

基于马尔科夫链模型的脉动装配线运行状态预测*

秦亚茹, 侯志霞, 吕瑞强, 刘毅萍

(中国航空制造技术研究院数字化制造技术航空科技重点实验室, 北京 100024)

[摘要] 分析了影响脉动生产线运行的因素和状态预测的必要性, 提出了基于马尔科夫链模型的状态预测方法。基于马尔科夫链预测理论, 对脉动生产线运行状态划分, 确定状态转移矩阵, 建立马尔科夫预测模型, 并对模型准确性进行验证。

关键词: 脉动生产线; 状态划分; 状态转移矩阵; 马尔科夫预测模型

Prediction on Pulse Production Line Running State Based on Markov Forecast Model

QIN Yaru, HOU Zhixia, LÜ Ruiqiang, LIU Yiping

(Aeronautical Key Laboratory for Digital Manufacturing Technology, AVIC Manufacturing Technology Institute, Beijing 100024, China)

[ABSTRACT] The current application of pulse production line is reviewed. The factors influencing pulse production line and the necessity of state prediction are analyzed. Then, a method of state prediction based on Markov chain model is proposed. Based on Markov chain forecast theory, the running states of pulse production line are classified. The state transition matrix is determined and Markov forecast model is established. Lastly, the validity of the model is estimated.

Keywords: Pulse production line; State division; State transition matrix; Markov forecast model

DOI: 10.16080/j.issn1671-833x.2017.05.092

传统飞机总装存在装配生产周期长、质量低、成本高、应变能力低等问题, 要解决此问题, 需要在总装配环节采用先进的技术和生产模式, 国外航空公司在飞机总装生产模式方面率先进行探索, 并取得一定应用成果。

波音公司 1999 年建立波音 737 移动式装配生产线, 总装时间由 22 天减少为 11 天, 最终缩短至 8 天, 并且工作流程产品存货降低 50%, 储备存货降低 59%。随着飞机总装移动式装配生产线的优势日益明显, 波音公司在波音 737、波音 757、波音 777 等机型移动式装配生产线的基础上, 为波音 787 建立了先进的脉动装配线。波音公司的成功经验带动了各大飞机制造商, 其中洛克希德·马丁飞机制造公司在 2003~2004 年建成了 F-35 总装脉动装配线, 大大缩短飞机总装时间, 到 2016 年 F-35 的总装脉动装配线已达到每天交付一架的水平^[1]。

脉动式移动装配生产线简称脉动装配线, 是连续式移动装配线的过渡阶段。与传统装配方式相比, 脉动装配线有以下优势^[2]: (1) 改善装配现场环境, 缩小厂房占用面积, 减少飞机生产投入; (2) 装配作业标准化、专业化的分工更易于保证飞机的质量和性能的稳定; (3) 规

范和拉动整个飞机的生产供应链的协同; (4) 采用新工艺方法和新装备技术, 能够促进飞机装配技术的发展。

根据我国航空企业的管理模式、工艺水平、技术现状以及产能需求, 目前脉动装配线更加符合我国国情。中航工业西安飞机工业(集团)有限责任公司建立了国内第一条总装脉动装配线, 从立项到最终建成, 经历了 5 年的调研论证与动工建设, 2010 年 5 月基本建成并投入使用, 实现国内飞机制造水平的显著提升。目前各大主机厂对脉动装配线规划和管控系统进行设计, 但是对脉动装配线状态预测技术还未开展研究。状态预测技术能够提前发现问题, 减少脉动装配线出问题的频率, 降低对装配线关键指标的影响, 因此对脉动装配线状态预测技术进行研究十分必要。

1 基于脉动装配线的状态预测必要性研究

1.1 脉动装配线特点和影响因素分析

脉动装配线具有以下特点: (1) 工人位置固定、分工明确精细、动作标准化, 同时可移动、自动化程度高的轻便工装代替传统固定式工装, 生产效率较高; (2) 配备专业供给线, 物料配套准确、配送精确到位; (3) 对设备实现集中、高效、安全智能管理; (4) 收集和回馈生产线

