Base types selection of PSS based on a priori algorithm and knowledge-based ANN

Zaifang Zhang\(^1\), Nana Chai\(^1\), Yuan Liu\(^1\), Beixin Xia\(^2\)

\(^1\)School of Mechatronic Engineering and Automation, Shanghai University, Shanghai, People's Republic of China
\(^2\)School of Management, Shanghai University, 99 Shangda Road, Shanghai 200444, People's Republic of China
E-mail: bxxia@shu.edu.cn

Abstract: Manufacturers tend to bundle a product with its related services as a product service system (PSS), to create more values for customers and gain competitive advantages for themselves. Configuration design is the key process of PSS development. Configuring a PSS involves selecting and combining appropriate product and service components, to satisfy individual customer requirements. This study studies the mapping relationship between customer requirement attributes and PSS base types in PSS for CNC machine tools, which provides a great reference value for engineers in configuration design. Owing to the high complexity and non-linearity between customer requirement and product, an integrated intelligent learning method based on a priori algorithm and knowledge-based artificial neural network (ANN) is proposed in this study. First, the data of historical configuration instance data sets are processed and then a priori algorithm is used to extract the effective rules as domain knowledge. Domain knowledge is used to build the initial structure of ANN. Moreover, data sets are used to further optimize the network. The knowledge-based ANN is used to realize the mapping between customer requirement attributes and PSS base types. The proposed method is validated in the selection of the PSS base type for CNC machine tools.

1 Introduction

In response to homogeneous competition, manufacturing enterprises achieve the differentiation of operation by attaching services to products in the customer-centric business pattern. In industrial production, manufacturing refers to the process of producing a product and the final product produced. Broadly speaking, the design is a preliminary stage in the whole process of manufacturing a product. Narrowly speaking, the design is the conception and drawing of a product. Moreover, manufacturing refers to the actual production activities according to the design drawings, which is the linkage between people and things (equipment, raw material power etc.), and finally, the production of the product. A product service system (PSS) is an integrated combination of products and services [1]. The design of PSS is a frontier problem in the current manufacturing field. It is designed to build a product and service portfolio that can meet the individual needs of customers to form a holistic solution [2]. This paper studies the mapping relationship between customer requirement attributes and PSS base types in PSS for the computer numerical control (CNC) machine tool. The PSS base-type learning model is established and the base-type selection algorithm model based on a priori and knowledge-based artificial neural network (ANN) integration is constructed. The method is applied to solve the problem of selecting PSS base types for the CNC machine tool. PSS configuration is a design method and a manufacturing process. It is difficult to effectively implement configuration activities based solely on the experience of engineers, and new configuration methods must be studied [3].

PSS configuration refers to the process of converting customer requirements into technical characteristics in highly complex and non-linear multidimensional mapping space. The design solution spaces of the PSS are complex; the main reason is the diversification of products and services in the PSS. Realising the mapping is difficult from customer requirements to whole solutions of the PSS by using the traditional configuration design method [4]. This paper mainly studies the PSS base-type selection. The base-type selection is the mapping of customer requirements to PSS base types and plays an important role in the early design phase [5, 6]. The base types of PSS for CNC machine tools are the products of the design space division. The base types of CNC machine tools can be obtained through customer requirements. The design scope of scheme configuration can be reduced to the design space expressed by base types. Moreover, the configuration scheme can be obtained by instantiating the relevant features in the base types. The selection of base types is a typical multi-input and multi-output problem with customer requirement attributes as input and the PSS base types as output. There are complex and non-linear relationships between customer requirement attributes and base types of the CNC machine tool. Moreover, this non-linear relationship is difficult to express with the exact mathematical model. Neurones are widely interconnected and run in parallel in neural networks, making the entire network highly non-linear [7]. Neural network based on system input and output sample study can approximate any complex and non-linear mapping with arbitrary accuracy [8, 9]. These advantages of neural networks can be used as a general mathematical model of multidimensional and non-linear relationships. Therefore, the non-linear relationship model is established between customer requirement attributes and base types for the CNC machine tool through a neural network in this paper.

The traditional neural network can use the try–error method to get high-performance network structure through time-consuming and complex experiments [10]. Compared to the traditional neural network, knowledge-based ANNs can use domain knowledge to construct relatively reasonable network structure, usually with higher fitting speed and better performance [11, 12]. Simsek applied knowledge-based modelling to the engineering modelling of reconfigurable five finger microstrip patch antennas. Knowledge-based ANNs were used to obtain more accurate results and required less time consumption and even less training data through the coarse model efficiency [13]. Simsek also developed a three-step modelling strategy that exploits knowledge-based techniques to improve some properties of conventional ANN modelling such as accuracy and data requirement [14]. Yu developed a hybrid learning-based model for online intelligent monitoring and diagnosis of the manufacturing processes. The integration of GARule and knowledge-based artificial neural network (KBANN) provided additional operational guidelines for operators or engineers to search for the assignable causes and to
adjust process parameters to bring the out-of-control process back to the normal state [15]. In this paper, the knowledge-based ANN is used to fit non-linear relationships between customer requirement attributes and base types for the CNC machine tool, so as to improve the accuracy of base-type selection for the CNC machine tool.

