

# 基于 MATLAB GUI 的研究生学业评估系统设计

徐增伟, 曾黄麟, 江 泌, 李利杰

(四川理工学院自动化与电子信息学院, 四川 自贡 643000)

**摘 要:**为了设计一个研究生学业评估系统,提出了一种基于 MATLAB GUI 的系统设计方法。Matlab 不仅具有的强大科学计算功能,而且具有满足一般要求的数据处理以及界面设计开发功能。利用 Matlab 这些功能,实现一个具有科学计算功能、简单易用、具有开放式可扩展环境的研究生学业评估系统设计。

**关键词:**MATLAB GUI; 学业评估; 系统设计

**中图分类号:**TP311

**文献标识码:**A

## 引 言

Matlab 是 MathWorks 公司开发的一种科学计算软件。Matlab 拥有强大的面向不同领域的工具箱,是广大科研人员进行科学研究的得力助手。Matlab 不仅具有强大的科学计算功能,还具有满足一般要求的数据处理以及界面设计开发功能。Matlab GUI(Graphic User Interface)是内置于 Matlab 的进行图形界面开发的模块。特别是所附带的 30 多种面向不同领域的工具箱支持,使得它在许多科学领域中成为计算机辅助设计和分析、算法研究和应用开发的基本工具和首选平台<sup>[1]</sup>。

Matlab 不仅具有强大的科学计算功能,还具有满足一般要求的数据采集以及界面设计开发功能。Matlab GUI(Graphic User Interface)是内置于 Matlab 的进行图形界面开发的模块<sup>[2]</sup>。直观地说,GUI 就是由各种图形对象组成的用户界面,在这种用户界面下,用户的命令和对程序的控制是通过“选择”各种图形对象来实现的。基本图形对象分为控件对象和用户界面菜单对象,简称控件和菜单。用户界面是指人与机器之间交互作用的工具和方法。图形用户界面(GUI)则是由窗口、光标、按键、菜单、文字说明等对象(Object)构成的一个用户界面。

GUI 有十分广泛的应用, MATLAB 最具特色的 Demo 演示部分中绝大部分都是 GUI 实例, GUI 的使用可以使我们要传达的信息变得更为简便、高效。此外编程语言也十分严谨,每个句柄的属性都十分详细,无论从界面属性还是回调函数等功能属性上讲, MATLAB 的 GUI 编程都是十分紧凑的。在进行 GUI 设计时,我们还可以通过 GUIDE 对界面各控件进行窗口性直观设计,这样就方便了许多。

## 1 研究生学业评估办法

### 1.1 学业成绩要求

根据某高校硕士研究生学业评估办法,硕士研究生学业评估的要求如表 1 所示。

### 1.2 综合测评计算方法

一年级:成绩积分占 90%、学科评定占 10%,另加科研积分;二年级:成绩积分占 10%、科研积分占 80%、学科评定占 10%。

成绩积分:公共学位课程积分求和占 50%,专业学位课程积分平均占 30%,选修课程积分平均占 20%。一年级成绩积分大于零方可参评。课程积分:课程得分 = 课程成绩 - 该门课程平均成绩。学科评定工作组织,二级学院需组成考评小组,成员由分管研究生工作院领

收稿日期:2011-05-27

基金项目:四川省科技厅应用基础研究专项课题(2011JY0051);四川省白酒及生物技术重点实验室重点专项课题基金(NJ2010-01);研究生创新课题基金(y2010013)

作者简介:徐增伟(1983-),男,山东东明人,硕士生,主要从事图像处理、模式识别及智能信息处理方面的研究。

导、辅导员、导师代表、学生干部代表、学生党员代表、学生代表组成,一般以 7 人为宜。考评小组对申请人作出

表 1 硕士研究生学业评估要求

成绩要求	一年级			二年级		
	优秀	良好	一般	优秀	良好	一般
学位课成绩	单科 $\geq 80$	80 > 单科 $\geq 75$	75 > 单科 $\geq 70$	单科 $\geq 80$	80 > 单科 $\geq 75$	75 > 单科 $\geq 70$
选修课成绩	平均 $\geq 80$	80 > 平均 $\geq 75$	75 > 平均 $\geq 70$	平均 $\geq 80$	80 > 平均 $\geq 75$	75 > 平均 $\geq 70$
GET 成绩	单科 $\geq 75$	75 > 单科 $\geq 70$	70 > 单科 $\geq 65$	单科 $\geq 75$	75 > 单科 $\geq 70$	70 > 单科 $\geq 65$
CET6 成绩	单科 $\geq 425$	$\geq 0$	$\geq 0$	$\geq 425$	$\geq 0$	$\geq 0$
综合评定成绩	$\geq 20$	前 25% 以内	前 40% 以内	$\geq 25$	前 25% 以内	前 40% 以内
论文成绩	$\geq 0$	$\geq 0$	$\geq 0$	核心 $\geq 3$	$\geq 1$	$\geq 1$

## 2 系统设计与实现

### 2.1 GUI 的建立

一般地, MATLAB 中 GUI 的建立有两大方式<sup>[3]</sup>。第一种是直接通过程序编写的产生对象,即利用 uicontrol、uimenu、uicontextmenu 等函数以编写 M 文件的方式来开发整个 GUI。此种方式的优点在于 GUI 菜单的建立比较齐全,并且不会额外产生一个 .fig 文件,并且程序代码的通用性非常高,因此当完成一个 GUI 后,该程序代码就可以复制到一般的 MATLAB 的 M 文件中使用,也可以 GUIDE 的 M 文件中使用,如已通过程序编写的方式编写一系列的 uimenu 菜单选项,因此就可以直接将这些程序复制输入到 GUIDE 的 Opening Function 中,直接在 GUIDE 中运用这些菜单选项,如此就可以节省额外开发相同或类似对象的时间。程序编写来建立 GUI 对象的最大缺点就是 GUI 对象位置的配置,若不是非常熟悉的用户可能会比较难以控制。

另一种方式是直接发通过 MATLAB 的 GUI 编辑界面——GUIDE 来建立 GUI。这个界面当前 (MATLAB 2009a 版) 已改善得非常实用,对于使用过窗口程序的用户或新手而言,是一个非常不错的选择,它有点类似 Visual Basic 这一类的开发环境,只要直接通过鼠标将对象拖拽到目的地,就可以快速地建构出 GUI;此外,这种方式在 M 文件的管理上也比较好,因此如果在日后要修改部分程序代码,可以快速且容易地找到修改部分的内容。本文就是选用这种方式来设计和实现研究生奖学金评价系统的。下面重点介绍一下使用 GUIDE 来建立 GUI 设计过程。

第 1 步:使用 Matlab 的 GUIDE 编辑器建立 GUI,直接在命令窗口输入 GUIDE,或由 MATLAB 的 Start 菜单,或用工具栏的建立 GUI 按钮,都可以打开 MATLAB 的 GUIDE 编辑器。第 2 步:建立 GUI 后,系统会自动生成 fig 以及 .m 文件,犹如我们常用的界面资源文件和后台代码文件,关于事件处理的代码可以在 m 中编写。第 3 步:通过 Compiler 编译为执行文件或由 MATLAB Runt-

ime Server 封装文件,即可在没有 MATLAB 的环境中使用。

### 2.2 系统设计与编译

先由 MATLAB 的 GUIDE 编辑器建立整个 GUI。如图 1 所示。再根据需要设置建立的 axes、uicontrol 等对象的相关属性<sup>[4]</sup>及 Callback 来实现相关的控制操作。研究生学业评价系统只适用于每个学生对自己学业成绩等级做出评估。在使用时用户需要自己输入相应的数据,按下系统设置的相应按钮即可完成操作。系统会对用户的学业成绩等级做出判定以及执行用户其他的操作。



图 1 四川理工学院研究生学业评估系统界面

为了对程序代码进行保密,以及方便用户发布可以在不同平台运行的可执行程序<sup>[5]</sup>, Matlab 提供了 Matlab 程序编译与发布功能。在进行 Matlab 程序编译之前,需要设置编译工具。在 Matlab 命令窗口中,输入运行 mbuild - setup,进行编译工具设置。命令运行后,Matlab 会检测已经安装在本机上的编译器,用户可以自行选择任意一个编译器,选择完毕后,运行编译命令,Matlab 将使用用户选定的编译器对程序进行编译。本文选择 Lcc - win32 C 2.4.1 编译器。

设置完毕编译工具后,即可以在命令窗口中运行编译命令,将 Matlab 程序编译成可执行程序。命令格式为:mcc - options mfile1 mfile2 ... mfileN;options 为编译

参数; mfile1 为待编译的 Matlab 文件, 如果包含多个 Matlab 文件, 可以直接列在 mfile1 后面。按照以下编译方法, 在命令窗口中输入 `mcc -m GREawards` 即生成独立执行文件。

### 2.3 发布程序

Matlab 文件编译完毕后, 生成的可执行程序还需要 Matlab 环境的支持, 要发布到没有安装 Matlab 的机器上运行, 还需要进行最后一项工作, 即打包 Matlab 组件运行环境<sup>[6]</sup> (Matlab Component Runtime, MCR)。将 MCR 与可执行程序一起打包, 拷贝到其他没有安装 Matlab 的机器上, 在该机器上安装 MCR, 安装完毕即可以直接在该机器上运行编译的 Matlab 可执行程序。Matlab2009a 中, 已经内置了 Matlab 编译发布工具箱, 在 Matlab 命令窗口中运行 `deploytool` 命令, 即可以调出该工具窗口, 通过该工具可以方便、快捷地编译发布 Matlab 程序。具体方法是在该工具中导入要打包发布的工程, 然后进行打包即可生成可执行程序 `GREawards.exe`。

### 3 结束语

Matlab 具有强大的科学计算功能和图形显示功能, 选用 Matlab 作为数据处理与分析工具是合适的, 但其界面开发和数据采集功能相对较弱, 不如 VisualC++ 或者

LabView 强大。不过, 对于一般的界面开发, Matlab 是足以胜任的, 同时也可以直接利用其强大的数据分析与处理功能。本文基于 Matlab GUI 的研究生学业评估系统的编程实现表明利用 Matlab GUI 可以进行包含一般界面元素的界面开发, 在此基础上, 加入人机对话模式, 实现数据的实时输入与分析处理, 并将评价结果进行显示。

### 参考文献:

- [1] 李显宏. MATLAB 7.X 界面设计与编译技巧[M]. 北京: 电子工业出版社, 2006.
- [2] 陈焱光. 精通 MATLAB GUI 设计[M]. 北京: 电子工业出版社, 2008.
- [3] 徐俊文, 王强. MATLAB 环境下的 GUI 编程[J]. 内蒙古民族大学学报, 2006, 21(6): 640-643.
- [4] 姚秀芳, 崔松菲. 基于 MATLAB GUIDE 的程序设计[J]. 电脑知识与技术, 2009(27): 7768.
- [5] 王战军, 沈明. 基于 Matlab GUI 的串口通信编程实现[J]. 现代电子技术, 2009(9): 38-40.
- [6] 陈安宇, 陈伟. Matlab 图形用户界面的应用研究[J]. 机电工程技术, 2008, 16(10): 16-21.

## An Implementation of Graduate Achievement Assessment System Based on MATLAB GUI

XU Zeng-wei, ZENG Huang-lin, JIANG Bi, LI Li-jie

(School of Automation and Electronic Information, Sichuan University of Science & Engineering, Zigong 643000, China)

**Abstract:** In order to design a graduate achievement assessment system, a method of system design based on MATLAB GUI is proposed. Matlab not only has the powerful scientific computing functions, but also has the capabilities meeting the general requirements of the data processing, interface designing and program developing. A graduate achievement assessment system is implemented by Matlab's functions and capabilities which is of scientific computing, easy accessibility and open extensible environment.

**Key words:** MATLAB GUI; achievement evaluation; system design

# 基于.NET的高校学生党员信息管理系统设计与实现

韦崇岗<sup>1</sup>,梁小晓<sup>2</sup>,兰静<sup>1</sup>

(1. 四川理工学院自动化与电子信息学院,四川 自贡 643000; 2. 四川理工学院计算机学院,四川 自贡 643000)

**摘要:**为了实现高校学生党员信息化管理,分析了高校学生党员管理工作中的现状,提出基于 B/S 结构架构的信息管理系统方案。系统采用 .NET 技术框架,ADO.NET 数据库访问技术和 SQL 数据库技术,依托校园网实施,成功地实现了高校学生党员的信息化管理。

**关键词:** .NET;高校学生党员;信息管理系统

**中图分类号:** TP311

**文献标识码:** A

计算机和信息技术的高速发展,对教育产生着深刻的影响,计算机管理已成为高校信息管理的重要工具和方法。在当今校园信息化建设日趋完善的背景下,高校学生党组织逐渐向社区型、社团型和群簇型等多种组织形式转变,学生在学习和生活的空间、时间上变得更加自由,使得学生党员在相互交流、学习讨论、培养发展、教育管理和组织生活开展等方面面临着许多新情况和新问题,亟待我们用信息化管理的方法实现信息化管理,以规范传统管理模式,创新管理方法。因此,适应客观环境和主观世界改变的新情况,使用计算机和网络技术,设计开发基于 .NET 的高校学生党员信息管理系统,是做好新时期高校学生党建工作的迫切需要。

## 1 需求分析

在设计高校学生党员信息管理系统时,一方面要充分考虑到不同用户角色在学生党建活动中的具体功能和作用;另一方面要考虑到学生党员在培养、教育、考察、发展过程中,各种角色的相互关系。通过这两个方面的考虑对系统进行分析。

### 1.1 系统流程分析

高校学生党员信息管理的一切常规工作都是围绕学生党员发展这个核心过程进行的,主要经历了培养、教育、考察、发展、转正和组织关系接转这几个过程,由

此形成的系统的综合业务流程<sup>[1]</sup>,如图 1 所示。

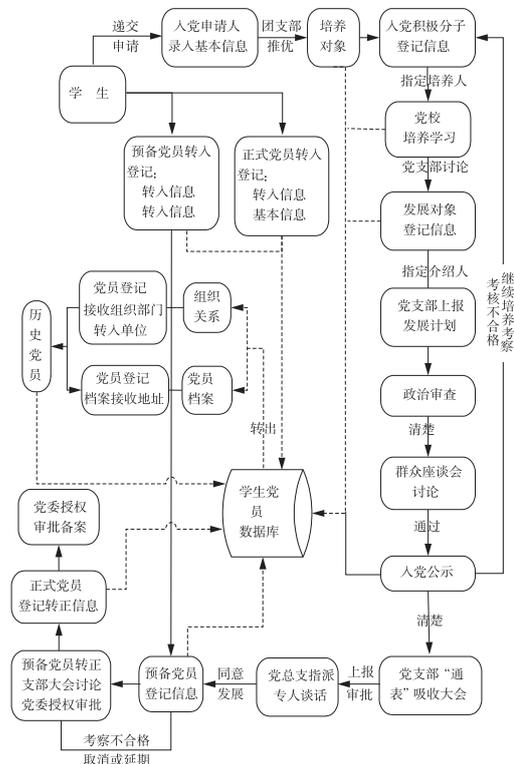


图 1 综合业务流程图

### 1.2 用户群分析

本系统的设计立足于以党委组织部统一监控管理、院(系)进行具体管理、党支部进行组织介绍人、培养人

收稿日期:2011-04-19

基金项目:四川省教育发展研究中心课题(CJF06059);四川理工学院科研基金项目(2006ZR013)

作者简介:韦崇岗(1978-),男,四川宜宾人,助理研究员,主要从事信息管理与思想政治教育方面的研究。

和入党人进行基础信息分散录入的管理流程。建立了一个基于不同角色,对入党人员进行培养、教育、考察和管理的相互独立,又相互关联的信息管理系统。涉及的用户有学生、团支部、党支部、二级学院、组织部等五种用户。

### 1.3 功能分析

高校学生党员信息管理系统分为前台管理和后台管理。前台管理包括党的政策理论宣传,党员培养考察吸收过程信息管理,党员档案信息管理,党员组织关系转接信息管理,党员信息管理和党组织活动管理等六大管理功能。后台管理包括用户管理、院(系)党总支管理、党支部管理、数据管理、系统信息管理模块。

### 1.4 系统数据流分析

根据用户群、功能和 workflows 分析,得出系统数据流程图如图 2 所示。

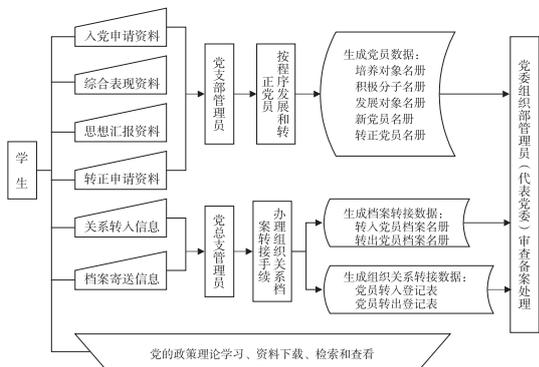


图 2 系统数据流图

## 2 系统设计

### 2.1 系统系统总体架构

系统依托校园网,应用于互联网,采用 B/S 的多层技术结构,利用 .NET 开发工具集、ADO .NET 数据库访问技术和 SQL Server 数据库存储技术架构。服务器端运行 Microsoft IIS,客户端利用各种浏览器实现操作,各级用户根据权限进行管理,管理层次分明,结构清晰,如图 3 所示。

