

华北电力大学

硕士学位论文

基于 SVM 的智能家居辅助决策系统研究

Research on Assistant Decision System of Smart Home Based on SVM

刘晨

2017 年 3 月

国内图书分类号：TP391.4
国际图书分类号：004.9

学校代码：10079
密级：公开

硕士学位论文

基于 SVM 的智能家居辅助决策系统研究

硕 士 研 究 生： 刘晨
导 师： 袁和金 副教授
申 请 学 位： 工学硕士
学 科： 计算机科学与技术
专 业： 计算机应用技术
所 在 学 院： 控制与计算机工程学院
答 辩 日 期： 2017 年 3 月
授 予 学 位 单 位： 华北电力大学

Classified Index: TP391.4
U.D.C: 004.9

Thesis for the Master Degree

**Research on Assistant Decision System of Smart
Home Based on SVM**

Candidate:	Liu Chen
Supervisor:	Prof. Yuan Hejin
Academic Degree Applied for:	Master of Engineering
Speciality:	Computer Application Technology
School:	School of Control and Computer Engineering
Date of Defence:	March, 2017
Degree-Conferring-Institution:	North China Electric Power University

华北电力大学硕士学位论文原创性声明

本人郑重声明：此处所提交的硕士学位论文《基于 SVM 的智能家居辅助决策系统研究》，是本人在导师指导下，在华北电力大学攻读硕士学位期间独立进行研究工作所取得的成果。据本人所知，论文中除已注明部分外不包含他人已发表或撰写过的研究成果。对本文的研究工作做出重要贡献的个人和集体，均已在文中以明确方式注明。本声明的法律结果将完全由本人承担。

作者签名：

日期： 年 月 日

华北电力大学硕士学位论文使用授权书

《基于 SVM 的智能家居辅助决策系统研究》系本人在华北电力大学攻读硕士学位期间在导师指导下完成的硕士学位论文。本论文的研究成果归华北电力大学所有，本论文的研究内容不得以其它单位的名义发表。本人完全了解华北电力大学关于保存、使用学位论文的规定，同意学校保留并向国家有关部门或机构送交论文的复印件和电子版本，同意学校将学位论文的全部或部分内容编入有关数据库进行检索，允许论文被查阅和借阅。本人授权华北电力大学，可以采用影印、缩印或扫描等复制手段保存、可以公布论文的全部或部分内容。

本学位论文属于（请在以上相应方框内打“√”）：

保密□，在 年解密后适用本授权书

不保密□

作者签名：

日期： 年 月 日

导师签名：

日期： 年 月 日

摘要

随着生活水平的提高及物联网、机器学习等技术的发展，如何使家居具备自主“学习”能力，根据用户个人习惯自主控制照明灯、窗帘等家居的开闭成为当前智能家居研究的热点之一，对智能家居的发展及推广具有重要意义。本文对智能家居辅助决策中的异常数据处理、获取用户决策规则集、用户类别判定及优化决策等问题进行了深入的研究。

本文主要工作如下：

1) 分析和研究了智能家居辅助决策中的环境信息特点、决策方法、评估标准等，给出了一种支持向量机和 K 近邻算法相结合的智能家居辅助决策模型。该模型构建主要包括个人环境信息建模、异常数据处理、相似用户查找及基于支持向量机的用户决策规则集获取等。通过实验，与一般决策系统、基于 BP 神经网络的决策系统进行决策准确度的对比，结果表明，该方法具有耗时少、精度高的特点。

2) 给出了一种基于自适应步长和增强随机扰动的改进萤火虫算法，用于优化决策系统，通过引入亮度值和欧几里得距离公式改进了原算法中的位置更新公式，并利用 matlab 中的测试函数对该算法进行了测试，实验证明该算法在避免收敛早熟及无法收敛问题上有一定的优势。

3) 设计并实现了该智能家居辅助决策原型系统，研究内容主要包括智能家居系统总体架构设计、个人环境信息的采集及用于查看和控制家居状态的智能家居 APP 的设计与开发。

关键词：智能家居；辅助决策；支持向量机；K 近邻算法；萤火虫算法

Abstract

With the improvement of living standards, Internet of Things, development of machine learning and other technology, how to make smart home with independent "learning" ability, self-control of lighting, curtains and other home opening and closing according to users' own habits, is becoming one of hot spots in the current smart home research, which has great significance of the smart home's development and promotion. In this paper, we study deep processing of the abnormal data, set of the user decision rule, decision of the user classification and the decision's optimization in the assistant decision-making system of intelligent home.

The main work of this paper is as follows:

1) Analyze and study the characteristics of environment information, decision-making method and evaluation standard in assistant decision-making system, and give a decision model based on support vector machine and K-nearest neighbor algorithm. This model mainly includes personal environment information modeling, exception data processing, similar user search and user decision rule set acquisition based on SVM. Compared with general decision system and the decision system based on BP neural network in the experiment, the result shows that this method has the characteristics of less time consuming and high precision.

2) An improved firefly algorithm based on adaptive step size and enhanced stochastic perturbation is proposed to optimize the decision system. By introducing the luminance value and the Euclidean distance formula, the position updating formula in the original algorithm is improved. The test function in matlab has tested this algorithm, which proves that the algorithm has some advantages in avoiding the problem of premature of convergence and unable to convergence.

3) Design and implement of the assistant decision-making prototype system of intelligent home. The research contents include the overall architecture design of smart home system, the collection of personal environmental information and the design and development of smart home APP for viewing and controlling the home.

Keywords: intelligent home; decision support; support vector machine; K-nearest neighbor algorithm; firefly algorithm

目录

摘要	I
Abstract	I
第 1 章 绪论	1
1.1 研究背景及意义	1
1.2 研究现状	2
1.2.1 国外研究现状	2
1.2.2 国内研究现状	4
1.3 主要研究内容及创新点	5
1.4 本文组织结构	5
第 2 章 智能家居辅助决策相关技术	7
2.1 智能家居辅助决策系统	7
2.1.1 智能家居	7
2.1.2 辅助决策系统	8
2.2 决策分类技术	9
2.2.1 BP 神经网络	9
2.2.2 支持向量机	10
2.2.3 决策树算法	11
2.2.4 K 近邻分类算法	12
2.3 支持向量机参数优化算法	13
2.3.1 网格搜索算法	14
2.3.2 遗传算法	15
2.3.3 粒子群算法	16
2.4 本章小结	17
第 3 章 基于 SVM 的辅助决策模型	18
3.1 辅助决策系统总体架构	18
3.2 个人环境信息建模	21
3.3 异常数据处理	22
3.4 相似用户查找算法	25
3.5 SVM 分类模型构建	26
3.6 本章小结	28
第 4 章 辅助决策模型优化	29
4.1 群智能与萤火虫算法	29
4.1.1 进化计算和群智能	29
4.1.2 人工萤火虫算法	30
4.1.3 萤火虫算法	32
4.2 改进萤火虫算法及其在决策优化中的应用	34
4.2.1 改进萤火虫算法	34

4.2.2 算法流程.....	35
4.2.3 改进萤火虫算法测试.....	36
4.3 混合核函数及其在决策优化中的应用.....	39
4.4 本章小结.....	42
第5章 辅助决策系统设计与实现.....	43
5.1 家居环境信息采集.....	43
5.2 硬件系统介绍.....	44
5.3 软件系统介绍.....	45
5.4 辅助决策系统准确度验证.....	48
5.5 本章小结.....	50
第6章 总结与展望.....	51
6.1 总结.....	51
6.2 展望.....	52
参考文献.....	53
攻读硕士学位期间发表的论文及其他成果.....	56
致谢.....	57

第 1 章 绪论

1.1 研究背景及意义

当今社会，随着机器学习、数据库等传统技术的发展与创新以及诸如大数据、云计算、无线传感网络^[1]等新兴技术的出现与应用，物联网技术再次进入快速发展时期，其研究成果更是层出不穷，且得到了人们的广泛认可与应用，因此又被人们称为第三次科技浪潮。物联网技术可将物理世界中的任何物品连入互联网中，借助于某些信息传感设备，最终实现物物互联，实现人与物、物与物的信息交换和通信，提高人类对物理世界的感知及交互能力^[2]。随着射频识别技术（RFID）、纳米技术、智能嵌入技术^[3]及传感器技术得到越来越广泛的应用，在 2009 年，圆桌会议上有人首次提出“智慧地球”的概念，提倡建设新一代的智慧型基础设施。随着越来越多的设备被连入互联网中，如监视运动的智能手环、智能建筑中的温湿度传感器等，导致数据量大、数据信息复杂的情况时常产生，人们开始将大数据、云计算、数据挖掘等技术应用于物联网中。

智能家居^[4]作为物联网技术的一个典型应用，通过无线传感器感知环境信息，并经过数据预处理，特征提取、信息决策等一系列过程来产生具体的操作命令，并最终反馈给智能电器或者执行器（如智能开关、智能窗帘、智能电视等）进行相应操作。其起源于上世纪 80 年代初，发展至今已有近 40 年的历史，其按照组成及技术分类共经历了四代的发展，其中，市面上普遍安装的大部分为第三代智能家居，其借助于物联网技术，将家居中的各种设备连接到一起，为用户提高综合、高效的家居服务，包括灯光等远程控制、环境监测、安防等。而最新的第四代家居定义为基于全 IP 技术，为用户提供“云”服务的同时，满足用户实现定制化、个性化的需求，正因其重点为关注用户个性化，体现家居环境的因人而异，其日益成为当今科学家研究的重点^[5]。

随着我国经济的发展，人们的生活水平越来越高，仅仅是实现远程控制等功能的智能家居渐渐不能满足人们的需求。且室内家居的日益增多，若均要求人们常常查看并控制所有设备的运转变得不太现实，也无端耗费了人们的大量精力。人们希望在实现个人家居环境“个性化”的同时，智能家居能像人一样具有自主的正确决策能力，如室内温度高时，系统自动打开风扇或空调进行降温等，将人从繁杂的家居控制中解放出来。近年来机器学习得到了广泛的发展及

其在解决分类、模式识别以及回归分析问题方面具有的明显优势，人们开始尝试将云计算^[6]、支持向量机（SVM）^[7]、神经网络^[8]、普适计算^[9]等算法应用于智能家居的信息决策中，还针对决策算法的选取、如何提高决策精确度等问题进行了一系列的研究，并取得了一定的成果。基于机器学习等方法^[10]的智能家居辅助决策系统具有自主“学习”与自动控制能力，可有效改善居住环境，使人们生活更加便捷化、舒适化、智能化。其对室内异常灾害的实时监测与提前预防、采取控制措施也有助于预防室内异常情况的发生，对推广和应用智能家居具有重要意义。但部分算法存在训练时间久、决策准确率低等缺点。

本文主要针对智能家居如何“学习”用户习惯实现自动控制的问题进行了研究，提出了一种基于支持向量机的智能家居辅助决策模型，“学习”某用户一段时间内的行为，可实现满足该用户需求的自主控制，即对窗户、窗帘、照明灯开闭的控制，且可为不同用户提供不同的决策规则。其成果解决了不同用户具有不同环境要求的问题，真正做到了家居环境的因人而异，提高了智能家居的智能化、个性化。

1.2 研究现状

关于智能家居的控制方法主要分为三种：一是全部由人操作，即用户通过移动终端等设备远程控制家居的状态；二是人与家居共同参与家居管理，即为家居的各环境信息设置阈值，通过对家居的实时监测和阈值大小调节家居状态；三是人与辅助决策系统共同参与家居管理，将机器学习、云计算等技术引入到家居决策中，家居通过“学习”用户一段时间的行为习惯，得到该用户不同环境下的行为习惯，利用学习经验操控家居状态。前两种方法实现简捷，但或智能化不足、或个性化不足、或操控家居占用用户太多精力，不能满足用户对家居日益增长的需求，渐渐失去市场竞争力。第三种方法真正体现出了智能家居的个性化、智能化，渐渐吸引了国内外大部分学者的研究目光。

1.2.1 国外研究现状

世界上的第一幢智能建筑于 1984 年首先在美国出现，随后，欧洲的部分发达国家先后提出了一系列的智能家居解决方案，都有了一定的应用实例，如：微软自主研发的“梦幻之家”等产品都日趋成熟并迅速抢占智能家居市场。此外，市面上出现的智能家居控制系统有美国的 X-10 系统、新加坡的 8X 系统等。其

中新加坡当地政府推广了一系列的智能城市规划项目,预期在 2025 年完成智慧国家计划和理解部分,也就是对系统加入机器学习。

英国阿尔斯特大学的 Michael 等人基于对大部分用户家居环境数据的统计分析,为各信息设置合适阈值,如温湿度、光照等,来控制家居状态,部分减少了人为对家居的干预,但家居个性化不足^[11]。此外,英国伯明翰大学的 Suyang Zhou 等人研究分析了实时家用能源的控制方法。包括热水器、空调的控制,提出一种滚动优化与模糊逻辑控制相结合的能量控制方法,具有较强的实用性。上述方法实现较为简单,但在运行速率及反应时间上存在较大差距,影响用户体验^[12]。因此,为增加系统实时性,有学者分别提出了基于网络物理系统(CPS)^[13]、Arduino Uno 微控制器^[14]的智能家居系统,可实时感知传感器信息并对设备做出决策,主要使用三个执行器(空调、窗和窗帘)监视和保持室温,提高了感测精度,使得设备具有了计算、自治等功能,但无法实现不同人对环境的不同要求,体现用户的个性化需求。

为满足用户对家居的个性化需求,国内外学者对其实现方法进行了一系列研究,并取得了一定成果。Jan Vanus 等人^[15]对室内温度和相对湿度进行统计分析,提出了一种基于决策树回归法的二氧化碳精度预测方法,并利用 BACnet 技术组件控制智能家居中的通风、制冷等,其成果应用于服务系统优化、减少建筑物运行成本等。Partha 等人^[16]提出了一种基于凸二次规划建模控制室温的方法,使每个空调通过线性回归模型学习传感器中数据,实现受控区域温度的自动化控制,该系统可应用于大型环境中,可一次接受多个输入,提高了室内温控的灵活性和效率。文献[17]针对智能家居如何学习居民的使用模式和偏好问题进行了研究,提出了一种基于机器学习的智能家居事件的数据挖掘算法,用于在满足用户需求的前提下尽可能的减少家居能源消耗,为家居控制自动提供建议,虽然有用的建议率仅达到了 10%,但其具有运算快、存储器消耗小的特点。文献[18]通过研究机器学习如何为给定的家庭建立模型和自动特征选择方法,提出了一种包装方法和遗传算法相结合的时间序列数据生成方法,来为每个家庭的环境信息进行单独设置。

国外关于智能家居决策方法的研究很多,且为满足人们日益增长的需求,方法渐渐由简单变为复杂,发展至今,将机器学习应用于智能家居决策中,使得智能家居拥有“自主学习”和自动调节能力,丰富和扩宽了家居决策的方法和思路。

1.2.2 国内研究现状

我国目前智能家居市场潜力巨大，起步虽晚，但发展较快。到 2015 年，我国在最新的“互联网+”行动计划中明确提出要发展大数据与物联网，此外，政府为推动信息化，大力发展宽带普及等，这些为物联网行业、智能家居的发展奠定了坚实的基础。国内各大厂商也纷纷涉足智能家居领域，腾讯、小米、海尔等大型企业均纷纷涉足物联网行业，智能家居经过一段时间各产业磨合，已处于爆发前夜。

智能家居的研究方向很多，目前研究较多的有智能决策及室内行为识别等。针对实现家居的智能决策问题，国内学者尝试将机器学习算法引入到辅助决策中，有人利用 BP 神经网络具有良好的非线性映射和自适应的特点，使用 BP 神经网络建立辅助决策系统，并结合 KNN 算法实现不同用户的分组。文献[19]提出了一种基于 SNS 智能家居系统用户意图推理研究的方案，即根据用户在 SNS 社交网络上发布的内容、环境、天气等信息，利用爬虫系统、TF-IDF 算法、贝叶斯算法等进行用户控制家居设备的意图推测。文献[20]针对环境信息复杂特性，提出了基于 SVM 数据挖掘算法的辅助决策模型，在一定程度上提高了 CPS 智能照明系统智能化。在研究领域，有使用神经网络进行室内火灾预警^[21]的先例。文献[22]针对室内用电浪费问题，提出了一套基于多信息融合的室内智能控制系统方案，选择人工神经网络作为系统的智能决策核心，并结合极限学习机对分类效率和学习速度进行了改善。此外，针对智能家居数据量大、信息复杂、扩展性差等特点，人们尝试将云计算技术^{[23][24]}应用于智能家居系统中来改善系统处理及存储数据的能力，并实现了基于云计算的智能家居管理系统提供的一项服务-智能家居花卉管理，通过为花卉的温湿度提前设定阈值并实时监控来满足花卉的基本环境需求，但本方法也并未实现系统的“自主学习”性能，有待于后续的进一步研究。

目前国内对智能家居决策方法的研究虽然很多，但多为基于统计分析等阈值设定的简单方法或某些家居控制系统，较少人将机器学习方法应用于其中，本文将 SVM 分类方法与智能家居决策相结合，在对室内温度、湿度、光照上的自动控制进行了一定的探索。

1.3 主要研究内容及创新点

本文针对如何实现室内温度、湿度、光照自动控制的问题，提出了一种基于支持向量机（support vector machine, SVM）的智能家居辅助决策模型，研究内容主要包括以下三点：

1) 如何确保训练数据的准确度，从而提高 SVM 训练结果的精度，除了尽量避免信道干扰外，本文给出了一种基于相似度估值的异常数据处理算法用于确保训练数据的准确度。

