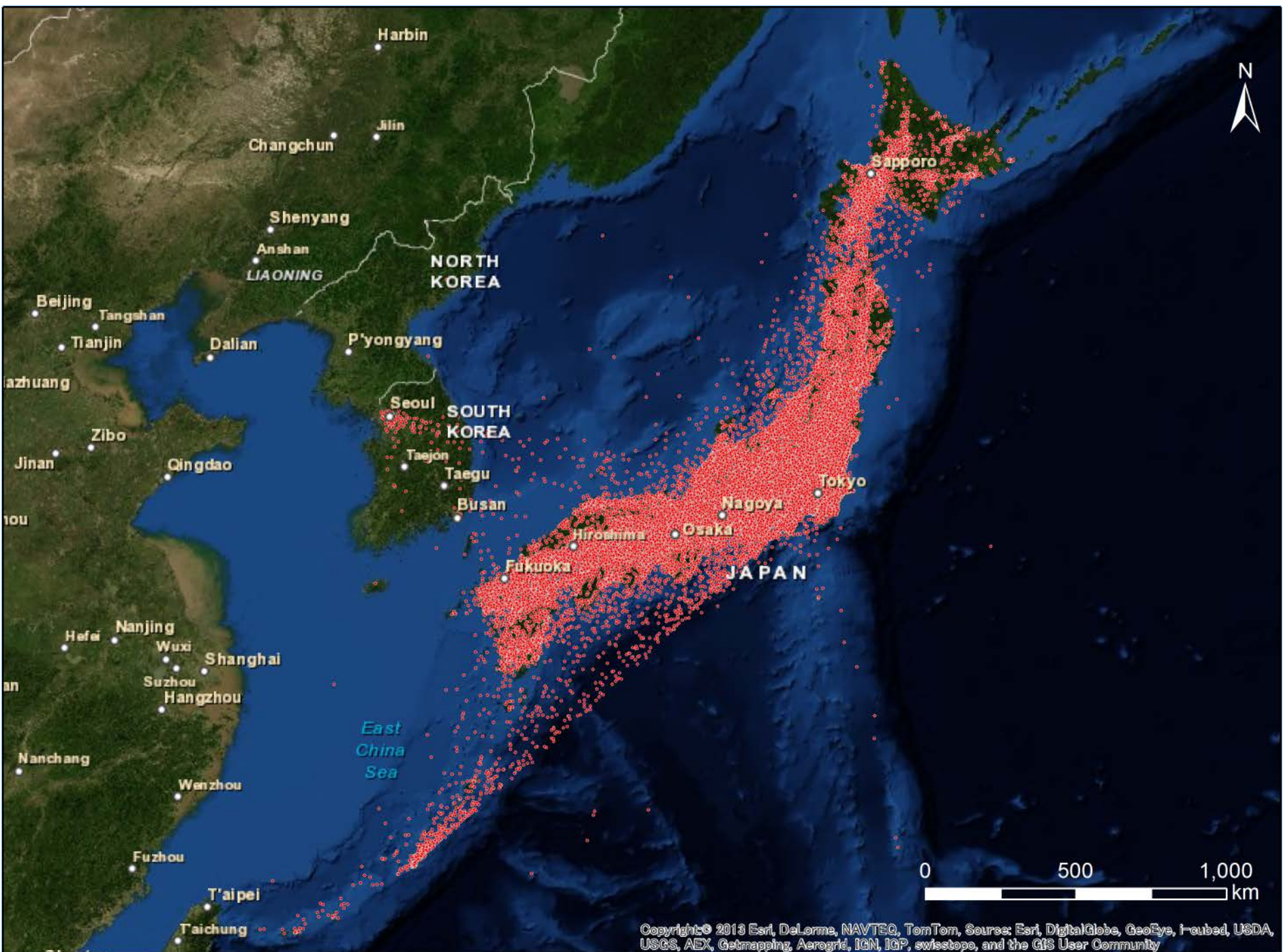


位置情報付きツイッター投稿データ にみるユーザー行動の基本的特徴 —観光行動分析への利用可能性—

桐村 喬

東京大学・日本学術振興会PD

kirimura@csis.u-tokyo.ac.jp



背景と研究目的

- 従来の観光行動の把握方法
 - アンケート調査
 - IC乗車券
 - GPS

} 大規模に調査できない

ツイッターデータは全世界のデータを誰でも入手可能
- 位置情報つきツイッターデータの利活用を図る
 - ユーザーの行動がどの程度わかるのか
 - 観光行動の分析はどの程度可能なのか

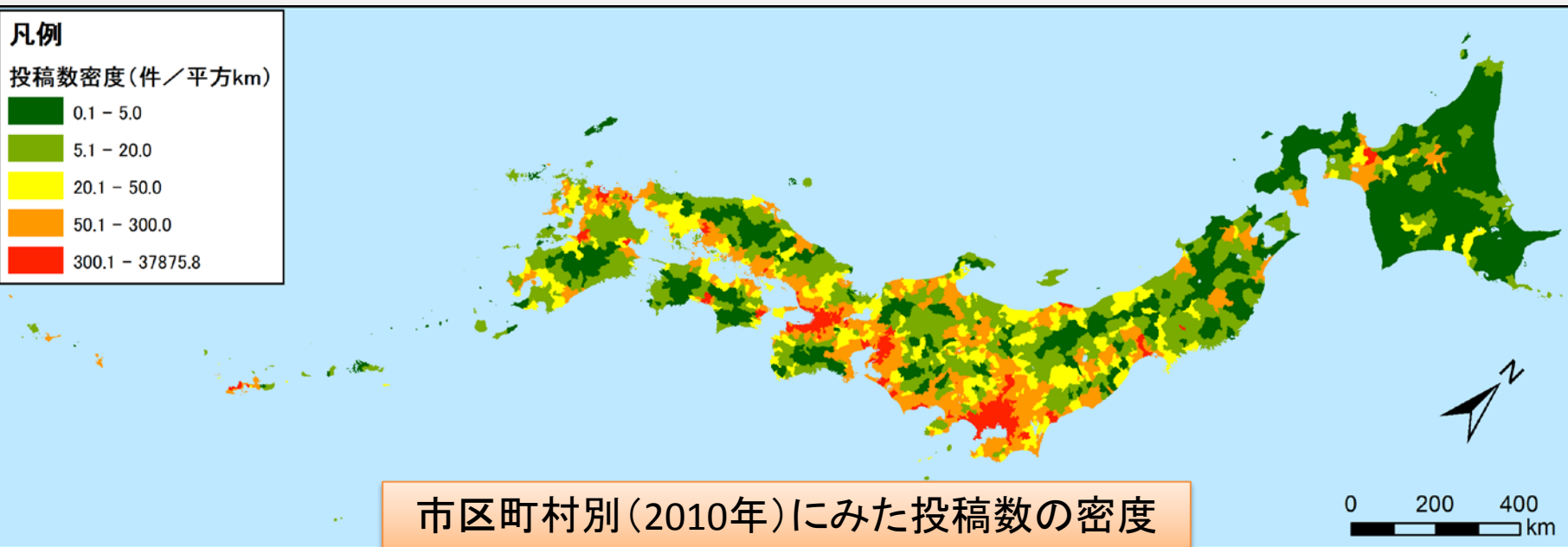
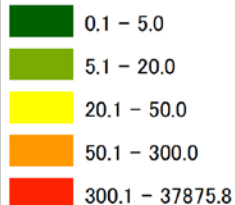
データの入手方法と概要

