

# “Pstay”～クラウドソーシングを活用した 店舗等の滞在者数データ開発の試み～

秋山祐樹 (aki@iis.u-tokyo.ac.jp)

マイクロジオデータ研究会

東京大学 地球観測データ統融合連携研究機構

空間情報科学研究センター

愛知大学 三遠南信地域連携研究センター

運営委員長

特任助教

客員研究員

研究員



# 発表の流れ

1. 本研究の背景
2. データの収集と開発
3. データの検証
4. 今後の展開

# 発表の流れ

## 1. 本研究の背景

2. データの収集と開発
3. データの検証
4. 今後の展開

# 1. 本研究の背景

まずは学術的な背景から…

これまでマイクロジオデータ研究会では

- ・電話帳データ
- ・大規模な人流データなど・・・

様々なマイクロジオデータ（MGD）を活用して、店舗や商店街の分布、またそこに滞在する人々の数を時系列的に推定する研究を行ってきた。

これまでの研究では**ある地域や商店街といった「面的」な範囲**の滞在者数は、ある程度推定出来るようになってきた。

しかし、**ある店舗や施設といった「点」**の滞在者数を把握・推定することは、既存データの集計単位や測位精度などの制約により困難だった。

# 1. 本研究の背景・目的

	位置情報	時系列的分布	滞在者数
店舗	○ (デジタル電話帳によつて個店単位で特定可能)	○ (Webから情報収集・解析することで特定・推定が可能 (岡本ほか, 2012) )	?
商業地域・商店街	○ (商業集積統計により 特定可能 (秋山ほか, 2013年) )	○ (Webから情報収集・ 解析することで 特定・推定可能 (岡本ほか, 2012) )	○ (大規模人流データより 推定可能 (秋山ほか, 2013) )

店舗単位の滞在者数をいかにして知るか？

# 1. 本研究の背景・目的

## 店舗単位の滞在者数をいかにして知るか？

### 調査会社の利用

#### 内訳（通行料調査の例）

- ・調査員日給 0.7～1万円
- ・調査員募集に掛かる労務費
- ・カウンターや調査票、地図等準備に掛かる労務費
- ・現地の下見にかかる労務費
- ・道路上で観測する場合には道路使用許可証の手数料
- ・調査員の教育にかかる費用
- ・データ打ち込みと集計に掛かる費用

単純だがトータルで高コストに

#### 通行量（カウント）調査

Count Pack

新規出店などの情報として必要な通行量（カウント）調査をスピーディかつ低価格にてご提供するパックです

最短3日  
で納品

1スポット

70,000円 (別途消費税)

##### [仕様]

- 対象エリア：首都圏（東京20km圏）
- 計測地点：1スポット（1地点）※1
- 計測日数：2日間（1週間内の任意2日間）※2
- 計測時間：ご指定の開始時間から10時間※3 毎時30分計測30分休憩の交互で実施
- 計測区分：男女2区分または方向別2区分※4
- 納品データ：調査日・時間帯・男女または方向別の数値結果

\*1.計測地点が複数の場合は更にお安くなる場合がございますので、お問い合わせください。

\*2.計測日数は、1週間の中から2日間でお選びください。（例）平日の水曜日と休日の日曜日

\*3.開始時間は7時～11時の範囲でご指定ください。（例）朝6時～夜19時

なお、計測時間は毎時30分計測のため各時間帯のデータを倍に補正した上でご納品します。

\*4.計測区分は公道上を覆面で実施できることを想定してます。場所によってはお受けできない場合もございます。また、調査許可をいただける施設または敷地内である場合は、最大で10区分まで対応することも可能です。

上記以外は別途お見積させていただきますのでお問い合わせください

<http://www.facecom.co.jp/price.html>

# 1. 本研究の背景・目的

店舗単位の滞在者数をいかにして知るか？

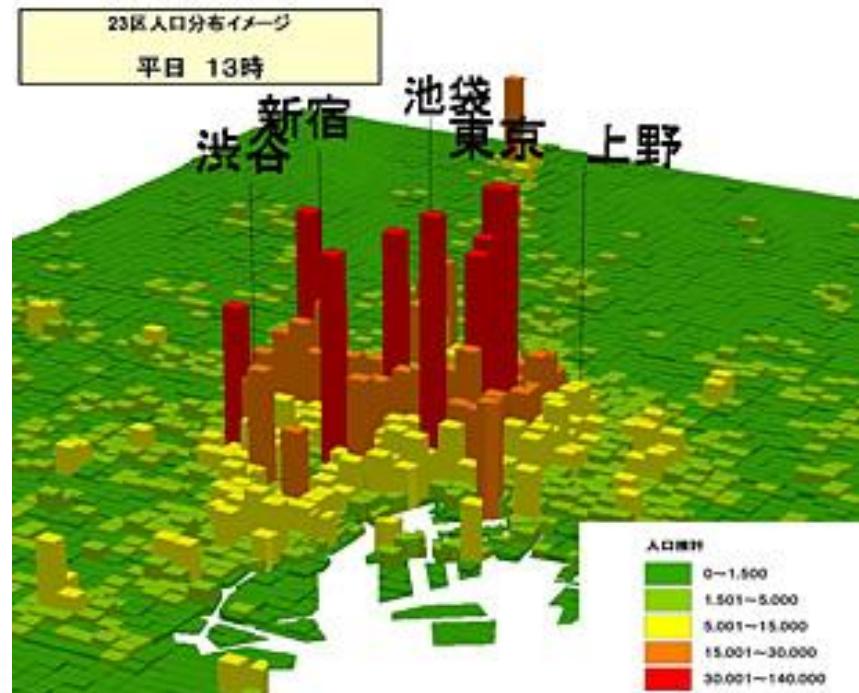
## 大規模人口統計の活用

例：NTTドコモ モバイル空間統計  
(2013年10月販売開始)

- ドコモの基地局利用者から人口を推計したデータ。性別・年齢層別・居住エリア別の人口の構成(人口構成)等を属性として保存。
- 価格： 数百万円～数千万円程度  
(2013年9月現在の公開情報)

継続的かつ広域をカバー。しかし高コスト。

モバイル空間統計イメージ：東京23区周辺の人口分布



<http://www.rbbtoday.com/article/2013/09/09/111503.html>

# 1. 本研究の背景・目的

店舗単位の滞在者数をいかにして知るか？

		範囲	精度	人手・時間	コスト	アクセス
既存データ 利用	大規模人流データ・ プローブデータ等	広域	面・ライン (店舗単位は 困難)	小さい	極めて 大きい	困難 (一部購入可 能)
	パーソントリップ 調査	広域	面・ライン (店舗単位は 困難)	大きい	小さい	利用可能 (研究用途)
	センサ・カメラ・ 屋内測位 (IMES) 等	狭域	点 (店舗単位も 可能)	小さい	小～大	困難
調査会社の利用		任意	点 (店舗単位も 可能)	大きい	大きい	容易
現地調査		狭域	点 (店舗単位も 可能)	大きい	小さい	容易

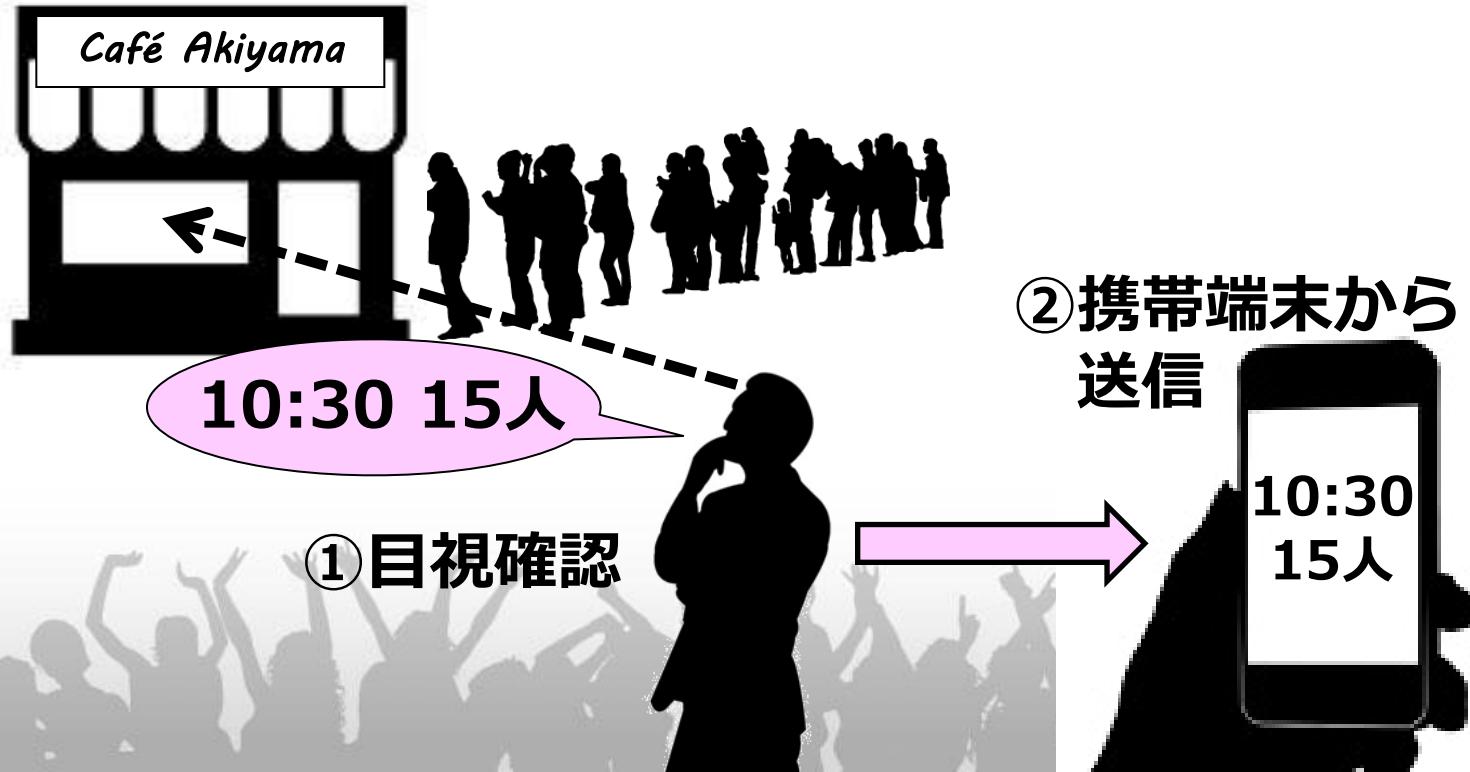
広域を点 (店舗単位) で 人手・手間を出来るだけ小さく抑えて  
低成本で調査し、しかもみんなで共有できるデータ  
を実現出来ないか？

