

近未来旅日記：行動支援のための情報収集と提示

間瀬 哲也^{*†} 大向 一輝[†] 濱崎 雅弘[¶] 武田 英明^{†‡§} 中山 泰一^{*}

^{*} 電気通信大学 情報工学科 [†] 国立情報学研究所 [‡] 総合研究大学院大学
[§] 東京大学 [¶] 産業技術総合研究所

1. はじめに

人がどこかを訪れる時には、何らかの目的が存在する。だが、目的が明確な場合もあれば、曖昧な場合もある。本研究では、目的が曖昧な場合の行動を支援するために、Web上に公開された文書を利用して行動の提案と関連する情報の提示を行うシステムを提案する。

情報収集では、行動するとき重要な要素である「位置」と「目的」に着目する。位置で検索することで、行程のプランニングが可能となる。さらに目的でフィルタリングすることで、望まない行動の提案がなされなくなる。プランニングされた場所は、文書収集のためのキーワードとする。収集された文書は、地名や位置情報を用いて話題の中心となる場所を推定し、位置的に近いと判断したものを選択し提示する。このように位置に依存したシステムとすることで、行程のプランニングと関連した場所に関する情報検索を組み合わせたシステムを実現する。

収集された情報は、旅日記風にして提示する。旅日記風に提示をすることで、行動の順序や内容を理解するための支援につながる。

2. 行動支援のための情報

2.1 行動に基づいた文書

昨今、多くの人々がWeblogを利用して情報を発信するようになってきている [1]。様々な内容が記述される中で、日記として利用され、感想や体験談などを記述したWeblogが多く見られる。体験談のような文書は、書き手の行動に基づいて記述されている。人の行動した結果には、他者の行動の手がかりとなる事柄が含まれている可能性が高い。それを利用することで、他者の行動を支援することができる

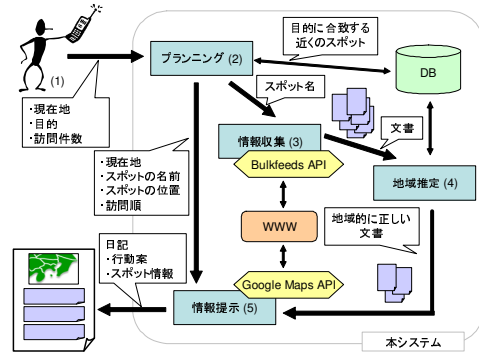


図 1: システムの概要

えた。また、行動に基づく文書は、場所を表す語が含まれている。そこで本研究では、行動を支援するための情報として、Weblogで公開されている場所について記述された文書に注目した。

2.2 文書が表す地域の推定

検索エンジンを用いてキーワード検索を行なった場合、キーワードとなっている語が含まれていれば文脈に関係なく結果として提示されてしまうという問題がある。Weblogでは多種多様な内容が記述される。そのため、文書中に地名や場所の名前が出現したとしても、実際にはその場所について記述されていない場合もある。本システムでは、文書中に出現した地名をもとに地域を絞り込み、さらに場所や駅の緯度経度を利用して話題の中心となっている場所の推定を行なう。

文書の表す地域を推定・分類する手法として、自然言語処理によるものがあるが [2]、本システムでは主に位置情報である緯度経度を利用して地域を推定する。この推定により、各文書の位置的な信頼度を高める。

3. 近未来旅日記システム

3.1 システムの概要

本システムは、プランニング、情報収集、地域推定、情報提示の4コンポーネントから構成される (図 1)。動作概要を以下に示す。

Travel Diary for Near Future: Information Gathering and Presentation to Support Users' Actions

Tetsuya Mase^{*†}, Ikki Ohmukai[†], Masahiro Hamasaki[¶], Hideaki Takeda^{†‡§}, and Yasuichi Nakayama^{*}

^{*}Department of Computer Science, University of Electro-Communications
[†]National Institute of Informatics [‡]The Graduate University for Advanced Studies [§]The University of Tokyo [¶]National Institute of Advanced Industrial Science and Technology

