Location Identification by WiFi Signal Strength

M. BALL, 9.23.19

Objective

The task at hand dictates determining feasibility for location detection *indoors* by using wireless access point signal strength (also known as WiFi fingerprinting)

The primary objective for this task is to determine if a model can be developed using the Indoor Locationing Dataset that can predict *room-level* location with a high degree of accuracy

This project is a re-visit in Python of initial analysis performed for the same objective in R

Data Overview

The provided dataset includes 19,937 rows and 530 total columns

- 520 of the 530 columns consist of wireless access points (WAPs) which are the variables of primary interest

There are a total of 905 unique locations in this dataset

Due to prior issues in determining room-level accuracy, the new approach split out the desired location into a combined identifier and kept the room as a separate identifier:

- Relative position, building ID, and floor were combined into Location
- Space ID was kept separate

The only predictor features used in modelling were the 520 WAPs

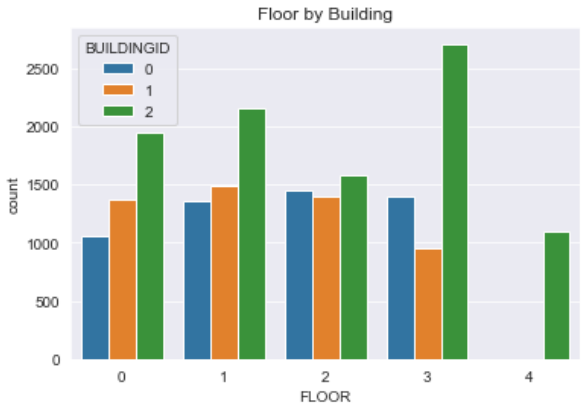
All data was scaled for analysis

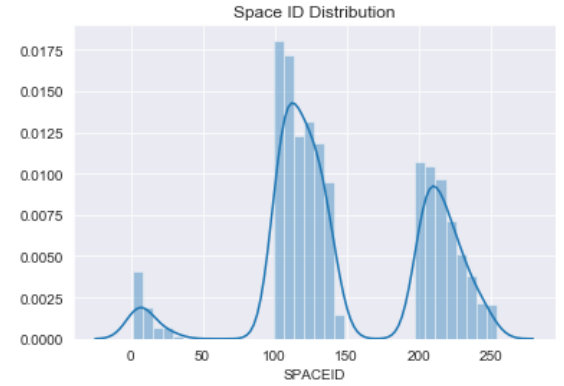
Data Insights

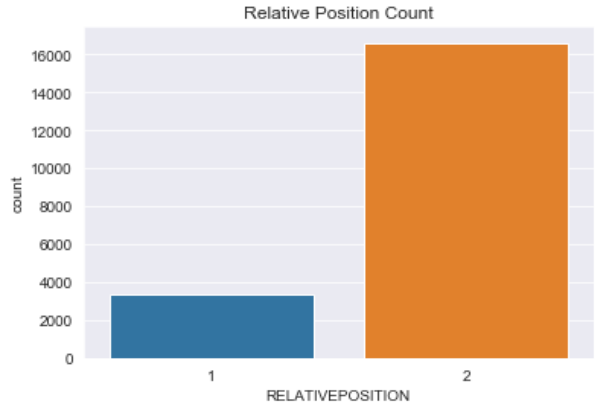
Space ID is not a continuous variable and is mostly grouped into three groups: between 0 and 50, 100 and 150, and 200 and 250

Only building 2 has a fifth floor (Floor 4 in data set)

Much more of the data was collected outside of the room (1) instead of inside of the room (2)







Summary of Models

Multiple models were tested for both the location identifier and also the space identifier

The best performing model type for location was also the best performing model type for the space identifier – random forest

Other models were tested (KNN, SVM, Decision Tree) but their accuracies were not as high

Dependent Var	Model Type	Accuracy	Kappa
Location	Random Forest	95.3%	0.95
	SVM	82.4%	0.81
	KNN	89.1%	0.88
SpaceID	Random Forest	80.2%	0.80
	SVM	58.5%	0.58
	KNN	65.2%	0.65

Analysis & Conclusions

When applying the models to the entire data set, the predictions were very accurate

Although it requires more work and added complexity, you can detect with a high degree of accuracy room-specific locations using WiFi signal strength

Component	Accuracy
Building ID	99.8%
Floor ID	99.7%
Relative Position	98.4%
Space ID	91.7%

Recommendations

Overall, the model performed very well based on the WiFi fingerprinting technique

However, for ease of use, it would be beneficial to get high accuracy with a single dependent variable instead of splitting between two

The best method to increase this accuracy is to increase the sample size for each single location by space ID, and to collect more of this data inside the room instead of outside

Although additional cost, reviewing specific indoor locations with multiple obstructions and providing additional WAPs in those locations may be helpful

- The incorrect space ID predictions in the dataset can help to pinpoint some of these locations
- This should also help address some device-to-device signal detection differences by increasing the overall signal strength