# 1. 本研究の背景・目的

店舗単位の滞在者数をいかにして知るか？

## クラウドソーシングの活用

店舗等の滞在者数を目視調査・共有し、サンプル数が**統計的に一定量**を越えると一日の時間帯別滞在者数の推計値を算出できるようにする。



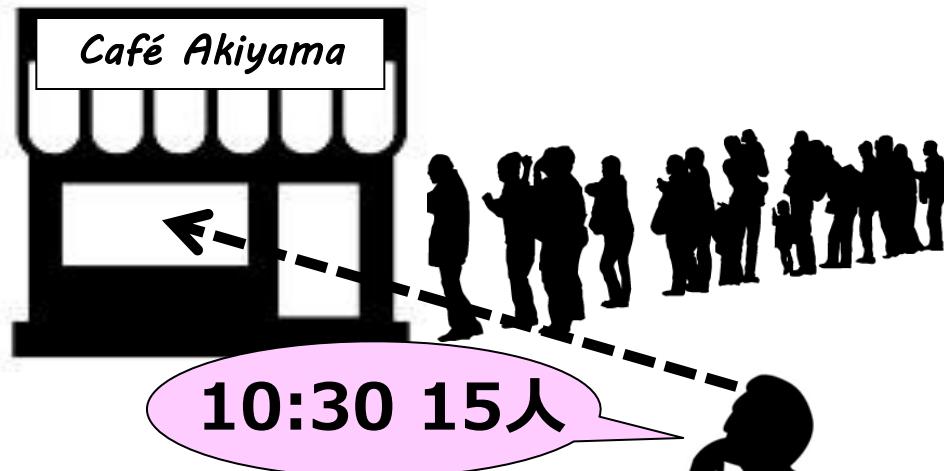
# 1. 本研究の背景・目的

店舗単位の滞在者数をいかにして知るか？

## クラウドソーシングの活用

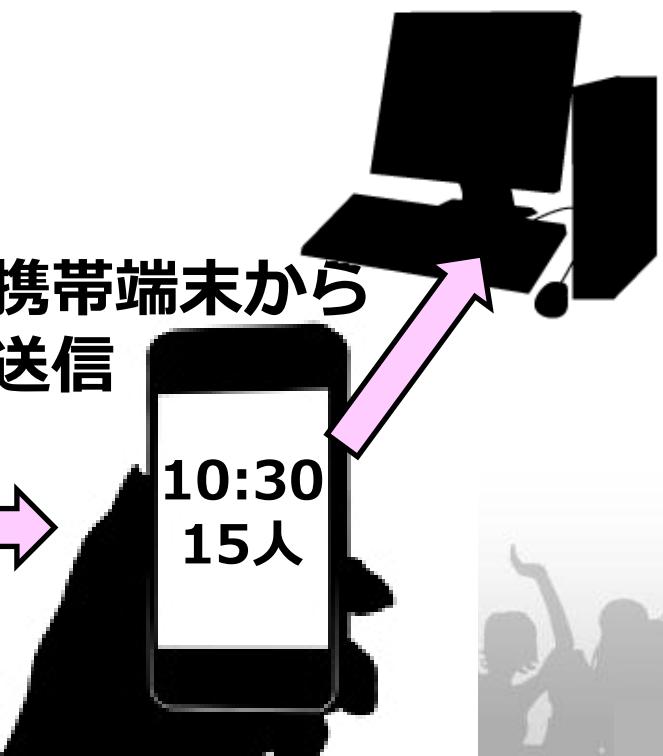
店舗等の滞在者数を目視調査・共有し、サンプル数が**統計的に一定量**を越えると一日の時間帯別滞在者数の推計値を算出できるようにする。

