

# 网络计划问题的度量、评价和算法选择

王国磊, 林琳, 钟诗胜

(哈尔滨工业大学 机电工程学院, 哈尔滨 150001, wanggl\_hit@163.com)

**摘要:** 为了研究不同算法处理网络计划问题的效果与网络计划问题的特征之间的关系, 在综合考虑网络图的拓扑结构、资源可替换性、截止日期等因素的基础上, 提出了柔性资源约束下网络计划问题的复杂度特征指标, 以及网络计划优劣的评价指标. 利用多元回归分析技术研究了不同算法处理网络计划问题的效果与网络计划特征之间的关系. 实验结果表明, 提出的特征指标能很好地度量网络计划问题的复杂性, 得到的回归方程也能比较准确的估计出算法的结果, 为根据网络计划问题的特征选择合适的算法提供可靠依据.

**关键词:** 网络计划; 复杂度; 回归分析; 算法选择

中图分类号: TP274

文献标志码: A

文章编号: 0367-6234(2010)03-0418-06

## Measurement, evaluation and algorithm selection on network planning problem

WANG Guo-lei, LIN Lin, ZHONG Shi-sheng

(School of Mechanical Engineering, Harbin Institute of Technology, Harbin 150001, China, wanggl\_hit@163.com)

**Abstract:** To investigate the treatment effect of different algorithms on network planning problem, the coefficients of network complexity (CNC) and network evaluation indices are proposed, taking the topological structure, interchangeability of resource and due date of the network into consideration. And then the relationship between the effects of different algorithms and the characteristics of network is analyzed based on multiple regression technology. The experimental results show that the proposed coefficients of network complexity are reasonable and the regression equations can accurately estimate the results of the algorithms, so that useful information can be provided to select a proper algorithm based on the characteristics of network planning problem.

**Key words:** network planning; complexity; regression analysis; algorithm selection

目前, 有关计划调度问题复杂性的研究主要有: Webster 等人分析了工件族调度问题的复杂性<sup>[1]</sup>; Yang 等人研究了面向订单的多机调度问题的复杂性<sup>[2]</sup>; Wu 等人从结构和操作两个角度对制造系统的复杂性进行了分析<sup>[3]</sup>. 有关网络计划问题复杂性的研究则有: Herroelen 最早提出了复杂度系数 (Complexity Index, *CI*) 的概念并得到了广泛应用<sup>[4-5]</sup>, 但是 *CI* 只适用于双代号网络图且算法比较复杂; 我国学者白思俊也提出了受资源约束的网络复杂性系数 *CRNC*<sup>[6-7]</sup>. 上述研究仅

仅考虑了计划任务本身的复杂性, 而对于其他外界因素, 例如交货时间、启动时间、资源互换性等则缺乏考虑.

本文针对有限柔性资源约束下的、带截止时间的多网络计划问题, 在综合考虑网络计划的结构、资源、时间等因素的基础上, 提出了度量网络计划的 4 种特征值. 进而, 拟定了网络计划结果的 3 种度量指标, 并基于此给出了计划结果度量公式. 最后, 利用多元线性回归方法分析了网络计划特征与不同算法结果之间的联系, 提出了根据不同网络计划问题的特征选取合适调度算法的策略.

### 1 问题描述

网络计划即为基于网络图所进行的计划工

收稿日期: 2008-12-23.

基金项目: 国家高技术研究发展计划资助项目(2008AA04Z401);

哈尔滨电机厂生产计划系统开发项目.

作者简介: 王国磊(1982—), 男, 博士;

钟诗胜(1964—), 男, 教授, 博士生导师.

