

パーソナライズされた店舗推薦システムの提案

A Proposal of a Personalized Restaurant Recommendation System

学籍番号 47-126781

氏 名 少路 健太 (Shoji, Kenta)

指導教員 木實 新一 准教授

1. はじめに

近年、クチコミサイトや SNS、ブログなどの CGM(Consumer Generated Media)と呼ばれる個人の情報発信をデータベース化した Web サイトが発展してきたことで、消費者は従来のクチコミ情報や生産者からの直接的な情報に加えて、新しい参考情報源を得ることができるようになったといえる。いま、これらの情報源からいかに多くの情報を取得するか、また正しい情報はどれであるかを判断するということが必要となっているため、そこに情報格差があるといっても過言ではない。新しいサービスやシステムをつくる上で、ユーザと直接的接点のない情報発信者や、そこから得られる情報に対する信頼性をどのように評価するかということが課題となっている。

既存の身近な店舗推薦システムの例として、「食べログ」について述べる。このシステムは、レビュー（店舗の評価者）のクチコミによって店舗ごとに点数がつけられており、ユーザは、エリアやジャンル、価格等を指定して店舗を検索することができる。しかし、対象のコンテンツが飲食店であるということもあり、ジャンルによっては、主観的要素が強く点数に表れてしまうという問題点がある。よって、ユーザは自分と嗜好の一致する人を探すためにクチコミ内

容を細かく読まなくてはならず、それをしていないユーザは自分が満足のいく可能性の高い店を特定することが容易ではない。ユーザの店舗検索結果から、他の店舗を表示することはできるが、いずれにしてもその店舗が本当にユーザに合っているかどうかを判断することは容易ではない。また、会員登録により、ユーザ自身による、検索インタフェースのカスタムや、お気に入りのレビューの登録などはできるが、嗜好の類似性の高いレビューの情報を自動取得することはいまのところできない。

2. 店舗推薦システム

本研究では、実際の行動履歴を用いて情報を推薦するシステムの一つとして、飲食店の推薦手法について検討する。本研究で提案するシステムを 4 つの要素に分けて述べる。

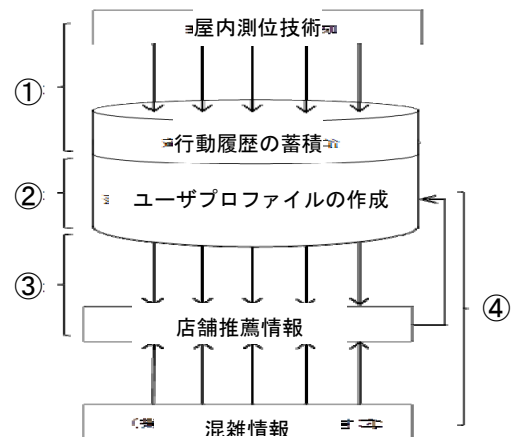


図 1 システム全体図

① ユーザの行動履歴の取得と蓄積

リアルタイムに詳細な行動履歴を自動的に取得することのできる方法を用いることが望ましい。現在、広く用いられている GPS による位置情報の取得は、飲食店のような屋内の位置情報を取得することができない。そこで、近年、屋内の位置情報を取得する方法として注目されている「屋内測位技術」を用いる。屋内測位技術を利用することにより、ビルや地下街、アーバンキャニオンにおいても店舗への来店イベントを高精度で推定することが可能になる。

屋外測位に基づく LBS(Location-Based Services)の研究は盛んに行われているが、屋内測位に基づいて、フロア情報を含む詳細な来店パターンから嗜好を推定し、きめの細かな飲食店の推薦を行うことのできるシステムの研究は、比較的未開拓な分野である。なお、本研究では、設備投資・手法の確立度合の観点から、Wi-Fi を用いた屋内測位の技術を用いる。具体的には、スマートフォンと施設内に設置された無線 LAN のアクセスポイントの電波状況から位置情報を取得し、店舗へ自動的に「チェックイン」を行う。それによって蓄積された細かい来店履歴のデータを用いて、ユーザにとって未開拓の店の推薦を行う。

屋内測位技術による行動履歴の取得が可能であるかどうか検証する実験を行った。本研究のシステムに実装するサービスの一例として、Walkbase というサービスを試験的に用いた。通常、位置情報サービスでは、地図と組み合わせて利用するものが多いが、Walkbase は緯度経度情報を特に意識せずに、例えば「総合研究棟の 435 号室」といった、記号的な場所の情報を扱うことがで