③サーバに集約



①目視確認

②携帯端末から  
送信



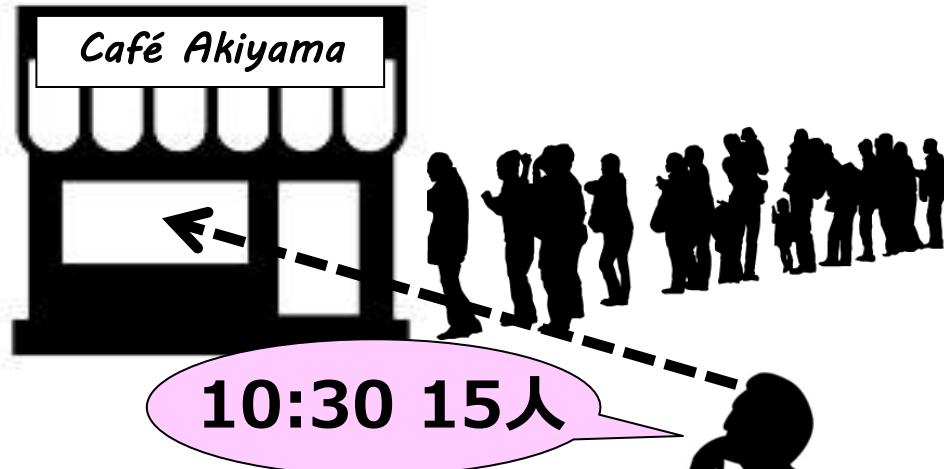
# 1. 本研究の背景・目的

店舗単位の滞在者数をいかにして知るか？

## クラウドソーシングの活用

店舗等の滞在者数を目視調査・共有し、サンプル数が**統計的に一定量**を越えると一日の時間帯別滞在者数の推計値を算出できるようにする。

③サーバに集約



①目視確認

②携帯端末から  
送信



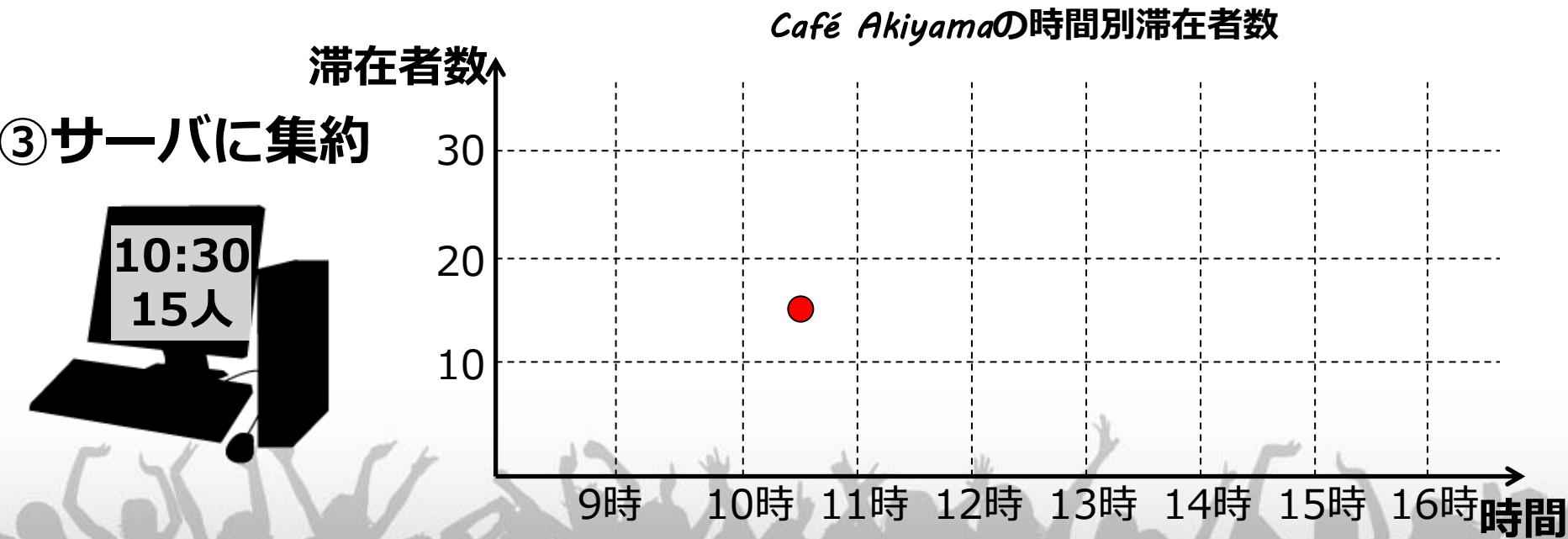
# 1. 本研究の背景・目的

店舗単位の滞在者数をいかにして知るか？

## クラウドソーシングの活用

店舗等の滞在者数を目視調査・共有し、サンプル数が**統計的に一定量**を越えると一日の時間帯別滞在者数の推計値を算出できるようにする。

### ③サーバに集約



# 1. 本研究の背景・目的

店舗単位の滞在者数をいかにして知るか？

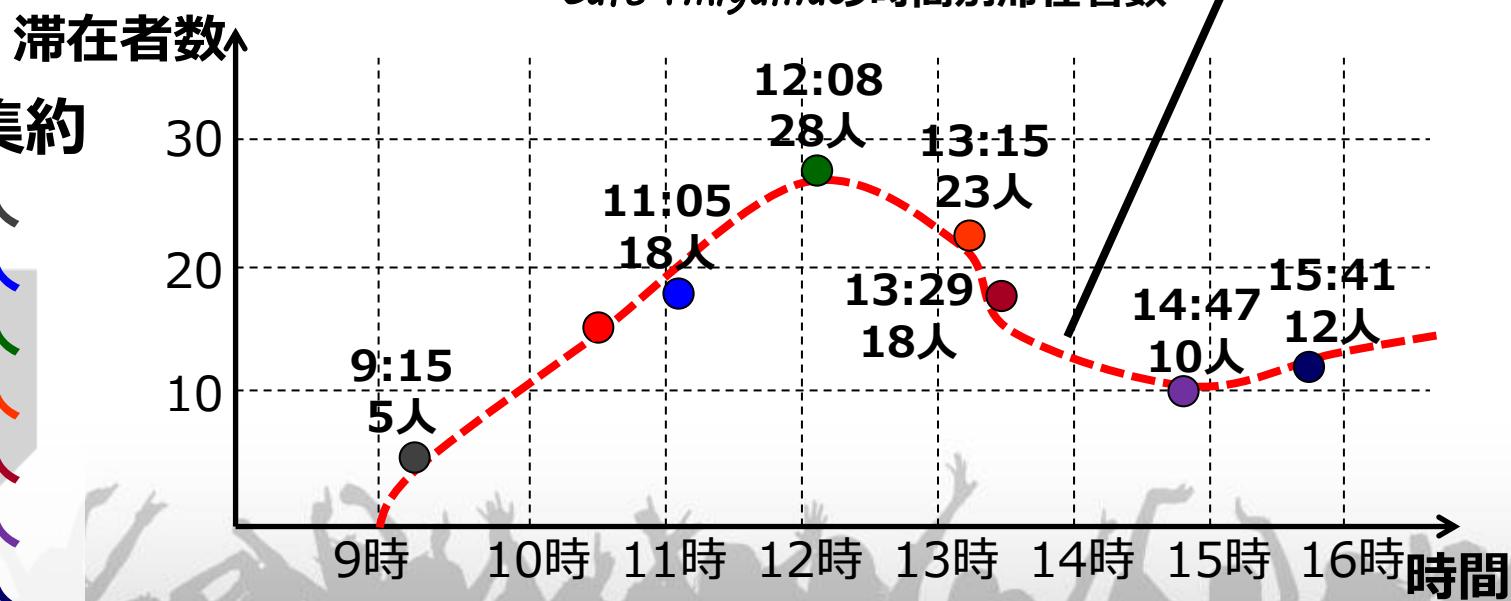
## クラウドソーシングの活用

店舗等の滞在者数を目視調査・共有し、サンプル数が**統計的に一定量**を越えると一日の時間帯別滞在者数の推計値を算出できるようにする。

④平滑化処理を行い  
任意の時間の滞在者  
数を推定

### ③サーバに集約

09:15	5人
11:05	18人
12:08	28人
13:15	23人
13:29	18人
14:47	10人
15:41	12人
...	



