Proceedings of CECNet 2022
A.J. Tallón-Ballesteros (Ed.)

doi:10.3233/FAIA220516

A Method for Scientific Cultivation Analysis Based on Knowledge Graphs

Aiyan WU[1] , Yongmei ZHANG and Shang YANG
School of Information Science and Technology, North China University of Technology, 100144, China

Abstract. The development of big data and artificial intelligence has improved the intelligence and informatization of scientific planting. A scientific cultivation analysis method based on knowledge graphs is proposed in this paper. First, the logical representation and the ontological representation are combined to realize the access of static cultivation information and dynamic cultivation experience, as well as their representation with graphs. Second, according to the characteristics of plant cultivation information, knowledge extraction is realized via relational computing. A relationship determination method based on the first derivative and a multi-level classification retrieval method based on a tree structure are proposed to extract cultivation experience from the experimental data. Then, multimedia technology is embedded in the RDF framework and implemented, which further realizes the display of decision suggestions after scientific cultivation information analysis. Finally, taking perennial flowers as an example, the realization and application performance of cultivation knowledge graphs are demonstrated.

Keywords. Scientific cultivation, knowledge graph, ontology, cross-media semantic

1. Introduction

With the development of big data and artificial intelligence, the informatization of the planting industry has developed rapidly. Especially in the past five years, planting informatization and sales informatization platforms have been widely used. At present, scientific cultivation driven by large databases has become popular in the field of planting. More and more scholars are interested in how to use big data methods to better diversify cultivation and disease control [1-3]. Therefore, how to control cultivation with advanced information technology has become an important research topic. Many researchers can acquire new knowledge about diverse cultivation and disease control with innovative experiments and analysis of data samples. However, this cultivation knowledge has been in the form of artificial experience. No intelligent analysis system has been formed [4]. Therefore, the universal application of scientific cultivation knowledge is poor. At the 6th International Plant Phenotyping Symposium, big data analysis was specifically discussed for scientific cultivation, promoting the application of big data technology in this area. Among many big data technologies, the knowledge graph has more advantages in mining and displaying the correlations of complex big data

[1]Corresponding Author, E-mail: wuaiyan@ncut.edu.cn . The paper is supported by the R&D Program of the Beijing Municipal Education Commission (KM202110009002).

[5]. Therefore, it is used to analyze and display scientific cultivation knowledge, which is helpful to better understand and apply the knowledge.

The main work is as follows:

(1) In order to integrate static cultivation information and dynamic cultivation experience, this paper proposes a dynamic knowledge representation based on logical reasoning and a structured ontology model.

(2) In order to extract empirical knowledge from cultivation information, a relationship determination algorithm based on the first-order derivative is proposed, and a multi-level classification retrieval method based on a tree structure is also proposed.

(3) In order to realize the graphical display of cultivation knowledge, multimedia technology is embedded in the RDF framework and implemented.

(4) The validity of the above methods is demonstrated by taking perennial flowers as an example.

2. Framework Structure of Knowledge Graphs

A knowledge graph is a special graph structure. It uses an ontology to represent natural objects or abstract concepts and uses relationships to simulate the interaction between ontologies [6]. It is represented via the RDF framework, and the core structure is shown in Fig. 1. Its core storage mode is a triple. For example, there is a relationship r between object h and object t, which can be expressed as (h, r, t). Here, h and t are ontologies in the structural model, and there is a relationship r between them. Many relationships can connect ontologies together. A triple is a subgraph by RDF and finally a complete industry semantic graph is achieved. An industry knowledge graph is a directed graph which contains both semantic information and graph structure information [7-8]. It can be represented by formula (1).

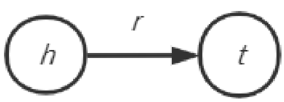

Figure 1. The core structure of the RDF framework.

G is an industry knowledge graph; H is the head ontology set; T is the tail ontology set, and R is the relationship set between ontologies. E is the collection of all ontologies of an industry graph, which contains H and T. In the knowledge graph, any two elements in the triple are known, and the other element can be known by reasoning. Here, the reasoning is actually the correlation effect generated by the various relations of the ontologies. So, the unknown relationships can be obtained by the inference of the known relationships.

$$G = (H, R, T) \qquad \begin{cases} H = \{h_1, h_2, \cdots, h_n\} \\ R = \{r_1, r_2, \cdots, r_m\} \\ T = \{t_1, t_2, \cdots, t_s\} \end{cases} \qquad (1)$$

$$E = H \bigcup T \qquad H \bigcap T \neq \phi$$

3. Knowledge Representation of Scientific Cultivation

A knowledge graph of scientific cultivation should show not only the basic attributes of plants, but also the cultivation experience of plant scholars. Plant attributes are static information which can be easily expressed by graphs. But the experience of scholars is dynamic, and this is difficult to express in existing knowledge structures. In this paper, a logical reasoning-based structured ontology model is proposed to represent the dynamic knowledge.

