

文章编号: 2095-2163(2019)02-0016-06

中图分类号: TP391.41

文献标志码: A

基于数据的多目标作业车间的复杂网络模型关键节点的寻找

韩佳蓉

(齐鲁工业大学(山东省科学院) 计算机科学与技术学院, 济南 250353)

摘要: 大数据时代下工业制造开始转向智能制造, 车间生产过程变得更为智能化和自动化, 基于数据信息的多目标作业车间复杂网络模型, 已然成为智能制造研究的新领域之一。本文研究的主要内容有两点, 其一是利用数据信息搭建一个多目标作业车间复杂网络模型; 其二是在搭建的模型基础上利用模糊网络分析法寻找该模型的关键节点, 选择模糊网络分析法使得关键节点评价的结果更具有客观性。最后进行仿真实验, 结果表明该模型能较好地应用于实际工业生产, 模糊网络分析法能有效挖掘该模型中的关键节点, 证明了方法的合理性和有效性。

关键词: 智能制造; 工业大数据; 多目标作业车间问题; 复杂网络; 模糊网络分析法

Searching for key nodes in complex network model of multi-objective job shop based on data

HAN Jiarong

(School of Computer Science and Technology, Qilu University of Technology(Shandong Academy of Sciences), Jinan 250353, China)

【Abstract】 In the era of big data, industrial manufacturing has begun to turn to intelligent manufacturing. The workshop production process has become more intelligent and automated. The complex network model of multi-objective workshop based on data information is one of the new fields of intelligent manufacturing research. There are two main points in this paper. One is to build a complex network model of multi-objective job shop based on data information; the other is to use the fuzzy network analysis method to find the key nodes of the model, and choose the fuzzy network analysis method to make the evaluation results of key nodes more objective. Finally, the simulation results show that the model can be better applied to actual industrial production, and the fuzzy network analysis method can effectively mine the key nodes in the model, which proves the rationality and effectiveness of the method.

【Key words】 intelligent manufacturing; industrial big data; multi-objective job shop problem; complex network; fuzzy network analysis

0 引言

近年来,生产过程日渐复杂、精细,产生的数据量越来越多,这就使得处理难度也随即增加,如此规模的数据量使生产的细节趋于量化与可控,并能为生产过程的分析提供丰富的资源^[1]。作为制造过程优化中至关重要的一部分,多目标柔性作业车间生产问题因在提高生产效率、降低生产成本等方面发挥强大作用而受到大多数企业学者的关注和重视^[2]。但传统流程工业模型优化受到发展的限制,面临着更为复杂的情形,研究者为寻求突破便利用复杂网络建模,将生产数据作为模型节点,从多目标制造的数据特征出发,对数据展开深入分析。有关复杂网络的建模方法有:文献[3]探讨了适应性供应链的概念,指出学习是提高供应链适应能力的一个重要途径,提出了适应性供应链的一个初步模型

框架。然后建立了供应链的复杂网络演化模型,导出了其基本统计规律。文献[4]将统计方法、非线性系统理论、控制理论以及矩阵理论等理论和方法应用到复杂网络的研究中,对复杂网络的动力学性质和加权复杂网络的建模两个方面进行了研究。而文献[5]中,结合复杂网络理论在复杂系统评价上的优势,运用统计物理、图论、运筹学及计算机模拟等方法,将实际网络特性与复杂网络理论进行关联,建立复杂产品制造过程网络演化模型,从而将复杂网络成功应用于智能制造,且为后续工作增加了一个新的研究思路。

综合前文所述,本文设计了一个基于生产数据信息的多目标作业车间复杂网络模型,该模型以车间实际运转过程中产生的数据作为模型的节点,数据间关系作为复杂网络的边,关系之间权重利用线性函数公式求出。而后利用模糊网络分析法

