Management of Basic Scientific Research Achievements Based on Knowledge Supernetwork

Wei Sun¹, Chunhui Huo² and Yang Yu*¹,²

¹,²Business School, Liaoning University, Shenyang, 110136, China
Institute of Systems Engineering, Northeastern University, Shenyang, 110819, China
E-mail: lnhappysw@126.com, iwuzetian@126.com, yuyang@ise.neu.edu.cn

Abstract

Basic scientific research achievements management involves multiple knowledge resources. We propose a knowledge supernetwork model to integrate multiple knowledge resources. First, person network, material carrier network, and knowledge element network are established. Subsequently, the knowledge supernetwork is constructed to integrate the above three networks. Second, the characteristics of knowledge supernetwork are investigated from the perspective of the quantitative analysis. Based on the knowledge supernetwork, some methods for solving real problems are developed. Finally, we construct a knowledge supernetwork model using real basic scientific research achievement files. The methods of knowledge representation, basic scientific research proposal grouping, and knowledge classification are proposed to illustrate how to manage basic scientific research achievement from the perspective of knowledge.

Keywords: knowledge supernetwork, knowledge management, basic scientific research achievements, knowledge resources

1. Introduction

Basic research [1-3], as a kind of scientific research, is the research on basic theory and is a knowledge-intensive work. Many countries around the world have increased investment in basic research. Basic scientific research achievement management (shorten as BSRAM) includes storing, statistics, evaluation, and application of the achievements. BSRAM aims to promote the development of basic research and to enhance knowledge innovation. Basic scientific research achievements are mostly expressed [4] in forms of academic paper, technical report, monograph, dissertation, proposal application, progress report, conclusive report, researcher report, etc.

At present, most of BSRAM are based on information management. Achievement attributes (such as institution of achievement owner, achievement type, and published year of achievement) are taken as the management information. For example, statistics and analysis of achievements are performed based on achievement type or published year of achievement.

Basic research achievements contain abundant valuable knowledge and it should be shared broadly. Therefore, BSRAM manages not only the achievement information, but also the knowledge in achievements.

In achievements, there are three types of knowledge resources [5]: achievement knowledge, person knowledge carrier (achievement owner), and material knowledge carrier (achievement files). We use knowledge supernetwork to integrate the three knowledge resources. Subsequently, we investigate the efficient method to manage achievements based on knowledge supernetwork. Supernetwork [6-8] is the “above and beyond” network and composed of multiple networks. Some researches on supernetwork focus on the network optimization using variational inequality. For example, Nagurny
used supernetwork to study the optimization balance of supply chain supernetwork, traffic supernetwork, and finance supernetwork [9-11].

The reminder of this paper is organized as follows. Section 2 states the knowledge supernetwork of BSRAM and the construction. Section 3 analyzes the characteristics of the knowledge supernetwork. Section 4 illustrates how to operate BSRAM based on the knowledge supernetwork. Section 5 gives the conclusions and future problems.

2. Knowledge Supernetwork

2.1. Knowledge Element Network

Knowledge element network is the sub-network and kernel network of the knowledge supernetwork. We use the knowledge element network to represent knowledge of basic scientific research achievements. Knowledge elements are regarded as nodes. Two knowledge elements are considered to be connected if they have the same word. Knowledge element network can be expressed as: \( G_k = (K, E_k) \). \( K = \{ke_1, ke_2, ..., ke_n\} \) is the set of all knowledge elements. \( E_k = \{(r_{ij}, r_{i+1,j}) | r_{ij} = (ke_i, ke_j)\} \) is the set of edges, where \( r_{ij} = 1 \) means there is a relation between \( ke_i \) and \( ke_j \).

Knowledge element [12-14] refers to the knowledge representation unit which can represent an item or a kind of knowledge. Knowledge element can be expressed as \( ke = w_1w_2...w_m \), where \( w_i \) expresses the \( i^{th} \) words in the knowledge element. For example, if \( ke = \{\text{knowledge management}\}, \text{then} w_1 = \text{knowledge} \text{and} w_2 = \text{management} \).

