[Array 15 (2022) 100197](https://doi.org/10.1016/j.array.2022.100197)

Contents lists available at [ScienceDirect](http://www.sciencedirect.com/science/journal/25900056)

Array

journal homepage: [www.sciencedirect.com/journal/array](https://www.sciencedirect.com/journal/array)

Research on multi-factory combination optimization based on DOSTAR☆

Sen Chen [a,](#_bookmark0)[\*, Jian Wang](#_bookmark2) a[, Manting Yan](#_bookmark0) a[, Chuntao Yang](#_bookmark0) b[, Huihui Han](#_bookmark1) a

a *Computer Integrated Manufacturing System Research Center, College of Electronics and Information Engineering, Tongji University, No.4800 Cao’an Road, Shanghai, 201804, China*

b *Global R&D Platform, Qingdao Economic & Technology Development Zone Haier Water-Heater Co., Ltd., Qingdao, 266101, China*

A R T I C L E I N F O

*Keywords:*

Combinatorial optimization Knowledge discovery Reinforcement learning AHP

Domain ontology Six-tuple

Ternary data fusion CPS

A B S T R A C T

With the development of industrial big data, it has become an important research direction to use combinatorial optimization to coordinate multi-objective problems in complex manufacturing scenarios with multiple factories. At present, most of the multi-objective problems are decomposed into single-objective solutions. However, it is difficult to resolve the contradiction between multiple goals. There are many participants in multi-objective problems and complex data types, so there is no suitable research method at present. Based on big data, this paper integrates various aspects of supply chain management of multiple factories, and proposes a DOSTAR combined model. On the one hand, it conducts knowledge discovery based on the fusion of human-cyber- physical ternary data, on the other hand, it conducts multi-objective optimization through knowledge struc- ture. Among them, the most important thing is to establish the six-tuple as the basic model. Then the space weight, time weight and decision weight are obtained through the weight sub-model. Finally, the improved reinforcement learning algorithm is used to extract relevant new knowledge and complete multi-objective co- ordination. This article takes the supply chain management of Haier water heaters as an example, using the above-mentioned combined model, and the experimental results show that the purpose of improving perfor- mance has been achieved.

# Credit author statement

Sen Chen: mainly responsible for modeling and algorithm debugging of combinatorial optimization. Jian Wang: PhD supervisor, generally guiding the article. Manting Yan: mainly responsible for scene research, data collection and research of relevant literature. Chuntao Yang: mainly involved in scene research and sorting out business data. Huihui Han: participated in relevant literature research, data sorting and al- gorithm debugging.

# Introduction

The coordination of multiple factories has been a hot topic in recent years. With the development of new technologies such as artificial in- telligence, the rapid changes in the market environment and the diver- sification of user needs. The multi-factory manufacturing environment becomes more and more complex, and relevant intelligent decision- making requires more and more data. Among them, the importance of

human as data sources has become increasingly obvious. The fusion of human-machine-physical ternary data [[1](#_bookmark15)] provides a solid foundation for knowledge discovery. Domain ontology is used to associate big data, then extract knowledge from it, and build a corresponding combination model. Finally, correct decisions through knowledge management and optimization could be made easily.

Different decision preferences and different levels of data belong to different goals. There are often conflicts between multiple goals. This paper constructs a DOSTAR model for combinatorial optimization (CO) to coordinate the benefits between multiple goals. In the DOSTAR model proposed in this paper, “DO” stands for domain ontology, “S" stands for

spatial data, “T" stands for time data, “A" stands for decision data based

on Analytic Hierarchy Process (AHP), and “R" stands for reinforcement

learning (RL) algorithm. Ternary data fusion requires massive amounts of data as a basis. When solving some specific problems, part of the data is missing or the problem to be solved is too complicated. Therefore, in the calculation, these specific data are abstracted as weights, which can greatly optimize the calculation process and ensure the accuracy of the

☆ Jian Wang is the professor of Computer Integrated Manufacturing System Research Center. Sen Chen, Huihui Han and Manting Yan are the doctor students of CIMS Research Center. Chuntao Yang is working in Haier Water-Heater Co., Ltd.

\* Corresponding author.

*E-mail address:* [1910067@tongji.edu.cn](mailto:1910067@tongji.edu.cn) (S. Chen).

<https://doi.org/10.1016/j.array.2022.100197>

Received 27 February 2022; Received in revised form 26 May 2022; Accepted 29 May 2022

Available online 7 June 2022

2590-0056/© 2022 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license ([http://creativecommons.org/licenses/by-](http://creativecommons.org/licenses/by-nc-nd/4.0/) [nc-nd/4.0/](http://creativecommons.org/licenses/by-nc-nd/4.0/)).

results. As one of the classic methods in the field of systems engineering, AHP is still widely used in recent years. It is because AHP will produce a series of relatively accurate decision weights in complex supply chain [[2](#_bookmark16)]. In addition, when calculating spatial and temporal data, this article imitates the human mindset and converts the specific values of spatio- temporal data into relative weight values. Therefore, spatiotemporal data can be easily incorporated into the final decision-making basis. These weighted data avoid the troubles caused by overly complex spe- cific data for decision-making. It should be explained that the temporal data represents the time point at which the process business occurs, but it is mainly used for the sequence of events in this study. Spatial data represents the business department, that is, the responsible subject of the event, to reflect the flow of business processes between different departments. The line chart with spatiotemporal data is used to visualize multiple event processes and facilitate the comparison before and after optimization.

Knowledge is the foundation of combinatorial optimization. The process of knowledge discovery is consistent with the combinatorial optimization model. There are many methods of knowledge discovery at present, among which the main keywords are data mining. In literature [[3](#_bookmark17)], a framework with multiple modules is proposed for knowledge discovery in the processing process. This article also uses a multi-module framework to study knowledge discovery. The premise of accurate reasoning is that the data meets certain specifications, which requires the form of knowledge expression to be more rigorous, such as extending from triples to six-tuples. Reinforcement learning is an intelligent al- gorithm that is more suitable for this scenario. This paper proposes an improved reinforcement learning algorithm to calculate and derive the associated knowledge in the complex manufacturing environment. It provides an associative knowledge discovery method that integrates multi-source spatiotemporal data and multiple sub-models.

In summary, based on the fusion of human-machine-physical ternary data, this paper proposes the DOSTAR method for combinatorial opti- mization. The purpose of this method is to model practical problems in complex manufacturing networks. The following content will focus on these points.

# Related work

The purpose of this study is to improve the efficiency of quality traceability of water heaters. Business scenarios have the characteristics of multiple data sources, multiple decision makers, multiple spatial and temporal constraints, and strong correlation between data. Therefore, data fusion, decision weights based on AHP, spatiotemporal data, ontology and domain knowledge need to be studied separately. There- fore, relevant research also focuses on these topics.

* 1. *Combinatorial optimization*

There are many studies on combinatorial optimization. Generally, the multi-objective optimization problem is transformed into sub- objectives of multi-level or multi-stage, and solved relatively sepa- rately. Then the subsets are combined in a certain way to achieve the goal of optimization. In order to better support the proposed combina- tion model, this paper will analyze the literatures of other researchers from the aspects of group decision-making, knowledge discovery, rein- forcement learning, graph computing, production scheduling, and service-oriented manufacturing.

The literature proposes a combinatorial optimization model for group decision-making [[4](#_bookmark18)]. Each expert is assigned a weighting coeffi- cient, which makes it easy to adjust the differences between the experts’

knowledge and experience. In the literature [[5](#_bookmark19)], non-dominated sorting

genetic algorithm (NSGA-II) is used to realize multi-objective combi- natorial optimization in MATLAB. Literature [[6](#_bookmark20)] and literature [[7](#_bookmark21)] are both researches that combine combinatorial optimization and knowl- edge discovery. Literature [[6](#_bookmark20)] proposes a multi-level combinatorial

optimization model for different decision-making preferences for pro- duction and distribution plans. Literature [[7](#_bookmark21)] first discovers knowledge from the solution, and then integrates the knowledge into the strategy. In addition, the combination of combinatorial optimization and graph computing is also a hot spot. Literature [[8](#_bookmark22)] explores the combinatorial optimization of graph, which uses a machine learning (reinforcement learning) model. In the literature [[9](#_bookmark23)], a multi-center variable-scale search algorithm is proposed to solve single-objective and multi-objective combinatorial optimization problems. Literature [[10](#_bookmark24)] has developed a set of software systems for exploring multiple combi- natorial optimization problems in complex networks. In Refs. [[11](#_bookmark25),[12](#_bookmark26)], the process of solving multi-objective production planning problems is hierarchical. Literature [[13](#_bookmark27)] integrates combination optimization and collaborative filtering in a complex service network to improve service efficiency.

