Construction of Knowledge Base for Industrial Park Planning of Waste Disposal Circular Economy

Wei XUN¹,a and Xu SUN²,b

¹School of Management, Wuhan University of Technology, Wuhan, China
²School of Safety Science and Emergency Management, Wuhan University of Technology, Wuhan, China

¹15927146889@163.com, ᵃᵇ1801071431@qq.com

Keywords: Circular economy; Industry; Waste disposal; Planning; Knowledge base.

Abstract. The problem of municipal solid waste has become a major social problem that the government urgently needs to solve, involving the country’s prosperity and people’s peace life. The development of circular economy industrial park is the focus of ecological civilization construction, and plays an important role in improving the ability of waste governance. This paper constructs a knowledge base for circular economy industrial park planning, including geo-database, main database and running database. Through text semantic analysis and correlation analysis, the key factors of circular economy industrial park planning for waste disposal are refined, so as to guide the planning of circular economy industrial park.

Introduction

In the early 1990s, some scholars led by Professor Ernest Lowe of Cornell University in the United States proposed the concept of Eco-Industry Park (EIP) based on industrial ecology ¹. The eco-industrial park complies with the 3R principle of reduction, reuse, and recycle of the circular economy, and optimizes various dispersed enterprises and industries into industrial ecological chains with the goal of echelon and closed-circuit recycling of material and energy, forming an industrial cluster and ecologicalization. The establishment of a new economic development model of natural resources-products-renewable resources, can transform pollution externalities into internal recycling eco-industrial parks, which might solve the problem of market failure and system failure, and achieve rational allocation of resources and total social benefits maximization ². Mirata³ et al studied the evolutionary effect of enterprise symbiosis on the regional environment in the eco-industrial network. Tan⁴ et al solved network planning problems with integer programming model and heuristic algorithm. In the research process of EIP, most of the researches focus on the analysis of location optimization, material circulation, energy flow, such as the site selection of industrial parks⁵, the construction model of park’s water network⁶, the park’s planning of thermal energy utilization⁷, industrial park management mode with Chinese characteristics, evaluation indexes of operational efficiency, and access rules of the park⁸–¹¹.

With the continuous expansion of China’s urban scale and the rapid increase of urban population, the total amount of municipal solid waste is growing at an alarming rate. More and more cities are trapped in the dilemma of waste siege, which seriously threatens the orderly construction of the country, the healthy development of the city and the normal life of the people. Under the weight of waste invading cities, human beings have a deeper understanding of waste. How to reduce waste generation, how to treat waste harmlessly, and how to further improve the utilization rate of waste resources, have become major problems facing the government. Developing the vein industry has become an important path to solve the dilemma of waste disposal.

The “vein industry” is an international metaphor for the comprehensive utilization of waste, also known as “vein economy” or “fourth industry”. The vein industry transforms the “resource-product-waste” economic model of traditional industry into the closed-loop economic model of resources-products-renewable resources, realizing the recycling of waste into treasure.
How to scientifically and rationally plan the vein industrial park becomes an urgent problem to be solved.

**Development Model for Circular EIP**

Six modes for the construction of circular economy industrial park was concluded: ecological development, sharing development, integration development, recycling, logistics, technological innovation. The construction of the vein industrial park is an important measure to develop a circular economy. The construction of circular economy industrial park is an important stage in the development of the park. In the process of the construction, planners need to make full use of and explore practices and work that are conducive to energy conservation, emission reduction, environmental protection, improvement of resource and energy use efficiency and green development. The development of circular economy industrial park is promoted by the government, the market and the enterprise together. The practice level can be divided into four levels: enterprise, industrial cluster, park and society.

The core of the development of circular economy industrial park is to maximize the efficiency of resource utilization and minimize the cost of resource consumption by the effect of industrial cluster, the facilities and services of the park and continuing technological innovation, so as to achieve the goal of green development. In order to promote the construction and development of circular economy industrial parks, it is necessary to build an industrial park database serving for the planning and operation of the park.

**Design of Reverse Logistics Network Knowledge Base Framework for Vein Industrial Park**

**Vein Industry Park Knowledge Base**

To build a vein industrial park that can maximize the use of resources, all kinds of information should be considered in the construction of the park. The vein industrial park is a complex system problem, and the knowledge base can assist decision makers in industrial layout, network design and park planning. The reverse logistics network knowledge base for the vein industry park mainly includes the basic database, the logistics network library and the operation model of both. The basic framework of the knowledge base is shown in Fig. 1.

