

基于 ANN 及 CBR 的产品需求映射实现方法^①

梁春霞

(西门子(杭州)高压开关有限公司 IT&IS 部门, 杭州 310018)

摘要: 针对基于实例推理的产品需求映射实现过程中, 指标权重设定的随机性及主观性导致无关的噪音特征影响结论的准确性等状况, 提出了基于神经网络的 CBR 权值学习方法(即通过特征灵敏度、活跃度、凸度、相关度进行权值定义), 以及输入特征的转换机制, 在此基础上给出了基于 CBR 的产品需求映射过程模型, 并将上述模型在移动工作台需求映射实现中进行了应用, 系统实现表明了提出方法的有效性, 为产品需求映射提供了一种新的有效实现途径。

关键词: ANN; CBR; 需求映射

Product Requirement Mapping Method Based on ANN&CBR

LIANG Chun-Xia

(Siemens High Voltage Circuit Breaker Co. Ltd., Hangzhou 310018, China)

Abstract: During the case-based reasoning procedure of requirement mapping, the randomness and subjectiveness of weight setting always lead to the poor accuracy of results affected by irrelevant noise. A new method of CBR weight setting has been put forward to solve it based on ANN, which defined the relevant weight from the aspects of feature sensitivity, activity, saliency, relevance and the transforming mechanism of features input also been described. Based on it, a new requirement mapping procedure module has been put forward oriented to CBR and verified with the design of the hydraulic workbench, which gives a new method for product requirement mapping.

Key words: ANN; CBR; requirement mapping

在基于实例推理的产品需求映射过程, 首先在产品知识库进行实例检索。当用户输入产品需求参数进行检索, 系统按照 k-NN 方法进行相似性案例匹配即根据下述公式进行运算:

$$D(x, q) = \sqrt{\sum_{f=1}^n \omega_f \times \text{difference}(x_f, q_f)}$$

其中, ω_f 为用户需求输入特征 x_f 的权重, n 为用户需求特征集合的维数, $\text{difference}(x_f, q_f)$ 为用户需求输入特征与实例库中实例相应特征 q_f 间的差分距离, 其通常采用欧式距离法进行运算。由于上述各指标相似度的加权和作为案例相似程度的评判依据, 因此各指标的权重分布可以影响, 甚至可以决定检索出的最匹配的案例。而在实际应用中, 指标的权重往往由专家凭

个人偏好设定, 具有一定的主观性和随机性。这种权重的获取方法可能会让多余或无关的噪音特征影响最后的实例推理结论。为了削弱权重分布对案例相似度的影响, 下面给出了应用神经网络(back propagation neural network 简称 BPNN)作为案例推理检索算法中权重获取实现, 在此基础上给出了基于 CBR 的产品需求映射实现过程, 并以移动工作台为实例进行了上述过程的模拟应用。

1 基于神经网络的 CBR 权值学习实现

人工神经网络(artificial neural network 简称 ANN)是用大量神经元的互连以及对各连接权值分布来表示特定的领域概念或知识。在知识获取的过程中, 它通过专家提出范例及相应的解即通过特定的学习算法对样

