

モバイル端末における旅行記の理解支援のための行程抽出と地図化

永澤 勇樹[†] 吉田 京平[†] 服部 峻[†]

[†] 室蘭工業大学大学院 〒050-8585 北海道室蘭市水元町 27-1

E-mail: †{14043031,13054053}@mmm.muroran-it.ac.jp, hattori@csse.muroran-it.ac.jp

あらまし 旅行記を読むことで様々な旅行を自宅に居ながら仮想体験したり、モバイル端末を携帯しながら旅行記の作者（有名人や歴史上の人物など）と同様の行程を実際に辿り、同じ体験をしたりすることができる。そのため、モバイル環境においても、旅行記を読む機会があると考えられる。しかしながら、モバイル機器は画面が小さく、また、旅行記は文字数が多いため、旅行記が読みにくいという問題がある。この問題を解決するために、旅行記から行程を抽出し、地図に表示することによって、旅行記の理解支援を行うモバイルアプリケーションを提案する。また、これらについての評価実験、及び、モバイルに適したインターフェースについての検討も行う。

キーワード 旅行記, モバイル端末, 理解支援, 行程抽出

A Study on Travel Path Extraction and Mapping for Understanding Support of Travelogues in Mobile Devices

Yuki NAGASAWA[†], Kyohei YOSHIDA[†], and Shun HATTORI[†]

[†] Muroran Institute of Technology Mizumoto 27-1, Muroran, Hokkaido 050-8585, Japan

E-mail: †{14043031,13054053}@mmm.muroran-it.ac.jp, hattori@csse.muroran-it.ac.jp

Abstract We can virtually experience various trips at home by reading travelogues. By carrying a mobile device, we can also trace the same travel paths in a travelogue that its author (e.g., a celebrity or historical figure) traced, and have the same experience that s/he had. Therefore, in mobile computing environments, there are some opportunities to read travelogues. However, there are some problems that mobile devices have too small screen to read articles, and that travelogues are hard to read because they tend to be too wordy. To solve these problems, this paper proposes a mobile application that supports understanding travelogues by extracting travel paths from travelogues and mapping them. In addition, the evaluation about travel path extraction and mapping, and some considerations about interface suitable for mobile devices have been carried out.

Key words Travelogues, Mobile Device, Understanding Support, Travel Path Extraction

1. ま え が き

旅行記には多数の旅行に有用な情報があり、旅行記を読むことで旅行の仮想体験を行うことができる。例えば、有名人や歴史上の人物などの行程から、仮想的に旅行するといった具合である。また、実際に旅行記を持ち出して実際に辿り、旅行記の作者と同じ体験をしたりすることもできる。そのため、モバイル環境においても、旅行記を読む機会があると考えられる。

しかしながら、多くの情報を伝えようとした結果、文字数が多くなった旅行記が多い。旅行中に長い文章を読むことはあまり望ましいことではないだろう。また、写真を本文と一緒に載せていることが多いため、モバイル機器のような小画面においてはレイアウトが崩れてしまい、読みにくくなることもある。

従って、情報を見落としやすいことも考えられる。

そこで本研究では、自然言語処理によって旅行記の行程を抽出して、それを地図にプロットする方法で旅行記の理解支援を行うことを考えた。ここで言う行程とは、旅行記における作者の移動経路である。最終的には、抽出した行程がマッピングされた地図を思い浮かべながら読んでいる状態になれば良いと考えている。そのための手順として、旅行記から行程を抽出する手法を提案し、評価実験を行った。また、モバイル端末に適したインターフェースの検討も行った。

2. 関 連 研 究

郡宏志らの「ブログからのビジターの代表的な行動経路とそのコンテキストの抽出 [1]」は、多数の京都観光に関するブログ

から代表的な経路を抽出して、それを地図に表示するという研究である。この研究では、旅行者の代表的な経路を求めているため、各ブログの厳密な行程を求める必要は無い。一方、本研究ではブログを含む個々の Web 記事から、行程を正確に抽出する必要がある。また、関連研究では自宅の最寄り駅等、旅行の開始地点から目的地までの途中経過には触れていない。その上、抽出対象の内容は京都観光に限られている。本研究では途中経過も考慮に入れて、旅行場所を限定することがなく、旅行記から正確に行程を抽出する手法を提案する。

