Modified Kruskal's Algorithm with Euclidean Steiner Tree in Electricity Grid

Hongzhen Zhu*
Yiran Hu†
Jialu Jiang‡
Yiyang Wang§
Shuze Zhuang¶
Jiachen Weng‖

This paper discussed an algorithm of finding electricity network on given vertices with different costs. And we also consider about their geographical feature might change the topology of network.

1. Introduction

1.1 Electricity grid

Definition 1.1.1. *An **electricity grid** or electricity network is an interconnected network for delivery electricity to demand points(cities, communities)*

Remark 1.1.1. *Electrical grid consist of power station(s), high voltage line to transfer the power between cities, and some step down transformer to convert the high voltage into low voltage.*

Example 1.1.1. *Here is an example from Wiki*

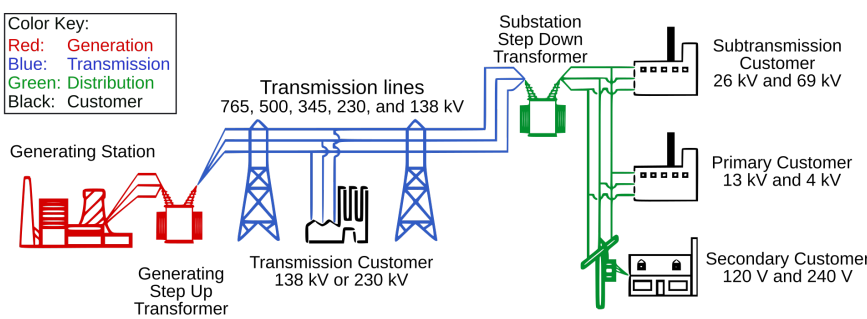

Fig. 1 Electrical grid

Remark 1.1.2. *In this paper, we will introduce how to minimize the total cost of electricity network of given vertices if there is some cost for edges and vertices.*

1.2 Centralized and Decentralized System

Given an electricity grid we have two types of systems to provide power.

*Department of Mathematics, University of Illinois Urbana-Champaign, 601 East John St., Champaign, IL, 61820, zhu990211@gmail.com
†Department of Mathematics, University of California, Berkeley, 970 Evans Hall, Berkeley, CA, 94720, the United States, yiranhu@berkeley.edu
‡Department of Mathematics and Statistics, Mount Holyoke College, 50 College St South Hadley MA, 01075, United States, jiang37j@mtholyoke.edu
§Faculty of Arts and Sciences, Beijing Normal University, Zhuhai, 519087, China, 202411079899@mail.bnu.edu.cn
¶School of Physics and Information Engineering, Fuzhou University, 350108, China1786131208@qq.com
‖Suzhou North America High School268 tian e dang road, Wuzhong District, Suzhou, Jiangsu, 215000, China, 107414740@qq.com

Definition 1.2.1. ***Centralized system*** *[1] is a power system that cities are connected by high voltage line and among each city they have some internal network that transfer the high voltage into low voltage.*

Definition 1.2.2. ***Decentralized system****[2] is system that cities are powered by some facility nearby themselves, like solar energy or wind power.*

Definition 1.2.3. *[1] For vertices that belong to decentralized system, they have some fixed cost that determine by the city itself(life cost of building solar energy facilities), we call it as* ***decentralized system cost/DSC***

Definition 1.2.4. *[1] For vertices belong to centralized system, they have three types of costs*

- ***Internal grid cost/IGC****: For any city i belong to centralized system, i itself need some money to convert the high voltage electricity into low voltage.*
- ***External grid cost/EGC****: For any two cities (i, j), their distance times the unit cost of high voltage line will be the external grid cost*
- ***Power station cost/PSC****: For each centralized system, there is a fixed cost for building a power station, we defined it as power power station cost.*

Remark 1.2.1. *In this paper, we assume each centralized system has only one power station.*

