

サイバー空間における インビジブルシティの見える化方策 — つぶやきに基づく リアルデータ駆動型考驗から —

横田 尚己¹・谷口 航太郎²・大橋 瑞生³・谷口 守⁴

¹学生非会員 筑波大学大学院 システム情報工学研究科 (〒305-8573 茨城県つくば市天王台1-1-1)
E-mail: s1720524@s.tsukuba.ac.jp

²学生非会員 筑波大学大学院 システム情報工学研究科 (〒305-8573 茨城県つくば市天王台1-1-1)
E-mail: s1820472@s.tsukuba.ac.jp

³学生非会員 筑波大学 理工学群 (〒305-8573 茨城県つくば市天王台1-1-1)
E-mail: s1511237@s.tsukuba.ac.jp

⁴正会員 筑波大学教授 システム情報系 (〒305-8573 茨城県つくば市天王台1-1-1)
E-mail: mamoru@sk.tsukuba.ac.jp

現在、ビッグデータを活用したデータ駆動型社会の到来を迎えている。都市計画分野においてもその策定プロセスにおいてデータの利活用は必須になりつつある。特に“インビジブル”なサイバー空間上のコミュニティである SNS を通じて大量に人々の行動が、日々蓄積されている。そのような背景から、SNS 上の人々の言動を用いて行動と都市空間上における実際の活動との関係性を明らかにした研究が行われおり、SNS のデータから都市の特徴を抽出することに成功している。本研究では、サイバー空間上で人々の活動・行動が投影されやすい動詞を明らかにした上で、実際の都市空間の特性等との比較を行った。結果、一部の動詞で実際の都市空間の人口構成や空間特性が反映される特性がサイバー空間上からも見られることが明らかになった。

Key Words: invisible city, SNS, twitter, regional analysis, big data

1. はじめに

現在、国境を問わずビッグデータを収集・活用した効率的な社会、すなわちデータ駆動型社会を実現を目指した動きが高まっている。その中でも、都市計画分野は強い影響を受けると言われている。OECD のレポートによれば、データ駆動型の効率的な都市空間が実現することによる利益は、交通分野のみの観点であっても全世界で、年 7200 億米ドルから 9200 億ドルにも上ると言われている¹⁾。

特に、今後人口減少が進むわが国において、持続可能な社会を都市を構築する上でデータを活用した効率的な都市の構築・計画は、極めて重要であるといえる。

事実、わが国の第 5 期科学技術基本計画において“society5.0”という概念が提唱されており、サイバー空間とフィジカル空間を高度に融合させた超スマート社会の構築を目指している²⁾。その一例として平成 28 年熊本地

震が発生した際には、ソーシャルネットワークサービス（以下 SNS と略）の一種である Twitter から、災害情報を収集・要約を行うシステム DISAANA および D-SUMM が内閣府によって運用され、実際の復旧・復興活動に役立てられた³⁾。これはサイバー空間という見えない都市（以下インビジブルシティ）と実空間の都市との連動性が高まっていることの表れであるといえる。

以上のことを踏まえ、本研究は、本格的なデータ駆動型社会を迎える中で、インビジブルシティの一つである SNS 上に現れる大量の人々の言動から、都市空間内における行動・活動と連動する動詞を抽出し、さらにその動詞と共に現れる語を分析することによって、都市空間ごとの特徴を明らかにする。最終的に今後の都市計画を行う上での SNS を活用した新しい情報収集手法の有用性を示す。

2. 研究の位置づけ

(1) 既往研究

近年では、SNS の普及・発展に着目し SNS から得られる膨大なデータから都市分析を試みた研究がいくつか存在している。栗原・佐々木ら⁴⁾は Twitter に表出する地域に対する不満や、要望を都市ごとに抽出する手法を開発した。小池・林田ら⁵⁾は「食べる」と「買う」の類語表現をキーワードとした上で、そのキーワードが含まれるつぶやきを収集を行った。その上で収集したデータと実際の商業施設の施設数や、その面積・年間販売額との関係性について明らかにしている。

サイバー空間から地域の特徴を抽出した研究として小池・渡辺⁶⁾、谷口・星野ら⁷⁾の研究があげられる。同様に伊藤、荒川ら⁸⁾は Twitter を用いて地域ごとの特徴となる TF-IDF などの自然言語処理技術を活用し自動的に抽出している。さらに Lakshika・Sanjaya ら⁹⁾は Twitter のつぶやきの内容に着目し、ストリートギャングのつぶやきを抽出する基準を開発し、それらを防犯等に反映させることを提言している。また自治体の SNS の広報活動等における利活用に関する研究では河井¹⁰⁾や鬼塚・萩原ら¹¹⁾によって有用性と規模の小さな自治体での SNS 利活用の問題点が指摘されている。