The above-mentioned studies are all outstanding and close to the scope of this article. The key factors of the studies are: multiple evalu- ation criteria, multiple layers of constraints, different parameters. However, most of them either did not introduce the relationship be- tween the key factors into the model, or the model was not compre- hensive enough.

* 1. *Ternary data fusion*

Ternary data fusion is an upgraded version of Cyber-Physical Sys- tems (CPS) and human-in-the-loop. It is an important part of the development of combinatorial optimization problems. The relevant research will be listed below.

CPS and Human-in-the-loop are well-known concept of data fusion. With the development of technology, they gradually integrate more human knowledge into traditional data fusion.

The key to CPS is to realize the fusion of data of cyberspace and physical system through human-computer interaction, to realize the increasing fusion of manufacturing process and autonomous decision- making of manufacturing system. The literature [[14](#_bookmark28)] also applies CPS to product design with increasing complexity, from single-discipline products to mechatronic systems to network physics systems, and in- tegrates cross-domain and cross-layer data, interdisciplinary knowledge and new product design processes through CPS.

Human-in-the-loop simulates Human factors and integrates the data obtained with cyberspace data and Internet of Things data. Although there is no clear fusion of human-cyber-physical data, it emphasizes to quantify the human factor and integrate it into of cyber-physical sys- tems. In order to avoid human error and simplify management, self- managed CPS was proposed in the literature [[15](#_bookmark29)]. The human factors are further simulated in the mixed environment of machine and mate- rial. The literature [[16](#_bookmark30)] proposes an architecture for seamless integra- tion of factory workers in an industrial network physical production environment, using semantic fusion data, and real-time analysis of data for anomaly detection. The literature [[17](#_bookmark31)] put forward that in a manufacturing environment, human beings can supervise and adjust Settings to become the source of knowledge and ability, diagnose situ- ations, make decisions and other activities that affect manufacturing performance, providing additional degrees of freedom for the CPS sys- tem as a whole.

The related data source from human-cyber-physical space includes: social network, Internet and IoT [[1](#_bookmark15)]. CPS and Human-in-the-loop are the predecessor of ternary data fusion. The fusion of ternary data is the basis of associative knowledge discovery in complex networks.

* 1. *AHP*

AHP can quantify human decisions; that is, it has the potential to integrate human data with other data. After decades of research and development, AHP has evolved into a variety of AHP methods, including fuzzy AHP, grey AHP, extended FUZZY AHP and some improved or

compound methods. In literature [[18](#_bookmark32)], an advanced supply chain risk assessment model based on order of magnitude AHP (OM-AHP) was developed to compare the tangible and intangible factors that affect supply chain risk. An illustrative example is given to demonstrate the effectiveness of this assessment model. The evaluation method of machining process scheme based on AHP-GREY correlation analysis is proposed in the literature [[19](#_bookmark33)]. Analytic hierarchy process (AHP) is used to analyze the factors that affect the quality of the machining process plan, and the correlation degree is calculated by correlation coefficient and combination weight. Finally, the quality of the process plan is determined according to the correlation degree of the plan.

* 1. *Time and space governance*

As mentioned earlier, the spatiotemporal data governance studied in this article is mainly to convert specific spatiotemporal data into weight values. At present, the mainstream time alignment methods mainly include interpolation extrapolation, least square method, Taylor expansion method, etc. In the aspect of space governance, the origin of coordinates is not unified, and the common methods include Kalman filtering [[20](#_bookmark34)] and least square method [[21](#_bookmark35)]. In addition, there are sys- tematic errors for different descriptions of the same object. Since the benchmark of each description object is different, the results may also have errors. The commonly used methods include least squares, maximum likelihood and so on. In addition, spatiotemporal data governance is inseparable from data mining. A spatiotemporal data mining method based on ontology semantics is proposed in the literature [[22](#_bookmark36)]. Through the spatial data analysis method based on event-event and event-place, the information is mined from two aspects of space and time.

* 1. *Knowledge discovery*

Knowledge discovery is the process and method of extracting knowledge from massive amounts of big data. Within the scope of this article, knowledge discovery is closely related to three concepts: data mining, knowledge representation and reinforcement learning.

1. Data mining As mentioned earlier, ternary data fusion provides a solid foundation

for data mining and knowledge discovery. In literature [[23](#_bookmark37)], many data

mining methods have been used to extract knowledge from solutions generated during multi-objective optimization. These methods are (i) sequential pattern mining, (ii) clustering-based classification trees, (iii) hybrid learning, and (iv) flexible pattern mining. Each method uses a unique learning strategy to generate explicit knowledge in the form of patterns, decision rules and unsupervised rules. In literature [[24](#_bookmark38)], data mining and knowledge discovery are carried out together in order to solve complex problems in intelligent production.

1. Knowledge representation The representation forms of knowledge include Resource Description

Framework (RDF), ontology, and knowledge graphs. Most of their data

forms are triples or variants of triples. In literature [[25](#_bookmark39)], it is proposed a multi-agent algorithm able to automatically discover relevant regular- ities (knowledge) in a given dataset. Each agent operates independently by performing a Markovian random walk on a weighted graph repre- sentation. In literature [[26](#_bookmark40)], it is proposed a principled knowledge-based model in the form of a computational ontology. The literature [[27](#_bookmark41)] proposes a knowledge discovery method based on knowledge graph, which integrates heterogeneous data by introducing knowledge graph.

1. Reinforcement learning

There are many intelligent algorithms that can be used for knowl- edge discovery. This research believes that reinforcement learning is more suitable for the discovery of related knowledge. Reinforcement learning is one of the paradigms and methodologies of machine learning, which is used to describe and solve problems in which an agent interacts with the environment to maximize returns or achieve specific goals through learning strategies.

In recent years, reinforcement learning has been used to find the path in the knowledge map [[28](#_bookmark42)], as well as entity search and relationship search to construct ontology species [[29](#_bookmark43)]. Methods based on reinforce- ment learning and semantic fusion selection are proposed in the litera- ture [[30](#_bookmark44)] to give Suggestions for decision making. Reinforcement learning is used in the literature [[31](#_bookmark45)] to predict the flow of urban spatial and temporal data. The literature [[32](#_bookmark46)] studies the related problems of time series data in the IoTs and uses reinforcement learning to solve the problem of mutual information minimization of historical dependence. All these indicate that reinforcement learning has been gradually used in ontology correlation calculation, but due to the lack of in-depth research, the current reinforcement learning has not made significant progress in associative knowledge discovery.

Compared with other studies, the advantages of this article are: semantic-based data fusion, which can integrate a wider range of data sources, data types and values; six-tuple-based semantic model provides a data foundation for AHP and reinforcement learning sub-models; Manage the spatiotemporal data in the six-tuple, and align the multi- source data at the spatiotemporal attribute level. These advantages will be gradually introduced in detail later. Since this research involves multiple fields, in order to better explain the innovation and practicality of this article, the related work is summarized in [Table 1](#_bookmark3):

# DOSTAR model

The quality traceability of water heaters is complex, and work effi- ciency needs to be significantly improved. First, it involves multiple responsible parties such as users, after-sales outlets, retailers, and manufacturers; it is difficult for multiple responsible parties to coordi- nate efficiently. Secondly, the data type, data format and value of each responsible party are different; data fusion is more difficult. Third, there are also stakeholders with different goals within the manufacturer who is the most responsible party; the decision-making weight of stake- holders will seriously affect the outcome of the decision. Fourth, the key factors affecting work efficiency should be assigned to multiple sub- models for research, and the correlative knowledge among them should be found to effectively improve overall work performance.

* 1. *Overall framework*

Overall, the research span of this article is very large. The first step is to collect the human-cyber-physical ternary data in the complex manufacturing environment according to the characteristics of the complex network of multiple factories, multiple sales companies, and multiple after-sales service outlets. The second step is based on the integration of ternary data to establish the domain ontology of the complex manufacturing environment. The third step is parallel to the second step. The weight sub-model converts the collected decision basis and results into decision weights, and transforms spatiotemporal data

**Table 1**

Related research statistics table.

Research areas Corresponding literature

Combinatorial Optimization [[4–13](#_bookmark18)]

Ternary data fusion [[1](#_bookmark15),[14](#_bookmark28),[16](#_bookmark30),[17](#_bookmark31)]

AHP [[18](#_bookmark32),[19](#_bookmark33)]

Time and Space Governance [[20–22](#_bookmark34)]

Knowledge Discovery [[23–32](#_bookmark37)]

into time and space weight values; and these weight values are stored in the form of the adjacency matrix. The fourth step is to form the six-tuple, which is to integrate the results of the second and third steps to form a six-tuple data set. The fifth step is the improved reinforcement learning algorithm, which converts the six-tuple into a weighted graph; then the weighted graph is chess boarded; therefore, the reinforcement learning algorithm can run smoothly. The sixth step is the result calculated from the fifth step. This result is a subgraph of the weighted graph of the previous six-tuple. At the same time, it is also a streamlined solution to a specific problem; it shows that associative knowledge is discovered. The above is shown in [Fig. 1](#_bookmark4).

