

ユーザの検索行動に基づく嗜好推定を用いた 複数人 Web 検索における意見集約支援システムの検討

本田 博之[†] 岩田 麻佑[†] 原 隆浩[†] 西尾章治郎[†]

[†] 大阪大学大学院情報科学研究科 〒565-0871 吹田市山田丘1番5号

E-mail: †{honda.hiroyuki,iwata.mayu,hara,nishio}@ist.osaka-u.ac.jp

あらまし Web 上の膨大な情報の中から適切な情報を得る手段の一つとして、複数人による Web 検索が一般的に行われている。例えば、グループで食事するレストランを決定するといった状況が考えられる。このような場合、個人が挙げた候補を複数人でまとめ、最終的にグループ全員が納得するように意見を集約する必要がある。しかしこの際、各メンバの嗜好を考慮する必要がある、全員の満足を得ることは容易ではない。そのため、Web 検索におけるユーザの検索行動に基づいて嗜好の推定を行い、円滑に意見を集約できるような支援システムが有効であると考えられる。そのようなシステムの実現を目指し、本研究では、被験者実験を行いユーザの検索行動と嗜好の関係性について調査する。

キーワード 複数人 Web 検索, 検索行動, 嗜好推定, 意見集約

Hiroyuki HONDA[†], Mayu IWATA[†], Takahiro HARA[†], and Shojiro NISHIO[†]

[†] Graduate School of Information Science and Technology, Osaka University

1-5 Yamadaoka, Suita, Osaka 565-0871, Japan

E-mail: †{honda.hiroyuki,iwata.mayu,hara,nishio}@ist.osaka-u.ac.jp

1. はじめに

近年の Web 上における情報量の爆発的増加に伴い、ユーザは、膨大な情報の中から、適切なものを Web 検索によって選び出すことを求められている。適切な情報を得る手段の一つとして、複数人による Web 検索がある。例えば、グループで食事するために、複数人で協力して情報サイトを用いてレストランを探すといった状況が考えられる。複数人による Web 検索では、複数人が挙げた多数の選択候補の中から、グループに属するメンバ全員が満足できるような適切な候補を決定する必要がある。本稿では、選択候補として挙げられる個々のコンテンツをアイテムと定義する。情報サイトなどはアイテムの集合を対象とした検索サービスとみなせる。

ここで、口頭の議論において複数のメンバ全員が満足できるような意見の集約を目的とした手法 [1] や、参加者の発言を要約するシステム [4] が提案されている。しかし、これらはいずれも口頭で実際の議論を行うこと前提としており、議論のための時間拘束やユーザの負荷を考慮していない。個人での意思決定に比べ、グループにおける意思決定すなわち意見集約ははるかに複雑である [3]。メンバごとの嗜好の違いにより価値判断が異なる場合、メンバ間で相互理解を得たり互いに譲歩したりす

るなど複雑な課題が多く、意見を集約する上でユーザには精神的負荷がかかる。また実際にユーザが向き合っていく口頭の議論では、自らが無意識に重視している要素を提示できない、全員の意見が公平に反映されないことがあるといった問題が考えられる。そのため、ユーザの負荷を軽減しつつ、メンバ全員の満足を得るような意見集約システムの構築は有効であると考えられる。

メンバ全員が満足できる意見に集約するためには、メンバ個人の嗜好を考慮する必要がある。嗜好とは、ユーザがどのようなアイテムを好むのか、あるいはアイテムを構成するどのような要素を重視するのかという特徴を指す。たとえば、ユーザはファストフード店が好きなのかそれとも焼肉店が好きなのか、あるいはファストフード店のメニューが好きなのか安めの価格帯が好きなのかといった点で嗜好が異なると考えられる。個人による Web 検索においては、ユーザの Web 上での行動に基づき嗜好を推定しユーザの望むアイテムを提示するシステム [5], [7] が多く提案されている。

そこで本研究では、これらの嗜好推定の手法を参考にし、閲覧時間、ブックマークなどのユーザの検索行動を基に嗜好推定を行い、複数人の嗜好を集約してメンバ全員が満足する解を導くシステムについて検討する。本稿ではまず、意見集約を行う

ための、各ユーザの嗜好推定の手法を検討することを主な目的とし、被験者実験を通して Web 検索におけるユーザの行動と嗜好の関係性を中心に調査し、考察を行う。

2. 関連研究

2.1 Web 検索における個人の嗜好推定

特定ページや特定サイトの閲覧時間やクリックスルー、検索クエリなどユーザの検索行動に基づいて嗜好を推定する研究が行われている。

Morita ら [5] は、ニュース記事において閲覧時間がユーザの興味のレベルと強い相関関係にあることを示している。ニュース記事のように、ほぼ同等の規模のアイテムが複数存在するような状況では同じ傾向であることが期待できる。例えばレストラン検索サイトではほぼ同等の情報量を持つ店舗情報が複数存在するので、閲覧時間と嗜好に相関があると期待される。

奥ら [7] は、ユーザごとに重要視している評価基準を考慮した情報推薦システムを提案している。このシステムでは、ユーザの置かれている状況も考慮し、ユーザの嗜好と状況両方に応じた情報推薦を実現している。ただしこのシステムは、モデル構築のため、ユーザが本来の Web 検索に関係のない評価作業を行う必要があり、ユーザの負荷が大きい。

Agichtein ら [6] は、クリックスルー、閲覧時間、検索クエリを用いた嗜好推定手法を提案している。ユーザごとに事前の機械学習を前提としているが、多様なユーザ行動を考慮することで高精度な嗜好推定が可能となっている。

本研究では、これらの研究を参考に、閲覧時間やブックマークといったユーザの検索行動に基づき、嗜好推定を行う手法を検討する。複数人による意見集約を円滑に行うため、どのアイテムを重視しているのかだけではなく、アイテムのどのような要素を重視しているのかという多様性のある嗜好推定を目指す。