- データの入手
 - アカウントさえ取得すれば誰でも入手可能
 - 全データを取得することは不可能(サンプリング)
 - ただし、位置情報つきデータの大半は取得できる
(Morstatter et al., 2013)
- **1年4か月分** (2012年2月から2013年6月まで)
 - 日本を含む矩形の範囲 5,600万件
 - 日本国内の陸上 4,500万件

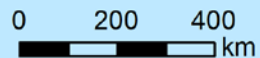
投稿データの位置情報

凡例

投稿数密度(件/平方km)



市区町村別(2010年)にみた投稿数の密度



大都市圏単位でみた投稿数の特徴

| 地域 | | 投稿数 (2012年2月-2013年5月) | 1,000人あたり投稿数 (2010年国勢調査人口) |
|-------|-----|--------------------------|-------------------------------|
| 三大都市圏 | 東京 | 16,603,053 | 454.5 |
| | 名古屋 | 2,763,866 | 303.5 |
| | 京阪神 | 8,034,126 | 415.4 |
| その他 | | 17,990,158 | 285.2 |
| 全体 | | 45,391,203 | 354.5 |

昼間人口1人あたり投稿数の上位20市区町村

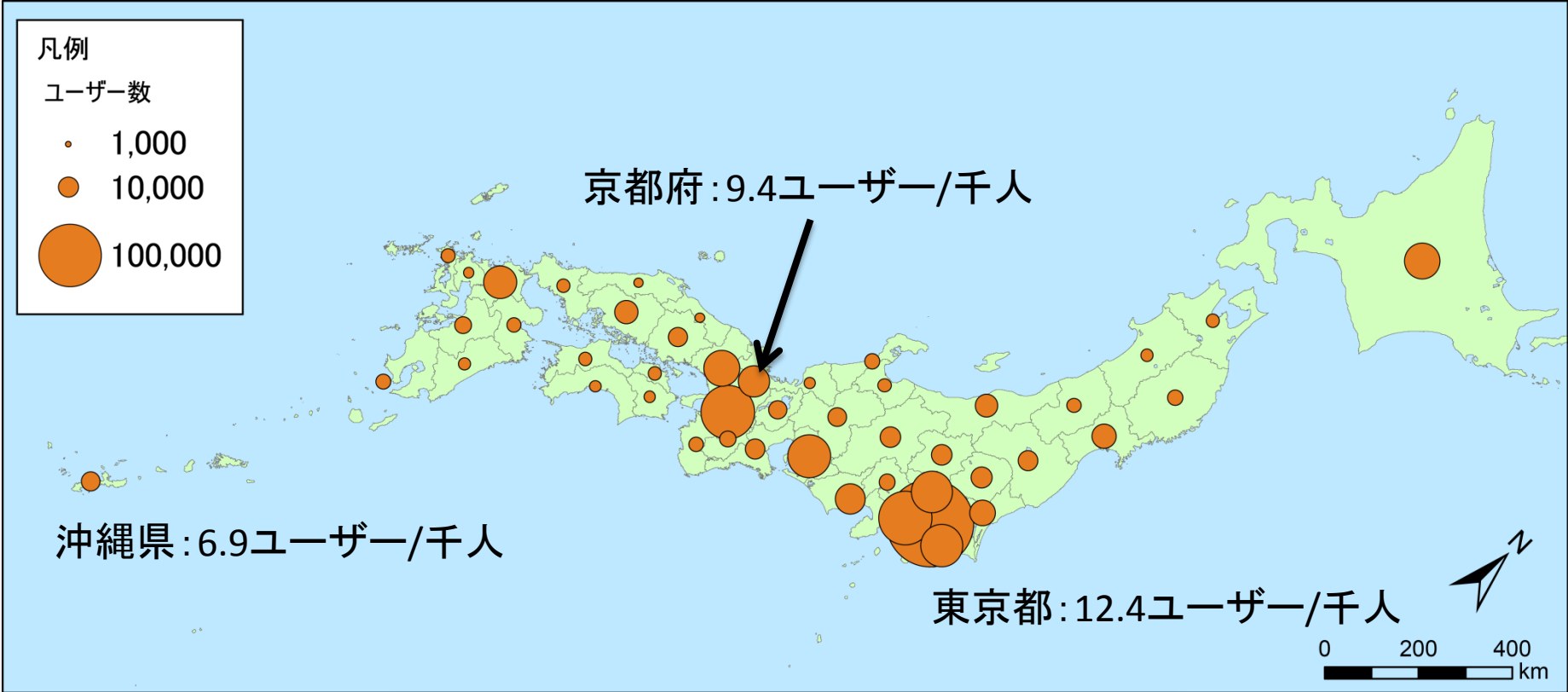
| No | 都道府県 | 市区町村 | 投稿数 2012/2～2013/5 | 昼間人口 2010年 | 昼間人口 1人あたり 投稿数 | 主な施設・観光資源 |
|----|------|------------|----------------------|---------------|----------------------|-----------|
| 1 | 岐阜県 | 大野郡白川村 | 47,232 | 1,772 | 26.655 | 白川郷(世界遺産) |
| 2 | 岩手県 | 西磐井郡平泉町 | 42,598 | 7,373 | 5.778 | 中尊寺(世界遺産) |
| 3 | 京都府 | 京都市伏見区 | 1,078,219 | 276,715 | 3.896 | 伏見稻荷大社 |
| 4 | 新潟県 | 南魚沼郡湯沢町 | 20,705 | 9,468 | 2.187 | スキー場 |
| 5 | 東京都 | 小笠原支庁小笠原村 | 6,217 | 2,844 | 2.186 | 自然(世界遺産) |
| 6 | 福島県 | 西白河郡泉崎村 | 13,464 | 7,556 | 1.782 | |
| 7 | 山梨県 | 南都留郡山中湖村 | 9,215 | 5,669 | 1.626 | 山中湖 |
| 8 | 山梨県 | 南都留郡道志村 | 2,452 | 1,559 | 1.573 | |
| 9 | 沖縄県 | 中頭郡嘉手納町 | 23,605 | 15,887 | 1.486 | 米軍施設 |
| 10 | 北海道 | 勇払郡占冠村 | 2,250 | 1,538 | 1.463 | スキー場・リゾート |
| 11 | 山梨県 | 南都留郡富士河口湖町 | 36,356 | 25,593 | 1.421 | 河口湖・西湖 |
| 12 | 東京都 | 渋谷区 | 738,015 | 520,698 | 1.417 | ターミナル駅 |
| 13 | 沖縄県 | 国頭郡金武町 | 15,438 | 11,056 | 1.396 | 米軍施設 |
| 14 | 長野県 | 北安曇郡白馬村 | 12,534 | 9,017 | 1.390 | スキー場 |
| 15 | 沖縄県 | 中頭郡北中城村 | 20,056 | 14,746 | 1.360 | 米軍施設 |
| 16 | 福島県 | 南会津郡檜枝岐村 | 866 | 666 | 1.300 | 尾瀬 |
| 17 | 沖縄県 | 島尻郡座間味村 | 1,140 | 890 | 1.281 | 離島 |
| 18 | 三重県 | 桑名郡木曾岬町 | 7,694 | 6,070 | 1.268 | |
| 19 | 青森県 | 三沢市 | 53,559 | 42,721 | 1.254 | 米軍施設 |
| 20 | 大阪府 | 大阪市浪速区 | 121,555 | 97,184 | 1.251 | 電気街(日本橋) |