Knowledge discovery in database (KDD) is a non-trivial process of identifying effective, novel, potentially useful, and ultimately understandable patterns from data sets [16]. The goal of KDD is to discover unknown, useful, and concise patterns. KDD is a cross-cutting field that has attracted researchers in related fields such as machine learning, pattern selection, databases, statistics, artificial intelligence (AI), and expert systems. In fact, data mining is a key step in KDD, which is a process of extracting information and knowledge from a large number of incomplete, noisy, fuzzy, and random data. This information and knowledge is implicit and not known in advance, but potentially useful [17]. The rule mining algorithm is used to mine the historical data sets, and domain knowledge can be constructed to initialise the knowledge-based ANN model. However, the rapid changes in the market, the old domain knowledge often cannot accurately reflect the current market situation. The rule mining algorithm is used to regularly excavate the new data sets. Moreover, it is easy to find new domain knowledge and then supplement and update the domain knowledge. Rule mining algorithms facilitate rapid response to market changes, making customers time sensitive to PSS base types mapping requirements. A priori algorithm is a more classic algorithm in rule mining [18, 19]. Therefore, this paper uses the a priori algorithm to extract domain knowledge.

In view of the above analysis, this paper constructs an integrated intelligent learning method based on the a priori algorithm and knowledge-based ANN. In this method, first, some examples are extracted from the historical design case as the database. Moreover, the association rule mining a priori algorithm is appropriately improved and applied to the classification rule mining between customer requirement attributes and base types for the CNC machine tool. The meaningful classification rules are extracted from the database to form domain knowledge. Then, the knowledge-based ANN can optimise the setting of the network structure by using domain knowledge. At the same time, it further learns the knowledge that is not fully expressed by domain knowledge from the database, and refines the network structure. Finally, the integrated intelligent learning method can establish complex and non-linear mapping relationships model between customer requirement attributes and base types for the CNC machine tool. The CNC machine tool base types can be selected. Moreover, the mapping of customer requirement attributes to the CNC machine tool base types can be realised.

The base type of PSS can be obtained directly from customer requirement attributes through the intelligent mapping method in this paper. The intelligent algorithm is used to realise the intellectualisation of design and manufacturing. The design is the preparation and prerequisite for the manufacturing process of the enterprise. Through the intelligent design of PSS configuration to provide engineers with reference, it can better and faster meet customer requirements. Moreover, it brings advantages to enterprises in the homogeneous competition. New-generation intelligent manufacturing represents an in-depth integration of new-generation AI technology and advanced manufacturing technology. It runs through every link in the full life cycle of design, production, product, and service. The design is the most important link to various product innovations. Intelligently optimised design, intelligently collaborated design, user-interactive intelligent customisation, and mass creation based on crowd intelligence are all major components of intelligent design [20]. Intelligent manufacturing systems should have the capabilities of flexibility, adaptability, and intelligence. These capabilities will require the control action to be distributed and integrated with different approaches including smart sensing, optimal design, and intelligent learning [21, 22]. The intelligent configuration method proposed in this paper is an important part of the intelligent design in intelligent manufacturing, which provides a new idea for an intelligent design. Cloud manufacturing (CM) is an emerging service-oriented business model to share manufacturing capabilities and resources on a cloud platform. Four key methods of how CM increases sustainability have been identified: collaborative design, computer automation, increase its resilience, and enhanced waste reduction, reuse, and recovery [23, 24]. Resource service matching is one of the CM's key issues. The improved genetic algorithms is proposed to search the best matching result for the request of customers, which assure the total cost and time of all the tasks are the lowest [25]. Manufacturing service configuration plays an important role in implementing CM [26]. In CM, the selected product attributes, service attributes, and sample results can be input into the proposed network using a priori algorithm and knowledge-based ANN method for learning. The learning network can be applied to the new customer requirement attributes, and the corresponding base type of PSS can be obtained. CM enables in-depth customisation [27]. The method may still have good performance in capability selection in CM with small data volume. However, as the amount of data increases, the network will become more complex and the ability of self-learning will be reduced. In the era of big data, deep learning networks can have better performance in CM.

The rests of this paper are organised as follows: Section 2 presents the base-type learning model of PSS. Knowledge discovery based on a priori is described in Section 3. Section 4 introduces the selection of PSS base types using a knowledge-based ANN. An application of the proposed method in the CNC machine of PSS is taken as an illustrative example in Section 5. Finally, the main contribution of this paper is summarised and future work is discussed in Section 6.

## 2 Base-type learning model of PSS

The object of this paper is the selection process of the PSS base types: there is a strong connection between the input parameters of the system (i.e. customer requirement attributes) and the output of the system (i.e. the PSS base types). As shown in Fig. 1, the mapping relationships between the customer requirement attributes and the output PSS base types are usually complex and non-linear. Moreover, it is very difficult to solve some complicated relations by mathematical methods such as linear regression. Non-linear relationship modelling is the core of PSS base-type selection.

This non-linear mapping is expressed as follows:

\[
Y = f(X_1, X_2, ..., X_n)
\]  

(1)

In the formula, \(X_i\) is the input variable of the system (i.e. customer requirement attributes) and \(Y\) is the output variable of the system (i.e. various base types of output PSS).

ANNs simulate the complexity and non-linear relationship between input space and output space [28]. However, it cannot provide effective and accurate for PSS base types. On the other hand, the a priori algorithm can extract useful rules to solve this problem in PSS base-type selection. Therefore, considering these
critical needs for computational efficiency and effectiveness, an intelligent learning model is proposed to implement PSS base-type selection by using the a priori algorithm and knowledge-based ANN. PSS base-type learning model is as shown in Fig. 2.

Knowledge-based ANN uses the training sets to learn the complex mapping relationships between customer requirement attribute parameters and PSS base types. Moreover, the output results are obtained by output neuron values, i.e. the PSS base types. Data mining a priori algorithm is used to extract explicit algorithm and knowledge-based ANN integrated intelligent results are obtained by output neurone values, i.e. the PSS base interval is mapped into a Boolean attribute.