本系统采用独特的 ActiveX 控件技术,使用者可以方便的对本系统结果与 office 结合起来,既可以对相应数据进行备份,又可以对数据进行二次分析<sup>[2]</sup>。

.NET 作为一个开发环境和可以作为平台支持下一代 Internet 的可编程结构<sup>[3]</sup>,其侧重点从传统的连接到互联网的单一网站或设备上,转移到计算机、设备和服务群组上,使其通力合作,提供更广泛更丰富的解决方

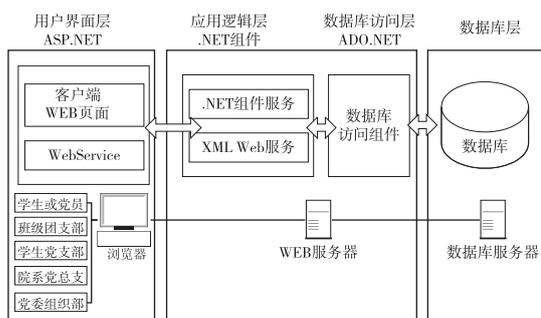


图 3 系统总体架构

案。ADO .NET 技术是解决 web 和分布式应用程序而开发的,它是采用不同程序语言创建的应用程序能互相通信,并考虑了伸缩性、无状态性和 xml 的问题。作为一组包含在 .NET 中的库,主要用于在 .NET 应用中各种数据存储间的通信。它包含了可与数据源连接、提交查询并处理结果的类<sup>[4]</sup>。在数据存储方面,较大的应用系统原则上选用 SQL Server 数据库。目前,SQL Server 2008 已经发布,但其性能还有待于实践的检验。因此,本系统采用已经被广泛成功应用的 SQL Server 2000 版本,它具有提供数据定义语言、交互式查询、应用程序的编译、事务运行管理、数据的组织与存取和数据库的维护等强大的功能。

### 2.2 数据库设计

#### (1) 系统 E-R 图

根据对系统的用户类型、主要功能、业务流程、数据流图 and 数据库设计原则,分析得到系统 E-R 图,如图 4 所示。

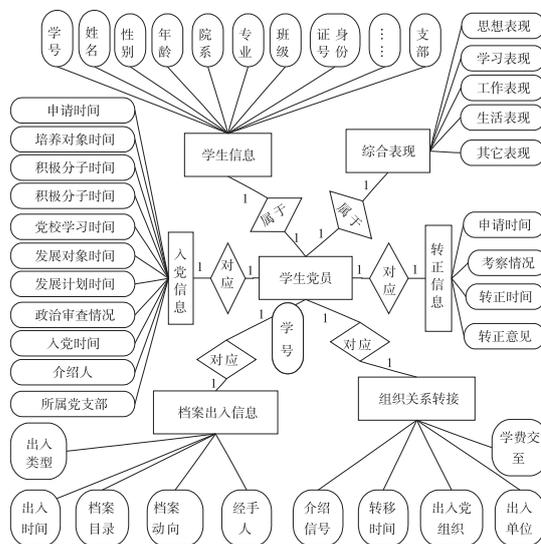


图 4 系统 E-R 图

## (2) 系统数据库表

数据库是系统信息管理的基础,数据结构的优劣直接影响着系统应用的效率及其扩展性。高效的数据库结构设计能大大提高数据存取效率,使数据完整性和一致性得到保证,同时,更有利于系统程序的实现。

在高校学生党员信息管理系统设计中,根据系统分析应该设计学生用户表、管理员用户表、档案目录信息表、组织关系出入信息表、培养对象过程评价信息表、入党积极分子培养过程信息表、发展对象过程信息表、党员发展过程信息表、党员转正过程信息表等。

## 2.3 系统功能模块设计

根据系统功能,系统设计了六个子系统和系统基础设置等七个功能模块。六个子系统分别为:党建新闻信息子系统、党员培养子系统、档案管理子系统、组织关系转接子系统、党员信息管理子系统、党组织生活信息管理子系统。其主要业务工作为:

**党建新闻信息子系统:**包括党的理论政策、党建新闻和党建工作通知公告三个板块。业务功能有新增、编辑、删除、检索、导出等五项功能。

**党员培养子系统:**包括党员的培养、管理和教育三个板块。业务功能根据不同用户角色有所不同:入党人向党组织提交自己的入党申请和综合表现情况信息,业务功能有提交入党申请、提交综合表现材料信息等功能;团支部根据入党申请人的综合表现情况向党支部推荐培养对象,业务功能有提交推荐意见等功能;党支部审查培养对象的综合表现材料,把培养对象列为入党积极分子、指定培养联系人、把入党积极分子列为发展对象、指定入党介绍人、上报发展计划、提交政治审查结论、提交群众座谈会意见、提交支部大会意见、上报拟发展党员名单、党员组织关系转入转出信息审查、党员档案转出信息审查、提交组织生活开展申请等;培养联系人向党支部提交入党对象的培养意见;入党介绍人向党组织介绍发展对象的综合表现情况;党总支审查入党人员的资格、审查党支部发展程序、组织发展工作审批和向党委上报申请备案;党委组织部代表党委审查发展结论。

**档案管理子系统:**包括档案形成、档案管理、档案转出三个板块。档案形成主要是形成党员档案的材料目录和综合表现情况,材料目录主要包括入党申请书、团组织推优材料、党校结业证书、入党积极分子登记表、群众座谈会记录、入党公示登记表、入党志愿书、预备党员

培养考察登记表、转正公示登记表和各个阶段的思想汇报;综合表现情况主要是生成由本人、培养联系人、团支部、党支部、入党介绍人和党总支在发展党员过程中形成的意见综合材料。档案管理主要是审查党员档案的完整性。档案转出主要是党员向党支部提交自己的转出档案单位和地位,党支部审核上报,党总支将党员档案寄送至档案接收单位。

**组织关系转接子系统:**包括转入和转出两个板块。组织关系转入主要由党员个人提交自己的转出组织部门、转出单位、介绍信号。组织关系转出主要由党员将确定自己组织关系接收组织部门、确定组织关系去向单位。由党支部负责接收和审查组织关系转入和转出信息,上报党总支;党总支形成报表,上报党委组织部,办理组织关系转入转出手续,更新党员数据库。

**党员信息管理子系统:**包括党员基本信息、入党过程中各个阶段的时间、形成党员名册。

**党组织生活信息管理子系统:**包括组织生活方案提交、修改审批、实施总结三个板块。主要由党支部提出申请、党总支修改审批和党支部实施总结。

**系统基础设置:**包括审批用户申请、院系管理、党支部管理、用户与管理人管理、系统数据管理、系统信息设置。

## 2.4 系统安全性考虑

系统安全性考虑主要基于三个方面:一个方面是基于对B/S结构本身的安全性因素,B/S架构的信息管理系统,其系统代码存放在服务器上,数据库存放在数据库服务器上,访问与管理在浏览器端实现的分层管理模块本身就具有非常高的安全性。二个方面是系统是基于校园网建设的,系统信息数据统一存放在学校的网络设备上,依托校园网的硬件和软件防火墙可以在非常大的范围内确保系统安全<sup>[5]</sup>。第三个方面是利用微软视窗的NTFS和DBMS的用户-角色所在不同层次分别对管理员权限进行分配限制,避免越权访问和交叉访问产生的系统威胁。另外,系统在设计时设计了系统登录日志管理功能,可以自动记录所有用户的登录和操作信息,监视系统运行<sup>[6]</sup>。

## 3 系统实现

### 3.1 数据库实现

#### (1) 创建数据库

创建学生党员信息数据库 DanJian\_info,用来保存

本系统的所有数据。创建数据库的脚本如下为:

```
create database DanJian_info
go
```

### (2) 数据表实现

本系统使用 SQL SEVER 2000 作为系统数据库服务器,采用一库多表组织<sup>[7]</sup>,数据库根据学生的学号是唯一的,数据表之间由学生学号作为主关键字建立联系关系。数据库表主要包括个人基本信息、入党申请人信息、培养对象信息、入党积极分子信息、发展对象信息、预备党员信息、正式党员信息、档案目录信息、组织关系转接信息、组织生活统计信息、用户信息、院系信息、支部信息等信息表。

### 3.2 用户登录实现

系统登录模块主要用于验证用户登录系统时输入的用户名、密码是否正确,如果正确,用户进入系统,否则将不能进入此系统,如图 5 所示。

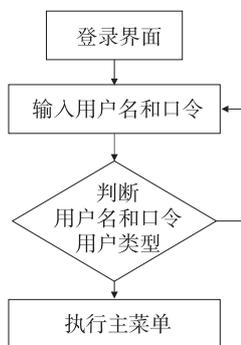


图 5 用户登录过程

在系统后台数据库中,建立 Users 用户表,包括用户主要信息和访问权限代码。访问权限控制着用户对数据库进行读取、插入、删除、修改等操作的权限组合。用户在系统提示下输入登录名、口令和校验码,并提交输入结果。Web 服务器在收到请求后,激活 .NET,通过 ADO .NET 驱动 ODBC 与用户表数据库建立连接,运行 .NET 处理程序,将输入信息与数据库记录进行比较,比较结果符合,允许其在权限范围内访问数据库,否则拒绝访问退回登录界面<sup>[8]</sup>。

### 3.3 系统设置与管理功能实现

系统用户和管理员的增加、修改、删除并分配相应权限;设置系统基本信息;党总支和党支部管理;导入学生基本数据,并进行编辑;对系统数据进行备份和恢复;发布党的政策、党建新闻和党建通知信息等。

### 3.4 党员培养信息管理实现

不同用户对学生入党过程信息的录入、查询、审核,生成及打印《党员公示名单》、《学生预备党员绝对增加名册》、《学生党员转正名册》、《学生党员发展计划名册》、《学生入党积极分子名册》、《学生培养对象名册》和《学生发展对象名册》等相关表格。

### 3.5 党员档案信息管理实现

对形成的党员档案进行编号、查询,生成及打印《学生党员档案目录》和《毕业生党党员档案转出登记表》,形成、接收和转出学生党员档案。

### 3.6 党员组织关系转接信息管理实现

对转入和转出党员信息进行登记、查询,生成及打印《中国共产党党员组织关系介绍信》、《毕业生党员组织关系接转登记表》,接收与发放介绍信。

### 3.7 组织生活信息管理实现

对各支部组织开展的组织生活进行申请、审批、总结进行管理,形成与打印《学生组织生活实施方案申请》、《学生组织生活总结》和《学生组织生活目录》。

### 3.8 系统数据与报表管理实现

本系统采用 .NET 和 SQL 数据技术实现信息管理系统,将调用报表处理的系统模块,实现报表生成。系统工作过程中,通过生成 HTML 文件或 XML 文件实现与 Office 软件连接<sup>[9]</sup>。目前,Office2000 和之后的版本都支持 HTML 作为本地的文档格式,文档在需要的时候以 HTML 格式或使用 XML 标记创建,之后用 MIME 类型传给客户端,客户端的文字将在 Office 中显示出来,从而实现对文档的编辑、保存,实现 .NET 与 Office 的有效连接<sup>[10]</sup>。

## 5 结束语

通过对高校学生党建工作的分析,设计了基于 B/S 的高校学生党员信息管理系统。系统基于校园网架构,具有安全性高,管理方便,扩展性和维护性好、人机界面友好的特点,大大地提高了高校党建工作效率和管理水平。

### 参考文献:

- [1] 雷 蕾,钟永强.高校学生党员信息管理系统的设计及实现[J].科技信息,2010(5):470-471.
- [2] 徐金梧,张晓彤.Pritchard J.COM 与 CORBA 本质和互用[M].北京:清华大学出版社,2002.

- [3] 眭碧霞,李春华.ASP.NET 程序设计与开发[M].西安:西安电子科技大学出版社,2008.
- [4] david sceppa.ado.net 技术内幕[M].北京:清华大学出版社,2005.
- [5] 汤永刚.基于安全 Web Services 应用系统中身份难服务的实现[J].计算机系统应用,2007(12):107-108.
- [6] 彭尚源,游君臣.基于 .NET Remoting 技术的 Web 安全服务应用[J].四川理工学院学报:自然科学版,2006,19(6):43-46.
- [7] 陈香生.基于 SQL 技术的数据库开发实践的探讨[J].信息与电脑:理论版,2010(11):103.
- [8] 王祥瑞,黄国华.ADO 数据库访问技术[J].长春大学学报,2010(2):62-63.
- [9] 胡开明,陈建华.用 .NET 实现对 SQL SERVER 数据库安全的动态监控[J].四川理工学院学报:自然科学版,2010,23(3):300-302.
- [10] 郝少华,韩 燮.基于 XML 技术的异构关系数据库集成模型[J].计算机工程与设计,2010(24):5285-5288.

## Design and Implementation of the MIS for College Party Members Based on .NET

WEI Chong-gang<sup>1</sup>, LIANG Xiao-xiao<sup>2</sup>, LAN Jing<sup>1</sup>

(1. School of Automation and Electronic Information, Sichuan University of Science & Engineering, Zigong 643000, China;

2. School of Computer Science, Sichuan University of Science & Engineering, Zigong 643000, China)

**Abstract:** In order to realize the university student party members' information management, after analyzing the college students' party member management work in present situation, framework based on B/S structure of information management system solutions is proposed. The .NET technology framework, ADO .NET database access technology and SQL database technology are used to rely on campus network implementation. The result shows that it is successfully realized the university student party members' information management.

**Key words:** .NET; college student party members; information management system

# 对一种基于口令和时间限制的用户身份认证协议的改进

叶俊, 付宇

(四川理工学院理学院, 四川 自贡 643000)

**摘要:**电子商务的应用越来越广泛, 高效且安全的电子商务协议研究是近期研究的热点。对 Luon - Chang Lin 等提出的基于口令和时间限制的用户身份认证的电子商务应用协议进行了分析。此协议的口令计算是基于模平方根难解问题的, 但是此阶段的方案存在漏洞并且运算效率低。对此提出了基于 Hash 链的改进方案, 并且对新方案进行了安全性分析, 此方案解决了原文献方案的漏洞并且比原文献中的方案更加高效和安全。

**关键词:**协议; hash 函数; 安全

**中图分类号:** TP309

**文献标识码:** A

## 引言

电子商务的应用越来越广泛, 如何保证电子商务的安全性和提高其效率是近年来研究的热点。在电子商务中, 身份认证<sup>[14]</sup>是一个很重要的环节, 在此方面已经有许多文献进行了研究, 许多基于口令的用户身份认证协议已经提出<sup>[5-9]</sup>。文献[7]给出了一种基于口令和时间限制的可计数的用户身份认证协议, 此协议基于模平方根的难解问题, 但是在协议的运算效率很低。因此, 本文提出了一种基于 Hash 链的改进方法, 从而提高了其效率。

## 1 Hash 链<sup>[10-12]</sup>简介

密码学中的 Hash 函数是一个压缩函数, 它将任意长度的消息压缩成固定长度的比特串。Hash 函数需要满足以下 3 个性质:

(1) 给定一个 hash 值  $h$ , 要找到一个  $x$  使得  $H(x) = h$  在计算上是不可行的。

(2) 给定一个  $x$ , 要找到一个  $y, y \neq x$  使得  $H(x) = H(y)$  在计算上是不可行的。

(3) 找到 2 个不同的消息  $x$  和  $y$  使得  $H(x) = H(y)$  在计算上是不可行的。

Hash 链是由一个公开的密码学 Hash 函数  $h$  进行递

归运算得到的。随机选取一个种子值  $s$ , 令  $x_0 = s$ , 在此基础上创建一条长为  $n + 1$  的 Hash 链:  $x_i = h(x_{i-1}), i = 1, 2, \dots, n$ 。

## 2 文献[7]用户口令方案重述

设  $j$  表示用户在允许的  $t$  次登陆机会中尚未登录次数,  $TK_j$  为表示第  $(t - j + 1)$  次登陆所需要的信息,  $TK_j: \{TID_j, ED_j, \alpha_j\}$ ,  $t$  为预先给定的允许用户  $U_i$  登录系统的次数, 系统计算

$$PW_i^{(j)} \equiv r_i^{2^j} \equiv (r_i^{2^{j-1}})^2 \pmod{n}$$

$$\alpha_j = MAC_p(ID_i \| TID_j \| ED_j \| PW_i^{(j-1)})$$

$$j = 1, 2, \dots, t$$

这里  $\|$  表示将两块数据连接到一起,  $TID_j$  表示  $TK_j$  的序列号,  $ED_j$  表示  $TK_j$  的过期时间,  $PW_i^{(j)}$  为用户第  $(t - j + 1)$  次登陆系统的口令,  $MAC_p(\cdot)$  表示生成消息认证码并且用  $p$  作为密钥用 DES 或者 FEAL 对消息认证码进行加密。