「観光」はユーザーの投稿頻度を高める要因？

ユーザー単位の行動の特徴

ユーザー単位の集計：生活圏の判定

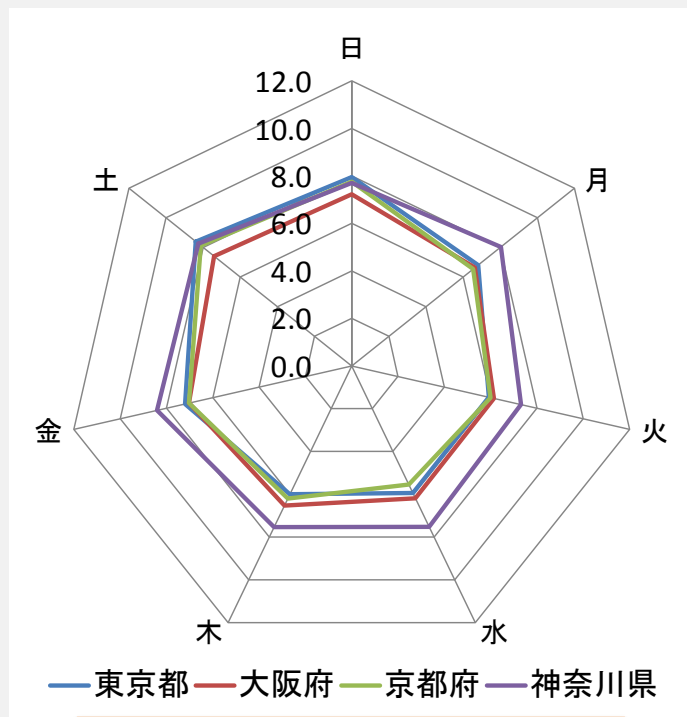
- 「観光」行動の地理的範囲と日常的な生活圏
 - ここでは都道府県単位で考える
- 生活圏の定義
 - 最も投稿数の多い都道府県を生活圏とする
 - 1kmメッシュ(3次メッシュ)単位で都道府県を判定
 - 計算を簡便にするため
- 全110万ユーザーのうち、日本国内を生活圏とするのは87万ユーザー



