

# 人流ビッグデータを用いたイベント参加前後の行動分析

小林稜介・宮澤聡・秋山祐樹・柴崎亮介

## Analysis of activities before and after events using Geo Big Data

Ryosuke KOBAYASHI, Satoshi MIYAZAWA,

Yuki AKIYAMA and Ryosuke SHIBASAKI

**Abstract:** Currently, there are more people going to events such as a professional baseball games and concerts. In addition, the Tokyo Olympic and Paralympic Games are coming up. In terms of economic, guests' activities before and after events are one of the important things in order to have a greater ripple effect. In consideration of this, we extracted dwell points before and after events using geo mobile phone big data. We then identified factors contributing to these activities.

**Keywords:** ツーリズム (tourism), 周遊行動 (excursion activities), 位置情報 (location information), 滞留点 (dwell point), 携帯電話 (mobile phone)

### 1. 背景と目的

近年、プロ野球や J リーグなどのスポーツ競技や音楽ライブイベントの動員数は増加傾向にあり、2018 年にはプロ野球公式戦及び音楽ライブイベントの動員数が過去最高を記録している (日本野球機構, コンサートプロモーターズ協会調べ)。また 2020 年には東京五輪の開催も控えており、様々なイベントに注目が集まっている。イベントに期待される経済効果は、会場での消費活動はもちろん、会場周辺における参加前後のショッピングや飲食などのイベントから派生した滞留行動 (立ち寄り) による波及効果も大きい。

しかし、これらの滞留行動に関する研究は現在少なく、音楽ライブイベント前後の行動分析を行った研究 (堀野ほか, 2018) が挙げられるが、アンケート調査の結果を利用しており、同様の調査を広範囲で定期的に行うことは困難である。

そこで本研究では、大量の携帯電話の移動履歴に基づく 1 年間の人の動きに関するビッグデータを用いて各イベントの参加者を抽出し、イベント参加前後の滞留点分布の可視化や滞留行動への影響要因を明らかにすることを目的とする。なお、本研究において滞留とは「一定範囲内に、一定時間以上留まること」を指し、イベントは日本イベント協会による定義である「何らかの目的を達成するための手段として開催される直接的なコミュニケーションメディア」に従うものとする。

### 2. データセット

#### 2.1 使用するデータセット

本研究で主に使用するデータセットは人流ビッグデータ、イベントデータ、Twitter 投稿データの計 3 つである。

#### 2.2 人流ビッグデータ

本研究で扱う人流ビッグデータは、株式会社 Agoop が提供するスマートフォンアプリからユーザーの同意を得て収集した、関東全域のポイント型流動人口データである。表 1 に同データの

概要及び統計量を示す．なお本研究では，位置精度が 365 m 以下(上位 88%)のデータのみを取り扱うものとした．

表-1 人流データの概要と基礎統計量

項目	詳細
対象者	関東地方全域の人々 (域内流入・域外流出含む)
対象期間	2016/01/01 ～ 2016/12/31
測位方法	GPS 測位, Wi-Fi 測位, 基地局 3 点測位
総ユーザー数	(延べ) 36,391,527
総レコード数	1,762,069,599

## 2.3 人流ビッグデータの特徴

人流ビッグデータには，2 つの特徴がある．まずユーザー ID は午前 0 時にリセットされるため，同一のユーザーを日を跨いで把握・分析することはできない．また，測位間隔が低頻度であるという特徴がある．測位頻度は携帯端末の OS によって異なるが，例えば Android 端末の場合 30 分に 1 回であり，低頻度な測位によるデータとなっている．

## 2.4 イベントデータ

本データは公式ウェブサイトをはじめとしたウェブサイト上で公開されている，2016 年に開催された様々なイベント情報を抽出したものである(表 2)．本研究では代表的なイベントととして，近年観客動員数が増加しているスポーツ競技と音楽イベントを分析対象とした．中でもスポーツ競技は日本プロ野球 12 球団による試合，J1 所属クラブによる試合を対象とする．また，音楽イベントはカテゴリや規模などが多岐に渡るため，2016 年の音楽ライブ年間観客動員ランキング(LiveFans 調べ)において，上位 30 位以内のアーティストによる単独ライブイベントのみを対象とする．