① 收稿时间:2010-10-27;收到修改稿时间:2010-12-09

本进行学习,通过网络内部自适应算法不断修改连接权值分布达到要求,并把专家求解实际问题的启发式知识和经验通过网络的互连及权值分布来体现。

根据网络输入特征对网络样本训练的不同影响方式,可以从输入特征的信息维及测度维来描述输入特征权重对网络训练的影响,其中信息维主要通过样本特征信号视图与样本特征结构视图来表达输入特征对网络训练的作用。通过对输入特征信号进行测试,可以获得不同输入特征值对网络样本训练结果的影响度,该测试方法包括特征灵敏度、特征活跃度计算方法。另外,可以通过研究 BP 网络的连接权与阈值变化来描述输入特征的结构视图,例如:具有较高连接权的网络输入节点可以被理解为重要输入特征,BP 网络调整则是通过删除低连接权的节点来实现,其通常通过特征凸度以及特征相关度计算方法来完成。输入特征测度维则通过样本特征输入对网络输出节点幅值及变化幅度的作用来表达输入特征对网络训练的影响。例如:若网络的某节点激励输入后导致网络输出会产生较大幅值响应,则认为该输入节点对网络样本训练作用较大,但是若通过该节点输入不同激励后,产生网络输出幅值变化较小时,则认为该节点输入对网络训练的重要性低,即 BPNN 设计时要求不同的特征激励输入能产生不同的响应输出,网络的节点的连接权能根据不同的特征值输入作出相应调整。下面则通过上述四种特征权重计算方法(特征灵敏度、特征活跃度、特征凸度以及特征相关度^[1]计算方法)获取 BPNN 训练后的特征输入权重集。

1) 特征灵敏度

输入特征的灵敏度是通过 BP 网络进行训练时,加入/移除某输入特征 x_i 对 BPNN 的输出差异相比较来获取的。其计算公式为:

$$Se_i = \frac{\sum_L \frac{|P^0 - P^i|}{P^0}}{n}$$

其中, P^0 为每一个样本输入在正常网络训练后得到的响应输出, P^i 为输入样本中某个节点特征 i 被移除后,通过网络训练后得到的响应输出,在 BPNN 训练时,输入节点特征的移除可通过将与该节点相关的所有连接权设置为 0 来实现。 L 为训练样本集合, n 为训练样本数量。

2) 特征活跃度

输入特征的活跃度是通过不同的样本经过网络训练后所得到网络响应输出变化幅度,若网络中某节点的响应输出幅度随不同的激励输入变化越大时,则认为该节点的特征活跃度越高。反之,若网络中某节点的响应输出不随激励输入的变化而变化时,则认为该节点的特征活跃度为 0。每一个隐层节点的特征活跃度计算公式为:

$$A_j = \sum_{i=1}^q ((w_{ij}^{(2)})^2 \cdot \text{var}(g(\sum_{i=0}^n w_{ji}^{(1)} x_i)))$$

根据上述隐层节点的特征活跃度计算公式可知,每一个输入节点 x_i 的特征活跃度为:

$$A_i = \sum_{j=1}^p ((w_{ji}^{(1)})^2 \cdot A_j)$$

其中 w_{tj} 表示网络隐层与输出层单元连接权, w_{ji} 网络输入层与隐层单元连接权, t 为隐层单元序列号, j 为输出层单元序列号, i 为输入层单元序列号,以下类似。

3) 特征凸度

输入特征的凸度是通过网络样本训练中节点连接权的循环迭代误差水平来表达的。在网络训练过程中,通过计算网络隐层或输出层节点连接权的迭代误差,删除较低凸度权重,循环迭代直至网络全局误差调整到合理的误差范围。网络节点的特征凸度与其相关连接权的平方成正比,其计算公式如下:

$$Sa_i = \sum_{j=1}^p \sum_{i=1}^q ((w_{ji}^{(1)})^2 \cdot (w_{ij}^{(2)})^2)$$

4) 特征相关度

根据文献[2]可知,输入特征的相关度与节点的连接权的方差相对应。根据此对应关系可知,网络隐层的节点相关度为:

$$R_j = \sum_{k=1}^q (w_{kj}^{(2)})^2 \cdot \text{var}(w_{ji}^{(1)})$$

根据隐层节点的相关度可知,输入节点 x_i 的相关度为:

$$R_i = R_j \cdot w_{ji}^{(1)}$$

根据上述特征权重计算方法可得:BPNN 输入特征权重为:

$$\omega_i = \sqrt{(Se_i)^2 + (A_i)^2 + (Sa_i)^2 + (R_i)^2} / 2$$

2 基于BPNN的CBR特征输入转换

在运用 BP 网络进行产品需求映射实现中，通过网络样本训练及特征灵敏度、特征活跃度、特征凸度、特征相关度计算方法获取基于 CBR 的产品需求权重。根据论文^[4]中基于事务特性表原理的需求信息元形式化表达可知，用户需求输入属性值的类型有数值型、字符型、布尔型等。然而对于 BP 网络输入而言，由于不同字符型属性值的差异难以用准确的公式计算来表达，故其要求所有的样本输入必须为数值型，因此在运用 ANN 进行样本训练时，必须对字符型及布尔型输入进行数据转换，如图 1 所示的特征转换机制。