For two different knowledge elements, if they contain the same word, then the relation will be established, as shown in Equation (1) [15].

\[
    r_{ij} = \begin{cases} 
        1, & \text{contain the same word in } ke_i \text{ and } ke_j \\
        0, & \text{donot contain the same word in } ke_i \text{ and } ke_j 
    \end{cases}
\]

2.2. Person Network

Person network [15] is the sub-network of the knowledge supernetwork and expressed as: \( G_p = (P, E_{p,p}) \). \( P = \{p_1, p_2, ..., p_m\} \) is the set of persons. \( E_{p,p} = \{(p_i, p_j) | i, j = 1, 2, ..., m\} \) is the set of edges, where \( (p_i, p_j) \) expresses the relationship between \( p_i \) and \( p_j \).

2.3. Material Carrier Network

Material carrier network [15] is the sub-network of the knowledge supernetwork and expressed as: \( G_m = (M, E_{m,m}) \). \( M = \{m_1, m_2, ..., m_n\} \) is the set of material carriers. \( E_{m,m} = \{(m_i, m_j) | i, j = 1, 2, ..., l\} \) is the set of edges, where \( (m_i, m_j) \) expresses the relation between \( m_i \) and \( m_j \).

2.4. Relations between Person and Material Carrier

There are two types of relationships between person and material carrier: the relationship from person to material carrier and the relationship from material carrier to person.

The relationship from person to material carrier is expressed as follows.

\[
    M(p_i) = \{m_j | m_j \in M, \theta(p_i, m_j) = 1\} \tag{2}
\]

\( M(p_i) \) expresses the set of material carriers owned by \( p_i \). \( \theta(p_i, m_j) = 1 \) expresses the relationship of person \( p_i \) owning material carrier \( m_j \).

The relationship from material carrier to person is expressed as follows.

\[
    P(m_i) = \{p_j | p_j \in P, \phi(m_i, p_j) = 1\} \tag{3}
\]

\( P(m_i) \) expresses the set of persons owning material carrier \( m_i \). \( \phi(m_i, p_j) = 1 \) expresses the relationship of material carrier \( m_i \) owned by person \( p_j \).
2.5. Relations between Material Carrier and Knowledge Element

There are two types of relationships between material carrier and knowledge element: the relationship from material carrier to knowledge element and the relationship from knowledge element to material carrier.

The relationship from material carrier to knowledge element is expressed as follows.

\[ K(m) = \{ke| ke \in K, \theta(m, ke) = 1\} \quad (4) \]

\( K(m) \) expresses the set of knowledge elements contained in material carrier \( m \). \( \theta(m, ke) = 1 \) expresses the relationship of \( m \) owning \( ke \).

The relationship from knowledge element to material carrier is expressed as follows.

\[ M(ke) = \{mj| mj \in M, \varphi(ke, mj) = 1\} \quad (5) \]

\( M(ke) \) expresses the set of material carriers owning knowledge element \( ke \). \( \varphi(ke, mj) = 1 \) expresses the relationship of knowledge element \( ke \) owned by \( mj \).

2.6. Relations between Person and Knowledge Element

There are two types of relationships between person and knowledge element: the relationship from person to knowledge element and the relationship from knowledge element to person.

The relationship from person to knowledge element is expressed as follows.

\[ K(p) = \{ke| ke \in K, \theta(p, ke) = 1\} \quad (6) \]

\( K(p) \) expresses the set of knowledge elements owned by the person \( p \). \( \theta(p, ke) = 1 \) expresses the relationship of \( p \) owning \( ke \). Equation (6) can be obtained by combining Equation (2) with Equation (4).

The relationship from knowledge element to person is expressed as Equation (7).

\[ P(ke) = \{pj| pj \in P, \varphi(ke, pj) = 1\} \quad (7) \]

\( P(ke) \) expresses the set of persons owning knowledge element \( ke \). \( \varphi(ke, pj) = 1 \) expresses the relationship of knowledge element \( ke \) owned by person \( pj \). The number of elements in \( P(ke) \) is expressed as \( |P(ke)| \). Thus \( |P(ke)| \) represents the number of the persons owning \( ke \).