# 1. 本研究の背景・目的

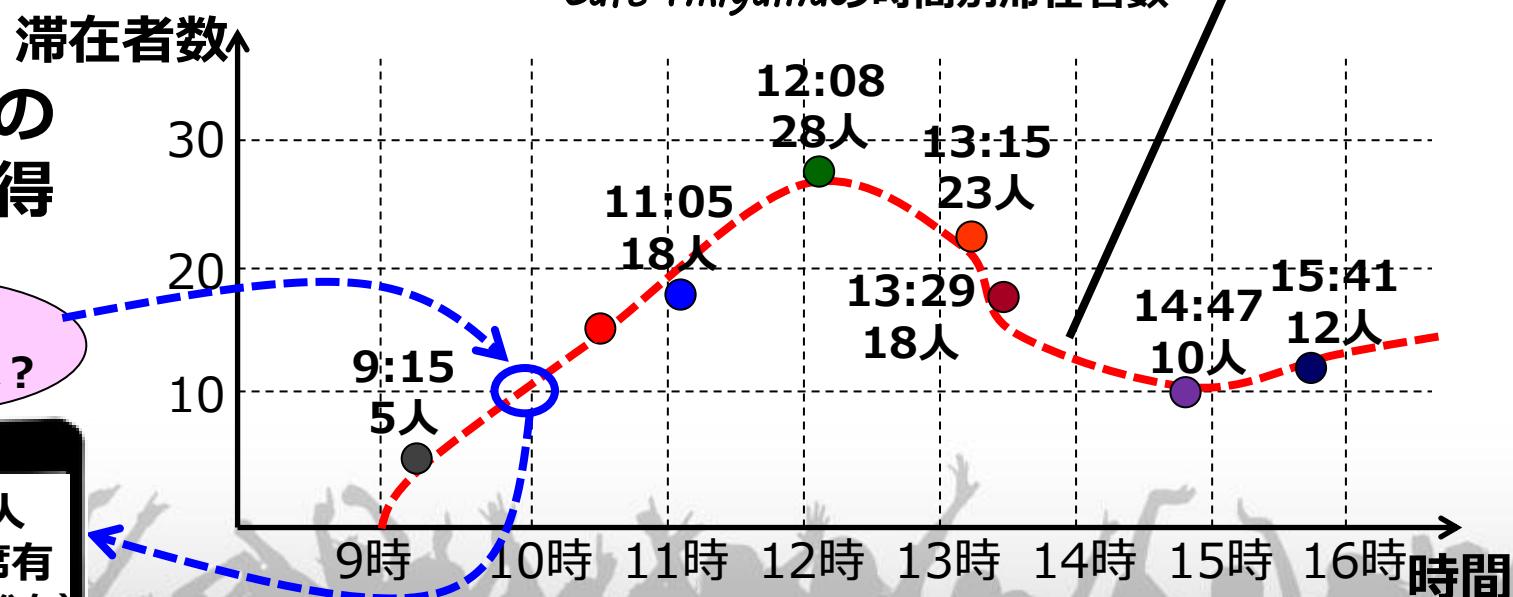
店舗単位の滞在者数をいかにして知るか？

## クラウドソーシングの活用

店舗等の滞在者数を目視調査・共有し、サンプル数が**統計的に一定量**を越えると一日の時間帯別滞在者数の推計値を算出できるようにする。

④平滑化処理を行い  
任意の時間の滞在者  
数を推定

⑤任意の時間の  
滞在者数を取得



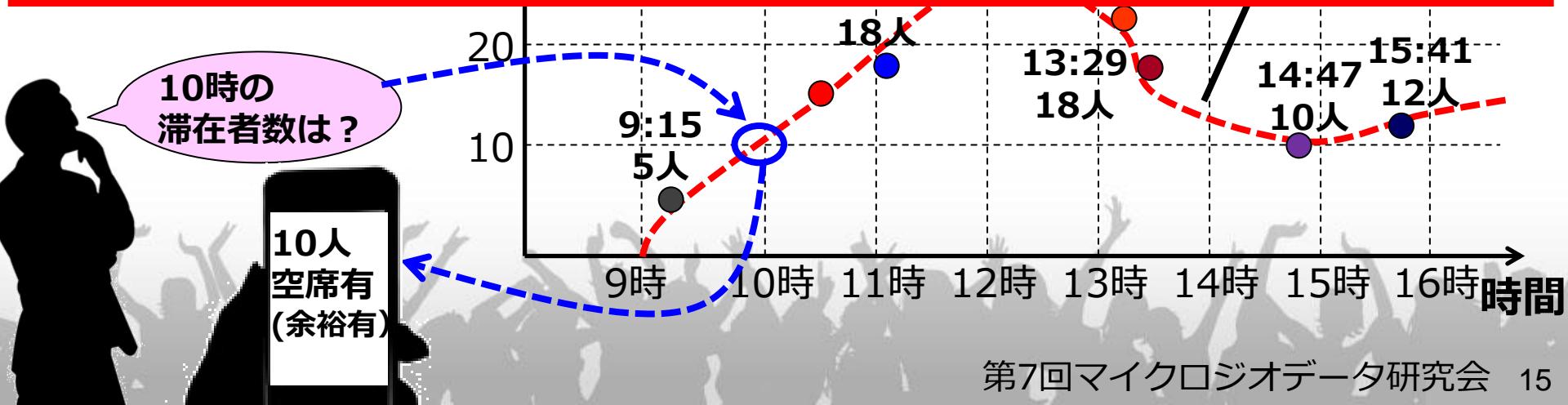
# 1. 本研究の背景・目的

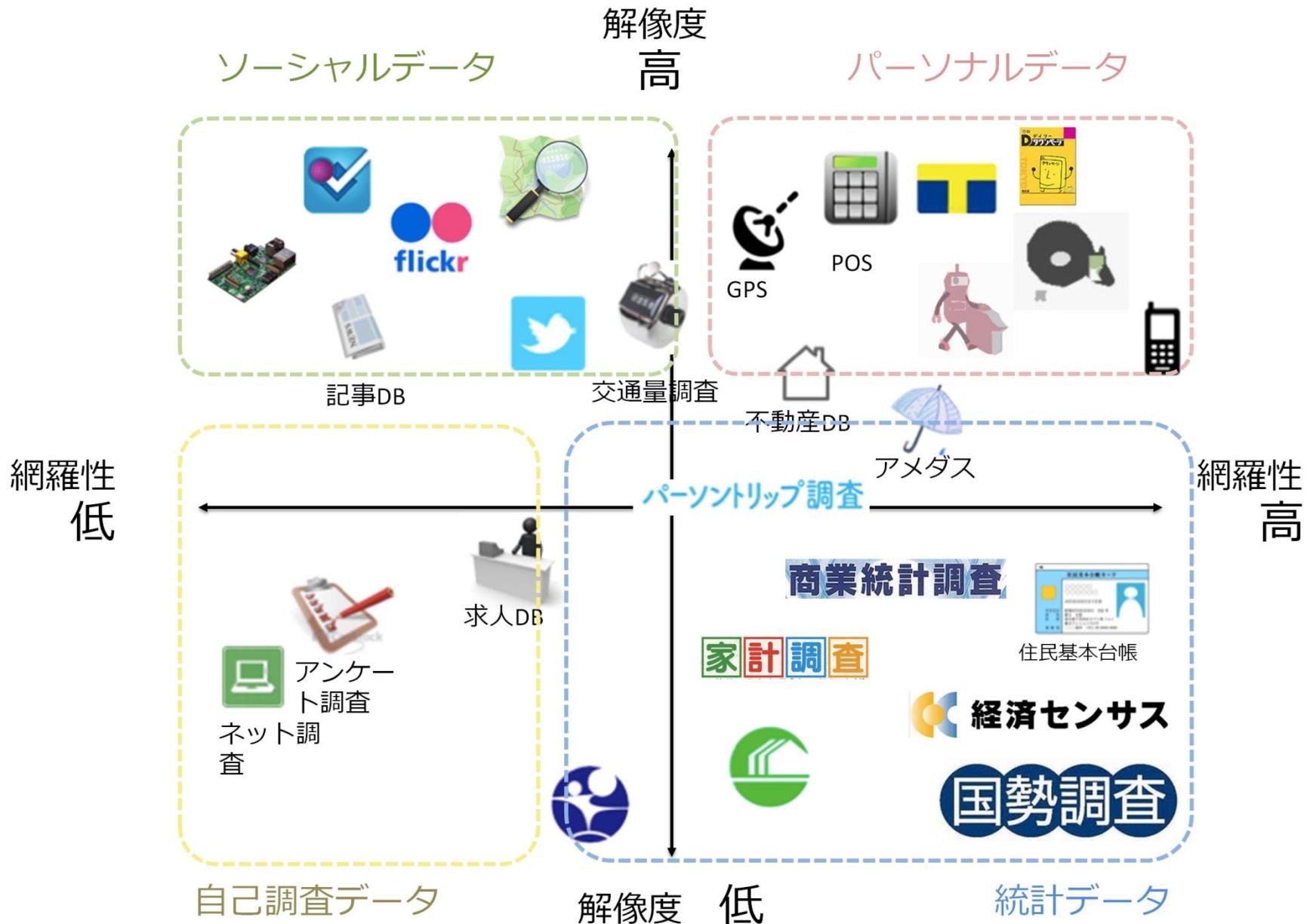
店舗単位の滞在者数をいかにして知るか？

## クラウドソーシングの活用

店舗等の滞在者数を目視調査・共有し、サンプル ④平滑化処理を行い

クラウドソーシングを活用した店舗等の滞在者数のデータを開発するプロジェクト  
= Pstay (Person stay)







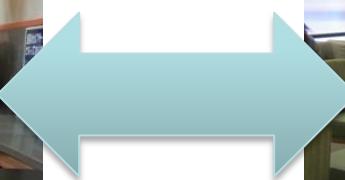
# 1. 本研究の背景

## 社会的課題への適用可能性1：混雑情報提供

混雑が分からないまま、消費者はただ来店するしかなく、業者は立地的に良いのか正確に分からぬまま出店。混雑状況が事前に分かれば、時間の浪費や非効果的な店舗営業を回避できる可能性がある。

混雑！

ガラガラ・・・



消費者  
目線

評判は良さそ  
うなので興味  
はあるが並び  
たくない

穴場？すい  
てるお店に  
入りたい

この近くに  
はまだ出店  
余地がある  
かも

この近くに  
店を出すの  
はやめた方  
がいいかも

業者  
目線

# 1. 本研究の背景

## 社会的課題への適用可能性2: 観光行動支援

Pstayデータが蓄積され、残席数などのリアルタイムな推定が可能になれば、観光地において観光客への飲食店の空席情報や空駐車場情報などの即時的な提供サービスが可能になる。

- 混雑情報がリアルタイムに共有できれば、観光客を極端に混雑している店舗や通りに集まることを回避させる事ができる。
- また店舗側も混雑状況、空席状況を発信することで観光客を呼びこむことが出来る可能性がある。

# 1. 本研究の背景

## 社会的課題への適用可能性3: 防災計画策定支援

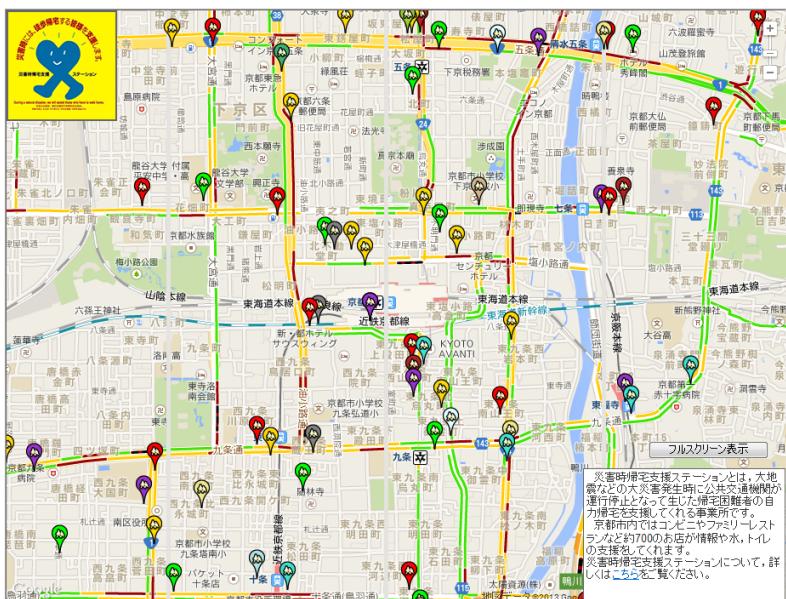
近年では災害時に飲食店など水やトイレを提供出来る店舗・施設を「災害時帰宅支援ステーション」と称し、帰宅困難者のための一時避難・支援施設として活用する試みも始まっている。自治体による上記政策の立案の際にも時間別の店舗滞在者数情報は重要である。



(写真提供:(社)日本フランチャイズチェーン協会)

災害時帰宅支援ステーションのステッカー

<http://www.kouiki-kansai.jp/contents.php?id=17>



災害時帰宅支援ステーションマップ (京都市)