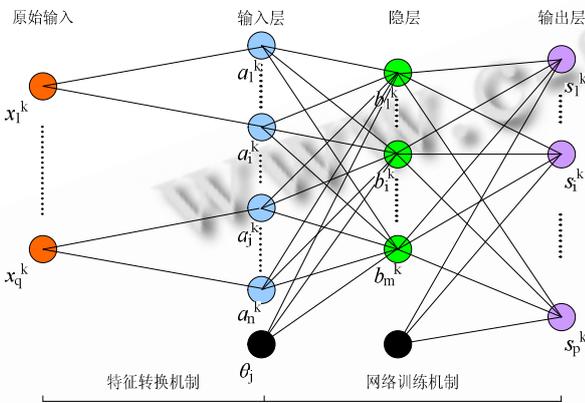


图 1 BP 网络样本转换及样本训练示意图

令 $(x_1, x_2, \dots, x_i, x_{i+1}, \dots, x_k, x_{k+1}, \dots, x_q)$ 为用户需求输入，其中 x_1, x_2, \dots, x_i 为字符型特征， x_{i+1}, \dots, x_k 为布尔型特征， x_{k+1}, \dots, x_q 为数值型特征。对于字符型特征 x_1, x_2, \dots, x_i ，设每个特征 x_p 对应的输入值域为 $(x_{1p}, mp, x_{2p}, mp, \dots, x_{ip}, mp)$ 则在 BP 网络构造过程中，原始输入 x_p 则对应若干个二进制输入 $(b_{1p}, mp, b_{2p}, mp, \dots, b_{ip}, mp)$ ，对于任何 b_{ip}, mp ，满足： $b_{ip}, mp \in \{0, 1\}$ 且 $b_{1p}, mp + b_{2p}, mp + \dots + b_{ip}, mp = 1$ ；对于布尔型特征输入 x_{i+1}, \dots, x_k ，由于每个特征 x_p 对应于输入值域 $\{0, 1\}$ ，故在 BP 网络构造时，原始输入与 BP 网络实际输入转换实现采用一对一方式即 true 为 1，false 对应 0；而对于数值型特征输入 x_{k+1}, \dots, x_q ，则原始输入与 BP 网络实际输入完全相同，直接参与 BP 神经网络样本训练。

3 基于CBR的需求映射过程模型

在基于实例推理的产品需求映射过程，首先在产品知识库进行实例检索。当用户输入产品需求参数进

行检索，系统按照 k-NN 方法进行相似性案例匹配即根据下述公式进行运算：

$$D(x, q) = \sqrt{\sum_{f=1}^n \omega_f \times \text{difference}(x_f, q_f)}$$

其中， ω_f 为用户需求输入特征 x_f 的权重， n 为用户需求特征集合的维数， $\text{difference}(x_f, q_f)$ 为用户需求输入特征与实例库中实例相应特征 q_f 间的差分距离，其通常采用欧式距离法进行运算。即当需求输入特征 x_f 值域为离散数值型时， $\text{difference}(x_f, q_f) = |x_f - q_f|$ ；然而当需求输入特征 x_f 值域为连续数值型时或为离散字符型时，上述特征距离计算方法则需要进行相应调整。

