G06Q 30/02	(2012.01)	G06F	17/60	326	5B084
G06F 21/24	(2006.01)	G06F	17/30	120A	
		G06F	17/30	412	

審査請求 未請求　請求項の数 6　ＯＬ　（全 23 頁）　最終頁に続く

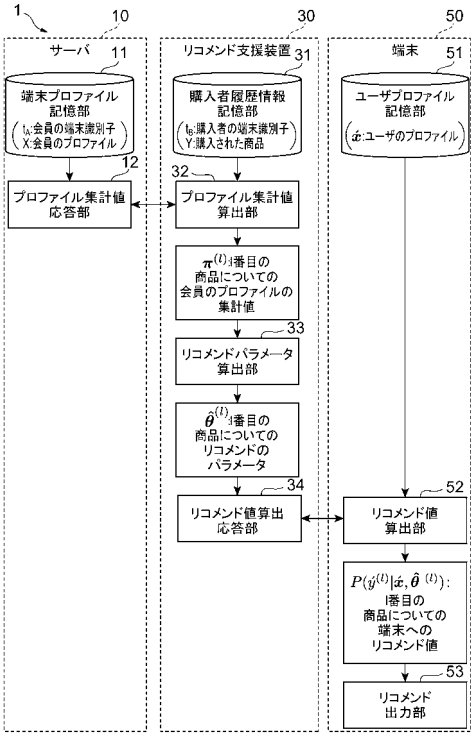
(21)出願番号　特願2010-269600 (P2010-269600)
(22)出願日　平成22年12月2日 (2010.12.2)

(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　山口　高康
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】リコメンドシステム及びリコメンド方法

(57)【要約】
【課題】ユーザのプライバシーに関する情報や、店舗が有する固有の履歴情報を他者に開示することなく、良好な処理効率により的確なリコメンドを実施する。
【解決手段】リコメンドシステム１では、サーバ１０及びリコメンド支援装置３０の間では、それぞれ所定の方式により暗号化された端末プロファイル及び購入者履歴情報が交換され、これらの情報に基づきリコメンド支援装置においてプロファイル集計値が算出される。また、リコメンド支援装置３０において、プロファイル集計値に基づき、リコメンドパラメータを商品ごとに得ることが可能となる。さらに、リコメンド支援装置３０において、所定の方式により端末により暗号化されたユーザプロファイルがリコメンドパラメータを用いてさらに暗号化されて端末に送信され、端末５０において、送信された情報に基づき商品ごとのリコメンド値が算出される。
【選択図】図１



【特許請求の範囲】
【請求項１】
複数の端末の端末識別子と該端末のユーザのプロファイルとが対応付けられた端末プロファイルを記憶している端末プロファイル記憶手段を備えるサーバと、
端末を識別する端末識別子と該端末のユーザにより購入された商品を示す情報とが対応付けられた購入者履歴情報を記憶している購入者履歴情報記憶手段を備えるリコメンド支援装置と、
当該端末のユーザのプロファイルであるユーザプロファイルを記憶しているユーザプロファイル記憶手段を備える端末と
を有するリコメンドシステムであって、
前記サーバは、
前記端末プロファイルを所定の方式により暗号化して生成した暗号化端末プロファイル、及び前記リコメンド支援装置により暗号化された前記購入者履歴情報である暗号化購入者履歴情報を所定に方式により暗号化して生成した２重暗号化購入者履歴情報を前記リコメンド支援装置に送信するプロファイル集計応答手段を備え、
前記リコメンド支援装置は、
前記購入者履歴情報を所定の方式により暗号化して生成した暗号化購入者履歴情報を前記サーバに送信し、該サーバから送信された暗号化端末プロファイル及び２重暗号化購入者履歴情報に基づき、商品を購入したユーザのプロファイルの集計値であるプロファイル集計値を当該商品ごとに算出するプロファイル集計値算出手段と、
前記プロファイル集計値算出手段により算出された前記プロファイル集計値に基づき、端末のユーザに対する当該商品に関するリコメンド情報を生成するためのリコメンドパラメータを所定の方式により算出するリコメンドパラメータ算出手段と、
前記端末により暗号化された前記ユーザプロファイルである暗号化ユーザプロファイル、及び前記リコメンドパラメータ算出手段により算出されたリコメンドパラメータを用いた所定の暗号化処理により、リコメンド値の算出のためのリコメンド値算出支援情報を生成し、該リコメンド値算出支援情報を前記端末に送信するリコメンド値算出応答手段とを備え、
前記端末は、
前記ユーザプロファイルを所定の方式により暗号化して生成した前記暗号化ユーザプロファイルを前記リコメンド支援装置に送信し、該リコメンド支援装置からのリコメンド値算出支援情報に基づき、商品ごとのリコメンド値を算出するリコメンド値算出手段と、
前記リコメンド値算出手段により算出されたリコメンド値に応じたリコメンド情報を出力するリコメンド出力手段とを備える
ことを特徴とするリコメンドシステム。
【請求項２】
前記プロファイル集計応答手段は、
前記端末プロファイルを、ハッシュ関数で暗号化し、当該プロファイル集計応答手段において生成した第１の乱数によりべき乗して前記暗号化端末プロファイルを生成し、前記暗号化購入者履歴情報を前記第１の乱数によりべき乗して２重暗号化購入者履歴情報を生成し、前記暗号化端末プロファイル及び前記２重暗号化購入者履歴情報を前記リコメンド支援装置に送信し、
前記プロファイル集計値算出手段は、
前記購入者履歴情報を、ハッシュ関数で暗号化し、当該プロファイル集計値算出手段において生成した第２の乱数によりべき乗して前記暗号化購入者履歴情報を生成し、前記暗号化購入者履歴情報を前記サーバに送信し、前記サーバから送信された前記暗号化端末プロファイル及び前記２重暗号化購入者履歴情報に基づき前記プロファイル集計値を算出する
ことを特徴とする請求項１に記載のリコメンドシステム。
【請求項３】
前記リコメンドパラメータ算出手段は、

前記プロファイル集計値算出手段により算出された前記プロファイル集計値に基づき、前記リコメンドパラメータをＭＡＰ推定により算出する
ことを特徴とする請求項１または２に記載のリコメンドシステム。
【請求項４】
前記リコメンド値算出応答手段は、
前記暗号化ユーザプロファイルに対して、前記リコメンドパラメータ算出手段により算出されたリコメンドパラメータを用いたべき乗処理することにより前記リコメンド値算出支援情報を生成し、
前記リコメンド値算出手段は、
加法準同型性を満たす前記端末の公開鍵で前記ユーザプロファイルを暗号化して前記暗号化ユーザプロファイルを生成し、加法準同型性を満たす前記端末の秘密鍵を用いて前記リコメンド値算出支援情報を復号化して前記リコメンド値を算出する
ことを特徴とする請求項１～３のいずれか１項に記載のリコメンドシステム。
【請求項５】
前記リコメンド値算出応答手段は、
当該前記リコメンド支援装置において任意に設定された調整パラメータを有し、所定の定数に対して前記リコメンドパラメータ算出手段により算出されたリコメンドパラメータを用いたべき乗処理をし、さらに、前記調整パラメータに基づき算出された値を乗じてリコメンド値算出支援調整情報を生成し、生成した前記リコメンド値算出支援調整情報を前記端末に送信し、
前記リコメンド値算出手段は、
加法準同型性を満たす前記端末の秘密鍵を用いて前記リコメンド値算出支援情報及び前記リコメンド値算出支援調整情報を復号化し、復号化された前記リコメンド値算出支援情報及び前記リコメンド値算出支援調整情報を用いて前記リコメンド値を算出する
ことを特徴とする請求項４に記載のリコメンドシステム。
【請求項６】
複数の端末の端末識別子と該端末のユーザのプロファイルとが対応付けられた端末プロファイルを記憶している端末プロファイル記憶手段を備えるサーバと、
端末を識別する端末識別子と該端末のユーザにより購入された商品を示す情報とが対応付けられた購入者履歴情報を記憶している購入者履歴情報記憶手段を備えるリコメンド支援装置と、
当該端末のユーザのプロファイルであるユーザプロファイルを記憶しているユーザプロファイル記憶手段を備える端末と
を有するリコメンドシステムにおけるリコメンド方法であって、
前記リコメンド支援装置が、前記購入者履歴情報を所定の方式により暗号化して生成した暗号化購入者履歴情報を前記サーバに送信する暗号化購入者履歴情報送信ステップと、
前記サーバが、前記端末プロファイルを所定の方式により暗号化して生成した暗号化端末プロファイル、及び前記暗号化購入者履歴情報送信ステップにおいて送信された前記暗号化購入者履歴情報を所定に方式により暗号化して生成した２重暗号化購入者履歴情報を前記リコメンド支援装置に送信するプロファイル集計応答ステップと、
前記リコメンド支援装置が、前記プロファイル集計応答ステップにおいて前記送信された前記暗号化端末プロファイル及び前記２重暗号化購入者履歴情報に基づき、商品を購入したユーザのプロファイルの集計値であるプロファイル集計値を当該商品ごとに算出するプロファイル集計値算出ステップと、
前記リコメンド支援装置が、前記プロファイル集計値算出ステップにおいて算出された前記プロファイル集計値に基づき、端末のユーザに対する当該商品に関するリコメンド情報を生成するためのリコメンドパラメータを所定の方式により算出するリコメンドパラメータ算出ステップと、
前記端末が、前記ユーザプロファイルを所定の方式により暗号化して生成した前記暗号化ユーザプロファイルを前記リコメンド支援装置に送信する暗号化ユーザプロファイル送

