Research on Management Decision Optimization Algorithm Based on Information Entropy in Cloud Computing Environment

Shiyun Sun¹,*
¹Yunnan Technology and Business University, Yunnan, China
*Corresponding author e-mail: shiyun.sun@stu.nida.ac.th

Abstract. In the cloud computing environment, information entropy is a false judgment on the composition of information and data inaccuracy and decision-making. By using information entropy to optimize the data information, under the Internet, the analysis of network user management data in the cloud computing environment can be quickly identify clustered data objects and reorganize management data in a cloud computing environment to improve the effectiveness of management decisions.

Keywords: Data Mining, Decision Support, Information Entropy Algorithm, Cloud Computing Environment

1. Introduction
The development of the background makes enterprises in the foundation Collect big data and emergency data, analyze its capabilities, and be able to complete tasks in challenges [1-3]. Big data-related database management systems can help companies obtain effective decision-making resources, which leads to scientific and effective optimization methods for companies’ organizational performance, increasing the competitiveness of enterprises in the industry [4-6].

The current information entropy, such as association rule mining, decision tree, genetic algorithm, etc., has solved the problems of the relationship and organizational structure of information data to a certain extent. However, on this basis, can data be more accurately and efficiently analyzed? Analyzing and using information is an objective requirement for the development of data mining. The uncertainty analysis of the decision support system based on information entropy can help the data used in the decision-making process to be further refined and sorted while using mining algorithms to strengthen the regularity of the processed data organization structure And rationality.

2. Optimization of information entropy algorithm based on information entropy in decision support system

2.1. Information entropy algorithm process
Suppose the data object is: \( y=(y_1, y_2, ..., y_m, i=1,2,N) \), the algorithm initializes the parameters at the starting position and initializes the pheromone distributed. Here, the initial value of the pheromone is set to 0, that is, \( T_i(0)=0 \), and each parameter is set in sequence at the starting position, such as: r is
the radius of the cluster, and ε is the statistical error. The distance between $y_i$ and $y_j$ is the weighted Euclidean distance, where the weighted Euclidean distance is represented by $d_{ij}$. Here, the weighted Euclidean distance between $x_i$ and $x_j$ can be expressed as:

$$d_{ij} = \sqrt{\sum_{k=1}^{m} P_k (Y_{ik} - Y_{jk})^2}$$

(1)

$$\tau_{ij}(t) = \begin{cases} 1, & d_{ij} \leq r \\ 0, & d_{ij} > r \end{cases}$$

(2)

After the ants distinguish the characteristics and attributes of the pile, they will choose to pick up or put down the behavior, then these two behaviors will be carried out according to a certain probability rule, and different spatial densities will form different probabilities, and $y_i$ will be merged into $y_j$. The probability is:

$$p_{ij}(t) = \frac{\tau_{ij}^{\alpha}(t)U_{ij}^{\beta}(t)}{\sum_{u \in U} \tau_{uj}^{\alpha}(t)U_{uj}^{\beta}(t)}$$

(3)

The process of information entropy algorithm can be shown in Figure 1.

![Figure 1. Information entropy algorithm process diagram.](image)

2.2. Information entropy algorithm optimization ideas based on information entropy

If the information entropy of the object $T_i$ before being picked up is greater than the information entropy after being picked up, the picking behavior will occur. For example, if the information entropy before $T_i$ is $E_1$ greater than the information entropy $E_2$ of the area after being picked up, then pick up. Similarly, an ant loaded with $T_i$ needs to calculate the information entropy of the object in the $d \times d$ area in the blank space. If the information entropy of the object in the $d \times d$ area is not put down, it
will be put down. For example, if the information entropy $E_1$ before putting it down is greater than the $E_1$ after putting it down, let it go.

If any object has $n$ mutually independent attributes $M_1, M_2, ..., M_n$ with value ranges $x_1, x_2, ..., x_n$, then the information entropy of the object in the area can be expressed by the following formula:

$$E(x^2) = -\sum_{i=1}^{n} \sum_{x \in X_i} P(x) \log p(x)$$  \hspace{1cm} (4)

$$p(x) = \frac{x}{y}$$  \hspace{1cm} (5)

3. Cluster analysis

In the calculation process, the types with higher entropy calculation values are clustered and then eliminated. The parts with lower entropy results are aggregated with higher efficiency. Each type of customer browsing path left after the elimination has similar attributes. And the difference in customer characteristics between classes is significant (Figure 2).

![Figure 2. Classification of 7 types of customers.](image)

By categorizing the attribute characteristics of the customer browsing path, the type of each type of browsing path is described, and according to the description, the various customer browsing path information that needs to be processed in the future can be classified and analyzed. There is a big difference between as shown in Table 1.

| Category | Type of customer                      |
|----------|--------------------------------------|
| 1        | Low value no call type               |
| 2        | General value customers              |
| 3        | Roaming                              |
| 4        | Public office                         |
| 5        | Highly loyal mass office             |
| 6        | High value                            |
| 7        | High-value business travel type      |

After classifying the customer browsing path information, in order to ensure that the information or data processing process is better than the algorithm before optimization to a certain extent, it is necessary to compare the misclassification error rate and the number of iterations of the entropy-based ant colony algorithm. Through the comparison of the two factors, more objective information on the efficiency of the algorithm can be obtained, and the process helps to further confirm the operability and superiority of the algorithm. Table 2 shows that the information entropy algorithm more accurately divides the customer path types into 7 categories after taking 100,000 repeated pick-up and drop-off
activities. However, a large number of repetitive activities slows down the calculation speed and affects the processing efficiency.