2.2 グループの意思決定

小柴ら [8] は、グループの意思決定の際「譲歩」の量がメンバー間で均衡すればグループ全体の満足度が向上するという「互恵性」を根拠にしたグループ意思決定支援機能を提案している。また福野らの研究 [9] によれば、相手よりも有利な解を得ることよりも、適度な譲歩によってグループの意思決定に貢献するような解のほうが、満足度が向上することが分かっている。これらの研究に基づく、グループの意思決定においてグループ全体の譲歩量を調整することは重要な要素であることが分かる。

グループによる意思決定支援のために、Saaty [13] や山田ら [14] は AHP (Analytic Hierarchy Process) を用いる手法を提案している。AHP とは、Saaty [10] [11] [12] によって提案された、複数の要素から構成されるアイテムの集合から、ユーザが望む最適なアイテムを選び出す意思決定手法である。AHP では、アイテム同士、及びアイテムを構成する要素同士で対比較を行い、アイテムの重要度を決定する。AHP のグループの意思決定への応用として、山田らは一対比較で決定する重要度に区間値を用いた手法を提案している。重要度に区間幅を持たせることで、多様なメンバーの意見を反映させている。本研究では嗜好を推定するため、嗜好を正確な数値で表現するのが困

難である。そのため、山田らが提案する、区間値を用いる手法を参考にすることが適切であると考えられる。

なお本稿においては、個人の嗜好推定に必要な要素を検討の対象とし、意見集約の手法は今後の検討事項とする。

3. 実験

複数人 Web 検索において、メンバ全員の満足を得る意見集約を実現するためには、個人レベルでの嗜好推定とそれに基づいたグループとしての嗜好を決定する必要がある。本稿では、まず、個人レベルでの嗜好推定に重要な要素を調査するために、被験者を用いた実験を行った。

3.1 実験設定

3.1.1 実験システム

実験では、一般的な情報検索サイトにおけるユーザの自然な検索行動を観察するため、Web アプリケーションとして実験用システムを構築し、以下のような機能を実装した。本稿ではレストラン検索を想定し、検索 API にはぐるなび API^(注1)を用いた。

- クエリ検索 (地域名, ジャンル, フリーワード)
- 検索結果の並べ替え (五十音順, 平均予算順, クーボンの有無)
- 店舗情報の詳細閲覧
- ブックマーク追加機能: フリーワードのタグ (任意), 5 段階のレーティング
- ブックマーク表示機能: 各ユーザが自分でブックマークした店舗のみを表示。
- ブックマーク共有機能: ブックマークした店舗を他ユーザに提示 (議論フェーズ (後述) でのみ使用)

システムのインタフェースを図 1, 2 に示す。画面左のサイドバーにはブックマークした店舗名が、ユーザが設定したタグおよび 5 段階のレーティングと共に表示される。ブックマークの追加機能において、ユーザがブックマークの整理などの目的で店舗ごとにタグを自由に付与できる機能を提供し、5 段階のレーティングでは、一般的な情報サイトなどでも多く見られる星印を用いた 5 段階評価を行う。ブックマーク共有機能は、議論フェーズ (後述) において、各ユーザがブックマークした店舗をグループ全体で共有する機能である。議論の最中「この店舗はどうだろうか」と提案する際に使用する。共有された店舗の情報は、全てのユーザのブックマークに追加される。実験では、以下に示すユーザの検索行動に関する情報を記録した。

- 検索クエリ
- 訪問時刻 (店舗情報のページにアクセスした時刻)
- 閲覧時間 (ページにアクセスし、離脱するまでの時間)
- ブックマーク (タグ, レーティング)
- 閲覧した店舗の情報 (店名, 平均予算, 徒歩時間, ジャンル)

3.1.2 実験方法

実験は被験者 3 名を 1 グループとし、被験者に 3.1.1 項で述

(注1): ぐるなび Web サービス <http://api.gnavi.co.jp/api/service.htm>

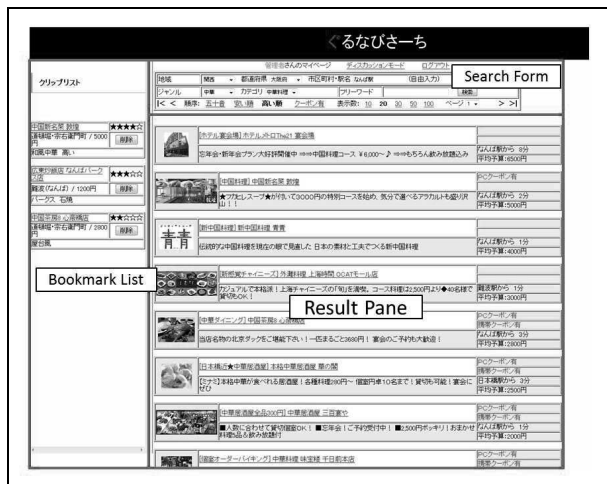


図 1 検索画面

べた Web アプリケーションを用いて自由に店舗を検索してもらった。被験者は、計 11 名からメンバを入れ替えて 5 グループを作成して実験を行った。被験者の内訳は 20 代男性 9 名、20 代女性 2 名である。

タスクは「週末の夜にグループの 3 名で食事を行うための店舗探し」と指定した。具体的には、まず店舗情報を個人で検索しブックマークとして追加してもらい（単独検索フェーズ）、その後、ブックマークを共有して口頭で議論を行ってもらいグループとしての解を決定してもらった（議論フェーズ）。

(1) 単独検索フェーズ 30 分：3.1.1 項で説明したアプリケーションを用いて、各個人の PC 端末で店舗情報を検索してもらった。この時、ユーザ間のコミュニケーションは禁止した。

(2) 単独検索フェーズに関するアンケート
単独検索フェーズ終了時には、以下のアンケートに回答してもらった。

- どのような点を重視して（どのようなお店を）検索しましたか？[自由記述]
- 満足行くまでお店の検索ができましたか？[5 段階]
- 自分の意見はまとまりましたか？[5 段階]
- 自分の意見は議論で積極的に主張するつもりですか？[5 段階]
- 各設問における理由 [自由記述]