<http://www.city.kyoto.lg.jp/gyoza/page/000108904.html>

# 発表の流れ

1. 本研究の背景・目的
- 2. データの収集と開発**
3. データの検証
4. 今後の展開

## 2. データの収集と開発

2013年10月～現在まで小田急線代々木上原駅周辺を中心に、店頭や店内にて滞在者数を目視で観測し収集した。

主に飲食店（カフェ・ファーストフード・居酒屋等）、小売店（スーパーマーケット・コンビニエンスストア・ドラッグストア等）、書店を対象とした。

### <データ収集方法>

- ・Googleフォームによる簡易的なデータ送信フォームを作成。
- ・店舗の名称と滞在人数を送信することで、店舗名称とその滞在者数、送信時間を記録出来るようにした。
- ・スマートフォン（iPhone, Android）、ガラケー、タブレット等、インターネットに接続可能な端末なら利用可能。

店舗調査

\*必須

人数 \*

場所 \*

送信

Google フォームでパスワードを送信しないでください。

Powered by  
Google Forms

データ送信フォーム

## 2. データの収集と開発

### 収集したデータ（一部）

調査日時	調査者	滞在者数	場所(店舗名称等)
6/6/2014 22:07:38	yoshio	7	サンクス 神田北口店
6/9/2014 19:30:35	akiyama	1	IL TRAMEZZINO
6/9/2014 19:42:02	akiyama	13	世田谷珈琲游
6/9/2014 19:42:44	akiyama	10	せい家 梅ヶ丘店
6/9/2014 19:46:18	akiyama	9	カルディコーヒーファーム梅ヶ丘店
6/9/2014 19:47:02	akiyama	19	名物串カツ田中
6/9/2014 20:03:15	akiyama	5	東秀 梅ヶ丘店
6/10/2014 11:03:02	akeyama	9	三菱東京UFJ銀行 笹塚支店 ATMコーナー
6/10/2014 11:07:17	akeyama	4	サンクス 笹塚二丁目店
6/10/2014 21:19:32	akeyama	3	サークルKサンクス 駒場4丁目店
6/10/2014 21:23:01	sengoku	9	まいばすけっと 東北沢店
6/10/2014 21:35:47	sengoku	6	ローソン 東北沢店
6/10/2014 21:39:40	sengoku	2	ローソン 東北沢南店
6/10/2014 21:56:29	sengoku	12	千里眼 駒場店
6/10/2014 21:56:51	sengoku	6	サークルKサンクス 駒場4丁目店
6/11/2014 13:08:16	akeyama	7	なか卯 笹塚店
6/11/2014 13:16:37	akeyama	3	サンクス 笹塚二丁目店
6/11/2014 21:10:04	akeyama	7	サークルKサンクス 駒場4丁目店
6/12/2014 0:58:05	akeyama	3	サークルKサンクス 駒場4丁目店
6/12/2014 15:26:53	nishimura	0	サンクス 柏駅前店
6/12/2014 22:07:21	nishimura	7	サンクス 神田駅北口店
6/13/2014 17:08:06	akiyama	6	サブウェイ代々木上原店
6/13/2014 17:08:24	sengoku	47	サンマルクカフェ 代々木上原店
6/13/2014 17:08:41	sengoku	4	なか卯 代々木上原店
6/14/2014 9:44:29	akiyama	7	セブンイレブン世田谷梅ヶ丘店
6/14/2014 12:42:37	yoshio	7	新潟濃厚味噌 弥彦
6/14/2014 12:46:16	yoshio	15	ラーメン 不如帰
6/14/2014 16:29:00	akiyama	4	IL TRAMEZZINO

## 2. データの収集と開発

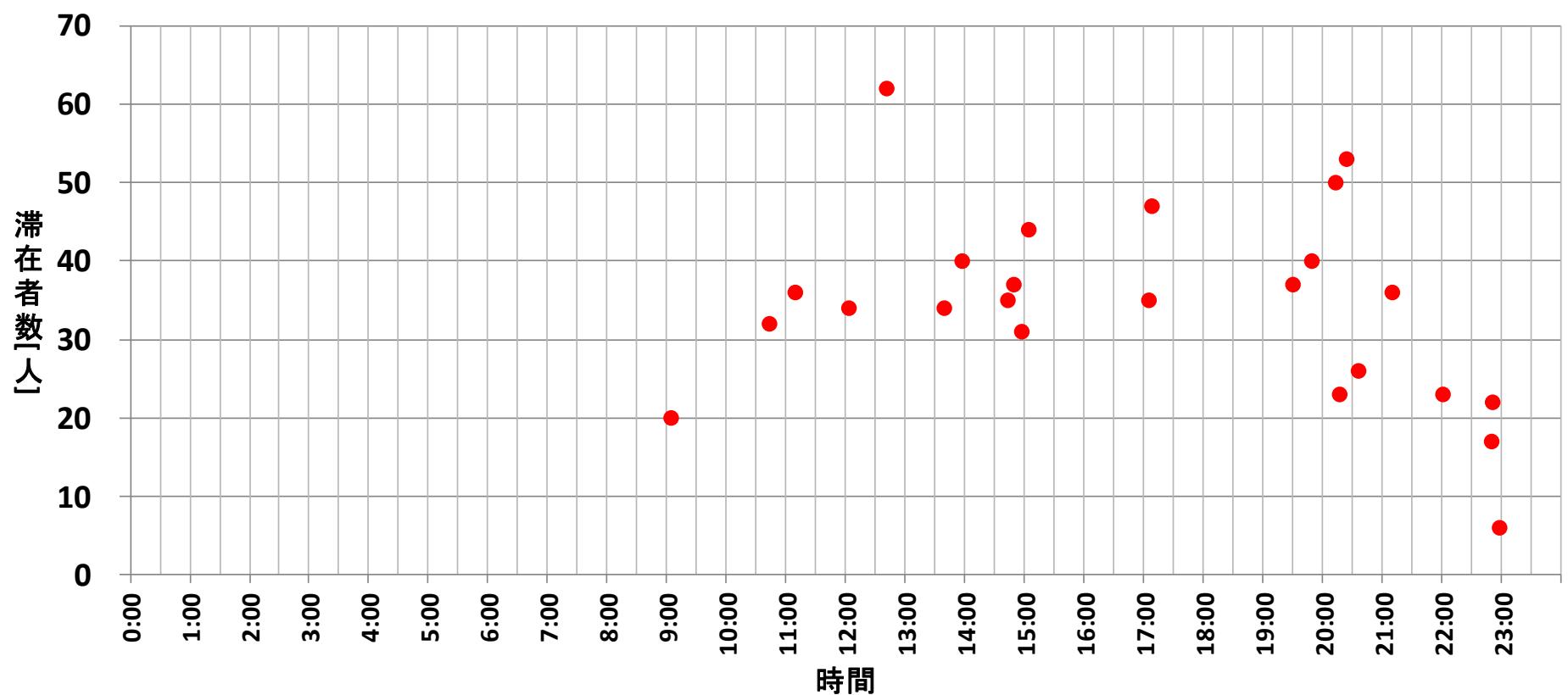
### 収集したデータ：サンマルクカフェ代々木上原店の場合

調査日時	調査者	滞在者数	場所(店舗名称等)
5/10/2014 9:53:05	sengoku	18	サンマルク 代々木上原店
9/26/2013 14:43:08	sengoku	35	サンマルク 代々木上原店
4/24/2014 9:04:53	sengoku	20	サンマルク 代々木上原店
4/24/2014 10:43:56	sengoku	32	サンマルク 代々木上原店
5/9/2014 14:49:52	sengoku	37	サンマルク 代々木上原店
5/10/2014 11:09:17	sengoku	36	サンマルク 代々木上原店
5/10/2014 20:17:51	sengoku	23	サンマルク 代々木上原店
5/13/2014 19:30:40	sengoku	37	サンマルク 代々木上原店
5/13/2014 20:13:33	sengoku	50	サンマルク 代々木上原店
5/13/2014 22:51:22	sengoku	22	サンマルク 代々木上原店
5/15/2014 13:39:18	sengoku	34	サンマルク 代々木上原店
5/15/2014 13:57:56	sengoku	40	サンマルク 代々木上原店
5/17/2014 21:10:07	sengoku	36	サンマルク 代々木上原店
5/17/2014 22:58:07	sengoku	6	サンマルク 代々木上原店
5/19/2014 12:03:17	sengoku	34	サンマルク 代々木上原店
5/20/2014 17:05:42	sengoku	35	サンマルク 代々木上原店
5/21/2014 15:04:06	sengoku	44	サンマルク 代々木上原店
5/23/2014 14:57:56	sengoku	31	サンマルク 代々木上原店
5/25/2014 20:36:56	sengoku	26	サンマルク 代々木上原店
5/31/2014 22:01:10	sengoku	23	サンマルク 代々木上原店
5/31/2014 22:50:02	sengoku	17	サンマルク 代々木上原店
6/4/2014 19:49:19	sengoku	40	サンマルク 代々木上原店
6/13/2014 17:08:24	sengoku	47	サンマルク 代々木上原店
6/19/2014 20:24:50	sengoku	53	サンマルク 代々木上原店
7/4/2014 12:41:20	sengoku	62	サンマルク 代々木上原店
7/18/2014 11:07:48	sengoku	45	サンマルク 代々木上原店
7/22/2014 12:07:07	sengoku	47	サンマルク 代々木上原店
7/22/2014 19:20:32	sengoku	60	サンマルク 代々木上原店
8/14/2014 18:52:52	sengoku	49	サンマルク 代々木上原店

