

データ分析およびデータ利活用した新しいシステムの提案

研究の概要

私達の研究グループの特徴は、データ分析を行うだけでなく、既存データを応用し、現場の人を交えて新しいシステムを提案し、実際にそのプロトタイプを構築し、評価をすることができる点です。これまでも、複数の企業と共同研究を進めてきました。

最初にデータ分析の例を紹介します。図1は、人流データと駅の乗降客のデータとを組み合わせ、乗降客の居住地の推定を行った例です。この分析の特徴は、終着駅のデータを使っていない点です。単純に終着駅の入流データを利用するとその終着駅周辺在住の人の情報も含まれます。そこで、終着より前の駅のデータを利用することにより、乗降客の居住地の推定を行いました。図2は、道路を通過する車番データの分析例です。車番の分析の結果から、宗教と地域との関係性の可能性を見出すことができました。図3は、POSデータを用いた可視化システムです。POSデータは商品がいつ買われたかの情報が記載されています。いくつかの仮定を持ち込むことで、POSデータを利用して、いかにも人が購入しているような可視化システムを実現しました。図4は、図3の結果を利用した、混雑状況の可視化システムです。人の行動をある程度想定すると、どこに人が滞留しやすいかなども可視化することができます。

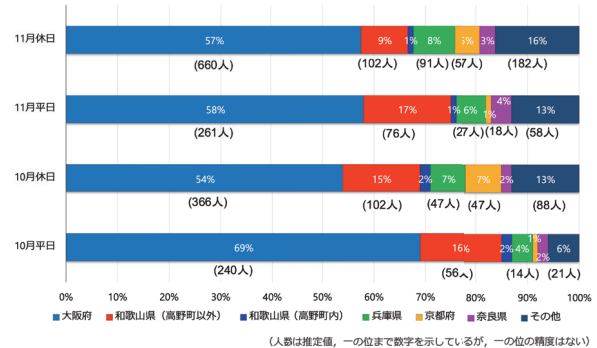


図1 高野山駅の降車する人の居住地推定

- 結果
 - 下記の府県で約90%
 - 和歌山 40%、大阪 30%、奈良 7%、兵庫 5%、愛知 2%、京都 2%、三重 1.5%
 - 近隣県からは多い
 - 東京、神奈川、愛知、埼玉、千葉、北海道、福岡
 - 人口比を考えると多いところ
 - 徳島県、香川県、岡山県、高知県、岐阜県、愛媛県
 - 四国が多いのは、四国のお道路と高野山との関係の強さ
 - 人口比を考えると少ないところ
 - 九州、東北は少ない (距離が遠いため)
 - 茨城、東京が人口比的には少ない (もって来て良いのでは?)
 - 各都道府県から高野山までの距離を考えた場合
 - 特徴的な結果はみられない
 - 九州、東北は少ない (距離が遠いところは少ない)

図2 都道府県別車両台数の割合のまとめ

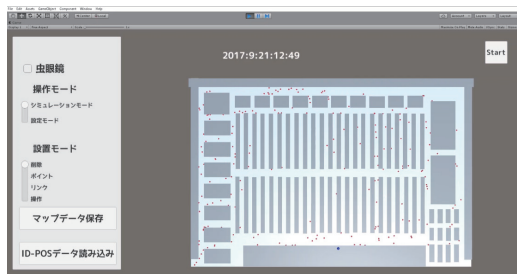


図3 POSデータを用いた人の動きの可視化

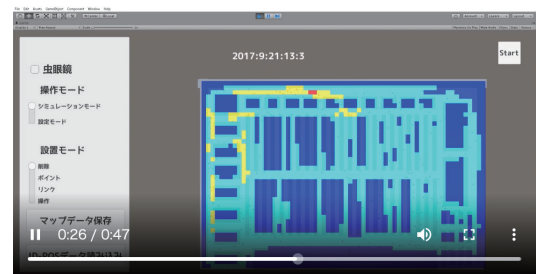


図4 POSデータを用いた混雑状況の可視化

研究の特徴

私たちの研究グループは、必要に応じて、その分野に精通した有識者と共同でデータ分析を進める柔軟性に特徴があります。また、既存データの利活用だけでなく、データ収集段階やその応用システムについて、提案や試作を行える点にも特徴があります。

実用化が想定される分野

観光分野、行政分野、農業分野

研究者からのメッセージ

様々なデータの新たな使い方の創出を目指しています。既存の問題を解決する方法を模索し、新しい展開について、一緒に取り組んでいきませんか? [研究紹介ホームページ] <https://web.wakayama-u.ac.jp/~yoshino/lab/>

研究分野 : ビッグデータ, データ分析, データマイニング

研究者の所属部局・職位・氏名 : 和歌山大学システム工学部 情報学領域・教授・吉野孝

本件に関するお問い合わせ : liaison@ml.wakayama-u.ac.jp