分类号: 密级:

UDC:

学 位 论 文

电力交易平台推荐系统的研究与移动端的实现

|  |  |  |  |
| --- | --- | --- | --- |
| 作者姓名： | 徐振康 | | |
| 指导教师： | 焦明海 副教授 | | |
|  | 东北大学计算机科学与工程学院 | | |
| 申请学位级别： | 硕士 | 学科类别： | 专业学位 |
| 学科专业名称： | 计算机技术 | | |
| 论文提交日期： |  | 论文答辩日期： |  |
| 学位授予日期： |  | 答辩委员会主席： |  |
| 评阅人： |  | | |

东北大学

2016年12月

**A Thesis in Computer Technology**

**The Research and Implementation of Electric Trading Recommend System and Mobile Terminal**

By Feng Xue

Supervisor: Associate Professor Jiao Minghai

**Northeastern University**

**Dec 2016**

独创性声明

本人声明，所呈交的学位论文是在导师的指导下完成的。论文中取得的研究成果除加以标注和致谢的地方外，不包含其他人己经发表或撰写过的研究成果，也不包括本人为获得其他学位而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示谢意。

学位论文作者签名：

日 期：

学位论文版权使用授权书

本学位论文作者和指导教师完全了解东北大学有关保留、使用学位论文的规定：即学校有权保留并向国家有关部门或机构送交论文的复印件和磁盘，允许论文被查阅和借阅。本人同意东北大学可以将学位论文的全部或部分内容编入有关数据库进行检索、交流。

作者和导师同意网上交流的时间为作者获得学位后：

半年□一年□一年半□两年□

学位论文作者签名： 导师签名：

签字日期： 签字日期:

# 摘 要

推荐系统是一种能快速帮助用户找到增值信息的工具，通过分析用户的历史行为信息给用户的兴趣建模，从而根据用户的兴趣模型主动向用户推荐满足他们需求的事物，挖掘出潜在的增值信息。系统将提供给用户的兴趣事物推荐列表，或者是预测用户对每个推荐事物的喜爱程度，这些系统帮助用户决定在事物集合中选择适当的事物，快捷推出首选事物。到目前为止，推荐系统是互联网以及电子商务应用中关键技术之一，已经在互联网领域广泛应用。

当前，电力领域进入深化改革阶段，电力改革中最先启动的部分是电力市场交易模式，即电力交易向全面市场化过渡，允许大企业用户直接与发电企业进行电力交易。大用户直购电的意义是对现有电力销售机制的探索，引入大用户直购电模式可以使电力市场同时出现多个卖家（发电企业）和多个买家(企业用户)。由此背景，本论文设计一个基于企业用户侧的电力需求特征的协同过滤推荐算法的电力交易推荐系统，将符合需求的发电企业推荐给购电企业用户。目前，在电力工业领域电力交易推荐移动平台尚属研究热点，本论文研究目的旨在通过开发推荐平台原型，针对企业用户与售电企业的双向推荐，同时对电力市场成员推荐相关的电力交易量。根据以上分析，本文研究并实现如下内容：

(1) 实现粒子群（PSO）优化BP神经网络算法。算法用于预测发电企业的阶段电力负荷，并将预测信息发布到电力交易移动平台，根据预测的电力负荷，对用电企业用户与发电企业之间进行供需电量平衡的匹配。

(2) 提出电力交易用户评分推荐算法。传统的电力交易模式下，用户很难获得企业用户对发电企业的偏好信息。基于移动端电力交易模式，购电企业用户可以自主选择发电企业。本文将用户与发电企业的交易次数设定为用户对发电企业的评分数据，由于传统推荐存在着评分数据稀疏问题，本文提出了基于用户需求属性相似度的方法实现评分矩阵填充，填充的评分矩阵针对电力市场直购电新增用户进行推荐，最终将推荐列表展示给购电大用户，将满足大企业用户需求的Top-K个发电企业推荐给大用户。实现售购双方供需利润互赢。

(3)设计基于Android的电力市场移动交易平台原型。搭建并实现了电力交易移动平台，实现了公告发布模块、用户（发电企业和大企业用户）信息发布模块、交易模块、推荐模块等，研究的电力负荷预测算法与推荐算法在平台逻辑层编程实现，成为原型系统的理论方法支撑。

关键词：电力负荷预测；协同过滤推荐；矩阵填充；用户相似度；交易移动端

# Abstract

Recommendation system is a tool to help users find the value-added information, through the analysis of user behavior history information, it is active to recommend users to meet their needs according to the user model, whichdigs out the potential value of information. The system will provide users interested in things or predict the recommendation list, like user recommendation for each of these things, the system can help users in selecting the appropriate set of things in things, also quickly launch of preferred things. So far, the recommendsystem is one of the key technologies in the internet and e-commerce applications, widely used in the internet field.

At present, the electric field enters the stage of deepening reform, the reform of electric power first launch is the electric power market transaction mode, namely power trading to market transition, allowing large enterprise users direct power exchange and power generation enterprises. The significance of direct business of large consumers is to explore the existing power sales mechanism, the direct trading mode can make the electricity market at the same time with a number of sellers (power generation enterprises) and with a number of buyers (business users). From the background, this paper designs a power trade recommendation system based on the collaborative filtering recommendation algorithm based on the power demand characteristics of enterprise users. And now, the field of power industry recommended mobile power trading platform is still a hot topic, the purpose of this study aims to develop a recommended platform prototype, recommended for two-way business users and the sale of electricity companies, while the recommended volume related to electric power market members. According to the above analysis, this paper studies and realizes the following contents:

(1)Particle swarm optimization (PSO) optimized BP neural network algorithm. The algorithm is used to predict the different phase power load of power generation enterprises, and the prediction power trading information released to the mobile platform.According to the power load forecasting of electricity, supply and demand balance of power is matched between enterprise users and power generation companies.

(2)Power transaction user rating recommendation algorithm is secondly presented in paper. Under the traditional electricity trading mode, it is difficult for users to obtain the preference information of enterprise users to power generation enterprises. Based on the mobile terminal electricity trading model, the electricity users can choose the power generation enterprises. And the number of transactions from the users and power companies is set to the score data of users of power enterprises.Because the score data sparseness problem is needed to research by the traditional recommendation, this paper puts forward a method of similarity score matrix of user demand attribute based on filling with direct purchase of electricity power market, recommendation is made to new users,also is recommendation list shown to purchase users. Top-K will meet a power enterprise user needs of large enterprises to recommend to users. Achieve the purchase and sale of supply and demand sides win each other.

(3)Design of mobile trading platform prototype based on Android open system. Set up and implement the power trading mobile platform, the announcement module, user (power generation enterprises and large enterprise users) information dissemination module, transaction module, recommendation module, algorithm and recommendation algorithm is implemented in the platform of programming logic layer power load forecast, as supporting for the theory and method of prototype system.

