**TELKOMNIKA**, Vol. 11, No. 9, September 2013, pp. 5224~5228
ISSN: 2302-4046



# Agricultural Knowledge Grid Construction

**Tan Cuiping*, Zheng Huaiguo, Zhang Junfeng, Sun Sufen, Li Guangda**
Institute of Information on Science and Technology of Agricultural, Beijing Academy of Agricultural and Forestry Sciences, Beijing, P. R. China
*Corresponding author, e-mail: tcpspring@163.com

***Abstract***

*In order to eliminate the ambiguity in semantic understandings during the retrieval of the users, as well as mining the relationship between the concept of agricultural knowledge, the association of knowledge among heterogeneous databases needs to be set up, which enable users to discover useful knowledge clues, and gradually form solution for the ultimate question. Based on the characteristics of agricultural knowledge and the achievements of knowledge grid research, with a combination of traditional agricultural thesauri and ontology technology, the agricultural knowledge grid was constructed with the resource layer, semantic layer and user level. It has been applied for semantic extension on retrieval, knowledge links, and knowledge reasoning diagnosis. It gains some achievements, which provide technical support and experience for the deeply agricultural knowledge services.*

***Keywords****: knowledge grid, agriculture, ontology*



## 1. Introduction

The term Grid of the information technology field sources from Power Grid [1]. The Grid technology is an increasingly powful technology, and it has been successfully used to grid computing [2-3], virtual organizations [4], etc. The applicaiton for knowledge management and sharing produced a new concept of Knowledge Grid.

Fran Berman earlier proposed the concept of Knowledge Grid, Knowledge Grid is an intelligent interconnection environment that enables users or virtual roles to effectively capture, publish, share and manage knowledge resources, and other services for the users and to provide the required knowledge services, support for knowledge innovation, and work together [5].

After that, several authors have given the definition of the Knowledge Grid, such as Lianghong Ding [6], Jing Li [7], and so on. Among that, The definition by H. Zhuge is more clear. The Knowledge Grid is an intelligent and sustainable interconnection environment that enables people and machines to effectively capture, publish, share and manage knowledge resources. It also provides appropriate on-demand services to support scientific research, technological innovation, cooperative teamwork, problem solving, and decision making. It incorporates epistemology and ontology to reflect human cognitive characteristics; exploits social, ecological and economic principles; and adopts techniques and standards developed during work toward the future web [8-10].

Along with the development of agricultural informationization in China, agricultural information resources and services platform construction has made remarkable achievements. A number of national, provincial and municipal agricultural information shared resource center have been established, such as the China Agricultural Science Data Center, Beijing agriculture digital information resource center, Guangdong agricultural information resources sharing platform, Sichuan agricultural science and technology literature information resources sharing platform. Data content covers plantation, breeding, aquaculture, biotechnology, biosafety, food safety, resources and environment, quality standards, agricultural zoning, microbial science, and so on.

It is worth noting that the concept of agricultural knowledge has obvious ambiguity, such as the synonymous relationship (cucumber, cuke), no matter the " cucumber " or " cuke" as search terms, you want to find the same content, but in general, the both could not be combined by the system. In addition, for the same name, different people have different comprehensions,

*Received February 24, 2013; Revised June 8, 2013; Accepted June 20, 2013*

such as "cabbage", can be considered to be " Chinese cabbage " or "cole". If using simply keywords for retrieval, the results are not they want usually. The relationship of the concept of agricultural knowledge is very rich and complex, like the ralationship of similarity, parent-child, host, feeding, breeding, constitution, introduction, with reference to, etc. It is very valuable for users, which is important knowledge clues to solve the problem. But because of the heterogeneity of agricultural sciences database, each scientific database is an "information island", which makes the establishment of the knowledge relationship between the heterogeneous databases very difficult.

In order to solve the ambiguity problem on the retrieval of the users, and to mine useful knowledge relationship, agricultural knowledge grid was constructed based on the characteristics of agricultural knowledge. It can provide knowledge clues, and gradually led the users to the ultimate solution.

