

基于 VB.NET + MapObjects 的人口统计 GIS 研究

李军利¹, 蔡爱民¹, 查良松², 赵 来¹

(1. 滁州学院, 安徽 滁州 239012; 2. 安徽师范大学, 安徽 芜湖 241003)

摘 要: 组件技术是当前 GIS 软件业的潮流。系统以安徽县域人口统计 GIS 为案例, 基于 .NET 环境, 将数据库技术、数学模型与 GIS 技术应用到人口统计信息管理当中, 开发了重心模块、洛伦茨模块、人口金字塔模块与 GM(1,1) 人口预测等模块, 完成了人口统计信息与地理空间信息的集成, 实现了地图浏览、统计查询、专题制图、模型分析等功能。其强大的空间分析和可视化功能增强了人口统计信息的表现力, 将进一步拓展人口统计 GIS 在相关领域的应用。

关键词: 组件技术; MapObjects; VB.NET; 人口统计 GIS

中图分类号: TP311; P208

文献标识码: A

文章编号: 1673-629X(2008)06-0243-05

Research of Population Statistic Geographic Information System Based on VB.NET and MapObjects

LI Jun-li¹, CAI Ai-min¹, ZHA Liang-song², ZHAO Lai¹

(1. Chuzhou University, Chuzhou 239012, China; 2. Anhui Normal University, Wuhu 241003, China)

Abstract: The component developing technology is the tidal current of the GIS software industry. Based on the case study of the Counties' Population Statistic Geographic Information System in Anhui province, the system applies the relationship database technology, the mathematical model, the GIS technology to population statistic information management. Under the .NET framework, develops Gravitation module, Lorenz module, Population Pyramid module and module of GM(1,1) population forecast. Thereby, it achieves the integration of the population statistic information and the geography spatial information, realizes the functions of map browsing, statistics inquiry, thematic mapping and model analysis, and enhances the population statistic information's expressiveness based on the powerful spatial analysis and the visualization functions, which will further promote more applications of population statistic GIS in the correlated domains.

Key words: component technology; MO; VB.NET; population statistic geographic information system

0 引 言

当前在面向特定专题或领域的地理信息系统平台开发中, 主要采用集成二次开发技术, 即以通用可视化开发工具为开发平台, 利用专业的 GIS(如 MapInfo、ArcGIS 等) 工具软件实现基本的 GIS 功能^[1]。目前最新的一些人口 GIS 软件工具包, 大多是基于 GIS 组件技术开发, 如 U. S. Census Bureau 开发的 CSPro(Census and Survey Processing System) 是人口普查软件包^[2]; World Health Organization 开发的用于监测评估全球人类或牲畜传染病发生发展情况的 HealthMapper 软件包^[3]; 联合国开发的 MORTPAK for Windows (United Nations Software Package for Demographic Mea-

surement) 是专门用于发展中国家人口统计分析的软件包^[4]。澳大利亚昆士兰州开发的 POPSTAR (Queensland Center for Population Research) 人口统计软件模块^[5]; 美国国家癌症研究中心的 Martin Kulldorff 研制的 SaTScan 软件包^[6]。诸如此类国外开发的软件包还很多, 这些专门性软件包的研制成功, 标志着复合数据源的人口信息与地理位置关联, 以专业模型与 GIS 集成模式正成为人口 GIS 研究热点。

1 GIS 组件——MO

MaPOjects(简称 MO) 是全球最大的 GIS 软件供应商 ESRI 公司在业界最早推出的一个 GIS 功能组件集合, 这些组件遵守微软的对象模型(COM) 或者 ActiveX 软件规范^[7]。软件开发可以通过在任何 OLE/COM 软件开发环境中嵌入 MO 来创建应用程序。它由一个称为 Map 的 ActiveX 控件和近 50 个具有事件、方法和属性的可编程 OLE 对象组成, 这些对象为应用

收稿日期: 2007-09-02

基金项目: 安徽省自然科学基金资助项目(050450103); 安徽省高校自然科研项目(2006kj018B)

作者简介: 李军利(1976-), 男, 安徽无为为人, 硕士, 主要从事空间数据挖掘、GIS 应用与开发研究。

开发人员提供了有力的制图与 GIS 功能支持^[8,9]。目前主要有 MO-Java Edition、Windows Edition 与 MO LT 三种版本,最高版本为 2.4。在标准 Windows 编程环境下,它能与图形、多媒体和数据库开发等技术较好融合,是一种基于前端应用业务的良好地图开发环境。