By integrating person network, material carrier network, and knowledge network, the relations between the three networks and knowledge supernetwork can be constructed.

3. Characteristics of Knowledge Supernetwork

3.1. Kernel Network of Knowledge Supernetwork

Knowledge element network is the kernel network of the knowledge supernetwork. The following characteristics can be obtained based on complex network theory [16].

(1) There are many cliques in the knowledge element network. Clique [16] is a subset of network, where points connect with each other. Equation (1) means that the knowledge elements contain the same word with one clique.

(2) Clique can be obtained by clustering. We use the Bron and Kerbosch algorithm [17] to find out all cliques. A clique can be expressed as: \( C(n) = \{ke1, ke2...ke_n\} \), where \( C \) is the clique name and \( n \) is the number of the knowledge elements in the clique.

(3) Domain network. A clique is regarded as a domain, and so the network obtained by clique clustering is called domain network.

3.2. Knowledge of Knowledge Carrier

The knowledge of material knowledge carrier and person knowledge carrier can be obtained by Equation (4) and Equation (6) respectively. Subsequently, the knowledge can be represented using knowledge element network or domain network. If the edges are ignored, then the knowledge can be expressed by knowledge element vector or domain
vector. Based on domain network, the knowledge of a knowledge carrier can be easily expressed by a tree structure.

3.3. Similarity of Knowledge Carrier

The most popular similarity algorithm is the one based on VSM [18-20]. We compute the similarity of two knowledge carriers based on VSM algorithm.

(1) Knowledge carrier similarity based on knowledge elements

All knowledge elements in all knowledge carriers can be represented as:

\[ K = (ke)_{i \times n} = [ke_1, ke_2, \ldots, ke_n] \]  

Then the knowledge elements owned by a knowledge carrier \( M_i \) can be represented by the vector:

\[ M_i = (v_{ij})_{1 \times n} = [v_{i1}, v_{i2}, \ldots, v_{in}] \]  

\( v_{ij} \) is judged if knowledge element \( ke_j \) is in knowledge carrier \( M_i \). If yes, \( v_{ij} \) is 1; otherwise, it is 0.

Subsequently, knowledge carrier \( M_i \) can be represented as:

\[ M_i = (w_{ij})_{1 \times n} = [w_{i1}, w_{i2}, \ldots, w_{in}] \]

\( w_{ij} \) is the weight of knowledge element \( ke_j \) in knowledge carrier \( M_i \). \( w_{ij} \) can be calculated by TF-IDF:

\[ w_{ij} = n/N * \log(M_i/M_k) \]

where \( n \) is the occurrence frequency of \( ke_j \) in knowledge carrier \( M_i \), \( N \) is the total number of knowledge elements in \( M_i \), \( |M| \) is the total number of knowledge carriers, and \( M_k \) is the number of knowledge carriers in which \( ke_j \) occurs. Thus the similarity of two knowledge carriers can be measured by vectors cosine.

\[ Sim(M_i, M_j) = \cos(\vec{m}_i, \vec{m}_j) = \frac{\vec{m}_i \cdot \vec{m}_j}{\|\vec{m}_i\| \|\vec{m}_j\|} \]  

Although, based on the similarity algorithm, two knowledge carriers are thought to be similar according to knowledge elements, they are possibly not always similar in domain. Consequently, we propose knowledge carrier similarity algorithm based on domains.

(2) Knowledge carrier similarity based on domains

1) Domain matrix

After the clique clustering of the knowledge element network, all cliques can be obtained. When a clique is regarded as a domain, then all domains can be represented by the matrix as follows:

\[ D = (s_{ij})_{n \times m} = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1m} \\ s_{21} & s_{22} & \cdots & s_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ s_{n1} & s_{n2} & \cdots & s_{nm} \end{bmatrix} \]  

\( s_{ij} \) expresses the relation between \( ke_i \) and \( D_j \). If \( ke_i \) is in \( D_j \), then \( s_{ij} \) is 1; otherwise, it is 0.