**Keywords:** power load forecasting; collaborative filtering; matrix completion; user similarity; trading mobile terminal

目 录

[独创性声明 I](#_Toc470247370)

[摘 要 II](#_Toc470247371)

[Abstract IV](#_Toc470247372)

[第1章 绪 论 1](#_Toc470247373)

[1.1 研究背景 1](#_Toc470247374)

[1.2 研究意义 2](#_Toc470247375)

[1.3 国内外研究现状 3](#_Toc470247376)

[1.3.1 国内外电力市场交易模式 3](#_Toc470247377)

[1.3.2 电力预测研究现状 4](#_Toc470247378)

[1.3.3 推荐系统研究现状 6](#_Toc470247379)

[1.4 论文章节安排 7](#_Toc470247380)

[第2章 相关技术介绍 9](#_Toc470247381)

[2.1 推荐系统相关技术 9](#_Toc470247382)

[2.2 移动端推荐系统 12](#_Toc470247383)

[2.3 相关预测算法的改进 13](#_Toc470247384)

[2.4 本章小结 15](#_Toc470247385)

[第3章 基于用户相似度的协同过滤推荐算法 17](#_Toc470247386)

[3.1 电力交易推荐系统模型 17](#_Toc470247387)

[3.1.1 基于电力交易推荐系统拟解决的问题 17](#_Toc470247388)

[3.1.2 电力交易推荐系统的架构模型 19](#_Toc470247389)

[3.2 基于用户相似度的矩阵填充 20](#_Toc470247390)

[3.2.1 问题描述 20](#_Toc470247391)

[3.2.1 用户相似度计算 21](#_Toc470247392)

[3.2.3电力交易稀疏矩阵的填充 24](#_Toc470247393)

[3.3 基于协同过滤算法的发电企业推荐 27](#_Toc470247394)

[3.3.1 电力交易系统新增用户推荐集的产生 27](#_Toc470247395)

[3.3.2 电力交易系统已交易用户的推荐集的产生 29](#_Toc470247396)

[3.4 实验结果及分析 30](#_Toc470247397)

[3.4.1评分数据标准化及实验数据获取 30](#_Toc470247398)

[3.4.2 实验评价标准 31](#_Toc470247399)

[3.4.3 系统参数的确定 32](#_Toc470247400)

[3.4.4 推荐系统准确率与召回率 34](#_Toc470247401)

[3.5 本章小结 35](#_Toc470247402)

[第4章 基于粒子群神经网络电力负荷预测 37](#_Toc470247403)

[4.1 BP神经网络算法 37](#_Toc470247404)

[4.1.1 BP神经网络模型 37](#_Toc470247405)

[4.1.2 BP神经网络算法 38](#_Toc470247406)

[4.2 PSO优化BP神经网络算法 39](#_Toc470247407)

[4.2.1 BP神经网络算法的缺点 40](#_Toc470247408)

[4.2.2 粒子群基本算法 41](#_Toc470247409)

[4.2.3 粒子群优化神经网络算法步骤 42](#_Toc470247410)

[4.3 电力负荷预测实验结果分析 44](#_Toc470247411)

[4.3.1 数据集选取 45](#_Toc470247412)

[4.3.2 数据规范化 45](#_Toc470247413)

[4.3.3 BP神经网络的结构设计与参数设置 46](#_Toc470247414)

[4.3.4 PSO-BP预测电力负荷实验结果 47](#_Toc470247415)

[4.4 本章小结 48](#_Toc470247416)

[第5章 电力交易推荐系统的移动端实现 49](#_Toc470247417)

[5.1 系统架构设计 49](#_Toc470247418)

[5.2 电力推荐系统的移动端的具体实现 51](#_Toc470247419)

[5.2.1 系统架构层次设计与实现 51](#_Toc470247420)

[5.2.2 系统数据库的设计与实现 52](#_Toc470247421)

[5.2.3 系统功能模块的设计与实现 53](#_Toc470247422)

[5.2.4系统界面设计与实现 54](#_Toc470247423)

[5.3 本章小结 54](#_Toc470247424)

[第6章 总结与展望 57](#_Toc470247425)

[6.1 总结 57](#_Toc470247426)

[6.2 展望 57](#_Toc470247427)

[参考文献 59](#_Toc470247428)

[致 谢 63](#_Toc470247429)

[攻读硕士学位期间的论文和项目情况 65](#_Toc470247430)

# 第1章 绪 论

## 1.1 研究背景

推荐系统是由几个相关的学科领域为支撑而组成的，例如近似理论、认知科学和信息检索等相关学科，但是在90年代中期，由于推荐的发展和重要性，推荐系统已经逐渐演变成为一个独立的研究领域[1,2,3]。目前，由于通讯网络的快速发展和便捷，使信息能够快速且容易获取。但随之带来的“信息过载”也成为亟待解决的问题。比如说如果一个消费者想购买数码相机，在做出购买的决定之前，需要阅读和比较所有对数码相机的在线评论，这会是一个令人沮丧的经历。那么通过推荐系统自动处理这些过载的信息，推荐给可能符合用户的兴趣的事物。准确的推荐可以使用户能够快速地找到理想的物品而不被无关的信息干扰，与此同时，供应商推荐这些用户感兴趣的匹配产品，希望每个访问网站的用户满意并作为回头客带来更大的利益。这也不难理解，Netflix比赛中,推荐精度提高了10%的参赛者被授予100万美元。

电力市场的一个基本职能是通过市场来调节电力生产经营活动，以达到并实现电力资源的优质合理利用。几十年来，传统的电力市场交易模式一直保持“自然垄断行业”状态，在未来的电力市场竞争中，打破垄断，放开竞争，是必然的发展趋势。目前，官方渠道发布的关于电力改革文件[4]中突出强调了电力体制改革的重要性和紧迫性：从根本上改变传统电力市场中发电厂与电网一体，政企不分的状态，争取形成电力市场自由竞争的多元化格局。在电力改革文件中，最引人瞩目的一个内容是“管住中间、放开两侧”，即电力交易模式的改变(如电力的市场化交易)，这可能是整个电力改革中最先启动的部分。传统电力市场中，发电企业并不能将电力直接传输给用户，而是需要国家电网来过渡。如今新电力改革方案启动后，发电企业如果能有售电牌照在手，就可直接实现点对点交易，售电主体和用户被赋予了自主选择权，他们可以相互自主确定双方交易的用户、双方交易的电量和双方交易的价格，其中，过网费的多少是按照国家的关于电价的输配方案的规定来决定的，最终向电网企业支付，最终直接洽谈合同，实现多方直接交易，比如短期的交易或者即时的交易，然后通过调度和交易机构实现，为工商业用户和企业用户以及其他等各类用户提供性价比更高、更加经济、更加优质的电力保障。改革后，电网企业的收入来源不再是以上网电价和销售电价价差为主要的收入，而是通过政府核实确定的输配电价来收取过网费，从而确保电网企业收入来源和收益水平达到稳定的状态，以规范电网企业资产管理和投资行为。电力系统负荷预测和电价预测是在充分考虑现有的条件下，研究和利用一套系统的方法，在满足一定精度要求的意义下，确定某个特定时刻的负荷数值与电价范围。随着对电力负荷预测和电价预测的需求的增加，它对电力系统的相关部门的重要性不言而喻。电力负荷预测在时间周期上包括中长期电力负荷预测和短期电力负荷预测两种类型，其中中长期预测又分为月度预测和年度预测两种。这两种时长的预测是电力系统规划部署的重要工作内容，使发电厂在宏观发展上有一个重要依据。短期电力预测一般是对周，天，数小时的电力负荷变化情况进行合理的估计，从而能够准确制作出发电计划和定价功能。电力负荷预测的准确性是发电企业科学合理规划的基础和重要指标，与此同时，由于电网不断的大规模扩展和电力市场目前的改革体制，要求电力负荷预测的准确度更加准确，加上电力数据的大规模累积必然会形成数据量大，以及影响因素繁多等特点，因而对电力负荷准确且高效的预测成为了电力领域的一个热点问题。国内外的学者长期以来对电力预测做了大量的研究，提出过许多关于电力负荷预测的模型。但是这些模型在准确度方面不尽人意，主要因为电力预测受到多种因素的影响，如天气，国家的经济，大型企业的产能等。