MO 包含一个地图控件和一组功能强大的对象,MO2.2 中对对象分为六组^[8,9]:地址匹配对象、数据访问对象、地图显示对象、几何对象、投影对象、实用对象。MO 开发步骤分为^[10,11]:(1)选择数据模型,构建数据库。构建空间数据库与属性数据库。(2)图层加载与显示、删改。在 GIS 中,地图上的所有对象一般都是分多个图层进行存储。在 MO 的地图控件中,点、线与面的 Shapefile 文件是分层加入,不同图层要设置相同的投影和系统,使用 MO 中的 GeoDataset、DataConnetion 对象与 FindGeoDataset 方法来连接和读取地理数据,最先加入面图层,然后依次加入线图层,点图层。(3)地图操作与专题分析。MO 通过它的几何对象执行缓冲区、差、交、并与异或等空间分析。通过 Symbol 对象与 Renderer 对象进行地图着色、美观显示与专题制图,通过 SearchExpress(表达式查找)、SearchByDistance(距离查找)与 SearchShape(图形查找)来实现特征查找。

2 人口统计 GIS 设计

2.1 需求与目标

针对“五普”后安徽省积累的大量人口普查资料,加上运用 GIS 技术研究安徽人口统计相对较少,在进行安徽县域人口统计 GIS 设计之前,项目组多次走访居巢区、无为县、枞阳县与芜湖县的统计局、计生委等部门,进行了用户信息需求调查与座谈。界定系统目标主要完成以下任务:实现以县区行政单位为界的人口统计数据的管理、检索、可视化查询与人口电子地图相关功能;直观显示人口的分布和增长趋势;进行人口重心、人口数量与土地面积关系、人口性别年龄的结构分析与分区域类别人口预测。

2.2 人口统计指标与模型

2.2.1 指标选取

人口统计指标是描述人口信息的重要标志,它是通过登记、统计和初步汇总人口信息得出的相关数字,通过指标可以从不同角度反映人口问题的实质^[12]。依据人口学、人口地理学、人口统计学的人口统计指标和 2000 年 11 月 1 日零时进行的第五次全国人口普查中长表统计 49 项内容、短表统计 19 项内容,结合中国自然资源数据库中的人口统计表与需求分析,选取 11 大类共 80 多个指标,来反映区域的人口数量、性别、户籍状况、年龄、民族、出生、死亡、迁移、职业、受教育程

度和住房情况。上述指标都将作为人口属性数据库中的数据表与关键字段。

2.2.2 人口统计 GIS 模型设计

任何 GIS 都是为一定的应用而建立的,而应用模型是联系 GIS 与常规专业研究的纽带。系统中综合利用应用模型,开展数据分析与应用^[13]。PSGIS 使用人口问题研究中的相关专业模型开展定量分析。系统设计以下 GIS 专业模型,对人口信息的时空分布特性进行分析。

(1)重心模型:这个概念源于物理学。人口分布重心最早由美国学者朗西斯·沃尔克于 1874 年提出,指在该地域空间平面上力矩达到平衡的一点就是人口分布中心,常指经度与纬度。从重心在一定时间段内的移动轨迹,可以看出人口信息分布的变动情况和变动速度^[14],可对人口在时间上、空间上的迁移变化进行分析。

(2)洛伦茨模型:是用来鉴定和表明社会收入分配平均程度的一条曲线,它是由美国统计学家洛伦茨提出,也叫频率累积曲线^[15]。系统使用它制作区域内某一特定人口统计指标占总量的百分比,将百分比数由低向高累加,再求相对于该指标的另一指标累积百分比序列,以此来分析人口统计指标的分散聚集程度。

(3)人口金字塔模型:是指在人口的性别结构与年龄结构,以图形来呈现人口年龄和性别的分布情形,以年龄为纵轴,以人口数量占总人口数的百分比为横轴,按左侧为男、右侧为女,形象直观地用几何图形的形式表示出来^[16]。通过人口金字塔反映区域的过去人口性别年龄情况结构,以及今后人口结构可能出现的趋势。