## 2.5 Twitter 投稿データ

イベント前後の行動を把握するためには，イベントが行われた開始・終了時刻の情報が必要であ

るが，音楽イベントについては終了時刻が一般的には公開されていない．そこで，SNS サイト Twitter に投稿されたキーワード「」(アーティスト名) AND “ライブ”」のツイート投稿数を元に終了時刻の補完を行った．参考として，図 1 にあるイベントにおけるツイート投稿数を表す時系列グラフを示す．開始前後は該当ツイート投稿数が増えるが，開始後徐々に投稿数が減少し，イベント終了後に感想などを含めた該当ツイート数が大きく増加する傾向があり，他のイベントにも同様に見られる．そこで開始時刻から 5 時間以内における，該当ツイート投稿数の直近 30 分間移動平均の変化率が最大となった時刻を終了時刻とみなす(図 1 の場合：20:10)．一方，推定開催時間が一般的な単独イベント開催時間と大きく異なる場合や十分なツイート投稿数がないイベントに関しては，他の推定開催時間の中央値(170 分間)を用いて終了時刻補完を行った．

表-2 取得したイベントデータの概要

項目	詳細
開催期間	2016/01/01 ～ 12/31 (252 日間)
イベント数	751 件(プロ野球：353 件，サッカー：182 件，ライブ：216 件)
データ項目	会場(計 33 施設)緯度/経度， 開始/終了時刻，開催時間 など

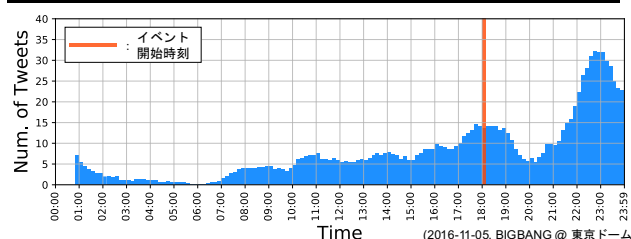


図-1 ツイート投稿数移動平均グラフ

## 3. 分析手法

### 3.1 滞留点抽出

イベント参加者及び滞留点の抽出フローを図 2 に示す．まず点(緯度・経度)の線形補間などの前処理の後，イベントの参加者候補の ID を抽出するため，会場で開催時間前後 2 時間に 1 つ以上ログを残した ID をイベント参加者候補と

して抽出した。なお本研究における「会場」とは、各会場施設ポリゴンの重心を中心点とした半径 500 m（位置誤差閾値 365 m と会場の大きさを考慮した距離）の円内とする。

その後、DBSCAN を用いて密度準拠型クラスタリングによる分散点の統一化を行った。DBSCAN におけるパラメータである閾値距離 (*eps*) は位置誤差として設けた閾値 365 m、最小クラスタサイズ (*min samples*) は 2 とした。これにより得られた会場クラスタでの滞留時間と滞留時間帯を元にイベント参加者を抽出し、各参加者のイベント当日の滞留点を得た。

ただし単に滞留点を抽出した場合、自宅を抽出する可能性も高い。そこで、自宅の可能性が高い点を滞留点から除外するため、一日における最初のログと最後のログに着目し、両ログが同一クラスタである場合は自宅と判断した。また両ログが位置誤差の影響や帰宅が日を跨ぐことにより、異なるクラスタに割り当てられている場合がある。それらに対しては、クラスタ中心間距離が距離 5 km 以内の場合、クラスタの統一化を行い両ログを自宅とみなした。上記の処理後も両ログのクラスタが異なる場合には会場からより距離が大きいログのみを自宅とみなした。