* 1. *Domain ontology sub-model*

The domain ontology model in this article is mainly established based on factors such as domain knowledge, expert experience, and data relationships. There are many places to study in complex manufacturing environment. This article is mainly based on the analysis of the actual situation of the complex manufacturing environment of Haier water heaters.

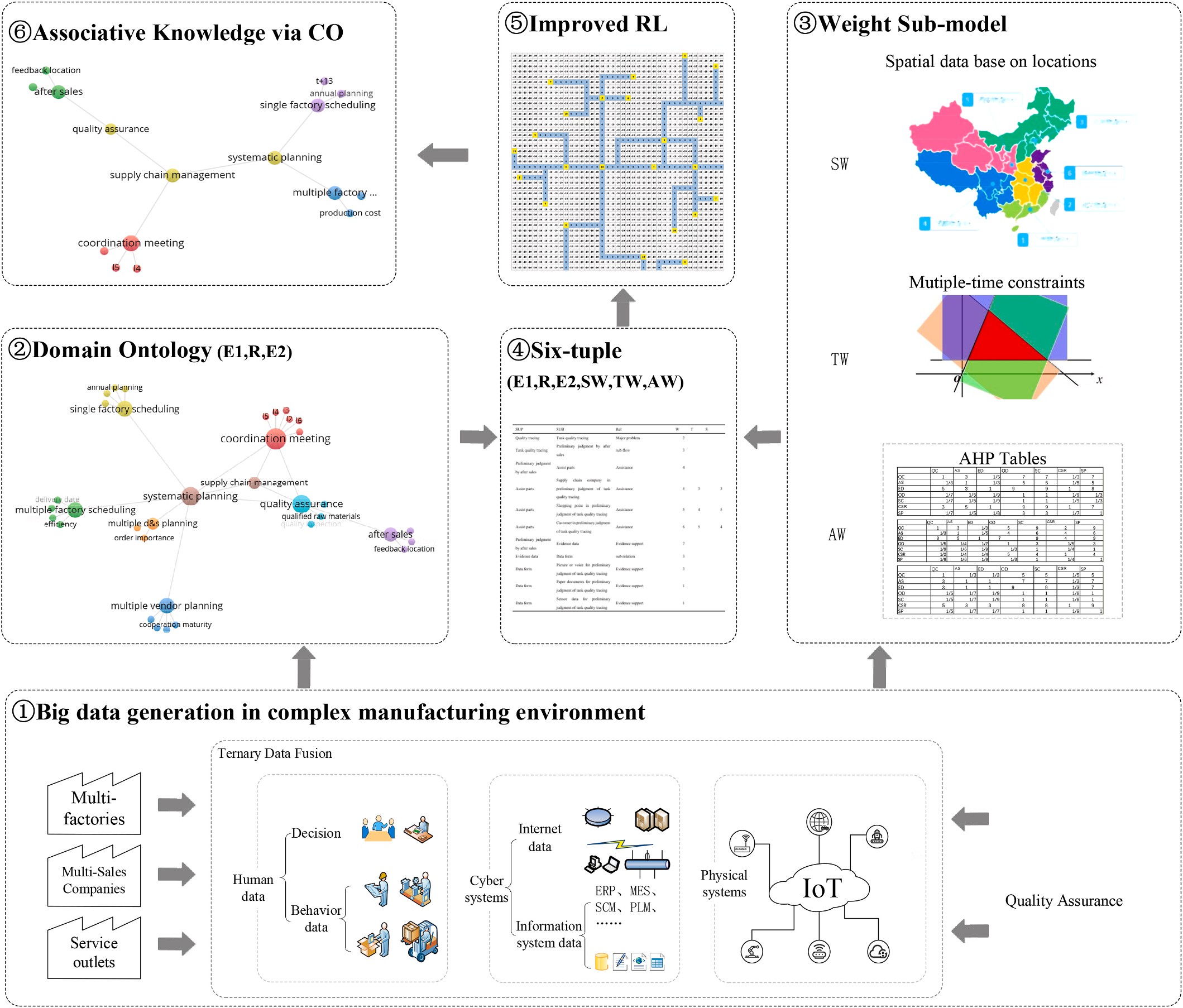
The main work content of the operation of the entire water heater

manufacturing environment includes: systematic planning, quality assurance and coordination meetings. Systematic planning is divided into single factory scheduling, multi factory scheduling, multi-vendor planning, and multi-D&S (delivery and sales) planning. The sched- uling of a single factory mainly refers to the annual plan. In addition,

there are also a rolling 13-day production plan (T+13), a production and delivery plan for the next day (T+1), and a supply plan for one day in advance (T- 1). Multi factory scheduling needs to consider the distance

between the delivery point and the shipping factory, which is strongly related to the cost and efficiency of transportation, and it is often considered to hand the order to the factory near the delivery point for production. In terms of quality assurance, after the after-sales service outlets receive feedback from users, they need to conduct quality tracing and determine the most suitable maintenance plan. Therefore, we also need to consider regional issues here, that is, considering spatiotemporal data and its weight data. Coordination meeting is an important mani- festation of human data in the entire model. The participants in the coordination meeting are senior experts from important factories, sales companies and after-sales service departments. They will discuss various uncertain factors in order to make correct decisions on specific issues.

There are many methods for constructing ontology models, most of



**Fig. 1.** The overall framework of associative knowledge discovery base on DOSTAR.

which are based on text mining. This article believes that decision- making structure, work flow, management specifications and encyclo- pedia can all be used as the basis for ontology construction. Domain ontology is an important step of this research, but the method of con- structing ontology is not the focus. Therefore, it will not go into too much detail here.

* 1. *wt sub-model*

Weight refers to the degree of importance of a certain factor or in- dicator relative to a certain thing. Different levels of importance should be represented by different values. The weight value is a relative value, which mainly indicates the order of the importance of different factors or indicators.

1. AHP

Decision weights are often calculated using the AHP method. The construction of domain ontology and reinforcement learning is based on the top-level data fusion framework here.

Since the reward matrix only needs to be transformed by the weight value of the decision factor, there is no need to compare the solutions. Therefore, the AHP model in this paper is limited to the criterion level, but not the scheme level. The specific modeling is as follows:

In the first step, in order to integrate human data from multiple perspectives, this study adopts the decision preferences of factory di- rector, QA director and Sales director throughout the traceability pro- cess, which are defined as follows:

*A* = *(aij)* represents the factory expert’s decision matrix;

*B* = *(bij)* represents the decision matrix of QA director;

*C* = *(cij)* represents the decision matrix of Sales director.

In the reward matrix of reinforcement learning, these decision values

need to be input, and different subsets of the matrix are selected ac- cording to the attributes of the six-tuples of different nodes, which are defined as follows:

⎧⎨ *aij*, *sub* ∈ { *factory* exp *ert* }

*V*(*s*) = {11, 12, 13, 14, 15, 16} (3)

*SW* = {1, 2, 3, 4, 5, 6, 7, 8, 9} (4)

*V*(*t*) = {11, 12, 13, 14, 15, 16} (5)

*TW* = {1, 2, 3, 4, 5, 6, 7, 8, 9} (6)

*L1*~*l6* means different locations. These locations are where different

factories, sales companies, and after-sales service outlets are located. In the weighting graph of time and space, these locations are nodes, and the weight value is the weight between any two nodes. For example,

companies. In the respective adjacency matrices *SW(l4,l2)* = *2, SW(l4, l6)* = *5, TW(l4,l2)* = *5, TW(l4,l6)* = *6*. This shows that the distance from products are produced in the *l4* factory and shipped to the *l2* and *l6* sales

the factory of *l4* to *l2* (the sales company) is relatively recent. Although

the production time is the same, it takes longer to transport to *l6*.

* 1. *Six-tuple sub-model*

The six tuples are combined as parent node, relationship, child node, space weight, time weight and AHP weight. The parent node, child node and relationship are derived from the triples of the ontology (entity 1, relationship, entity 2). Three weights values are from weight sub-model. By adjusting the weights of similarly related data for different targets and different dimensions, the fusion calculation can be smoother. Based on the above analysis, the six-tuple model is expressed as follows:

*F* = 〈*E*1, *R*, *E*2, *SW*, *TW*, *AW*〉 (7)

Among them:

“*F*” is for six tuples; “*E1*” is for father nodes; “*E2*” is for child nodes “*R*” is for relations;

“*SW*” is for space weight value; “*TW*” is for time weight value; “*AW*” is for AHP weight value.

