

「Wi-Fi 人口統計データ」の位置精度に関する初期検証 村上広史*・森下直哉**・忽滑谷優里***

A Preliminary Study on the Positional Accuracy of “Wi-Fi Demography Data”

Hiroshi Murakami*, Naoya Morishita**, Yuri Nukariya***

Abstract: The widespread use of smartphones has enabled the emergence of many applications that take advantage of positional information of smartphones. While there are different ways to acquire or estimate smartphone’s position, using the access logs of Wi-Fi access points looks promising as they allow seamless analyses of human movement between indoors and outdoors. This paper conducted a preliminary study on the accuracy of estimated positions of access points provided in “Wi-Fi Demography Data,” one of the positional information products based on Wi-Fi access logs of smartphones. The coordinates (latitude and longitude) of 189 access points in Shibuya, Tokyo were surveyed in the field and compared with those estimated in the logs, and their differences were calculated. The average of these differences was 17 m. Then the comparison was made between the differences and the standard deviation (SD) of estimated coordinates internally calculated in the logs for each access point. The result suggests that SD can be used to eliminate access points that have large positional errors, and hence to improve the overall positional accuracy of “Wi-Fi Demography Data.”

Keywords: 携帯端末 (smartphone), アクセスログ (access log), アクセスポイント (access point), 位置精度 (positional accuracy)

1. はじめに

携帯端末の普及により、その位置情報を活用して特定の場所の人流を分析する取組が拡大している。例えば、COVID-19 感染拡大を抑制するための人流抑制効果については、携帯端末の位置情報を用いた人流の分析結果が政府のウェブサイトでも随時公開されている (<https://corona.go.jp/dashboard/>)。また、イベント参加者の行動分析や観光地間の人の流れの分析など多様な分野での活用が進んでいる (小林ほか, 2019, 相原, 2017)。

一方、携帯端末の位置情報の活用については、個人情報やプライバシーとの関係が懸念されるが、総務省によって法的な整理も含めて検討が進んでいる (総務省, 2014)。また、関連する民間企業も位置情報を活用したマーケティングやサービス施策の促進を目的とする事業者団体を設立し、位置情報等の利活用に関するガイドラインを定めている (LBMA Japan, 2020)。このため、携帯端末の位置情報は、今後プライバシーへの適切な配慮のもとで、その活用分野や範囲が一層拡大していくと考えられる。

携帯端末の位置情報の中でも利用が進んでいるも

のとしては、携帯電話基地局を活用したもの、GPS 位置情報及び Wi-Fi 位置情報の 3 種類があり、位置精度や活用可能範囲など、それぞれの種類の特性を活かした活用が行われている (総務省, 2014)。最近ではこれらに加え、Beacon の活用も進んでいる。このうち、Wi-Fi 位置情報は、都市部を中心とした無料 Wi-Fi などのアクセスポイント (AP) の増加により、今後活用の拡大が想定されることから、その特性を理解することは今後の適切な活用を図る上で有用と考えられる。特に、無料 Wi-Fi の AP への自動接続アプリにより収集されたアクセスログのデータは、屋内外を問わず特定店舗への利用者の来店を推定することが可能であるため、その位置精度の分析は、今後のアクセスログデータの活用により有用な知見を提供すると考えられる。

そこで、本研究ではアクセスログデータの一つとして国際航業株式会社から販売されている「Wi-Fi 人口統計データ」のうち、個々の利用者の位置が推定されているポイントデータ (以下、「ログデータ」) に含まれる AP の推定位置 (緯度経度) と現地調査で得られた該当する AP の位置の比較により、AP の

* 正会員 青山学院大学地球社会共生学部 (Aoyama Gakuin University)

〒252-5258 神奈川県相模原市中央区淵野辺 5-10-1 E-mail : hiroshi_murakami@gsc.aoyama.ac.jp

** 学生会員 青山学院大学地球社会共生学部 (Aoyama Gakuin University)

*** 非会員 国際航業株式会社

位置情報の精度検証を行った。

2. ログデータと現地調査

使用したログデータの範囲は図-1に示す研究対象範囲（東京都の渋谷駅，表参道駅，明治神宮前駅によって囲まれた地域で，6次メッシュで164メッシュ分）である。



図-1 研究対象範囲（背景に地理院地図を使用）

2.1. ログデータ

ログデータには，Wi-Fi 自動接続アプリの利用者が特定されないようにハッシュ化された利用者ごとのID，APにアクセスした日時，APのSSID（Service Set Identifier），APの推定位置（緯度及び経度），利用者の年代及び性別などが記録されている。このうち，SSIDはWi-Fiサービスを提供する店舗等のAP全体の総称であるため，例えばコンビニチェーン店のWi-FiサービスのSSIDが分かっても，個別のコンビニ店舗を特定することはできない。

