

印刷企业生产数据的智能演化分析与应用

刘永军

(石家庄经济学院 信息工程分院 河北 石家庄 050031)

摘要: 印刷生产企业的资源管理和生产流程管理基于 ERP 系统,但缺乏对大量数据的分析和挖掘。采用智能演化思想,运用先进的算法模型、MATLAB 强大的数据处理能力及 MATLAB 相关工具箱的强大数据分析能力,建立数据挖掘的手段和分析的模型,进行数据的演化和全局性直接搜索,在企业数据分析的宏观层面和特定领域探索其良好的应用,获得了较好的方法和途径。

关键词: ERP 系统; 数据分析; 智能计算; 全局搜索; 遗传算法专业工具箱

中图分类号: TP311.11 **文献标志码:** A **文章编号:** 2095-0373(2014)01-0095-05

0 引言

印刷企业的企业资源计划 ERP 在企业印刷业务管理方面为企业提供了现代化的信息管理手段,企业借助 ERP 系统实现生产信息管理的现代化,做到各类数据能够及时、准确的传递,大大减少人为主观因素和人为差错,并提高了企业生产率。ERP 系统实现了企业生产信息的采集输入、存储、传送、应用等,使用多年并形成并积累了大量的历史数据数据。

但利用 ERP 系统积累的数据进行分析,发挥数据的更大作用,是对企业各种数据信息的进一步应用,通过企业大量数据的搜集形成数据仓库,对各类历史数据信息的智能化分析、挖掘、处理,找到数据信息内在的关联、变化,实现对企业的生产进行更好地控制和管理,提高企业的生产能力、管理能力、决策水平、资金使用效率,减少生产的盲目性、主观性,从而避免影响企业生产和效率的一些错误。

以遗传算法为基础的演化计算对数据的操作分析,并综合运用演化计算方法和传统数据处理手段、人工经验、Matlab 平台与专业工具箱等,实现数据挖掘和智能处理的目标。

1 数据挖掘技术与遗传演化方法

数据挖掘是从大量的、不完全的、有噪声的、模糊的、随机的实际应用数据中,提取隐含在其中的、人们事先不知道的、但又是潜在有用的信息和知识的过程。数据的分析与挖掘主要侧重分类、聚类、关联、预测等问题的解决^[1]。尤其在以数据库为基础的数据关联和预测,用于把世界上存在着内在的某种关联以及对未来数据趋势的预测作为发现的重点。实现这些功能的技术有比较成功的经验^[2],方法涉及关联规则方法、MBR 方法、聚集检测、连接分析、决策树和规则推理、神经网络、遗传算法、多目标线性规划方法、支持向量机方法等。

遗传算法是模仿自然界生物进化机制发展起来的随机全局搜索和优化方法,其思想源于生物遗传学和适者生存的自然规律,原理借鉴基因遗传变异和自然选择淘汰的生物进化过程,是具有“生存+检测”的迭代过程的搜索算法,用计算机模拟大自然的这一演化过程,特别是生物的进化过程,来求解复杂问题的一类计算模型^[3]。

DOI: 10.13319/j.cnki.sjztdxxbzb.2014.01.19

收稿日期: 2013-01-07

作者简介: 刘永军 男 1970 年出生 副教授

基金项目: 石家庄市科技局科技攻关项目(1011335A); 河北省科技支撑计划项目(14211810D)

1.1 遗传算法的基本原理和过程

遗传算法以一种群体中的所有个体为对象,并利用随机化技术指导对一个被编码的参数空间进行高效搜索。其中,选择、交叉和变异构成了遗传算法的遗传操作;参数编码、初始群体的设定、适应度函数的设计、遗传操作设计、控制参数设定等 5 个要素组成了遗传算法的核心内容^[4]。

1.1.1 遗传算法的基本过程

遗传算法中有两类运算,即遗传运算(交叉和变异);进化运算(淘汰和选择)。运用基本遗传算法进行问题求解的过程如下:

(1) 编码。在进行搜索之前先将解空间的解数据表示成遗传空间的基因型串结构数据,这些串结构数据的不同组合便构成了不同的可行解。

(2) 随机产生初始群体。随机产生 N 个串结构数据,每个串结构数据称为一个个体, N 个个体构成了一个初始群体。

(3) 适应度评估检测。针对特定问题定义适应性函数,根据适应性函数计算每个个体的适应度。适应度是反映个体或解的优劣性,遗传算法就是要寻得适应度最大的个体^[5]。

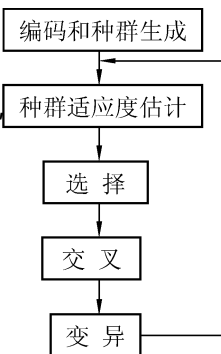
(4) 选择。通过选择从当前群体中选择出优良的个体,使它们有机会作为父代为下一代繁殖子孙,遗传算法通过选择算子作用于群体,进行选择的原则是适应性强的个体为下一代贡献一个或多个的概率大,体现了进化论的适者生存原则。

(5) 交叉。交叉操作体现了信息交换的思想,通常取值在 0.85 ~ 0.95^[6];通过交叉算子操作得到新一代个体,新个体继承了其父辈个体的特性。

(6) 变异。变异操作是首先在群体中随机选择一个个体,对于选中的个体以一定的概率随机地改变串结构数据中某个位置的值,通常取值在 0.001 ~ 0.05,这样变异为新个体的产生提供了机会。

1.1.2 遗传算法的一般步骤

遗传算法的一般步骤如图 1 所示。遗传算法模拟基因重组与进化的自然过程,把待解决问题的参数编成二进制或十进制码,即基因,若干基因组成一个染色体,许多染色体进行类似于自然选择、配对交叉和变异的运算,经过世代遗传,即多次重复迭代,直至得到最后的优化结果。



1.2 遗传算法实现数据的分析处理

Matlab 具有矩阵、数组等高效数据运算、处理能力,加上 Matlab 优化工具箱中的 GA 求解器和直接搜索工具箱,为演化设计提供了有力的支持。选择库存有关的订单、入库、出库金额情况,并按月汇总,对生产数据运用智能计算与分析的手段,在宏观层面进行企业资金有效利用方面的数据挖掘和决策支持。

(1) 数据的读取与规律分析。数据以 ERP 系统导出的出库、入库、订单等金额信息数据为依据,使用三年的历史数据作为分析对象,把第四年的数据作为分析结果的检验、预测和比对。先把三年的历史数据按月进行清洗、汇总,把当月各个单品材料或单个订单的价值金额等数据作为演化分析的对象。

在用函数 xlsread() 读取数据基础上,对入库金额、出库金额、订单金额这些反应整体情况的数据,及一些单品消耗材料的数据,如液化气、扎捆绳、黑墨等进行图形化显示分析,找到数据变化的规律。通过 Matlab 对各种数据进行折线图的绘制,如材料入库、材料出库、订单、单品材料的出入库等数据呈现一定的波动性和周期性,采用傅里叶级数函数能够很好地反映这些数据的综合特性。

(2) 数据的拟合。傅里叶级数函数的三角函数表达式

$$f(t) = A_0 + \sum_{n=1}^{\infty} (A_n \cos(n\omega_0 t) + B_n \sin(n\omega_0 t)) \quad (1)$$

式中, $A_0 = \frac{1}{T_0} \int_t^{t+T_0} f(t) dt$; $A_n = \frac{2}{T_0} \int_t^{t+T_0} f(t) \cos(n\omega_0 t) dt$; $B_n = \frac{2}{T_0} \int_t^{t+T_0} f(t) \sin(n\omega_0 t) dt$; $\omega_0 = 2\pi/T_0$, T_0 为数据信息的周期。

