

潜在的な嗜好に基づくユーザ誘導と コンテンツ流通システムの安定化 に関する研究開発 (082107003)

研究代表者：大阪府立大学 本多 克宏
研究分担者：大阪経済法科大学 呉 志賢
大阪府立大学 青木 真吾
大阪府立大学 野津 亮

1

研究開発の目的

高度コンテンツ創造流通のための基盤技術確立

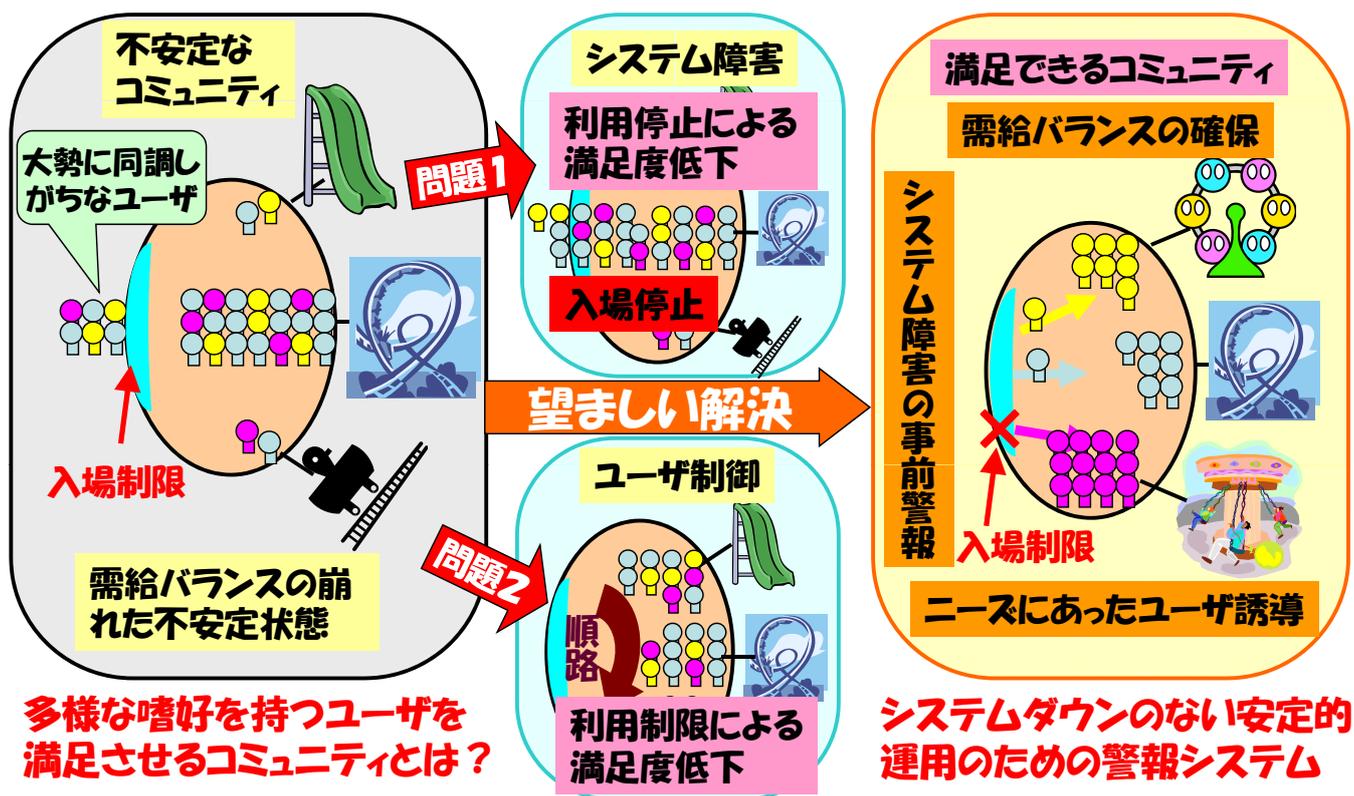
- ユーザの多様なニーズに合致したコンテンツの容易な検索と提示を可能にする
- ユーザ誘導に基づく一極集中の回避によるコンテンツサーバの安定運用を支える