*Rsub* =

*bij*, *sub* ∈ {*quality*}

⎩ *cij*, *sub* ∈ { *after sales* }

(1)

* 1. *Improved reinforcement learning algorithm*

Reinforcement learning models can find strategies or solutions that

In the second step, decision preferences are transformed into values in the reward matrix through data reduction. It is defined as follows:

*R(subi)* is for the values of decision preference at children nodes

“*subi*”

*R*′*(subi)* is for the values of decision preference at children nodes

“*subi*” after data reduction.

After many adjustments, the following formula is most suitable for integrating decision preferences into the reward matrix:

*R*′ (*sub* ) = *ROUNDUP*(*R*(*subi*) \* *C*, 0) (2)

*i*

C is a constant here, and its value should be determined by trial and

*R’(subi)*∈*{1,2,3,4,5,6,7,8,9,10}*. And according to repeated calculations, it comes *R(subi)*∈*(0.01, 0.5)*. After many trials, the value of C here is error. According to the setting of the reward matrix in this research,

selected as 20.

1. Time and space governance

Spatiotemporal data has a lot of specific data, and in this study, in order to better integrate multiple models, they are transformed into adjacency matrix of weighted graph. It is defined as follows:

maximize the return of the objective function in a live maze of chess- boards. In this paper, the reinforcement learning model is used for associative knowledge discovery. Taking the multi-factory management as an example, we create the corresponding domain ontology based on actual work. Since the composition of this domain ontology is relatively complicated, this article shows the core graph, as shown in [Fig. 2](#_bookmark5).

Obviously, the domain ontology in [Fig. 2](#_bookmark5) is difficult to calculate directly for reinforcement learning. So, we need to transform it. There are several steps to achieve that. Step 1, domain ontology is needed to transform onto “chessboard”. In this step, the domain ontology is put

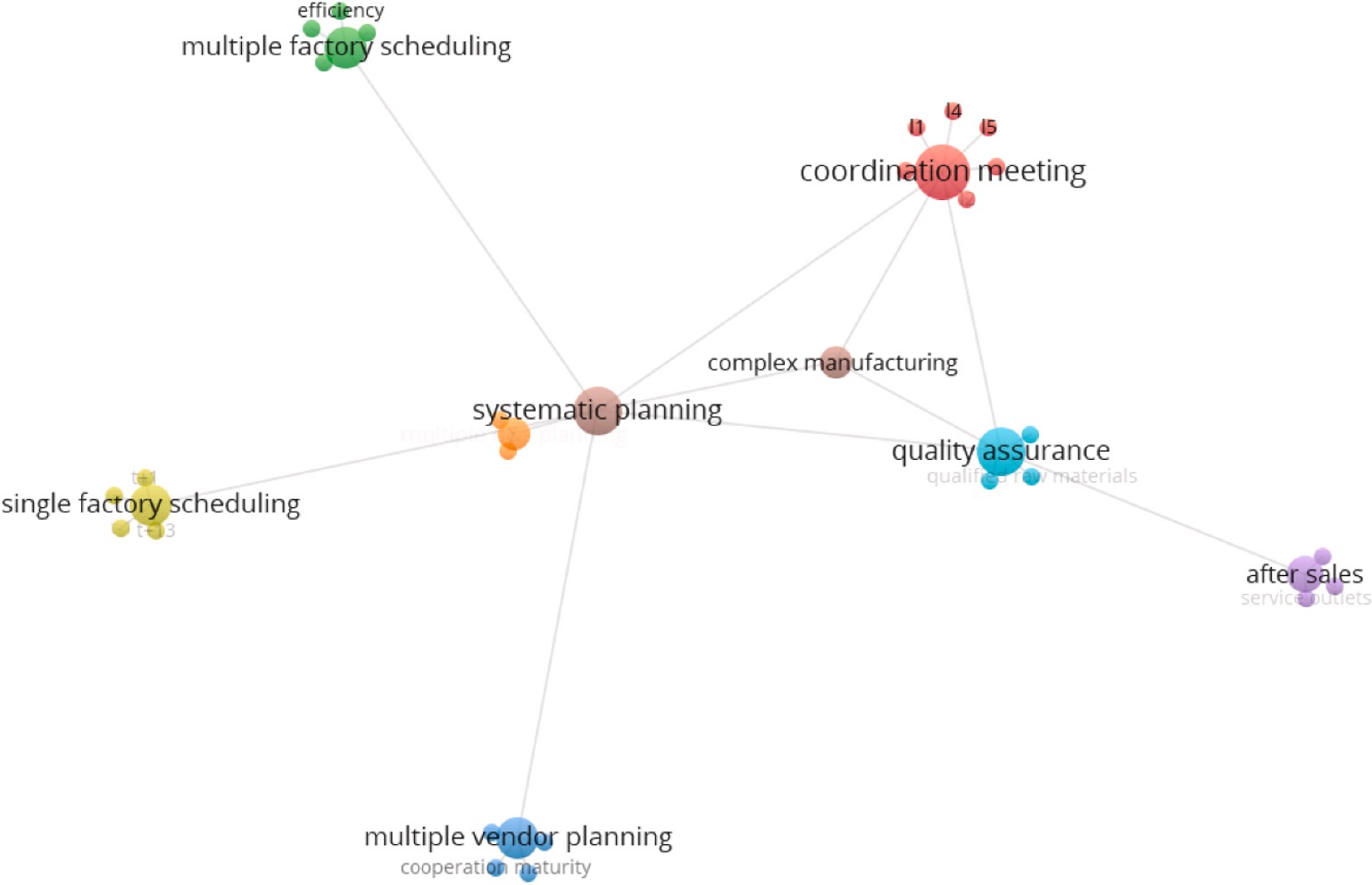
onto the chessboard. The nodes on the chessboard have specific posi-

tions. But the angle between the relationships (i.e., the connection) of each two nodes of ontologies can be arbitrary. It is not suitable for the standard policy of reinforcement learning. Therefore, the graph of domain ontology should be further processed.

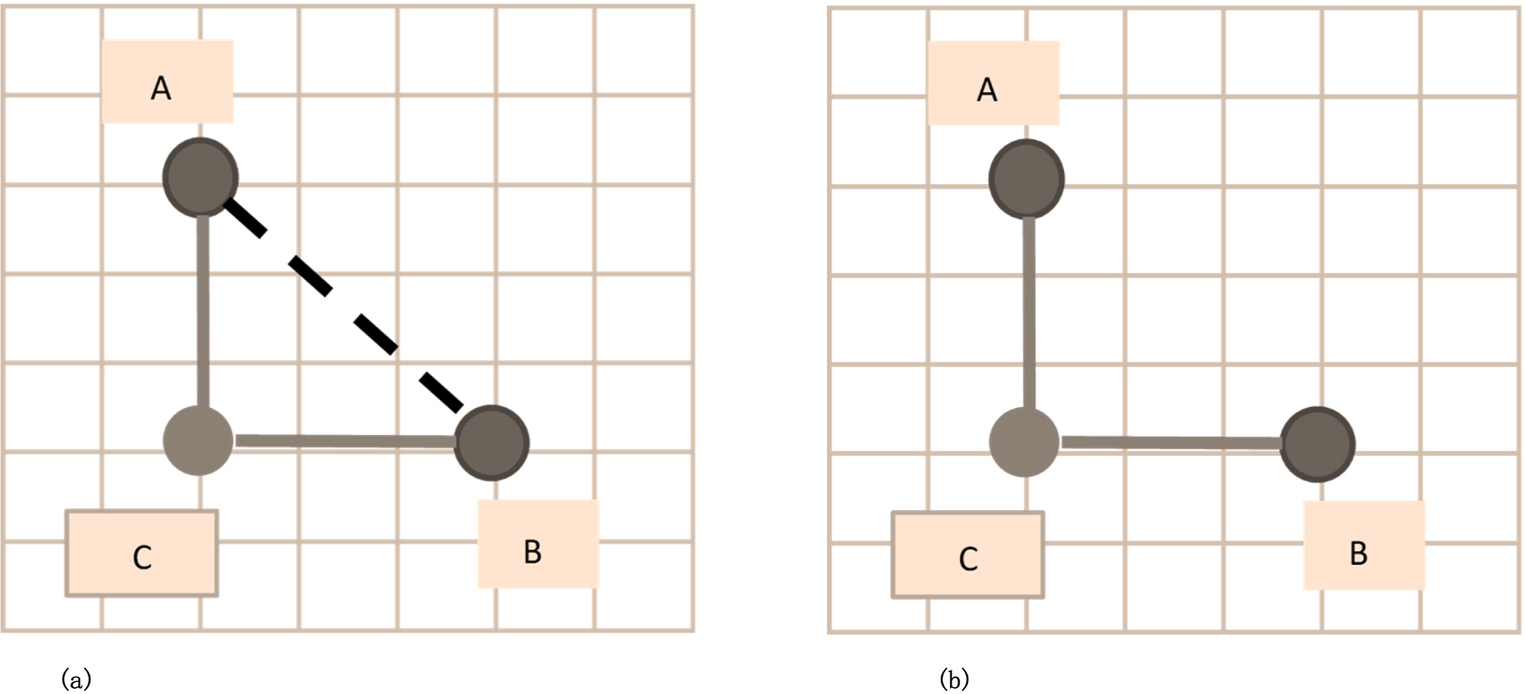
The step 2, adding the “Right angle” inside the basic graph. Between

