

大数据背景下山东省农作制信息查询优化平台

姜红花¹,王秀丽¹,白鹏³,李光忠¹,宁堂原^{2*}

1. 山东农业大学 信息学院, 山东 泰安 271018

2. 山东农业大学作物生物学国家重点实验室, 山东 泰安 271018

3. 山东农业大学 机电学院, 山东 泰安 271018

摘要: 山东省是农业大省, 农业信息量大, 需要运用大数据的数据存取、基础架构、数据处理、统计分析等技术, 根据全省不同经济带的区域生态地理及环境特征与国内外市场动向, 把人口—资源—环境—发展作为一个复合系统。本文详细介绍了农作制信息查询与种植优化平台的开发过程、数据存取、基础架构、数据挖掘、统计分析及具体的实现策略; 该平台由登录注册模块、信息查询模块、种植优化模块和后台管理模块构成, 通过表格、柱状图和 ARCGIS 地图的方式将数据直观动态的展示给用户; 动态查询的设定, 使查询操作更便捷有效; 以劳动力, 肥料, 资金等作为农业种植规划的约束条件, 结合单纯形法, 在现有资源的前提下, 给予用户最优的种植策略。

关键词: 信息查询; ARCGIS地图; 种植优化; 大数据

中图分类号: TP391

文献标识码: A

文章编号: 1000-2324(2017)01-0150-07

Information Query Optimization Platform about Shandong Province Farming System on Hadoop Background

JIANG Hong-hua¹, WANG Xiu-li¹, BAI Peng³, LI Guang-zhong¹, NING Tang-yuan^{2*}

1. College of Information Science and Engineering/Shandong Agricultural University, Tai'an 271018, China

2. State Key Laboratory of Crop Biology/Shandong Agricultural University, Tai'an 271018, China

3. College of Mechanical and Electronic Engineering/Shandong Agricultural University, Tai'an 271018, China

Abstract: Shandong province is a big agricultural province. And its agricultural information is vast. So it need some hadoop's technology to complete an integrated system, such as data's saving and taking, basic frame, data's processing, statistical analysis and so on. It integrates population, resource, environment and development, basing on regional ecological geography of different economic zone in Shandong province. This paper introduces in detail the development of the platform about farming system information query and planting optimization. In addition, it tells about systemic data's saving and taking, basic frame, data mining, statistical analysis and specific implementation strategy. This platform is made of login registration, information search, planting optimization and content manage system. It may show lots of data for user by the form of table, histogram, ARCGIS map visually and dynamically. It can give user the most optimal planting strategy, integrating simplex method and agricultural plant planning's constraint condition, such as the labor, fertilizer, fund, etc.

Keywords: Information query; ARCGIS map; planting optimization; hadoop

目前山东省尚处在半集约半自给半商品农作制水平, 东部沿海和鲁中农业区正在向现代农作制迈进, 初步形成了粮食、饲料、经济等作物协调发展的新格局。山东省农作制运行中存在着一系列不容忽视的非持续性制约因素^[1]。一方面粮食生产附加值低, 缺少园艺、经济等高附加值作物, 另一方面第二、三产业发展薄弱, 缺少农产品加工业。随着信息技术的发展应用, 农作制度信息化管理的作用和意义日趋重要。大数据背景下山东省农作制信息查询优化系统可以实现区域农作制度信息化管理, 根据全省不同经济带的区域生态地理及环境特征与国内外市场动向, 把人口—资源—环境—发展作为一个复合系统, 运用大数据的数据存取、数据处理、地理信息、模型模拟及专家系统统计分析等技术, 对区域农业资源、种植制度及相关产业和农产品市场供需与价格等方面的动态数据进行资源管理与分析, 对其发展前景预测与辅助决策, 实现农作制度的地区化、多元化、信息化, 实现高产优质高效的农业可持续化^[2,3]。利用山东省区域优势, 科学规划, 实现农作物种植效益的最大化^[4,5]。

