

# プライベート仮想ライブラリPVLのためのP2P型推薦機構

上田 真由美, 近藤 育雄, 上島 紳一  
関西大学大学院総合情報学研究科

近年, インターネット上の情報資源を個別化して利用するための枠組みに関する研究が盛んである. 我々は, 利用者の対話的操作をエージェントの信頼度として得点化し, フィードバック利用することによって, システムの個別化を段階的に進めるプライベート仮想ライブラリPVLを提案している.

本稿では, P2P型ネットワークを基盤としたPVLの推薦機構を提案する. PVLは, 単一利用者の利用を目的とし, 利用者は自分にとって重要な情報だけをBookShelfに保存している. 本推薦機構では, 各利用者がBookShelfに保存している情報をPVLコミュニティで共有することにより, 他の利用者が持つ専門的な情報を提供することを可能にする.

本手法により, (1)PVLを繰り返し利用することによって生じた偏りを補填し, 見逃す情報が少なくなる. (2)初めてPVLを利用する場合でも, PVLコミュニティの支援によりよい情報を得ることが可能となる, などの利点を生み出す. 本手法の有効性を検証するために, プロトタイプシステムを構築し, 表示に関する適合率, 格納に関する適合率に注目しシステムの評価実験を行った. 実験結果より, 単一利用者がPVLを用いた場合と比較して, 格納に関する適合率が65.5%, 表示に関する適合率が83.8%となり, ともに大幅に上昇することがわかった.

本手法を用いることにより, P2P型ネットワークを基盤として, PVLコミュニティをボトムアップに構築することが可能である.

キーワード: 個人化技術, メタ検索エンジン, 推薦機構, 協調フィルタリング, P2P, エージェント技術

## P2P-based Recommendation Mechanism for Private Virtual Library PVL

Mayumi UEDA, Ikuro KONDO, Shinichi UESHIMA  
Graduate School of Informatics, Kansai University

Recently, personalization has been one of central research issues for utilizing large volume of WEB contents over the Internet for personal use. We've been proposing a Private Virtual Library PVL, which is a combination of meta search engine and information repositories using BookShelf metaphor. PVL has an embedded mechanism that maps users' interactive operations to scores of access agents to information sources, and users can personalize his/her BookShelf. PVL is designed as a single user system, and every user stores his/her favorite books in BookShelf, and he/she might be interested in them.

In this paper, we introduce P2P-based recommendation mechanism into PVL to share those personal data among PVL users' communities. For this purpose, we build a basic unit as a two-tiered structure of PVL on a peer, which implies peer works as a communication kernel, while PVL works as a principal of each peer. We construct a prototype system using Peer Alliance, as a communication network, and evaluate query results from other PVL users on Peer Alliance by precision factor. The numerical results show that users keep those obtained data from PVL community by 65.5% of probability, which gives higher probability than the one in the case for a single PVL use. Furthermore, we have the advantage that users can construct P2P-based PVL communities in a bottom-up fashion, and that PVL users can join PVL community anytime.

keyword: personalization, meta-search engine, recommendation, collaborative filtering, P2P, agent technology

## 1 はじめに

近年、インターネット上の情報資源を個別化して利用するための枠組みに関する研究が盛んである [1]。我々は、利用者の対話的操作をエージェントの信頼度として得点化し、フィードバック利用することによって、システムの個別化を段階的に進めるプライベート仮想ライブラリ PVL を提案している [2]。

本稿では、PVL とピア・アライアンス・アプリケーション (以下 PA と記す) [4] を基本単位として扱うことによって、PVL の通信手段として P2P 型ネットワークを用いることを可能にする。これにより、PVL を利用するコミュニティにおいて、相互に情報交換をすることができる。一般的に、利用者は自分にとって重要な情報だけを保存する傾向にあり、BookShelf に格納されている情報は個別化されて、専門的な情報であると考えられる。すなわち、P2P 型ネットワークを通信手段とすることで、PVL 利用者は、他の利用者が持つ専門的な情報を共有することが可能になる。

