エッジ型情報収集支援における 個人クエリ相互利用の 利用者視野拡大効果 - 実験報告 -

金沢工業大学情報工学科 小林 暉 (3年)
佐治寿一 (3年)
風尾勇侑 (3年)
芦川拓実 (3年)
指導教官 金道敏樹

はじめに

他人のクエリを利用することで、**情報アクセスへの利便性**と**視野拡大**を並行できるかもしれない。

目次

- 構想について
- 作ったシステム
- 使用した技術
- クエリ利用とは
- 実験結果
- まとめ

構想の前提

- 各個人にカスタマイズされた個人情報フィルタ
その中にある個人クエリは、ユーザの個人視野を表す。

構想-1

- 1. ユーザの興味(ユーザの視野)を広げるには、他者のクエリを利用すると便利なのではないか。
- 2. あるコンテンツに対して、興味の似通った他者のクエリを利用した検索が行えれば、ユーザの興味範囲が広がる(かも)。

他者のプライバシーにどう配慮するか

構想- 2 プライバシーを保つために

- プライバシーをユーザがコントロールするには、エッジ型の方が良い。
 - 今回はサーバクライアント（≠エッジ型）で実装
- 個人クエリをユーザ端末に置ける技術で実現しよう
 - 大きな辞書を使わない情報フィルタINSOPを活用

作ったシステム

おもしろ？映画検索支援サービス

まず、いくつかの映画にチェックをつけてください

名前：

学習

シャッフル

CheckBox	Image	Title	Abst
☐ ○ ☐ ×		ID:197 「アンティーク西洋骨董洋菓子店」	...
☐ ○ ☐ ×		ID:215 「ウタモノガタリ CINEMAFIGHTERSproject 「ファンキー」	...
☐ ○ ☐ ×		ID:20 「パニックインベガスエイリアン襲来」	...
☐ ○ ☐ ×		ID:132 「フローズンライター」	...

作ったシステム

名前 :

学習

(済) 表示

(済) 非表示

シャッフル

CheckBox	Image	Title	Abst	Score
☐ ○ ☐ × (済)		ID : 28 「アンドロメダ…」		117.6
☐ ○ ☐ × (済)		ID : 24 「ディープインセクト」		85.91
☐ ○ ☐ × (済)		ID : 266 「ウォーキングZ」		52.88
☐ ○ ☐ ×		ID:1 「プラネットオブロボット」		15.88

スコアの大きい順に並べている

システム構成

CGIプログラムにアクセスする

実行結果をユーザに返す

CGIプログラム

Apache
サーバ

プログラム実行

実行結果 (スコア、
HTMLファイル)

ユーザが興味の
有無をチェック
する

使用した技術 1 - プロファイリング技術 : INSOP

- INSOP

- ユーザプロファイルを生成する個人情報フィルタ

使用した技術 1 - INSOP

使用した技術 1 - INSOP

ユーザごとに小さな辞書を自動生成し、「価値」のあるキーワードを登録

Positive key words: $q(w_i) \gg p$

Negative key words: $q(w_i) \ll p$

Neutral key words: $q(w_i) \doteq p$

興味有り = 正

興味なし = 負

p : ユーザが興味ありとする確率

$q(w_i)$: キーワード w_i が付いた情報をユーザーが興味ありとする確率

個人辞書へは、キーワードが運ぶ個人の興味の情報量の多い

Positive キーワードと **Negative** キーワードが登録される。

使用した技術 1 - INSOP

番号	必要性	信頼性	要約
1	94.7	92.8	OECD関係協会合終了 電子商取引規制で溝 欧州 民間
2	55.7	46.4	世界情報基盤機構の年次総会 共同宣言採択し閉幕 電子商
3	44.1	41.1	総合経営再構築 日立、東芝、三菱電機が変わる(4) =テ
4	43.1	42.0	東芝 米英でデジタルテレビ発売 米国には来春65型プロ
5	39.7	39.0	岐阜県できょう 電子商取引で総合窓口 岐阜統合モール発
6	27.9	26.9	テレコムサービス協会 シンガポールとEC実験 口座振替
7	24.0	23.1	電子商取引で郵政相 消費者団体などと話し合いの場作る
8	22.4	22.1	日本商品取引員協会 電子取引の定義 小委員解を白紙に
9	17.0	16.0	NEC インターネット検索システム開発 文書タイプや数
10	16.6	16.1	サイバーマップ・ジャパン ネット上の地図情報検索 リニ
11	16.0	8.6	G I I Cの98年度年次総会 EC促進へ共同宣言採択
12	15.0	12.5	世界情報基盤機構総会 電子商取引促す共同宣言を発表
13	6.8	6.1	キヤノン デジタル写真をTV上で編集、加工ができるデジ
14	2.2	1.5	日本語ワープロ年末用に新製品出揃う 二極分化が鮮明
15	1.8	1.0	ビジネスTODAY=苦境の半導体 球が突破口 米ボール
16	1.5	1.1	APECの主要議題明らか 貿易自由化 例外分野を削除
17	1.0	0.9	トヨタ自動車 出資「モバイル放送」 本田技研工業も参加
18	1.0	0.8	デジタル放送で “シンボル論争” 過熱 新東京タワー さ