图 1 遗传算法的一般步骤

对各种数据在 Matlab 上,运用其函数编程分别对各种数据进行分析、拟合等操作,得到其拟合图形和表达式函数。

从式(1)计算得出 2010 年 10 个月的预测值,与实际发生的数值比较,平均误差约为 10%,对实际数据的偏差在预测的合理范围,能够有效指导企业生产计划和活动。同样方法可以得到其他数据不同参数的曲线、函数表达式以及相关的预测参考数据。拟合对应的函数曲线如图 2 所示,其中纵坐标为函数值大小(单位:元),横坐标为 2010 年以前 3 年的月份顺序。

2 智能演化分析

印刷加工企业合同订货和销售、材料总出入库以及各种单品材料的采购与使用等数据是相互关联的,建立相应的分析模型,便于分析资金的最优使用和生产的规律。

2.1 数据的分析模型

斯塔科尔伯格(Stackelberg)均衡模型是一个领导者-跟随者模型,领导者决定自己的数量,然后跟随者可以观察到这个数量,并根据观察到的这个数量决定自己的数量,领导者决定自己的数量的时候,充分了解跟随者会如何行动,即跟随者的反映函数,因此,领导者自然会预期到自己的决定对跟随者的影响。在

印刷企业的生成中涉及到的数据的实际,考虑以保障生产所需最佳库存保障为优化目标,生产所需出库(领导者)的函数 f 成为材料入库反应函数 g 的约束,从而得到库存最优资金占有的模型

$$E = \int_a^b g(V) dV - \int_a^b f(V) dV \quad (2)$$

式中, g 和 f 是根据观察的历史数据拟合得到,找到使得 E 为最小值即为保障生产所需的最小库存金额,是生产资金宏观的最优资金使用模式。当然,根据具体问题的需要,也可以获得其最大值,即传统决策论中的“最小最大”准则,最小最大定理所体现的基本思想是“抱最好的希望,做最坏的打算”。按照斯塔科尔伯格均衡模型,材料总库存金额与出库总金额关联,得到总库存分析模型

$$F_{\text{库}}(x) = F_{\text{总入}}(x) - F_{\text{总出}}(x) \quad (3)$$

同样可以得到其他数据的分析模型。

2.2 智能的演化

采用 Matlab 平台上的遗传算法直接搜索优化工具箱模拟达尔文生物进化论的自然选择和遗传学机理的操作,建立基于企业生产过程宏观的数据分析的计算模型。在 Matlab 命令窗口输入 gatool,即可打开、进入遗传算法工具的 GUI 界面,针对解决的问题需要编写待优化函数的 M 文件,作为欲求最小值的目标函数,在 GA GUI 界面中的适应度函数文本框输入@ 目标函数名后,输入选择变量个数以及各种演化参数,诸如种群的类型、尺度、创建函数、初始种群、初始得分、初始范围,以及适应度测量、选择、复制、变异、交叉、迁移、混合函数、停止准则、输出函数、命令窗口显示、向量化等,在相应的窗格设置参数。

程序运行得到结果如下:演化代数 100;种群规模 50;交叉概率 0.900;变异概率 0.050;找到的最大值 $[x \ f(x)]$ 为 [17.254 0, 16.972 043.025 6];演化时间 6.046 s。

编程实现对演化过程和结果的图形化显示,如染色体初始位置、染色体最终位置、以及最优、平均函数值变化趋势等信息,如图 3 所示,其中(a)和(b)纵坐标为函数值大小(元),横坐标为 2010 年以前 3 年的月份顺序,(c)纵坐标为函数值大小(元),横坐标为演化代数。

图 3(a)显示了演化初始时随机生成的 50 个种群染色体的位置,按照程序和参数的设置进行 100 代演化,动态变化形成图 3(b)的演化最终位置,得到最优的结果,从图 3(c)观察到目标函数值随演化的进行的平均值变化趋势和逐步趋于稳定的最优值。

如果把表达式取反进行相同的演化,所得结果再取反就可以得到最小值。程序运行得到结果:演化代数 100;种群规模 50;交叉概率 0.900;变异概率 0.050;找到的最小值 $[x \ f(x)]$ 为 [1.563 8,

