

## Research on Hierarchy Structure Generation Method of Ontology Knowledge Pan-concept in Agriculture

Yang Yang, Chen Kong, Chao Liu and Shao-wen Li\*

Department of Computer Science, Anhui Agriculture University, Shushan, Anhui, Hefei, China

### Abstract

In this paper, a hierarchy building method of an uncertain ontology concept based on cloud transformation is presented. First, extracting a qualitative concept from the database through cloud transformation; second, improving the original concept leaping algorithm to obtain a synthetically formalized expression of the coarser-grained uncertain concept, which makes the algorithm-output concept hierarchy more realistic; Finally, analyzing the tea science data with this method to extract a qualitative concept and build the concept hierarchy. The experimental results show that the qualitative concept can be more accurately extracted and characterized with this method, and the concept based on the cloud model can simultaneously express its randomness and fuzziness, which makes the ontology built based on these concepts to describe the concept model more accurately and helps to objectively build agricultural ontology.

**Keywords:** Agricultural ontology; Uncertainty; Cloud model; Cloud transform; Pan-concept hierarchy

### Introduction

The knowledge in agriculture field is a high-quality production factor, which significantly improves the efficiency of agricultural labor and capital production. The ontology modeling of the knowledge can enhance its sharing and reusability. However, in the previous research of building the agricultural ontology, it was found that the attributes of many objects and descriptive objects were very uncertain, such as the depth of leaf color, the leaf size, the shoot growth situation, the grain yield, the suitable or unsuitable planting crops in some regions, and the size and color of certain pest lesion etc. Such things can be seen everywhere. Now, a more general solution introduces fuzzy logic to the ontology to deal with uncertain information [1-5]. Professional knowledge is uncertain itself, and certain knowledge is a special expression form of uncertain knowledge. Therefore, how to extract the uncertain concept from the domain knowledge, how to build the hierarchy of uncertain concept, and how to form a set of ontology-based uncertain knowledge expression methods are the urgent tasks of ontology research.

Uncertain knowledge system in agriculture field is composed of the cloudy concepts and the pan-concept hierarchy. Cloudy concepts are the foundation of gaining pan-hierarchy. The concept cloud extraction method is an objective imagining description of the distribution of uncertain data in agricultural field, in case of preserving its uncertainty. The method of continuous data discretization, which is acquired by cloud transformation, gives full consideration of the distribution of the actual data and can extract qualitative concepts from the distribution of the continuous data better. Based on the distribution characteristics of some attribute of a frequency, cloud transformation can automatically create several different overlaying cloud particles, which is a kind of method in cloud concept expression [6]. According to different needs of users, the common extraction methods of the concept cloud are reverse cloud and cloud transformation; and after data are pretreated and extracted from the concept cloud with cloud transformation, several concepts of atomic cloud are gained. These concepts are too tiny and disperse, the range of the concept expression is too overlapping, and the concept hierarchy isn't clear, all of which are not good for understanding knowledge. Therefore, it is necessary to leap the concept of atomic cloud to achieve the aim to coarsen the concept.

The leaped concepts express more independently and more widely, while the attaching hierarchy of the concept is improved in the process of merging. After leaping hierarchy by hierarchy to a predetermined hierarchy or forming a unique concept, the pan-hierarchy structure of ontology knowledge pan-concept in agricultural is finally obtained.

Based on above reasons, this article will introduce a method of building uncertain ontology pan-concept tree based on cloud transformation. The problem of expressing uncertain concept in building agricultural ontology is resolved by selecting the characteristic attributes of the representation concept and extracting the comprehensive digital feature based on the cloud model concept. The qualitative concept and conceptual hierarchy extracted with this method are helpful in building the agricultural ontology and reflecting the uncertainty of concept better.

### Cloud model and cloud transformation

Cloud model, proposed by Li De-yi [7] in 1995, has been successfully used in predicting, association rules mining, comprehensive evaluating, intelligent controlling [8-18], and many other fields. Cloud model is a model to solve the uncertainty of concepts in natural language processing. It doubly associates randomness and fuzziness and is a qualitative and quantitative uncertainty transformation model [19], which can effectively deal with the uncertain problems in expressing and analyzing concepts [20]. Based on fuzzy mathematics and random mathematics, the theory gets rid of the traditional mode of the "precision mathematics" on the "uncertainty". It is expressed by three numerical characteristics: expectation  $E_x$ , entropy  $E_n$ , and hyper entropy  $H_e$ .  $E_x$  is the central value of the universe of discourse, and best represents the qualitative concept. Entropy  $E_n$  is the range of qualitative concepts that

\*Corresponding author: Shao-wen Li, Department of Computer Science, Anhui Agriculture University, Shushan, Anhui, Hefei, China, Tel: +86 15956905421; E-mail: [arabain@126.com](mailto:arabain@126.com)

