Evaluation of Agricultural Economic Information Based on Kruskal Algorithm and Principal Component Analysis

Chenzhong Lyu(✉)

College of Rural Revitalization, Sichuan Open University, Chengdu 610000, China
lvchengzhong2022@163.com

Abstract. Kruskal algorithm is a data reduction method Kruskal algorithm and PCA with different component numbers. Through the comprehensive evaluation of the existing rural economic information system, it can provide governments at all levels and relevant departments with scientific evaluation basis for the construction status, service level and benefits of the economic information system. On the one hand, it can avoid inefficient operation and resource waste caused by repeated construction; On the other hand, we found the weak links in the construction and development of rural economic information system through evaluation, and strengthened the software and hardware construction of the system itself.

Keywords: Principal component analysis · Kruskal algorithm · Evaluation of agricultural economic information

1 Introduction

In the process of agricultural production, agricultural information technology does not exist out of thin air. It can be used in agricultural production. whose main function is o provide information services to the outside world (including farmers, governments, etc.), so as to realize the penetration of information technology into the agricultural production process and the integration with agriculture. The establishment of rural economic information system has solved this problem well [1].

The rural economic information system refers to a whole composed of institutions, personnel, information resources, information infrastructure and necessary information technologies engaged in rural economic information work. Its main function is to collect, sort, process, store, transmit and feedback rural economic information. On the one hand, it provides scientific basis for decision-makers at all levels to formulate agricultural development strategies and guide the development of rural economy; on the other hand, it provides services for the production and operation of farmers, thus promoting the development of rural economy towards market economy and industrialization. The transparent and orderly information environment provided by the rural economic information system, a sound and smooth information network and comprehensive and high-quality information services are of great [2].

J. C. Hung et al. (Eds.): FC 2023, LNEE 1132, pp. 223–228, 2024.
https://doi.org/10.1007/978-981-99-9538-7_33

China's rural economic information system is still in the development stage. Although the construction of the national, provincial, municipal (regional) and county backbone networks has been basically completed, the extension of the backbone networks below the county level still needs further development, which means that the "last kilometer" problem we usually refer to has not been properly solved. The information system below the county level. Another obvious problem is the low efficiency caused by the lack of "roads", "vehicles", "goods" and "drivers" in the construction and development of economic information system ' The road is the rural economic information backbone network, the car is the various application systems used in the network system, the goods is the data information loaded on the application system, and the driver is the information service personnel within the system [3]. How to organize them so that they can play a full role directly affects the efficiency of the rural economic information system, which requires a comprehensive evaluation of the construction and development of the rural economic information system to find out the problems and propose targeted solutions.

2 Related Work

2.1 Research Status at Home and Abroad

The research on the comprehensive evaluation of the rural economic information system is relatively few in China at present. Its research level is limited to the research on economic significance and the construction of the rural economic information system. In addition, because the comprehensive evaluation of rural economic information system belongs to the research scope of system evaluation, the information system evaluation involving the research scope of system engineering also has a strong guiding role in the research of the comprehensive evaluation of rural economic information system [4].

As for the evaluation of information systems, China has also carried out relatively in-depth research within the scope of system engineering. Some of the research focuses on the determine their degree of advantage through comparison and calculation. Cost benefit analysis method is generally adopted. There is also the comprehensive evaluation of information systems, which is a vexing problem. The reason is that this is a fuzzy problem with no gray elements. Multiple Objective Decision Making (MODM), Grey Evaluation, Delphi, HGF, etc.

2.2 Rural Economic Information

Rural economic information refers to the description and true reflection of the essential attributes, movements and characteristics between various elements in the agricultural economic system and between these elements and the economic system. It is reflected by a large amount of data and information through the mutual connection and transformation between various industries and departments of material, energy, capital and people flow. Generally speaking, rural economic information is the reflection of the essential attribute, movement state and characteristics of rural economic activities. The development and utilization of rural economic information is not only conducive to the correct decision-making and planning of the rural economy, but also conducive to strengthening the

communication between the rural economic system and the external economic system, so that the rural economic system can operate in an orderly manner.