收稿日期: 2015-10-23

修回日期: 2015-11-08

基金项目: 公益性行业(农业)科研专项(201103001、201503121-05)

作者简介: 姜红花(1971-),女,博士,副教授,主要从事农业信息化研究。E-mail:j_honghua@163.com

***通讯作者:** Author for correspondence. E-mail:ningty@163.com

1 开发工具

农作制信息查询与种植优化平台选用 Oracle 数据库服务器, ArcGIS Server 地图服务器。采用 B/S 模式, 通过 BlazeDS 实现服务器与客户端通信, 以 Hibernate 框架与数据库进行交互, Java 为后台服务器开发语言^[6], flex 和 ActionScript 为前台客户端开发语言^[7], 用 ARCGIS 技术发布 GIS 资源, 实现数据展示的多元化, 通过调用 ARCGIS API for Flex 将地图资源和其它资源(ArcGIS Online)嵌入到 Web 中^[8]。以 MyEclipse+FlashBuilder+Apache Tomcat 为平台开发环境^[9]。整个框架包括表现层、应用层和数据层 3 个部分^[10], 如图 1 所示。

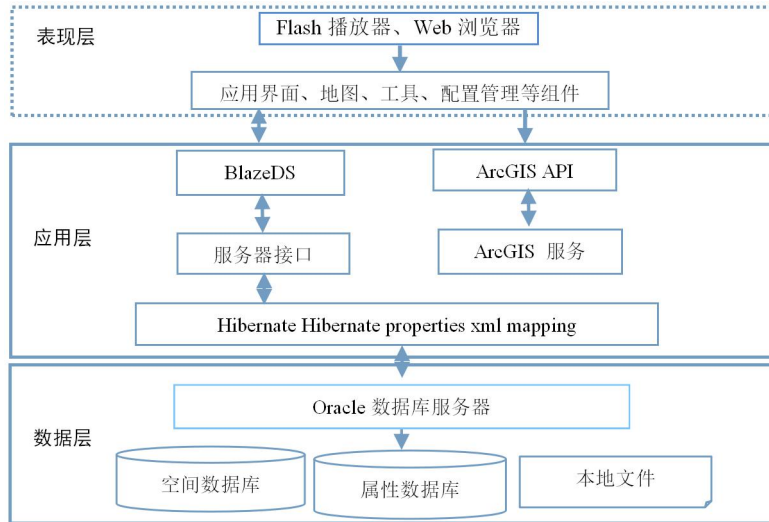


图 1 系统总体框架结构

Fig.1 The total framework of the system

客户端基于组件划分, 具有很强的扩展性和重用性, 基于 FLEX 的多种地图服务相结合以及通信技术使空间信息发布和浏览速度大大提高^[11]。

2 平台基础架构

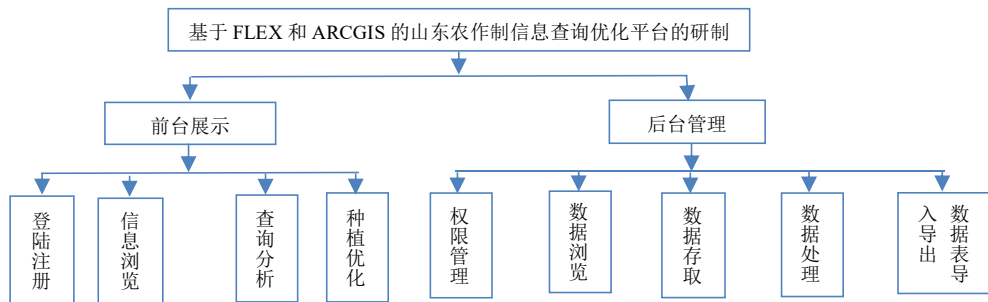


图 2 平台功能模块

Fig.2 Platform function module

农作制信息查询与种植优化平台分为前台展示和后台管理两个部分, 如图 2 所示。前台展示包括登录注册、农作制信息查询分析, 信息浏览以及作物种植优化模块; 后台管理包括权限管理、数据表浏览、数据维护、数据批量处理模块^[12]。其中信息浏览由地图显示和数据浏览构成; 作物种植界面由优化模块和数据备份模块构成; 后台管理由用户信息管理模块和数据库维护等模块构成。