(3) 議論フェーズ 30 分：各ユーザがブックマークした店舗情報をブックマーク共有機能によって共有し、口頭での議論も交えてグループとしての候補の上位 3 店舗を選出してもらった。

(4) 議論フェーズに関するアンケート
議論フェーズ終了時には、以下のアンケートに回答してもらった。

- 結論（決定したお店）には満足していますか？[5 段階]
- 自分の意見は十分に主張できましたか？[5 段階]
- 自分の意見は十分に反映されましたか？[5 段階]
- 議論は円滑に進みましたか？[5 段階]
- 議論のイニシアチヴを執ったメンバはいますか？またそれは誰ですか？[択一式]
- 各設問における理由 [自由記述]

単独検索フェーズについてのアンケートは主に検索行動自体の満足度と、ユーザの意識上の嗜好を把握するためのものであ

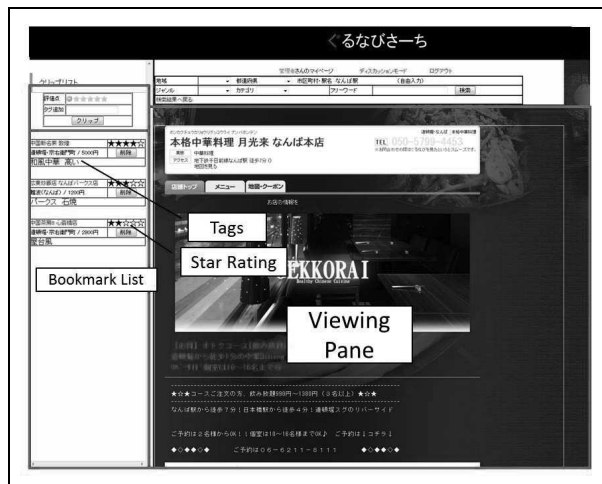


図 2 店舗情報の詳細画面

る。議論フェーズについてのアンケートに関しては、口頭での議論の特徴とシステムによって支援すべき点を把握することが目的である。また、アンケート内容については、最初に実験を行った 3 グループの結果を踏まえ、それ以降のグループでは、一部アンケート内容の変更・追加を行った。

実験は 3 名の被験者が円卓に互いに向かい合って着座し、自分の端末画面のみが閲覧できる形で行い、全てのフェーズが終了するまで離席は禁止した。実験時間は説明などを入れて 1 時間半程度（単独検索フェーズ 30 分+アンケート 5 分程度+議論フェーズ 30 分+アンケート 5 分程度）である。

3.2 実験結果と考察

本節では、実験によって得られたユーザの検索行動のデータとアンケート結果に基づき、ユーザの嗜好を推定する際に重要な要素について考察する。まず、ユーザのアイテムに対する評価について、どのような判断で評価を行なっているかを、アンケート結果を中心に検証する。次にユーザの検索行動の一つとして閲覧時間を取り上げ、アイテムに対する重要度との関係性を検証する。その後、検索行動、アンケート結果に基づき、具体的な嗜好推定の方法の検討と意見集約に必要な要素の調査を行う。

3.2.1 アイテムに対する評価

まず、ユーザが Web 検索においてどのような視点でアイテムを評価しているかについて検証する。表 1 に、アンケート「どのような点を重視して（どのようなお店を）検索しましたか？」に対する回答を示す。このアンケートは、グループ 1~ 3 に対して行い、グループ 4,5 に対しては行っていない。結果を見ると、飲食店検索において重視した点について、ユーザ間である程度共通項があることが分かる。特に「値段（予算）」、「駅の近さ」「雰囲気」「店のジャンル」などは多くのユーザで共通している。この結果より、ユーザは飲食店を検索する際いくつかの評価軸を基準に店舗を選定しており、さらに、その評価軸は全てのユーザである程度共通なものであるといえる。したがって、嗜好推定を行うためにユーザ間で共通の評価軸を用い、さらに、その評価軸はシステム側で静的に設定しても十分であると考えられる。

表 1 検索時に重視した点

グループ 1	
ユーザ A	飲み屋, 値段, 駅周辺, いろんな種類の店
ユーザ B	チェーン店でない, 値段, 雰囲気, メンバにあったところ
ユーザ C	個室の有無, 値段重視, クーポンやキャンペーンの有無
グループ 2	
ユーザ A	居酒屋, 値段 (安め)
ユーザ D	がっつり食べられる, 飲み屋, 安い, 雰囲気
ユーザ E	駅近い, 安い, 普段食べられないもの, 雰囲気
グループ 3	
ユーザ C	食べ放題, 安い, 雰囲気, 飲み屋
ユーザ D	安い, 駅近い, お酒が飲める, 隠れ家
ユーザ F	いろいろな種類の店, コスパが良い

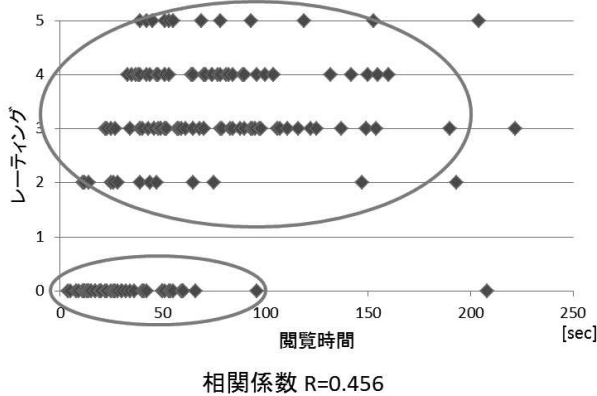


図 3 閲覧時間とレーティングの関係

3.2.2 閲覧時間と嗜好の関係

続いて、先行研究 [5] [6] においても示されている、閲覧時間と嗜好の関係性を検証する。図 3 に、各店舗情報の詳細ページにおける閲覧時間とユーザが付与したレーティングの関係を示す。閲覧時間とレーティングの相関係数は 0.456 であり、中程度の正の相関が見られた。レーティングは明示的なユーザの評価であるが、暗示的なパラメタである閲覧時間とある程度関連していることが分かる。特に、レーティングを与えていない（レーティングが 0 の）アイテムと、レーティングを与えている（レーティングが 1 以上の）アイテムの間には閲覧時間の差が見られる。しかし、レーティングが 1 以上のアイテムのみに注目した場合、レーティングと閲覧時間にはほとんど相関が認められない（相関係数 0.01 以下）。このことから、閲覧時間の長さは嗜好の有無を表す指標に用いることができる一方で、嗜好の強さを示す値として直接的に用いるべきではないといえる。今後、他に嗜好の強さを示す検索行動について、さらに調査する必要がある。

