

クラウドソーシングを活用した歩行者通行量推定技術の開発

秋山祐樹*、仙石裕明*、上山智士**、中西航***、笹尾知世****、西村隆宏*****、杉田暁*****

*東京大学空間情報科学研究センター、**東京大学地球観測データ統融合連携研究機構

東京大学大学院工学系研究科社会基盤学専攻、*東京大学大学院新領域創成科学研究科

*****株式会社リクルートホールディングス、*****中部大学国際 GIS センター

1. はじめに

我が国では全国の道路交通量の把握のために、全国道路・街路交通情勢調査（道路交通センサス）の一環として、「一般交通量調査」を行い、日本全国の道路の交通量と道路現況を収集してきた。同調査で収集されたデータは道路の計画・建設・維持修繕その他の管理などについての基礎資料を得ることを目的に実施されてきた。また日本全国の自治体では「商店街通行量調査」も実施されている。同調査は特に商店街とその周辺地域における「歩行者通行量」の調査に力を入れており、中心市街地・商店街の顧客の吸引力の現況把握と商店街活性化の資料とすることを目的に実施されている。

以上のように通行量調査は道路の開発・維持管理や中心市街地・商店街の活性化支援のための政策立案を行う上で重要な役割を果たしてきた。しかし現状の通行量調査には以下の課題がある。

①調査を実施する際の労力が大きいこと

現状の手法では調査地点を複数設定し、それらの地点に調査員を調査の全期間に渡って配置し続けることで、データを収集し続ける必要がある。そのため広域で大規模な調査を連続的に実施することは物理的に非常に困難である。

②更新頻度を上げることが困難なこと

①でも指摘した通り、これまでの調査方法は労力が大きいため、同調査を高頻度を実施することは困難である。実際に既存の調査は特定の日時を対象としているため、通行量の曜日・季節での違いや、平時とイベント発生時（例：お祭りの開催・交通障害や災害の発生など）の違いなどの把握が十分に出来ない。

③公開時のデータフォーマットが統一されていないこと

多くの自治体で公開されている通行量調査の結果は PDF 形式であり、また調査地点ごとに調査シートが分かれている場合もある。そのため分析を始める前にまずこれらのデータを利用可能なフォーマット（Excel・csv・dbf 形式等）に整形する必要がある、同データを用いた研究に着手する際の障壁となる。

このような課題に対して、近年では携帯電話の GPS ログや基地局との通信記録を用いて、大量の人々の流動・滞留を把握する試みも始まっており、同データを用いた商店街の時間別滞在者数を推定しようとする試みも見られる（文献：1）。しかし携帯電話の位置情報は個人情報保護の観点から利用制約が多く、当面は同データを公開・流通させることは困難であると考えられる。

では歩行者通行量を「低コスト」かつ「高頻度」に把握するにはどのような方法が考えられるだろうか。この課題に対して本研究では、スマートフォンなどのデバイスを用いて「クラウドソーシング」により大量の歩行者通行量情報を「離散的に収集」し、収集した離散的データを平滑化して連続的なデータに加工することで「任意の時間断面の歩行者通行量を連続的に推定」する手法の開発を目的とする。

クラウドソーシングを活用した歩行者通行量や店舗等の滞在者数を取得する試みは東田・櫻木（文献：2）で実施されているが、同研究はデータ取得方法が中心であり、データの平滑化手法に関する議論や平滑化に堪えるデータ量の検証は不十分である。また離散的な滞在者数から全体を補間・推定する試みは森田・須川（文献：3）や笠原・島川（文献：4）などで行われているが、いずれも信頼性の高い結果を得るにはより多くのデータ収集が必要であり、その実現には多大な労力・時間がかかる点が課題であるとしている。クラウドソーシングはこの「多くのデータ収集」という課題を解決出来る可能性がある技術といえる。

なお研究代表者ほかはこれまでに店舗の滞在者数については、既に離散的に収集した滞在者数を平滑化することで任意の時間断面の店舗滞在者数をある程度推定できることを確認している（文献：5）。そこで本研究では同手法を応用・発展させて歩行者通行量にも応用することを試みた。特に本研究では研究代表者ほか