3. 提案手法

本章では、提案手法（システム）の概要を述べる。

3.1 行程抽出

行程を地図化する前に、まずは旅行記の本文から行程を抜き出す必要がある。具体的には、旅行記から作者が移動した経路を抽出する。各経路の起点または終点には、鉄道駅を利用する。ここで言う鉄道駅とは、Wikipedia の「日本の鉄道駅一覧 [2]」の項目に書かれている駅名、及び、国土交通省国土政策局国土情報課の「駅別乗降客数データ [3]」の属性情報の駅名である。この鉄道駅から駅名リストを作成している。

抽出した駅名を繋げることで、行程（有順序経路）を得る。順序については、実際に通った駅名の順序と、旅行記に書かれている駅名の順序は一致することが多いため、抽出した駅名を旅行記での出現順に繋ぐことで、行程（有順序経路）を得る。但し、同じ駅名が連続した場合は、一つに集約する。

駅名の抽出に用いる提案手法は以下の二つである。

3.1.1 提案手法 1

鉄道駅での発着を描写した表現にマッチする厳密なパターンをできる限り網羅的に用意する手法である。

「乗車」「到着」「乗換」「出発」「移動」等、発着を表す 5 つの動詞から 30 文字前をパターンマッチする。それぞれの動詞において、異なるパターンを用意して、マッチした駅名を通った駅とする。この時に文を跨いでも構わない。パターンには駅名の他に鉄道に特化した情報として、路線名や列車名、時刻表現も活用する。これらを用いて、旅行記から通ったと考えられる駅を抽出する。用いたパターンの一例を以下に記す。

- 「時刻」発の「列車」に（乗車）
- 「時刻」に「駅」に（到着）
- 「駅」に（到着）
- 「駅」で「列車」に（乗り換え）
- 「時刻」に「駅」を（出発）
- 「駅」を「時刻」に（出発）
- 「駅」を（出発）
- 「列車」で「駅」に（移動）

ここで言う「時刻」は、時刻表記を正規表現にしたものの他、「定刻」や「一分遅れ」等も含まれる。「駅」は駅名リストの駅であり、この情報から乗降駅を求める。「列車」は例えば、「函館本線札幌行き最終特急列車」や「特急ずらん 2 号」等、列車に関する情報である。

3.1.2 提案手法 2

形態素解析を行った後、「(駅名) ... (発着)」というパターンで緩く広く取ってから、例外処理でフィルタリングする手法である。(発着) は、駅に到着または出発したことを示す表現である。また、提案手法 1 と異なり、文を跨がない。

例外処理については以下の通りである。

- (駅名) と (発着) の間に、(発着) 以外の動詞が含まれるパターンは除く
- (発着) より前に格助詞「が」が含まれるパターンは除く（旅行記の作者以外の行動に関する描写を除くためである。但し、通過駅としては正しい場合もある）
- (発着) に否定の助動詞が含まれるパターンは除く（旅行記の作者の行動の否定形を除くためである）
- (発着) の直後に助詞「の」が含まれるパターンは除く

3.2 地図化

行程を抽出した後は行程を地図にプロットする。まずは、位置情報として駅の緯度と経度を取得して、抽出した駅を地図にプロットする。しかし、プロットする際に同名駅問題という問題がある。駅名が同じでも場所が異なる駅があり、その場合はどちらの駅かを判断しなければならない。どちらの駅なのかについての情報が文章中に無い旅行記が大半を占める。それでも、判断の基準として文章中の情報で判断することは可能であるが、実際に旅行記を読む場合において、前後に通る駅の位置からどちらの駅かを判断していることが多い。そのため、直前に通った駅から地理的に最も近い駅を次に通る駅とする。各同名駅はお互いに遠く離れており、それに加えて、一度の乗車で一気に遠くへは行かないという前提である。