興味あり

興味なし

キーワードが運ぶ「興味度」は統計的
キーワード評価値(SKIC)で計算される。

各情報に対する「興味度」は
SKICの総和で推定できる。

SKICの**総和の大きい順**に情報を並べる。

クエリ利用：プライバシーへの配慮

個人クエリの中身を比較するのではなく、**スコアの類似性**の計算をしているだけ！

直接比較はユーザのプライバシー保護が難しい

使用した技術2 - コンテンツスコアによる類似性の推定

正の相関

負の相関

使用した技術2 - コンテンツスコアによる類似性の推定

正の相関

負の相関

あるユーザ α とユーザ β の各コンテンツに対する**興味の類似性**を、相関により表す。

実験結果 - 1

- 金沢工業大学の学生並びに職員にアプリケーションを利用してもらい、個人クエリを収集した。

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S15	S16	S18	S19	S21
S1	10	-1	2	4	1	2	0	2	1	0	0	4	1	1	0	2	1	2
S2	-1	10	1	2	1	0	1	-1	1	2	-1	-1	0	2	3	-1	0	-2
S3	2	1	10	2	-1	0	-2	0	0	4	-1	-1	3	-1	0	-2	-2	2
S4	4	2	2	10	2	1	1	2	2	1	1	3	1	1	2	2	0	3
S5	1	1	-1	2	10	2	0	0	1	0	3	-1	1	0	1	2	0	1
S6	2	0	0	1	2	10	1	2	2	1	2	0	1	3	2	2	2	3
S7	0	1	-2	1	0	1	10	1	0	-1	1	2	0	2	2	1	2	-1
S8	2	-1	0	2	0	2	1	10	3	1	2	3	0	2	0	3	1	2
S9	1	1	0	2	1	2	0	3	10	3	2	0	0	3	2	1	1	2
S10	0	2	4	1	0	1	-1	1	3	10	1	0	2	0	1	-2	1	2
S11	0	-1	-1	1	3	2	1	2	2	1	10	2	1	1	1	4	3	2
S12	4	-1	-1	3	-1	0	2	3	0	0	2	10	-1	0	-1	3	3	2
S13	1	0	3	1	1	1	0	0	0	2	1	-1	10	0	1	-1	-1	1
S15	1	2	-1	1	0	3	2	2	3	0	1	0	0	10	1	2	1	0
S16	0	3	0	2	1	2	2	0	2	1	1	-1	1	1	10	-1	1	1
S18	2	-1	-2	2	2	2	1	3	1	-2	4	3	-1	2	-1	10	2	1
S19	1	0	-2	0	0	2	2	1	1	1	3	3	-1	1	1	2	10	1
S21	2	-2	2	3	1	3	-1	2	2	2	2	2	1	0	1	1	1	10