2) 分析研究智能家居辅助决策系统及相关算法，给出了一种支持向量机和 K 近邻算法相结合的智能家居辅助决策模型，实现家居系统的自主学习和自动控制，减少人为参与。

3) 如何提高 SVM 分类精度，使得分类最优。文中从混合核函数和参数优化两个方面进行考虑，构造了一个包含局部核函数和全局核函数的线性函数，通过调节二者比例使得分类最优；其次，提出了改进的萤火虫算法应用 SVM 参数选择中，进一步提高 SVM 分类性能，并通过实验验证了其可行性。

4) 如何为具有相似环境要求的用户分组，分组的准确与否直接影响到了新用户组别的判定。文中采用相似用户查找算法来确定用户的组别，并不断对用户行为进行监听，实时更新用户组别。

本文创新点如下：

(1) 针对标准萤火虫算法存在收敛早熟及无法收敛的问题，提出了一种基于自适应步长和增强随机扰动的改进萤火虫算法，通过引入亮度值和欧几里得距离公式改进了原算法中的位置更新公式，加快收敛速度。

(2) 将改进后的萤火虫算法应用于 SVM 参数选择中，增加了萤火虫算法的适用范围。

(3) 将改进后的 SVM 分类器应用于智能家居辅助决策系统中，提高了系统决策的精度。

1.4 本文组织结构

本文共分为六章，内容安排如下：

第 1 章：介绍了本文研究内容的背景、意义及现状，并总结和陈述了本文的主要内容。

第 2 章：介绍了研究课题的相关技术，包括智能家居、辅助决策、决策分类技术及常用的支持向量机参数优化算法。

第 3 章：详细介绍了基于 SVM 的辅助决策模型构建，包括辅助决策系统模型总体架构设计、个人环境信息建模、异常数据的处理、相似用户查找及基于 SVM 的分类模型构建。

第 4 章：介绍了萤火虫算法的相关概念及优缺点，提出了一种基于自适应步长和增强随机扰动的改进萤火虫算法，并进行了算法的测试；最后，提出了一种混合核函数用于优化辅助决策模型。

第 5 章：主要对辅助决策系统的设计与实现进行了介绍。包括家居环境信息的采集、软硬件系统的介绍及本文决策系统与其他决策系统系统的对比。

第 6 章：总结与展望。对整篇文章进行了总结并对下一步研究方向进行了探讨。

第 2 章 智能家居辅助决策相关技术

2.1 智能家居辅助决策系统

2.1.1 智能家居

智能家居的本质是利用计算机技术、先进的网络通信技术以及综合布线技术,将人们日常家居生活中相关的各个子系统联合起来,实现对其的统筹管理。通过使用智能家居的相关技术^[25],人们的生活将会更加的环保、舒适、智能、安全、节能,近年来,随着物联网技术、机器学习技术的发展,人们希望智能家居中的设备能够“自主学习”人们的生活习惯进而自动控制家居的运转或关闭,减少居民在家居控制方面投入的时间和精力。例如夏天,决策系统根据住宅的温度以及居民平时下班时间等情景信息进行决策,自动发送控制命令,在居民回家之前将空调打开,提供必要的照明灯,这样即使居民在下班前忘记了远程控制,也能在回家后享受舒适的居家环境。

智能家居的运作可以分为三个过程,一是环境数据的采集,即所谓的“环境感知”过程,将物理世界的自然信息转变为可用于网络传输和计算的数字信息;二是根据“感知”的数据进行数据分析,然后做出决策的过程。决策的形式有以下三种情况:一是完全由人进行,即人根据展示的信息来自己做出决策,进行控制。二是决策完全交给智能家居中的辅助决策系统来完成,即辅助决策系统根据感知层采集的数据通过数据挖掘的形式,获得有用信息,进而做出相应决策。三是以上两种决策方式相结合,一般由辅助决策系统进行控制,然而当系统所做出的决策与用户的需求不一致或者相矛盾时,用户也可以直接决策进行干预和纠正。如今的智能家居环境中,往往采用第三种决策形式,而研究的重点在于如何提高辅助决策系统的准确率,使决策更加人性化、智能化;第三个过程是将人或决策系统发出的控制命令传给相应的设备,执行相应的操作。

辅助决策系统的功能包括预防火灾、煤气泄漏、“学习”与智能控制等。但现阶段的也有诸多不足之处,辅助决策系统根据环境状况进行决策以控制设备的运转,然而这些决策系统采用的算法和规则过于简单,忽略了人与人之间的差异性以及环境的多样性,决策不够准确。

2.1.2 辅助决策系统

智能家居系统包括中央控制管理系统、家庭网络、终端（如传感器、控制器）、信息中心、外联网络等，其中信息中心中包括辅助决策系统，本系统将嵌入式系统和机器学习结合起来，既实现了远程监控的功能，同时利用分布式深度学习算法使系统能够适应用户的行为，动态的做出决策，在很大程度上弥补了传统智能家居系统在对家居电气设备智能控制上的不足。该系统的总体架构如图 2-1 所示。

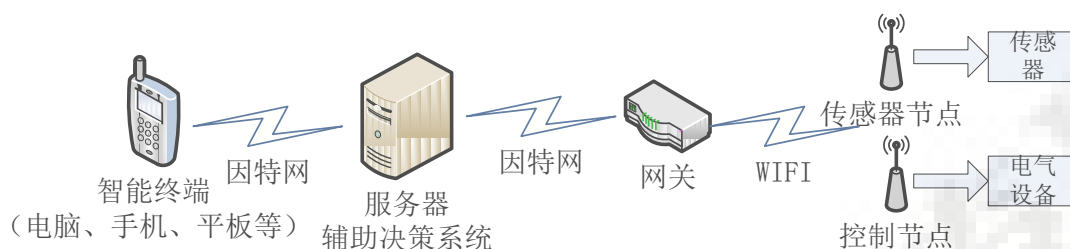


图 2-1 系统总体架构

数据的传输流程为：用户使用智能手机通过 WiFi 和辅助决策系统的计算机相连接，并与之交互。而智能家居环境中的无线传感器、执行器以及基站共同组成无线传感器网络，这些节点通过 ZigBee 协议进行连接和通讯。而这里的基站，是能够实现以太网和 ZigBee 相互转换的 ZigBee 以太网网关，它将以太网和 ZigBee 之间的数据进行透明的转化，实现辅助决策系统和无线传感器网络的通讯。智能家居中的数据传输方向如图 2-2 所示。

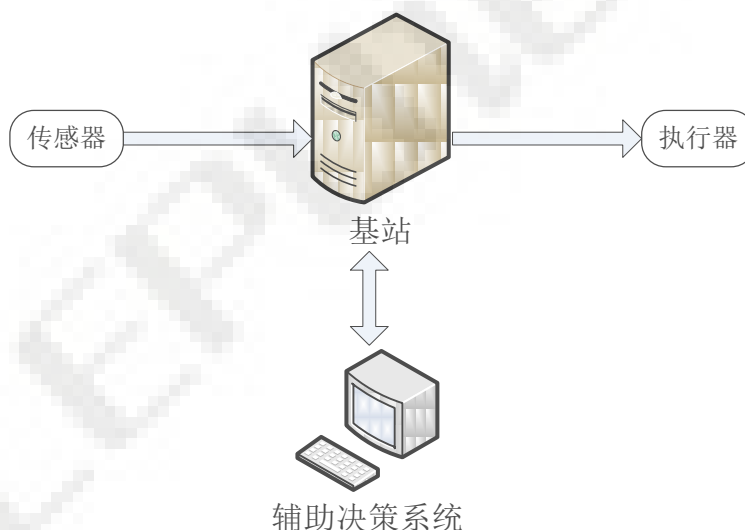


图 2-2 数据传输方向示意图

决策的评估标准包括决策的准确度、新用户分组的准确度等，即判断辅助决策系统给出的控制命令是否与人给出的控制命令一致来判断决策的正确与否，新加入用户的行为习惯与系统为其设定的行为习惯是否一致。

2.2 决策分类技术

2.2.1 BP 神经网络

BP(Back Propagation)神经网络^[26]是 1986 年由 Rumelhart 等人提出的，是一种信号前向传播、误差反向传播的多层前馈型神经网络，其不需事先确定输入-输出的数学关系。其以梯度下降法作为其学习规则，其中层次一般包括输入层（input layer）、隐含层（hidden layer）、输出层（output layer），输入层和输出层均只有一层，而隐含层可包括多层；而前馈指误差的反向传导。此外，其是一种有监督的算法，即有导师学习。该算法的思想是输入信号从输入层输入，向后层层传递，且前一层的输入只影响后一层的输出，当输出未满足预期时，则误差反向传导，同时根据误差调整网络的权值和阈值，这样根据不断的输入一次次的进行权值和阈值的更新、调整至结束。其学习过程如图 2-3 所示。

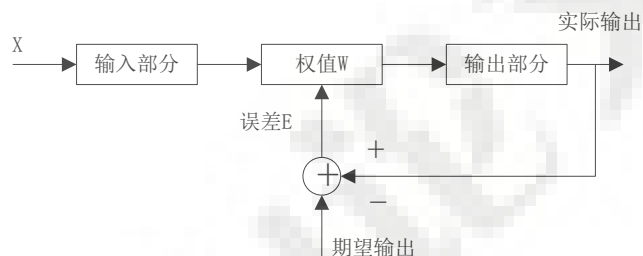


图 2-3 BP 学习过程

BP 神经网络算法具有较强的非线性映射能力、样本数据要求精度高数据量大等特点。一个典型的 3 层 BP 神经网络的拓扑结构如图 2-4 所示。

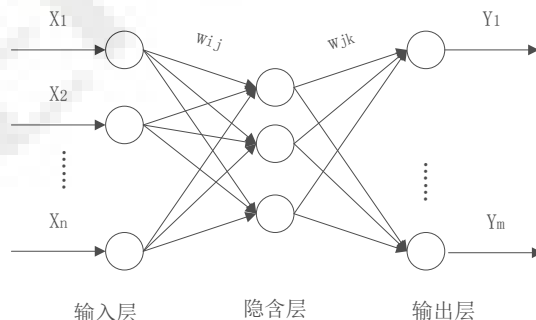


图 2-4 三层 BP 神经网络拓扑结构图

图 2-4 中, X_1 至 X_n 表示该神经网络的输入, Y_1 至 Y_m 表示输出, w_{ij} 和 w_{jk} 为层间的权值。

BP 神经网络具有较好的自适应和自组织能力,在学习过程中调节权重和阈值,来适应训练集。虽然 BP 神经网络有很广泛的应用,但也有一些缺陷和不足。

(1) 其学习速率需预先设定,而一般为避免错过最优解,其一般设置较小,这便导致了训练时间长等问题,而智能家居中的环境数据较为复杂、实时性要求高,使用 BP 神经网络决策可能耗时太久,影响用户体验。

(2) BP 算法的学习规则为梯度下降法,很可能得到一个局部最小值,易陷入局部最优解。

(3) BP 神经网络的学习和记忆具有不稳定性,若增加了新样本,则需重新计算权重和阈值。而智能家居中的环境数据具有多样性,用户规则集可能需要时常更新,使用 BP 神经网络无疑增加了决策的复杂性。

2.2.2 支持向量机

支持向量机 (support vector machine, SVM) [27] 是由 Vapnik 等人在 20 世纪末提出的一种基于 SLT 的机器学习算法。其基本思想是寻找最优的分类超平面,该平面尽可能的距离两类样本间的间隔最大。其最初设计用于解决线性空间内的二类分类问题,但实际中大部分问题具有非线性、多维化等特点。为此, SVM 引入了核函数 [28] 的概念,核函数用于将低维非线性的样本映射到高维空间,使得其可分。核函数不同,支持向量机算法也不同。目前常用的核函数主要有四种:

(1) 线性核函数: $K(x_i, x_j) = x_i \cdot x_j$;

(2) 径向基核函数: $K(x_i, x_j) = \exp(-\gamma |x_i - x_j|^2)$;

(3) 多项式核函数: $K(x_i, x_j) = \sigma(\lambda x_i \cdot x_j + \theta)$;

(4) 多层感知器 (Sigmoid) 核函数: $K(x_i, x_j) = \tanh(\lambda x_i \cdot x_j + \theta)$;

在四种核函数中径向基核函数的应用最为广泛,但也应具体问题具体分析。SVM 具有可靠的理论基础,模型简单,其在小样本、非线性方面具有无与伦比

的优势，通过核函数的引用，在一定程度上避免了维数灾难，部分克服了“过学习”的问题，被广泛应用于模式识别、回归预测、图像识别中。支持向量机的体系结构如图 2-5 所示。

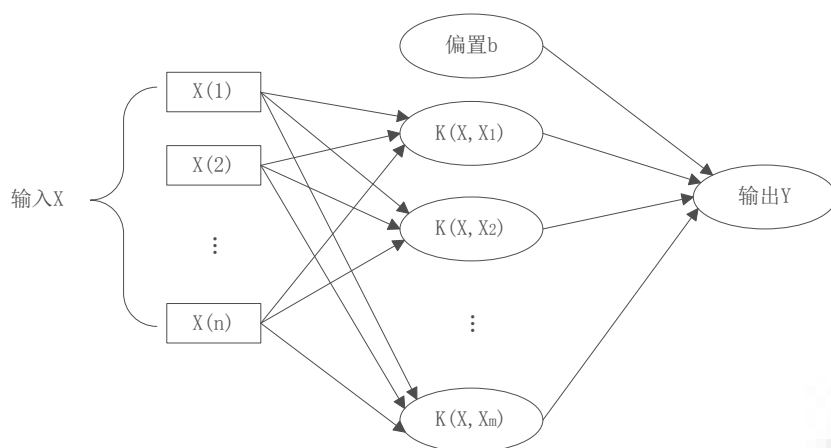


图 2-5 支持向量机体系结构图

此外，现实生活中的分类问题大部分为多类分类问题^{[29][30][31]}，针对这一情况，需要将 SVM 改造为多类分类器，目前常用的方法多为将多个二分类器组合来实现多分类器的构造，常用方法为一对多和一对一方法。第一种的思想为将某一类归为一类，其余的均归为另一类，若样本数为 k ，则构造的分类器数也为 k ，即 k 个 SVM。第二种方法的思想是：任意两类构造一个支持向量机，假设样本数为 k ，则需构造 $k(k-1)/2$ 个支持向量机。当有测试样本时，利用所有分类器进行投票，得票数最多的类别设定为未知样本的类别。其中，LIBSVM 工具箱中的多类分类函数即是基于第二种思想。

本次实验中采用的多类分类问题的方法即为一对一法，一对多方法中训练集为 $1:M$ ，这种情况下存在偏颇，因而不是很实用。此外，因为本实验中分类数目较小且一定，即 model 数较少，代价在可接受范围内，因此多类分类器的构造方法选用一对一方法。

2.2.3 决策树算法

决策树算法^[32]通过归纳学习来得到树形的分类模型。它通过对一组无序、无规则数据的训练，推理出树形分类规则，用于对未知样本的分类。它的实现主要分为两步：第一步是通过训练样本集来建立并优化一棵决策树。该过程实际是对数据集进行学习来了解数据集的规则，进行机器学习，该过程通常分为

两个阶段：建树和剪枝。第二步是利用建立好的决策树模型进行数据分类预测。决策树的建树流程如图 2-6 所示，其中 A 表示样本集合， S 表示训练集， N 表示一个分类叶节点。

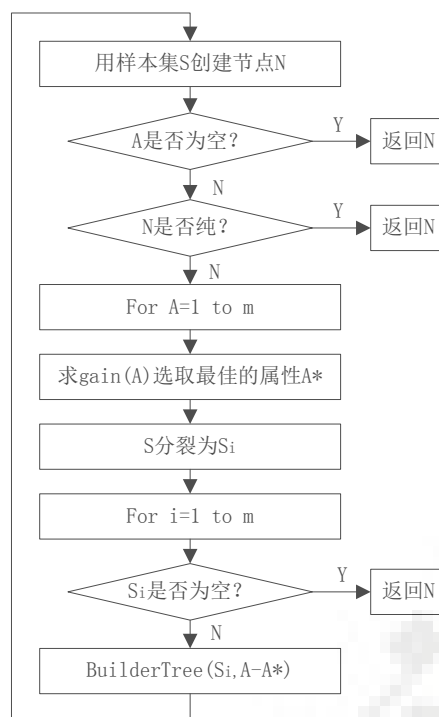


图 2-6 决策树算法流程图

剪枝是决策树建立过程中一个必不可少的阶段，是一种减少训练样本噪声影响的基本技术，在准确理解分类特征描述和防止噪声影响的前提下，对生成的决策树按照一定的方法进行剪枝优化，来达到更好的修剪效果，既确保了分类精度，也提高了可理解性。各种决策树算法中最经典的当属 ID3 算法，其应用最为广泛，后来人们又提出了基于改进的 ID4.5 算法。后来，为应对分类数据量大的特点，人们又提出了诸如 PUBLIC 算法、SPRINT 等专门针对大数据的改进算法。

决策树算法具有结构简单、分类效率高、计算量小、分类精度高等优点，得到了广泛的应用，但该算法在处理缺失数据、过度拟合等问题时也表现出了较差的分类效果。

2.2.4 K 近邻分类算法

K 近邻分类(K nearest neighbour)算法最早由 Hart 等人在 70 年代提出，其

基于临近规则，即其类别判定仅依靠与它临近的几个样本，KNN 算法基本思想为：利用选定的距离函数选取与待分类样本距离最近的 K 个样本，选择 K 个样本中出现次数最多的一类作为该待分类样本的新类别^{[33][34]}。如图 2-7 所示。