一方，APの推定位置は，個々のAP機器固有のIDであるBSSID（Basic Service Set Identifier；MACアドレスと同一の場合が多い）ごとに，利用者がアクセスした際の携帯端末の位置情報（GPSによる緯度経度）から統計処理により計算されている。このBSSIDとAPの推定位置及び該当するSSIDの参照表は内部データベースとして随時更新され，ログデータの作成に用いられている。このため，複数のBSSIDを持つWi-Fiサービスが同一店舗から同一SSIDで提供されている場合，そのSSIDに対して位置が必ずしも一致しない複数の推定位置のデータが記録されることになる。また，同一SSIDを使用する複数のチェーン店が近接して存在する場合，SSIDだけでどの店舗へのアクセスかを判断することが困難な場合も発生し得る。

したがって，ログデータの位置精度を検証するには，個々のAPのBSSIDとその位置の情報を現地で

取得し，ログデータ作成の過程で 사용되는内部データベースに記録されたBSSIDと照合することで，個別APごとの推定位置と現地調査で取得した位置から位置精度を分析する必要がある。

2.2. 現地調査

APごとのBSSIDを調査するには実際に現地でAPからの信号に含まれるBSSIDを記録するとともに，その場所の位置を計測する必要がある。一方，APには，コンビニチェーン店のように店舗単位で設置されていて現地での位置が明確なものと，携帯電話会社等が契約者向けに設置しているAPで現地での位置が特定しにくいものと大きく2種類がある。このため，現地調査においては，BSSIDの調査とともに，現地での位置確認の可否を記録した。位置精度の検証は位置確認が可能なAPに対して行った。

現地調査は2021年4月1日，17日及び24日の3日間に，それぞれ1人，8人，4人で徒歩により実施した。複数で行った場合は，同一SSIDが発する複数のBSSIDの情報を捕捉できるように，全員が一緒に行動した。BSSID情報の取得には携帯端末アプリ“Net Analyzer”を活用し，該当する店舗の位置をWeb地図上で確認して緯度経度を記録した。

現地で収集したユニークなBSSIDは329点であり，このうち携帯電話会社や大学により設置されたものを除き，現地でAPの位置が確認できたものは189点であった。これらの中にはビル内や入口が地下にあるものなど，GPSの電波が届きにくい店舗が3点あったが，他は上空視界の確保が可能な道路に直結する店舗であった。現地調査の軌跡と取得したBSSIDの位置を図-2に示す。

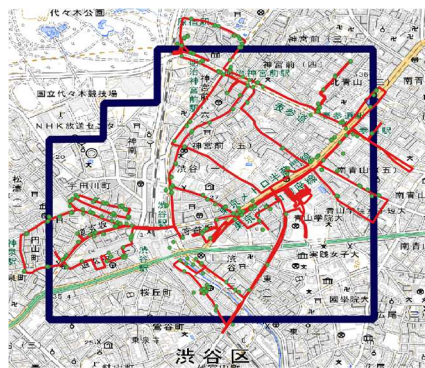


図-2 現地調査軌跡：赤線は現地調査の軌跡；緑点は位置を特定したAP（背景に地理院地図を使用）

3. 位置精度の検証

3.1. 現地調査による位置精度検証結果

現地で場所が確認できた 189 点の BSSID に対して、ログデータの AP 推定位置（緯度経度）と現地調査で確認した当該 AP の位置（緯度経度）の距離（以下、「推定位置誤差」）を国土地理院の「距離と方位角の計算」サイトで計算した結果のヒストグラムを図-3 に示す。これら 189 点の AP のうち推定位置誤差が 260m 以上の点が 2 点確認されたが、他は全て 60m 未満となった。189 点の推定位置誤差の平均は 17m、標準偏差は 28m である。推定位置誤差が大きな 2 点を除いた 187 点の平均は 15m、標準偏差は 10m である。