登录注册: 实现用户的申请、注册和审批筛选。用户分普通用户和管理员用户。普通用户能进行信息的查询浏览及作物种植优化操作; 管理员用户能进行数据库管理、用户权限分配等。用户注册权限由管理员进行审批筛选发送激活邮件给用户。登陆密码采用 MD5 加密, 以保证系统的安全性。

运行环境: 运行在 windows 系统上, 服务器内存建议 2 G 左右, 处理器硬盘容量在 30 G 以上, 网络局域网广域网均可^[13,14]。

3 数据存取

山东省农作制的大量信息，需要采用大数据的存取技术。本系统数据的完整性和一致性相比较，为了使系统具有较快的响应速度，系统所有数据存在一个数据库中。数据库中的数据主要来源于农业部门统计的《山东省农业统计年鉴》及相关文献，数据时段为 1 年，涉及山东省 17 地级市农林牧副渔产量产值数据、播种面积、养殖面积、灌溉面积等农林用地、建设用地等共约 50 个数据指标。各类数据通过分析处理、优化运算，可得到各区域农作物的最优化种植生产潜力。

3.1 数据库结构

系统数据库共有 29 个数据表，分别是地区基本信息表、产值类别表、各市历年农林牧渔产值基础表、各市历年农林牧渔产值统计一级类别表、各市历年农林牧渔产值统计二级类别表、全省历年农林牧渔总产值表、全省历年粮棉油生产情况表、全省历年畜牧业肉产量存栏量情况表、全省历年渔业产量养殖面积情况表、土地类别表、各市地类面积情况表、农业主要产品类别表、农业主要产品生产情况表、水产品养殖捕捞类别表、水产品生产情况表、水产品养殖面积情况表、畜牧业主要产品类别表、畜牧业主要产品生产情况表、茶叶水果类别表、茶叶水果产量表、茶叶水果用地情况表、各市灌溉面积情况表、林业营林类别表、林业营林情况表、林业主要产品类别表、林业主要产品生产情况表、林业用地情况表、用户信息表。部分表结构如下：

表 1 产值类别

Table 1 Output value category

编号 No.	字段名称 Field name	数据类型 Data type	标识 Logo	性质 Property
1	CZLBH	INT	产值类别号	主键
2	MC	VARCHAR	产值名称	

表 2 地区基本信息

Table 2 Regional basic information

编号 No.	字段名称 Field name	数据类型 Type of data	标识 Logo	性质 Property
1	DQBH	INT	地区编号	主键
2	DQMC	VARCHAR	地区名称	

表 3 各市历年农林牧渔产值基础表

Table 3 Output basic list of agriculture, forestry, animal husbandry and fishery in every city every year

编号 No.	字段名称 Field name	数据类型 Data type	标识 Logo	性质 Property
1	ID	INT	序列号	主键
2	CZLBH	INT	产值类别号	外键
3	NF	DATE	年份	
4	DQBH	INT	地区编号	外键
5	CZ	DOUBLE	产值	

表 4 农业主要产品生产情况表

Table 4 Production of major agricultural products

编号 No.	字段名称 Field name	数据类型 Type of data	标识 Logo	性质 Property
1	ID	INT	序列号	主键
2	CPLBH	INT	产品类别号	外键
3	NF	DATE	年份	
4	DQBH	INT	地区编号	外键
5	MJ	DOUBLE	播种面积	
6	ZCL	DOUBLE	总产量	
7	CL	DOUBLE	单产	