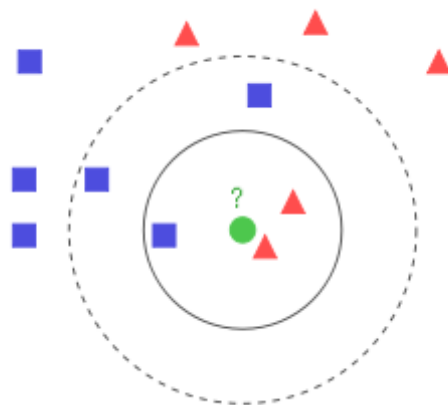


图 2-7 KNN 实例图

判断图中绿色圆形属于红色三角一类还是蓝色方框一类？当 K 为 3 时，红色三角多于蓝色方框，其属于红色三角一类；当 K 为 5 时，蓝色方框多于红色三角，其属于蓝色方框一类。

该算法具有理论成熟、思路简单、实现快速等优点，但同时也存在着一些不足：

(1) K 值的设定。研究表明，设定合适的 K 值对 KNN 算法的分类精度具有很大的影响， K 值过大或过小均会影响分类精度。至今的研究中，仍未提出有效的方法来进行 K 值的选取，常通过设置不同 K 值对比的方法来寻找合适 K 值。

(2) 针对不同类别样本差距较大的情况，KNN 算法精度常常不能保证。当待分类样本属于训练集中样本数较少的类别，而其他类别样本数又较多，即使怎样增大 K 值，均不会得到预期的效果。

(3) 计算量有时很大。当训练集样本数较多时，需计算测试样本至每个训练样本的距离，无疑显得笨重、繁琐了很多，也降低了算法的执行效率。人们常在训练前删除无关数据的方法来精简训练集，提高算法效率。

2.3 支持向量机参数优化算法

SVM 模型参数优化主要为核函数参数 g 和惩罚参数 c 的选择，二者是影响支持向量机分类性能的主要参数。参数 g 则决定了数据映射到新的特征空间后

的分布。参数 c 为惩罚因子，其值越大意味着越不允许误差的出现，因为该参数代表着误差的容忍度，其作用是调节置信范围和经验风险的比例，使得机器学习的能力最优。 c 的大小与算法复杂度成正比，与经验风险值成反比^{[35][36]}。本文中选取高斯核函数作为 SVM 分类器的核函数。其参数 σ 为径向基半径，高斯核函数公式如式（2-1）所示。

$$K(x_i, x_j) = \exp \left[-\frac{|x_i - x_j|^2}{\sigma^2} \right] \quad (2-1)$$

由公式可知，当 σ 无穷大时，错误率为 0，但正确分类率也较低，即学习能力低，分类能力差；当 σ 无穷小时，得到一个常数，正确分类率低，从而引发“过度拟合”。

若记 $g = \frac{1}{\sigma^2}$ ，则优化 σ ，就是优化 g 。参数 g 对高斯核函数性能优劣有很大影响。

近年来，人们开始利用各种智能优化算法来自动优化惩罚因子和核函数参数的值^{[37][38][39]}。常用的参数选择优化算法概括为三大类：序列最小优化算法、网格搜索算法和群智能优化算法。而常用的群智能算法包括遗传算法、粒子群算法等。

2.3.1 网格搜索算法

网格搜索算法（Grid Search Algorithm, GSA）基于交叉验证的思想，是一种较为基本的 SVM 参数优化算法^[40]。其原理是空间范围一定，将待搜索参数按照一定步长划分为一系列网格，最后遍历测试网格中的点来寻找使得模型性能最优的参数值对。基本步骤是首先设定参数 g 的范围 $g \in [g_1, g_2]$ ，变化步长是 g_s ，参数 c 的范围 $c \in [c_1, c_2]$ ，变化步长是 c_s 。从两组范围中取数分别配对如 (c', g') 进行训练。由此引发的缺点是计算量大，效率低。一种改进算法是首先设定较大的范围和步长，粗略选择出一组最优的 (c', g') 。然后再在附近范围设置较小的范围和步长进行精确查找。

网格搜索算法是一种实用的数据搜索方法，其具有简单方便、寻优速度快等优点，但同时还具有参数较多时训练时间长、参数需提前预估、参数空间区域较小等缺点。

2.3.2 遗传算法

遗传算法（Genetic Algorithm,简称 GA）是一种随机全局搜索优化算法，通过借鉴生物进化论和遗传学说，模拟生物界中生物的发展进化方式得出^[41]。由 Michigan 大学的 Holland 等人在 1967 年提出，主要内容有包括复制、交叉、变异等遗传算子，使用双倍体的编码方法等。此外，它将适者生存规则和群体中染色体的随机交换机制相结合使用，并对当前最好假设进行变异和重组，得到后续假设。遗传算法的流程图如图 2-8 所示。

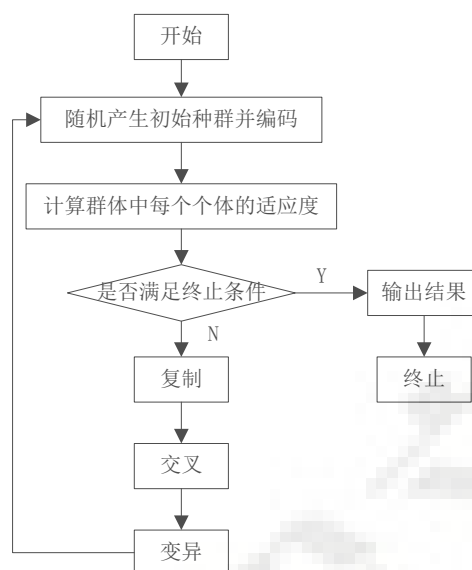


图 2-8 遗传算法流程图

步骤1: 编码，用遗传空间的基因型串结构数据表征解空间的所有数据，解空间中的不同点即用这些串结构数据的不同组合来表示。

步骤2: 初始化群体的生成，随机初始化 m 个串结构数据，一个表示一个生物个体，设置代数计数器 t 为 0，最大进化次数 T ，初始群体用 $P(0)$ 表示。

步骤3: 个体适应度的计算，主要为适应度函数的确定。

步骤4: 复制，也称为选择。根据个体适应度，从当代中选择出最优个体用于下一代的繁殖。

步骤5: 交叉，是新个体产生的主要方式，也是遗传算法中的主要操作。即将群体中各个个体随机配对，并交换部分染色体，产生新个体。

步骤6: 变异，即随机选择个体，对其基因进行改变，在一定概率下发生突变。

步骤7: 结束条件判断：若 $t \leq T$ ，则 $t=t+1$ ，并转到步骤 4 继续执行；若 $t > T$ ，

将得到的具有最大个体适应度的个体输出，运算终止。

遗传算法具有搜索范围广、不易陷入局部最优等优点，但其也有其缺点，如 GA 编码不准确、效率低、易过早收敛等缺点。

2.3.3 粒子群算法

粒子群算法(Particle Swarm Optimization,PSO)是由 J.Kennedy 等人于 1995 年提出，受鸟群捕食行为的启发而获得，其基本思想为模拟鸟群运动，得到问题的最优解^[42]。发展至今，该算法已相当成熟。该算法中，用搜索空间表征鸟类的搜索空间；用食物表征问题最优解；而每个解就是空间中的一只鸟，称其为“粒子”。每个粒子均有一个适应度值(fitness value)和一个速度，该速度决定了方向和距离，最后在解空间中，粒子们随着最优粒子的移动而移动。其算法流程如图 2-9 所示。

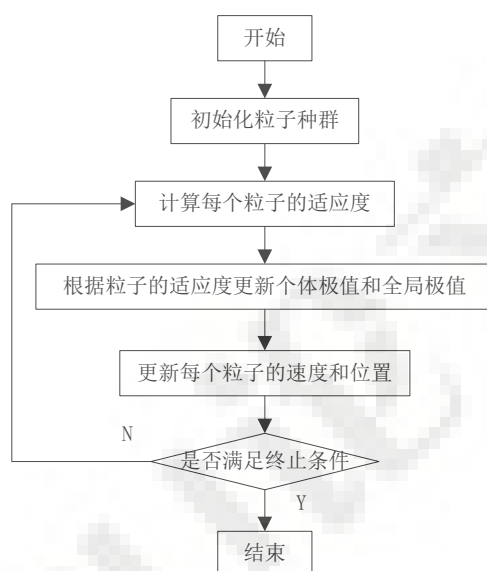


图 2-9 粒子群算法流程图

步骤1: 初始化，包括粒子群的位置、速度、种群规模、最大迭代次数等。

步骤2: 计算各个粒子的适应度值。

步骤3: 比较各个粒子的适应度值，得到最优的值 best。

步骤4: 将每个粒子的适应度值与种群最优值对比，若当前值更优，则更新 best 的值及下标。

步骤5: 更新粒子的速度、位置。

步骤6: 结束条件判断，是否达到迭代次数，若是，寻优结束，否则， $t=t+1$ ，并转至步骤 2。

粒子群及改进算法因其收敛速度快，全局优化性能高等特点在这一领域应用较广，但其局部性能较差，且精度易浮动。

2.4 本章小结

本章介绍了智能家居辅助决策的相关技术。简要陈述了智能家居及辅助决策系统的概念及功能；随后，对各种决策分类技术进行了介绍分析，包括 BP 神经网络、支持向量机、决策树算法、K 近邻分类算法；最后，分析对比了网格搜索算法、遗传算法、粒子群算法在支持向量机参数优化中的应用与优缺点对比分析，为本文算法的提出做铺垫。

第3章 基于 SVM 的辅助决策模型

3.1 辅助决策系统总体架构

智能家居是一种融合了多种技术于一体的智能化家居控制系统，综合管理家居生活中的安防、灯光控制、窗帘控制等。此外，还有基于机器学习的分析系统，用于对用户行为的自动化分析、数据挖掘，用以提醒用户注意节能环保及智能控制家居终端，制定优化策略算法等。智能家居的目标是如何以少量的投入、简便的实现方法来获取最大的功效，实现便捷、安全、智能化的生活。本文的智能家居系统主要包括传感器、智能家居设备、服务数据处理端、PC 和手机 APP，其中辅助决策系统位于数据处理层。其主要的连接关系及系统层次关系如图 3-1、3-2 所示。

硬件方面包括：继电器、无线通信、红外、PC 等。

软件方面包括：通信协议、有线无线网络协议的集成与实现；针对传感器的查询、添加、删除、控制等软件的实现。对窗户、窗帘、照明灯的远程控制。

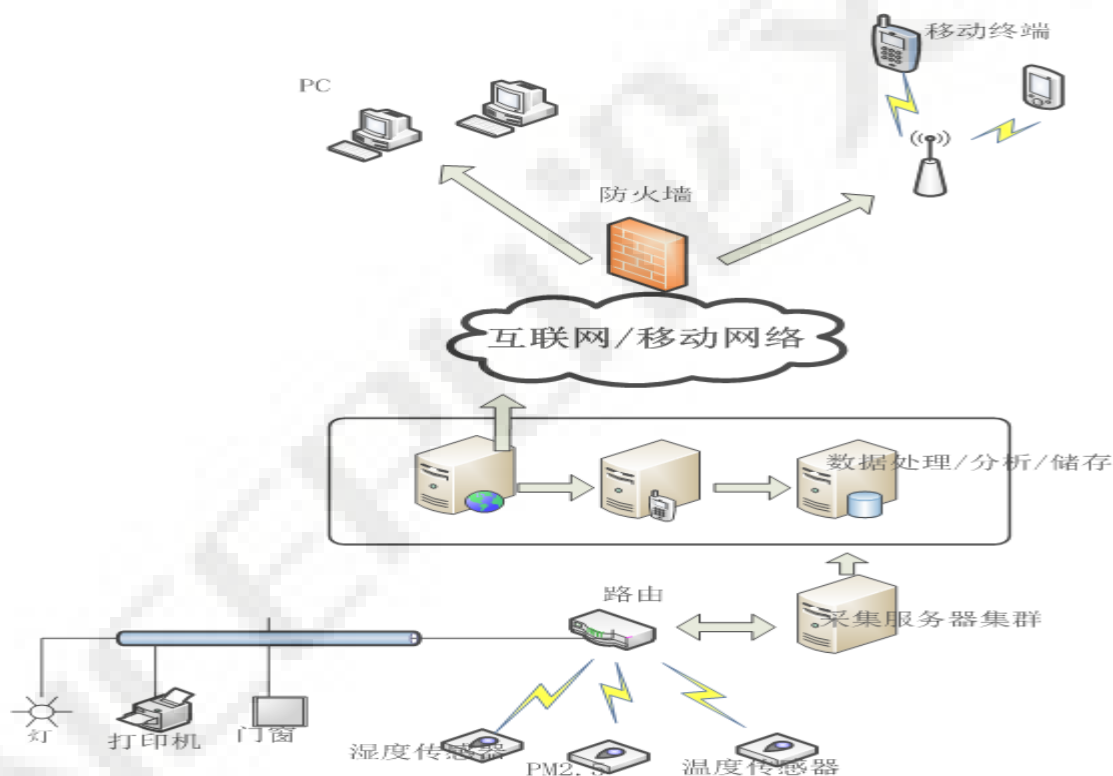


图 3-1 智能家居系统结构图

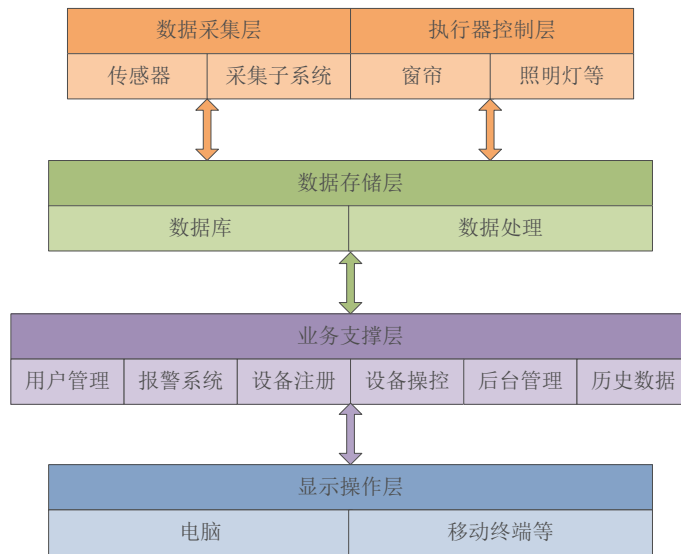


图 3-2 系统总体层次关系图

数据采集层主要包括传感器及智能家居终端等。系统通过传感器等设备采集家居的状态信息，通过内部的 WIFI 模块上传采集数据至路由器，并由路由器将数据上传到后台服务数据处理端，后台服务器对数据进行存储、分析、处理等操作，将家居控制状态反馈到家居控制器，智能控制家居状态，用户可通过手机 APP 或网站进行直观的查看和控制家居环境信息及家居状态，环境信息包括各项数据指标和变化曲线。

本文将 SVM 算法作为辅助决策的分类算法，通过传感器采集数据，对数据进行预处理，利用数据构建个人环境信息模型，借助于 LIBSVM 工具箱对样本进行训练得到分类模型，进而得到该用户的决策规则集，再利用 KNN 算法确定用户是否属于数据库中原有类别或是新类别，最后利用数据库中的规则集实时对家居进行智能决策。下面详细介绍本文基于 SVM 的辅助决策算法的步骤：

步骤1: 用户登录，对用户环境信息进行预处理，即异常数据处理。数据的精确与否对机器学习算法的分类精确度具有很大影响，因此文中采用一定的算法对训练数据进行预处理，包括缺失数据的填充和异常数据的修改等。

步骤2: 个人环境信息建模。选取某用户的编号及同一时间的执行器状态、室内室外的温湿度、光照等信息作为 SVM 算法的一个样本。

步骤3: 判断用户类别，若为老用户，则根据用户 ID 从数据库中选定规则集控制；若为新用户，利用系统的通用决策规则（该规则集为用户数最多的一组类别的决策规则集）进行智能控制，并监听用户行为。若一段时间内用户无

控制指令，则表明该规则适用于该用户，否则，转步骤 4。

步骤4: 利用 KNN 算法查找与该用户行为相似的一组类别，将其规则集重新赋予该用户，并继续监听用户行为，若一段时间内用户无控制指令，则表明该规则适用于该用户，否则，转步骤 5。

步骤5: 表明数据库中无该用户适用的规则集，则利用 SVM 算法为用户单独训练决策规则集。构建、优化、完善个人的决策规则集。选取适当的训练集和测试集，利用 SVM 算法和 LIBSVM 工具箱对样本进行训练，得到不同的模型用于测试，并测试分类的准确度。通过对核函数选取的改进及参数的选取来改进分类的精确度。对某用户新加入的与原来不同的环境数据重新构建决策模型，并训练，最终加入其决策规则集中。

其流程图及时序图如图 3-3、3-4 所示。

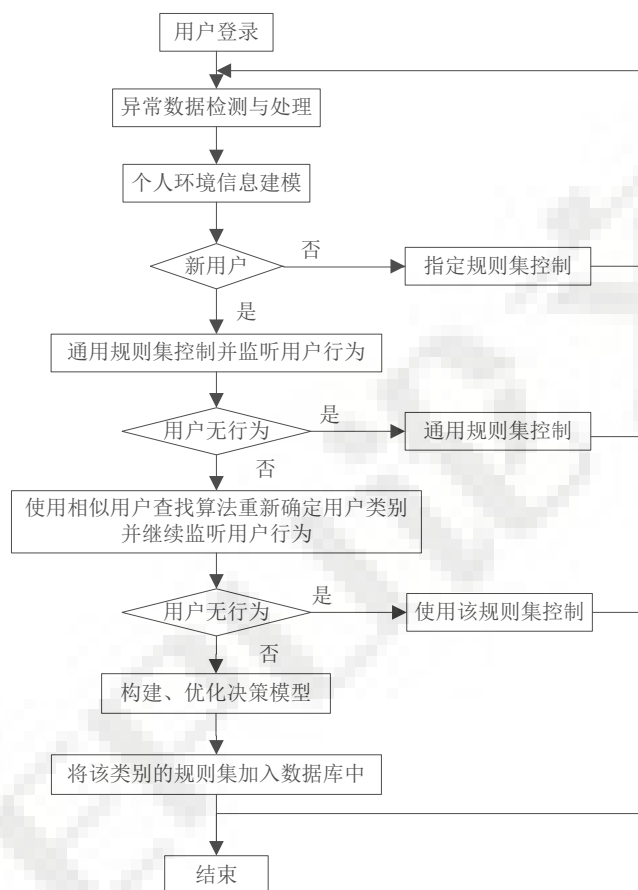


图 3-3 辅助决策算法流程图

如图 3-3 所示，其流程中最重要的步骤主要包括异常数据检测与处理、个人环境信息建模、合适的相似用户查找算法及分类模型的构建、优化。前三项及分类模型的构建在本章后续章节进行分析研究与实验，而分类模型的优化在