Example 1.2.1. *Here is an example of two systems and their costs:*

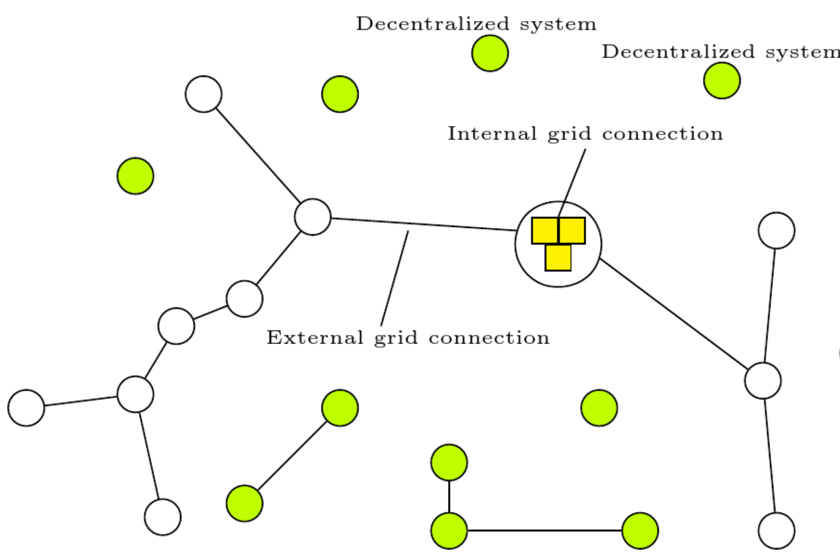

Fig. 2 Centralized system and decentralize system

Remark 1.2.2. *[2]In real life, some cities with large demand of power might want to build a centralized system. However, the emissions from usage of fossil fuels from centralized system make people consider some sustainable energy like solar power, wind power. Those needs are mainly the reason why people study decentralized system. The decentralized systems could be connected with some other decentralized systems or just be isolated vertices.*

Remark 1.2.3. *Noted the electricity systems in this paper are trees, and for each system we need one power station to provide energy to whole system.*

1.3 Euclidean Steiner tree

Definition 1.3.1. ***Euclidean Steiner problem*** *[3] is to find shortest network connecting a given set of points(terminal) T in Euclidean plane if we are allowed to add extra junctions that is not belong to T*

Example 1.3.1. ***Euclidean Steiner problem when*** *$|T| = 3$, recall Fermat's point of triangle ABC is the point F that minimize the distance $|FA| + |FB| + |FC|$, therefore, Fermat's point is the solution for $|T| = 3$*

Definition 1.3.2. *[3] Given a Euclidean Steiner problem, the tree we find out is defined as* ***Euclidean Steiner tree (EST)***

Definition 1.3.3. *An EST is full if it has $|T| - 2$ steiner points, such EST we defined it as* ***full steiner tree (FST)****.*

Example 1.3.2. *Here is an example of FST when $|T| = 3, 4$*

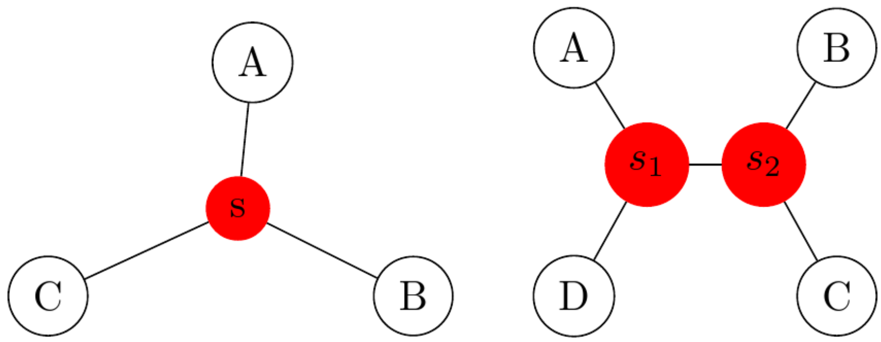

Fig. 3 Example of FST when $|T| = 3, 4$

Theorem 1.3.1. *Degree and angle properties[3]*

- *The Steiner tree must have three edges incident to each Steiner point,*
- *Every Steiner points have three incident edges, any two those three edges have 120 degree.*

