Empirical Instances Knowledge Mining Method for Ship Design Domain

Tao Xie(✉)

Institute of Logistics Science and Engineering, Shanghai Maritime University, Shanghai, China
Xietao.199708@163.com

Abstract. Knowledge mining based on ship instances information has an important support role for ship design. In order to obtain easy to use knowledge from ship design instances composed of natural language text, this paper proposes an empirical case knowledge mining method for ship design domain. The method uses natural language processing techniques to structure the empirical instances according to the characteristics of the ship design domain and the structural characteristics of the empirical instances; and then combines knowledge clustering and variable precision rough set theory to obtain ship design rules from the instances. The experimental analysis of the data set provided by shipyards shows the feasibility and effectiveness of this method to obtain ship design knowledge rules.

Keywords: Knowledge mining · ship design domain · empirical instances · natural language processing

1 Introduction

Ship design is a knowledge-intensive job that requires close collaboration among designers with multidisciplinary backgrounds in structure, materials, mechanics, electronics, etc. Designers can only grasp the most basic design knowledge through design manuals and codes specified by classification societies, and struggle to cope with the complex design issues that arise in specific situations. Therefore, it is important to investigate knowledge mining methods for ship design to extract knowledge from past cases and assist designers to complete design tasks with specific requirements to improve the design effectiveness of shipyards.

Knowledge mining is an important step in the domain of knowledge management, a non-trivial process of identifying valid knowledge from data. In recent years, knowledge mining has also been gradually reused in the field of ship design and widely applied to the design stages of ship structure design [1], ship equipment design [2], ship energy system design [3], ship cabin layout [4], etc. The literature [2, 5] and [6] cite structured numerical data from tables, which are more normative and often can only obtain decision rules within specific values, and lack the adaptability to specific scenarios. The literature [7] and [8] cite unstructured textual data derived from design cases to obtain generative

X. Yuan et al. (Eds.): ICEKIM 2023, AHCS 13, pp. 2003–2010, 2023.
https://doi.org/10.2991/978-94-6463-172-2_223

rules and knowledge triples, respectively, which are more useful to assist designers in making decisions about design tasks than triples that only represent relationships. In existing research, knowledge mining still focuses on explicit knowledge mining, and automatic mining of tacit knowledge of designed empirical instances is not perfect. Combining the strengths and weaknesses of the aforementioned literature, in this paper, we propose an unstructured text-based empirical instance knowledge mining method that uses deep learning-based natural language processing techniques to unify the structured representations of Chinese empirical instance text infographics and extract ship design rules that can be easily applied. Knowledge mining method for empirical instances.

2 Knowledge Mining Methods for Empirical Instances

2.1 Structured Framework and Normalized Representation

Experience instance knowledge structuring is the process of instance knowledge symbolization, which is the basis of instance knowledge management and specification of research objects. Considering the key features required to generate empirical instances in the process of ship design problem solving, a three-level multivariate group instance knowledge structuring framework is established, as showed in Fig. 1.

In order to filter out meaningless information in the text and achieve complete unification of the structure, a normalized representation of nine fine-grained elemental information is required. The use of Backus-Naur Form (BNF) to describe empirical instance knowledge can effectively adapt to the complexity of ship terminology in the following form.

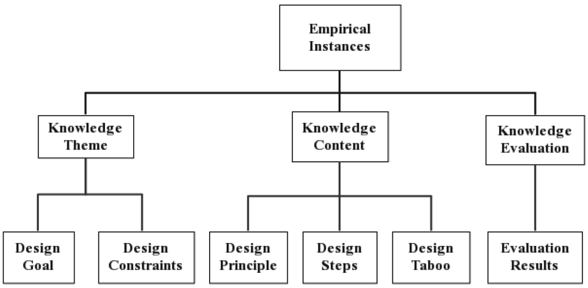

Fig. 1. Structured framework for knowledge of three-level multivariate group instances

Design Instance::=<Problem Requirement><Solution>{Knowledge Evaluation}
Problem Requirement::=<Design Goal>{Design Constraint}
Design Goal::=[<object>[attribute]]<action><object>[attribute]
Object::=<term>{term}
Attribute::=<name><type>[value][unit]
Action::=optimize |implement |increase |decrease |avoid
Name::=legal character
Type::=literal |value |Boolean |fuzzy
Design Constraint::=<object>[attribute][<logical-relationship><object>[attribute]][<action><object>[attribute]]
Logical Relation::= equals |not equal |includes |does not include |is less than |is greater than
Solution::={Design Principle}<Design Step>{Design Taboo}
Design Principle::=<Name><Principle Content>
Principle content::=<url>[line number][column number]
Design Step::=<Design Goal>{Design Goal}
Design Taboo::=<Design Step>{ Undesirable result}
Undesirable result::=<object><property>
Knowledge Evaluation::=<evaluator><Evaluation Result>
Evaluation Result::=strongly positive |weakly positive |neutral |weakly negative |strongly negative