如今，随着互联网与大数据时代的到来，推荐系统也被广泛应用于电子商务、图书推荐、音乐推荐等领域，不仅给用户免去在互联网中的海量数据里寻找满意的商品的烦恼，而且给商品提供者带来更多的经济利益。针对电力市场交易模式的改变，并结合以上改革内容以及导师电力相关项目，本论文首先是根据电力改革中交易方式的改变开发一个移动平台以服务发电厂和大型用户在该平台进行交易，并在该移动平台实现电力交易推荐系统。电力交易业务是电力市场交易的核心业务，电力移动平台交易中心根据预测发电厂市场需求，利用市场实际供需情况组织双方市场成员开展交易。

## 1.2 研究意义

在当前互联网技术以及市场竞争的大趋势下，各行各业的交易模式发生变化，开始将线下传统交易发展至网上交易，电力改革前，发电厂只是负责生产电，国家电网是电力交易的枢纽，发电厂和用户之间的交易需要通过国家电网参与，因此，随着电力的改革，电力交易方式也在改变，电是市场上的一种特殊商品，在经济上不可储藏，因此需要在生产者和消费者之间保持发电量的平衡稳定。设计一个直接交易的网上平台对电力点对点式直接交易模式至关重要。同时，将现如今的推荐系统应用到电力行业，对交易双方都是有利而无害的。结合以上内容，本文旨在电力交易方式的改变下，通过更精确的预测算法预测出发电企业下一个阶段的发电量，并提出基于用户侧属性特征的协同过滤推荐算法，将符合大企业用户需求的发电厂推荐给用户，为电力市场提供一个较好的移动平台进行双方交易匹配，实现移动平台上的交易结算。该平台对发电企业以及大用户双方都是有利而无弊的，促进双方的合作和经济的发展。

除此之外，电力市场中，火力发电行业与水力发电、核电和其他能源发电并列属于发电行业，中国的电力市场依旧以火力发电为主要供电主体，火力发电厂主要以煤炭为燃料运作。经调查显示，截止到2020年，非化石能源占一次能源损耗比重达到15%，天然气的比重可以到达10%以上，煤炭损耗的比重至少要节制在62%范围内。同时，目前存在的“弃风”、“弃光”等抛弃能源现象，对电网企业，有电网建设上的难题，也有经济性的考虑。因为电网企业过去是以赚取差价作为主要盈利手段，所以会更加倾向于接纳低价的火电、水电等。相比较之下，风电、光伏发电等价格更贵，对电网冲击更大。若输配电价改革落实到位，火力发电的低价格将不再成为一项很强的竞争力，要优先安排风能、太阳能、生物质能等可再生能源保障性发电，热电联产及燃气发电、水电、核电兼顾资源或调峰等需求安排发电，高效节能、超低排放煤电机组被放在最后。虽然有观点认为，这一顺序根据各地实际情况，具体实施时可能存在较大出入，但已经可以看出国家对于燃煤发电机组的定位愈发明确。因此，本论文旨在通过电力市场交易方式的改变，通过主动向大企业用户推荐发电厂，促进新能源发电企业的交易比重，既能促进经济发展，又可以优化能源使用率。

## 1.3 国内外研究现状

### 1.3.1 国内外电力市场交易模式

电力行业被认为是垄断行业[5]，因为电力行业有单一的生产者承担，其目的是可以利用规模的经济模式达到提高效率、节约资源。在第二次世界大战结束后，因为经济的发展和科学技术的进步，电力行业也加快了发展的步伐，在20世纪下半叶许多国家电力工业的基本结构以垂直一体化国有垄断模式存在[6]，如在1946年和1947年，法国电力公司和英国中央发电局分别成立了国家范围内的垄断一体化的国有大型电力企业，美国虽然没有直接的在全国范围内成立电力国有化和垄断集中化，但是也在企业间进行了较大范围的合并。但是，电力行业的垄断经营模式虽然获得规模性的经济，却带来了资源配置效率降低的问题。因为垄断的企业是具有市场化的，其可以按照自身获得利润最大化来制定其垄断价格，由于垄断的价格超出边缘成本，势必会造成资源配置效率降低问题。为了控制由于自然垄断而带来的超额利润，各个国家对电力工业进行了电力改革制度，采取了价格管控。但是，价格管制难以解决因为垄断而带来的资源配置效率低下等问题，另一个原因是因为电力行业的垄断经营使其避免了市场竞争的压力，这势必会导致因为降低成本而缺乏提高生产的动力。因此，由于垄断经营模式，使电力行业的弊端远远多过由于规模经济导致的高效率。除此之外，电力行业由发电方、输配电方和售电主体三个环节组成，每个环节的规模经济模式特征并不是相同的，将电力行业笼统的认为是垄断经营模式是不太准确的。把电力行业各个环节分开管理，在发电方和售电方这两个环节引入市场竞争的模式，输电环节可以采取政府管控下的垄断经营。

通过以上内容的分析，在20世纪80年代末到90年代初，世界上许多的国家开始执行电力市场改革，改革内容主要体现在放松管制、引入市场竞争上。其中，英国最先启动大规模的电力市场化改革[7]，将原有的12个地区的电力局完全私有化，将电力池作为电力市场的交易中心，发电公司和地区电力公司与5000多个大用户通过该交易中心进行批发电力交易，通过电力市场的改革，逐步形成了全国统一的电力交场市场。并且通过电力改革，英国的电力市场不仅解决了以往供不应求的情况还将电价大幅度下降。在英国电力改革取得一定成果后，许多国家也相继展开了电力市场不同程度的改革，改革内容是解除市场垄断，引入市场竞争，优化资源配置和效率，增强电力行业的活力等也成为了现代电力行业的发展趋势[8]。