第 4 章进行了详细阐述。

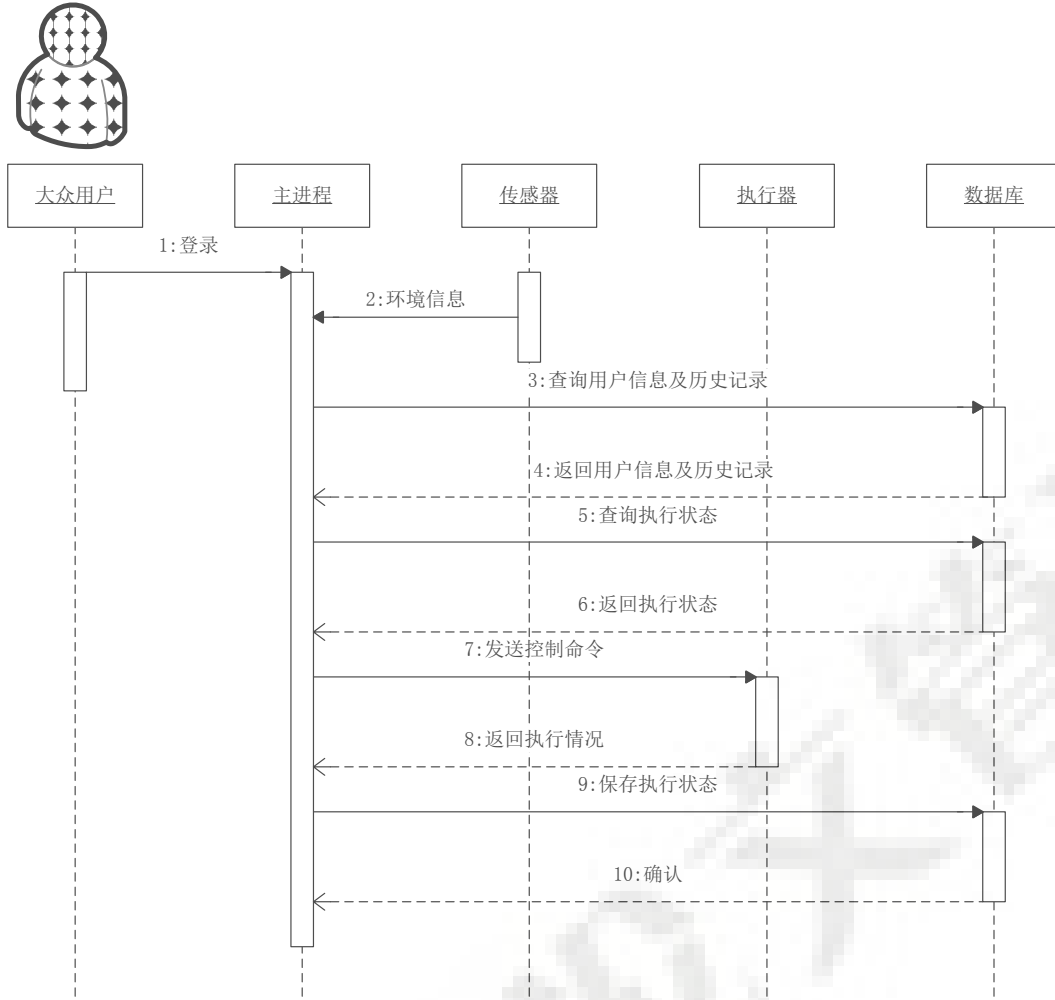


图 3-4 系统时序图

3.2 个人环境信息建模

智能家居环境信息主要包括温湿度、光照及执行器状态等，本文中特征选取为一组室外环境信息和一组室内环境信息，包括温度、湿度和光照。其中，为确保测量精度、减小误差，室内、室外分别设置两组温湿度及光照传感器。用 $\overrightarrow{In_1}$ 和 $\overrightarrow{In_2}$ 分别表示两组室内传感器， $\overrightarrow{Out_1}$ 和 $\overrightarrow{Out_2}$ 分别表示两组室外传感器，则有：

$$\begin{aligned}\overrightarrow{In_1} &= \{L_{I1}, T_{I1}, H_{I1}\} & \overrightarrow{In_2} &= \{L_{I2}, T_{I2}, H_{I2}\} \\ \overrightarrow{Out_1} &= \{L_{O1}, T_{O1}, H_{O1}\} & \overrightarrow{Out_2} &= \{L_{O2}, T_{O2}, H_{O2}\}\end{aligned}$$

其中 L_{i1} , T_{i1} , H_{i1} 分别表示 In_i 这组室内传感器采集的光照、温度、湿度信息, 另外三组类似。为降低维数、减小复杂度及确保测量数据的精确度, 室内、室外的环境信息分别取平均值, 即:

$$\overline{In} = (\overline{In_1} + \overline{In_2}) / 2, \quad \overline{Out} = (\overline{Out_1} + \overline{Out_2}) / 2;$$

此外, 用 \overline{S} 表示控制窗户、窗帘、照明灯开闭的执行器的状态, 则有:

$$\overline{S} = \{s_1, s_2, s_3\}$$

其中 s_1 表示窗户的状态, s_2 表示窗帘的状态, s_3 表示照明灯的状态, s_1, s_2, s_3 值要么为 0, 要么为 1, 0 表示状态为关, 1 表示状态为开。同时, 为区别不同用户, 为每个用户编号, 用 ID 来表示不同用户, 若为 0 则表示该用户为新用户, 当 ID 大于 1 时, 表示该用户为老用户。此外,

综上所述, 该系统的输入数据格式如下:

$$\overline{I} = \{\overline{Out}, \overline{In}, \overline{S}, ID\}$$

当输入以上格式的数据后, 系统判断用户类别进而选择相应的决策; 将数据中的 ID 去掉, 作为决策算法的输入, 进而产生决策结果, 即执行器状态, 用 \overline{C} 表示。 \overline{C} 的格式与执行器状态 \overline{S} 一致, 即用 $\overline{C} = \{c_1, c_2, c_3\}$ 。当 \overline{C} 的结果与 \overline{S} 不一致时, 决策系统发出命令通过执行器控制开关状态, 最后将结果保存至数据库中。

3.3 异常数据处理

机器学习算法普遍对数据的要求比较高, 然而当输入数据发生异常或缺失达到一定比例时, 会严重影响分类预测的精确度^[43]。受环境因素、设备因素、网络因素等的影响, 无线传感器网络数据传输的可靠性降低, 导致数据异常或丢失的现象时有发生, 基于此, 我们可以考虑从硬件和软件两个角度来提高数据的准确率。从硬件及网络传输方面, 可通过选择合适的路由算法、定期查看无线传感器状态等方法来提高数据传输的可靠性; 其次, 数据采集完成后, 可通过一些特殊的数据预处理算法来进一步提高数据的准确率, 来尽可能降低异

常数据对支持向量机算法的影响，提高决策的准确性。

传感器异常数据处理流程如图 3-5 所示。

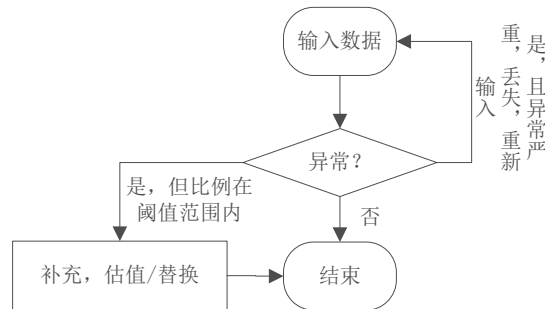


图 3-5 异常数据处理流程图

首先对采集到的数据进行初步检验，查看输入数据中是否存在异常，检测异常的方式分为以下三个步骤：

- 1) 查看是否有丢失数据，若有，记录丢失值的个数。
- 2) 检测采集到的所有数据是否在各自的正常范围内，若不在，记录异常数值的个数。
- 3) 计算丢失和异常值在输入数据中所占的比例 α 。

若 α 的值为 0，表明无异常数据，则处理结束；当 α 大于设定阈值时，此时数据异常严重，丢失该部分数据；若 α 小于设定阈值，则选取适当方法对异常数据处理。本文异常数据处理方法为从训练集中选取 K 个与测试数据最相似的数据，并计算每组数据的权值，相似度越高则权值越大。例如当前输入数据为 $\bar{X}(x_1, x_2, \dots, x_i, \dots, x_n)$ ，其中 x_i 为异常项，通过查找到与之相似的两条数据 $\bar{A}(a_1, a_2, \dots, a_i, \dots, a_n)$ ， $\bar{B}(b_1, b_2, \dots, b_i, \dots, b_n)$ ，那么我们令 $x_i = \alpha a_i + \beta b_i$ ，其中， $\alpha + \beta = 1$ 。其中，在本实验中，重点在于异常值估值的初始样本集的选取、 K 值的选取。考虑到计算量的大小、算法复杂、本实验的数据特点等因素，本文中样本集的选取为选取相邻几天内异常项附近几个时间点的数据， K 值设为 10。步骤如下：

- 步骤1: 根据异常项的时间选取初始的样本集并设定 K 的初始值。
- 步骤2: 利用欧基里德距离公式等确定 K 个与异常项最接近的数据。
- 步骤3: 确定各自的权值，然后确定替换值。

具体实现如下：定义 $\bar{S}_1(s_1, s_2, \dots, s_i, \dots, s_n)$ 表示一个样本， s_i 为异常项，可采用欧式距离计算相似度，如式 (3-1) 所示。

$$dis(\vec{S_i}, \vec{S_j}) = \sqrt{\sum_{k=1}^n (v_{ik} - v_{jk})^2} \quad (3-1)$$

由上式可看出，dis 越小，两个样本的相似度越大，将不同距离的邻居对该样本产生的影响赋予不同的权值(weight)，则权值的计算如式（3-2）所示：

$$W_i = \frac{(1/d_{i,j})}{\sum_{j=1}^k (1/d_{i,j})} \quad (3-2)$$

其中， Dis_i 表示与未知样本 X 最近的 K 个样本中的第 i 个样本的距离，上式反应出离样本 X 距离越大，权值越小。

使用加权平均的方法来计算异常项的估计值，计算公式如式（3-3）所示：

$$s_i = \sum_{j=1}^k w_j \cdot v_{jr} \quad (3-3)$$

其中 r 表示第 r 个属性值异常。

文中称上述方法为相似度估值法。还有另外两种异常值处理的方法，一是选取用户前一天的数据作为估计值，称为正常值替换法；二是选取用户相邻几天的数据取平均值，称为临近估值法。

实验中选取几条正常的的数据，将其中的某项改为异常值，然后用这三种估值算法进行估值，并与准确值进行对比，选出最佳方法。例如：测试某室外温度的估值准确性，将某条正常的室外温度由正常值设为异常值，随后取出训练集中的所有正常数据的温度值分别利用上述三种方法进行估值，通过实验对比三种方法的优劣。一条输入数据格式如图 3-6 所示：

```
>> data(1,:)
ans =
Columns 1 through 6
    0.2000    1.0000   50.0000         0   25.0000   50.0000
Columns 7 through 9
         0         0    1.0000
```

图 3-6 数据格式图

第 1-6 项为室外及室内的光照强度、温度、湿度。第 7-9 项为执行器的开关状态，但最后三项的执行器状态不作为实验数据。

三种方法的估值实验结果如表 3-1 所示。

由表中数据可以看出，每次估值后，相似度估值均取得较满意的结果。而后两种方法受时间、天气、气候等影响较大，数据并不一定与异常数据当天数

据相似。例如第 2 条数据中因为前后两天温差变化较大，用前天数据替代后天数据造成数据的极大错误。第 6 条数据中因为取的是中午和黄昏时候的数据，光照强度相差较大，因此，也不应作为可替换数据。虽然，相似度估值法复杂度较高，但在可接受范围内，所以本文采用基于相似度的估值方法对异常数据进行估值替换。

表 3-1 三种估值方法实验结果表

序号	正常值	相似度估值法	正常值替换法	临近估值法	异常数据
1	27.6513	25.4724	28.8723	27.4612	室内温度
2	4.5626	4.8825	15.4226	7.8745	室外温度
3	32.5686	33.4512	30.5627	31.4524	室内湿度
4	43.5691	43.2573	44.4562	43.9873	室外湿度
5	3567.8533	3345.6945	3456.2831	3539.8773	室外光照
6	2857.6725	2766.1956	1124.9267	2231.8927	室外光照

3.4 相似用户查找算法

前文对单个用户的统计数据训练与预测，得到了分类模型，即单个用户的决策规则集，单个用户何种环境（温湿度、光照）下执行器状态的不同。但系统没必要为所有用户都要训练分类模型，即可将具有相似环境要求的用户归为一组，共用一套决策规则集，文中使用基于 KNN 的相似用户查找算法对相似用户进行查找，并从理论和实验两方面验证该方法的可行性。

定义用户组用 \vec{U} 表示， $\vec{U} = \{U_1, U_2, U_3, \dots, U_n\}$ 表示具有不同环境要求的 n 个用户组号。投票计数器用 \vec{V} 表示， $\vec{V} = \{V_1, V_2, V_3, \dots, V_n\}$ ，使用 KNN 算法查找与该用户环境要求相似的用户，并将计数器中最大项的组别作为该用户的组别。步骤如下：

步骤1: 获取该用户的一组输入 $\vec{I} = \{Out, In, S, P\}$ ，此处去掉了用户 ID。P 为决策系统推测的用户开关状态。

步骤2: 使用 KNN 算法找出与向量 \vec{I} 最相似的 k 个向量的类别号 $\{g_1, g_2, g_3, \dots, g_k\}$ 。

步骤3: 统计每个向量的组别。对 $0 \leq i \leq n, 0 \leq j \leq k$ ，若 $U_i = g_k$ ，则 $V_i = V_i + 1$ 。

步骤4: 选择 \bar{V} 中最大的数的组号作为当前用户的组别, 继续监听用户行为。

步骤5: 监听到用户新的行为, 转步骤 1。

根据前面的描述, 选择具有较大差距的两个组别 A、B, 分别从中取出一个样本 a、b, 分别选取某一时间段两个成员的 10 条行为数据 (环境数据、设备状态、用户命令等), 然后使用基于 KNN 的用户决策规则判定算法进行分组, 设定 KNN 算法中的 K 为 7, 得到的 a 和 b 属于哪个小组的累计投票结果如表 3-2 所示。

表 3-2 投票结果表

每个行为后的累计投票结果											
	小组	1	2	3	4	5	6	7	8	9	10
a	A	3	6	9	12	23	26	27	31	35	42
	B	4	8	12	16	12	16	22	25	28	28
b	A	5	9	12	13	12	12	16	22	25	26
	B	2	5	9	15	23	30	33	34	38	44

由表 3-2 可看出, 用户 a 在前 4 次投票中错分到了 B 组, 后面其余分组均分到了 A 组, 与预期分类符合。用户 b 前三次错分到了 A 组, 后面分组均为 B 组, 符合实际情况。从而验证了本文分组算法的合理性。但存在算法分类速度慢、数据量大时查找速度慢等不足, 可通过某些方法来改进查找算法。如可比较部分属性的相似度, 再从得到的数据中进一步查找其他的相似属性; 或为数据项添加采集日期时间、天气、季节、地域等环境信息索引, 通过外在环境信息缩小查找范围。

因本文研究重点为 SVM 及如何提高 SVM 分类的准确性, 对 KNN 算法的改进未做进一步研究。

3.5 SVM 分类模型构建

根据前文介绍, 一条用户数据的输入格式为: $\vec{I} = \{\overline{Out}, \overline{In}, \overline{S}, ID\}$, 去掉其中的 ID, 剩余数据作为 SVM 算法的一条训练数据输入。首先从数据集中根据用户编号 ID 取出某一用户的 240 条数据, 其中, 数据选取时, 为保证实验的可靠性, 应分别从不同年龄段、不同地域、不同环境喜好的用户中分别选取一定数据, 并进行异常数据处理, 当异常数据在训练集中的比例在可接受阈值内。再从中选取 120 条数据作为训练数据, 剩余的 120 条数据作为测试数据, 因为

根据执行器的状态分类共有 8 种情况，在数据选择时应尽量选择执行器状态相同，但环境数据相差较大的数据，因此，也造成了选择的训练数据量偏少，这样，模型训练才更精确。训练流程如图 3-7 所示。