信ステップと、
前記リコメンド支援装置が、前記暗号化ユーザプロファイル送信ステップにおいて送信された前記暗号化ユーザプロファイル、及び前記リコメンドパラメータ算出ステップにおいて算出されたリコメンドパラメータを用いた所定の暗号化処理により、リコメンド値の算出のためのリコメンド値算出支援情報を生成し、該リコメンド値算出支援情報を前記端末に送信するリコメンド値算出応答ステップと、
前記端末が、前記リコメンド値算出応答ステップにおいて送信された前記リコメンド値算出支援情報に基づき、商品ごとのリコメンド値を算出するリコメンド値算出ステップと、
前記端末が、前記リコメンド値算出ステップにおいて算出されたリコメンド値に応じたリコメンド情報を出力するリコメンド出力ステップと
を有することを特徴とするリコメンド方法。
【発明の詳細な説明】
【技術分野】
【０００１】
本発明は、リコメンドシステム及びリコメンド方法に関する。
【背景技術】
【０００２】
ユーザがモバイル環境で有用な情報をタイムリーに手に入れるために、情報提示機能の高度化に対する要求が高まっている。モバイル環境で使用される端末では、画面サイズに制限があるため、効率の良い情報提示機能が必要である。情報提示機能の高度化の実現方法として、ユーザの属性や商品の購入履歴に基づいて、ユーザの趣向に合う商品を推薦するリコメンド技術が提案されている。しかしながら、ユーザ属性や商品の購入履歴といった推薦の元となるデータは、複数の企業にわたって管理されている場合が多い。このため、ユーザのプライバシーや企業のデータを保護するために、組織の枠を超えてリコメンドを行うことは難しかった。
【０００３】
かかる事情のもと、近年、暗号プロトコルや統計的開示制御技術の導入により、組織の枠を超えてリコメンドを提供する方式が注目されている。リコメンドの方法としては、協調フィルタリングとコンテンツベースの方法が存在する。協調フィルタリングは、計算機が、あるユーザと購買履歴の似ているユーザを探し出し、その購買履歴の似ているユーザが購入した商品をリコメンドする。コンテンツベースの方法は、計算機が、あるユーザが好む商品と似ている商品を探し出し、その商品をリコメンドする。
【０００４】
コンテンツベースの方法は、協調フィルタリングに比べて新商品のような履歴が少ない商品でもリコメンドし易いという利点を有する。両方法ともユーザのプライベートな情報を取り扱うため、ユーザのプライバシーを保護する必要がある。協調フィルタリングでは、秘匿積集合プロトコルを利用してプライバシーを保護する方法が提案されている。また、コンテンツベースの方法では、ユーザのプライバシーを保護するべく、ユーザの好みと商品のマッチングを、例えば洋服の寸法をＳ，Ｍ，Ｌといったサイズや、料理を和、洋、中というカテゴリで行う方法がある（例えば、特許文献１参照）。
【先行技術文献】
【特許文献】
【０００５】
【特許文献１】特開２００８－２８２３５３号公報
【発明の概要】
【発明が解決しようとする課題】
【０００６】
上記のように、新商品のような履歴が少ない商品でも的確なリコメンドを実施すべく、コンテンツベースの方法を採用できることが好ましい。しかしながら、コンテンツベース

の方法において、来店者の属性や履歴を開示することなく、顧客数や商品数が多い場合であっても、効率的なリコメンドが可能な方式は提案されていない。

【０００７】

例えば、特定の事業者が有し当該事業者に属する会員のプロファイルを管理しているサーバ、特定の店舗が有し当該店舗において商品を購入した端末の情報を管理しているリコメンド支援装置、及びあるユーザが有し当該ユーザのプロファイルを記憶している端末の３者間において各々が有する情報が他者に開示されることなく、端末のユーザに適切なリコメンド情報を好適な処理効率により提供することが望まれている。

【０００８】

そこで、本発明は、上記問題点に鑑みてなされたものであり、ユーザのプライバシーに関する情報や、店舗が有する固有の履歴情報を他者に開示することなく、良好な処理効率により的確なリコメンドを実施することが可能なリコメンドシステム及びリコメンド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】

【０００９】

上記課題を解決するために、本発明のリコメンドシステムは、複数の端末の端末識別子と該端末のユーザのプロファイルとが対応付けられた端末プロファイルを記憶している端末プロファイル記憶手段を備えるサーバと、端末を識別する端末識別子と該端末のユーザにより購入された商品を示す情報とが対応付けられた購入者履歴情報を記憶している購入者履歴情報記憶手段を備えるリコメンド支援装置と、当該端末のユーザのプロファイルであるユーザプロファイルを記憶しているユーザプロファイル記憶手段を備える端末とを有するリコメンドシステムであって、サーバは、端末プロファイルを所定の方式により暗号化して生成した暗号化端末プロファイル、及びリコメンド支援装置により暗号化された購入者履歴情報である暗号化購入者履歴情報を所定に方式により暗号化して生成した２重暗号化購入者履歴情報をリコメンド支援装置に送信するプロファイル集計応答手段を備え、リコメンド支援装置は、購入者履歴情報を所定の方式により暗号化して生成した暗号化購入者履歴情報をサーバに送信し、該サーバから送信された暗号化端末プロファイル及び２重暗号化購入者履歴情報に基づき、商品を購入したユーザのプロファイルの集計値であるプロファイル集計値を当該商品ごとに算出するプロファイル集計値算出手段と、プロファイル集計値算出手段により算出されたプロファイル集計値に基づき、当該商品に関する端末のユーザに対するリコメンド情報を生成するためのリコメンドパラメータを所定の方式により算出するリコメンドパラメータ算出手段と、端末により暗号化されたユーザプロファイルである暗号化ユーザプロファイル、及びリコメンドパラメータ算出手段により算出されたリコメンドパラメータを用いた所定の暗号化処理により、リコメンド値の算出のためのリコメンド値算出支援情報を生成し、該リコメンド値算出支援情報を端末に送信するリコメンド値算出応答手段とを備え、端末は、ユーザプロファイルを所定の方式により暗号化して生成した暗号化ユーザプロファイルをリコメンド支援装置に送信し、該リコメンド支援装置からのリコメンド値算出支援情報に基づき、商品ごとのリコメンド値を算出するリコメンド値算出手段と、リコメンド値算出手段により算出されたリコメンド値に応じたリコメンド情報を出力するリコメンド出力手段とを備えることを特徴とする。

【００１０】

また、上記課題を解決するために、本発明のリコメンド方法は、複数の端末の端末識別子と該端末のユーザのプロファイルとが対応付けられた端末プロファイルを記憶している端末プロファイル記憶手段を備えるサーバと、端末を識別する端末識別子と該端末のユーザにより購入された商品を示す情報とが対応付けられた購入者履歴情報を記憶している購入者履歴情報記憶手段を備えるリコメンド支援装置と、当該端末のユーザのプロファイルであるユーザプロファイルを記憶しているユーザプロファイル記憶手段を備える端末とを有するリコメンドシステムにおけるリコメンド方法であって、リコメンド支援装置が、購入者履歴情報を所定の方式により暗号化して生成した暗号化購入者履歴情報をサーバに送信する暗号化購入者履歴情報送信ステップと、サーバが、端末プロファイルを所定の方式

により暗号化して生成した暗号化端末プロファイル、及び暗号化購入者履歴情報送信ステップにおいて送信された暗号化購入者履歴情報を所定に方式により暗号化して生成した２重暗号化購入者履歴情報をリコメンド支援装置に送信するプロファイル集計応答ステップと、リコメンド支援装置が、プロファイル集計応答ステップにおいて送信された暗号化端末プロファイル及び２重暗号化購入者履歴情報に基づき、商品を購入したユーザのプロファイルの集計値であるプロファイル集計値を当該商品ごとに算出するプロファイル集計値算出ステップと、リコメンド支援装置が、プロファイル集計値算出ステップにおいて算出されたプロファイル集計値に基づき、当該商品に関する端末のユーザに対するリコメンド情報を生成するためのリコメンドパラメータを所定の方式により算出するリコメンドパラメータ算出ステップと、端末が、ユーザプロファイルを所定の方式により暗号化して生成した暗号化ユーザプロファイルをリコメンド支援装置に送信する暗号化ユーザプロファイル送信ステップと、リコメンド支援装置が、暗号化ユーザプロファイル送信ステップにおいて送信された暗号化ユーザプロファイル、及びリコメンドパラメータ算出ステップにおいて算出されたリコメンドパラメータを用いた所定の暗号化処理により、リコメンド値の算出のためのリコメンド値算出支援情報を生成し、該リコメンド値算出支援情報を端末に送信するリコメンド値算出応答ステップと、端末が、リコメンド値算出応答ステップにおいて送信されたリコメンド値算出支援情報に基づき、商品ごとのリコメンド値を算出するリコメンド値算出ステップと、端末が、リコメンド値算出ステップにおいて算出されたリコメンド値に応じたリコメンド情報を出力するリコメンド出力ステップとを有することを特徴とする。

