

列車の混雑情報に基づく参加型行動シナリオ推薦システムの提案

大野航・木實新一・笹尾知世・少路健太

Proposal of a participatory activity scenario recommendation system based on congestion information of trains

Wataru OHNO, Shin'ichi KONOMI, Tomoyo SASAO and Kenta SHOJI

Abstract: Today, a system to obtain the train congestion such as "KOMIREPO" is being used. However, it has not prompted behavioral changes of the user, to reduce congestion and improve comfort. In this paper, we propose a participatory recommender system that can collect possible activity scenarios from citizens and field research volunteers, existing information on the web etc. Thus, we aim to achieve a new ride form of citizens by the utilization of real-time congestion information and log of the past.

Keywords: 列車混雑情報 (congestion information of train), 行動シナリオ (activity scenario), 情報推薦システム (recommender system), 利用者参加型システム (user participatory system)

1. はじめに

列車の混雑度測定は、各鉄道会社が独自に行なっており、主に目視測定や自動改札データを用いている¹⁾。また、近年普及したスマートフォンのアプリケーションによって、乗客自身が手軽に混雑度を受信・提供できる参加型のサービスも提案されている。しかし、混雑情報に合わせて乗車変更を促す行動シナリオを提供するには到っていない。適当な駅で食事をしてから後続の電車に乗れば座れるといった情報を提供する必要がある。

そこで本稿では、列車の混雑情報に基づき、乗車変更を促す行動シナリオ（食事・ショッピング・休憩・散歩等）を Web やボランティア等によって把握した各駅の特徴や利用者のプロフィール等に基づいて、推薦するシステムを提案する。

2. 関連研究

2.1 列車混雑度測定

鉄道会社や個人による、列車の混雑度測定には様々な方法が提案されている。清水¹⁾は各鉄道会社が行なっている測定方法として、目視測定や応荷重装置、自動改札データ等があるとまとめている。我々²⁾は、各種センサ (CO₂・加速度) 用いて、混雑度を自動推定するための実験を実際の電車内で行った。その結果、CO₂濃度と混雑度には比例関係が見て取れ、加速度センサによって姿勢が判別できる。これらを用いた汎用的なツールによって混雑度を推定する可能性を示した³⁾。

また、近年普及したスマートフォン等の小型デバイスを用いた手法も提案されている。Navitime社は「こみれぽ」⁴⁾により、乗客自身が提供側・受信側となる参加型のアプリケーションを配信している。中野ら⁵⁾は、無線 LAN アクセスポイントの検索するためのプローブ要求を用いることで、各停車駅間の混雑度を推定する手法を提案している。

大野航 〒277-8568 千葉県柏市柏の葉 5-1-5

東京大学空間情報科学研究センター

Phone: 04-7136-4291

E-mail: wataru_ohno@csis.u-tokyo.ac.jp

2.2 乗客行動変化

時間帯や車両毎に列車の混雑度には、ばらつきが見られる。このばらつきの平準化が利用者の快適性の向上や列車の定時運行確保に有効である。松田・轟ら⁶⁾は、路面電車において、後続の電車の混雑状況をリアルタイムに情報板で表示させ、利用者に提供し、乗車変更を促す社会実験を行った。この混雑情報の提供により、混雑を分散させる可能性がある事を示した。

2.3 情報推薦システム

情報推薦の基本方式として、内容ベースフィルタリングと協調フィルタリングの2種類がある。前者は、推薦する情報の内容に基づき、情報の取捨選択を行う。後者は、同じ好みを持ったコミュニティを発見し、そのコミュニティが共通して好む情報を選択する。また、推薦を行う上で重要な要素である嗜好抽出技術には、明示的手法と暗黙的手法の2種類ある。前者は、ユーザから直接に、興味に関する情報を入力してもらう方法である。後者は、ユーザの Web 等の閲覧時の挙動から、ユーザの興味に関する情報を取得する方法である⁷⁾。本稿では、利用者の嗜好の取得に SNS（ソーシャルネットワークサービス）の一つである「Twitter」⁸⁾を使用する。早川ら⁹⁾はユーザ自身の投稿とユーザの友人の投稿を関連付けて *tf-idf* 法によりユーザの嗜好分析を試みた研究を行なっている。渡邊ら¹⁰⁾は Twitter に寄せられる投稿を分析し、語の共起関係と共に逆文書出現頻度を用いて作成した関連語の辞書により、語の関連性を考慮したユーザ嗜好の分析を行い、ユーザの嗜好を表す興味語を抽出する研究を行なっている。