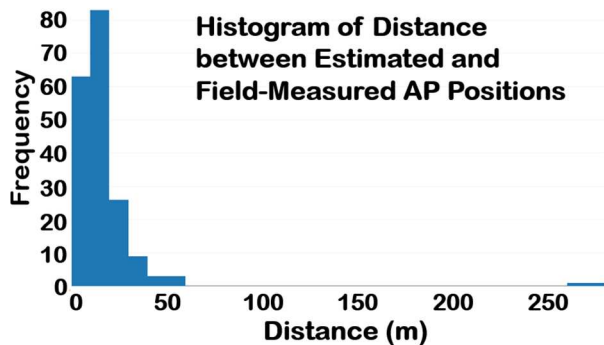


図-3 ログデータの AP 推定位置と現地調査による AP 位置の距離（推定位置誤差）のヒストグラム

3.2. 推定位置の標準偏差

一方、ログデータの作成の際には、各 AP の推定位置を計算する際に用いた複数の携帯端末の位置情報の標準偏差（以下、「推定位置 SD」）を計算し、ログデータの位置精度の目安として内部利用されている。ただし、携帯端末の位置情報が 1 件しか得られなかった AP に対しては推定位置の標準偏差は計算されていない。現地調査で BSSID が取得された AP のうち、推定位置 SD の計算が行われなかった 6 点の AP を除いた残りの 183 点に対する AP の推定位置 SD のヒストグラムは図-4 のとおりである。これら 183 点の推定位置 SD の平均は 17m、標準偏差は 21m である。これらのうち、推定位置精度が 100m を超えた点は 2 点で、212m と 120m であり、他は全て 81m 以下である。また、これら 2 点は推定位置誤差が 260m 以上あった上述の 2 点と同一の AP である。推定位置誤差が大きな 2 点を除いた 181 点の推定位置 SD の平均は 15m、標準偏差は 13m である。

推定位置 SD が計算された 183 点について、推定位置誤差との間を散布図で示したのが、図-5 である。そして、推定位置誤差が 260m 以上あった 2 点を除いた場合の同様な散布図は図-6 のとおりである。

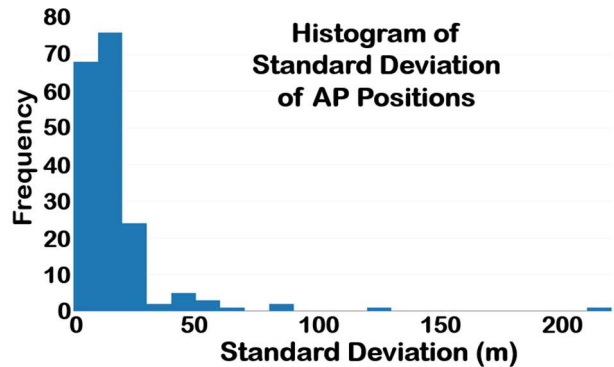


図-4 現地調査で BSSID が取得された AP の推定位置精度のヒストグラム

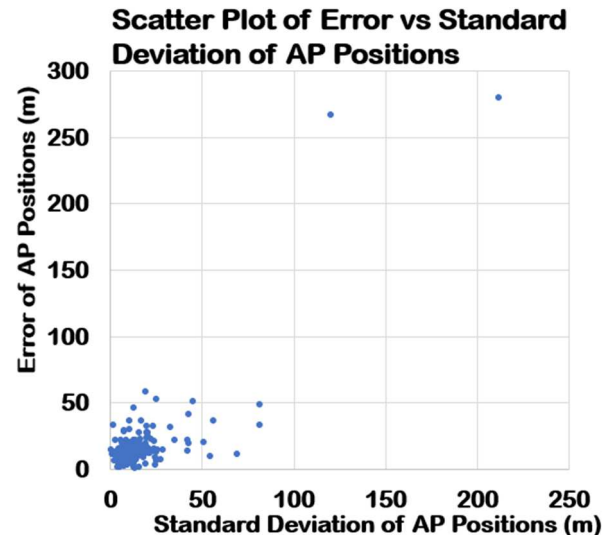


図-5 推定位置誤差と推定位置 SD の散布図

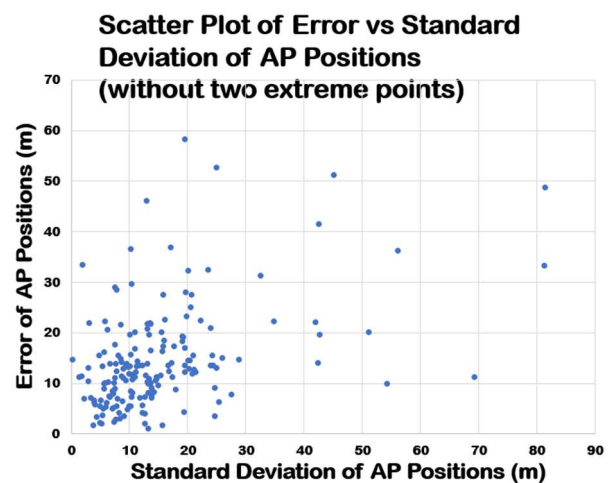


図-6 誤差の大きな 2 点を除いた推定位置誤差と推定位置 SD の散布図

また、推定位置誤差と推定位置 SD の空間分布をそれぞれ図-7 と図-8 に示す。高層ビルが多い地域や道路幅が狭い地域では、衛星測位の精度が低下すると考えられるが、定性的な評価ではあるものの、推定位置誤差と推定位置 SD に地域的な影響や互いの空間分布における明確な類似性は認め難い。