【００１１】

本発明のリコメンドシステム及びリコメンド方法によれば、サーバ及びリコメンド支援装置の間では、それぞれ所定の方式により暗号化された端末プロファイル及び購入者履歴情報が交換され、これらの情報に基づきリコメンド支援装置においてプロファイル集計値が算出される。これにより、サーバが管理する端末プロファイルがリコメンド支援装置に漏洩されず、また、リコメンド支援装置が管理する購入者履歴情報がサーバに認識可能な状態で渡されることなく、リコメンド支援装置はプロファイル集計値を得ることができる。また、リコメンド支援装置において、プロファイル集計値に基づき、リコメンド支援装置において商品を購入する可能性の高いユーザの属性の傾向をリコメンドパラメータとして商品ごとに得ることが可能となる。さらに、リコメンド支援装置において、所定の方式により端末により暗号化されたユーザプロファイルがリコメンドパラメータを用いてさらに暗号化されて端末に送信され、端末において、送信された情報に基づき商品ごとのリコメンド値が算出される。これにより、端末は、端末が管理するユーザプロファイルを、リコメンド支援装置やサーバに認識可能な状態で取得されることなく、商品に関するリコメンド情報を得ることができる。

【００１２】

また、本発明のリコメンドシステムでは、プロファイル集計応答手段は、端末プロファイルを、ハッシュ関数で暗号化し、当該プロファイル集計応答手段において生成した第１の乱数によりべき乗して暗号化端末プロファイルを生成し、暗号化購入者履歴情報を第１の乱数によりべき乗して２重暗号化購入者履歴情報を生成し、暗号化端末プロファイル及び２重暗号化購入者履歴情報をリコメンド支援装置に送信し、プロファイル集計値算出手段は、購入者履歴情報を、ハッシュ関数で暗号化し、当該プロファイル集計値算出手段において生成した第２の乱数によりべき乗して暗号化購入者履歴情報を生成し、暗号化購入者履歴情報をサーバに送信し、サーバから送信された暗号化端末プロファイル及び２重暗号化購入者履歴情報に基づきプロファイル集計値を算出することを特徴とする。

【００１３】

この構成によれば、購入者履歴情報は、加法準同型性を満たす方式であるハッシュ化及びべき乗処理を用いてリコメンド支援装置により暗号化され、サーバに送信され、さらにサーバにおいてべき乗処理された後にリコメンド支援装置に送信される。また、端末プロファイルは、ハッシュ化及びべき乗処理されてリコメンド支援装置に送信される。これに

より、リコメンド支援装置において、購入者履歴情報及び端末プロファイルの共通集合としてのプロファイル集計値を適切に得ることが可能となる。

【００１４】

また、本発明のリコメンドシステムでは、リコメンドパラメータ算出手段は、プロファイル集計値算出手段により算出されたプロファイル集計値に基づき、リコメンドパラメータをＭＡＰ推定により算出することを特徴とする。

【００１５】

この構成によれば、プロファイル集計値に基づき、好適なリコメンドパラメータを得ることが可能となる。

【００１６】

また、本発明のリコメンドシステムでは、リコメンド値算出応答手段は、暗号化ユーザプロファイルに対して、リコメンドパラメータ算出手段により算出されたリコメンドパラメータを用いたべき乗処理することによりリコメンド値算出支援情報を生成し、リコメンド値算出手段は、加法準同型性を満たす端末の公開鍵でユーザプロファイルを暗号化して暗号化ユーザプロファイルを生成し、加法準同型性を満たす端末の秘密鍵を用いてリコメンド値算出支援情報を復号化してリコメンド値を算出することを特徴とする。

【００１７】

この構成によれば、ユーザプロファイルは、加法準同型性を満たす公開鍵で端末により暗号化、及びリコメンド支援装置に送信され、更にリコメンド支援装置において、リコメンドパラメータを用いてべき乗処理された後に端末に送信される。これにより、ユーザプロファイルとリコメンドパラメータとの秘匿内積計算が実現されるので、端末において、ユーザプロファイルとリコメンドパラメータとの内積に基づくリコメンド値を商品ごとに得ることができる。

【００１８】

また、本発明のリコメンドシステムでは、リコメンド値算出応答手段は、当該リコメンド支援装置において任意に設定された調整パラメータを有し、所定の定数に対してリコメンドパラメータ算出手段により算出されたリコメンドパラメータを用いたべき乗処理をし、さらに、調整パラメータに基づき算出された値を乗じてリコメンド値算出支援調整情報を生成し、生成したリコメンド値算出支援調整情報を端末に送信し、リコメンド値算出手段は、加法準同型性を満たす端末の秘密鍵を用いてリコメンド値算出支援情報及びリコメンド値算出支援調整情報を復号化し、復号化されたリコメンド値算出支援情報及びリコメンド値算出支援調整情報を用いてリコメンド値を算出することを特徴とする。

【００１９】

この構成によれば、リコメンド支援装置は、端末のユーザプロファイルだけではなく、リコメンド支援装置を有する店舗側の意向を加味して、端末へのリコメンドを行うこともできる。

【発明の効果】

【００２０】

ユーザのプライバシーに関する情報や、店舗が有する固有の履歴情報を他者に開示することなく、良好な処理効率により的確なリコメンドを実施することが可能となる。

【図面の簡単な説明】

【００２１】

【図１】リコメンドシステムの機能的構成を示すブロック図である。

【図２】リコメンド支援装置のハードブロック図である。

【図３】リコメンドシステムを構成する各装置と処理に用いられる各種情報との関係を示す図である。

【図４】リコメンドシステムにおけるリコメンド方法の処理内容を示すタイミングチャートである。

【図５】実用性における従来手法と本実施形態の手法との比較を示す図である。

【発明を実施するための形態】

【００２２】
まず、本実施形態のリコメンドシステムの説明に先立って、本実施形態の方式の背景となる従来方式に関する説明をする。以下に、従来方式として、秘匿内積計算を用いたナイーブベイズ識別器、及びセキュアマッチングを用いたクロス集計を説明する。これらの方式は、いずれもサーバ、リコメンド支援装置及び端末間において、各装置が有するデータを他の装置に対して秘匿したままで、端末に対してリコメンドを実施することを目的としている。
【００２３】
[秘匿内積計算を用いたナイーブベイズ識別器]
・サーバ及びリコメンド支援装置間の処理
Vaidyaらは、分割したデータベースを互いに秘匿したままでナイーブベイズ識別器を実現する方式を提案している（Jaideep Vaidya, and Chris Clifton. Privacy preserving naive bayesclassfier for vertically partitioned data. In Society for Industrial andApplied Mathematics, 2008.）。属性情報を保持するサーバとクラス情報を保持するリコメンド支援装置とが協働して、ナイーブベイズの計算を行う。ただし、本方式では、サーバ及びリコメンド支援装置がそれぞれ持っているデータは同期していることが前提とされる。
【００２４】
ナイーブベイズの独立性の仮定から、式（１）に示すように、$x_v^{(l)}$のそれぞれの条件下における$y^{(l)}$の条件付確率Ｐが求められ、未知パラメータが学習される。
【数１】

$$P(y^{(l)}|x_v^{(l)}) = \frac{P(y^{(l)},x_v^{(l)})}{P(x_v^{(l)})} = \frac{y^{(l)}\cdot x_v^{(l)}}{N}\frac{N}{|x_v^{(l)}|} \qquad (1)$$

【００２５】
サーバ及びリコメンド支援装置がそれぞれ持っているデータは同期していることが前提であるで、式（１）に示す条件付確率の計算式では、Ｎが打ち消され合う。また、リコメンド支援装置は、単独で$|x_v^{(l)}|$を計算できる。よって、$y^{(l)}\cdot x_v^{(l)}$の内積計算さえできれば、リコメンド支援装置は、未知パラメータを学習できる。この内積計算は、いわゆる秘匿内積計算プロトコルで実現される。ここで、秘匿内積計算プロトコルの一般的な記述を以下に示す。秘匿内積計算では、原理的にべき乗計算が必要となるので、計算効率を高めることは困難である。
【００２６】
［秘匿内積計算］
【００２７】
ここで説明する秘匿内積計算では、サーバが有する$x=(x_1,\cdots,x_n)$及びリコメンド支援装置が有する$y=(y_1,\cdots,y_n)$を入力とする。なお、ｘ及びｙはｎ次元のベクトルである。そして、ｘ・ｙ（ｘとｙの内積）を出力とする。
１．リコメンド支援装置は、加法準同型性を満たすリコメンド支援装置の公開鍵でｙを暗号化して、サーバに$E(y_1),\cdots,E(y_n)$を送る。なお、$E(y_n)$は、暗号化されたy_nを表す。
２．サーバは、送信された情報を各々ｘでべき乗処理して$E(y_1)^{x_1},\cdots,E(y_n)^{x_n}$を得る。そして、これをシャッフルしてリコメンド支援装置に送信する。
３．リコメンド支援装置は、加法準同型性を満たすリコメンド支援装置の秘密鍵で、サーバから受信したメッセージに基づき、
$x\cdot y=\Sigma_{i=1}^{n} D\{E(x_i y_i)\}=\Sigma_{i=1}^{n} D\{E(y_i)^{x_i}\}$
といった計算処理により、出力を得る。
【００２８】
・リコメンド支援装置及び端末間の処理
二者間秘匿回路計算プロトコル（A. C. Yao. How to generate and exchangesecrets (extended abstract). In IEEE FOCS '86, pp. 162-167, 1986.）を実行して、最大尤度の