系统将  $ID_i, TK_1, TK_2, \dots, TK_t$  和  $PW_i^{(t)}$  发送给用户  $U_i$ 。

若用户  $U_i$  登录成功, 将使得第(7)步成立的  $R_a^{(j)}$  作为下一次用户登录的口令  $PW_i^{(a-1)}$  发送给用户。

### 3 对文献[7]中协议的分析

文献[7]中协议是建立在模平方根难解的基础之上的,并且是运用到电子商务中。但是由上述协议可知,在身份认证阶段的第(5)步,需要计算模平方根  $x^2 = PW_i^{(a)} \bmod n$ ,但是在实际运算中,计算模平方根的效率很低。并且模平方根可能出现循环的现象。例如,

$$7^2 = 18 \bmod 31, 18^2 = 14 \bmod 31$$

$$14^2 = 10 \bmod 31, 10^2 = 7 \bmod 31$$

而用户的下一次登陆口令为上一次口令的模平方根,这样口令可能出现循环,一旦出现循环用户就可以提前知道下一次的口令。系统存在一定的漏洞。

为了填补此漏洞并且提高运作的效率,本文提出了以下的基于 Hash 链的改进方案。

### 4 对文献[7]中协议的改进

#### 4.1 注册阶段

(1)对于用户  $U_i$ ,系统随机选择  $m$  组数据  $r_j = (x_j, y_j) (j = 1, 2, \dots, m)$ ,经过这  $m$  个点  $(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)$  构造  $m-1$  次插值多项式  $A(x) = a_0 + a_1x + \dots + a_{m-1}x^{m-1}$ ,使之满足  $y_j = A(x_j) (1 \leq j \leq m)$ ,则  $A(x)$  的系数为  $(a_0, a_1, \dots, a_{m-1})$ ; 然后令  $r_i = a_0 \| a_1 \| \dots \| a_{m-1}$ , ( $\|$ 表示连接符),这就是系统生成的随机数  $r_i$ ;

(2)计算用户口令

$$PW_i^{(j)} = h^j(r_i)$$

并且保存(其中  $h(\cdot)$  为一 Hash 函数,  $h^j$  表示连续用 Hash 函数作用  $j$  次);

(3)系统计算

$$\alpha_j = MAC_{r_i}(ID_i \| TID_j \| ED_j \| PW_i^{(j-1)})$$

$$j = 1, 2, \dots, t$$

(4)系统将  $ID_i, TK_1, TK_2, \dots, TK_t$  和  $PW_i^{(1)}$  发送给用户  $U_i$ , 其中  $TK_j$  包含信息  $\{TID_j, ED_j, \alpha_j\}$ 。

#### 4.2 登录阶段

(1)用户  $U_i$  从授时中心获取当前时间  $TS$ ;

(2)用户  $U_i$  计算  $\beta_a = MAC_{\alpha_a}(ID_i \| TS)$ ;

(3)  $U_i$  将  $ID_i, PW_i^{(a)}, TID_a, ED_a, TS$  和  $\beta_a$  以及登录请求发送给系统。

#### 4.3 身份认证阶段

(1)设系统收到用户的登录请求时间为  $TS'$ , 此时系统计算是否有  $TS' - TS > \Delta TS$ , 其中  $\Delta TS$  为系统预先设定的所允许的用户发送登录请求和系统接收到用户登录请求的时间差。若  $TS' - TS > \Delta TS$  成立,则系统拒绝用户请求;

(2)系统验证  $ID_i$  与  $TID_a$  是否匹配,若不匹配的拒绝请求;

(3)系统验证  $TID_a$  是否过期,若过期则拒绝请求;

(4)系统验证是否  $TS < ED_a$ , 若不成立,则拒绝服务;

(5)系统验证  $\beta_a = MAC_{\alpha_a}(ID_i \| TS)$  是否成立,若不成立,则拒绝服务;

(6)若(1)~(5)均验证通过则用户  $U_i$  登录成功,这时将  $h^{i-1}(r_i)$  作为下一次的口令  $PW_i^{(a-1)}$  发送给用户。

### 5 改进方案的安全性和效率分析

(1)新方案的对抗重复登录的攻击和防止修改  $TK_j$  的有效时间攻击强度与原方案的安全性一样。

(2)  $r_i$  是由系统产生  $m$  组随机数,然后将这  $m$  组随机数联合起来生成的一个新的随机数,攻击者想去猜测  $r_i$  是不可能的;

(3)  $PW_i^{(j)} = h^j(r_i)$ , 实际形成一 Hash 链,则每次的口令是不同的,且不知道  $r_i$  是得不到每次登录的口令的,由于 Hash 函数是一个单向函数,知道前一次的口令想得到下一次的口令,在计算上是不可行的;

(4)文献[7]中基于模平方根的难解问题,其运算速度慢,而改进的方案中使用的是 Hash 函数,计算速度明显比计算模平方根更快。方案的比较见表1。

表1 运算效率的比较

	文献[7]	本方案
模指数运算(次)	$n$	0
模平方根运算(次)	$n$	0
Hash 函数运算(次)	0	$n$

### 6 结论

在电子商务中安全和效率都是很重要的,要在安全的前提下提高效率。本文对文献[7]的协议进行分析,该协议是基于模平方根难解问题的。根据原协议的算法,虽然攻击者几乎不可能计算得到模平方根,但是其运算效率低。本文提出了基于 hash 函数的改进方案,并且对安全性进行了分析,可以看出新的方案比原文献中的方案更加高效和安全。

#### 参考文献:

- [1] Kyungah Shim. Efficient identity-based authenticated key agreement protocol based on Weil pairing[J]. IEEE Electronics Letters, 2003, 39(8): 653-654.
- [2] 肖文, 王移芝, 沈旭昆. 基于 ECC 的 PostgreSQL 口令认证的研究与改进[J]. 计算机工程与设计, 2009, 30

- (7):1603-1604.
- [3] 张利华,章丽萍,张有光,等.基于口令的远程身份认证及密钥协商协议[J].计算机应用,2009,29(4):924-927.
- [4] 朱月珍.基于身份认证的网上证券交易系统的安全性研究[J].计算机应用与软件,2009,6(4):269-270.
- [5] Lee W B,Wu C C,Tsaur W J.A novel deniable authentication protocol using generalized Elgamal signature scheme[J].Information Sciences2007,177(6):1376-1381.
- [6] 雷文,赵攀,张弘.一种动态口令认证方案的研究与改进[J].四川理工学院学报:自然科学版,2009,22(5):47-50.
- [7] Lin Iuon-Chang,Chang Chin-Chen.A countable and time-bound password-based user authentication scheme for the applications of electronic commerce[J].Information Sciences,2009,179(9):1269-1277.
- [8] Nam J,Lee Y,Kim S,et al.Security weakness in a three-party pairing-based protocol for password authenticated key exchange[J]. Information Sciences, 2007, 177 (6): 1364-1375.
- [9] 郑丽萍,易虹.一种PKI体系下的私钥安全存取方案[J].四川理工学院学报:自然科学版,2011,24(1):62-65.
- [10] 贺蕾,甘勇,李娜娜,等.一种基于逆hash链的RFID安全协议[J].计算机应用与软件,2009,26(2):87-88.
- [11] 李章林,卢桂章,辛运伟.基于Hash链的可扩展RFID验证协议[J].计算机工程,2008,34(2):173-175.
- [12] 孟健,杨阳.基于PayWord的自更新Hash链微支付协议[J].计算机工程,2009,35(3):63-65.

## Improvement of Time-bound and Password-based User Authentication Protocol

YE Jun, FU Yu

(School of Science, Sichuan University of Science & Engineering, Zigong 643000, China)

**Abstract:** Electronic commerce is widely used now, and the resent research focus on the efficiency and security of electronic commerce protocols. A countable and time-bound password-based user authentication scheme for the electronic commerce promoted by Iuon-Chang Lin and Chin-Chen Chang is discussed. User's password of the protocols is based on the hard problem modular square root. However, there's a bug in this scheme, and the efficiency of the scheme is low. A new scheme based upon Hash chain is advanced, and its security is discussed. The new scheme has solved the bug in the original scheme, and it is more efficient and safer.

**Key words:** protocols; hash chain; security

# IPTV 数字版权管理解决方案研究

蒋华龙<sup>1</sup>, 刘智广<sup>2</sup>

(1. 四川理工学院计算机学院, 四川 自贡 643000; 2. 成都三零凯天通信实业有限公司, 成都 610041)

**摘要:**随着三网融合步伐的加快,网络电视(IPTV)的普及已成为必然。如何保证 IPTV 在制作、分发、传播和收看各个环节的安全,建立有效可行的访问控制和版权保护措施是内容创作者和版权所有者普遍关注的焦点。在详细讨论了实现 IPTV 数字版权管理相关技术的基础上,给出了一种 IPTV 数字版权管理的解决方案,基于该方案设计并实现了 IPTV 数字版权管理的演示系统,演示结果表明该方案能对 IPTV 中音视频媒体的版权进行有效保护。

**关键词:** IPTV; DRM; 解决方案; 身份认证

**中图分类号:** TP309

**文献标识码:** A

## 引言

随着计算机技术、多媒体技术和通信技术的发展,特别是宽带网络技术的发展,以数字化和网络化为核心的新一代媒体传播技术—交互式网络电视(IPTV)已成为电视传媒的重要发展方向。IPTV 以宽带网作为基础设施,以家用电视机或计算机作为主要终端设备,集互联网、多媒体、通信等多种技术于一体,通过 IP 协议向家庭用户提供包括数字电视在内的多种交互式数字媒体服务的全新技术。IPTV 可以实现媒体内容提供商和消费者之间互动,能根据用户的选择提供广泛的多媒体服务功能,包括视频广播、视频点播、网页浏览、电子邮件以及娱乐、教育及电子商务等功能<sup>[1]</sup>。但是,在因特网上非法获取和传播数字多媒体内容也十分容易,直接损害了作品版权人的利益,严重影响了创作人的创作积极性,必将阻碍和制约 IPTV 的健康发展<sup>[2]</sup>。因此,如何保证 IPTV 在制作、分发、传播和收看各个环节的安全,建立有效可行的访问控制和版权保护措施是内容创作者和版权所有者普遍关注的焦点。只有解决了版权保护问题, IPTV 才可能真正实施,才能保证 IPTV 应用市场正常有序的发展。

本文在详细讨论 IPTV 数字版权管理相关技术的基础上给出一种 IPTV 数字版权管理的解决方案,基于该

方案设计并实现 IPTV 数字版权管理的演示系统,演示结果表明该系统能对 IPTV 中音视频媒体的版权进行有效保护。

## 1 DRM 技术概述

数字版权管理(Digital Rights Management, 简称 DRM)是一项涉及到技术、法律和商业各个层面的系统工程。它作为一种新型的数字内容保护技术,对数字产品分发、传输和使用等各个环节进行控制,使得数字产品只能被授权使用的人按照授权的方式在授权的期限内使用。DRM 实施的指导思想是对数字内容进行加密和增设附加使用规则来判断用户是否具有使用对应数字内容的授权或者权限<sup>[3]</sup>,为数字媒体的商业运作提供了一套完整的实现手段。DRM 技术的出现,使得版权提供者可采取更灵活的节目销售方式,能较容易地确保数字媒体内容被合法的使用,同时有效地保护知识产权<sup>[4]</sup>。

DRM 不仅仅指版权保护,同时也提供了数字媒体内容的传输、管理和发行等一套完整的解决方案,它包含数字版权信息使用、受版权保护的数字媒体内容的管理和分发。

在 IPTV DRM 系统的实现方案中,涉及的技术非常广泛,主要包括密码学技术、身份认证技术、数字签名技

收稿日期:2011-04-06

基金项目:核高基科技重大专项子课题(2009ZX01039-003-001-05)

作者简介:蒋华龙(1968-),男,四川广安人,讲师,主要从事计算机应用技术方面的研究。

术等<sup>[5]</sup>。

### 1.1 内容加密技术

密码技术是信息安全领域的主要技术之一,现有的数字内容保护多采用加密方法实现,即首先将多媒体数据文件加密成密文后发布,使得非法攻击者无法从密文中获取重要信息,从而达到版权保护和信息安全的目的。

目前的密钥体制从原理上可分为两大类,即对称密钥(私钥)体制和非对称密钥(公钥)体制。对称密钥体制与非对称密钥体制是两种不同的加密机制,它们解决不同的问题。对称密码算法速度极快,算法设计简单,并且对选择密文攻击不敏感。目前使用较多的对称加密算法包括:DES、3DES、AES、RC4、Blowfish等。非对称加密算法不要求通信双方事先传递密钥或有任何约定就能完成保密通信,并且密钥管理方便,可实现防止假冒和抵赖,因此,更适合网络通信中的保密通信要求。但非对称加密算法的加密和解密速度慢,因此只适用于对少量数据进行加密。目前使用较多的对称加密算法包括:RSA、ElGamal、ECC、Rabin等<sup>[6]</sup>。

在本方案设计中,使用对称加密算法来加密媒体内容,使用非对称加密算法来保护证书和其他敏感数据,如用来加密内容密钥等。

### 1.2 身份认证技术

身份认证技术是计算机网络中确认操作者身份的一种技术。如何保证以数字身份进行操作的操作者就是这个数字身份合法拥有者,在IPTV DRM系统中是非常重要的。利用公钥证书可以实现有效的身份认证、设备认证以及权限分配、管理和吊销等功能,对相关设备和存储媒质,如机顶盒、硬盘、PC机等分配公钥证书,用于在播放节目前进行身份认证<sup>[7]</sup>。

本方案设计中,通过X.509数字证书和PKI认证体系确立准入控制、验证双方身份、使用交互的不可抵赖性和可审计性,建立IPTV DRM系统中各方的信任体系。

### 1.3 数字签名技术

数字签名是通过密码算法对数据进行加、解密变换来实现的,用DES算法、RSA算法都可实现数字签名。数字签名能够实现对原始信息的鉴别与验证,保证信息的完整性、真实性、机密性和发送者对所发报文的不可抵赖性<sup>[8]</sup>。目前,应用广泛的数字签名方法主要有三种,即RSA签名、DSS签名和Hash签名。这三种算法可单独使用,也可综合在一起使用。

## 2 IPTV DRM 系统方案设计

设计中,所使用的流媒体文件格式为当前比较常用

的MP4文件格式。系统能对基于MP4视频格式的点播流进行加密和打包,并由内容发布服务器进行发布。用户接收到加密流后,不能直接解码收看,必须由认证服务器对用户进行身份认证和授权,然后从密码服务器获得解密密钥后,才能正常播放。

图1为IPTV DRM系统的总体框架,该系统主要包括六个功能模块:内容打包模块,内容分发模块,密钥管理模块,身份认证模块,终端解密模块和终端身份认证模块。根据系统的具体实现,不同的功能模块在系统中扮演不同的角色。各功能模块是逻辑上分离的,但不要各自对应于独立的物理节点(如服务器)。根据具体实现时的系统配置,不同的功能模块可以在相同或不同的物理节点上实现,不同的DRM系统具体实现可以由部分或全部的功能模块组成,这依赖于业务模型对DRM系统的具体要求和配置。

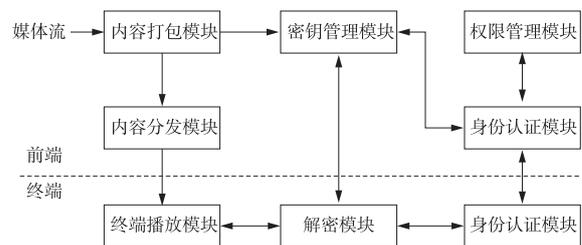


图1 IPTV DRM 系统总体架构

#### · 内容打包模块

主要负责对IPTV播控系统提供的媒体流进行打包和加密处理,并将处理结果发送到内容分发系统。

#### · 内容分发模块

主要负责监听终端用户的播放请求,对请求的音视频内容进行分发。

#### · 密钥管理模块

主要负责维护和管理整个IPTV DRM系统的密钥信息并向版权管理系统提供查询服务。

#### · 证书管理模块

主要负责证书的生成、发放和管理,对终端身份进行认证,提供对证书状态查询等功能,并与权限管理模块交互,通过查询接口获得用户对流媒体的访问权限,实现身份与权限的统一。

#### · 终端解密模块

主要负责利用解密密钥对经过加密处理的视频数据进行解密操作,将解密后的音视频数据送给终端播放模块进行播放。

#### · 身份认证模块

主要负责实现终端与认证系统服务端的双向身份认证,获得认证服务器分发的解密密钥,将获得的解密密钥提供给数据解密模块。

### 3 IPTV DRM 系统实现流程分析

#### 3.1 IPTV DRM 系统流程

IPTV DRM 系统流程如图 2 所示。具体流程描述如下。

第一步、内容提供者将要发布的媒体文件送到打包模块,对媒体文件使用对称加密技术加密和打包,根据一定的规则生成相应的内容 ID,将加密后的媒体文件送内容发布平台,采用非对称加密技术加密内容密钥并将其与内容 ID 一起存入密钥管理数据库。

第二步、用户通过点播平台点播想要收看的节目后,终端接收到来自内容发布服务器推送的节目数据。若节目数据被加密,则调用 DRM 系统的身份认证模块向身份认证服务器发起认证请求,实现认证服务器、终端的双向身份认证。