3.2.3 評価軸の重要度の推定可能性

ここまでで、ユーザが複数の評価軸を用いて店舗を評価しており、またその評価の良し悪しは閲覧時間に反映される可能性があることを示した。しかし、ユーザがそれぞれの店舗に対して、どの評価軸に基づいて高い評価を与えたのかは分からない。そこで次に、ユーザが各評価軸をどの程度重視して店舗の評価を行なっているか（各評価軸へのこだわりの強さ）について考察を行う。これ以降、飲食店を評価する際に一般的によく用い

られ、かつパラメタをデータとして比較的取得しやすい「平均予算」「徒歩時間」「ジャンル」の 3 つを、評価軸として扱う。他の評価軸については、今後の拡張における検討対象とする。

グループ 3~ 5 に対しては、単独検索フェーズ終了時のアンケートにおいて、以下の設問を追加した。

- 次の要素をどの程度重要視しましたか？また、理想とする値も教えて下さい。
 - － 平均予算（理想値は区間値で回答）
 - － 徒歩時間（理想値は区間値で回答）
 - － ジャンル（理想値：任意のジャンル列挙）

この設問では、3 つの評価軸（平均予算、徒歩時間、ジャンル）をそれぞれどれほど重視したかについての 5 段階評価、および、ユーザ自身が望む理想値を回答してもらった。これらの結果に基づき、ユーザが各評価軸をどの程度重視しているかという度合いを推定する方法について、ユーザが閲覧した店舗の持つ情報（ぐるなび API で取得した平均予算、徒歩時間、ジャンル）、ユーザが店舗に付与したレーティング、およびアンケート結果を踏まえて考察する。具体的には、ユーザがある評価軸のある値へ強いこだわりをもっているほど、閲覧したり高いレーティングを与える店舗の種類に偏りが生じ、結果的にユーザの閲覧した店舗の持つ情報の値のばらつきが、その評価軸に対するこだわりと相関関係にあるという仮説に基づいて検証を行う。例えば、あるユーザが他のユーザに比べて安い店舗にこだわりがあるのなら、そのユーザは「平均予算」の低い店舗ばかりを優先して閲覧し、高いレーティングを与えると予想される。結果として、安さを重視するユーザの全閲覧履歴の店舗の「平均予算」のばらつきは、他のユーザの閲覧した店舗の「平均予算」のばらつきよりも小さくなると考えられる。

以下では、各評価軸でユーザのこだわりの強さを推定する方法について順に検討する。

(1) 平均予算：

「平均予算」にこだわりをもつユーザは、自らが望む一定程度の範囲の予算額にあてはまる店舗を多く閲覧・ブックマークしたり、高いレーティングを与えると考えられる。そこで、ユーザの閲覧した店舗における平均予算のばらつき（標準偏差）を計算し、その標準偏差値が「平均予算をどの程度重要視しますか」というアンケートへの回答とどのような関係にあるかを検証する。標準偏差を求めるための平均値には、ユーザが与えたレーティングの値で重みを与えて算出する。ユーザ u が閲覧した店舗 $T_{ui}(i = 0, 1, 2, \dots, n-1 | n \text{ はユーザ } u \text{ の総閲覧数})$ の「平均予算」の値を b_{ui} 、その店舗にユーザが付与したレーティングの値を $r_{ui}(r = 1, 2, 3, 4, 5, \text{レーティングなしは } 1 \text{ として計算})$ とした時、重みを考慮した平均 A_u 、分散 V_u 、標準偏差 S_u を以下のように算出する。

$$A_u = \frac{\sum_{i=0}^{n-1} b_{ui} \cdot r_{ui}}{\sum_{i=0}^{n-1} r_{ui}} \quad (1)$$

$$V_u = \frac{\sum_{i=0}^{n-1} A_u - b_{ui}}{n} \quad (2)$$

$$S_u = \sqrt{V_u} \quad (3)$$

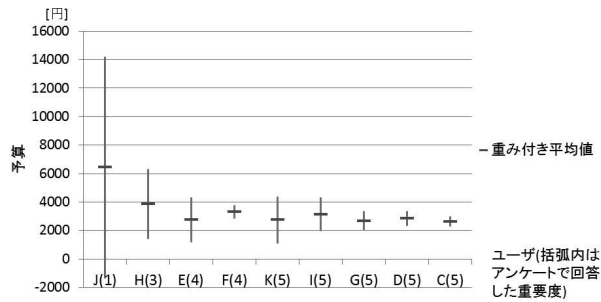


図 4 平均予算に対する重要度と閲覧した店舗の平均予算のばらつきの関係

図 4 に、各ユーザの閲覧結果から導出した「平均予算」の平均値と標準偏差、アンケートの回答の関係を示す。横軸はユーザを表し、縦軸は平均予算 [円] を表す。括弧内はそのユーザがアンケートで回答した「平均予算」の重要度を示し、ユーザは重要度の値の大きさにソートしてある。プロット点は各ユーザの重み付き平均値を表し、平均値から上下に伸びる棒グラフは標準偏差の大きさを示す。ユーザ F を除き、アンケートで回答された重要度が高いほど、閲覧した店舗の「平均予算」の平均値が小さく、かつ標準偏差が小さい傾向にあることが分かる。アンケートで回答された重要度の、平均値との相関係数は -0.944 、標準偏差との相関係数は -0.915 である。これより、「平均予算」を重要視するユーザほど、重要度が高いほど、ユーザが理想とする「平均予算」の値は小さくなっているといえる。またこの時、高いレーティングを与えた店舗の「平均予算」が一定値に偏っているといえる。逆に、「平均予算」に強いこだわりがないユーザは、他の評価軸に注目して閲覧を行なっているため、高いレーティングを与えられた「平均予算」が一定値付近に集中することがなく、ばらつきが大きくなると考えられる。