クラスを求めることにより、データの識別が行われる。しかしながら、このプロトコルによる計算コストは高い。

【００２９】

以上のように、上記方式による処理は、計算効率、計算コストにおいて、好適ではなく、顧客数や商品数が多い場合において、処理が困難となる。後述する本実施形態では、計算効率の向上が図られており、処理速度が向上されている。

【００３０】

[セキュアマッチングを用いたクロス集計]

・サーバ及びリコメンド支援装置間の処理

Agrawalらは、可換な一方向性関数を用いて、ハッシュ値の生成にO(n)、照合にO(log n)のコストで計算できる照合タグ方式を提案した（R. Agrawal, A. V. Ev_mievski, and R. Srikant. Information sharingacross private databases. In ACM SIGMOD 2003, pp. 86-97, 2003.）。また、千田らは、ランダムオラクルモデルの上で照合タグ方式の安全性を証明し（千田，五十嵐，高橋．照合タグを用いた秘匿共通集合計算プロトコルとその応用．コンピュータセキュリティシンポジウム2009，IPSJ，2009.）、さらに計算量と通信量を削減して、クロス集計表を効率的に作成する方式を提案している（千田，寺田，山口，五十嵐，濱田．セキュアマッチングを用いた組織間クロス分析．コンピュータセキュリティシンポジウム2010，IPSJ，2010.）。

・リコメンド支援装置及び端末間の処理

クロス集計表から、安全にデータを識別する方法は提案されていない。

【００３１】

以上のように、上記方式による処理では、端末のユーザは、好適なリコメンド情報を、自らのプロファイルを公開することなく得ることはできない。後述する本実施形態では、ユーザのプロファイルが秘匿されたままで、端末がリコメンド情報を得ることが可能である。

【００３２】

続いて、本発明に係るリコメンドシステム及びリコメンド方法の実施形態について図面を参照して説明する。なお、可能な場合には、同一の部分には同一の符号を付して、重複する説明を省略する。

【００３３】

図１は、リコメンドシステム１の機能的構成を示すブロック図である。図１に示すように、リコメンドシステム１は、サーバ１０、リコメンド支援装置３０及び端末５０を有する。

【００３４】

サーバ１０は、例えば、特定の事業者が有するサーバであって、当該事業者に属する会員のプロファイルを管理している。リコメンド支援装置３０は、特定の店舗が有する装置であって、当該店舗において商品を購入した端末の情報を商品と対応付けて管理している。端末５０は、ユーザが有する端末であって、当該ユーザのプロファイルを記憶している。

【００３５】

ここで、本実施形態の説明において用いられる記号及び用語の定義を以下に述べる。

【００３６】

t_Aは、サーバ１０において管理されている会員の端末識別子である。

【数２】

$$\boldsymbol{t}_A \in \mathbb{Z}^{N_A}$$

N_Aは、サーバ１０において管理されている会員数である。なお、Ｚは整数の集合を表す（以下、同様）。即ち、上記の表記は、t_AがN_A個の整数であることを示す。

【００３７】

Ｘは、サーバ１０において管理されている会員のプロファイルであって、Ｖ行Ｎ$_{A}$列のマトリクスである。

【数３】

$$X \in \mathcal{R}^{V \times N_A}$$
$$(x_{v,n_A} > 0)$$

Ｖは、プロファイルの項目数である。ｘ$_{ｖ，ｎ A}$は、Ｘに含まれる一の要素である。ｖは、一のプロファイル項目を示し、ｎ$_{A}$は、一の会員を示す。なお、Ｒは実数の集合を表す（以下、同様）。即ち、上記の表記は、ｘが実数であることを示す。即ち、一の会員のプロファイルは、当該会員の属性を示し、ｖ次元のベクトルデータとして表されることができる。

【００３８】

ｔ$_{B}$は、リコメンド支援装置３０において管理されている購入者履歴情報における購入者の端末識別子である。

【数４】

$$t_B \in \mathcal{Z}^{N_B}$$

Ｎ$_{B}$は、リコメンド支援装置３０において管理されている延べの購入者数である。即ち、同一ユーザが商品を２回購入した場合には、当該購入者は、数Ｎ$_{B}$の集計において２回数えられている。

【００３９】

Ｙは、リコメンド支援装置３０において管理されている購入された商品を示す情報であって、Ｌ行Ｎ$_{B}$列のマトリクスである。

【数５】

$$Y \in \{0,1\}^{L \times N_B}$$
$$\left(\sum_{l=1}^{L} y_{l,n_B} = 1\right)$$

Ｌは、商品の種類数である。ｙ$_{ｌ，ｎ B}$は、Ｙに含まれる一の要素である。ｙにおけるｌ番目の要素の値が１である場合には、ｎ$_{B}$番目の購入者により商品ｌが買われたことを示す。括弧内の式は、１人の購入者は、１回の購入において１個の商品を購入することを示す。

【００４０】

ｘ′は、端末５０のユーザのプロファイルであって、ｖ次元のベクトルデータである。なお、文字ｘに付されたダッシュは、当該データが当該端末５０のユーザに関するものであることを意味し、以降で引用する式においても同様である。

【数６】

$$\acute{x} \in \mathcal{R}^{V}$$

【００４１】

πは、ｌ番目の商品についての、サーバ１０が管理する会員のプロファイルの集計値であって、ｖ次元のベクトルである。

【数７】

$$\pi^{(l)} \in \mathcal{R}^{V}$$

【００４２】

θは、ｌ番目の商品についてのリコメンドパラメータであって、ｖ次元の確率ベクトルである。

【数８】

$$\hat{\boldsymbol{\theta}}^{(l)} \in \mathcal{R}^{V}$$

このθは、（$\Sigma^{V}_{v=1}\theta_{v}^{(l)}=1$）の式により正規化されている。なお、文字θに付されたハット記号（＾）は、当該データがサーバ１０に管理される会員に関するデータ及びリコメンド支援装置３０の店舗における購入者に関するデータに基づくものであることを意味し、以降で引用する式においても同様である。

【００４３】

Ｐは、ｌ番目の商品についての、端末５０に対するリコメンド値である。

【数９】

$$P(\acute{y}^{(l)}|\acute{\boldsymbol{x}},\hat{\boldsymbol{\theta}}^{(l)})$$

即ち、Ｐは、（ｘ´，θ＾$^{(l)}$）といった条件下におけるｙ´$^{(l)}$の条件付確率として表される。

【００４４】

図２は、リコメンド支援装置３０のハードウエア構成図である。リコメンド支援装置３０は、物理的には、図２に示すように、ＣＰＵ１０１、主記憶装置であるＲＡＭ１０２及びＲＯＭ１０３、データ送受信デバイスである通信モジュール１０４、ハードディスク、フラッシュメモリ等の補助記憶装置１０５、入力デバイスであるキーボード等の入力装置１０６、ディスプレイ等の出力装置１０７などを含むコンピュータシステムとして構成されている。図１に示した各機能は、図２に示すＣＰＵ１０１、ＲＡＭ１０２等のハードウエア上に所定のコンピュータソフトウェアを読み込ませることにより、ＣＰＵ１０１の制御のもとで通信モジュール１０４、入力装置１０６、出力装置１０７を動作させるとともに、ＲＡＭ１０２や補助記憶装置１０５におけるデータの読み出し及び書き込みを行うことで実現される。また、サーバ１０及び端末５０も、図２に示したリコメンド支援装置３０と同様のハードウエア構成を有するコンピュータシステムとして構成される。

【００４５】

再び、図１を参照してリコメンドシステム１を説明する。サーバ１０は、端末プロファイル記憶部１１（端末プロファイル記憶手段）及びプロファイル集計応答部１２（プロファイル集計応答手段）を備える。リコメンド支援装置３０は、購入者履歴情報記憶部３１（購入者履歴情報記憶手段）、プロファイル集計値算出部３２（プロファイル集計値算出手段）、リコメンドパラメータ算出部３３（リコメンドパラメータ算出手段）及びリコメンド値算出応答部３４（リコメンド値算出応答手段）を備える。端末５０は、ユーザプロファイル記憶部５１（ユーザプロファイル記憶手段）、リコメンド値算出部５２（リコメンド値算出手段）、及びリコメンド出力部５３（リコメンド出力手段）を備える。以下、図１を参照しながら、各機能部の機能を説明する。

【００４６】

端末プロファイル記憶部１１は、サーバ１０が管理する複数の会員の端末の端末識別子と該端末のユーザのプロファイルとが対応付けられた端末プロファイルＸを記憶している記憶手段である。

【００４７】

プロファイル集計値応答部１２は、端末プロファイルＸを所定の方式により暗号化して生成した暗号化端末プロファイル、及びリコメンド支援装置３０により暗号化された購入者履歴情報Ｙである暗号化購入者履歴情報を所定に方式により暗号化して生成した２重暗号化購入者履歴情報をリコメンド支援装置３０に送信する部分である。

【００４８】

購入者履歴情報記憶部３１は、リコメンド支援装置３０を有する店舗において購入された商品を示す情報と、その商品を購入した端末を識別する識別子とが対応付けられた購入

者履歴情報Ｙを記憶している記憶手段である。

【００４９】

プロファイル集計値算出部３２は、購入者履歴情報Ｙを所定の方式により暗号化して生成した暗号化購入者履歴情報をサーバ１０に送信する部分である。ここで、暗号化の所定の方式は、例えば、ハッシュ化及び乱数によるべき乗処理である。また、プロファイル集計値算出部３２は、サーバ１０から送信された暗号化端末プロファイル及び２重暗号化購入者履歴情報に基づき、商品を購入したユーザのプロファイルの集計値であるプロファイル集計値πを当該商品ごとに算出する機能を有する。