**Table 2.** 5 corresponding customer types.

| Category | Structure | Customer Type           |
|----------|-----------|-------------------------|
| 2        | Single    | General value customers |
| 4        | Double    | Public office           |
| 5        | Single&Double | Highly loyal mass office |
| 6        | Double    | High value               |
| 7        | Single&Double | High-value business travel type |

**Table 3.** Comparison of two clustering methods.

|                        | Information entropy algorithm | Adjusted information entropy algorithm |
|------------------------|-------------------------------|----------------------------------------|
| Misclassification error rate/% | 5.2                           | 10.3                                   |
| Number of iterations   | 100000                        | 46000                                  |

On the whole, the entropy-based ant colony algorithm improves the efficiency of the algorithm in the whole process, but the complexity and time consumption of its algorithm are slightly higher than that of the ant colony algorithm. The algorithm process makes certain changes to the ants and the objects that are picked up and put down. The data objects are subjected to the uncertainty calculation of information entropy before being selected. In this way, the range of the objects that the ants pick up and put down is reduced. This process can further improve the efficiency of clustering and also improve the accuracy of customer segmentation to a certain extent. It provides a more accurate basis for companies to focus on high-precision customers and make more effective marketing decisions.

### 4. Measures to improve management decisions

#### 4.1. Building an enterprise-level big data integration system

Enterprise management decision-making may be somewhat complicated due to a large amount of information and changing factors. Therefore, it is necessary to establish a large-scale enterprise data integration system and use big data technology in the form of cloud computing to enable it to effectively analyze the value of information, thereby helping enterprises better adapt to the ever-changing information environment. Due to the difference in scale and level between enterprises, enterprises should build according to their own different specifications to adapt to the development of enterprise data integration systems. This system should be comprehensive, practical and malleable, and can help enterprises achieve balanced development of different levels. In addition, The data integration system can be used for user information feedback to better understand the main needs of the market.

#### 4.2. Organize the big data technical team within the enterprise

The background of big data brought by enterprises is mainly raw data, which is not only used directly. Therefore, if you want to make good use of these data, at the same time building a big data platform, building a big data platform is a technical team within the enterprise, and the original data is used for identification. Management and interpretation. When recruiting talents, companies may be biased towards recruiting talents. Big data brings together big data and corporate talents, so as to better improve the company's competitive advantage in personnel business.

#### 4.3. Create a big data-oriented corporate culture

The use of big data and new technologies will inevitably cause changes in corporate culture. The use of big data in corporate culture to update traditional culture helps employees fall into the habit of using
data to make decisions, build a learning corporate culture, and enable employees to have basic Data analysis and processing capabilities, through the establishment of a sound incentive mechanism, encourage employees to continue to learn, analyze data while increasing communication between colleagues, forming a shared environment and speeding up corporate development.

5. Conclusion
To a certain extent, the clustering algorithm has been further improved, especially in the further refinement and optimization of the data to be processed in the clustering process. In the past application fields of clustering, such as pattern recognition and mathematical statistics, we have made a certain reference, especially on the basis of the principle of data aggregation in the information entropy algorithm, and the rules of ant picking and dropping. The data to be processed is huge and the result of data processing is not ideal.

References
[1] A, W. W., A, D. W., B, V. P. S., A, Y. W., A, J. W., & C, L. W., et al. (2018). Optimization of rainfall networks using information entropy and temporal variability analysis. Journal of Hydrology, 559, 136-155.
[2] Giagkiozis, I., Purshouse, R. C., & Fleming, P. J. (2014). Generalized decomposition and cross entropy methods for many-objective optimization. Information Sciences, 282, 363-387.
[3] Pourshahabili, S., Nikoo, M. R., Raei, E., & Adamowski, J. F. (2018). An entropy-based approach to fuzzy multi-objective optimization of reservoir water quality monitoring networks considering uncertainties. Water Resources Management, 32(13), 4425-4443.
[4] Zuo, L., Dong, S., Zhu, C., Shu, L., & Han, G. (2018). A cloud resource evaluation model based on entropy optimization and ant colony clustering. The Computer Journal, 58(6), 1254-1266.
[5] Nakagawa, Y., James, R. J. W., Rego, César, & Edirisinghe, C. (2014). Entropy-based optimization of nonlinear separable discrete decision models. Management Science, 60(3), 695-707.
[6] Hennig, P., & Schuler, C. J. (2011). Entropy search for information-efficient global optimization. Journal of Machine Learning Research, 13(9), 1809-1837.