第三步、认证服务器向播控平台的用户权限管理服务器发起查询请求,确认特定用户对特定视频资源的访问权限。

第四步、认证成功且用户有播放权限,身份认证模块则向密钥管理服务器请求相应节目的解密密钥,终端获得密钥后则调用数据解密模块进行解密,将解密后的数据交给 IPTV 播控平台的视频播放器,从而实现对整个播放过程的有效保护。

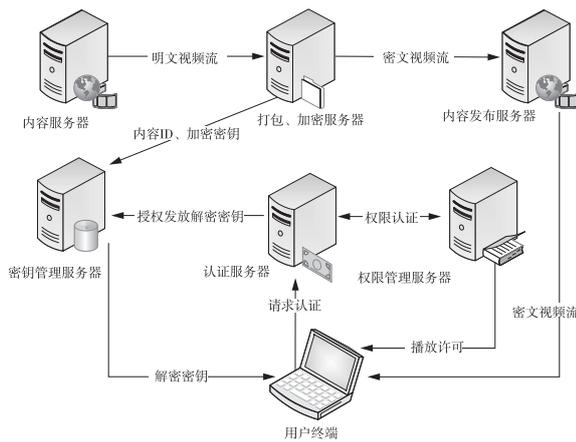


图 2 IPTV DRM 系统流程

#### 3.2 密钥管理流程

密钥管理系统负责产生、存储、分发和管理系统使用的各种密钥,并存储和管理用户证书。密钥管理系统维护节目与内容加密密钥的映射关系,并为节目的加密提供内容加密密钥。密钥管理系统不直接与用户终端系统交互,服务端通过授权管理系统发送权利对象为用户发布密钥。

当内容打包模块将生成的节目信息、内容 ID 信息和内容密钥发送给密钥管理系统后,用户通过业务系统

选定指定节目,请求权限管理系统获取对该节目的权限使用对象;权限管理系统向密钥管理系统查询指定节目的加密密钥;密钥管理系统根据节目的内容 ID 向数据库查询指定节目的加密密钥,并在响应消息中将加密密钥返回给业务系统。权限管理系统向业务系统返回权限使用对象,其中包括指定节目的加密密钥。

#### 3.3 节目发布流程

节目发布系统主要负责打包、加密媒体文件并进行流化发布,其中最重要的两个组成部分是内容发布系统和内容打包、加密系统。内容发布系统是 IPTV 播控系统的关键部分,它承担了将视频信息推送给终端的任务,本系统使用 Darwin Streaming Server 作为内容发布服务器。打包、加密系统主要负责对节目内容进行加密处理,它使用随机生成的密钥对媒体流进行加密,加密后的媒体流发送给内容发布系统,而加密密钥和节目信息将传送给密钥管理系统,由密钥管理系统负责管理该密钥。

#### 3.4 认证授权流程

内容密钥和权限对象的安全传输对 DRM 系统的实现至关重要,认证授权系统包括两个主要组成部分:身份认证系统和权限管理系统。身份认证系统主要负责证书的生成、发放和管理,对终端身份进行认证,提供对证书状态查询等功能,并与 IPTV 播控平台的权限管理系统交互,通过查询接口获得用户对流媒体的访问权限,实现身份与权限的统一。身份认证系统同时负责向密钥管理系统请求获取指定节目的内容密钥,并将该信息返回给已经通过身份认证的终端用户。

IPTV 终端用户在播放节目之前首先应向认证系统进行注册,使用其拥有的用户信息向身份认证系统请求获取 X.509 的数字证书,该证书是身份认证系统的 CA (Certificate Authority) 根据终端用户的用户名、密码、RSA 公钥等信息生成;身份认证系统收到 IPTV 终端的证书申请后,向业务系统进行查询以验证用户的身份,利用其终端身份信息生成对应的 X.509 数字证书,并以 PKCS12 存储结构的形式发送给 IPTV 终端并加以存储;当用户点播节目时,IPTV 终端向内容发布系统请求获取节目,收到请求的媒体流数据后,若数据被加密,终端则利用自己的数字证书向身份认证系统进行身份认证以获取视频数据的解密密钥;身份认证系统向权限管理系统请求获取该用户对节目的权限信息;权限管理系统向数据库查询指定用户拥有的权限信息并返回;身份认证系统查看用户拥有的权限信息,确认其拥有观看指定节目的权限后,向密钥管理系统请求获取指定节目的内容密钥;密钥管理系统响应内容密钥请求并返回,然后身份认证系统将权限信息和内容密钥以许可证的方式返

回给 IPTV 终端, IPTV 终端利用内容密钥对视频数据进行解密播放。

### 3.5 终端点播流程

终端点播是 IPTV DRM 系统整个流程的最终体现。用户点播节目时, 首先由 IPTV 终端向业务系统请求获取节目单; 业务系统收到 IPTV 终端的节目单请求后, 向终端发送指定节目的详细信息和节目所在的内容分发服务器的地址; IPTV 终端根据获得的分发服务器地址向内容发布系统请求获取节目数据; 内容发布系统向终端发送加密的节目视频流; 由于节目是加密的, IPTV 终端的播放器将无法播放, 此时终端将使用获得的数字证书向身份认证系统进行身份认证以获取节目数据的解密密钥; 身份认证系统收到 IPTV 终端的身份认证请求后, 返回包含权限信息和内容密钥的许可证书; IPTV 终端使用许可证书中包含的内容密钥对加密节目进行解密播放。

## 4 结束语

有效保护了 IPTV 节目的版权, 也就保护了节目创作者人和节目版权拥有者的利益, 必将有力促进 IPTV 产业的健康有序发展。本文通过对 DRM 技术的分析, 给出了 IPTV DRM 系统的设计方案, 详细阐述了 IPTV DRM 系统中各主要模块的实现流程。使用对称加密算法加密媒体内容, 使用非对称加密算法保护证书和加密内容密钥等敏感数据, 通过 X. 509 数字证书和 PKI 认证体系确立准入控制、验证双方身份、使用交互的不可抵赖性和可审计性, 建立 IPTV DRM 系统中各方的信任体

系。在此基础上, 设计并实现了 IPTV DRM 演示系统, 演示结果表明该系统能对 IPTV 中的音视频媒体的版权进行有效的保护, 对进一步研究和开发实用的 IPTV DRM 系统具有一定的参考价值。

### 参考文献:

- [1] Sohn D. Understanding DRM[J]. Queue, 2007, 5(7): 32-39.
- [2] 邝代英. DRM 数字权限管理系统的设计与实现[D]. 北京: 北京邮电大学, 2007.
- [3] 俞银燕, 汤帆. 数字版权保护技术研究综述[J]. 计算机学报, 2005, 28(12): 1957-1968.
- [4] 陈江涛. 关于数字版权管理(DRM)技术与应用[EB/OL]. <http://info.broadcast.hc360.com/HTML/001/002/017/67745.htm>, 2004.11.15.
- [5] Ivan I, Toma C, Popa M, et al. Secure platform for digital rights management distribution[J]. WSEAS Transactions on Computers, 2007, 6(3): 478-485.
- [6] Nash A, Duane W, Joseph C, et al. 张玉清等译. 公钥基础设施(PKI)实现和管理电子安全[M]. 北京: 清华大学出版社, 2002.
- [7] 裴庆祺, 高铭鼎, 范科峰. 数字电视领域的数字版权保护技术标准综述[J]. 信息技术与标准化, 2007(4): 33-37.
- [8] Li Jiguo, Huang Xinyi, Mu Yi, et al. Certificate-based signature: Security model and efficient construction[J]. Lecture Notes in Computer Science, 2007, 4582: 110-125.

## Research and Realization of IPTV Digital Copyright Management Solution

JIANG Hua-long<sup>1</sup>, LIU Zhi-guang<sup>2</sup>

(1. School of Computer Science, Sichuan University of Science & Engineering, Zigong 643000, China;

2. Chengdu 30 Kaitian Communication Industry Co. Ltd, Chengdu 610041, China)

**Abstract:** With quick pace of integration of broadcast, TV network, telecommunication network and internet, Internet Protocol television (IPTV) becomes popular. The content originator and copyright owner focus on how to ensure safety of production, distribution, transmission, and watch for IPTV and establish effective and feasible access control and copyright protection measures. Based on detailed discussion of realizing relevant technologies of IPTV digital copyright management, this paper gives a resolution of IPTV digital copyright management. Based on this solution, a presentation system of IPTV digital copyright management is designed and realized. The result of presentation shows that this solution can effectively protect copyright of IPTV's audio and video media.

**Key words:** IPTV; DRM; solution; identity authentication

# 改进的 Q 学习算法及其 RoboCup 中的应用

周燕艳<sup>1,2</sup>

(1. 合肥工业大学计算机与信息学院, 合肥 230001; 2. 铜陵学院数学与计算机系, 安徽 铜陵 244000)

**摘要:**传统的 Q 学习已被有效地应用于处理 RoboCup 中传球策略问题,但是它仅能简单地离散化连续的状态、动作空间。文章提出一种改进的 Q 学习算法,提出将神经网络应用于 Q 学习,系统只需学习部分状态—动作的 Q 值,即可进行 Q 学习,有效的提高收敛的速度。最后在 RoboCup 环境中验证这个算法,对传球成功率有所提高。

**关键词:**RoboCup;神经网络;Q 学习;智能体

**中图分类号:**TP18

**文献标识码:**A

## 引言

Robocup 就是机器人的世界杯足球锦标赛的简称。该比赛和计算机多个学科关联,主要与分布式处理、人工智能等领域有着非常重要的联系,同样对于相关的学科的发展也具有重要的影响。Robocup 目前拥有如下的比赛模式:用软件进行模拟比赛和用具体的实物的机器人进行足球比赛。这里,我们进行对仿真比赛系统予以研究。Robocup 比赛主要进行的是多 Agent 的对抗问题予以研究,模拟正式的足球比赛,在该比赛系统中,各 Agent 类似于一个比赛球员,实时获取比赛场上的各种参数,同时可以根据比赛实时的场景发出相应的动作由 Server 来执行,同时为了加大与真正的比赛的相似程度,加入了许多实时的噪声来提高比赛的不确定性。

RoboCup 使用的是 Client/Server 模式,Server 系统由 RoboCup 联合会来提供,各个参赛队的任务就是设计自己的 Client 模块。

在比赛中,通过 UDP/IP 协议来实现各个 Client 与统一的 Server 之间的数据通信,每个 Client 都可以向 Server 发送命令来表达自己在下个比赛周期里需要执行的动作,同时可以从 Server 端接受一系列的信息来获取

场上的实时状态,对每个球员来说,信息是相对有限和可能存在误差的。Server 在一个比赛周期中接受比赛双方队伍成员的命令,然后通过有效命令来对场上所有物体的位置和速度进行计算,并对有关信息进行发送,另外,Server 还承担着裁判的任务。和现实的足球比赛一样,如何赢取比赛,取决于很多因素,其中传球的准确度是一个非常关键的条件。本文就是从试图提高球员的传球准确度来进行研究。传统的 Q 学习方法已经被用来解决这方面的问题并取得了一些成功,但是该方法本身具有着不足,主要表现在输入的是离散的状态数据,结果输出也是离散的情况,而在 RoboCup 比赛中,场上各个对象的状态本身是连续的,动作又是离散的情况,使用传统的 Q 学习方法来处理,就会引起相关的信息的误差度高,在传球这一问题上会引起传球准确率的下降<sup>[1]</sup>。在借鉴传统的 Q 学习方法的基础上,本文引入了 BP 神经网络方法,将二者结合起来组成新的混合模型,该模型首先从实际的比赛场景中直接获取成功的数据,提供给 BP 神经网络系统进行学习,通过一定周期的学习,达到网络收敛的最优状态,再提供给 Q 学习模型进行处理,这样有效的解决了传统 Q 学习状态空间的离散而导致传球准确率相对不高的问题。

收稿日期:2011-03-11

作者简介:周燕艳(1978-),女,安徽枞阳人,讲师,硕士生,主要从事人工智能与图像处理方面的研究。

## 1 传统 Q 学习和 BP 神经网络简介

### 1.1 传统 Q 学习

Q 学习是强化学习(reinforcement learning)的一种方法。强化学习是一种重要的机器学习方法,它是从参数扰动自适应控制、动物学习等理论发展而产生而出现。该方法监督学习技术不同,不是使用正例和反例来决定需要执行的动作,而是使用称为试错(trial-and-error)的方法,来寻求最优行为策略<sup>[2,3]</sup>。对于环境模型,它的采取的方法不是去事先予以判断、估计,而是直接优化一个函数,该函数可以进行迭代处理,该函数被称为 Q 函数。它的定义为:在状态  $s_t$  时则执行动作  $a_t$ ,以后按最优动作序列执行时的折扣累计强化值来决定后续的动作。Q 函数的表达式表示如(1):

$$Q(s_t, a_t) = (1 - \alpha) Q(s_t, a_t) + \alpha(\gamma + V_{t+1}) \quad (1)$$

在这里,  $Q$  是在状态  $s_t$  下进行动作  $a_t$  的预期值;  $s_t$  是状态向量;  $a_t$  是动作向量;  $r$  是执行动作  $a_t$  获得的立即回报;  $\alpha$  是控制收敛的学习率;  $\gamma$  为折扣因子;  $V$  是在状态  $s_{t+1}$  下  $Q$  的最大值,即:

$$V = \max_{a \in A} Q(s_{t+1}, a) \quad (2)$$

智能体如果想要获取较大的强化值,那么在每个状态中,选择动作的依据必须是当就前拥有最高  $Q$  值的动作。但在任何状态学习的初始时段,对由于 Agent 拥有的经验相对较少,这时的  $Q$  值并不能表示的是正确的强化值。如果一味的只是选择最高  $Q$  值的动作,这样很有可能导致 Agent 总是沿着相同的路径计算,也就不可能获得理想的结果。所以,在 Agent 的选择动作时候,需要人为添加随机因素,最常用的方法是引入 Boltzmann 分布:

$$P(s_t, a) = \frac{\exp\left(\frac{Q(s_t, a)}{T}\right)}{\sum_{a \in A} \exp\left(\frac{Q(s_t, a)}{T}\right)}, a \in A \quad (3)$$

因此,调整后的 Q 学习算法运算步骤如下:

- (1) 初始化  $Q$  值;
- (2) 得到  $t$  时刻的状态  $s_t$ ;
- (3) 计算每个动作的  $Q(s_t, a_t)$ , 根据  $Q(s_t, a_t)$  选择一个动作  $a_t$  执行, 选择一个动作  $a_t$  执行, 计算获取新的状态  $s_{t+1}$  以及回报  $r$ ;
- (4) 利用公式(1)更新  $Q$  值, 转向(2), 直至  $Q$  表收敛。

$Q$  表为存储状态-动作对以及对应  $Q$  值的表格。

Q 学习是一个应用很广泛的强化学习算法,但它存在一定的问题。首先, Q 学习算法不能适用于连续状态空间和动作空间的学习,其次,当状态空间很大时,  $Q$  表在内存中的存储以及查询都变得十分困难。本文试图将神经网络和 Q 学习相结合来解决 Q 学习中存在的上述问题<sup>[4]</sup>。

### 1.2 BP 神经网络

BP(Back Propagation)网络是 1986 年由 Rumelhart 和 McClelland 为首的科学家小组提出,是一种按误差逆传播算法训练的多层前馈网络,是目前应用最广泛的神经网络模型之一。BP 网络能学习和存储大量的输入-输出模式映射关系,而无需事前揭示描述这种映射关系的数学方程。它的学习规则是使用最速下降法,通过反向传播来不断调整网络的权值和阈值,使网络的误差平方和最小。BP 神经网络模型拓扑结构包括输入层(input)、隐含层(hidden layer)和输出层(output layer)。BP 网络的特点是信号由输入层单向传输到输出层,同一层神经元之间不传递信息,每个神经元与邻层所有神经元相连,连接权重用  $\omega_{ij}$  表示,各神经元的作用函数为 Sigmoid 函数:  $f(x) = 1/(1 + e^{-x})$ 。同时它正向传播信号,方向传播误差。BP 网络隐含层节点个数选择到目前为止还没有确定的规则,根据经验公式  $\sqrt{m \times n} + 1 \sim 10$  ( $m, n$  表示输入输出的节点个数)<sup>[5,6]</sup>。

## 2 基于神经网络和 Q 学习的传球策略

### 2.1 RoboCup 中传球问题分析

Soccer Server 中,为了方便,用一个圆来表示来队员的身体和足球,并且规定,二者之间位置不允许有重叠的部分。当球和队员的距离(即两圆心的距离)小于某设定值(称之为传球半径)时,这个队员可以向 Server 发出一个 kick(即踢球)命令,该命令还包含两个参数(分别为角度和力量),如果该命令操作成功,对球而言就获取一个矢量加速度值,也则改变了球的运动状态。注意,对任一球员来说,只有当球处于在某个范围(称之为控球范围)内时, kick 命令才能有效<sup>[7]</sup>。

当本队的队员获得对足球的控制权后,这里称该队员为 Mate1, 此时, Mate1 可以选择带球(Dribble)、传球(Pass)和射门(Shoot)等动作,在这些动作中,考虑到动作应用的概率,其中传球(Pass)是最多的动作,因而,这