【００５０】

なお、一般に暗号化とは、復号化可能な変換処理を指す。一方、ハッシュ化は、不可逆であり復号化不可能な変換処理である。しかし、本発明及び本実施形態の説明においては、ハッシュ化を「ハッシュ関数により暗号化」として、暗号化の一種として取り扱うこととする。

【００５１】

リコメンドパラメータ算出部３３は、プロファイル集計値算出部３２により算出されたプロファイル集計値πに基づき、端末のユーザに対する当該商品に関するリコメンド情報を生成するためのリコメンドパラメータθを所定の方式により算出する部分である。

【００５２】

リコメンド値算出応答部３４は、端末５０により暗号化されたユーザプロファイルｘ´である暗号化ユーザプロファイル、及びリコメンドパラメータ算出部３３により算出されたリコメンドパラメータθを用いた所定の暗号化処理により、リコメンド値Ｐの算出のためのリコメンド値算出支援情報を生成し、該リコメンド値算出支援情報を端末に送信する部分である。

【００５３】

端末５０のユーザプロファイル記憶部５１は、当該端末５０のユーザのプロファイルであるユーザプロファイルｘ´を記憶している記憶手段である。

【００５４】

リコメンド値算出部５２は、ユーザプロファイルｘ´を所定の方式により暗号化して生成した暗号化ユーザプロファイルをリコメンド支援装置３０に送信する部分である。ここで、暗号化のための所定の方式は、例えば、いわゆる秘匿内積計算の一部の処理である。また、リコメンド値算出部５２は、リコメンド支援装置３０からのリコメンド値算出支援情報に基づき、商品ごとのリコメンド値Ｐを算出する機能を有する。

【００５５】

リコメンド出力部５３は、リコメンド値算出部５２により算出されたリコメンド値Ｐに応じたリコメンド情報を出力する部分である。

【００５６】

リコメンド支援装置３０のプロファイル集計値算出部３２及びサーバ１０のプロファイル集計値応答部１２は、会員の端末識別子t_Aと会員のプロファイルＸを持つサーバ１０と、購入者の端末識別子t_Bと購入された商品の情報Ｙを持つリコメンド支援装置３０とが、互いのプライバシーを保護しつつ情報を交換し、商品を購入した会員のプロファイル集計値πをリコメンド支援装置３０だけが得る手段を構成する。リコメンド支援装置３０は、サーバ１０とのいわゆるセキュアマッチング処理により、ｌ番目の商品についての会員のプロファイル集計値$\pi^{(l)}$を得る。一般的なプロトコルではサーバ１０及びリコメンド支援装置３０の両方がプロファイル集計値を得るが、本方式では、リコメンド支援装置のみがプロファイル集計値πを得ることができる。

【００５７】

リコメンド支援装置３０のリコメンドパラメータ算出部３３は、プロファイル集計値πを取得したリコメンド支援装置３０が、商品ｌについてのリコメンドパラメータθを得る手段を構成する。即ち、リコメンドパラメータ算出部３３は、ｌ番目の商品についての会員のプロファイル集計値πを用いて、店舗のノウハウであるｌ番目の商品についてのリコ

メンドパラメータθを学習する。

【００５８】

端末５０のリコメンド値算出部５２及びリコメンド支援装置３０のリコメンド値算出応答部３４は、リコメンドパラメータθを持つリコメンド支援装置３０と、自機ユーザのプロファイルｘ′を持つ端末５０とが、互いのプライバシーを保護しつつ情報を交換し、リコメンド支援装置３０の店舗の商品の端末５０に対するリコメンド値Ｐを端末５０だけが得る手段を構成する。端末５０は、リコメンド支援装置３０とのいわゆる秘匿内積計算により、端末５０のプロファイルｘ′と学習済みのｌ番目の商品についてのリコメンドパラメータθとを用いて、ｌ番目の商品についての端末５０へのリコメンド値Ｐを算出する。

【００５９】

続いて、本実施形態のリコメンド情報の提供におけるユースケースと、当該ユースケースにおけるサーバ１０、リコメンド支援装置３０及び端末５０それぞれのメリットについて説明する。即ち、本実施形態におけるユースケースは以下のとおりである。

【００６０】

会員のユーザ属性（プロファイル）を管理している事業者のサーバ１０と、店舗に備えられたリコメンド支援装置３０と、その店舗に買い物にやってきたユーザの端末５０との三者がいて、店舗の経営者は、端末５０のユーザへのリコメンドにより、店舗の売り上げを上げたいと考えているとする。また、事業者は、サーバ１０に管理している、顧客から預かったユーザ属性を漏洩したくない、といった事情が存在する。店舗の経営者は、当該店舗の売り上げ情報と、店舗で商品を購入したユーザ属性の統計値とを漏洩したくない、といった事情が存在する。店舗は、事業者の会員情報を持つ事業者のサーバ１０とのコミュニケーションを通じて、店舗で商品を購入したユーザ属性（プロファイル）の統計値が得られる。ただし、ユーザ属性の統計値は、ある人数以上の顧客を集計対象とした場合にのみに限られる。

【００６１】

端末５０のユーザは、自身のユーザ属性（プロファイル）、及び当該ユーザが閲覧する商品の履歴を漏洩したくない、といった事情がある。端末５０のユーザは、上記店舗に初めて来店したとしても、面倒な会員申し込み手続きなどをせずにスムーズにリコメンドを受けたい。ただし、端末５０のユーザは、上記事業者の会員であるとは限らない。

【００６２】

上記のユースケースを実現するには、サーバ１０、リコメンド支援装置３０及び端末５０の三者が以下のようなメリットを享受できるようにしなければならない。

１．サーバ１０は会員から預かっているユーザ属性（プロファイル）の情報漏えいを防止できる。

２．リコメンド支援装置３０を備える店舗は、当該店舗で販売した商品の情報をサーバ１０に渡さずに、当該店舗の商品を購入する可能性の高いユーザ属性の傾向を知ることができる。

３．端末５０のユーザは、ユーザのプライバシー情報をサーバ１０及びリコメンド支援装置３０に渡さずに、ユーザが欲する商品を、当該ユーザが欲する可能性が高い順に商品を閲覧できる。

【００６３】

次に、図３を参照して、本実施形態における、サーバ１０を有する事業者、リコメンド支援装置３０を有する店舗、及び端末５０を有するユーザ間の関係と、３者が有する情報及び３者間でやり取りされる情報を説明する。

【００６４】

ここでは、店舗にやってきた客は、１度の買い物でＬ種類の商品の中から１種類の商品だけを購入することとする。精度良いリコメンドを行うため、端末５０へのリコメンド情報を算出する際には古くから実績のあるナイーブベイズ（Charu C. Aggarwal, and Philip S. Yu. Privacy-Preserving Data Mining:Models and Algorithms. Springer, 2008.）を用いることとする。ナイーブベイズのアプローチでは分布の仮定を置く必要があり、

人の属性を表すプロファイルは多項分布に従うこととする。すなわち、プロファイルをＶ個の項目からなるベクトルで表して、各項目の値を０以上とする。また、リコメンドパラメータθの分布は、一般に用いられている共役事前分布を用いる（渡部．ベイズ統計学入門．福村出版,1999.）。多項分布の共役事前分布はディレクレ分布であるので、リコメンドパラメータθをＶ個の項目からなるベクトルで表して、各項目の値の和を１とする。

【００６５】

リコメンドパラメータθを求める際には、事後確率を最大にする観点から、ｌ番目の商品についてのリコメンドパラメータ$\theta_v^{(l)}$をＭＡＰ推定で求めることとする。本実施形態で用いられるベイズアプローチ識別器を一般的な表現で記述すると以下のようになる。

【００６６】

［ベイズアプローチ識別機］

ｄ´の事後の予測分布は式（２）により表される。

【数１０】

$$p(\acute{\boldsymbol{d}}|\mathbb{D}) = \int_{\Theta} m(\hat{\boldsymbol{d}}|\Theta)p(\Theta|\mathbb{D})d\Theta \qquad (2)$$

リコメンドパラメータ算出部３３は、学習データＤを用いて、未知パラメータマトリクスθを事後確率最大の観点で学習する。ここで、学習データＤは、会員のプロファイルであるユーザ属性マトリクス$X \in R^{V \times N}$と購入された商品の情報である購買動向マトリクス$Y \in \{0, 1\}^{L \times N}$との組である。すなわち、Ｄ＝｛Ｘ，Ｙ｝である。

【００６７】

学習データＤを得た後の未知パラメータθの事後確率はベイズの定理より、式（３）に示す比例関係が得られる。

【数１１】

$$p(\Theta|\mathbb{D}) = \frac{p(\mathbb{D}|\Theta)p(\Theta)}{p(\Theta)} \propto p(\mathbb{D}|\Theta)p(\Theta) \qquad (3)$$

上記式（３）の比例関係の対数をとっても大小関係は変わらないので、上記式（３）の比例関係の対数の式で未知パラメータマトリクス$\theta \in R^{V \times L}$を変化させ、事後確率を最大とする未知パラメータθ＾を下記式（４）により求める。

