DOI: 10.54254/2753-8818/2025.DL29229

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

Hongzhen Zhu[1*†], Yiran Hu[2†]

[1]Department of Mathematics, University of Illinois Urbana-Champaign, Champaign, USA
[2]Department of Mathematics, University of California, Berkeley, USA
**Corresponding Author. Email: zhu990211@gmail.com*
[†]These authors contributed equally to this work and should be considered as co-first author.

Abstract. Designing electricity supply network is a classic topic in graph theory, it could be considered as network problem of optimizing supplies and demands. However, electricity network could also be solved in a way of optimizing building costs. This paper discussed a modified Kruskal's algorithm of finding electricity network on given vertices with different types of costs: internal grid cost, external grid cost, decentralized system cost, power station cost. The paper also introduces how the Euclidean Steiner Tree works when we worry about the geographical features of network might change the topology of designing. We compared the modified Kruskal's algorithm and modified Kruskal algorithm with Euclidean Steiner Tree and eventually demonstrate the modified Kruskal's algorithm with Euclidean Steiner Tree will reduce the total cost by lowering power station costs despite increasing external grid costs for some cases.

Keywords: Graph Theory, Euclidean Steiner Tree, Kruskal's Algorithm, Electricity Grid

1. Introduction

With rapid development of global demand for energy, different types of electricity grid are shown up everywhere. Two types of grid system will be discussed in this paper: centralized system and decentralized system [1]. In developed countries like The U.S. relies heavily on a centralized electricity system [2], which the electrical power is generated in large scale with resource like fossil fuel plants, nuclear reactors. However, the usage of those resources causes high emissions and some other environmental issues. On the other hand, some developed countries might not have condition to develop a large uniformly centralized system due to economical or geographical reasons. This brings attention to developing some localized electricity infrastructure which runs with renewable energy sources like solar energy, wind power, and water conservancy. Such systems are called decentralized systems.

This paper will use a modified Kruskal's algorithms with Euclidean Stiener Trees to present a reasonable electrical grid that optimize the total cost under certain condition.

DOI: 10.54254/2753-8818/2025.DL29229

2. Problem statement

2.1. Electricity grid

Definition 2.1.1 An electricity grid or electricity network is an interconnected network for delivery electricity to demand points (cities, communities)

Definition 2.1.2 Electrical grid consists 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 2.1.1 Here is an example from Wiki

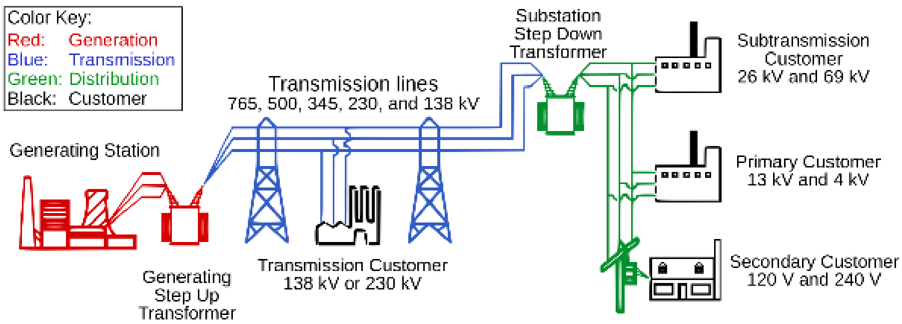

Figure 1. Electrical grid

Remark 2.1.1 In this paper, we will introduce an algorithm to minimize the cost of building an electricity network of given cities(vertices) if there is some cost for edges and vertices.

2.2. Centralized and decentralized system

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

Definition 2.2.1 Centralized system [3]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 2.2.2 Decentralized system [4] is system that cities are powered by some facility nearby themselves, like solar energy or wind power.

Definition 2.2.3 [3]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 2.2.4 For vertices belong to centralized system, they have two types of costs [3]. 1)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. 2) 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. 3) Power station cost/PSC: for each centralized system, there is a fixed cost for building a power station, we defined it as power station cost.