3.3 インターフェース

図 1 は PC 環境での使用を想定したシステムのインターフェースであり、モバイル端末での使用は考慮していない。モバイル端末の画面の大きさを考慮すると、旅行記、地図、行程表の 3 つを同時に表示するより、一つずつ表示して、それぞれをユーザのタッチ操作で切り替えられるようにする必要がある。さらに、全体として画面遷移するだけでなく、経路毎に関連する箇所へと互いにリンク巡航したり、行程抽出結果に基づく索引付けによって、モバイルユーザの地理的位置に応じた自動スクロールや強調表示、モバイル検索機能などが有用である。

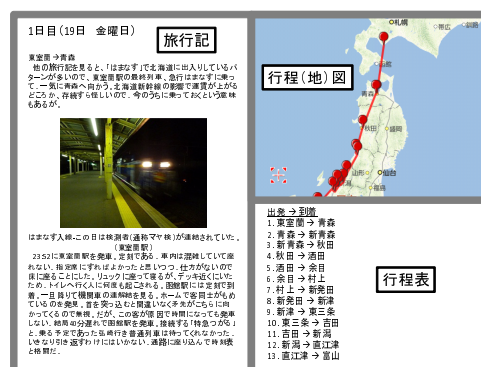


図 1 システムイメージ

Fig. 1 An overview of the proposed system.

4. 評価実験

本章では、提案手法のうち行程抽出について評価実験を行う。まず、4種類の比較手法を定義し、行程抽出精度に関して2種類の提案手法との比較実験を行う。また、提案手法の詳細分析も行い、改善の手掛かりを探る。

4.1 評価実験用の旅行記および正解セット

ウェブから100件の旅行記を無作為に選び、各々の旅行記から正解の行程（有順序経路）を読み取り、著者らのクロスチェックにより検証することで正解セットを作成した。基本的には、本文中から読み取っているが、旅行記の本文に、表やテキスト等によってまとめられた列車情報がある場合は、その列車情報を正解の経路とする。また、一列車ごとの移動記録となるが、実際には乗り換えが必要な経路でも、乗り換え無しで移動したように書かれている場合は、それで一つの経路とする。さらに、旅行記の最初または最後に、作者が列車に乗っている状態である場合は、例外として、乗車している列車の確認できる途中駅を経路とする。

正解の経路については、駅から駅まで乗車した経路を正解の経路とする。但し「空港駅」の場合は、駅に通らずにその空港を通った場合でも正解の経路とする。駅までの移動手段については問わない。通過しただけの駅など、下車または乗り換えしていない駅が含まれている場合は、正解の経路としない。

評価実験については、抽出した経路と実際の経路を比較して、再現率と適合率を評価する。抽出した結果の駅の組み合わせと、正解の経路の駅の組み合わせが一致したときを正しい経路とする。例え途中に通ったとしても、間に誤った駅が抽出された場合は正解としない。

4.2 比較手法との比較実験

まず、提案手法との比較実験のために用意した4つの比較手法について述べる。

- 比較手法1：本文中に出て来た駅名リスト上の駅名（○駅 or ○○）を文書中の出現順に繋ぐ
- 比較手法1.5：本文中に出て来た駅名リスト上の駅名（○駅のみ）を文書中の出現順に繋ぐ
- 比較手法2：（複合）名詞「○○駅」を文書中の出現順に繋ぐ。駅名リストに依存しない方法である。
- 比較手法3：駅名リスト上の駅名（○○駅 or ○○）、助詞、発着を表す単語が接続するパターンを切り出し、その駅名を文書中の出現順に繋ぐ

但し、提案手法と同様に、接尾の「駅」を除去しても同一の駅名が連続して出現した場合には一つに集約している。また、「森（駅）」「関（駅）」「フェリーターミナル（駅）」など、接尾に「駅」が付かない場合、必ずしも駅名（固有名詞）としてではなく一般名詞として使用されている可能性があるため、比較手法1や比較手法3においては接尾に「駅」が付くパターンのみを抽出している。

表1は、提案手法2種類と比較手法4種類の順序付き経路抽出としての適合率・再現率およびF値、行程の起点および終点の正解率、無順序地点（駅名）抽出としての適合率・再現率およびF値を示している。失敗は行程が抽出できなかった旅行記であり、再現率と適合率が0%として計算する。