2) Domain vector of knowledge carrier

Based on Equation (9), all knowledge carriers can be represented as the following matrix:

\[ M = (v_{ik})_{n \times n} = \begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1n} \\ v_{21} & v_{22} & \cdots & v_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ v_{n1} & v_{n2} & \cdots & v_{nn} \end{bmatrix} \]  

By multiplying matrix \( M \) by matrix \( D \), all knowledge carriers can be represented as the following matrix:
\[ T = (c_{ij})_{k \times m} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1m} \\ c_{21} & c_{22} & \cdots & c_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ c_{k1} & c_{k2} & \cdots & c_{km} \end{bmatrix} \] (13)

\[ c_{ij} = M \times D_j \times (v_{ik})_{1 \times n} \times (s_{kj})_{n \times 1} = \sum_{k=1}^{n} v_{ik} \times s_{kj} \] . \( c_{ij} \) denotes the number of \( D_j \)'s knowledge elements which are contained in \( M \). After \( c_{ij} \) is divided by the total number of knowledge elements in knowledge carriers, matrix \( T \) can be normalized.

(3) Similarity based on domain vector

According to Equation (13), the similarity of two knowledge carriers based on domain vector can be calculated.

\[ \text{Sim}(M_i, M_j) = \text{Sim}(T_i, T_j) = \cos(\vec{t}_i, \vec{t}_j) = \frac{\vec{t}_i \cdot \vec{t}_j}{\| \vec{t}_i \| \| \vec{t}_j \|} = \frac{\sum_{k=1}^{m} c_{ik} \times c_{jk}}{\sqrt{\sum_{k=1}^{m} c_{ik}^2} \sqrt{\sum_{k=1}^{m} c_{jk}^2}} \] (14)

3.4. Classifying Knowledge

Knowledge can be classified into exclusive knowledge and common knowledge by exclusive degree of knowledge (ked).

\[ \text{ked}(ke_i) = \frac{1}{|P(ke_i)|} \] (15)

\( |P(ke_i)| \) is the number of the persons owning \( ke_i \). If \( |P(ke_i)|=1 \), then \( \text{ked}(ke_i) = 100\% \). This indicates that \( ke_i \) is owned by only one person. So \( ke_i \) must be exclusive knowledge [21]. If \( |P(ke_i)|=n \), then \( \text{ked}(ke_i) = 1/n \ (n \ is \ the \ total \ number \ of \ persons) \). This means that \( ke_i \) is owned by all persons. So \( ke_i \) must be common knowledge.

(1) Exclusive knowledge

All exclusive knowledge can be obtained by the following equation.

\[ E(K) = \{ ke_i | kpd(ke_i) \geq \text{thr}(ek) \} \] (16)

\( \text{thr}(ek) \) is the threshold for determining exclusive knowledge. If \( \text{ked}(ke_i) \geq \text{thr}(ek) \), then \( ke_i \) is considered as exclusive knowledge.

(2) Common knowledge

All common knowledge can be obtained by the following equation.

\[ C(K) = \{ ke_i | kpd(ke_i) \leq \text{thr}(ck) \} \] (17)

\( \text{thr}(ck) \) is the threshold for determining common knowledge. If \( \text{ked}(ke_i) \leq \text{thr}(ck) \), then \( ke_i \) is regarded as common knowledge.

(3) Exclusive domain

For a domain, only when most of knowledge elements are exclusive knowledge, the domain is thought as an exclusive one. Let \( E(D_i) \) be the set of exclusive knowledge elements in domain \( D_i \), then exclusive domain can be determined by domain knowledge exclusive degree (ded).

\[ \text{ded}(D_i) = \frac{|E(D_i)|}{|D_i|} \] (18)

\( |E(D_i)| \) is the number of exclusive knowledge elements in domain \( D_i \) and \( |D_i| \) is the number of knowledge elements in domain \( D_i \).

