

途上国における社会経済インフラの計画・運営
管理にかかるビッグデータの活用手法：
都市交通インフラ計画を例として

2021年12月24日

高橋君成、有田禎之

1. 研究の概要

研究タイトル 途上国における社会経済インフラの計画・運営管理にかかるビッグデータの活用手法：都市交通インフラ計画を例として

研究者 高橋君成、有田禎之、佐々木邦明教授（早稲田大学理工学部）

研究の背景と目的 (1)途上国の都市交通インフラ計画の課題

- ・急速な都市の成長と交通問題の顕在化
- ・自然災害・新型コロナ等社会経済活動への多大な影響
- ・マスタープランの陳腐化、予測の困難さ

(2) JICAによるマスタープラン調査の課題

- ・サンプル調査と母集団拡大をベースにした旧来型調査
- ・高い調査費と自主財源による更新が困難

(3) 交通インフラ計画へのビッグデータ活用の可能性と課題

- ・安価なローデータ
- ・ローデータの信頼性、有効性の高いモデル構築の課題

⇒都市交通インフラを例にして、インフラの計画、整備・運営管理に利活用可能なビッグデータの特定・構築やモデル開発を行うことを目的とする

1. 研究の概要

研究活動の内容

【活動1】 全世界における都市交通インフラの計画にかかるビッグデータの利活用の実態把握

【活動2】 ラオス及びカンボジアの都市交通マスタープランの概要とベースラインデータの調査手法（家庭訪問調査、交通量調査等）、データの質、モデルの妥当性の検証

【活動3】 ラオス及びカンボジアにおける都市交通インフラの計画に資するビッグデータ（携帯位置情報、交通管制データ、バス運行データ等）の特定とベースラインデータの構築及びモデルの有効性の検証

【活動4】 都市交通インフラの計画におけるビッグデータの活用の主流化にかかる検討（技術面や費用面での実現可能性調査、コロナ禍における副次的効果）

期待される成果

- ・ 途上国においてインフラ計画・運営管理とその持続性の向上に資する安価で信頼性の高いビッグデータの利活用の促進
- ・ 同手法のJICA業務への適用による競合他社との技術面での差別化、優位化
- ・ 研究技術協力、現地調査協力に関わる研究機関とのネットワークの構築とJICA業務等への協働参画と研究成果の展開

2. 研究の進捗（主に研究活動3）

- 1) **GPSデータ入手** ラオス、カンボジアの3か年（2019～2021年各3か月）の携帯端末のGPSデータを入手
- 2) **GPSデータ基礎分析** ラオスのGPSデータの分析を行い、その特性を把握
- 3) **文献レビュー** 我が国、先進国、途上国のデータ活用事例や活用手法にかかる文献をレビュー
- 4) **GPSサンプルデータ解析** カンボジア、ラオスのGPSデータを活用し、トリップデータへの変換手法を検証
- 5) **GPSデータ・モデル精度検証** 同国のGPSデータから得られたトリップデータとサンプル調査（家庭訪問調査）によるトリップデータを比較し、データやモデルの精度を検証

(図解) 従来型交通調査 (上段) とビックデータ (下段) によるトリップデータ



3. 研究成果（GPSデータ入手）

- ・ GPSデータのソースは携帯アプリの位置情報
- ・ サプライヤは有象無象

| 会社名 | 国 | 保有期間 | 端末数 | 費用* |
|------------|-------|------|-------|-------------|
| Lifesight | 106カ国 | 1年間 | 500百万 | 3,000米ドル/月 |
| Quadrant | 170カ国 | 2年間 | 500百万 | 2,000米ドル/月 |
| Onemata | 226カ国 | 1年間 | 700百万 | 11,000米ドル/月 |
| Datastream | 163カ国 | 2年間 | 不明 | 15,000米ドル/月 |

- ・ 携帯アプリの位置情報の他、個人属性も提供するサプライヤ有
- ・ その他、位置情報の分析（マーケティング他）もサービス提供
- ・ 契約には1）目的外利用禁止、2）二次利用禁止などの条項
- ・ 本研究ではQuadrant社（シンガポール）からGPSデータを購入（1,000ドル×9か月×50%割引=4,500ドル）

（参考）プノンペン都市交通情報収集確認調査では、約5,000万円で交通調査を実施

3. 研究成果（GPSデータ入手）

- 主な情報は携帯ID、座標とその時間
- 携帯の複数のアプリから位置情報を収集
- 個人情報や入手先アプリ情報は秘匿

| Data Field | Type | Description |
|---------------------|--------|--|
| Device_ID | String | Unique advertising identifier from the device (un-hashed) |
| ID_Type | String | Advertising ID type: IDFA (iOS) and ADID (Android) |
| Latitude | Float | Latitude of the event |
| Longitude | Float | Longitude of the event |
| Horizontal_Accuracy | Float | GPS accuracy in meters as reported by the device OS |
| Timestamp | Long | Unix timestamp of the event (milliseconds) since “the epoch” |
| IP_Address | String | IP address of the event generated from the device |
| Device_OS | String | Operating system of the device: "iOS" or "android" |
| Device_OS_Version | String | Version of the device operating system |
| User_Agent | String | Web browser version and operating system |
| Country_Code | String | ISO2 2-digit alpha country code for the event |
| Source_ID | String | Quadrant unique identifier for the data source |
| Publisher_ID | String | Unique developer identifier |
| App_ID | String | Unique application identifier |
| Location_Context | String | 0 = foreground, event captured when the app is open 1 = background, event captured when the app is not open |
| Geohash | String | Unique alphanumeric string to express a location |