3. 提案手法

3.1 推薦システム概要

まず、本稿で提案する推薦システムについて説明する。混雑度に基づいて、算出された空き時間（以下、行動可能時間とする）をどのように過ごすかは、駅自体の特性やその利用者自身の嗜好によって様々である。そこで本稿では、それらを考慮した食事や買い物や休憩、散歩など（以下、行動シナリオとする）を利用者の嗜好に合わせて推薦を行うシステムを提案する。図1に本稿で提案する行動シナリオ推薦システムの概念を例と共に示す。各詳細情報をそれ以下の節に記述する。

3.2 列車混雑度検索

列車の混雑度は時刻によって、大きく異なるため、リアルタイムに把握する必要がある。本稿では、2節で述べた「こみれば」のような参加型混雑度投稿アプリケーションにより容易に測定できる混雑レベル（表1）の取得を行うことを前提とする。また、その情報を参考にして行動可能時間 t を算出する。例えば、自分が乗ろうとしている列車の混雑度が Lv.5 という投稿があった場合、過去の履歴や他の投稿を参考にすることで、その電車を避けようとするため、Lv が下がる時間まで行動可能時間が発生する。

表1 混雑レベル⁴⁾

Lv.1	Lv.2	Lv.3
余裕で座れる	席はいっぱい	普通に立てる
Lv.4	Lv.5	Lv.6
圧迫される	身動き不可	乗れない

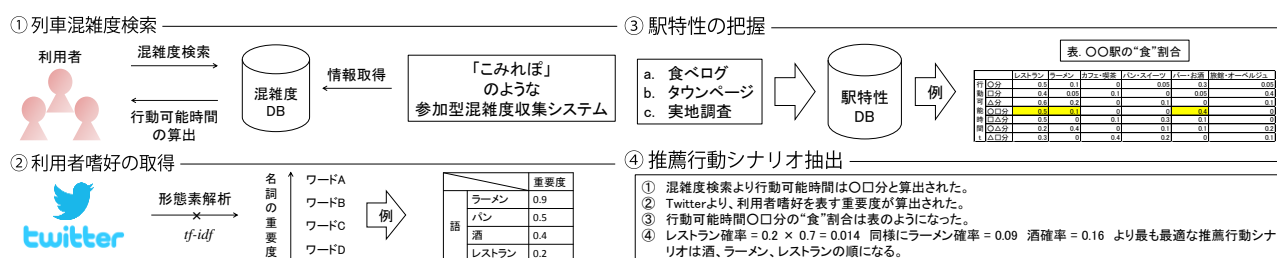


図1 推薦システム概要

3.3 利用者嗜好の取得

利用者嗜好の取得方法として、形態素解析と *tf-idf* 法を用いる。形態素解析とは、ある言葉が変化・活用しない部分を最小単位と捉え、その単位ごとに言葉を分解する手法のことで、最小単位が何であるのかを用意した日本語辞書と比較することで判別する。*tf-idf* 法とは、文書に出現する語に重み付けを行う手法である⁹⁾。Twitter API を用いて取得した利用者の投稿を形態素解析のフリーソフトである「茶筌」を用いて、名詞に限定して抽出し、*tf-idf* 法により重み付けを行う。その重みを重要度とすると、ユーザ投稿に含まれる語 i の重要度 w_i は以下の(1)式によって与えられる。

$$w_{i,u} = tf_{i,u} \cdot \log \frac{|U|}{df_i} \quad (1)$$

$tf_{i,u}$ はユーザ u の投稿における語 i の出現頻度を表し、Twitter ユーザの集合を U 、またその総数を $|U|$ 、 U のうち語 i を含む投稿を行ったユーザ数を df_i とする。

この手法を用いることで、各ユーザの投稿内容の名詞の重要度が測定され、重要度の高いワードに関連した特徴を持つといえる。

3.4 駅特性の把握

駅によって異なる特性を把握するため、本稿では主に3つの手法を用いる。表2にそれらの手法のためのカテゴリを示す。

表2 駅特性把握のためのカテゴリライズ

食	レストラン、ラーメン、カフェ・喫茶、パン・スイーツ、バー・お酒、旅館・オーベルジュ
買い物	アウトレットモール、お菓子・スイーツ、贈り物・みやげ、おもちゃ・ゲーム、靴・かばん・小物、コンビニ・スーパー・デパート・酒・飲料品、CD・DVD・ビデオ・レコード、食料品、スポーツ用品、制服・作業服、チケット、通販、文房具、ベビー・子ども用品、本、洋服、リサイクルショップ
実地調査	散歩、喫煙、休憩、娯楽