作,本文中的网络图为单代号网络图,用圆圈表示任务,箭线表示任务之间的约束关系.需要特别指出的是:1)任务有自己的执行时间和所需资源,而任务有多种资源可以选择使用,称为资源柔性;2)每个网络图代表一个项目,因此具有自己的开始时间和截止时间;3)同一时间可能有多个项目存在,称为多网络图计划问题;4)由于资源柔性的存在,关键路径  $L_{KPL}$  的计算不同于普通网络图,而是取节点所有可用资源中,加工时间最短的一个,因为这是排产中最希望的结果.而多个项目存在时,关键路径长  $L_{KPL}$  显然为所有网络图合并成的大网络图的关键路径长,即

$$L_{KPL} = \max_{1 \leq i \leq N} \{L_i\} .$$

其中:  $L_i$  为网络图  $i$  的关键路径长.

## 2 网络计划的特性指标设计

为了分析网络计划算法的有效性和效率,人们一直希望能够精确地度量网络计划的复杂程度,即找到网络计划的复杂性指标 ( $C_{CNC}$ ).但是,由于网络图中存在着多种数量和质量因素, $C_{CNC}$  的精确度量方法一直没有定论.本文从4个不同方面对  $C_{CNC}$  进行衡量,除了网络图的拓扑结构因素外,还考虑了其他因素,对  $C_{CNC}$  的衡量更全面.

### 2.1 网络拓扑结构系数 $C_{CTS}$

网络的拓扑结构是衡量其复杂度的最重要因素之一.  $C_{CTS}$  与两个因素有关,一是关系数,二是节点数,而且具有以下特征:1)如果网络1和网络2同构,则二者复杂度相同,即  $C_{CTS}(1) = C_{CTS}(2)$ ;2)如果节点数相同,网络2比网络1约束多,则  $C_{CTS}(2) > C_{CTS}(1)$ ;3)由于仅由一条链组成的网络为最简单的网络,因此在节点数相同的情况下其复杂度最小;4)由于全连通网络为最复杂的网络,因此在节点数相同的情况下其复杂度最大.

由以上特征可知,  $C_{CTS}$  可以采用比较的方式引出,即将网络的关系数  $A$  与同节点数全连通网络的关系数  $A_{max}$  相比较.因此,对于  $N$  个节点的网络图,定义其拓扑复杂度为

$$C_{CTS N} = \frac{\ln A}{\ln A_{max}} = \frac{\ln A}{\ln(\sum_{i=1}^{N-1} (N-i))} . \quad (1)$$

其中:

$$A_{max} = \sum_{i=1}^{N-1} (N-i) .$$

式(1)显然满足  $C_{CTS}$  的全部特征,而且使全连通网络的复杂度为1.但式(1)不能衡量节点数不同的网络图的复杂度,所以还需要对式(1)进

行修正.根据信息论的观点,复杂度可以定义为系统表明自身方式数目的对数<sup>[8]</sup>,所以修正后的  $C_{CTS}$  为

$$C_{CTS} = \ln N \cdot \frac{\ln A}{\ln(\sum_{i=1}^{N-1} (N-i))} . \quad (2)$$

实际工程应用中,网络图节点通常不会超过100,根据式(2)计算100个节点之内的网络图复杂度并绘制图像,如图1所示.可以看出,  $C_{CTS}$  在0~2之间,且随网络图节点数和关系数的增加而增加.因此式(2)的定义是合理、可行的.

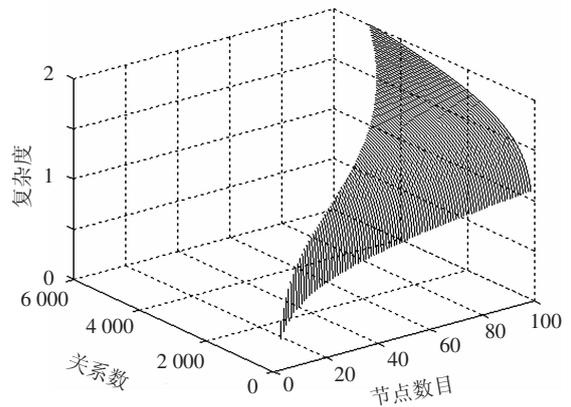


图1 100个节点内的网络拓扑复杂度对比

### 2.2 资源选择系数 $C_{CRS}$

对于网络计划问题,每个活动的可选资源越多,则排产的复杂性越高.本文定义式(3)为网络计划的资源选择系数 ( $C_{CRS}$ ),并将其作为衡量  $C_{CNC}$  的重要指标之一.

$$C_{CRS} = \ln \prod_{i=1}^N S_i / \ln N . \quad (3)$$

其中:  $S_i$  为节点  $i$  的可选资源数.显然,每个节点仅有一个可选资源时的资源选择度为0.