里将传球策略作为本文的研究重点。

不失一般性,下面分析传球的几种简单场景:

(1) 如图 1(a) 所示在一定范围里没有对方球员, 定义 Mate1 将要传球的队员为 Mate2, Mate1 和 Mate2 均用白色圆圈表示, 球用黑色小圆点表示;

(2) 如图 1(b) 或 1(c) 所示, 在 Mate1 和 Mate2 之间有了一个对方队员 Opp1, 黑色圆圈表示, 传球路线需要经过 Opp1;

(3) 如图 1(d) 所示, 在 Mate1 和 Mate2 之间有了两个对方队员 Opp1、Opp2, 传球路线需要经过 Opp1 和 Opp2 之间。

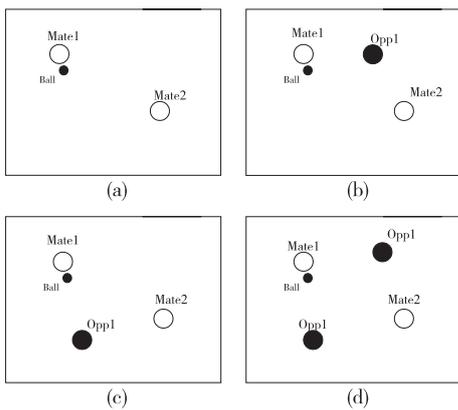


图 1 传球场景示意图

下面, 仅仅考虑最复杂的图 1(d) 图的情形, 而(a)、(b)和(c)均可认为是(d)的特殊情况。现在的问题变为在已知 Mate1、Mate2、Opp1 和 Opp2 的位置的前提下, Mate1 以什么样的速度和角度将球可以成功地传给 Mate2。为了减少参数的个数, 假设 Mate1 的位置不考虑, 仅仅考虑 Mate2、Opp1 和 Opp2 相对于 Mate1 的位置, 角度就为 Mate2 相对于 Mate1 的方向, 那么就可以只有三个位置参数输入, 一个速度值的输出。

### 2.2 基于 BP 神经网络模型

这时 BP 神经网络的节点为 3 个输入量, 1 个输出量 (即  $m=3, n=1$ ), 隐含层选择了 8 个节点, 如图 2 所示, 设网络输入为  $x_1, x_2, x_3$ , 输入为  $y$ 。

输入层各神经元的激发函数选用比例系数为 1 的线性函数, 则网络输入层的输出分别是  $x_1, x_2, x_3$ , 隐含神经元的输入为:

$$I_i = \sum_{j=1}^3 \omega_{ij} x_j \quad (i=1, 2, 3, \dots, 8) \quad (4)$$

神经元的输出为:

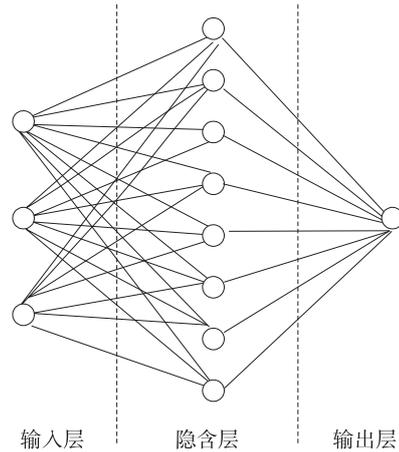


图 2 BP 神经网络示意图

$$O_i = 1 / (1 + e^{-I_i}) \quad (5)$$

$v_i$  为输出层神经元与隐含层神经元  $i$  的连接权, 则网络输出为:

$$y = \sum_{i=1}^8 v_i O_i \quad (6)$$

在由  $\omega_{ij}, v_i$  组成的连接权向量  $W$  初始化后, 就可以在给定的一组网络输入后, 由上述式子求出网络的输出  $y$ , 此为正向信号传播。

对于某样本  $(x_{1p}, x_{2p}, x_{3p}; t_p)$ ,  $p$  为样本数, 由正向计算得到  $y_p$ , 定义网络输出误差为:

$$d_p = t_p - y_p \quad (7)$$

误差函数为:

$$e_p = 1/2 * d_p^2 \quad (8)$$

在反向计算中, 沿着误差函数  $e_p$  随着  $W$  变化的负梯度方向对  $W$  进行修正值为  $\Delta W$ :

$$\Delta W = -\eta \frac{\partial e_p}{\partial W} \quad (9)$$

$\eta$  为学习率, 取 0 - 1 间的数 (通过大量试验, 取 0.56)。最后采用迭代式  $W + \Delta W \rightarrow W$  对原  $W$  进行修正, 得到新的连接权向量  $W$ 。

对于所有的学习样本, 均按照样本排列顺序进行计算, 从而求出学习样本的能量函数值:

$$E = \sum_{p=1}^P e_p \quad (10)$$

利用  $E$  值对网络计算精度进行评价, 当  $E$  值满足  $E < 0.00001$  时, 停止迭代, 否则, 进行新一轮的迭代计算。

训练时, 我们构造各种 Mate2、Opp1 和 Opp2 相对于 Mate1 的位置, Mate1 使用各种速度传球, 当 Mate2 成功地接到球, 就将该场景的相关数据记录下来。

### 2.3 结合神经网络和 Q 学习的传球学习过程

神经网络和 Q 学习结合的算法过程如下<sup>[8]</sup>:

- (1) 设定随机的场景中的参数  $X$ ;
- (2)  $Y = V_{BP}(X)$ ;
- (3)  $V_{t+1} = Y$ ;
- (4)  $Q(s_t, a_t) = (1 - \alpha) * Q(s_t, a_t) + \alpha * (\gamma + V_{t+1})$ ;
- (5)  $S_i = S_{i+1}$ , 转(2); 直至 Q 收敛。

这里  $X$  为状态—动作对  $(s, a)$ ,  $V_{BP}$  为神经网络输出,  $Q(s_i, a_i)$  为强化学习得到的 Q 值,  $V$  是在状态  $S_{t+1}$  下 Q 的最大值。Q 学习和 BP 神经网络予以结合, 目的是通过 BP 神经网络的强大存储能力以及函数估计能力。对于这一类的学习系统, 一般来说神经网络在其中的主要功能为: 首先主动设定外界环境的完全或不完整状态集, 把这些状态集提交给神经网络, 作为 BP 计算的输入, 使用神经网络进行计算, 然后把计算的结果作为 Q 学习系统所需的 Q 值或 V 值。

### 2.4 实验结果及分析

将上述算法应用到创建的训练传球场景中。在试验中表示 Q 值的神经网络为三层前馈神经网络, 隐藏层的单元数是 8 个, 网络的各个节点的权值在  $[-0.5, 0.5]$  范围内随机取值。设置更新 Q 值中的学习率  $\alpha = 10^{-5}$ 。图 3 显示的 5000 个周期的训练结果, 可以看到经典 Q 算法到 3500 的时候才达到最好的结果, 但是比引入 BP 神经网络的 Q 学习的效果要差。因为其使用 Q 表存储 Q 值的时候, 要将状态空间离散化, 导致丢失了大量的状态信息。

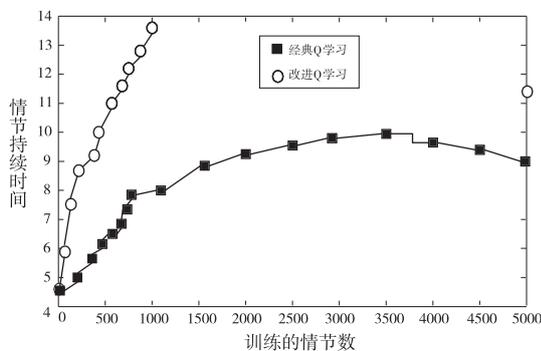


图 3 5000 个情节训练结果比较图

最后, 将该方法应用于传球训练的统计, 总共进行 1000 个训练周期, 每个训练周期为 5 个服务器周期, 对同一样本空间进行训练, 和经典 Q 学习方法相比实验结果统计如图 4 所示。

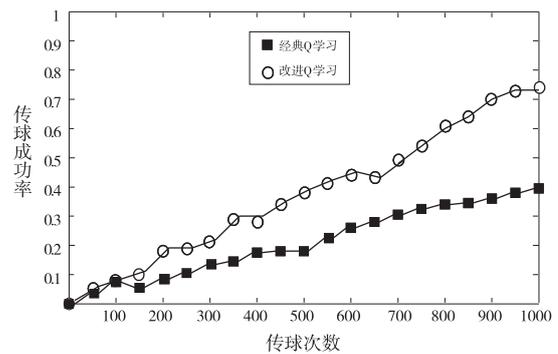


图 4 传球试验结果比较示意图

从仿真实验结果来看, 传球的成功率相比以往有了较大的提高, 基于神经网络的强化学习算法在一定程度上提高了传球的成功率。用神经网络来存储 Q 值函数与常见的 Q 表方法相比在时空性能上有了一定程度的提高。首先, 从时间性能上来看, 由于神经网络具有泛化特性, 只需学习部分状态—动作的 Q 值, 就可以获得其它近似状态—动作的 Q 值, 因此其学习速度要比 Q 表方法快。其次, 从空间性能上来看, 神经网络所需的存储空间只和网络的连接权个数有关, 而连接权的个数一般要远远小于状态的个数, 所以采用神经网络来存储 Q 值函数的方法所占用的存储也要小于后者。

### 3 结束语

提出将神经网络应用于 Q 学习, 有效地提高了系统泛化能力, 实现了 RoboCup 中底层传球策略的优化。实验结果说明了与神经网络结合的 Q 学习在一定的训练周期结束后收敛, 最终得到了优化的传球行为序列的值, 根据最后的实验结果分析可知, 取得了一定的预期目标。今后将对上层踢球策略以及改进神经网络的 Q 学习算法需要进一步研究。

### 参考文献:

- [1] Yan X W. Fuzzy Advantage Learning[J]. IEEE, 2000: 865-870.
- [2] Kaelbling L P, Littman M L, Moore A W. Reinforcement learning: A survey[J]. Journal of Artificial Intelligence, 1996, 4: 237-285.
- [3] Sutton R S, Barto A G. Reinforcement Learning[M]. Cambridge, MA: The MIT Press, 1998.
- [4] 高建清, 王浩, 于磊, 等. 一种模糊强化学习算法及其在 RoboCup 中的应用[J]. 计算机工程与应用, 2006

- (6):52-54.
- [5] 许盛中,蔡乐才.基于本体的图书智能检索系统的模型研究[J].四川理工学院学报:自然科学版,2009,23(5):55-58.
- [6] Philipp W K,Shie M,Doina P. Automatic basis function construction for approximate dynamic programming and reinforcement learning[C]//Proceedings of the 23rd International Conference on Machine Learning. Cambridge:MIT Press,2006:1103-1115.
- [7] 高春涛.一种新的混沌神经网络及其应用[J].哈尔滨理工大学学报,2010(2):35-38.
- [8] 周勇,刘锋.基于改进的 Q 学习的 RoboCup 传球策略研究[J].计算机技术与发展,2008,18(4):63-66.

## Improved Q-learning Algorithm and Its Application in RoboCup Environment

ZHOU Yan-yan<sup>1,2</sup>

(1. School of Computer and Information, Hefei University of Technology, Hefei 230001, China;

2. Department of Mathematics and Computer, Tongling College, Tongling 244000, China)

**Abstract:** Q-learning has traditionally been used effectively in dealing with RoboCup ball tactics, but it is only a simple discretization of continuous state and action space. Proposed a modified Q learning algorithm, neural network applied to Q learning, the system only need to learn some of the state-action Q value, you can get a continuous approximation of Q value, and can effectively improve generalization ability. Finally, in the RoboCup environment, the algorithm is proved to achieve optimal playing strategy, and effectively improves the success rate of passing ball.

**Key words:** RoboCup; neural network; Q learning; Agent

# JDBC 数据库连接池中事务处理的研究与实现

何爱华<sup>1</sup>, 戚晓明<sup>2</sup>

(1. 蚌埠学院计算机科学与技术系, 安徽 蚌埠 233030; 2. 中国矿业大学地理信息与遥感科学系, 江苏 徐州 221116)

**摘 要:** JDBC 驱动程序提供的连接类实现了事务控制机制, 为了实现多任务多用户的操作, 本文提出了一种禁用自动提交模式, 该模式把对事务的控制权转交给程序开发人员, 由程序员决定在适当的时候调用连接对象的适当方法来完成任务。

**关键词:** 连接池; 事务处理; ACID; 单事务; 嵌套事务

**中图分类号:** TP115

**文献标识码:** A

数据处理是数据库系统的最大功能, 为了保证数据的安全有效和正确可靠, 数据库管理系统必须提供统一的数据保护功能, 事务处理技术是提供数据保护功能的有效措施。现在流行的 JDBC 数据库连接池技术是一种高效的数据库连接技术, 它负责分配、管理和释放数据库连接, 允许应用程序重复使用一个现有的数据库连接, 这项技术能明显提高对数据库操作的性能, 事务处理是连接池技术的关键部分, 它保证了数据库操作过程中事务性单元内的所有操作都成功完成。

## 1 事务的 ACID 性质

事务是一组不同于一般程序的逻辑单元, 它使数据从一种状态转变成另外一种状态, 它的执行要受到某些特殊性质的约束。区分事务和一般程序的特性简称为 ACID 特性, ACID 是指原子性(Atomicity)、一致性(Consistency)、隔离性(Isolation)和持久性(Durability), ACID 特性描述了事务处理的规则。

### (1) 原子性

事务必须是原子工作单元, 当对数据库中的数据操作时, 必须确保事务一直执行完成, 或如果没有全都执行, 则该操作不会对数据产生任何影响。

### (2) 一致性

事务在独立执行完成时, 必须保持数据库一致状

态。数据库必须满足所有完整性约束, 设计者应该保证事务执行完成后, 数据库仍然满足所有的完整性约束。

### (3) 隔离性

并发执行将原子性复杂化且破坏了事务的一致性, 这就必须为并发执行指定相关的约束, 以保证一致性和原子性。事务的隔离性是这种约束的体现, 通常可以从抽象的数据库系统上采取隔离性, 如数据库中的每个数据项都有唯一的名称, 使用名称来进行读写操作; 也可以从关系数据库系统上采取隔离性, 如使用 SQL 语句访问数据。

### (4) 持久性

事务执行完成后, 其执行结果必须反映到数据库中, 它对于系统的影响是永久性的, 这种特性称为事务的持久性。即使系统发生 CPU 崩溃、磁盘错误等意外时, 已提交的数据仍能保留在系统中。

## 2 数据库连接池中的事务处理机制

使用数据库连接进行普通的数据库访问相对简单, 但对于事务处理, 操作起来则复杂很多。由于事务的原子性特征, 要求数据库的操作符合“All - All - Nothing”原则, 即要么全部执行, 要么什么都不做。

### 2.1 单事务处理

JDBC 中的 Connection 提供了对事务的支持。在

JDBC 的数据库操作中,一项事务由若干条表达式组成,它是一个不可分割的逻辑单元。打开一个 Connection 对象时,AutoCommit 属性缺省值为自动提交方式(auto-commit)。为了能将多个 SQL 语句组合成一个事务,可以通过设置 Connection 的 AutoCommit 属性值为 false 来禁止 JDBC 事务的自动提交方式,然后显式的调用 commit() 方法来提交 JDBC 事务,调用 rollback() 方法来结束事务。为了安全有效地进行 Connection 复用,每个事务独占一个连接,从而降低了事务处理的复杂度。由于隶属于该事务的所有数据库操作都是通过这一个连接完成的,独占连接的方式不会妨碍连接的复用,并且事务方法也复用了其他一些数据库的方法。下面以 JavaBean 中使用 JDBC 方式进行 delete 事务处理为举说明事务处理方式。

```
//userId 为用户标识,由用户所在的线程来标识
public int delete(int userId) {
    dbconn = new DataBaseConnection();
    Connection conn = dbc.getConnection();
    try {
        // 禁止自动提交方式
        conn.setAutoCommit(false);
        //数据库更新操作 1
        dbconn.executeUpdate("delete from bylaw where ID
        = " + userId);
        //数据库更新操作 2
        dbconn.executeUpdate("delete from bylaw _content
        where ID = " + userId);
        //提交 JDBC 事务
        conn.commit();
        //恢复 JDBC 事务的默认提交方式
        conn.setAutoCommit(true);
        dbconn.close();
        return 1;
    }
    catch(Exception e) {
        conn.rollback();//撤销 JDBC 事务
        e.printStackTrace();
        dbconn.close();
        return -1;
    }
}
```

以上程序段中两个数据库更新的操作要么全部执行,要么两者均不成功。

## 2.2 嵌套事务处理

### 2.2.1 嵌套事务

嵌套事务作为扩展事务模型的一种,应用较为广泛。嵌套事务将复杂的父事务分解为多个子事务,但仍然保留全局的隔离性和原子性。父事务可以依次创建子事务,当子事务全部执行完成后才能执行父事务。子事务集可以并发执行,当一个子事务执行完成后才能执行另外一个子事务。