Received June 24, 2017; Accepted July 22, 2017; Published August 02, 2017

Citation: Yang Y, Kong C, Liu C, Li SW (2017) Research on Hierarchy Structure Generation Method of Ontology Knowledge Pan-concept in Agriculture. Adv Robot Autom 6: 166. doi: [10.4172/2168-9695.1000166](https://doi.org/10.4172/2168-9695.1000166)

Copyright: © 2017 Yang Y, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

can be measured, reflecting the fuzziness of qualitative concepts. Super entropy  $He$  is the entropy of the entropy, it associates fuzziness and randomness and represents the uncertainty of the qualitative concept. This provides a good method for the expression of uncertain knowledge in the field, introduces the cloud model into the agricultural ontology modeling, and finally creates a new form of agricultural ontology--cloud ontology, representing the uncertain knowledge.

Cloud transformation [21] is a process of recovering the concept description from the actual data distribution of a certain universe of discourse and is a transformation process from a quantitative description to a qualitative description. Based on an exact quantitative dataset, a concept variable can be considered as a set of atomic concepts:

$$A \{A_1(Ex_1, En_1, He_1), \dots, A_k(Ex_k, En_k, He_k)\}$$

where  $A_i (i=1, 2, \dots, k)$  is the atomic concept represented by the cloud model.

It has been proved that any kinds of probability distribution can be disintegrated into several normal distribution. Because the normal cloud is ubiquitous, the frequency distribution of the data may be treated as the superposition of the expected curves of normal clouds with different magnitudes during the process of transform. Mathematically, cloud transform is:

$$\left\{ \begin{array}{l} g(x) = \sum_{i=1}^n (a_i * f_i(x)) + \varepsilon(x) \\ s.t. \quad 0 < \max \left( \left| g(x) - \sum_{i=1}^n (a_i * f_i(x)) \right| \right) < \varepsilon \end{array} \right.$$

In the expression:  $g(x)$  as the data distribution function;  $f_i(x)$  as the expected function of the cloud model;  $a_i$  as the amplitude coefficient;  $n$  as the number of the superposition clouds;  $\hat{\alpha}$  as the error threshold, used to control the entropy construction and iteration fitting number, and  $\hat{\alpha}(x)$  as the error curve.

Here, cloud transformation aims to extract a more fine-grained concept from the actual data distribution of a certain attribute of a certain coarse-grained concept. High-frequency appearing data values more to the qualitative concept than low-frequency appearing data values. The local maximum point in the data frequency distribution is the convergent center of the data. It can be taken as the center of the concept, i.e. the expectation of the cloud model. The data reflected by this concept are subtracting the corresponding numerical value of the qualitative concept from the original distribution, then searching for the local maximum point, and the local maximum is searched iteratively until the asked preciseness is satisfied. The algorithm of cloud transformation [21] is described in detail in the document.

### A leaping method of concept cloud hierarchy

**The traditional concept leaping method:** A concept leaping algorithm [21] is presented and used to leap atomic concept extracted from the age data distribution curve to generate the corresponding pan-concept tree. Its concept synthesis algorithm is described as follows:

(1) Define function  $MEC'_{ci}(x)$ :

$$MEC'_{ci}(x) = \begin{cases} MEC_{ci}(x), & MEC_{ci}(x) \geq MEC_{ci} \\ 0, & other \end{cases}$$

(2) Calculate the cutting entropy  $En'_i$ :

$$En'_i = \frac{1}{\sqrt{2\pi}} \int MEC'_{ci}(x) dx, i = 1, 2$$

(3) Calculate the mathematical characteristics of the synthesized cloud:

$$\begin{cases} Ex = \frac{Ex_1 En'_1 + Ex_2 En'_2}{En'_1 + 521 En'_2} \\ En = En'_1 + En'_2 \\ He = \frac{He_1 En'_1 + He_2 En'_2}{En'_1 + En'_2} \end{cases}$$

The concept merging process of the above algorithm is to select the center value of the concept constantly, the calculation method is  $|Ex_1 - Ex_2|$ , which merges two closest clouds. However, in the agricultural data set, some closest clouds are not necessarily suitable for merging, and the above algorithm does not provide the merging process of several clouds. In the light of the shortcomings of the existing concept cloud merging algorithm, the above algorithm is presented in this article from the perspective of daily concept understanding, the more duplicate contents the two concepts express, the fitter the two concepts are. Therefore, in this article, the concept of cloud fitting degree  $N$  (referred to as the cloud fitting degree) as a criteria is presented to quantify the two groups of concepts if they are suitable for cloud merging.