However, because the historical configuration example contains a lot of numerical attributes, the values of these attributes in the database, the association rules are divided into quantitative association rules. Among them, the Boolean attributes in the database contains many data types, First, it is necessary to use the a priori algorithm and knowledge-based fuzzy C-means (FCM) algorithm to classify the data. Then, the fuzzy clustering, and partition matrix and the centre of C classes so that the given objective function [formula (2)] is minimal

\[ J = \sum_{i=1}^{c} \sum_{j=1}^{n} (u_{ij})^m d(x_j, v_i) \]  

\[ d(x_j, v_i) = \sqrt{\sum_{k=1}^{d} (x_{jk} - v_{ik})^2} \]  

\[ M_{c} = \{ u \in R^{n} | 1 \leq j \leq n, 1 \leq i \leq c \} \]

In formula (2), \( u_{ij} \) is the degree of the target data \( x_j \) belonging to the \( i \)th class, and satisfies 0 \( \leq u_{ij} \leq 1 \) and \( u_{ij} + u_{ik} + \ldots + u_{i(k-1)} = 1 \). Here, \( m \) is the weighting index, used to control the fuzzy degree of fuzzy clustering, and \( m \in [1, \infty) \). In formula (4), \( U \) is a membership matrix, \( R \) is a real field, \( R^m \) is a collection of \( c \) rows and \( n \) columns on all real fields.

The iterative process of the FCM algorithm is as follows (dividing the target data sets into C classes):

**Step (1):** Set the number of categories parameter \( c \), 2 \( \leq c \leq n \), take the \( m \), initialise the membership matrix \( U^{(0)} \in M_{c}\), and set the number of cycles \( s = 0 \), 1, 2, ….  

**Step (2):** Use \( U = U^{(0)} \) to calculate the vectors \( v_i = v_i^s \), where  

\[ v_i = \frac{\sum_{j=1}^{n} (u_{ij})^m x_j}{\sum_{j=1}^{n} (u_{ij})^m} \]  

**Step (3):** Modify \( U = U^{(s+1)} \in M_{c} \). Remember \( d_{ij} = d(x_j, v_i) \), for each fixed \( i \), if for all the samples \( 1 \leq j \leq n \), there is \( d_{ij} > 0 \), then  

\[ u_{ij} = \frac{1}{\sum_{i=1}^{c} d_{ij}(d_{ij})^{m-1}} \]  

**Step (4):** Take the appropriate matrix norm \( \parallel \parallel \), take \( c \), \( \varepsilon \) is any small real number, if \( \parallel U^{(s+1)} \parallel \leq \varepsilon \), then stop the cycle; otherwise, set \( s = s + 1 \), and return to step (2).

### 3 Knowledge discovery based on a priori

Association rule discovery is one of the most important tasks in data mining. Its goal is to discover all the frequent patterns and strong association rules in the data sets. According to the types of attributes in the database, the association rules are divided into Boolean association rules, category association rules, and quantitative association rules. Among them, the Boolean attributes refer to the value of only 0 and 1 attributes. The a priori algorithm is the most classical Boolean association rule mining algorithm proposed by Agrawal [16]. The data type used in this paper is a most typical and widely used algorithms. On the one hand, the use of fuzzy sets can better achieve the conversion between continuous and discrete amounts, and finally, soften the dividing boundary of the attribute domain. On the other hand, the FCM algorithm can effectively reflect the actual distribution of the data in the processing of high bias data.

Assume that the target data set is \( X = \{ x_1, x_2, \ldots, x_n \} \), where \( x_i = \{ x_{i0}, x_{i1}, \ldots, x_{in} \} \), \( i = 1, 2, \ldots, n \). The basic idea that the FCM algorithm gathers the target data into class C is to find a fuzzy partition matrix and the centre of C classes so that the given objective function [formula (2)] is minimal

3.1 Interval partitioning based on FCM algorithm

Since the data in the history configuration instance database contains a lot of numerical attributes, the values of these attributes are relatively wide, not just 0 and 1. It is necessary to convert the numerical attribute into a Boolean attribute, which is divided into two cases: (i) when the full value of the numerical attribute is a few, only each attribute is mapped into a Boolean attribute and (ii) when the range of numerical attributes is wide, the range of these attributes must be divided into several intervals, and then each interval is mapped into a Boolean attribute.

For type attributes, according to the first transformation method described above, the one-to-one mapping can be made to Boolean attributes. The value of the category attribute in each set of data is expressed by its maximum boundary value, which is converted to numerical type. For the above-mentioned second case, the clustering algorithm is generally used. Since the clustering algorithm can divide the value of the numerical attribute into several intervals, which can effectively reflect the actual distribution of the data. FCM clustering algorithm is one of the most typical and widely used algorithms. On the one hand, the use of fuzzy sets can better achieve the conversion between continuous and discrete amounts, and finally, soften the dividing boundary of the attribute domain. On the other hand, the FCM algorithm can effectively reflect the actual distribution of the data in the processing of high bias data.
mining are used as evaluation indicators. The greater the degree of support and confidence, the higher the frequency of the rule in the data sets, and the better the quality of the rule.

\[
\text{Support}(R) = \frac{\text{SUP}(X \& Y)}{|D|}
\]

(7)

\[
\text{Confidence}(R) = \frac{\text{SUP}(X \& Y)}{\text{SUP}(X)}
\]

(8)

In formulas (7) and (8), SUP \( X \& Y \) represents the number of instances in the database \( D \) that satisfy the rules \( X \) and \( Y \) related conditions, i.e. the number of \( X \) \& \( Y \) support, \( \text{SUP}(X) \) represents the number of instances in the database \( D \) that satisfies the rules \( X \) related conditions, \(|D|\) represents the total number of instances in database \( D \).