ユーザのニーズに合致した誘導とその評価が必要

1. 誰もが自在にコンテンツを探し出せる環境
2. 情報の信頼を確保しつつ使える環境

2

研究開発の目的のイメージ



本発表ではユーザ誘導にかかる成果を紹介

ユーザの誘導に関する研究背景

- 情報氾濫の昨今、ユーザをいかにして嗜好に合致したコンテンツと結び付けるか、が大きな問題（高度コンテンツ創造流通への課題）。
- ヒトごとに異なる嗜好に合致した「**個人好みの推薦**」を行う協調フィルタリング。
⇒通販サイト・アマゾンなどで実用化。
- 情報コンテンツが多種・大量・多様化する中で、「個人好みの推薦」機構のより広範な普及が必要かつ期待されている。

情報フィルタリングとe-commerce

•通販サイト・アマゾンのおすすめモデル



Amazon プライム
この商品を土曜日(7月12日)までに受け取りたい場合は、16時間 26分以内にご注文ください。
Amazonプライムへの会員登録はお済みですか? ご登録はこちら
お急ぎ便送料無料で今すぐ注文する

ドヴォルザーク交響曲第9番「新世界より」
～カラヤンヘルベルトフォン(アーティスト、指揮)、ドヴォルザーク(作曲)、スズキナ(作曲)、ベルリンフィルハーモニー管弦楽団(演奏)
★★★★☆ (3) 1,500 (税込)
価格: ¥ 1,500 (税込) この商品は1500円以上国内配送料無料を利用して配送されます。詳細
ポイント: 15pt (1%) 詳細はこちら
在庫状況(詳しくはこちら): 在庫あり。この商品は、Amazon.co.jp が販売、発送します。キフ包装を利用できます。

2008/7/12 土曜日 にお届けします! 今から16時間と26分以内にお申し込み、「お急ぎ便」オプション(有料)を選択して注文を確定されたご注文が対象です。詳しくはこちら
一部アイテムが不具合商品の場合、上記の日付が適用されない場合があります。配送オプションを選択する画面、もしくは注文確定するときに必ず配達予定日をご確認ください。

6点の新品/中古商品を見る: ¥ 1,349より
マーケットプレイスからの新品中古品購入は簡単&安全! Amazonポイント、Amazon.co.jp が販売する商品のみご利用、獲得できます。
この商品のURLをメールで共有する

あわせて買いたい
このCDとホルスト組曲「惑星」～ベルリンフィルハーモニー管弦楽団 カラヤン(ヘルベルトフォン)をあわせて買う
参考価格合計: ¥ 3,300
一括注文: ¥ 3,166
ポイント合計: 31pt (1%)
同方ともカートに入れる

この商品を買った人はこんな商品も買っています

ホルスト組曲「惑星」～ベルリンフィルハーモニー	ベートーヴェン交響曲第9番「合唱」～ベルリンフィルハーモニー	モーツァルト「リアーナ」[ソング] 近江トビネ～ベルリンフィルハーモニー	ベートーヴェン交響曲第5番～カラヤン(ヘルベルトフォン)	ドヴォルザーク交響曲第9番～ハンスライオン(ナード)	ベートーヴェン交響曲第5番～ハンスライオン(ナード)
★★★★☆ (3) ¥ 1,666	★★★★☆ (3) ¥ 1,709	★★★★☆ (1) ¥ 1,500	★★★★☆ (1) ¥ 960	★★★★☆ (10) ¥ 1,596	★★★★☆ (3) ¥ 960