![Vein Industry Park Knowledge Base](image)

Figure 1. The Basic Knowledge Base Framework for the Vein Industry Park.

**Vein Industrial Park Database**

The Vein Industry Park database is the base of the entire knowledge base, which stores factual knowledge and provides a data basis for subsequent calculation and reasoning. The database consists of three databases, and its structure shown in Fig. 2.
The geo-database mainly contains geographic information and planning information, and stores the spatial distribution information of the park. The main database of the industrial park stores industrial symbiosis data and information of enterprises entering the park. Industrial symbiosis data refers to the symbiotic relationship of enterprises. The running database of industrial park mainly includes resource indexes and status indexes to monitor the operation of the park in real time.

Reverse Logistics Network Database

The reverse logistics network database mainly includes reverse logistics network design, related planning and goal solving model. Generally, the objectives of reverse logistics network operation are divided into four categories: optimal location, lowest transportation cost, maximum resource utilization, and minimum environmental pollution.

Enterprises in industrial parks should have a symbiotic relationship as much as possible, so that energy, materials and information can be fully utilized in the park, and the balance boundary between supply and demand in the park should be maximized. The problem, relationship and object multi-relationship in marketing opportunity discovery can be described by the related concepts of the hypergraph path, thereby inferring a portfolio of enterprises with symbiotic relationships.

Reverse logistics is a necessary condition for enterprise portfolio. At present, there are few enterprises specializing in reverse logistics. Reverse logistics mainly adopts a shared channel with forward logistics. The service function is single and the operation is difficult. The venous enterprise recycles and reuse waste materials for a long term, and has the business of undertaking reverse logistics processing technology, recycling network, warehousing and transportation, and market operation.

After each part is manufactured into a molded product through assembly, processing, etc, it enters the storage area of the product manufacturer, is distributed to each distributor after the completion of mass production, and is finally used by the user after purchasing. This phase belongs to the forward logistics phase. In the reverse logistics phase, due to quality, use time, use experience and other reasons, the products of users will inevitably be damaged or abandoned. At this time, the products can be repaired or recycled and remanufactured, etc. The wastes will enter into the vein industrial park, be re-manufactured into reusable products through disassembly, primary process and other processes. In the context of multi-objective optimization, solutions that meet actual needs are selected.

Vein Industrial Park Knowledge Base Construction Model

Corpus Preprocessing

Select the literature about the vein industry park planning, and mark the text which belongs to the target domain. Take the complete text as the sample corpus from which to extract the subject vocabulary of the target domain. The preparatory work of relevant corpus mainly includes the following three parts: word segmentation, vocabulary statistics and vocabulary filtering.
Vocabulary Construction

According to the corpus preprocessing results, the domain vocabulary set \( W = \{ w_i | i = 1, 2, 3, \ldots, n \} \) and the domain text set \( D = \{ d_j | j = 1, 2, 3, \ldots, m \} \) are gained, the theme vocabulary’s candidate words are selected using the TF-IDF algorithm.

Salton et al. proposed the TF-IDF algorithm in 1973. The main idea is that if a vocabulary appears frequently in a certain type of article while less frequently found in other types of articles, the vocabulary is considered to have a better category distinguishing ability. In the field vocabulary extraction, the TF-IDF method can be used to quantify the domain distinguishing ability of the vocabulary.

The specific calculation of the TF-IDF algorithm is as follows:

\[
TFIDF_{ij} = TF_{ij} \times IDF_{i} = \frac{f_{ij} \times \log_{2} \frac{N}{n_{i}}}{\max_{k} f_{kj}}.
\]

Where \( f_{ij} \) represents the frequency of the vocabulary \( w_i \) in the text \( d_j \), \( \max_{k} f_{kj} \) represents the maximum value of all vocabulary frequencies in the text \( d_j \), \( n_{i} \) represents the text frequency of the vocabulary \( w_i \) in the text, and \( N \) represents the total number of texts in the sample corpus.