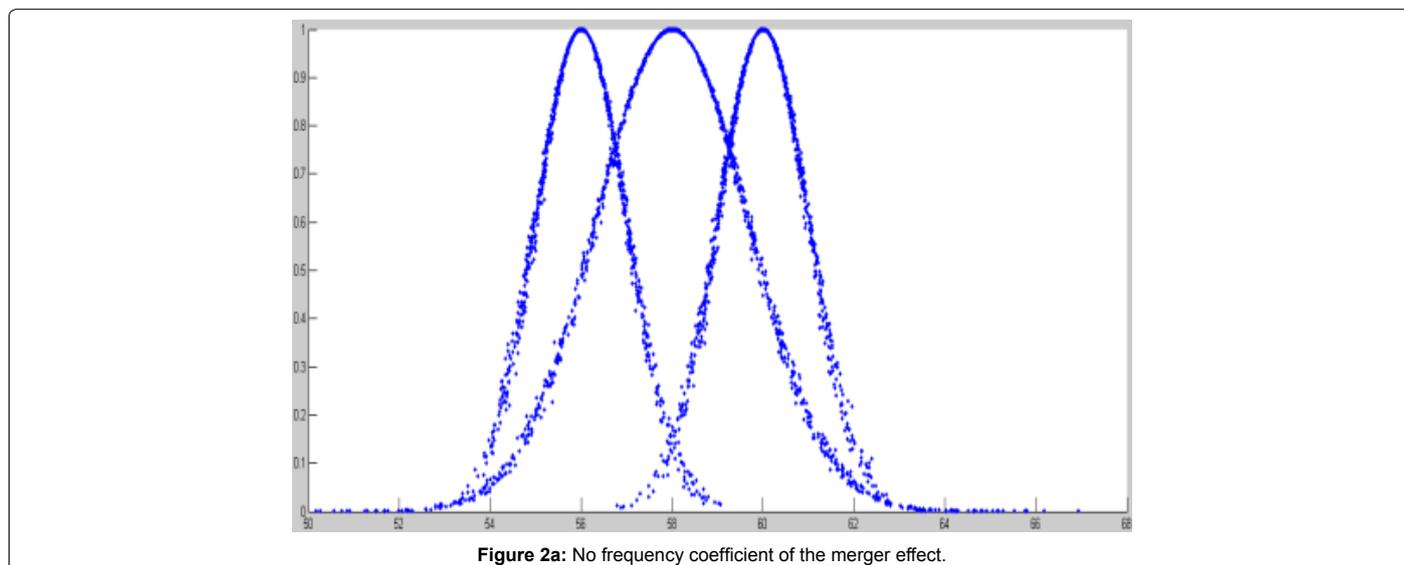
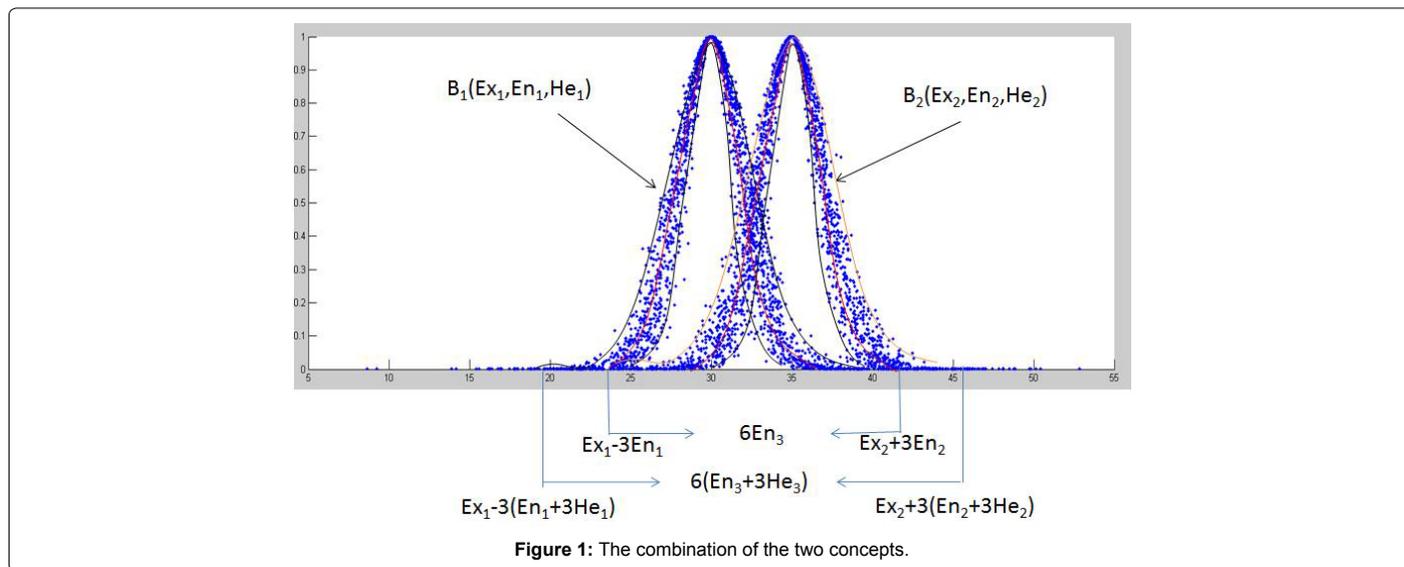
**Concept merging algorithm with considering concept cloud distance:** Suppose there are two clouds  $B_1(Ex_1, En_1, He_1), B_2(Ex_2, En_2, He_2)$  and the concept of adjacent normal cloud on the universe of discourse, and  $Ex_1 < Ex_2$ .

