

文章编号:1001-9081(2007)04-1009-04

超短期汇率的预测研究

黄巧玲, 谢维波

(华侨大学 信息科学与工程学院, 福建 泉州 362021)

(huangql7@hqu.edu.cn)

摘要:提出了一种适合超短期汇率预测的模型方法。实验数据通过网络获取,模型采用的是相空间重构与卡尔曼滤波计算的方法来对超短期汇率数据进行建模和预测,并与 BP 神经网络模型进行了比较。实验结果表明,所建立的模型方法能很好地跟踪即时汇率变化趋势,预测精度比较高,且算法运行速度比 BP 神经网络模型快得多。最后,给出了在 .NET 环境下实现了汇率在线预测的全部过程。

关键词:超短期汇率预测;数据获取;相空间重构与卡尔曼滤波;在线预测

中图分类号: TP39; TP182 **文献标识码:** A

Prediction research of the supper-shortterm exchange rate

HUANG Qiao-ling, XIE Wei-bo

(College of Information Technology and Engineer, Huaqiao University, Quanzhou Fujian 362021, China)

Abstract: A new method which is fit to the prediction of supper-shortterm exchange rate was proposed. The data for experimentation was got from Internet, and the model was established with the reconstructed phase space and Kalman. Compared with the neural network predicting model, the results show that the proposed model can track the instant exchange rates change well, and it is better than the neural network model both in accuracy rate and run-time length. Finally, the whole procedure of prediction on line based on .NET was introduced.

Key words: prediction of the supper-shortterm exchange rate; data getting; the phase space reconstruction and Kalman; predicting on line

0 引言

超短期汇率预测是指预测一天内汇率的变化趋势,它对外汇市场上的日常交易来说是非常必须的。目前,汇率交易是全球全天 24h 通过网络进行的,不同时段交易获利不同。因此,对企业经营和个人炒汇来说,汇率即时数据的跟踪是十分有意义的。就目前汇率预测研究方法来看,最热门的工具是神经网络方法,神经网络具有很强的非线性逼近能力,是非线性系统研究的好方法。但是,神经网络是在学习输入输出样本的基础上获得的,灵活性高,但缺乏可靠的数学表达形式,而且现有的学习算法收敛速度比较低,难以满足在线学习的要求^[1]。卡尔曼滤波是一种可用于非线性系统的滤波算法,具有最优估计性能,其递推计算形式及算法实现主要是矩阵的加减、乘除及求逆等计算量不大的特征,使其适合实时处理的需要。

根据以上的分析,本文利用卡尔曼滤波方法,提出了一种适合超短期汇率预测的模型。模型所采用的卡尔曼滤波器的初始状态通过相空间重构技术得到。文中实现了银行网站即时汇率的接收及存储,为本文预测模型提供实验数据来源。

1 实验数据的获取

1.1 数据获取及存储

本文的数据是从某银行网页获取的。数据获取的目的就是获取此页面的汇率报盘数据并存入数据库中。获取的整个

过程如图 1 所示。

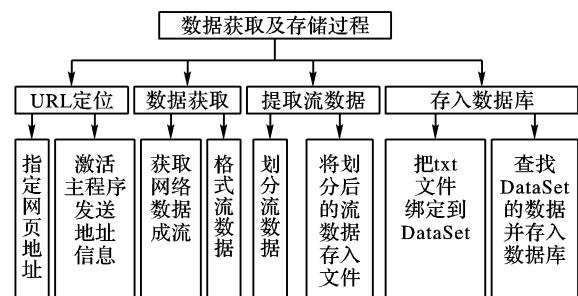


图 1 数据获取及存储过程

1.2 数据获取的界面及主要实现代码

数据获取的页面如图 2 所示,主要实现代码(节选)如下:

```

public void GetData()
{ //获取外汇数据到 TXT 文件
stringFilePath = System. Web. HttpContext. Current. Server.
MapPath( "data. txt" );
stringPageUrl = "http://61. 154. 12. 228/html/read1. txt";
//防止网络连接超时
try
{ //调用方法获取网络数据
string Str = GetHtml. getPageInfo( PageUrl );
//往文件中加内容
StreamWriter w = File. CreateText( FilePath );
//调用方法往 TXT 文件中加内容
}
}
  
