Optimisation of Cell Selection for a Network Provider

I. Optimal Selection of Network Cells for Population Access

A front-end network provider has to rent bandwidth on existing 5G cells. They must choose the best combination of cells within their financial constraints. This subsystem defines the 'best' combination as the combination capable of providing coverage for the highest number of people. A model was created for UK population density from which population access could be calculated for each cell. As data could not be found on cell location, a random set of locations was generated to create a set of possible cells. Cells were split into three categories: Macro, Micro and Pico cells [1]. Femtocells were ignored as they are for indoor use.

A. Optimisation formulation

The problem formulation can be seen as follows, and can be reduced to the bounded 'Knapsack' problem.

$$\boldsymbol{min}: \; f(\boldsymbol{p}) = -\sum_{i=1}^{n} p_i x_i \quad (1)$$

p_i is the population accessed by the cell. $\boldsymbol{p}$ is the vector of these values.

$$\boldsymbol{s.t.} \; h(\boldsymbol{c}) = \sum_{i=1}^{n} c_i x_i - c_{max} \leq 0 \quad (2)$$

Where $\boldsymbol{c}$ is the vector of costs associated with each cell. And c_{max} is the maximum the network provider can spend on cells. For this model this is chosen to be £500,000.

$$x_i \in \{0,1\} \quad (3)$$

x_i is the cell selection variable.

$$c_i \in \{c_{macro}, c_{micro}, c_{pico}\} \quad (4)$$

c_i is the cost of each cell.

The population reached for each cell is a function of its location and range.

$$r_i \in \{r_{macro}, r_{micro}, r_{pico}\} \quad (5)$$

$$for\ \{cxi, cyi\} \; \forall \; \varepsilon - 1000\ km < 0 \quad (6)$$

r_i is the range of each cell. Chosen to be the range in coverage values for each cell (*see Figure 1)*. cx_i, cy_i are the coordinates for the location of each cell. sx_i, sy_i are the coordinates for the centre of each settlement.

Base station type	Number of users	Coverage (km)	Bandwidth (MHz)	RF (W)	Location	Users
Femtocell	1 to 30	0.01 to 0.1	10	0.001 to 0.25	Indoor	Homes or small offices
Picocell	30 to 100	0.1 to 0.2	20	0.25 to 1	Indoor/ outdoor	High-rise buildings, hotels, office buildings or parks
Microcell/ metrocell	100 to 2,000	1 to 2	20 to 40	1 to 10	Indoor/ outdoor	Shopping centers, transportation hubs, city blocks, stadiums, temporary events
Macrocell	>2,000	5 to 32	60 to 75	10 to >50	Outdoor	Suburban, city and rural areas

Figure 1: Cell Categories. Source [1]

B. Model Approach

The model formulation can be split into four sections: Knapsack Formulation, Data processing, Population Density Map Formulation and Cell Population Access.

1) Knapsack Formulation

The network provider will have limited available capital to lease bandwidth on cells. They must find the optimal combination of cells for the money they are spending. This can be reduced to a bounded 'knapsack' problem.

2) Data Processing

Data processing was required to create two different data sets. One data set was settlement population and location data for the UK. The other was cell location, range and cost data.

The settlement population data had to be created from two different data sets. One contained longitude and latitude for every settlement within the UK. Further processing was required to complete calculations with this data. The spherical coordinates of longitude and latitude were projected onto a Cartesian plane, preserving the distance between each point. See (7) (8). (It was assumed that the Earth's diameter was constant for this calculation.)

$$\mathrm{sx_i} = (\mathrm{long_i} - 50)\ 2\mathrm{r_{earth}}\pi/180 \quad (7)$$

$$sy_\mathrm{i} = (\mathrm{long_i} + 7)\ 2\mathrm{r_{earth}}\pi/180 \quad (8)$$

Population data only existed for a limited set of settlements. Therefore the set of settlements was reduced to the size of this set and the data was concatenated.