なお、ユーザ F について、アンケートで回答された重要度が 4 であるにもかかわらず、重要度が 5 と回答したユーザ C などと同程度に標準偏差、すなわちばらつきが小さい。ユーザ F は単独検索フェーズ後の自由記述のアンケートの設問「満足行くまでお店の検索ができましたか？」に対して 5 段階評価で 4 と回答しており、その理由として「(グループのメンバが) ある程度価格帯がはっきりしているメンバだから」と答えている。また、別の設問「自分の意見はまとまりましたか？」に対しては 5 段階評価で 5 と回答しており、これには理由として「価格も内容もちょうど良いお店が多くあったから」と答えている。これらの結果より、ユーザ F は主にグループのメンバを考慮し、閲覧したり高いレーティングを与える価格帯を固定していたことが分かる。このことから、「平均予算」の重要度は 4 ではあるが、その価格帯が検索開始時点から一貫して一定値に定められていたために、その標準偏差すなわちばらつきが小さくなったと考えられる。

以上の結果より、「平均予算」という評価軸については、ユーザの閲覧した店舗、店舗に付与したレーティングに基づき、「平均予算」のばらつきを求めることで、ユーザのこだわりを数値化できる可能性があることを確認した。

(2) 徒歩時間：

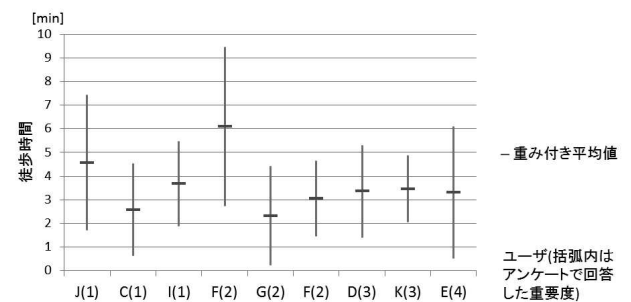


図 5 徒歩時間に対する重要度と閲覧した店舗の徒歩時間のばらつきの関係

「徒歩時間」についても、「平均予算」と同様、強いこだわりがあるユーザほど、閲覧した店舗の「徒歩時間」のばらつきが小さくなると予想される。「平均予算」と同様レーティング値で重み付けして分散、標準偏差を求めた。平均値、標準偏差とアンケートの回答の関係を図 5 に示す。「平均予算」と異なり、標準偏差の大きさとアンケートとの間にほとんど相関が見られず、「平均予算」と比べるとユーザ間で標準偏差の違いは小さい。アンケートで回答された重要度の、平均値との相関係数は -0.104 、標準偏差との相関係数は 0.021 である。ここで、アンケート「理想とする徒歩時間」に対する回答に対して、全てのユーザが「徒歩時間」はおよそ 15 分以下ならばいくらかでもよいと回答している。一方で、各ユーザが閲覧した店舗の「徒歩時間」は最大でも 10 分であり、ユーザの望む条件を満たしている。このことから、「徒歩時間」は、ユーザの店舗に対する評価にほとんど影響を与えず、閲覧履歴における店舗の「徒歩時間」が一定値に偏ることはなかったと考えられる。よって、単に「徒歩時間」のばらつきを重要度の推定に用いるのは、今回の場合においては適当ではない。今後、「徒歩時間」を数分単位でこだわるユーザを含めた環境で実験・検証を行い、重要度の推定手法について検討する。

(3) ジャンル：

「ジャンル」は数値データではないので、「平均予算」のように、値の平均や分散、標準偏差を求めることができない。そこで、各店舗のジャンルをベクトルで表現し、その分散値を求めることにした。具体的には、まず、ぐるなび API で定義されている 135 の「ジャンル」を、手動で 12 のカテゴリに分類した。分類の際、Wikipedia の Category の記事^(注2)、及び Yahoo!JAPAN カテゴリ^(注3)の分類を参考にした。例えば、「鉄板焼き」、「焼き鳥」、「串焼き」、「炭火焼・ろばた焼き」などのジャンルは「鉄板料理」というカテゴリに分類した。そして、各店舗のジャンルを 12 次元のベクトル空間で表現する。例えば、「焼き鳥」「和風居酒屋」というジャンルに該当する店舗は $(0,0,0,1,1,0,0,0,0,0,0,0)$ と表現する。

以上の方法で各店舗にベクトルを与えた上で、ユーザのレーティングの重みを考慮した分散値を求める。計算方法は「平均予算」「徒歩時間」と同様とし、対象をベクトルに拡張する。このように算出した分散値を「ジャンル」の分散値とする。図 2