作者简介: 韩佳蓉(1992-),女,硕士研究生,主要研究方向:工业大数据与智能制造。

收稿日期: 2018-12-20

(FANP) 求出关键节点, 该方法是网络分析法在不确定性和含糊性问题上的延伸, 是一种能将 2 种特性复杂问题的定量化方法。最后进行仿真分析, 证明该方法的有效性和合理性, 进而表明这是能够适用于多目标作业车间复杂网络模型。

1 多目标作业车间复杂网络模型的搭建

多目标车间作业的复杂网络模型搭建需要经过数据预处理、建立边、确定权重等关键研究步骤。本文用 $G = (R, E, W)$ 来表示该复杂网络模型, 其中 R 表示节点集合, E 表示边的集合, W 表示权重集合, 这是一个有向加权的网络。本文将给出研究论述如下。

1.1 数据预处理

生产中的数据直接来源于分布于车间各个生产单元中的传感器, 例如温度传感器、压力传感器、速度传感器等, 按照设定的参数, 传感器每隔一段时间便返回当前数据到服务器。由于实际传输过程中传感器无响应导致数据丢失, 由传感器信号失真导致的数据错误和异常是不可避免的, 为此就要对数据进行预处理: 将传感器传回来的一串数据作为一个按时间排列的数据序列, 序列的第一个数据编号为 0, 每个序列的第一个数据不能为空。某一时刻的某序列值为 null (传感器未响应请求或捕获数据失败), 则所有序列在这一时刻数据重写为 null。接下来根据序列生成逻辑序列: 开始处理的第一个数据所处的时刻为 0 时刻, 该时刻对应逻辑序列值为 0。从第二个数字开始, 若数据序列中与前一个数据相比增加了, 则逻辑序列生成的数值记为 1, 不变为 0, 减小为 -1。若遇到 null 数据, 该数据与下一条数据不生成逻辑序列项, 从第三条开始重复执行上述规则, 直至遍历完整个数据序列。

在预处理阶段会出现 2 种特殊节点。一类是持续恒增的数据, 另一类是恒定不变的数据。恒定不变的数据记为 N/A (not available), 不参与后续边的建立和边权的设置; 持续增量数据生成增量阵列 IQA (Incremental Quantity Array), IQA 生成时处理 null 类数据与正常序列相同, 记录与上一条数据相比数据增值量。根据 IQA 生成逻辑序列, 求出 IQA 中的众数, 并将 IQA 中的每个数据与之比较后生成逻辑序列, 若大于该值记为 1, 等于记为 0, 小于记为 -1。

1.2 复杂网络模型的建立

在基于数据的复杂网络模型中, 模型的节点 R

不再是具体的某个生产环节, 而是不断产生数据序列的数据点集合。数据节点与工序间的关系如图 1 所示, 由图 1 中可以看出一个流程实体包含了多个数据。

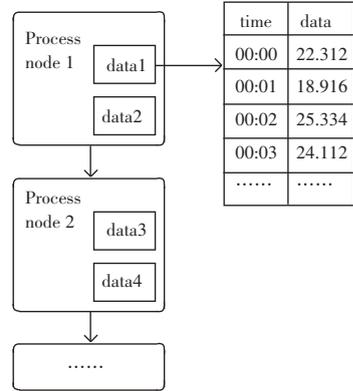


图 1 工序与节点表示图

Fig. 1 Process and node diagram

在数据预处理阶段, 本文已将数据处理作为逻辑序列, 在此基础上即利用 Apriori 算法挖掘这些逻辑序列之间的关系。设定存在 A, B 两个节点, 若 A, B 同增、同减用 $A \rightarrow B$ 表示; A 增大、 B 减小, 或 A 减小、 B 增大时用 $A \rightarrow -B$ 表示; 2 个节点不存在关联用 $A \rightarrow \neg B$ 表示。节点 A, B 之间的关系, 只有 $A \rightarrow B$ 、 $A \rightarrow -B$ 以及 $A \rightarrow \neg B$ 三种情况事件, 则所有事件逻辑表示及概率见表 1。