Remark 1.3.1. *Clearly EST is very rare in real life, since its degree and angle requirement might not be feasible due to some geographical reason.*

1.4 Flat Region

Definition 1.4.1. ***Flat region*** *is a region where allow the EST exist.*

Remark 1.4.1. *We define flat region because in real life some cities tend to gather in the flat plane. We will allow EST exist only in flat region.*

2. Methodology

2.1 Hexagonal coordinate system (HCS)

We introduce a method for finding the Euclidean Steiner Tree [4], if the topology of tree is given.

From Theorem 1.3.1 Steiner points have lines only three direction with 120° apart. Therefore we could use *hexagonal coordinate system (HCS)*. This subsection Introduce an idea from F.K. Hwang [4]

Let U, V, W be three axes going through the origin and cutting the plane into six 60° cones. Points in this plane are represented by vector (u, v, w)

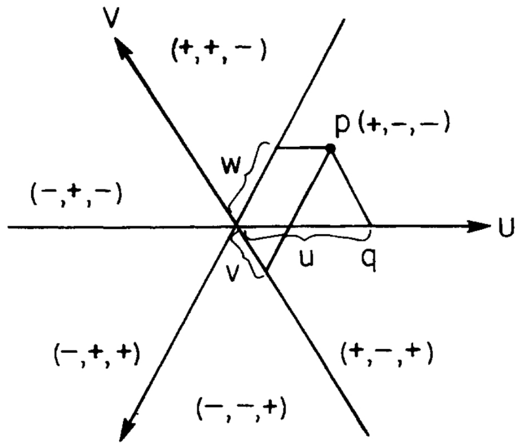

Fig. 4 Hexagonal system

Lemma 2.1.1. *$u + v + w = 0$ for any point $P = (u, v, w)$*

Lemma 2.1.2. *Let P has Cartesian coordinates (x, y). Then $w = -2y/\sqrt{3}, u = x + y/\sqrt{3}, v = -x + y/\sqrt{3}$, in particular,*

$$x = \frac{u - v}{2}, y = -\frac{\sqrt{3}}{2} w$$

Proposition 2.1.1. *Suppose $s_3 = (u_3, v_3, w_3)$ is connected to $a_1 = (u_1, v_1, w_1)$ and $a_2 = (u_2, v_2, w_2)$ and the lines at a_1 and a_2 are parallel to U axis and the V axis, respectively. Then*

$$w_3 = w_1, u_3 = u_2, v_3 = -w_3 - u_3 = -w_1 - u_2$$

T is bigraph with partite sets N_1, N_2, we assign direction to each edge and assume that an edge always starts from a vertex in N_2 and ends in a vertex in N_1

Theorem 2.1.3. *Define $\epsilon_i = 1$ if a_i is in N_1 and $\epsilon_i = -1$ otherwise. Define $d_i = u_i(v_i, w_i, respectively)$ if the line at a_i is parallel to the $V(W, U, respectively)$ axis. Then*

$$\sum_{i=1}^{n} \epsilon_i d_i = 0, called\ the\ \textbf{\textit{characteristic equation of T}}.$$

Next, we would use some mathematics in linear algebra and trigonometry.

Proposition 2.1.2. *The first term in the characteristic equation can be arbitrarily set to be u_1, v_1, w_1*

Theorem 2.1.4. *We now consider the general case that the lines of an FST are parallel to the axes only after a clockwise rotation of angle θ. Define $l = \cos\theta, k = \sin\theta/\sqrt{3}$. Then the new coordinates can be obtained from the original coordinates through the transformation*

$$\begin{pmatrix} u' \\ v' \\ w' \end{pmatrix} = \begin{bmatrix} l & k & -k \\ -k & l & k \\ k & -k & l \end{bmatrix} \begin{pmatrix} u \\ v \\ w \end{pmatrix}$$