* 基金项目: 中国航空工业集团公司技术创新基金项目(KC371503)。

的生产管理信息及报警信息,建立问题快速响应机制。基于对脉动装配线特点的分析,影响脉动装配线运行状态因素包括:人员、物料、设备、工序、工具等,例如物料配套准确率、配送准时率、设备故障率、工序交检合格率等都会对脉动装配线节拍、产能等关键指标产生影响。

1.2 状态预测的必要性

脉动装配线在运行过程中经常会出现问题,为了不影响节拍、产能等关键指标,发挥脉动装配线优势,需要预先综合分析人员、物料、设备、工序等因素对运行状态造成的影响,对脉动装配线进行状态预测。状态预测的目的是根据脉动装配线状态变化规律和当前状态,预测未来运行状态,提前发现问题,利用或改进未来状态,最大程度降低对脉动装配线造成的影响。本文利用马尔科夫方法对脉动装配线运行状态进行预测,马尔科夫方法基于状态变化的规律性和关联性,利用事物过去的变化规律推测出未来的变化趋势^[3],预测的关键是建立马尔科夫链模型。

2 基于马尔科夫链模型的脉动装配线状态预测研究与应用

2.1 马尔科夫链模型的适用性

影响因素之间存在错综复杂的联系,很难运用结构化的模型解释,数据之间的相互依存关系是研究对象的重要特性,因此根据数据的变动规律建立时间序列模型是一种有效的方法。马尔科夫链是一种时间序列,通过分析数据之间的关系,形成状态转移矩阵,进而分析事件的发展趋势。马尔科夫链具有无后效性,将来取值只与现在的取值有关,即某个阶段的状态一旦确定,此后过程的演变不再受此前各种状态及决策的影响。脉动装配线的运行状态是一个有限非平稳时间序列,具有趋势性与随机性,生产线在每个时段所处的状态是随机的,状态之间按照一定的概率进行转移,下个时段的状态由当前状态以及转移概率决定,状态指标参数量值在时间轴上是离散的,因此构成了典型的马尔科夫链^[4]。在脉动装配线运行状态变化的过程中,通过输入某一时刻状态分布,利用马尔科夫链模型预测出下一周期状态分布。

2.2 马尔科夫链模型概述

通常情况下,一个马尔科夫链模型由3元素表示,即 $\lambda = (X, \pi, A)$,其中: X 用来表示模型中的状态,状态的数量用 N 表示,即 $X = \{x_1, x_2, \dots, x_N\}$, N 个状态之间根据具体的实际问题而存在一定联系; $\pi = \{\pi_i\}$, $i = 1, 2, \dots, N$, π 表示状态空间的初始状态分布,具体的含义是初始条件下处于各个状态的数量;用 $A = (p_{ij})_{N \times N}$ 来表示状态转移矩阵,其中 p_{ij} 表示在时刻 $t-1$ 时,模型的状

态为 x_i ,在时刻 t 时模型的状态转移到 x_j ,此时 $p_{ij} = P(x_j | x_i)$ 且转移矩阵 A 的每一行元素之和等于1,即为 $\sum_{j=1}^N p_{ij} = 1 (i=1, 2, \dots, N)$,由于整个模型中有 N 种可能的状态,因此转移矩阵是 N 阶的方阵。

马尔科夫链模型如图1所示。

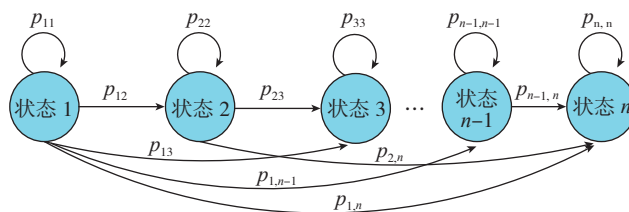


图1 马尔科夫链模型

Fig.1 Model of Markov chain

时间和状态都是离散的马尔科夫过程称为马尔科夫链,即 $X_n = X(n)$, $n = 0, 1, 2, \dots, K$ 。马尔科夫链是随机变量 $X_1, X_2, X_3, \dots, X_n$ 的一个序列,这些变量所有可能取值的集合构成状态空间,则 X_n 的值是时间为 n 的状态。马尔科夫性质描述为 X_{n+1} 对于过去状态的条件概率分布仅是 X_n 的一个函数, x 为过程中的某个状态,如式(1)所示。

$$P(X_{n+1} = x | X_0, X_1, X_2, \dots, X_n) = P(X_{n+1} = x | X_n) \quad (1)$$

马尔科夫链满足两个条件:(1) $t+1$ 时刻系统状态的概率分布只与 t 时刻的状态有关,与 t 时刻以前的状态无关;(2)从 t 时刻到 $t+1$ 时刻的状态转移与 t 值无关。