最後に、得られたクラスタから、信号待ちや乗り換えなどの一時的な滞留を除くため 30 分以上滞留した点（自宅を除く）を滞留点として抽出した。図 3 にある個人におけるクラスタ及び滞留点のイメージを示す。

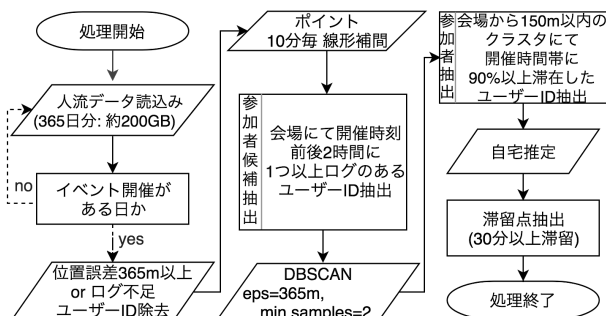


図-2 参加者・滞留点の抽出フロー

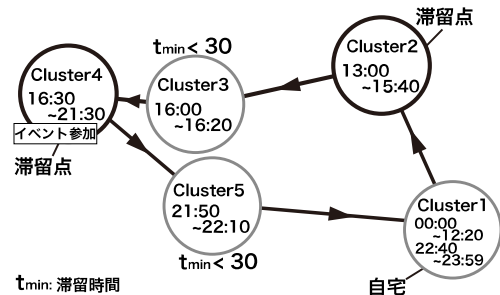


図-3 クラスタ・滞留点イメージ

### 3.2 滞留行動の多変量解析

得られた滞留点を用い、重回帰分析によって、イベント参加前後の会場付近での滞留行動に影響を及ぼす要素を明らかにした。なお「会場付近」とは、日常生活で歩いていける範囲の最頻値である半径 1 km（内閣府調べ）の円内（中心点は各会場施設ポリゴンの重心）とする。重回帰分析に際して、滞留時間が長いほど店舗の売上が上がることから（IHL Group, 2015）、目的変数を会場付近での滞留時間とし、説明変数は施策によって介入できるような空間的要素（総務省 経済センサス、国土数値情報を利用）を設定した（表 3）。

表-3 説明変数と定義

変数名	定義
周辺観光資源数	会場から半径 5 km 以内の観光資源数
交通アクセス	最寄り鉄道駅からの所要時間（分）
最寄り駅の発展度	最寄り鉄道駅の乗降客数（人/日）
飲食店数	会場が位置する市区町村の飲食店数

## 4. 分析結果

### 4.1 結果・考察

今回上述の手法で抽出できたイベント参加者は延べ計 10,570 人、滞留点は 19,578 点である。表 4 にイベント開催時間外における会場での滞留時間を示す。列名「イベント前比率」とは前後の滞留時間における、イベント前の滞留時間の割合であり、前後どちらにより長く滞留したかを表す。なお、各カテゴリーのイベント平均開始時刻は 17 時前後、平均終了時刻は 19 時前後である。表 4 から、前後合わせた滞留時間は音楽イベントが最も長く、時間比率は音楽イベントが特に高いと

わかる。これは会場にて主にイベント前に行われる限定グッズの物販が要因の一つと考えられる。

表-4 会場における開催時間外の滞留時間

カテゴリ	前後滞留時間 (分)	イベント前比率 (%)
音楽	160	93.8
サッカー	125	80.6
野球	83	87.7

※いずれも中央値

## 4.2 滞留点分布

図 4 にイベント前後の滞留点分布を示す。イベント前の滞留点は 6,198 点、イベント後の滞留点は 2,020 点である。イベント前は開始まで時間があることから会場以外にも分布しており、自宅から会場に直接来るのではなく都心をはじめとした、様々な場所に滞留していることがわかる。一方で、イベント後の滞留点は会場と隣接したメッシュ内で多く、会場最寄り駅周辺での飲食などを目的とした滞留が示唆される。