相関係数が最も大きい値（黄色）を利用することで、ユーザにとって便利な推薦になる。

※本来の数値を10倍し、整数で表示している

実験結果 - 1

- 金沢工業大学の学生並びに職員にアプリケーションを利用してもらい、個人クエリを収集した。

相関係数が最も大きい値（黄色）を利用することで、ユーザにとって便利な推薦になる。

※本来の数値を10倍し、整数で表示している

実験結果 - 1

- 金沢工業大学の学生並びに職員にアプリケーションを利用してもらい、個人クエリを収集した。

相関係数が最も大きい値（黄色）を利用することで、ユーザにとって便利な推薦になる。

※本来の数値を10倍し、整数で表示している

実験結果 - 2 ユーザ α の推薦結果

他人のスコア利用前

○ ₀ ○ ₀ ×		ID:60「明日に向かって撃て！」
○ ₀ ○ ₀ ×		ID:187「夕陽の群盗」
○ ₀ ○ ₀ ×		ID:184「GiveandGo-ギブアンドゴー-」
○ ₀ ○ ₀ ×		ID:294「オリーブの林をぬけて」
○ ₀ ○ ₀ ×		ID:30「ゴジラ キング・オブ・モンスターズ」
○ ₀ ○ ₀ ×		ID:34「劇場版仮面ライダーエグゼイドトゥルーエンディング」
○ ₀ ○ ₀ ×		ID:109「マードーオブキャット」
○ ₀ ○ ₀ ×		ID:191「ハレルの楽園」
○ ₀ ○ ₀ ×		ID:85「コララインとボタンの魔女」
○ ₀ ○ ₀ ×		ID:24「ディープインセクト」
○ ₀ ○ ₀ ×		ID:283「最低で最高のサリー」
○ ₀ ○ ₀ ×		ID:173「ひろしま」
○ ₀ ○ ₀ ×		ID:297「ファントム・スレッド」
○ ₀ ○ ₀ ×		ID:293「ういらぶ」

ユーザ α のスコアを見ると、ミステリー要素、少しのラブコメ要素を含むコンテンツに興味があることが分かった。

実験結果 - 3 推薦結果の変化

自分のスコアのみ

○ ×		ID:60「明日に向かって撃て！」
○ ×		ID:187「夕陽の群盗」
○ ×		ID:184「GiveandGo-ギブアンドゴー-」
○ ×		ID:294「オリーブの林をぬけて」
○ ×		ID:30「ゴジラ キング・オブ・モンスターズ」
○ ×		ID:34「劇場版仮面ライダーエグゼイドトゥルーエンディング」
○ ×		ID:109「マダーオブキャット」
○ ×		ID:191「ナルトの楽園」
○ ×		ID:85「コララインとボタンの魔女」
○ ×		ID:24「ディーブインセクト」
○ ×		ID:283「最低で最高のサリー」
○ ×		ID:173「ひろしま」
○ ×		ID:297「ファントム・スレッド」
○ ×		ID:293「ういらぶ」

他人のスコア利用後

○ ×		ID:286「セカンドバージン」
○ ×		ID:294「オリーブの林をぬけて」
○ ×		ID:89「きみと、波にのれたら」
○ ×		ID:293「ういらぶ」
○ ×		ID:96「ONE PIECE FILM GOLD」
○ ×		ID:287「おとなの恋はまわり道」
○ ×		ID:290「夏ノ日、君ノ声」
○ ×		ID:217「心が叫びたがってるんだ。(2017)」
○ ×		ID:53「烈車戦隊トッキュウジャー-THEMOVIEギャラクシーラインSOS」
○ ×		ID:87「君の隣を食べたい(2018)」
○ ×		ID:296「想影」
○ ×		ID:125「マスカレードホテル」
○ ×		ID:75「ディリリとパリの時間旅行」
○ ×		ID:78「夏目友人帳うつせみに結ぶ」

実験結果 - 3 推薦結果の変化

自分のスコアのみ

0.0 x		ID:60「明日に向かって撃て！」
0.0 x		ID:187「夕陽の群盗」
0.0 x		ID:184「GiveandGo-ギブアンドゴー-」
0.0 x		ID:294「オリーブの林をぬけて」
0.0 x		ID:30「ゴジラ キング・オブ・モンスターズ」
0.0 x		ID:34「劇場版仮面ライダーエグゼイドトゥルーエンディング」
0.0 x		ID:173「ひろしま」
0.0 x		ID:297「ファントム・スレッド」
0.0 x		ID:293「ういらぶ」

見落としていたカテゴリーの映画が、**上位**にリストアップすることが確認できた。
(例：ういらぶ、オリーブの林をぬけて)

他人のスコア利用後

0.0 x		ID:286「セカンドバージン」
0.0 x		ID:294「オリーブの林をぬけて」
0.0 x		ID:89「きみと、波にのれたら」
0.0 x		ID:293「ういらぶ」
0.0 x		ID:96「ONE PIECE FILM GOLD」
0.0 x		ID:287「おとなの恋はまわり道」
0.0 x		ID:290「夏ノ日、君ノ声」
0.0 x		ID:217「心が叫びたがってるんだ。(2017)」
0.0 x		ID:53「烈車戦隊トッキュウジャー-THEMOVIEギャラクシーラインSOS」
0.0 x		ID:87「君の隣を食べたい(2018)」
0.0 x		ID:296「想影」
0.0 x		ID:125「マスカレードホテル」
0.0 x		ID:75「ディリリとパリの時間旅行」
0.0 x		ID:78「夏目友人帳うつせみに結ぶ」