## 2. Research Method

In this paper, we established three-level agricultural knowledge grid architecture, and made heterogeneous resources, semantic web and user environment to be an organic whole. Based on traditional agricultural thesaurus, we used ontology technology to build conceptual relationship dictionary, and improved new word discovery mechanism, enriched the existing vocabulary structure and relations. The knowledge grid was applied for semantic extension on retrieval, knowledge links, and knowledge reasoning diagnosis. In the process of building, the job of indexing is so heavy. To that effect, Visual Studio 2005 and Protégé computer application software was used.

### 2.1. Agricultural Knowledge Grid Architecture

The architecture of agricultural knowledge grid is illustrated in Figure 1. The resource layer is used to achieve the physical connection between the agricultural resources database. Heterogeneous database integration engine is charge with the sharing and management of heterogeneous database. Through it, the difference of information resource service platform is shielded.

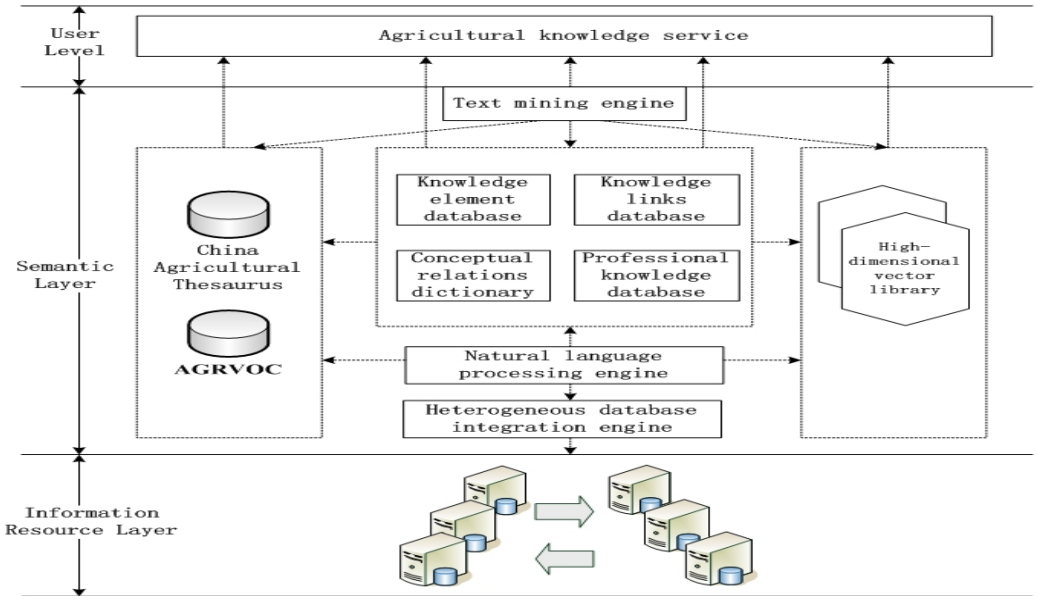


Figure 1. Agricultural Knowledge Grid Architecture

In the semantic layer, several dictionary and knowledge database is found with a combination of traditional agricultural thesauri and ontology technology, which is used for text mining, retrieval, and intelligent reasoning. Based on vocabulary and vector space model

(VSM), encoding identification, segmentation, POS tagging, syntax analysis for farmers' local vernacular character is realized by natural language retrieval engine. Conceptual relationship dictionary reveals the basic relationship of using of, on behalf of, belong, divided with, reference with. Meanwhile, the other relationships of being suitable to, damaging on or hosting in, description, pictures and video are created. Agricultural professional knowledge database for reasoning diagnosis is made of the experts' experience and knowledge. Knowledge links database is used to reveal the reference relationship between heterogeneous databases.

In the user-level, text mining engine is used to mine the text content, and the relationship between texts, knowledge elements, and knowledge links of heterogeneous databases, which provides users the depth of agricultural knowledge services.