Cell location data for 5G cells is not widely available for the UK. Therefore to demonstrate the effectiveness of the optimization techniques, a set of different cell types at randomly chosen locations was generated. These locations were limited by the size of the map, which was 1000km x 1000km. Two different sets were create for two different optimization tasks. One was created using the maximum coverage for each antenna type. The other was using average coverage. (18.5km for a macro cell). The cost of leasing was bandwidth on a cell was given an arbitrary value as information on this was not available. Arbitrary numbers of cells were added:

Cell Type	Quantity	Cost / £
Macro	400	50,000
Micro	800	10,000
Pico	5000	2,000

Table 1: Cell Parameters

3) Population Density Map Formulation

The population density map was required to calculate the population access for each cell. The population map was created by modeling population density as a function of distance from the settlement center. The first suggestion for a model, by Clarke in 1950, was an exponential decay[2]:

$$\rho(r) = A_0 e^{-\beta r} \quad (9)$$

r being distance from city center. A_0 being a trivial population density at $r = 0$ and β the decay rate. This was later refined to form *Sherrat's Model*; A Gaussian function for population density as function of distance from a city center[3].

$$\rho(r) = \rho_0 e^{-\frac{r^2}{2r_0^2}} \quad (10)$$

ρ_0 being the maximum population density, r_0 being the characteristic radius, ($r_0 = \frac{1}{\beta}$). This was the model chosen for creating of a UK-wide population density map. This was converted from Polar to Cartesian form as city locations were given as Cartesian coordinates.

$$\rho(x, y) = \rho_0 e^{-\frac{(x-x_j)^2 + (y-y_j)^2}{2r_0^2}} \quad (11)$$

Where (x_j, y_j) are the coordinates for the settlement center.

A model was created for each settlement with estimates being made for maximum population density by categorizing each settlement on population.

Settlement Type	Population Range	ρ_0/km^2	Based On
Small	< 25,000	4,000	Whitehaven
Medium	25,000 < P < 250,000	11,000	Basingstoke
Large	250,000 < P < 5,000,000	20,000	Liverpool
London	> 5,000,000	35,000	London

Table 2: City Categorization and Max Population Density Estimates.

These estimations were used to calculate a suitable r_0 for each settlement using:

$$Total\ Population = \iint_{-\infty}^{\infty} \rho_0 e^{-\frac{(x-x_j)^2 + (y-y_j)^2}{2r_0^2}} dxdy \quad (12)$$

$$Total\ Population = 2\pi\rho_0 {r_0}^2 \quad (13)$$

With r_0 calculated, a Gaussian Function could be generated for each settlement. These were then summed together to create a population density map for the UK, $\rho_{UK}(x, y)$.

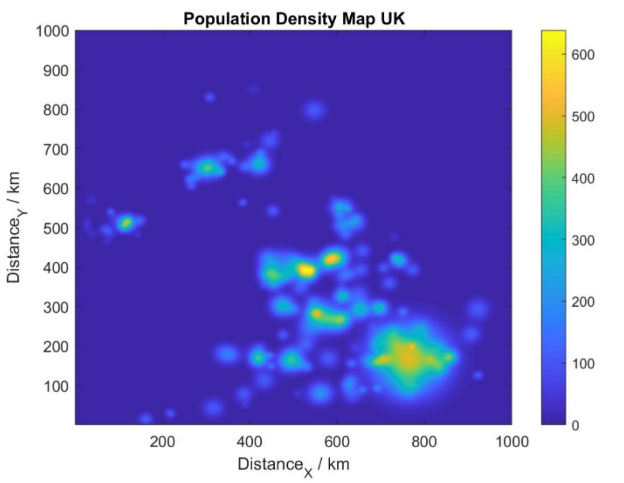

Figure 2: UK Population Density Heat Map

SETTLEMENT DATA ASSUMPTIONS

The settlement data available only included data from settlements with populations above 10,000. The total population covered by these settlements is approximately 28,500,000. This is only 40% of the total population of the UK. The rest of the population could not be added as the area of the map occupied by the UK has not be defined due to lack of UK coastline data. Therefore the population was assumed to be approximately 28,500,000 for the system.

GAUSSIAN MODEL ASSUMPTIONS

Three assumptions exist within this model. One states that all settlements of similar population size follow have a similar population density distribution. The second is that modern settlements follow a Gaussian Function for population distribution. The final is the assumption that population density is constant with respect to time. These assumptions were made to reduce the problem complexity.

4) *Cell Popualtion Access*

The population accessed by each cell was calculated by finding the volume under the surface within the coverage of the cell (D).

$$p_i = \iint_D \rho_{uk}(x, y) dA \quad (14)$$

$$p_i = \int_{-r_i+cx_i}^{r_i+cx_i} \int_{-r_i+cy_i}^{-r_i+cy_i} \rho_{uk}(x, y) dxdy \quad (15)$$

where p_i is the population accessed by the cell, r_i is coverage radius and (cx_i, cy_i) are the location coordinates of the cell.