relationship and turns the segment AB to polyline ABC, and AC⊥BC. And two nodes of ontologies, one kind of right angle is set to transition their then, the original segment is removed. This is shown as in [Fig. 3](#_bookmark6)(a). And

then, the original segment is removed, such as segment AB in [Fig. 3](#_bookmark6)(b). The step 3, it is possible to perform a chessboard simulation of the reward matrix for reinforcement learning model. The values of reward matrix come from weight sub-model. [Fig. 4](#_bookmark7) shows very clearly that all



**Fig. 2.** Domain ontology for complex manufacturing environment of water heater.



**Fig. 3.** (a) Adding right angle for the notes of ontologies**. (b)** Removing the original segment of notes.

the nodes and path are transformed into matrix or a kind of chessboard.

Step 4, the calculation of the reinforcement learning algorithm on the chessboard.

The formula for Bellman equation of Q function is as follows:

The related pseudocode for Q function (8) is as follows:

self.q\_table.ix [state, action] + = self.alpha \* (next\_state\_reward + self.gamma \* next\_state\_q\_values.max () - self.q\_table.ix [state, action])

Four policies of up, down, left and right are stood as *u*, *d*, *l* and *r*. The corresponding pseudocode is as follows:

*π* ∑ *a* [ *a*

∑ *π* ′ ′ ]

def get\_next\_state (self, state, action):

*Q* (*s*, *a*) =

Where.

*Pss*′

*s*′

*Rss*′ + *γ*

*a*′

*Q* (*s* , *a* )

(8)

if action = = ‘u’ and state next\_state = state - self.MAZE\_C elif action = = ‘d’ and state next\_state = state + self.MAZE\_C

elif action = = ‘l’ and state % self.MAZE\_C ! = 0:

next\_state = state - 1

“*π*" stands for policy,

“*γ*" represents the state value function,

“*Q*" stands for action value function,

“*P*" stands for state transition probability,

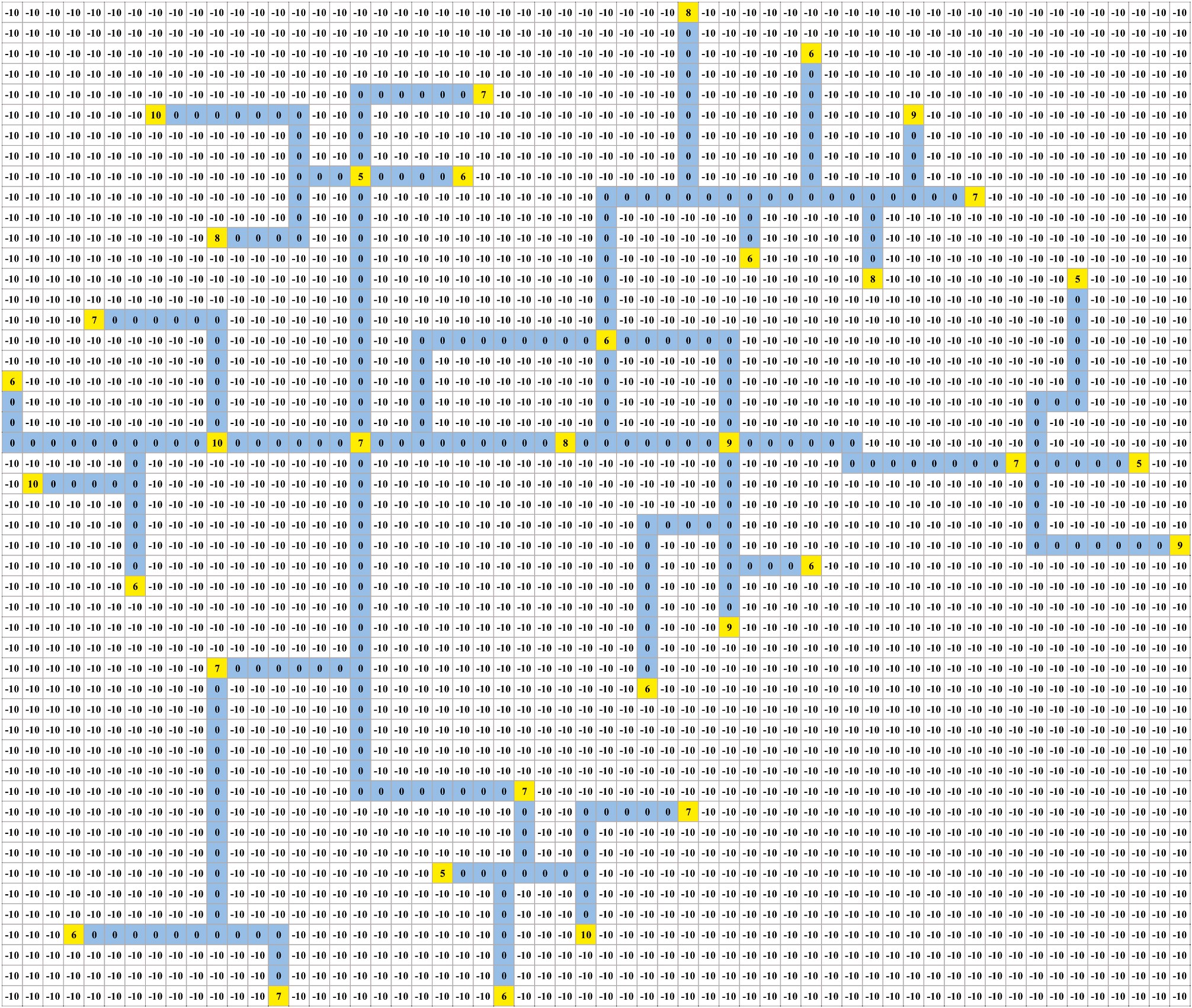
“*R*" stands for reward, “*s*" stands for state, and “*a*" stands for action.

elif action = = ‘r’ and state % self.MAZE\_C ! = self.MAZE\_C - 1: next\_state = state + 1

else:

next\_state = state return next\_state

In order to facilitate the calculation, it is necessary to set the co- ordinates of the starting point and the ending point (for example, take the top-level node as the starting point and the four corners of the “chessboard” as the ending point). And in order to prevent the agent of



**Fig. 4.** Reward matrix in chessboard.

reinforcement learning from going around too much meaninglessly, it is needed to set a pioneering ratio (such as 30%), that is, the policies in Q- table are adopted in 70% of cases, but there is a 30% probability that the agent is randomly selected an action to have a try. In addition, for the purpose of preventing the agent from missing important “passing

points”, this model limits the minimum value of the cumulative dis-

counted return (reward), such as 80 or even higher. That is, the total value of the cumulative discounted return of the model must be greater than 80, otherwise it is not considered to find a suitable result. But the calculation steps increase rapidly when the number of this constraint is larger. After more than 5000 steps of training (due to the existence of random values, sometimes tens of thousands of steps, and more complex cases requiring even millions or more), the optimal cumulative dis- counted return (reward) can be found for one kind of tracing.

# Case study

In this part, it takes specific events in the complex manufacturing environment of Haier water heaters as an example to verify the above model. Orders for “A" products need to be produced in advance and

shipped to the sales company “*l2*” (location “*l5*′′). Although these orders

are in the annual plan of a certain factory (location “*l4*′′), the delivery date has been advanced. At the same time, product A received some fault

feedback at after-sales outlets. Now it is expected to deliver on time under the premise of ensuring product quality.

* 1. *Six-tuple*

The number of records in the six-tuple data will be relatively large, and it will continue to be updated. Only important data related to the case are listed here. The six-tuple is shown in [Table 2](#_bookmark8).