同样的,为了验证公式的正确性,计算并绘制100个节点内的网络图资源选择度曲线,如图2所示.

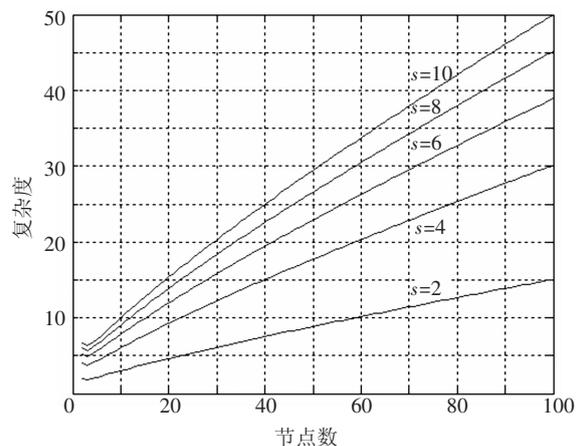


图2 100个节点内的资源选择度对比

图 2 中  $S$  是指每个活动的可选资源数量, 本文取  $S = 2 \sim 10$ , 基本覆盖了实际工程中可能的取值范围. 可以看出, 100 个节点以内的网络图资源选择度为  $0 \sim 50$ , 图 2 是假设所有任务节点都有  $S$  个可选资源的情况下计算得出的, 而实际工程应用中不可能每个任务都具有如此大的资源柔性, 因此, 实际工程应用中的资源选择度要更小一些, 一般为  $0 \sim 10$ . 由于式(3) 随可选资源种类的增多而增大, 而且取值范围不大, 便于实际应用, 所以, 资源选择度的定义也是较为合理、可行的.

### 2.3 松弛系数 $C_{CBT}$

在过去的研究中, 无论是从何种角度出发定义的网络图复杂度, 都仅就网络图本身出发, 没有考虑到更高层次的信息. 而实际上, 在计划时间比较宽松时和计划时间很紧张时的排产难度是不同的. 因此, 本文提出了松弛系数( $C_{CBT}$ ) 的定义, 用  $C_{CBT}$  作为衡量  $C_{CNC}$  的一项指标.

$$C_{CBT} = \frac{t_{dd} - t_{at} - L_{KPL}}{L_{KPL}}$$

其中:  $t_{dd}$  为网络计划的截止时间,  $t_{at}$  为网络计划的开始时间.

$C_{CBT}$  和截止时间有关, 因此, 此处不对  $C_{CBT}$  的大小详加讨论, 但是通常情况下都有  $C_{CBT} > 0$ , 否则说明网络计划的可用时间短于关键路径长, 该网络计划无法按期完成.

对于多个网络图的网络计划, 可以通过添加总的开始/结束节点将多个网络图整合成一个大网络. 但是, 在考虑计划截止时间时, 整合的方法因为每个项目的截止时间不同而失效. 因此, 本文给出了  $n$  个网络图的整体松弛系数计算公式:

$$C_{TCBT} = \frac{1}{n} \cdot \sum_{i=1}^n C_{CBT_i}$$

### 2.4 资源负荷系数 $C_{CRL}$

很多学者都对资源约束的网络计划问题的复杂度度量进行了研究, 认为通过网络资源总需求量与总供应量之间的比值来作为复杂性系数是比较合理的. 然而, 资源往往还具有柔性, 这给资源复杂度的度量带来了新的困难, 因此, 本文提出了资源负荷系数 ( $C_{CRL}$ ) 的概念.