表 1 所有事件逻辑表示及概率表示

Tab. 1 All practical logical representations and probabilistic representations

事件的逻辑表示	概率的逻辑表示	概率
AB	$P(A \cup B)$	$P(1, 1) + P(-1, -1) + \xi$
$A - B$	$P(A \cup -B)$	$P(1, -1) + P(-1, 1) + \xi$
$A \neg B$	$P(others)$	$P(1, 0) + P(0, 1) + P(0, -1) + P(-1, 0) + \xi$

综合上述分析, 就能得到每个事件的支持度计算公式可表示为:

$$\begin{cases}
 s(AB) = s(1, 1) + s(0, 0) + s(-1, -1); \\
 s(A - B) = s(1, -1) + s(0, 0) + s(-1, 1); \\
 s(others) = s(1, 0) + s(0, 1) + s(0, -1) + s(-1, 0).
 \end{cases}
 \tag{1}$$

当 $\xi = 0$ 时, $P(others) = P(1, 0) + P(0, 1) + P(0, -1) + P(-1, 0) + \xi$, 若 $P(others)$ 的值大于或等于 30% (节点 A, B 间不存在关联), 计算 $c(others) = 1/3(c(1, 0) + c(0, 1) + c(-1, 0))$ 的值, 其值大于 44% 则认定 A, B 之间没有关联。若 $P(others)$ 不满足最小支持度或最小置信度,

则 A, B 之间可能存在关联, 验证当 $\xi = 0$ 时 $P(A \cup B)$ 和 $P(A \cup -B)$ 的值, 若其值大于 40%, 则 A, B 之间存在关联。当确定 A, B 存在何种关联后, 研究推得对应的置信度的数学公式如下:

$$\begin{cases} c(AB) = \frac{1}{3}(c(1, 1) + c(0, 0) + c(-1, -1)), \\ c(A - B) = \frac{1}{3}(c(1, -1) + c(0, 0) + c(-1, 1)). \end{cases} \quad (2)$$

若 2 个节点最小支持度为 40% 且满足最小置信度为 60%, 则这 2 个节点间存在强关联规则, 在 A, B 两节点之间建立边。若不满足最小支持度 40% 或不满足最小置信度 60%, 则 A, B 之间不存在强关联规则, 不在 A, B 之间建立边。

为了准确地描述数据之间的关系, 需要在模型中为每一条边加入权重。边的权重集可表示为 $W = \{w_{ij} = f(r_i, r_j) \mid i, j \in (1, 2, 3, \dots, r)\}$, 其中, w_{ij} 表示从节点 i 指向节点 j 的边的权重。若 A, B 两节点是有关联的, 则上游节点产生的数据和下游节点产生的数据间将存在一定的函数关系。设节点 r_i 与 r_j 为节点集 r 中 2 个相邻的节点, 边的方向由 r_i 指向 r_j , 则在某一时刻节点 r_i 与 r_j 的值存在函数关系 $r_j = f(r_i)$, 该函数表达式可由节点 n_i 与节点 n_j 的数据序列求得。那么 r_i, r_j 之间的权值可表示为 $W_{ij} = f'(r_i)$ 。当且仅当两节点间为线性变化关系时, 权值 W_{ij} 才为常数。不难理解, 在实际生产过程中, 该函数关系并不总是线性的, 因此在很多情况下, 该值是变化的。研究中假设当节点 r_i 的值为 x 时, 节点 r_j 的值为 y , 则权值 W_{ij} 的数学表述如下:

$$W_{ij} = \frac{\partial(f(r_i))}{\partial r_i}. \quad (3)$$

当两点间的关系为非线性关系时, 得到的权值将是一个函数值随上游节点数据变动的函数表达式, 反映了上游节点对下游节点影响力的大小。由此可以构建一个基于数据的有向加权复杂网络模型。

2 模糊网络分析法

模糊网络分析法是基于网络分析法与模糊集理论相结合的一种系统决策方法^[6]。对此可做探讨分述如下。

2.1 三角模糊数

模糊数的模糊集合可以表示为 $F = \{x \in R \mid \mu_f(x)\}$, $R: -\infty < x < +\infty, \mu_f(x)$ 是从 R 到区间 $[0, 1]$ 上的一个连续的映射。三角模糊数 M 常用