ユーザ満足度の向上 & 購買機会の向上

協調フィルタリングシステム

- ✓膨大な量の情報から利用者にあった情報を推薦する技術
- ✓データ行列中の未評価値, すなわち欠測値を推定する

•評価値行列・・・アイテムに対するユーザの評価値

	囲碁	将棋	園芸	テニス
太郎	5		5	4
花子	2	5		3
雄三		2	1	2
純一	5	1	5	

欠測要素
予測値が大きければ推薦する

- (1) 全ユーザと類似度算出
- (2) 似たユーザを重視した平均化

友達の探索
友達の意見の参照

疑似的なくちコミの再現

従来技術の問題点

既に実用化されているものには、通販サイト・アマゾンによる「おすすめモデル」等があるが、

大量の履歴データを保持する必要

推薦探索に大規模な計算時間が必要

個人履歴の保持による情報セキュリティ

高性能・高コストな専用サーバが必要

という問題があり、安価で利便性の高いサービスとして普及するまでには至っていない。

7

共クラスター構造推定による 協調フィルタリングモデルの開発

- ユーザとアイテムのグループ化による推薦

	item1	item2	item3	item4	item5	item6
user a	○	○				○
user b	○	○			○	
user c		○	○			
user d		○	○	○		
user e					○	○
user f		○		○		

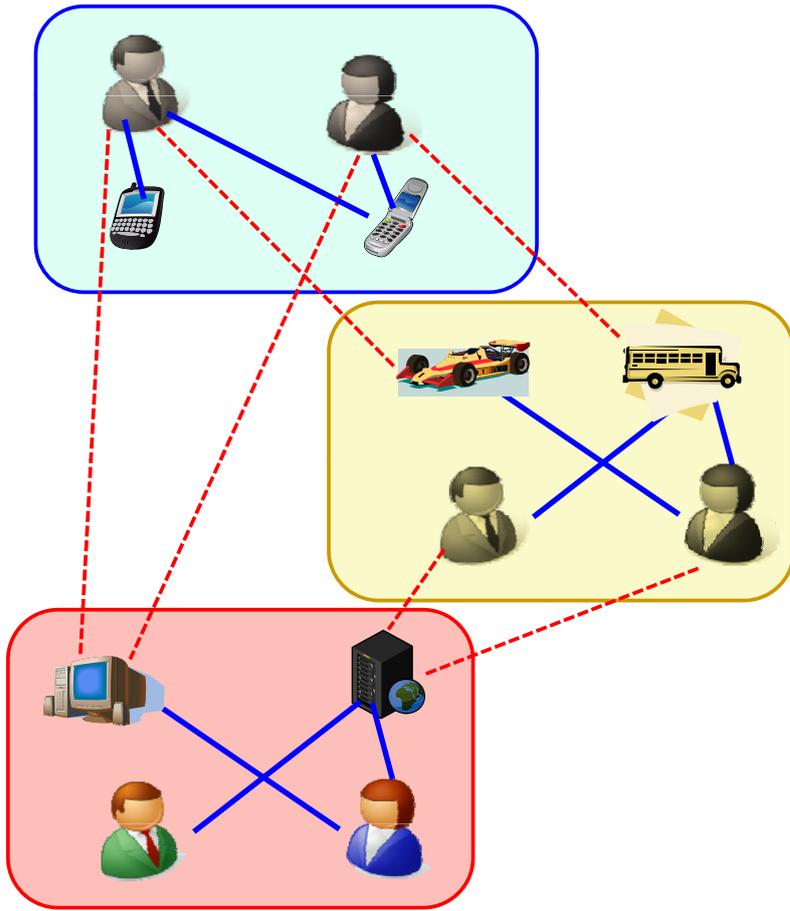
	item1	item5	item6	item3	item4	item2
user a	○	○	○			○
user b	○	○				○
user e		○	○			
user c				○		○
user d				○	○	○
user f					○	○