1. GPS 機能付携帯電話で位置情報を取得し、行動の目的と訪問件数を選択してサーバに送信する。

2. 1.をもとに、目的に合致するスポットを検索する。訪問件数が2件以上の場合、出発点を直前に検索したスポットに設定して同様の処理を行なう。

3. DB から検索された各スポットの名前と一般的に呼ばれているであろう名前をキーワードとしてRSS¹検索を行なう。

4. 3.で検索された情報から各スポット1文書ずつ選択し、地域推定を行ない、正しくその場所について記述されているか調べる。

5. 4.で正しいと判断された文書を、旅日記テンプレートを用いて旅日記を作成し、提示する。

3.2 プランニング

現在地の緯度経度(始点)・目的・訪問件数を受け取り、目的に適合するスポットをDBから検索し、最も近いスポットを選出する。訪問件数が2件以上であれば、選出されたスポットを視点とし、同様の処理をする。3件目以降についても同様の処理を行うことで、目的に適合したスポットを順に巡るプランが作成できる。

3.3 情報収集

行動に基づいた文書の収集には、Bulkfeeds API²を利用してRSS検索をする。プランニングで選出されたスポット名を検索クエリとしたキーワード検索を行なう。出来る限り多くの情報を得るため、検索クエリには正式名称だけでなく一般的に呼ばれているであろう名称も利用する。収集したRSSはスポットごとにファイルに保存して管理する。

3.4 地域推定

収集されたRSSから1文書をランダムに選択し、Mecab³を利用して形態素解析を行う。形態素解析により地名と判断された名詞と、日本郵政公社⁴が配布している郵便番号データの地名を比較し、都道府県を推定する。

連続する名詞(または未知語)から、考えられる全ての組合せで複合語を作成し、スポット名の候補群とする。推定された県から候補群と一致するスポットや駅をDBから検索し、一致したスポットや駅の出現回数を記録しておく。このとき、隣接県からも検索するようにした。

¹RDF Site Summary

²<http://bulkfeeds.net/app/developer.html>

³<http://chasen.org/~taku/software/mecab/>

⁴<http://www.post.japanpost.jp/index.html>

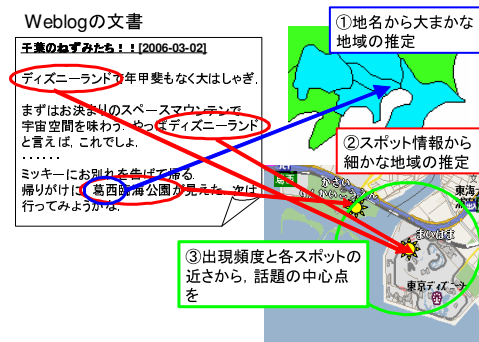


図 2: 地域推定の概要

話題の中心となっている場所の決定には、名前の出現回数と出現したスポット同士の距離によってポイントを加算していく手法を用いた。距離によるポイントの加算は、10km以内にある文書中出现したスポットに対して距離に応じたポイントを加算するようにした。この処理の結果、ポイントが最大となったスポットが、その文書における話題の中心であると推定される(図2)。

3.5 情報提示

先に述べた3コンポーネントによって処理された情報を、あらかじめ作成しておいた旅日記テンプレートに埋め込む。収集された情報は他人が書いた文書であるため、引用文として扱う。また、文書の原文へのリンクを生成し、現在地と提案したスポットの位置を記した地図をGoogle Maps API⁵を利用して生成する。

4. おわりに

本研究では、目的が曖昧な場合のための行動支援システムを提案し、実装した。行動を支援するために、他人の行動履歴を用いることで、行動の手がかりとなる情報を含んだ支援が可能となった。また旅日記として提示し、ストーリーをもたせることで、行動の順序や内容の理解の支援も可能となった。

参考文献

- [1] 総務省: ブログ・SNSの現状分析及び将来予測, http://www.soumu.go.jp/s-news/2005/050517_3.html (2005).
- [2] 金木 雄太, 山田 剛一, 絹川 博之, 中川 裕司: 地名辞書を利用した地名の曖昧性解消と文書の地域分類, 第19回人工知能学会全国大会論文集, 2E1-03 (2005).

⁵<http://www.google.com/apis/maps/>