

(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2011-154591

(P2011-154591A)

(43) 公開日 平成23年8月11日(2011.8.11)

(51) Int.Cl.	F I	テーマコード (参考)
G06Q 10/00 (2006.01)	G06F 19/00 100	5B075
G06Q 30/00 (2006.01)	G06F 17/60 326	
G06F 17/30 (2006.01)	G06F 17/30 340A	
G06N 5/04 (2006.01)	G06N 5/04 550N	
	G06N 5/04 580A	
審査請求 未請求 請求項の数 10 O L (全 22 頁)		

(21) 出願番号 特願2010-16336 (P2010-16336)
 (22) 出願日 平成22年1月28日 (2010.1.28)

(71) 出願人 000004237
 日本電気株式会社
 東京都港区芝五丁目7番1号
 (74) 代理人 100095407
 弁理士 木村 満
 (72) 発明者 小阪 勇氣
 東京都港区芝五丁目7番1号 日本電気株式会社内
 (72) 発明者 広瀬 俊亮
 東京都港区芝五丁目7番1号 日本電気株式会社内
 Fターム(参考) 5B075 PR04 PR08

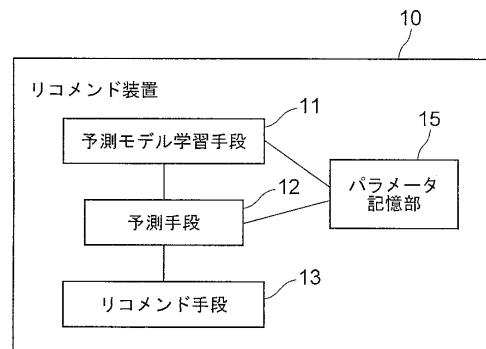
(54) 【発明の名称】 リコメンド装置、方法、及び、プログラム

(57) 【要約】

【課題】 アイテムに対してアクションを起こしたユーザに、そのアイテムに関連があるアイテムをリコメンドできるリコメンド装置を提供する。

【解決手段】 予測モデル学習手段 11 は、パラメータ記憶部 15 から予測モデルのパラメータを読み出し、読み出したパラメータと、複数のユーザのアイテムに対するアクションの履歴を示すシーケンスデータとに基づき、シーケンスデータにおけるユーザがアクションを起こしたアイテムの順序と、リンクの密接度合いで表現されたアイテム間の関連性との双方を考慮して予測モデルを学習する。予測手段 12 は、学習された予測モデルのパラメータと、リコメンド対象のユーザのシーケンスデータとに基づいて、リコメンド対象のユーザが次にアクションを起こすアイテムを予測する。リコメンド手段 13 は、予測されたアイテムに基づいて、ユーザにリコメンドするアイテムを決定する。

【選択図】 図 1



【特許請求の範囲】**【請求項 1】**

ユーザがあるアイテムに対してアクションを起こしたとき、次にアクションを起こすアイテムが特定のアイテムである確率を表す予測モデルのパラメータをパラメータ記憶部から読み出し、該読み出したパラメータと、複数のユーザのアイテムに対するアクションの履歴を示すシーケンスデータとに基づき、前記シーケンスデータにおけるユーザがアクションを起こしたアイテムの順序と、リンクの密接度合いで表現されたアイテム間の関連性との双方を考慮して前記予測モデルを学習し、該学習した予測モデルのパラメータを前記パラメータ記憶部に記憶する予測モデル学習手段と、

前記パラメータ記憶部から前記パラメータを読み出し、該読み出したパラメータと、リコメンド対象のユーザのシーケンスデータとに基づいて、リコメンド対象のユーザが次にアクションを起こすアイテムを予測する予測手段と、

前記予測されたアイテムに基づいて、ユーザにリコメンドするアイテムを決定するリコメンド手段とを備えるリコメンド装置。

【請求項 2】

前記予測モデル学習手段は、リンク構造を表現するKernel法を用いて得られるアイテム間の密接度合いを表すモデルと、複数のユーザのアイテムに対するアクションの順序を表す隠れマルコフモデルとを統合して前記予測モデルを構築し、該構築した予測モデルを学習する、請求項 1 に記載のリコメンド装置。

【請求項 3】

前記予測モデル学習手段は、前記予測モデルの学習では、前記ユーザがあるアイテムに対してアクションを起こしたとしたとき、当該アイテムとリンク構造が密接するアイテムに対してユーザが次にアクションを起こす確率を、当該アイテムとリンク構造が密接しないアイテムに対してユーザが次にアクションを起こす確率よりも高くする、請求項 1 又は 2 に記載のリコメンド装置。

【請求項 4】

ユーザから、閲覧を希望するWebページのリクエストを受け付け、該受け付けたリクエストをユーザリクエスト記憶部に記憶するリクエスト受付手段と、

前記リクエストを行ったユーザの過去のWebページ訪問履歴を前記リクエスト記憶部から読み出す読み出し手段と、

ユーザがあるWebページを閲覧しているとき、次にユーザが閲覧するWebページが特定のWebページである確率を表す予測モデルであって、複数のユーザのWebページ訪問履歴におけるユーザが閲覧したWebページの順序と、リンクの密接度合いで表現されたWebページ間の関連性との双方を考慮して学習された予測モデルのパラメータをパラメータ記憶部から読み出し、該読み出したパラメータと、前記リクエストを行ったユーザの過去のWebページ訪問履歴とに基づいて、ユーザが次に訪問するWebページを予測する予測手段と、

前記予測手段で予測されたWebページに基づいて、ユーザにリコメンドするWebページを決定するリコメンド手段と、

Webページを記憶するWebページ記憶部から、前記ユーザがリクエストしたWebページと、前記リコメンド手段が決定したWebページに関する情報とを読み出し、前記リクエストを行ったユーザに配信するWebページ配信手段とを備えるリコメンドシステム。

【請求項 5】

前記Webページ配信手段は、前記Webページ記憶部から、前記リコメンド手段が決定したWebページへのリンク情報を読み出し、前記ユーザがリクエストしたWebページに前記読み出したリンク情報を付加して前記ユーザに配信する、請求項 4 に記載のリコメンドシステム。

【請求項 6】

前記リクエスト記憶部から読み出された複数のユーザのWebページ訪問履歴と、前記

10

20

30

40

50

パラメータ記憶部に記憶された前記予測モデルのパラメータとに基づいて、前記Webページ訪問履歴におけるユーザが閲覧したWebページの順序と、リンクの密接度合いで表現されたWebページ間の関連性との双方を考慮して前記予測モデルを学習し、該学習した予測モデルのパラメータを前記パラメータ記憶部に記憶する予測モデル学習手段を更に備える、請求項4又は5に記載のリコメンドシステム。

【請求項7】

前記予測モデル学習手段は、リンク構造を表現するKernel法を用いて得られるWebページ間の密接度合いを表すモデルと、複数のユーザのWebページ閲覧順序を表す隠れマルコフモデルとを統合して前記予測モデルを構築し、該構築した予測モデルを学習する、請求項6に記載のリコメンドシステム。

10

【請求項8】

前記予測モデル学習手段は、前記予測モデルの学習では、前記ユーザがあるWebページを訪問したとき、当該Webページとリンク構造が密接するWebページをユーザが次に閲覧する確率を、当該Webページとリンク構造が密接しないWebページをユーザが次に閲覧する確率よりも高くする、請求項6又は7に記載のリコメンドシステム。

【請求項9】

コンピュータが、ユーザがあるアイテムに対してアクションを起こしたとき、次にアクションを起こすアイテムが特定のアイテムである確率を表す予測モデルのパラメータをパラメータ記憶部から読み出すステップと、

コンピュータが、前記読み出したパラメータと、複数のユーザのアイテムに対するアクションの履歴を示すシーケンスデータとに基づき、前記シーケンスデータにおけるユーザがアクションを起こしたアイテムの順序と、リンクの密接度合いで表現されたアイテム間の関連性との双方を考慮して前記予測モデルを学習し、該学習した予測モデルのパラメータを前記パラメータ記憶部に記憶するステップと、

20

コンピュータが、前記パラメータ記憶部から前記パラメータを読み出し、該読み出したパラメータと、リコメンド対象のユーザのシーケンスデータとに基づいて、リコメンド対象のユーザが次にアクションを起こすアイテムを予測するステップと、

コンピュータが、前記予測したアイテムに基づいて、ユーザにリコメンドするアイテムを決定するステップとを有するリコメンド方法。

【請求項10】

30

コンピュータに、

ユーザがあるアイテムに対してアクションを起こしたとき、次にアクションを起こすアイテムが特定のアイテムである確率を表す予測モデルのパラメータをパラメータ記憶部から読み出す処理と、

前記読み出されたパラメータと、複数のユーザのアイテムに対するアクションの履歴を示すシーケンスデータとに基づき、前記シーケンスデータにおけるユーザがアクションを起こしたアイテムの順序と、リンクの密接度合いで表現されたアイテム間の関連性との双方を考慮して前記予測モデルを学習し、該学習した予測モデルのパラメータを前記パラメータ記憶部に記憶する処理と、

前記パラメータ記憶部から前記パラメータを読み出し、該読み出したパラメータと、リコメンド対象のユーザのシーケンスデータとに基づいて、リコメンド対象のユーザが次にアクションを起こすアイテムを予測する処理と、

40

前記予測されたアイテムに基づいて、ユーザにリコメンドするアイテムを決定する処理とを実行させるプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、リコメンド装置、方法、及び、プログラムに関し、更に詳しくは、ユーザのアイテムに対するアクションの履歴を用いて、ユーザにリコメンドするアイテムを決定するリコメンド装置、方法、及び、プログラムに関する。