(4)灰色模型:是用某一指标的过去行为来预测未来,预测结果是该指标在未来各个时刻的具体化数值^[17]。通过时间序列历史数据揭示现象随时间变化的规律,将这种规律延伸到未来,从而对该现象的未来做出预测,系统借用此模型预测各地人口情况。

2.3 系统结构与功能

在微软 Windows(2000/NT/XP)操作平台上,使用以 VB.NET+MO+ArcSDE for SQL Server 2000 为解决方案,采用 C/S 三层结构:数据访问层,提供了系统数据操作与维护功能;业务逻辑层,定义应用系统的对象的行为以及它们之间的关系,即组织人口统计及相关数据,通过行政区划编码和时间序列编码,提取、检索人口统计信息,向表现层提供各类统计报表、时空分析的专业人口 GIS 模型和调用接口;可视化表现层即客户应用平台,描述业务对象的表现形式并接受用户的输入,通过人口统计 GIS 表达人口专题信息。

人口统计 GIS 系统应具有管理人口统计属性信息与空间特征的双重能力。

系统主要功能设计如下:

(1)系统管理。具备常见 MIS 系统功能。

(2)地图浏览。放大、缩小、漫游、全屏、定位查询、双向关联查询、鹰眼导航等。

(3)统计查询。提供安徽多年份人口的教育、民族、行业、性别、年龄与住房信息,及县区人口概况与县区基本空间信息等各种指标查询,实现人口统计信息化管理,执行常规操作,如更新、修改、保存与输出等。

(4)专题图表。单项统计报表、复合统计报表、点密度图、柱状图、饼图、分级图、四分位数、标准差、单值图、按查询指标字段进行标注等功能,反映具体区域的人口统计指标的单要素、多要素的特征和相互,以可视化形式再现人口信息的时空分布特征和人口结构。

(5)模型分析。多年份人口重心、人口数量与区域面积的洛伦茨曲线、78 个县区的人口金字塔分析、各类人口数量的 GM(1,1)预测,针对大量人口统计数据,开展更高层次的专业分析,为研究人口规律、制定人口政策提供辅助决策。

(6)系统帮助。方便用户使用系统,提供帮助或答疑。

2.4 数据库设计

人口统计 GIS 中主要进行系统的空间数据库与属性数据库设计。

空间数据库设计,根据系统需要,使用的图形资料为 1:75 万安徽省政区图。将纸质地图通过高分辨率扫描仪扫描后输入计算机,再经过图像软件处理后导入 ArcGIS 软件作为矢量化与建拓扑关系。以 Shapefile 格式转存到 ArcSDE 中。其中在 Shapefile 文件的属性表中,增加一个县区编码字段,作为与属性表连接的空间索引,使用国家统计局网站公布的截止 2005 年 12 月 31 日最新的县及县以上行政区划代码。系统的属性数据主要包括描述空间数据的点、线与面各要素特征的属性数据与人口统计信息的特征数据。通过采集多年份人口资料,分别构建教育、民族、行业、性别、年龄、住房条件、县区基本空间关系表,以此来描述人口统计的各项指标之间关系。

3 人口统计 GIS 实现

系统实现了人口统计信息与地理空间信息的集成,具备地图浏览、统计查询、专题制图、模型分析等功能。以下给出一些核心模型算法与程序实现。

3.1 专题渲染

制作专题地图是根据某个特定的专题对地图进行

渲染的过程,通常也称为地理对象进行着色技术,同一地理信息可以有不同的表示方法和符号系统来传递,同一表示方法也可以传递不同的地理信息。系统提供多个专题制图。图 1 为 2000 年各县市第三产业人口占总人口比重。以下这段程序是实现对人口统计信息的某个指标字段制作独立值专题地图,其它专题图层渲染方法与此相似,换用不同的 Renderer 对象执行。