3.1. Dynamic knowledge representation based on logical reasoning and structured ontology model

Plant growth parameters include phenotypic parameters and physiological parameters. These will change with the change of the external environment. The growth environment can directly affect the physiological metabolism of plants and then change their growth behavior and shape. So, plant growth parameters can change with empirical measures, either increase or decrease. In a scientific cultivation experiment, if one phenotypic parameter and one physiological parameter increase or decrease together, there is a same direction relationship between them. On the contrary, there is a reverse relationship.

In order to express a dynamic knowledge system, the structured method of ontology is used to construct the cultivation knowledge graph, and logical reasoning is used to realize the representation of scientific cultivation experience. The ontology is used to express the entity concepts of a knowledge graph regarding cultivation, and the dynamic relationship between ontologies can be expressed by formulas (2) to (5). Here, f represents empirical measures; e and i are phenotypic parameters and physiological parameters, respectively. They are entities expressed by ontologies. R represents the change of ontology, including increase and decrease, which are represented with Δ and ∇, respectively. Accordingly, the same direction relationship is represented as Δ' and the reverse relationship is represented as ∇'. Triples are used to represent the relationship between ontologies. A dynamic knowledge graph is a series of triple patterns.

$$\exists f_1 \vee f_2 \vee \cdots \vee f_n \rightarrow R(e) \tag{2}$$

$$R = \{\Delta, \nabla\} \tag{3}$$

$$\exists \Delta(e) \vee \exists \Delta(i) \Leftrightarrow (e, \Delta', i) \tag{4}$$

$$\exists \Delta(e) \vee \exists \nabla(i) \Leftrightarrow (e, \nabla', i) \tag{5}$$

3.2. Cultivating knowledge extraction

The knowledge in scientific cultivation mainly includes two parts: the influence of scientific measures on physiological parameters, and the relationship between physiological parameters and phenotypic parameters. It is usually necessary to take measures according to the objectives and record the implementation degree and corresponding parameter values. From a scientific cultivation experiment, it is not difficult to find that the parameter values generated under different measures either increase or decrease, and generally show a linear trend. Therefore, a relational calculation method based on the first derivative is proposed, and the relational patterns are represented by triples, as shown in formulas (6) and (7) below.

$$\exists \frac{f'(x_s)}{f'(x_s')} < 0 \Rightarrow (e_s, \nabla', i_s) \;\; x_s \in [e_{s-1}, e_s], x_s' \in [i_{s-1}, i_s] \tag{6}$$

$$\exists \frac{f'(x_s)}{f'(x_s')} < 0 \Rightarrow (e_s, \nabla', i_s) \;\; x_s \in [e_{s-1}, e_s], x_s' \in [i_{s-1}, i_s] \quad (7)$$

Under the same measure, the value of a phenotypic parameter is expressed by e_s , and the value of a physiological parameter is expressed by i_s . The degree of the measures is gradually increased and new parameter values are constantly generated.

4. Realization and Application of the Scientific Cultivation Knowledge Graph

4.1. Definition of ontology conceptual model

Taking perennial flowers as an example, the conceptual model of a cultivation knowledge graph is given. In general, there are nine phenotypic parameters and nine physiological parameters in perennial flowers, and these physiological parameters are mainly photosynthetic parameters. Control measures can be roughly divided into four categories: light, shading, spray chemical reagent, and water control. Therefore, in order to clearly express the correlation between the internal and external parameters of different varieties of perennial flowers, the ontologies are defined with multiple layers in the cultivation knowledge graph. The ontologies and their attributes are defined as shown in Table 1.

Table 1. Definition of the ontologies conceptual model

ontology	sub ontology	attributes	constraints
plant	non	species	character
		variety	character
phenotype	Plant height	unit	character
		value	float
	Stem diameter	unit	character
		value	float
	Number of blades	unit	character
		value	float
	Leaf yellowing number	unit	character
		value	float
	Leaf area	unit	character
		value	float
	Specific leaf weight	unit	character
		value	float
	Bud length	unit	character
		value	float
	Bud diameter	unit	character
		value	float
	Number of flowers	unit	character
		value	float
physiology	Pn	unit	character
		value	float
	Gs	unit	character
		value	float
	Ls	unit	character
		value	float
	Ci	unit	character
		value	float
	LCP	unit	character
		value	float
	LSP	unit	character
		value	float
	Pmax	unit	character

		value	float
	AQY	unit	character
		value	float
	Rd	unit	character
		value	float
measure	light	unit	character
		value	float
		effect	character
	shade	unit	character
		value	float
		effect	character
	spray chemical reagent	unit	character
		value	float
		effect	character
	control water	unit	character
		value	float
		effect	character

4.2. Implementation of knowledge extraction

Different plants have different growth needs. So, the effect of similar measures on different plants is quite different. Also, the same variety with different measures will get different growth shapes. For example, the shading method can be 75% or 50%, and different implementations can achieve different results. Considering the actual access cost and stability, this research uses a relational database to store cultivation information and triplesl.