ユーザ・
アイテムの
対を推定

共クラスター構造として情報を圧縮・利用

8

認知科学の知見を取り入れたクラスタリングモデル



- いくつかのグループに分割することで、グラフ構造(自己の認知体系)が均衡化
- 購買履歴のような0-1型のデータ(0が負の関係と、関係性未知の両方を含む)の取り扱いに有利

(例)購買の有無

$$s_{ij} = \begin{cases} 1: \text{正の関係} \\ 0: \text{負の関係or未知} \end{cases}$$

9

新規技術における計算アルゴリズム

(1) 関係性行列からの固有値の算出

関係性行列 $S = \{s_{ij}\}$

	u-1	...	u-n	i-1	...	i-m
u-1	0	...	0	1	...	1
:	:	...	:	:	...	:
u-n	0	...	0	1	...	0
i-1	1	...	1	0	...	0
:	:	...	:	:	...	:
i-m	1	...	0	0	...	0

→
最大固有値と固有ベクトルの算出
(べき乗法, ヤコビ法など)

固有ベクトル $w_k = \{w_{ki}\}$

	w_{ki}
u-1	0.8
:	:
u-n	0.2
i-1	0.7
:	:
i-m	0.3

閾値より値が大きいユーザとアイテムを第k番目の近傍としてグループ化

(2) 関係性行列からの抽出済みの近傍を削除

グループに割り当て済みのユーザを排除するため、関係性行列の対角要素を変形する。

$$s_{ii} = \begin{cases} -\frac{1}{k-1} \sum_{t=1}^{k-1} \beta_t w_{it}^2 & i \text{ がユーザの場合} \\ 0 & i \text{ がアイテムの場合} \end{cases}$$

重み β は、グループの塊度合いから決める

(1)へ戻り、次の近傍(k+1番)を探索

	u-1	...	u-n	i-1	...	i-m
u-1	s_{ii}	...	0	1	...	1
:	:	...	:	:	...	:
u-n	0	...	s_{ii}	1	...	0
i-1	1	...	1	s_{ii}	...	0
:	:	...	:	:	...	:
i-m	1	...	0	0	...	s_{ii}

10

推薦性能についての検証実験結果

- 購買履歴データ（モニタ調査による実世界のデータ）

世帯数：996世帯

製品数：18種類（各製品を所有しているか否か）

半分をテスト世帯とし、購買の有無を予測

- 推薦性能（ROC感度）の比較

ROC感度：推薦性能の総合的指標。大きいほど良い。

製品	ピアノ	パソコン	ワープロ	VD	オーブン	珈琲メーカー
従来法	0.577	0.580	0.528	0.600	0.569	0.515
新技術	0.644	0.647	0.544	0.547	0.575	0.587
改善率(%)	11.6	11.6	3.0	-8.8	1.1	14.0