图 3-7 SVM 训练流程图

利用 LIBSVM 工具箱进行训练，得到的实验结果如图 3-8 所示。

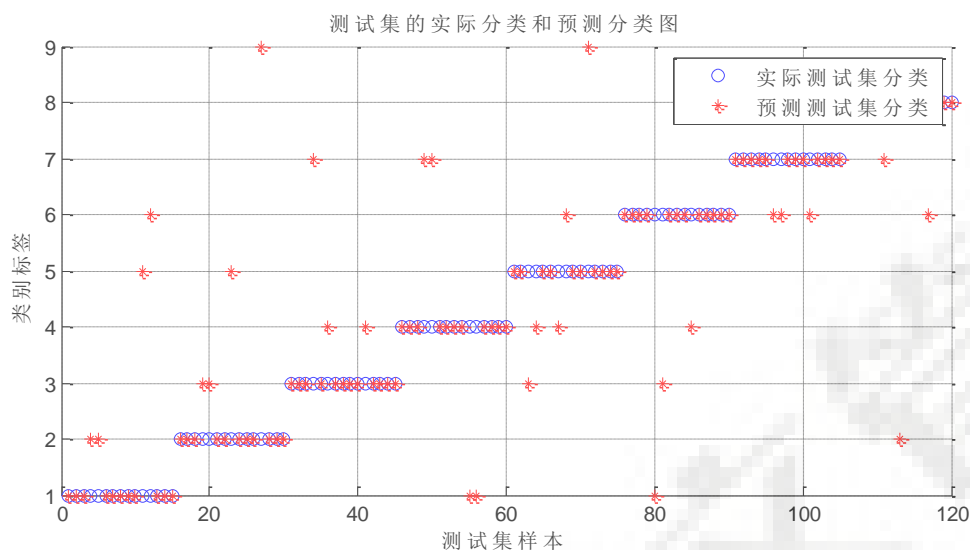


图 3-8 预测结果图

最终得到的分类准确率为： $\text{Accuracy}=75.8333\%(91/120)$ 。实验中通过 SVM 分类模型预测样本测试样本的正确数与测试样本的个数的比值计算准确率，其公式为： $\text{准确率}=\text{分类正确样本数}/\text{测试样本数}\times 100\%$ 。

图 3-8 中，x 轴为 120 个测试样本，y 轴为 8 个类别标签，由实验可看出，利用 SVM 算法实现了用户不同家居环境下的不同决策，实现了用户家居的个性化需求，也进一步提高了家居的智能性。

选用 SVM 方法作为辅助决策系统的决策算法主要有以下几点原因：

(1) SVM 算法具有坚实的理论基础。SVM 算法基于 SLT 理论、VC 维、结构风险最小化等理论，模型简洁明了。

(2) 在辅助决策系统中，个人环境信息的分类需要增量变化，在对某人进行统计后进行分类预测，得出该用户的决策规则。但样本量的有限性，导致环境信息表达并不完整，期间可能还会有新的属于该用户的决策规则加入，而 SVM 具有强有力的增量学习能力，能找出新的个人的决策规则。

(3) SVM 算法分类函数是由少量的支持向量所决定，因此在分类决策时并

不需要计算所有的样本向量，很大的减少了时间复杂度。

(4)用户居住环境的室内和室外环境数据具有明显的相关性，即室内温湿度等的高低会受到室外温湿度等的影响，而 SVM 对特征相关性不敏感，其他算法较为敏感。

(5)在该问题中，无论是用户的个人决策规则集的构建还是多用户的分类，均为多类分类问题，虽然原始的支持向量机是针对二类分类问题的算法，但通过间接采用一对一的多类分类方法，利用 LIBSVM 工具箱中函数较为实用的实现了多类决策分类问题。

(6)SVM 算法在处理非线性分类问题时，通过选取适当的核函数将低维的不可分样本向高维空间映射，进而在高维空间建立线性学习机，由核函数的性质可知，我们不需知道映射的具体表达式，而且虽然增加了维数但几乎不增加计算的复杂性，且在一定程度上避免了“维数灾难”。

但从实验还可看出，75%的分类准确率远远不能满足人们实际生活的需要，因此，本文后续章节从多角度考虑如何提高辅助决策的准确率来减少人们对家居设备的干预，真正使得智能家居设备能够自主学习、自主决策，且拥有较高的决策准确率。调查研究发现，SVM 算法的核函数选取及参数选择对分类准确率具有较大影响，本文在第四章分别从以上两个角度对智能家居辅助决策模型进行了优化。

3.6 本章小结

本章详细介绍了基于 SVM 辅助决策模型的整体架构及流程，包括模型总体架构设计、算法流程及详细实现，其中包括如何构建个人环境信息模型、异常数据的处理、相似用户查找算法等，并利用 LIBSVM 工具箱实现了分类模型的构建，实现了家居的自主学习性、个性化功能，但实验结果并不是很理想，分类效果较差，因此，下一章节重点讨论与研究如何优化辅助决策模型，使其决策更准确。

第 4 章 辅助决策模型优化

4.1 群智能与萤火虫算法

4.1.1 进化计算和群智能

进化算法一般以随机分布的临时解作为起始，以特定启发式的规则对解空间进行不断的搜索和迭代操作，从而不断改进解的质量，这类算法的典型代表有：模拟退火（Simulated Annealing,SA）、遗传算法（Genetic Algorithm,GA）和粒子群算法等。

进化计算作为新兴的求解最优化的方法，一经提出就引起了科学界的兴趣和高度重视。进化算法起源于我们生存的自然环境-自然。在自然界中存在着各种各样的生物，在残酷的自然环境中，为获得食物、交配或躲避天敌的袭击，各种生物养成了各种各样的行为方式来适应自然，获得物种的延续。在整个生命演化过程中，保留下来的生物不可能仅仅是被动的接受环境，它们也学着慢慢地适应环境、改变部分环境，并且通过不断的学习和进化来不断提高其适应环境的能力。而人们在观察环境及其中的生物时，时不时被其中的某种生物强大的生存能力和毅力所折服，赞叹其奇迹般的生存形式，而人们在观察过程中，这种种生命现象正是我们进化算法的思想源泉，激励着人们不断地探索最佳生命结构和形式。

与传统优化算法相比，进化算法有以下一些共同的特点：

- 1) 不以达到某个最优性条件或找到理论上的精确最优解为目标，而对于计算的速率和效率更加看重。
- 2) 对约束函数和目标函数的要求较宽松。
- 3) 算法思想来自于对自然规律的模仿，具有人工智能的特点。
- 4) 大多数智能算法包括一个或多个个体组成种群，将参数寻优的过程演化为种群进化的过程。
- 5) 通常这些算法的理论较薄弱，并不能保证收敛到最优解。

在进化算法中，萤火虫算法属于群智能算法，因为它是受自然界中群居动物生活习性启发而发明的^[53]。群居生物不会以单个个体出动，通常依靠集体的力量，进行觅食、繁衍、御敌等，表现出类似于人类社会的小型社会的特点。

在这个小型社会中，每个个体都有明确的分工。为达到某目的（类似于优化问题的解），群居动物能很好的对成员进行分配，通过成员的分工协作从而达到目的。其智能行为表现在：群体在不确定的动态环境中能够做出适当的反应，而这种反应最终被证明有助于其实现行为目标。群智能定义如下：由具有若干个有限的非智能体所组成的系统表现出集体智能行为的一种特性。群智能算法可以说是建立在人工生命的基础上对群居生物社会系统的模拟。

4.1.2 人工萤火虫算法

人工萤火虫算法（Glowworm Swarm Optimization, GSO）是一种新型的群智能算法，由 Ghose 等人在 2005 年提出^{[44][45]}。该算法具有局部寻优、全局寻优的能力。由于其优势，近年来越来越多的学者将精力投入到该算法的研究中，随着研究的不断深入与发展，已将该算法成功应用于多模态函数优化、声源定位、多信号源定位问题等众多领域。

GSO 算法基于萤火虫发光吸引异性的原理，其算法步骤为：首先在解空间随机设置 n 只萤火虫，并根据位置调节其自身亮度值，且各萤火虫均具有各自独立不同的视野空间，在该视野空间中查找亮度比自己亮的萤火虫，从中选取最亮的一只并向其移动，移动完毕后更新亮度、视野空间及位置，继续寻找下一个萤火虫直至结束。整个算法的迭代流程分为：亮度更新阶段、位置更新阶段及动态决策域更新阶段。

1) 亮度更新阶段

亮度值按公式（4-1）进行更新。

$$l_i(t) = (1 - \rho)l_i(t-1) + \gamma J_i(t) \quad (4-1)$$

$l_i(t)$ 表示第 t 次时的亮度值， ρ 是荧光素消失率，表示荧光的变化速率，

$J_i(t)$ 为第 t 次时的适应度值。

2) 位置更新阶段

在该阶段中，萤火虫先找出比自己亮度亮的所有萤火虫形成邻居集合，如公式（4-2）所示：

$$N_i(t) = \{j: d_{ij}(t) < r_d^i(t); l_i(t) < l_j(t)\} \quad (4-2)$$

$d_{ij}(t)$ 表示在第 t 次迭代时萤火虫 i 和 j 的距离, j 表示萤火虫 i 邻居中的一员, $r_d^i(t)$ 表示萤火虫 i 在 t 时刻的视野范围。

根据轮盘概率公式从邻居集合中选择出要靠近的萤火虫, 如公式 (4-3) 所示:

$$P_{ij} = \frac{l_j(t) - l_i(t)}{\sum_{k \in N_i(t)} l_k(t) - l_i(t)} \quad (4-3)$$

P_{ij} 表示萤火虫 i 向萤火虫 j 移动的概率, 由公式可看出, 萤火虫自身亮度越高, 其被选中的概率就越大。当选择好要靠近的个体后, 根据式 (4-4) 开始向其移动, 并进行位置更新:

$$x_i(t+1) = x_i(t) + s \left\{ \frac{x_j(t) - x_i(t)}{\|x_j(t) - x_i(t)\|} \right\} \quad (4-4)$$

式中 s 表示移动的步长, $\|x_j(t) - x_i(t)\|$ 描述了两只萤火虫间的欧式距离。

3) 动态决策域更新阶段

萤火虫位置更新后, 需更新其视野范围, 即动态决策域, 该范围应适当, 过大或过小均不利于寻找最优解。其更新公式如 (4-5) 所示:

$$r_d^i(t+1) = \min \left\{ r_s, \max \left\{ 0, r_d^i(t) + \beta(n_i - |N_i(t)|) \right\} \right\} \quad (4-5)$$

其中, β 控制决策域的大小, r_s 是初始的视野最大值, n_i 是邻居数的阈值。

GSO 算法实现步骤:

步骤1: 初始化各参数值, 包括萤火虫个数、最大迭代次数、各位置初始化、动态决策域初始化等。

步骤2: 按照公式 (4-1) 进行亮度更新。

步骤3: 按照公式 (4-2) 求得每只萤火虫的邻居集合。

步骤4: 依据轮盘概率公式 (4-3) 在其邻居集合中选择出所要靠近的萤火虫。

步骤5: 按照公式 (4-4) 进行位置更新。

步骤6: 按照公式 (4-5) 进行动态决策域更新。

步骤7: 回步骤 2 再循环, 若次数达最大循环次数, 则循环终止, 输出结果。

GSO 算法的流程图如图 4-1 所示。

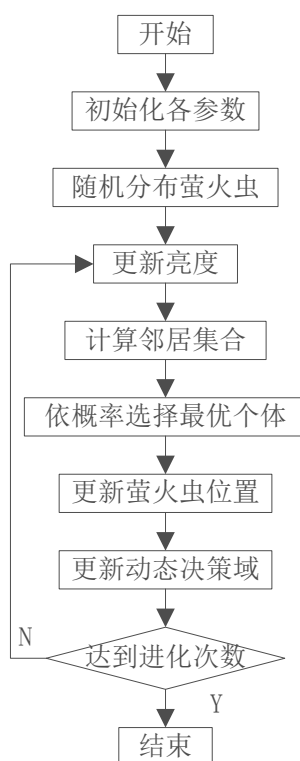


图 4-1 GSO 算法流程图

4.1.3 萤火虫算法

萤火虫算法（Firefly Algorithm, FA）对发光机制做了理想化处理，两只萤火虫间的吸引力只取决于自身的亮度和相互间的吸引力。而当前位置的适应度值决定了自身的亮度，位置越好则亮度越高。亮度越高，则吸引力越大，并按照亮度低的萤火虫向亮度高的萤火虫移动的原则。若亮度相同，二者均随机移动。FA 算法的数学表述形式如下所示：

定义 1 萤火虫的相对荧光亮度计算公式如（4-6）所示。

$$I(r) = I(o) \times e^{-\lambda r_{ij}} \quad (4-6)$$

其中 $I(o)$ 表示初始亮度，其与位置有关，位置越好，则亮度越高； λ 为光线消失率，当距离增加或受到传播媒介的影响， λ 会减小。 r_{ij} 是萤火虫 i 和 j 之间的欧式距离。

定义 2 萤火虫之间的相对吸引度公式如（4-7）所示。

$$\varphi = \varphi_0 \times e^{-\lambda r_{ij}^2} \quad (4-7)$$

λ 为光照吸收系数, φ_0 表示初始吸引度, 即光源处 ($r=0$) 的吸引度, 一般为 1。

定义 3 位置更新公式如 (4-8) 所示。

$$x_j = x_i + \varphi(x_j - x_i) + \alpha \times rand \quad (4-8)$$

其中 x_i 、 x_j 分别表示萤火虫 i 、 j 的空间位置, 此处 i 被吸引向 j 。 α 是步长因子, 为区间 $[0,1]$ 上的常数, $rand$ 在 $[0,1]$ 上随机分布。FA 算法实现步骤:

步骤1: 初始化各参数值, 包括萤火虫个数、最大迭代次数、位置初始化等。

步骤2: 对所有萤火虫位置进行随机初始化, 并计算亮度的初值 I_0 。

步骤3: 计算萤火虫之间的相对亮度与吸引力, 进而确定个体的移动方向。

步骤4: 更新萤火虫的位置, 若某萤火虫已处于最佳位置, 则对其进行随机扰动来扩大搜索范围以寻求全局最优解。

步骤5: 重新计算亮度值。

步骤6: 当达到最大搜索次数后则转到步骤 7; 否则转回步骤 3, 进行下一次的搜索。

步骤7: 输出全局最优解及最优个体。

FA 算法的流程图如图 4-2 所示。

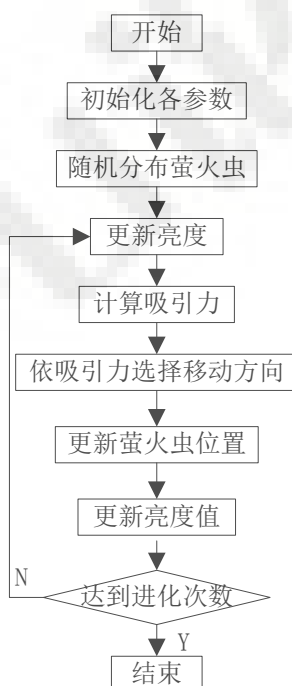


图 4-2 FA 算法流程图

4.2 改进萤火虫算法及其在决策优化中的应用

4.2.1 改进萤火虫算法

研究表明,标准 FA 算法存在收敛早熟及无法收敛的问题,部分原因是参数过早设置或设置不当^[46]。而收敛的快慢与步长有关,固定步长造成算法早期萤火虫个体不能最快速度的靠近亮度高的同伴,同时在算法后期萤火虫之间的亮度差值减小,位置比较接近,原先的步长使得萤火虫个体始终无法向最亮的那只进一步靠近而造成不稳定现象,影响了算法精度^[47]。如何调整步长使得算法具有最优解成为人们研究的热点内容之一^{[48][49]}。

目前研究人员对 FA 算法的改进主要从以下三个方面进行改进:基于自适应策略的 FA、基于参数调节的 FA 及混合萤火虫算法^{[50][51]}。

自适应策略主要针对固定参数提出,自适应策略主张将参数变为动态可变,比如固定步长和扰动因子,通过对步长及扰动因子的改进,实现当两萤火虫相差较远时能快速靠近,当两者距离较近时又可慢速移动,避免了 FA 算法的早熟及最终无法靠近最优解的问题^[52]。

基于参数调节策略主要是通过如混沌思想、自动学习机等调节参数 α 和 γ ,避免算法陷入局部最优,但也同时存在算法过于复杂等问题。

混合萤火虫算法主要通过利用其它算法的优势与 FA 算法结合,如遗传算法等,虽提高了精度,但同时算法存在了复杂化、效率低等缺点。

本文基于自适应策略,对 FA 算法提出了如下两条改进方法:

1) 为使亮度相差较大,距离较远的萤火虫向更亮的萤火虫大踏步的前进,本文借鉴 GSO 算法中的亮度计算公式来改进 FA 算中的位置更新公式。

2) FA 在宽搜索区域和高维度优化方面存在吸引力较弱的问题,较难进行位置更新,本文提出了根据 r_{ij} 调节随机参数的方法,用 $2r_{ij}(1/2 - rand)$ 取代 FA 算法中位置公式中的随机部分。

针对以上两点公式,最终改进的 FA 算法的位置更新公式如式(4-9)所示:

$$x_j = x_i + l_j(t) \varphi_0 e^{-\lambda r_{ij}^2} (x_j - x_i) + 2r_{ij} \left(\frac{1}{2} - rand \right) \quad (4-9)$$

其中 r_{ij} 表示两只萤火虫之间的欧几里得距离,距离计算公式如式(4-10)

所示：

$$r_{ij} = \|x_j - x_i\| \sqrt{\sum_{k=1}^d (x_{i,k} - x_{j,k})^2} \quad (4-10)$$

e 为数学常量，值大约为 1.71828， ϕ_0 表示初始时的吸引力， λ 为光强吸收系数。

$2r_{ij}\left(\frac{1}{2} - rand\right)$ 为在移动完毕后，在附近区域内做一个随机扰动的扰动因子，

其随 r_{ij} 的变化而变化。

亮度越大，移动步长也会越大，能够使得算法在早期加速收敛，后期随着相对亮度的衰减，移动步长也会相应减小，从而能够在最佳位置处进行微调，避免了收敛不稳定现象发生。最后，本文将基于上述方法改进的 FA 算法称为改进萤火虫算法，记为 IFA。