また、単一利用者で PVL を利用する場合に起こりうる、過度の個別化を修正することも同時に可能になる。さらに、個別化が進んでいない利用者に対して、専門化された情報を提供することも可能になる。

ここでは、PVL の P2P 型推薦機構の実現手法と評価について述べる。

2 節では、PVL の基本構成と特徴について述べる。3 節で PVL のための推薦機構の概要について述べ、4 節で、評価実験を行い本システムの有効性を検証する。5 節で関連研究との比較を行い、最後に 6 節で本稿のまとめを行う。

## 2 PVL とは

本節では、PVL の基本構成と特徴について述べる。

### 2.1 基本構成

PVL は利用者の対話的操作をエージェントの信頼度として得点化し、フィードバック利用することによって、システムの個別化を段階的に進める枠組みである。

図 2 に PVL の基本ソフトウェア構成を示す。PVL は個人データを蓄積する利用者プロフィールと BookShelf、情報収集や利用者インタフェースの

個別化を制御するエージェント群、さらに、エージェント群を統合するメディアータから構成される。また、インターネット上の情報資源を統合的に利用するために、各情報資源に対応するアクセスエージェントを備えている。これらのアクセスエージェントは、ラッパーの働きを兼ねており、キーワードで構成されるブール演算を URL に変換している。また、利用者の対話的操作をフィードバック利用するために、利用者プロフィールを用いる。

さらに、メディアータがエージェント群を統合して利用するため、情報の形式に応じて他のアプリケーションを起動することが可能である (図 1 左下)。

### 2.2 特徴

PVL の特徴を以下に示す。

- エージェントの得点付けによる個別化  
PVL では、利用者の対話的操作をアクセスエージェントの信頼度として得点化する。この得点を利用してクライアントでの個別化を可能にする。
- 対話的得点付け  
個別化サービスを受けるために特別な操作は必要ではなく、対話的操作とアクセスエージェントの得点の対応付けにより、利用者は、意識せずに個別化することができる。
- フィードバック型システム  
PVL は利用者の対話的操作をフィードバックして用いるため、システムを繰り返し利用することで、段階的に個別化が進む。
- BookShelf のボトムアップな構築  
利用者の対話的操作によって、BookShelf をクライアントに構築することが可能である。BookShelf は個別化したインターネットのビューに相当する。
- 利用者インタフェースの切り替え  
PVL では、利用者インタフェースの 1 つとして Virtual Book を用いる (図 1 右下)。Virtual Book は本の形状をした 3D ブラウジングインタフェースで、本のページをめくる動作と同じように、ページをめくりながらブラウジングを進めることが可能である [3]。