```

收稿日期:2006-09-27;修订日期:2006-12-20 基金项目:福建省自然科学基金资助项目(A0540005)

作者简介:黄巧玲(1981-),女,福建漳州人,硕士研究生,主要研究方向:图像处理与模式识别; 谢维波(1964-),男,福建泉州人,副教授,主要研究方向:信号处理、模式识别、智能信息处理。

```

w. WriteLine( GetHtml. FormatString( Str ) );
//关闭文件流
w. Close();
}
catch( System. Exception e )
{ Console. WriteLine( e. ToString() ); }
}
#region 获取网络上指定的数据流
public static string getPageInfo( string pageUrl )
{ WebRequest myReq = WebRequest. Create( pageUrl );
  WebResponse myRes = myReq. GetResponse();
  Stream resStream = myRes. GetResponseStream();
  StreamReader sr = new StreamReader ( resStream, Encoding.
  Default );
  string pageInfo = ( sr. ReadToEnd() ). ToString();
  myRes. Close();
  string b = "pinzhong, A, kpi, jgj, jdj, zrz, zd, mlj, mcj, data, 0/1, 1"
  + pageInfo;
  return b;
}
#endregion
#region 格式化流
public static string FormatString( string Str )
{ Str = Str. Replace( "|", "\n" ); //当有“|”号转换成换行
  return Str;
}
#endregion

```

外汇名称	时间	开盘价	最高价	最低价	昨收价	涨跌	买入价	卖出价	变动
美元对欧元	13:52:43	1.8847	1.8833	1.8825	1.8835	-0.0015	1.882	1.8852	0
英镑对美元	13:52:43	7.7584	7.758	7.7576	7.7581	-0.0001	7.7501	7.7661	0
美元对港币	13:52:43	111.78	111.67	111.52	111.6	-0.08	111.51	111.75	0
美元对日元	13:52:43	1.1033	1.1013	1.0999	1.101	-0.0023	1.0986	1.102	0
欧元对美元	13:52:43	0.753	0.7517	0.7503	0.7519	-0.0002	0.7511	0.7533	0
美元对加元	13:52:43	1.2075	1.2069	1.2054	1.2074	-0.0005	1.2041	1.2075	0
欧元对英镑	13:52:43	1.295	1.2932	1.2919	1.2921	-0.0004	1.2931	1.2955	0
美元对欧元	13:52:43	1.5748	1.5731	1.572	1.5736	-0.0004	1.5692	1.5762	0
欧元对港币	13:52:43	14.621	14.61	14.604	14.612	-0.011	14.595	14.631	0
港币对日元	13:52:43	14.41	14.39	14.37	14.38	-0.03	14.35	14.43	0
加元对港币	13:52:43	7.0536	7.0444	7.032	7.0464	-0.0092	7.0409	7.0609	0
澳元对港币	13:52:43	5.8418	5.8317	5.8209	5.8333	-0.0017	5.8266	5.8446	0
港币对欧元	13:52:43	6.4362	6.428	6.4245	6.4255	-0.0007	6.424	6.444	0
英镑对欧元	13:52:43	2.2743	2.273	2.2707	2.2741	-0.0014	2.268	2.2744	0
英镑对日元	13:52:43	210.47	210.31	210.1	210.2	-0.17	210.1	210.44	0
欧元对日元	13:52:43	84.11	83.94	83.76	83.91	-0.04	83.87	84.07	0
港币对日元	13:52:43	92.64	92.53	92.46	92.43	-0.05	92.48	92.68	0
加元对日元	13:52:43	101.54	101.4	101.18	101.36	-0.03	101.35	101.55	0
英镑对加元	13:52:43	2.078	2.0741	2.0719	2.0737	-0.0036	2.0694	2.0756	0
英镑对美元	13:52:43	2.5107	2.5054	2.5008	2.505	-0.0053	2.502	2.5062	0