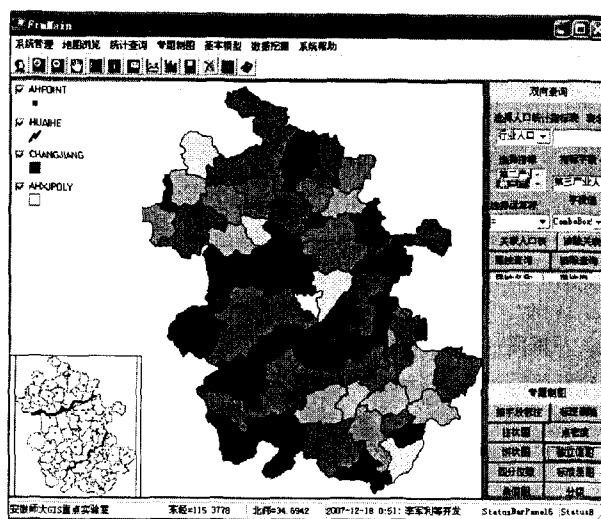


图 1 专题地图

Try

```
Dim pRecs As ESRI. MO2. Core. Recordset = player. Records
'建立图层的记录集
```

```
Dim strings As ESRI. MO2. Core. Strings = New ESRI.
MO2. Core. Strings
```

定义一个字符串变量

```
Dim i As Integer
```

```
Dim sZBMC_field As String'定义一个变量存放专题地图人
口指标字段名
```

```
sZBMC_field = TextField.Text.Trim.ToString'用于获取独
立值专题地图的字段
```

```
Do While Not pRecs.EOF
```

'使用循环,获取所有的字段的记录值

```
strings.Add(pRecs.Fields.Item(sZBMC_field).Value-
AsString)
```

```
pRecs.MoveNext()
```

```
Loop
```

```
Dim pRend As New ESRI. MO2. Core. ValueMapRenderer'定
义一个渲染对象
```

```
pRend.Field = sZBMC_field
```

```
pRend.ValueCount = strings.Count'记录值数量
```

```
For i = 0 To strings.Count - 1
```

```
pRend.Value(Convert.ToInt16(i)) = strings.Item(i)'渲
染实例对象得到赋值
```

```
Next i
```

```
player.Renderer = pRend
```

```
Me.Map1.CtlRefresh()
```

```
Catch ex As COMException '出错部分以下省略
```

```
Finally
```

```
End Try
```

3.2 人口重心模型

人口重心的计算需要与几何重心坐标关联,如果不使用行政重心作为经纬度坐标计算,可使用 MapObjects 中的多边形对象的 Centroid(几何重心)属性,先用一个二维数组通过程序产生全省每个县区多边形的几何重心,再通过人口重心公式产生重心坐标。

在 Map1 - AfterTrackingLayerDraw 函数中加入以下程序代码显示人口重心:

```
q.Set(a(k, 0), a(k, 1)) '读入计算生成的人口重心经纬度
```

坐标,k 与年份关联

```
pts.Add(q) '加入到点集 pts 当中
```

If flag = True Then '与主窗体显示人口重心按钮单击事件中的逻辑变量一致,用于控制程序是否显示人口重心

```
If Not pts Is Nothing Then
```

Map1.DrawShape(pts, sym1) '显示重心,Year 为每次选择的年份变量

```
Map1.DrawText(Year, q, TextSym1) '显示年份
```

```
End If
```

```
End If
```

```
pts.Remove(0) '清除重心坐标,以免下次重绘重叠
```

3.3 洛伦茨模型与人口型金字塔

GDI+ 是 .NET 中基于类的应用程序编程接口,它可以用与设备无关的方式执行二维图形操作。运用 Graphics 类,根据模型的数学运算方法,写出算法,通过 DrawCurve 函数绘图。图 2 为 2000 年安徽县市人口数量与区域面积的洛伦茨曲线。图 3 为 2000 年合肥市人口金字塔模型。

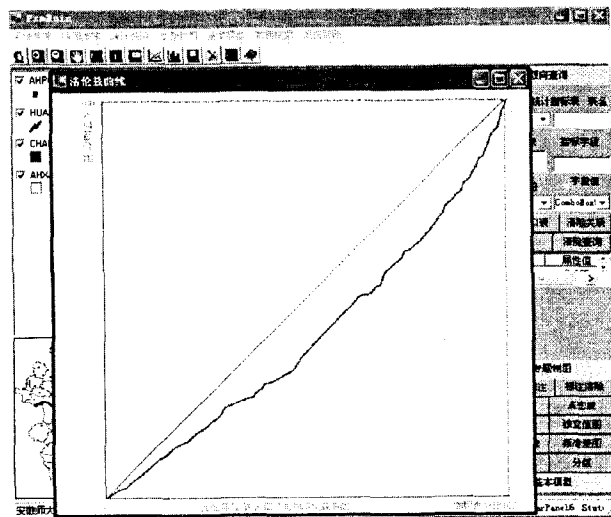


图 2 洛伦茨曲线

3.4 GM(1,1)模型

该模型在实现时主要涉及一些矩阵的运算。系统

运用 VB.NET 开发了灰色模型算法,包括矩阵复合运算、数组换维、误差检验等类模块的实现程序,图 4 为系统中运行界面,利用 GM(1,1)模型预测来安县总人口与非农人口数,目前系统基础数据只录入了 1997~2003 年统计数据。