Proposition 2.1.3. $l^2 + 3k^2 = 1$

Example 2.1.1. *Given five vertices with fixed coordinates and the topology of EST, and we find a Euclidean Steiner tree for them.*

Assuming we have location of five points as below

$$a_1 = (2.08341,\ 1.08757961) \quad a_2 = (1.63409,\ 1.41168165) \quad a_3 = (1.61438, 1.41334364)$$
$$a_4 = (1.01234, 1.41849322) \quad a_5 = (1.69556, 0.56043465)$$

Their topology is in Fig.5

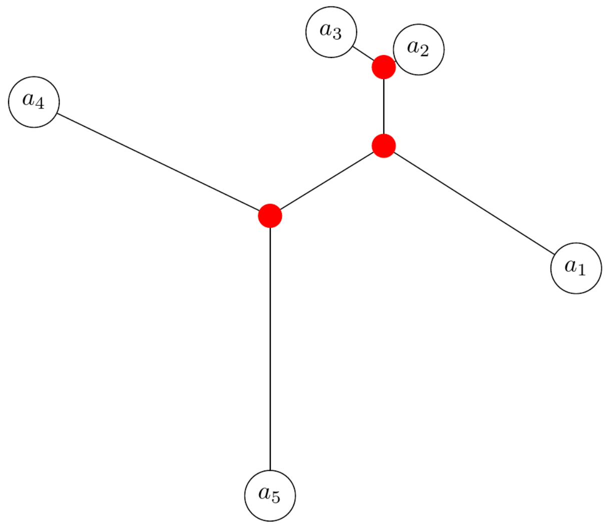

Fig. 5 Topology

By Lemma 2.1.2. we gain their hexagonal coordinate and denoted them as follow:

$$a'_1 = (2.71132438, -1.45549562, -1.25582876) \quad a'_2 = (2.44912478, -0.81905522, -1.63006956)$$
$$a'_3 = (2.43037433, -0.79838567, -1.63198866) \quad a'_4 = (1.83130744, -0.19337256, -1.63793488)$$
$$a'_5 = (2.0191271, -1.3719929, -0.64713419)$$

By Theorem 2.1.4 the rotation matrix $T = \begin{pmatrix} l & k & -k \\ -k & l & k \\ k & -k & l \end{pmatrix}$ and we obtain the rotated hex coordinate for each vertices by $a''_i = Ta'_i$

$$a''_1 = (2.71132438l - 0.19966686k,\ -3.96715314k - 1.45549562l,\ 4.16682k - 1.25582876l)$$
$$a''_2 = (2.44912478l + 0.81101434k,\ -4.07919434k - 0.81905522l,\ 3.26818k - 1.63006956l)$$
$$a'_3 = (2.43037433l + 0.83360299k,\ -4.06236299k - 0.79838567l,\ 3.22876k - 1.63198866l)$$
$$a'_4 = (1.83130744l + 1.44456232k,\ -3.46924232k - 0.19337256l,\ 2.02468k - 1.63793488l)$$
$$a'_5 = (2.0191271l - 0.72485871k,\ -2.66626129k - 1.3719929l,\ 3.39112k - 0.64713419l)$$

Then, by Theorem 2.1.3 we have characteristic function

$$-v_1 + u_2 + v_3 + v_4 + w_5 = 0.63768217k + 2.26572798l = 0$$

Noted by Proposition 2.1.3. $l^2+3k^2 = 1$, so we solve for l, k and get $l = 0.16038977570631482$, $k = -0.5698757524960676$. Now just plug l, k into each a''_i and then using Proposition 2.1.1 we get coordinate of each Steiner point as follow:

$$s_1 = (-0.06936283, 2.18698927, -2.11762644)$$
$$s_2 = (0.09028867, 2.02733777, -2.11762644)$$
$$s_3 = (0.09028867, 1.9460221, -2.03631077)$$

Finally, we need to multiply s_1, s_2, s_3 with T^{-1} and use Lemma 2.1.2 to find their coordinate in Cartesian plane.