の既存研究（文献：5）で実施されていなかったクラウドソーシングを活用して歩行者通行量等を推計するにあたり、必要な「計測時間」を明らかにすることに取り組んだ。また本研究で実施する通行量調査にあたっては、自治体・商工会議所が実施する既存の通行量調査との連携を行うことで全数調査を実施し、同調査結果との比較を行うことで本研究の手法による推定結果の信頼性を明らかにした。

2. 方法

本研究では歩行者通行量の計測において、専用のアプリを開発し、同アプリを用いて調査員が1時間に2～3回、1回5分間の通行量および入店者情報等を投稿することで、データの収集を行った（図1）。

既存の手法（文献：5）では計測した一定期間の通行量から1時間分の通行量を拡大推計していたが、本研究では通行量計測における必要な計測時間を確かめることが目的であるため、平滑化等を実施せず単純な倍数により通行量を推計した。例えばある地点において「11:00～11:05」の通行量が39人の場合、「11:00～11:30」は39(人)×6=234人とし、「11:30～11:35」の通行量が44人の場合、「11:30～12:00」は44(人)×6=264人とし、合計(11:00～12:00)498人の通行量があったと推計する。

そして推計した結果を自治体が実施した全数調査結果との比較することで、以上の推計方法の精度および必要な計測時間を明らかにする。

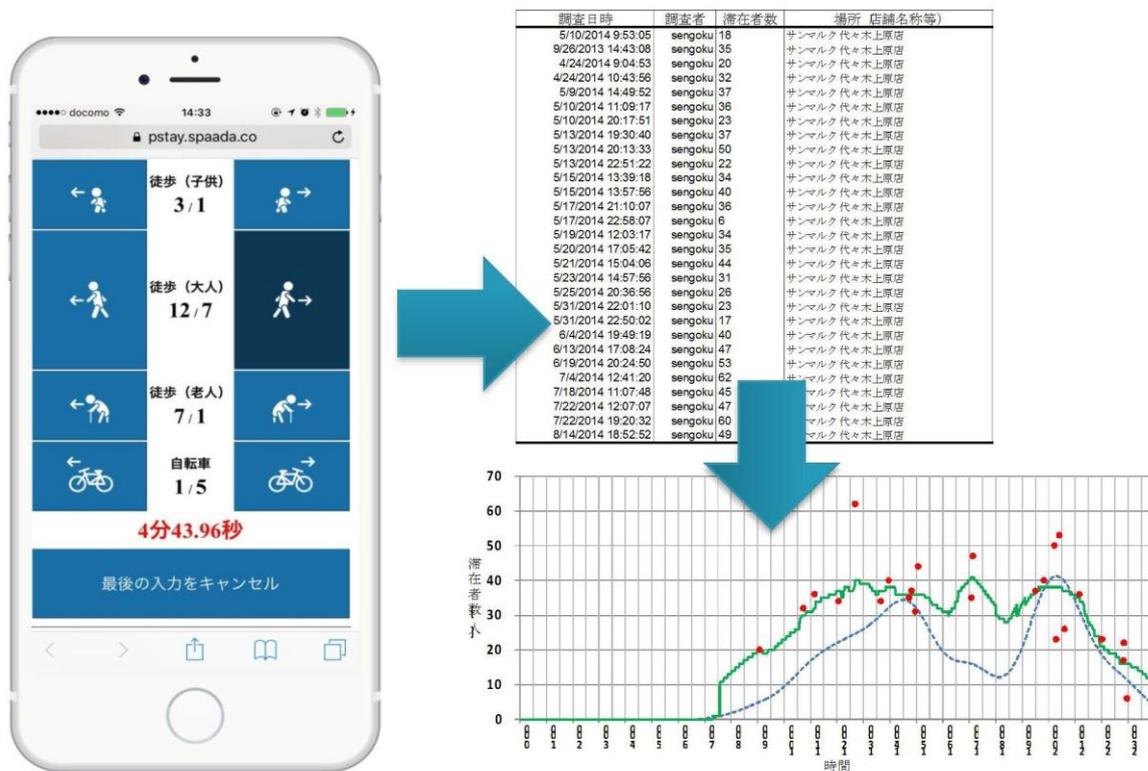


図1 歩行者計測方法と推計の流れ
(通行量計測アプリは株式会社マイクロベースが開発・提供)