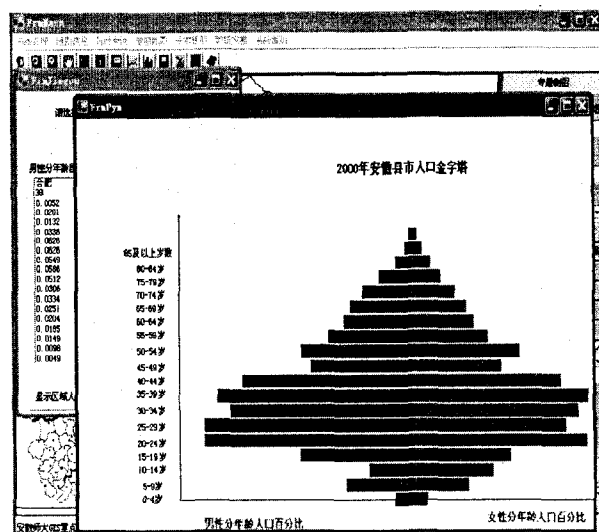


图 3 人口金字塔模型

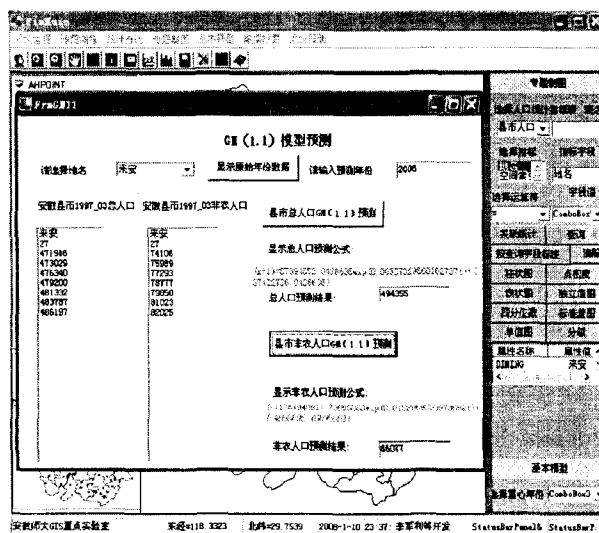


图 4 GM(1,1)模型

4 结 语

系统实现了人口信息与空间信息集成,使得每一个人口统计指标都能用地图语言表达,能进行统计信息与空间信息双向查询;系统将人口应用领域的重心、洛伦茨、人口金字塔与灰色模型有机地嵌入到传统 MIS 当中,运用模型对人口问题所依赖的规律或过程进行抽象或模拟,能有效地帮助相关人员理解和解决实际问题,提高人口信息化管理水平。笔者只从技术上探讨了系统的集成方法与功能实现,对其人口问题成因分析未能深入探讨。由于国内人口统计 GIS 整体

研究仍处于初级阶段,诸多关键技术还有待完善。特别是多源空间数据的有效集成、人口空间数据的不确定性、人口信息的不确定和不规范对系统建设影响巨大,一些先进的技术和方法,如人口空间数据仓库等技术,未能很好地集成到人口统计 GIS 当中。

参考文献:

- [1] 刘 光. 地理信息系统——组件开发篇[M]. 北京:中国电力出版社,2003.
- [2] U. S. Census Bureau. CSPro Getting Started[EB/OL]. 2007. <http://www.census.gov/ipc/www/cspro/index.html>.
- [3] World Health Organization. Public Health Mapping and GIS [EB/OL]. 2007. <http://www.who.int/health-mapping/tools/healthmapper/en/index.html>.
- [4] United Nations Population Division Department of Economic and Social Affairs. MORTPAK for Windows[EB/OL]. 2003. <http://www.un.org/esa/population/publications/mortpak/MORTPAKwebpage.pdf>.
- [5] Queensland Centre for Population Research. Population Projections for aState/Territory And its Regions [EB/OL]. 2005. <http://www.gpa.uq.edu.au/qcpr/Homepage/documents/POPSTAR.pdf>.
- [6] Kulldorff M. SaTScan Technical Documentation[EB/OL].