4.2.2 算法流程

步骤1: 初始化各参数值，包括萤火虫个数、最大迭代次数、位置初始化及动态决策域初始化等。

步骤2: 对萤火虫的位置进行随机初始化，并计算初始的亮度 l_0 。

步骤3: 计算萤火虫之间的相对亮度与吸引力，确定萤火虫的移动方向。

步骤4: 依据亮度公式计算亮度。

步骤5: 对萤火虫进行位置更新，若其已处于最优位置，则对其进行随机扰动，扩大其搜索范围来寻求全局最优解。

步骤6: 重新计算亮度值。

步骤7: 当达到最大搜索次数后则转到步骤 7；否则转回步骤 3，进行下一次的搜索。

步骤8: 输出全局最优值和最优个体。

IFA 算法的实现流程图如图 4-3 所示。

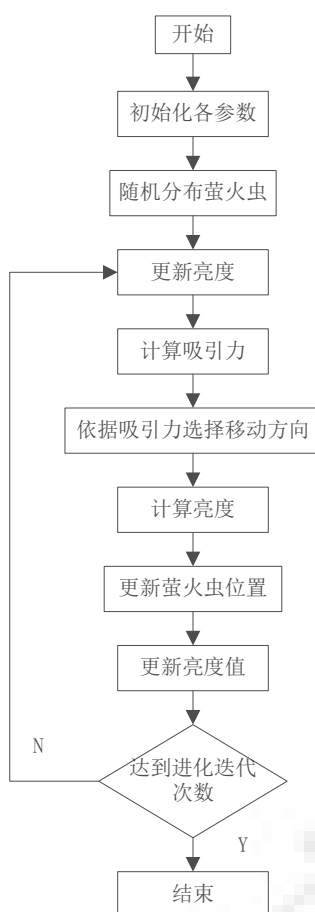


图 4-3 IFA 算法流程图

4.2.3 改进萤火虫算法测试

为验证本文提出的 IFA 算法的有效性，本文采用 4 个常用的标准测试函数进行测试，测试函数如下：

1) Sphere 单峰函数

$$f_1(x) = \sum_{i=1}^n x_i^2 \quad n=30, |x_i| \leq 100$$

sphere 作为经典的测试函数之一，具有函数简单、寻找全局最优值较容易等优点，常用来测试算法的寻优精度。

2) Rosenbrock 单峰函数

$$f_2(x) = 100 \times (x_1^2 - x_2)^2 + (1 - x_1)^2 \quad 1.5 \leq x_i \leq 3$$

该函数为极小化困难的非凸单峰函数，变量之间的关联性较大，全局最优值寻找较为困难。该函数常用来评价算法的执行效率。

3) Rastrigin 多峰函数，其最优点 $x=0, f(x)=0$.

$$f_3(x) = \sum_{i=1}^n [x_i^2 - 10\cos(2\pi x_i) + 10] \quad n = 30, |x_i| \leq 5.12$$

4) Griewank 多峰函数

$$f_4(x) = \frac{1}{4000} \sum_{i=1}^n x_i^2 - \prod_{i=1}^n \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1 \quad n = 30, |x_i| \leq 600$$

查阅相关资料并分析可得最优点为 $x_i = 0, f(x) = 0$ 。本次实验的实验平台选取 win7 的 64 位系统，8G 内存，软件版本为 Matlab 2013a。其中迭代次数为 300 次，算法独立运行 20 次，分别在 10、20、30 维进行测试，实验结果记录如表 4-1 至 4-4 所示。

表 4-1 Sphere 函数测试结果

维数 ID	算法	平均值	最大值	最小值
10	IFA	0.000476	0.000754	1.305e-007
	FA	0.00685	0.0653	0.00306
20	IFA	0.1627	0.7124	1.198e-003
	FA	0.6073	0.9081	0.0795
30	IFA	9.7865	47.224	0.002319
	FA	17.986	99.785	0.00301

表 4-2 Rosenbrock 函数测试结果

维数 ID	算法	平均值	最大值	最小值
10	IFA	0.327	1.422	0.0654
	FA	4.3309	6.8708	2.3543
20	IFA	15.502	21.805	5.309
	FA	50.5539	83.2938	15.2319
30	IFA	29.9863	65.1672	1.5034
	FA	59.3421	125.4564	18.5063

表 4-3 Rastrigin 函数测试结果

维数 ID	算法	平均值	最大值	最小值
10	IFA	0.05023	0.1332	0.0081
	FA	1.2435	3.6023	0.9869
20	IFA	1.5053	7.3301	0.1201
	FA	8.2453	23.5672	3.0142
30	IFA	5.4942	17.3482	1.2169
	FA	19.0302	30.5765	5.0305

表 4-4 Griewank 函数测试结果

维数 ID	算法	平均值	最大值	最小值
10	IFA	2.1439	6.0075	0.6521
	FA	10.3256	34.2493	5.4634
20	IFA	7.4532	17.2194	0.7685
	FA	20.6745	55.0021	12.3984
30	IFA	9.5673	27.1209	4.5683
	FA	24.3563	87.3845	9.0989

在 10 维时这四个函数测试波形图如图 4-4 至 4-7 所示。由下述图表可看出，IFA 算法较 FA 算法具有更好的适应度，其中最重要的是，其在迭代次数上远少于 FA 算法，这表明本文提出的改进 FA 算法具有更好的收敛性，在一定程度上避免了 FA 算法无法收敛的问题，大大提高了算法的执行速度，从实验结果可看出，改进后的 FA 算法达到了预期目的。

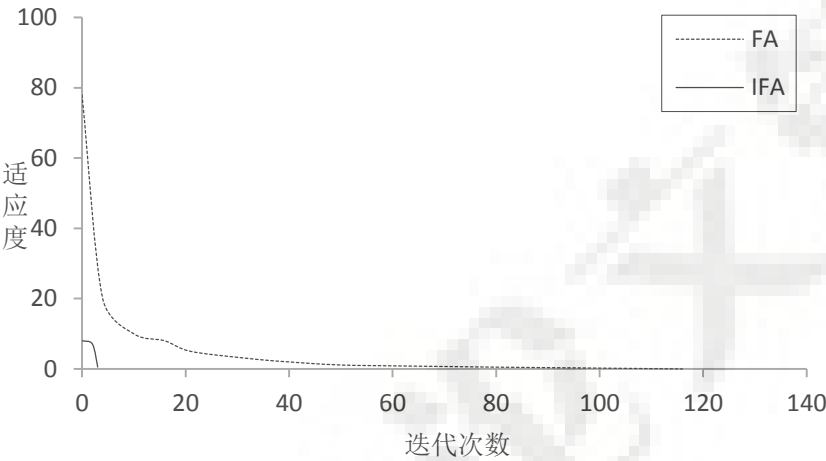


图 4-4 Sphere 函数的收敛曲线对比

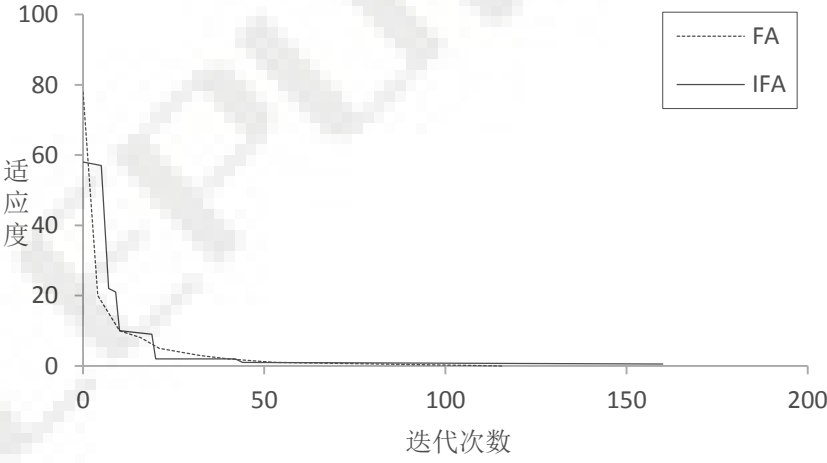


图 4-5 Rosenbrock 函数的收敛曲线对比

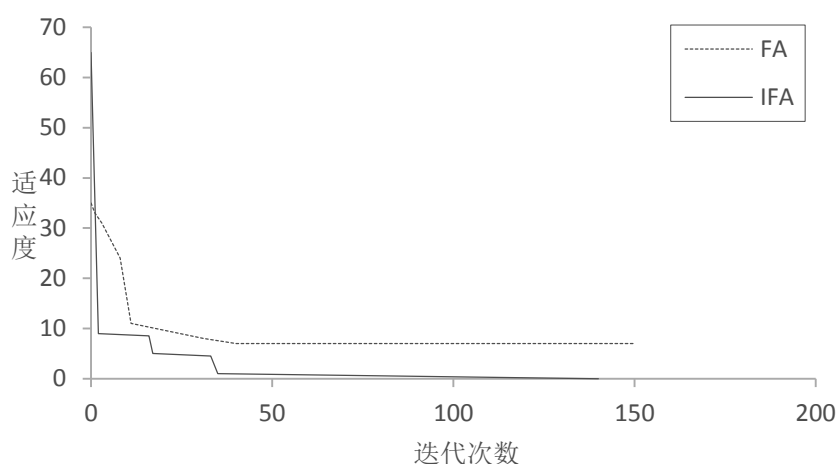


图 4-6 Rastrigin 函数的收敛曲线对比

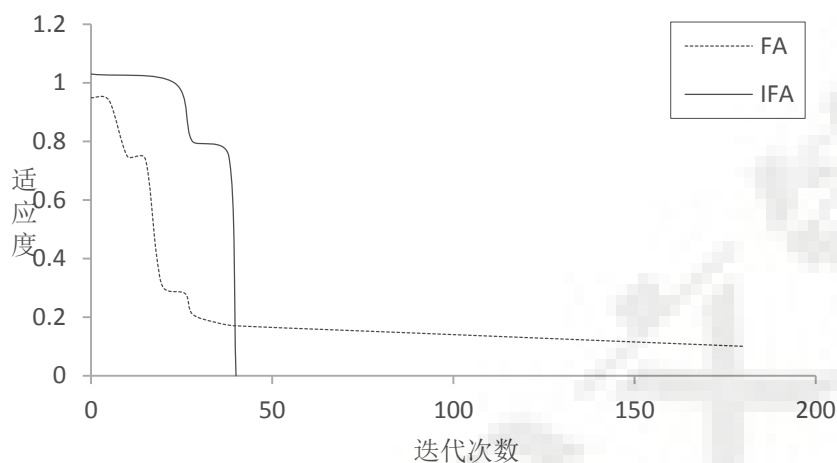


图 4-7 Griewank 函数的收敛曲线对比

4.3 混合核函数及其在决策优化中的应用

SVM 算法通过核函数将低维非线性不可分问题转为高维线性可分，并在一定程度上避免了“维数灾难”。目前常用的核函数包括：线性核函数、多项式核函数、RBF 核函数、sigmoid 核函数。研究成果表明，核函数参数 g 和惩罚因子 c 的选取对支持向量机分类准确率的影响很大，因此选取合适的核函数对 SVM 分类问题很重要^{[53][54]}。

统计学习理论指出：若函数满足 Mercer 条件，则该函数可作为 SVM 的核函数，则存在积分算子 $T_k : L^2(\Omega) \rightarrow L^2(\Omega)$ ，使得

$$(T_k f)(\cdot) = \int_{\Omega} K(\cdot, x) f(x) d_x$$

为正，其中 K 是连续对称函数， Ω 是有限维欧式空间中的有界闭集，即对

于任意的 $f \in L^2(\Omega)$ 可以得到

$$\int_{\Omega \times \Omega} K(z, x) f(x) f(z) d_x d_z \geq 0$$

这里的函数 $K(x_i, y_j)$ 就是核函数，其实现了样本从低维非线性到高维线性的转化。

核函数分为局部核函数和全局核函数。其中 RBF 核函数常用作局部核函数，而多项式核函数为常见的全局核函数。局部核函数的学习能力较强，但泛化能力较弱，而全局核函数泛化能力较强，而学习能力较差^{[55][56]}。本文构造了一个基于局部核函数和全局核函数的混合核函数，并通过调节比例因子来调节二者的比例，使之具备各自的优点，提高 SVM 分类准确率。混合核函数的构造形式如式 (4-11) 所示：

$$K_{mix} = \lambda K_g + (1 - \lambda) K_l \quad (4-11)$$

其中局部核函数 K_l 一般选取 RBF 核函数，全局核函数 K_g 一般选取多项式核函数。 λ 为比例因子，调节二者的比重使得分类最优，使得混合函数既有一定的学习能力，又有较强的泛化能力。通过实验发现，当 λ 的范围在 0.01-0.45 之间的时候，混合核函数的性能最好，既继承了局部核函数学习能力强的优点，又继承了全局核函数泛化能力强的优点。选取指定的多项式函数、RBF 函数及混合核函数，在同一测试点 ($x=10$) 处进行测试，对每个函数分别设置 $\gamma=1$ 、 $\gamma=2$ 、 $\gamma=5$ ，因此每个图中包含三条曲线。图 4-8、4-9、4-10 分别为 RBF 核函数、混合核函数及多项式核函数在同一测试点处的函数特性图。

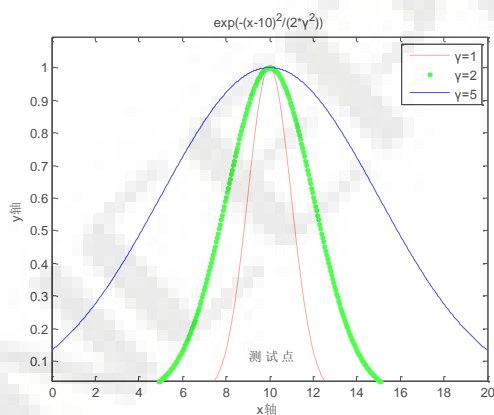


图 4-8 RBF 核函数

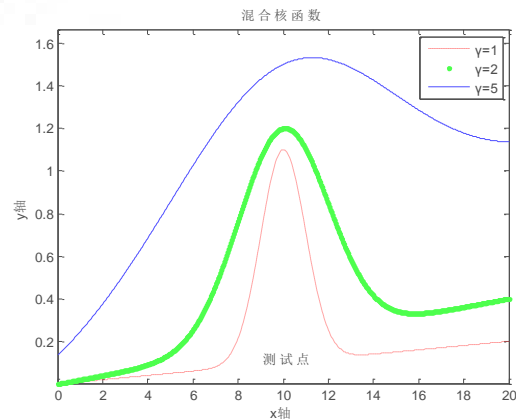


图 4-9 混合核函数

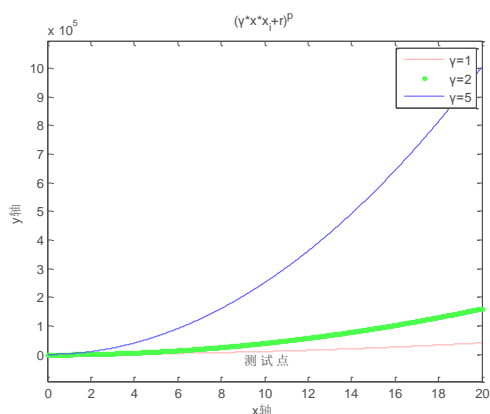


图 4-10 多项式核函数

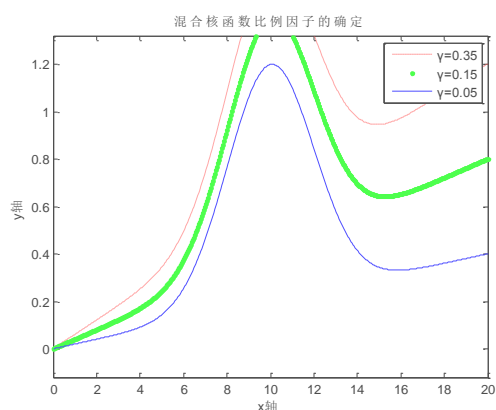


图 4-11 核函数比例因子变化图

结合图 4-8、4-9、4-10 可知，① γ 越大，函数的泛化能力越强；② 混合核函数不仅具有较好的全局泛化性，又具有局部核函数较强的学习能力的优点。

上述实验在一定程度上证明了混合核函数的优势，但如何确定混合核函数中的比例因子仍是一个亟待解决的问题。本文从核函数选取的各种情况综合考虑并通过实验对比，将 γ 设为固定，通过改变比例因子 λ 来观察函数特性曲线变化来确定合适的 λ 。通过查阅文献的值， λ 的范围在 0.01 至 0.45 时，混合核函数的性能最好。本次实验 λ 分别取 $\lambda = 0.35$ 、0.15、0.05，对应的实验结果如图 4-11 所示。

观察图 4-11 可知，当 λ 取值为 0.05 时，相对而言，混合核函数具有较强的泛化能力和学习能力，因此，本文中混合核函数的比例因子确定为 0.05。

随后，利用本章第 2 节中的实验数据测试采用混合核函数的 SVM 辅助决策分类算法的分类精确度，结果如图 4-12 所示。

单独使用某一核函数与采用混合核函数分类精确度的对比，如表 4-5 所示：

表 4-5 单一核函数与混合核函数对比表

核函数类型	分类精度(%)	支持向量率(%)	训练时间(s)
多项式核函数	69.87	2.47	0.65
RBF 核函数	75.83	31.25	0.87
混合核函数	80.83	11.84	0.93