本文研究的是离散状态空间的马尔科夫模型,则状态转移矩阵即为条件概率:

$$p_{ij} = P(X_{n+1} = j | X_n = i) \quad (2)$$

对于一个离散状态空间,对于一步状态转移,得到下一步状态为:

$$P\lambda = \lambda^* \quad (3)$$

2.3 基于马尔科夫模型的生产线状态预测应用

通过对某飞机制造厂某型号飞机现场调研,在生产过程中通过对装配过程中采集到的工序完工率、交检合格率以及物料配套正确率等数据进行综合分析,预测脉动线运行状态。对工序进度、工序完成情况、工序报警等信息进行统计与存储,统计出一定时间内的工序完工率以及交检合格率。根据物料报警信息,可以统计一定时间内的物料配套正确率。本文研究所用算例将脉动装配线分成13个站位,基于现场调研和结合生产情况,通过采集6个周期的运行数据,进行状态划分,求取状态转移矩阵,进而对运行状态进行预测。

2.3.1 状态划分

工序的按时完工率、物料配套的正确率、工序质量合格率3个指标参数的平均值作为脉动装配线运行状态的指标参数 y_j ,同时也作为马尔科夫链状态划分的依据。依据指标参数 y_j 的数值,将状态分为 S_1, S_2, S_3, S_4 4种, S_1 表明运行状态很好,状态指标参数范围为 $[0.85,1]$; S_2 表明运行状态较好,状态指标参数范围为 $[0.7,0.85]$; S_3 表明运行状态一般,状态指标参数范围为 $[0.6,0.7]$; S_4 表明运行状态较差,状态指标参数范围为 $[0,0.6]$ 。

2.3.2 脉动装配线运行状态预测的马尔科夫链模型

本文建立具有4个状态的马尔科夫链模型,任一个状态只能向其自身或之后的状态转化,如图2所示。13个站位状态指标参数如表1所示。

以第5周期状态为初始分布状态,对前5周期数据进行统计,利用3种方法求取状态转移矩阵,即基于频数统计的方法、基于平均值方法和基于优化的方法。根据状态指标参数,得到相邻两个周期之间状态转移频数,如表2所示。

(1) 基于频数统计方法的状态转移矩阵^[5]。

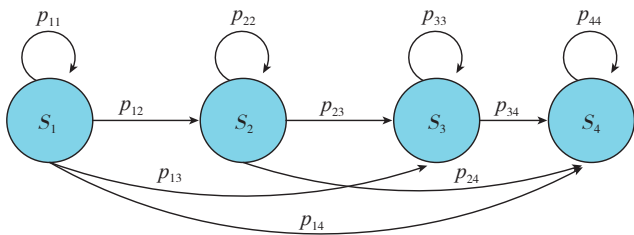


图2 脉动装配线运行状态预测马尔科夫链模型
Fig.2 Model of Markov chain about prediction on pulse production line running state

表1 不同站位状态指标参数 %

站位	第1周期	第2周期	第3周期	第4周期	第5周期	第6周期
站位1	69.54	62.55	84.31	87.95	73.58	81.27
站位2	67.85	69.93	63.56	84.57	75.67	83.75
站位3	68.32	65.32	68.29	80.92	80.35	78.46
站位4	75.48	59.67	83.45	90.73	92.10	59.72
站位5	68.27	68.27	65.67	69.34	83.49	83.39
站位6	84.56	80.99	58.33	90.24	90.34	92.21
站位7	92.82	64.86	69.05	69.11	82.94	83.55
站位8	63.33	67.44	83.29	80.79	84.59	67.62
站位9	68.12	69.53	68.24	63.80	68.56	58.72
站位10	90.07	94.01	57.80	58.32	69.37	59.07
站位11	91.86	87.45	59.68	69.72	67.55	82.75
站位12	66.94	69.50	67.88	69.88	66.73	69.74
站位13	90.78	87.22	59.63	68.39	68.92	69.66

