

タクシーの実車データマイニングに基づく、 待ち/流し位置レコメンデーション

松本謙治† 高山毅† 熊谷彩夏† 村田嘉利† 佐藤永欣†

†岩手県立大学ソフトウェア情報学部ソフトウェア情報学科

1. はじめに

近年、タクシー業界に対する情報システムの研究が活発化している。一般に、タクシーが顧客を確保するための方式には、車を停車させて顧客を待つ方式である「待ち」、車を走らせて顧客を探す方式である「流し」、顧客から場所の指定を受け、そこへ向かう方式である「無線(注文)」の三つがある。既存システムとして代表的な AVM(自動車両位置管理)システムは主に無線(注文)方式を支援する。しかし、待ち/流し方式の支援に関する検討は十分とは言えない。

そこで本研究では、実車記録の分析によって適切な待ち/流しの位置のレコメンデーションを行い、顧客の効率的な確保、利益の拡充を目指す。

2. 先行研究

文献[1]では、運転日報を用いてタクシー交通の行動実態を分析している。具体的には、日本国内の地方都市と大都市郊外部の都市について、乗客の移動事象であるトリップの特性を調べている。そして、タクシー車両の配置の重要性を指摘している。また、文献[2]は、タクシーテレマティクスシステムから収集した位置履歴データから、タクシーの乗車パターンを分析している。ここでタクシーテレマティクスシステムとは、AVMシステムのような移動体通信とインターネットを利用した、自動車向けの次世代情報提供サービスである。これらの先行研究では、タクシーが具体的にどこで待ち/流しをすべきかをレコメンデ(推薦)するには至っていない。

3. 実車データのマイニングに基づく待ち/流し位置レコメンデーション

本節では、著者らが行った四回の実験と取り組みを述べる。

3.1 四回の実験を通しての共通事項

3.1.1 分析方法

タクシー乗務員が記入する紙媒体の実車記録を電子化したものを分析し、傾向を抽出する。

3.1.2 レコメンデーション方法

待ち/流し位置としてレコメンデする地点の情報を、情報提供シートと呼ぶ紙媒体で乗務員に配布する。具体的には、実車地点単位で実車回数ランキングを提示する。

3.1.3 評価方法

以下を尋ねるアンケートを実験後に実施する:

- (i) 情報提供シートが役に立ったかの 5 段階主観的評価
 - (ii) 情報提供シートに基づく、待ち/流し位置の変更の有無
 - (iii) 待ち/流し開始から実車になるまでの平均時間
- (ii),(iii)より、レコメンデーションに従って待ち/流し位置を変更するか否かによる差異を分析する。

3.1.4 協力依頼先

県庁所在地であり地方都市の I 県 M 市のタクシー会社 H 社とその乗務員 150 名にデータ提供、実験協力を依頼する。

3.2 第一次実験: 時間の適切な粒度の検討

3.2.1 方法

表 1 一次実験対象期間

実験対象期間	7月第3金曜日から1週間
過去のデータの分析範囲	2008/7/18(金)~24(木)
レコメンデーション実施	2009/7/17(金)~23(木)

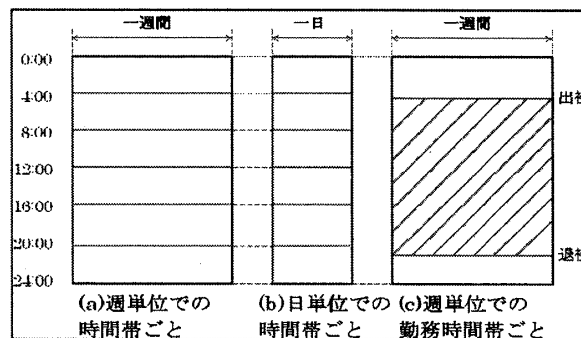


図 1 時間の粒度の三つの案。

図 1(a)~(c)の三通りの粒度で、実車地点単位で実車回数ランキングを各上位 3 件まで作成し、情報提供する。

3.2.2 結果

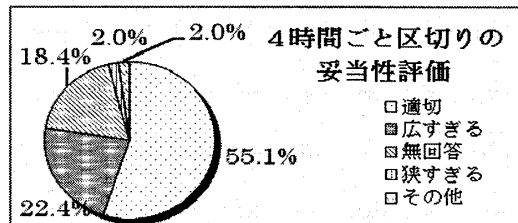


図 2 時間の区切りの妥当性。

図 2 は、図 1(a),(b)で設定した「4時間区切り」の妥当性の評価結果であり、4 時間ごとにすることに最も多数の回答が得られた。また、期間については、再現性を確保するためにタクシー会社側との協議に基づき、以降週単位(図 1(a))を採用する。

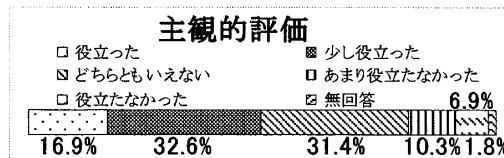


図 3 一次実験における主観的評価。

図 3 は一次実験における情報提供シートの主観的評価である。また、情報提供シートの効果としては、待ちで 4.1%、流しで 16.7%の時間短縮が見られた。

Waiting/Cruising Position Recommendation Based on Mining of Occupied Taxi Data

K.Matsumoto†, T.Takayama†, A.Kumagai†, Y.Murata†, N.Sato†

†Faculty of Software and Information Science, Iwate Prefectural University

3.3 第二次実験: 実車回数ランキングの適切な表示順位位の検討

3.3.1 方法

表 2 二次実験対象期間

実験対象期間	8 月第 2 日曜日 から 1 週間
過去のデータの分析範囲	2008/8/10(日)~16(土)
レコメンデーション実施	2009/8/9(日)~15(土)