きる。Walkbase を用いた位置情報の取得実験手順は以下のとおりである。

(1)実験対象となる商業施設内の店舗において、アクセスポイントの電波強度パターンを反復測定し、ロケーションリストを登録する。

(2)アクセスポイントの情報を定期的に自動取得し、現在の日時・場所名・緯度経度・ロケーションの ID をサーバから取得するためのアプリケーションを開発した。このアプリを起動して店舗を訪れ、位置情報の取得状況を評価した。

② 行動履歴を用いた嗜好の抽出とユーザプロフィールの作成

屋内測位技術によって取得・蓄積した行動履歴から、ユーザの食の嗜好を把握し、ユーザプロフィールを作成する。飲食店の利用に関する行動と個人の食嗜好の関連性を評価するために、日常のなかでの飲食店の利用についてクラウドソーシングサイト「Lancers」と Web アンケート作成ツール「SurveyMonkey」を用いて、「外食嗜好に係わるアンケート」と称した調査を行い、191 件の有効な回答を得た。一般に、人は外食をする際、1 人であれば妥協の有無にかかわらず、利用する店舗を自ら決定しているといえる。しかし、グループになると、様々な理由で、選択権が他人に移ることがある。よって、外食をする人数によって、店舗の選択権がどこにあるのかを把握するために、「外食をする際、最終的な店舗の決定権を自分が握っていると思いますか？」という質問を、外食をする際の人数を 4 段階に分けて尋ねた。回答は、「全く自分では決めない」「ほとんど自分では決めない」「どちらかといえば自分以外が決めることが多

い」「どちらかといえば自分が決めることが多い」「ほとんど自分が決めている」「いつも自分が決めている」の6段階で、それぞれ順に1~6点とし、食事をする人数ごとに平均点を算出した。

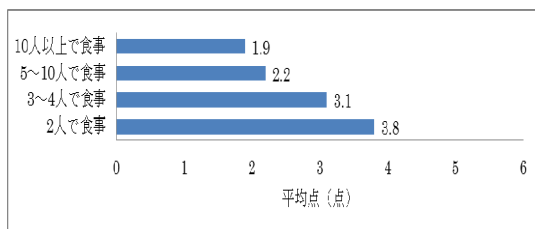


図2 自己選択度の分布

この平均点を「自己選択度」と呼ぶ。「どちらかといえば自分が決めることが多い」の得点である、3点を超えるのが「3~4人で食事」までである。つまり、この人数までは、個人の嗜好を用いて飲食店の推薦をすることが比較的可能である。自らの嗜好に一致していると自覚している飲食店を「行きつけの店」と定義し、アンケートから、人はなにをもって「行きつけ」であると判断しているのかということを分析した結果について述べる。1週間の平均外食回数と行きつけの店の利用頻度から、1週間あたりどのぐらいの割合で行きつけの店を訪れているかという値を求め、これを「行きつけ度」とした。行きつけ度 FL (= Favorite Level) は、1週間の平均外食回数 a と、行きつけの店の1週間当たりの利用回数 f を用いて以下の式により算出する。

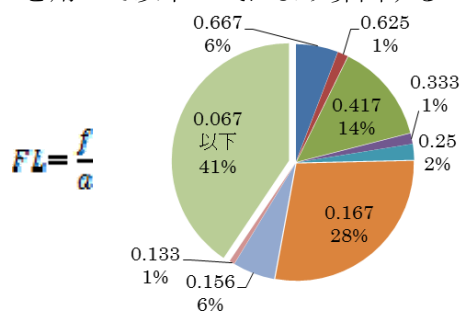


図5-11 行きつけ度 FL の分布

約6割が $FL \geq 0.133$ である ($1/7 < 0.1333 < 1/8$) ことから、ある一定の期間内に全ての飲食店外食した回数のうち、7~8回に一回以上訪れているような飲食店は行きつけの店であると判断できる可能性がある。

ユーザは利用する店舗を決定するとき、意識的または無意識的に総合的に判断して1つの飲食店に決定するが、そのバランスは人によって様々で、例えば、メニューの料金を最も重視するユーザもいれば、味を重視するユーザもいる。味といっても、その嗜好は人それぞれである。また「給料日前だから、安い店にしよう」とか「昨日は焼肉を食べたから、今日はカロリーを気にしてうどんにしよう」といったように、店舗を選ぶ条件は、同じユーザでも時と場合によって異なることも少なくない。よって、飲食店に関する嗜好は、長期的なものと、短期的なものを考慮する必要があるといえる。長期的・短期的に様々な状況に応じた行きつけの店を推測することができれば、そのユーザの嗜好を抽出したといえる。