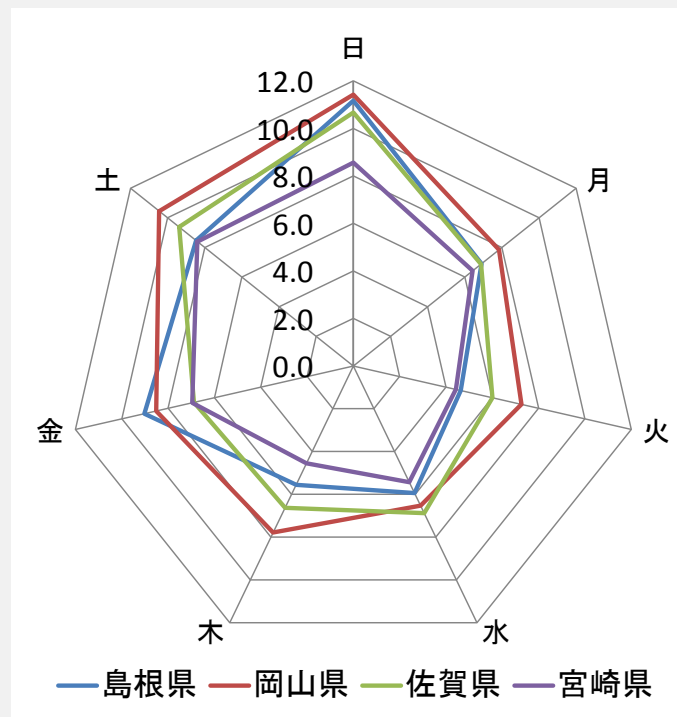
生活圏(都道府県単位)別のユーザー数

ユーザー単位の集計：行動の曜日差

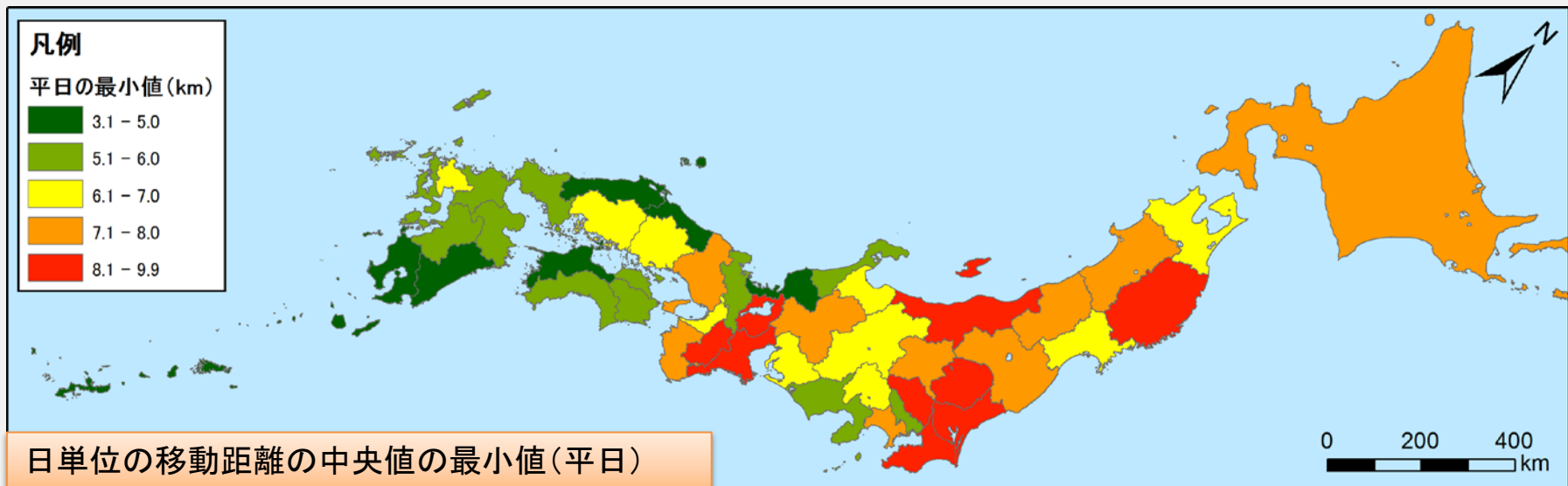
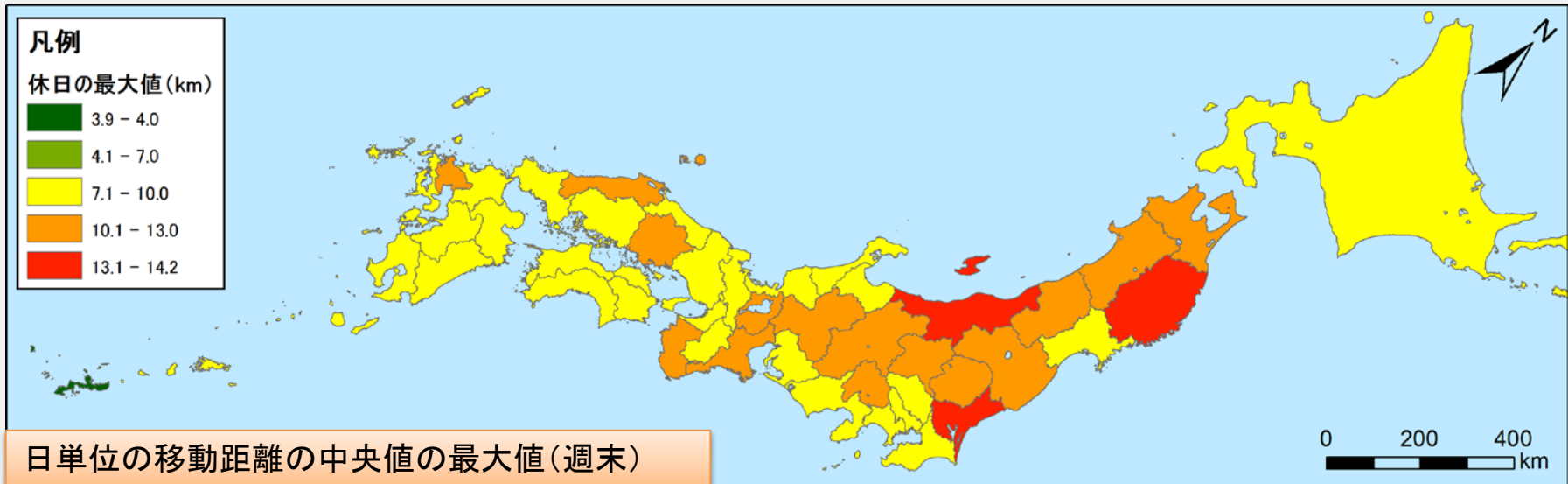
- 11万ユーザーを対象に曜日別の移動距離の中央値を算出