$$s_1 = (1.62922803, 1.40771215)$$
$$s_2 = (1.6548345, 1.25012754)$$
$$s_3 = (1.59184586, 1.19870119)$$

The graph is as below:

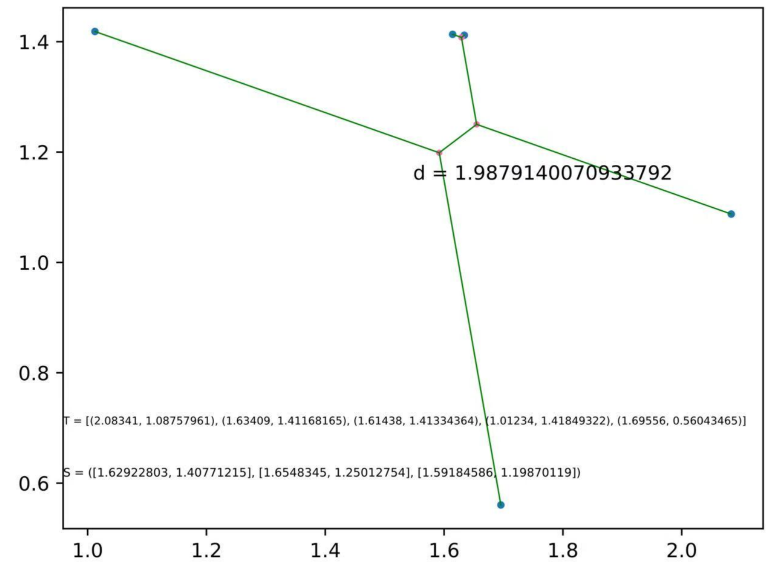

Fig. 6 An example

2.2 Modified Kruskal's method (MK)

We introduce a new variable [5]

$$MVmax = \frac{DSC - IGC}{\text{Unit cost of medium voltage line}}$$

then in each iteration of Kruskal's algorithm [6], we add a new constraint that if the $MVmax$ of both vertices is greater than their distance, then this edge should be added to the network.

Note that, we might end up with a forest then the largest tree will be considered as the main centralized system, the rest pieces could either be smaller centralized system or some decentralized systems.

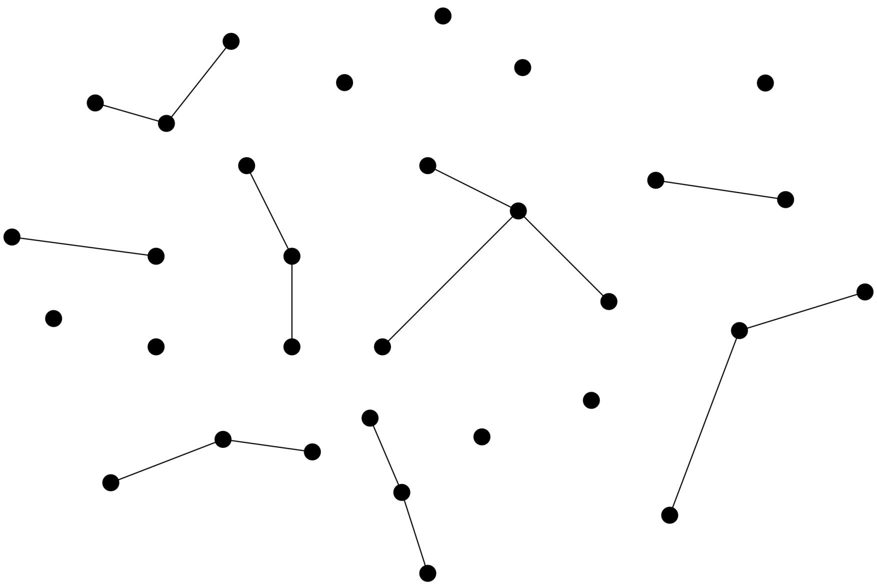

Fig. 7 MK method with flat region

Assuming some vertices belong to a flat region.