(注2) : Wikipedia <http://ja.wikipedia.org/wiki/Category:飲食店>

(注3) : Yahoo!JAPAN カテゴリ <http://dir.yahoo.co.jp/>

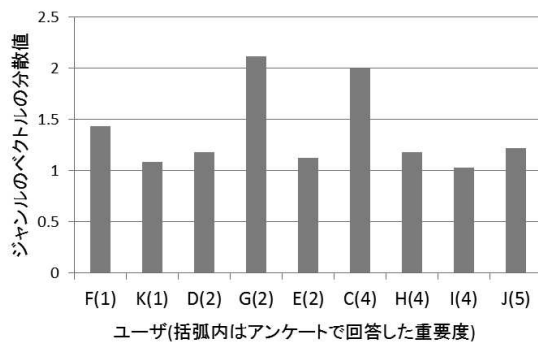


図 6 ジャンルに対する重要度と店舗のジャンルの分散値の関係

に、各ユーザの閲覧結果から導出した「ジャンル」の分散値と、アンケートの回答の関係を示す。分散値とアンケート回答の重要度にほとんど相関が見られないことが分かる。

この原因を明らかにするために、ユーザにアンケートで回答してもらった理想の「ジャンル」とユーザの閲覧した店舗のジャンルの関係を調査した。ユーザには、自分がどのようなジャンルを望んで検索を行ったか自由に回答してもらったため、以下、ユーザが自由に回答したジャンルのことを「希望ジャンル」と呼び、単に「ジャンル」と呼ぶものはぐるなび API で定義されている店舗のジャンルのこととする。アンケート回答を基に、ユーザ毎に閲覧した店舗のジャンルを、希望ジャンルに関連する「ジャンル」とそうでない「ジャンル」に手動で振り分け、それぞれのレーティングの総和を求めた。これにより、ユーザが実際にはどの程度希望ジャンルに偏って検索を行ない、高いレーティングを付与していたかを把握できると考えられる。この時、希望ジャンルに関連する「ジャンル」は実験者の判断で決定した。ユーザが閲覧した店舗が持つ「ジャンル」の希望ジャンルへの偏りと、ユーザの回答したジャンルの重要度の関係性を調査するために、希望ジャンルに関連する「ジャンル」のレーティングの総和を G_t 、それ以外の「ジャンル」のレーティングの総和を G_f とし、その比率 $P = \frac{G_t}{G_f}$ を求めた。 P の値が大きいくほど、希望ジャンルに関連した店舗を多く閲覧し、またレーティングを高く付与していると言える。図 7 に、各ユーザのレーティング総和の比率 P と、アンケート回答の重要度の関係を示す。アンケートで「ジャンル」の重要度が高いと回答したユーザほど、希望ジャンルに関連する「ジャンル」の比率が高く、希望ジャンルに偏って閲覧およびレーティングの付与を行なっていることが分かる。このことから、ユーザは自らが望むジャンルに偏って検索を行なっており、重要度が高いほどその偏りは強いといえる。

一方で、先に示したぐるなび API で提供されるジャンルに基づいたベクトルを用いた分散を求める手法では、重要度が高いほど一定値に偏るといった傾向は見られなかった。この結果の理由として、同一の「ジャンル」がユーザ毎に異なる意味で捉えられることが考えられる。例えば、ユーザ J が回答した理想ジャンルである「洋食」には「洋食居酒屋」が関連する「ジャンル」として合致する。一方、ユーザ E が回答した理想ジャンルである「居酒屋」にも「洋食居酒屋」が関連する「ジャンル」

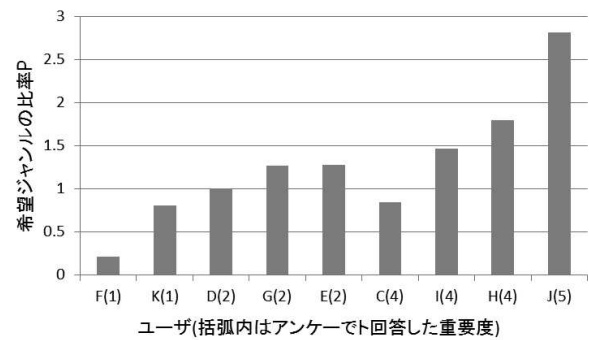


図 7 希望ジャンルとその他のジャンルの比、及び重要度の関係

として合致する。この例における「洋食居酒屋」のように、同一の「ジャンル」を異なる意味合いで扱うべき状況があるために、ユーザ間の「ジャンル」への認識の違いを考慮しなければならない。また、ぐるなび API で店舗に割り当てられた「ジャンル」とユーザの考えるジャンルとが一致していない可能性も考えられる。ぐるなび API で取得できる全 135 の「ジャンル」は、「洋食屋」という比較的抽象的なものから、「ステーキ」「パスタ」といった具体的な料理名まで存在するが、抽象度に応じた階層構造などを用いることなく、全て同列で扱われている。ここで、各店舗には、1 つもしくは 2 つの「ジャンル」が割り当てられている。事前にいくつかの店舗データの「ジャンル」について手動で確認したところ、店舗のページから受けた印象と、ぐるなび API でその店舗に割り当てられた「ジャンル」の内容がユーザ視点では適合していない例や、最大 2 つの「ジャンル」では店舗の特徴の表現に不十分と感じる例が見られた。例えば、焼鳥やもんじゃ焼きなど、幅広いメニューを提供している居酒屋 X があるとする。居酒屋 X の代表メニューは鉄板焼きであり、「和風居酒屋」「鉄板焼き」というジャンルが割り当てられている。ここでユーザが、「焼鳥」「お好み焼き」というメニューに魅力を感じて高いレーティングを付与した場合、「和風居酒屋」「鉄板焼き」というジャンルにレーティングが付与されたことになってしまい、ユーザの嗜好を的確に反映できない。

以上のように、ユーザ間の「ジャンル」の認識の違いや、ぐるなび API の「ジャンル」とユーザの「ジャンル」の認識の不一致のため、今回のように統一的なジャンルを利用する方法では、「ジャンル」の重要度を推定することが難しいと考えられる。ただし、アンケート結果で回答されたユーザの希望ジャンルを用いることで、「ジャンル」の重要度に応じてユーザの閲覧した店舗の「ジャンル」に偏りが現れることは示唆されたため、「ジャンル」をユーザの意図通りに適切に扱う方法について今後検討する必要がある。

3.2.4 各評価軸でユーザが望む値の推定可能性

評価軸の重要度を推定するだけでなく、各評価軸において、ユーザが求める具体的な値についても決定する必要がある。ここでは、「平均予算」および「ジャンル」の真値について検討する。「徒歩時間」は、3.2.3 項で述べたとおり、ユーザが理想とする値の範囲が広く、真値を推定してもその妥当性を検証するのが困難であるため、ここでは議論しない。

