CPE613 Optimization Methods

Term project proposal

Supawit Somsakul 65070800402

Project name: Cellular network station location planning

Introduction:

In cellular network, planning a location of station is considered a difficult task as many factors need to be considered. The location of the station must be chosen so that it can provide service to the largest number of people while maintaining the construction and operating cost within company's limit. As well as other factors such as accessibility, availability of power sources, risk of natural disasters and impact to local communities.

The problem can get more complicated if we factor in the fact that the subscribers are mobile. They can move to other places throughout a day. This resulted in solution that only considered single pattern of subscribers will be overfitted and perform poorly on other patterns. To address this issue, the solution built for this problem must be able to learn subscriber's pattern dynamically.

In this study, two objectives are considered. One being to maximize satisfaction of users. Which will be in terms of average signal strength of all users in the area. The other is the construction and operating cost of the station. These two objectives are opposite, and it is not possible to satisfied both at the same time. Thus, this problem can be considered as a multi objective optimization problem.

To optimize this problem, a multi-objective evolutionary algorithm (MOEA) is used. It is an optimization algorithm that is used to find the best solution for a problem that has multiple objectives. The goal is to find the best tradeoff between average signal strength and costs to get the most satisfiable results for the company.

For the issue of moving subscribers, a Monte Carlo simulation is integrated into MOEA to presents different patterns of subscribers to tentative solution pool. In each generation, a population will be presented with different pattern of subscribers. By using this approach, solution from MOEA will become less overfit to specific pattern. Making solution more generalized.

Assumption:

There are few assumptions made in this study, which include:

1. Area grid used in this study will be a 1600 x 1600 2-dimension grid with scale of 1:12.
2. There are 10000 subscribers in the area.
3. Subscribers are not randomly scattered over the area. Instead, there are a well-defined cluster scattered across the grid.

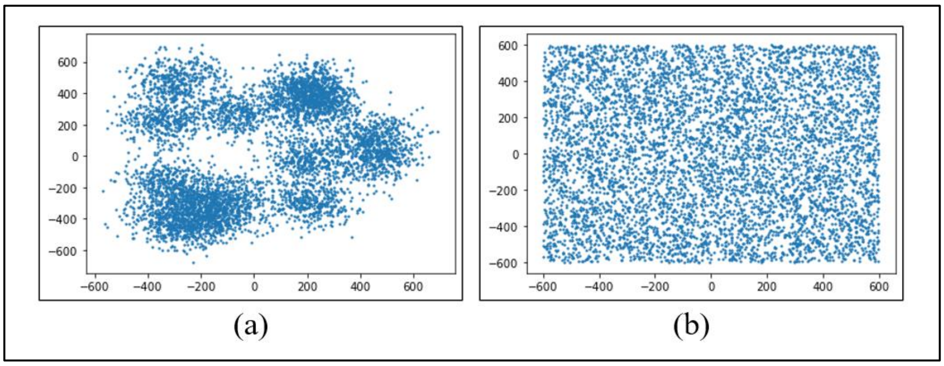

Figure 1: An example of subscriber's pattern
(a) With clusters
(b) Without cluster (randomly scattered)

4. There is a probability associated with each subscriber's patterns indicates how likely each pattern will occurred. The summation of probability of all patterns is equals to 1.

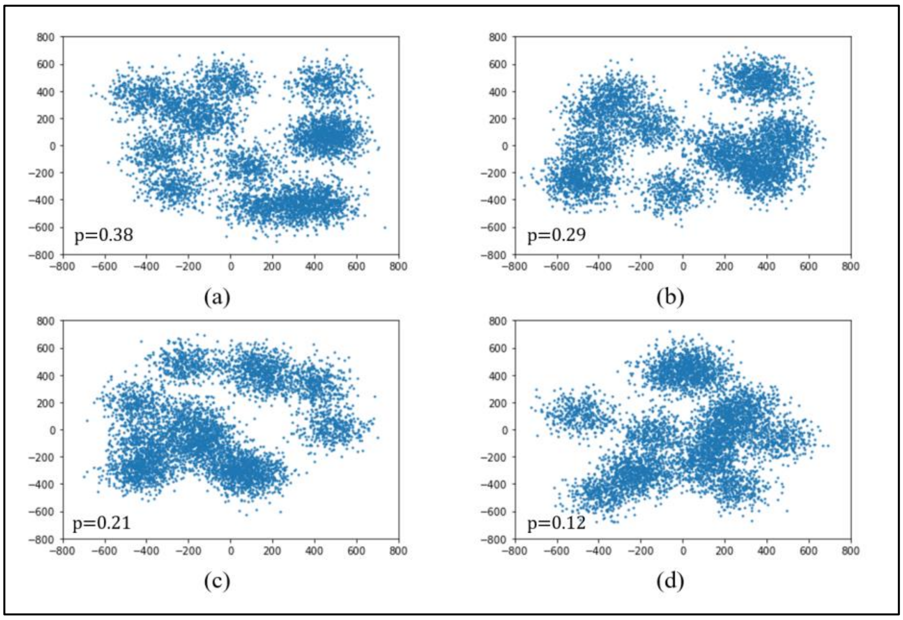

Figure 2: An example of subscriber's pattern with probability
(a) A pattern with probability of 0.38
(b) A pattern with probability of 0.29
(c) A pattern with probability of 0.21
(d) A pattern with probability of 0.12