当需求输入特征 x_f 值域为连续数值型时，比如，用户要求移动工作台的工作运行速度为 8-24mm/s，回程速度为：120mm/s 等等，则传统的欧式距离计算方法则无法准确给出用户需求特征值与实例相应特征值域的相似度。根据两个非负的确定值 a 和 b 之间的相似性定义^[4]：

$$\text{sim}(a, b) = 1 - \frac{|a - b|}{\max(a, b)}$$

可以求得实例特征值 c 与用户需求特征值域区间 $[\alpha, \beta]$ 相似度：

$$\text{sim}(c, [\alpha, \beta]) = \begin{cases} \frac{c(\ln \beta - \ln \alpha)}{\beta - \alpha} & c \leq \alpha \\ 1 - \frac{1}{\beta - \alpha} \left(\frac{(\alpha - c)^2}{2c} + \frac{(\beta - c)^2}{2\beta} \right) & \alpha < c < \beta \\ \frac{\alpha + \beta}{2c} & c \geq \beta \end{cases}$$

那么，此时 $\text{difference}(x_f, q_f) = |x_f - q_f| = 1 - \text{sim}(x_f, q_f)$ 。

当需求输入特征 x_f 值域为离散字符串型时，用户需求特征输入 x_f 与实例库中实例相应特征 q_f 的差分距离则通过值差分矩阵来获取，其具体为：

$$\text{difference}(x_f, q_f) = \sum_{i=1}^n \left| \frac{D(f_i = x_f \cap g = c_i)}{D(f_i = x_f)} - \frac{D(f_i = q_f \cap g = c_i)}{D(f_i = q_f)} \right|$$

其中 $D(*)$ 表示实例库中满足条件*的实例数， f_i 为用户需求输入特征， g 为实例库中实例分类特征， l 为实例库中实例分类数目， c_1 为实例库中某一类实例。

$\frac{D(f_i = x_f \cap g = c_1)}{D(f_i = x_f)}$ 表示满足用户需求输入特征为 x_f 且实例属于 c_1 分类的概率分布。

当用户需求输入 f_i 与实例库中某分类特征的概率分布类似，则 $\text{difference}(x_f, q_f) \rightarrow 0$ 。

在给出了用户需求输入特征与实例库中样本的相

似度计算方法后,下面则给出基于 CBR 的需求映射实现流程,其具体如图 2 所示:

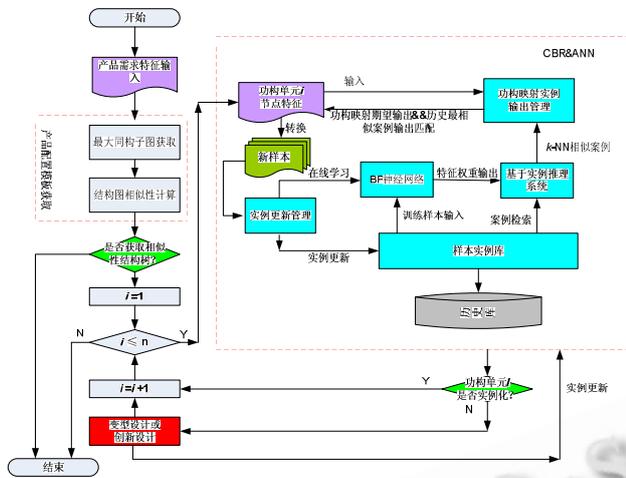


图 2 基于实例的产品需求映射流程图

1) 根据用户需求输入,获取所有产品功能结构树中与用户需求输入相匹配的最大同构子图。

2) 按照结构图相似性计算方法进行范例结构图相似性比较,并根据相似性比较结果,获取与用户需求输入相匹配的相似性结构树^[4]。

3) 若与用户需求输入相匹配的相似性结构树不存在,则转向步骤 7,流程结束。设计人员根据用户需求重新设计并构建与其对应的功能结构树。

4) 根据相似性结构树节点分布,按节点权重系数从高到低依次顺序进行功构节点 i 实例化。

5) 当 $i \leq n$ 时,根据论文^[5]实例库中样本数据所示方法构建 BP 网络,通过对实例库中的实例进行基于 BPNN 的网络样本培训,获取实例库中的功构节点 i 需求特征权重集 $\{\omega_{i1}, \omega_{i2}, \dots, \omega_{im}\}$ 。并根据动态计算获取的需求特征权重按 k-NN 方法获取与用户需求特征匹配的最相似案例,在此基础上通过对其进行维护实现与用户需求特征输入相对应的节点实例,将维护后的实例存入样本库中,并通过神经网络在线学习对功构节点 i 的需求权重集进行修正;同时将样本库中废弃或过时噪音样本数据存入历史样本库。