* 1. *AHP tables*

Based on the cases of water heater quality traceability, the AHP sub- model is used to calculate the decision-making factors of three decision makers (factory director, QA director, and sales director). Here, the level

of AHP’s decision factors and the hierarchical structure of the knowl- edge subgraph are different. The advantage of this is that it not only

simplifies the calculation of each sub-module, but also makes use of the advantages of each sub-module. Quality assurance (QA), feedback location (AF1), repaired factory (AF2), SFS (single factory scheduling), multi-factory planning (MFP), participants “l2” (CM\_l2), participants

“l4” (CM\_l4) and participants “l5” (CM\_l5) are selected as the calculation

objects of AHP.

Their decision matrix is as in [Tables 3–5](#_bookmark10). In order to ensure the

**Table 2**

A six-tuple when the order is advanced and the quality needs to be improved.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| E1 | R | E2 | SW | TW | AW |  |  | QA | AF1 | AF2 | SFS | MFP | CM\_l2 | CM\_l4 | CM\_l5 |
| complex | sub-level | Quality | 8 | 9 | 6 |  | QA | 1 | 3 | 3 | 9 | 1 | 7 | 1 | 7 |
| manufacturing |  | assurance |  |  |  |  | AF1 | 1/3 | 1 | 1 | 3 | 1/3 | 6 | 1/3 | 6 |
| environment |  |  |  |  |  |  | AF2 | 1/3 | 1 | 1 | 3 | 1/3 | 6 | 1/3 | 6 |
| management |  |  |  |  |  |  | SFS | 1/9 | 1/3 | 1/3 | 1 | 1/7 | 1/4 | 1/7 | 1/4 |
| complex | sub-level | Systematic | 8 | 7 | 6 |  | MFP | 1 | 3 | 3 | 7 | 1 | 6 | 1 | 6 |
| manufacturing |  | Planning |  |  |  |  | CM\_l2 | 1/7 | 1/6 | 1/6 | 4 | 1/6 | 1 | 1/5 | 1 |
| environment |  |  |  |  |  |  | CM\_l4 | 1 | 3 | 3 | 7 | 1 | 5 | 1 | 5 |
| management |  |  |  |  |  |  | CM\_l5 | 1/7 | 1/6 | 1/6 | 4 | 1/6 | 1 | 1/5 | 1 |
| complex | sub-level | Coordination | 7 | 6 | 6 |  |  |  |  |  |  |  |  |  |  |

**Table 5**

The matrix of sales director’s decision. *CR3* = 0.0706.

manufacturing environment management

Quality assurance influencing factors

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Systematic Planning | sub-level | Single factory scheduling | 7 | 9 | 8 | [Table 6](#_bookmark11) shows the basis and results of the combined weight calcu-  lation. The final weight (defined as *FWij*) is equal to the sum of the factor |
| Systematic Planning | sub-level | Multiple factory | 7 | 9 | 7 | weight of each decision maker multiplied by the integrated weight. It is |

meeting

After sales 6 7 6

represented by *IW1*, *IW2*, and *IW3* respectively. The specific values of multiple decision makers are shown in [Table 6](#_bookmark11). Based on the above calculations, the final comprehensive weights are shown in [Table 7](#_bookmark12).

defined as [formula (9)](#_bookmark9).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | | scheduling |  | | |
| Single factory | combination | Annual planning | 7 | 7 | 8 |
| scheduling | calculation |  |  |  |  |
| Single factory | combination | T+13 | 5 | 5 | 6 |
| scheduling | calculation |  |  |  |  |
| Multiple factory | combination | Production cost | 6 | 5 | 8 |
| scheduling | calculation |  |  |  |  |
| Multiple factory | combination | Load vs | 8 | 5 | 6 |
| scheduling | calculation | Capacity |  |  |  |

*FWij* =

∑ *aij* × *IW*1 + ∑ *bij* × *IW*2 + ∑ *cij* × *IW*3

(9)

Through the above steps, the knowledge and experience of different decision makers can be reflected in the final decision matrix. The weight values of these decision matrices will become part of the reward matrix

3

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Coordination meeting | influencing | L2 | 9 | 6 | 6 | of the reinforcement learning algorithm. |
|  | factors |  |  |  |  |  |
| Coordination meeting | influencing | L4 | 7 | 6 | 8 |  |
| Coordination meeting | factors  influencing | L5 | 9 | 6 | 8 | *4.3. Results of improved RL* |
|  | factors |  |  |  |  |  |

**Table 3**

The matrix of factory director’s decision. *CR*1 = 0.091.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | QA | AF1 | AF2 | SFS | MFP | CM\_l2 | CM\_l4 | CM\_l5 |
| QA | 1 | 7 | 5 | 5 | 1/3 | 3 | 3 | 3 |
| AF1 | 1/7 | 1 | 1/5 | 1/5 | 1/7 | 1/7 | 1/7 | 1/7 |
| AF2 | 1/5 | 5 | 1 | 1 | 1/5 | 1/5 | 1/5 | 1/5 |
| SFS | 1/5 | 5 | 1 | 1 | 1/5 | 1/7 | 1/7 | 1/7 |
| MFP | 3 | 7 | 5 | 5 | 1 | 1 | 1 | 1 |
| CM\_l2 | 1/3 | 7 | 5 | 7 | 1 | 1 | 1 | 1 |
| CM\_l4 | 1/3 | 7 | 5 | 7 | 1 | 1 | 1 | 1 |
| CM\_l5 | 1/3 | 7 | 5 | 7 | 1 | 1 | 1 | 1 |

**Table 4**

The matrix of QA director’s decision. *CR*2 = 0.0721.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | QA | AF1 | AF2 | SFS | MFP | CM\_l2 | CM\_l4 | CM\_l5 |
| QA | 1 | 2 | 2 | 7 | 1/3 | 5 | 2 | 7 |
| AF1 | 1/2 | 1 | 2 | 5 | 1/2 | 4 | 1/3 | 6 |
| AF2 | 1/2 | 1/2 | 1 | 1 | 1/5 | 5 | 1 | 7 |
| SFS | 1/7 | 1/5 | 1 | 1 | 1/7 | 1/3 | 1/3 | 3 |
| MFP | 3 | 2 | 5 | 7 | 1 | 3 | 3 | 9 |
| CM\_l2 | 1/5 | 1/4 | 1/5 | 3 | 1/3 | 1 | 1/3 | 3 |
| CM\_l4 | 1/2 | 3 | 1 | 3 | 1/3 | 3 | 1 | 5 |
| CM\_l5 | 1/7 | 1/6 | 1/7 | 1/3 | 1/9 | 1/3 | 1/5 | 1 |

essentially experts representing the interests of location. The individual decision preferences of these three experts will be obtained through AHP first, and then integrated according to the method of cooperative games. After calculation, the subgraph representing the associated knowledge is shown in [Fig. 5](#_bookmark13).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| After sales | influencing factors | Feedback location | 7 | 6 | 8 | The subgraph for the above case is calculated. This shows that the  most detailed factors that need to be considered are: feedback location, |
| After sales | influencing factors | Rework factory | 5 | 9 | 5 | maintenance plant, annual plan, T+13, cost, load vs capacity. It should  be noted that the l2, l4, and l5 related to the coordination meeting are |

In order to facilitate the calculation, it is necessary to set the co- ordinates of the starting point and the ending point (for example, take the top-level node as the starting point and the four corners of the “chessboard” as the ending point). And in order to prevent the agent of

reinforcement learning from going around too much meaninglessly, it is

needed to set a pioneering ratio (such as 20%), that is, the policies in Q- table is adopted in 80% of cases, but there is a 20% probability that the agent is randomly selected an action to have a try. In addition, for the

purpose of preventing the agent from missing important “passing points”, this model limits the minimum value of the cumulative dis- counted return (reward), such as 150 or even higher. That is, the total

value of the cumulative discounted return of the model must be greater than 150, otherwise it is not considered to find a suitable result. But the calculation steps increase rapidly when the number of this constraint is larger. After more than 20,000 steps of training (due to the existence of random values, sometimes tens of thousands of steps, and more complex cases requiring even millions or more), the optimal cumulative dis- counted return (reward) can be found.

In order to compare the effects of different algorithms with this case, their advantages and disadvantages are listed in [Table 8](#_bookmark14). And the

consistency of decision-making results, it is necessary to ensure that CR is less than 0.1. In addition, there are certain conflicts between multiple decision-making preferences. In order to further optimize this situation, different decision preferences need to be weighted to obtain a result that

**Table 6**

The weight of decision makers.

