Transformer Location Analysis Based on Poi Information

Yi Yan*, Jianyi Xiao*, Chaolin Zengb, Yongjiao Yangc, Lichao Zhand and Bo Li*
Information Center Guangdong Power Grid Company Limited, Guangdong, China

*Corresponding author e-mail: yanyi@gdxx.csg.cn, 67754750@qq.com, lin_tree@qq.com, yangyongjiao124@163.com, zlc4work@163.com, libo@gdxx.csg.cn

Abstract. The location of the transformer is an important part of the power transmission and transformation project. This paper first summarizes and summarizes the predecessors' rules for transformer location, and draws the following six rules: close to the load center, making the regional power grid layout reasonable, high and low pressure side access lines convenient, station area topography, station area geological. And the application of screening 1451 original poi data to obtain 368 poi data in accordance with the rules of 3 and above, and then introduce the ap clustering model, and apply the data to 9 ap clustering models by setting different parameters. In the optimal result, six clustering points are given, and the scores of the six locations are calculated and normalized to obtain the ranking of the transformers. The first one is the textile street.

1. Introduction
The decision to build a power transmission and transformation project is a complicated task. The proposal of a power transmission and transformation project shall be subject to project feasibility, investment and construction through technical demonstrations such as feasibility study, preliminary design and construction design. The choice of transformer address is an important part of the construction of power transmission and transformation project. Whether the site selection is correct and the overall layout is reasonable, it is very important for infrastructure investment, construction speed, economic and safety of operation. Even a decisive role. Practice has proved that all the work that attaches importance to the preliminary work, the site selection is good, and the overall layout is reasonable and compact, the investment is saved, the construction is fast, and the economic benefits are high. On the contrary, it will cause loss and waste to the power construction and even affect the safe power supply. The selection and determination of the site must be based on scientific and objective reality. It is necessary to comprehensively consider various factors and select the best solution. To this end, through the previous site selection rules of transformers, the site selection rules were summarized, and the AP clustering theory was introduced to make the construction of power transmission and transformation projects scientific and standardized.

In the face of several options for the choice of transformer address, which one can best meet the requirements, it is necessary to comprehensively consider the extent to which the judged plan satisfies all aspects of the conditions and draw scientific conclusions. This decision-making process, due to a variety of factors, must consider various factors to make the final decision. The better method is to use AP clustering. The basic idea and mathematical model of the method and the specific application in transformer address selection will be discussed below.
2. Transformer address selection
Since the index problem of transformer location has been given the relevant standards by the predecessors, we have carried out the following summary:

1. Close to the load center: The choice of the substation site must be adapted to the requirements of the power system development planning and layout, as close as possible to the main users and close to the load center. In this way, the investment and power loss of the transmission and distribution lines are reduced, the probability of causing the accident is reduced, and other problems caused by the site being away from the load center can be avoided.

2. Saving land: saving land for engineering is our national policy. We need to follow the principle of technical and economic rationality, rationally arrange, and improve the utilization rate of land as much as possible. Any land that can be used by wasteland cannot occupy cultivated land, and any land can be used. Do not occupy good fields. In particular, it is necessary to avoid occupying land with high economic benefits such as vegetable fields. The land should be compact and adapted to local conditions. The use of inferior land as a site selection option is one of the main conditions for determining the quality of a design.

3. Terrain and landform of the station area: When selecting the site, the construction policy based on agriculture should be implemented. It is necessary not only to implement the spirit of saving land, occupying or occupying less farmland, but its topography and landform should be conducive to the layout of substations.

4. Transportation: Sites should be selected as close as possible to existing or planned railways, highways and other transportation lines to reduce transportation investment, speed up construction and reduce transportation costs. The choice of site should also consider the transportation of equipment materials during construction. Special consideration should be given to large equipment.

5. Make the regional power grid layout reasonable: the original power supply, new power supply and planned power supply should be considered. The regional power supply and substation are not concentrated on one side, so that the power supply layout is dispersed, which can reduce the investment and network loss of the secondary network, and can reach the day of safe power supply.

6. Environmental Protection: With the development of industrial and agricultural production, the issue of environmental pollution has been increasingly valued by countries all over the world. Because of the continuous emergence of new industrial bases, it will inevitably bring a series of new pollutant discharge problems to the environment. Protecting the environment and improving the environment are crucial.

7. High and low pressure side inlet and outlet lines are convenient: When selecting the site, the corridors of voltage inlet and outlet lines should be considered to make the entry and exit lines convenient, reduce the crossover of the transmission lines, and reduce the line corner.