### 2.2. New Word Discovery Methods for the Agricultural Knowledge Grid

Because of the long production cycles, agricultural concepts and relationships between concepts are lagging behind in traditional agricultural thesauri. In order to solve it, based on natural language processing and intelligent aggregation technology, this paper extracted keywords from users' questions, search terms, and labels, found network hot words through automatically aggregating, and add the new words and the conceptual relationships to the dictionary.

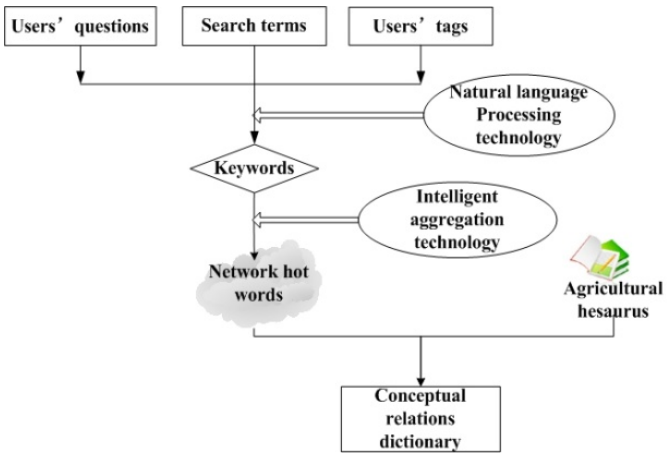


Figure. 2 New Words Discovery Methods

## 3. Results and Analysis

As a result, the contents of this agricultural knowledge grid have covered vegetables, fruit trees, crops, livestock, poultry, and aquiculture, and it has revealed 18 mainly semantic relationships. 200 agricultural heterogeneous database resources are associated in the semantic layer. As an important component of the agricultural knowledge service platform, agricultural knowledge grid has been used to three sectors. Firstly for the extension of search terms, the recall rate and precision rate has been improved, Secondly for the detailed information display, the knowledge links were connected to one net, which can recommend the knowledge clues, and lead the user to form the end solution. Thirdly for the vegetable disease and pest diagnostic system, the reasoning ability was significantly improved.

### 3.1. Search Term Central Extensions

Search term central extensions can effectively eliminate the retrieval ambiguity problems, such as the user input the key words of "green food", the system will recommend the similar words "natural", "organic food", the subdivision words of "Grade A green food "and "Grade AA green food", and the associated words like "banned pesticides", "limited pesticides". It not only extends near-synonyms and synonyms relationship of search terms, but also will guide the user to further explicit retrieval needs, construct a more rational search strategy, and enhance user retrieval efficiency.


Figure 3. Search Term Central Extensions

### 3.2. Knowledge Links between Heterogeneous Databases

Knowledge association between heterogeneous databases is realized by Agricultural knowledge grid, which can effectively improve the user's knowledge discovery capabilities. For example, "breeding units" of the vegetable varieties database is related with agricultural institutions database, and "experts of the field" of agricultural institutions database is related with agricultural experts database. "the application of pesticides" of the vegetable varieties database is related with registered pesticide database, "manufacturer" of registered pesticide database is related with agricultural enterprises database, and so on. Every knowledge point is not isolated, and through any knowledge node, the users can enter into the whole agricultural knowledge network to find the solution they really needs.

### 3.3. Vegetable Disease and Pest Diagnostic System based on the Agricultural Knowledge Grid

Based on professional knowledge database and conceptual relations dictionary, it is reverse reasoning through the different damage characteristics of the five parts of the vegetables, root, stem, leaf, flower, and fruit. Finally system gives the diagnosis by judge and the value of possibility, and the user can confirm it by diseases and pests' pictures.