图2 数据获取页面

2 超短期汇率模型的建立

2.1 相空间重构技术

2.1.1 相空间重构理论

在时间序列的分析中,决定序列的可观测因素很多。而且相互作用的动力学方程往往是非线性的,甚至是混沌的。同时,因测量精度的实际限制、计算的复杂性,以及可能存在的本质上的非确定性因素等多方面的困难,严重制约着人们对时间序列内在机制的理解。20世纪80年代以来,由于Takens^[2]对Whitney早期在拓扑学方面工作的发展,使得深入分析时间序列的背景和动力学机制成为可能。在确定性的基础上,对序列动力学因素的分析,目前广泛采用的是延迟坐标状态空间重构法。一般来说,非线性系统的相空间可能维数很高,甚至无穷,但在大多数情况下维数并不知道。在实际问题中,对于给定的时间序列,通常是将其扩展到三维甚至更高维的空间中去,以便把时间序列中蕴藏的信息充分地显露出来,这就是延迟坐标状态空间重构法,其具体描述如下:

设动力系统实测数据为: $\{y_i, i = 1, 2, \dots, N\}$ 是一个 n 维系统的某一状态输出。嵌入维数为 m , 时滞为 τ , 建立一个

“(m, τ)”窗口。它使时间序列 $\{y_i, i = 1, 2, \dots, N\}$ 中的 m 维的所有元素同时出现, 当时间序列顺序移过窗口时, 可得到一个 m 维状态向量:

$$X_i = [y_i, y_{i+\tau}, y_{i+2\tau}, \dots, y_{i+(m-1)\tau}]^T \quad (1)$$

这里 $N_m = N - (m-1)\tau$, 为重构向量的个数。由Tankens定理可知, 只要满足 $m \geq 2n + 1$ (n 为原相空间维数), 则重构后的状态向量能恢复原动力学系统的本质特征。

2.1.2 参数确定方法

相空间重构的关键在于嵌入维数 m 和时滞 τ 的确定, 目前, 确定嵌入维数 m 的常用方法有伪最近邻法、奇异值分解法; 确定时滞 τ 的方法主要有自相关函数法及互信息量法。但 these 方法存在一定的缺陷。伪最近邻法的结果会受到阈值的影响; 奇异值分解的结果依赖于超空间维数的选择, 利用奇异值分解法确定的 m 不仅可能会随着超空间维数的不同而不同, 而且可能会随着数据长度 N 的不同而不同; 在相关维计算法中, 取数据矢量间的欧氏距离值的第一个到达平稳时的 m 估计嵌入维数。但矢量间的欧氏距离值可能达不到平稳, 这时就无法求得 m ; 另一方面, 确定时滞的典型方法自相关和互信息法虽然具有计算量小等优点, 但计算结果具有不一致性^[3]。

文献[4]提出一个基于样本时间序列及其替代数据的相空间的微熵率方法, 同步确定 τ 和 m 。该方法的主要优点是用一个简单的测度同时优化 m 和 τ , 避免了互信息量法和错误最近邻法的不一致性。该方法的物理意义明显, 实际效果较好, 故本文采用此方法估计这两个嵌入参数的最优值: m_{opt} 和 τ_{opt} 。

给定信号 $x(t)$ ($t = 1, 2, \dots, N$) 的 N_s 个替代数据 $x_{s,i}(t)$, $i = 1, \dots, m$, 定义熵率 (ER) 为:

$$R_{ent}(m, \tau) = I(m, \tau) + \frac{m \ln n}{n} \quad (2)$$

其中 n 是延迟矢量数, $I(m, \tau)$ 为:

$$I(m, \tau) = \frac{H(x, m, \tau)}{\langle H(x_{s,i}, m, \tau) \rangle_i} \quad (3)$$