$$N = \frac{(Ex_1 + 3En_1) - (Ex_2 - 3En_2)}{6 \min(En_1, En_2)}$$

The span of the cloud model expresses the acceptable scope of number field of the concept, from "the rules", 99.74% of the data elements in the concept can use  $[Ex-3En, Ex+3En]$  to express [22] the commonly accepted number field scope of two concepts  $(Ex_2+3En_2)-(Ex_1-3En_1)$ , and  $6 \min(En_1, En_2)$  express their smaller universe of discourse value, therefore  $N$  express the proportion of the intersecting part of the acceptable universe of discourse scope of the two concepts to the smaller universe of discourse scope of the concept (Figure 1). The concept with the smaller universe of discourse scope is entirely subordinate to the cloud with the larger universe of discourse scope, the intersecting part of universe of discourse scope is equal to or larger than the one, that is to say, if one concept is a sub-concept of another concept, then it is not suitable for merging, when the intersecting part of the universe of discourse scope of the two concepts is null, it represents there is no intersection in the two concepts, then the concepts are not suitable for merging. Therefore, the two concepts are suitable for concept merging only when the two concepts have different and common intersection.

In the process of actual gaining concept hierarchy of agricultural field, often need to make the concept leaping to a certain level according to needs, therefore, we set the threshold  $\alpha$ , when the  $N$  reaches the threshold  $\alpha$ , stop merging the concepts. In this article, the merging process adopts the concept leaping method based on the rules [23]. The concept resulting from the cloud synthesis should still include the cloud parts with the main contribution, so the expected position of the concept is the midpoint of the scope, that is:

$$\begin{aligned} Ex_3 &= \frac{(Ex_1 - 3En_1) + (Ex_2 + 3En_2)}{2} \\ &= \frac{Ex_1 + Ex_2}{2} + \frac{3(En_2 - En_1)}{2} \end{aligned}$$



The acceptable concept scope after B3 cloud synthesizing should be unchanged, it is also  $6En_3 = (Ex_2 + 3En_2) - (Ex_1 - 3En_1)$ , so it is:

$$En_3 = \frac{Ex_2 - Ex_1}{6} + \frac{En_2 + En_1}{2}$$

and simultaneously it is:

$$6(En_3 + 3He_3) = (Ex_2 + 3(En_2 + 3He_2)) - (Ex_1 - 3(En_1 + 3He_1))$$

$$En_3 + 3He_3 = \frac{Ex_2 - Ex_1}{6} + \frac{En_2 + En_1}{2} + \frac{3(He_1 - He_2)}{2} = En_3 + \frac{3(He_1 + He_2)}{2}$$

So, it is concluded

$$He_3 = \frac{He_1 + He_2}{2}$$

This method gives the reason for the two clouds merging. However, when acquiring the agricultural ontology concept hierarchy, most of the objects processed are real-time collected data sets. In the process of concept merging for this data set, the sample frequency need to be considered, therefore, it can be expressed  $f \times C(Ex, En, He)$  associated with the sample frequency cloud. For example, 100 high plants and 50

medium plants are synthesized to 150 medium and high plants, that is to say,  $100 \times$  Cloud high and  $50 \times$  Cloud medium plants generate new concept,  $150 \times$  Cloud medium and high plant, with the cloud synthesizing method. In this article, from Figure 2a and 2b, it is clearly seen that the new concept acquired with the cloud synthesizing method leaps to a higher hierarchy on the basis of the two original concepts, which is more in line with the accuracy of the concept expression when transforming the quantitative data into the qualitative concepts in agriculture.

If the normal cloud model set which we got after the cloud transformation is the number of cloud models after the cloud transformer, the pan-concept hierarchy is built as follows:

Step 0: Selecting the data in the field for pre-processing, removing the data with larger error to form the domain concept data set.

Step 1: Using the cloud transforming algorithm to extract the concept and get the atomic concept.