8. Station area geological conditions site: The selection should fully consider the geological conditions of the station area, avoiding faults, landslides, subsidence areas, caves and other geological conditions, and should also avoid mining areas.

It can be seen from the above that there are conceptual overlaps between the eight indicators and considering the feasibility of applying substation location indicators. It can be seen that the land use indicators and the site topography are duplicate concepts, which can be merged into station terrain features; transportation, To make the regional power grid layout reasonable, environmental protection is also a repetitive concept, which can be replaced by rationalizing the regional power grid layout. Therefore, the five indicators selected after summarization and streamlining are as follows:

(1) Close to the load center
(2) Make the regional power grid layout reasonable
(3) High and low pressure side access lines are convenient
(4) Topographical features of the station area
(5) Geological conditions in the station area
3. Ap algorithm principle
After screening the region's poi by the above selection indicators, after obtaining the geographic location data, we adopt the AP clustering algorithm because the algorithm has the following advantages:

(1) There is no need to specify the number of clusters in advance
(2) The clustering result is very stable
(3) Applicable to asymmetric similarity matrix
(4) The initial value is not sensitive
Therefore, it is suitable for clustering of this data.

3.1. Basic idea
The basic idea of the AP algorithm is to regard all samples as nodes of the network, and then calculate the cluster center of a sample through the message transmission of each side of the network. In the clustering process, there are two kinds of messages transmitted between the nodes, which are the degree of attraction and the degree of attribution. The AP algorithm continuously updates the attraction and attribution of each point through an iterative process until m high-quality Exemplars (equivalent to centroids) are generated, and the remaining data points are allocated to the corresponding clusters.

3.2. Algorithm step
At the beginning of the AP algorithm, all nodes are regarded as potential cluster centers, and then the most suitable cluster centers are found through communication between nodes.

The input is the similarity matrix S between nodes:

Where S(i,j) represents the similarity between node i and node j, and also indicates the appropriate degree of j as the clustering center of i. Usually S(i,j) takes the negative value S(k,k) of the Euclidean distance of i,j to indicate the appropriate degree of the node k as the clustering center of k, that is, the appropriate degree of the node k becoming the clustering center. In the beginning, this value is the value given by the user at initialization, which will affect the final number of clusters. Usually, the minimum or median of the entire matrix is taken. The larger the value is, the more the number of classes will eventually be generated. Many.

There are two types of messages transmitted between AP nodes: attraction and attribution.

First, the degree of attraction is the information of node i to node k, and the degree of attraction of node k to node i is transmitted, denoted as r(i,k). We have a similarity matrix to record the appropriateness of the cluster center where k becomes i. S(i,k), then we only need to prove that k is more suitable than other nodes. For other nodes k', we have S(i,k') to represent node k' as the cluster center of node i. Suitability.

Then define an a(i,k') to indicate the acceptance of i for node k'

Adding these two values, a(i,k')+S(i,k') can calculate the appropriate degree of node k' as the cluster center of node i. Here, find the largest a(i,k')+S(i,k') in all other node nodes k', ie max {a(i,k')+S(i,k')} S(i,k)-max {a(i,k')+S(i,k')} can get the attractiveness of k for i: r(i,k)=S(i,k)- Max {a(i,k')+S(i,k')}

Next, the degree of attribution a(i, k) mentioned above is calculated, indicating the degree to which node i selects node k as its clustering center. If the degree of suitability of the node k as the cluster center of the other node i' is large, the degree of suitability of the node k as the cluster center of the node i may also be large. Then we can first calculate the attraction r(i',k) of node k to other nodes and then do an accumulation and indicate the attraction of node k to other nodes: \( \sum \max \{0,r(i',k)\} \) and then Add r(k,k) to why we add r(k,k). According to the attractiveness formula, we can see that other r(k,k) reacts to how much node k is not suitable for being classified into other aggregates. Class center. That is, a(i,k)=min{0,r(k,k)}+ \( \sum \max \{0,r(i',k)\} \), a(k,k)=max {0,r(i',k)} mainly reflects the ability of k as a clustering center

Convert the above text language into a mathematical language as follows

Update the attractiveness matrix:

\[
r_{t+1}(i,k) = \begin{cases} 
S(i,k) - \max\{a_t(i,j) + r_t(i,j)\}, & i \neq k \\
S(i,k) - \max\{S(i,j)\}, & i = k 
\end{cases}
\]
Update attribution matrix

\[ a_{t+1}(i, k) = \begin{cases} 
\min\{0, r_{t+1}(k, k) + \sum_{j \neq i, k} \max\{r_{t+1}(i, k), 0\}\}, i \neq k \\
\sum_{j \neq i} \max\{r_{t+1}(i, k), 0\}, i = k 
\end{cases} \]