表 2 「平均予算」に関する真値とアンケートの関係

	アンケートの重要度	アンケートの理想値 [円]	真値 [円]
グループ 3			
ユーザ C	5	0-3500	2628
ユーザ D	5	無回答	2841
ユーザ F	4	無回答	3290
グループ 4			
ユーザ E	4	2000-4000	2738
ユーザ G	5	2500-3500	2662
ユーザ H	3	2000-6000	3854
グループ 5			
ユーザ I	5	3000-5000	3129
ユーザ J	1	0-5000	6432
ユーザ K	5	1000-4000	2741

まず、「平均予算」について検討する。「平均予算」については、ばらついた値の重心はその集合の代表値の一つであることから、3.2.3 項で求めた、重み付き平均を真値として算出することを考えた。表 2 に、各ユーザの「平均予算をどの程度重要視しましたか」というアンケートへの回答、ユーザが理想とする「平均予算」の区間値、推定した真値の関係を示す。真値がユーザの理想とする「平均予算」の区間の中央付近の値であることが分かる。このことから、「平均予算」に関しては、重み付き平均を真値とし、一定幅の前後値も含めた区間値をユーザの望む値の推定値とみなすことができると考えられる。また、アンケートへの回答の重要度が大きいユーザほど、理想とする値の区間幅が小さい。すなわち、ユーザの望む値を区間値として推定する際、その区間幅は重要度を用いて決定できると考えられる。

次に「ジャンル」の真値について検討する。「ジャンル」毎にユーザが付与したレーティングの総和を求め、総和が最も大きい上位 3 つの「ジャンル」を真値として決定した。表 3 に、各ユーザの「ジャンルをどの程度重要視しましたか」というアンケートへの回答、アンケートの回答による希望ジャンル、及び推定した真値の関係を示す。ユーザ C を除く全てのユーザで、推定による真値の「ジャンル」に希望ジャンルに関連のあるものが含まれている。このことは、レーティングの総和を求める上記の手法が妥当である可能性を示す。また、ユーザは希望ジャンルに関連する「ジャンル」を持つ店舗に偏って閲覧している傾向があることも示している。このように、ジャンルごとのレーティングを総和し、総和した値が大きいものを選ぶ手法は、「ジャンル」の真値を求めることに有用であるといえる。

3.2.5 グループの意思決定における議論の効果

複数人での議論による意思決定では、個人が単独で行う意思決定とは異なり、他のメンバの主張する意見に影響を受けながら意思決定を行うことが考えられる。本節では、議論フェーズ終了後の、議論内容とその結果に関するアンケートに対するユーザの回答について考察を行う。アンケート「結論（決定したお店）には満足していますか?」、「自分の意見は十分に反映されましたか?」および「自分の意見は十分に反映されましたか?」に対し、高い評価点（4 または 5）を回答したユーザに着目する。以上の 3 つの設問に対する回答の理由には、大きく分

けて 2 つの傾向が見られた。一つは、“美味しそう”、“食べたい”という、店舗や料理に対する好みを理由とするもの（6/11 人の被験者が回答）で、もう一つは、“自分が提案した店舗が採用された”という、自らの意見が活かしたことを理由とするもの（5/11 人の被験者が回答）である。飲食店や料理に対する好みは、ユーザがそれぞれ持つ個人の基準で完結する判断である。他のメンバが提案したものであっても、自分の好みに合致していれば、満足度は高くなる。一方で、自分の提案した店舗が採用される時には、自分の意見がメンバに認められたことに対する満足が大きいと考えられる。メンバに認められることによる満足は、個人での意思決定では発生せず、複数人の議論において発生する。ユーザが、他のメンバに意見を認められたり、また、「譲歩」の量がメンバ間で均衡することによってグループ全体の満足度が向上するという「互恵性」[8] [9] に基いて、自分が他のメンバの意見を認めたことが他のメンバに伝わるように意見を集約し提示することで、グループ全体の満足を得られると考えられる。

今後、意見集約の手法を検討する上で、メンバ各個人の嗜好とともに、各メンバの意見をどの程度ずつ反映させるかも重要な要素となる。例えば、グループの意見を集約して実際に店舗を提示する時、あるユーザにとって 1 位は好みと異なる店舗であるが、2 位、3 位が自分の意見であったために結果的に高い満足度を得るといったことが考えられる。

4. 嗜好推定と意見集約の手法の検討

3. までの実験で得られた知見を基に、複数人の意見集約を視野に入れた嗜好推定の手法を検討する。複数人で行う Web 検索の対象は予備実験と同様に飲食店の情報とする。

(1) 複数の評価軸：ユーザの飲食店に対する評価を複数の評価軸から構成されると定義する。現時点においては、Web 検索で飲食店を選択・評価する際の基本的な要素として以下の 3 項目を用いる。

- 平均予算
- (徒歩時間)
- ジャンル

「徒歩時間」については、3.2.3 項で述べたとおり、今後は 15 分以下の小さい値を望むようなユーザを対象に検討を行う。また、これらの評価軸以外についても、今回用いたような重要度の推定方法が適用できるかどうかも含めて検討し、拡張していく。例えば「最寄り駅」「平均メニュー単価」「雰囲気」などが考えられる。「平均メニュー単価」や「雰囲気」などは、ぐるなび API などの情報サイト側から明示的には提供されていない情報なので、店舗のページ解析が必要となる。

(2) 評価軸の重要度：各評価軸をどの程度重視しているか（重要度）は値のばらつき（分散・標準偏差）を用いることで数値化できる可能性が示された。今後、重要度の具体的な推定手法を検討する。数値化することで、評価軸同士の比較も可能となり、あるユーザがどの軸をより重視しているかということも把握できるようになると考えられる。