Step 2: Calculating the cloud fitting degree among the concept

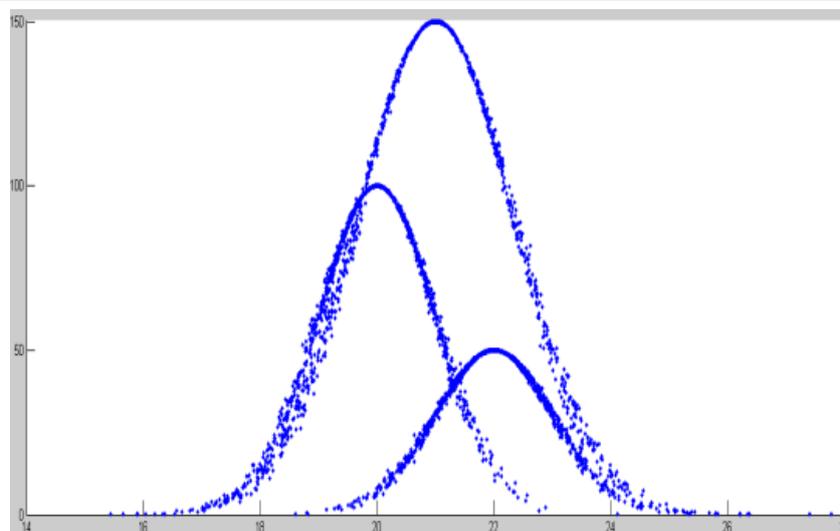


Figure 2b: With the frequency coefficient of the merger effect.

clouds and comparing them, then can do the next step. Otherwise it represents it isn't suitable for concept merging or leaping and exit.

Step 3: Merging the concept clouds according to the cloud fitting degree, from big to small, and using the cloud synthesizing concept to express the merged scope.

Step 4: Repeat steps (3) and (4) until the cloud fitness of the remaining concepts is less than the threshold  $\alpha$ , then stop merging. Each cloud represents the concept of a corresponding number domain, different hierarchy of the concept cloud is combined to the concept hierarchy of cloud concept.

From the above building process, the number of the pan-concept tree which does not necessarily leap to the top hierarchy of the normal cloud model, is one. The stop condition of the pan-concept hierarchy building is that the fitness degree of all normal cloud models in the current hierarchy is less than the given threshold value  $T$ , which indicates that the concept of normal cloud model expression in the current hierarchy is less fuzzy, continuing to merge the normal cloud model will make a serious deviation to the concept expression. Thus it is not appropriate to continue to merge.

## Experiment and analysis

Selecting 4092 meteorological data, collected in the tea plantation in Mount Huang area, China in September, 2016 to analyze, selecting the characteristic attributes for 15 indicators and selecting the characteristic attribute which can mostly express the uncertain ontology model, such as rain fall, sunshine duration, air temperature, air humidity, soil moisture, soil temperature, and evaporation respectively. Through the method of cloud transformation, extracting a concept for its price attribution, leaping the atomic concept through concept leaping method mentioned above, and generating a virtual pan-concept tree to build the concept hierarchy.

**Uncertain concept extraction:** Taking the attribute "soil moisture" as an example to explain the detailed process of the experiment (Figure 3a).

After the soil moisture data cloud-enabled [24], five fuzzy atomic concept cloud maps of soil moisture can be obtained. It can be seen that the expected curve of the cloud model which corresponds to the

qualitative concept obtained by cloud transformation basically reflects the actual data distribution (Figure 3b).