2.2 A Framework for Knowledge Mining Methods Based on Deep Learning Algorithms

The knowledge mining method in this paper includes three parts: semantic role identification of instance fragments, knowledge attribute-value relationship extraction and knowledge clustering and rule mining, as shown in Fig. 2.

2.3 Specific Process

A. Semantic role recognition of instance fragments

In Chinese instances, text data is often not easily recognized by computers as important information due to the lack of formatting constraints. We preprocess the ship design data to obtain a perceptual ensemble of instances in terms of posts. These sentences are then classified by the deep neural network model Ernie and matched with the corresponding elements of the structured representation framework to achieve a uniform semi-structuring of the unordered text.

B. Knowledge attribute-value relationship extraction

Attribute-value extraction is mainly divided into two parts: attribute entity extraction and relationship extraction. We define fine-grained entity types and relationship categories applicable to ship design, and use Ernie model-based named entity identification techniques and relationship extraction techniques to achieve a normalized representation

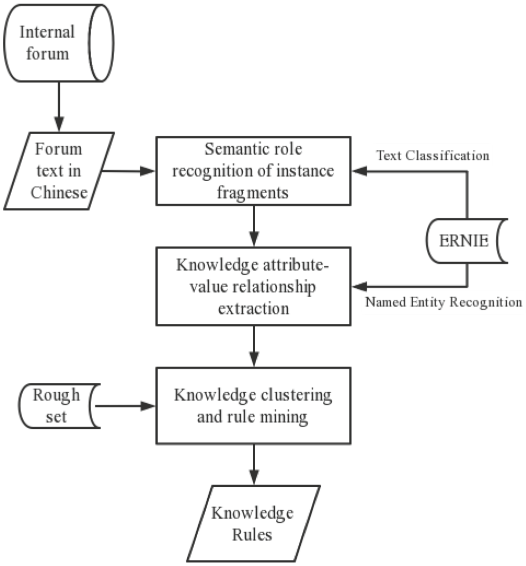

Fig. 2. Knowledge mining framework

of structured empirical instance knowledge. In this paper, seven categories of attribute entities are defined based on the terminology vocabulary of ship design, and two major categories of relationships are defined using knowledge triads considering the correlation between attribute entities. The annotation methods of entity attribute category and entity relationship category are shown in Table 1 and Table 2.

C. Knowledge clustering and rule mining

The instance knowledge is clustered by using semantic relevance calculation based on subject keywords, and then the rule mining of the clustered instance knowledge is performed by using the variable precision rough set algorithm, and a measurement index

Table 1. Types of entity property annotation in the field of ship design

Classification Description	Example
Object	Deck, ship's side
Property	Length, Area
Value	200, 15
Unit	m, m^2
Action	Calibrate, enhance
Relation	Greater than, equal to
Reference	Steel Code, Part 4, CCS Marine Code 2.3.5.5

Table 2. Types of relationships in the knowledge triad

Type 1	<Entity1, Relationship, Entity2>
Entity 1	Object, Value + Object
Relationship	Action, Relation, Value + Action…
Entity 2	Object, Property, Reference…
Example	<Deck, Replacement, Board>, <Cover, Not less than, Base plate>
Type 2	<Entity, Attribute, Attribute Value>
Entity 1	Object, Value + Object
Attribute	Property, Action + Property…
Attribute Value	Value, Value + Unit, Relation + Value…
Example	<Ballast tube, outside diameter, 50 mm>

is introduced to evaluate the effectiveness of the mining rules. Considering the complexity of empirical case text elements and the diversity of attributes, the simplification of redundant attributes can effectively improve the reliability of mining rules. In this paper, we use the rough set method to refine the set of case knowledge attributes from same-topic cases to obtain several decision rules, which are represented by a generative system of IF-THEN. The expression is as follows.