Construction of the Word Co-occurrence Characteristic Matrix

The main idea of the word co-occurrence model is that in some domain corpora, when some words often appear in the same natural language window unit, they are considered to have semantic relevance in the corpus environment. All the semantically related words in the corpus are clustered together, and the vocabulary correlation or correlation strength is used as the weight to obtain the co-occurrence feature matrix of the topic candidates in the field. The calculation model is as follows:

\[
C(w_i, w_j) = \frac{1}{2} \left( \frac{f(w_i, w_j)}{f(w_i)} + \frac{f(w_j, w_i)}{f(w_j)} \right).
\]

\( f(w_i) \) is the number of occurrences of the vocabulary \( w_i \) in the field text set \( D \), and \( f(w_i, w_j) \) is the number of times the vocabulary \( w_i \) and \( w_j \) co-occur in the same natural language window unit in the domain text set \( D \).

Select the Cluster of Words

Suppose \( W_i = \{ w_i | i = 1, 2, 3, \ldots, s, s \leq t \} \) is a connected cluster of words, that is, a subtopic in the target domain theme. The semantic relevance \( O(w_i) \) between a certain word \( w_i \) and the word cluster \( W_i \) in the definition cluster is as follows:

\[
O(w_i) = TFIDF(w_i) + \sum_{j=1}^{s} Q(w_i, w_j),
\]

Through normalization, the semantic contribution of the vocabulary \( w_i \) to the cluster of words \( W_i \) can be described as follows:

\[
O(w_i | W_i) = \frac{O(w_i)}{\sum_{w_{i}' \in W_i} O(w_{i}')},
\]

Thus, the vocabulary with the highest semantic contribution to the cluster of words \( W_i \) is selected as the cluster center \( w_i^* \), namely:

\[
w_i^* = \max_{w_{i} \in W_i} O(w_i | W_i)
\]

The other words in the cluster are the connected words of the central vocabulary of the cluster.

Through above handling, domain thesaurus can be generated. (1) Selecting any initial vocabulary node in the subject cluster \( W \), set to \( w_0 \); (2) Traversing the cluster of words in which the vocabulary \( w_0 \) is located, setting it as \( W0 \in W \), and calculating the cluster center of the cluster of words \( w_0 \), set to \( w_0^* \); (3) Taking the vocabulary \( w_0^* \) as the central vocabulary, and using the other vocabulary in the word cluster \( W0 \) as the connected vocabulary, construct the target domain vocabulary \( w_0^* \), \( O(w_0^*) \); (4) If the subject cluster is \( W \cap W_0 = \emptyset \), it exits; otherwise, select the vocabulary node, set it as \( w_j \neq w_0 \), and repeat steps (2) and (3) above.
Screening Algorithm for the Index System of the Vein Industrial Park

Correlation coefficient method can be used to initially screen the indexes, thus eliminate the indexes with multiple collinearity and avoid different indexes reflecting the same information. The correlation of indexes refers to the consistency of the meaning of different indexes or the consistency of the impact on the evaluation results, resulting in the repeated use of information and reducing the scientific and rational evaluation. The index correlation coefficient is calculated as follows:

\[ r_{ij} = \frac{\sum_{k=1}^{m} (y_{ki} - \text{ave}(y_i))(y_{kj} - \text{ave}(y_j))}{\sqrt{\sum_{k=1}^{m} (y_{ki} - \text{ave}(y_i))^2 \cdot \sum_{k=1}^{m} (y_{kj} - \text{ave}(y_j))^2}} \] (6)

where \( r_{ij} \) is the correlation coefficient between the index i and the index j; \( y_{kj} \) is the score of the i-th index of the kth evaluation object; \( m \) is the total number of samples of all the evaluation objects of a certain index.

According to the construction of the knowledge base of the vein industrial park, it can assist in the screening of planning reference indexes, the determination of the relationship between enterprises supply and demand, the focus of policy concerns, and the main modes of industrial park operation. For example, the circular EIP with domestic waste as the main part, there are many concerns about employment, propaganda, sorting area and franchise, which are different from the technologies, hazardous waste treatment, and qualifications of enterprises that are involved in the recycling industry of electronic and electrical products.

Conclusion

As a vital part of the supply chain, reverse logistics is becoming increasingly valuable. The vein industry is an important direction for the construction of ecological civilization in the future. By constructing the domain vocabulary, professional vocabulary and index system of the vein industry, this paper constructs a knowledge base for the planning and operation of the auxiliary vein industrial park, and proposes a knowledge base framework with knowledge reasoning for the reverse logistics network of the vein industrial park, which has important practical significance for optimizing the development pattern of vein industry parks.

Acknowledgement

This work is partially supported by “the Fundamental Research Funds for the Central Universities” (WUT: 195203001)

References