CELL POPULATION ACCESS ASSUMPTIONS

One assumption exists within this stage of model formulation. It assumes that all the population within the coverage radius can access the antenna, ignoring the propagation problems generated by geography. This assumption was made to reduce problem complexity.

C. *Optimisation Methods and Results*

Two different optimization techniques were used to solve the problem. The first is a stochastic method, Simulated Annealing. The second, a dynamic programming method is used to validate the result of the Simulated Annealing program. Both methods are non-deterministic due to the NP-Complete nature of the bounded Knapsack problem.

1) *Simulated Annealing*

Simulated Annealing is used to find the combinatorial optimal for this Knapsack problem. The simulated annealing program was written by *Benjamin Misch, 2012* [4]. Certain parameters must be chosen for Simulated Annealing:

Parameter	Value
Cost Constraint	£500,000
β step	0.01
Iteration No.	1500

Table 3: Simulated Annealing Parameters

Suitable β step and number of iterations were found using experimentation. They were chosen when the function gave most consistent outputs.

GLOBAL MINIMUM

The 'global minimum' output of (1) ranges $-78{,}450$ < GM < -72,500 people. This range exists due to the non-deterministic nature of Simulated Annealing.

SELECTED CELLS

The program provides a list of selected arrays. This is a translation of the selection variable values for each cell. If the cell was selected, $x_i = 1$. If the cell is not selected, $x_i = 0$. A visualization of selected cells can be seen in *Figure 3*.

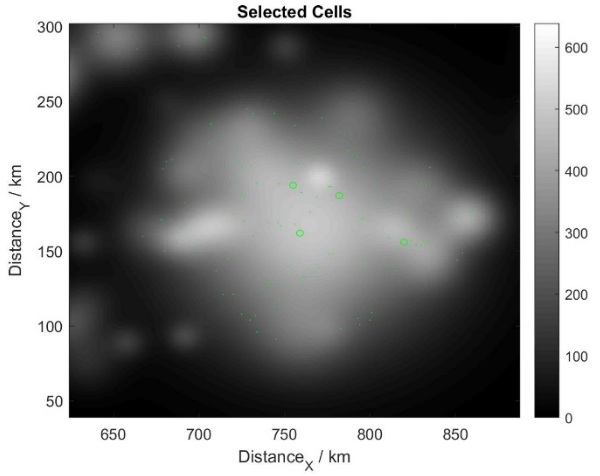

Figure 3: Selected Cells London chosen by Simulated Annealing

2) Dynamic Programming

Dynamic Programming is used as a method to validate the solution found using simulated annealing. A dynamic program method written by Petter Strandmark is used[5].

GLOBAL MINIMUM

The 'global minimum' output of (1) for the objective function is found to be $-79{,}620$ people. No range appears for maximum population access using dynamic programming.

SELECTED CELLS

There are minor difference in the list of selected cells. These are more prominent in cells with lower population access.

3) Results Discussion

The results appeared similar for both Simulated Annealing and Dynamic Programming. It seems to be a quirk of our current model that very few Macro Cells are selected. This is likely due the parameters assigned to Macro Cells. The four cells that are chosen using Simulated Annealing are likely due to the starting point in the selection being Macro Cells.

D. Model Limitations

One urgent limitation currently exist within this model. The model currently does not include the coastline of the UK as a constraint for cell location or population. This means both cells and population can be found at coordinates that would be in sea. This is a development that should be made to increase the validity of this model.

E. Conclusion

This report has covered optimisation formulation, the modelling approach and the optimisation methods and results.

The optimization formulation is reduced to a bounded knapsack problem. *What is the optimal combination of network cells to maximize population access within a cost constraint?*

The model is formulated in three sections: data processing, population density map production and cell population access. Data processing includes concatenation of data sets and random cell location generation. Population density map production uses the Sherrat model to model population and a function of distance from city center. Cell population access uses the integral under this surface over the coverage area of the cell.

