

研究指導 中澤真 准教授

モバイル端末向けの位置情報を利用した情報推薦システム

笹川 可菜

1. はじめに

近年、インターネットを利用した情報推薦サービスが広く展開されている。例えばAmazon¹のようなショッピングサイトでは、ユーザの過去の閲覧・購買履歴から嗜好を分析し、これに適した商品をユーザ別に提供する機能が実装されている。

このユーザの特性抽出や嗜好分析に欠かせないのが、ユーザの行動を記録したライフログであり、中でも位置情報を利用するサービスに注目が集まっている[1]。位置情報は携帯端末のGPS²などを使った情報であり、実世界での行動履歴を取得することができる。このため、インターネット上での行動履歴である閲覧・購買履歴と同じように特性を把握する判断材料の一つとして利用でき、情報推薦の精度を高めるためにも効果があると考えられる。

位置情報を用いたユーザ行動の分析方法としては、位置情報からユーザの生活圏を推定して広告を無駄なく配信できることを目的とした小谷の研究がある[2]。しかし、この生活圏の推定方法を本研究の実証実験用データに適用したところ、生活圏が極端に狭くなってしまったり、逆に普段あまり行かない場所も範囲に含まれたりするなど、適切な結果を得られない場合があることが明らかになった。これを情報推薦に用いた場合、範囲が狭すぎて推薦対象が存在しないケースや、普段全く訪問しないエリアの店舗などを推薦してしまうなどの問題が生じる可能性が高い。このため、精度の高い情報推薦サービスを実現するには、他の計算方法を利用して生活圏を推定することが必要になる。

そこで本研究では、従来の情報推薦サービスの精度を上げるためモバイル端末向けの位置情報を利用した情報推薦システムを提案する。そのために、タイプの異なる被験者の位置情報をモバイル端末によって一定期間取得する実証実験を行い、このデータに基づいて理想的な生活圏の算出方法を明らかにする。そして、この新たに求めた生活圏を情報推薦システムの中でどのように活用すべきかを検討する。

2. 情報推薦サービスにおける位置情報

2.1. 情報推薦サービスとライフログ

ユーザー一人ひとりに適した情報推薦をするシステムでは、それぞれのユーザの特性や嗜好を正確に抽出することが何よりも重要である。この嗜好を把握するための情報として利用されるのがライフログである。ライフログとはユーザの生活行動を記録したデジタルデータのことで

あり、閲覧履歴などインターネット上での行動記録から食事や移動距離など実際の生活記録まで生活に関する記録のほとんどを指す。これらはユーザの特性・嗜好を分析するための豊富な材料となるため、用いるべきライフログの種類、収集方法、解析法を検討することが精度の高い情報推薦サービスの実現のために必要となる。

2.2. 情報推薦における位置情報の可能性

インターネット上の情報推薦では、ライフログとしてユーザの閲覧・購買履歴が利用されることが多い。これに対して、現実世界でのユーザの位置情報を用いた情報推薦サービスは少なく、位置の履歴を使わず現在地周辺の店舗情報を表示する程度の機能に留まっている。

インターネットという仮想空間上の行動履歴である閲覧・購買履歴に対して、位置情報は現実世界におけるユーザの行動記録であるため、リアル店舗の情報推薦などには不可欠なライフログである。にもかかわらず、位置情報が利用されない理由としては、個人情報である位置情報を収集されることに抵抗を感じる人がいること[3]、ライフログを個人情報としてどのように扱うべきであるかという制度が明確に定まっていなかったことが挙げられる。しかし、これらは位置情報の匿名化、収集間隔の調整など技術的に解決できる可能性が高い。またプライバシーに関するガイドラインを作成する法人も現れたことにより[4][5]、位置情報を活用したサービスは今後広まってくる。期待できる。