このような既往研究がある一方で、その多くは分析対象地域が限定されていたり、分析対象期間に限られている。また、既往研究では収集したテキストデータに対して品詞ごとに単語に分解（以下形態素解析）を行った後、単語の単純集計を用いた分析が多く行われているが、単語の単純集計のみでは、必ずしも頻出語が都市空間内の行動や特徴を反映しているとは言い切れない。

(2) 本研究の特長

本研究は以下のような特長を有している。

- 1) データ駆動型の都市計画を考える上で都市空間とつぶやきの連動性を、行動・活動を表す動詞から明らかにする新規性。
- 2) 都市との連動性の高い動詞と共に用いられる名詞を明らかにすることで、各都市におけるインビジブルシティ上のコミュニティの見える化を行える有用性。
- 3) 事件やイベントなどの影響により、瞬間的に投稿（1 回の投稿では 140 文字まで投稿可能で、以下投稿をつぶやきと表記する）件数が伸びるという特徴をもつ Twitter のつぶやきを 1 ヶ月に渡って大規模に収集することにより、事件やイベントの影響を最小化し、定量的な分析を可能にする信頼性。
- 4) 都市空間内に現れない、人々の行動に紐づく動詞が明らかになることで、今後のインビジブルシティ研究に寄与する発展可能性。

3. 分析内容について

(1) 収集・分析するデータについて

表-1 は 2016 年以降日本国内において月間アクティブユーザー数を公表している SNS の特性を示したものである。SNS には多種多様なものが存在している。各種 SNS に共通している機能は不特定多数の人・企業・団体など SNS を介してつながることができる情報を受け取るだけでなく、誰もが特定もしくは不特定多数の人々に情報発信することができる。中でも Twitter は以下のような特徴を持つ。

- 1) アクティブユーザーの年齢に偏りが最も小さい SNS の 1 つであり、かつ月間アクティブユーザーが国内で 2 番目に多い SNS である。
- 2) 日本国内から毎秒 5~10 ツイートの高頻度でつぶやきがなされており、人々のライフスタイルと強く結びついている。
- 3) 基本的に匿名でつぶやきを行うことができ、ユーザーの本音が表出しやすくと考えられる SNS である。
- 4) つぶやきに位置情報が紐付いている。

以上のような特徴を持つため Twitter は都市情報を定量的に抽出するツールとして適していると考えられる。

(2) つぶやきの取得方法と分析対象期間

本研究では Twitter Stream API を用いて 2017 年 11 月に日本国内からつぶやかれた 1% のつぶやきをランダムに¹²⁾収集した。収集したデータ例を表-2 に示す。なお、表-2 では Screen_name は投稿者の Twitter 上の名前であり、個人の特定につながる可能性があるため、A・B・C に置き換えている。今回のデータ収集はつぶやき全体の 1% に限定されるが、その量は約 358 万件、クリーニング済み段階でも 256 万件を上回っており、十分にその量は担保されている。

本研究におけるつぶやきの収集期間の設定理由は、10・11 月は連休等が少なく、日常的な行動が現れやすい時期であり、国勢調査や全国都市交通特性調査等の各種統計調査が行われる時期とも重なるからである。ただし、2017 年 10 月には衆議院選挙が行われているため、日常的なつぶやきの割合が他の期間と比べ、少ないと考えら

表-1 SNS の特質(表¹²⁾を参考に筆者作成)

SNS	特質	特徴					日本国内の月間アクティブユーザー数*
		主要機能	位置情報	利用頻度	拡散性	匿名性	
LINE	a, b, c, d	無	高	有	無	無	7000万人以上 (2017年9月)
Twitter	a, c, d	有	高	有	高	有	4500万人 (2017年10月)
Facebook	a, b, c, d	有	低	有	無	無	2800万人 (2017年10月)
Instagram	a, b, c, d	有	低	有	中	無	2000万人 (2017年10月)

a=交流, b=記録(ライフログ), c=写真・動画, d=連絡ツール, e=ゲーム

*各SNSでのHPの最新情報に基づく

表-2 Twitter 取得データ例

screen name	投稿場所	つぶやき	時刻
A	東京 千代田区	ついでで	2017/11/1 12:01
B	福岡 福岡市中央区	心はいつも平常心	2017/11/1 12:01
C	千葉 千葉市中央区	アベンジャー適性もあるのか	2017/11/1 12:01