情報提供シートに記載する実車回数ランキングの表示範囲を上位 3 件から実車回数が 10 件以上の全てに変更する。そして、乗務員として有用なのは、1~10 位, 11~20 位, 21~30 位, 31 位以降のいずれかを実験後のアンケートにて調査する。

3.3.2 結果

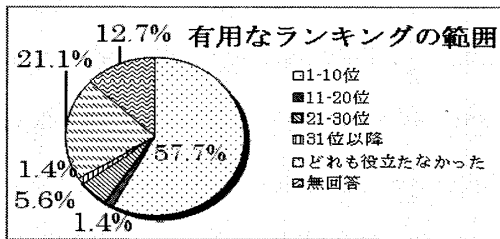


図 4 有効なランキングの範囲。

図 4 は、提供して有用な順位位の範囲の評価である。閑散期だったため、実用性としては上位 10 件程度であるということが分かったが、時間帯によっては 11 位以降も有用である場合も考えられるため、今後の実験では掲載範囲の変更は行わない。

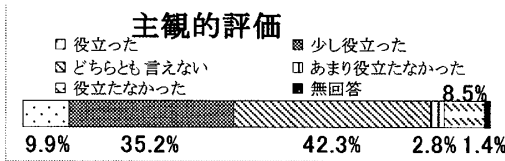


図 5 二次実験における有用性評価。

図 5 は、二次実験における情報提供シートの主観的評価である。情報提供シートの効果としては、待ちの短縮時間は 3.6%、流しの短縮時間は 2.6%であった。

3.4 第三次実験: 実車回数情報と料金実績情報の重要性の比較

3.4.1 方法

表 3 三次実験対象期間

実験対象期間	11 月第 1 日曜日 から 1 週間
過去のデータの分析範囲	2007/11/4(日)~10(土), 2008/11/2(日)~8(土)
レコメンデーション実施	2009/11/1(日)~7(土)

三次実験では地点ごとの実車回数ランキングに料金実績(平均額)を追加し、実車回数と料金のどちらが有用であるかをアンケートで評価する。

3.4.2 結果

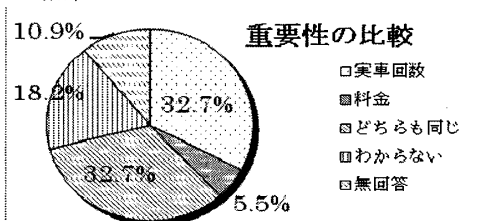


図 6 実車回数実績と料金実績いずれが重要か。
図 6 の通り、料金よりも実車回数が重要との評価であった。

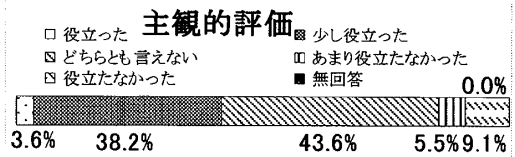


図 7 三次実験における有用性評価。

図 7 は、三次実験での情報提供シートの評価である。情報提供シートの効果としては、三次実験での時間短縮は待ちで 3.4%、流しで 6.7%である。

3.5 第四次実験: 実車地点の適切な粒度の検討

3.5.1 方法

表 4 四次実験対象期間

実験対象期間	11 月第 4 日曜日 から 1 週間
過去のデータの分析範囲	2007/11/25(日)~12/1(土), 2008/11/23(日)~29(土)
レコメンデーション実施	2009/11/22(日)~28(土)

四次実験では、今まで住所・施設名単位であった実車地点の粒度を 7 つのより粗粒度なエリア区分に変更する。そして、実験後アンケートにてエリア単位、住所・施設名単位のどちらが有用であったかを評価する。

3.5.2 結果

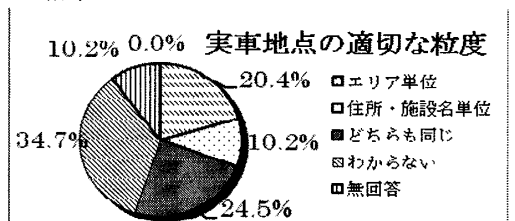


図 8 実車地点の適切な粒度。

図 8 の通り、住所・施設名単位より粒度の粗いエリア単位での分析が有効であるという結果が得られた。四次実験における有用性の評価をまとめたものが図 9 である。

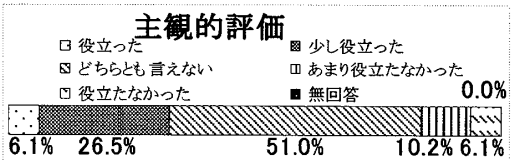


図 9 四次実験における有用性評価。

情報提供シートの効果としては、四次実験での時間短縮は待ちで 22.8%、流しで 12.6%である。

4. 結論と今後の展望

本稿では、タクシーの実車データを分析して、実車が期待できる待ち/流しの位置を乗務員へレコメンデーションを試みた。評価実験の結果によれば、提案手法は有用である。今後の展望として、以下の二点が挙げられる: (i)現在の状況をリアルタイムで取り込み、動的に需要を予測してのレコメンデーション精度の向上, (ii)AVM システムや、配車オペレータとの効果的な連携。

参考文献

- [1] 塚田悟之ほか: 運転日報によるタクシー交通の行動実態分析, 第 20 回交通工学研究発表会, pp.213-216, 2000.
- [2] S.Lee, G.L. Park: "Analysis of the Passenger Pick-Up Pattern for Taxi Location Recommendation", IEEE 4th Intl. Conf. on Networked Computing and Advance Information Management, pp.199-204, 2008.