【数１２】

$$\hat{\Theta} = \underset{\Theta}{\operatorname{argmax}} P(\Theta|\mathbb{D}) = \underset{\Theta}{\operatorname{argmax}} \left[\mathcal{L}(\mathbb{D};\Theta) + \log P(\Theta)\right] \qquad (4)$$

【００６８】

Ｄの事後の予測分布にθ＾を代入すれば、未知の属性ｘ＾をＬ個のクラスに識別するベイズ識別器は式（５）のようになる。

【数１３】

$$\acute{y} = \underset{l}{\operatorname{argmax}} \left[\mathcal{L}(\acute{\boldsymbol{x}};\hat{\boldsymbol{\theta}}^{(l)}) + \log P(\hat{\boldsymbol{\theta}}^{(l)}|\mathbb{D})\right] \qquad (5)$$

【００６９】

ナイーブベイズは説明変数の独立性を仮定する手法である。そのため、サーバ１０が有する会員のプロファイルのＶ個の項目は独立であるとみなされるとする。すなわち、ｌ番目の商品を購入した会員のｖ番目のプロファイルの値を独立に集計できれば良い。スケールメリットを得るには、サーバ１０にプロファイルが管理されている会員数N_Aやリコメンド支援装置３０の店舗での購入者数N_Bが多くなっても破綻なく処理されることが要求されるため、サーバ１０がリコメンド支援装置３０に会員のプロファイル集計値πを開示

するのにセキュアマッチング（千田，寺田，山口，五十嵐，濱田．セキュアマッチングを用いた組織間クロス分析．コンピュータセキュリティシンポジウム2010，IPSJ，2010.）を用いることとする。

【００７０】

多項分布は指数分布族であるので、ｌ番目の商品についての端末５０へのリコメンド値Ｐを算出する際に、端末５０のユーザプロファイルｘ´とリコメンドパラメータθの対数との内積計算が必要となる。端末５０のユーザのプライバシーと、リコメンド支援装置３０の店舗のノウハウとを保護するため、リコメンド支援装置３０が端末５０にリコメンド情報を提供するのに秘匿内積計算（Jaideep Vaidya, and Chris Clifton. Privacy preserving naive bayesclassier for vertically partitioned data. In Society for Industrial and AppliedMathematics, 2008.）を用いることとする。

【００７１】

次に、図４のタイミングチャートを参照して、本実施形態のリコメンド方法を説明する。

【００７２】

ステップＳ１において、リコメンド支援装置３０は、プロファイル集計値算出部３２を用いた処理により、ｌ番目の商品についての会員のプロファイル集計値πをいわゆるセキュアマッチングにより取得する。ここで、ｌ番目の商品に関する属性ｖのプロファイル集計値は、式（６）により表される。

【数１４】

$$\pi_v^{(l)} = |x_{n,v} \cap y^{(l)}|_n = \sum_{n=1}^{N^{(l)}} x_{n,v}^{(l)} \qquad (6)$$

【００７３】

具体的には、リコメンド支援装置３０のプロファイル集計値算出部３２及びサーバ１０のプロファイル集計値応答部１２は、購入者の端末識別子t_B及び購入された商品の情報Ｙと、会員の端末識別子t_A及び会員のプロファイルＸとを秘匿性を維持しつつ交換し、商品を購入した会員のプロファイル集計値πをプロファイル集計値算出部３２だけが得る。

【００７４】

さらに具体的には、プロファイル集計値算出部３２は、購入者履歴情報Ｙを、ハッシュ関数で暗号化（ハッシュ化）し、当該プロファイル集計値算出部３２において生成した乱数によりべき乗して暗号化購入者履歴情報を生成し、暗号化購入者履歴情報をサーバ１０に送信する。次いで、プロファイル集計値応答部１２は、端末プロファイルＸを、ハッシュ関数で暗号化し、当該プロファイル集計値応答部１２において生成した乱数によりべき乗して暗号化端末プロファイルを生成し、暗号化購入者履歴情報を同乱数によりべき乗して２重暗号化購入者履歴情報を生成し、暗号化端末プロファイル及び２重暗号化購入者履歴情報をリコメンド支援装置３０に送信する。そして、プロファイル集計値算出部３２は、サーバ１０から送信された暗号化端末プロファイル及び２重暗号化購入者履歴情報に基づきプロファイル集計値πを算出する。以上説明したセキュアマッチングによる秘匿集計の処理を具体的に説明する。

【００７５】

［セキュアマッチングによる秘匿集計］

ここでは、入力をサーバ１０が有する$x=(x_1, \cdots, x_{N_A})$、及びリコメンド支援装置３０が有する$y=(y_1, \cdots, y_{N_B})$として、出力として、$|x \cap y|$を得る処理を説明する。

１．プロファイル集計値算出部３２は、乱数$r_B \in Z_q$を選び（Ｚはｒが整数であることを示す）、位数ｑの巡回群Ｇと、Ｇを値域とするハッシュ関数Ｈで、ｙをハッシュして、サーバ１０に$H(y_1)^{r_B}, \cdots, H(y_{N_B})^{r_B}$を送る。

２．プロファイル集計値応答部１２は、乱数$r_A \in Z_q$を選び、リコメンド支援装置３０へ$H(x_1)^{r_A}, \cdots, H(x_{N_A})^{r_A}$と、$H(y_1)^{r_B r_A}, \cdots, H(y_{N_B})^{r_B r_A}$とをシャッフルして送る。
３．プロファイル集計値算出部３２は、プロファイル集計値応答部１２から送られた$H(x_1)^{r_A}, \cdots, H(x_{N_A})^{r_A}$を乱数$r_B$によりべき乗して$H(x_1)^{r_A r_B}, \cdots, H(x_{N_A})^{r_A r_B}$を得る。そして、プロファイル集計値算出部３２は、$H(x_v)^{r_A r_B} = H(y_v)^{r_B r_A}$を満たす個数$|x \cap y|$を取得する。この個数からプロファイル集計値πを得ることができる。

【００７６】

次に、ステップＳ２において、リコメンドパラメータ算出部３３は、ｌ番目の商品についての会員のプロファイル集計値$\pi^{(l)}$を用いて、店舗のノウハウであるｌ番目の商品についてのリコメンドパラメータθを学習する。具体的には、リコメンドパラメータ算出部３３は、プロファイル集計値算出部３２により算出されたプロファイル集計値πに基づき、リコメンドパラメータθをいわゆるＭＡＰ推定により算出する。ここで学習、算出されるリコメンドパラメータθは、式（７）により表される。

【数１５】

$$\hat{\theta}_v^{(l)} = \frac{\pi_v^{(l)} + (\xi_v^{(l)} - 1)}{\left(\sum_{v=1}^{V} \pi_v^{(l)}\right) + V(\xi_v^{(l)} - 1)} \qquad (7)$$

【００７７】

リコメンドパラメータθの学習、算出は、例えば、以下に示す、最大の事後確率を得る未知パラメータの推定処理によって、実施される。

【００７８】

［最大の事後確率を得る未知パラメータの推定］

【００７９】

この推定処理に必要な、多項分布を仮定した尤度関数は、以下の式（８）～（１１）のように表される。

【数１６】

$$Multi\{N, \boldsymbol{x}^{(l)}; \boldsymbol{\theta}^{(l)}\} = \frac{N^{(l)}!}{\prod_v x_v^{(l)}!} \{\theta_v^{(l)}\}^{x_v^{(l)}} \qquad (8)$$

【数１７】

$$P(\boldsymbol{x}^{(l)}|\boldsymbol{\theta}^{(l)}) \propto \prod_{v=1}^{V} (\theta_v^{(l)})^{x_v^{(l)}} \qquad (9)$$

【数１８】

$$\log P(\boldsymbol{x}^{(l)}|\boldsymbol{\theta}^{(l)}) \propto \sum_{v=1}^{V} x_v^{(l)} \log \theta_v^{(l)} \qquad (10)$$

【数１９】

$$\mathcal{L}(X^{(l)}; \boldsymbol{\theta}^{(l)}) \triangleq \log P(X^{(l)}|\boldsymbol{\theta}^{(l)}) \propto \sum_{n=1}^{N^{(l)}} \sum_{v=1}^{V} x_{n,v}^{(l)} \log \theta_v^{(l)} \qquad (11)$$

【００８０】

また、この推定処理に必要な、ディレクレ分布を仮定した事前確率の対数は、以下の式（１２）～（１４）のように表される。

【数２０】

$$Dirichlet\{\boldsymbol{\theta}^{(l)};\boldsymbol{\xi}^{(l)}\} = \frac{\Gamma(\sum_v \xi_v^{(l)})}{\prod_v \Gamma(\xi_v^{(l)})} \prod_v \{\theta_v^{(l)}\}^{\xi_v^{(l)}-1} \quad (12)$$

【数２１】

$$P(\boldsymbol{\theta}^{(l)}) \propto \prod_{v=1}^{V} \{\theta_v^{(l)}\}^{\xi_v^{(l)}-1} \quad (13)$$

【数２２】

$$\log P(\boldsymbol{\theta}^{(l)}) \propto \sum_{v=1}^{V} (\xi_v^{(l)}-1)\log\theta_v^{(l)} \quad (14)$$