(l, m, s) 来表示, 即 $M = (l, m, s)$, 其中 $l \leq m \leq s$ 。三角模糊数 M 的隶属函数 $f_M(x): R \rightarrow [0, 1]$, 如图 2 所示, 进而推得其数学定义可写作如下形式:

$$f_M(x) = \begin{cases} (x - l)/(m - l), & l \leq x \leq m; \\ (s - x)/(s - m), & m \leq x \leq s; \\ 0, & x < l \text{ or } x > s. \end{cases} \quad (4)$$

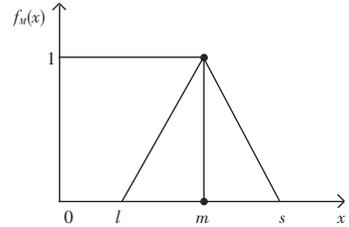


图 2 三角模糊数 M

Fig. 2 Triangular fuzzy number M

其中, l 和 s 分别表示为 M 评价小组所确定的下界和上界最小最大值, m 为隶属度最可能值。 $s - l$ 越大、越模糊; $s - l$ 越小、模糊度越低, $l = m = s$ 时说明判断是非模糊的, 三角模糊数的取值参见表 2。

表 2 重要程度定量表

Tab. 2 Quantitative table of importance

a_{ij}	指标 i 比指标 j 重要程度
1	$i = j$
3	$i > j$
5	$i >> j$
7	$i >>> j$
9	$i >>>> j$

2.2 模糊网络分析法基本步骤

首先确定复杂网络关键节点的评语集以及评价因素集: 评判者要对评判对象做出各种可能的评判结果, 这些结果组成一个集合 $V = \{V_1, V_2, \dots, V_m\}$ 。复杂网络的因素集 $U = \{U_1, U_2, \dots, U_N\}$, 其中 $U_i = \{U_{i1}, U_{i2}, \dots, U_{in}\}$, ($i = 1, 2, \dots, N$)。获得对复杂网络的指标进行单因素评价, 建立 U 到 V 的模糊关系 $F \times R$, 下一步即需求得 FANP 的权重。在本文中, ANP 模型分为控制层和网络层两个部分, 如图 3 所示。由图 3 可知, 控制层是一级指标, 即设备自身属性和网络特征参数, U_1 包括设备价值、生产能力、故障频率、维修费用; 网络层是二级指标, 二级指标 U_2 包括度、介数、聚类系数、节点特征向量。此后, 就是研究运算得到复杂网络关键节点评价的各个指标权重, 利用各个指标权重带入实际节点的真实数据, 最终得到排名靠前的节点为关键节点。研究中, 各步骤设计内容可详述如下。

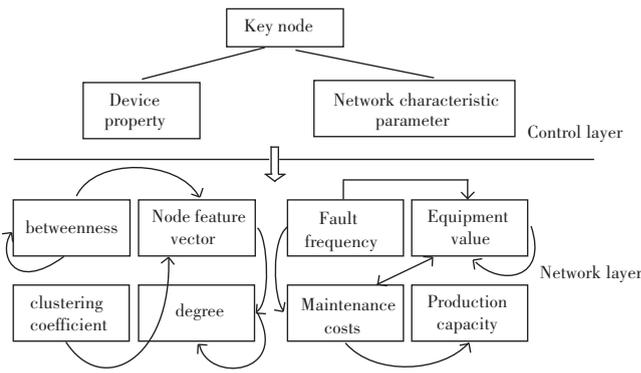


图 3 ANP 结构模型

Fig. 3 ANP structure model

(1)应用三角模糊数构造模糊判断矩阵。假设节点组 U_1 中某一节点 $U_{li}(i = 1, 2, \dots, n)$ 对复杂网络关键节点的影响程度为次准则,运用三角模糊数的性质来构造 U_1 中各个节点间两两互补判断矩阵,记 $p_{ij} = (l_{ij}, m_{ij}, s_{ij})$, 共有 n 个。对任意的 p_{ij}, p_{ji} 都有 $l_{ji} + s_{ij} = m_{ji} + m_{ij} = s_{ji} + l_{ij} = 1$ 。