6) 当 $i=n$ 时,产品功能结构树上的所有节点均实例化,则基于需求特征输入的可配置产品功构映射实现完毕,转向步骤 7。

7) 结束。

4 基于ANN&CBR的移动工作台需求映射实现

首先对实例库中所有移动工作台的特征需求输入及移动工作台功构映射输出特征进行转换,从而构造移动工作台功构映射实现的 BPNN 训练样本。移动工作台的用户需求输入包括:电机功率 W (离散数值型)、承载重量 T (离散数值型)、台面尺寸 L (离散数值型)、工作台移动速度 v (离散数值型)、运行方式(离散字符型),输出为从动自由轮(离散字符型)、移动工作台(离散数值型)、传动轴(离散数值型)、减速机(离散数值型)、传动轴连接套(离散字符型)、齿轮轴(离散字符型)、轴承座(离散字符型)、主动轴(离散字符型)、从动定位轮(离散字符型)等单元实例。根据文献[4]中单约束线结构产品功构实例中介绍的用户需求输入及映射输出可得移动工作台训练样本:输入样本特征输入数目为 6,输出样本特征数目为 8,如表 1 及表 2 所示。其最左侧为样本序号,共 28 个训练样本。表 1 中运行方式特征输入通过样本特征转换机制,将运行方式特征列(MB、MC、MBC)转换为二进制列((0,1)、(0,1)、(0,1)),按类似方法将移动工作台输出特征(传动轴连接套、齿轮轴、轴承座、主动轴、从动定位轮、从动自由轮)亦转化为二进制列,如表 2 所示。

由于样本输出中减速器、移动工作台、传动轴等特征与输入特征电机功率、台面尺寸存在显性关系,即具有一一对应的线性关系,为提高神经网络训练速度,对 BPNN 输出特征中关于减速器、移动工作台、传动轴等特征部分不进行训练。根据样本特征转换机制可知,神经网络的输入特征为 8,输出特征为 37,样本数为 28,按文献[4]中 BP 网络中隐层单元构造方法令隐层单元数目为 15~35,如图 3 所示。

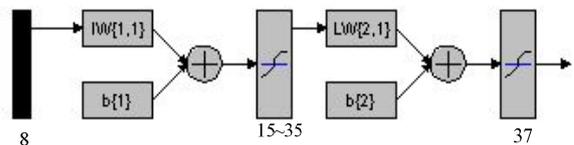


图 3 移动工作台功构映射特征网络示意图

运用 MATLAB 神经网络工具箱进行 BPNN 模拟不同隐层单元数对应不同的收敛过程,其对应网络训练目标与实际功构映射特征输出误差如表 3 所示,根

据不同隐层单元数对应的训练误差可知,当隐层单元数为30时,其对应的训练误差最小,同时考虑网络训练速度,将网络隐含层的神经元数目设定为30。当网络隐含层神经元的数目设定为30,并以表1中移动工作台功能需求参数作为输入特征,以表2中功构映射实例参数作为输出特征,进行功构映射特征网络训练,网络在迭代276次后,收敛于 $9.58562e-5$ 。此时输入层与隐层的连接权 ω_{ij} 及隐层与输出层单元连接权 ω_{jk} 分别如表4及5所示。

表1 移动工作台功构映射网络训练样本输入特征:

	电机功率 W	电机速度 n	承载重量 T	台面尺寸 L	工作台移动速度 v	运行方式
1	4.5	3600	50	4000	240mm/s	0 0 1
2	4.5	3500	55	4000	225mm/s	0 0 1
3	4.5	3200	60	4000	200mm/s	0 0 1
4	4.5	3000	65	4000	180mm/s	0 0 1
5	4.5	2800	70	4000	175mm/s	0 0 1
6	4.0	3600	50	3900	225mm/s	0 0 1
7	4.0	3500	55	3900	200mm/s	0 0 1
8	4.0	3200	60	3900	180mm/s	0 0 1
9	4.0	3000	65	3900	175mm/s	0 0 1
10	4.0	2800	70	3900	160mm/s	0 0 1
11	3.5	2500	70	3800	150mm/s	0 0 1
12	3.5	3500	50	3800	210mm/s	0 0 1
13	3.0	2500	60	3700	160mm/s	0 0 1
14	3.0	3000	50	3700	200mm/s	0 0 1
15	2.5	2800	70	3600	100mm/s	0 0 1
16	2.5	3500	50	3600	140mm/s	0 0 1
17	2.25	2500	70	3500	90mm/s	0 1 0
18	2.25	2800	60	3400	105mm/s	0 1 0
19	2.25	3000	55	3300	115mm/s	0 1 0
20	2.25	3200	50	3200	128mm/s	0 1 0
21	2.25	3500	45	3100	140mm/s	0 1 0
22	2.0	1500	50	3000	56mm/s	1 0 0
23	2.0	2000	45	2900	60mm/s	1 0 0
24	2.0	2500	40	2800	70mm/s	1 0 0
25	2.0	2800	35	2700	80mm/s	1 0 0
26	2.0	3000	30	2600	95mm/s	1 0 0
27	2.0	3200	25	2500	110mm/s	1 0 0
28	2.0	3500	20	2400	140mm/s	1 0 0

表2 移动工作台功构映射网络训练输出样本

实例编号	实例特征						实例输出					
	1	2	3	4	5	6	1	2	3	4	5	6
1	1	1	1	1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	1	1	1	1	1	1
3	1	1	1	1	1	1	1	1	1	1	1	1
4	1	1	1	1	1	1	1	1	1	1	1	1
5	1	1	1	1	1	1	1	1	1	1	1	1
6	1	1	1	1	1	1	1	1	1	1	1	1
7	1	1	1	1	1	1	1	1	1	1	1	1
8	1	1	1	1	1	1	1	1	1	1	1	1
9	1	1	1	1	1	1	1	1	1	1	1	1
10	1	1	1	1	1	1	1	1	1	1	1	1
11	1	1	1	1	1	1	1	1	1	1	1	1
12	1	1	1	1	1	1	1	1	1	1	1	1
13	1	1	1	1	1	1	1	1	1	1	1	1
14	1	1	1	1	1	1	1	1	1	1	1	1
15	1	1	1	1	1	1	1	1	1	1	1	1
16	1	1	1	1	1	1	1	1	1	1	1	1
17	1	1	1	1	1	1	1	1	1	1	1	1
18	1	1	1	1	1	1	1	1	1	1	1	1
19	1	1	1	1	1	1	1	1	1	1	1	1
20	1	1	1	1	1	1	1	1	1	1	1	1
21	1	1	1	1	1	1	1	1	1	1	1	1
22	1	1	1	1	1	1	1	1	1	1	1	1
23	1	1	1	1	1	1	1	1	1	1	1	1
24	1	1	1	1	1	1	1	1	1	1	1	1
25	1	1	1	1	1	1	1	1	1	1	1	1
26	1	1	1	1	1	1	1	1	1	1	1	1
27	1	1	1	1	1	1	1	1	1	1	1	1
28	1	1	1	1	1	1	1	1	1	1	1	1

表3 移动工作台功构映射特征网络训练误差表

隐层单元数	15	16	17	18	19	20	21
误差(Y-T)	0.1429	0.1526	0.1435	0.112	0.1541	0.1636	0.1118
隐层单元数	22	23	24	25	26	27	28
误差(Y-T)	0.2009	0.157	0.1689	0.179	0.1651	0.1673	0.1331
隐层单元数	29	30	31	32	33	34	35
误差(Y-T)	0.1524	0.0738	0.1683	0.1317	0.1452	0.1453	0.1491