在 JDBC 的事务处理机制中,嵌套事务的层次结构相对于外部来说是封装的,单个的子事务不能决定自身是提交还是异常终止,而是由父事务来决定是否提交。以下几点说明了嵌套事务的 ACID 特性。

(1)子事务是隔离执行的一个工作单元,子事务的提交和持久性取决于父事务的提交。当子事务所有的父事务提交后,该子事务才能提交并保持持久性,若一个父事务异常终止,该事务所有的子事务也终止,即使子事务已经提交了。如果一个子事务异常终止,它的执行对数据库的结果没有任何影响,但是可能会改变父事务的执行路径,对数据库的状态产生影响。

(2)嵌套事务保持整体的一致性,但子事务不一定保持一致性。

(3)每个嵌套事务必须是隔离的,所以相对于其他嵌套事务而言它可以串行化。子事务也具备这种特性。

### 2.2.2 基于 JDBC 的嵌套事务处理

JDBC 中处理嵌套事务的方法采用引用计数法,引用计数指的是嵌套的层次。以下代码说明了 JDBC 的嵌套事务处理的方式。

```
ParentClass {
    /* *
     * 事务属性配置为
     * PROPAGATION_REQUIRED
     */
    void invoke () {
        try {
            ClassA. invoke ();
        }
        catch(Bussiness. ClassA. Exception) {
            abort();
        }
    }
}
```

```

try {
ClassB. invoke ();
}
catch ( Bussiness. ClassB. Exception) {
ClassC. invoke ();
}
catch( Bussiness. ClassC. Exception) {
abort();
}
try {
ClassD. invoke ();
}
catch( Bussiness. ClassD. Exception) {
ClassE. invoke ();
}
catch( Bussiness. ClassE. Exception) {
abort();
}
}
}
}
ClassA {
/* *
* 事务属性配置为 PROPAGATION_NESTED
* 该事务需要被嵌套
*/
...
void methodA() {
}
}
/* *
* 以上子事务的方法分别为 methodB、methodC、
methodD、methodE
*/

```

以上程序段中出现了 A - B - D、A - B - E、A - C - D、A - C - E 的嵌套。通过这样的嵌套约定,可以满足数据完整性的需求,嵌套事务的约束如下:

(1) 如果子事务处于活动状态,则 Parent Class 事务只能提交事务、撤销事务或创建其它子事务。

(2) 如果子事务 ClassA 被提交,则 methodA() 的执行结果对于其他子事务来说是隐藏的,直到父事务提交为止,而 ParentClass 事务则能够检测到子事务做出的任

何动作。若子事务 ClassA 被撤销,则它不会对父事务有任何影响。

(3) 父事务提交或撤销决定子事务的提交和撤销。比如若 ClassA 提交,而其他子事务都无法成功执行,则 ClassA 也会撤销。JDBC SavePoint 提供了解决相应问题的技术。

(4) 子事务提交后,它对数据的加锁并没有释放,这些锁的控制权被转交给父事务,直到父事务提交或撤销后,锁才会被释放。

JDBC SavePoint(保存点)技术能够将已经提交的事务状态恢复到一个事务提交以前的任意保存点。假设以上代码没有执行到保存点,若 ClassA 执行成功而 ClassB 执行失败,在 JDBC 的 SavePoint 技术支持下,无需撤销 ChlidA 事务,在保留 methodA 结果的同时仅仅撤销 methodB,然后继续执行 methodD。实现伪代码如下:

```

Connection conn = dbc.getConnection();
Statement st = conn.createStatement();
st.executeUpdate("INSERT INTO TAB1 (COL1)
VALUES" + "'FIRST'");
//设置保存点
Savepoint savept1 = conn.setSavepoint("SAVEPOINT
_1");
stmt.executeUpdate("INSERT INTO TAB1 (COL1)
"+"VALUES ('SECOND')");
...
conn.rollback(savept1);
...
conn.commit();

```

### 3 结束语

事务是现代数据库理论中的核心概念之一,事务管理机制是 JDBC 规范的重要组成部分,禁止 JDBC 的自动提交事务方式后,就可以把多个数据库操作作为一个事务,这样做可以保持多次更新操作后,相关数据的一致性。应用系统中与数据库的所有交互都是一个事务的一部分,随着 JDBC 平台数据处理业务多方向发展,深入了解事务管理机制是优化应用设计的关键。

### 参考文献:

- [1] 魏海萍,KEVIN Mukhar.Java 数据库应用程序编程指南[M].北京:电子工业出版社,2002.

- [2] Philip M Lewis. Databases and Transaction Processing An Application-Oriented Approach[M]. 北京:机械工业出版社,2005.
- [3] 祖荫柏. 基于 Web 服务的分布式应用事务处理探讨[J]. 电脑与电信,2010(1):71-72.
- [4] 梁雄友. 基于分布式事务流的动态可串行调度算法[J]. 计算机工程与应用,2010,46(8):144-147,152.
- [5] 李贤华,彭真. 基于 JDBC-ODBC 的数据库访问方法的设计与实现[J]. 计算机与现代化,2008(5):28-31.
- [6] 邓勇涛. 关于 JDBC 事务操作的探讨[J]. 长江大学学报,2009,6(4):268-269.
- [7] 龚松杰. Web 服务事务处理模型研究[J]. 计算机工程,2007,33(10):283.

## Study and Implementation of the Transaction in the JDBC Database Connection Pool

HE Ai-hua<sup>1</sup>, QI Xiao-ming<sup>2</sup>

(1. Department of Computer Science and Technology, Bengbu College, Bengbu 233030, China;

2. Department of Remote Sensing and Geographic Information Science, China University of Mining and Technology, Xuzhou 221116, China)

**Abstract:** The connection class offers a transaction control mechanism in the JDBC drivers, and the JDBC drivers will automatically starts a transaction. By this way, the drivers will automatically end the transaction. It applies to any single user database, for many users of the database application, must change this automatic the way to achieve much more users. To solve the problem, a mode disables automatic submitted, the mode of the transaction control over to the application developer. Programmer in an appropriate time to call the connection object to the proper way to accomplish the task.

**Key words:** connection pooling ;transaction; ACID; single transaction; nested transaction

# 基于模拟退火遗传算法的 IIR 数字滤波器参数优化设计

谭 焜<sup>1</sup>, 刘自山<sup>2</sup>, 李凌宇<sup>3</sup>

(1. 北京地铁运营有限公司机电公司, 北京 100043; 2. 四川理工学院理学院, 四川 自贡 643000;  
3. 崇州市建筑勘察设计研究院, 四川 崇州 611230)

**摘 要:**结合模拟退火算法的思想和遗传算法的思想,提出模拟退火遗传算法,并用此算法进行滤波器参数整定与优化,同时使用自适应交叉率和变异率,以及适应度拉伸方法对传统遗传算法进行改进。该算法有效抑制早熟,又具有收敛性快、全局寻优与局部寻优能力。仿真结果表明,基于此算法寻优设计的滤波器控制器具有更好的滤波特性。

**关键词:**模拟退火遗传算法;IIR 数字滤波器;参数整定;MATLAB 仿真

**中图分类号:**TP391;TP273

**文献标识码:**A

## 引 言

遗传算法<sup>[1-3]</sup>简称 GA(*Genetic Algorithms*),它是根据自然界的遗传机理而设计的一种全新概念的优化方法,是一种建立在生物界进化规律(适者生存,优胜劣汰遗传机制)演化而来的随机搜索方法,它模拟了生物界中的生命进化机制,在人工系统中能够实现特定目标的优化。它是由美国 Michigan 大学的 J. Holland 教授在《*Adaptation in Natural and Artificial Systems*》一书中于 1975 年首先提出。该算法的特点主要是直接对结构对象进行操作,不存在求导和函数连续性的限定;具有内在的隐并行性和更好的全局寻优能力;采用概率化的寻优方法,能自动获取和指导优化的搜索空间,自适应地调整搜索方向,不需要确定的规则。已被人们广泛地应用于组合优化、机器学习、信号处理、自适应控制和人工生命等领域,尤其适应于处理传统方法难于解决的复杂和非线性问题,所以越来越成为人们解决高度复杂问题的一个新方法和新思路。

模拟退火算法<sup>[4]</sup>简称 SAA(*Simulated Annealing Algorithms*),由 Kirkpatrick 于 1982 年根据固体退火思想为解决大规模组合优化问题,特别是 NP 问题完全组合优化而设计的有效近似算法。它的设计思想是将固

体加热至充分高,然后让其徐徐冷却,加温时,固体内部粒子随温升变为无序状,内能增大,而徐徐冷却时粒子渐趋有序,在每个温度都达到平衡态,最后在常温时达到基态,内能减为最小,使之凝结成规整晶体的热力学过程,从统计物理学的观点看,随着温度的降低,物体的能量将逐渐趋于一个较低状态,并最后达到某种平衡,是基于蒙特卡罗迭代求解法的一种启发式随机搜索过程。

到目前为止, MATLAB 信号处理工具箱<sup>[5-6]</sup>提供了多种设计和实现 IIR 滤波器与 FIR 滤波器的方法。如设计 IIR 滤波器的变换法,方法比较简单,方便,但也存在一些不足,如设计过程繁琐不易掌握,设计结果往往并不理想,虽然设计参数已经表格化,但查表取得的数据不能保证最优的,表中不可能提供关于不同的抽样点数、通带边缘频率、阻带边缘频率、过渡样本数的全部数据。若表中查不到数据,也只能近似估计过渡样带本值,这样也不一定能保证全局最优。这些方法都有一定的局限性,而本文采用自适应模拟退火遗传算法对数字滤波器参数进行整定与优化,是一种先进的优化技术,弥补了以上方法的不足,设计出更好的数字滤波器。

## 1 算法思想与流程

模拟退火遗传算法的基本思想是:在寻优时先对参

收稿日期:2011-04-26

基金项目:四川省高等教育人才培养质量和教学改革项目(P09252);四川省教育厅自然科学基金重点项目(10ZA136);人工智能四川省(高校)重点实验室开放基金重点项目(2009RZ001);四川理工学院理学院重点科研项目(09LXYA03)

作者简介:谭焜(1974-),男,重庆垫江人,工程师,主要从事地铁机电系统集成自动化控制优化设计方面的研究。

数进行编码,给出初始化温度,按一定规模初始化一个种群,种群中的每一个体代表一个可能的解。然后根据适应度值函数,计算每一个体的适应度值并进行适应度拉伸,判断是否达到系统要求,如不满足减低温度,依此控制再生操作之后按自适应概率对种群进行交叉、变异操作。这样种群不断进化,直至寻优结束。自适应在线模拟退火遗传算法结合了模拟退火算法和遗传算法的思想,并在每个采样时间都对参数进行遗传寻优和对遗传算法参数中的交叉概率与变异概率进行自适应调整<sup>[7]</sup>,并对适应度进行拉伸,降低温度,直到达到平衡为止。在采样时间  $i$  的具体计算步骤为:

(1) 编码:编码的方法很多,二进制编码,浮点数编码,字符编码,变成编码等。而二进制编码是目前遗传算法中最常用的编码方法。其中本文采用的是浮点数编码。

(2) 初始群体的生成:初始群体是根据设置的群体大小而随机产生的。它作为进化的初始代,即第一代。

(3) 适应度评估检测:遗传算法在搜索时不需要外部的信息,只是根据适应度函数来评估个体的优劣,并作为以后操作的依据。在对适应度评价之前,要对适应度进行拉伸,其拉伸公式为:

$$f'_i = \frac{e^{f_i/T}}{\sum_{i=1}^M e^{f_i/T}} \quad (1)$$

$$T = T_0 (0.99)^{gen-1}$$

其中:  $f'_i$  为第  $i$  个个体的拉伸适应度,  $f_i$  为第  $i$  个个体的适应度,  $M$  为种群大小,  $gen$  为当前遗传代数,  $T$  为温度,  $T_0$  为初始温度。

设组合优化问题的一个解  $i$  及其目标函数分别与固体的微观状态  $i$  及其能量等价。随着算法进程递减其值的温度控制参数  $T$ , 控制参数  $T$  的每一个取值,算法持续进行“产生新解—判断—接受/舍弃”的迭代过程就对于固体在某一恒温下趋于热平衡的过程。从统计物理学获得的 Metropolis 接受准则应用于确定从当前解  $i$  到新解  $j$  转移的概率  $p_k$ :

$$p_k(i \Rightarrow j) = \begin{cases} 1 & f(i) \leq f(j) \\ \exp\left(\frac{f(i) - f(j)}{T}\right) & \text{否则} \end{cases} \quad (2)$$

开始时  $T$  的取值比较大,在进行足够多的状态转移后,缓慢减小  $T$  的值,如此反复,直到满足停止条件时算法终止。

(4) 选择与复制:从群体中选择优胜的个体,淘汰劣质个体的操作叫选择。选择算子有时又称为再生算子,选择操作是建立在群体中个体的适应度评估基础上的,目前常用的选择算子有以下几种:适应度比例方法、随机遍历抽样法、局部选择法。其中轮盘赌选择法

(roulette wheel selection)是最简单也是最常用的选择方法。在该方法中,各个个体的选择概率和其适应度值成比例。选择的目的是把优化的个体(或解)直接遗传到下一代或通过配对交叉产生新的个体再遗传到下一代。

(5) 交叉操作:在自然界生物进化过程中起核心作用的是生物遗传基因的重组(加上变异)。同样,遗传算法中起核心作用的是遗传操作的交叉算子。交叉模拟了生物进化过程中的繁殖现象,通过两个父代个体的部分结构加以替换重组而生成新的优良个体。交叉算子根据交叉率将种群中的两个个体随机地交换某些基因,能够产生新的基因组合,期望将有益基因组合在一起。根据编码表示方法的不同,可以有以下的算法:实值重组包括离散重组、中间重组、线性重组、扩展线性重组;二进制交叉包括单点交叉、多点交叉、均匀交叉、洗牌交叉、缩小代理交叉等。最常用的交叉算子为单点交叉。具体操作是:在个体串中随机设定一个交叉点,实行交叉时,该点前或后的两个个体的部分结构进行互换,并生成两个新个体。同时本文的交叉概率运用下列公式进行自适应调整:

$$p_{c1} = 0.99e^{-gen/\mu}; p_{c2} = 0.4e^{-gen/\mu}$$

$$p_c = \begin{cases} p_{c1} - \frac{(p_{c1} - p_{c2})(f' - f_{avg})}{f_{max} - f_{avg}} & f' \geq f_{avg} \\ p_{c1} & f' < f_{avg} \end{cases} \quad (3)$$

其中:  $\mu$  为控制概率变化参数(一般取 100 - 500),  $f_{max}$  为群体的最大适应值,  $f_{avg}$  为群体的平均适应值,  $f'$  为参与交叉的两个个体的较大适应值。

(6) 变异:变异算子是对群体中的个体串中的某些基因座上的基因值作变动。变异运算用来模拟生物在自然的遗传环境中由于各种偶然因素引起的基因突变,它以很小的概率随机地改变遗传基因(表示染色体的符号串的某一位)的值。也能产生新的个体。依据个体编码表示方法的不同,有实值变异和二进制变异等。本文的变异概率运用下列公式进行自适应调整:

$$p_{m1} = 0.3e^{-gen/\mu}; p_{m2} = 0.01e^{-gen/\mu}$$

$$p_m = \begin{cases} p_{m1} - \frac{(p_{m1} - p_{m2})(f_{max} - f)}{f_{max} - f_{avg}} & f \geq f_{avg} \\ p_{m1} & f < f_{avg} \end{cases} \quad (4)$$

其中:  $f$  为变异个体的适应值。

(7) 解码:在寻找优化结束后,对其结果进行解码,使之表示成要求的结果。

## 2 IIR 数字滤波器设计步骤

首先确定一种最优准则,如设计出的频率响应幅值  $|H(e^{j\omega})|$  与所要求的理想频率响应幅值  $|H_d(e^{j\omega})|$  的均方误差最小准则,或它们的最大误差最小准则等;然

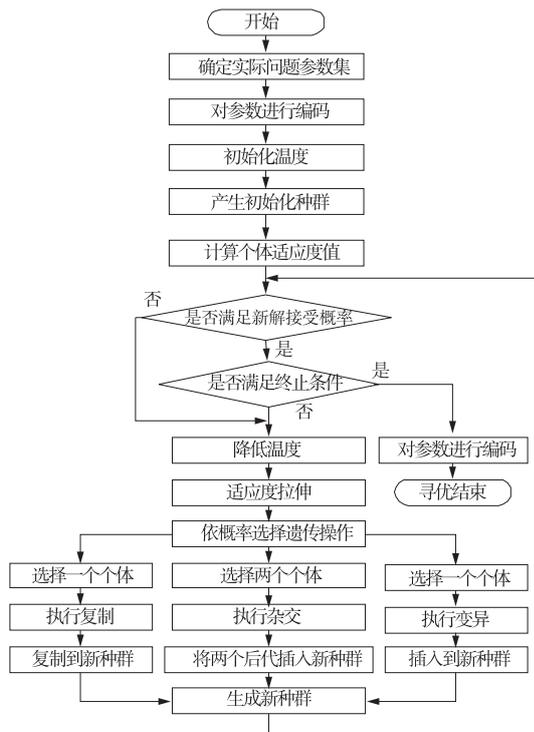


图1 模拟退火遗传算法流程图

后求在此最佳准则下滤波器系统函数的系数。根据均方误差最小准则,在一组离散点  $\omega_i (i = 1, 2, \dots, N)$  上所要求的频率响应  $|H_d(e^{j\omega})|$  的值为  $|H_d(e^{j\omega_i})|$ , 假定实际求出的频率响应为  $H(e^{j\omega})$ , 那么在给定的离散频率点处,所要求的频率响应的幅值与求出的实际频率响应的幅值的均方误差为:

$$E = \sum_{i=1}^N [ |H(e^{j\omega_i})| - |H_d(e^{j\omega_i})| ]^2$$

优化设计的目的是调整各  $H(e^{j\omega})$ , 即调整  $H(e^{j\omega})$  的系数,使  $E$  最小。这样即可以将得到的  $H(e^{j\omega})$  作为  $H_d(e^{j\omega})$  的逼近值。