1) 对于具有  $N$  个节点的网络图, 定义资源  $i$  的最大资源负荷系数为

$$C_{CRL_{max}}^i = \frac{\sum_{j=1}^N t_j^i}{L_{KPL} \cdot u_i}$$

其中:  $t_j^i$  为任务  $j$  使用资源  $i$  时的持续时间, 且如果任务  $j$  不能使用资源  $i$ , 则  $t_j^i = 0$ . 那么  $\sum_{j=1}^N t_j^i$  就表示

所有可能使用资源  $i$  的任务的总量, 即资源  $i$  的最大可能负荷; 而资源  $i$  单位时间可用量  $u_i$  与整体关键路径长  $L_{KPL}$  的乘积则表示资源  $i$  的总可用量. 那么二者的比值显然就可以作为资源  $i$  的最大资源负荷系数.

2) 定义资源  $i$  的最小资源负荷系数为

$$C_{CRL_{min}}^i = \frac{\sum_{j=1}^N ot_j^i}{L_{KPL} \cdot u_i}$$

其中:  $ot_j^i$  表示任务  $j$  使用资源  $i$  时的持续时间, 且如果任务  $j$  还可以使用其他资源, 则  $ot_j^i = 0$ . 那么  $\sum_{j=1}^N ot_j^i$  就表示所有必须使用资源  $i$  的任务总量, 即资源  $i$  的最小可能负荷.

3) 基于以上两种定义, 给出网络计划的整体资源负荷系数为

$$C_{TCRL} = \frac{1}{2M} \cdot \sum_{i=1}^M (C_{CRL_{max}}^i + C_{CRL_{min}}^i)$$

其中  $M$  是网络计划中用到的资源种类.

### 2.5 归一化

通过以上 4 种系数, 可以从不同角度刻画网络计划排产的困难度和复杂度, 然而由于 4 种系数的出发角度不同, 因此量纲也不同, 在使用前需要归一化:

$$x' = 0.1 + (0.9 - 0.1) \times \frac{x - \min(x)}{\max(x) - \min(x)}$$

## 3 计划优劣评价指标设计

### 3.1 长度系数 $C_{CL}$

由于不同计划问题的长度不一致, 因此采用式(4) 所示的计划长度系数作为评价指标.

$$C_{CL} = 1 - \frac{L_{KPL}}{L_{RL}} \tag{4}$$

式中:  $L_{RL}$  为计划实际长度. 可以看出, 当计划的实际长度等于关键路径长度时,  $C_{CL} = 0$ , 且随着计划实际长度的增大,  $C_{CL}$  也增大. 显然,  $C_{CL}$  越小越好.

### 3.2 拖期系数 $C_{CT}$

$$C_{CT_i} = \begin{cases} 0, & \text{if } L_{RL_i} < (t_{dd_i} - t_{at}); \\ \varphi_{TC_i} \cdot \left(1 - \frac{t_{dd_i} - t_{at}}{L_{RL_i}}\right), & \text{if } L_{RL_i} > (t_{dd_i} - t_{at}). \end{cases}$$

式中:  $\varphi_{TC_i}$  为网络计划  $i$  的拖期惩罚, 可以理解为权重. 显然, 式(4) 的含义是网络计划不拖期时, 拖期系数为 0, 而拖期越大, 拖期系数越大, 且有  $0 \leq C_{CT_i} \leq \varphi_{TC_i}$ .

当共有  $L$  个网络图存在时, 定义所有项目的

平均拖期系数为

$$C_{CT} = \sum_{i=1}^L C_{CT_i} / \sum_{i=1}^L \varphi_{TC_i}$$

显然,  $0 \leq C_{CT} \leq 1$ ,  $C_{CT}$  越大, 任务整体拖期越大, 计划的评价越差.