Example 2.2.1 Here is an example of two systems and their costs:

DOI: 10.54254/2753-8818/2025.DL29229

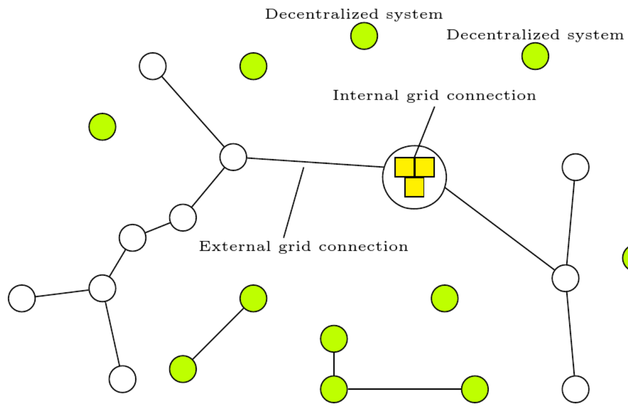

Figure 2. Centralized system and decentralized system

Remark 2.2.1 In real life, some cities with large demand of power might want to build a centralized system [4]. 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 systems. The decentralized systems could be connected with some other decentralized systems or just be isolated vertices.

Remark 2.2.2 Noted the electricity systems in this paper are trees, and for each system we need one power station to provide energy to whole system.

2.3. Euclidean Steiner tree

Definition 2.3.1 Euclidean Steiner problem [5] is to find shortest network connecting the given set of points(terminal) T in Euclidean plane if it is allowed to add extra junctions that is not belong to T .

Example 2.3.1 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 2.3.2 [5]Given a Euclidean Steiner problem, the tree we find out is defined as Euclidean Steiner tree (EST)

Definition 2.3.3 Defining an EST is full if it has $|T| - 2$ steiner points, such EST we called it as Full Steiner tree (FST).

Example 2.3.2 Here is an example of FST when $|T| = 3,\ 4$

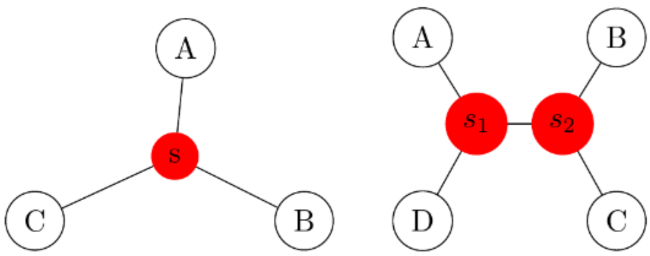

Figure 3. FST for three and four vertices

Theorem 2.3.1 Degree and angle properties [5] 1) Every Steiner point must have three edges incident to it. 2) All Steiner points have three incident edges; any two those three edges have 120

DOI: 10.54254/2753-8818/2025.DL29229

degrees.

Remark 2.3.1 Clearly EST is very rare in real life, since its degree and angle requirement might not be feasible due to some geographical reason.

2.4. Flat region

Definition 2.4.1 Flat region is a region where allow the EST to exist.

Remark 2.4.1 We define flat regions because in real life some cities tend to gather in the flat plane. We will allow $EST\ to\ exist$ ist only in flat regions.

3. Methodology

3.1. Hexagonal Coordinate System (HCS)

We introduce a method for finding the Euclidean Steiner Tree [6], From Theorem 1.3.1 Steiner points have lines only three directions with 120° apart. Therefore, we could use hexagonal coordinate system (HCS). This subsection Introduce an idea from F.K. Hwang [6]

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 $(u,\ v,\ w)$.