实际滤波器  $H(e^{j\omega})$  常采用二阶节的级连形式表示,因为这种结构的频率响应对系数变化的灵敏度低,便于调整频率响应。设:

$$H(z) = A \prod_{k=1}^M \frac{1 + a_k z^{-1} + b_k z^{-2}}{1 + c_k z^{-1} + d_k z^{-2}} \quad (5)$$

则滤波器的频率响应为:

$$H(e^{j\omega}) = A \prod_{k=1}^M \frac{1 + a_k e^{-j\omega} + b_k e^{-j2\omega}}{1 + c_k e^{-j\omega} + d_k e^{-j2\omega}} = AP(e^{j\omega})$$

其中

$$P(e^{j\omega}) = \prod_{k=1}^M \frac{1 + a_k e^{-j\omega} + b_k e^{-j2\omega}}{1 + c_k e^{-j\omega} + d_k e^{-j2\omega}}$$

滤波器在各个频率采样点处的均方误差可表示为:

$$E = \sum_{i=1}^N [ |A| |P(e^{j\omega_i})| - |H_d(e^{j\omega_i})| ]^2$$

在式中共有  $(4k + 1)$  个未知参数,将  $A$  以外的  $4k$  个参数

表示成向量为:

$$\vec{\lambda} = (a_1, b_1, c_1, d_1, a_2, b_2, c_2, d_2, \dots, a_k, b_k, c_k, d_k)^T$$

最佳增益  $A$  可以用  $E$  对  $A$  求微分的方法求出,将  $E$  对  $A$  求微分并令其为零得:

$$\frac{\partial E}{\partial |A|} = \sum_{i=1}^N \{ 2 [ |A| |P(e^{j\omega_i})| - |H_d(e^{j\omega_i})| ] \cdot$$

$$|P(e^{j\omega_i})| \} = 0$$

于是可以得到最佳增益  $A$  :

$$|A| = \frac{\sum_{i=1}^N |P(e^{j\omega_i})| |H_d(e^{j\omega_i})|}{\sum_{i=1}^N |P(e^{j\omega_i})|^2}$$

由于只考虑滤波器的幅度响应误差,所以  $A$  的正负值对结果没有影响。为了保证所设计出的滤波器是稳定的,需要对优化参数的取值范围进行限定,同时也可以使参数的取值范围得以缩小以提高精度。由式(5)可知,只要使每个二阶节的极点都在  $Z$  平面的单位圆内,即使  $1 + c_k z^{-1} + d_k z^{-2} (k = 1, 2, \dots, M)$  的极点都满足条件  $|z_k| < 1$ , 就可得参数  $c_k, d_k$  的取值范围为:

$$-2 < c_k < 2, -1 < d_k < 1 (k = 1, 2, \dots, M) \quad (6)$$

如果将  $a_k, b_k$  的取值范围也限定在  $Z$  平面的单位圆内,则所设计出的滤波器是具有最小相位的滤波器。

### 3 设计实例与比较

设计一低通 IIR 滤波器,性能指标为:

$$|H_d(e^{j\omega})| = \begin{cases} 1 & 0 \leq \omega \leq 0.4\pi \\ 0 & 0.5\pi \leq \omega \leq \pi \end{cases}$$

选取滤波器的阶数为 6, 频率采样点取 46 个,待优化的参数共 12 个,为保证设计的滤波器是稳定的且为最小相位,参数的取值范围选取如式(6)所示。采用模拟退火遗传算法进行优化,在 MATLAB<sup>[6]</sup> 中编程并运行,所得到的 IIR 滤波器的传递函数为:

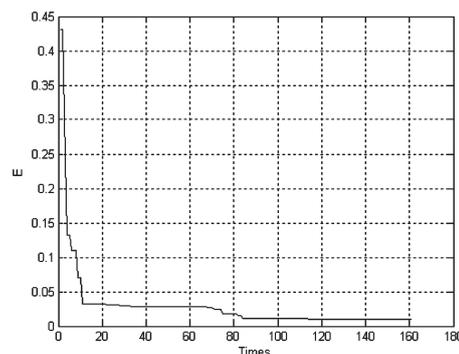


图2 误差寻优过程

$$H(z) = 0.0443 \times \frac{1 + 0.4124z^{-1} + 0.8972z^{-2}}{1 - 0.6539z^{-1} + 0.0762z^{-2}} \times$$

$$\frac{1 + 1.8557z^{-1} + 0.5007z^{-2}}{1 - 0.7967z^{-1} + 0.4891z^{-2}} \times \frac{1 + 0.0206z^{-1} + 0.1182z^{-2}}{1 - 0.5082z^{-1} + 0.8222z^{-2}}$$

其最小均方误差为:  $E = 0.0103$ 。

在设计 Butterworth 滤波器、Chebyshev I 型滤波器和椭圆滤波器时,通带频率不超过 3dB,阻带频率不低于

40 dB 时,运用 MATLAB 内部滤波器设计函数,需要对所设计的滤波器进行最优阶数的选择,再进行设计滤波器,所设计的 Butterworth 滤波器、Chebyshev I 型滤波器与椭圆滤波器的阶数分别为 15 阶、7 阶与 6 阶,其幅值响应曲线分别如下:

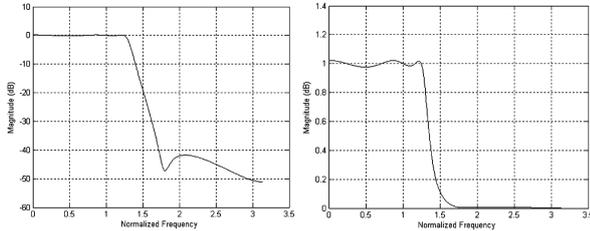


图 3 所设计低通 IIR 滤波器的幅值特性曲线

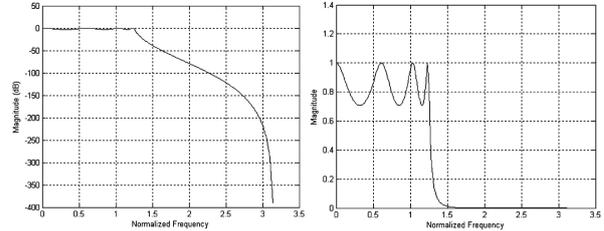


图 5 Chebyshev I 型滤波器幅值特性曲线

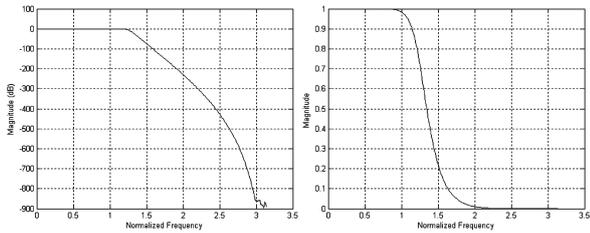


图 4 Butterworth 滤波器幅值特性曲线

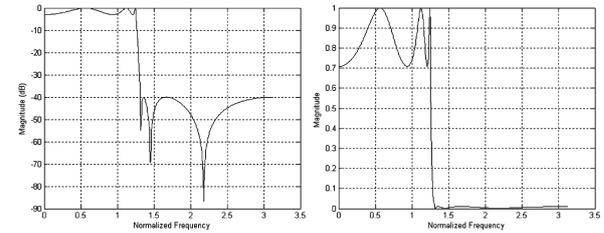


图 6 椭圆滤波器幅值特性曲线

Butterworth 滤波器的传递函数为:

$$H(z) = \frac{(0.113 + 1.692z^{-1} + 11.842z^{-2} + 51.315z^{-3} + 153.95z^{-4} + 338.68z^{-5} + 564.47z^{-6} + 725.74z^{-7} + \dots)}{1 - 2.9949z^{-1} + 6.0533z^{-2} - 8.317z^{-3} + 8.9085z^{-4} - 7.4744z^{-5} + 5.0897z^{-6} - 2.8082z^{-7} + \dots}$$

$$\dots + 725.74z^{-8} + 564.47z^{-9} + 338.68z^{-10} + 153.95z^{-11} + 51.315z^{-12} + 11.842z^{-13} + 1.692z^{-14} + 0.113z^{-15}) \times 10^{-4}$$

$$\dots + 1.2648z^{-8} - 0.4604z^{-9} + 0.1342z^{-10} - 0.0306z^{-11} + 0.0053z^{-12} - 0.0006z^{-13} + 0.00005z^{-14} - 0.000002z^{-15}$$

Chebyshev I 型滤波器的传递函数为:

$$H(z) = \frac{0.0005 + 0.0038z^{-1} + 0.0113z^{-2} + 0.0188z^{-3} + 0.0188z^{-4} + 0.0113z^{-5} + 0.0038z^{-6} + 0.0005z^{-7}}{1 - 3.7899z^{-1} + 7.7551z^{-2} - 10.289z^{-3} + 9.4209z^{-4} - 5.9209z^{-5} + 2.3852z^{-6} - 0.4905z^{-7}}$$

椭圆滤波器的传递函数为:

$$H(z) = \frac{0.0410 + 0.0169z^{-1} + 0.0933z^{-2} + 0.0393z^{-3} + 0.0933z^{-4} + 0.0169z^{-5} + 0.0410z^{-6}}{1 - 2.656z^{-1} + 4.6732z^{-2} - 5.0466z^{-3} + 3.8439z^{-4} - 1.8265z^{-5} + 0.4947z^{-6}}$$

从以上所设计的几种滤波器与模拟退火算法所设计的滤波器相比较,后者有更好的滤波特性。而前者中 Butterworth 滤波器能在通带内达到最大限度的平坦,在截止频率之后下降斜度迅速增加;Chebyshev I 型滤波器在截止频率处的下降斜度较小,在截止频率之后下降斜度迅速增加;椭圆滤波器在截止频率处的下降斜度较大。在这三种滤波器中,Chebyshev I 型滤波器在通带内为等波纹,而椭圆滤波器在通带和阻带内均为等波纹。在后者中在通带和阻带虽然也有波纹,但波纹较小。

### 4 结 论

由以上的分析可知,自适应模拟退火遗传算法利用自适应调整,对性能较差的个体采用较大的交叉率和变异率,而对于性能优良的个体则根据适应度的大小采用适当的交叉率和变异率。随着繁衍代数的增加,交叉率和变异率将随之下降,以利于算法的收敛。同时对适应度适当的拉伸,使得在温度较高时适应度相近的个体产生后代概率相近,而在温度不断下降后,拉伸作用增强,

使适应度相近的个体适应度差异放大,从而使得优秀的个体优势更加明显。在此基础上运用此算法所设计的滤波器有更好的滤波特性,也可以运用此算法设计其他类型的滤波器。但是,由于自适应模拟退火遗传算法所处理的数据量的增加,在适应调整中,温度下降比较慢,遗传代数增加,导致运行时间较长,若减少温度达到的精度,会导致结果结果的精度降低,这些是它的不足之处。

#### 参考文献:

- [1] 周明,孙树栋.遗传算法原理及应用[M].北京:国防工业出版社,1999.
- [2] 李敏强.遗传算法的基本理论与应用[M].北京:科学出版社,2002.
- [3] 李凌宇,郭贵法,许锦标.基于模拟退火遗传算法的PID参数整定与优化[J].广东工业大学学报,2010,27(2),80-83.
- [4] 王凌.智能优化算法及其应用[M].北京:清华大学出版社,2004.
- [5] 李娜.MATLAB仿真技术与应用[M].北京:清华大学出版社,2003.
- [6] 郭仕剑.MATLAB7.X数字信号处理[M].北京:人民邮电出版社,2006.
- [7] 巩敦卫,潘凤萍.自适应遗传算法理论及应用[M].徐州:中国矿业大学出版社,2003.

## Parameter Optimization of IIR Digital Filter Based on Adaptive Simulated Annealing Genetic Algorithms

TAN Xiao<sup>1</sup>, LIU Zi-shan<sup>2</sup>, LI Ling-yu<sup>3</sup>

(1. Beijing Metro Corporation Co. Ltd, Beijing 100043, China;

2. School of Science, Sichuan University of Science & Engineering, Zigong 643000, China;

3. Chongzhou Construction Survey and Design Institute, Chongzhou 611230, China)

**Abstract:** Combining the thoughts of Simulated Annealing Algorithms and Genetic Algorithms, an improved genetic algorithm (GA)-Simulated Annealing Genetic Algorithms is proposed to solve the problem of the parameter optimization of Digital Filter. The new GA has many improvements in algorithm like adaptive crossover and mutation operators, and scaling or stretching the adaptability. This algorithm avoids premature convergence and has a quick convergence performance. Additionally it has both the capability of optimization in the global area and the capability of optimization of the direct search methods in the local area. The simulation result shows that the Digital Filter designed upon the improved GA is of good filter characteristics.

**Key words:** Simulated Annealing Genetic Algorithms; IIR digital filter; optimization tuning; MATLAB simulation

# 三维地震数据体的切片播放算法

汪在荣<sup>1</sup>, 刘益和<sup>2</sup>

(内江师范学院计算机科学学院, 四川 内江 641110)

**摘要:**目前三维数据可视化技术已经广泛应用于地震解释中,而绝大多数的三维地震解释,是通过数据体的切片实行的。切片显示步骤是:确定切面多边形、对数据体的重采样和图像合成、切片的图像绘制。并且通常需要对切片实现播放、缩放等操作,此时,对于有限的内存采用传统的方法实现庞大的三维地震数据体的切片播放是相当的困难的。文章提出了一种将三维地震数据进行分块存储及显示的技术,即三维数据的砖块组织结构,并建立了快速索引机制,及砖块结构的调度算法。实验证明,该算法对三维地震数据切片的播放具有实时性及高效性。

**关键词:**三维数据可视化;切片播放;砖块结构;三维地震数据切片;调度算法

**中图分类号:**TP393.17

**文献标识码:**A

## 引言

三维地震数据的最主要特点之一就是数据量非常庞大,通常是 1GB 以上。面对如此庞大的三维数据,可以采用纹理贴图<sup>[1]</sup>和光线投射算法<sup>[2]</sup>等实现体绘制。但在绝大多数的三维地震解释中,专业人员希望能够更清晰地了解这些数据真正的构成形态,从而准确地判断油气藏的存储位置,提高生产效率,降低生产费用,这些都是通过对数据体的切片显示来完成。所谓切片就是以平行于三维地震体的三个方向 X、Y、Z 中某一方向的平面去切三维数据体得出的切面,它包括两个垂直剖面纵测线与横测线,和水平切面。三组正交切片显示的基本过程是:确定切面多边形、数据重采样与图像合成、切片的图像绘制<sup>[3]</sup>。在地震解释中用户要求对三维体数据沿 X、Y、Z 方向上进行相应的动态显示操作,即切片的播放。而但传统的三维地震数据可视化软件并没有提供切片的播放功能或是由于存在反复从外存中重采样数据使得切片的显示无法跟上播放速度。因此,本文基于实际项目开发过程中,提出了一种砖块组织结构及其快速的索引调度机制,并以此为基础,将原始三维地震数据转化为砖块结构文件,实现了三维地震数据体三组

正交切片任意速度的播放。

## 1 三维地震数据的砖块结构表示

对于海量的三维数据体在切片播放时要求切片的更新显示与速度同步,但由于播放速度使得在某一位置切片只要能够反映数据的分布情况即可,对其具体的精度要求并不高,但是当播放暂停时,此时显示的切片应该能准确反映在当前位置的数据图形。如果不采取相应的简化策略,直接对切片进行绘制,处理起来极为不便,严重影响运行效率和显示速度<sup>[4]</sup>。故要想实现实时显示与播放,必须进行数据简化(抽稀),简化后的数据能够描绘出原切片数据的分布趋势。但是,当切片播放暂停或停止时,只有重新调用未经采样前的数据进行切片绘制,才能保证图形的精度。因此,不失一般性,本文提出当切片正在播放时,则调用经过抽稀后的数据,这样提高了切片的绘制速度,当播放暂停或停止时,立即调用此切片位置的原始数据,这样则保证了图形的精度。本文在参考文献[5]的砖块结构算法上,进行内存优化和调度算法上的创新,从而更快速的实现三维地震数据体切片的播放。