对比西方的许多国家，中国的电力行业的国有垄断模式经营也持续了一段很长的时间。国家电网公司掌握的全国大部分地区的发电、输配电和售电等业务。通过借鉴西方国家的电力改革经验，同时结合我国实际情况，1998年，中国的电力改革首先在东北电网的辽宁，吉林，黑龙江以及山东、浙江、上海作为电力改革试点地区，采用发电厂与电网分开的模式经营，目标是建立发电侧的开放市场。随即，2002年开始，6个省级的电力市场也开始按照不同的经营模式运营。到2002年，电力体制出台了相关的改革文件，方案制定了关于电力改革的大方向，改革提出重组发电资产，建立独立的发电公司。同时也要重组电网资产，分别设立了电网公司等其他改革方案，大大缩小了电力行业垄断经营的模式。通过改革，我国的电力市场逐渐趋于完善，截止到2015年，新的电力改革文件出台，全国大部分地区的发电企业都可以实施竞价上网的模式运行，符合条件的大用户可以直接向发电企业购电，实现发电企业与大用户点对点直接交易，这将是目前我国电力市场建设改革的重点。

### 1.3.2 电力预测研究现状

随着当下电力工业的快速发展，电网的优化管理越来越重要。准确的预测电力负荷趋势是优化管理的先决条件，同时也是当今电力系统实现自动化调度和保持稳定的理论依据。电力负荷预测如果按照预测的期限分类可以分为：长期负荷预测，中期负荷预测，短期负荷预测。长期负荷预测一般是指时间为以年为单位的期限，通过多年度电量，年度电力等指标为内容进行预测。中期负荷预测一般是指时间为月为单位的期限，通常以月度各指标进行预测。短期则是以周、天甚至以小时为单位，来预测某个高峰期的电力负荷。本文将在小数据量时，预测中长期的相应的指标，同时在数据量大时，预测出短期的电荷的相应指标。前者一般使用的方法有类比法和专家估算法等，后者一般可采用灰色模型，回归分析预测，基于时间序列的预测方式等。但是这些方法的实验结果会受到多变量，非确定等因素影响，因而单一的预测方式不能适应更加复杂的情况，需要能将所有因素综合考虑的组合方式进行建模预测。随着机器学习算法的蓬勃兴起，电力负荷预测早就从单一变量延伸到了对多因素多变量问题的解决。依靠电力工作人员长期积累的经验来预测电力负荷的方法，早已经不适合目前电力市场的快速发展。进入二十一世纪，随着电力交易的更加开放，信息化程度高，电力系统更加复杂化，预测方法需要更加科学，预测的准确度需求更高，这也对预测电力负荷的理论知识的研究展开更深入的讨论。当前，国内外研究电力负荷预测的论文比较多，这其中使用的预测模型和分析方法也多种多样，总结起来包括两大类的方法：传统负荷预测，新兴的预测。

传统负荷预测传统的电力负荷预测一般挖掘出电力负荷本身的规律或者电力负荷与一些简单的变量之间的关系，从而运用这些规律对未来进行负荷预测。但这种方式所预测出的结果并不是让人十分满意。概括常见的方法有以下几种：

(1)时间序列法是在七十年代提出来的。首先需要获取电力负荷的历史数据，将这些历史数据当做为时间为变量的增长序列。因而可以对历史数据的统计进行系统的分析和研究，找出电力负荷随时间变化的关系，以此来建立时间序列模型。最后通过这样一个模型来对未来的某地区的电力负荷进行趋势预测。

(2)回归分析法属于统计学方法的一种。通常情况下，回归分析首先需要确定预测值和影响因子之间的关系进而预测出所需值。因为回归模型是对历史数据的从新构造，因而此方法能很好的对历史数据进行拟合。得到了较好的预测结果。

(3)曲线回归法又可称之为趋势外推法。当电力负荷变化与季节因素相关联很小时，则就可以根据已知的历史数据拟合得到一条曲线。曲线回归法的优点在于：适用于数据量少，有效的预测短期负荷。其缺点是：不能随动态的负荷变化进行预测，会增大误差影响结果。

新兴负荷预测

(1)人工神经网络神经网络[9]最早的研究是从40年代开始，有心理学家Mcclloch和数理逻辑学家Pitts首先提出了一个简单的神经网络模型。它是模拟动物神经网络行为特征的一种方法。这种预测方法能够从样本数据中自动学习以前的经验，通过样本数据建立一个规律模型，然后进行预测。由于神经网络具有多个神经原结构，因而它适用于分布式并行处理，显然具有较快的处理速度和较高的容错性。人工神经网络可自主学习，在学习的过程中进行优化，因而能得到一个较好的结果。但是人工神经网络也有其局限性比如：容易陷入局部极小值，当网络建立的过于庞大时收敛速度会出现缓慢的情况，同时他所需要的样本容量比较大等。这就出现了以神经网络为主体的改进算法。

(2)小波分析法[10]：其应用是数学研究成果转向工程学的另一个代表。它是对时域和频域分析的一种方法，使得在时域和频域中都达到良好的局部化性质。小波变化能将各种频率糅合在一起，并且将这些频率拆分成不同频带上的块信息。在进行短期负荷预测时，可以对负荷序列进行小波变换，这样就可以将序列映射到不同的尺度上，各个尺度上的子序列分别代表原序列中不同频域的分量，从而表现负荷序列的周期性。在经过分解过后的负荷子序列，分别建立相应的模型，得到各个负荷子序列预测情况，最后对子序列进行重构，完成预测结果。

(3)组合模型法：由于对电力负荷的影响因素多种多样，我们很难从单一模型中建立起对结果的精确预测。这是由于单个模型的或多或少尤其自身的局限性，使得得出的结果难以让人满意，这就进一步的促进了组合模型的发展。在国内外相当多的文献都运用了组合模型的方式进行算法改进。例如可以运用启发式算法（模拟退火算法，粒子群算法，遗传算法等）对神经网络进行迭代次数的优化。这样能综合各种算法的优点，提高预测的精确度和效果。

### 1.3.3 推荐系统研究现状

推荐系统是[11,12]目前互联网中最常见的智能技术，无论是电子商务、电影网站、音乐网站以及目前的在线广告和在线应用推荐，推荐系统都扮演着至关重要的角色。在20世纪90年代早期，利用网络上数百万人的意见帮助人们寻找自己感兴趣的商品和内容非常具有诱惑力。目前为止，这个简单想法确实在不同的领域也证明了其有效性，这个简单的想法就是推荐系统的前身。1992年，帕洛阿尔托研究中心在Tapestry系统中引入了协同过滤推荐的概念和算法，向世人展现了如果将显示的标签数据以及隐式的用户行为数据存入数据库中，以及用户如何通过这些数据进行过滤选择出自己感兴趣的商品或内容。