(3) 評価軸の真値：各評価軸の真値について、本研究では

表 3 「ジャンル」に関する真値とアンケートの関係

	アンケートの重要度	アンケートの希望ジャンル	真値
グループ 3			
ユーザ C	4	お好み焼き・焼鳥	活魚・海鮮/インドネシア・東南アジア料理/和風居酒屋
ユーザ D	2	飲み屋/しっかり食べられる	焼鳥/炭火焼き・ろばた焼き/和風居酒屋
ユーザ F	1	刺身/魚介	郷土料理/活魚・海鮮/和風居酒屋
グループ 4			
ユーザ E	2	居酒屋	和風居酒屋/洋風居酒屋/ダイニングバー
ユーザ G	2	洋食/和食/他にないメニュー	創作料理/ブッフェ・バイキング/和風居酒屋
ユーザ H	4	イタリアン/洋食/ステーキ/チーズフォンデュ	イタリア料理/ステーキ/和風居酒屋/洋風居酒屋
グループ 5			
ユーザ I	4	それなりにボリュームを取れるジャンル	串揚げ・揚げ物/焼鳥/焼肉・ホルモン/カジュアルレストラン
ユーザ J	5	洋食/おしゃれ/今食べたいもの	創作料理/ステーキ/洋風居酒屋
ユーザ K	1	スタミナ・好きなもの	割烹/焼肉・ホルモン/ラーメン/韓国料理

主に閲覧時間とレーティング、およびユーザの閲覧した店舗情報に基いて推定できる可能性が示された。今後、真値の推定の具体的な方法を構築する。

(4) 意見集約：評価軸の重要度を基にグループとしての評価軸の重要度を算出する手法の一つとして、本稿でも紹介したグループ AHP を適用する方法が考えられる。グループとしての評価軸の重要度を決定した後は、グループとしての真値も算出する必要がある。グループ AHP の具体的な適用手法については今後検討する予定であるが、ユーザ間で同じ評価軸における重要度が異なるため、これを重みとして各ユーザの真値を集約（重み付き平均など）することなどが考えられる。

5. おわりに

本稿では複数人による Web 検索における意見集約の支援システムの構築を目指し、まず、ユーザの嗜好と検索行動との関係性を被験者実験によって検証した。検索行動のデータを、アンケート結果と対応させることで以下のような知見が得られた。

- ユーザは情報サイトなど、互いに同等程度の情報を持つアイテム群（飲食店など）を検索し選定する際、複数の評価軸で対象を評価しており、またその評価軸はユーザ間である程度共通している。
- ユーザがどの評価軸をどの程度重視しているかについて、ユーザのレーティングのばらつきにある程度反映されている。平均予算は特に、その分散値が小さいほどこだわりが強いと考えられる。
- 閲覧時間とユーザの嗜好には中程度の正の相関があり、閲覧時間を嗜好推定の指標に用いることの妥当性が示された。これを用いることで、ユーザの明示的な評価なしで、アイテムに対するユーザの評価を得られることが期待される。
- ユーザのアイテムに対する明示的な評価を重みとした平均値は、ユーザが望む真値を傾向的に表すものとして扱うことができる。

今後は、ユーザの検索行動を用いた嗜好推定の手法を確立させる予定である。また、意見集約に有効と思われる要素について、被験者実験を行い、意見集約のためにグループ AHP の導入方法について検討する予定である。

文 献

- [1] 高野伸栄, 鈴木聡士, 代替案修正ベクトル法による合意形成システムに関する研究, 土木学会論文集, Vol.716, pp.11-20, 2002.
- [2] 藤井敦, Opinion Reader : 意思決定支援を目的とした主観情報の集約・可視化システム, 電子情報通信学会論文誌, Vol.J91-D, No.2, pp.459-470, 2008.
- [3] 加藤直孝, 合意形成プロセスにおける参加者の視点情報の共有に基づくグループ意思決定支援システムの研究, 博士論文, 1998.
- [4] 古田一雄, 前原基芳, 高嶋亮祐, 中田圭一, 知的支援機能を備えた電子会議システム, 社会技術研究論文集, Vol.1, No.0, pp.299-306, 2003.
- [5] M.Morita, Y.Shinoda, Information Filtering Based on User Behavior Analysis and Best Match Text Retrieval, Proceedings of SIGIR '94, pp. 272-281, 1994.
- [6] E.Agichtein, E.Brill, S.Dumais, R.Ragno, Learning User Interaction Models for Predicting Web Search Result Preferences, Proceedings of SIGIR '06, pp.3-10, 2006.
- [7] 奥健太, 中島伸介, 宮崎純, 植村俊亮, 加藤博一, 情報推薦におけるユーザの価値判断基準モデルに基づくコンテキスト依存型ランキング方式, 情報処理学会論文誌データベース, Vol.2, No.1, pp.57-80, 2009.
- [8] 小柴等, 加藤直孝, 國藤進, 互惠性を用いたグループ意思決定支援機能の提案, 情報処理学会論文誌, Vol.50, No.1, 268-277, 2009.
- [9] 福野光輝, 大淵憲一, 最終提案交渉における受け手の拒否動機分析: 同一性保護の観点から, 社会心理学研究, Vol.16, No.3, pp.184-192, 2001.
- [10] Saaty T.L., A Scaling Method For Priorities In Hierarchical Structures, Journal of Mathematical Psychology, Vol.15, No.3, pp.237-281, 1977.
- [11] Saaty T.L., The Analytic Hierarchy Process, McGraw-Hill, 1980.
- [12] Saaty T.L., Decision Making With The Analytic Hierarchy Process, International Journal of Services Sciences, Vol.1, No.1, 2008.
- [13] Saaty T.L., Group Decision Making And The AHP, The Analytic Hierarchy Process, Springer-Verlag, pp.59-67, 1989.
- [14] 山田善靖, 杉山学, 八巻直一, 合意形成モデルを用いたグループ AHP, Journal of the Operations Research Society of Japan, Vol.40, No.2, pp.236-244, 1997.