3.2 数据表之间的关系

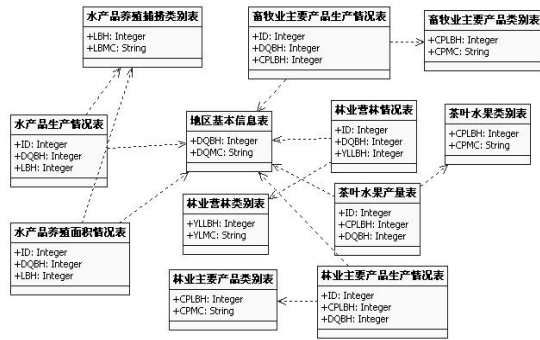


图 3 表间关系 1
Fig.3 Relationship 1 between tables

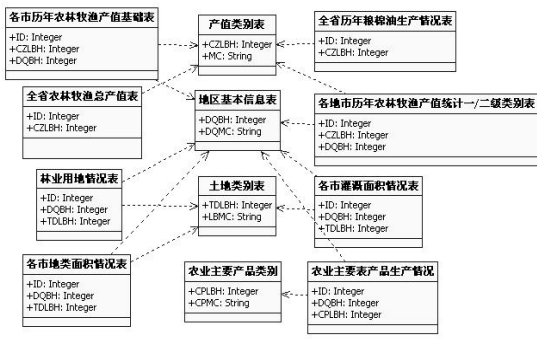


图 4 表间关系 2
Fig.4 Relationship 2 between tables

4 数据查询浏览模块

4.1 数据浏览

数据浏览是对全省历年农林牧副渔的产量、生产情况、种植面积（养殖面积）及历年总产值构成等数据的展示，数据纵览部分又分为全省历年数据纵览和各市历年数据纵览。单击农作制信息查询与浏览界面的左侧或选中地图中要浏览的城市，即可出现相应的数据。

4.2 查询分析

数据查询功能主要是对农林牧副渔业产量产值和播种、养殖等数据的查询，分为数据纵览和详细查询两部分，查询的内容主要采用三种方式进行数据展示：表格，柱状图/折线图，ARCGIS 地图。

表格：客户端设置查询条件后系统以表格的形式展现数据，表格列标题即用户查询设定的条件，并可对表格数据按条件排序，是通过公用函数 public function table Show()实现。

柱状图/折线图：将查询的数据生成相应柱状图或折线图的形式进行对比分析，并为图表加标签。柱状图/折线图可以显示单个城市多年份单个关键字、单个年份多个城市单个关键字和单个城市单个年份多个关键字的数据。当用户选择多城市多年份多关键字时，用户需选择柱状图的纵横坐标(地区-年份，地区-关键字)，比如用户选择地区-年份，需再次选择某一城市多个年份或某一年份多个城市来显示数据。其中柱状图如图 5 所示。



图 5 数据浏览柱状图
Fig.5 Data browsing histogram



图 6 鼠标选择截图区域
Fig.6 Screen shot area of mouse selection

详细查询模块由四个查询条件组合实现数据查询。四个查询条件分别为城市、年份、农林牧渔具体分类和对作物查询具体数据的限定，比如产量，产值，播种面积等。城市，年份以及字段查询均可进行多选设定，例如（淄博市、济南市）+（2006,2008）+农业小麦+（产值、产量）。

作物分类分为四大类（农林牧渔），大类下设二级分类及更详细的分类，例如农业下分粮食，油料，麻类等，粮食又分夏收粮和秋收粮，夏收粮又有具体的分类，以此类推。

ARCGIS地图标签可以单城市单年份不同类别对比。地图数据展示方式为在相应的城市区域上加数据标签，显示年份及关键字数据。通过ARCGIS API MapPoint方法确定城市所在区域后，在该区域添加图层，并通过InfoSymbol对象添加要显示的数据信息。

4.3 截图保存功能

截图功能用来保存浏览、查询的图表，方便用户制作报表等用途,如图 6 所示。用户点击截图按钮时前端调用服务器端截图接口，用户通过鼠标选取矩形截图区域后双击选择保存路径，将截图保