50

【背景技術】

【0002】

リコメンド装置は、ユーザに対してアイテムをリコメンドする装置である。リコメンド装置は、ユーザのアイテムに対するアクションの履歴を用いて、ユーザにリコメンドするアイテムを決定する。リコメンド技術は、Webページのリコメンドや、商品のリコメンドなど、様々な分野で利用されている。リコメンドするアイテムがWebページの場合、リコメンド装置は、ユーザが訪れたWebページの訪問履歴から、ユーザが次に訪れるWebページを予測し、そのWebページをユーザにリコメンドする。リコメンド装置を用い、ユーザに対して予測したWebページをリコメンドすることで、ユーザは次に訪れるWebページに迅速にアクセスすることができ、ユーザの利便性が向上する。

10

【0003】

非特許文献1及び非特許文献2は、リコメンド技術が記載された文献である。非特許文献1及び非特許文献2に記載のリコメンド装置は、蓄積されたユーザのWebページ訪問履歴から、Webページ間の遷移をマルコフモデルで学習する。リコメンド装置は、ユーザがポータルサイトを訪れると、学習したマルコフモデルを用いて、次にどのWebページを訪れるかを予測する。リコメンド装置は、予測したWebページをユーザに推薦する。ユーザは、推薦されたWebページを迅速に訪問することができ、ユーザの利便性を向上させることができる効果がある。ポータルサイトを訪れたユーザが、いつ、どのWebページを訪れたかは、ポータルサイトのWebサーバに蓄積されている。

20

【0004】

非特許文献1及び非特許文献2に記載のリコメンド装置は、主に、ユーザが訪れた各Webページの訪問回数と、Webページから次のWebページを訪れた遷移の回数との特徴を学習する。言い換えれば、非特許文献1及び非特許文献2では、学習モデルに、蓄積されたユーザのWebページ訪問履歴の各Webページの頻度と遷移の頻度とを学習するモデルを用いている。このモデルは、過去に多くのユーザが連続して訪れるWebページ間の遷移確率は高くなり、連続して訪れるユーザの数が少ないWebページ間の遷移確率は低くなる。このようなモデルを用いて、あるWebページを訪問したユーザに対し、次に訪れるWebページを予測すると、過去に多くのユーザが訪れたWebページをユーザに推薦することができる。

【先行技術文献】

30

【非特許文献】

【0005】

【非特許文献1】Link prediction and path analysis using Markov chains, Sarukkai, R.R, WWW9, 2000

【非特許文献2】Using Markov Chains for Link Prediction in Adaptive Web Sites, Jianhan Zhu, Jun Hong, and John G. Hughes, LNCS 2002

【非特許文献3】Kernel Trick Embedded Gaussian Mixture Model, Jingdong Wang and Jianguo Lee and Changshui Zhang, In Lecture Notes in Artificial Intelligence, 2003

【発明の概要】

40

【発明が解決しようとする課題】

【0006】

非特許文献1及び非特許文献2に記載された手法では、過去に多くのユーザがWebページや連続して訪れたWebページなど、ユーザの訪問頻度が高いWebページを、ユーザが次に訪れるWebページとして、そのユーザに推薦できる。しかしながら、ユーザは、訪問頻度が高いWebページを欲しているとは限らない。例えば、ユーザは、今訪れているWebページに関連した情報を求めていることもあり、その場合に訪問頻度が高いWebページを推薦しても、ユーザの利便性は向上しない。ユーザの利便性向上のためには、過去に多くのユーザが訪れたWebページや連続して訪れたWebページばかりでなく、今訪れているWebページと関連する内容のWebページを推薦することが実用上重要

50

である。

【0007】

訪問頻度が高いWebページと関連する内容のWebページとが一致する場合もある。しかし、近年のポータルサイトは、どのページからも「ポータルサイトのトップページ」を訪れることができるなど、関連しないページへの頻度が高い。例えば、蓄積されたユーザのWebページ訪問履歴に、「アニメ」の特集ページを訪れた後に、「ポータルサイトのトップページ」のページを訪れた履歴が多くあり、特集ページに関連のある別の「アニメ」のページを訪れた履歴が少ない場合を考える。この場合、非特許文献1及び非特許文献2の手法では、「アニメ」の特集ページの次に「ポータルサイトのトップページ」を訪れると予測するモデルを学習する。

10

【0008】

非特許文献1及び非特許文献2の手法では、上記の学習を行う結果、「アニメ」の特集ページを訪れたユーザに、「ポータルサイトのトップページ」を推薦することになる。しかし、「アニメ」の特集ページの次に、「ポータルサイトのトップページ」ではなく、別の「アニメ」のページを推薦すると、ユーザの利便性を高められる場合もある。非特許文献1及び非特許文献2の手法では、ユーザのWebページ訪問履歴に、訪問した「アニメ」の特集ページに関連のある別の「アニメ」のページを訪れた履歴が少ない場合は、ユーザに別の「アニメ」のページを推薦することができず、ユーザの利便性向上を図ることができないという問題がある。

【0009】

本発明は、上記の問題点を解消し、アイテムに対するアクションの頻度が低い場合でも、アイテムに対してアクションを起こしたユーザに、そのアイテムに関連があるアイテムをリコメンドできるリコメンド装置、方法、及び、プログラムを提供することを目的とする。

20

【課題を解決するための手段】

【0010】

本発明は、上記目的を達成するために、ユーザがあるアイテムに対してアクションを起こしたとき、次にアクションを起こすアイテムが特定のアイテムである確率を表す予測モデルのパラメータをパラメータ記憶部から読み出し、該読み出したパラメータと、複数のユーザのアイテムに対するアクションの履歴を示すシーケンスデータとに基づき、前記シーケンスデータにおけるユーザがアクションを起こしたアイテムの順序と、リンクの密接度合いで表現されたアイテム間の関連性との双方を考慮して前記予測モデルを学習し、該学習した予測モデルのパラメータを前記パラメータ記憶部に記憶する予測モデル学習手段と、前記パラメータ記憶部から前記パラメータを読み出し、該読み出したパラメータと、リコメンド対象のユーザのシーケンスデータとに基づいて、リコメンド対象のユーザが次にアクションを起こすアイテムを予測する予測手段と、前記予測されたアイテムに基づいて、ユーザにリコメンドするアイテムを決定するリコメンド手段とを備えるリコメンド装置を提供する。

30

【0011】

本発明は、ユーザから、閲覧を希望するWebページのリクエストを受け付け、該受け付けたリクエストをユーザリクエスト記憶部に記憶するリクエスト受付手段と、前記リクエストを行ったユーザの過去のWebページ訪問履歴を前記リクエスト記憶部から読み出す読み出し手段と、ユーザがあるWebページを閲覧しているとき、次にユーザが閲覧するWebページが特定のWebページである確率を表す予測モデルであって、複数のユーザのWebページ訪問履歴におけるユーザが閲覧したWebページの順序と、リンクの密接度合いで表現されたWebページ間の関連性との双方を考慮して学習された予測モデルのパラメータをパラメータ記憶部から読み出し、該読み出したパラメータと、前記リクエストを行ったユーザの過去のWebページ訪問履歴とに基づいて、ユーザが次に訪問するWebページを予測する予測手段と、前記予測手段で予測されたWebページに基づいて、ユーザにリコメンドするWebページを決定するリコメンド手段と、Webページを記

40

50

憶する Web ページ記憶部から、前記ユーザがリクエストした Web ページと、前記リコメント手段が決定した Web ページに関する情報とを読み出し、前記リクエストを行ったユーザに配信する Web ページ配信手段とを備えるリコメンドシステムを提供する。

【 0 0 1 2 】

本発明は、ユーザがあるアイテムに対してアクションを起こしたとき、次にアクションを起こすアイテムが特定のアイテムである確率を表す予測モデルのパラメータをパラメータ記憶部から読み出すステップと、前記読み出したパラメータと、複数のユーザのアイテムに対するアクションの履歴を示すシーケンスデータとに基づき、前記シーケンスデータにおけるユーザがアクションを起こしたアイテムの順序と、リンクの密接度合いで表現されたアイテム間の関連性との双方を考慮して前記予測モデルを学習し、該学習した予測モデルのパラメータを前記パラメータ記憶部に記憶するステップと、前記パラメータ記憶部から前記パラメータを読み出し、該読み出したパラメータと、リコメンド対象のユーザのシーケンスデータとに基づいて、リコメンド対象のユーザが次にアクションを起こすアイテムを予測するステップと、前記予測したアイテムに基づいて、ユーザにリコメンドするアイテムを決定するステップとを有するリコメンド方法を提供する。

10

【 0 0 1 3 】

本発明は、コンピュータに、ユーザがあるアイテムに対してアクションを起こしたとき、次にアクションを起こすアイテムが特定のアイテムである確率を表す予測モデルのパラメータをパラメータ記憶部から読み出す処理と、前記読み出されたパラメータと、複数のユーザのアイテムに対するアクションの履歴を示すシーケンスデータとに基づき、前記シーケンスデータにおけるユーザがアクションを起こしたアイテムの順序と、リンクの密接度合いで表現されたアイテム間の関連性との双方を考慮して前記予測モデルを学習し、該学習した予測モデルのパラメータを前記パラメータ記憶部に記憶する処理と、前記パラメータ記憶部から前記パラメータを読み出し、該読み出したパラメータと、リコメンド対象のユーザのシーケンスデータとに基づいて、リコメンド対象のユーザが次にアクションを起こすアイテムを予測する処理と、前記予測されたアイテムに基づいて、ユーザにリコメンドするアイテムを決定する処理とを実行させるプログラムを提供する。