All exclusive domains can be obtained by the following equation:

\[ E(D) = \{ D_i | \text{ded}(D_i) \geq \text{thr}(ed) \} \] (19)

\( \text{thr}(ed) \) is the threshold for determining exclusive domain. If \( \text{ded}(D_i) \geq \text{thr}(ed) \), then \( D_i \) is considered as an exclusive domain.
(4) Common domain
Similarly, common domain can be determined by domain knowledge common degree 
\((dcd)\).

\[
dcd(D_i) = \frac{|C(D_i)|}{|D|}
\]  
(20)

\(|C(D_i)|\) represents the number of common knowledge elements in domain \(D_i\).

All common domains can be obtained by the following equation:

\[C(D) = \{D_i \mid dcd(D_i) \geq thr(cd)\}\]

(21)

\(thr(cd)\) is the threshold for determining common domain. If \(dcd(D_i) \geq thr(cd)\), then \(D_i\) is thought as a common domain.

4. BSRAM based on Knowledge Supernetwork

4.1. Case Background
We take the BSRAM in a certain country as the case study. Abundant achievements are obtained in the basic scientific research every year, and the effective management of these achievements is very significant. Therefore, the following issues should be solved effectively: 1) how to represent knowledge in BSRAM; 2) how to group proposals based on knowledge in the proposals; and 3) how to classify achievements knowledge.

4.2. Data
We use the achievement files of Management Science as the experimental data. There are 30 persons and 301 achievement files. Table 1 summarizes the data. Columns “Subject code” and “Subject name” states the subject information of funded proposals. The data in column “Proposals” are the number of funded proposals within the subject.

<table>
<thead>
<tr>
<th>Subject code</th>
<th>Subject name</th>
<th>Proposals</th>
<th>Persons</th>
<th>Achievements files</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>Decision-Making Theory</td>
<td>5</td>
<td>5</td>
<td>38</td>
</tr>
<tr>
<td>02</td>
<td>Organization Theory</td>
<td>5</td>
<td>5</td>
<td>49</td>
</tr>
<tr>
<td>03</td>
<td>Industrial Engineering</td>
<td>4</td>
<td>4</td>
<td>46</td>
</tr>
<tr>
<td>04</td>
<td>Management Information System and Decision Support System</td>
<td>8</td>
<td>8</td>
<td>100</td>
</tr>
<tr>
<td>05</td>
<td>Quantitative Economic Analysis</td>
<td>4</td>
<td>4</td>
<td>41</td>
</tr>
<tr>
<td>06</td>
<td>Complexity Science</td>
<td>4</td>
<td>4</td>
<td>27</td>
</tr>
</tbody>
</table>

The part of the knowledge supernetwork built by the above data is shown as Figure 1.
4.3. Knowledge Representation

According to the statement of section 3.2, we can represent the knowledge of a proposal. Based on knowledge element network, domain network, and tree structure, a proposal’s knowledge can be represented as Figs. 2a, 2b and 2c, respectively.
4.4. Proposal Grouping

Based on the subject code, we can group proposals as Figure 3a. However, the result fails to show the similarity between proposal and subject. Based on our domain-based similarity algorithm, we group proposals as Figure 3b. The figure shows the similarity between proposal and subject.

4.5. Achievements Knowledge Classification

By setting \( thr(ek) \) as 100% and \( thr(ck) \) as 1/6, we obtain exclusive knowledge and common knowledge according to Equation (16) and Equation (17) respectively. The knowledge element networks of exclusive knowledge and common knowledge are shown in Figure 4.
By setting $thr(ed)$ as 0.4 and $thr(cd)$ as 0.1, we obtain exclusive domains and common domains respectively, as demonstrated in Figure 5.

To investigate valuable information, the above results are analyzed in detail.