IF condition holds **THEN** conclusion holds

3 Experiment and Analysis

3.1 Data Source

The empirical instance text data mainly comes from Chinese discussion posts in an internal forum of a large shipbuilding company in Shanghai. The data set includes 310 posts on shipbuilding repair, marine engineering technology, and ship software technology.

3.2 Experience Instance Structured

The dataset is imported into the trained ERNIE model for text classification and attribute-value relationship extraction respectively. In this section, the version of the pre-trained model used is ERNIE-1.0 Chinese, and the word vector is set to 768 dimensions. The semi-structured and structured empirical examples can be obtained in this way in turn, as in Fig. 3 and 4.

The transition to semi-structured posts in Fig. 3 shows that our approach effectively classifies individual sentences according to the structured framework and removes irrelevant sentences. Figure 4 shows that the structured instances formed after attribute-value relation extraction retain only the most core entity words in the statement. The experiments in this chapter demonstrate that our approach can effectively transform everyday textual data in the ship design domain into efficient structured empirical instances for subsequent processing.

NO.	Speaker	Content
1	Fater	[hull] About the plate replacement of thin plate Would like to ask you next master, how to deal with deck board? 6MM thick. He used to have one side that was bonded. Is it all right to lap now? This plate can not be added at will, all the plate seams have a special position. How to deal with it?
2	black	That's weird. What kind of structure do you have? You better get a picture of it.
3	zdbzhang	Better send a more comprehensive picture to see.
4	Fater	The CAR DECK of a car boat.
5	xuehu	It depends on the structure: whether the original lap is flat or inclined deck. You're not talking about where the ramp is, are you?
6	yzhang	Hey, hey! You're the right person to ask. I'm sure a regular boat won't do it. Car carriers are a possibility. Some of the decks of car boats do not bear the total load, but only bear the partial load (the weight of the car). These decks are particularly thin, and the sides are not welded to the main hull, or are fixed by spot welding. Only a ring around the deck is attached to the structure and acts as a side girder. In this case, lap bonding is also possible. Specific basis to see the structure of the calculation book. But if the shipowners and ship inspectors do not understand this structure, they will certainly not agree with you. There's nothing I can do about it.
7	hhll	Learn.

Design Goal	how to deal with deck board?
Design Constraint	6MM thick.He used to have one side that was bonded. Is it all right to lap now? This plate can not be added at will, all the plate seams have a special position. The CAR DECK of a car boat.
Design Principle	Specific basis to see the structure of the calculation book.
Design Step	Some of the decks of car boats do not bear the total load, but only bear the partial load (the weight of the car). These decks are particularly thin, and the sides are not welded to the main hull, or are fixed by spot welding. Only a ring around the deck is attached to the structure and acts as a side girder. In this case, lap bonding is also possible.
Design Taboo	A regular boat won't do it. But if the shipowners and ship inspectors do not understand this structure, they will certainly not agree with you.
Evaluation Result	Learn.

Fig. 3. A role recognition example

Design Goal	[<deck>[deal with]]<changing>
Design Constraint	<thick>[6][MM];[<side>[one]]<bonded>;<plate><not be added><plate seams>;[<plate seams>[all][special position]];<car boat><CAR DECK>
Design Principle	<the structure of the calculation book>
Design step	<car boats><decks><not bear>[the total load]<bear>[the partial load];[<decks>[thin][particularly]][<to><the main hull><not welded><or><fixed by spot welding>];[<deck>[a ring around]][<to><the structure><attached>]<side girder>[act];<lap bonding>[possible]
Design Taboo	<shipowners and ship inspectors>[<boat>[regular]][won't do];<shipowners and ship inspectors><not understand><structure><not agree>
Evaluation Result	<weakly positive>

Fig. 4. Structured empirical instance

3.3 Rules Mining

In this section, we take the problem of "whether the pipe is galvanized" as an example, and get a total of 13 relevant examples by clustering, and extract the attribute entities from the examples to get a total of 8 conditional attributes such as common pipe and pipe diameter, and the decision attribute is whether the pipe is galvanized. The set of rough sets is shown in Fig. 5.