The optimization method uses two techniques to find global minimum: Simulated Annealing and Dynamic Programming. This methods find that the max population access is approximately 80,000 people within a constraint of £500,000. *(These numbers are trivial as the little un-informed assumptions on antenna prices were made).*

[1] N. Lemieux and M. Zhao, 'Small Cells, Big Impact: Designing Power Solutions for 5G Applications', p. 7, 2019.
[2] C. Clark, 'Urban Population Densities', *J. R. Stat. Soc. Ser. Gen.*, vol. 114, no. 4, pp. 490–496, 1951.
[3] Y. Chen, 'A new model of urban population density indicating latent fractal structure', *Int. J. Urban Sustain. Dev.*, vol. 1, no. 1–2, pp. 89–110, 2010.
[4] B. Misch, 'Simulated Annealing and the Knapsack Problem', p. 5.
[5] P. Strandmark, *Dynamic Programming 1-0 Knapsack Problem*. 2009.

F. Appendix

Cell Selection Table Simulated Annealing (Cell_id):

457	2018	2795
872	2019	2797
951	2022	2834
953	2025	2872
1205	2029	2910
1215	2187	2927
1250	2189	2958
1296	2213	2994
1319	2214	3001
1330	2241	3023
1365	2265	3055
1407	2268	3067
1428	2291	3093
1439	2297	3094
1463	2309	3113
1473	2310	3165
1488	2349	3174
1512	2360	3217
1541	2377	3251
1552	2381	3280
1589	2406	3289
1623	2418	3322
1636	2426	3338
1652	2428	3361
1673	2480	3379
1677	2491	3384
1805	2499	3429
1807	2501	3438
1812	2507	3479
1817	2515	3487
1819	2526	3581
1825	2549	3634
1848	2641	3645
1862	2648	3684
1883	2684	3695
1893	2691	3710
1905	2733	3724
1912	2751	3754
1923	2758	3768
1927	2765	3781
1933	2766	3784
1974	2771	3838
1978	2784	3847
2015	2787	3867

3869	4576	5485
3873	4610	5521
3946	4652	5575
3948	4663	5586
3954	4666	5590
3973	4732	5601
3989	4742	5605
3990	4745	5632
4029	4799	5634
4032	4804	5680
4042	4809	5695
4058	4851	5733
4104	4864	5739
4200	4897	5757
4202	4906	5766
4250	4920	5772
4275	4923	5792
4291	4940	5802
4318	4976	5819
4321	5026	5825
4352	5053	5839
4381	5065	5848
4382	5081	5867
4383	5129	5886
4391	5144	5911
4392	5160	5957
4394	5227	5971
4410	5233	6015
4425	5331	6039
4435	5332	6050
4437	5349	6099
4463	5383	6106
4485	5418	6124
4562	5428	6139

Cell Selection Table Dynamic Programming(Cell_id):

1205	1673	2019
1215	1677	2022
1250	1805	2025
1296	1807	2029
1319	1812	2112
1330	1817	2184
1365	1819	2187
1407	1825	2189
1428	1848	2213
1439	1862	2214
1463	1883	2241
1473	1893	2265
1492	1905	2268
1512	1912	2291
1541	1923	2297
1552	1927	2309
1589	1933	2310
1621	1974	2349
1623	1978	2360
1636	2015	2377
1652	2018	2381

2406
2416
2418
2426
2428
2480
2491
2499
2501
2507
2515
2526
2549
2553
2578
2589
2641
2648
2684
2691
2733
2751
2758
2765
2766
2771
2784
2787
2795
2797
2834
2872
2893
2910
2927
2958
2994
3001
3013
3023
3055
3067
3093
3094
3113
3165
3174
3217
3233
3251
3277
3280
3289
3322
3338
3361
3379
3382
3384
3413
3429
3438
3479
3487
3581
3634
3645
3684
3695
3710
3724
3754
3768
3781
3784
3838
3847
3867
3869
3873
3946
3948
3954
3973
3989
3990
4029
4032
4042
4058
4104
4106
4200
4202
4250
4275
4291
4318
4321
4352
4381
4382
4383
4391
4392
4394
4410
4425
4435
4437
4463
4485
4562
4576
4610
4652
4663
4666
4732
4742
4745
4799
4804
4809
4851
4864
4897
4906
4920
4923
4940
4964
4976
5026
5053
5065
5081
5129
5144
5160
5227
5233
5331
5332
5349
5383
5412
5418
5428
5485
5521
5575
5586
5590
5601
5605
5615
5632
5634
5680
5695
5733
5739
5757
5766
5772
5792
5802
5819
5825
5839
5848
5867
5886
5911
5952
5957
5971
6015
6039
6050
6052
6099
6106
6124
6139
6176