3. 研究成果（GPSデータ基礎分析とデータ特性） ～ラオスのGPSデータを例として～

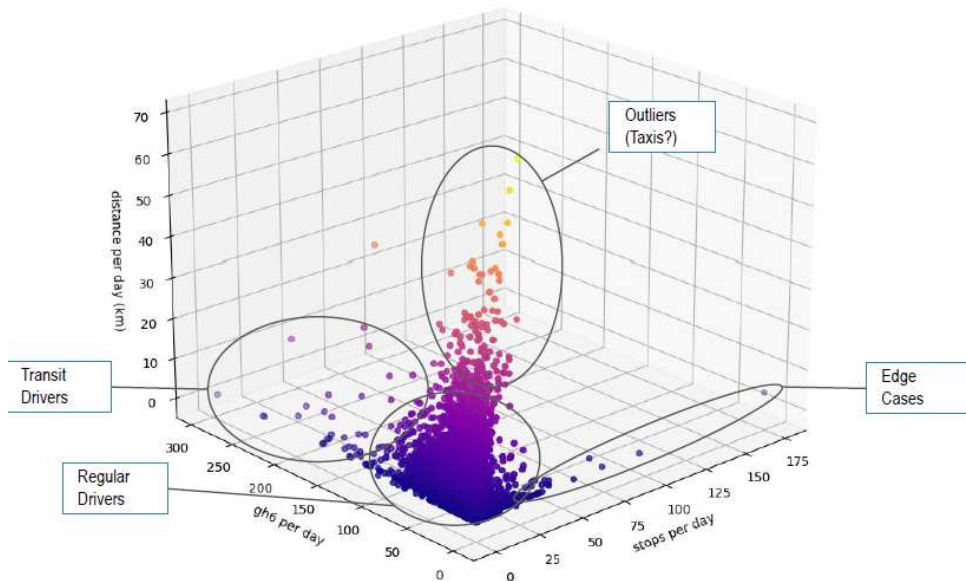
| | |
|----------------------|---|
| データ数 | 端末数（日当たり9,000～15,000）、レコード数（日当たり約300,000） * 2021年6月の調査日3日間 |
| 経年変化 | 3カ年の同月の端末数は安定しておらず、経年変化を分析するには不適 |
| 曜日変化 | 曜日による端末数も、倍半分になる日が非周期的・突発的に現れ、曜日変動の解析を行うには精度が不安 |
| 記録回数 | 一日の記録回数が低頻度（1ないし2回/日）の端末が端末全体の3～4割。100回/日を越える端末も5～8%存在。データの精度（ここでは位置情報記録回数）のばらつきをどう処理するかが課題の一つ。 |
| トリップへの変換とその処理 | 記録回数が十分な端末については、移動の開始地点と目的地以外の位置情報記録地点（GPSの誤差や敷地内の移動など）が多いため、これらを判別し排除する作業の自動化が必要 |

3. 研究成果 (GPSサンプルデータ解析) ～カンボジアのGPSデータを例として～

GPSデータの処理方法

過去のGPSデータのノイズの処理の類似業務経験から (1) 重複データ、(2) 特異なスピード・移動回数、(3) データのスミージングにより、クリーンデータを作成

James教授 (ノッティンガム大学)



Summary of whole Pre-processing operation:



22,636,158 valid points for stop detection usage in Sprint 3 (i.e. both purified "static" and "non-static" points, but excluding detected bounce)

3. 研究成果（GPSサンプルデータ解析） ～カンボジアのGPSデータを例として～

GPSデータのトリップへの変換方法

楠瀬、佐々木教授
(早稲田大学)

「東京都市圏における長期のGPSデータを用いた移動経路の推定に関する研究」（大野ら）を参考に、ステイポイントを特定し、トリップエンド（移動の起終点）を定義
2時間以内で300m以上移動しているステイポイントをトリップエンドとし、その間をトリップとして抽出（現在、閾値の妥当性を検証中）

ステイポイント抽出結果例



A地点 (6ポイント)

B地点 (1ポイント)

| ステイポイント検出時間 | A地点 (6ポイント) | B地点 (1ポイント) |
|-------------|---------------------------|---------------------------|
| | 01/01 20:40 - 01/02 06:47 | 01/24 09:43 - 01/24 11:52 |
| | 01/03 20:43 - 01/06 08:39 | |
| | 01/06 22:54 - 01/07 23:09 | |
| | 01/10 22:47 - 01/11 07:11 | |
| | 01/24 14:48 - 01/25 14:33 | |
| | 01/26 16:59 - 01/27 08:33 | |

考察

- 何度も滞在、すべて夜間を含んでいる >自宅?
- 昼間に一度だけ >外出先?

(参考) ステイポイント
・一定範囲 (150m) 内に一定時間 (20分) 以上滞在したポイント。このようなポイント同士の重心をとってトリップエンドとみなす
(長所) 滞在中にGPSの精度によって細かくプロットが動いているように見える問題を解消
(短所) 自宅・勤務先など長時間の滞在は検出できるが、買い物など短時間の滞在はとれない

4. 今後の研究活動と予定

【活動1】全世界におけるビックデータの利活用の実態把握

【活動2】ラオス及びカンボジアの都市交通マスタープランの調査手法、データの質、モデルの妥当性の検証

【活動3】ラオス及びカンボジアにおける都市交通インフラの計画に資するビックデータ（携帯位置情報、交通管制データ、バス運行データ等）の特定とベースラインデータの構築及びモデルの有効性の検証

■ GPSデータ特性の把握とクリーニング手法の検証

■ GPSデータのトリップデータへの変換手法の検証

■ トリップデータの母集団への拡大手法の確立

■ PTデータとの比較検証

【活動4】都市交通インフラの計画におけるビックデータの活用の主流化にかかる検討（技術面や費用面での実現可能性調査、コロナ禍における副次的効果）