由实验结果可知，采用混合核函数在一定程度上提高了分类的精确度，且支持向量率较低，但训练时间比单一核函数时间较长，分析及查阅资料可知，使用混合核函数虽能在一定程度上提高分类精度，但由于使用多个单一核函数和引入比例因子，从而加大了算法或计算的复杂程度，因此基于此考虑，当样

本量较小时，模型训练时间较短，可采用混合核函数；当样本数据量较大，在分类精度差别不大的情况下，模型训练时间过长会严重影响用户体验，此时，建议采用单一的 RBF 核函数即可。

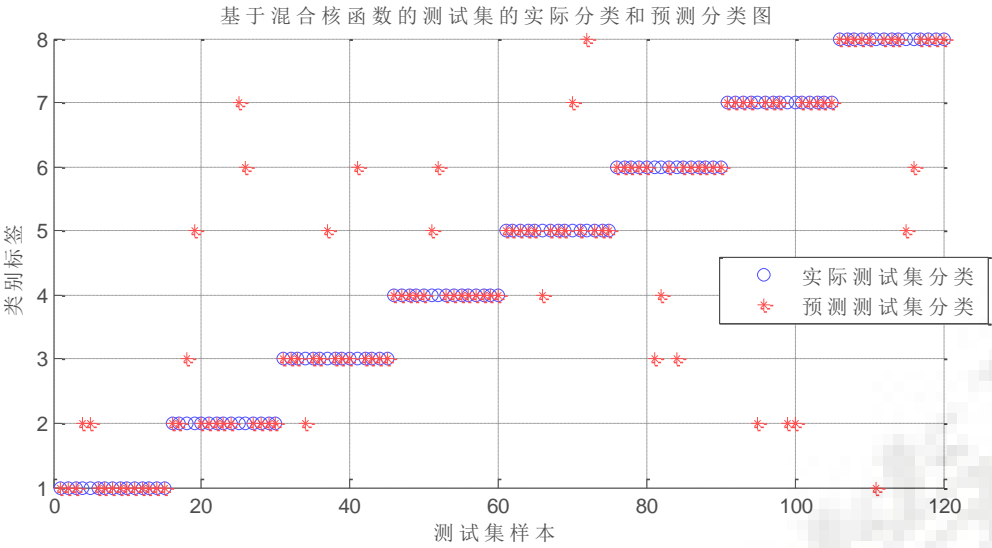


图 4-12 基于混合核函数的测试集的实际分类和预测分类图

4.4 本章小结

本章通过采用核函数和使用改进的萤火虫算法改进参数选择，从两个角度来提高智能家居辅助决策的准确性，并通过实验对两种方法的可行性及精确度进行了验证。

第5章 辅助决策系统设计与实现

5.1 家居环境信息采集

智能家居的整个辅助决策系统由传感器、开关、决策中心和执行器组成，其中，为了更精确的确定环境信息和防止误差过大，传感器分为室内传感器和室外传感器两组，且室内室外均包括两组传感器，室外传感器正对太阳光。而执行器有可控窗户、可控窗帘和可控照明灯开关三种，平面图各符号对应的元素及功能描述如图 5-1 所示。

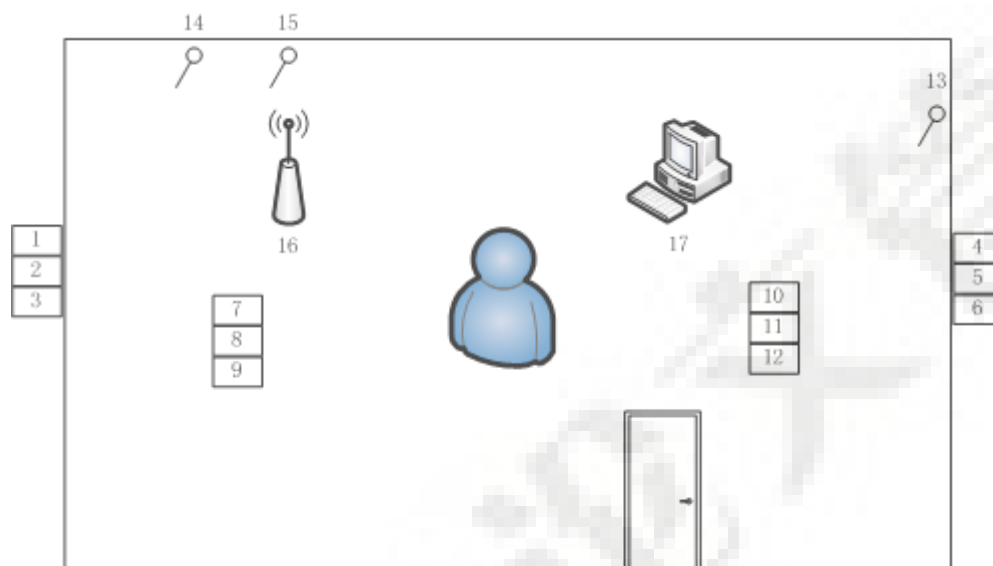


图 5-1 系统平面图

上图中编号 1、4 分别为室外的光照传感器；编号 2、5 分别为室外的温度传感器；编号 3、6 分别为室外的湿度传感器；编号 7、10 分别为室内的光照传感器；编号 8、11 分别为室内的温度传感器；编号 9、12 分别为室内的湿度传感器；编号 13 分别为无线控制的照明灯开关，该开关既可以接收来自基站的控制命令控制照明灯的开关，也可手动控制其打开和关闭；编号 14 为无线控制的窗户力臂开关，控制窗户的开闭；编号 15 为无线控制的窗帘力臂开关，控制窗帘的开闭；编号 16 为无线基站，既可接收来自无线传感器采集的数据，也可发送执行器的命令控制执行器的开关，其通过串口与辅助决策系统相连接，进行数据交互；编号 17 为辅助决策系统，通过串口与基站相连，进行数据交互，同时可与局域网中的用户进行数据交互。

5.2 硬件系统介绍

系统本地端通过 WIFI 和无线路由器搭建家居内部控制网络，并使用网关作为主控端。同时网关通过因特网与服务器通讯，在服务器上进行存储和机器学习等操作。辅助决策系统利用机器学习算法和采集的用户数据对用户行为进行分类预测，智能控制家居设备的开关。用户也可以通过手机等智能终端与服务器相接实现对家庭环境的监测和设备的控制，且用户的优先级高于辅助决策系统的优先级。

网关起的是一个枢纽和控制的作用，外部通过因特网连接到服务器，内部则通过工作在 AP 模式的 WIFI 模块与节点通信。其次，网关通过 Flash 存储器保存一些配置信息和数据信息。网关如图 5-2 所示：

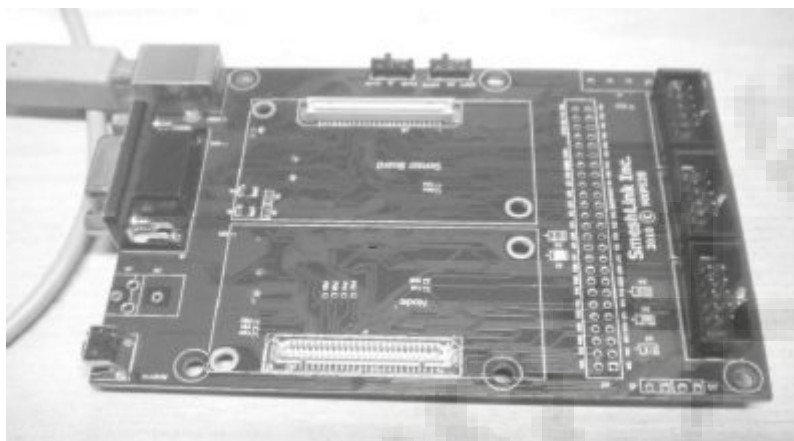


图 5-2 网关

由于节点需要与网关通信，所以采用了工作在 STA 模式的 WIFI 无线通信模块。另外，节点分为控制节点和传感器节点两类。其中控制节点通过 PWM、SPI 等信号提供驱动电器设备，传感器节点则通过相应通讯协议获取传感器数据。同时，利用 LCD 显示屏来显示一些设备相关的信息，并通过 EEPROM 存储器来存储配置信息。传感器如图 5-3 所示。



图 5-3 传感器节点

实验使用的无线执行器窗户可控开关、窗帘可控开关和照明控制开关如图所示,这些执行器都基于 ZigBee 技术,可以方便的通过各类终端对其进行控制。另外,智能家居决策系统运行在 PC 上,其操作系统为 windows7 64bit 版本。执行器如图 5-4 所示。



图 5-4 执行器

5.3 软件系统介绍

本文中的软件系统为基于 Myeclipse 等开发工具开发的智能家居 APP 系统,该系统包括后台接口及 APP 的开发。后台接口开发工具为 MyEclipse Professional 2014,还用到了 Spring、Mybatis 等框架, Spring 的版本为 spring-framework-4.0,持久层框架 Mybatis 的版本为 Mybatis 3,其较其他持久层框架具有更好的动态耦合性,其所涉及实现的主要功能为实现数据库各表中数据的增删改查及部分数据逻辑分析功能,用于支持用户在终端对环境数据的查看及对家居的远程操控。该 APP 开发工具为 Eclipse,其版本号为 eclipse MARS.2,配置 ADT、SDK 等开发包来配置开发环境,通过调用 myeclipse 等工具开发的后台接口实现与数据库及辅助决策系统中数据的通信,实时查看环境信息及操控家居状态。该手机 APP 系统的主要功能如图 5-5 所示,其系统类图如图 5-6 所示。

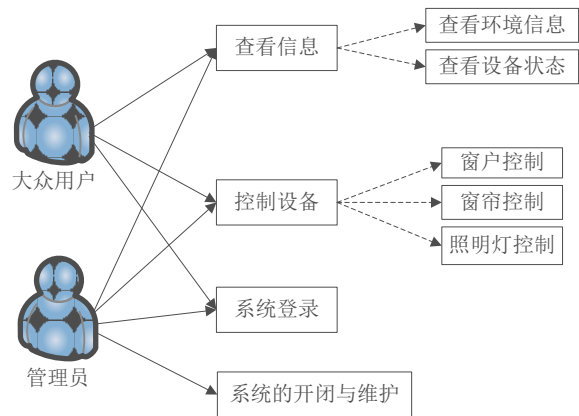


图 5-5 系统功能图

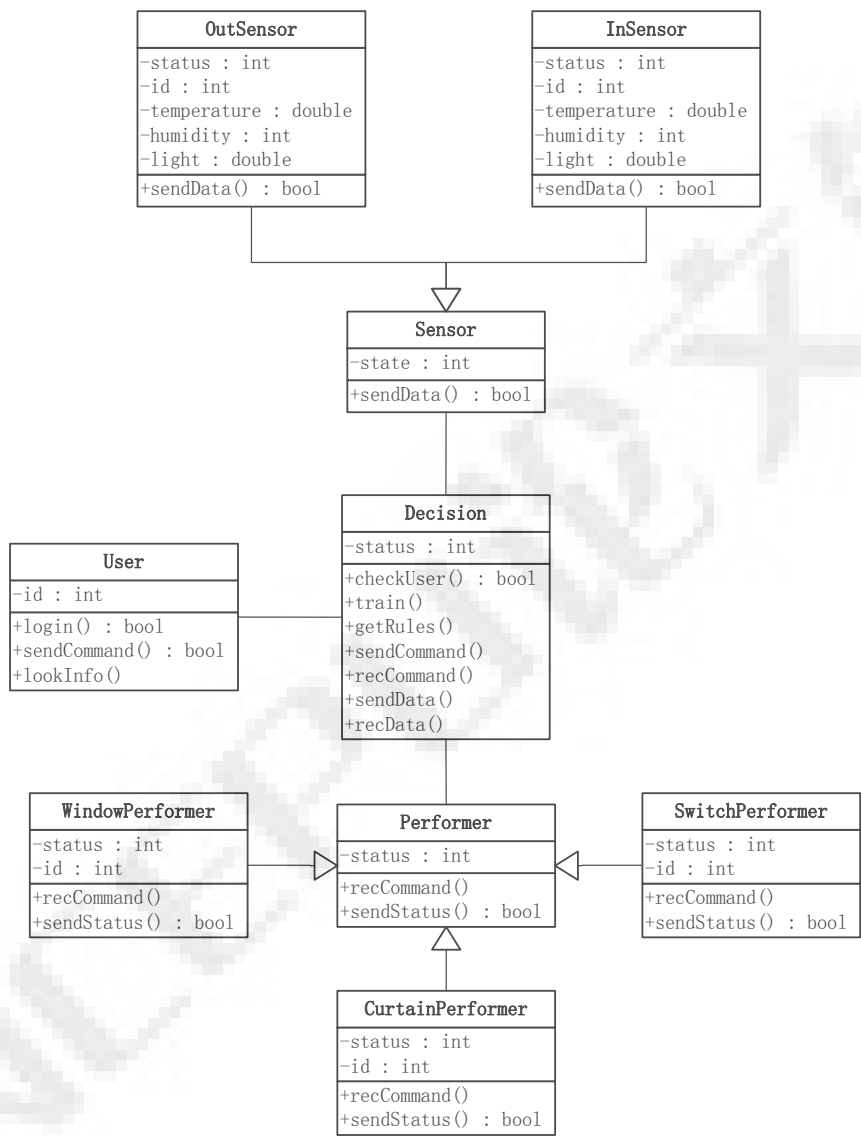


图 5-6 智能家居辅助决策系统类图

由图 5-5 可看出，系统用户分为大众用户和管理员，大众用户可实现系统的注册与登录，查看家居环境信息及传感器基本状态信息，如室内温度、光照、湿度等。还包括设备控制，如窗帘、窗户、照明灯的开闭等。管理员除具备上述功能外，还包括对系统的开闭及利用机器学习算法对系统进行训练等。

图 5-6 系统类图中，室内、室外传感器类继承自一般传感器类，其私有成员包括传感器状态、编号、温度、湿度、光照等，功能函数主要为发送数据等。窗户、窗帘、照明灯执行器类继承自一般执行器类，其私有成员包括执行器状态及编号，功能函数主要为命令的发送与接收等。用户类的私有成员包括用户编号等基本用户信息，功能函数包括登录、命令发送、信息查看（温度、湿度、光照）。决策类为该系统的核心类，其私有成员为用户状态，功能函数包括用户核验、用户规则集的获取、训练用户决策规则集、发送和接收数据、发送和接收命令等。

该 APP 主要包括三大模块：个人中心模块、传感器模块、系统设置模块。个人中心模块主要包括用户登录与注册、个人信息修改、家庭住址设置和个人传感器设备清单等。页面如图 5-7、5-8、5-9 所示：



图 5-7 登录页面



图 5-8 注册页面



图 5-9 我的家页面

传感器模块：主要包括设备注册和设备清单。设备注册为用户绑定自家已安装的传感器，然后用户即可在移动端实时查看家居环境信息。设备清单为该用户已绑定的所有传感器的清单，包括传感器名称、位置、编号及添加时间。若用户想要查看某条传感器的采集数据信息，包括历史数据查询等，可单击该项传感器进入相应页面进行查看，结果如图 5-10、5-11、5-12 所示。



图 5-10 温度展示图 图 5-11 数据记录页面 图 5-12 设备清单

系统设置模块功能主要包括消息与报警、开关寻线、操控音效、动作间隔等选项。其中消息与报警主要实现对温湿度、光照等环境因素的实时监控，为上述因素设置阈值，当超过或低于阈值时，手机通过震动或响铃方式提醒用户或智能决策控制执行器开闭，防止室内灾害的发生。开关寻线目的是实现用户通过手机远程操控智能家居设备状态的目的。部分页面展示如图 5-13、5-14、5-15 所示。



图 5-13 地址设置页面 图 5-14 消息与报警 图 5-15 设置

5.4 辅助决策系统准确度验证

实验中选取用户 A 和用户 B 的各 300 条真实数据，其中，数据选取时，为保证实验的可靠性，增强实验说服力，应分别从不同年龄段、不同地域、不同环境喜好的用户中分别选取一定数据，将通用决策系统、智能决策系统与本文

决策系统进行对比。其中通用决策系统决策规则为系统中大部分用户的决策规则或根据统计提前设定的决策规则集，其不需训练，只需个选取 50 条用作测试即可。智能决策系统各取两用户的 250 条数据作为训练集，各自剩余的 50 条数据作为测试集。本文决策系统分别对两个用户训练出两条决策规则，然后再各自测试，计算准确率，最后求二者的平均值作为本文决策系统的准确率。三种决策系统决策结果如表 5-1、图 5-16 所示所示。

表 5-1 三种决策系统决策结果

		第一次	第二次	第三次	第四次	第五次	平均
通用决策系统		0.62	0.65	0.61	0.59	0.63	0.62
智能决策系统	BP 神经网络	0.73	0.76	0.74	0.72	0.74	0.74
	SVM	0.71	0.72	0.73	0.73	0.71	0.72
本文决策系统	用户 A	0.84	0.82	0.86	0.83	0.84	0.838
	用户 B	0.86	0.87	0.85	0.86	0.88	0.864
	平均	0.85	0.845	0.855	0.845	0.86	0.851