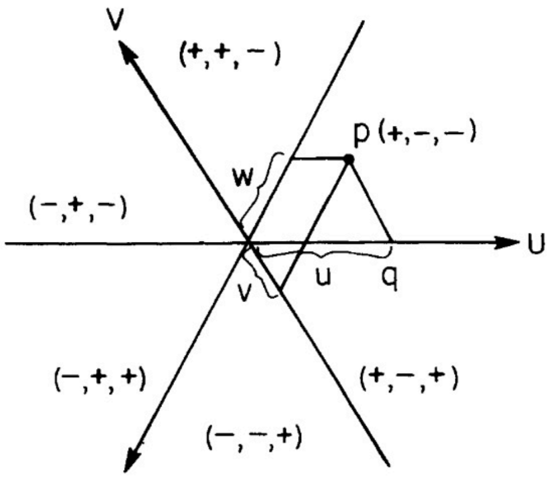

Figure 4. Hexagonal system

Lemma 3.1.1 $u+v+w=0$ for any point $P=(u,\ v,\ w)$

Lemma 3.1.2 Let P has Cartesian coordinates $(x,\ y)$. Then $w=-\frac{2y}{\sqrt{3}}$, $u=x+\frac{y}{\sqrt{3}}$, $v=-x+\frac{y}{\sqrt{3}}$, in particular, $x=\frac{u-v}{2}$, $y=-\frac{\sqrt{3}}{2}w$.

Proposition 3.1.1 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 $w_3=w_1,\ u_3=u_2,\ v_3=-w_3-u_2$

T is bipartite graph with partite set N_1, N_2 , we define direction to each edge and assume an edge starts from a vertex in N_2 and ends in a vertex in N_1 .

Theorem 3.1.1 Define $\epsilon_i=1$ if $a_i\in N_1$ and $\epsilon_i=-1$ otherwise. Define $d_i=u_i\,(v_i,w_i)$, when an edge at a_i is parallel to the $V\,(W,U)$ axis, then $\sum_i^n \epsilon_i d_i=0$, called the characteristic equation of T .

DOI: 10.54254/2753-8818/2025.DL29229

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

Proposition 3.1.2 The characteristic equation can have the first term been arbitrarily set to be u_1, v_1, w_1 .

Theorem 3.1.2 Assume edges of an FST are parallel to the axes only after a clockwise rotation of angle θ . Define $l = \cos\theta,\ k = \sin\frac{\theta}{\sqrt{3}}$. Then the new coordinates can be obtained from the original coordinates through the transformation matrix $\begin{pmatrix} u' \\ v' \\ w' \end{pmatrix} = \begin{bmatrix} l & k & -k \\ -k & l & k \\ k & -k & l \end{bmatrix} \begin{pmatrix} u \\ v \\ w \end{pmatrix}$

Proposition 3.1.3 $l^2 + 3k^2 = 1$

Example 3.1.1 See Appendix

3.2. Modified Kruskal's algorithm (MK)

We introduce a new variable [7] $MVmax = \frac{DSC-IGC}{\text{Unit cost of medium voltage line}}$, then in each iteration of Kruskal's algorithm [8], we add a new constraint that if the $MVmax$ of both vertices is larger than their distance then this edge should be added to the network.

Noted that, we might end up with a forest (a collection of trees) then each piece can be considered as a smaller centralized system or some decentralized systems.

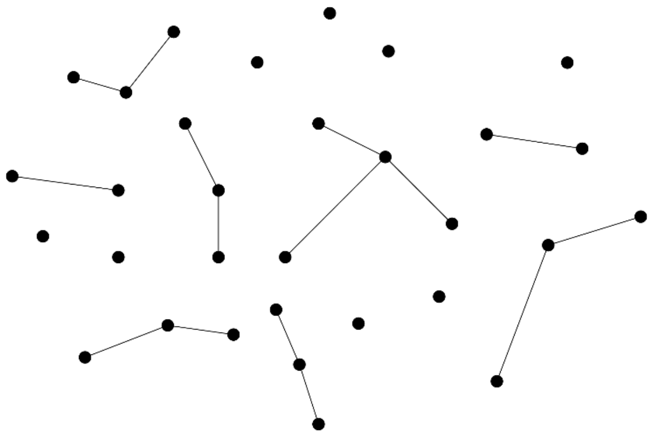

Figure 5. Graph generated by MK

Assuming some vertices belong to a flat region.