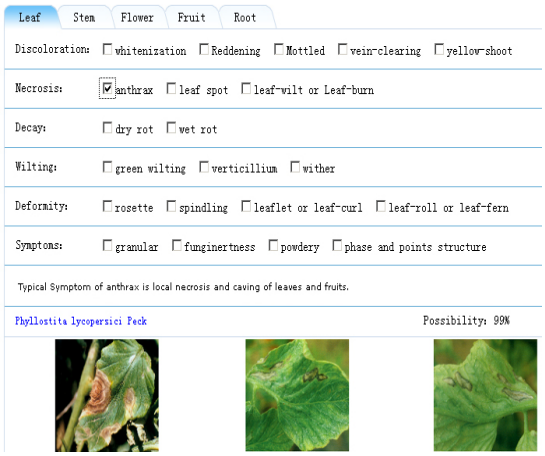


Figure 4. Tomato Disease Diagnostic System

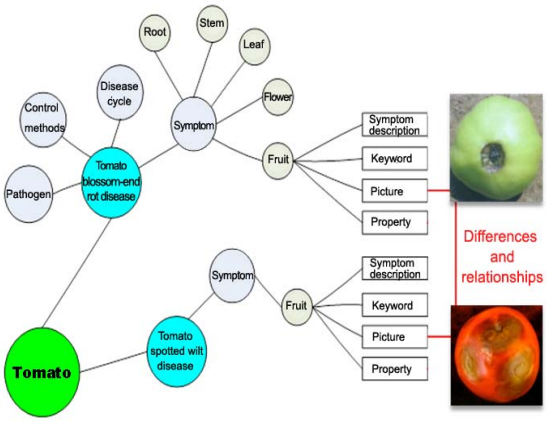


Figure 5. Tomato Disease Diagnostic Mechanism

## 4. Conclusion

Agricultural knowledge grid is one example of the application of grid technology in information processing and knowledge acquisition. It is still in the preliminary stage of development due to the lack of indexing. But with the development of the agricultural informationization in China, agricultural knowledge grid research will be more deeply. In the future, it may be used in cloud computing, pest and disease monitoring, early warning of the biological invasion information monitoring, and so on.

## References

[1] I Foster.*The Grid: A New Infrastructure for 21st Century Science.* Physics Today. 2002; (2): 42-47.
[2] Chetty M, Buyya R. Weaving computational grids: how analogous are they with electrical grids? *Computing in Science and Engineering.* 2002; (4): 61–71.
[3] SeyedElyar Hashemseresht, Ali Asghar Pourhaji Kazem. RDVBT: Resource Distance Vector Binary Tree Algorithm for Resource Discovery in Grid. *IAES International Journal of Artificial Intelligence.* 2012; 1(2): 45-53.
[4] Deng Guang, Liu Qingwang, Li Zengyuan, Zhang Xu, Huang Zhenchun. A Kind of Lidar Application Grid Based on eScience's View. *TELKOMNIKA Indonesian Journal of Electrical Engineering.* 2012; 10(5): 1147-1150 .
[5] Fran Berman. From TeraGrid to Knowledge Grid. *Communications of the ACM.* 2001; 44(11): 27-28.
[6] Lianhong Ding and Xiang Li and Yunpeng Xing. Pushing Scientific Documents by Discovering Interest in Information Flow within E-Science Knowledge Grid. *Grid and Cooperative Computing.* 2005; 498-510.
[7] Jing Li. *A Memory Based Model for Knowledge Organization and Sharing in Knowledge Grid.* IFIP TC8 Publications. 2007; 1295-1299.
[8] H Zhuge. China's E-Science Knowledge Grid Environment. *IEEE Intelligent Systems.* 2004; 19(1): 13-17.
[9] H Zhuge. Communities and Emerging Semantics in Semantic Link Network: Discovery and Learning. *IEEE Transactions on Knowledge and Data Engineering.* 2009; 21(6): 785-799.
[10] H Zhuge. Semantic Linking through Spaces for Cyber-physical-socio Intelligence: A Methodology. *Artificial Intelligence.* 2011; 175: 988-1019.