纵观推荐系统的发展历程，主要有四个相互交叠的发展阶段。在早期阶段基本上就属于推荐系统的探索阶段，不仅包括了协同过滤还包罗了基于知识的推荐系统，在初始阶段，例如FindMe系统证实了推荐系统的有效性，为日后人们在该领域展开科研以及互联网商业实践方面奠定了基础。在这一阶段中关键的事件是1996年3月份在伯克利举办的推荐系统协同过滤专利研讨会，研讨会将不同领域的中的人们聚集在一起，主要围绕推荐系统这一主题进行商讨最终达成共识。接下来的发展阶段，推荐系统随着快速发展的互联网行业而展现其功能，1995年MIT的Pattie Maes研究组创立了Agents公司，后期更名为荧火虫网络，同期还有许多其他公司兴起并奋起直追，在实验过程中，工作人员为了证明推荐系统能够提供有效的推荐而遇到各种挑战。需要处理上百万的用户和物品以及每秒成千上百的交易，开发人员为了降低在线计算时间而开发许多新的推荐算法，包括至今还在应用的降维方法和基于物品的关联算法。在这一时期，推荐系统的研究领域非常广泛，例如冷启动问题、隐式评分及可信度等用户体验相关的问题。但是，随着2000年到2005年互联网泡沫的破灭，许多推荐系统公司由于无法将推荐利用到更全面的商业主流互联网公司而倒闭，但是推荐系统作为一种技术依旧广泛应用在电子商务、互联网推荐中。同时，推荐系统吸引了各个学科的研究人员加入，如机器学习，人工智能，数据挖掘等，他们将本学科的方法引入到推荐系统中，为推荐系统注入了新鲜血液，提供了更多新的方法和成果。由于可以获得海量的数据，推荐算法的研究成果取得了很大的提升，在2006年Netflix大奖赛上将预测精度提高10%被奖励100万美元更是将推荐算法推上了高峰。

在我国，互联网行业快速发展，中国的网民数量已经达到世界首位，网络消费正在悄然改变着人们的消费观念，成为最有潜力的消费主流。许多国内大型电子商务平台都引入了推荐系统，如淘宝网、京东商城、当当网等针对不同用户的购物行为分析其偏好，使用户体验到个性化定制的消费体验，在海量商品中更加容易获得自己感兴趣的商品。除此之外，一些社区网站的兴起使电子商务平台网站结合用户的社区网络进行个性化推荐，如人人网、朋友圈、微博等社交网络的兴起，以及以陌陌、微信等为代表的基于移动互联网的社交网络快速发展社交网络的蓬勃发展推动了个性化推荐技术的进步，也对新形势下个性化推荐提出了新的要求，加速了电子商务与社交网络的深度融合。

## 1.4 论文章节安排

本文主要部分分为六个章节进行，下面描述每个章节的研究内容。

第1章，绪论介绍了论文的研究背景和意义，阐述了论文的主要内容是基于电力改革的背景而提出的交易平台设计，并在系统中实现电力负荷预测及推荐算法。同时，介绍了电力交易的国内外研究现状以及推荐系统的研究现状。

第2章，相关技术介绍主要是详细介绍推荐系统中协同过滤算法，同时介绍了传统互联网推荐系统与移动端推荐系统的概念和区别，除此之外，还介绍了推荐系统的相关技术和电力预测技术。

第3章，基于用户属性相似度的协同过滤推荐算法主要是针对本文电力交易方式改变这一背景，提出了基于用户相似度的矩阵填充算法，填充后的矩阵实现基于用户相似度的协同过滤推荐，以列表的形式将Top\_k发电企业推荐给大用户，促进双方交易。

第4章，基于粒子群神经网络电力负荷预测主要通过粒子群优化神经网络算法，精确预测发电企业的电力负荷，以服务与推荐算法中关于发电企业推荐的电量匹配中。

第5章，电力交易推荐系统的移动端的实现主要介绍了移动端电力推荐系统的整体架构以及架构实现，并将预测算法与推荐算法在移动端平台实现，实现整体电力交易推荐系统。

第6章，总结与展望主要内容是总结整篇论文的思路与设计，并针对不足对下一步工作提出展望。

# 第2章 相关技术介绍

随着信息技术和互联网时代的到来，我们已经从信息匮乏的时代逐渐演变到了信息过载的时代，无论是对消费者还是对生产者，都需要面对很大的挑战。对于消费者而言，从过载的信息中搜索寻找到满足自己需求的信息或者商品是一件很困难的事情；对于生产者而言，被消费者选中自己生产的信息或者商品，并且受到广泛的关注也是一件困难的事情。因此，随着信息时代的发展，推荐系统应运而生，推荐系统的三个参与方分别是：用户，物品提供者和提供推荐系统的网站[11,12]，如图2.1所示。其任务就是联系消费者和生产者，既能帮助消费者发现自己感兴趣的信息，又能帮助生产者可以在对其感兴趣的消费者面前展示自己的信息，最终使消费者和生产者实现共赢的状态。同时，推荐系统也是一种关于信息的过滤技术，并且也是帮助消费者进行决策的支持系统，可以帮助消费者在购买商品的各个阶段中辅助消费者减少决策资源、提升决策的质量。

图2.1 推荐系统参与者

Fig. 2.1 Recommended system participants

## 2.1 推荐系统相关技术

推荐系统是一种能快速帮助用户找到对自己有用信息的工具，无需用户提供准确的需求信息，通过分析用户的历史行为信息而给用户的兴趣建模，从而根据用户的兴趣模型主动向用户推荐满足他们需求的商品或信息。推荐系统一般给用户提供一个可能用户感兴趣的事物的推荐列表，或者是预测用户对每个推荐的事物的喜爱程度，这些系统帮助用户决定在事物集合中选择适当且轻松的找到首选的事物。

通常来说，有两种推荐算法：协同过滤推荐算法和基于内容的推荐算法。协同过滤算法可以进一步分为基于模型协同过滤和基于邻域的协同过滤算法[1,2,3]。基于模型的方法是使用了user-item的评价信息学习出的一个预测模型，其整体思想是通过建模user-item的交互信息反应出用户和事物的潜在特征，例如用户的偏好类别和物品的种类类别；与其相反的是，基于邻域的协同过滤推荐方法是应用系统中的user-item的评价分数直接预测对新事物的评价分数，下面主要介绍一下关于邻域的协同过滤推荐算法，分为两个主要的部分：一类是基于用户的协同过滤算法；另一类是基于物品的协同过滤算法。

基于用户的协同过滤算法大体上是通过以下两个过程实现：

(1) 通过计算相似度找到与用户兴趣相似的用户集合。

(2) 通过这个相似兴趣用户集合，找到目标用户感兴趣的且没有交易过的物品推荐给目标用户。

步骤(1)中用户之间的兴趣相似度是计算得关键，协同过滤算法中计算用户相似度主要采用用户行为的相似性计算用户间的兴趣相似度。如果用户分别用u和v表示，表示用户u正反馈的物品集合，N(v)表示用户v正反馈的商品集合，可以通过Jaccard简单地计算u和v的相似度。

(2.1)

或者余弦相似度计算：

(2.2)

但是对于上述的余弦相似度公式过于粗糙，接下来有学者通过改进该公式来提高基于用户协同过滤算法的推荐性能。John S.Breese在其论文中提出改进的公式，他认为两个用户对冷门的物品表达了同样的行为才能说明他们的兴趣是相似的。公式如下：