Rural economic information resources refer to economic information service organizations, information networks, service personnel, information facilities, information physical media, agricultural information centers, policy information institutions and rural information industries in rural areas. These resources are processed, processed, organized, managed and developed to become economic information products or serve economic information, and are reasonably distributed and utilized through the market to create benefits for agricultural production. Rural economic information resources exist in the whole process of agricultural economic activities, including production, exchange, distribution, consumption and other aspects, all over the rural social levels. The flow of information is criss cross. The top-down economic information flows from the agricultural management departments and higher level social service organizations of the government to the grassroots agricultural operators. The bottom-up economic information flow mainly comes from the grassroots agricultural operators and relevant non-agricultural operators. Its process is just the opposite of the "top-down" process, and flows to high-level government agricultural organizations. The horizontal economic information process mainly refers to the information communication between various relevant agricultural production and socio-economic organizations. Under the condition of market economy, all kinds of agricultural economic activities are inseparable from the role of market mechanism. The communication of most agricultural economic information should also be realized through the market, which makes the market information become an important information resource focused by organizations at all levels, from the central government to farmers and farmers' enterprises.

3 Kruskal Algorithm and Principal Component Analysis

Kruskal algorithm was first proposed by Joseph Bemaid Kruskal in 1956. It is a classical greedy algorithm used to solve and which does not form a cycle with the selected edge, and add it to the edge set E; Otherwise, select the next edge until there are n - 1 edges in edge set E (n vertices in the graph).

The basic calculation steps are: given an initial connected graph with n vertices, the initial total vertex set is the vertex of the connected graph, the selected vertex set and the selected path set are empty sets, and the candidate path set is the branch set of the connected graph; After the above initialization, an independent tree with each vertex as the root node is formed; According to the weight of each path, select the path with the lowest weight in the candidate path set. If the two vertices of the path do not belong to the same tree, move the path into the selected path set, and move the new vertices at the end of the path into the selected vertex set; Update the candidate path set, and repeat the above path selection process until the set of selected points is equal to the total vertex set, that is, finally form the minimum number of n vertices, that is, form a tree graph with n-1 edges.

When using the standard K algorithm to solve the minimum tree of a connected graph, first give the weight values of the connected graph, the paths to be selected, and each path, as shown in Fig. 1.

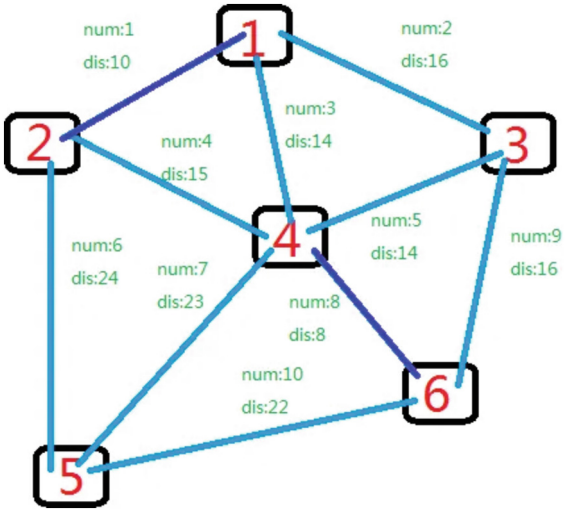

Fig. 1 Kruskal Search Process Chart

By extracting the principal component features, eliminating the correlation between data and reducing data redundancy, the dimension of the original data can be greatly reduced, while maintaining most of the information of the original image, which solves the bottleneck problem that the image dimension is too high to process or too slow to process.

At present, digital images are growing explosively. Because of the high similarity between adjacent pixels, the image representation process contains redundancy; At the same time, the dimension of image in feature representation is very high. In the process of processing and computing, large storage space must be allocated, and a large amount of computing time must be consumed. Therefore, high-dimensional data processing has become the bottleneck of the problem. The effective processing of high-dimensional data by principal component analysis enables it to be used for preprocessing, redundancy removal, feature extraction, data compression, etc. Principal component analysis is one of the most successful techniques in image recognition and compression.

4 Evaluation of Agricultural Economic Information Based on Kruskal Algorithm and Principal Component Analysis

As a tool of discrete pattern classification, graph theory has the advantages of intuitiveness and clarity. More importantly, it is able to transform a pattern classification problem into a very simple graph, and then analyze the classification complexity of the original pattern through the characteristics of the graph. In the process of research, Yucheng administrative division map was registered according to the spliced topographic map, and the longitude and latitude coordinates of the centroid point were extracted as the coordinates of the evaluation unit.