最後に、提案手法2種類と比較手法4種類の適合率・再現率の散布図を図2から図7に示す。

4.3 各手法の詳細分析

比較手法1については、再現率が高く、一経路も抽出できなかった場合は無い。しかしその一方で、適合率が非常に低い。一部の駅は抽出時に制限が掛かっているが、それが無い場合はさらに適合率が低くなると考えられる。図2の結果からは、右下の方に固まっており、再現率100%となる旅行記が多い。

比較手法1.5は、比較手法1と同様に駅名リストに依存し、かつ、比較手法2と同様に「○○駅」のみを使い、駅名のパターンを制限しているため、適合率が上がるはずである。しかしながら、上昇はわずかであり、また、失敗件数が全ての手法の中で一番多い。これは駅を表すのに「○○駅」と書かれていない旅行記が多いからだと考えられる。図3の結果からは、比較手法1と比べて再現率が大幅に減り、分布が広がっている。

比較手法2の特徴は、抽出結果が駅名リストに依存しないことである。これはリストに無い廃駅や新駅、通称なども抽出し得る反面、駅名ではないものも多く抽出してしまう危険性を伴う。そのため、ノイズが多くなってしまった結果、比較手法1や1.5よりも悪くなっている。100件中53件は、再現率または適合率が0%である。形態素解析を行い、接尾「駅」に限定する手法も考えられるが、単純なパターンマッチではなくってしまう。図4からは、比較手法1では右下に集まっていた結果が左へと移行したようである。

比較手法3の特徴は、他の3つの比較手法の中で最も結果が良い。しかし、再現率はあまり高くないが、他の比較手法よりも厳密なパターンを用いているためである。図5からは、適合率が50%以上の結果が他の比較手法よりも多いことがわかる。

表1 比較実験の結果

	行程（有順序経路）抽出			起点 正解率	終点 正解率	無順序地点（駅名）抽出			失敗
	再現率	適合率	F 値			再現率	適合率	F 値	
比較手法1	81.9874%	14.5878%	0.247	36.6336%	39.6039%	97.4359%	20.6467%	0.340	0
比較手法1.5	22.3420%	18.9137%	0.204	43.5643%	31.6831%	45.4639%	47.2499%	0.463	20
比較手法2	19.0385%	9.9005%	0.130	46.5346%	29.7029%	46.2959%	34.7658%	0.397	10
比較手法3	24.3781%	32.5455%	0.278	28.7128%	58.4158%	52.4903%	67.1300%	0.589	11
提案手法1	31.5581%	37.4831%	0.343	20.7921%	59.4059%	55.1431%	65.0106%	0.597	12
提案手法2	51.8108%	48.1059%	0.498	50.4950%	66.3366%	71.2966%	68.6850%	0.699	6

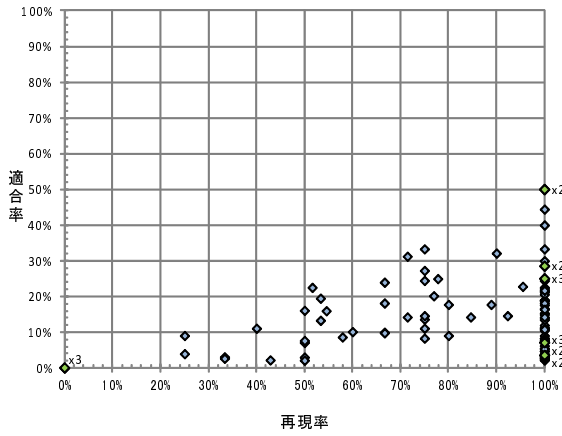


図 2 比較手法 1 (駅名 or 駅名+「駅」)

Fig.2 Comparative method 1 (station-name only or station-name + 「駅」).

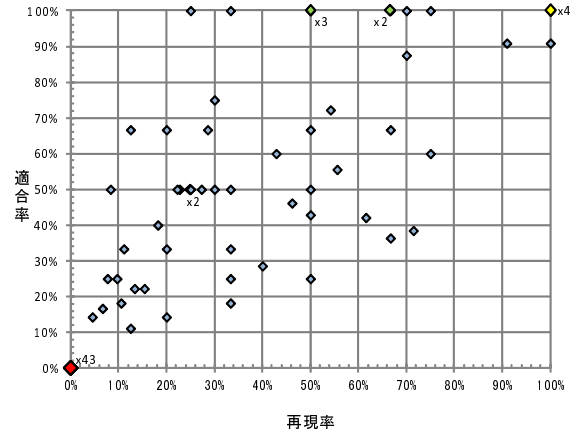


図 5 比較手法 3 (駅名+助詞+発着)

Fig.5 Comparative method 3 (station-name + particle + verb).