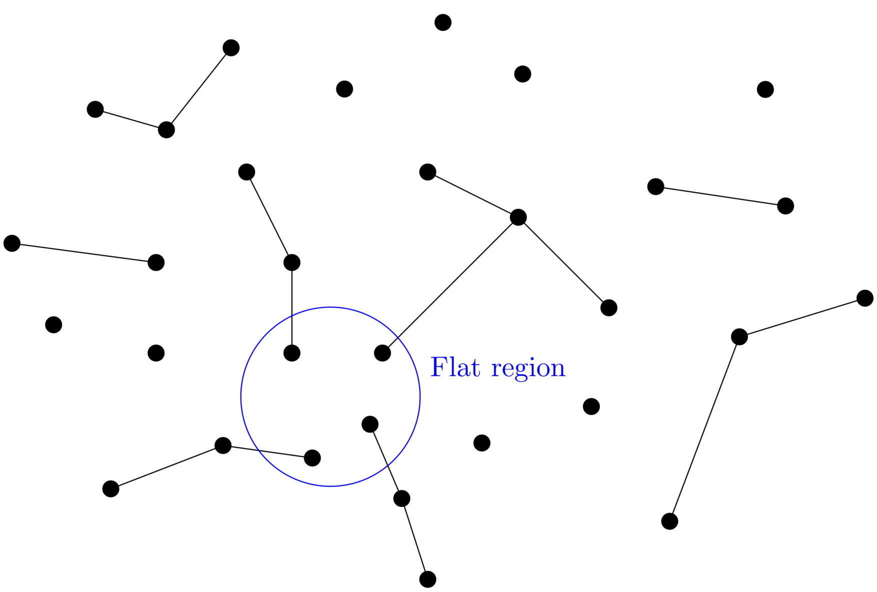

Fig. 8 MK method with flat region

We now check those vertices in flat region, and find Euclidean Steiner tree by HCS.

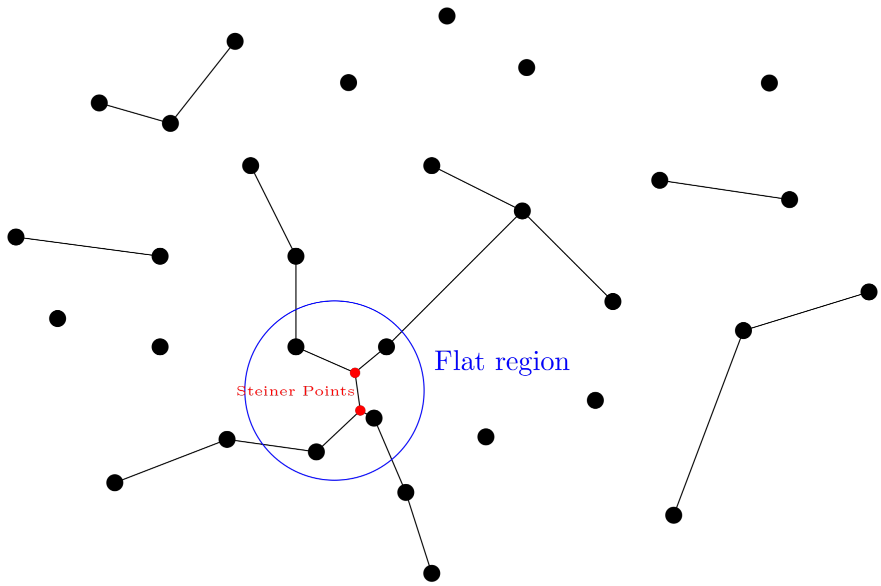

Fig. 9 MK method with flat region and EST

The new tree will have additional red vertices (Steiner points) and additional edges so these points and edges will increase EGC. But the new system will shorten the PSC since we decrease the number centralized system.

3. Result

We do a simulation based on a data generated by ourselves.