In view of the above experimental characteristics, the experimental information is first stored according to the species of perennial flowers, and then stored according to the varieties in the same species. In the same variety database, the experimental information is first stored according to the category of measures, and then stored according to the implementation method in the same category. The cultivation knowledge mined from the information is stored in a separate database, which is still stored in the same multi-level storage structure. Obviously, this is a multi-level storage structure. For multi-level storage, the most efficient retrieval algorithm is based on a tree structure. Therefore, a multi-level classification retrieval method based on a tree structure is proposed to access the experimental data and extract relationships. The algorithm is described as follows:

```
Traversal(species.variety) /* access database by species.variety */
{
Node=measure;/* Locate the root node according to the measures */
Traversal(node) /* access database by measure */
{
   Visit(node.method) /* then access parameter values by method of implementation
*/
   While(not null)
   {
     X=the value of a phenotypic parameter;
     X'=the value of a physiological parameter;
If(f'(x)/f'(x')>0)
       Assign Δ' to triples /* represented by triples */
    else Assign ∇' to triples
    }
```

```
    }
}
```

4.3. Implementation and application of the knowledge graph

The cultivating knowledge graph is dynamic. With the production of new experimental data, it will change. Therefore, the realization of the cultivating knowledge graph mainly includes information updating, knowledge mining, and knowledge retrieval displayed as a graph. The data is updated or accessed according to the above multi-level storage structure. The attribute of the stored information is consistent with the definition of the ontology. With the continuous updating of data, the amount of data will become larger and larger.Therefore, a temporary data table mechanism is set. These new data are first stored in the temporary data table for analysis. After the knowledge is extracted, they are stored in the fixed database. At the same time, the temporary data tables are released.

And the cross-media semantic representation of the knowledge graph is further explored. For example, after analysis, the proposed cultivation method and the main influencing parameters need to be highlighted in the cultivation knowledge graph. Cross-media representation can be achieved by defining related properties. In the realization of the cultivation knowledge graph, the color of the relevant nodes is automatically changed to highlight relevant ontologies. The cross-media semantic representation of knowledge graphs is researched to better display cultivation suggestions that meet the needs of cultivation goals in a unique graphical way. Finally, the Oracle database is used to store experimental information and triples, and Java is used to realize the development of the cultivation knowledge graph.

Taking Marco Polo as an example, the application effect of the cultivation knowledge graph is described. The aim is to have buds about five centimeters long and three flowers per plant. After retrieval and analysis with the cultivation knowledge graph, the recommended cultivation suggestion graph is shown in Fig.2.

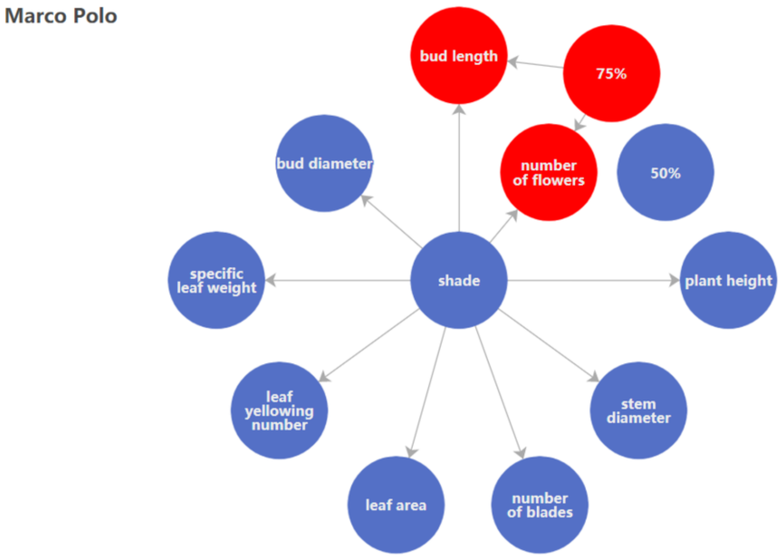

Figure 2. The cultivation analysis graph for Marco Polo.

