# Wi-Fi Score

## Introduction

## Key KPIs

Poor\_rssi and poor\_phyrate, discussed in introduction, are defined as below:

1. Each day, multiple records of Wi-Fi performance (station history) are collected for each device or station.
2. Standards or thresholds are established to categorize each record as either having good or poor Received Signal Strength Indicator (RSSI) and Physical Rate (PHY Rate).
3. The historical data for each station is aggregated to determine the percentage of records classified as good or poor RSSI and PHY Rate for that specific **device or station.**

### poor\_rssi

two features are used to obtain poor\_rssi, different connect types have different threshold.

* connect\_type: what kind of STAs is being used, e.g. 2.4GHz, 5G, MoCA, …
* signal\_strength: RSSI

RSSI follows that

* 2.4Ghz - sample below -78dbm classified as poor
* 5.0Ghz - sample below -75dbm classified as poor
* 6.0Ghz - sample below -70dbm classified as poor

### poor\_phyrate

Phy Rate follows that:

* Phy Rate < 65Mbps classified as poor

### Weights:

Weights for each station/device is defined majorly concerning volume, also includes stationarity and daily average phyrate.

* Volume = byte\_send + byte\_received
* Stationary or Non-stationary is defined considering if there is significant daily variation of RSSI
  + For each home, find P90 and median of RSSI
  + | P90 - median | > 5 dB, classify as non-stationary
  + Map stationary to 1, and non-stationary to 0
* average phyrate: average all phyrate of daily records.

Each device is categorized as one of four categories (*Figure 1*):

1. stationary + high average phyrate
2. stationary + high average phyrate
3. non-stationary + low average phyrate
4. non-stationary + low average phyrate

for first three categories, we use the proportion of log volume of specific device to all log volume of all devices in this home, as weight:

For last category, loT device characterized as stationary and low phyrate, we want to penalize its influence by further reduce its weight

\* is threshold of low average phyrate comes from statistical inference

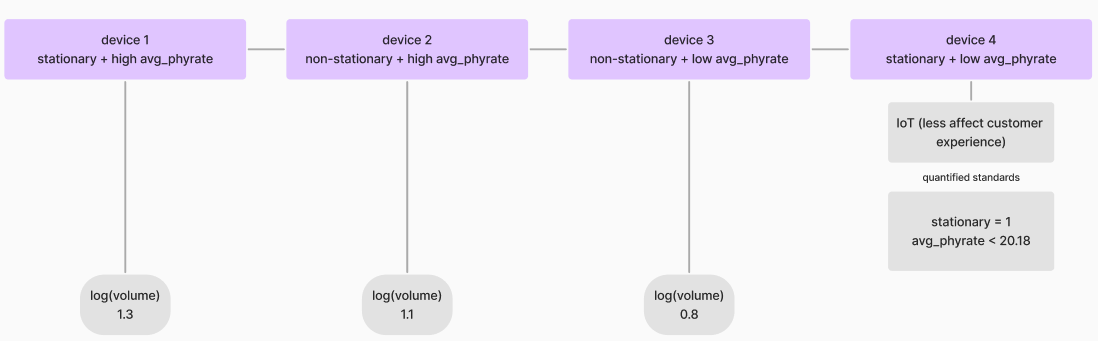


Figure 1 four weight categories of device

## Data Filtering

Before start data process, we firstly filter some irrelavent data

* Drop wired connect\_type so that we can focus on wireless WiFi.
* Exclude records where Phy Rate is not inside the range given in WiFi Score 1.0.
  + 6 < Phy Rate < 2,500
  + Phy Rate not in [12,24]
* Exclude the outliers in signal\_strength by comparing the number dropped using Interquartile range method (exclude the upper bound; use 10% as Q1, and 90% as Q3), and 3% of data when arranged in increasing order, and choose whichever is smaller.

## Extender identification

* Use router serial number (inside rowkey) to decide devices per home.
* Match router serial number to find different models of routers and extenders in Device Groups data.
* Match parent id in Station History data with station mac in Device Groups data and check the parent mac in Device Group data to see if the device is connected to a router or an extender, map router with 1 and extender with 0.

## Final Output

Columns:

* date: date of the data
* mdn, cust\_id: user information
* serial\_num: the router serial number
* poor\_rssi: RSSI poor percentage of all devices in one home (calculate with weights)
* poor\_phyrate: Phy Rate poor percentage of all devices in one home (calculate with weights)
* num\_station
* dg\_model\_indiv: device used as router or extender, such as ASK-NCQ1338FA
* home\_score: final score of all devices