而微熵 $H(x) = \sum_{j=1}^N \ln(N\rho_j) + \ln 2 + C_E$, N 是数据长度, ρ_j

是第 j 个延迟矢量与其最近邻点之间的欧氏距离, 欧拉常数 $C_E \approx 0.5772$ 。熵率图上的最小值在 m 和 τ 轴上分别对应 m_{opt} 和 τ_{opt} ; 然后, 重构相空间:

$$X(k) = [x(k), x(k + \tau_{opt}), \dots, x(k + (m_{opt} - 1)\tau_{opt})]^T \quad (4)$$

2.2 基于单步预测的卡尔曼滤波计算过程

卡尔曼滤波是一种统计估算方法, 主要思路是: 预测方程中的回归系数是随时间变化的。预测每向前延伸一步, 都将预测结果与观测结果进行比较, 其差别 (预测误差) 将以适当的方式反馈到回归系数的变化方程中去。通过利用前一时间刻预测误差的反馈信息及时修正预测方程, 以提高下一时刻的预测精度。卡尔曼滤波方法不论预测次数 (或量测次数) 如何增加, 不需要存储大量历史的量测数据, 减少了计算机的存储, 而且只进行矩阵的加、减、乘、除和求逆运算, 通常计算量不大, 从而满足了应用滤波的实时性要求。

卡尔曼滤波预测公式为^[5]:

$$G(n) = K(n|n-1)C^T(n)[C(n)K(n|n-1)C^T(n) + R(n)]^{-1} \quad (5)$$

$$a(n) = Y(n) - C(n)X(n|n-1) \quad (6)$$

$$X(n|n) = F(n, n-1)X(n-1|n-1) + G(n)a(n) \quad (7)$$

$$K(n|n-1) = F(n, n-1)K(n-1|n-1)F^T(n, n-1) + Q(n-1) \quad (8)$$

$$K(n|n) = (I - G(n)C(n))K(n|n-1) \quad (9)$$

$$X(n+1|n) = F(n+1, n)X(n|n) \quad (10)$$

$$Y(n+1|n) = C(n+1)X(n+1|n) \quad (11)$$

式中: $X(n+1|n)$ 表示给定观测值 $Y(1), Y(2), \dots, Y(n)$ 在 $n+1$ 时刻状态的预测估计。 $G(n)$ 为滤波增益矩阵, $F(n, n-1)$ 是从 $n-1$ 时刻到 n 时刻的转移矩阵, $K(n|n-1)$ 为 $x(n+1|n)$ 中误差的相关矩阵, $C(n)$ 为 n 时刻的测量矩阵, $Q(n)$ 为过程噪声的相关矩阵, $R(n)$ 为测量噪声的相关矩阵。 $Y(n+1|n)$ 为 $n+1$ 时刻的预测估计值。

2.3 预测模型的建立

在卡尔曼的应用中, 输入向量过程: 观测值 = $\{y(1), y(2), \dots, y(n)\}$ 。状态向量为 $x(n)$, 滤波器是通过状态向量的递推计算得到未来状态的估计值。因此, 滤波器初始状态向量的选取对预测结果的影响具有关键的意义。大量的研究表明, 经济时间序列是混沌的。可以用混沌时间序列的知识来确定滤波器初始状态向量的维数及向量分量的构成。对给定的汇率时间序列, 求得序列的最优嵌入维数和最佳时滞, 把最优嵌入维数作为滤波器初始状态向量的维数, 把求得的最佳时滞作为初始状态向量分量的时间间隔。