実験結果 - 3 推薦結果の変化

自分のスコアのみ

○ ○ ×		ID:60「明日に向かって撃て！」
○ ○ ×		ID:187「夕陽の群盗」
○ ○ ×		ID:184「GiveandGo-ギブアンドゴー-」
○ ○ ×		ID:294「オリーブの林をぬけて」
○ ○ ×		ID:30「ゴジラ キング・オブ・モンスターズ」
○ ○ ×		ID:34「劇場版仮面ライダーエグゼイドトゥルーエンディング」
○ ○ ○		ID:293「ういらぶ」
○ ○ ○		ID:173「ひろしま」
○ ○ ×		ID:297「ファントム・スレッド」
○ ○ ×		ID:293「ういらぶ」

見落としていたカテゴリーの映画が、**上位**にリストアップすることが確認できた。
(例：ういらぶ、オリーブの林をぬけて)

いい感じ！

他人のスコア利用後

○ ○ ×		ID:286「セカンドバージン」
○ ○ ×		ID:294「オリーブの林をぬけて」
○ ○ ×		ID:89「きみと、波にのれたら」
○ ○ ×		ID:293「ういらぶ」
○ ○ ×		ID:96「ONE PIECE FILM GOLD」
○ ○ ×		ID:287「おとなの恋はまわり道」
○ ○ ○		ID:290「夏ノ日、君ノ声」
○ ○ ○		ID:217「心が叫びたがってるんだ。(2017)」
○ ○ ○		ID:53「烈車戦隊トッキュウジャー-THEMOVIEギャラクシーラインSOS」
○ ○ ×		ID:87「君の隣を食べたい(2018)」
○ ○ ×		ID:296「想影」
○ ○ ×		ID:125「マスカレードホテル」
○ ○ ×		ID:75「ディリリとパリの時間旅行」
○ ○ ×		ID:78「夏目友人帳うつせみに結ぶ」

実験結果 - 3 推薦結果の変化

自分のスコアのみ

○ ○ ×		ID:60「明日に向かって撃て！」
○ ○ ×		ID:187「夕陽の群盗」
○ ○ ×		ID:184「GiveandGo-ギブアンドゴー-」
○ ○ ×		ID:294「オリーブの林をぬけて」
○ ○ ×		ID:30「ゴジラ キング・オブ・モンスターズ」
○ ○ ×		ID:34「劇場版仮面ライダーエグゼイドトゥルーエンディング」
○ ○ ×		ID:173「ひろしま」

見落としていたカテゴリーの映画が、**上位**にリストアップすることが確認できた。
(例：ういらぶ、オリーブの林をぬけて)

他人のスコア利用後

○ ○ ×		ID:286「セカンドバージン」
○ ○ ×		ID:294「オリーブの林をぬけて」
○ ○ ×		ID:89「きみと、波にのれたら」
○ ○ ×		ID:293「ういらぶ」
○ ○ ×		ID:96「ONE PIECE FILM GOLD」
○ ○ ×		ID:287「おとなの恋はまわり道」
○ ○ ×		ID:290「夏ノ日、君ノ声」
○ ○ ×		ID:217「心が叫びたがってるんだ。(2017)」
○ ○ ×		ID:53「烈車戦隊トッキュウジャー-THEMOVIEギャラクシーラインSOS」
○ ○ ×		ID:87「君の隣を食べたい(2018)」
○ ○ ×		ID:296「想影」
○ ○ ×		ID:125「マスカレードホテル」

ユーザ α の気づきの範囲が大きくなる方向に可能性が見られた

まとめ

ユーザ β のクエリをユーザ α が利用することで、
ユーザ α が見落としていた映画をすくい上げることができた。

新しいコンテンツを目に入れることができた。
これは視野拡大につながるはず

今後、視野拡大への影響を調べたい。
個人クエリの変化で計測できると考えている。

お願い

- 具体的な計算方法については次の芦川が説明しますので、よろしくお願いします。

Fin

ご清聴ありがとうございました