(2) 确定超矩阵的局部权重向量 W_{11} 及其它。 W_{11} 是一个矩阵,这是 U_1 中的某一节点 $U_{li}(i = 1, 2, \dots, n)$ 对复杂网络关键节点影响程度的次准则,判断 U_1 中各个节点两两相比较的重要性。对其设计过程可阐述如下。

Step1 计算节点 U_{li} 的综合重要程度 C_{li} , 研究推得的计算公式为:

$$C_{li} = \sum_{j=1}^n p_{ij} \otimes \left[\sum_{i=1}^n \sum_{j=1}^n p_{ij} \right]^{-1} (i = 1, 2, \dots, n, j = 1, 2, \dots, n), \tag{5}$$

其中, C_{li} 是三角模糊数,可以表示为 $C_{li} = (l_{ij}^{li}, m_{ij}^{li}, s_{ij}^{li})$, $(i, j = 1, 2, \dots, n)$ 。研究中还将用到的其它计算公式见如下:

$$\sum_{j=1}^n p_{ij} = \left[\sum_{j=1}^n l_{ij}, \sum_{j=1}^n m_{ij}, \sum_{j=1}^n s_{ij} \right] (i = 1, 2, \dots, n, j = 1, 2, \dots, n), \tag{6}$$

$$\sum_{i=1}^n \sum_{j=1}^n p_{ij} = \left[\sum_{i=1}^n \sum_{j=1}^n l_{ij}, \sum_{i=1}^n \sum_{j=1}^n m_{ij}, \sum_{i=1}^n \sum_{j=1}^n s_{ij} \right] (i = 1, 2, \dots, n, j = 1, 2, \dots, n), \tag{7}$$

$$\left[\sum_{i=1}^n \sum_{j=1}^n p_{ij} \right]^{-1} = \left[\sum_{i=1}^n \sum_{j=1}^n l_{ij}, \sum_{i=1}^n \sum_{j=1}^n m_{ij}, \sum_{i=1}^n \sum_{j=1}^n s_{ij} \right]^{-1} = \left[\frac{1}{\sum_{i=1}^n \sum_{j=1}^n s_{ij}}, \frac{1}{\sum_{i=1}^n \sum_{j=1}^n m_{ij}}, \frac{1}{\sum_{i=1}^n \sum_{j=1}^n l_{ij}} \right], \tag{8}$$

Step 2 计算 $C_{li} \geq C_{lk}$ 的可能程度,其计算公式为:

$$(C_{li} \geq C_{lk}) =$$

$$\left\{ \begin{aligned} &1, && m_{ij}^{li} \geq m_{ij}^{lk}; \\ &\frac{s_{ij}^{li} - l_{ij}^{lk}}{s_{ij}^{li} - l_{ij}^{lk} + m_{ij}^{lk} - m_{ij}^{li}}, && m_{ij}^{li} < m_{ij}^{lk}, l_{ij}^{lk} \leq s_{ij}^{li}; \\ &0, && \text{其它.} \end{aligned} \right\} \tag{9}$$

其中, $i = 1, 2, \dots, n, k = 1, 2, \dots, n \ \& \ k \neq i, j = 1, 2, \dots, n$ 。

Step 3 计算 U_1 中的节点 $U_{li}(i = 1, 2, \dots, n)$ 相对于其它节点的可能性重要程度,具体公式如下:

$$d'(U_{li}) = \min V(C_{li} \geq C_{lk}, C_{lh}, i = 1, 2, \dots, n, k = 1, 2, \dots, n \ \& \ k \neq i, h = 1, 2, \dots, n \ \& \ k \neq i \neq h), \tag{10}$$