【００８１】

この推定処理では、以下の式（１５）によりリコメンドパラメータを推定する。

【数２３】

$$\hat{\boldsymbol{\theta}}^{(l)} = \underset{\boldsymbol{\theta}^{(l)}}{\operatorname{argmax}}\{\mathcal{L}(X^{(l)};\boldsymbol{\theta}^{(l)}) + \log P(\boldsymbol{\theta}^{(l)})\} \quad (15)$$

続いて、最大化したい上記式（１５）の中括弧の中に、尤度関数と事前確率の対数を代入して、下記式（１６）を得る。

【数２４】

$$\mathcal{L}(X^{(l)};\boldsymbol{\theta}^{(l)}) + \log P(\boldsymbol{\theta}^{(l)}) = \left(\sum_{n=1}^{N^{(l)}}\sum_{v=1}^{V} x_{n,v}^{(l)} \log\theta_v^{(l)}\right) + \left(\sum_{v=1}^{V}(\xi_v^{(l)}-1)\log\theta_v^{(l)}\right) \quad (16)$$

目的関数をＪとおき、未知パラメータを求める。即ち、式（１７）～（１９）に示すように、ラグランジュの未定係数法により、$\Sigma^{V}_{v=1}\theta_v^{(l)}=1$のディレクレ分布の条件下でＪの最大化を行う。

【数２５】

$$J = \left(\sum_{n=1}^{N^{(l)}}\sum_{v=1}^{V} x_{n,v}^{(l)} \log\theta_v^{(l)}\right) + \left(\sum_{v=1}^{V}(\xi_v^{(l)}-1)\log\theta_v^{(l)}\right) + \lambda\left(\sum_{v=1}^{V}\theta_v^{(l)}-1\right) \to Max \quad (17)$$

【数２６】

$$\frac{\partial J}{\partial\theta_v^{(l)}} = \left(\frac{1}{\hat{\theta}_v^{(l)}}\sum_{n=1}^{N^{(l)}} x_{n,v}^{(l)}\right) + (\xi_v^{(l)}-1)\frac{1}{\hat{\theta}_v^{(l)}} + \lambda = 0 \quad (18)$$

【数２７】

$$\hat{\theta}_v^{(l)} = \frac{\left(\sum_{n=1}^{N^{(l)}} x_{n,v}^{(l)}\right) + (\xi_v^{(l)} - 1)}{-\lambda} \qquad (19)$$

条件より、$\Sigma^{V}_{v=1}\theta_v^{(l)} = 1$であるので、未知パラメータは以下の式（２０）～（２２）のようになる。

【数２８】

$$\sum_{v=1}^{V} \hat{\theta}_v^{(l)} = \sum_{v=1}^{V} \frac{\left(\sum_{n=1}^{N^{(l)}} x_{n,v}^{(l)}\right) + (\xi_v^{(l)} - 1)}{-\lambda} = 1 \qquad (20)$$

【数２９】

$$\lambda = -\left(\sum_{v=1}^{V}\sum_{n=1}^{N^{(l)}} x_{n,v}^{(l)}\right) - V(\xi_v^{(l)} - 1) \qquad (21)$$

【数３０】

$$\hat{\theta}_v^{(l)} = \frac{\left(\sum_{n=1}^{N^{(l)}} x_{n,v}^{(l)}\right) + (\xi_v^{(l)} - 1)}{\left(\sum_{v=1}^{V}\sum_{n=1}^{N^{(l)}} x_{n,v}^{(l)}\right) + V(\xi_v^{(l)} - 1)} \qquad (22)$$

【００８２】

続くステップＳ３において、端末５０のリコメンド値算出部５２は、リコメンド支援装置３０との秘匿内積計算により、端末５０のプロファイルと、学習済みのｌ番目の商品についてのリコメンドパラメータθとを用いて、ｌ番目の商品についての端末５０へのリコメンド値を算出する。ここで、端末５０のユーザプロファイルｘ´及び端末５０が閲覧する商品の履歴がリコメンド支援装置３０に知られることはないため、端末５０のユーザのプライバシーは保護される。

【００８３】

具体的には、リコメンド値算出部５２は、加法準同型性を満たす端末５０の公開鍵でユーザプロファイルｘ´を、暗号化して暗号化ユーザプロファイルを生成する。次いで、リコメンド値算出応答部３４は、暗号化ユーザプロファイルに対して、リコメンドパラメータ算出部３３により算出されたリコメンドパラメータθを用いたべき乗処理をすることによりリコメンド値算出支援情報を生成する。そして、リコメンド値算出部５２は、加法準同型性を満たす端末５０の秘密鍵を用いてリコメンド値算出支援情報を復号化して、商品ごとのリコメンド値Ｐを算出する。

【００８４】

ここで、リコメンド値は、下記式（２３）のように表される。

【数３１】

$$P(\acute{y}^{(l)}|\acute{x}, \hat{\boldsymbol{\theta}}^{(l)}) = \left(\sum_{v=1}^{V} \acute{x}_v \log \hat{\theta}_v^{(l)}\right) + \left(\sum_{v=1}^{V} (\xi_v^{(l)} - 1)\log \hat{\theta}_v^{(l)}\right) \qquad (23)$$

なお、式（２３）は、式（５）に式（１０）及び式（１４）を適用することにより導出可能である。

【００８５】

以下に、リコメンド値算出部５２及びリコメンド値算出応答部３４により実施される秘

匿内積計算について説明する。ここで説明する秘匿内積計算では、リコメンド値算出応答部３４が有するリコメンドパラメータθ及びリコメンド値算出部５２が有するユーザプロファイルｘ´を入力とする。これらの入力情報はいずれもＶ次元のベクトルである。ただし、リコメンド値算出応答部３４は、リコメンドパラメータθの対数値を入力とする。そして、式（２３）の右辺第１項の内積を出力とする。

１．リコメンド値算出部５２は、加法準同型性を満たす端末５０の公開鍵でユーザプロファイルｘ´を暗号化して、リコメンド値算出応答部３４にＥ（x_1），・・・，Ｅ（y_v）を送る。

２．リコメンド値算出応答部３４は、送信された情報を各々リコメンドパラメータθの対数でべき乗処理して、その結果得られた情報

Ｅ（x_1）$^{\log\theta}$，・・・，Ｅ（x_v）$^{\log\theta_v}$を得る。そして、これをシャッフルした情報（リコメンド値算出支援情報）をリコメンド値算出部５２に送信する。

３．リコメンド値算出部５２は、加法準同型性を満たす端末５０の秘密鍵で、リコメンド値算出応答部３４から受信したメッセージに基づき、

$x' \cdot \log\theta^{(l)} = \Sigma^{v}_{i=1} D\{E(x_i \log\theta_i)\} = \Sigma^{v}_{i=1} D\{E(x_i)^{\log\theta i}\}$といった計算処理により、出力を得る。

【００８６】

なお、式（２３）における右辺第２項は、x'_vを含まないので、リコメンド値Ｐに端末５０のユーザプロファイルに依存しない値を加える項である。この項の値は、リコメンドパラメータθの対数及びディレクレ分布のハイパーパラメータ（調整パラメータ）（$\xi_v^{(l)}$）の大きさに応じて大きくなる。ディレクレ分布のハイパーパラメータはリコメンド支援装置３０において任意に指定することができる。例えば，ディレクレ分布のハイパーパラメータξの値をｖ及びｌによらずに全て２とした場合には、（$\xi_v^{(l)}-1$）の値が１となるので、リコメンドパラメータθの対数のみからこの項の値が算出される。また、特定の商品のディレクレ分布のハイパーパラメータξの値を大きくすると，該当する商品についての（$\xi_v^{(l)}-1$）の値が１より大きくなるので、特定の商品に関するこの項の値が大きくなる。即ち、リコメンド支援装置３０は、端末５０のユーザプロファイルだけではなく、リコメンド支援装置３０を有する店舗側の意向を加味して、端末５０へのリコメンドを行うこともできる。なお、この項におけるΣの計算は、上述した秘匿内積計算において、ユーザプロファイルｘ´の値を１（所定の定数）としてリコメンド支援装置３０から端末５０に送信される情報（リコメンド値算出支援調整情報）に基づき実施可能である。

【００８７】

そして、ステップＳ４において、リコメンド出力部５３は、リコメンド値算出部５２により算出されたリコメンド値Ｐに応じたリコメンド情報を出力する。

【００８８】

次に、図５を参照して、本実施形態の手法と従来の手法とを対比して、本実施形態の手法のメリットを説明する。本実施形態の手法は従来の秘匿内積計算を用いたナイーブベイズ識別器が問題としていた、スケールの問題を解決できる手法である。なお、スケールとは、処理対象とするデータ量が大量であっても破綻なく、又は妥当な処理負荷により処理可能なことをいう。当然ながら、実用においては、プライバシーを保護しつつリコメンドを提供する際には、精度及びスケールのいずれもが欠けてはならない。そこで、本実施形態の手法が実用に耐えうるバランスを持つ事を述べるべく、精度、スケールの観点で、本実施形態の手法と従来手法を比較する。