3.2.2 A priori-based classification rule mining algorithm: The a priori algorithm was originally applied to association rule mining, which decomposes association rule mining into two sub-problems: (i) find all the frequent item sets that exist in the transaction database, i.e. the project set whose support is greater than the support threshold (i.e. the minimum support, \( \text{minsup} \)). (ii) Generate strong association rules based on the found frequent item sets, i.e. the association rules that produce support and confidence, respectively, greater than the support threshold and the confidence threshold (i.e. \( \text{minimum confidence, minconf} \)). The pseudocode of the algorithm is shown in Fig. 3.

This paper, the mapping rules from customer requirement attribute to base types for the CNC machine tool are classified rules. When using a priori for rule mining, the mining rules need to be screened. The following questions need to be considered in the screening: (i) The attributes contained in the preceding item \( X \) must be customer requirement attributes and cannot contain base types of the CNC machine tool. (ii) The rules of \( Y \) can only contain base types of the CNC machine tool, and cannot contain customer requirement attributes. (iii) The requested category item cannot be empty, i.e. the rules ‘\( X \Rightarrow Y \)’, the latter part of \( Y \neq \emptyset \). After the screening, removing the rules that do not satisfy the three conditions, the rules are obtained as required classification rules.

According to the above description, after extracting the rules from the data sets, it is necessary to decode the obtained classification rules represented by Boolean type. At the same time, because the rules are often redundant, the extracted rules need to be reduced and summarised. When the two rules conflict, delete all two of these rules. When there is a relationship between the preceding items of the two classification rules, the included rule is deleted. For example, rule 1: \( 1 < x_1 < 2 \Rightarrow y_1 \) and rule 2: \( 1 < x_1 < 3 \Rightarrow y_1 \); it is clear that interval \([1, 3]\) contains interval \([1, 2]\), so rule 1 is deleted, while rule 2 is retained. When only one item in the preceding attribute of the two classification rules is different, the two rules can be grouped into a rule. For example, rule 3: \( x_2 = 2 \Rightarrow y_2 \) and rule 4: \( x_3 > 2 \Rightarrow y_2 \). Then, a new rule can be obtained by induction: \( (x_2 = 2) \) or \( (x_3 > 2) \Rightarrow y_2 \).

4 Selection of PSS base types using a knowledge-based ANN

The knowledge-based ANN is a method of constructing network topology and giving initial connection weights on the basis of acquired expert domain knowledge. In the previous section, the a priori-based classification rule mining algorithm has extracted the domain knowledge of the relationships between customer requirement attribute parameters and base types for the CNC machine tool. The domain knowledge can be effectively used for the construction of a knowledge-based ANN model, and effectively modelled the relationships between customer requirement attribute parameters and base types for the CNC machine tool. Therefore,
the a priori algorithm and knowledge-based ANN algorithm can be seamlessly integrated, as shown in Fig. 4.

The back-propagation (BP) neural network was proposed by scientists led by Rumelhart and McClelland in 1986. It is a multilayer feed-forward neural network trained by the error BP algorithm. BP is a typical neural network model because of its excellent ability in arbitrary complex pattern classification and multidimensional function mapping, as well as its mature training method [29, 30]. The BP neural network algorithm repeats two phases: forward propagation and BP. In the forward propagation phase, the input signals are propagated forward from the input layer to the hidden layer for processing, and then to the output layer. The output is then compared with the expected value and the difference between the expected value and the output is referred to as the error. In the BP phase, the loss function is formulated by using the error. Moreover, the weight of the neural network is updated by the optimisation method to minimise the loss function [31, 32]. The BP neural network algorithm systematically solves the learning problem of connection weight of the hidden layer in the multilayer neural network [32].

Knowledge-based ANN learning includes the following two steps.

The first step is to transform rules into network topologies. In fact, the approximate and correct domain knowledge (rules) is transformed into the initial network topology of the neural network. The domain knowledge (rules) comes from the rules of a priori knowledge discovery. The initial network topology of this step is based on the mapping method proposed by Towell and Shavlik [33] and Osório and Amy [34]. In this way, compared to the conventional neural network algorithm, the learning of neural networks with an initial topology is no longer zero based, but on the basis of certain domain knowledge. The method can shorten the training time of the neural network and improve learning efficiency.

The second step is to use the BP learning algorithm [35], using the training data sets extracted from the historical configuration example to train the network constructed in the first step. In this step of the study, the weights of the neural network are further refined and revised, thus mapping accuracy of customer requirement attributes to base types for the CNC machine tool is improved.

### 4.1 Customer requirement attributes data processing

The mapping between customer requirement attributes and base types for the CNC machine tool is established by using a knowledge-based ANN. The inputs of the knowledge-based ANN are the customer requirement attributes in the CNC machine tool, and base types of the CNC machine tool corresponding to the customer requirement attributes are outputs. The corresponding knowledge of the initial topology of the neural network is the domain knowledge (rules).

In general, when using neural networks for type selection, the choice of input attributes has a large impact on their selection performance [36]. Through the market survey and the analysis of historical data, the core requirements of customers are extracted. The core requirements of each customer are represented by requirement attributes \( CR_i \), \( i \in \{1, 2, ..., N\} \), \( N \) represents the number of customer requirement attributes, and for any requirement attribute \( CR_i \), the value is expressed by \( CR_{ij} \), \( j \in \{1, 2, ..., m\} \), where \( m \) represents the number of possible values of the attribute \( CR_i \). Then, the customer requirement attributes can be expressed by a set of requirement vectors \( \{CR_{ij}\}, \forall i \in \{1, 2, ..., N\}, \forall j \in \{1, 2, ..., m\} \), and the customer requirement attribute vector is used as the knowledge-based ANN input vector.