存为 JPG 格式文件。截图功能是使用 java.util.Robot 类的 public void snapshot()来实现对屏幕矩形区域的捕获。

4.4 地图查询功能

地图查询功能指用户在选定某一城市时地图对应区域突出显示，颜色加深，可以直观的看到选择的的城市。用户选取城市时，Action Script 根据查询选定的城市名调用 ARCGIS API 的 get FID()方法获得该城市 FID 地图属性，由该属性创建 query task 任务，确定该城市在地图上的区域，在该区域上添加一图层，设定该图层边界线加粗，并加深该区域的颜色，以达到突出显示的目的。

4.5 查询业务和种植优化业务

查询业务和种植优化业务的流程图如图7、图8所示。

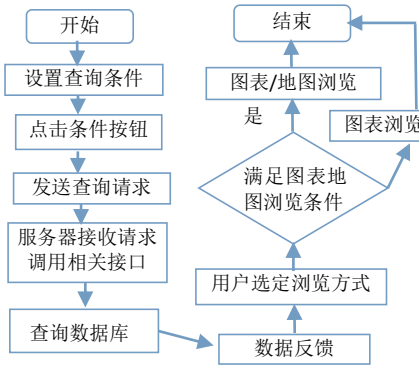


图 7 查询业务流程图

Fig.7 Query the business flow chart

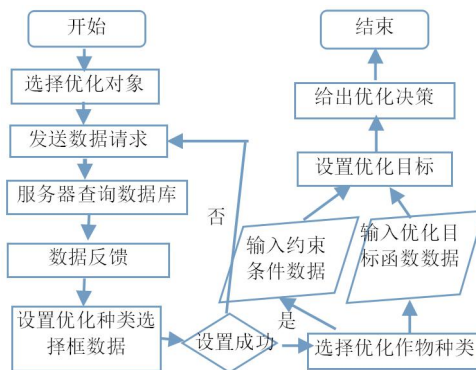


图 8 种植优化业务流程图

Fig.8 Planting optimization business flow chart

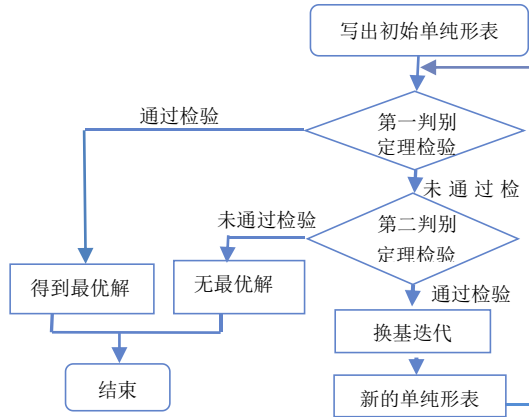


图 9 单纯形法解题流程图

Fig.9 Solving Problem Flowchart of simplex method

5 种植优化模块

5.1 线性规划基本思想

理论根据^[15]: 线性规划问题的可行域是 n 维向量空间 R_n 中的多面凸集，其最优值如果存在必在该凸集的某顶点处达到。顶点所对应的可行解称为基本可行解^[16]。

把线性规划问题约束方程组表达成典型型方程组，先找出基本可行解，鉴别判断是否是最优解；若不是，则按照一定法则转换到另一改进基本可行解再鉴别；若仍不是，则再转换,按此重复进行。基本可行解的个数有限，经有限次转换必能得出问题最优解。一般解题流程如图9所示^[17]。

5.2 线性规划的数学模型

线性规划的标准形式。

目标函数: $S=c_1x_1+c_2x_2+\dots+c_nx_n$

$$\text{约束条件: } \begin{cases} a_{11}x_1+a_{12}x_2+\dots+a_{1n}x_n=b \\ a_{21}x_1+a_{22}x_2+\dots+a_{2n}x_n=b \\ \vdots \\ a_{m1}x_1+a_{m2}x_2+\dots+a_{mn}x_n=b_m \\ x_1 \geq 0, x_2 \geq 0, \dots, x_n \geq 0 \end{cases}$$