20

【 発明の 効果 】

【 0 0 1 4 】

本発明のリコメンド装置、方法、及び、プログラムは、アイテムに対するアクションの頻度が低い場合でも、アイテムに対してアクションを起こしたユーザに、そのアイテムに関連があるアイテムをリコメンドすることができる。

30

【 図面の簡単な説明 】

【 0 0 1 5 】

【 図 1 】 本発明のリコメンド装置の概略構成を示すブロック図。

【 図 2 】 本発明のリコメンドシステムの概略構成を示すブロック図。

【 図 3 】 本発明の一実施形態のリコメンドシステムを示すブロック図。

【 図 4 】 ユーザ端末を示すブロック図。

【 図 5 】 Web サーバを示すブロック図。

【 図 6 】 リコメンド装置を示すブロック図。

40

【 図 7 】 予測モデル学習の動作を示すシーケンス図。

【 図 8 】 コンテンツ訪問履歴の具体例を示す図。

【 図 9 】 リコメンドページ決定の動作を示すシーケンス図。

【 図 1 0 】 リコメンドページへのリンクを付け加えた Web ページを示す図。

【 発明を実施するための形態 】

【 0 0 1 6 】

本発明の実施形態の説明に先立って、本発明の概要を説明する。図 1 は、本発明のリコメンド装置の概略構成を示している。リコメンド装置 1 0 は、予測モデル学習手段 1 1 と、予測手段 1 2 と、リコメンド手段 1 3 とを有する。リコメンド装置 1 0 は、パラメータ記憶部 1 5 を参照して処理を行う。パラメータ記憶部 1 5 は、ユーザがあるアイテムに対

50

してアクションを起こしたとき、次にアクションを起こすアイテムが特定のアイテムである確率を表す予測モデルのパラメータを記憶する。ここで、アイテムとしては、例えば、Web ページや商品などが考えられ、アイテムに対するアクションとしては、例えば Web ページの閲覧や商品の購入などが考えられる。なお、パラメータ記憶部 15 は、リコメンド装置 10 の内部にあっても、外部にあってもよい。

【0017】

予測モデル学習手段 11 は、パラメータ記憶部 15 から予測モデルのパラメータを読み出し、読み出したパラメータと、複数のユーザのアイテムに対するアクションの履歴を示すシーケンスデータとに基づき、シーケンスデータにおけるユーザがアクションを起こしたアイテムの順序と、リンクの密接度合いで表現されたアイテム間の関連性との双方を考慮して予測モデルを学習する。シーケンスデータは、例えばアイテムが Web ページであれば Web ページの訪問履歴とすることができ、アイテムが商品であれば商品の購入履歴とすることができる。予測モデル学習手段 11 は、学習した予測モデルのパラメータをパラメータ記憶部 15 に記憶する。

10

【0018】

予測手段 12 は、パラメータ記憶部 15 から予測モデルのパラメータを読み出す。予測手段 12 は、読み出したパラメータと、リコメンド対象のユーザのシーケンスデータとに基づいて、リコメンド対象のユーザが次にアクションを起こすアイテムを予測する。予測手段 12 は、例えばリコメンド対象のユーザの Web ページ訪問履歴と、読み出したパラメータ（予測モデル）とに基づいて、ユーザが次にどんな Web ページを閲覧するかを予測する。リコメンド手段 13 は、予測手段 12 で予測されたアイテムに基づいて、ユーザにリコメンドするアイテムを決定する。

20

【0019】

本発明のリコメンド装置 10 では、予測モデル学習手段 11 は、複数のユーザがアクションを起こしたアイテムの順序と、リンクの密接度合いで表現されたアイテム間の関連性との双方を考慮して予測モデルを学習する。予測手段 12 は、予測モデル学習手段 11 で学習された予測モデル（パラメータ）を用いて、ユーザが次にアクションを起こすアイテムを予測する。本発明では、アイテム間の関連性をリンクの密集度合いで表現し、各ページの関連性と、シーケンスデータとして蓄積されたアクションの履歴における各アイテムの頻度及びアイテム間の遷移の頻度の特徴とから、次にユーザがアクションを起こすアイテムを予測する。このように予測したアイテムに基づいて、ユーザにリコメンドするアイテムを決定することで、頻度が高いアイテムだけでなく、ユーザがアクションを起こしたアイテムに関連があるアイテムを、ユーザにリコメンドすることができる。

30

【0020】

図 2 は、本発明のリコメンドシステムの概略構成を示している。リコメンドシステム 20 は、リクエスト受付手段 21 と、読み出し手段 22 と、予測手段 23 と、リコメンド手段 24 と、Web ページ配信手段 25 とを有する。リクエスト受付手段 21 は、ユーザから、閲覧を希望する Web ページのリクエストを受け付ける。リクエスト受付手段 21 は、受け付けたリクエストをユーザリクエスト記憶部 26 に記憶する。ユーザリクエスト記憶部 26 は、リコメンドシステムの内部にあっても、外部にあってもよい。読み出し手段 22 は、リクエストを行ったユーザの過去の Web ページ訪問履歴を、ユーザリクエスト記憶部 26 から読み出す。

40

【0021】

パラメータ記憶部 27 は、予測モデルのパラメータを記憶する。ここで、予測モデルは、ユーザがある Web ページを閲覧しているとき、次にユーザが閲覧する Web ページが特定の Web ページである確率を表すモデルである。予測モデルは、予測モデルの学習を行う手段（予測モデル学習手段）が、複数のユーザの Web ページ訪問履歴におけるユーザが閲覧した Web ページの順序と、リンクの密接度合いで表現された Web ページ間の関連性との双方を考慮して学習することで生成される。パラメータ記憶部 27 は、リコメンドシステム 20 の内部にあっても、外部にあってもよい。また、リコメンドシステム 2

50

0 は、予測モデル学習手段を備えていてもよい。

【0022】

予測手段23は、パラメータ記憶部27から、予測モデルのパラメータを読み出す。予測手段23は、読み出したパラメータと、読み出し手段22が読み出したWebページ訪問履歴とに基づいて、ユーザが次に訪問するWebページを予測する。リコメンド手段24は、予測手段23で予測されたWebページに基づいて、ユーザにリコメンドするWebページを決定する。Webページ記憶部28は、Webページを記憶する。Webページ配信手段25は、Webページ記憶部28から、ユーザがリクエストしたWebページと、リコメンド手段24が決定したWebページに関する情報とを読み出す。Webページ配信手段25は、リクエストを行ったユーザに、リクエストされたWebページと、リコメンドするWebページに関する情報とを配信する。

10

【0023】

本発明のリコメンドシステムでは、予測手段23は、複数のユーザのWebページ訪問履歴におけるWebページの閲覧順序と、リンク(ハイパーリンク)の密接度合いで表現されたWebページ間の関連性との双方を考慮した予測モデル(パラメータ)を用いて、ユーザが次に閲覧するWebページを予測する。本発明では、Webページ間の関連性をリンクの密集度合いで表現し、各Webページの関連性と、Webページ訪問履歴における各Webページの頻度及びWebページ間の遷移の頻度の特徴とから、次にユーザが閲覧するWebページを予測する。このように予測したWebページに基づいて、ユーザにリコメンドするWebページを決定することで、閲覧頻度や遷移の頻度が高いWebページアイテムだけでなく、ユーザが閲覧したWebページに関連があるWebページを、ユーザにリコメンドすることができる。

20

【0024】

以下、図面を参照し、本発明の実施の形態を詳細に説明する。図3は、本発明の一実施形態に係るリコメンドシステムを示している。リコメンドシステムは、Webサーバ103とリコメンド装置104とを含む。ユーザ端末101は、ユーザに対して入出力のインターフェースを提供する。ユーザ端末101は、インターネット102を介して、Webサーバ103に所望のページをリクエストする。Webサーバ103は、ユーザからのリクエストの履歴(ページ訪問履歴)を保持する。

【0025】

リコメンド装置104は、ユーザのページ訪問履歴に基づいて、予測モデルを学習する。また、リコメンド装置104は、予測モデルを用いて、ユーザにリコメンドするWebページを決定する。Webサーバ103は、ユーザがリクエストしたWebページと、リコメンド装置104が決定したページ(リコメンドページ)に関する情報とを、ユーザ端末101へ配信する。

30

【0026】

図4は、ユーザ端末101の構成を示している。ユーザ端末101は、例えば携帯電話機やパーソナルコンピュータなどの情報処理装置である。ユーザ端末101は、プログラム制御で動作する。ユーザ端末101は、Webページリクエスト手段201と、Webページ表示手段202とを有する。

40

【0027】

Webページリクエスト手段201は、ユーザが訪問したいWebページを、Webサーバ103へリクエストする。例えば、Webページリクエスト手段201は、ユーザがスポーツのコンテンツを訪問したい場合は、Webサーバ103へスポーツのコンテンツをリクエストする。Webページ表示手段202は、Webサーバ103が配信するコンテンツ(Webページ)を表示する。

【0028】

図5は、Webサーバ103の構成を示している。Webサーバ103は、プログラム制御で動作するサーバシステム(コンピュータシステム)である。Webサーバ103は、Webページ配信手段301、ユーザリクエスト記憶部302、Webページ記憶部3