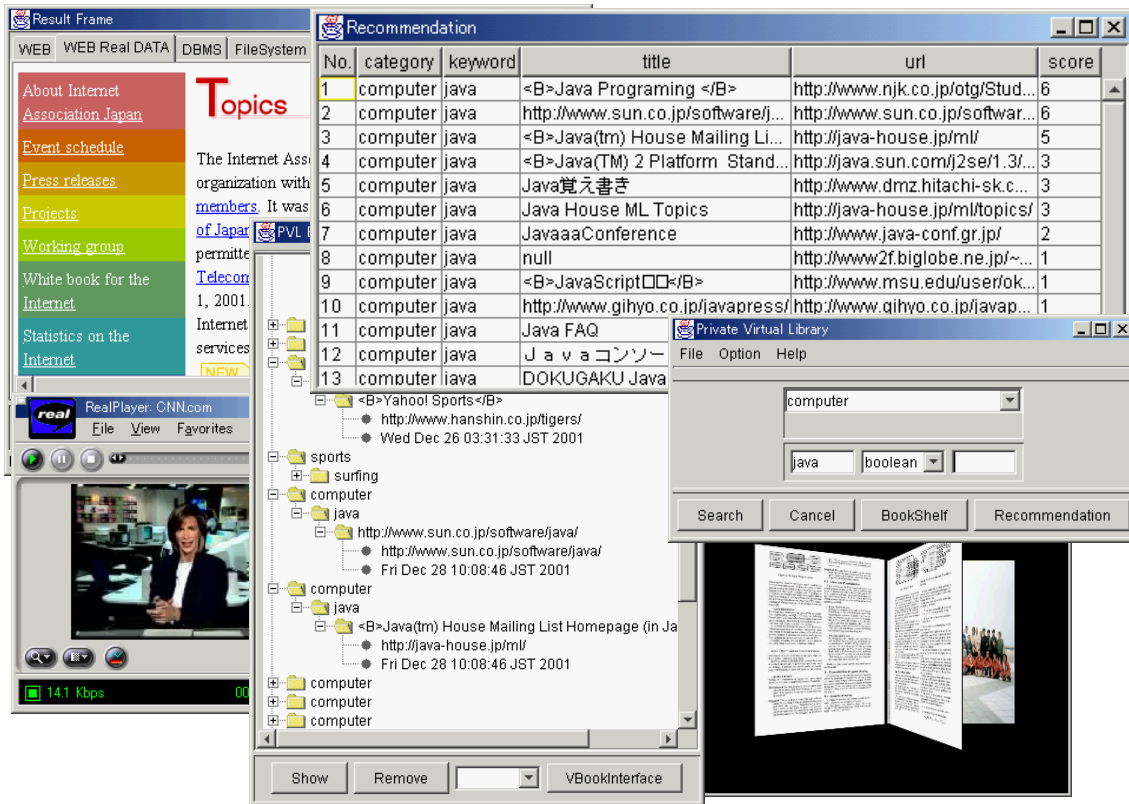


図 1: PVL のプロトタイプシステムの画面例

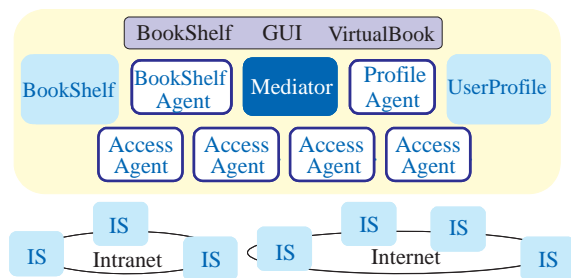


図 2: PVL の基本ソフトウェア構成

### 3 P2P 型推薦機構

本節では、PVL のための推薦機構について述べる。

PVL では、ネットワーク上の情報資源を利用者側で個別化して利用するため、専用のサーバや情報資源を構築する必要がない。推薦機構も、P2P 型ネットワークを通信手段に用いることで、専用のサーバを構築せずに、情報を共有し交換することが可能になる。すなわち、本推薦機構は、ネットワーク上に分散した不特定多数の PVL 利用者が

情報の利用者になり、同時に情報の提供者になり推薦情報を提供する仕組みである。

#### 3.1 PVL とピアの関係

PVL は、単一利用者がインターネット上の情報資源を個別化して利用することを目標とした個人ベースの情報システムである。しかし、ここでは、P2P 型ネットワークを通信手段に用いて PVL コミュニティを形成し、コミュニティベースの情報提供を行う (図 3)。

図 4 に各ピアの構成を示す。

各ピアは PVL と PA で構成される。PVL は図 4 上部に示すように、BookShelf と利用者プロフィールなどの個人データの格納部と、情報収集や利用者インタフェースの個別化を制御するエージェント群によって構成する。

通信手段として用いる PA は図 4 下部に示すように、ピア間接続に用いるコミュニケーション・カーネルと、PA のサービスの一つである PVL Joint サービス、リクエスト文の伝播先を記述したアライアンス表によって構成する。