其中, $b_i \geq 0 (i=1,2,\dots,m)$, $a_{ij}, b_i, c_j (i=1,2,\dots,m; j=1,2,\dots,n)$ 均为常数, $x_j (j=1,2,\dots,n)$ 为未知量。

5.3 优化设计示例

设某种植基地要种植粮食、棉花和油料3类农作物, 现有资源水资源3350 m³, 肥料470 kg, 劳动力575个, 粮食、棉花、油料所需资源情况如下表所示:

表 7 粮食、棉花、油料每公顷所需资源情况表
Table 7 Resources of grain, cotton and oil per hectare

资源 Resource 类型 Type	水资源 Water resources	肥料 Fertilizer	劳动力 Labor force	收益 Gain income
粮食	3780m ³	428kg	600 个	8203
油料	2350m ³	548kg	525 个	9439
棉花	3330m ³	520kg	550 个	15812

土地资源约束: $x_1+x_2+x_3=1$

水资源约束: $3780 x_1+2350 x_2+3330 x_3 \leq 3350$

肥料约束: $428 x_1+548 x_2+520 x_3 \leq 470$

劳动力约束: $600 x_1+525 x_2+550 x_3 \leq 575$

目标收益最大即: $8203 x_1+9439 x_2+15812 x_3$

由此建立的种植优化模型如下:

$$\text{max } \begin{cases} S=8203 x_1+9439 x_2+15812 x_3 \\ x_1+x_2+x_3=1 \\ 3780 x_1+2350 x_2+3330 x_3 \leq 3350 \\ 428 x_1+548 x_2+520 x_3 \leq 470 \\ 600 x_1+525 x_2+550 x_3 \leq 575 \\ x_1 \geq 0, x_2 \geq 0, x_3 \geq 0 \end{cases}$$

5.4 种植优化界面设计

种植优化界面分为四部分, 第一部分为优化对象, 作物种类设置, 数据来源于数据库; 第二部分为目标函数, 约束条件设置; 第三部分为优化目标设置及结果显示; 第四部分为历史操作记录; 用户还可以选择将操作数据与结果存入文件中保存备用。种植优化界面如图 10 所示:

6 农业统计数据后台数据管理系统

6.1 后台管理模块

后台管理模块是实现数据维护及用户管理, 用户管理包括用户的启用、禁用、筛选激活及权限的管理; 数据维护主要是各数据表的管理, 包括: 数据表浏览、维护、数据批量删除及数据表整体导入导出。数据浏览表的界面如图 11 所示。管理主界面左侧点击相应的项可浏览表内数据。

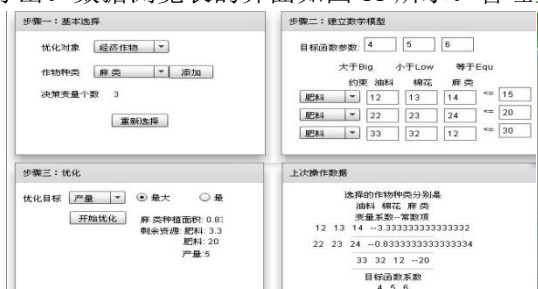


图 10 种植优化操作界面图

Fig.10 Planting optimization interface



图 11 数据浏览

Fig.11 Data browse

6.2 数据库维护

数据库维护提供数据基本的增、删、改、查等功能。通过农产品类别号查询定位到目标数据项进行修改或删除操;也可根据共性条件查询删除批量数据或批量导入数据,并设有“取消”按钮。查询字段有年份和产品类别号,其中年份是精确查询,类别号是匹配文本开头的模糊查询。

6.3 数据表的整体导入和导出

数据添加有两种方式:一是逐条数据依次添加;另一种是多条数据批量添加。数据的导出:将表数据全部导出为 excel 文件或根据 excel 文件中相应数据表结构,按数据字段的顺序导出。