(2.3)

可以通过以上公式看出，惩罚了用户u和用户v在他们喜欢的热门事物上相似度的影响。

基于物品的协同过滤推荐算法的核心思想是找到与用户喜欢的物品相似的物品，将这些相似的物品推荐给用户。该算法主要分为两个部分：

(1) 计算物品之间的相似度。计算下列关于物品i和物品j之间的相似度,其中N(i)和N (j)分别表示喜欢物品i和j的用户集合。

(2.4)

(2) 根据用户的历史行为数据和物品之间的相似度给用户产生推荐列表。

推荐算法的准确性通常用召回率和准确率计算，在top-k排名列表中，对于命中或者召回的部分可以被定义为用户u在所有测试集中的相关物品的选中的物品的top-k列表集合，N（k，u）表示测试集上用户top-k的排名列表，N（u）表示测试集上所有用户u对相关物品的评分。在测试集上所有用户的平均召回率的公式如下：

(2.5)

准确率的定义描述的是最终推荐的列表集合到底有多少比例发生过用户-物品的评分记录。公式如下：

(2.6)

在线社交网络[13,14]（online social networks）对于进一步提高推荐系统的准确度提供了一个新的机会，在现实生活中，人们在购买产品或者消费服务之前通常会在他们的社交网络中询问朋友的建议，社会学领域和心理学表明，人们倾向于两个或多个与自己兴趣相投的人的态度，被称为同质性。由于稳定和持久的社会绑定，人们更与朋友分享他们的个人意见，相比陌生人和供应商，用户通常更信任他们朋友的推荐[14,15,16]。目前流行的在线社交网络如Twitter，YouTube等为人们提供了新的虚拟交流社区。在线社交网络不仅方便用户分享他们彼此的意见，也作为一个平台开发新的推荐算法，通过一定的方式将用户与物品联系起来，如：好友，用户历史兴趣以及自身注册信息，如图2.2所示。一个社交推荐系统对传统的推荐系统在准确性上进行了改进。

图2.2 推荐系统联系用户和物品的方式

Fig. 2.2 Recommend system to contact users and items

## 2.2 移动端推荐系统

移动设备对于目前存在的计算机技术拓展了检索数据的方法，因为移动设备的可移动性可以使用户在任何地方任何时间给用户推荐其感兴趣信息[17,18]。在本论文中，将电力交易推荐系统实现在移动环境中。

表2.1 移动推荐系统与互联网推荐系统对比

Table 2.1 Recommendation system comparison between mobile and Internet

|  |  |  |
| --- | --- | --- |
| 性能 | 移动网络推荐系统 | 传统互联网推荐系统 |
| 推荐结果显示 | 列表方式、基于评价的方式、基于其他用户的体验方式、基于地图的方式 | 列表方式、基于评论方式、基于其他用户体验的方式 |
| 推荐服务网络获取方式 | 需要用户移动通信网络连接和无线网络接入 | 需要用户接入有线网络 |
| 用户设备的移动性 | 手机、PAD和平板电脑等可以随用户移动 | 无 |
| 推荐实时性要求 | 对推荐实时性要求较高 | 对推荐实时性要求较低 |

与传统的互联网推荐系统相比，移动推荐系统将推荐系统延伸到移动网络环境中，又一步缩小了用户在现实环境与虚拟网络环境的距离。与传统互联网用户比较，移动用户处在更复杂的移动网络环境中，需要精确且充分提取移动用户的各种因为移动网络环境产生的各种类型移动信息内容后，才能更好地提供移动推荐。同时，随着移动通信技术的发展与无线网络的介入，用户可以通过连入4G移动网络和访问无线网络随时随地进行网络交易等。移动推荐系统在推荐算法上与传统互联网推荐系统的主体思想基本是一致的，但是移动网络推荐系统需要考虑用户所处的移动网络环境带来更复杂的行为信息。下列表格描述了移动推荐相同与传统互联网推荐系统的差异分析[19]。

对于移动推荐系统的普适性和移动个性化推荐，目前在我们的生活中已存在大量移动推荐系统的应用，下表展示了一些移动推荐系统的经典应用[20]。

表2.2 移动推荐系统应用

Table 2.2 Mobile recommendation system application

|  |  |  |  |
| --- | --- | --- | --- |
| 移动推荐系统应用 | 经典应用 | 移动设备类型 | 推荐方法 |
| 移动推荐新闻 | Daily Learner  Moners | 手机、平板电脑等 | 基于内容推荐方法  基于内容推荐方法 |
| 移动推荐旅游 | Compass  Cyberguide  MTRS | 手机、平板电脑等 | 混合推荐  基于知识的推荐  协同过滤推荐方法 |
| 移动应用程序 | Appjoy  Appazaar | 手机 | 协同过滤推荐方法  协同过滤推荐方法 |
| 电影和音乐 | CoFoSIM  MOBICO  RS-Movie | 手机 | 协同过滤推荐方法  混合推荐 |

## 2.3 相关预测算法的改进

人工神经网络是一个信息处理系统。这种系统中，信息是由几个相互关联的被处理的简单的元素，被称为神经元，它们被定为在分离的网络层，最好的神经网络多层感知器由三个独立的层组成：输入层、输出层和隐藏层。其中，确定隐藏层和神经元的数目是问题的难点，每一层的每一个神经元连接到下一层的所有神经元，但是这一层中不存在相互连接，在每个连接中，会分配到相应的权重并乘以传输信号[9]。

神经网络的性能优劣依赖于网络的拓扑结构或体系结构，包括隐藏层的数目和隐藏层中的神经元的数目。在训练网络中应该有足够的输入输出模式，被称为训练对。在训练结束后，一旦错误误差达到了预先指定的误差范围内，就会被指定为是最佳模型。反向传播（BP）算法是神经网络中最强大的技术，在1986年由Rumelhart和McCelland为首的科学家小组提出的一个人工神经网络的分支，在前馈BP人工神经网络中，神经元被相关层组织且向前发送信号，在该算法中，基于预测和实际网络输出之间的差异，对神经元间连接的权重进行调整，这个过程被称为学习或训练，预测值和实际输出之间的差异被称为网络误差，所产生的误差通过网络传播并更新每个连接之间的权重，这被称为传播，重复此过程直到误差收敛到一个定义的水平，如均方根误差（RMSE）。BP是一种按误差反向传播的算法达到训练多层消息向前反的神经额网络，是目前应用最广泛的神经网络模型之一。BP神经网络最为重要的优点是它具有自学习和自适应性的特性，且具有鲁棒性和泛化性，因此目前BP神经网络应用于多个领域，可以用作分类、聚类、预测等.其中，BP神经网络较多被采用神经网络需要有一定量的历史数据，通过训练大量的历史数据，神经网络可以学习到数据中隐含的内容。但是随着BP神经网络在多种应用中的实现，也产生了一些特有的关于BP神经网络的缺点，例如在寻优过程中容易陷入局部极小值、收敛的速度较慢、迭代时间长等，而且神经网络的结构，包括输入变量的格式及隐含层数目大小的确定还需要通过实验来确定。随着人们对感知机制兴趣的减退，有关于神经网络的相关研究也停滞了一段很长的时间，1982年和1984年，美国的物理学家Hopfield在美国科学院院刊上发表了关于人工神经网络研究的两篇论文，在学术界又产生了巨大的反响。人们又将注意力重新放在了神经网络的研究上并且将神经网络应用到的实际中。