In the customer requirement attribute vector, each attribute \( CR_i \) may be Boolean and may be numerical, textual, or in a numerical range, i.e. category type. Since knowledge-based ANNs cannot recognise so many types of data, various types of customer requirement attributes need to be preprocessed to a form of input data acceptable to knowledge-based ANNs.

Boolean attribute refers to the value of only 0 and 1 attributes. For example, whether the sunroof of the car is available can be 0 to indicate that there is no sunroof, and 1 to indicate that there is a sunroof. Numerical attribute refers to the use of a discrete real number to express the attributes such as the power of the hair dryer, expressed in numerical attributes, i.e. 1000, 1200, and 1500 W. Type attribute refers to the attribute values in the form of text expression. This range of attribute value is generally defined in advance such as the colour of a laptop, and its value can be selected from the white, red, black, and blue. The category attribute whose value is not a value, but a range such as the life of energy-saving lamps, the value of collection can be \([20,000–30,000, 45,000–70,000, 40,000–80,000, 50,000–90,000]\) (unit/h).

For these different types of customer requirement attributes, it is necessary to preprocess data to convert it into knowledge-based ANN acceptable input form. Different types of attributes have different preprocessing methods, as follows:

(i) For a Boolean attribute, two input neurones are assigned to them. If 0 is taken, the corresponding input neurone is 1 0, and if it is 1, the corresponding input neurone is 0 1.

(ii) For a numerical attribute, each numerical customer requirement attribute corresponds to an input neurone of the knowledge-based ANN. Before input, the domain value of attribute needs to be normalised, as shown in formula (9). After the normalisation process, the range of the attribute will be mapped to \([0, 1]\) range.

\[
CR_i' = \frac{CR_i - \min_{j \in J} \{CR_j\}}{\max_{j \in J} \{CR_j\} - \min_{j \in J} \{CR_j\}}
\]

In formula (9), \( CR_i' \) represents the input neurone of the customer requirement attribute \( CR_i \), \( J \) represents the number of options for the customer requirement attributes.

(iii) For the type attribute, each attribute value corresponds to an input neurone. For example, the colour of the laptop is corresponding to the four input neurones, if the customer chooses ‘red’, the values of the four input neurones are 0 1 0 0, respectively.

(iv) For the category attribute, take the maximum of the interval and normalise to the \([0, 1]\) interval, and then obtain the set of boundary values according to the arrangement from small to large. Each value in the set of boundary values corresponds to an input neurone. For example, the life of the energy-saving lamp, whose boundary value set is its value set, can be \([0, 0.444, 0.556, 1]\), corresponding to the four input neurones.

As described above, the customer requirement attribute and its correspondence with the knowledge-based ANN input neurones are shown in Table 1.

### 4.2 Knowledge transformation and initial structure of the knowledge-based ANN

Domain knowledge expresses the relationships between customer requirement attributes and base types for the CNC machine tool. Converting it into a knowledge-based ANN structure is a very critical step in the knowledge-based ANN algorithm. Since the a priori algorithm has extracted domain knowledge (i.e. mapping rules), the knowledge-based ANN can use this domain knowledge...
to implement rules as network structure transformation steps. In this process, the preamble of the rule is transformed into the input unit of the neural network. The consequent of the rule is transformed into the output unit of the neural network. Moreover, the intermediate conclusion of the rule is transformed into the hidden layer of the neural network. Correspondingly, the correlation among the front rule, the back rule, and the intermediate conclusion are converted into the weight of each node of the neural network, as shown in Fig. 5.

Summarise the transformation methods proposed by Towell and Osorio from the rules to the network structure, and get seven main steps of network structure initialisation-based domain knowledge, as shown in Fig. 6:

(1) **For zero-order production rules**: The method proposed by Towell and Shavlik [33] is only directed toward zero-order production rules, and is transformed into a network structure.

① **Take the rules**: Take rule If A and B and Not(C) and Not(D) then X as an example, where A, B, C, and D are binary inputs, X is the output, and the rule is transformed into the neural network structure as shown in Fig. 7a.

② **Extraction of the rules**: For the rule: If A or B or C or D then X, the rule is transformed into the neural network structure method as shown in Fig. 7b.

(2) **Production rules for zero order over (0+)**: Osório and Amy [34] extended the types of rules that knowledge-based ANN can use rules of zero-order production rules, i.e. rules with real numbers as input features.

**Rule**: If Greater Than (A, a) Then X, If Less Than (B, b) Then X, where A and B are input characteristics and X is the corresponding conclusion. The process of mapping such rules into a network structure is shown in Fig. 8.

If the rule is very complex, it can be transformed into several simple rule forms, then use the above methods to transform the network structure.

On the basis of these transformation methods, the data invisible knowledge contained in the rule is mapped into the topology, initial weights, and biases of the neural network. Therefore, before the network starts training, an optimised network structure is provided, which avoids a complex process of obtaining the relatively appropriate neural network structure. At the same time, the initial weights and biases of the optimisation, not only improve the performance of the network, but also accelerate the fitting process, and save training time.