50

03、リコメンドページ選択手段304、リクエスト受付手段306、読み出し手段307、出力手段308、及び、入力手段309を有する。

【0029】

リクエスト受付手段306は、ユーザから、閲覧を希望するWebページのリクエストを受け付ける。ユーザリクエスト記憶部302は、ユーザからのWebページに対するリクエストの履歴を記憶する。ユーザリクエスト記憶部302は、ユーザが、どのような順序でWebページを訪問したかを示すWebページ訪問履歴を、履歴情報として記憶する。リクエスト受付手段306は、ユーザからのリクエストを受け付けると、受け付けたリクエストをユーザリクエスト記憶部302に記憶する。

【0030】

読み出し手段307は、ユーザリクエスト記憶部302からWebページ訪問履歴を読み出す。出力手段308は、Webページ訪問履歴を、リコメンド装置104へ出力する。リコメンド装置104は、出力手段308から出力されたWebページ訪問履歴を用いて、予測モデルを学習し、ユーザにリコメンドすべきWebページを決定する。入力手段309は、リコメンド装置104から、ユーザに対してリコメンドすべきWebページの識別情報を含むリコメンド結果を受け取る。

【0031】

Webページ記憶部303は、ユーザに対して配信すべきWebページを記憶する。リコメンドページ選択手段304は、入力手段309を介して、リコメンド装置104からリコメンド結果を受け取る。リコメンドページ選択手段304は、Webページ記憶部303から、ユーザがリクエストしたWebページを読み出す。また、リコメンドページ選択手段304は、Webページ記憶部303から、リコメンド結果に含まれるリコメンドページに関する情報を読み出す。Webページ配信手段301は、ユーザからリクエストされたWebページと、リコメンドページに関する情報とを、リクエスト元のユーザのユーザ端末101へ配信する。

【0032】

図6は、リコメンド装置104の構成を示している。リコメンド装置104は、プログラム制御で動作するサーバシステム(コンピュータシステム)である。リコメンド装置104は、予測モデル学習手段402、予測手段403、記憶部404、入力手段405、出力手段406、及び、リコメンド手段407を備える。なお、リコメンド装置104とWebサーバ103とは個別の装置として構成されている必要はなく、これらは、同じサーバ上に構成されていてもよい。

【0033】

入力手段405は、Webサーバ103から、ユーザのWebページ訪問履歴を受け取る。予測モデル学習手段402は、入力手段405を介して、複数のユーザのWebページ訪問履歴を入力し、予測モデルを学習する。ここで、予測モデルは、ユーザがあるWebページを閲覧しているとき、次にユーザが閲覧するWebページが特定のWebページである確率を表すモデルである。記憶部(パラメータ記憶部)404は、予測モデルのパラメータを記憶する。予測モデル学習手段402は、複数のユーザのWebページ訪問履歴と、記憶部404に記憶された予測モデルのパラメータとに基づき、Webページ訪問履歴におけるユーザが閲覧したWebページの順序と、リンクの密接度合いで表現されたWebページ間の関連性との双方を考慮して予測モデルを学習する。

【0034】

予測モデル学習手段402は、リンク構造を表現するKernel法を用いて得られるWebページ間の密接度合いを表すモデルと、複数のユーザのWebページ閲覧順序を表す隠れマルコフモデル(HMM:Hidden Markov Models)とを統合して予測モデルを構築し、構築した予測モデルを学習する。予測モデル学習手段402は、予測モデルの学習では、ユーザがあるWebページを訪問したとき、そのWebページとリンク構造が密接するWebページをユーザが次に閲覧する確率を、そのWebページとリンク構造が密接しないWebページをユーザが次に閲覧する確率よりも高くする。予測モデル学習手段402は、

10

20

30

40

50

学習した予測モデルのパラメータを記憶部404に記憶する。

【0035】

予測手段403は、入力手段405を介して、Webページの閲覧をリクエストしたユーザの過去のWebページ訪問履歴を入力する。予測手段403は、記憶部404から、予測モデルのパラメータを読み出し、読み出したパラメータと、Webページの閲覧をリクエストしたユーザの過去のWebページ訪問履歴とに基づいて、ユーザが次に訪問するWebページを予測する。リコメンド手段407は、予測手段403で予測されたWebページに基づいて、ユーザにリコメンドするWebページを決定する。リコメンド手段407は、決定したWebページを識別する情報を含むリコメンド結果を、出力手段406を介して、Webサーバ103へ出力する。

10

【0036】

なお、図6では、同じリコメンド装置104上で、予測モデルの学習と、学習した予測モデルを用いたリコメンドページの決定とを行っているが、これらは別個の装置で行ってもよい。すなわち、予測モデル学習手段402を有する装置(モデル学習装置)と、予測手段403及びリコメンド手段407を有する装置(リコメンド装置)とを分けることも可能である。予測モデルのパラメータを記憶するパラメータ記憶部は、モデル学習装置とリコメンド装置との双方に設けられていてもよく、或いは、モデル学習装置とリコメンド装置とで、共通のパラメータ記憶部を用いてもよい。

【0037】

本実施形態の動作は、大きく分けて、予測モデルの学習処理と、学習した予測モデルを用いてユーザにWebページをリコメンドする処理との2つがある。まず、予測モデルの学習の処理の手順を説明する。図7は、予測モデル学習の動作を示すシーケンス図である。Webサーバ103の読み出し手段307(図5)は、ユーザリクエスト記憶部302から、過去にリクエストを送信したことがある複数のユーザのWebページ訪問履歴を読み出す(ステップA1)。読み出し手段307は、例えばユーザからのリクエストがあったとき、当該ユーザを含む複数のユーザのWebページ訪問履歴を読み出す。或いは、読み出し手段307は、ユーザリクエスト記憶部302にある程度の履歴が溜まったときに、複数のユーザのWebページ訪問履歴を読み出してもよい。

20

【0038】

図8は、Webページ訪問履歴の具体例を示している。Webページ訪問履歴は、各ユーザが閲覧したWebページのURL(Uniform Resource Locator)を、閲覧順に並べたシーケンスデータである。例えば、ユーザAのWebページ訪問履歴は、ユーザAが、スポーツ(<http://sports>)、アニメ(<http://anime>)、ニュース(<http://news>)の順で、Webページを閲覧したことを示している。読み出し手段307は、ステップA1では、各ユーザのWebページ訪問履歴を読み出す。

30

【0039】

読み出し手段307は、出力手段308を介して、リコメンド装置104へ、読み出した複数ユーザのWebページ訪問履歴を出力する(ステップA2)。リコメンド装置104の予測モデル学習手段402(図6)は、入力手段405を介して、複数ユーザのWebページ訪問履歴を受け取る。予測モデル学習手段402は、予測モデルのパラメータを記憶部404から読み出す(ステップA3)。予測モデル学習手段402は、まだ一度も予測モデルを学習していないとき、つまり、学習済みの予測モデルのパラメータが記憶部404に記憶されていないときは、記憶部404からパラメータの初期値を読み出すものとする。

40

【0040】

予測モデル学習手段402は、Webページ訪問履歴と、記憶部404から読み出したパラメータとを用いて、予測モデルを学習する(ステップA4)。予測モデル学習手段402は、学習した予測モデルのパラメータを記憶部404に記憶する(ステップA5)。

【0041】

本実施形態では、予測モデル学習手段402は、予測モデルに隠れマルコフモデルを用

50

いる。隠れマルコフモデルは、シーケンスデータを、データの発生モデルを持つ隠れ状態の遷移として表現するデータモデルである。ここで、“隠れ”という用語は、実際には観測できないことを意味する。隠れマルコフモデルは、実際には観測されていない隠れ状態が遷移することで、シーケンスデータが生成されると仮定したモデルである。隠れマルコフモデルは、文字認識、自然言語処理、音声認識、動作認識など、幅広く利用されている。隠れマルコフモデルでは、データをシンボルの列として表現する。本実施形態の場合、各シンボルは、Web ページを表す。

【0042】

予測モデル学習手段402が読み出し手段307から受け取る複数ユーザのWeb ページ訪問履歴（入力データ）をXとする。ユーザ数をMとし、各ユーザのWeb ページ訪問履歴をX_i（i = 1 ~ M）とすると、X = (X₁, . . . , X_M)である。X_iは、i番目のユーザのWeb ページ訪問履歴を表す。各ユーザのWeb ページ訪問履歴は、各ユーザが訪問したWeb ページを時系列で並べたシーケンスデータである。X₁はX₁ = x₁₁, x₁₂, x₁₃, . . . で表される。例えば、ユーザが、「アニメ」、「ニュース」、「天気予報」の順にWeb ページをリクエストしたとすれば、x₁₁ = 「アニメ」のWeb ページ、x₁₂ = 「ニュース」のWeb ページ、x₁₃ = 「天気予報」のWeb ページとなる。

10

【0043】

隠れマルコフモデルは、下記式1でモデル化する統計モデルである。

【数1】

20

$$\begin{aligned}
 P(X, Z, S | \theta) &= \prod_{m=1}^M \prod_{k=1}^K \pi_k^{z_{mk}} P(X_m, S_k | \theta_k)^{z_{mk}} \\
 &= \prod_{m=1}^M \prod_{k=1}^K \pi_k^{z_{mk}} P(s_1^k | \Gamma^k)^{z_{mk}} \left[\prod_{n=2}^{N_m} P(s_n^k | s_{n-1}^k, A^k)^{z_{mk}} \right] \prod_{n=1}^{N_m} P(x_{mn} | s_n^k, B^k)^{z_{mk}}
 \end{aligned}
 \tag{1}$$