3. 結果

本研究では下記の2回に分けて調査を実施した。

3.1 第1回通行量調査（茨城県水戸市）

同調査は茨城県水戸市において2015年7月5日（日）・6日（月）の2日間を対象に実施された。調査地点は水戸市内の12地点とした（図2）。また専用アプリによる計測時間は5分間に固定した。なお同調査を実施する際、水戸市役所の協力により同市の通行量調査に並行して同調査を実施し、同市の全数調査結果と比較することで検証を行った。



図2 第1回通行量調査の調査対象地点（茨城県水戸市）

同調査から得られた結果を図3に示す。水戸駅前の中村ビル前（No.2）では日中の時間帯は推計結果と水戸市による全数調査結果が近い値となることが確認できる。一方、歩行者通行量の変動が大きいとみられる帰宅ラッシュの時間帯（18時-19時）では誤差が大きい傾向にある。水戸駅から約1.7km離れた地点にある旧京成百貨店前（No.11）においても同様の傾向が確認された。同調査の対象外となっている7時から9時にかけての通勤ラッシュ時においても同様の傾向となるものと考えられる。

以上の結果から、比較的乗降客数の多い駅やバス停周辺のような局所的な変化が頻繁に起こる地点では、より詳細な通行量推移を把握する必要があることが分かった。そこで同調査では計測時間を5分間に固定していたが、この結果を踏まえて通行量をカウントした時点の時刻を保持することで、自由に再集計・検証できるように調査方法を改善した。

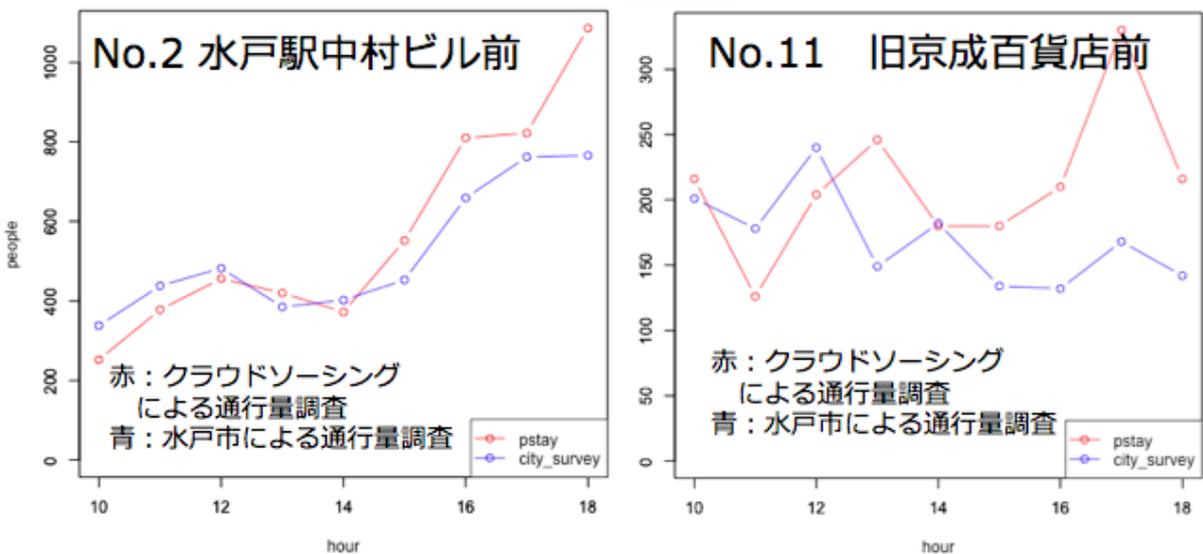


図3 第1回通行量調査における推計結果（茨城県水戸市、7月6日 10:00~19:00）

3.2 第2回通行量調査（東京都立川市）

同調査は東京都立川市の JR 立川駅北口において 2015 年 11 月 30 日（月）・12 月 6 日（日）・12 月 14 日（月）・12 月 19 日（土）の 4 日間を対象に実施された。調査地点は JR 立川駅北口側の 4 地点とした（図 4）。また同調査では茨城県水戸市の実験結果から得られた成果に基づき、専用アプリによる計測時間の固定は行わないものとした。なお同地域は大規模商業施設「ららぽーと立川立飛」が 2015 年 12 月 7 日（月）に開店したため、開店前後で通行量が変わる可能性が高いと考えられる地点である。