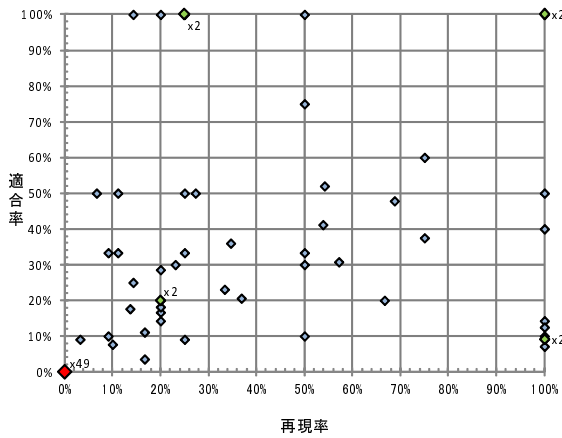


図 3 比較手法 1.5 (駅名+「駅」)

Fig.3 Comparative method 1.5 (station-name + 「駅」).

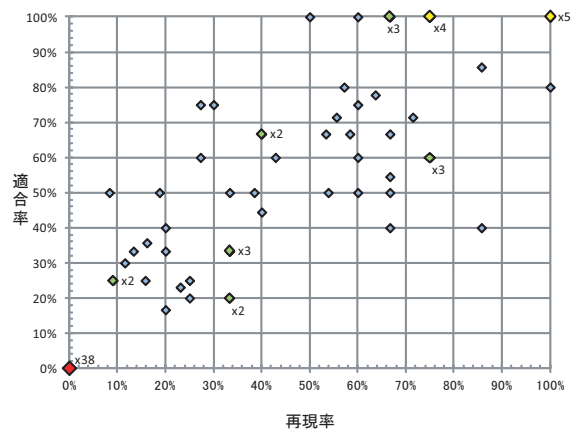


図 6 提案手法 1 (複雑なパターン)

Fig.6 Proposed method 1 (complex patterns).

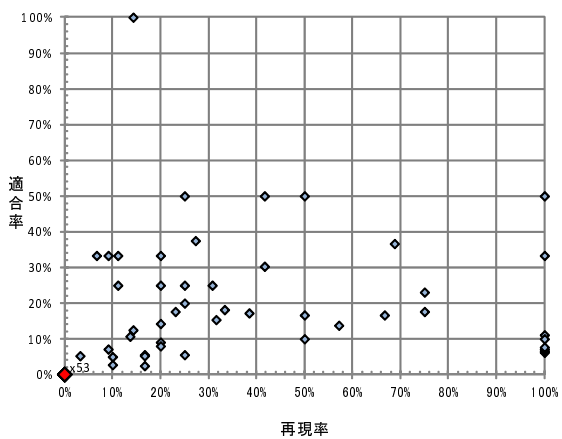


図 4 比較手法 2 (任意名詞+「駅」)

Fig.4 Comparative method 2 (noun + 「駅」).

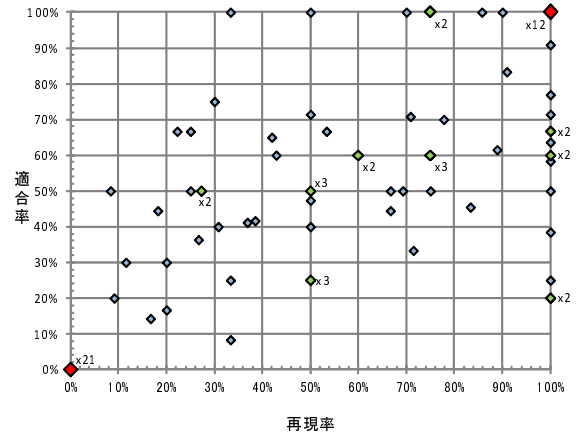


図 7 提案手法 2 (駅名+...+発着 → フィルタ処理)

Fig.7 Proposed method 2 (station-name + ... + verb → filtering).