Mは、ユーザ数、つまり、シーケンスデータの数をわらし、N_mは、m番目のシーケンスデータの長さを表す。Kは、HMMの混合数である。HMMを混合しない場合は、K = 1とする。X = {X₁, . . . , X_M}は、シンボル列を表す。X_mは、m番目のシンボル列を表す。x_{m n}は、m個目のシーケンスデータのn個目のシンボルを表す。S_kは、コンポーネントk番目の隠れ状態集合を表す。S = {S₁, . . . , S_k}である。ここで、コンポーネントとは、モデルの混合番号である。

30

【0044】

π_kは、各コンポーネントの混合確率を表す。0 ≤ π_k ≤ 1であり、

【数2】

$$\sum_{k=1}^K \pi_k = 1$$

である。zは、k次元の2値の確率変数であり、z_{m k} ∈ {0, 1}である。P(z_{m k} = 1) = π_kであり、

【数3】

40

$$\sum_{k=1}^K z_{mk} = 1$$

である。zは、indicator functionと呼ばれる。Zは、zの集合を表す。s₁^kは、コンポーネントk番目の時刻1における隠れ状態を表す。s_n^kは、コンポーネントk番目の時刻nにおける隠れ状態を表す。π_i^k = P(s_{1 i} = 1)は、時刻1のときに隠れ状態がiの確率を表す。S_i^kは、i^kを要素に持つ集合であり、k番目のコンポーネントのS_i^kを表す。ここで、

【数 4】

$$\sum_{i=0}^{I^k} \gamma_i^k = 1$$

である。。

【0045】

A^k は、コンポーネント k 番目の隠れ変数の遷移確率のパラメータを表す。 $P(s_{n^k} | S_{n-1}^k, A^k) Z^{m^k}$ は、パラメータ A^k のもと、 k 番目のコンポーネントの、 $n-1$ 個目の隠れ状態から n 個目の隠れ状態への遷移確率を表す。 B^k は、コンポーネント k 番目のシンボルの生起確率のパラメータを表す。 $P(x_{m_n} | S_n^k, B^k) Z^{m^k}$ は、パラメータ B^k のもと、 k 番目のコンポーネントにおいて、 n 個目の隠れ状態から、 m 個目のシーケンスデータの n 個目のシンボルが生起した確率を表す。

10

【0046】

通常の HMM では、複数のユーザの Web ページの閲覧順序をモデル化することができるものの、リンクの密接度合いで表現される各 Web ページ間の関連性も併せてモデル化することは困難である。本実施形態では、予測モデルに、複数のユーザの Web ページの閲覧順序と、リンク（ハイパーリンク）の密接度合いで表現された各 Web ページ間の関連性との双方を考慮し、ユーザが次に訪れる確率が高い Web ページを予測する予測モデルを用いる。このような予測モデルを得るために、本実施形態では、HMM の $P(x_{m_n} | S_n^k, B^k) Z^{m^k}$ に、リンク構造を表現するカーネル関数を導入する。そのような HMM を用いることで、複数のユーザの Web ページの閲覧順序と、リンクの密接度合いで表現される各 Web ページ間の関連性との両方を合わせた予測モデルが得られる。

20

【0047】

以下では、 $P(x_{m_n} | S_n^k, B^k) Z^{m^k}$ に Commute Time Kernel を導入する例を説明する。ただし、この部分は、リンク構造を表現するカーネル関数であればよく、その他の Kernel を導入することも可能である。例えば、 $P(x_{m_n} | S_n^k, B^k) Z^{m^k}$ に、Diffusion Kernel を導入してもよい。

【0048】

Commute Time Kernel は、グラフラプラシアン of the 擬似逆行列として定義できる。グラフラプラシアンは、 $R \times R$ の行列である。 R は、 $X = \{X_1, \dots, X_M\}$ に含まれる Web ページの種類の数を表す。グラフラプラシアンの例を下記式 2 に示す。

30

【数 5】

$$R = \begin{pmatrix} -2 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix} \quad (2)$$

上記式 2 において、グラフラプラシアンの各行及び各列は、それぞれ Web ページ 1 ~ Web ページ 3 に対応している。グラフラプラシアンの 1 行目は、Web ページ 1 と Web ページ 2、及び、Web ページ 1 と Web ページ 3 のそれぞれにリンクがある旨を示している。1 行 1 列目には、Web ページ 1 とリンクでつながる Web ページの総数にマイナスをかけた値が入る。グラフラプラシアンの 2 行目は、Web ページ 2 と Web ページ 1 とにリンクがあることを表す。2 行 2 列目には、Web ページ 2 とリンクでつながる Web ページの総数にマイナスをかけた値が入る。グラフラプラシアンの 3 行目は、Web ページ 3 と Web ページ 1 とにリンクがあることを表す。3 行 3 列目には、Web ページ 3 とリンクでつながる Web ページの総数にマイナスをかけた値が入る。

40

【0049】

本実施形態では、ユーザがある Web ページを訪問した後に、別の Web ページを訪問したとき、そのある Web ページと別の Web ページとにリンクがあるとしている。例えば、ユーザが Web ページ 1 の次に Web ページ 2 を訪れた場合に、Web ページ 1 と Web ページ 2 とにリンクがあるとした。予測モデル学習手段 402 は、 $X = \{X_1, \dots$

50

・ , X_M }のシーケンスデータ中にWebページが連続していれば、連続したWebページ間にはリンクがあるとして、モデルを学習する。

【0050】

Commute Time Kernelは、Webページ間の関連の強さをリンクの密集度合いで測ることができる距離関数である。Commute Time Kernelは、Webページ間を行き来するルートが多い場合は、その2つのWebページは距離が近く、Webページ間を行き来するルートが少ない場合は、その2つのWebページは距離が遠いことを表す距離関数である。

【0051】

本実施形態では、HMMのP(x_{mn} | S_n^k, B^k) Z^{m_k}に、ガウス分布を用いて、Commute Time Kernel Kを導入する(下記式3)。

10

【数6】

$$P(x_{mn} | s_n^k, B^k)^{z_{mk}} = N(x_{mn} | \mu_k, \Sigma_k) \quad (3)$$

Kを導入するために、上記式3を変形する。

【数7】

$$P(x_{mn} | B^k)^{z_{mk}} = N(s_i = S(x_{mn}) | B_k) = \frac{1}{(2\pi)^{\frac{\sigma}{2}} \prod_{e=1}^{\sigma} (\lambda_{ke}^{\frac{1}{2}})} \exp\left(-\frac{1}{2} \sum_{e=1}^{\sigma} \frac{y_{ke}^2}{\lambda_{ke}}\right) \frac{1}{(2\pi\rho_k)^{\frac{D-\sigma}{2}}} \exp\left(-\frac{\varepsilon_k^2(s_i)}{2\rho_k}\right) \quad (4)$$

上記式4において、S(x_{mn})は、Webページの種類を表す。s_i = S(x_{mn})は、x_{mn}のWebページの種類がWebページiであることを表す。y_{ke}は、y_{ke} = (x_{mn} - μ_k) V_{ke}である。V_{ke}は、_kのe番目に大きい固有値を表す。_kは、コンポーネントkの、_kのe番目に大きい固有値を表す。ε_k²は、

20

【数8】

$$\varepsilon_k^2 = \sum_{i=\sigma+1}^D \lambda_{ki}$$

である。Dは、_kの固有値の全個数を表す。_kは、D個の固有値のうち、上位個ということを表す。

【数9】

$$\rho_k = \frac{1}{D-\sigma} \sum_{i=\sigma+1}^{\sigma} \lambda_{ki}$$

である。

【0052】

ここで、

【数10】

$$\sum_k V_{ke} = \lambda_{ke} V_{ke} \Leftrightarrow \tilde{K}_k \beta_{ke} = \lambda_{ke} \beta_{ke}$$

が成り立つことを利用すると、_{ke}とy_{ke}とは、μ_kと_kとを用いずに、K~_kとK~_k'で計算できる。_{ke}は、K~_kのe番目に大きい固有値。y_{ke} = _k^T K~_k(i, j)'である。_k = [_{k1}, ..., _{kD}]。K~_kとK~_k'は、それぞれ、

40

【数11】

$$\tilde{K}_k = K_k - W_k K_k - K_k W_k + W_k K_k W_k$$

$$\tilde{K}_k' = K_k' - W_k' K_k' - K_k' W_k' + W_k' K_k' W_k'$$

である。ここで、

50

【数 1 2】

$$W_k = \omega_k \omega_k^T, W_k' = 1_D \omega_k^T$$

$$K_{k,(i,j)} = (\omega_{ki} \phi(s_i) \bullet \omega_{kj} \phi(s_j)), K_{k,(i,j)}' = (\tilde{\phi}(s_i) \bullet \omega_{kj} \tilde{\phi}(s_j))$$

である。(・)は、内積を表す。

【数 1 3】

$$\omega_{ki} = \sqrt{\gamma_i^k}$$

$$\mu_k = \sum_{i=1}^{N \times M} \phi(s_i) \omega_{ki}^2$$

$$\tilde{\phi}(s_i) = \phi(s_i) - \mu_k$$

とする。

【0053】

本実施形態では、 $K_k, (i, j)$ を Commute Time Kernel とすることで、Web ページ間の関連性をモデルに組み込む。パラメータ B^k は、

【数 1 4】

$$B^k = \{ \tilde{K}_k, \tilde{K}_k' \}$$

とする。これについては、例えば、非特許文献 3 に記載されている。

【0054】

モデルのパラメータは、 $\theta = (\theta^1, \theta^2, \dots, \theta^K)$ 、 $\theta^k = (\theta^k, \theta^k, A^k, B^k)$ とする。予測モデル学習手段 402 は、ステップ A3 では、記憶部 404 から $\theta = (\theta^1, \theta^2, \dots, \theta^K)$ を読み出す。予測モデル学習手段 402 は、ステップ A4 で、Web ページ訪問履歴と、読み出したパラメータとを用いて、予測モデルを学習する。予測モデル学習手段 402 は、ステップ A5 で、学習したモデルのパラメータを、記憶部 404 に記憶する。