### Table 1 Customer requirement attributes and their correspondence with knowledge-based ANN input neurones

<table>
<thead>
<tr>
<th>Customer requirement attribute</th>
<th>Attribute type</th>
<th>Range of neurones (or set of values)</th>
<th>Corresponding input neurones values</th>
</tr>
</thead>
<tbody>
<tr>
<td>CR1</td>
<td>numerical type</td>
<td>CRV,1,2,3,4</td>
<td>1 normalised value [0, 1] interval</td>
</tr>
<tr>
<td>CR2</td>
<td>description type</td>
<td>CRV,2,3,4</td>
<td>0, 1 (value CRV,2,2)</td>
</tr>
<tr>
<td>CR3</td>
<td>Boolean type</td>
<td>CRV,3,4</td>
<td>0, 1</td>
</tr>
<tr>
<td>CR4</td>
<td>category type</td>
<td>CRV,4,1,2,3,4</td>
<td>0, 0, 1, 0 (the target values are normalised in the interval [CRV,2,2, CRV,4,3])</td>
</tr>
</tbody>
</table>

### 4.3 Knowledge-based ANN training

PSS base-type design space is: $\Omega_{nS} = \{PSSM_1, PSSM_2, PSSM_H\}$. Each output neurone of the knowledge-based ANN corresponds to a PSS base type, and each output neurone has a value of 0 or 1. The knowledge-based ANN has a total of H output neurones. In this paper, H is 3. If the target of the CNC machine tool base type is PSSM_2, the knowledge-based ANN output vector is \{0, 0, 1\}.

After transforming domain knowledge into the initial topology of neural networks, the network is further trained by training data sets extracted from the historical configuration instance library. The training data sets consist of customer requirement attribute – PSS base-type data pairs, expressed as follows:

$$D_{TRAIN} = \{(CRV_i), PSSM_q)\}, \quad q = 1, 2, ..., H$$

On the basis of BP, knowledge-based ANN improves its network structure by adjusting the connection weight. Finally, the
knowledge-based ANN will be fitted into an optimised set of network connection weights. Moreover, the network structure described by the weight set expresses the non-linear relationship between customer requirement attributes and PSS base type. The trained knowledge-based ANN is solidified. Under the new customer requirement attributes, it can be used as the base-type selection of the PSS that adapts to the requirements.

5 Case study

This paper is about the design process of PSS for CNC machine tools. This paper studies how to use the a priori and knowledge-based ANN to realise the mapping between customer requirement attributes, the CNC machine tools base types. Moreover, the network structure described by the weight set expresses the non-linear relationship between customer requirement attributes and CNC machine tools base types. The base-type PSSM is solidified. Under the new customer requirement attributes, it can be used as the base-type selection of the PSS that adapts to the requirements.

The initial data sets have 183 sets of data, as shown in Table 2 (each row represents a set of data). The 129 groups of data were randomly selected as the initial training data sets, the base-type PSSM had 40 sets of data, the base-type PSSM had 47 sets of data, and the base-type PSSM had 42 sets of data. In addition, 54 sets of data were used as initial test data sets including 22 groups of CNC machine tool base types from the historical configuration case, 16 groups of PSSM data, and 24 groups of PSSM data. In this case, the set of optional values for each attribute is: CR1 ∈ {500, 750, …, 7500, 8000}, CR2 ∈ {S, M, L}, CR3 ∈ {10–30, 30–50, 50–70, 70–90}, CR4 ∈ {1000, 2000}, CR5 ∈ {N, P}, CR6 ∈ {S, 5, 6, 7, …, 12}, CR7 ∈ {N, P}, CR8 ∈ {N, P}.

5.1 Knowledge extraction based on a priori

Before using a priori to rule mining data sets, the initial data sets need to be processed. First, the value of the category attribute in each set of data is expressed by its maximum boundary value, which is transformed into a numerical type. Moreover, the numerical attributes in the customer requirement attributes are clustered by FCM algorithm. In this section, the values of the spindle speed (CR1), tool storage capacity (CR3), the machining accuracy (CR5), tool storage capacity (CR3), and the reliability (CR6) are discrete numerical attributes, tool storage capacity (CR3) is a category attribute, and the other four attributes are type attributes. At the same time, the enterprise has designed three base types PSSM (i = 1, 2, 3) to realise configuration activities of the CNC machine tool. Different base types have different topologies. Each base type contains a number of the CNC machine tool technology features. Moreover, the specific CNC machine tool configuration scheme can be obtained by instantiating these technical features.

According to the a priori and knowledge-based ANN described the above statement, the configuration design engineer needs to obtain customer requirement attributes and the corresponding CNC machine tool base types from the historical configuration case, form the initial data sets, and deal with the data accordingly. The data can be used as the training data sets of a priori knowledge extraction and neural network training. Then, the structure of neural network is constructed by using a priori's domain knowledge extracted from data sets (i.e. classification rules). On this basis, the data set is used to train the network, and the trained neural network is used to identify the corresponding PSS base type of the CNC machine tool in the new customer requirement attributes.

