

機械学習を用いた無線 LAN 屋内測位における精度評価

竹内 裕幸[†] 平川 豊[‡] 大関 和夫[‡]
 芝浦工業大学大学院理工学研究科[†] 芝浦工業大学工学部[‡]

1. 研究背景

近年、位置情報を用いたサービスが普及しており、ナビゲーションシステムや SNS や在室管理といった場面で活用されている。位置測位では一般的に GPS が用いられているが、屋内での利用においては大きな誤差が生じてしまう問題がある。そのため、屋内では GPS 以外のデータを用いた位置測位が求められるが、高い位置測位精度を出すことが課題となっている。本研究では、屋内測位において利用者が特定の部屋に在室しているかどうかを判定する問題に着目し精度評価を行う。

2. 関連研究

屋内測位はこれまでも多くの研究がなされており、Bluetooth 発信機を用いた研究[1]や出席管理を目的として無線 LAN アクセスポイントのデータを利用した研究[2]がある。これらは受信信号強度(以下, RSSI)のデータを利用することで位置を算出するものである。RSSI 値は測定場所により変動するため、アルゴリズムを適応させることで位置測位の精度向上を行うことができる[3]。

3. 提案手法

本研究では、無線 LAN アクセスポイントからの識別子(以下, SSID), MAC アドレス(以下, BSSID), RSSI 値のデータを取得することにより無線 LAN 屋内測位を行う。データの解析にあたっては、近年の計算機技術の向上に伴い注目を集めている機械学習を用いて、取得したデータから在室か不在かの分類を行う。その際、効率的な精度向上に繋がるデータ整形手法を提案する。本研究の進め方を以下に示す。

- ① 在室判定を行いたい部屋(以下, 判定部屋)とそれ以外の部屋において、スマートフォンを用いて無線 LAN アクセスポイントからのデータを取得する。なお、データ取得は芝浦工業大学豊洲キャンパス研究棟で行う。
- ② 取得したデータに対して機械学習を適応させるためにデータを整形する。
- ③ R 言語[4]を用いてニューラルネットワークによる分類推定を行う。
- ④ ②のデータ整形手法を変更し、再度③を行うことで分類精度の違いを検証する。

4. 実験

4.1. データ取得

無線 LAN アクセスポイントのデータを計測する手法として、Android スマートフォンを用いた。アプリケーションの開発には、「Android Studio(Ver2.2.2)」[5]を用いた。使用した端末は「Xperia Z5 Compact(SO-02H, Android 7.0)」であり、取得データは SSID, BSSID, RSSI である。これらのデータを部屋ごとに取得する。なお、データ取得の際、データ整形を効率的に行うために RSSI 値で降順に並び替える処理を行っている。取得したデータ例を図 1 に示す。

```
[1times]
SSID_NAME_0, xx:xx:xx:yy:yy:yy, -42
SSID_NAME_1, yy:yy:yy:zz:zz:zz, -43
...
SSID_NAME_30, xx:zz:zz:zz:yy:yy, -87
[2times]
SSID_NAME_2, zz:xx:xx:xx:zz:zz, -42
SSID_NAME_0, xx:xx:xx:yy:yy:yy, -44
...
[1000times]
...
SSID_NAME_25, xx:zz:zz:xx:xx:yy, -86
```

図 1. 取得データ

4.2. データ整形

取得データに対して以下の整形を行った。

- A) 各部屋における RSSI 値の上位からデータを抽出し、判定部屋には在室ラベル(Status 1)、それ以外の部屋には不在ラベル(Status 0)をそれぞれ付与する。整形したデータセット例(以下, データセット A)を表 1 に示す。

表 1. データセット A

Sta.	(SSID, BSSID, RSSI)			(SSID, ...)			(..., RSSI)
1	nwLab SSID	nwLab BSSID	nwLab RSSI	sitLan SSID	itLab RSSI
0	sitLan SSID	sitLan BSSID	sitLan RSSI	nwLab SSID	seLab RSSI
...
1	nwLab SSID	nwLab BSSID	nwLab RSSI	itLab SSID	seLab RSSI

表 1 の最上位データは、判定部屋でデータ取得した際、RSSI 値の順が「nwLab, sitLan, ..., itLab」の場合に作成されるデータである。2 番目のデータ行は、判定部屋以外でデータ取得した際、RSSI 値の順が「sitLan, nwLab, ..., seLab」の場合に作成されるデータである。データ取得時の RSSI 値により作成されるデータセットの並びは異なる。SSID, BSSID, RSSI 値の 3 列で 1 セットとなっており、これを 10 セット作成した。なお、分類推定の際には SSID, BSSID をそれぞれ固有の整数値に変換させてから行う。

Evaluation of Accuracy on Wireless LAN Indoor Positioning Using Machine Learning

[†]Hiroiyuki Takeuchi, [‡]Yutaka Hirakawa, [‡]Kazuo Ohzeki

[†]Electrical Engineering and Computer Science, Shibaura Institute of Technology, Tokyo, Japan