### 3.3 资源使用系数 $C_{CRU}$

为了评价算法对资源选择的优劣, 本文提出了资源使用系数  $C_{CRU}$  的概念,

$$C_{CRU} = \frac{R_{RU} - R_{RUmin}}{R_{RUmax} - R_{RUmin}}$$

其中  $R_{RU}$  表示网络计划排产后的实际资源使用总量,  $R_{RUmin}$  表示网络计划的最小资源使用量, 即网络中所有活动全部选择持续时间最短的资源时的资源使用总量, 而  $R_{RUmax}$  则表示网络计划的最大资源使用量, 即网络中所有活动全部选择持续时间最长的资源时的资源使用总量. 同样的,  $C_{CRU}$  也是越小越好, 而且有  $0 \leq C_{CRU} \leq 1$ .

### 3.4 评价函数

在评价指标多且不易横向对比时, 可以采用模糊综合评价法、层次分析法等<sup>[9-10]</sup>, 然而, 以上3种指标都与计划的评价成反比, 且均在  $0 \sim 1$  之间, 因此, 可以直接定义整体评价  $E$  为

$$E = 1 - \alpha_1 \times C_{CL} + \alpha_2 \times C_{CT} + \alpha_3 \times C_{CRU} \quad (5)$$

式中:  $\alpha_1, \alpha_2, \alpha_3$  是3种评价指标的权值, 可以根据不同偏好设定, 但必须满足式(6).

$$\sum_{i=1}^3 \alpha_i = 1 \quad (6)$$

式(5)从3个不同的角度对计划进行衡量, 能够比较全面地反映出计划的优劣, 而且计划评价价值满足  $0 \leq E \leq 1$ , 越大越好, 便于实际应用.

## 4 回归分析

### 4.1 回归方程的建立

首先建立如下回归分析模型:

$$Y = a + b_1 \times X_1 + b_2 \times X_2 + b_3 \times X_3 + b_4 \times X_4$$

其中:

$$Y = f_0(E), X_1 = f_1(C_{CTS}),$$

$$X_2 = f_2(C_{CRS}), X_3 = f_3(C_{TCBT}), X_4 = f_4(C_{TCRL}).$$

这样处理的原因是特征值与评价值之间不一定是线性关系, 要借用线性回归模型进行分析, 就要进行一定的处理, 化非线性回归为线性回归. 本文采用最常用的14种函数对自变量加以处理, 如表1所示.

为了确保回归方程有意义, 首先进行定性分析: 网络拓扑结构系数  $C_{CTS}$  越高, 排产时的约束越

表1 14种实验用变换函数

序号	函数 $f(x)$	序号	函数 $f(x)$
1	$y = x$	2	$y = x^2$
3	$y = x^3$	4	$y = x^4$
5	$y = x^{1/2}$	6	$y = x^{1/3}$
7	$y = x^{-1}$	8	$y = x^{-2}$
9	$y = x^{-3}$	10	$y =  x ^{-2}$
11	$y =  x ^{-3}$	12	$y = e^x$
13	$y = \ln x$	14	$y = e^{-x}$

多, 则得到计划的评价系数  $E$  就越小, 即  $C_{CTS}$  与  $E$  成反比; 松弛系数  $C_{TCBT}$  越小, 任务的完成期限越短, 那么得到的计划拖期度越大, 则评价系数  $E$  越小, 即  $C_{TCBT}$  与  $E$  成正比; 资源负荷系数  $C_{TCRL}$  越大, 说明设备的压力越大, 则得到计划的拖期度也越大, 评价系数  $E$  越小, 即  $C_{TCRL}$  与  $E$  成反比. 因此, 可以确定  $C_{TCBT}$  的回归系数  $> 0$ , 而  $C_{CNC}$  和  $C_{TCRL}$  的回归系数小于  $0$ . 此外, 为了保证回归方程中各参数足够明显, 限定回归系数绝对值大于  $0.05$ .