Factory director (*IW1*)

QA director (*IW2*)

Sales director (*IW3*)

satisfies every decision maker. After the negotiation of the decision makers, the integrated weights of the three decision makers are

Integrated weight

40% 40% 20%

**Table 7**

Weight results.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Factory director | QA director | Sales director | Final weight |
| QA | 24.22% | 19.54% | 23.88% | 24.09% |
| AF1 | 1.84% | 13.49% | 10.80% | 9.42% |
| AF2 | 4.16% | 10.59% | 10.80% | 9.52% |
| SFS | 3.87% | 4.26% | 2.43% | 3.72% |
| MFP | 20.42% | 30.55% | 22.45% | 27.14% |
| CM\_l2 | 15.16% | 5.75% | 3.98% | 7.20% |
| CM\_l4 | 15.16% | 13.73% | 21.67% | 15.40% |
| CM\_l5 | 15.16% | 2.10% | 3.98% | 3.49% |

abilities of interpretability and fusing cross-domain and cross-layer data are quite innovative and remarkable for real business use. If the same data is used, other algorithms cannot perform calculations directly, and some data preprocessing work is also required. In addition, single al- gorithms can only solve sub problems; they cannot solve all the prob- lems corresponding to the algorithms proposed in this article.

In this section, first six-tuple is used to integrate multi-source data and perform the governance of spatiotemporal data. Then, the decision weights are determined by AHP. Then, through reinforcement learning

algorithms, calculations are performed on a “chessboard” based on the ontology of the business scenario and fused with the calculation results

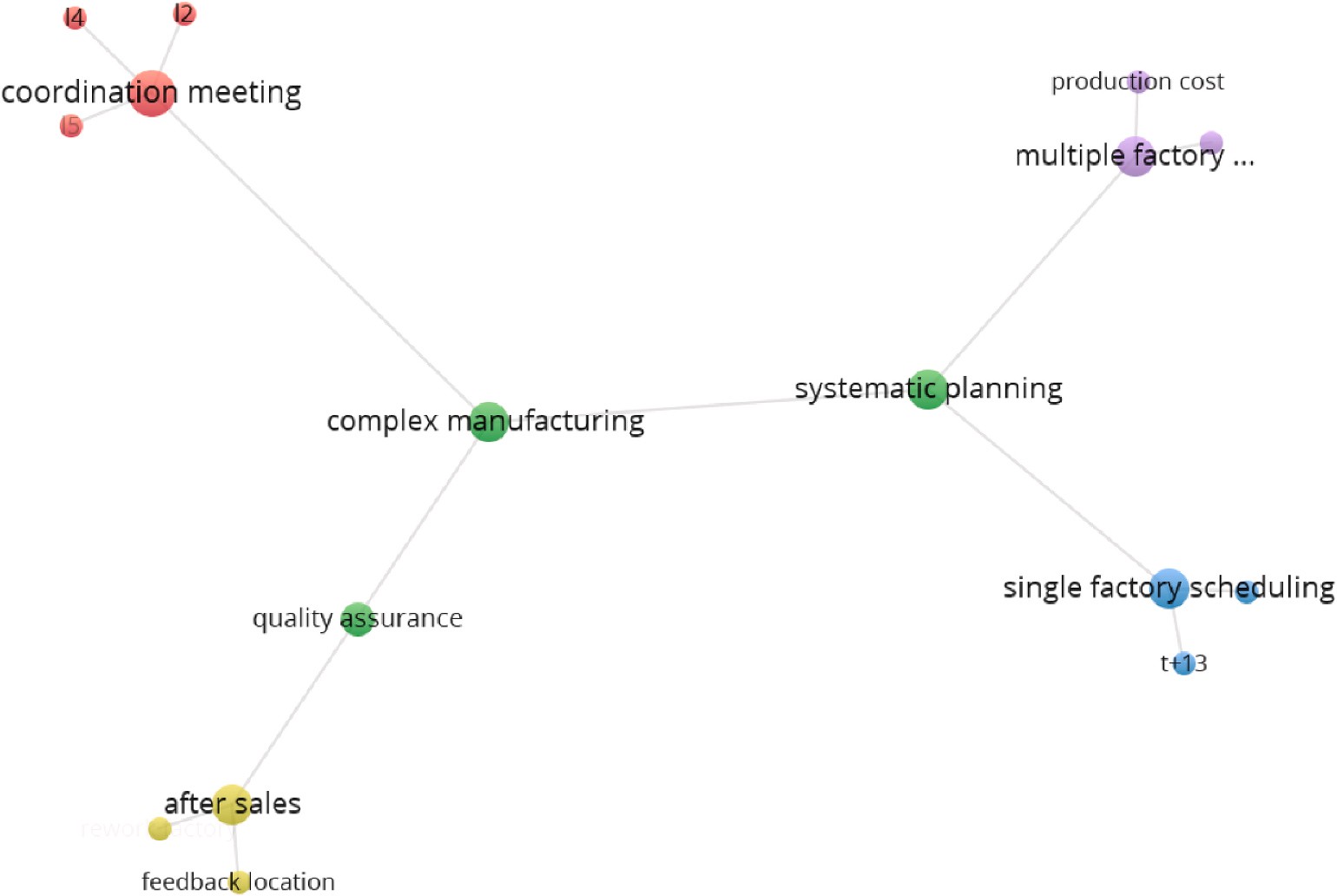
of other sub-models; associative knowledge is discovered for decision- making. Through data verification, the calculation results of this model improve the overall work efficiency of water heater quality

traceability.

# Conclusion

Based on the multi-factory case of Haier electric water heater, this paper constructs an DOSTAR fusion model for associative knowledge discovery. This fusion model is divided into four parts: Weight sub- model, domain ontology, six-tuple and improved reinforcement learning. This research integrates various types of big data from multiple dimensions, multiple perspectives, cross-regions, and across time hori- zons throughout the whole process. These data include structured, semi- structured and unstructured data. These data are connected through AHP and domain ontology. The inclusiveness of these connections is very good. In particular, this study fuses human data through the AHP sub-model. The improved reinforcement learning sub-model shows that this research uses artificial intelligence algorithms for associative knowledge discovery. Finally, through the case study, it is obviously that this method can effectively optimize the entire manufacturing network to achieve the purpose of reducing costs and increasing efficiency. It provides innovative ideas for solving related problems.

Due to the relatively short time, there is no time to debug the multi- agent reinforcement learning model. I believe this will greatly shorten the calculation time of the reinforcement learning model. In addition, follow-up research work will further expand the data. More detailed and large knowledge discovery is expected.



**Fig. 5.** The subgraph representing the associative knowledge.

**Table 8**

Comparison of the effects of different algorithms.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Method comparison | Advantage | Disadvantage | Interpretability | Cross-domain and cross layer |
| Deterministic planning model | Fast and accurate | Can’t solve the problem of | Yes | No |
|  |  | uncertainty |  |  |
| Ontology | Suitable for the fusion of ontology and relationship at | Unable to calculate environmental | Yes | No |
|  | the level of natural language | feedback |  |  |
| Machine learning | Can handle massive amounts of data | Prone to fitting problems and | No | No |
|  |  | dimensional disasters |  |  |
| New method in this article (reinforcement | Can handle massive amounts of data and | Initial modeling takes time | Yes | Yes |
| learning + ontology) | interpretability |  |  |  |

# Declaration of competing interest

The authors declared that they have no conflicts of interest to this work. We declare that we do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted.

# Acknowledgement

This work was supported by the National Science and Technology Innovation 2030 Next-Generation Artificial Intelligence Major Project, Data-Driven Tripartite Collaborative Decision-Making and Optimiza- tion, under Grant 2018AAA0101800.

# References

1. Chen S, Wang J, Li H, Wang Z, Liu F, Li S. Top-down human-cyber-physical data fusion based on reinforcement learning. IEEE Access 2020;8:134233–45. [https://](https://doi.org/10.1109/ACCESS.2020.3011254) [doi.org/10.1109/ACCESS.2020.3011254](https://doi.org/10.1109/ACCESS.2020.3011254).
2. Chand P, Thakkar JJ, Ghosh KK. Analysis of supply chain complexity drivers for Indian mining equipment manufacturing companies combining SAP-LAP and AHP.

Resour Pol 2018;59:389–410. <https://doi.org/10.1016/j.resourpol.2018.08.011>.

1. Amouzgar K, Bandaru S, Andersson T, Ng AHC. A framework for simulation-based multi-objective optimization and knowledge discovery of machining process. Int J

Adv Manuf Technol 2018;98(9–12):2469–86. [https://doi.org/10.1007/s00170-](https://doi.org/10.1007/s00170-018-2360-8) [018-2360-8](https://doi.org/10.1007/s00170-018-2360-8).