【0055】

次いで、リコメンドページの決定する処理の手順を説明する。図 9 は、学習した予測モデルを用いたリコメンドページ決定の動作を示すシーケンス図である。ユーザ端末 101 の Web ページリクエスト手段 201 は、ユーザからの指示に基づいて、Web サーバ 103 に、閲覧を希望する Web ページをリクエストする (ステップ B1)。Web サーバ 103 のリクエスト受付手段 306 は、ユーザ端末 101 からのリクエストを受け付ける (ステップ B2)。リクエスト受付手段 306 は、受け付けたリクエストを、ユーザリクエスト記憶部 302 に記憶する。

【0056】

Web ページ配信手段 301 は、読み出し手段 307 に、リクエストを受け付けたユーザの情報を出力する (ステップ B3)。読み出し手段 307 は、ユーザリクエスト記憶部 302 から、リクエストを受け付けたユーザの過去の Web ページ訪問履歴を読み出す (ステップ B4)。読み出し手段 307 は、読み出したユーザの過去の Web ページ訪問履歴とを、出力手段 308 を通じて、リコメンド装置 104 へ出力する (ステップ B5)。

【0057】

リコメンド装置 104 の予測手段 403 は、入力手段 405 を通じて、対象ユーザの過去の Web ページ訪問履歴を受け取る。予測手段 403 は、記憶部 404 から、学習した予測モデルのパラメータを読み出す (ステップ B6)。予測手段 403 は、対象ユーザの

10

20

30

40

50

Web ページ訪問履歴 $X_n = x_{n1}, \dots, x_{nh}$ から、ユーザが h 番目の Web ページの次に、 $\{1, \dots, R\}$ のどの Web ページを訪れる可能性が高いかを示す尤度を、下記式 5、式 6 を用いて計算する (ステップ B 7)。ここで、 x_{nh+1} を、ユーザが x_{nh} の次に訪れる Web ページとする。尤度は、式 5 を用いて算出できる。

【数 1 5】

$$P(X_n, x_{nh+1}, Z, S | \theta) = \prod_{k=1}^K \pi_k^{z_{nk}} P(X_n, x_{nh+1}, S_k | \theta_k)^{z_{nk}} \quad (5)$$

$$= \prod_{k=1}^K [\pi_k^{z_{nk}} P(s_1^k | \Gamma^k)^{z_{nk}} \prod_{n=2}^{N_n} P(s_n^k | s_{n-1}^k, A^k)^{z_{nk}} \prod_{n=1}^{N_n} P(x_{mn} | s_n^k, B^k)^{z_{nk}}]$$

10

【数 1 6】

$$P(X_n, x_{nh+1}) = \sum_{Z, S} P(X, Z, S | \theta) \quad (6)$$

式 6 は、Forward-backward アルゴリズムを用いて、計算できる。

【0058】

リコメンド手段 407 は、予測手段 403 で算出した $P(X_n, x_{nh+1})$ の中で、値が最も大きくなる x_{nh+1} の Web ページを、ユーザにリコメンドするページとして決定する (ステップ B 8)。リコメンド手段 407 は、リコメンド結果、すなわち、ユーザにどの Web ページをリコメンドするかを示す情報を、出力手段 406 を通じて、Web サーバ 103 へ送信する (ステップ B 9)。

20

【0059】

Web サーバ 103 のリコメンドページ選択手段 304 は、入力手段 309 を通じて、リコメンド結果を受け取る。リコメンドページ選択手段 304 は、Web ページ記憶部 303 から、リコメンドする Web ページに関する情報として、リコメンドページへのリンク情報を読み出し、Web ページ配信手段 301 へ出力する (ステップ B 10)。また、リコメンドページ選択手段 304 は、Web ページ記憶部 303 から、ユーザがリクエストした Web ページを読み出し、Web ページ配信手段 301 へ出力する (ステップ B 11)。なお、リコメンドページへのリンク情報の読み出し、及び、ユーザがリクエストした Web ページの読み出しは、Web ページ配信手段 301 が行ってもよい。

30

【0060】

Web ページ配信手段 301 は、ユーザがリクエストした Web ページに、リコメンドする Web ページへのリンクを付け加える。Web ページ配信手段 301 は、Web ページをリクエストしたユーザのユーザ端末 101 に、リコメンドする Web ページへのリンクを付け加えた Web ページを配信する (ステップ B 12)。

【0061】

ユーザ端末 101 の Web ページ表示手段 202 は、リコメンドする Web ページへのリンクを含む Web ページを受け取り、表示画面上に表示する (ステップ B 13)。図 10 は、画面表示例を示している。Web ページ 801 は、通常コンテンツに加えて、リコメンドされた Web ページへのリンク 802 を含む。ユーザは、このリンク 802 をクリックすることで、リコメンドされた Web ページの閲覧が可能である。

40

【0062】

本実施形態では、予測モデル学習手段 402 (図 6) は、複数のユーザの Web ページ訪問履歴におけるユーザが閲覧した Web ページの順序と、各ページ間の関連性がリンク (ハイパーリンク) の密接度合いで表現された Web ページ間の関連性との双方を考慮して予測モデルを学習する。本実施形態では、あるページ間にリンクがある場合に、そのページ間に関連性があるとする。蓄積されたユーザの訪問履歴に、1 度以上連続してユーザが訪れていたページ間には、リンクがあるとする。リンクの関連の強さは、リンクの密集度合いで表現する。Web ページ間を行き来するリンクのルートの種類が多い場合は関連性が強く、ルートの種類が少ない場合は関連性が小さい。

50

【 0 0 6 3 】

予測手段 4 0 3 は、各ページ間の関連性をリンクの密集度合いで表現し、各ページの関連性、及び、蓄積されたユーザの訪問履歴の各 Web ページの頻度と遷移の頻度との特徴から、次にユーザが訪れる Web ページを予測する。Web ページ配信手段 3 0 1 (図 5) は、予測された Web ページ (リコメンドページ) に関する情報を、ユーザがリクエストした Web ページに付け加えて配信する。ユーザは、リコメンドされた Web ページに関する情報を参照し、必要に応じて、リコメンドページを閲覧することができる。

【 0 0 6 4 】

ユーザは、現在訪れている Web ページに関連した情報を求めていることもあるため、訪問頻度が高い Web ページばかりを推薦すると、ユーザの利便性の低下を招くことがある。そのため、過去に多くのユーザが訪れた Web ページや連続して訪れた Web ページの他に、現在訪れている Web ページと関連する内容の Web ページを推薦することが、実用上重要である。本実施形態では、ページ間の関連性も考慮した予測モデルを用いてリコメンドページを決定し、ユーザにリコメンドするため、訪問頻度が高い Web ページばかりでなく、Web ページを閲覧するユーザに、その Web ページに関連性のある Web ページをリコメンドすることができる。

【 0 0 6 5 】

例えば、ポータルサイト内に蓄積されたユーザの Web ページ訪問履歴に、「アニメ」の特集ページを訪れた後に、「天気予報」の Web ページを訪れた履歴が多く存在する場合を考える。この場合に、蓄積されたユーザの Web ページ訪問履歴の各 Web ページの頻度と遷移の頻度との特徴から次にユーザが訪れる Web ページを予測すると、「アニメ」の特集ページの次に「天気予報」の Web ページを訪れると予測することになる。本実施形態では、各ページの関連性も考慮して、ユーザが次に訪れる Web ページを予測する。「アニメ」の特集ページと「天気予報」の Web ページとは、ページの内容の関連性が低く、ページを提供するポータル内では、この両者間を行き来するルートは、一般的に少ないと考えられる。従って、本実施形態では、予測モデルの学習において、このページの関連性は低いと計算される。一方、蓄積されたユーザの訪問履歴に、「アニメ」の特集ページからアニメの購入ページを訪れた履歴が 1 つ以上存在した場合、両者の間にリンクが存在すると仮定する。この両者間のページの内容は、関連性も高く、ページを提供するポータルサイト内には、両者を行き来するリンクが多数存在すると考えられる。従って、本実施形態では、予測モデルの学習において、このページの関連性が高いと計算される。このようにページ間の関連性を計算し、蓄積されたユーザの Web ページ訪問履歴の各 Web ページの頻度と遷移の頻度との特徴と組み合わせることで、「アニメ」の特集ページを訪れたユーザに対して、訪問頻度と遷移の頻度とが高い「天気予報」のページではなく、アニメ購入ページを推薦することができる。なお、上記の例は、一例であり、訪問頻度、遷移の頻度、関連性の組み合わせから、次にユーザが訪れる Web ページを予測した結果、訪問頻度と遷移の頻度が高い「天気予報」のページを推薦することも考えられる。

【 0 0 6 6 】

本実施形態では、ポータルサイトに蓄積された Web ページの訪問履歴から抽出した、訪問頻度、遷移の頻度、及び、ページ間の関連性の特徴から、Web ページを閲覧するユーザに対し、その Web ページに関連した Web ページをリコメンドできる。本実施形態では、リンクの密接度合いで表現された Web ページ間の関連性を、蓄積された Web ページ訪問履歴から得ているので、予測モデルの学習に際して、各ページに関連する情報をあらかじめ準備しておく必要がないという効果もある。