## 2. データの収集と開発

収集したデータ：サンマルクカフェ代々木上原店の場合

- 滞在者数（実測値）



## 2. データの収集と開発

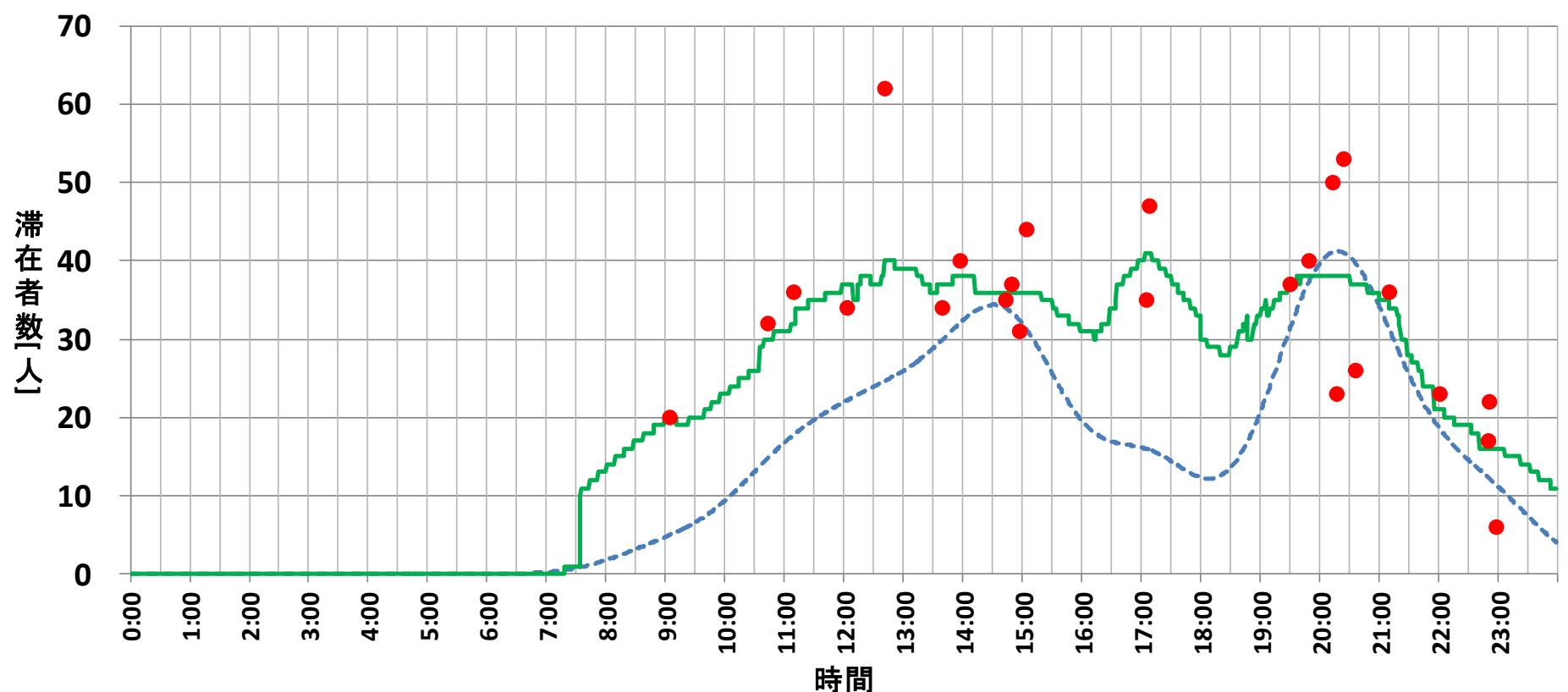
### 非連續な実測値の平滑化

- 各時間の滞在者数をその時間にその店舗に滞在者がいる確率と見なし、00:00から23:59の間、1分毎にカーネル密度推定を行った（カーネル関数はGaussian関数、バンド幅は可変（Sheather and Jones (1991)））。
- カーネル密度が最も大きくなる時刻 $t_{max}$ からバンド幅以内にある観測値の平均値（ただし $t_{max}$ からの時間差により影響力を重み付けする）で、カーネル密度を拡大する。
- 毎分の推定値においてその周辺（バンド幅分）の観測値の値との総和の平均を計算して推定値を補正する。