最大14%（平均5.4%）の精度改善 ⇒ 0-1型履歴に強い

11

新技術の利点

- データ保持コストの軽減

グループ構造の情報のみを保持して推薦

⇒ メモリの所要量が少ない

- 計算量の軽減

所属グループを探索するのみなので、計算負荷が小

⇒ 計算CPUにかかる負担が少ない

- 情報の匿名化によるセキュリティ向上

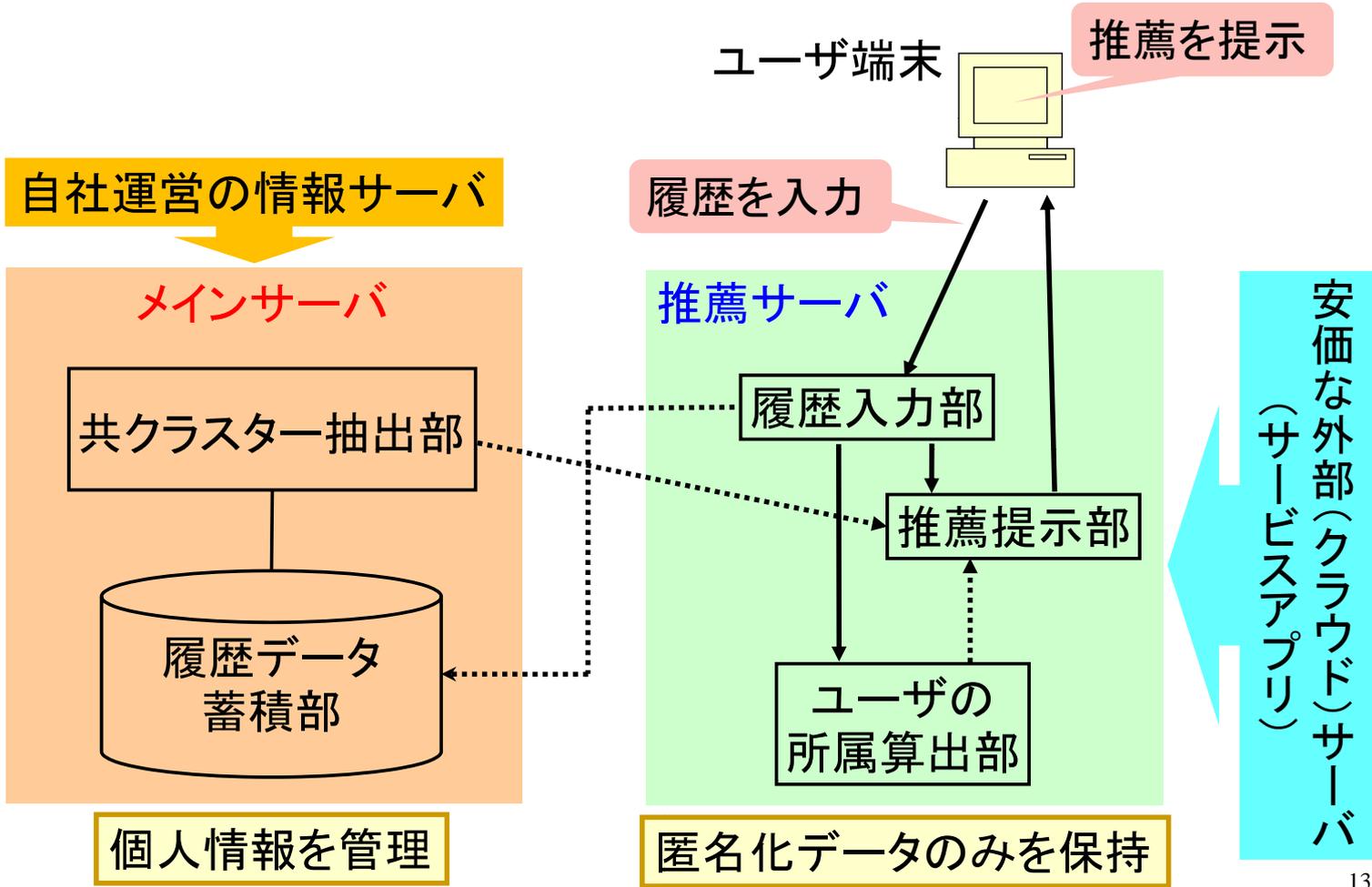
多数のユーザをグループ化 ⇒ 匿名性の高い情報管理

⇒ 履歴情報から個人が特定される危険を回避

クラウドサーバなどの外部ソースの利用に最適

12

開発システムの構成図(例)



デモシステム・旅行先おすすめ「ワムトラ」の紹介

旅先案内 全国版 (やり直しはブラウザの更新ボタンを押してください)