表-3 分析対象都市一覧

分析対象都市			
札幌市	上尾市	会津若松市	三豊市
仙台市	足利市	川西市	美濃加茂市
堺市	習志野市	石垣市	交野市
姫路市	弘前市	延岡市	東温市
倉敷市	下関市	伊東市	新庄市
相模原市	流山市	敦賀市	熊本県小国町
豊中市	磐田市	軽井沢町	



図-1 分析対象となる動詞の選定フロー

表-4 分析対象の動詞選定のための名詞リスト

種別	キーワード
交通機関	「車」「バス」「バイク」「自転車」「電車」「新幹線」「飛行機」「航空機」「駅」「空港」
場所性を表す単語	具体的な地名のほかに、「地元」「近所」「会社」「仕事」「職場」「学校」「中学」「高校」「大学」「部活」「バイト」「家」「店」「カラオケ」「劇場」
イベントに関連する単語	「ライブ」「イベント」「祭」「大会」「花火」「試合」

$$P_{ac} = \frac{N_{ac}}{N_{av}} \quad (1)$$

$$\overline{P_C} \geq P_{ac} \quad (2)$$

N_{av} : 動詞Aの出現回数

N_{ac} : 動詞Aと表-4に含まれる名詞との共起回数

P_{ac} : 動詞Aの出現回数に対する動詞Aと表-4に含まれる名詞との共起回数の割合

$\overline{P_C}$: P_{ac} の平均

れる。したがって本研究では 2017 年 11 月に日本国内から寄せられたつぶやきを分析対象とした。

(3) 分析対象都市と分析方法

分析対象都市については様々な属性を持った都市が抽出されていると考えられる谷口・星野ら⁷⁾の分析対象都市と同じ都市を選んだ(表-3 参照)。そして収集したデータを図-1 のフローと、表-4 の行動選定名詞リストを用いてクリーニングを行い、都市空間での行動が強く反映される動詞(以下行動共起動詞)の特定を行う。なおテキストの形態素解には樋口が開発した KH Coder¹³⁾を用いた。

また表-4 のリストは動詞の目的語、つまり名詞が都市空間の施設名や行動を表すと考えられるものをリスト化した。具体的には都市空間内での移動行動が見られる「交通機関」、そして都市空間内での具体的な行動を起こす場所が現れる「場所性」を表すキーワード、そして都市空間内で発生する「イベント」に関連するキーワードを選んだ。ただし、都市空間における場所性を表したイベントに関する名詞は無数にあり、表-4 のリストが必ずしもすべてを網羅できていない点を留意する必要がある。

最後に同じつぶやき内で出現した単語どうし(共起)を線(ノード)で結び、グラフ理論に基づいた KH Coder の共起ネットワーク作成機能を用いて、選定した動詞と、共起回数が上位 100 位名詞との関係性を図-3~図-6 に各単語から発生するノードの多さ(以下次数中心性)からつぶやき内での単語の重要性を評価する。

4. Twitter を用いた分析結果

(1) 行動共起動詞に関する分析

図-2 に各分析対象都市一覧とつぶやき数、および動詞の出現回数を示す。続いて全分析対象都市のつぶやきの頻出上位 30 位の動詞と表-4 の行動選定名詞リストと次数中心性を見る。表-5 はその関係性の強さを表すものである。特に背景が黄色になっている動詞は、平均より表-4 の名詞と同じつぶやき内で用いられている回数(以下共起回数)の割合が多いものである。併せて図-3 では動詞のなかで頻出上位 1 位、2 位である「行く」および「思う」との共起回数が上位 100 の名詞との次数中心性について明らかにしたものである。表-5 および図-3 より明らかになったことは以下の通りである。

- 1) 「行く」に着目すると表-4 の名詞のうち 31 種類と共起されている。特に「京都」「東京」のように具体的な地名 8 種類と共起している。
- 2) 「思う」と共起されている表-4 の名詞は 14 種類あ