1. Korsemov D, Borissova D, Mustakerov I. Combinatorial optimization model for group decision-making. Cybern Inf Technol 2018;18(2):65–73. [https://doi.org/](https://doi.org/10.2478/cait-2018-0028) [10.2478/cait-2018-0028](https://doi.org/10.2478/cait-2018-0028).
2. Martinez S, Gonzalez C, Hospitaler A, Albero V. Sustainability assessment of constructive solutions for urban Spain: a multi-objective combinatorial optimization problem. Sustainability 2019;11(3). [https://doi.org/10.3390/](https://doi.org/10.3390/su11030839) [su11030839](https://doi.org/10.3390/su11030839).
3. Sitek P, Wikarek J. A multi-level approach to ubiquitous modeling and solving

constraints in combinatorial optimization problems in production and distribution. Appl Intell 2018;48(5):1344–67. <https://doi.org/10.1007/s10489-017-1107-9>.

1. Elorza A, Hernando L, Mendiburu A, Lozano JA. Estimating attraction basin sizes of

combinatorial optimization problems. Prog Artif Intell 2018;7(4):369–84. [https://](https://doi.org/10.1007/s13748-018-0156-6) [doi.org/10.1007/s13748-018-0156-6](https://doi.org/10.1007/s13748-018-0156-6).

1. Vesselinova N, Steinert R, Perez-Ramirez DF, Boman M. Learning combinatorial optimization on graphs: a survey with applications to networking. IEEE Access

2020;8:120388–416. <https://doi.org/10.1109/ACCESS.2020.3004964>.

1. Lu H, Zhou R, Cheng S, Shi Y. Multi-center variable-scale search algorithm for combinatorial optimization problems with the multimodal property. Appl Soft Comput 2019;84. <https://doi.org/10.1016/j.asoc.2019.105726>.
2. Chalupa D, Hawick KA. GraphCombEx: a software tool for exploration of

combinatorial optimisation properties of large graphs. Soft Comput 2019;23(14): 5715–24. <https://doi.org/10.1007/s00500-018-3230-x>.

1. Chaabani A, Bechikh S, Ben Said L. A co-evolutionary hybrid decomposition-based

algorithm for bi-level combinatorial optimization problems. Soft Comput 2020;24 (10):7211–29. <https://doi.org/10.1007/s00500-019-04337-0>.

1. Jiang T, Zhang C. Application of grey wolf optimization for solving combinatorial

problems: job shop and flexible job shop scheduling cases. IEEE Access 2018;6: 26231–40. <https://doi.org/10.1109/ACCESS.2018.2833552>.

1. Zhao S, Zhang Q, Peng Z, Lu X. Personalized manufacturing service composition

recommendation: combining combinatorial optimization and collaborative

filtering. J Combin Optim 2020;40(3):733–56. [https://doi.org/10.1007/s10878-](https://doi.org/10.1007/s10878-020-00613-0) [020-00613-0](https://doi.org/10.1007/s10878-020-00613-0).

1. [Merlo C, et al. Proposal of a user-centred approach for CPS design: pillbox case](http://refhub.elsevier.com/S2590-0056(22)00048-0/sref14) [study. IFAC-PapersOnLine 2019;51(34):196–201](http://refhub.elsevier.com/S2590-0056(22)00048-0/sref14).
2. Zhou P, et al. A comprehensive technological survey on the dependable self-

management CPS: from self-adaptive architecture to self-management strategies. P. Zhou et al. Sensors 2019;19(5). <https://doi.org/10.3390/s19051033>.

1. Barz M, Barz M, Poller P, Schneider M, Zillner S, Sonntag D. Human-in-the-Loop

control processes in gas turbine maintenance. Indus Appl Holonic Multi-Agent Syst 2017;10444:255–68. <https://doi.org/10.1007/978-3-319-64635-0_19>.

1. P. Fantini et al., “Exploring the integration of the human as a flexibility factor in

CPS enabled manufacturing environments: Methodology and results,” pp. 5711–5716, doi: 10.1109/IECON.2016.7793579.

1. Dong Q, Cooper O. An orders-of-magnitude AHP supply chain risk assessment framework. Int J Prod Econ 2016;182:144–56. [https://doi.org/10.1016/j.](https://doi.org/10.1016/j.ijpe.2016.08.021) [ijpe.2016.08.021](https://doi.org/10.1016/j.ijpe.2016.08.021).
2. Wang K, Ji W. Evaluation of machining process scheme based on hierarchical analysis and grey correlation analysis. Mod Manuf Eng 2019;2:24–30. [https://doi.](https://doi.org/10.16731/j.cnki.1671-3133.2019.02.005) [org/10.16731/j.cnki.1671-3133.2019.02.005](https://doi.org/10.16731/j.cnki.1671-3133.2019.02.005).
3. Helmick RE, Rice TR. Removal of alignment errors in an integrated system of two 3-D sensors. IEEE Trans Aero Electron Syst 1993;29(4):1333–43. [https://doi.org/](https://doi.org/10.1109/7.259537) [10.1109/7.259537](https://doi.org/10.1109/7.259537).
4. Mukhopadhyay SC, Leung H. Advances in Wireless Sensors and Sensor Networks, vol. 64. Springer Berlin Heidelberg; 2010. [https://doi.org/10.1007/978-3-642-](https://doi.org/10.1007/978-3-642-12707-6) [12707-6](https://doi.org/10.1007/978-3-642-12707-6).
5. Wang S, Liu X, Wang H, Hu Q. A case study on spatio-temporal data mining of urban social management events based on ontology semantic analysis. Sustainability 2018;10(6). <https://doi.org/10.3390/su10062084>.
6. Bandaru S, Ng AHC, Deb K. Data mining methods for knowledge discovery in multi-objective optimization: Part B - new developments and applications. Expert

Syst Appl 2017;70:119–38. <https://doi.org/10.1016/j.eswa.2016.10.016>.

1. Cheng Y, Chen K, Sun H, Zhang Y, Tao F. Data and knowledge mining with big data towards smart production. J Ind Inf Integr 2018;9:1–13. [https://doi.org/10.1016/](https://doi.org/10.1016/j.jii.2017.08.001) [j.jii.2017.08.001](https://doi.org/10.1016/j.jii.2017.08.001).
2. Bianchi FM, Maiorino E, Livi L, Rizzi A, Sadeghian A. An agent-based algorithm

exploiting multiple local dissimilarities for clusters mining and knowledge discovery. Soft Comput 2017;21(5):1347–69. [https://doi.org/10.1007/s00500-](https://doi.org/10.1007/s00500-015-1876-1) [015-1876-1](https://doi.org/10.1007/s00500-015-1876-1).

1. Sanfilippo EM, Belkadi F, Bernard A. Ontology-based knowledge representation for additive manufacturing. Comput Ind 2019;109:182–94. [https://doi.org/10.1016/j.](https://doi.org/10.1016/j.compind.2019.03.006) [compind.2019.03.006](https://doi.org/10.1016/j.compind.2019.03.006).
2. Zeng X, Tu X, Liu Y, Fu X, Su Y. Toward better drug discovery with knowledge graph. Curr Opin Struct Biol 2022;72:114–26. [https://doi.org/10.1016/j.](https://doi.org/10.1016/j.sbi.2021.09.003) [sbi.2021.09.003](https://doi.org/10.1016/j.sbi.2021.09.003).
3. [Chen H, Li G, Sun Y, et al. A multi-hop link prediction approach based on](http://refhub.elsevier.com/S2590-0056(22)00048-0/sref28) [reinforcement learning in knowledge graphs. In: 2018 11th International](http://refhub.elsevier.com/S2590-0056(22)00048-0/sref28)

[Symposium on Computational Intelligence and Design (ISCID); 2018. p. 165–9](http://refhub.elsevier.com/S2590-0056(22)00048-0/sref28).

1. Z. Li, X. Jin, S. Guan, Y. Wang, and X. Cheng, “Path reasoning over knowledge graph: a multi-agent and reinforcement learning based method,” 2018 IEEE International Conference on Data Mining Workshops (ICDMW) IEEE*,* pp. 929–936, doi: 10.1109/ICDMW.2018.00135.
2. [Gohari FS, Tarokh MJ. New recommender framework: combining semantic](http://refhub.elsevier.com/S2590-0056(22)00048-0/sref30) [similarity fusion and bicluster collaborative filtering. Comput Intell 2016;32(4):](http://refhub.elsevier.com/S2590-0056(22)00048-0/sref30)

[561–86](http://refhub.elsevier.com/S2590-0056(22)00048-0/sref30).

1. Xie P, Li T, Liu J, Du S, Yang X, Zhang J. Urban flow prediction from spatiotemporal data using machine learning: a survey. Inf Fusion 2020;59:1–12. <https://doi.org/10.1016/j.inffus.2020.01.002>.
2. Erdemir E, Dragotti PL, Gündüz D. Privacy-aware time-series data sharing with deep reinforcement learning. IEEE Trans Inf Forensics Secur 2021;16:389–401. <https://doi.org/10.1109/TIFS.2020.3013200>.