采用14种函数对各参数处理后, 回归分析可以看作是如下优化问题:

$$\begin{aligned} \max \quad & \{R, F\}, \\ \text{s. t.} \quad & |b_i| > 0.05, \quad i = 1, 2, 3, 4, \quad (7) \\ & b_1 < 0, b_3 > 0, b_4 < 0. \end{aligned}$$

其中:  $R$  为复相关系数,  $F$  为回归方程显著性  $F$  检验的方差检验值.

下面给出建立最优回归方程的算法步骤:

- 1) 采集网络图样本并计算网络特征值  $C_{CTS}$ ,  $C_{CRS}$ ,  $C_{TCBT}$ ,  $C_{TCRL}$ ;
- 2) 选择排产算法对所有样本进行排产, 并计算结果评价价值  $E$ ;
- 3) 选择处理函数  $f_0, f_1, f_2, f_3, f_4$ ; 如果所有组合都已尝试, 结束; 否则, 转步骤4);
- 4): 建立线性回归方程, 计算回归参数;
- 5) 如果回归参数符合式(7)约束要求, 转6); 否则, 转3);
- 6) 如果回归参数好于现有最优结果(拟和优度  $F$ 、复相关系数  $R$  最大), 则记录结果; 否则, 转3).

### 4.2 基于回归方程的算法选择

由于根据回归方程取得的估计值肯定与真实值之间存在一定的误差, 因此如何根据回归方程选择合适的算法也值得深入探讨, 本文提出了一种实验统计相结合的办法.

- 1) 用建立回归方程时所使用的样本进行试验, 即用回归方程得到各算法结果的预测值, 与各算法求得的实际评价价值进行比较, 得到对于估计

值取得最优的前  $\omega$  种算法,其实际值取得最优的概率  $P_\omega$ ,然后根据这个统计量对算法进行选择.例如,如果选择估计值最优的前 3 种算法,那么就能以  $P_3$  的概率保证取到实际最优算法.

2) 构造如式(8)所示的统计量  $Q$  来表示样本观察值与回归方程所得的估计值之间的偏差,

$$Q = \sum_{i=1}^n (D_i - \hat{D}_i)^2 \text{ 且 } Q/\delta^2 \sim \chi^2(n - m - 1). \tag{8}$$

由概率学可知,当样本  $n$  较大时,  $D_i - \hat{D}_i$  近似地服从正态分布  $N(0, \delta^2)$ ,其期望为 0,方差为

$$\delta^2 = Q/(n - m - 1).$$

将方法 1)和方法 2)相结合,就可在考虑偏差值  $\delta$  的基础上,通过估计评价值,选取估计值中最小的前若干种算法加以调用,大大减少计算量.

## 5 实验

### 5.1 生成样本

按照表 2 参数随机生成 200 个网络图,并按顺序编号.

表 2 网络图参数设定

参数	网络计划节点数 $N$	网络计划关系数	活动持续时间	活动的可用资源种类	网络计划缓冲时间
范围	5 ~ 100	$(N - 1) \sim 5(N - 1)$	1 ~ 10	1 ~ 5	$L_{KPL} \sim 1.5L_{KPL}$

### 5.2 建立回归方程

为了确保取得从最简单到最复杂的情况的所有数据,进行 200 次实验,其中第  $i$  次实验取前  $i$  个网络图作为计划输入数据,并根据表 3 所列 6 种排列算法分别安排项目进度,计算评价值.

表 3 试验用排产算法

算法	说明
MI	基于权重规则的启发式方法
SPT	基于短任务优先原则的启发式方法
FIFO	基于先到先服务原则的启发式方法
EDD	基于最早交货时间原则的启发式方法
MST	基于最短松弛时间原则的启发式方法
MSB	基于最多后续活动原则的启发式方法

得到样本数据后,按照 4.1 节所述方法建立最优线性回归方程,记录相关结果于表 4、表 5 中.表 4 记录的是最优回归方程中曲线函数的选择.以 EDD 算法为例,  $f_1 = 6$  为 EDD 算法的回归方程中,变换函数  $f_1$  取表 1 中的函数 6,即  $y = x^{1/3}$ .