PA のサービスの一つである PVL Joint サービ

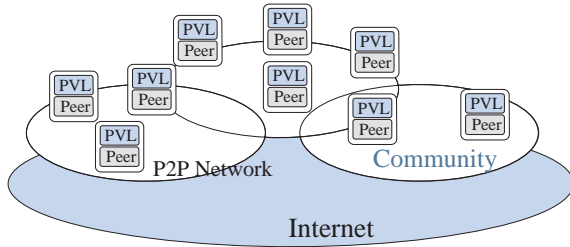


図 3: PVL コミュニティの形成

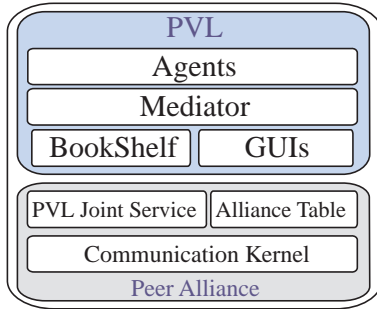


図 4: 各ピアの構成

スには、各ピアに存在する PVL の BookShelf ファイルを検索する PVL コマンドと、操作結果を返し、指定したファイルに書き出す PVLRecomm コマンドがある。

PVL コミュニティはアライアンス表に従って、動的に形成され、利用者は複数のコミュニティに属することが可能である。これにより、目的に応じてコミュニティを使い分け、より専門化された情報を得ることが可能になる。

### 3.2 推薦機構の特徴

2.2 で述べたように、PVL はフィードバック型システムのため、繰り返し利用することによって個別化したサービスを受けることが可能である。しかし、個別化を進めることにより偏りが生じる。この偏りが個人にとって好ましい場合と、好ましくない場合がある。

ここでは、好ましくない場合として、(1) 偏りにより利用されないアクセスエージェントが出てきた結果、利用者にとって必要な情報が提供されない可能性、(2) 他の利用者が個別化を進めた結果、よい情報をたくさん得ている、(3) PVL の利用度が少なく、個別化が進んでいない、などが考えられる。本システムでは、他の PVL 利用者群を PVL コミュニティのメンバーと考え、PVL コミュニティ

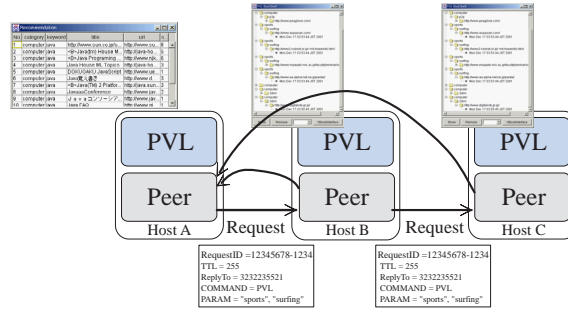


図 5: PVL コミュニティへのリクエストの伝播

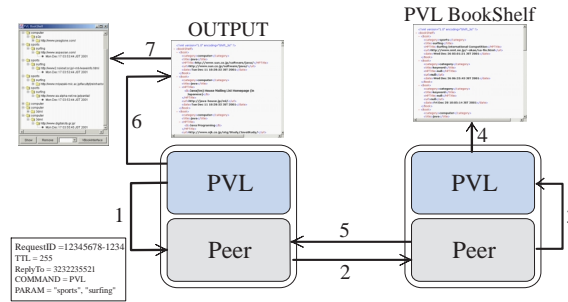


図 6: PVL コミュニティに属する各ピアの動作

による P2P 型推薦機構によって、偏りによって不足した情報の補填を行う。

本システムでは、各利用者が重要と判断し BookShelf に保存した情報は、他の利用者にとっても価値のある情報であると判断し推薦に用いる。

### 3.3 推薦機構の動作

図 5 に PVL コミュニティへのリクエストの伝播を示す。図 6 には、リクエスト伝播時の各ピアの動作を記述する。

ここでは、図 5 と図 6 により、ピア間のリクエストの伝播とピア内の処理について述べる。

PVL から PA へのデータ受け渡しは、ローカルホストの所定のポート番号（ここでは 9153）へ UDP 形式でパケットを送信する。また、PA から PVL への推薦結果の受け渡しは、XML ファイルへの書き出しによって行う。

- リクエストの生成

図 5 左の Host A の PVL の利用者インタフェースにカテゴリとキーワードを入力し、推薦情報を要求すると、ローカルホストの PA に対し PVL コマンドを記述したリクエスト文を渡す (図 6 矢印 1)。

[PVL コマンド記述時のリクエスト文]

```
RequestID = 12345678-1234-1234-1234
TTL = 255
ReplyTo = 3232235521
COMMAND = PVL
PARAM = "sports", "surfing"
```