[‡]Information Science and Engineering, Shibaura Institute of Technology, Tokyo, Japan

B) 判定部屋での全取得データに対して、事前に SSID ごとの RSSI 値の平均を求める。そして、判定部屋における RSSI 値が平均的に大きい方から 10 個の SSID の並び順を決める。次に、測定したデータから、並び順を決めた SSID の RSSI 値を格納したデータを作成し、在室のラベル(Status 1)を付与する。

判定部屋以外でのデータに対しても、判定部屋で並び順を決めた SSID の RSSI 値を格納したデータを作成し、不在のラベル(Status 0)を付与する。なお、並び順を決めた SSID が取得できなかったデータは、RSSI 値が 0 として埋めた。整形したデータセット例(以下、データセット B)を表 2 に示す。

表 2. データセット B

Status	nwLab SSID	sitLan SSID	...	seLab SSID
1	-43	-42	...	-48
0	-68	-69	...	-47
...
1	-40	0	...	-75

表 2 は、判定部屋における RSSI 値の平均が「nwLab, sitLan, ..., seLab」の順に大きかった場合に作成されるデータセットである。どの部屋のデータに対しても、事前に判定部屋で求めた SSID の並び順に沿って RSSI 値が格納される。

4.3. 機械学習

4.2 章で整形したデータセットそれぞれに対して、ニューラルネットワークによる分類[6]を行う。表 4 に学習環境を示す。

表 3. 学習環境

言語	R version 3.4.1
在室データ数	100
不在データ数	100
隠れ層のユニット数	4
最大繰り返し演算数	500

分類の問題を解くニューラルネットワークでは、出力ユニットからの出力信号にロジスティックシグモイド関数が適用されており、予測値は在室グループに属する確率値として得られる。本研究では 2 値分類を扱っているため、50%以上を在室と判定させた。この分類推定を一定数行い、それぞれの分類精度の平均を求めた。なお、データセット B の分類推定では SSID 上位 10 個まで(B-10)、5 個(B-5)まで、3 個(B-3)までの 3 つのパターンを作成した。表 4~7 にそれぞれのデータセットでの結果を示す。

表 4. 分類結果平均(データセット A)

	不在データ(0)	在室データ(1)
不在(0)と判定	76.2	7.6
在室(1)と判定	23.8	92.4

表 5. 分類結果平均(データセット B-10)

	不在データ(0)	在室データ(1)
不在(0)と判定	97.9	0
在室(1)と判定	2.1	100

表 6. 分類結果平均(データセット B-5)

	不在データ(0)	在室データ(1)
不在(0)と判定	99.1	0.2
在室(1)と判定	0.9	99.8

表 7. 分類結果平均(データセット B-3)

	不在データ(0)	在室データ(1)
不在(0)と判定	98.2	9.8
在室(1)と判定	1.8	90.2

データセット A では、不在のデータを在室していると誤判定してしまっている確率が高くなり、76.2%の精度となった。

データセット B では分類精度が高くなる結果となった。上位 SSID 数が多いほど分類精度が高い傾向にあるが、データセット B-5 が約 99%となり、最も高い分類精度となった。

5. まとめと今後の課題

本研究では、無線 LAN 屋内測位において、アクセスポイントのデータから在室判定を行う問題に対して機械学習を用いて精度評価を行った。実験の結果より、取得データの整形は、判定部屋での RSSI 値が大きい SSID の並び順を決定し、上位 5 個のデータセットで学習する手法が有効となった。

今後の課題としては、データセット B での SSID 数別の分類において、検証データ数を増やすことで信頼区間やより詳細な傾向分析をする必要がある。

参考文献

- [1] 古館達也, 堀川三好, 菅原光政, “受信信号強度を用いた屋内測位手法の提案”, 情報処理学会研究報告, Vol.2014-MBL-73, No.21, pp.1-8, 2014-11
- [2] 田中成典, 中村健二, 北川悦司, 井上晴可, 野村一葵, “無線通信を用いた出席管理アプリケーションの開発”, 情報処理学会第 75 回全国大会講演論文集, Vol.2013, No.1, pp.589-590, 2013-3
- [3] Weixing Xue, Weining Qiu, Xianghong Hua, and Kegen Yu, Senior Member, IEEE, “Improved Wi-Fi RSSI Measurement for Indoor Localization”, IEEE Sensors Journal, Vol.17, No.7, pp.2224-2230, April 1 2017
- [4] R: The R Project for Statistical Computing, <https://www.r-project.org/>, 2017-12
- [5] Android Studio: Android の公式 IDE, <https://developer.android.com/studio/index.html?hl=ja>, 2017-10
- [6] 横内大介, 青木義充, 『イメージでつかむ機械学習入門 ~豊富なグラフ, シンプルな数学, R で理解する~』, 技術評論社, 2017