目前，为了克服BP神经网络的缺陷，在分类预测中，各种形式的或改进的神经网络方法被应用，如将模糊理论与神经网络相结合，将小波分析法与神经网络结合，利用遗传算法优化神经网络等。

(1) 遗传算法优化BP算法[21]：遗传算法属于进化算法的一种，寻找最优解的方式是通过模仿自然界的选择与遗传，其基本的三个算子：选择、交叉和变异。其中，该文作者提出应用遗传算法于神经网络算法结合，组合算法主要分两步分组成。第一步利用遗传算法优化初始化权值，因为遗传算法的全局搜索能力强，第二步是根据BP算法的局部搜索速度快的特点获得最佳的权值。相对传统的BP神经网络算法相比，收敛速度加快，误差减小很多。虽然遗传算法的优点是其具有良好的全局搜索能力，可以快速搜索出所有解的解空间，而不会快速下降得到局部最优解；并利用其固有的并行性分布式计算，以加快搜索速度。但遗传算法的局部搜索能力较差，导致简单遗传算法耗时较多，在进化后期搜索效率较低，在现实应用中，遗传算法也容易呈现早熟收敛问题。遗传算法一直难以解决即要保持个体良好，又要保持种群的多样性的问题。

(2) 小波分析法优化BP算法[22]：本文第一章已将小波分析法的原理介绍完毕，针对BP神经网络算法的缺点，许多学者利用小波分析法结合神经网络算法使预测效果更准确。小波分析优化神经网络方式为两种，主要有辅助式和嵌套式。辅助式主要是通过小波分析法对提取的信号进行特征预处理，然后利用神经网络训练，可以通过神经网络对小波分析法的参数与系数的最优解求出，或者利用小波变换对信号进行预处理以用来提取其特征，再将特征数据传入神经网络中进行训练，处理后可以大大减少神经网络输入层的神经元数目；嵌套式的方法主要是将小波分析的运算与神经网络融合，形成了众所周知的小波神经网络。BP神经网络算法众所周知存在的问题是收敛速度慢，因此，提出了小波网络模型使用了随机梯度算法，近年来，又提出了针对小波神经网络算法的改进方法，如模糊小波神经网络的区间学习算法，基于离散小波的学习算法等，主要研究的一点是怎样选择最优的代价函数以及最小的均方根误差的标准。虽然小波神经网络算法在网络的学习方法、拓扑结构设计和鲁棒性不断有研究成果出现有效的优化算法，但是对于结构庞大和鲁棒性较差的问题还没有得到有效的解决。

## 2.4 本章小结

本章主要介绍了推荐系统的相关技术，具体描述了协同过滤算法的相关技术，同时分析了移动端推荐系统和传统推荐系统的异同点以及目前移动端推荐系统的应用，还介绍了相关算法对BP神经网络算法的改进，并分析优化算法的优缺点。

# 第3章 基于用户相似度的协同过滤推荐算法

为了开发针对电力交易的推荐系统，提出了基于用户属性相似度的协同过滤推荐算法[23]，下一章内容中预测的电力负荷是为了针对本章电力交易推荐时大用户与发电厂进行电量匹配时使用。本文推荐算法首先针对现有交易模式下形成的大用户-电厂矩阵进行评分矩阵填充，填充方法采用重新定义的用户属性相似度以及各用户对发电企业的评分拟合而成。然后通过填充的评分矩阵针对大企业用户进行协同过滤推荐。分析该算法的合理有效性，通过实验确定算法的准确性。

## 3.1 电力交易推荐系统模型

协同过滤推荐算法是目前应用最广的推荐算法，在电子商务及音乐、电影推荐系统中广泛应用，为本文奠定了理论基础[24]。本文用大用户表示企业用户，指的是可以接入较高的电压等级，具有一定购电能力的电力用户。在当前电力改革体制中，大用户直购电的意义是对现有电力销售机制的探索，引入大用户直购电模式可以使电力市场同时出现多个卖家（发电企业）和多个买家(企业用户)。对于发电企业来说，以前的电力交易体制中，电网是企业用户的唯一买家，发电企业只负责发电，并不能参与的销售环节，其电价的决定权并不在发电企业的掌握之中，因此在发电企业的积极性、对发电技术的改良以及降低成本的动力都会大大降低。现在，电力改革后，大用户可以向发电企业直接购电，电网企业不再是唯一的卖家，发电企业随着面对销售对象的增多以及势必带来的竞争模式，对发电技术的动力和定价的灵活性大大提高；对于大用户来说，他们属于电力交易的终端，引入直购电机制，使大用户有了与发电企业“讨价还价”的权利，可能根据其是大用户这一优势获得比以往更低的电价。

### 3.1.1 基于电力交易推荐系统拟解决的问题

本文主要研究电力交易的推荐系统。在现有的电力交易市场中，由于几十年来的电力自然垄断经营模式，企业用户与发电厂并不能直接交易，而是通过电力交易平台进行调度，企业用户需要与国家电网指定的发电厂进行交易，导致交易的选择强制性。而改革之后，企业用户可以通过自身需求与一个或者多个发电企业进行直接的自主选择交易，这种直购电模式已经在试点地区运行了一段时间，本文推荐系统的最终目的是针对电力改革后对电力市场中的新的企业用户推荐令其满意的发电企业。针对本文电力交易的推荐系统主要解决以下几个问题。

(1) 处理大用户-发电企业评分矩阵数据稀疏化

众所周知，大部分的协同过滤推荐算法通过用户-物品评分矩阵进行基于用户相似度推荐或者基于物品相似度推荐，其中矩阵的评分数据对推荐算法的精确度有至关重要的作用。通过分析用户的购物行为以及用户对购买过的物品的评分进行相应的推荐，如果用户积极主动评分，则矩阵的每一行都会产生多个评分数据；如果用户并不积极主动评分，相应行便会缺失评分数据，导致数据稀疏。在本文的电力交易应用中，由于之前的电力交易模式，使企业用户与发电企业在之前的交易模式中一直被“强制”要求的交易关系对，无法确定企业用户对发电企业的评价信息，电力改革后，企业用户可以对发电企业自主进行选择，他们的选择必定是考虑到了自身的需求以及各方面属性，因此将他们与发电企业的交易次数设定为初始评分矩阵中的评分，用户越满意与其交易的发电企业，则交易次数就会相应增加，反之，则交易次数减少。因此将交易次数转化为大用户-发电企业矩阵中的评分是可行的。但是由于电力改革是刚刚启动的，并没有大量的用户与发电企业进行过直接交易，矩阵中相应的评分便会是缺失值。就产生了上述所描述的矩阵稀疏问题，针对本文拟解决问题产生的稀疏矩阵，我们提出了采用基于大用户需求属性相似度的方法进行评分的矩阵填充，确定评分矩阵中的缺失值，填充后的相对完整的矩阵再针对已有用户和新用户进行协同过滤推荐。