表4 移动工作台功构映射特征网络输入层与隐层单元连接权 ω_{ij} 列表

	1	2	3	4	5	6	7	8
1	-1.704	-0.0086	-0.7017	0.0223	-0.0842	1.9878	-1.4876	-0.4495
2	-1.0274	-0.0033	0.166	-0.0039	0.0759	2.9203	-0.3369	1.712
3	-0.9536	-0.0051	-1.7064	0.0417	-0.1984	-1.4615	1.1271	1.9349
4	0.1996	-0.0142	-2.27	0.0459	-0.0744	-1.929	-0.0943	-0.0226
5	0.2712	-0.0197	-0.3104	0.011	0.2247	-1.5829	3.3097	-0.7426
6	0.8481	0.0013	-0.0989	0.0022	-0.0615	-0.4275	1.0206	-2.2322
7	-0.8609	0.0023	0.1486	-0.0071	0.0233	-0.054	-1.5333	-1.8545
8	-0.8448	-0.0012	-0.1426	0.0027	0.0074	1.5772	-2.4954	-1.2429
...
29	-0.321	-0.0061	-0.1557	0.0089	-0.0166	-2.2678	1.1828	2.7792
30	0.313	0.0018	0.2338	-0.0056	0.015	-1.1684	-1.9642	-1.3446

表5 移动工作台功构映射特征网络隐层与输出层单元连接权 ω_{jk} 列表

	1	2	3	4	5	6	...	30
1	0.0514	0.4809	0.0796	0.0367	-0.4455	-0.145	...	0.3246
2	0.3117	-0.1112	-0.0034	0.1082	0.2963	-0.4853	...	-0.1667
3	-0.7967	-0.6919	0.4557	-0.451	0.048	0.3679	...	0.1206
4	-0.0672	0.2834	0.0852	0.193	-0.4044	-0.0637	...	-0.2985
5	0.1371	-0.1358	0.0401	0.3164	0.0554	-0.5273	...	0.1204
6	-0.869	-0.4233	0.3592	-0.3412	-0.1068	0.2481	...	-0.2832
7	-0.0492	0.0462	-0.0186	0.4127	0.0167	-0.6934	...	-0.6027
8	-0.0423	0.0578	0.124	-0.5626	-0.351	0.2916	...	0.014
...
35	0.8859	-0.5198	-0.7685	-1.0087	-1.0051	0.9275	...	-0.088
36	-0.4208	0.2426	1.5287	0.3969	1.5766	0.1177	...	-0.2171
37	-0.4459	0.2468	-1.1031	0.0337	-0.8115	0.8602	...	-0.9071

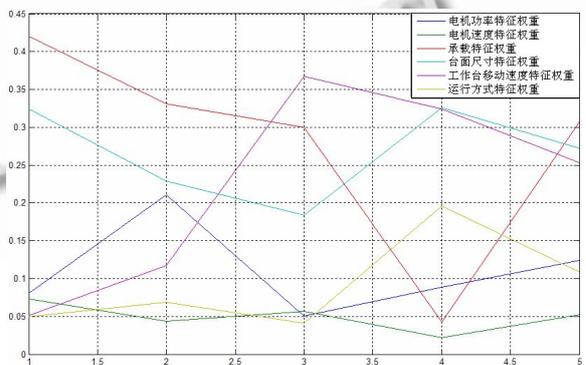


图4 移动工作台功构映射特征权重列表

表6 实例库样本与用户需求特征相似性列表

	电机功率 W	电机速度 n	承载重量 T	台面尺寸 L	工作台移动速度 v	运行方式	Σ
C	(2.5-4)kw	(3000-4500)	50-60T	3000-3500mm	150-200mm/s	MBC	
1	0.7222	0.9067	0.9116	0.8125	0.7291	1	1.8645
2	0.7222	0.9033	0.9541	0.8125	0.7778	1	1.9149
3	0.7222	0.8712	0.9167	0.8125	0.875	1	1.9378
4	0.7222	0.8109	0.8461	0.8125	0.94	1	1.921
5	0.7222	0.7586	0.7857	0.8125	0.96178	1	1.8895
6	0.8125	0.9067	0.9116	0.8333	0.7778	1	1.9229
...
28	0.6277	0.9033	0.3646	0.7399	0.8055	0.3333	1.3583