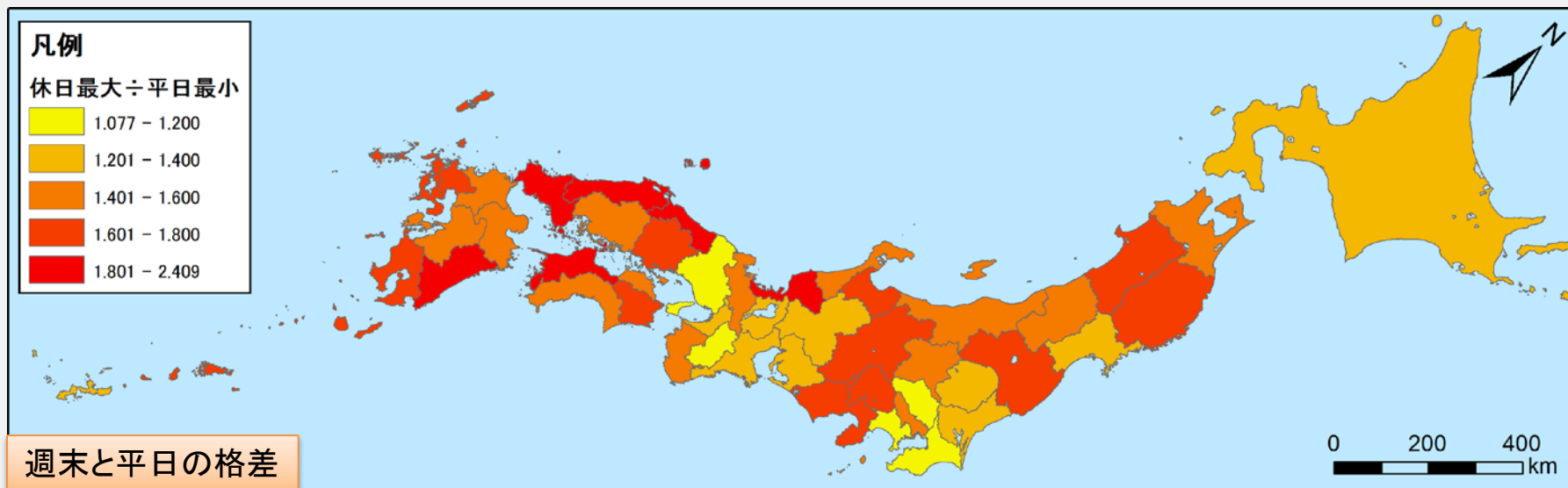


東京・大阪・京都・神奈川
単位 (km)



島根・岡山・佐賀・宮崎
単位 (km)