美元港币

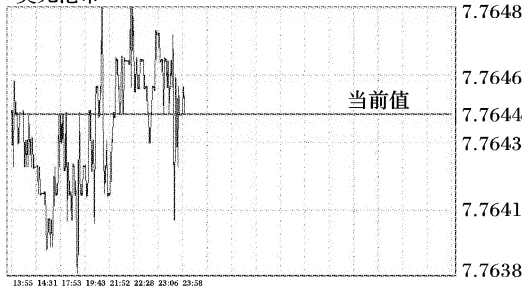


图3 2006年5月26日的即时汇率曲线图

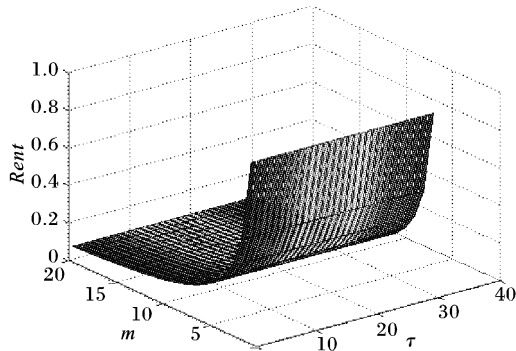


图4 即时汇率熵率图

对 1.1 获取的 2006 年 5 月 26 日 13:55 至 23:58 的 1066 条美元港币即时汇率, 采用微熵率方法求解时间序列的最优嵌入维数和最佳时滞。求得熵率图如图 4 所示。求得的最优嵌入维数和最佳时滞为 (20, 3)。根据以上的分析, 得到实验数据 $\{y(1), y(2), \dots, y(n)\}$ 进行卡尔曼预测的初始状态向量输入为 $x(1) = [y_3(1), y_3(4), \dots, y_3(80)]^T$ 。

3 模型预测结果分析

以下为基于相空间重构与卡尔曼滤波对 2006 年 5 月 26

日 13:55 至 23:58 的 1066 条美元港币即时汇率的跟踪预测结果。为了对比研究, 本文实现了 BP 神经网络模型, 并把本文算法的预测结果与 BP 神经网络的预测结果进行了比较。图 5 为汇率真实值、本文算法预测值及 BP 神经网络预测值的对比图; 图 6 为本文算法预测残差; 图 7 为 BP 神经网络预测残差。为了更好地反映预测模型的效果, 本文采用以下几个性能评价, 其中 e_i 为汇率实际值与预测值的绝对误差, s_i 为汇率实际值, $i = 1: 1006, N = 1066$ 。

$$\text{平均绝对误差 (MAE): } MAE = \frac{\sum |e_i|}{N}$$

$$\text{均方差误差 (MSE): } MSE = \frac{\sum (e_i)^2}{N}$$

$$\text{均方根误差 (RMSE): } RMSE = \sqrt{MSE}$$

$$\text{平均相对误差 (MAPE): } MAPE = \frac{1}{N} \sum \left| \frac{e_i}{s_i} \right|$$