【 0 0 6 7 】

なお、上記実施形態では、Web ページを閲覧するユーザに Web ページをリコメンドする例を説明したが、これには限定されない。例えば、ユーザの商品購入の履歴を用いて、商品を購入するユーザに、その商品に関連がある商品をリコメンドしてもよい。また、上記実施形態では、予測モデルの学習において、Web ページ閲覧履歴に含まれる Web ページをシンボル列として HMM を学習する例を説明したが、シンボル列は、Web ペー

10

20

30

40

50

ジ（アイテム）そのものを並べたものには限定されない。例えば、あらかじめWebページをカテゴリに分けておき、シンボルとしてカテゴリを用い、ユーザが閲覧したWebページのカテゴリを並べたシンボル列に対してHMMを学習してもよい。その場合、リコメンド手段407（図6）は、ユーザが次に閲覧すると予測されたカテゴリに属するWebページの中から、ユーザにリコメンドするWebページを決定すればよい。Webページをカテゴリ分けしておく構成では、HMMの計算コストを低減できる効果がある。

【0068】

以上、本発明をその好適な実施形態に基づいて説明したが、本発明のリコメンド装置、システム、方法、及び、プログラムは、上記実施形態にのみ限定されるものではなく、上記実施形態の構成から種々の修正及び変更を施したのも、本発明の範囲に含まれる。

10

【0069】

（付記1）

コンピュータが、ユーザがあるアイテムに対してアクションを起こしたとき、次にアクションを起こすアイテムが特定のアイテムである確率を表す予測モデルのパラメータをパラメータ記憶部から読み出すステップと、

コンピュータが、前記読み出したパラメータと、複数のユーザのアイテムに対するアクションの履歴を示すシーケンスデータとに基づき、前記シーケンスデータにおけるユーザがアクションを起こしたアイテムの順序と、リンクの密接度合いで表現されたアイテム間の関連性との双方を考慮して前記予測モデルを学習し、該学習した予測モデルのパラメータを前記パラメータ記憶部に記憶するステップと、

20

コンピュータが、前記パラメータ記憶部から前記パラメータを読み出し、該読み出したパラメータと、リコメンド対象のユーザのシーケンスデータとに基づいて、リコメンド対象のユーザが次にアクションを起こすアイテムを予測するステップと、

コンピュータが、前記予測したアイテムに基づいて、ユーザにリコメンドするアイテムを決定するステップとを有するリコメンド方法。

（付記2）

前記予測モデルを学習するステップでは、コンピュータは、リンク構造を表現するKernel法を用いて得られるアイテム間の密接度合いを表すモデルと、複数のユーザのアイテムに対するアクションの順序を表す隠れマルコフモデルとを統合して前記予測モデルを構築し、該構築した予測モデルを学習する、付記1に記載のリコメンド方法。

30

（付記3）

前記予測モデルを学習するステップでは、コンピュータは、前記ユーザがあるアイテムに対してアクションを起こしたとしたとき、前記予測モデルにおける、当該アイテムとリンク構造が密接するアイテムに対してユーザが次にアクションを起こす確率を、当該アイテムとリンク構造が密接しないアイテムに対してユーザが次にアクションを起こす確率よりも高くする、付記1又は2に記載のリコメンド方法。

（付記4）

コンピュータが、ユーザから、閲覧を希望するWebページのリクエストを受け付け、受け付けたリクエストをユーザリクエスト記憶部に記憶するステップと、

コンピュータが、前記リクエストを行ったユーザの過去のWebページ訪問履歴を前記リクエスト記憶部から読み出すステップと、

40

コンピュータが、ユーザがあるWebページを閲覧しているとき、次にユーザが閲覧するWebページが特定のWebページである確率を表す予測モデルであって、複数のユーザのWebページ訪問履歴におけるユーザが閲覧したWebページの順序と、リンクの密接度合いで表現されたWebページ間の関連性との双方を考慮して学習された予測モデルのパラメータをパラメータ記憶部から読み出し、該読み出したパラメータと、前記リクエストを行ったユーザの過去のWebページ訪問履歴とに基づいて、ユーザが次に訪問するWebページを予測するステップと、

コンピュータが、前記予測したWebページに基づいて、ユーザにリコメンドするWebページを決定するステップと、

50

コンピュータが、Web ページを記憶する Web ページ記憶部から、前記ユーザがリクエストした Web ページと、前記リコメンドすると決定した Web ページに関する情報とを読み出し、前記リクエストを行ったユーザに配信するステップを有するリコメンド方法。

(付記 5)

前記 Web ページを配信するステップでは、コンピュータは、前記 Web ページ記憶部から、前記リコメンドすると決定した Web ページへのリンク情報を読み出し、前記ユーザがリクエストした Web ページに前記読み出したリンク情報を付加して前記ユーザに配信する、付記 4 に記載のリコメンド方法。

(付記 6)

コンピュータが、前記リクエスト記憶部から読み出された複数のユーザの Web ページ訪問履歴と、前記パラメータ記憶部に記憶された前記予測モデルのパラメータとに基づいて、前記 Web ページ訪問履歴におけるユーザが閲覧した Web ページの順序と、リンクの密接度合いで表現された Web ページ間の関連性との双方を考慮して前記予測モデルを学習し、該学習した予測モデルのパラメータを前記パラメータ記憶部に記憶するステップを更に有する、付記 4 又は 5 に記載のリコメンド方法。

(付記 7)

前記予測モデルを学習するステップでは、コンピュータは、リンク構造を表現する Kernel 法を用いて得られる Web ページ間の密接度合いを表すモデルと、複数のユーザの Web ページ閲覧順序を表す隠れマルコフモデルとを統合して前記予測モデルを構築し、該構築した予測モデルを学習する、付記 6 に記載のリコメンド方法。

(付記 8)

前記予測モデルを学習するステップでは、コンピュータは、前記ユーザがある Web ページを訪問したとき、前記予測モデルにおける、当該 Web ページとリンク構造が密接する Web ページをユーザが次に閲覧する確率を、当該 Web ページとリンク構造が密接しない Web ページをユーザが次に閲覧する確率よりも高くする、付記 6 又は 7 に記載のリコメンド方法。

(付記 10)

コンピュータに、

ユーザがあるアイテムに対してアクションを起こしたとき、次にアクションを起こすアイテムが特定のアイテムである確率を表す予測モデルのパラメータをパラメータ記憶部から読み出す処理と、

前記読み出されたパラメータと、複数のユーザのアイテムに対するアクションの履歴を示すシーケンスデータとに基づき、前記シーケンスデータにおけるユーザがアクションを起こしたアイテムの順序と、リンクの密接度合いで表現されたアイテム間の関連性との双方を考慮して前記予測モデルを学習し、該学習した予測モデルのパラメータを前記パラメータ記憶部に記憶する処理と、

前記パラメータ記憶部から前記パラメータを読み出し、該読み出したパラメータと、リコメンド対象のユーザのシーケンスデータとに基づいて、リコメンド対象のユーザが次にアクションを起こすアイテムを予測する処理と、

前記予測されたアイテムに基づいて、ユーザにリコメンドするアイテムを決定する処理とを実行させるプログラム。

(付記 11)

前記予測モデルを学習する処理では、リンク構造を表現する Kernel 法を用いて得られるアイテム間の密接度合いを表すモデルと、複数のユーザのアイテムに対するアクションの順序を表す隠れマルコフモデルとを統合して前記予測モデルを構築し、該構築した予測モデルを学習する、付記 10 に記載のプログラム。

(付記 12)

前記予測モデルを学習する処理では、前記ユーザがあるアイテムに対してアクションを起こしたとしたとき、前記予測モデルにおける、当該アイテムとリンク構造が密接するア

10

20

30

40

50

アイテムに対してユーザが次にアクションを起こす確率を、当該アイテムとリンク構造が密接しないアイテムに対してユーザが次にアクションを起こす確率よりも高くする、付記 10 又は 11 に記載のプログラム。

(付記 13)

コンピュータに、

ユーザから、閲覧を希望する Web ページのリクエストを受け付け、受け付けたリクエストをユーザリクエスト記憶部に記憶する処理と、

前記リクエストを行ったユーザの過去の Web ページ訪問履歴を前記リクエスト記憶部から読み出す処理と、

ユーザがある Web ページを閲覧しているとき、次にユーザが閲覧する Web ページが特定の Web ページである確率を表す予測モデルであって、複数のユーザの Web ページ訪問履歴におけるユーザが閲覧した Web ページの順序と、リンクの密接度合いで表現された Web ページ間の関連性との双方を考慮して学習された予測モデルのパラメータをパラメータ記憶部から読み出し、該読み出したパラメータと、前記リクエストを行ったユーザの過去の Web ページ訪問履歴とに基づいて、ユーザが次に訪問する Web ページを予測する処理と、

前記予測した Web ページに基づいて、ユーザにリコメンドする Web ページを決定する処理と、

Web ページを記憶する Web ページ記憶部から、前記ユーザがリクエストした Web ページと、前記リコメンドすると決定した Web ページに関する情報とを読み出し、前記リクエストを行ったユーザに配信する処理とを実行させるプログラム。

(付記 14)

前記 Web ページを配信する処理では、コンピュータは、前記 Web ページ記憶部から、前記リコメンドすると決定した Web ページへのリンク情報を読み出し、前記ユーザがリクエストした Web ページに前記読み出したリンク情報を付加して前記ユーザに配信する、付記 13 に記載のプログラム。

(付記 15)