RequestID フィールドには、リクエストを識別するための番号を記述し、TTL フィールドの値によって、リクエストの到達範囲を限定する。ReplyTo フィールドには、ローカルホストの IP アドレスを 10 進数に変換したものを記述する。COMMAND フィールドには、PVL の BookShelf ファイルを検索する PVL コマンドを、PARAM フィールドには、PVL の利用者インタフェースに入力されたカテゴリとキーワードを記述する。

- リクエストの送信

図 5 の Host A の PA は PVL からリクエストを受け取ると、アライアンス表に従って、Host B の PA にリクエストを伝播させる (図 6 矢印 2)。

- コマンドの実行

PVL コマンドを記述したリクエストを受け取った Host B の PA は、Host B にある PVL の BookShelf を検索し、“PARAM” フィールドに記述されたカテゴリ・キーワードと一致した Book の抽出を行う (図 6 矢印 3, 4)。

- 応答リクエストの生成と送信

[PVLRecomm コマンド記述時のリクエスト文]

```
RequestID = 12345678-1234-1234-1234
TTL = 255
ReplyTo = 3232235526
COMMAND = PVLRecomm
PARAM = <Book> <category>sports
</category> <KW>surfing</KW>
<title>NSA Homepage</title> <url>
http://www.surf-nsa.net</url> <D>
Wed Dec 26 05:01:21 JST 2001</D>
</Book> <Book> ... </Book>
```

該当する Book 情報の抽出後、PVLRecomm コマンドを記述したリクエスト文を生成し、PVL コマンドが記入されたリクエスト文の“ReplyTo”フィールドに記述されていた IP アドレスが示すピア (Host A) へ投げ返す (図 6

矢印 5)。リクエストの識別のために、受け取ったリクエスト文に含まれた RequestID を記述する。“PARAM”フィールドには、抽出した情報、“COMMAND”フィールドには PVL-Recomm を記述する。

- リクエストの伝播

Host B の PA はアライアンス表に従って、リクエストを Host C の PA に伝播し、同様の操作を行う。

- コマンドの実行

推薦要求を発信したピア (Host A) は、PVL-Recomm コマンドが記入されたリクエスト文を受け取ると、“PARAM”フィールドに記述された値を推薦情報として OUTPUT ファイルに書き出す (図 6 矢印 6)。

- 利用者への提示

PVL は OUTPUT ファイルに書き出された情報を、PVL コミュニティ内で評価の高い順に並べ、推薦情報フレーム (図 1 右上) に表示し利用者に提供する (図 6 矢印 7)。

ただし、個人の嗜好が他の利用者に知られることがないように、推薦情報フレームには、どの利用者の BookShelf から提供された情報かは表示しない。

PVL コミュニティから提供された情報の中には、複数の利用者が重複して保存している情報が存在する。複数の利用者が同じ情報を保存していた場合、重複回数を数え、その情報の得点として推薦情報フレームに表示する (右端)。推薦情報の重複回数を得点として表示することで、PVL コミュニティの中で評価の高い情報を見つけることが容易になり、利用者が有効な情報を見つけ出す手助けとなる。

## 4 評価実験

本節では、プロトタイプシステムを用いて、評価実験を行い、本手法の有効性について述べる。

### 4.1 準備

P2P 型ネットワーク上に 3 で述べた PVL のための推薦機構を実装した。

インターネット上に存在する複数のポータルサイトに対応するアクセスエージェントや、利用者プロ

ファイルなどPVL独自の情報資源に対するエージェント群を実装し、メディアータによって統合している。本プロトタイプシステムでは、Yahoo! JAPANとインフォシーク、Excite エキサイト、Lycos Japanに対応するアクセスエージェントを準備した。また、BookShelf ファイルは自動的にXML形式でタグ付けしておく。