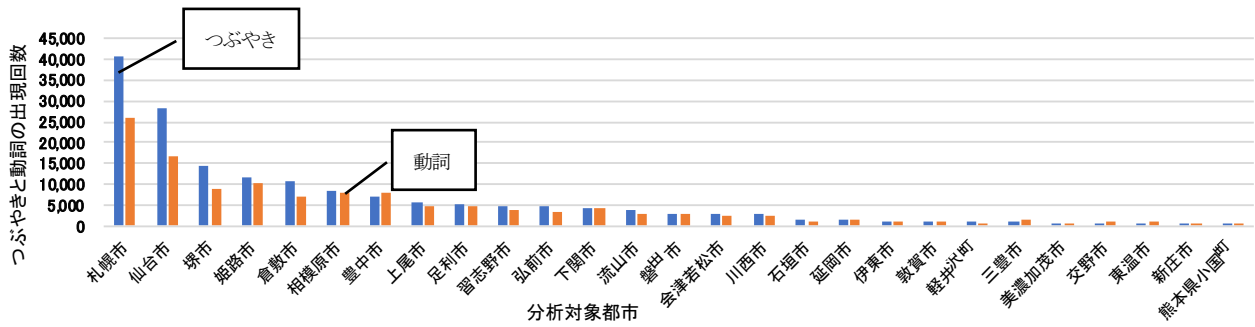


図-2 分析対象都市におけるつばやき数および動詞の出現回数

表-5 上位頻出 30 位の動詞の出現回数と表-4 のリストとの重複回数

動詞	動詞の出現回数	表-4との共起回数	共起割合	動詞	動詞の出現回数	表-4との共起回数	共起割合	動詞	動詞の出現回数	表-4との共起回数	共起割合
行く	4108	986	24%	帰る	1262	358	28%	会う	646	25	4%
思う	3535	182	5%	終わる	1182	361	31%	楽しむ	621	86	14%
見る	3133	172	5%	出来る	883	78	9%	行う	574	91	16%
言う	2537	71	3%	入る	864	38	4%	考える	561	5	1%
頑張る	1710	340	20%	知る	806	18	2%	撮る	557	20	4%
食べる	1604	58	4%	飲む	779	11	1%	乗る	552	153	28%
出る	1585	161	10%	使う	762	6	1%	分かる	541	32	6%
寝る	1573	93	6%	起きる	705	34	5%	遊ぶ	518	23	4%
来る	1535	282	18%	聞く	703	5	1%	走る	502	49	10%
買う	1335	59	4%	作る	697	18	3%	待つ	497	20	4%