- Randomly generate 125 points in a plane
- For vertex i, $DCS_i \sim unif(9000, 15000)$
- For vertex i, $IGC_i \sim unif(4000, 9000)$
- For each Steiner point s_i, its cost is generated by $unif(150, 600)$
- $PST \sim unif(6000, 10000)$
- The unit cost of high voltage line is random variable follow $unif(40000, 60000)$

The plot of 125 points and result of MK method is at next page

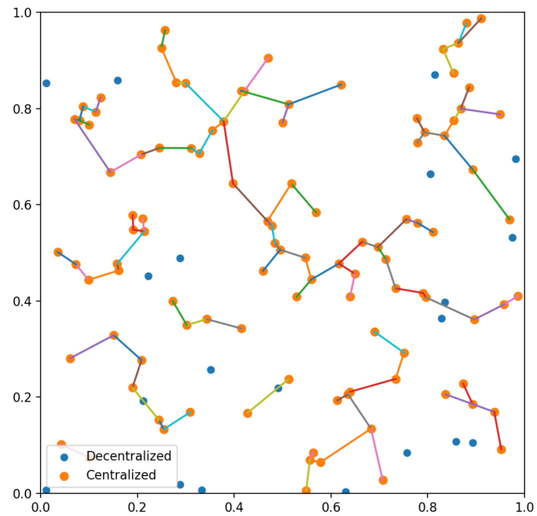

Fig. 10 125 points with graph from MK

Then we randomly generate two circle stand for the flat regions, and we will find EST that connect the components belong to these regions.

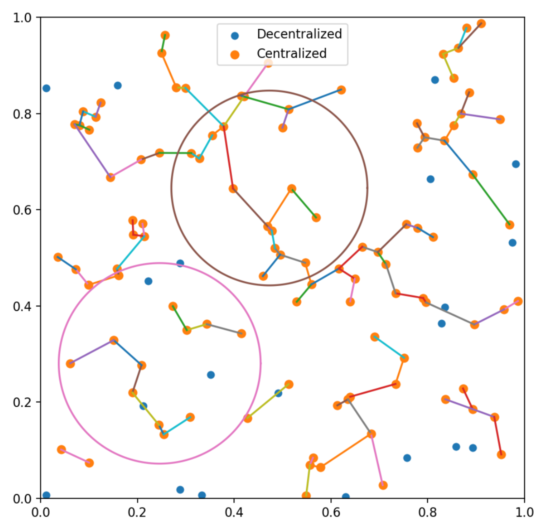

Fig. 11 125 **points with flat region after** MK

We now use HCS to find the EST connect those components. For each Steiner point we added, their cost is a random variable from $unif(150, 600)$

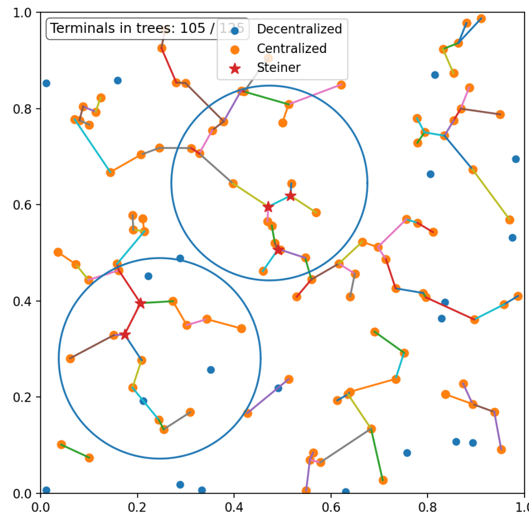

Fig. 12 Final network

4. Discussion

According to our graph, we end up with a network that cover 84% of total vertices, this percentage should be determined by their IGC, EGC, DSC which are generated by uniform distribution. Noted the vertices are randomly generated, in real life their location might depend on the geographical feature. If the components are connected by EST then we have extra EGC since we have additional points (Steiner points) and edges, but we decrease the number of power station, since for each system we need one power station. The potential issue we have here might be the cost of Steiner points, they are fixed cost that generated from uniform distribution with same upper and lower bound as IGC, they might be different since they are different from the demand points. Note the PST is also randomly generated from $unif(6000, 10000)$, but for those centralized system with more demand points, their cost might become larger. Because in real life when a centralized system contain large number of cities, its PST should get increase since more demand points require more energy.