提案手法1の特徴は、以上の4種類の比較手法と比べて、再現率と適合率が上がって来ていることである。また、図6から、比較手法3よりもやや右上に移行し、そこそ良い結果とあまり良くない結果の二つのグループに分かれている。図8は経路数毎の再現率と適合率のグラフである。

提案手法2は、図7から、提案手法1と比べて再現率と適合率が共に100%の点が増え、全体的にさらに右上に移行しているが、提案手法1では斜めに集まっていた点が散らばっている。

5. 考察

本章では、行程抽出、行程の地図化、行程表と行程(地)図を表示するインターフェースのそれぞれについて考察を述べる。

5.1 行程抽出に関する考察

行程抽出において、失敗の原因は主に二つある。

- 誤認…不正解の駅を誤認してしまうこと。適合率が低下する原因となる。
- 見逃し…実際の行程において、通った駅を認識しないこと。再現率が低下する原因となる。

これらについて、以下でそれぞれ詳しく考察する。

5.1.1 誤認の主な原因

誤認した駅は、「一応通っている駅」と「全く通っていない(可能性がある)駅」に分けられる。「一応通っている駅」には、移動途中に通った駅と実際の行動よりも先に述べられた通る予定の駅である。経路としては通っているのだが、本研究では、誤った経路と見なしている。また、正解経路の定義次第で結果が大きく変わる。しかし、文章による情報のみでは、一旦下車したが同じ列車に乗り続けた駅と列車を乗り換えた駅との区別が付かず、文章の意味を把握する必要がある。

また、予定について述べている文から駅名を抽出することも多く、この場合、正しい順序で抽出された後、再び直前の駅を抽出することが多い。表1によると、行程(有順序経路)抽出の再現率が低くても、無順序地点(駅名)抽出の再現率は高く、一度通った駅が抽出されてしまうことで再現率が低くなることから、同じ駅を通らなくすることで解決できると考える。同じ駅を通る旅行記の結果が上手く行かなくなる可能性があるが、調査対象の100件の旅行記中で同じ駅を通る旅行記は100件中25件(25%)と多少はあるが、有用だと考えられる。

とは言え、「一応通っている駅」の場合は、マッピングされるマーカーが増え、より正確な経路を描けることから、行程表お

よび行程(地)図の実用上の悪影響は無いと考えられる。

「全く通っていない(可能性がある)駅」については、作者ではなく、作者と同じ列車に乗っていた客が主語の文章から抽出されたため、誤認識した例がある。係り受け解析等を行い、主語が何なのかを解析する必要がある。

また、例えば列車の行き先や、列車について説明している時の地名等が抽出された例もある。列車が主語となる他は「○○行き」という言葉が使われている。最初に予定が述べられて急遽予定が変更になった場合、予定の部分で抽出された駅もある。

一般名詞を駅名と誤認した例もあり、例えば「駅前」や「フェリーターミナル」である。これらの駅は存在しているため、駅名リストにも含まれている。しかし、旅行記において、これらの具体的な駅を意味しないにも関わらず使われることも多いため、誤認の原因となる。対策としては、これらの駅を抽出対象から除外する、もしくは、これらの末尾に「駅」が付いている場合のみを抽出することで、解決できると考えた。

5.1.2 見逃しの主な原因

見逃してしまった駅は、パターンが合わずに抽出失敗したことが多い。しかし、闇雲にパターンを加えると誤認が多くなる可能性がある。また、動詞を元に抽出しているので、動詞が無い(省略されている)と駅名を抽出できずに見逃してしまう。そのため、助詞を用いて乗降した駅を示している文や、駅の写真を用いて乗降した駅を説明している文章も抽出できなかった。

パターンが合わない例に、普段使うことのない動詞が用いられていることもある。提案手法1で用いたパターンは、駅名の後に動詞が来る前提であるため、「次に降りた駅は、札幌駅です」のように、駅名の前に動詞がある文章も抽出できない。