位置情報の履歴は、ユーザの行動パターンや生活圏といった特性を推定することで、情報推薦に活用することが可能となる。生活圏とはユーザが普段消費行動をとっている地域のことであり、位置情報から得られる移動履歴・場所などの情報から推定することが可能である。ユーザの生活様式や行動範囲に合わせた情報推薦をするためには、生活圏を的確に把握できるかどうか重要な鍵となる。

3. 位置情報に基づくユーザの生活圏の推定

ユーザの位置情報の履歴から生活圏を推定する手法としては小谷の研究がある[2]。この研究では、蓄積された位置情報から生活圏をユーザ別に推定し、その生活圏を市場としている広告のみを限定配信するための方式が提案されている。本章では、この研究における生活圏の推定方法とその問題点について述べる。

3.1. 生活圏の推定方法

小谷の生活圏定義を、数式を使って再定義する。まずユーザの一定時間毎の位置情報の履歴を (X_1, X_2, \dots, X_n) とする。全位置情報の中から特定の位置 X_i を抽出し、他の位置 X_j との距離 $L(X_i, X_j)$ を求める。そし

¹ <http://www.amazon.co.jp/>

² Global Positioning System(全地球測位システム)

て X_i を基点とした全ての距離の総和 S_i を以下のように求める。

$$S_i = \sum_{j \in all} L(X_i, X_j)$$

この S_i を全ての位置に対して計算し、 S_i が最も小さい位置を X^* とする。次に X^* から他の位置情報までの平均距離を以下のように算出し、

$$\frac{\sum_{j \in all} L(X^*, X_j)}{n}$$

平均距離以上離れた位置情報を除外する。残った位置情報群から新たな中心点を求め、他の位置情報までの平均距離を半径とした円を生活圏として定義する。

3.2. 実データに基づく生活圏推定の検証

上述の推定方法の適切さを検証するため、2 人の被験者に Apple 社のタブレット端末「iPad2」を常時携帯させ、一週間分の位置情報を 5 分間隔で自動記録させた。位置情報の取得には iPad 用アプリ「My Tracks³」を利用した。また 2 点間の距離の計算には Web サイト「keisan⁴」、緯度経度から住所を取得するために Web サイト「Google MAP API V3⁵」のお勉強⁵」を用いた。

3.2.1. 生活圏の半径に関する検証

まず自宅から短大まで距離が離れており、位置情報が広範囲に分布している被験者 A について、前節で述べた方法で生活圏を推定した結果を図 1 に示す。行動範囲に対して、生活圏が自宅付近の半径 593m の非常に狭いものとなってしまったことが図から読み取れる。このような行動範囲と大きく異なる生活圏では適切な情報推薦できない可能性が高く、また、エリアが狭すぎることにより推薦対象となる情報が存在しないというケースも考えられる。このような結果となった原因を探るため、既存研究が 2 段階の計算により位置情報を絞り込んでいる点に注目し、絞り込みをせずに 1 回の計算で生活圏を推定した場合と、定義通り 2 回の計算で推定した場合とを比較した。この場合の生活圏の半径は 4.51km となり、2 段階の計算で求めた生活圏よりもかなり大きなエリアとなった(図 2)。



図 1 既存研究の方法での生活圏



図 2 計算回数による生活圏の比較



図 3 重心が中心点の生活圏

3.2.2. 生活圏の中心点に関する検証

先の結果は生活圏の大きさに関する検証であったが、今度は生活圏の中心点について検証する。図 2 における被験者の生活圏の中心点は自宅付近となっているが、学生であるこの被験者は自宅と短大を往復する生活をしているため適切な中心点とは言い難い。また、生活圏の範囲に通学先の短大が全く含まれていない点も問題である。短大の滞在時間にも左右されるが、本来であれば自宅と短大の中間付近が中心点となる生活圏が、この被験者の行動範囲を的確に示したものだと思われる。そこで、中心点の算出方法を変更することで生活圏がどのように変化するかを検証した。中心点を他の点との距離の総和が最小となる位置とするのではなく、全ての測定データの緯度経度の重心、すなわち平均値とした。この結果を図 3 に示す。生活圏の半径は 4.68km と、従来の方法と比較すると範囲が非常に広くなり、中心点も自宅と短大の中間地点に移動したことが分かる。したがって、被験者本来の生活圏を示すことが可能となるため、この推定方法が有用であると考えられる。