The EST problem is actually NP hard [7]. Generally it will be extremely tough to find the topology of smallest EST from given vertices since there are superexponential number of those[8]. In Fig.11 we don't have many components in flat region so we can just enumerate all cases and then find it out, but in real life this part might be time consuming.

5. Conclusion

We compared our algorithm with Modified Kruskal's algorithm that has no EST.

Algorithm/Costs	Power plant	PST	PST for all systems	Edges	EGC	IGC	DSC	Cost of each Steiner point	Total cost
Modified Kruskal	10	6979.981	69799.81	95	269180.9	2182885	205613	0	1208281
Modified Kruskal with EST	8	6979.981	55839.85	102	277738.6	2182885	205613	1345.719	1204224

From Table above we can conclude for this data the MK with EST seems decrease the cost of electricity grid compared with MK. This might because PST is too large so decrease the number of power plant will save a big cost of building it.

In this paper, we provide an algorithm about minimizing the total cost of electricity network. Each iteration is based on the costs of vertices. We also concluded that the topology of electricity network is mainly determined by the geographical feature, if the flat region is quite common on the map, then we can have more EST to connect more components.

References

[1] Bolukbasi, G., and Kocaman, A. S., "A prize collecting Steiner tree approach to least cost evaluation of grid and off-grid electrification systems," *Energy*, Vol. 160, 2018, p. 536–543. https://doi.org/10.1016/j.energy.2018.07.029, URL https://doi.org/10.1016/j.energy.2018.07.029.

[2] Kaundinya, D. P., Balachandra, P., and Ravindranath, N., "Grid-connected versus stand-alone energy systems for decentralized power—A review of literature," *Renewable and Sustainable Energy Reviews*, Vol. 13, No. 8, 2009, pp. 2041–2050. https://doi.org/https://doi.org/10.1016/j.rser.2009.02.002, URL https://www.sciencedirect.com/science/article/pii/S1364032109000483.

[3] Brazil, M., Graham, R. L., Thomas, D. A., and Zachariasen, M., "On the history of the Euclidean Steiner tree problem," *Archive for History of Exact Sciences*, Vol. 68, 2013, pp. 327 – 354. URL https://api.semanticscholar.org/CorpusID:9211593.

[4] Hwang, F., and Weng, J., "Hexagonal coordinate systems and steiner minimal trees," *Discrete Mathematics*, Vol. 62, No. 1, 1986, pp. 49–57. https://doi.org/https://doi.org/10.1016/0012-365X(86)90040-3, URL https://www.sciencedirect.com/science/article/pii/0012365X86900403.

[5] Parshall, L., Pillai, D., Mohan, S., Sanoh, A., and Modi, V., "National electricity planning in settings with low pre-existing grid coverage: Development of a spatial model and case study of Kenya," *Energy Policy*, Vol. 37, No. 6, 2009, pp. 2395–2410. https://doi.org/https://doi.org/10.1016/j.enpol.2009.01.021, URL https://www.sciencedirect.com/science/article/pii/S0301421509000561, china Energy Efficiency.

[6] Kruskal, J. B., "On the shortest spanning subtree of a graph and the traveling salesman problem," 1956. URL https://api.semanticscholar.org/CorpusID:120068278.

[7] Garey, M. R., Graham, R. L., and Johnson, D. S., "The Complexity of Computing Steiner Minimal Trees," *SIAM Journal on Applied Mathematics*, Vol. 32, No. 4, 1977, pp. 835–859. URL http://www.jstor.org/stable/2100193.

[8] Gilbert, E. N., and Pollak, H. O., "Steiner Minimal Trees," *SIAM Journal on Applied Mathematics*, Vol. 16, No. 1, 1968, pp. 1–29. URL http://www.jstor.org/stable/2099400.