背景黄色は共起回数の割合が平均以上のもの=行動共起動詞

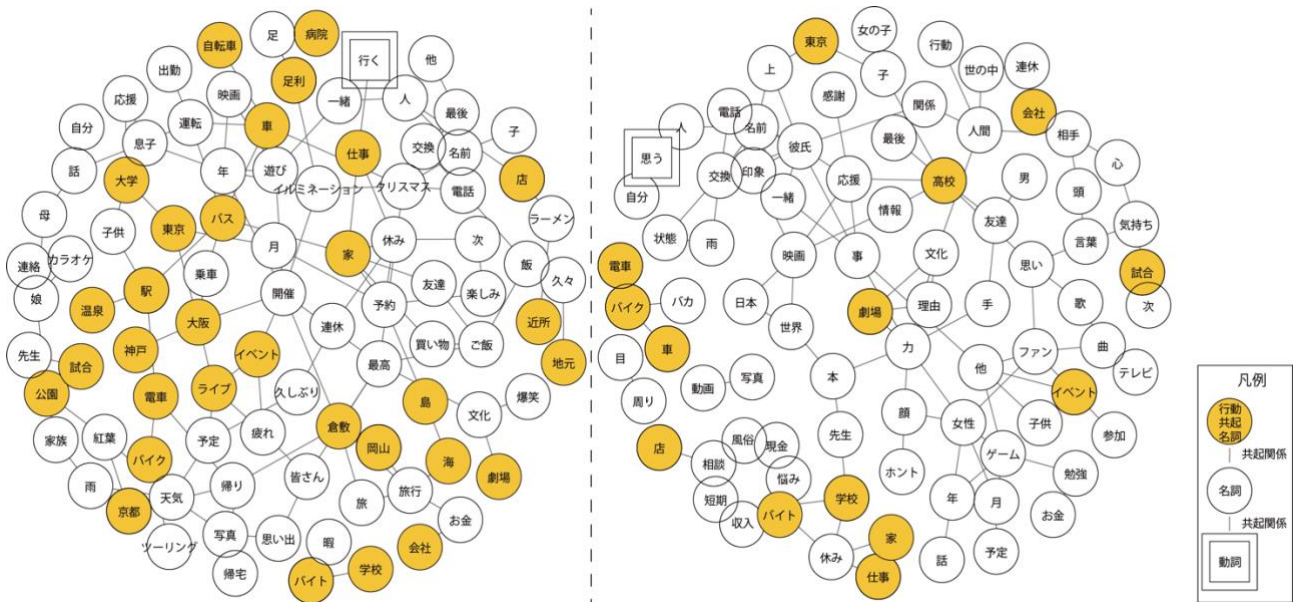


図-3 「行く」(左)と「思う」(右)における名詞の共起の比較

り、表-4の名詞との共起の様子は、「行く」と比べて都市との関係性が弱いことが分かる。一方で「高校」が6つの名詞と共起していることや「人間」「関係」などに表されるように、人間関係に関連した名詞が多い。従って「思う」という動詞は都市空間内での人々の行動が投影されにくいといえる。

(2) 分析対象都市での行動共起動詞と次数中心性

分析対象都市ごとに分析した際に、顕著な特徴が現れた札幌市・仙台市・相模原市の分析結果を図-4、図-5、図-6に示す。なお、これらの図は行動共起動詞とその共起回数上位100の名詞をプロットし、その次数中心性について明らかにしたものである。

しかし、3.(3)で前述したように本研究ではすべての都市施設等を網羅できていない。今後、表-4の「場所性を表す単語」等の名詞の客観的な収集手法について再度検討する必要がある。

謝辞

本研究ではJSPS科学研究費(16K14315)の助成を受けた。御礼申し上げる。

参考文献

- 1) OECD (大磯一, 入江晃史監訳, 斎藤長行, 田中絵麻訳) : OECD ビッグデータ白書 データ駆動型イノベーションが開く未来社会, p419, 明石出版, 2018.
- 2) 内閣府 : 第5期科学技術基本計画, p11, 2016.
- 3) 国立研究開発法人情報通信研究機構 : DISANAA とは, <http://disaana.jp/rtime/search4pc.jsp>(最終閲覧日 2018.07)
- 4) 栗原理聡, 佐々木彬, 松田耕史, 岡崎直観, 乾健太郎, Twitter を利用した地域毎の要望抽出, 人工知能学会全国大会論文集人工知能学会全国大会論文集 29, 1-4, 2015.
- 5) 小池 太輔, 林田和人, 渡辺仁史, ソーシャルメディアに現れる行動と実空間における生活行為の関係, 日本建築学会学術講演梗概集 F-1 都市計画 建築社会システム, 巻 2011, pp573-574, 2011.
- 6) 小池 太輔, 渡辺仁史, ソーシャルメディアのテキスト情報に現れる都市イメージの抽出, 研究報告 82(2), 457-460, 一般社団法人日本建築学会, 2012.
- 7) 谷口守, 星野奈月, 富永透「つぶやき」から見る都市についての一考察 -都市名に着目して-, 日本都市計画学会論文集 Vol.29 No.3, pp.939-944, 2014.
- 8) 伊藤晶, 荒川豊, 田頭茂明, 福田晃, Twitter からの地域特徴語の自動検出に関する一検討, 第75回情報処理学会全国大会講演論文集 2013(1), 101-103, 2013.
- 9) Lakshika Balasuriya, Sanjaya Wijeratne, Derek Doran, Amit Sheth, Finding Street Gang Members on Twitter, IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), pp685-692, 2016.
- 10) 河井考仁, 地方自治体によるソーシャルメディアの活用について, 社会情報学会(SSII)学会大会研究発表論文集, 123-126, 2013.
- 11) 鬼塚健一郎, 萩原和, 星野敏, 清水夏樹, 橋本禪, Facebook を活用したワークショップの効果と課題 -京丹後市五十河地域における取り組みを通じて-, 農村計画学会誌 2014 年 32 巻 4 号 p. 507-516, 2014.
- 12) Twitter Developers : Sample realtime Tweets, https://developer.twitter.com/en/docs/tweets/sample-realtime/overview/GET_status_sample.html, (最終閲覧 2018 年 6 月)
- 13) 樋口耕一, KH Coder2.x リファレンスマニュアル (2013 年 8 月 5 日)
- 14) 文部科学省, 平成 29 年学校基本調査
- 15) 環境省 : 3.21 都市公園の現況 (都道府県別), https://www.env.go.jp/doc/toukei/contents/tbldata/h29/2017_3_21.xls(最終閲覧 2018 年 7 月)
- 16) 仙台市 : 公園イベント紹介, <http://www.city.sendai.jp/aoba-koen-somu/aobaku/koen/shokai.html>(最終閲覧 2018 年 7 月)