調査地点 1（サンサンロード）における全数調査の結果を表 1 に示す。同地点は立川駅北口と立川市役所・立川税務署等の行政機関、IKEA 立川をつなぐ歩行者道となっており、人通りが多い地点である。同地点において 1 分、3 分、5 分、8 分おきにアプリで収集した通行量を再集計し、1 時間に拡大推計を行った結果と全数調査結果との 2 乗平均平方根誤差（RMSE）をとった値を図 5 に示す。なお、スミノルフグラブス検定を行い、外れ値となる観測値（図 5 中の青い点）については、異常値として外すこととした。この結果から計測時間が長くなるにつれて、誤差が小さくなっていくことが確認できる。同地点の場合では計測時間が 5 分になった時点で、8 分の場合と差が少なくなることが分かる。計測時間と推計精度はトレードオフの関係にあると言え、同結果からその程度を定量的に示すことができた。



図 4 第2回通行量調査の対象地域

表 1 調査地点 1 サンサンロードの全数調査結果

歩行者 (大人)	ららぽーと開店前	
	12/6 (日)	
	晴れ	
	IKEA 方面へ	IKEA 方面から
9時 - 10時	108	123
10時 - 11時	220	194
11時 - 12時	207	168
12時 - 13時	344	314
13時 - 14時	1123	772
14時 - 15時	565	382
15時 - 16時	586	512
16時 - 17時	524	546
17時 - 18時	201	317
18時 - 19時	233	320
合計	4111	3648