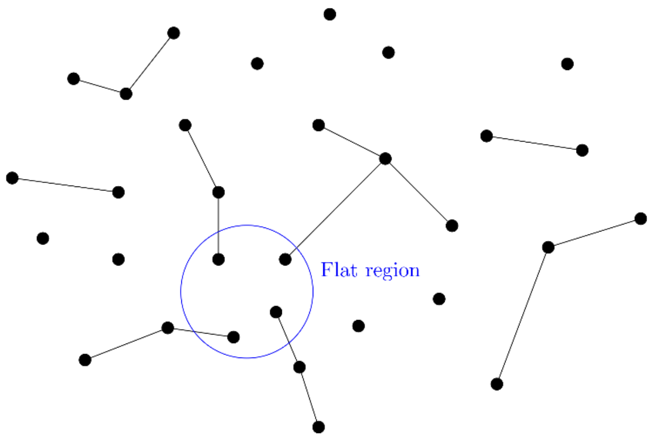

Figure 6. MK method with flat region

DOI: 10.54254/2753-8818/2025.DL29229

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

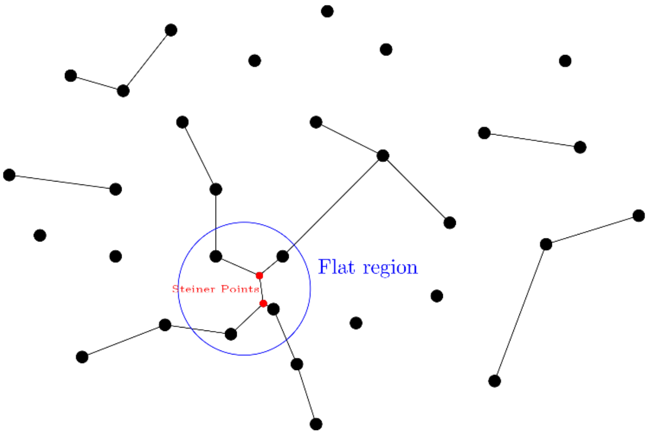

Figure 7. MK method with flat region and EST

The new tree will have additional red vertices (Steiner points), these points will create extra EGC since we have extra edges. However, the network will shorten the PSC since we decrease the number centralized system.

4. Result

We do a simulation based on data generated by ourselves.

- Randomly generates 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 MK method as follow:

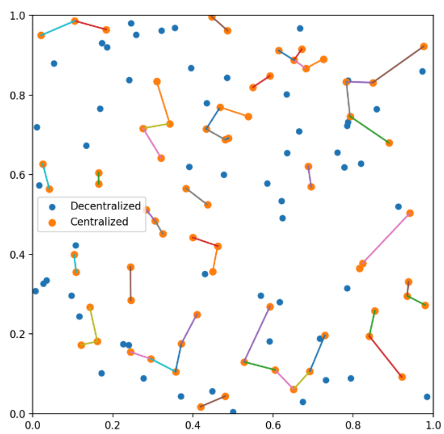

Figure 8. 125 points and graph of MK

DOI: 10.54254/2753-8818/2025.DL29229

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

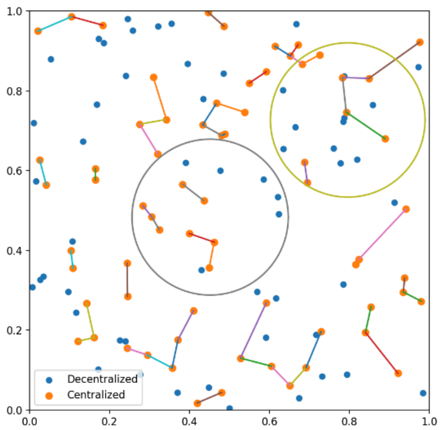

Figure 9. 125 points with flat region after MK

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

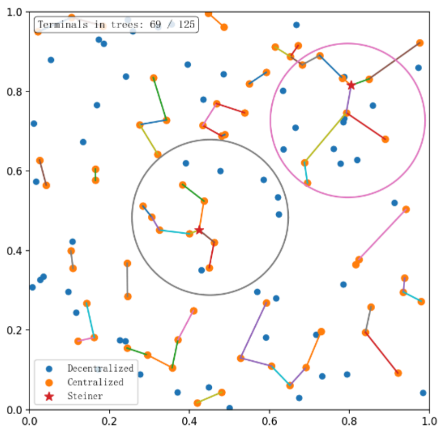

Figure 10. Final network

5. 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

DOI: 10.54254/2753-8818/2025.DL29229

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 contains large number of cities, its PST should get increase since more demand points require more energy.

The EST problem is actually NP hard [9], and sometimes it will be extremely tough to find the topology of smallest EST from given vertices since there are super exponential number of those [10]. In Figure 9 we don't have many components in flat regions so we can just enumerate all cases and then find out, but in real life this part might be time consuming.

6. Conclusion

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

Table 1. Results of two methods

Algorithm/Co sts	Power plant	PST	PST for all systems	Edge s	EGC	IGC	DSC	Cost of each Steiner point	Total cost
MK	10	6979.98 1	69799.81	95	269180. 9	218288 5	20561 3	0	120828 1
MK with EST	8	6979.98 1	55839.85	102	277738. 6	218288 5	20561 3	1345.719	120422 4

From Table 1 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.

Acknowledgement

Hongzhen Zhu and Yiran Hu contributed equally to this work and should be considered as co-first author.

References

[1] AnhNguyen, "Senecaesg, " 7 May 2025. [Online]. Available: https: //senecaesg.com/insights/centralized-vs-decentralized-energy-systems-2025-esg-guide/. [Accessed 24 10 2025].
[2] K. Verclas, "American-German Institute, " 27 April 2012. [Online]. Available: https: //aicgs.org/publication/the-decentralization-of-the-electricity-grid-mitigating-risk-in-the-energy-sector/. [Accessed 24 10 2025].
[3] G. a. K. A. S. Bolukbasi, "A prize collecting Steiner tree approach to least cost evaluation of grid and off-grid electrification systems, " Energy, Vols. Vol. 160, , p. 536–543, 2018.
[4] D. P. B. P. a. R. N. Kaundinya, "“Grid-connected versus stand-alone energy systems for decentralized, " Renewable and Sustainable Energy Reviews, Vols. Vol. 13, No. 8, 2009, p. 2041–2050, 2009.
[5] M. G. R. L. T. D. A. a. Z. M. Brazil, "On the history of the Euclidean Steiner tree problem, " Archive for History of Exact Sciences, vol. 68, p. 327 – 354, 2013.
[6] F. a. W. J. Hwang, "Hexagonal coordinate systems and steiner minimal trees, " Discrete Mathematics, Vols. Vol. 62, No. 1, p. 49–57, 1986.
[7] L. P. D. M. S. S. A. a. M. V. Parshall, "National electricity planning in settings with low pre-existing grid, " Energy Policy, Vols. 37, No. 6, p. 2395–2410, 2009.

[8] J. B. Kruskal, "On the shortest spanning subtree of a graph and the traveling salesman problem, " Mathematics, 1956.
[9] M. R. G. R. L. a. J. D. S. Garey, "The Complexity of Computing Steiner Minimal Trees, " SIAM Journal on Applied Mathematics, Vols. 32, No. 4, pp. 835-859, 1977.
[10] E. N. a. P. H. O. Gilbert, "Steiner Minimal Trees, " SIAM Journal on Applied Mathematics, Vols. 16, No. 1, p. 1–29, 1968.