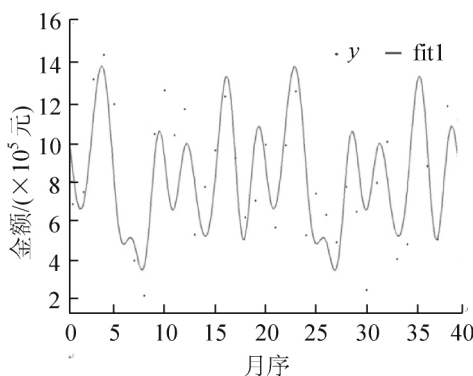


图 2 材料入库金额数据拟合曲线

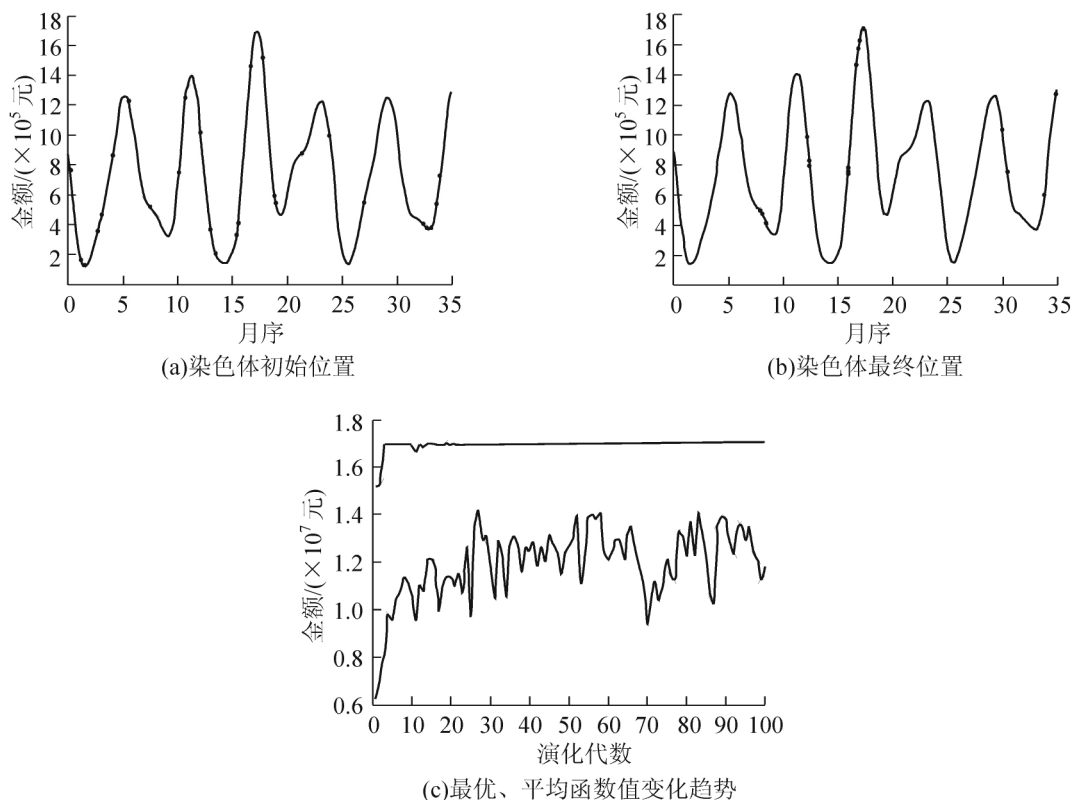


图 3 GA 程序执行的过程与图形输出结果窗口

1 258 927.665 5]; 演化时间 6.129 s。

所得结果 [1 258 927.665 5, 16 972 043.025 6] 为 $F_{\text{订-入}}(x) = F_{\text{订货}}(x) - F_{\text{总入}}(x)$ 的最值范围, 是当月合同订货金额与材料总入库金额差的范围, 为当月生产的计算利润范围。采用同样方法, 根据其他数据的分析模型, 可以得到库存情况、单品材料情况、采购、资金利用等结果, 为宏观决策提供一定的依据。