## 2. データの収集と開発

収集したデータ：サンマルクカフェ代々木上原店の場合

- 滞在者数（実測値）
- カーネル密度に基づく推定滞在者数
- 補正後の推定滞在者数

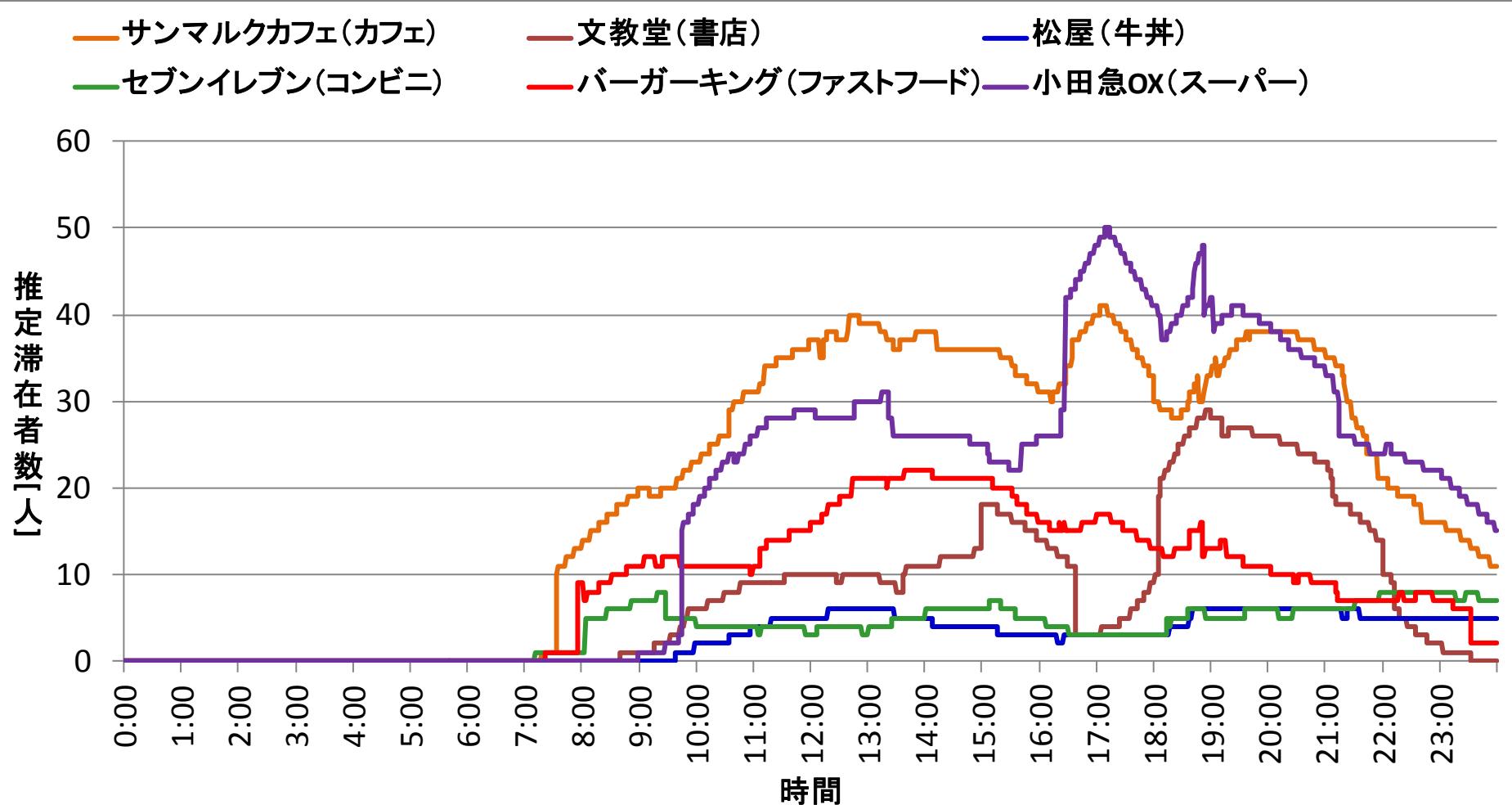


## 2. データの収集と開発

推定滞在者数：代々木上原駅周辺の例

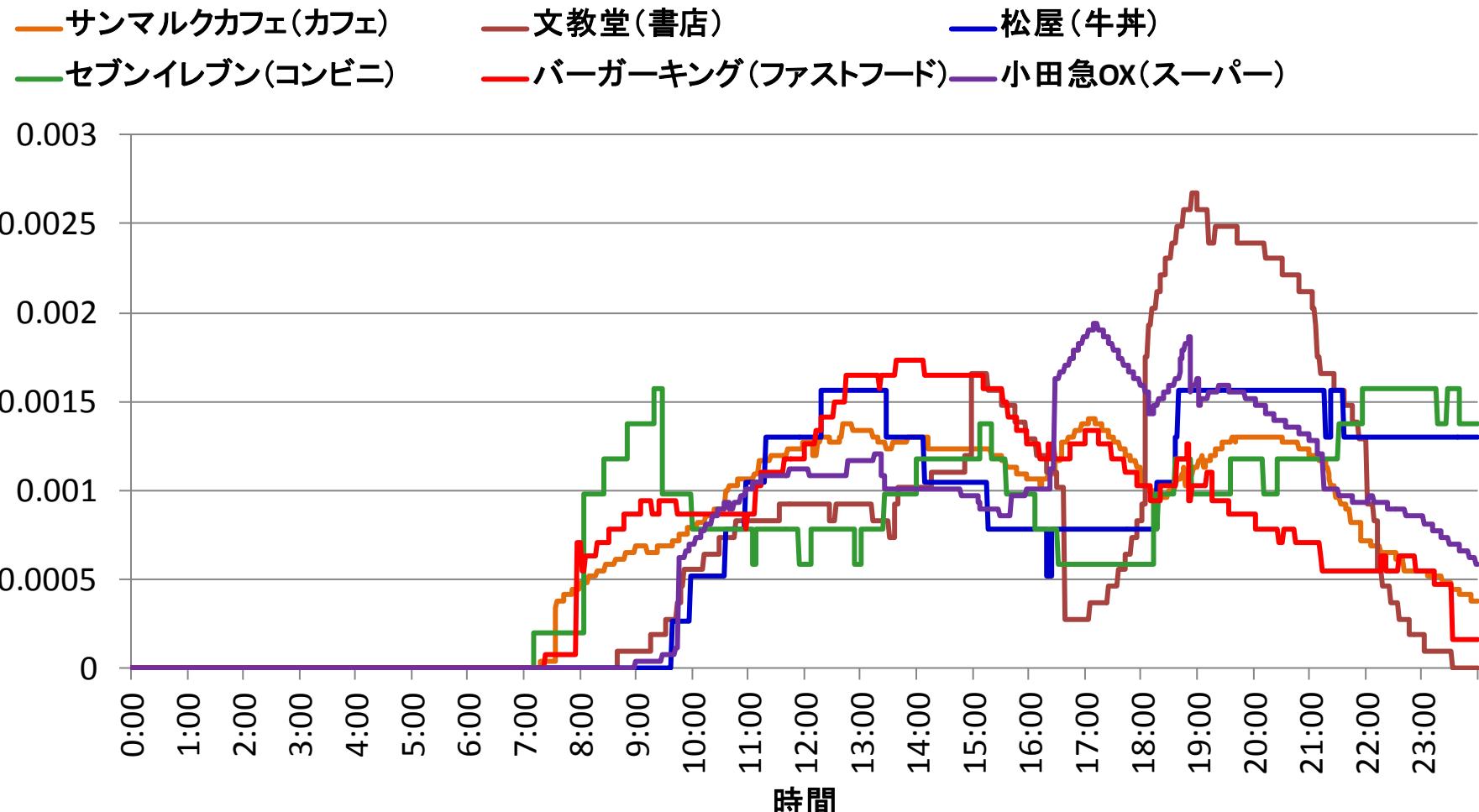
## 2. データの収集と開発

### 推定滞在者数：代々木上原駅周辺の例



## 2. データの収集と開発

### 推定滞在者数密度：代々木上原駅周辺の例



# 発表の流れ

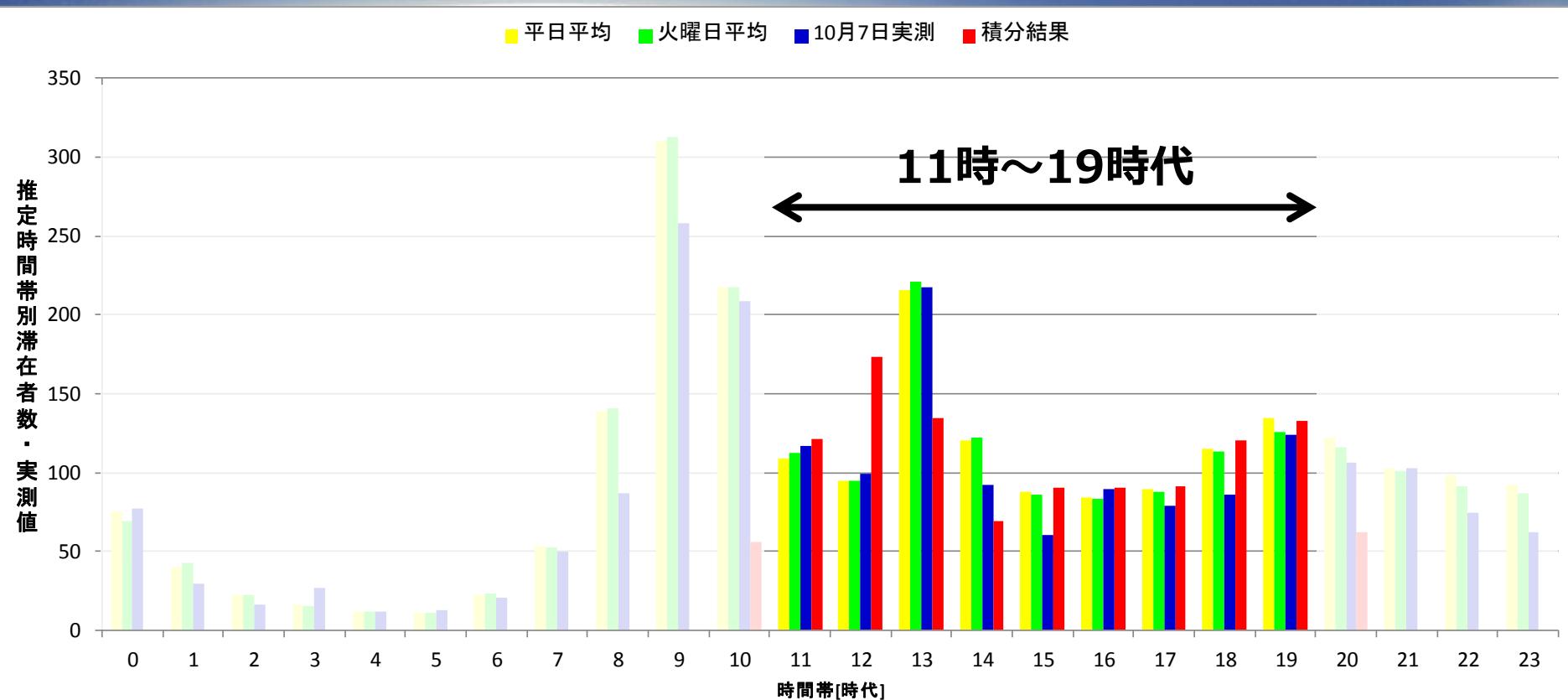
1. 本研究の背景・目的
2. データの収集と開発
- 3. データの検証**
4. 今後の展開

### 3. データの検証

- ・時間別推定滞在者数の信頼性を検証するために、株式会社サークルKサンクスよりご提供頂いた店舗（都内某駅周辺9店舗）のPOSデータとの比較を実施した。
- ・POSデータを頂いた店舗を対象に、2014年10月7日（火）午前11時～午後7時に調査員9人体制で対象店舗を巡回・調査した。
- ・各調査員は概ね1時間で対象9店舗を2回調査。 =各店舗1時間に概ね18サンプルを収集。
- ・同時に各店舗で調査員が任意で選んだ客を対象に、入店から退店までの時間を計測した（分単位）。
- ・POSデータからは時間帯別（1時間毎）のレジ利用者数（買い物客数）を把握できる。
- ・一方、本研究で得られるデータは任意の時間の店舗内滞在者数であるため、同データを1時間毎に積分することで比較できる状態にした。
- ・積分の幅は店舗の平均滞在時間を採用した。 =今回は2分とした。