地域によって差はあるが、週末には行動範囲が大きく広がる

より広い意味での観光行動が投稿データに表れている

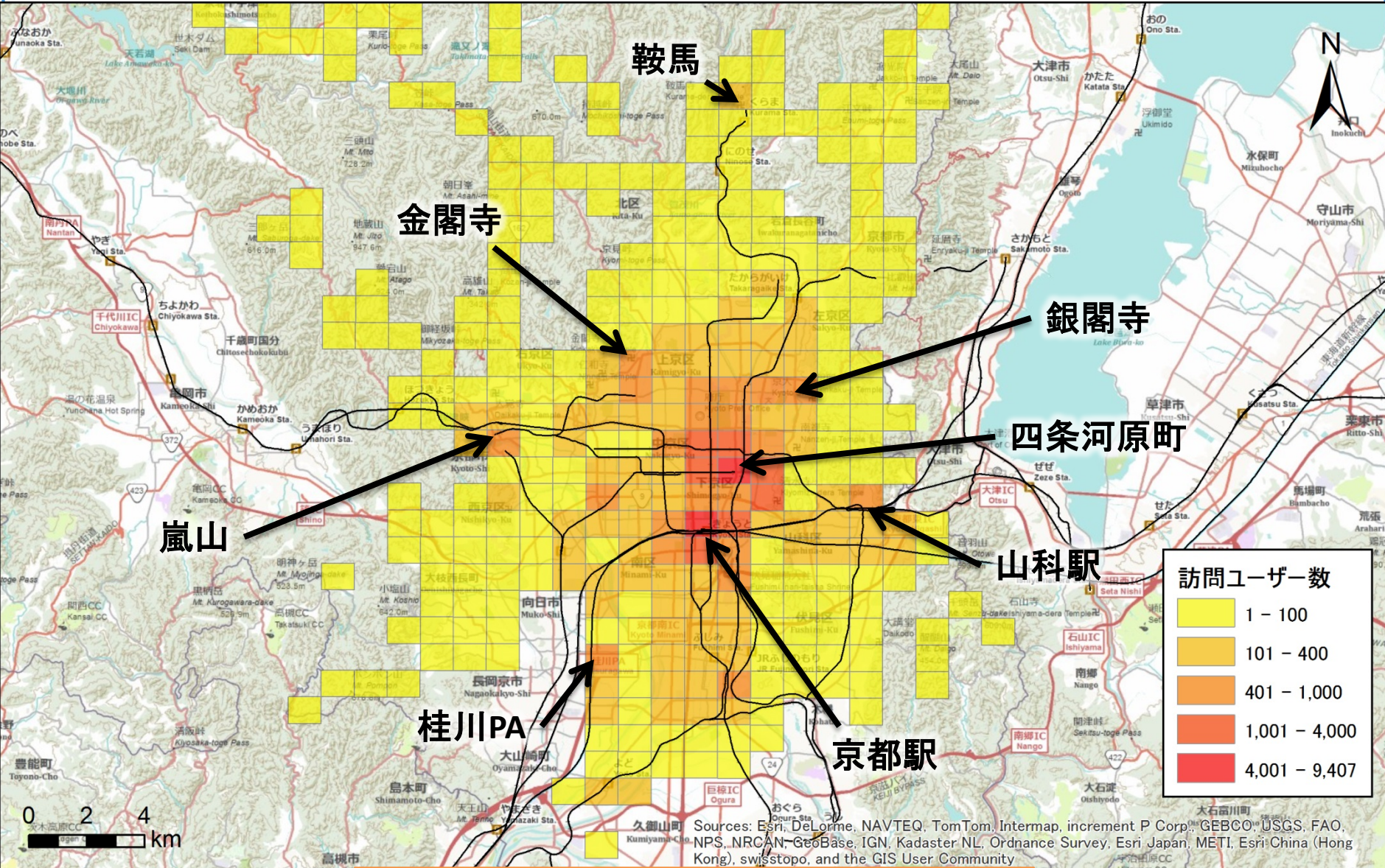
ユーザー行動の分析例ー京都市ー

京都市への訪問ユーザー

- 訪問ユーザーの定義
 - 京都市内で1度でも投稿したことがある
 - 京都市内各区を生活圏としていない
 - 京都市内での投稿数が全投稿の1割以下

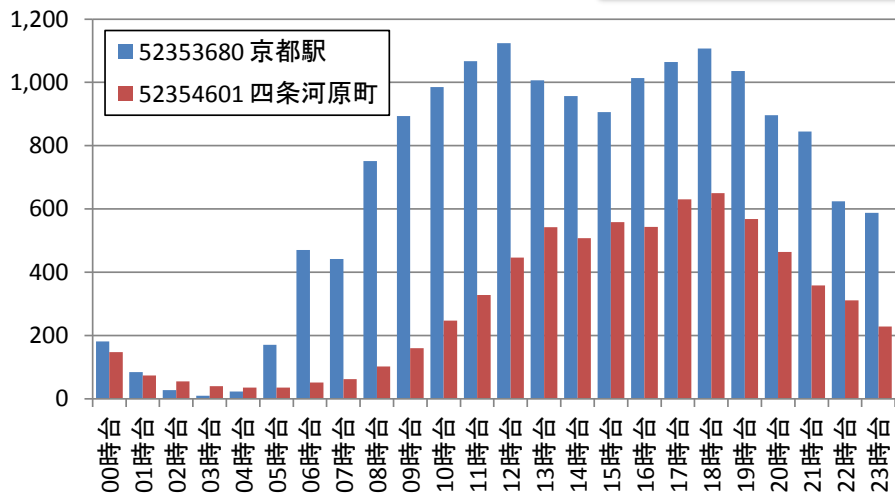
関東地方への偏り

| | 投稿データに基づく 訪問ユーザー (2012年2月～2013年5月) | | 京都市観光調査 (2010年) 入込客 |
|-----|--|--------|---------------------------|
| | ユーザー数 | 構成比 | 構成比 |
| 北海道 | 283 | 1.3% | 1.1% |
| 東北 | 408 | 1.8% | 1.3% |
| 関東 | 8,240 | 36.8% | 14.0% |
| 中部 | 3,136 | 14.0% | 11.5% |
| 近畿 | 8,490 | 37.9% | 61.2% |
| 中国 | 682 | 3.0% | 4.9% |
| 四国 | 331 | 1.5% | 2.4% |
| 九州 | 819 | 3.7% | 3.6% |
| 全体 | 22,389 | 100.0% | 100.0% |

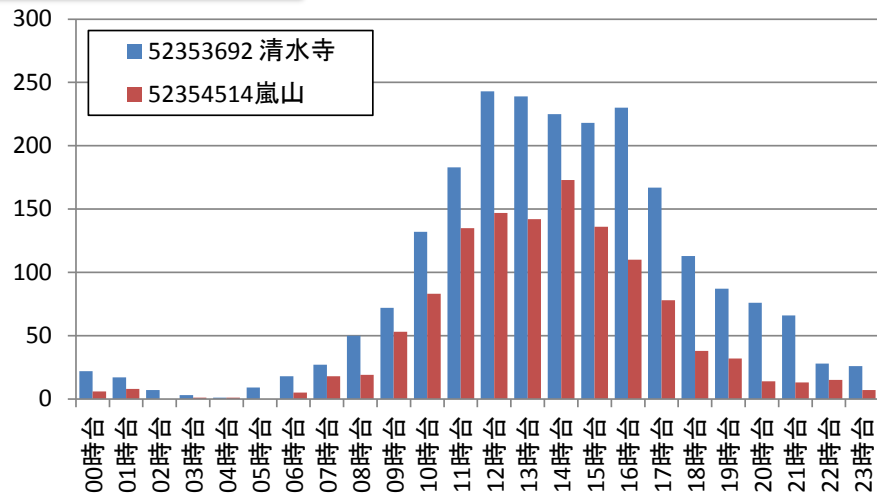


訪問ユーザーの数(3次メッシュ単位)