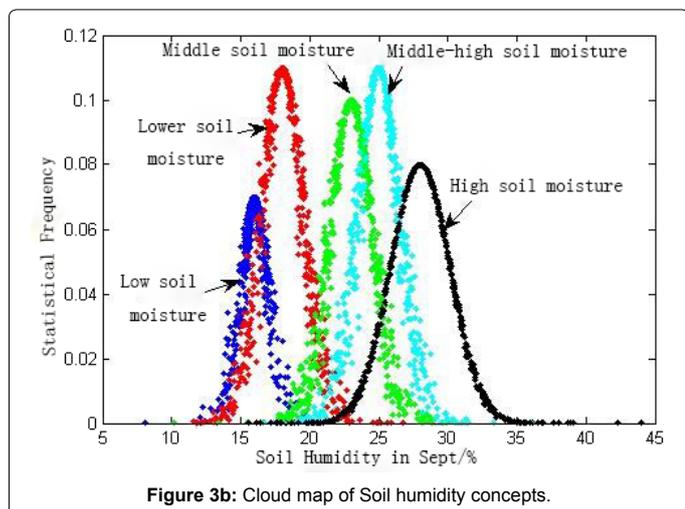
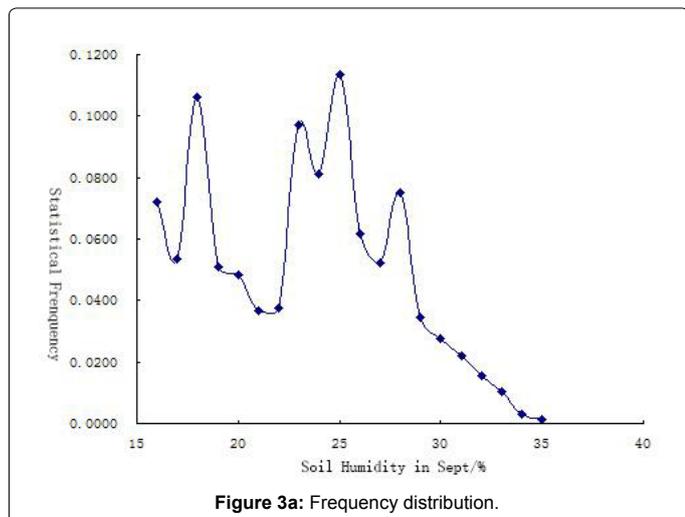
**Concept leaping and concept hierarchy construction:** In order to achieve the concept leaping and get a broader concept in high hierarchy, it is necessary to extract the fine crushing concepts of bottom hierarchy from the cloud transformation to merge. While by the traditional "soft/or" method [23], it needs to select the two cloud model with smaller distance (ie, Ex value gap) to merge, so the cloud  $C_1$  and  $C_2$  merge,  $C_3$  and  $C_4$  merge,  $C_5$  has a bye, and so on, it doesn't stop until the merging leaps to one concept. Merging process diagram is as Figure 3a, with the concept leaping method based on "3En" rule, it still takes the distance between two clouds as the merging base, so the merging process is the same, but because the merge algorithm is different, the three calculated eigenvalues of the high-hierarchy cloud are also different. Merging process diagram is as Figure 3b.

Using the cloud synthesis method given above, it is needed to calculate the cloud fitness between the clouds firstly, as follows: ( $C_1, C_2$ ): 0.996; ( $C_2, C_3$ ): 0.44; ( $C_3, C_4$ ): 0.78; ( $C_4, C_5$ ): 0.88. So the merge order is different from the first two methods,  $C_1$  and  $C_2$ ,  $C_4$  and  $C_5$  are merged firstly,  $C_3$  has a bye, and so on, it doesn't stop until the merging leaps to one concept. Merging process diagram is as Figure 3c.

From the comparison of Figure 3c, it can be seen that when the decision conditions whether they can be merged are different, the three methods begin to make a difference from the process of merging, and the high hierarchy concept eigenvalues obtained after the hierarchy leaping are changed accordingly to Figure 3d is the simulation experiment effect of merging method to achieve hierarchy leaping in this article.

## Analysis of Experiment Results

Because of the superiority of the cloud model itself, it can better express the uncertainty of a concept. In the process of leaping from the initial 5 concepts to the final 1 concept, the concept hierarchy construction is completed at the same time. The digital eigenvalues of a concept and cloud maps can be used to replace the original concept in building an ontology, which is more intuitive.