根据上述运算所得特征权重及用户需求输入, 按基于实例的可配置产品结构映射过程中需求特征与实例特征相似性计算方法, 可得用户功能需求特征与实例库样本特征相似性如表 6 所示, 并根据 $\sum \omega * d$ 可得与用户需求特征最相似的样本实例 3。设计人员根据样本实例 6 的结构特征, 按直接转换、置换法、改造法或基于框架的方案变换对样本 6 进行修正。设计人员完成样本修正后, 将修正的实例加入实例库中, 从而实现系统的自我学习任务。实例的学习过程包括实例的评价和修正, 即评价和确认是否进行实例更新。具有较高评价的实例为成功实例, 具有较低评价的实例为失败实例。实例的更新依赖于当前实例与库中实例的相似性, 其主要包括更新信息内容确定、更新形式确定、实例索引及实例存储。

5 结论

(1) 分析了目前基于实例推理方法中权重制定的主观性对推理结论的影响。

(2) 提出了基于神经网络的 CBR 权值学习实现方法及输入特征转换机制, 在此基础上给出了基于神经网络算法和 CBR 的需求映射实现过程。

(3) 将上述方法应用于移动工作台功能特征映射取实现过程中, 实践证明上述方法的有效性。

参考文献

- 1 Kwang HI, Sang CP. Case-based reasoning and neural network based expert system for personalization. *Expert Systems with Applications*, 2007, 32(1): 77-85.
- 2 Segee BE, Carter MJ. Fault tolerance of pruned multilayer networks. *Proceedings of international joint conference on neural networks*, 1991, II: 447-452.
- 3 王世伟. 基于知识的产品配置建模、演化及其应用研究[博士学位论文]. 杭州: 浙江大学, 2004.
- 4 谢清. 面向方案设计的可配置产品功能--结构映射原理、方法及关键技术研究[博士学位论文]. 杭州: 浙江大学, 2007.
- 4 Badrul S, Karypis G, Konstan J, Riedl J. Analysis of recommendation algorithms for e-commerce. *Proc. of EC'00*. Minneapolis, 2000.
- 5 Mobarshel B, Cooley R, Srivastava J. Automatic personalization based on Web usage mining. *Communications of ACM*, 2000, 43(8): 142-151.
- 6 Baeza-Yates RA, Ribeiro-Neto B. *Modern Information Retrieval*. MA: Addison-Wesley, 1999.
- 7 Aas K, Eikvil L. *Text Categorisation: A Survey*. Norwegian: ACM Computing Center, 1999.
- 8 Xu R, Wunsch D. Survey of clustering algorithms. *IEEE Trans. on Neural Networks*, 2005, 16(3): 645-678.
- 9 Sheng ZH, Wu Y, Jiang JH, Shou LD, Chen G. InfoSigs: A Fine-Grained Clustering Algorithm for Web Objects. *Journal of Computer Research and Development*, 2010, 47(5): 796-803.
- 10 Zhao Y, Wang XL, Liu BQ. Fusion of clustering trigger-pair features for POS tagging based on maximum entropy model. *Journal of Computer Research and Development*, 2006, 43(2): 268-274.
- 11 Theobald M, Siddharth J, Paepcke A. SpotSigs: Robust and efficient near duplicate detection in large Web collections. *Proc. of the 31st SIGIR Conf on Research and Development in Information Retrieval*. New York: ACM, 2008. 563-573.
- 12 Shannon CE. A Mathematical Theory of Communication. *The Bell System Technical Journal*, 1948, 27: 379-423, 623-656.
- 13 Gu L, Baxter R. *Decision models for record linkage*. LNCS 3755: Data Mining AusDM. Berlin: Kluwer Academic Publishers, 1998.

(上接第 30 页)