Examples of experience	Main Body	Tubing	Pipe Diameter	Installation Location	Part	Accessories	Ship type	Circulation medium	Whether galvanized
X1	Ballast tube	Grey cast iron pipe	<DN200	None	Inner surface	None	Bulk Carrier	Seawater	Yes
X2	Ballast tube	Grey cast iron pipe	DN100	Ballast tanks	Surface	Flange connection	Bulk Carrier	Seawater	No
X3	Ballast tube	Grey cast iron pipe	DN150	None	Surface	Expansion joint	Ore Ships	Seawater	Yes
X4	Ballast tube	Glass fiber reinforced plastic tube	DN50	Ballast tanks	None	None	Bulk Carrier	Seawater	No
X5	Ballast tube	Grey cast iron pipe	DN150	Ballast tanks	None	None	Oil tanker	Seawater	Yes
X6	Oxygen acetylene tube	Stainless Steel Tubes	DN15	Machine shop	Inner and outer walls	None	Bulk Carrier	Oxygen	No
X7	Miscellaneous Air Hose	Stainless Steel Tubes	DN20	Pneumatic control circuit	Inner and outer walls	None	Bulk Carrier	Air	No
X8	Hot water pipes	PVC pipe	DN20	Other	Inner and outer walls	None	Bulk Carrier	Fresh Water	No
X9	Measuring tube	Stainless Steel Tubes	DN15	Fresh water tank	Surface	None	Oil tanker	Fresh Water	No
X10	Oil pipe	Purple copper tube	DN100	Bunker	Outer Wall	None	Bulk Carrier	Diesel	No
X11	Oil pipe	Purple copper tube	DN100	Water chamber	Outer Wall	None	Bulk Carrier	Diesel	Yes
X12	Oil pipe	Purple copper tube	DN100	Water chamber	Inner Wall	None	Bulk Carrier	Diesel	No
X13	Breathable tube	Stainless Steel Tubes	DN10	Bunker	Inner and outer walls	None	Bulk Carrier	Air	No

Fig. 5. Rough set of “whether the pipe is galvanized”

To obtain accurate and error-free empirical rules, we introduce two indicators of confidence (CF) and coverage (CV), and set criteria from a systematic and objective perspective: $CF(A \rightarrow B) \geq 0.85$, $CV(A \rightarrow B) \geq 0.3$

Finally, two valid knowledge rules are obtained.

– IF pipe material = “grey cast iron pipe” AND installation location = “other” THEN galvanized
– IF pipe = “stainless steel pipe” THEN not galvanized

4 Conclusions

In this paper, we propose an empirical instance knowledge mining method applicable to the ship design field to transform unstructured text into usable knowledge rules in response to the problems of low utilization of text-based instances and complex design language structure in the ship design field. Through the experiments in the self-built ship design dataset, the feasibility of the method and the correctness of the obtained knowledge are proved, which can effectively serve the downstream ship design activities such as recommendation and retrieval, and improve the intelligent design of shipyards.

References

1. H Z Yang, J F Chen, N Ma, D Y Wang. Implementation of knowledge-based engineering methodology in ship structural design[J]. Computer-Aided Design, 2012, Computer-Aided Design, 2012, 44(3):196–202.

2. Y L Wang, Y B Jiang, G Guan, J P Xing, G L Yu. 3D layout design of ship cabin equipment based on knowledge engineering[J]. Journal of Shanghai Jiaotong University, 2021, 55(10): 1219–1227. DOI: https://doi.org/10.16183/j.cnki.jsjtu.2020.136.
3. K Zhang. Design of ship life-saving system arrangement based on knowledge engineering [J]. Ship Standardization Engineer, 2021, 54(02): 91–95. DOI: https://doi.org/10.14141/j.31-1981.2021.02.018.
4. Y L Wang, Z P Wu, G Guan, K Li, S H Chai. Research on intelligent design method of ship multi-deck compartment layout based on improved taboo search genetic algorithm[J]. Ocean Engineering, 2021(225):108823.
5. J Liu, W J Liu, C B Liu. Inverse data acquisition method of curve-like graphs in ship shaft system design specification[J]. Ship and Sea Engineering, 2017, 46(04):140–143.
6. Q Zheng, H C Chang, B W Feng, Z Y Liu, C S Zhan. Research on knowledge-extraction technology in optimization of ship-resistance performance[J]. Ocean Engineering, 2019(179):325–336.
7. C.Q Feng, C G Liu. A rough set knowledge modeling-based design method for ship welding production[J]. Shipbuilding Technology, 2015(01):58–62+78.
8. C Y Wang, Z H Jiang, Y J Ji, F H Wang, Y W Huang, M Xue. An automatic construction method of empirical knowledge graph for ship outfitting design[J]. Mechanical Design and Research, 2021, 37(04):163–169+181. DOI: https://doi.org/10.13952/j.cnki.jofmdr.2021.0158.