空间数据导入到数据库前,要对数据进行预处理,包括数据格式的转换、图层的筛选、数据完整性的检查、拓扑关系的检查等,以确保空间数据的数据质量及整个系统运行的稳定性。其次,根据应用的需要,按照数据库数据表设计原则,对数据表结构进行设计;按照数据层的数据组织结构,建立数据集和要素类,利用 Arc Catalog 导入工具或编写的导入程序将处理好的空间数据导入到空间数据库中。属性数据和元数据,则将整理好的文字资料导入到 ORACLE10g 数据库对应表空间的数据表中。

数据层利用 Arc SDE 进行空间数据的管理,实现了空间数据和属性数据的无缝集成,为 WEBGIS 的信息查询、分析、决策提供了强大的数据支持。

7 总结与展望

农作制信息查询与优化平台利用 Oracle 数据库,结合 Flex 技术和 ARCGIS 地图服务开发的农业综合服务性平台,界面丰富,交互性强,动态数据查询、直观多样的数据浏览及种植优化决策,将数据与地图更好的结合在一起,实现了数据柱状图地图显示、地图排序及不同作物数据对比等功能。许多农业生产问题借助线性规划思想来解决,运用单纯形法数学模型表达,求出问题最优解,给予用户最优的种植策略建议。

目前平台数据只是局限在山东省 17 地市,今后会将范围具体到县区,让数据更加详细;采用更高效的优化算法取代单纯形法,充分利用农业资源,挖掘现有耕地潜力,实现种植业、养殖业及加工业的协调发展,有效解决经济发展与资源、环境间的矛盾。

参考文献

- [1] 李增嘉,王芸,韩宾.山东省现代农作制发展现状、问题与对策[M].南京:东南大学出版社,2006:232-240
- [2] 刘巽浩.论中国农业的集约化与持续化[J].农业现代化研究,2000,21(1):1-8
- [3] 刘巽浩,陈阜,高旺盛.我国东中西片农作制特征与战略优先序[J].农业现代化研究,2004,25(5):321-329
- [4] 韩惠芳,周勋波,宁堂原,等.山东现代农作制特征及其发展策略[J].作物杂志,2010(1):1-5
- [5] 刘金爱.我国农业信息化发展的现状、问题与对策[D].青岛:青岛农业大学,2009
- [6] 麻清源,张兵,张超.基于组件式GIS的数字农业空间信息管理平台开发研究[D].上海:华东师范大学,2007
- [7] 邬群勇,王钦敏,汪小钦.农业空间信息资源共享服务平台[D].福州:福州大学,2012
- [8] 乔珂.Action Script 3.0权威指南[M].北京:电子工业出版社,2008
- [9] Christian Bauer, Gavin King. Java Persistence with Hibernate: Hibernate 实战[M].第2版.杨春花,彭永康,俞黎敏,译.北京:人民邮电出版社,2008
- [10] 吴信才.基于Flex的webGIS开发[M].北京:电子工业出版社,2010
- [11] 山东省统计局.山东省统计年鉴2010[M].北京:中国统计出版社,2010
- [12] 兰天,曲鹏东,孙高飞,等.Flex企业应用开发实战[M].北京:机械工业出版社,2010
- [13] Doug Winnie. Fundamentals of Action Script 3.0[M].USA:Peachpit press,2011
- [14] Abdulsalam Z, Akinola MO, Buwanhot YY. Problems and Prospects of Information and Communication Technologies Application in Agriculture in Nigeria[J]. The information manage,2008,8(1):7-16
- [15] 徐文君,冀德刚,李红智.单纯形法在农作物布局中的应用[D].保定:河北农业大学,2010
- [16] 吴祈宗.运筹学与最优化方法[M].北京:机械工业出版社,2003
- [17] 朱春江,唐德善.基于线性规划模型的农业种植业结构优化研究[D].南京:河海大学,2006