時間帯別訪問ユーザー数

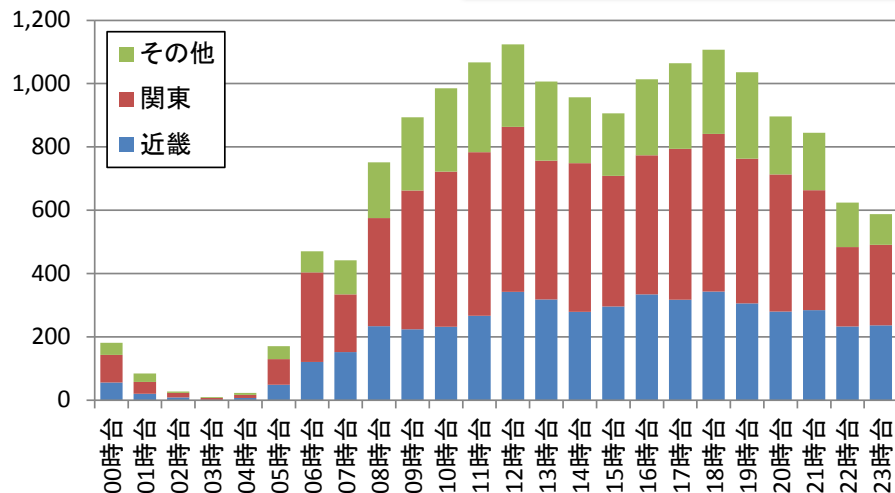


京都駅と四条河原町の比較

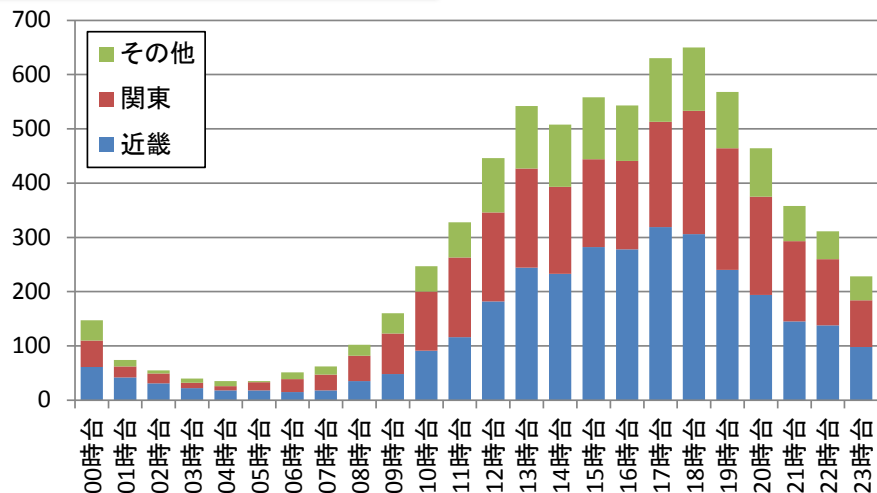


清水寺と嵐山の比較

時間帯・生活圏(地方)別訪問ユーザー数



京都駅



四条河原町

日帰り訪問ユーザーの市内流動

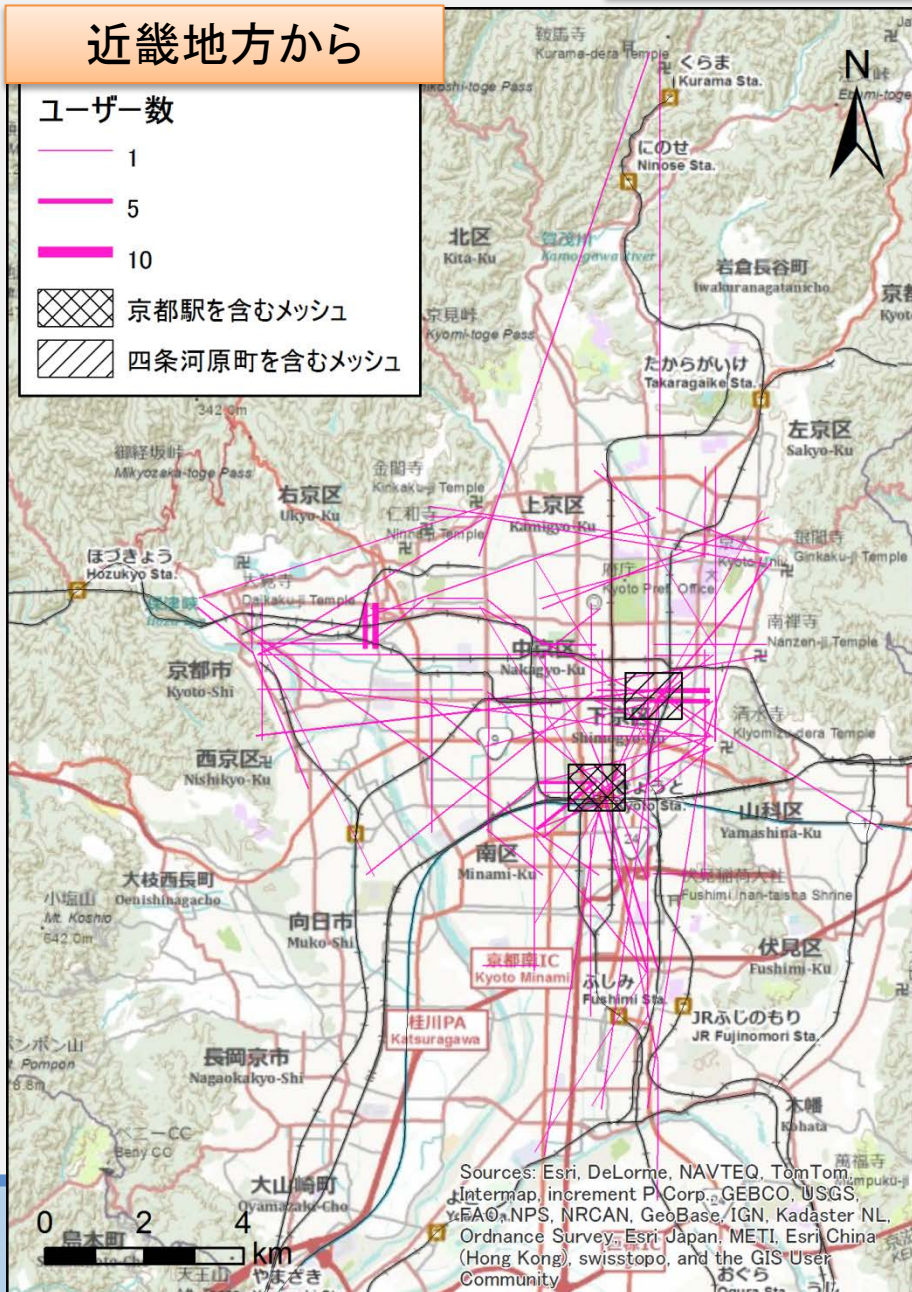
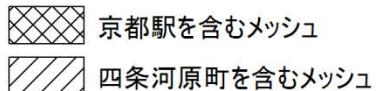
- 日帰り訪問ユーザーの定義
 - 京都市内での投稿が全期間で1日のみ
 - その日に10回以上投稿
 - **180ユーザー**

| 生活圏(地方) | ユーザー数 |
|---------|-------|
| 北海道 | 1 |
| 東北 | 7 |
| 関東 | 86 |
| 中部 | 28 |
| 近畿 | 43 |
| 中国 | 2 |
| 四国 | 4 |
| 九州 | 9 |
| 合計 | 180 |