表2 相邻周期状态转移频数

状态	1,2 周期				2,3 周期				3,4 周期				4,5 周期			
	S_1	S_2	S_3	S_4	S_1	S_2	S_3	S_4	S_1	S_2	S_3	S_4	S_1	S_2	S_3	S_4
S_1	3	1	0	0	3	0	0	0	1	2	0	0	0	1	0	0
S_2	0	7	0	0	0	6	2	0	0	4	2	0	0	4	2	0
S_3	0	0	1	1	0	0	0	1	0	0	1	2	0	0	3	0
S_4	0	0	0	0	0	0	1	0	0	0	0	1	0	0	1	2

通过对前5周期状态转移频数求和,得到前5周期状态转移频数,如表3所示。

表3 前5周期状态转移频数

状态	S_1	S_2	S_3	S_4
S_1	7	4	0	0
S_2	0	21	6	0
S_3	0	0	5	4
S_4	0	0	2	3

通过统计各个状态间转移的频数,得状态转移概率矩阵:

$$P_{count} = \begin{pmatrix} 7/11 & 4/11 & 0 & 0 \\ 0 & 7/9 & 2/9 & 0 \\ 0 & 0 & 5/9 & 4/9 \\ 0 & 0 & 2/5 & 3/5 \end{pmatrix}$$

(2) 基于平均值方法的状态转移矩阵。

通过表2统计出相邻周期的状态转移矩阵:

$$P_{i1} = \begin{pmatrix} 3/4 & 1/4 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$P_{i2} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 3/4 & 1/4 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

$$P_{i3} = \begin{pmatrix} 1/3 & 2/3 & 0 & 0 \\ 0 & 2/3 & 1/3 & 0 \\ 0 & 0 & 1/3 & 2/3 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$P_{i4} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 2/3 & 1/3 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 2/3 & 1/3 \end{pmatrix}$$

通过求取平均值得平均状态转移矩阵:

$$P_{mean} = \begin{pmatrix} 0.5208 & 0.4792 & 0 & 0 \\ 0 & 0.7708 & 0.2292 & 0 \\ 0 & 0 & 0.4583 & 0.5417 \\ 0 & 0 & 0.3333 & 0.6667 \end{pmatrix}$$

(3) 基于优化方法的状态转移矩阵^[6]。

现假定 $q(t) = (p_i(1), p_i(2), \dots, p_i(n))$ 是在 t 时刻的状态分布, 其中 $t=0, 1, 2, \dots, m$, 由于外在因素的影响, 一步状态转移矩阵在相邻时间不完全相同, 因此得到的 $q(t+1)$ 和 $q(t)$ 之间必然存在误差, 要使误差降到最低, 可构造下面的优化模型:

$$f(p) = \min \sum_{t=0}^{m-1} \|q(t+1) - q(t)\|^2 = \min \sum_{t=0}^{m-1} (q(t+1) - q(t))(q(t+1) - q(t))^T \quad (4)$$

以状态转移概率矩阵与理论误差平方和最小为原则, 建立状态转移矩阵的优化模型。设一步状态转移矩阵为:

$$P_0 = \begin{pmatrix} x_1 & x_2 & 0 & 0 \\ 0 & x_3 & x_4 & 0 \\ 0 & 0 & x_5 & x_6 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