PAのサービスの一つとして、受け取ったリクエスト文をもとにBookShelfファイルの検索を行うPVLサービスを実装した。

#### 4.2 実験手順

図1にプロトタイプシステムの画面例を示す。

利用者がログイン画面にユーザー名を入力すると、利用者プロフィールを参照し、図1右中のオペレーションフレームを開く。ここでカテゴリを選択しキーワードを入力後、検索要求を送ると、各アクセスエージェントはインターネット上に対して検索を行い、利用者プロフィールに記述された得点に従って、表示量や順序の個別化が行われ、利用者インタフェースに返される。まず、一次情報としてURLやWebサイトのタイトルが提供される。一次情報の中から興味のある情報をクリックすると実データを表示する。一次情報と実データは結果表示フレーム(図1左)のタブを切り替えることで扱う。実データを見て、興味があるデータとしてSAVEボタンを押すと、BookShelfに保存される。保存した情報は、BookShelfビューワーを用いて図1中下のツリー形式で利用者に提供する。

また、オペレーションフレームから検索要求を送らずに、推薦要求を送ることも可能である。カテゴリを選択し、キーワードを入力後、推薦要求を送ると、PAはアライアンスをたどってPVLからのリクエストを伝播する。PVLコミュニティから返された推薦情報は、XML形式でタグ付けされたOUTPUTファイルに書き出される。推薦情報は図1右上の形式で利用者に提供される。利用者は、この推薦情報フレームの右端の得点を見て、PVLコミュニティで評価の高い情報を知ることができる。ここでも、興味のある情報をクリックすると、結果表示フレームに実データが表示される。

#### 4.3 評価パラメータ

上で述べたプロトタイプシステムを用いて研究室の学生10名にPVLを実際に使用してもらい、以下の3種類の実験を行った。実験1,2では、PVL

のエージェントの得点付け手法の有効性を確認し、実験3でP2P型推薦機構の有効性について述べる。

1. PVL利用時の格納に関する適合率 の推移
2. PVL利用時の表示に関する適合率 の推移
3. PVLコミュニティにより推薦された情報に関する適合率 , の推移

実験1,2は単一利用者がPVLを繰り返し利用し、個別化を進めていく場合の実験で、エージェントの得点付けの有効性の検証に用いる。また、実験3はP2P型ネットワークを用いた、PVLコミュニティによる推薦の有効性を示すのに用いる。

実験では、カテゴリをあらかじめ“computer”に設定し、キーワードはPVLの利用を進めながら、自由に入力してもらった。実験1,2ではPVLと検索エンジン(インフォシークとLycos Japan)を使って同じキーワードにより繰り返し検索を行った。ただし、PVL利用時の検索に用いたキーワードを第1回目の検索から覚えておき、検索エンジン(インフォシーク・Lycos Japan)利用時の検索では、同じキーワードを用いた。

検索結果の中から、独自の判断で、要・不要を判断し、表示する実データの決定と、BookShelfに保存するデータの決定を行った。

評価に用いたパラメータは、(1)検索結果の数(A)、(2)一次情報の中から実際に選択し実データを見た数(B)、(3)BookShelfに格納した数(C)、(4)各操作開始時のアクセスエージェントの得点などで、表示率に関する適合率 と、格納率に関する適合率 について調査した。適合率 , は、それぞれ次式とする。