1. **Exclusive knowledge.** The first 6 exclusive knowledge elements are “maximum likelihood estimation”, “organization structure reform”, “intelligent traffic system”, “layered genetic algorithm”, “export credit insurance”, and “evolutionary genetic algorithm”.

2. **Common knowledge.** The first 6 common knowledge elements are listed in Table 2.

**Table 2. The First Six common Knowledge Elements**

<table>
<thead>
<tr>
<th>knowledge element</th>
<th>computer</th>
<th>complexity</th>
<th>certainty</th>
<th>management science</th>
<th>database</th>
<th>systems engineering</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(ke_i)$</td>
<td>38</td>
<td>34</td>
<td>33</td>
<td>32</td>
<td>30</td>
<td>29</td>
</tr>
</tbody>
</table>

Figure 4. Knowledge Classification

Figure 5. Knowledge Domain Classification
③ Exclusive domain. The detailed information about exclusive domains is displayed in Table 3. If exclusive knowledge and exclusive domains are important, then the government should provide sufficient fund to support the researches related to them.

**Table 3. Detailed Information of Exclusive Domains**

<table>
<thead>
<tr>
<th>domain on name</th>
<th>organization</th>
<th>system structure</th>
<th>economy theory</th>
<th>technology information</th>
<th>mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>ded(Di)</td>
<td>0.56</td>
<td>0.52</td>
<td>0.50</td>
<td>0.48</td>
<td>0.45</td>
</tr>
<tr>
<td>E(Di)</td>
<td>85</td>
<td>149</td>
<td>71</td>
<td>122</td>
<td>73</td>
</tr>
<tr>
<td>D_i</td>
<td>151</td>
<td>287</td>
<td>173</td>
<td>252</td>
<td>161</td>
</tr>
</tbody>
</table>

④ Common domain. The detailed information about common domains is illustrated in Table 4. From Table 4, hot researches in different times are revealed.

**Table 4. Detailed Information of Common Domains**

| domain on name | competition application method question science product computation decision |
|----------------|-------------------|-----------------|-------------------|-------------------|----------------|
| dcd(Di)        | 0.26              | 0.26            | 0.25              | 0.24              | 0.23            |
| C(Di)          | 14                | 11              | 30                | 14                | 14              |
| D_i            | 53                | 42              | 118               | 58                | 60              |

From all above results, we can observe that there are more exclusive knowledge and exclusive domains than common knowledge and common domains respectively. The result means that the basic scientific research focuses on new knowledge and new domains.

5. Conclusion and Open Problems

The paper uses the knowledge supernetwork to integrate different knowledge resources in BSRAM. First, we define the knowledge supernetwork and study the main components. Then, the paper analyzes the characteristics of the knowledge supernetwork and proposes some methods based on the knowledge supernetwork. At last, we use the knowledge supernetwork and the proposed methods to manage basic scientific research achievements from the perspective of knowledge.

There are still some problems not resolved yet in this paper, such as other types of knowledge elements, relations among knowledge elements, and the weight of $c_{ij}$ in domain-vector-based similarity algorithm.

Acknowledgments

This research is supported by the National Natural Science Foundation of China (61203182), the National Social Science Foundation of China (13CGL045), and Liaoning Education Science Program (JG15EB061).

References


Authors

Wei Sun, she received Ph.D. degree in management science and engineering, Dalian University of Technology, Dalian, China, in 2008. From December 2008 onwards she has been serving as an associate professor in the Business School, Liaoning University, Shenyang, China. Her major researches include complex network, Management Science, Operations Research, etc.

Chunhui Huo, he received Ph.D. degree in management science, Liaoning University, Shenyang, China, in 2006. From 2007 onwards he has been serving as a faculty member in Business School, Liaoning University. His major researches include Strategic Management, Management Science, etc.

Yang Yu, he received Ph.D. degree in management science and engineering, Dalian University of Technology, Dalian, China, 2009. From 2009 June onwards he has been serving as a faculty member in Information Science and Engineering School, Northeastern University, Shenyang, China. His major researches include Complex
Network, Supernetwork, Operations Research, Management Science, etc.