### 3. データの検証

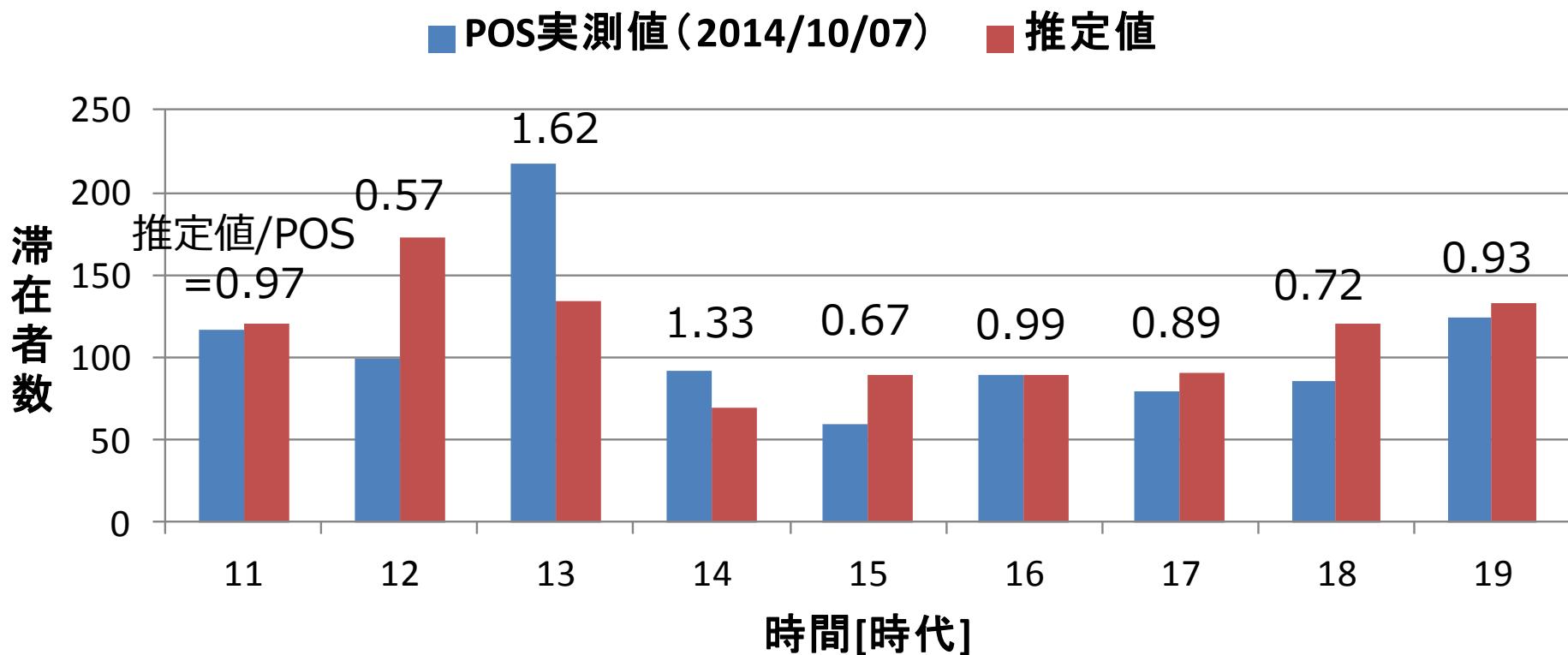
#### 某A店における推定値とPOS実測値の比較



- 多くの時間帯で本研究で得られた滞在者数とPOS実測値が近い値となる。
- 調査日当日、火曜日平均、平日平均のPOS実測値の値に大きな差は無い。

### 3. データの検証

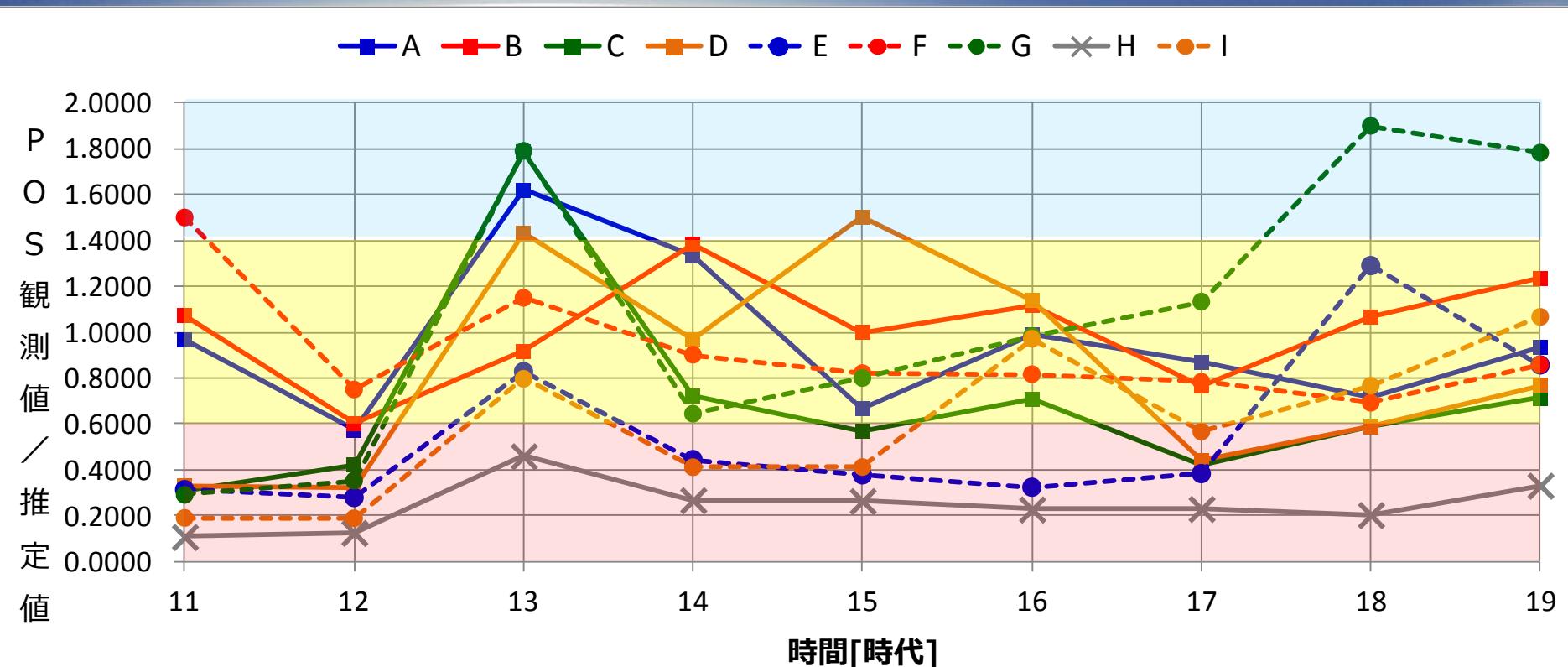
#### 某A店における推定値とPOS実測値の比較



- 12時、13時は誤差が大きくなりがちであった。
  - > 12時：推定値の方が大きい=レジ待ち、または買物無しで退店する人が多い。
  - > 13時：推定値の方が小さい=平均的な店舗滞在時間が2分よりも短い。

### 3. データの検証

#### 全調査対象店舗における推定値とPOS実測値の比較



- ・推定値の方が大きい=レジ待ち、または買物無しで退店する人が多い。
  - ・推定値の方が小さい=平均的な店舗滞在時間が2分よりも短い。
- \* H店はイトーイン、喫煙コーナーがあるため推定値が常に大きい。

# 発表の流れ

1. 本研究の背景・目的
2. データの収集と開発
3. データの検証
- 4. 今後の展開**

# 4. 今後の展開

- ・本研究ではクラウドソーシングを活用して、店舗等の滞在者数を離散的に収集し、それらを時間方向に平滑化することで、任意時間の滞在者数を店舗ごとに推定する手法を提案した。
- ・コンビニのPOSデータで取得された店舗利用者数との比較を行い、本研究で得られた結果がコンビニの店舗利用者の実測値とある程度一致することを明らかにした。

## <今後の展開>

- ・他業種、業態における推定結果の信頼性検証
- ・有意な推定結果を得るために必要なサンプル数の検討
- ・クラウドソーシングを活用して効率よく情報を収集するためのアプリケーション開発とアプリユーザーの獲得
- ・複数ユーザーで収集した結果が日本全国でリアルタイムに閲覧・共有出来る環境作り

# アプリ開発状況



そこで・・・

# 皆様のご協力をお願い致します

本プロジェクトの更なる推進には、全国の店舗、通り、駐車場などの滞在者数が大量に蓄積されていく必要があります。

そこで皆様の力を借りできればと考えています。アピリューザーを募集しております。アピリューザーの方は全国のみんなで集めた情報を閲覧・取得出来ます。

研究・講義・演習などで使ってみたい方も大歓迎です。興味のある方は秋山（aki@iis.u-tokyo.ac.jp）までご一報下さい。

# ご清聴頂きありがとうございました

<Contact>

秋山祐樹

マイクロジオデータ研究会 運営委員長

東京大学地球観測データ統融合連携研究機構 特任助教

東京大学空間情報科学研究センター 客員研究員

愛知大学三遠南信地域連携研究センター 研究員

Email: [aki@iis.u-tokyo.ac.jp](mailto:aki@iis.u-tokyo.ac.jp)

URL: <http://shiba.iis.u-tokyo.ac.jp/member/akiyama/>

(本日の資料もアップする予定です)

マイクロジオデータ研究会ホームページ

<http://geodata.csis.u-tokyo.ac.jp/wp/>



「秋山祐樹」・”akiyama.yuuki”で検索