これら 2 つの検証は、行動範囲が相対的に小さい被験者 B についても同じ結果となった。

3.3. 定義方法における問題点

検証の結果、計算回数や中心点を変える方法が効果的であることを示したが、時間に応じた生活圏というものをもさらに考える必要がある。時間を考慮せずに 1 つの生活圏を求めるだけでは、範囲が狭くなりすぎて推薦する情報がなかったり、また逆に広くなりすぎて該当する店舗が膨大な数になったりする可能性がある。一般的

³ GPS を使って移動履歴を記録する無料アプリ。

⁴ <http://keisan.casio.jp/>

⁵ <http://www.cocoaliz.com/GoogleMapsApiV3/index/8/>

に、時間帯によって行動範囲に差異のあるユーザが多いことが予想されるため、時間帯で分けた複数の生活圏を推定することについて検討する必要がある。

4. モバイル端末向け情報推薦システムの提案

前節の検証結果では、ユーザの生活圏を推定するために中心点として重心を使用し、生活圏の半径の算出にも中心点から全ての観測点に対する平均距離を用いるべきであることを明らかにした。ここでは、従来よりもユーザの行動特性が反映された生活圏を算出するために、ユーザが時々訪れるエリアまでを含む準生活圏と時間帯別の生活圏の推定方法について議論する。

4.1. 準生活圏

ユーザが時々訪れるエリアを対象範囲に含め、推薦情報を増やすために、生活圏を拡大した準生活圏となるものを定義する。具体的には、通常的生活圏の半径に重心から他の位置情報までの距離の標準偏差を加えたものを半径とする範囲を準生活圏と定める。生活圏の半径に標準偏差を2回3回と加算することで、準生活圏を階層的に定義することも可能であるが、本研究では1標準偏差分を加えた場合だけを扱うこととする。

この準生活圏の振る舞いを前節と同じ実証実験データによって検証した結果を図4、図5に示す。行動範囲が狭いユーザの場合は、平均から離れた複数の場所に位置情報が集中するため平均値と比べて標準偏差は大きくなり、準生活圏の半径は生活圏の半径の倍近くとなる。一方範囲が広い人の場合は、平均値付近すなわち生活圏の円周上近くに位置情報が集まる傾向にあるため平均値と比べて標準偏差が小さくなり、生活圏と準生活圏の差も小さくなっている。この方法により、ユーザの主要な行動範囲である生活圏と、時々訪れる場所までを含む準生活圏というような行動パターンに応じたエリア設定が可能になる。しかし、いずれの生活圏も中心点を重心としている点は共通であるため、半径の拡大と共に普段全く行かない場所も含まれてしまうという弊害も生じてしまう。行動範囲が広いユーザであればあるほど、その傾向は顕著となる。この問題に対処するため、ユーザの生活パターンに応じて時間帯別に複数の生活圏を推定する方法について次に述べる。

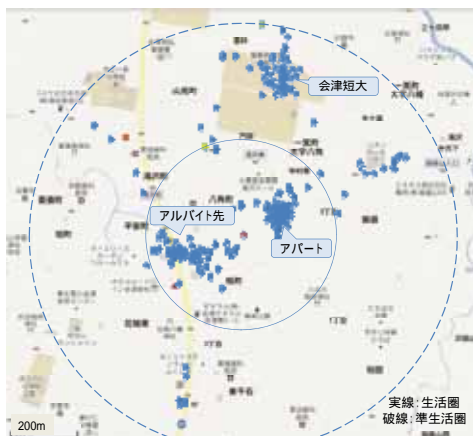


図4 行動範囲が狭い人の生活圏と準生活圏



図5 行動範囲が広い人の生活圏と準生活圏

4.2. 時間帯を考慮した生活圏

時間帯別に生活圏を推定する場合、時間区分をどのように設定するかが問題となる。そこで、「午前と午後」の二つの時間帯に分けた場合の生活圏の変化について検証した。図6に示したように、時間区分によって中心点がずれ、生活の範囲に差異が生じることが確認できた。しかし重複する部分も広いため、さらに時間帯を細分化することについての効果を検証する必要がある。