$$\text{適合率} = \frac{\text{BookShelfに格納した数}(C)}{\text{検索結果の数}(A)} \quad (1)$$

$$\text{適合率} = \frac{\text{実データを表示した数}(B)}{\text{検索結果の数}(A)} \quad (2)$$

#### 4.4 実験結果

本実験により得た結果を示す。

10人の利用者の格納に関する適合率 の平均値を図7に示す。図8では、表示に関する適合率 の平均値、図9では、推薦機構利用時の適合率 , の動きを示す。

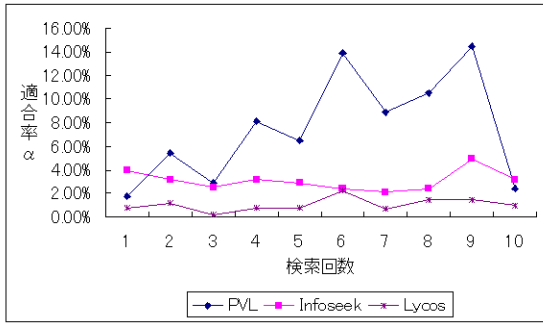


図 7: 格納に関する適合率  $\alpha$  の推移

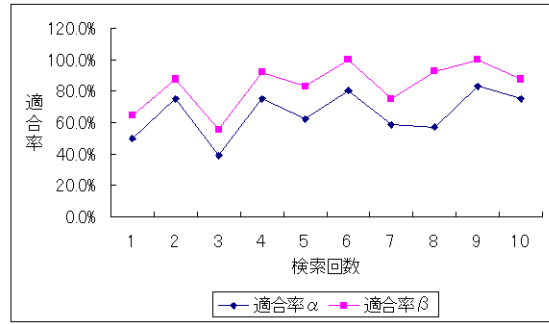


図 9: 推薦機構利用時の適合率  $\alpha$  ,  $\beta$  の推移

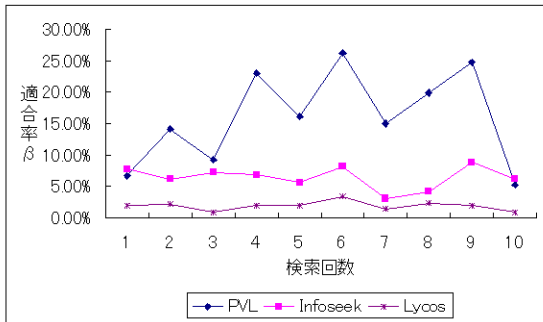


図 8: 表示に関する適合率  $\beta$  の推移

表 1: 各手法における適合率  $\alpha$  ,  $\beta$  の比較

	適合率 $\alpha$	適合率 $\beta$
PVL(単一利用)	7.47 %	16.02 %
PVL(推薦利用)	65.5 %	83.8 %
Infoseek	3.08 %	6.40 %
Lycos Japan	1.02 %	2.07 %

がわかる。

## 5 関連研究

インターネット上の情報量の増加に伴い、推薦システムが重要視されている。これらの多くは、サーバーで一括して収集した情報を用いて推薦を行う推薦システムである。これに対し、本システムは P2P 型通信機構を用いて、不特定多数の利用者間で推薦を行う。不特定多数の PVL 利用者と情報交換を行うために、動的に PVL 利用者内でコミュニティを形成することが可能である。

また、各 PVL におけるエージェントの信頼度を交換することで、くちこみ的に情報の入手先を探索する機構も同様に実現可能である。

本システムの推薦機構による情報収集機構は、内容フィルタリング方式ではなく、協調フィルタリング方式に相当する。この点に関して、不特定多数の利用者を対象とした方式となっている。

[5] では、偵察型エージェントシステムとして、Letizia と PowerScout について述べている。Letizia は、利用者の閲覧行動から好みをリアルタイムに解析し、利用者の好みにあった情報を推薦するシステムである。また、PowerScout は、利用

表 1 は、従来型 PVL、既存の検索エンジン、本推薦機構に関する、適合率  $\alpha$  ,  $\beta$  の値を示す。(値はそれぞれ 10 回検索を行った後の平均値を用いた。)

また、図 7, 8 では、 $\alpha$  が本システム利用時、 $\beta$  が各々検索エンジン (インフォseek, Lycos Japan) 単独利用時の適合率の推移を表す。

図 9 では、 $\alpha$  が適合率  $\alpha$  ,  $\beta$  が適合率  $\beta$  の推移を表す。

図 7, 8 により、PVL では検索回数の増加に従って、適合率  $\alpha$  ,  $\beta$  が段階的に増加していることがわかる。これにより、エージェントの得点付け手法の有効性が示される。また、図 9 から、本推薦機構は利用回数に関わらず、適合率  $\alpha$  ,  $\beta$  とともに高い数値を保っていることがわかる。従って、PVL の利用度が少なく個別化が進んでいない利用者に対しても、有効な情報提供が可能となる。

また、表 1 により、本推薦機構を利用すると従来の PVL 利用と比較して、適合率  $\alpha$  が約 8.7 倍上昇し、適合率  $\beta$  の場合は約 5 倍上昇していること