表 4 变换函数选择

算法	$f_0$	$f_1$	$f_2$	$f_3$	$f_4$
MI	13	1	12	13	13
SPT	13	1	2	6	5
FIFO	13	4	3	6	13
EDD	13	6	2	13	13
MST	13	5	12	6	6
MSB	5	6	2	6	13

表 5 各方法的回归参数值

算法	$R^2$	$F$	回归系数	$1 - \alpha$ 置信区间	算法	$R^2$	$F$	回归系数	$1 - \alpha$ 置信区间
MI	0.928	916.3	$a = 0.495$	$[0.196, 0.795]$	SPT	0.990 8	3 476.1	$a = 0.186$	$[0.136, 0.236]$
			$b_1 = -0.258$	$[-0.419, -0.098]$				$b_1 = -0.374$	$[-0.444, -0.305]$
			$b_2 = -0.451$	$[-0.549, -0.353]$				$b_2 = -0.194$	$[-0.247, -0.142]$
			$b_3 = 0.192$	$[0.150, 0.233]$				$b_3 = 0.152$	$[0.085, 0.219]$
			$b_4 = -0.156$	$[-0.203, -0.110]$				$b_4 = -0.603$	$[-0.681, -0.525]$
FIFO	0.980	1 522.6	$a = -0.770$	$[-0.981, -0.559]$	EDD	0.989 0	2 893.6	$a = -0.253$	$[-0.420, -0.086]$
			$b_1 = -0.453$	$[-0.766, -0.140]$				$b_1 = -0.266$	$[-0.421, -0.112]$
			$b_2 = -0.205$	$[-0.275, -0.134]$				$b_2 = -0.416$	$[-0.476, -0.356]$
			$b_3 = 0.346$	$[0.212, 0.480]$				$b_3 = 0.073$	$[0.050, 0.097]$
			$b_4 = -0.265$	$[-0.346, -0.185]$				$b_4 = -0.254$	$[-0.290, -0.217]$
MST	0.992	4 124.0	$a = 0.836$	$[0.762, 0.909]$	MSB	0.990 1	3 233.4	$a = 0.795$	$[0.732, 0.858]$
			$b_1 = -0.281$	$[-0.388, -0.173]$				$b_1 = -0.159$	$[-0.215, -0.103]$
			$b_2 = -0.218$	$[-0.267, -0.169]$				$b_2 = -0.103$	$[-0.126, -0.080]$
			$b_3 = 0.143$	$[0.053, 0.233]$				$b_3 = 0.091$	$[0.056, 0.127]$
			$b_4 = -1.130$	$[-1.260, -0.994]$				$b_4 = -0.115$	$[-0.128, -0.103]$

而表5中记录的则是各算法回归方程的回归系数以及相关参数值.以SPT算法为例,其相关函数如下:

$$\ln(Eva) = 0.1859 - 0.3741 \times C_{CTS} - 0.1944 \times C_{CRS^2} + 0.1518 \times C_{TCBT}^{1/3} - 0.6032 \times C_{TCRL}^{1/2}$$

可以看出,表5中各方法的复相关系数 $R^2$ 均大于0.9,表明回归方程高度显著.采用 $F$ 检验,查 $F$ -分布临界值表可得, $F_{\alpha}(4,200) = 3.41$ ,可见对于置信度为99%( $\alpha = 0.01$ ),各算法的 $F$ 值都能满足条件 $F > F_{\alpha}$ .所以可以认为回归方程是高度显著的,但这不能说明回归方程中所有的自变量都对因变量有显著影响,因此还要对回归系数进行检验:由表5中回归系数的 $1 - \alpha$ 置信区间可以看出,各个方法的置信区间都不包含0,这说明整体评价与各回归系数的线性关系是显著的.