Step 4 重复 Step1~Step3 n 次,可以得到权重向量 W_{11}^{li} 。这个过程中得到了 n 个 $d'(u_{li})$, 继而得到 $W_{11}^{(1)'} = (d'(U_{11}), d'(U_{12}), \dots, d'(U_{1n}), \dots, d'(U_{n1}), \dots, d'(U_{nn}))^T$, 然后将其归一化就可以得到权重向量 $W_{11}^{(1)} = (d(u_{11}), d(u_{12}), \dots, d(u_{1n}), \dots, d(u_{n1}), \dots, d(u_{nn}))^T$ 。

Step 5 重复 Step1~Step4 n 次就可以得到 n 个 $W_{11}^{(i)}$, 获得超矩阵局部权重向量 W_{11} , 即: $W_{11} = (W_{11}^{(1)}, W_{11}^{(2)}, \dots, W_{11}^{(n)})^T$ 。

(3) 计算 $W_{ij}(i, j = 1, 2, \dots, N)$ 。以节点组 U_i 中各个节点对复杂网络关键节点的影响程度为次准则,将 U_j 中的节点两两比较或者模糊判断矩阵并且进行一致性检验,再用相同计算得出 $W_{ij}(i \neq j)$ 。

(4) 确定复杂网络关键节点的超矩阵 W 和加权超矩阵 \bar{W} , 加权超矩阵可根据式(11) 计算得到, 即:

$$\bar{W} = A \cdot W = \begin{bmatrix} a_{11} W_{11} & a_{12} W_{12} & \dots & a_{n1} W_{n1} \\ a_{21} W_{21} & a_{22} W_{22} & \dots & a_{n2} W_{n2} \\ \dots & \dots & \dots & \dots \\ a_{n1} W_{n1} & a_{n2} W_{n2} & \dots & a_{nn} W_{nn} \end{bmatrix}. \tag{11}$$

(5) 确定 FANP 的权重 Q 。FANP 的权重 Q 是 \bar{W} 极限超矩阵的列详细, 利用 Matlab 计算出 \bar{W}^∞ , 便可确定权重向量。

3 仿真实验分析

3.1 复杂网络模型的搭建

本文以某无碱窑炉工艺生产玻璃纤维的加工流程为实例, 其中的无碱窑炉工艺生产玻璃纤维的加工流程如图 4 所示。该生产流程包括 12 个不同环节, 总共有 139 个数据传感器接收点。选取一条生产线的的数据, 利用本文提出的方法建立基于数据的复杂网络模型, 其可视化效果如图 5 所示。由图 5 可以见

到,图中包含了部分孤立节点,这些节点中有一些是监控型节点,这是由节点属性导致。在该模型中,边反映了是否存在关联,边的权重代表了数据之间在何种关联。利用复杂网络相关公式计算出不同节点的度、介数、聚类系数和节点特征向量。计算结果见表3,由于计算数据较多,文中仅节选了部分数据。

表3 复杂网络评价指标和重要度

Tab. 3 Evaluation indicators and importance of complex networks

节点	度	介数	聚类系数	节点特征向量
R_1	0.002 264	0.053 420	0.005 740	0.026 480
R_2	0.036 580	0.018 990	0.003 240	0.034 730
...
R_{139}	0.013 930	0.220 230	0.001 130	0.112 500