While considering the spatial adjacency, it is more important to consider the internal similarity between the zoning units of each township. Therefore, the weighted connected graph generated by connecting lines is used to represent the zoning. Finally, the undirected graph shown in Fig. 2 is obtained. In the undirected graph, the weight values of edges are sorted from small to large, so as to obtain the weight value increasing sequence of edges, and then the set of edges is selected by increasing the weight value sequence in the graph. If the newly selected edge and the confirmed edge form a loop (that is, a ring edge sequence connected end to end), discard the edge and continue to select the next edge in the edge sequence with increasing weight until the last edge in the sequence. The meaning of the blue line in Fig. 2 shows the spatial adjacency and internal similarity between each point (township), where spatial adjacency represents geographical location information, and internal similarity represents the similarity of agricultural economic information data or conditions. The directly connected villages and towns indicate that the conditions of agricultural economic information indicators for evaluation are similar, and large-scale comprehensive planning can be formed accordingly when determining agricultural economic zoning.

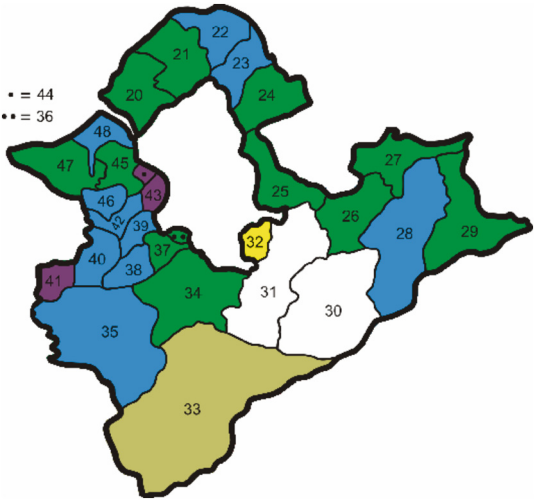

Fig. 2 Kruskal algorithm cluster diagram of 11 townships in Yucheng City

Combined with the minimum spanning tree Kruskal algorithm clustering results of agricultural economic information of 11 villages and towns, the comprehensive evaluation results of agricultural economic information of 11 villages and towns were obtained, and the following comprehensive ranking was obtained: municipal central office>Fangsi Town>Lun Town>Ju Town Township>Liangjia Town>Xinzhai Town>Xindian>Zhangzhuang Town>Shiliwang Township>Litun Township>Anren Town.

5 Conclusion

The two-dimensional graph theory clustering analysis graph based on the spanning tree Kruskal algorithm considers the geographical location, but does not distinguish the size comparison of the same subtree location, while the comprehensive score of the principal component analysis does not consider the coordination of the spatial location. Therefore, the two methods can complement each other and organically combine. The comprehensive evaluation results of agricultural economic information obtained by integrating Kruskal algorithm and principal component analysis are more convincing because they make the evaluation process more scientific. The comprehensive evaluation result intensity of agricultural economic information of 11 towns and townships in Yucheng City is: municipal central office>Fangsi Town>Lun Town>Ju Town>Liangjia Town>Xinzhai Town>Xindian>Zhangzhuang Town>Shiliwang Township>Litun Township>Anren Town.

References

1. Guo, X., Pei, L., Sui, J., et al.: Agricultural support policies and China's cyclical evolutionary path of agricultural economic growth. Sustainability **13**(11), 6134 (2021)
2. Starke, F., de Rijk, A., Wegewitz, U.: Effect, economic and process evaluation of a mental health return to work intervention: study design. Eur. J. Public Health 32(Supplement_3), ckac130–211 (2022)
3. Xc, A., Feng, L.B., Yh, C., et al.: Two-dimensional evaluation model of electrical equipment based on combined weighting and rating algorithm. Energy Rep. **7**, 443–448 (2021)
4. Zedan, A., Helmey, M.: The impact of agricultural economic policies on the current status of the area and productivity of grain crops in Egypt. Egypts Presidential Specialized Council for Education and Scientific Research. Fayoum J. Agric. Res. Dev. **35**(2), 380–390 (2021)