### 5.3 回归方程的准确性试验

为了验证回归方程的准确性,采用前述方法再生成50个测试样本,以SPT方法为例,首先算出计划评价的估计值,然后与SPT方法算出的实际评价价值相比较.如图3所示.可以看出,由回归方程估计得到的评价价值比较贴于真实值,最大误差不超过0.05.

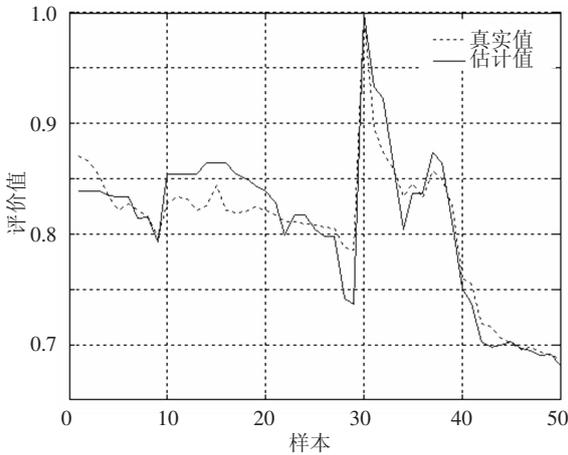


图3 回归方程的准确性检验

### 5.4 算法选择的可靠性试验

对于上述50个测试样本,根据4.2节的方法对算法进行选择,得到统计值如表6所示.例如,根据建立起来的回归方程选择估计值最优的前3种算法,能以92%的概率保证取到实际最优算法.这在算法较多的情况下可以节省大量的计算时间.

表6 测试样本统计值 %

$P_1$	$P_2$	$P_3$	$P_4$	$P_5$	$P_6$
66	82	92	96	98	100

## 6 结 论

1) 拓宽了对网络计划复杂度的研究.除了网络图本身的拓扑结构外,还考虑了资源可替换性、截止日期等因素,提出了4种更有实际应用价值的网络计划问题的特征指标,以及3种计划优劣的度量指标.

2) 提供了一种基于回归分析选择计划算法的思路和步骤.即根据回归方程预测、估计不同算法处理问题的效果,从而选择合适算法.

3) 下一步有待研究在特征值增多的情况下可能出现的多重共线性问题,以及其他预测、拟和方法的引入等.

### 参考文献:

- [1] WEBSTER S T. The complexity of scheduling job families about a common due date[J]. Operations Research Letters, 1997, 20(2):65-74.
- [2] YANG J. The complexity of customer order scheduling problems on parallel machines[J]. Computers & Operations Research, 2005, 32(7):1921-1939.
- [3] WU Y, FRIZELLE G, EFSTATHIOU J. A study on the cost of operational complexity in customer-supplier systems[J]. International Journal of Production Economics, 2007, 106(1):217-229.
- [4] REYCK B D, HERROELEN W. On the use of the complexity index as a measure of complexity in activity networks[J]. European Journal of Operational Research, 1996, 91(2):347-366.
- [5] AKKAN C, DREXL A, KIMMS A. Network decomposition-based benchmark results for the discrete time-cost tradeoff problem[J]. European Journal of Operational Research, 2005, 165(2):339-358.
- [6] 白思俊. 活动网络计划约束的复杂度度量及其应用[J]. 航空学报, 1994, 15(7):891-894.
- [7] 白思俊. 资源有限网络计划启发式方法的评价(下)——启发式方法与网络特征的相关性分析[J]. 运筹与管理, 1999, 8(4):52-56.
- [8] 黄欣荣. 复杂性究竟有多复杂?——论复杂性的测度[J]. 系统辩证学学报, 2005, 13(4):49-55.
- [9] 杜义贤,田启华. 基于神经网络的模糊综合评价方法[J]. 系统工程与电子技术, 2005, 27(9):1583-1586.
- [10] AL-HARBI K M A. Application of the AHP in project management[J]. International Journal of Project Management, 2001, 19(1):19-27.

(编辑 杨波)