5. A 4G LTE station transmitted standard frequencies band. Which include 700, 900, 1800, 2100, 2300 and 2600 MHz band [1]. In this study, only 900, 1800 and 2600 MHz band are considered.

6. There are 2 types of station: normal bandwidth (Type I) and high bandwidth (Type II). Each has different construction cost for different frequencies band. Table 1 gives information about the capacity and cost of each station type.

Table 1 Construction cost of station

Station type	Capacity	Construction cost (baht)
900 MHz Type I	800	1150000
900 MHz Type II	1200	1500000
1800 MHz Type I	850	880000
1800 MHz Type II	1250	1220000
2600 MHz Type II	800	950000
2600 MHz Type II	1300	1350000

7. There can be at maximum 20 station constructed.

8. Signal strength received at given distance are assumed to follow a Free-space path loss (FSPL) [2] formula with no obstacle. FSPL is defined as:

$$FSPL_{dB}(i,s) = 20\log_{10} d_{i,s} + 20\log_{10} f_s + 32.44 \qquad (1)$$

Where:

$FSPL_{dB}$ indicates signal strength in decibel
d indicates distance between station and subscriber in meters
f indicates transmitted frequency of station in megahertz

9. Distance measurement used in this study is Euclidean distance. Which defined by:

$$d(i,j) = \sqrt{\left(x_i - x_j\right)^2 + \left(y_i - y_j\right)^2} \qquad (2)$$

Where:

d indicates distance between two entities i and j
x_i, x_j indicates position on x axis of i and j
y_i, y_j indicates position on y axis of i and j

10. Attachment order of subscriber to station is random.

11. Subscriber will always pick station with strongest signal that is not fully utilized.

Model formulation:

In this study, a multi-objective optimization model is formed as follows.

Let:

n	Number of subscribers
m	Number of stations
$R = FSPL(i, s)$	Signal strength of subscriber i that attached to station s
$C = \sum_{i=1}^{6} c_i x_i$	Construction cost of station i
$S = \sum_{i=1}^{n} s_i y_i$	Number of subscribers attached to station s

Objective:

$Max\colon R_a = \frac{\sum_{i=1}^{n} R_i}{n}$	Maximize average signal strength of all subscribers
$Min\colon C_a = \sum_{i=1}^{m} C_i$	Minimize construction cost

Subject to:

$m \geq 1$	At least 1 station needs to be constructed
$m \leq 20$	At most 20 stations can be constructed
$\sum_{i=1}^{m} x_i \leq 20$	At most 20 stations can be constructed
$\forall S_i \leq S_i^{max}$	Number of subscribers per station should not exceed maximum of each station
$x_i \in \{0,1\}$	Selection flag for station type of each station
$y_i \in \{0,1\}$	Selection flag for station attached of each subscriber

Input dataset:

As it is hard to find real-world cellular network dataset with multiple patterns online, dataset used in this study was obtained from a simulation. The simulation assumption is that the subscribers are not randomly scattered across the area. Instead, they are scattered as a well-defined cluster across the grid.

A simulation was done via Python's scikit-learn library. There are 10000 subscribers scattered in a cluster across 1600 x 1600 grid.

Problem size:

There are 1600 x 1600 possible position for each station, there can be as much as 20 stations constructed and each stations have 6 possible configurations. Total possible solutions of this problem are:

$$\sum_{i=1}^{20}(1600 \times 1600) \times \binom{6}{1}^{i} \approx 1.123 \times 10^{22}$$

Algorithm and parameters:

Algorithm used in this study is simple multi-objective evolutionary algorithm (MOEA) with Monte Carlo simulation on subscriber's pattern. Parameters used in simulation are:

Population size: 1000
Crossover rate: 0.2
Mutation rate: 0.05
Elitism ratio: 0.25

References:

[1] The National Broadcasting and Telecommunication Commission, "Telecommunications Business Licensing," 11 3 2020. [Online]. Available: https://telecom-license.nbtc.go.th/getattachment/Information/spectrumroadmap /spectrumroadmap/spectrumroadmap.pdf.aspx. [Accessed 30 10 2022].

[2] J. G. Proakis and M. Salehi, "Performance Analysis for Wireline and Radio Communication Systems," in *Digital Communications*, New York, McGraw-Hill, 2008, p. 262.