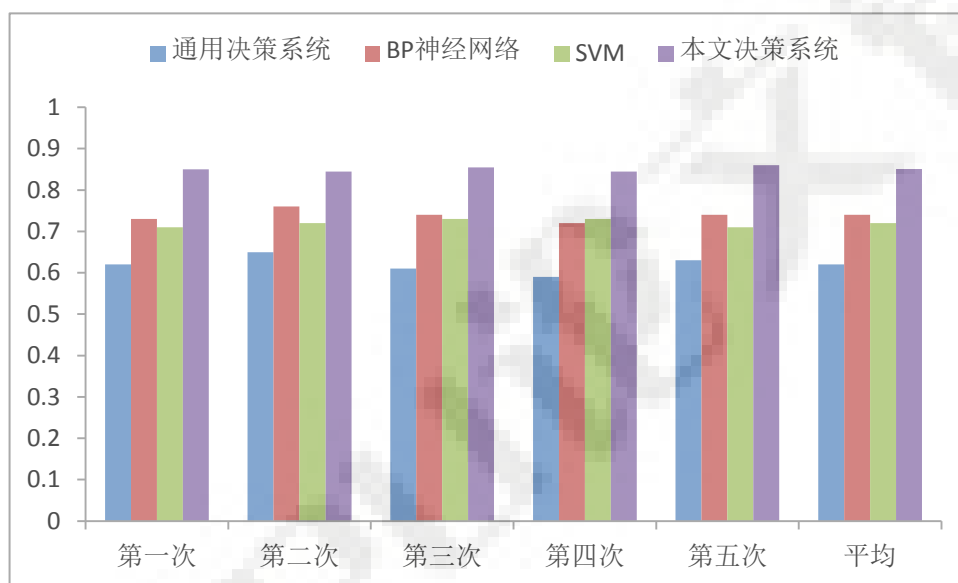


图 5-16 三种决策系统决策结果图

由表中数据看出，通用决策系统仅有 62% 的准确率，即三种执行器全部执行正确的概率只有 0.62，这就需要人们时常自动控制执行器的开闭，耗费太多精力。而引入机器学习算法的智能决策系统，通过对一段时间内个人行为的“学习”，决策准确率达到 70% 左右，大约提高了 10 个百分点，较通用系统具有更强的实用性。而本文决策算法的决策准确率达到 85% 左右，较智能决策系统又有了很大的进步，在于本文系统对环境对人环境的不同要求进行了区分，

训练不同环境要求的用户数据，产生多种类别的决策规则。而一般机器学习分类的准确率在 90% 以上，而本文中只有 80% 左右，说明人的行为具有一定的随意性，增加了研究的难度。也在一定程度上说明了在分类学习算法中，神经网络的主要问题是可解释性差，且当样本量较大时训练时间较长，而样本量较小时又可能遇到过学习问题，SVM 有效地改进了这些弱点。

5.5 本章小结

本章先后介绍了基于 SVM 的辅助决策系统的家居环境信息的采集方法及软硬件系统介绍，最后对本文中提出的决策算法与其他决策算法进行实验对比，证明了本文方法的可行性。

第 6 章 总结与展望

6.1 总结

本文对物联网、智能家居、辅助决策、SVM 等相关理论进行了实践和验证，将 SVM 算法应用于辅助决策系统中，设计了一个基于 SVM 的智能家居辅助决策系统，通过对环境数据的采集与训练，实现了对环境数据的实时感知和对家居设备的智能控制，并通过实验验证了本文方法的可行性，取得了如下结论：

- 1) 辅助决策系统采集的数据有时会存在异常，而异常数据对利用 SVM 算法训练模型有很大影响，异常数据过多会严重影响分类精度，因此，采取一定的算法对异常数据进行处理成为一项重要的内容。本文给出了一种基于相似度估值的异常数据处理算法，通过将该算法与其他算法进行实验对比，比较各自对异常项的估值与正常值的差异，本文算法取得了较好的效果，验证了本文方法在理论和实际应用中的可行性，为后文训练分类模型奠定了坚实的数据基础。
- 2) 文中通过理论与实验验证表明，SVM 能够很好的应用于智能家居的辅助决策系统中，通过将家居环境数据作为输入进行训练分类模型，该决策系统分析得出正确的执行状态来控制执行器的开闭，使得家居环境真正做到因人而异，同时也提升了智能家居的智能化程度。
- 3) 文中对提高辅助决策的精度进行了一定的尝试，并取得了较理想的效果。针对核函数对 SVM 算法分类精度存在较大影响，且单个核函数或学习能力不足，或泛化能力不足的问题，本文基于核函数的原理，提出了一种混合核函数的方法，该核函数同时具有了学习力强且泛化能力强的优点，同时进行了实验验证，通过在分类精度、支持向量率和训练时间等方面的比较，表明了本文提出的混合核函数的实用性，提高了辅助决策的精度。
- 4) 文中针对 SVM 参数对分类影响较大的问题，提出了一种改进的萤火虫算法来优化辅助决策算法。萤火虫算法具有模型简单、参数较少等优点，在 SVM 参数选择中具有很大优势，但由于 FA 算法中参数固定或设置不当等，导致算法过早收敛或无法收敛等问题。本文提出自适应步长的改进方法使得算法早期加速收敛，后期以较小的步长靠近最优值避免无法收敛的问题；同时，引入亮度公式和增强随机扰动改进位置更新公式中的扰动因子，加强算法的探索能力。最后通过标准测试函数对算法的可行性进行测试，实验表明这种改进提高了算

法的寻优精度和收敛速度，取得了较理想的结果。

5) 设计并实现了一个基于 SVM 的智能家居辅助决策原型系统，将改进后的系统与其他决策系统对比，验证了论文方法的有效性。

6.2 展望

本文利用 SVM 算法，实现了智能家居的辅助决策，即对窗户、窗帘、照明灯的智能控制，但在实际问题中，肯定会有更多的问题，本次实验只是对智能家居辅助决策系统研究的一次尝试，也需要不断的完善和优化，更好的提高其“智能化”程度。

1) 针对智能家居数据信息复杂、信息量大的特点，可将云计算或大数据技术应用到辅助决策系统中，大数据具有大量、高速、多样、价值的特点，将并行计算应用于辅助决策系统中可能会提高决策的速度等。

2) 从节能环保、室内灾害预测的角度考虑，当同一环境或相似环境面临多条决策时，应尽量减少能耗来达到节能的目的。其次，对火灾等室内灾害数据利用机器学习算法进行训练分析，得到相应的规律模型，对室内灾害进行提前预测。

3) 当室内同时出现多位用户且用户行为习惯差异较大时，应采取何种规则控制家居的状态，设想了或可通过为用户设置优先级或基于 K 近邻的多人投票的原则来选取合适的规则集等，有待于进一步研究。

参考文献

- [1] Kinney P. Zigbee technology: Wireless control that simply works[C]. Communications design conference. 2003, 2:1-7.
- [2] 陈海明,崔莉,谢开斌. 物联网体系结构与实现方法的比较研究[J]. 计算机学报,2013,01:168-188.
- [3] Wu G, Talwar S, Johnsson K, et al. M2M: From mobile to embedded internet[J]. Communications Magazine, IEEE, 2011, 49(4): 36-43.
- [4] 彭洪明. 智能家居的体系结构及关键技术研究[D].北京交通大学,2012.
- [5] 童晓渝,房秉毅,张云勇. 物联网智能家居发展分析[J]. 移动通信,2010,09:16-20.
- [6] 宋林,周媛,曾玉姗. 基于云计算和机器学习的智能家居系统设计[J]. 电子设计工程,2015,24:5-7.
- [7] 付韬. 基于 SVM 的 CPS 智能照明系统辅助感知决策[D].吉林大学,2013.
- [8] 陈飞彦. 以人为本的智能家居辅助决策系统的研究与实现[D].吉林大学,2015.
- [9] 李勇. 基于上下文感知计算的智能家居系统研究[D].重庆大学,2008.
- [10] Jiang L, Liu D Y, Yang B. Smart home research[C]. Proceedings of the Third Conference on Machine Learning and Cybernetics SHANGHAI. 2004: 659-664.
- [11] Michael P. Poland,Chris D. Nugent,Hui Wang,Liming Chen. Smart Home Research: Projects and Issues[J]. International Journal of Ambient Computing and Intelligence (IJACI),2009,14.
- [12] Suyang Zhou,Zhi Wu,Jianing Li,Xiao-ping Zhang. Real-time Energy Control Approach for Smart Home Energy Management System[J]. Electric Power Components and Systems,2014,423-4.
- [13] Cheng Z, Tan Y, Lim A O. Fitting method for hybrid temperature control in smart home environment[C]//Modelling, Identification & Control (ICMIC), 2014 Proceedings of the 6th International Conference on. IEEE, 2014: 300-305.
- [14] Behera A R, Devi J, Mishra D S. A comparative study and implementation of real time home automation system[C]//Energy Systems and Applications, 2015 International Conference on. IEEE, 2015: 28-33.
- [15] Vanus J, Martinek R, Bilik P, et al. New method for accurate prediction of CO₂ in the Smart Home[C]//Instrumentation and Measurement Technology Conference Proceedings (I2MTC), 2016 IEEE International. IEEE, 2016: 1-5.
- [16] Wang Y, Dasgupta P. Intelligent and adaptive temperature control for large-scale buildings and homes[C]//2016 IEEE 13th International Conference on Networking, Sensing, and Control (ICNSC). IEEE, 2016: 1-6.

- [17]Schweizer D, Zehnder M, Wache H, et al. Using Consumer Behavior Data to Reduce Energy Consumption in Smart Homes: Applying Machine Learning to Save Energy without Lowering Comfort of Inhabitants[C]//2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA). IEEE, 2015: 1123-1129.
- [18]Hausmann M D, Ziekow H. The Potential of Household Specific Feature Selection for Analysing Smart Home Time-Series Data[C]//Smart SysTech 2016; European Conference on Smart Objects, Systems and Technologies; Proceedings of. VDE, 2016: 1-6.
- [19]许永伟. 基于 SNS 智能家居系统的用户意图推理研究[D].北京邮电大学,2015.
- [20]付韬. 基于 SVM 的 CPS 智能照明系统辅助感知决策[D].吉林大学,2013.
- [21]赵亚琴. 基于模糊神经网络的火灾识别算法[J]. 计算机仿真,2015,02:369-373.
- [22]刘玉佳. 基于多信息融合的室内智能控制系统研究与实现[D].河北工业大学,2014.
- [23]宫辰. 基于云计算的智能家居管理系统设计与应用[D].南京邮电大学,2014.
- [24]岳敬华. 基于云服务的智能家居系统的研究与设计[D].杭州电子科技大学,2014.
- [25]R Harper. Inside the smart home[M]. London: Springer, 2003.
- [26]师洪涛,杨静玲,丁茂生,王金梅. 基于小波—BP 神经网络的短期风电功率预测方法[J]. 电力系统自动化,2011,16:44-48.
- [27]张学工. 关于统计学习理论与支持向量机[J]. 自动化学报,2000,01:36-46.
- [28]李红英. 支持向量分类机的核函数研究[D].重庆大学,2009.
- [29]Le Thi H A, Nguyen M C. Efficient Algorithms for Feature Selection in Multi-classSupport Vector Machine[M]//Advanced Computational Methods for KnowledgeEngineering. Springer International Publishing, 2013: 41-52.
- [30]郭创新,朱承治,张琳,彭明伟,刘毅. 应用多分类多核学习支持向量机的变压器故障诊断方法[J]. 中国电机工程学报,2010,13:128-134.
- [31]杨雪. 支持向量机多类分类方法的研究[D].哈尔滨工程大学,2006.
- [32]杨静,张楠男,李建,刘延明,梁美红. 决策树算法的研究与应用[J]. 计算机技术与发展,2010,02:114-116+120.
- [33]闭小梅,闭瑞华. KNN 算法综述[J]. 科技创新导报,2009,14:31.
- [34]崔江,王友仁. 基于支持向量机与最近邻分类器的模拟电路故障诊断新策略[J]. 仪器仪表学报,2010,01:45-50.
- [35]胡俊. 支持向量机参数优化问题的研究[D].哈尔滨工业大学,2009.
- [36]徐晓明. SVM 参数寻优及其在分类中的应用[D].大连海事大学,2014.
- [37]Ren Y, Bai G. Determination of optimal SVM parameters by using GA/PSO[J]. Journal of Computers, 2010, 5(8): 1160-1168.

- [38]Zhang X L, Chen X F, He Z J. An ACO-based algorithm for parameter optimization of support vector machines[J]. Expert Systems with Applications, 2010, 37(9):6618-6628.
- [39]丁勇,秦晓明,何寒晖. 支持向量机的参数优化及其文本分类中的应用[J]. 计算机仿真,2010,11:187-190.
- [40]李兵,姚全珠,罗作民,田元,王伟. 基于网格模式搜索的支持向量机模型选择[J]. 计算机工程与应用,2008,15:136-138.
- [41]马永杰,云文霞. 遗传算法研究进展[J]. 计算机应用研究,2012,04:1201-1206+1210.
- [42]黄太安,生佳根,徐红洋,黄泽峰. 一种改进的简化粒子群算法[J]. 计算机仿真,2013,02:327-330+335.
- [43]李晓菲. 数据预处理算法的研究与应用[D].西南交通大学,2006.
- [44]刘长平,叶春明. 一种新颖的仿生群智能优化算法:萤火虫算法[J]. 计算机应用研究,2011,09:3295-3297.
- [45]高伟明. 萤火虫算法的研究与应用[D].兰州大学,2013.
- [46]Feng Yuan,Shouqiang Chen,Hong Liu. Association Rules Mining on Heart Failure Differential Treatment Based on the Improved Firefly Algorithm[J]. Journal of Computers,2014,94.
- [47]Shuhao Yu,Shanlin Yang,Shoubao Su,Sabri Arik. Self-Adaptive Step Firefly Algorithm[J]. Journal of Applied Mathematics,2013,2013.
- [48]王迎菊,周永权. 一种基于荧光素扩散的人工萤火虫算法[J]. 计算机工程与应用,2012,10:34-38.
- [49]Mohammad Khajehzadeh,Mohd Raihan Taha,Mahdiyeh Eslami. A New Hybrid Firefly Algorithm for Foundation Optimization[J]. National Academy Science Letters,2013,363.
- [50]程美英,倪志伟,朱旭辉. 萤火虫优化算法理论研究综述[J]. 计算机科学,2015,04:19-24.
- [51]王吉权,王福林. 萤火虫算法的改进分析及应用[J]. 计算机应用,2014,09:2552-2556.
- [52]Krishnamoorthi Murugasamy. Hybrid Clustering Using Firefly Optimization and Fuzzy C-Means Algorithm[J]. Circuits and Systems,2016,0709.
- [53]丁世飞,齐丙娟,谭红艳. 支持向量机理论与算法研究综述[J]. 电子科技大学学报,2011,01:2-10.
- [54]郑小霞,钱锋. 高斯核支持向量机分类和模型参数选择研究[J]. 计算机工程与应用,2006,01:77-79.
- [55]刘明,周水生,吴慧. 一种新的混合核函数支持向量机[J]. 计算机应用,2009,(S2):167-168+206.
- [56]任彧,梅盛鑫. 基于混合核函数支持向量机和遗传算法的人脸识别[J]. 计算机应用与软件,2011,(04):260-263.

攻读硕士学位期间发表的论文及其他成果

(一) 发表的学术论文

- [1] 一种智能家居系统中的室温曲线自动判定方法[J]. 电脑编程技巧与维护, 2016, 16:85-87. 第一作者

(二) 申请的软著

- [1] 《知识问答系统 V1.0》
[2] 《智能家居 APPV1.0》
[3] 《电力用户停电管理系统 V1.0》

(三) 参与的科研项目

- [1] 数字化煤场网站开发
[2] 智能家居 APP 开发
[3] 康宝医生端 APP 优化

致谢

光阴似箭，日月如梭，不知不觉中，研究生生涯所剩无几。两年多的学习和生活，不仅使我学到了很多的专业知识，也学到了很多为人处世的道理，收获了很多的人生感悟。两年多的时间里，也曾沮丧、也曾迷茫，高兴过、悲伤过，但当经历过后，感觉人生阅历又上了一个台阶，内心变得更加坚定和充实，同时，也培养了我坚韧不拔的毅力和勇于尝试、探索的科研精神，开阔了视野，增长了专业知识，研究生期间收获颇多。

首先感谢我的导师袁和金老师，在攻读硕士期间，老师渊博的知识、严谨的态度和强烈的敬业精神时时刻刻影响着我，为我们创造了良好的学习和科研环境。袁老师平时态度和蔼、平易近人，在忙碌的时候也经常督促和指导我们，特别是在论文书写过程中，时时督促与指导，给出了许多宝贵的意见，使我取得了很大进步。今后，我将牢记袁老师的教诲，努力奋发，在此，再次致以崇高的敬意和真挚的感谢。

同样感谢实验室的小伙伴们，感谢他们在平时科研中给出的指导和意见，感谢他们在求职期间给出的意见和看法，给予我的无私的鼓励和支持。是他们让我在求职期间不再迷茫，也是他们不断督促和激励我在学习和科研的道路上不断进步。也同时感谢实验室的人们，感谢他们在日常生活中给予我的快乐，使实验室能够有好的科研氛围。感谢这两年来一直陪伴在我身边的朋友们，感谢他们在生活、学习、求职过程中给予的诸多帮助、鼓励和陪伴，感谢他们给予了我研究生期间那么多值得回忆的快乐时光。

最后还要感谢把我抚养成人的父母，他们总是毫无怨言的尽一切可能满足我在求学路上的所有要求，感谢他们对我一直的呵护，在今后日子里，我会更加努力，希望将来有能力报答父母的养育之恩。

研究生生涯即将画上了句号，感觉人生总是聚散匆匆，还没来得及好好珍惜却要面临着匆匆话别，想起心里难免有难过与不舍，在余下来的短短时光里，更应加倍珍惜。面对即将到来的离别，对我们来说，既是终点，也是起点，即将结束求学生活，马上步入社会中去，总是充满了很多期待，也有些迷茫，唯有脚踏实地，用心工作、生活，积极去迎接未知的挑战，努力实现自身价值的同时，也争取为社会尽一份绵薄之力。