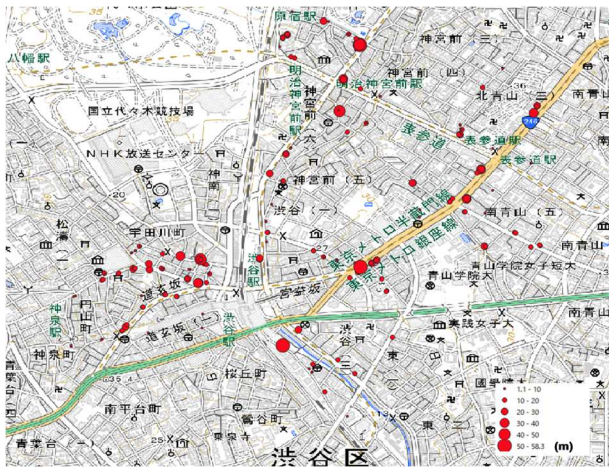


図-7 誤差の大きな2点を除いた推定位置誤差の空間分布（背景に地理院地図を使用）

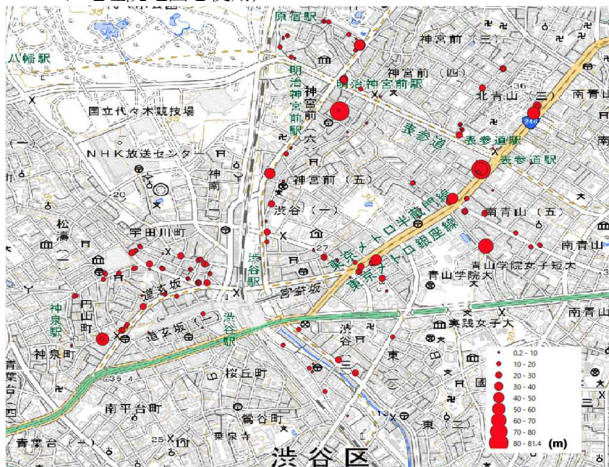


図-8 誤差の大きな2点を除いた推定位置 SD の空間分布（背景に地理院地図を使用）

4. 考察とまとめ

渋谷地域の現地調査により、現地で位置の確認が容易な AP についてログデータの推定位置精度を検証した結果、誤差の大きな AP は存在するが、全体として 30m 程度の誤差に収まっていることが確認できた。また、誤差が極端に大きな AP についても、内部データである推定位置 SD を活用することで取り除くことが可能であることが分かった。今回の精度検証は、現地で位置が確認できる AP のみを対象としたが、ログデータの推定位置計算方法は現地での位置確認の可否とは独立であることから、すべて

の AP に対して同様な位置精度を期待できるものと考えられる。

また、図-6 からは推定位置誤差と推定位置 SD の間には弱い正の相関がみられた（相関係数：0.44）。各 AP の推定位置の系統誤差が AP 間で大きく変わらない場合には、推定位置誤差と推定位置 SD が相関関係を持つことが期待できるため、そのことをある程度反映しているものと考えられる。このため、上述のように推定位置 SD を活用することで、誤差が極端に大きくない AP のみのログデータを抽出することは可能であると考えられる。

今回の現地調査で対象とした AP は、そのほとんどが屋外から直接アクセス可能で GPS での測位に大きな支障がないものであった。Wi-Fi による位置推定の長所である屋内 AP のログデータの位置精度検証については今後の課題としたい。

謝辞

本稿をまとめるにあたり、Wi-Fi 自動接続アプリ事業者の協力を得てログデータの分析を行った。また、現地調査においては青山学院大学地球社会共生学部の3年生有志の協力を得た。ここに謝意を表する次第である。

参考文献

- 相原健郎 (2017) ビッグデータを用いた観光動態把握とその活用：動体データで訪日外客の動きをとらえる、「情報管理」, 59(11), 743-754.
- 小林稜介・宮澤聡・秋山祐樹・柴崎亮介 (2019) 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, 第28回地理情報システム学会研究発表大会予稿集.
- 総務省(2014)位置情報プライバシーレポート～位置情報に関するプライバシーの適切な保護と社会的利活用の両立に向けて～, 緊急時等における位置情報の取扱いに関する検討会 報告書(平成26年7月).
- LBMA Japan (2020) 位置情報等の「デバイスロケーションデータ」利活用に関するガイドライン, <https://www.lbmajapan.com/guideline>.