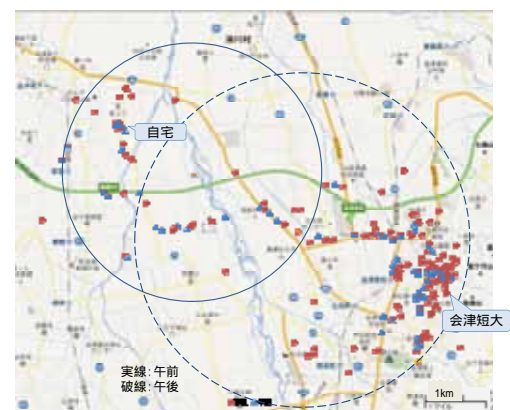


図6 午前と午後に分けた生活圏

そこで、「0～8時と8～16時と16～24時」と時間帯を8時間毎、3つに分けた場合について検証した。図7に示した結果では、先程と同じように重複する部分は広いが、細かく時間帯を分けることでユーザの活動時間帯の把握が可能であることが読み取れる。また行動範囲が狭いユーザほど時間帯による生活圏の差異が顕著になることが図8から明らかとなった。さらに、これらの結果は時間帯別の生活圏の半径を比較することによって、ユーザの主要な活動時間帯が推定できることを示唆している。最も半径の大きい生活圏を構成した時間帯が、そのユーザの最も活動的な時間帯であることを表しており、これにより朝型、夜型といったユーザの行動パターンの把握に活用することもできる。例えば、「夜型の人には居酒屋や娯楽施設、24時間営業の飲食店を推薦する」といったユーザのライフスタイルに合わせた情報推薦を行うことが可能となる。この他にも曜日で「平日と休日」に分けた場合の生活圏の差異について検証したが、時間帯を半日毎の2つに分けたケースとほとんど変化がなかった。

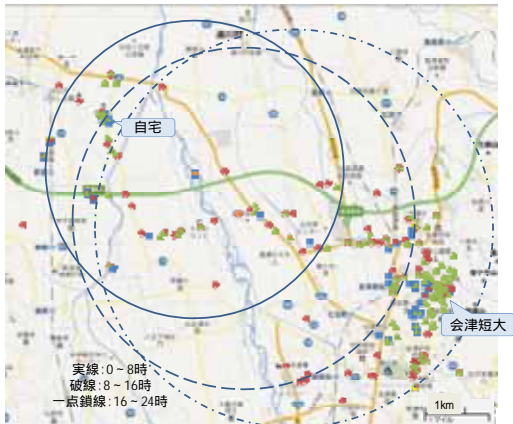


図7 時間帯を3つに分けた生活圏(行動範囲が広いユーザ)

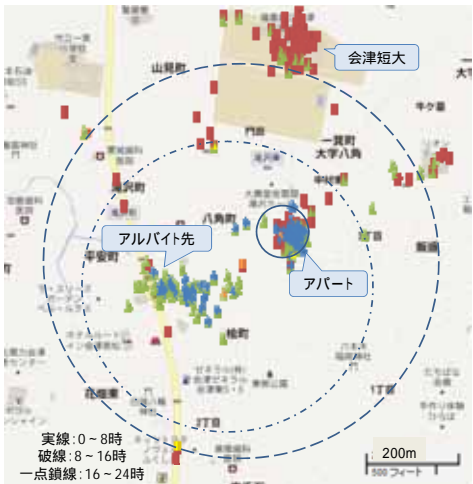


図8 時間帯を3つに分けた生活圏(行動範囲が狭いユーザ)