2005. <http://www.satscan.org/>.

- [7] Longley P A, Goodchild M F. 地理信息系统(上卷):原理与技术[M]. 第 2 版. 唐中实等译. 北京:电子工业出版社, 2004.
- [8] 芦东听,李典蔚,任 静,等. 基于组件式 GIS 的移动奥运智能交通系统[J]. 计算机技术与发展,2007,17(5):59-62.
- [9] ESRI. MapObjects getting.. started[EB/OL]. 2005. <http://www.esri.com/software/mapobjects/index.html>.
- [10] Ralston B A. Developing GIS Solutions With MO and Visual Basic[M]. New York: Onword press,2002.
- [11] 徐苏维,盛业华,黄家柱,等. 基于 MO 的 GIS 综合应用开发[J]. 计算机工程与应用,2006(1):194-197.
- [12] 刘 铮,邬沧萍,查瑞传. 人口统计学[M]. 北京:中国人民大学出版社,1981.
- [13] 吴信才. 地理信息系统设计与实现[M]. 北京:电子工业出版社,2004.
- [14] 陆 林. 人文地理学[M]. 北京:高等教育出版社,2004.
- [15] 张宗斌,王庆功. 现代西方经济学教程[M]. 北京:北京师范大学出版社,2002.
- [16] 刘德钦,刘 宇,薛新玉. 中国人口分布及空间相关分析[J]. 遥感信息,2004,29(6):76-79.
- [17] 邓聚龙. 灰预测与灰决策[M]. 武汉:华中科技大学出版社,2002.

(上接第 242 页)

通过大量实验表明,混合算法的表现明显优于基本效果,在时间消耗相近的情况下混合算法求解效果更好,混合算法不仅求解效率提升,并且具有更好的鲁棒性,虽然基本粒子群算法通过增加叠代次数、修改惯性常数 w 增加优化性能。但是对陷入局部最小改进性能有限,混合算法则能较好地避免这一问题,另外文中提出的单独对全局最有值局部搜索的策略也有利用减少普通混合算法执行时间过长的缺点。该混合策略并非专门针对 VRP 问题,可以应用到其他领域。

5 结束语

文中在结合禁忌搜索的基础上对基本粒子群算法提出了一种新的混合策略,该混合算法结合了粒子群算法在整体空间的全局搜索能力和禁忌搜索在局部空间个体邻域的独立寻优搜索能力。改进的粒子群优化算法通过实验分析可以得出以下结论:

- 1)充分利用了粒子群算法的群体并行搜索和禁忌搜索的串行领域搜索优势,实现了高效的搜索算法;
- 2)提升了优化性能,并加快了优化速度,尤其是对规模较大的复杂问题能够快速有效地求解;

3)具有鲁棒性好和参数要求不敏感,改善了禁忌搜索的对初始解依赖过强的缺点。

作为粒子群算法和禁忌搜索算法的尝试,为粒子群算法解决组合优化问题提出了一种可行的方法,该方法也可以应用到其他类似问题的求解。

参考文献:

- [1] Kennedy J, Eberhart R C. Particle swarm optimization[C]// In: Proc. of the IEEE Conf. on Neural Networks, IV. Perth: IEEE Press,1995:1942-1948.
- [2] 胡 旺,李志蜀. 一种更简化而高效的粒子群优化算法[J]. Journal of Software,2007,18(4):861-868.
- [3] 李 军,郭耀煌. 物流配送车辆优化调度理论与方法[M]. 北京:中国物资出版社,2001.
- [4] 张丽艳,庞小红,夏蔚军,等. 带时间窗车辆路径问题的混合粒子群算法[J]. 上海交通大学学报,2006,40(11):1890-1894.
- [5] 赵 波,曹一家. 电力系统无功优化的多智能体粒子群优化算法[J]. 中国电机工程学报,2005,25(5):1-7.
- [6] Ayed S, Imtiaz S, Sabah A M. Particle swarm optimization for task assignment problem[J]. Microprocessors and Microsystems, 2002,26:363-371.