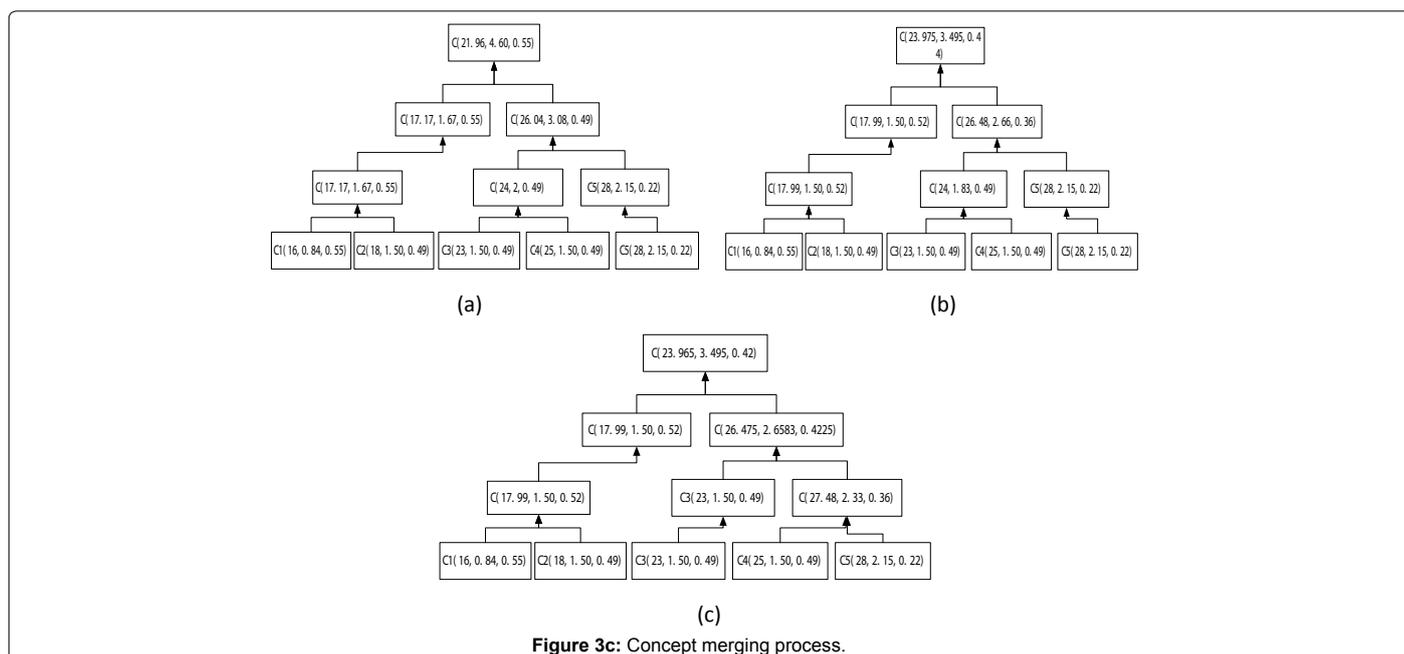
The superiority of the cloud model itself can better express the uncertainty of the concept. Table 1 provides the concept centers, which are gained with the traditional soft/or method and the "3En" cloud synthesis method. They are 21.96 and 23.975 respectively. The concept center 23.965, which is gained with the algorithm presented in this article, can express the overall hierarchy of soil moisture, which is closest to the average value of the actual statistical data, 23.51. Comparing with the original cloud transformation method, it can be seen that the concept centers obtained with the improved cloud transformation method are closer to the ones obtained with the soft/or method and "3En" cloud synthesis method. From Table 1, it can be seen that the concept error rates of the three algorithms are 6.60, 2.0 and 1.9 respectively. The above experimental results show that the method proposed in this article is more accurate than the two former methods, and the concept hierarchy construction can be clearly seen in the merging process.

### Conclusion

Firstly, the existing fuzzy ontology building method ignores the randomness and fuzziness of a concept. In this article, the cloud model is used to describe the qualitative concept, so as to better reflect the association between the randomness and fuzziness of the qualitative concept, extract atomic concepts from the domain database with the concept extraction method, and improve the existing concept merging-based concept leaping algorithm to consider the distance relationship among concepts when merging concepts, so as to generate pan-concept tree based on this method and build the concept hierarchy. Finally, the method is verified by an actual example. It is proved that the method can be used to extract the qualitative concept from the database more accurately, build the concept hierarchy, and provide a reference for

Algorithm	Concept Center	Actual mean	Conceptual Error Rate (%)
soft/or	21.96	23.51	6.60
"3En" cloud synthesis	23.975	23.51	2.0
This article method	23.965	23.51	1.9

**Table 1: Error rate of concepts.**



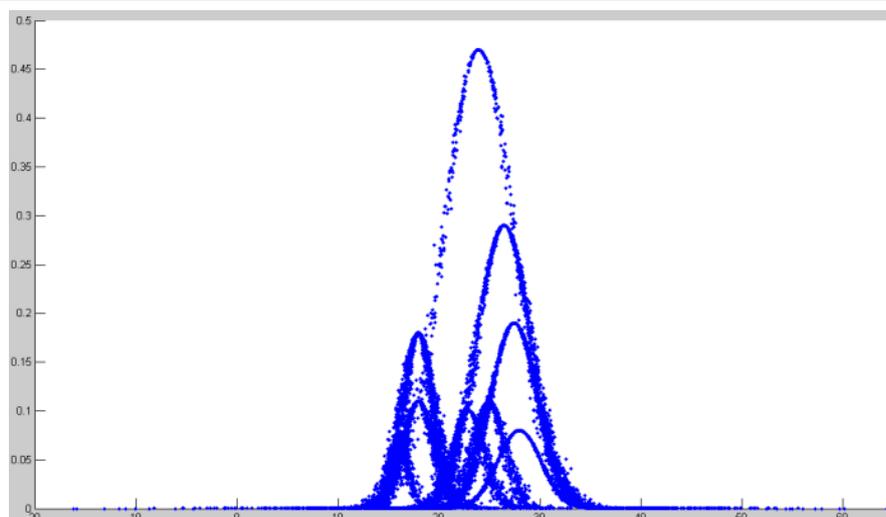


Figure 3d: Tea plantation soil moisture pan-concept hierarchical structure.

using the cloud transformation algorithm in ontology modeling. However, building agricultural ontology pan-concept hierarchy focuses on the uncertain classification relations among agricultural ontology, and the issue of non-classification relation pan concept hierarchy is also worthy of further studying.