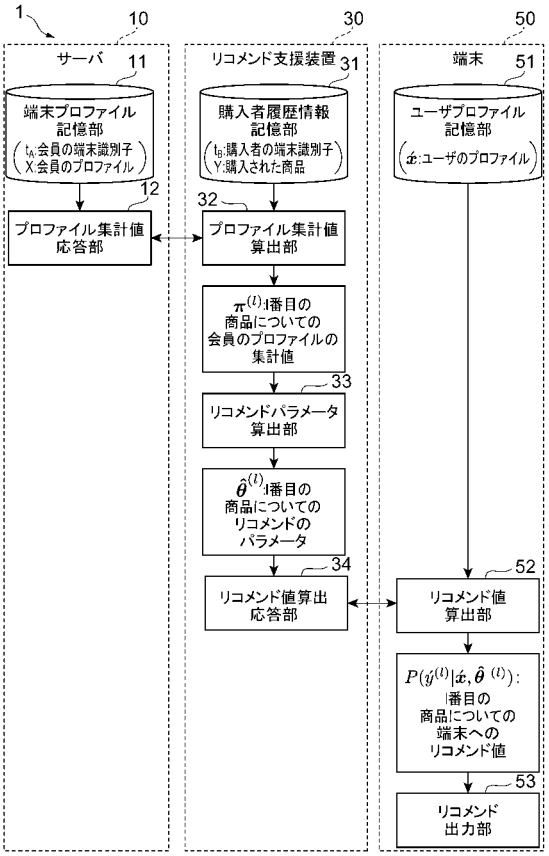
【００８９】

図５に示すように、精度の面では、秘匿内積計算を用いたナイーブベイズ識別器（図５中の「ナイーブベイズ識別器」）と本実施形態の方式が、実績のあるナイーブベイズの手法を実装できるので良好である。セキュアマッチングを用いたクロス集計（図５中の「クロス集計」）はクロス集計表からリコメンドに用いる識別結果を直接得られないので、精度は低い。

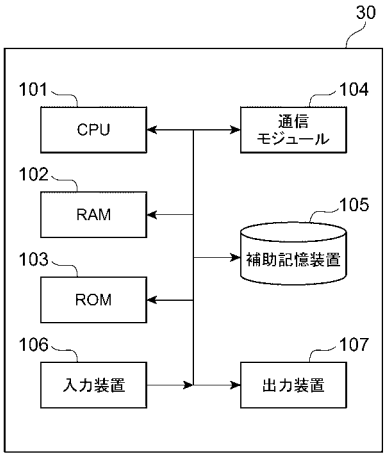
【００９０】
　また、スケールの面では、セキュアマッチングを用いたクロス集計と本実施形態の方式が、大きなクロス集計表を効率的に作成できるので良好である。秘匿内積計算を用いたナイーブベイズ識別器は、秘匿内積計算のコストが高いのでスケーラビリティに劣る。よって、本実施形態の手法によれば、顧客情報を持つサーバ１０と販売情報をもつリコメンド支援装置３０と端末５０との間における処理において、プライバシーを保護できる安全性と、実績のあるナイーブベイズの精度と、顧客数と商品数の増加に対するスケールメリットが得られる。
【００９１】
　本実施形態のリコメンドシステム１によれば、サーバ１０，リコメンド支援装置３０及び端末５０の三者間での処理において、従来と比較して新たにリコメンド支援装置３０内でナイーブベイズの学習パラメータを算出する処理を加えることによって、サーバ１０とリコメンド支援装置との間での処理を、スケーラビリティの低い秘匿内積計算ではなく、スケーラビリティの高いセキュアマッチングを用いることができるようになる。よって、この工夫を取り入れた本実施形態の手法によれば、顧客数と商品数の増加に対するスケールメリットが得られる。
【００９２】
　また、サーバ１０，リコメンド支援装置３０及び端末５０の三者は、それぞれ以下のメリットを享受できる。
１．サーバ１０は顧客から預かっているユーザ属性の情報漏えいを防止できる。
２．リコメンド支援装置３０は、リコメンド支援装置３０の店舗で販売した商品の情報をサーバ１０に渡さずに、リコメンド支援装置３０の店舗の商品を購入する可能性の高いユーザ属性の傾向を知ることができる。また、リコメンド支援装置３０の店舗側の意向をリコメンドに反映することもできる。
３．端末５０は、端末５０のユーザのプライバシー情報をサーバ１０やリコメンド支援装置３０に渡さずに、端末５０のユーザに、ユーザが欲する可能性の高い商品を、その可能性の高い順に閲覧できる。
【００９３】
　以上説明した本実施形態のリコメンドシステム１及びリコメンド方法では、サーバ１０及びリコメンド支援装置３０の間では、それぞれ所定の方式により暗号化された端末プロファイルＸ及び購入者履歴情報Ｙが交換され、これらの情報に基づきリコメンド支援装置３０においてプロファイル集計値πが算出される。これにより、サーバ１０が管理する端末プロファイルＸがリコメンド支援装置３０に漏洩されず、また、リコメンド支援装置３０が管理する購入者履歴情報Ｙがサーバ１０に認識可能な状態で渡されることなく、リコメンド支援装置３０はプロファイル集計値πを得ることができる。また、リコメンド支援装置３０において、プロファイル集計値πに基づき、リコメンド支援装置３０において商品を購入する可能性の高いユーザの属性の傾向をリコメンドパラメータθとして商品ごとに得ることが可能となる。さらに、リコメンド支援装置３０において、所定の方式により端末５０により暗号化されたユーザプロファイルｘ´がリコメンドパラメータθを用いてさらに暗号化されて端末５０に送信され、端末５０において、送信された情報に基づき商品ごとのリコメンド値Ｐが算出される。これにより、端末５０は、端末５０が管理するユーザプロファイルｘ´を、リコメンド支援装置３０やサーバ１０に認識可能な状態で取得されることなく、商品に関するリコメンド情報Ｐを得ることができる。
【００９４】
　以上、本発明をその実施形態に基づいて詳細に説明した。しかし、本発明は上記実施形態に限定されるものではない。本発明は、その要旨を逸脱しない範囲で様々な変形が可能である。
【符号の説明】
【００９５】
　１…リコメンドシステム、１０…サーバ、１１…端末プロファイル記憶部、１２…プロ

ファイル集計値応答部、３０…リコメンド支援装置、３１…購入者履歴情報記憶部、３２…プロファイル集計値算出部、３３…リコメンドパラメータ算出部、３４…リコメンド値算出応答部、５０…端末、５１…ユーザプロファイル記憶部、５２…リコメンド値算出部、５３…リコメンド出力部、Ｐ…リコメンド値、t_A…端末識別子、t_B…端末識別子、ｖ…属性、Ｘ…端末プロファイル、ｘ…ユーザプロファイル、Ｙ…購入者履歴情報、θ…リコメンドパラメータ、π…プロファイル集計値。

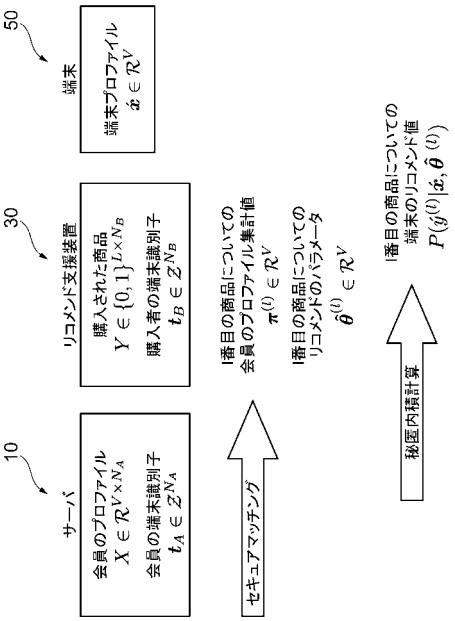
【図１】



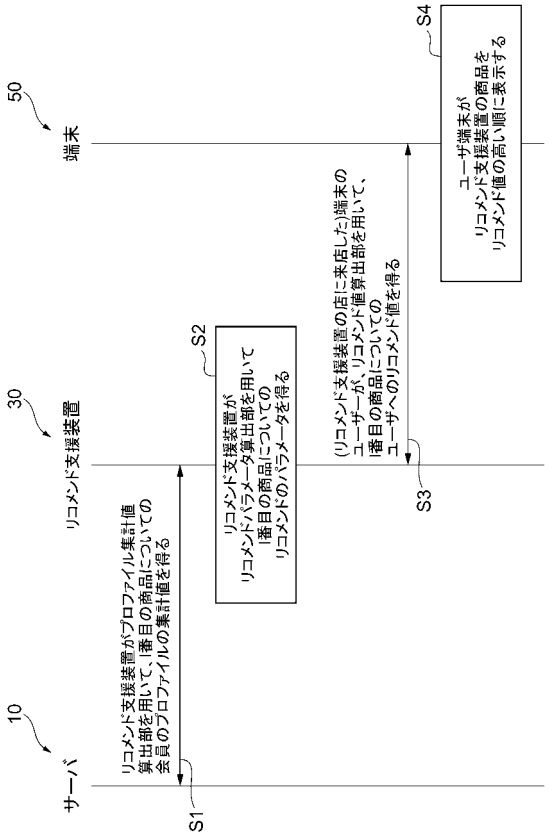
【図２】



【図３】



【図４】



【図５】

評価項目	ナイーブベイズ識別器	クロス集計	本実施形態方式
精度	○	×	○
スケール	×	○	○

フロントページの続き

(51)Int.Cl.	F I	テーマコード（参考）
	Ｇ０６Ｆ 12/14 ５６０Ｚ	

(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

Ｆターム(参考) 5B017 AA08 BA07 CA16
5B075 KK53 KK54 NK45 PR08
5B084 AA01 AA12 AB39 BA02 BB15 CA12 CA13 CB04 CB22 CE03
CE12