### Table 2 Initial data sets

<table>
<thead>
<tr>
<th>No.</th>
<th>CR1</th>
<th>CR2</th>
<th>CR3</th>
<th>CR4</th>
<th>CR5</th>
<th>CR6</th>
<th>CR7</th>
<th>CR8</th>
<th>PSSM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>900</td>
<td>S</td>
<td>10–30</td>
<td>P</td>
<td>11</td>
<td>600</td>
<td>M</td>
<td>N</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>6000</td>
<td>L</td>
<td>50–70</td>
<td>N</td>
<td>8</td>
<td>800</td>
<td>H</td>
<td>P</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>3250</td>
<td>S</td>
<td>30–50</td>
<td>P</td>
<td>10</td>
<td>500</td>
<td>L</td>
<td>N</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>7500</td>
<td>S</td>
<td>10–30</td>
<td>P</td>
<td>8</td>
<td>1200</td>
<td>H</td>
<td>N</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>4500</td>
<td>M</td>
<td>30–50</td>
<td>P</td>
<td>9</td>
<td>800</td>
<td>L</td>
<td>N</td>
<td>2</td>
</tr>
</tbody>
</table>

### Table 3 Boolean training data sets

<table>
<thead>
<tr>
<th>No.</th>
<th>CR1</th>
<th>CR2</th>
<th>CR3</th>
<th>CR4</th>
<th>CR5</th>
<th>CR6</th>
<th>CR7</th>
<th>CR8</th>
<th>PSSM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

This is an open access article published by the IET under the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0/)
Rule R3: IF \([0.70 \leq CR_{r1} < 1) AND (CR_{r2} = 0 0 1) AND (CR_{r4} = 0 1) AND ([CR_{r5} = 0.29) OR (0.74 < CR_{r6} < 1) OR (CR_{r7} = 0 0 1)]\), THEN PSSM = PSSM3.

5.2 Data processing, knowledge-based ANN construction and training

The initial training data sets and the initial test data sets are transformed according to Table 1, and random sorting is carried out to obtain the transformed training data sets and test data sets, as shown in Tables 4 and 5.

Using the rule sets obtained in Section 5.1, the rule sets are mapped to the structure of knowledge-based ANN using to Towell and Osorio's methods, as shown in Fig. 9. The input layer node 1 to the hidden layer node H1 maps the condition CR1 < 0.27 of the rule R1 and the input layer nodes 2–4 to the hidden layer node H2 maps the condition CR2 = 1 0 0 of the rule R1. In this way, the conditions of the rule sets are mapped in turn into the connection between the network layer and layer nodes, 107 connections (indicated by solid lines) are generated, and the initial weights and biases of the network are determined.

After determining the connection relation expressed in the rule sets, the connection relations between nodes and nodes are not expressed. These connections are represented by dotted lines, for each additional dotted line connection, connection weights are set to a random number that approaches 0. Thus, after the completion of the rule transformation, the topology of the knowledge-based ANN is finally determined, i.e. 14-23-10-3-3. In this topology, there is an input layer, three hidden layers, and an output layer. There are 14 nodes in the input layer, the first hidden layer has 23 nodes, the second hidden layer has 10 nodes, the third hidden layer has 3 nodes, and the output layer has 3 nodes.

Knowledge-based ANN training uses the BP algorithm to further learn the knowledge of the training data sets. The three hidden layers and output neurons of the network are, respectively, tansig, logsig, tansig, and purelin functions. Using momentum gradient descent algorithm traingdm, learning rate lr is set to 0.1, unstable in classification, especially in the case of a small sample size.

After 1000 iterations, the final mean square error of knowledge-based ANN is 0.0025 and BP's final mean square error is 0.0075. It can be seen that the knowledge-based ANN has better training results and faster fitting speed than BP networks.

5.3 Results comparison and analysis

The test results that identify the error in the test data sets are removed. Five representative data in the test data sets (such as nos. 1, 8, 31, 40, and 51) are selected optionally to compare the performance of the knowledge-based ANN and the standard BP network in the selection of base type for the CNC machine tools. Table 6 shows the test results of the two networks. It can be seen that the knowledge-based ANN and standard BP networks have the same network structure and associated training parameters. However, the weights and biases between all nodes in the standard BP network are not determined by the domain knowledge and take a random value that approaches 0. It is found that the mean square error of knowledge-based ANN training has little change. Moreover, the mean square error of standard BP network has large change. The result indicates that the standard BP network is unstable in classification, especially in the case of a small sample size.

Table 4 Transformed training data sets

<table>
<thead>
<tr>
<th>No.</th>
<th>CR1</th>
<th>CR2</th>
<th>CR3</th>
<th>CR4</th>
<th>CR5</th>
<th>CR6</th>
<th>CR7</th>
<th>CR8</th>
<th>PSSM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.053</td>
<td>1 0 0</td>
<td>0</td>
<td>1 0</td>
<td>1.000</td>
<td>0.250</td>
<td>1 0 0</td>
<td>1 0</td>
<td>1 0 0</td>
</tr>
<tr>
<td>2</td>
<td>0.867</td>
<td>0 0 1</td>
<td>1.000</td>
<td>0 1</td>
<td>0.286</td>
<td>1.000</td>
<td>1 0 0</td>
<td>0 1</td>
<td>0 0 1</td>
</tr>
<tr>
<td>3</td>
<td>0.413</td>
<td>0 0 1</td>
<td>0.667</td>
<td>1 0</td>
<td>0.429</td>
<td>0.625</td>
<td>0 0 1</td>
<td>1 0</td>
<td>0 1 0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>1 0 0</td>
<td>0 0 1</td>
<td>0.857</td>
<td>0.125</td>
<td>1 0 0</td>
<td>1 0</td>
<td>1 0 0</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>128</td>
<td>0.933</td>
<td>0 1 0</td>
<td>1.000</td>
<td>0 1</td>
<td>0</td>
<td>0.750</td>
<td>0 0 1</td>
<td>0 1</td>
<td>0 0 1</td>
</tr>
<tr>
<td>129</td>
<td>0.627</td>
<td>0 0 1</td>
<td>0.667</td>
<td>1 0</td>
<td>0.429</td>
<td>0.625</td>
<td>0 1 0</td>
<td>1 0</td>
<td>1 0 0</td>
</tr>
</tbody>
</table>