コンピュータに、前記リクエスト記憶部から読み出された複数のユーザの Web ページ訪問履歴と、前記パラメータ記憶部に記憶された前記予測モデルのパラメータとに基づいて、前記 Web ページ訪問履歴におけるユーザが閲覧した Web ページの順序と、リンクの密接度合いで表現された Web ページ間の関連性との双方を考慮して前記予測モデルを学習し、該学習した予測モデルのパラメータを前記パラメータ記憶部に記憶する処理を更に実行させる、付記 13 又は 14 に記載のプログラム。

(付記 16)

前記予測モデルを学習する処理では、リンク構造を表現する Kernel 法を用いて得られる Web ページ間の密接度合いを表すモデルと、複数のユーザの Web ページ閲覧順序を表す隠れマルコフモデルとを統合して前記予測モデルを構築し、該構築した予測モデルを学習する、付記 15 に記載のプログラム。

(付記 17)

前記予測モデルを学習する処理では、前記ユーザがある Web ページを訪問したとき、前記予測モデルにおける、当該 Web ページとリンク構造が密接する Web ページをユーザが次に閲覧する確率を、当該 Web ページとリンク構造が密接しない Web ページをユーザが次に閲覧する確率よりも高くする、付記 15 又は 16 に記載のプログラム。

【符号の説明】

【0070】

10：リコメンド装置

11：予測モデル学習手段

12：予測手段

13：リコメンド手段

15：パラメータ記憶部

10

20

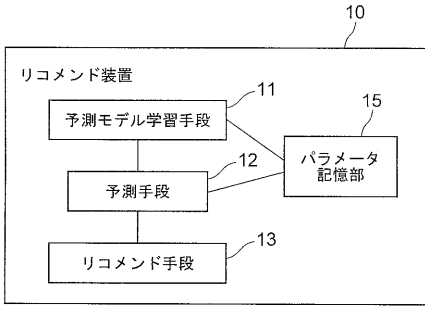
30

40

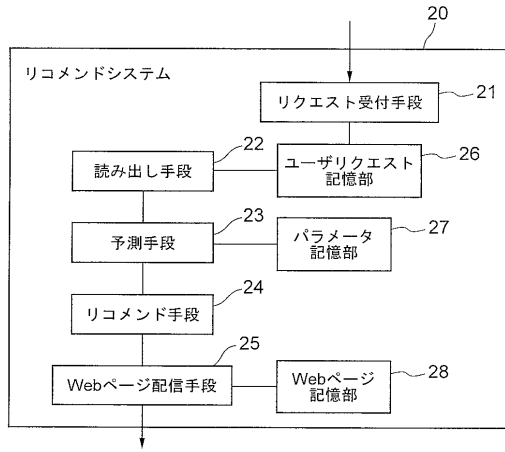
50

20	: リコメンドシステム	
21	: リクエスト受付手段	
22	: 読み出し手段	
23	: 予測手段	
24	: リコメンド手段	
25	: Web ページ配信手段	
26	: ユーザリクエスト記憶部	
27	: パラメータ記憶部	
28	: Web ページ記憶部	
101	: ユーザ端末	10
102	: インターネット	
103	: Web サーバ	
104	: リコメンド装置	
201	: Web ページリクエスト手段	
202	: Web ページ表示手段	
301	: Web ページ配信手段	
302	: ユーザリクエスト記憶部	
303	: Web ページ記憶部	
304	: リコメンドページ選択手段	
306	: リクエスト受付手段	20
307	: 読み出し手段	
308	: 出力手段	
309	: 入力手段	
402	: 予測モデル学習手段	
403	: 予測手段	
404	: 記憶部	
405	: 入力手段	
406	: 出力手段	
407	: リコメンド手段	

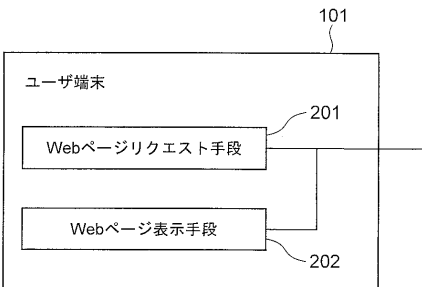
【 図 1 】



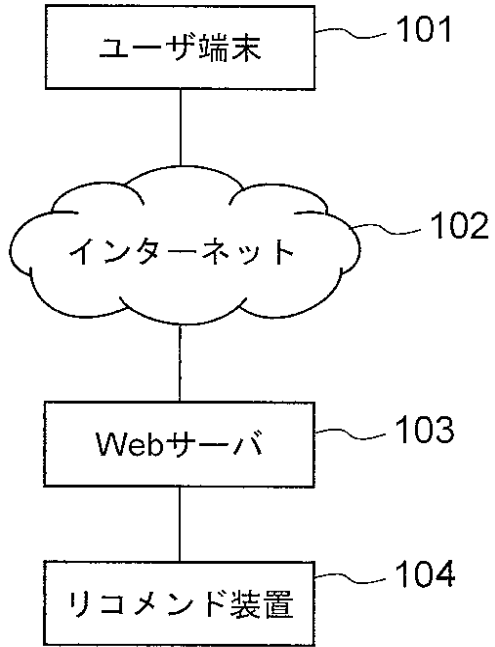
【 図 2 】



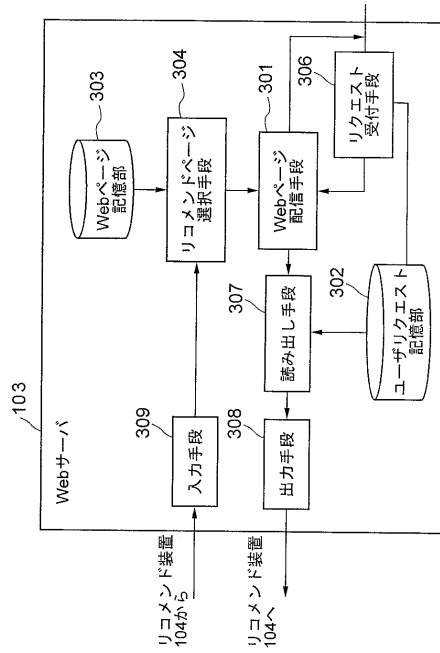
【 図 4 】



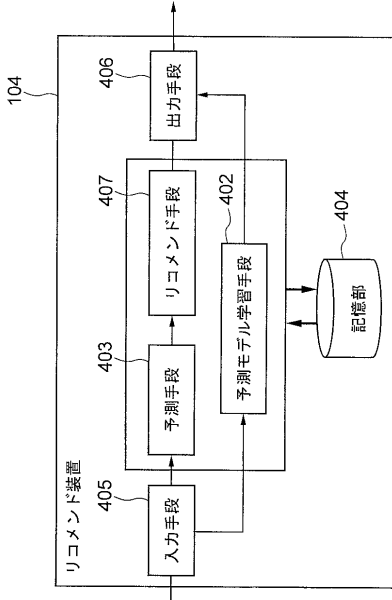
【 図 3 】



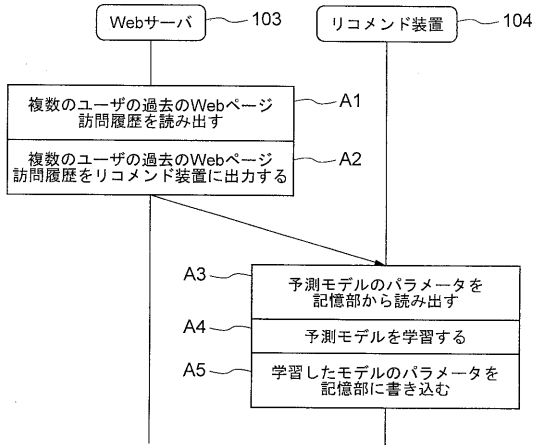
【 図 5 】



【図 6】



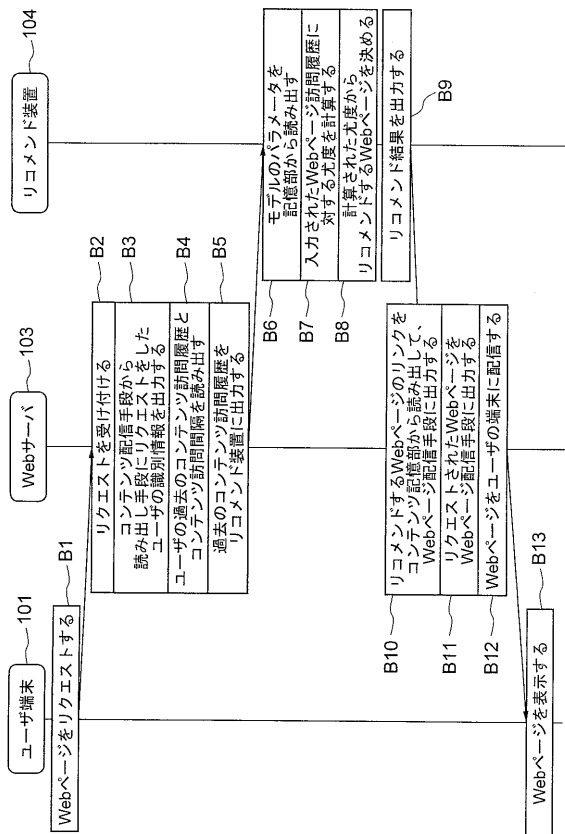
【図 7】



【図 8】

	Webページ訪問履歴
ユーザA	http://sports→http://anime→http://news→…
ユーザB	http://sports→http://anime→http://news→…
・	…
・	…
・	…
ユーザE	http://news→http://news→http://news→…

【図 9】



【図 10】