定义 $P_{01}, P_{02}, P_{03}, P_{04}$ 分别为:

$$P_{01} = (p_1(1), p_1(2), p_1(3), p_1(4))$$

$$P_{02} = (p_2(1), p_2(2), p_2(3), p_2(4))$$

$$P_{03} = (p_3(1), p_3(2), p_3(3), p_3(4))$$

$$P_{04} = (p_4(1), p_4(2), p_4(3), p_4(4))$$

目标函数为:

$$f(x_i) = \min \sum_{i=1}^4 \sum_{j=1}^4 (P_i(j) - P_0(j))(P_i(j) - P_0(j))^T \quad (5)$$

$$s.t. \begin{cases} x_1 + x_2 = 1 \\ x_3 + x_4 = 1 \\ x_5 + x_6 = 1 \\ x_1, x_2, x_3, x_4, x_5, x_6 \geq 0 \end{cases} \quad (6)$$

优化模型的状态转移矩阵为:

$$P_0 = \begin{pmatrix} 0.5076 & 0.4924 & 0 & 0 \\ 0 & 0.7852 & 0.2148 & 0 \\ 0 & 0 & 0.3363 & 0.6637 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

2.4 结果分析

第5周期的状态数 $\lambda_5 = (2, 6, 5, 0)$ 为初始状态, 利用公式(3)得到频数统计模型、平均值模型、优化模型对第6周期状态分布的预测结果 $\lambda_{count} = (1.2727, 5.3939, 4.1111, 2.2222)$ 、 $\lambda_{mean} = (1.0417, 5.5833, 3.6667, 2.7083)$ 、 $\lambda_{60} = (1.0152, 5.6960, 2.9703, 3.3185)$ 。

第6周期真实的状态数为 $\lambda_6 = (1, 6, 3, 3)$, 将预测结果与真实状态数进行对比, 如图3所示, 误差如表4所示。由图3得出, 利用优化模型得出的状态分布与真实状态分布相似度最高; 由表4得知, 优化模型的相对误差整体最小, 因此利用优化模型对生产线运行状态进行预测准确率相对较高。

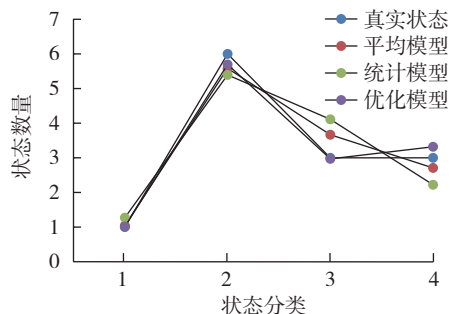


图3 预测结果分析

Fig.3 Analysis of prediction results

表4 预测模型误差比较

比较项	S_1	S_2	S_3	S_4
真实状态	1	6	3	3
统计模型误差 /%	27.27	-10.10	37.04	-25.93
平均模型误差 /%	4.17	-6.954	22.22	9.72
优化模型误差 /%	1.52	-5.07	0.99	10.62

2.5 建模方法评价

基于频数统计方法建立模型的前提是将频率等同于概率, 在基于大量试验数据的前提下, 频率近似等于概率, 由于本文研究数据样本较小, 利用频数统计方法建立模型结果误差偏大。通过对多个状态转移矩阵求平均值, 能够得到一步状态转移矩阵近似情况, 不能准确反映实际情况。在基于平均值模型的基础上, 以理论状态概率向量与实际状态概率向量平方和最小为原则, 建立优化预测模型, 发现优化预测模型相对误差较小, 预测结果相对准确。

3 总结与展望

论文利用马尔科夫方法建立马尔科夫链预测模型, 并对脉动装配线运行状态进行预测, 模型算法简单, 预测结果较为准确, 今后将开展进一步研究工作。

(1) 本文利用一个算例提出了脉动装配线运行状态预测方法, 在实际生产中要对该方法进行检验, 同时需要进一步研究状态划分方法, 更加准确地对脉动装配线运行状态进行划分, 并且根据运行状态的波动性, 对模型进行修正调整。

(2) 在实际生产中, 影响脉动装配线运行因素很多, 并且不同的用户关注的影响因素不同, 因此通过引入隐马尔科夫模型, 将影响因素组成可观测序列输入到模型中, 通过改变可观测序列的元素以及长度来实现输入不同影响因素的目的。脉动装配线运行状态作为隐状态, 影响因素作为可观测序列, 建立可观测序列与隐状态之间的对应关系以及各个状态模型。通过输入一

(下转第104页)