3 结语

针对印刷企业的管理数据信息, 提出遗传算法分析手段, 结合 Matlab 专业工具箱和经济领域的模型理论等对目标函数有很好的分析功能, 结合经济领域的斯塔科尔伯格模型形成的关于印刷企业各种数据模型的关联, 在材料采购、材料的使用、资金的准备、利用、库存的状态、合同的履行及生产组织等等方面可以有很好的应用, 为企业的宏观管理、决策角度实现合同订货、材料总采购入库、材料总出库、合同交货间的关联和制约, 材料总采购入库、单品材料的入库、单品材料的出库、材料总出库之间的关联和制约, 而库存占用资金、生产资金、各单品材料资金占用形成相应的关联和制约。

参 考 文 献

- [1] 邵峰晶. 数据挖掘原理与算法 [M]. 北京: 科学出版社, 2009: 55-103.
- [2] David M Smith. MATLAB 工程计算 [M]. 北京: 清华大学出版社, 2008: 30-46.
- [3] 黄竞伟, 朱福喜, 康立山. 计算智能 [M]. 北京: 科学出版社, 2010: 20-48.
- [4] 张军, 詹志辉. 计算智能 [M]. 北京: 清华大学出版社, 2009: 10-33.
- [5] A Baykasoglu. Gene expression programming based meta-modelling approach to production line design [J]. International Journal of Computer Integrated Manufacturing, 2008, 21(6): 657-665.
- [6] Adil Baykasoglu. MEPAR-miner: Multi-expression programming for classification rule mining [J]. European Journal of Operational Research, 2007, 183: 767-784.

(下转第 110 页)

Springer-Verlag ,1990.

[5] Arnold V I. Mathematical Methods of Classical Mechanics [M]. Berlin: Springer , 1978.

[6] Gu Zhuquan ,Zhang Junxian ,Liu Wei. Two new completely integrable systems related to the KdV equation hierarchy [J]. IL NUOVO CIMENTO 2008 ,123(5) : 606-622.

[7] Zhao Ye ,Gu Zhuquan ,Liu Yafeng. The Neumann system for the 4th-order eigenvalue problem and constraint flows of the coupled KdV-type equations [J]. Eur. Phys. J. Plus 2012 ,127: 77.

A New (2 +1) -dimensional Nonlinear Evolution Equation and the Involutive Representations of the Solutions

Liu Yafeng , Liu Wei

(Dept of Math & Phys , Shijiazhuang Tiedao University , Shijiazhuang 050043 , China)

Abstract: A new (2 +1) -dimensional nonlinear evolution equation is obtained based on the compatible condition of two linear spectral problems. By means of the constraint condition between the potentials and the eigenvector , the Bargmann system is generated. The Jacobi-Ostrogradsky coordinate system has been found through Euler-Lagrange equation and Legendre transformations. Using the nonlinearization approach of Lax pairs , a new finite-dimensional Hamilton canonical equations are obtained. Finally , it is proved completely integrable systems , and the involutive solutions of the evolution equations are given.

Key words: spectral problem; nonlinearization of Lax pairs; Bargmann system; integrable system; involutive representations

(责任编辑 车轩玉)

~~~~~

( 上接第 98 页)

## Analysis and Application of Intelligent Evolution in Printing Production Data

Liu Yongjun

( The Department of Information Engineering , Shijiazhuang University of Economics , Shijiazhuang 050031 , China)

**Abstract:** Print production enterprise resource management and the management of the production process based on the ERP system , but the lack of large amounts of data analysis and data mining , this paper adopts the intelligent evolution of thinking , the use of advanced algorithms model of MATLAB powerful data processing capability and MATLAB related to the toolbox of powerful dataanalysis capabilities and data mining tools and analytical models , the evolution of the data and global direct search , and explore its application in the analysis of the macro level and in specific areas of the enterprise data , better ways and means.

**Key words:** ERP system; data analysis; computational intelligence; global search; genetic algorithm professional toolbox

( 责任编辑 刘宪福)