算法运行时间 $RUNT$: 为预测算法开始到结束的时间间隔。

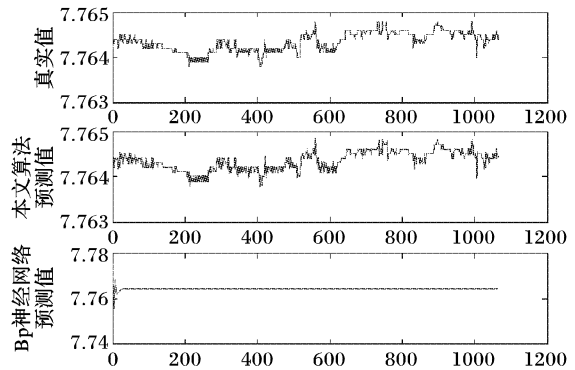


图5 真实值/预测值

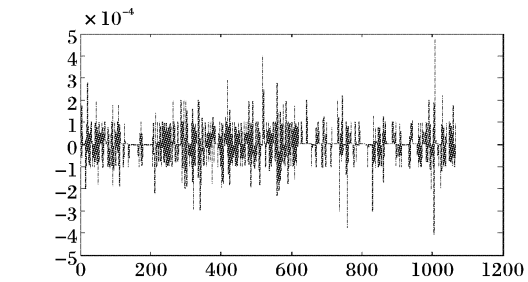


图6 本文算法的预测残差

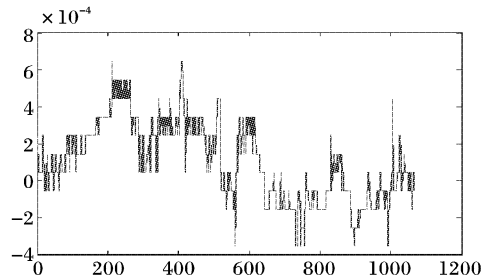


图7 BP神经网络的预测残差

表1 各项性能指标比较

各项性能指标	BP神经网络	本文算法
平均绝对误差 (MAE)	0.2192	0.0461
均方差误差 (MSE)	6.8491e-008	5.4191e-009
均方根误差 (RMSE)	2.6171e-004	7.3614e-005
平均相对误差 (MAPE)	2.6480e-005	5.5690e-006
算法运行时间 RUNT	400s	≈10s

根据以上的实验结果可知, 本文算法的预测效果良好, 能

很好地跟踪即时汇率的变化趋势。且算法运行速度快,能满足在线预测的实时性要求。卡尔曼滤波递推估计的方式使其在跟踪上具有很好的性能。而神经网络全局逼近的方式使其在超短期汇率预测中,实验精度比较低,收敛速度慢的缺陷也使神经网络预测模型不能实现即时汇率的在线预测。

4 在线预测的实现过程

汇率在线预测的思想是:用户在页面上输入预测的汇率种类如美元欧元,请求预测某个数据的未来值,如明日开盘价、今日最高价、下个时刻的即时汇率值等。系统根据用户的选择,查询数据库,并把查询结果传送给预测模块,由预测模块完成预测后,把预测结果返回给用户。

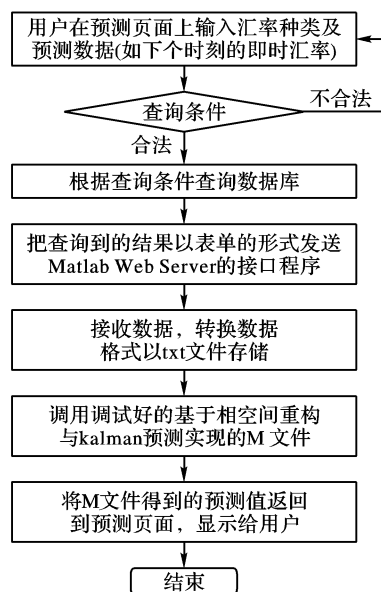


图8 汇率在线预测过程

实现上述在线预测的关键是解决 Matlab 与 .NET 的动态数据交换过程,本文采用 Matlab Web Server 接口程序 matweb.exe 来实现 Matlab 与 .NET 的动态数据交换。它是目前较新且简单可行的方法。Matlab Web Server 的核心是作为 CGI 程序的可执行程序 matweb.exe。动态数据的交换过程就

是由网关接口程序 matweb 和用户编制的 Matlab 应用程序协作完成的^[6]。

在进行动态数据的交换时,输入数据通过网页以表单的形式发送到通用网关接口程序。通用网关接口程序收到输入数据后,分析输入数据,调用与之对应的 Matlab 的 M 文件。这时,通用网关接口程序将利用系统的 Matlab 服务,进行 Matlab 运算,得到输出或绘制输出图形,并将结果返回。从而完成动态数据的交换过程。图 8 是基于 .NET 和 Matlab 动态数据交换的汇率在线预测过程。