Table 5 Transformed test data sets

<table>
<thead>
<tr>
<th>No.</th>
<th>CR1</th>
<th>CR2</th>
<th>CR3</th>
<th>CR4</th>
<th>CR5</th>
<th>CR6</th>
<th>CR7</th>
<th>CR8</th>
<th>PSSM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.267</td>
<td>1 0 0</td>
<td>0.333</td>
<td>1 0</td>
<td>1.000</td>
<td>0.250</td>
<td>1 0 0</td>
<td>1 0</td>
<td>1 0 0</td>
</tr>
<tr>
<td>2</td>
<td>0.200</td>
<td>1 0 0</td>
<td>0</td>
<td>1 0</td>
<td>0.714</td>
<td>0</td>
<td>1 0 0</td>
<td>1 0</td>
<td>1 0 0</td>
</tr>
<tr>
<td>3</td>
<td>0.467</td>
<td>0 1 0</td>
<td>0.667</td>
<td>1 0</td>
<td>0.571</td>
<td>0.500</td>
<td>0 0 1</td>
<td>1 0</td>
<td>0 1 0</td>
</tr>
<tr>
<td>4</td>
<td>0.733</td>
<td>0 0 1</td>
<td>0 1 0</td>
<td>0.714</td>
<td>0.625</td>
<td>0 0 1</td>
<td>0 1</td>
<td>0 0 1</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>53</td>
<td>0.413</td>
<td>0 1 0</td>
<td>0.667</td>
<td>1 0</td>
<td>0.428</td>
<td>0.500</td>
<td>0 0 1</td>
<td>1 0</td>
<td>0 1 0</td>
</tr>
<tr>
<td>54</td>
<td>0.053</td>
<td>1 0 0</td>
<td>0 1 0</td>
<td>0.571</td>
<td>0</td>
<td>0 0 1</td>
<td>1 0</td>
<td>1 0 0</td>
<td></td>
</tr>
</tbody>
</table>
with the standard BP networks with different structures. The selected standard BP network has the same number of network nodes as a knowledge-based ANN, but the number of layers in the hidden layer is different. The selected BP network hidden layer structure is 23-10-3, 23-13, 33-3, 36, and the training parameters of each network are consistent with the training parameters of the above knowledge-based ANN. The classification selection accuracy as shown in Table 7, knowledge-based ANN can identify 98.45% of the sample base types in the training data set and identify 92.59% of the sample base types in the test data sets. The accuracy of the knowledge-based ANN is much higher than other standard BP networks. From this comparison, it can be seen that knowledge-based ANN has better generalisation performance than standard BP networks, especially in the small sample of this example.

Compared to the standard BP neural network, the initial network structure of knowledge-based ANN is designed by domain knowledge. Especially in the case of small samples, the knowledge-based ANN has a faster fitting speed, shorter training time, and higher generalisation performance due to optimisation of network topology, initial weight and deviation, and further training. In the standard BP network, learning performance is greatly influenced by the network training parameters (such as learning rate, network structure, and the initial weights). It is usually necessary to carry out continuous experiments through try–error to train a better performance BP network, but this test is more complicated and time-consuming. In this paper, the a priori algorithm is used to extract the explicit classification rules (domain knowledge) from the historical configuration examples. Domain knowledge is transformed into the topology of the neural network, i.e. initial weights and biases, avoiding the complex process of continuous testing by try–error. The knowledge-based ANN has better performance in improving the selection accuracy of the CNC machine tool base type.

The trained knowledge-based ANN is solidified and can be used to map new customer requirement attributes to the PSS base types.
types. For example, for a new customer requirement attribute vector (3250, L, 50–70, N, 9, 800, L, N), after data processing and transformation, the customer requirement attribute vector becomes (0.3667, 0 1 0, 0.6667, 1 0, 0.5714, 0.5000, 0 1 0, 1 0). Then, the requirement vector is input to the solidified knowledge-based ANN, after the operation, the output is obtained, i.e. PSS base-type PSSM for CNC machine tools. In this base type, customer requirement attributes can be mapped to the next level.

6 Conclusions
Aiming at the problem of PSS base-type selection, a base-type selection method is proposed based on a priori and knowledge-based ANN integration. The ANN base-type learning model is established and base-type selection algorithm model based on a priori and knowledge-based ANN integration is constructed. First, data preprocessing of the historical configuration instance is transformed into the Boolean data sets that can be processed by the a priori algorithm. Moreover, then the algorithm is used to extract the effective rules as domain knowledge, i.e. the effective relationships between customer requirement attribute parameters and base types for the CNC machine tool. Knowledge-based ANN uses the domain knowledge to build initial network topology, and trains data sets to improve the accuracy of the mapping of customer requirement attributes to PSS base types. This method can be applied to solve the problem of selecting PSS base types of CNC machine tools, which can effectively extract the non-linear mapping relationships between customer requirement attributes and corresponding base types of the CNC machine tool. Through the comparison of the BP algorithm and knowledge-based ANN algorithm, it is proved that knowledge-based ANN has faster type selection and meet the market requirements.

Future work will consider other methods to excavate the relationships between customer requirement attributes and PSSs. A convolution neural network will be used to learn the relationships among customer requirement attributes and the PSSs, which can be used to realise configuration activities of the PSS and provide effective guidance for the product designers.

7 Acknowledgments
This project is supported by the National Natural Science Foundation of China (Nos. 51205242 and 71401098). The authors express sincere appreciation to the anonymous referees for their helpful comments to improve the quality of this paper.

8 References