駅名リストと完全に一致しなければ、駅として抽出できないため、ひらがなが混ざっていたり、誤字や括弧が含まれていたりの場合は駅名と認識されない。駅名にふりがなが振られている場合も見逃している。例えば「音威子府(おといねっふ)駅に到着した」では、(おといねっふ)があるため抽出できない。また、代名詞が用いられている場合も抽出できない。

別名だが同じ駅として扱っても問題が無い場合、よく片側が省略される。私鉄名が入っている駅で、私鉄名を略して書かれている場合は、誤認の原因となる。作者の表記ミスがあった場合も、訂正が行われずに見逃す。また、正式名称ではなく、通称等で書かれていると抽出できない。これは同名駅問題とは逆の問題として扱うことができる。

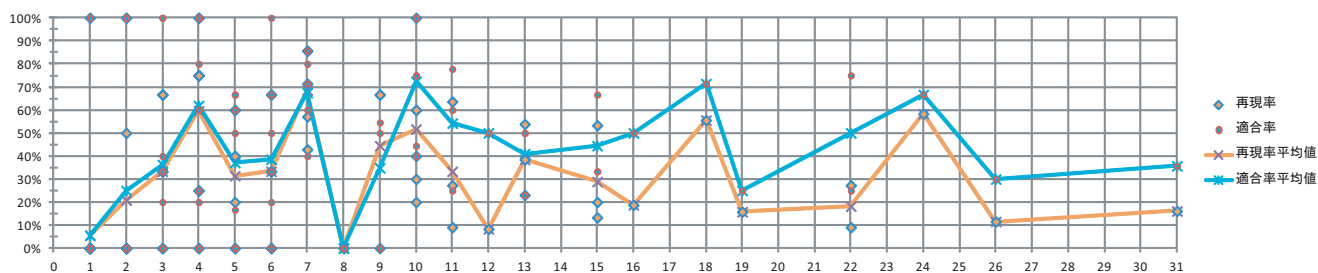


図8 提案手法1の経路数毎の再現率と適合率

Fig. 8 Recall and precision of the proposed method 1 per number of travel paths.

「いよいよ」や「やっと」等、感嘆詞等が入っているために抽出できなくなる場合もある。例えば、「駅に到着」では抽出できるが、提案手法1では「駅にやっと到着」は抽出できない。これらをパターンに加える手もあるが、何かしらの方法で感嘆詞を除去してからパターンマッチを行う方が効率的である。

本文中に駅名すら出てこない場合がある。特に経路数が多いと、乗り換えのためだけに降りた駅が本文に書かれず省略される傾向がある。それにも関わらず、列車名や路線名、本文以外に書かれている経路情報から、本文中には書かれていない駅を通ったことが分かる場合がある。そのため、列車情報や路線情報を用いて乗降した駅を推測する方法、または旅行記の経路情報を優先的に抜き出すという方法も有効ではないかと考える。

しかしながら、以上に挙げた情報すらない場合もあるが、地図にプロットされても文章中には殆ど出てこないため、抽出されなくても理解支援には影響が無いと考えられる。

5.2 地図化に関する考察

同名駅は互いにかなり離れていて、都道府県当たり一駅しか無いという前提で解決法を提案した。しかし、この前提に当てはまらない駅がある。それぞれの駅は互いに近くにあるとは言え、歩いて乗り換えると10分以上掛かる上、互いに連絡されていないため、同一駅として扱うには無理がある。

とは言え、どちらの駅が地図にプロットされても概ね行程を理解できるが、行程抽出に路線名等、位置情報以外の駅情報を利用することも考慮に入ると、区別することが望ましい。しかし、互いに近くにあるため、位置情報のみでどちらの駅を通ったかを判断すると、どちらの駅についても距離差が僅かとなり誤認する可能性が高い。この解決方法として、路線情報を活用して、前後の駅と同じ路線の駅を通った駅とすれば、どちらが通った駅かの判断は可能である。但し、抽出駅数が少ない場合も考慮に入れて、路線情報でも判断できない場合には、代表駅としてどちらか一方に絞るということも必要だろう。