Appendix

Example 7.1.1 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 Figure 11.

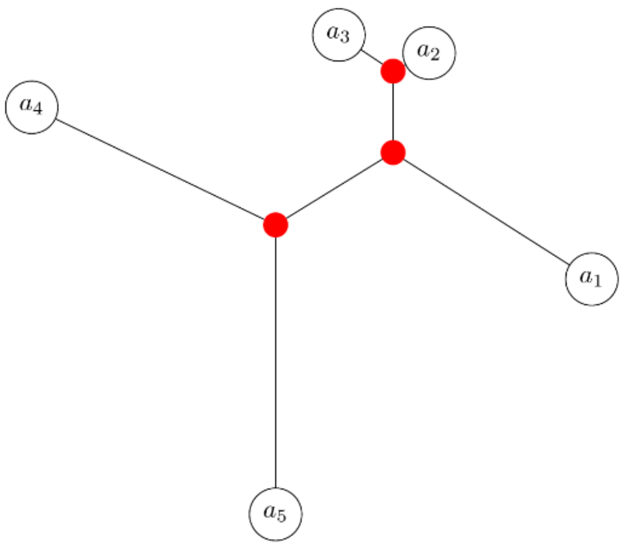

Figure 11. Topology

By Lemma 3.1.2 we gained their hexagonal coordinate and denoted them as follows:

$$a_1^{'} = (2.71132438,\ -1.45549562,\ -1.25582876) \quad a_2^{'} = (2.44912478,\ -0.81905522,\ -1.63006956)$$
$$a_3^{'} = (2.43037433,\ -0.79838567,\ -1.63198866) \quad a_4^{'} = (1.83130744,\ -0.19337256,\ -1.63793488)$$
$$a_5^{'} = (2.0191271,\ -1.3719929,\ -0.64713419)$$

By Theorem 3.1.2 we use rotation matrix T to obtain the rotated hexagonal coordinates for each vertex by $a_i^{''} = Ta_i^{'}$

DOI: 10.54254/2753-8818/2025.DL29229

$$a_1^{''} = (2.71132438l - 0.19966686k,\ -3.96715314k - 1.45549562l,\ 4.16682k - 1.25582876l)$$
$$a_2^{''} = (2.44912478l + 0.81101434k,\ -4.07919434k - 0.81905522l,\ 3.26818k - 1.63006956l)$$
$$a_3^{''} = (2.43037433l + 0.83360299k,\ -4.06236299k - 0.79838567l,\ 3.22876k - 1.63198866l)$$
$$a_4^{''} = (1.83130744l + 1.44456232k,\ -3.46924232k - 0.19337256l,\ 2.02468k - 1.63793488l)$$
$$a_5^{''} = ((2.0191271l - 0.72485871k,\ -2.66626129k - 1.3719929l,\ 3.39112k - 0.64713419l)$$

Then by Theorem 3.1.1 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 3.1.1 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 3.1.2 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:

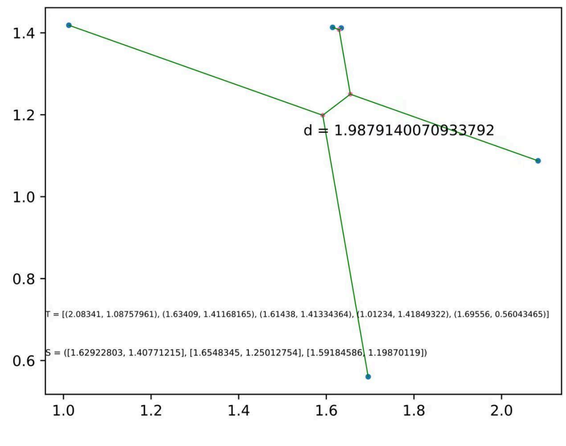

Figure 12. FinalEST