5 结语

汇率对国家经济、企业经营和个人投资来说都十分重要,汇率预测的研究一直是国内外学者研究的热点。随着越来越多人把汇率买卖当作一种投资,个人对汇率知识的获取及其行为规律的把握越来越强烈。因此,超短期汇率预测的研究对汇率日常交易具有十分重要的意义。针对目前流行的神经网络预测模型不能适应超短期预测的要求,本文提出的基于相空间重构与卡尔曼滤波计算的方法能很好地跟踪即时汇率的变化趋势,且运行时间上满足在线预测的实时性要求。

参考文献:

- [1] HUNT KJ, SBARBARO D, ZBIKOW R, *et al.* Neural networks for control system-a survey[J]. Automatica, 1992, 28(6): 1083 - 1112.
- [2] TAKENS F. Detecting strange attractors in turbulence[J]. Lecture Notes in Mathematics, 1982, 23(12): 198 - 366.
- [3] 杨绍清, 贾传炎. 两种实用的相空间重构方法[J]. 物理学报. 2002, 51(11): 2452 - 2453.
- [4] GAUTAMA T, MANDIC DP, VAN HULLE MM. A differential entropy based method for determining the optimal embedding parameters of a signal[A]. Proceedings of the Int Conf on Acoustics, Speech and Signal Processing[C]. Hong Kong, 2003, 35(6): 29 - 32.
- [5] (美) HAYKIN S. 自适应滤波器原理[M]. 郑宝玉, 译. 北京: 电子工业出版社, 2005.
- [6] 熊凌. 基于 .NET 的 Matlab 网络动态数据交换的研究与实现[J]. 微计算机信息, 2005, 8(21): 31 - 33.

(上接第 1008 页)

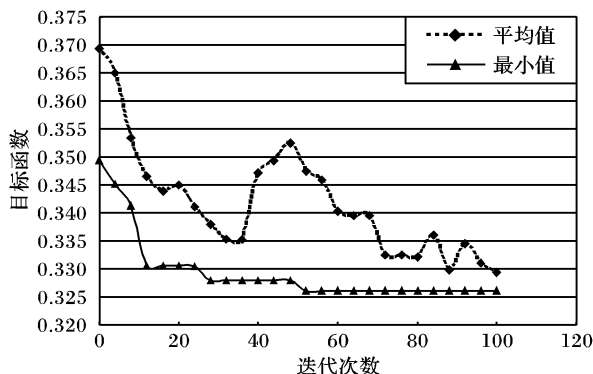


图3 自适应遗传算法的世代图

5 结语

本文根据任务完成的交货期、质量、成本和服务四个评价指标,利用总任务下各评价指标的权重以及在每个评价指标下各子任务之间的特性来建立基于任务分解结构的供应链伙伴选择多目标决策模型,并在该模型中考虑到任务分解模式

下的任务完成时间的精确计算。然后,利用自适应遗传算法求解模型,并通过算例分析表明,该方法可以用于求解基于任务分解结构的供应链合作伙伴优化选择问题。

参考文献:

- [1] WU NQ, MAO N, QIAN YM. An approach to partner selection in agile manufacturing[J]. Journal of Intelligent Manufacturing, 1999, 10(6): 519 - 529.
- [2] VAIRAKTARAKIS, GEORGE L, HOSSEINI, *et al.* Forming partnerships in a virtual factory[A]. Proceedings of the 27th Annual Meeting of the Decision Science Institute[C]. Atlanta, GA, USA, 1996. 1394 - 1396.
- [3] MINIS I. Optimal Selection of partners selection in agile manufacturing[EB/OL]. www.isr.umd.edu/labs/research.html, 2006 - 08.
- [4] 钱碧波, 潘晓弘, 程耀东. 敏捷虚拟合作伙伴选择评价体系研究[J]. 中国机械项目, 2000, 11(4): 397 - 401.
- [5] 于海龙. 协同制造[M]. 北京: 清华大学出版社, 2004. 56.
- [6] 马士华. 供应链管理[M]. 北京: 机械工业出版社, 2001. 124.
- [7] 代颖. 基于遗传算法的供应链联盟伙伴选择[J]. 西南交通大学学报, 2004, 39(4).