また現行のシステムでは、駅名から位置情報を取得するために、外部のWebAPIを使用しており、位置情報を取得できる駅名は限られている上、追加や編集ができない。そのため、抽出結果を増やすために、よく使われる駅の通称を駅名リストに追加しても、これらの位置情報が得られないため、地図上にプロットできない。対策としては、追加した駅名をAPIに問い合わせる際に、位置情報が得られる駅名に変換してから問い合わせる方法が有効であると考えられる。

5.3 インターフェースに関する考察

現行のシステムは、PC環境を想定したアプリケーションとして開発している。画面が小さいモバイル端末において、旅行記、行程表、地図を同時に表示するのは難しいであろう。これら3つをユーザの操作で切り替えられるようにする必要がある。

各項目を切り替える時は、旅行記からは、書かれている地名の場所を確認する時に行程表または地図に切り替える。行程表からは、行程表の経路を地図上で見たい時に地図に切り替える。また、ある経路から作者がどう感じたかを見る時に旅行記に切り替える。地図からは、ある経路について、作者がどう感じたかを確認する時に旅行記に切り替える。また、経路を表形式

で見たいときに行程表に切り替える。

各項目をどのように切り替えるかについては、旅行記から行程表または地図へは、文章または地名をタップしたら、その文章が示す行程（行程表または地図）を表示する。行程表から地図へは、行程をタップしたら、その行程が示す経路と駅名を表示する。地図から旅行記へは、地図上の経路をタップしたら、文章の該当箇所を強調表示する。行程表から旅行記へは、表示されている経路をタップしたら、文章の該当箇所を強調表示する。地図から行程表へは、地図上の経路をタップしたら、行程表の該当箇所を強調表示する。また、行程表と旅行記の相互切り替えと地図と旅行記の相互切り替えが似ているので、行程表と地図を同時に表示するのも良いと考えられる。

6. まとめと今後の課題

本稿では、モバイル環境で旅行記の理解支援を行うため、長文で複雑になりがちな旅行記の概要（サマリー）として、自然言語処理を用いて旅行記から行程を抽出して行程表を自動生成し、それを地図上にマッピングする手法を提案した。4種類の比較手法を用いて比較実験を行った結果、最良の提案手法はF値で0.498（再現率0.518、適合率0.481）の抽出精度となった。

また、各々のメリットを組み合わせたハイブリッドな行程抽出手法として、提案手法2の行程抽出結果をベースに、経路を1つも抽出できなかった場合（100件中6件）だけ、抽出失敗が無い比較手法1の結果を採用することで、F値で0.526（再現率0.568、適合率0.490）の抽出精度が得られた。

今後の研究課題としては、駅名の抽出については、パターンの見直しの他に、実際には通過していない駅に関するパターンも必要ではないかと考えている。旅行記から抽出した駅の通る順序については、同じ駅を通らない前提で経路をプロットする他、駅的位置情報や路線情報（ある路線上にどんな駅が在るか）を用いることで、正確な経路を抽出することを考えている。また、発着を表す動詞の時制も考慮して予定や仮定の描写から抽出されるのを抑制する手法などによって、抽出精度のさらなる改善を図ったり、形態素解析だけでなく係り受け解析を活用することによって「次に下車した駅は士別です」「やっと着いた八代駅」などの駅名と発着を表す単語の出現順序が逆のパターンの抽出にも対応できるであろう。さらに、地点（駅）間を繋ぐ経路そのものだけでなく、発着時間や移動手段（列車愛称や種別、路線名など）も旅行記から抽出したり、他のデータベースやウェブマイニングした知識データ（各経路で体験し得る風景の画像など）も活用したりすることで、行程表や行程（地図）に関連情報を拡充して行くことを考えている。

文 献

- [1] 郡 宏志, 服部 峻, 手塚 太郎, 田島 敬史, 田中 克己, “ブログからのビジターの代表的な行動経路とそのコンテキストの抽出,” 夏のデータベースワークショップ DBWS 2006, pp.35-42 (2006).
- [2] Wikipedia, “日本の鉄道駅一覧,” <http://ja.wikipedia.org/wiki/日本の鉄道駅一覧> (2013/7/25).
- [3] 国土交通省国土政策局国土情報課, “国土数値情報 駅別乗降客数データ,” <http://nlftp.mlit.go.jp/ksj/gml/datalist/KsjTmplt-S12.html> (2014/3/7).