对于海量的三维地震数据处理来说,显然不能将其

收稿日期:2011-05-27

基金项目:四川省教育厅自然科学重点科研项目(09ZA055)

作者简介:汪在荣(1975-),男,四川简阳人,讲师,硕士,主要从事通信网络及网络安全方面的研究。

全部调入内存中,只能在需要时才将所需的数据读入内存中进行处理。相对于CPU的速度来说,磁盘I/O是瓶颈之所在<sup>[6]</sup>,在切片播放或静态显示时,处理的数据只是一部分数据。

结合上述需求和三维地震数据的特点,本文在参考文献[5]提出的基础上,并结合I/O传输问题,对三维地震数据进行分块,这样既降低了地震数据体的操作难度又简化和提高了绘制效率。

三维地震数据的分块大小通常取2的幂次方,本文中经过实践分析,参见表1各2的幂次分块进行数据传输的耗时对比,从表中可以看出,对于256级的灰度图像,以64\*64\*64(256K)为最佳,按照这样的大小进行三维数据的分块,每次硬盘传送一块的数据耗时非常短。

表1 数据分块

分块大小 (KB)	传送数据大小为4000KB的I/O次数	冗余数据大小 (KB)
16	250	0
32	125	0
64	63	32
128	32	32
256	16	160
512	8	416
1024	4	928

从表1中可以看出,就分块而言,并无标准可循,可以按任意规则进行数据分块,但在应用中则必须结合索引、磁盘I/O、效率等进行考虑。不规则的分块将影响索引构建(索引用于查找定位数据)、索引操作及图形重构的效率。除此之外,分块太大或太小都将影响系统的有效性能,如果分块数据过大,则可能导致读入的冗余数据过多;反之,如果分块数据过小,则导致频繁的磁盘寻址和读写操作,使得所需数据的I/O访问时间增加<sup>[8]</sup>。

在介绍砖块组织结构前,先了解一下三维地震数据场。它是由横测线(xLine)、纵测线(inLine)和时间(time)确定的坐标系统。而(xLine, inLine)则构成了地震数据的道。如图1所示,显示了三维地震数据场和以块划分后砖块编号顺序。

根据前面的介绍,将原始三维地震数据转化为砖块结构文件来实现切片的播放,转换后的数据分为索引文件、采样文件以及砖块文件三部分。图2为砖块文件结构示意图。

索引文件由砖块数据整体描述信息块和道记录描述信息两部分组成。其中砖块数据整体描述信息块的主要作用是对砖块文件进行整体描述,它记录了原始数据中time、xLine、inLine三个方向的砖块数,三个方向标

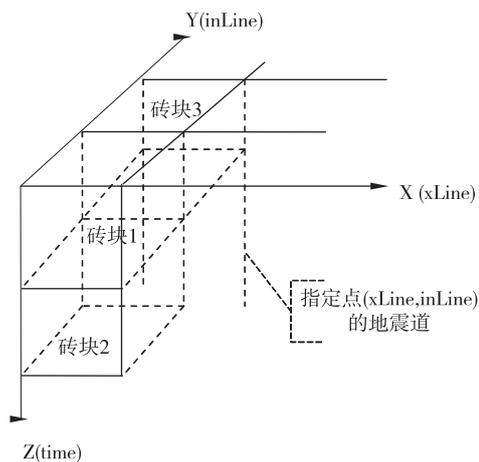


图1 三维数据场按块划分及砖块编号顺序

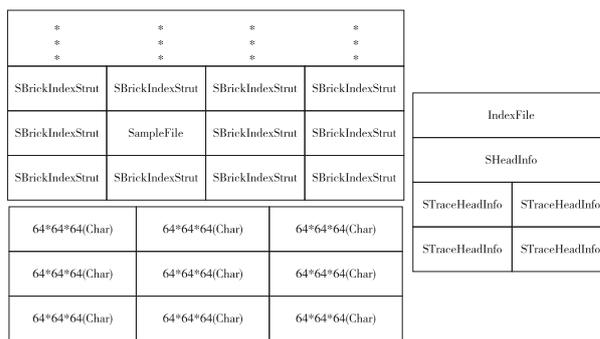


图2 砖块文件结构示意图

号的最值,速度(振幅)最值,大地坐标最值,本文中使用的结构体进行存储。索引文件的另一部分是原始数据中每一道记录的描述信息,包括每道的xLine号、inLine号及该道对应的大地坐标,它也是由相应的结构体进行存放管理,然后再利用一个三维数组(按照Time、Inline、xLine的顺序进行组织)存储所有道的数据结构。索引文件为砖块文件和采样文件数据提取提供参数,其主要作用是描述砖块结构。

砖块文件是砖块结构的主要部分,它存储了完整的绘图数据。砖块文件中存放的是依次排列的大小为64\*64\*64字节的小砖块,这些小砖块按照time、inLine、xLine顺序进行组织存放。考虑到不同原始数据的速度(振幅)数据类型可能不相同,占用的存储空间也不同,而且有的数据类型占用的存储空间较大(如double类型),因此砖块文件将原始数据值统一转化为Char类型的一个颜色值,由于颜色变化是用户关心的焦点,所以进行上述类型的转化不仅大大减小了存储空间,而且还不会影响绘图效果。

采样文件是根据采样间隔从原始数据中提取的数据,是原始数据经过抽稀后的文件,它存储了部分原始

绘图数据,虽没有原始数据详细,但通过它能够对原始数据进行定性和定量分析,能够反映原始数据的整体情况。砖块文件存储的是原始数据的所有数据点,数据量较大,不可能一次性调入内存进行处理,将数据存放在硬盘上,又受硬盘传输速率的限制,不可能实时的调入内存,也就不能进行实时动态播放。采样文件能够反映原始数据的整体情况且数据量很小,可以一次性调入内存,因此能够实时动态播放。采样文件由经过抽稀后的部分原始数据(实际中可利用一个三维数组进行存储)和每一个小砖块文件在整体砖块文件中的偏移量(实际中可利用结构体进行存储)两部分组成。

以上三个文件是在将原始三维地震数据转换为砖块文件时生成。对于 xLine 方向的砖块数为 maxXLine/64, inLine 方向的砖块数为 maxInLine/64, time 方向是由每道的采样点来确定的,故该方向砖块数为 sampleCount/64,因此,总砖块数为:

$$brickCount = \frac{\maxInLine}{64} \times \frac{\maxXLine}{64} \times \frac{\text{sampleCount}}{64} \quad (1)$$

而砖块的编号是按 time、inLine、xLine 方向顺序编号,并以索引文件结构体描述形式将其它数据填入索引文件中,最后生成的索引文件是 .index 格式。下一步是生成砖块文件(brick 格式),砖块文件中存放的是按照砖块编号将原始三维数据的颜色值顺序写入,与此同时,生成采样文件(sample 格式)并写该文件,以纵测线(inLine)数为外重循环,以横测线(xLine)为第二重循环,以 time 方向的每道采样点数(sampleCount)为内层循环,读取每一个原始数据点的颜色值写入 brick 文件中,同时,判断该点是否被采样,若是,则同时将该点写入 sample 文件中,直至所有的数据点全部转换结束。brick 文件和 sample 文件都是以划分后的砖块为单位进行组织的。

## 2 砖块结构的 FIFO 调度

至此,已经阐述了将原始三维地震数据的砖块结构表示。由于 sample 文件和 brick 文件存放的分别是采样后的砖块数据和原始砖块数据,这些数据都存放在磁盘上,在切片播放或停止时,需要确定当前切片数据在砖块文件或采样文件中的具体位置和大小,然后将其调入内存中。因此,砖块结构的调度是实现切片显示或播放的一个重要环节,也是本文研究的另一个重点。

无论切片的播放还是切片的显示都要求能够获取更快的速度,特别是实时显示。而磁盘数据的读取速度较内存数据要慢,为了加快数据的处理速度,减少对磁

盘的 I/O 操作,可以将当前需要的数据存放在事先分配的内存中,每次只处理内存中的数据。

本文根据系统内存的大小,对转换为砖块结构后的数据设置相应的高速缓存,高速缓存也按块分配,即将整个内存块划分为大小相同的单元,每块的大小与 brick 文件或 sample 文件大小相同,当缓冲区数据需要更新时,每次读取分块后的文件中的一块数据。这里 brick 文件是用于切片停止播放时显示准确的信息,而 sample 文件是用于切片播放时简化后的数据,两者用于不同的目的,但两者都是以砖块组织的,所以它们的调度方式是一样的。本文以 brick 文件的调度为例,描述缓冲区的调度方法。为了更好的进行数据调度,采用建立内存块索引的方法进行数据的调度管理。在缓冲数据更新时读取数据采用 FIFO 算法。

文中采用一个辅助数据结构来存放整个三维体数据的所有砖块的状态和信息。结构体表示为:

```
typedef struct _tagBrickInfo
{
    unsigned char status; //0:表示常驻内存,1:硬盘
    unsigned int index; //砖块在内存索引表中的索引位置
} SBrickInfo;
```

该辅助数据结构是用数组来存放,因为数组是一种简单的数据结构,而且其访问速度也最快。砖块的调度步骤如下:

- (1) 建立三维体数据的辅助数据结构;
- (2) 建立一个内存块索引表和开辟砖块数据缓冲区;
- (3) 计算切片所需要的砖块号,并结合 .index 文件取得砖块在 .brick 文件中的偏移位置;
- (4) 当某一砖块或某几个砖块数据需要被调入内存中时,修改辅助数据结构相应砖块号的 status 标志为 0,表示该砖块数据已经被调入内存,同时按 FIFO 调度算法计算出该砖块在内存块索引表中的索引位置(即查找一个内存缓冲区),若该索引位置为空,则将在该索引位置中写入存放的砖块号以及指向该砖块数据的指针,并将该数据读入内存中;若该索引位置已经有砖块数据,则修改该索引位置的砖块号所在的辅助数据结构信息(即将原先存放的砖块号的 SBrickInfo 结构体信息分别置 status 为 1, index 为初始值 MAX),再将内存索引表中该索引位置的 index 设置为需要被替换的新砖块号,同时将砖块数据指针指向的原来砖块数据进行释放,让该指针指向新的砖块数据。

图 3 反映了辅助数据结构与内存索引表的关系。

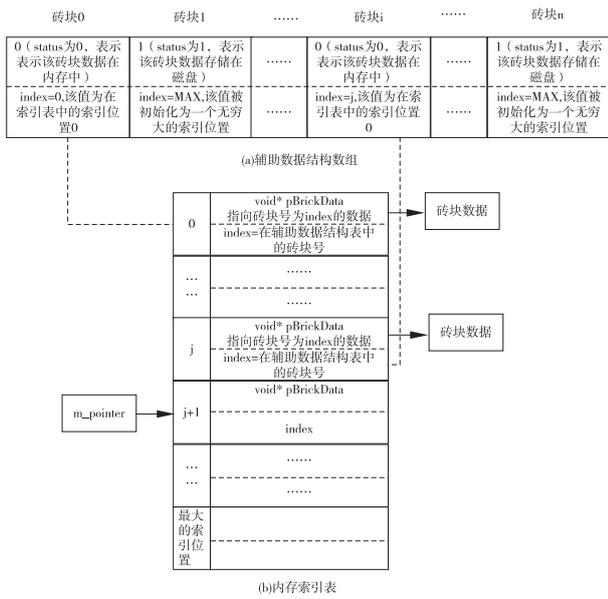
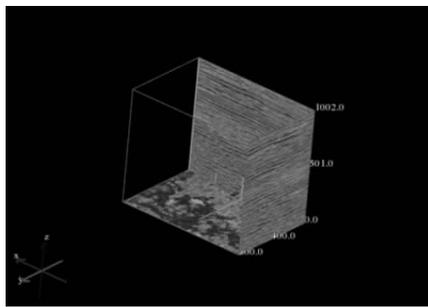


图3 辅助数据结构与内存索引表的关系

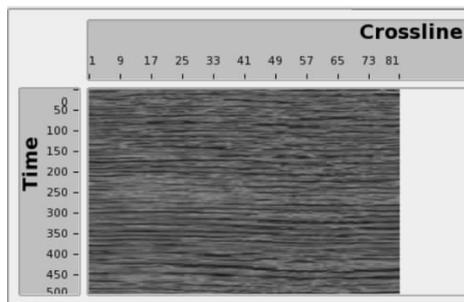
### 3 切片播放算法

利用前面介绍的砖块结构和砖块结构的调度, 本文通过以下3个步骤进行三维地震数据体的切片播放或显示: (1) 利用砖块结构对相应的数据结构进行初始化; (2) 根据当前图形状态获取切片图形数据。如是播放暂停或停止状态, 则计算出当前位置各切片相应的砖块号, 并根据砖块索引判断是否需要将该数据调入内存; 如是播放状态, 则计算出当前播放位置各个切片在采样文件中的位置, 并读取相应的采样文件中的数据。(3) 根据相应的数据运用一定的图形图像原理进行图形绘制。

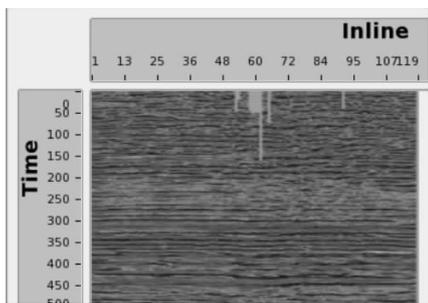
该方法运用于实际的地震数据处理软件开发中, 图4显示了使用该方法进行切片显示与播放的图形。图(a)中显示的是使用砖块结构算法进行切片播放时的三维立体图, 此时三个方向的切片分别是 inline 号为1的切片(图(b)所示), crossline 号为1的切片(图(c)所示), 采样时间 time 为0的切片(图(d)所示)。



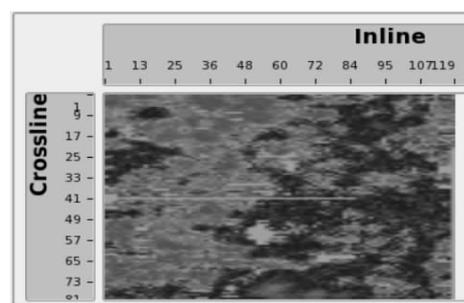
(a) (inline, crossline, time)的三维图



(b) inline=1的切片图



(c) crossline=1的切片图



(d) time=0ms的切片图

图4 切片的显示与三个方向的切片播放

### 4 结束语

采用以上的切片播放和调度算法, 使得在处理庞大的三维地震数据量时不受计算机内存容量的限制, 并且数据量的大小也不影响处理和显示的速度, 大大提高了海量三维地震数据在图形显示和切片播放时的速度, 可实现三维地震数据切片的实时播放。

### 参考文献:

- [1] 夏冰心. 基于纹理映射的三维地震数据可视化研究[D]. 南京理工大学, 2008.
- [2] 张二华, 高林, 马仁安, 等. 三维地震数据可视化原理及方法[J]. CT理论与应用研究, 2007, 16(3): 20-28.
- [3] Xue Daqing, Roger Crawfis. Efficient Splatting Using Modern Graphics Hardware[J]. Proceedings of the Computer Graphics, Imaging and Vision: New Trends, 2005, 8(3): 1-21.
- [4] David Laur, Hanrahan P. Hierarchical Footprint method: A

- Progressive Refinement Algorithm for Volume Rendering [C].New York,NY,USA:ACM SIGGRAPH Computer Graphics,1991:285-288.
- [5] 夏凡.地震数据处理系统中三维可视化的研究与实现[D].成都电子科技大学,2010.
- [6] Peter Lindstrom,Valerio Pascucci.Visualization of Large Terrains Made Easy[C].In:Proceedings of the IEEE Visualization'01,2001,363-371.
- [7] Dachsbacher C, Stamminger M. Rendering Procedural Terrain by Geometry Image Warping[C].In:Proceedings of the Eurographics'04,2004,103-110.
- [8] Hua Yi-Xin.Theory and Technology of Geographics Information System[M].Beijing:PLA Press,2001:1-2.
- [9] 钟晓霞.大范围复杂场景的简化与漫游技术研究[D].南京理工大学,2003.
- [10] 张伟.显微图像拼接系统设计与实现[D].北京邮电大学,2009.

## Play of Slices Algorithm of 3-D Seismic Data Volume

WANG Zai-rong<sup>1</sup>, LIU Yi-he<sup>2</sup>

(School of Computer Science, Neijiang Normal University, Neijiang 641110, China)

**Abstract:** At present, three-dimensional data visualization technology has been widely applied in seismic interpretation, however most of 3D seismic data is interpreted through the slices of the volume. The traditional display steps of slices were as follows: determine the section polygon, re-sample data and compound the image, rendering the slice images. Also, it needs to achieve operations with these slices such as playing, zooming, etc. Simultaneously, it is quite difficult for the limited memory to realize the play of slices of huge 3D seismic data. Therefore, a technology is proposed which stored and display 3D seismic data in blocks. It means the brick structure of 3D seismic data. The mechanism for fast indexing scheduling algorithm of brick structure is established. Experiments show that the algorithm is real-time playback and efficiency for three-dimensional slice of seismic data.

**Key words:** three-dimensional data visualization; play of slices; brick structure; seismic data volume; scheduling algorithm