日帰り訪問ユーザーの行動

近畿地方から

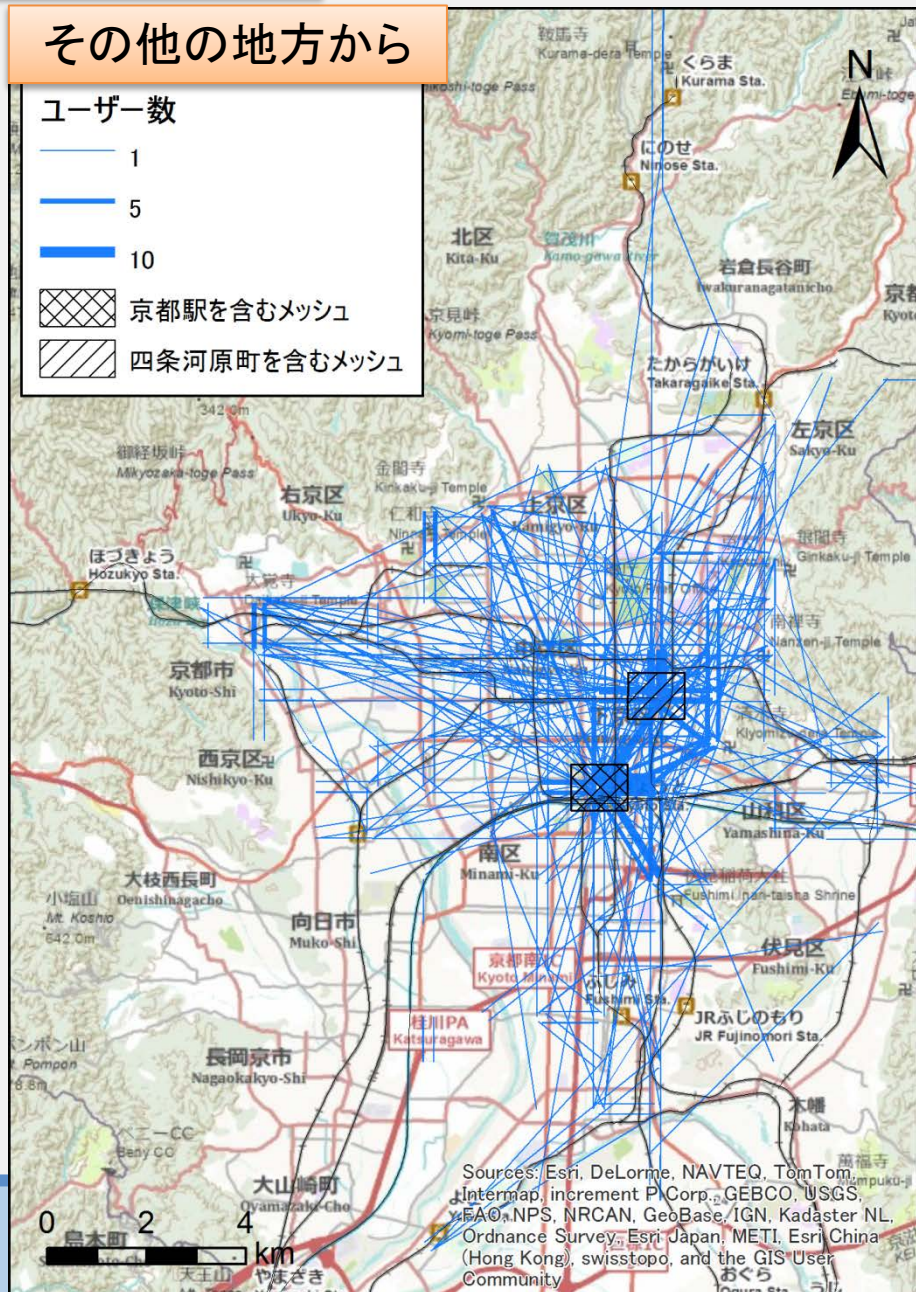
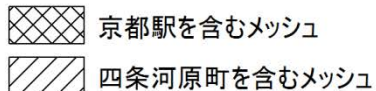
ユーザー数



Sources: Esri, DeLorme, NAVTEQ, TomTom, Intermap, increment P Corp., GEBCO, USGS, FAO, NPS, NRCAN, GeoBase, IGN, Kadaster NL, Ordnance Survey, Esri Japan, METI, Esri China (Hong Kong), swisstopo, and the GIS User Community

その他の地方から

ユーザー数



Sources: Esri, DeLorme, NAVTEQ, TomTom, Intermap, increment P Corp., GEBCO, USGS, FAO, NPS, NRCAN, GeoBase, IGN, Kadaster NL, Ordnance Survey, Esri Japan, METI, Esri China (Hong Kong), swisstopo, and the GIS User Community

まとめ

- ユーザー行動の基本的な特徴
 - 日常的な生活圏は二大都市圏に偏る
 - 行動範囲は週末に広がる
- 京都市を事例としたユーザー行動の分析
 - 訪問数は、観光統計よりも二大都市圏に偏る
 - 観光行動として解釈できる妥当な結果を確認
 - 時間別・生活圏別の訪問ユーザー数の把握
 - 訪問ユーザーの生活圏別の行動パターンの違い

位置情報付き投稿データの有用性

- 誰でも入手でき、分析できる
 - Twitterのアカウントと少々プログラムが必要
- ユーザー単位で集計すると行動が見えてくる
 - 各地への訪問履歴などもある程度把握できる
- 投稿内容など、場所に関する情報も多い
 - 今回は分析できなかった

課題

- ユーザーの属性
 - 年齢、性別、職業、行動目的、同行者数…
 - ⇒ 投稿内容やユーザー情報の解析が助けになる？
- 位置情報の精度、全行動に対する網羅性
 - GPSなどの性能、プログラムによる嘘の位置情報
 - 投稿数が少ないと、行動は把握できない
 - ⇒ 投稿数が一定数あるユーザーを対象にするしかない
- その他、整理・対処すべき課題は用途によって様々

参考文献・資料

- 京都市(2010):『京都市観光調査年報』
- 杉本興運 (2012): 観光者の視覚的体験情報に基づく回遊空間の評価, GIS-理論と応用, 20(1), 39-50.
- 矢部直人・倉田陽平 (2013): 東京大都市圏におけるIC乗車券を用いた訪日外国人の観光行動分析, GIS-理論と応用, 21(1), 35-46.
- Morstatter, F., Pfeffer, J., Liu H. and Carley, K. M., 2013. Is the Sample Good Enough? Comparing Data from Twitter's Streaming API with Twitter's Firehose. Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media, 400-408.