者が現在閲覧しているページからキーワードを抽出し、検索エンジンによって推薦ページを収集する。PVLのアクセスエージェントは、この偵察型エージェントに相当する。

Singhらは、P2Pパラダイムにおいてピア間の相互信頼度と各ピアの特性ベクトルを定義し、類似性をcosine尺度により評価し、ピア集合全体の動きを推定する研究を行っている[6]。我々は、実際にピアから得た情報を推薦情報として利用し、適合率を評価している。

Blink.com[7]では、個人のブックマークをBlink(サーバ)に登録することによって、インターネット上の他の利用者とブックマークを共有することを可能にする。また、登録されたブックマークはカテゴリ毎に分類し、他の利用者からの評価を同時に提供するため、多くの利用者に評価された情報、すなわち厳選された情報を容易に入手することが可能である。[8]も個人のブックマークを、インターネット上の他の利用者と共有することを可能にしている。ここでは、コンテンツに基づいて自動的にブックマークを分類し、リンク先のページが削除されている場合、自動的にブックマークを削除する。さらに、利用頻度の高いページを自動的にブックマークに追加するため、プロキシによって、利用者の行動を監視している。

## 6 まとめ

本稿では、P2Pネットワークを基盤としたPVLの推薦機構を提案した。また、プロトタイプシステムを用いて、本推薦機構の有効性を示すため評価実験を行った。

PVLは単一利用者の利用を目的としているが、本推薦機構は、各利用者が自分にとって重要であると判断しBookShelfに保存している情報をもとに、PVLコミュニティで共有することによって、他の利用者が持つ専門的な情報を提供することが可能になった。PVLコミュニティによる推薦は、協調フィルタリング方式と同様の情報提供を可能にする。

また、プロトタイプシステムを用いた評価実験により、本推薦機構を利用することで、格納に関する適合率が65.5%、表示に関する適合率が83.3%となり、単一利用者がPVLを用いた場合と比較して、大幅に上昇することを確認した。

今後の課題として以下の点が考えられる。(1)利

用者プロフィールや、P2P型ネットワークを用いるため、セキュリティ問題に関する検討が必要である。(2)ここでは、推薦要求をすべてのPVL利用者(PVLコミュニティ)に対して行ったが、今後は、カテゴリごとに専門化されたコミュニティを動的に形成し、専門化された情報提供を目指す。(3)各ピアに対する信頼度を考慮し、適合率の上昇を目指す。(4)さらに、実験人数を増加させ、リクエストの伝播範囲の違いによる適合率の変化を追跡したい。

## 参考文献

- [1] *Special issue on Personalization*, Communications of the ACM, Vol.43, No.8 (2000)
- [2] Ueda, M. Ueshima, S.: *Client-side Personalization Based on Agent Scoring in Private Virtual Library PVL*, 4th Asia-Pacific Symposium on Information and Telecommunication Technologies (APSITT2001), pp. 31–35 (2001)
- [3] Ueda, M. et al: *Browsing Interface of Private Virtual Library PVL*, 情報処理学会第63回全国大会論文集(3), pp. 9–10(2001)
- [4] Kondo, I. Ueshima, S.: *Peer Alliance: Probing the Internet for World Wide Database - Framework of Resource Sharing via Quilt Queries*, 4th Asia-Pacific Symposium on Information and Telecommunication Technologies (APSITT2001), pp. 269–273 (2001)
- [5] Henry Lieberman et al: *Exploring the Web with RECONNAISSANCE AGENTS*, Communications of the ACM, Vol.44, No.8, pp. 69–75 (2001)
- [6] Munindar P. Singh et al: *COMMUNITY-BASED SERVICE LOCATION*, Communications of the ACM, Vol.44, No.4, pp. 49–54 (2001)
- [7] Blink.com, <http://www.blink.com>
- [8] Wen-Syan Li, et al: *PowerBookmarks: A System for Personalizable Web Information Organization, Sharing, and Management*, Proceedings of the SIGMOD '99, pp. 565–567 (1999)