According to the analysis of the knowledge graph, the 75% shading method can be used to achieve the desired cultivation effect. Fifty Marco Polo plants are experimentally cultivated, of which 25 plants are shaded by 75%. When each plant sprouts to 5 cm long, 25 of them shall be subject to the 75% shading treatment with a black shading net. During the experiment, the management of temperature, humidity, and cultivation are consistent. In the experiment, the bud length and number of flowers are measured, as shown in Fig. 3. As can be seen from this figure, most Marco Polo lilies have flower buds longer than five centimeters and have more than four flowers. Their photosynthetic metabolism can be weakened with the 75% shading method, which reduces their bud length to about five centimeters and number of flowers to three. The expected cultivation results are obtained.

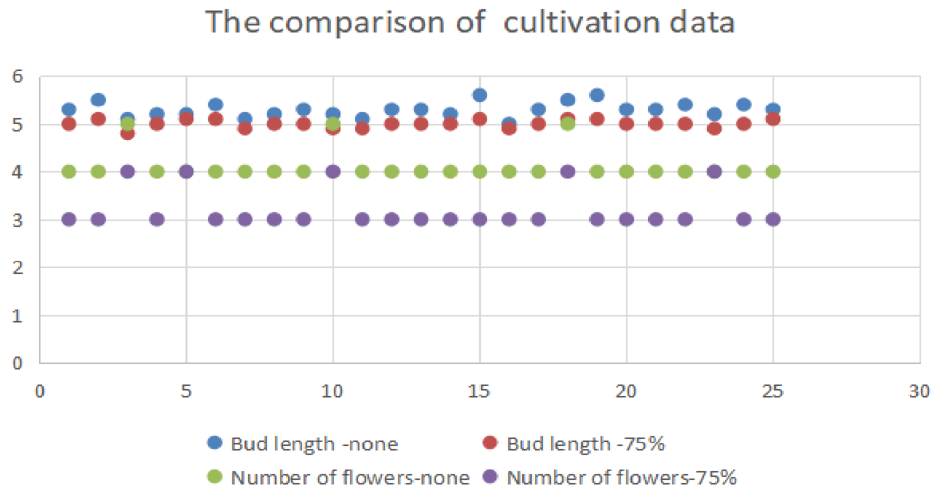

Figure 3. The comparison of cultivation data.

5. Conclusion

This paper presents a scientific cultivation analysis method based on a knowledge graph. It mainly addresses the extraction and display of dynamic experiential knowledge in the process of cultivation analysis. Finally, the scientific cultivation knowledge graph is realized which can provide intelligent analysis and diversified cultivation suggestions for flower cultivation. The effectiveness of this method is proved by the example of lilies cultivated in weak light. The research results will improve the intelligent development of scientific cultivation and the popularization of the cultivation experience.

In the future, the study will be improved from the following aspects. First, it will further discuss how to integrate the knowledge graph and the traditional knowledge representation to better meet the practical application needs of scientific cultivation. Second, the determination of complex relations will be further discussed to better realize intelligent search and accurate cultivation suggestions, and provide a comprehensive reference for the scientific cultivation of flowers with knowledge graph.

References

[1] Shichao Jin, Xiliang Sun, Fangfang Wu. Lidar Sheds New Light on Plant Phenomics for Plant Breeding and Management: Recent Advances and Future Prospects, ISPRS Journal of Photogrammetry and Remote Sensing. 2021, no.171, pp. 202-223.

[2] Patrick Hüther, Niklas Schandry, Katharina Jandrasits. Erratum to: ARADEEPOPSIS, an Automated Workflow for Top-View Plant Phenomics Using Semantic Segmentation of Leaf States, The Plant Cell. 2021, no.32, P. 3674-3688.

[3] Sunil K. Kenchanmane Raju, Addie M. Thompson, James C. Schnable. Advances in Plant Phenomics: From Data and Algorithms to Biological Insights, Applications in Plant Sciences. 2020, vol.8, no.8, pp. 11386.

[4] Luigi Gennaro Izzo, Matthew A. Mickens, Giovanna Aronne. Spectral Effects of Blue and Red Light on Growth, Anatomy, And Physiology of Lettuce, Physiologia Plantarum, 2021, pp. 1-12.

[5] Zhonggui Ma, Runyu Ni, Kaihang Yu. Recent Advances, Key Techniques and Future Challenges of Knowledge Graph, Chinese Journal of Engineering, 2020, vol.42, no.10, pp. 1254-1266.

[6] Aiyan Wu, Yongmei Zhang. Data Analysis Method of Talent Cultivation Based on Relational Graph, Journal of Computers. 2021, vol.32, no.6, pp. 168-175.

[7] Hogan, Aidan, Eva Blomqvist. Knowledge Graphs, ACM Computing Surveys, 2021, vol.54. no.4, pp. 1-37.

[8] Wiharja Kemas, Jeff Z.Pan. Schema Aware Iterative Knowledge Graph Completion, Journal of Web Semantics, 2020, no.65, pp. 100616.