#### References

- Formica A (2006) Ontology-based concept similarity in Formal Concept Analysis. *Information Sciences: An International Journal* 176: 2624-2641.
- Lee CS, Jian Z W, Huang LK (2005) A fuzzy ontology and its application to news summarization. *Systems Man & Cybernetics Part B Cybernetics IEEE Transactions on Systems* 35: 859-880.
- Widyantoro DH, Yen JA (2001) Fuzzy Ontology-based Abstract Search Engine and Its User Studies. *IEEE International Conference on Fuzzy Systems*, pp: 1291-1294.
- Wallace M, Avrithis Y (2004) Fuzzy relational knowledge representation and context in the service of semantic information retrieval. *IEEE International Conference on Fuzzy Systems* 3: 1397-1402.
- Maedche A, Staab S (2001) Ontology learning for the semantic web. *Special Issue on the semantic web* 16: 72-79.
- Fu Bin, Li Dao-guo, Wang Mu-hu (2011) Review and Prospect of Cloud Model Research. *Application Research of Computers* 28: 420-426.
- De-yi LI, Changyu LIU, Luying LIU (2005) Study on the universality of the normal cloud model. *Engineering Sciences* 3:18-24.
- Cheng T, Li Z, Li D (2006) An Integrated Cloud Model for Measurement Errors and Fuzziness. *Progress in Spatial Data Handling*. Springer Berlin Heidelberg, pp: 699-718.
- Li SB, Shao CX (2007) Classification of single trial EEG based on cloud model for brain-computer interfaces. *Lecture Notes in Computer Science, Life System Modeling and Simulation*, pp: 335-343.
- Fan J, Li D (1999) Mining Classification Knowledge Based on Cloud Models. *Pacific-Asia Conference on Methodologies for Knowledge Discovery and Data Mining*. Springer-Verlag, pp: 317-326.
- Yang L, Zhang Y, Li H (2008) Decorative pattern design of ceramic based on cloud model and fractal art. *Design, CAID/CD 2008. 9th Computer-Aided Industrial Design and Conceptual Design*. IEEE, pp: 890- 894.
- Cui W, Guan Z, Qin K (2008) A Multi-Scale Image Segmentation Algorithm Based on the Cloud Model. *Proceedings of the 8th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences*, Shanghai, P. R. China, pp: 270-276
- Zhang JC, Hu GY (2003) Application of uncertainty reasoning based on cloud model in time series prediction. *Journal of Zhejiang University Science* 4: 578-583.
- Yang B, Tian YQ, Zhu ZY (2004) Generalization-based discovery of spatial association rules with linguistic cloud models. *Journal of Systems Engineering and Electronics* 15: 728-734.
- Kang HY, Fan XZ (2006) User's Relevance of PIR System Based on Cloud Models. *Journal of Beijing Institute of Technology* 15:181-185.
- Mattioli V, Basili P, Bonafoni S (2009) Analysis and improvements of cloud models for propagation studies. *Radio Science* 44: 605-617.
- Tseng JL (2009) Shape-Sensitive Surface Reconstruction for Low-Resolution Point-Cloud Models. *International Conference on Computational Science and ITS Applications*. IEEE Computer Society, pp: 198-207.
- Hu S, Li D, Liu Y (2007) Mining weights of land evaluation factors based on cloud model and correlation analysis. *Geo-spatial Information Science* 10: 218-222.
- LI De-yi, Meng Hai-jun, SHI Xue-mei (1995) Following Cloud and Subordinate Cloud Generators. *Journal of Computer Research and Development* 32: 16-18.
- Li DY, Liu CY, Gan WY (2009) A New Cognitive Model: Cloud model. *International Journal of Intelligent Systems* 24: 357-375.
- LI De-yi, Liu Chang-yu, Du Wei (2005) *Uncertainty artificial intelligence*. Defense Industry Press.
- CHEN Hao, LI Bing, Liu Chang-yu (2015) An Inverse Cloud Algorithm with Uncertainty. *Small-sized Microcomputer System* 36: 544-549.
- Wang Jin-Zhi, Wang Guo-Yin, XU Chang-Lin (2013) Application of a New Cloud Integrated Method in Color Image Segmentation. *Journal of Intelligent Systems*, pp: 517-525.
- LI De-yi (2000) *Uncertainty in Knowledge Representation*. Chinese Engineering Science 2: 73-79.