地域ごとに、行ったことのある観光地を採点してください

(1) 観光地を評価

九州

中国
四国

近畿

東海

北陸

関東

前の画面に戻る

5点満点 (1:つまらない~5:お気に入り) で採点してください

茨城	(1) 霞ヶ浦	<input type="radio"/>	<input checked="" type="radio"/>	行ったことがない				
	(2) 水戸・偕楽園	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	行ったことがない
	(3) 花めぐり	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	行ったことがない
栃木	(1) 中禅寺湖	<input type="radio"/>	<input checked="" type="radio"/>	行ったことがない				
	(2) 日光・東照宮	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	行ったことがない
	(3) 鬼怒川温泉	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	行ったことがない
群馬	(1) 妙義高原	<input type="radio"/>	<input checked="" type="radio"/>	行ったことがない				
	(2) 水上	<input type="radio"/>	<input checked="" type="radio"/>	行ったことがない				
	(3) 碓氷峠	<input type="radio"/>	<input checked="" type="radio"/>	行ったことがない				
埼玉	(1) 大宮	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	行ったことがない
	(2) 川越	<input type="radio"/>	<input checked="" type="radio"/>	行ったことがない				
	(3) 上野	<input type="radio"/>	<input checked="" type="radio"/>	行ったことがない				
千葉	(1) 九十九里	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	行ったことがない
	(2) 佐倉・城下町	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	行ったことがない
	(3) 東京ディズニーランド	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	行ったことがない

(2) おすすめ提示

<http://www.cs.osakafu-u.ac.jp/hi/honda/tool/>
Google や Yahoo にて「ワムトラ」と検索ください

おすすめシステム構築ツール「ワムナビ」の紹介

自分のwebページ上に、誰でも簡単に、おすすめシステムを構築できるJAVAアプレット

(1) 履歴データを準備

ユーザID	寿司	天ぷら	そば	味噌汁
1	5	4		
2		3		5
3	2		5	
4	4		2	3
5		5	5	1
6	1	4	5	
7		1		2
8	3		4	
9	1	3		5
10	5		5	
11		2		4
12	4		1	3

ツールをサーバにコピー

食べたことのある料理を評価してください。

和食 洋食 中華

(2) 自分のwebページにおすすめシステム!

おすすめの表示へ 結果のサンプル表示へ

おすすめスポット・トップ5

- 第1位 ハンバーグ
- 第2位 ソフトクリーム
- 第3位 クレープ
- 第4位 ケーキ
- 第5位 シュークリーム

和食

前の画面に戻る

「お気に入り、普通、嫌い、食べたことがない」で評価

寿司	<input type="radio"/> お気に入り	<input type="radio"/> 普通
天ぷら	<input type="radio"/> お気に入り	<input type="radio"/> 普通
そば	<input type="radio"/> お気に入り	<input type="radio"/> 普通
味噌汁	<input type="radio"/> お気に入り	<input type="radio"/> 普通

<http://www.cs.osakafu-u.ac.jp/hi/honda/tool/>
「おすすめシステム ワムナビ」と検索ください

15

想定される活用の形態

● ITインフラ(サービス)提供業

インターネットプロバイダ, クラウドサービス会社
⇒ハードウェアの提供にソリューションを付加

提供サービスの付加価値の向上



● コンテンツ発信・提供業(ITインフラ利用者)

ケーブルTV, オンデマンド配信, 非営利ユーザ(観光協会)
⇒顧客の利便性に貢献, コンテンツ配信機会の増加

提供コンテンツの“質”と利便性の向上



16

- 各種のサービスコンテンツへの組み込み

- ネット通販での商品推薦

(**広告提示**)

販売機会の向上



あなたへの
おすすめは
...

- 特許検索の支援インタフェース

(関連性の**自動検索**)

ユーザ利便性の向上



- 携帯型ナビゲーションシステム

(掘出し物への**自動誘導**)

ユーザ満足度の向上



17

まとめ

- 高度コンテンツ創造流通を支える基盤技術として、**誰もが自在に好みのコンテンツを探索**できる協調フィルタリングのアルゴリズムを開発した。
- 情報セキュリティ(情報の匿名化)のレベルを**常に保証する機能**の開発が課題である。

k-匿名性の保証

少なくともk人の履歴情報を単一のグループ情報にまとめる
⇒履歴から個人が特定不能



18