4.3. 情報推薦時における生活圏活用法

これまでの生活圏に関する検証結果を踏まえ、これらを情報推薦システムの中でどのように活用するか検討する。

まず、生活圏を反映させて圏内にある店舗から優先的に店名・場所といった店舗情報を表示する方法が考えられる。位置情報とは別のライフログに基づくユーザの嗜好に沿ったランキングリストがある場合には、そのリストに生活圏を反映させて順番を変更させる。例えば、生活圏、準生活圏、圏外の順で重みをつけて計算し、値が高いものから上位に表示するように順番を入れ替えるなどのアルゴリズムが考えられる。また、時間帯別の生活圏を用いることで、ユーザのライフスタイルを詳細に分析した精度の高い情報推薦や、ランチタイムなどの生活リズムに合わせた情報推薦が可能になり、情報推薦の幅が広がると考えられる。生活圏内のほとんどの店舗が既知であるというユーザに対しては、準生活圏を利用する方法が効果的である。情報推薦ではセレンディピティ⁶が重要であり、生活圏よりも少し外側のエリアである準生活圏内の情報を提供することで、ユーザが意外性のある情報に接する機会を増やすことができる。このことにより生活圏内の店舗を行き尽くしてしまったユー

ザにとっても効果的な推薦ができると考える。

このように生活圏を考慮して情報推薦を行なうことで、より個々のユーザに特化した効果的な情報推薦が可能になるのである。

5. むすび

本研究では、従来の情報推薦サービスの精度を上げるため、位置情報を利用した情報推薦システムについて提案した。生活圏を時間帯で分けることで、それぞれのユーザの活動時間帯に合わせた情報を推薦できることが可能となり、その活用方法についても検討した。これにより、ユーザに特化した情報推薦を展開することが可能であると考えられる。

今回は検証に用いたデータが実験機器の関係で被験者数・取得期間共に十分ではなかったが、信頼性を高めるために、また旅行や出張などの一時的な遠出があった場合の振る舞いを検証するためにも、データを増やして検証する必要があるだろう。また、時間帯をさらに細分化して生活圏を推定した場合に、ユーザのライフスタイルの推定がどの程度まで可能なのかということについても調査していく必要がある。

参考文献

- [1] IT media プロフェッショナルモバイル, モバイル情報サービスは「2012~13年に大幅な成長」 シード・プランニングが予測, <http://www.itmedia.co.jp/promobile/articles/1109/28/news073.html>
- [2] 小谷翔一, “位置情報を用いた生活圏定義型広告配信方式の提案”, 平成18年度フロンティアプロジェクト修士学位论文, 2007.
- [3] Garbagenews.com, ケータイの位置情報サービス、抵抗感のある人は4割近く, <http://www.garbagenews.net/archives/1605505.html>
- [4] 一般社団法人インターネット広告推進協議会, 行動ターゲティング広告ガイドライン, 2010, http://www.jiaa.org/download/JIAA_BTAguideline2010_100603.pdf
- [5] 総務省, 利用者視点を踏まえた ICT サービスに係る諸問題に関する研究会 第二次提言, 2010, http://www.soumu.go.jp/main_content/000061991.pdf
- [6] 佐野正弘, 位置情報ビジネス, 毎日コミュニケーションズ, 2011.
- [7] 竹内雄一郎, 杉本雅則, “位置情報履歴を利用したユーザアダプティブな街案内システム”, 電子情報通信学会論文誌 2007/11, Vol.J90 D, No.11, 2007.
- [8] 上田太一郎, 上田和明, 淵上美喜, 高橋玲子 他, エクセルで学ぶデータマイニング, オーム社, 2005.
- [9] ロジカルアーツ研究所, 緯度経度から距離と方位を求めるプログラム, <http://www.logical-arts.jp/blog/?p=136>

⁶ 思いがけないものを偶然発見する能力。