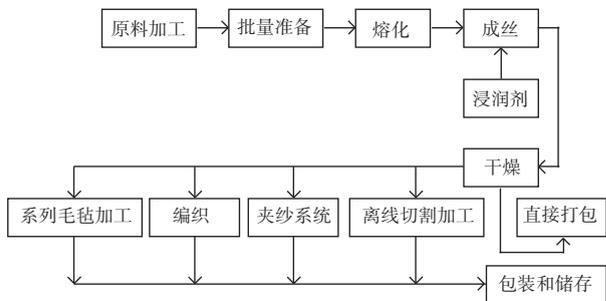


图4 玻璃纤维生产过程任务流程图

Fig. 4 Task flow chart of fiberglass production process

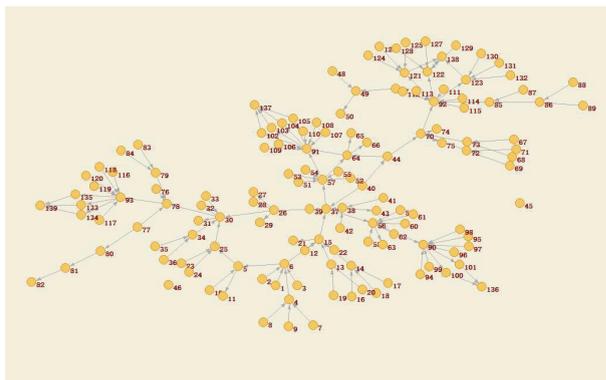


图5 基于数据的复杂网络模型图

Fig. 5 Data-based complex network model diagram

依据重要度评价体系中的评价属性对传感器所监测机器的设备价值、产能、故障频率、维修成本等数据收集汇总后进行归一化处理,处理结果见表4。由于数据较多,也只列举了其中一部分。

表4 资源节点本身属性

Tab. 4 The attributes of resource node

节点	设备价值	生产能力	故障频率	维修费用
R_1	0.199 0	0.657	0.33	0.23
R_2	0.599 0	0.341	0.10	0.60
...
R_{139}	0.000 3	0.829	0.07	0.14

3.2 复杂网络关键节点评价

为了模型描述的方便与精确起见,研究中把评价目标评语分为4个等级,分别是: $V = \{ \text{非常重要、较重要、较不重要、不重要} \}$,一级指标因素集 $U = \{ U_1, U_2 \} = \{ \text{设备自身属性,网络特征参数} \}$,二级指标因素集 $U_1 = \{ U_{11}, U_{12}, U_{13}, U_{14} \} = \{ \text{设备价值,生产能力,故障频率,维修费用} \}$; $U_2 = \{ U_{21}, U_{22}, U_{23}, U_{24} \} = \{ \text{度,介数,聚类系数,节点特征向量} \}$ 。选择智能制造领域专家、一线人员、工厂管理人员等对二级指标进行单因素评判,获得二级指标评判结果,详见表5。

表5 各指标的模糊权重集

Tab. 5 Fuzzy weight set of indicators

一级指标	权重	二级指标	权重	等级 (V)
设备自身属性	0.437	设备价值	0.21	0.42 0.53 0.05 0
		生产能力	0.39	0.51 0.39 0.10 0
		故障频率	0.30	0.46 0.10 0.07 0.04
		维修费用	0.10	0.23 0.51 0.10 0.16
网络特征参数	0.563	度	0.20	0.30 0.40 0.20 0.10
		介数	0.30	0.50 0.40 0.10 0
		聚类系数	0.30	0.40 0.50 0.10 0
		节点特征向量	0.20	0.23 0.50 0.20 0.07

利用表5拟将进行模糊综合评判,得到 $F \times R$ 。考虑到篇幅有限,故而文中省略了此后各步骤的运算结果。而在将得到的指标综合评价系数与各个节点相应的实际数据实现有机结合后,则逐个进行运算,最终得到排名靠前的节点,即为所求关键节点,完整序列对比输出见表6。该结果与传统的AHP层次分析法相比较,模糊网络分析法修正了AHP层次分析法的主观性和含糊性,识别度更高,识别结果更加准确。

表6 节点重要性排序结果

Tab. 6 Node importance sorting results

研究方法	节点排序
AHP 层次分析法	$R_{92} > R_{70} > R_{91} > R_{98} > R_{90} > R_{30} > R_{12} > R_{25} > R_{136} > R_{37}$
模糊网络分析法	$R_{92} > R_{56} > R_{15} > R_{70} > R_{91} > R_{98} > R_{90} > R_{30} > R_{12} > R_{25}$

4 结束语

本文利用数据信息通过 Apriori 算法挖掘数据关系,搭建了一个多目标作业车间复杂网络模型,该方法通过复杂网络的形式将生产中的数据组织起 (下转第 27 页)