Attenuate the two equations according to the attenuation coefficient \( \lambda \)

\[ r_{t+1}(i, k) = \lambda \cdot r_t(i, k) + (1 - \lambda) \cdot r_{t+1}(i, k) \]
\[ a_{t+1}(i, k) = \lambda \cdot a_t(i, k) + (1 - \lambda) \cdot a_{t+1}(i, k) \]

4. Algorithm case

The basic idea of the example is to obtain the poi information of the transformer location area, and then meet the conditions close to the load center according to the screening rules, so that the regional power grid layout is reasonable, the high and low pressure side access lines are convenient, the station terrain and geomorphology, and the station area geological conditions data. If it is selected, if it can satisfy almost none of the above five conditions, the screening conditions can be reduced. If three conditions or more can be satisfied at the same time, the selected data can be selected into the AP clustering model, and \( n \) can be obtained. Cluster points, near the cluster point is the transformer construction address we selected, the flow is shown in Figure 1:

![Figure 1. Algorithm Case Process](image)

4.1. Data preparation

Get 1451 data of a certain area, the field is (1) type of poi point (2) longitude (3) latitude, the data is as shown in Figure 2:

![Figure 2. Original poi information](image)
4.2. Data screening
Screening and judging each piece of data, since there are few points that can satisfy the conditions, the conditions for screening are lowered, and the poi point information satisfies three or more conditions, and the selected data is 368. The final data is as shown in Figure 3 below:

![Figure 3. Poi data after screening](image)

4.3. Model training
For the training of the model, the Euclidean distance is used as the similarity matrix. The parameter preference is even if the value of the diagonal of S is larger, the larger the value will gather more points, so the median of the similar matrix is adopted, and 0.44 is incremented as a multiplier, and a total of 9 sets of training are performed.

4.4. Model result
The results of training 9 sets of data are shown in Figure 4 below.

![Figure 4. AP clustering results](image)
It can be seen from the above figure that the p-values of each group of training are close to 0. Among them, the 7th group training is abnormal and should be discarded. The other 8 training models are within the normal orientation. The corresponding model can be selected for transformer selection according to the specific situation. Address, here select the last model for analysis. The preference value of the last model is -1.49E-04, and the number of clusters that the algorithm iterates out is six. The locations corresponding to the six cluster points are as follows:

Table 1. Cluster points

| Latitude and longitude | Location name          |
|------------------------|------------------------|
| 113.270646 23.109531  | Textile street         |
| 113.28295  23.10908   | Tianyu Company         |
| 113.267586 23.10902   | Building 15 of Datuntang |
| 113.275278 23.108979  | Bread chop             |
| 113.263378 23.108978  | Yimi Hotel             |
| 113.279113 23.108917  | Lingnan Animal Hospital |

After extracting the cluster points, the attractiveness of the six locations to the other 368 locations is calculated, and the scores are normalized and ranked, as shown in Table 2 below:

Table 2. Cluster Point Ranking

| Location name          | Score ranking |
|------------------------|---------------|
| Textile street         | 1             |
| Bread chop             | 0.916154      |
| Building 15 of Datuntang | 0.863204     |
| Lingnan Animal Hospital | 0.578857     |
| Yimi Hotel             | 0.422264      |
| Tianyu Company         | 0             |

References

[1] Zhang Suzhi, Yang Wei, Chen Xiaoni, Li Penghui. A coarse-grained parallel AP clustering algorithm based on intra-class and inter-class distance [J]. Journal of Central China Normal University (Natural Science), 2018, 52 (06): 781 -787+797.
[2] Wang Yi. Research on county distribution network planning based on hierarchical partitioning [D]. Shandong University, 2018.
[3] Zhuo Binbin, Shi Wendi, Zhu Yadong, Zhong Zhixiang. Research on the Model of Substation Site Selection and Constant Capacity Planning [J]. Electrical Engineering, 2018 (08): 46-50.
[4] Chen Yunfei, Du Taihang, Jiang Chundong, Wang Jingyu, Li Juanmei. Research on indoor positioning based on AP layout optimization and K-means clustering algorithm [J]. Application of Electronic Technique, 2018, 44 (03): 68-71.
[5] WANG Tianfang, ZHANG Linghua. Indoor Positioning Algorithm Based on Fuzzy Classification and AP Weighting [J]. Computer Technology and Development, 2017, 27 (04): 46-50.
[6] Xu Qing, Wu Jie. Application of Fuzzy Comprehensive Evaluation in Location Selection of Substation [J]. Electric Power Construction, 2004 (07): 24-26.