③ ユーザプロフィールを用いた店舗推薦

アンケートにより取得したファミリーレストラン・ハンバーガー屋・牛丼屋の来店頻度、時間帯毎の外食頻度のユーザプロフィール、行きつけの飲食店リストを用いて、店舗推薦をするための手法について述べる。

ユーザベースの協調フィルタリングを用いて、類似度の高いユーザを導き、そのユーザが行きつけとしている店舗を推薦する。類似度算出には、ユーザとアイテムのマトリクスを用いる。「来店頻度」による評価値行列を Q と「時間帯毎の外食頻度」による評価値行列を R とすると、 Q と R に基づく類似度は、ピアソンの相関関数の式を用い

て以下のように求めることができる。

$$sim_q(a, b) = \frac{\sum_{p \in P} (q_{a,p} - \bar{q}_a)(q_{b,p} - \bar{q}_b)}{\sqrt{\sum_{p \in P} (q_{a,p} - \bar{q}_a)^2} \sqrt{\sum_{p \in P} (q_{b,p} - \bar{q}_b)^2}}$$

$$sim_r(a, b) = \frac{\sum_{p \in P} (r_{a,p} - \bar{r}_a)(r_{b,p} - \bar{r}_b)}{\sqrt{\sum_{p \in P} (r_{a,p} - \bar{r}_a)^2} \sqrt{\sum_{p \in P} (r_{b,p} - \bar{r}_b)^2}}$$

$$sim_q(a, b), sim_r(a, b)$$

評価値行列 Q, R におけるユーザ a とユーザ b の類似度

$q_{i,j}, r_{i,j}$

$n \times m$ の評価値データベースの行/列目の評価値

この二つの値の比重を変えて加算した値に基づき、推薦する店舗を決定する。

④ 推薦結果に対する付加情報の提供

推薦された飲食店の現在の状況（混雑度など）の情報を付加することで、リアルタイムの情報とする。また、どの混雑度で訪れたかの情報も蓄積することで、ユーザプロフィールの精度の向上につなげる。

混雑情報は、クラウドソーシングにより取得する。取得した飲食店の混雑情報にリアルタイムの情報の信頼度を付与することで、情報の疎性を補う。信頼度の算出には、情報がサーバに送られてから、別のユーザが参照するまでに店内の空席率に対して影響を与えたとされる不確定性を u と仮定し、信頼度 r (%) を以下の式により定義する。

$$r(\%) = (1 - u) \times 100$$

3. 結論と今後の課題

本研究では「屋内測位技術」を用いて詳細な行動履歴を取得・蓄積するという手法と、ユーザの「行きつけの店」に着目した嗜好の抽出方法により、解決を試みた。「屋内測位技術」における位置情報の取得は、取得した情報の正確性の面で改善余地はあるものの、今後のサービスの改良などによ

って、より正確な情報を取得できるようになると考えられる。また、「行きつけの店」に関しては、行動と嗜好の相関性を分析するためのアンケート調査により、「行きつけ度」という指標を示すことで行動履歴からユーザの行きつけの店を自動的に推測できる可能性が示された。また、同アンケート調査により、飲食店を訪れる人数、頻度、時間帯といった要素が嗜好に大きく影響するとわかり、頻度と時間帯を用いてユーザに推薦リストを作成する手法の提案に至った。人数に関しては、どの程度の人数までは個人の嗜好を用いて推薦をすることが可能であるのかという予測までは出来たが、それ以上の人数になったときに、複数の人数の嗜好を組み合わせる推薦に用いる手法の提案までには至らなかった。

参考文献

- [1]稲葉真純，長野伸一，長健太，溝口祐美子，川村隆浩，株式会社東芝研究開発センター，CGM 分析技術の現状と課題-メタデータ，オントロジーの応用可能性について，人工知能学会研究会資料，SIG-SWO-A603-06
- [2]澁谷寛，ネット上の消費者情報検索とネット・クチコミのマーケティング利用，AD STUDIES Vol.20 2007.
- [3]共立出版株式会社，監訳者，田中克己，角谷和俊，情報推薦システム入門.
- [4]吉澤菜津子，遠藤貴裕，永見健一，屋内位置情報における推定技術の開発と新しいサービスの展開について，INTEC TECHNICAL JOURNAL 2013 第 13 号.
- [5]塩野崎敦，無線 LAN 屋内測位を用いた新体験型位置情報サービスと将来構想，Lisra 設立記念シンポジウム.
- [6]atmarkIT，2012 年 11 月 19 日，<http://www.atmarkit.co.jp/ait/articles/1211/19/news012.html>，Android で使える O2O 技術まとめ解説.