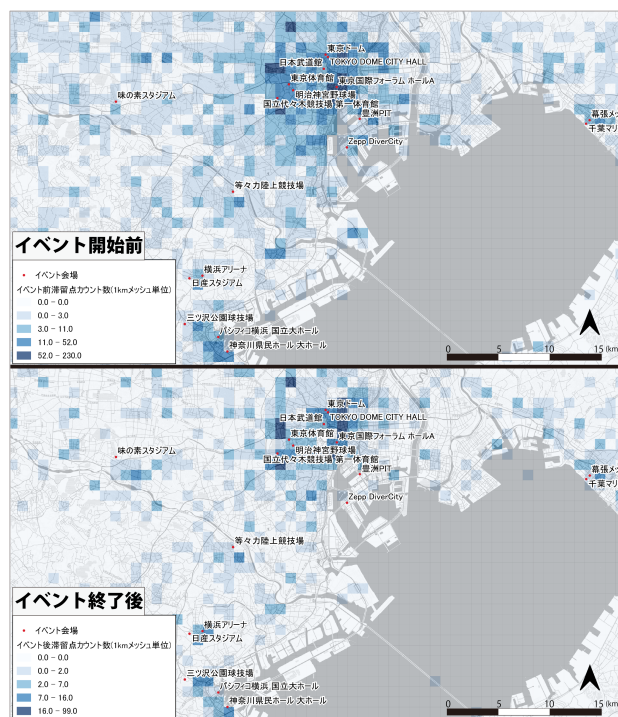


図-4 滞留点分布（都心部、1 km メッシュ単位）

## 4.3 重回帰分析

重回帰分析の結果は表 5 の通りである。決定係数が低いと相対的ではあるが、 $t$  値から滞留時間に最も影響する要素は「最寄り鉄道駅から

所要時間」とわかる。所要時間が増えるほど、イベント開始に遅れないよう会場付近の滞留時間が長くなることが考えられる。

表-5 重回帰分析結果 (n = 9,722)

変数名	偏回帰係数	$t$ 値	VIF
周辺観光資源数	-0.0089*	-6.375	4.152
交通アクセス	0.0219*	7.997	2.275
最寄り駅の発展度	0.0043*	3.819	3.608
飲食店数	0.0105*	4.622	2.133
定数項	4.4565*	-	-
Adj. $R^2$ = 0.025			

※\* $p < 0.1$

## 5. まとめ

本研究では人流ビッグデータを用いて、イベント参加前後の滞留行動を明らかにすることができた。今後、更なるモデルの改善を行うことでイベント参加前後の時空間的なニーズを把握でき、イベント参加前後に会場付近での消費行動を促し周辺地域の活性化に繋げることが可能である。

## 謝辞

本研究は株式会社日立製作所と国立大学法人東京大学による産学協創「日立東大ラボ」による研究の一貫として実施したものである。また、東大 CSIS 共同研究 No.794 の成果の一部として実施した。ここに記して謝意を表したい。

## 参考文献

- 堀野智寛, 伊藤香織, Burgess Andrew, 2018. ライブ・コンサート前の過ごし方にみる都市利用, 日本建築学会大会学術講演梗概集, 7158
- セントラル・リーグ 年度別入場者数(1950~2018), 日本野球機構, 2019. [http://npb.jp/statistics/attendance\\_yearly\\_cl.pdf](http://npb.jp/statistics/attendance_yearly_cl.pdf), アクセス日: 2019-08-26
- ライブ市場調査, コンサートプロモーターズ協会, 2017. <https://www.acpc.or.jp/marketing/transition/>, アクセス日: 2019-08-26
- 内閣府, 2009. 歩いて暮らせるまちづくりに関する世論調査, 世論調査報告書
- 2016 年 年間ライブ観客動員ランキング, Live-Fans, 2016. <https://www.livefans.jp/matome/17242>, アクセス日: 2019-08-26
- IHL Group, 2015. Impact of Store Networks and Wi-Fi on Customer Experience, 2015 Store Infrastructure Study