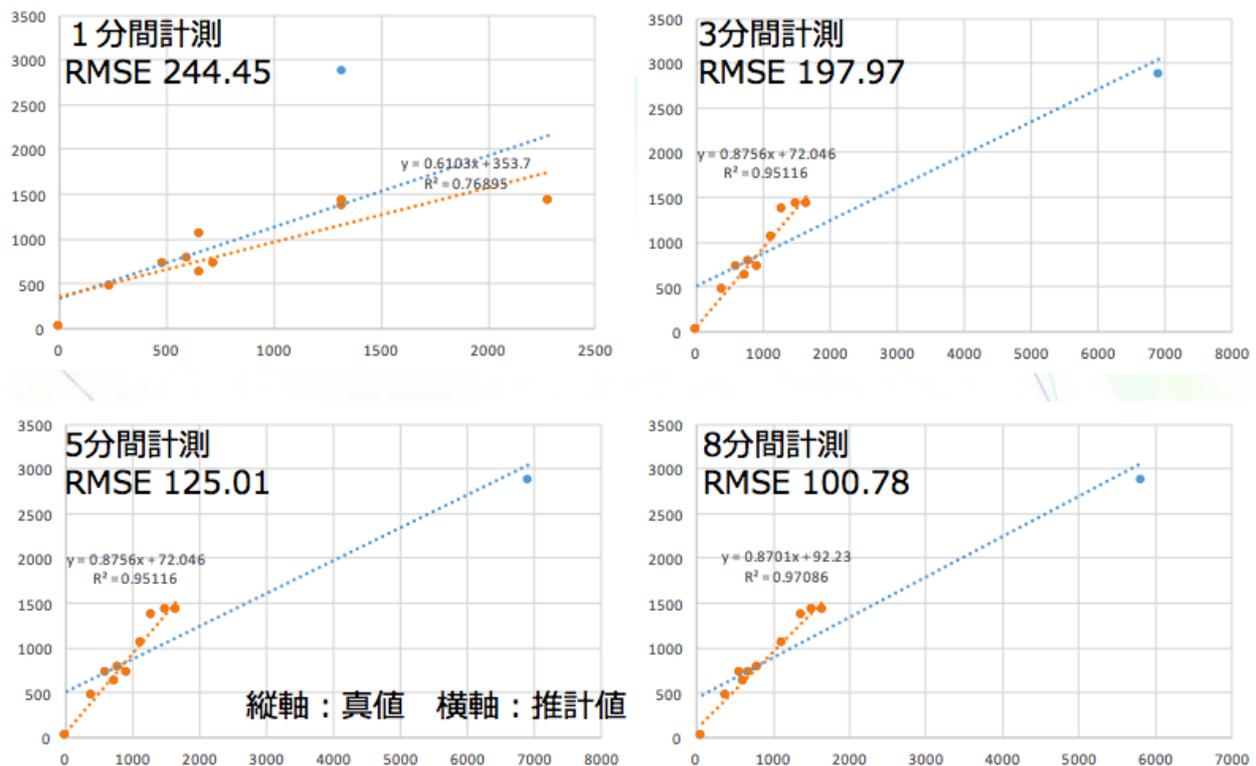


図5 調査地点1における計測時間帯別2乗平均平方根誤差 (RMSE)

4. 考察とまとめ

本研究ではクラウドソーシングを活用した歩行者通行量を推計するにあたり、必要な計測時間および推計精度について調査・分析を行った。その結果、局所的に通行量が急増する地域については、モデルから除外することで通行量を精度よく推計出来る可能性があることが分かった。一方、局所的な変化が少ない地域においては、計測時間が少ない場合でも通行量推計ができることが確認された。局所的な変化が発生すると考えられる地域を事前に想定できれば、少ない計測時間で対応できる可能性がある。

今後は今回調査した地点データを中心に地域の特性を分類し、分類ごとの必要な計測時間の関係性を明らかにしたいと考えている。また調査時には子供・大人・高齢者などの歩行者の属性や、自転車での移動も計測していた。そこで今後はこれらの情報も考慮した推定方法を検討することで、推定結果の信頼性向上に繋がりたいと考えている。

同技術は将来的にはクラウドソーシングによるリアルタイムなデータ収集が可能になるため、究極的には通行量や滞在者数のリアルタイムな把握が可能になる。様々な「ビッグデータ」が様々な「デバイス」により「リアルタイム」に収集され、蓄積され続けていく時代はそう遠くない未来に到来するものと考えられる。本研究はそんな「リアルタイムビッグデータ」時代の到来を見越した萌芽的研究であると考えている。本研究が今後さらに発展し、「デジタルアース」と連携することで、こうした時代の到来を見越したデータの処理・可視化・配信の技術開発を加速・発展させていけるものと期待したい。

5. 謝辞

本研究は中部大学問題複合体を対象とするデジタルアース共同利用・共同研究 IDEAS201507 の助成を受けたものです。また茨城県水戸市には同市における通行量調査でお世話になりました。ここに記して謝意を表します。

参考文献・データ

1. 秋山祐樹・Teerayut Horanont・柴崎亮介, ”大規模人流データを用いた商業地域における来訪者数の時系列分析”, 第22回地理情報システム学会講演論文集, C-5-4, 2013.

2. 東田圭介・櫻木伸幸, ”クラウドソーシングを用いた POI 情報収集”, 電子情報通信学会技術研究報告. AI, 人工知能と知識処理 111(447), 17-19, 2012.
3. 森田孝夫・須川慎哉, “小売市場等における買物客数と最大滞留人員の推定方法に関する研究”, 日本建築学会学術講演梗概集. E, 建築計画, 農村計画 1986, 679-680, 1986.
4. 笠原太郎・島川陽一, “徒歩交通量を推定する GIS の実装の検討”, 高等専門学校の教育と研究 : 日本高専学会誌 13(2), 43-48, 2008.
5. 秋山祐樹・仙石裕明・西村隆宏・桑田賢太郎・朱山裕宜・柴崎亮介, ”クラウドソーシングを活用した店舗等の滞在者数推定の試み”, 第 23 回地理情報システム学会講演論文集, E-5-1 2014.