一つ目は、つくばエクスプレス¹¹⁾の各駅において実地調査を行い、Web上では取得できない情報（パンフレットや写真等）を取得した。二つ目は、食べログ¹²⁾の「食」のカテゴリの件数割合より各駅の食におけるカテゴリ特性を把握する。三つ目は、タウンページ¹³⁾より「買い物」のカテゴリの件数割合より各駅の買い物におけるカテゴリ特性を把握する。また、それらの情報取得方法を以下と図2に示す。

- ① 各駅の名前で検索する。
- ② 各駅の周辺の範囲（300m, 500m, 1km, 3km, 5km）においてそれぞれ絞り込む。
- ③ その結果をジャンル毎でさらに絞り込む。
- ④ 検索結果の件数を数え、周辺範囲とジャンル毎で件数割合を算出する。



図2 食べログによる食におけるカテゴリ特性把握手順

そして、それぞれの割合 R_t を以下の(2)式で算出する。行動可能時間 t における全件数を N_t とし、各カテゴリそれぞれの件数を C_t とする。「買い物」に関しても同様の手順で行う。

$$R_t = \frac{C_t}{N_t} \quad (2)$$

また、行動可能時間 t に関しては、距離 80m につき 1 分間要するとして算出した。

3.5 推薦行動シナリオ抽出

語 i の重要度より最適な行動シナリオを抽出する値を表す $P_{i,u}$ を以下に示す式(3)で表す. 式(1)より算出した, Twitter より取得した利用者嗜好を表す重要度 w_i と式(2)より算出した混雑度検索より得られた行動可能時間毎の各駅のカテゴリの特性 R_t を掛け合わせた. その値によって最適な推薦行動シナリオを抽出する.

$$P_{i,u} = w_i \times R_t \quad (3)$$

ここで, 語 i とカテゴリが一致する R_t の積を求めている. 一般には語 i とカテゴリは完全には一致しない. その場合は, 語 i に対応する R_t を求めるためにカテゴリを記述する語を網羅した辞書を利用することができる. この辞書は, 食べログやタウンページの情報を元に作成する.

4. おわりに

本稿では, 列車混雑度に基づいた, 様々な行動シナリオを推薦する手法を既存の情報推薦手法を用いて行う体系的なシステムの提案を行った. 将来的には, 様々な条件によって変動する“混雑度”をログ等により, 信頼性の高いデータとして「乗換案内」のような時刻表アプリケーションに付与することを目的としたい. そして, 参加型の投稿システムにより, リアルタイム性や情報量の増加を追求し, 従来容易に知ることのできなかった多様な行動可能時間の過ごし方を推薦し, 利用者をより快適にし, 結果的に列車の混雑が緩和することを目的とする. また, 行動可能時間中に複数のサイトを利用することのできるルートの推薦についても今後検討を行う予定である.

参考文献

- 1) 清水英範: 都市鉄道の混雑率の測定方法, 第3回鉄道整備等基礎調査報告シンポジウム予稿集, 2005
- 2) 大野航, 木實新一: 参加型センシングによる電車混雑推定方法の提案, 地理情報システム学会第21回研究大会, 2012
- 3) 大野航, 少路健太, 木實新一: 混雑センシングツールキット, ISPJ インタラクション2013シンポジウム (インタラクティブ発表), 1EXB-43
- 4) 「こみれぽ」〈<http://www.navitime.co.jp/?ctl=0171>〉 (2013/08/25 アクセス)
- 5) 中野隆介, 沼尾雅之: 無線 LAN アクセスポイントへの検索要求を利用した鉄道車内混雑度推定, DEIM Forum 2012 A10-1
- 6) 轟朝幸, 松本修一, 松田博和: 路面電車利用者への混空情報提供の有用性の検証, 運輸政策研究, Vol.11, No.1, pp.15-22, 2008
- 7) 土方嘉徳: 嗜好抽出と情報推薦技術, 情報処理学会論文誌, Vol.47, No.4, 2006
- 8) 早川豪, 岡部誠, 尾内理紀夫: Twitter を利用したソーシャルニュース記事推薦システム, 情報処理学会研究報告, データベース・システム研究会報告, Vol.2011, No.16, pp.1-4, 2011
- 9) 渡邊恵太, 加藤昇平: Twitter における語の関連性に着目したユーザ興味語抽出手法の提案, 人工知能学会誌, 2012
- 10) 「つくばエクスプレス」〈<http://www.mir.co.jp/>〉 (2013/08/25 アクセス)
- 11) 「Twitter」〈<https://twitter.com/>〉 (2013/08/25 アクセス)
- 12) 「食べログ」〈<http://tabelog.com/>〉 (2013/08/25 アクセス)
- 13) 「iタウンページ」〈<http://itp.ne.jp/?rf=1>〉 (2013/08/25 アクセス)