(2) 通过用户间的相似度进行评分矩阵填充

根据文献的显示，在基于用户的协同过滤推荐系统中，通常采用Pearon相关系数的方法确定相似用户集合，在不同的领域中，相似度度量方法还有向量余弦法、调整的向量余弦方法、改进的Pearon相关系数法、均方差或者Spearman秩相关系数法等。本文研究的电力交易推荐系统中，针对上述的稀疏评分矩阵，采用企业用户间的属性相似度以及交易情况拟合为最终的用户相似度，利用相似度以及对发电企业的评分进行矩阵填充。

(3) 针对新用户产生推荐列表

最理想的推荐情况是用户可以通过注册网站是主动告我们其兴趣爱好，喜欢什么商品，但是这种方法存在三个问题：第一，用户的兴趣爱好并不能用现有的自然语言技术完全理解，很难描述；第二，用户的兴趣爱好是不断产生变化的，但是他们不能总是在注册信息里更新他们的喜好；第三，并不是每个用户都明确自己的兴趣爱好，他们本身很难用语言其描述自己的兴趣爱好[25,26]。对于推荐系统而言，用户的历史行为和兴趣偏好是其推测用户未来行为和爱好的关键点，因此，大量的用户数据是推荐系统的重要成员，许多例如阿里巴巴、百度、去哪儿网等这些大型网站，用户数据已经积累很多，并不会对其推荐功能产生困扰。而对于初始的网站或者纯做推荐系统的网站而言，用户数据并不是很多，可能会出现冷启动的问题（cold start）。在本文中，由于新用户加入到电力交易推荐系统中产生冷启动问题。新用户到来，我们并没有他们曾经的交易数据或者行为信息，从而无法提供相应的推荐给新用户。因此，本文针对新用户可以分为匿名用户和注册用户两类，匿名用户并未在该交易系统注册信息，只是随意浏览；新用户是指用户注册该交易系统，并填写相应的需求信息。针对这两种用户，需要不同方法产生推荐集。

### 3.1.2 电力交易推荐系统的架构模型

图3.1所示为整个电力推荐系统的整体架构图。根据电力交易推荐系统架构图可以看出，本文的推荐系统主要分为三个模块，数据采集模块、数据预处理模块以及推荐模块，针对电力推荐系统的三个模块进行简单的介绍。

图3.1 电力交易推荐系统架构模型

Fig. 3.1 Electricity trading recommendation system framework

(1) 数据采集模块：在该电力交易推荐系统中，数据采集的主要任务是采集大用户的注册信息，发电企业的注册信息以及双方交易的信息，注册信息是最容易获取到的，只需要在电力交易网站中获取，大用户的注册信息包括公司ID，企业性质，企业规模，需求电价，上次交易电价，需求电量，交易电量，所处位置等，对于发电企业，除了其注册信息，还要采集影响发电企业电力负荷的影响因素，如天气因素、煤炭价格、经济因素等。双方交易信息包括交易电厂，交易用户，交易时间，交易电价以及交易电量。

(2) 数据预处理模块：该模块主要针对采集到的数据进行筛选，分析大用户和发电企业的交易属性信息、自身属性信息、分析用户需求属性的相似度。同时，将现有的电力交易数据映射为大用户-发电企业初始评分矩阵，矩阵的中的评分数据代表了用户与相应发电企业的交易次数。

(3) 推荐模块：对数据预处理产生的初始矩阵进行评分填充，采用基于用户需求相似度以及其他用户对相应电厂的满意评分通过迭代处理使评分矩阵数据趋于丰富且稳定。最终得到的评分矩阵进行协同过滤推荐。

通过各模块的处理，最终将推荐结果以列表集的形式展示给用户。

## 3.2 基于用户相似度的矩阵填充

在许多推荐算法的应用中，数据以低秩矩阵的形式表示[27,28,29,30]。例如对于一些推荐系统，其矩阵的内容表示用户对事物的偏好度；对于交通网络分析，矩阵的内容表示始发点到终点的车流量。在许多实例中，已知的矩阵只有很小的百分比是有矩阵内容的，例如，在Netflix大奖比赛中有480K用户18K电影矩阵数据集，但在矩阵中只有1%的评价信息。这种得到矩阵中缺失数据的方法被称为矩阵填充，矩阵填充方法的目的是准确的推断出矩阵缺失项的值，形成一个假设的完整的矩阵。例如，对一个真实矩阵T，其中已知的矩阵数据只是一个位置集合C定义为矩阵TC，矩阵填充方法利用初始矩阵T中的已有数据（即矩阵TC）建立一个关于矩阵T的较好的估计矩阵Tu。在实践中，这种估计矩阵可能与真实矩阵有很大的不同，特别是当矩阵TC并不足够提供有效的信息来建立一个性能好的估计矩阵。在许多实例中，可以通过积极获取额外的观察信息来解决不足的矩阵TC。举例说明，在推荐系统中，users可以被要求给一定量的items评分；在交通分析中，可以安装额外的监测点。这些额外的信息C可以组成一个增广矩阵TC来获取更多的信息使估计的矩阵Tu更精确。本文针对发电企业-大用户矩阵中的稀疏评分矩阵，需要填充每个用户对相应的发电企业的评分数值，使整个矩阵评分更丰富，以用于接下来的推荐。

### 3.2.1 问题描述

本文的推荐系统中，需要确定大用户对其他发电企业的满意度来进行下一步的推荐功能，在3.1.1节的描述中，直购电模式已经在试点企业运行一段时间，所以我们拥有一个用电企业对不同发电企业的交易次数，我们将购电次数来确定初始矩阵的评分，因为直购电的交易模式是用户自主选择的，必然会综合考虑自身与发电企业的情况再进行交易，因此将与发电企业的交易次数映射为评分信息是可行的（如：A企业对甲电厂选购4次，对乙电厂选购8次，我们将矩阵中每一行评分总分定为100，分别计算A企业对甲乙电厂的评分，则对甲电厂和乙电厂的评分分别为：4/12\*100、8/12\*100），最终将每一个评分映射到（0,100）区间内。但是在直购电模式下也就是需要用电企业对所有的发电企业的大量评分数据才能进行接下来的推荐。可是进行试点的购电企业中，并没有大量的交易行为，如果使用企业用户选择发电厂家的现有数据进行推荐，就会因评分矩阵不完整而导致推荐准确度不高。为此一个完整评分矩阵对推荐系统的精确度起到十分重要的作用。

随着推荐系统的不断发展，矩阵填充已成为了推荐系统中十分关键也是必不可少的部分，如Natali[27]在2015年提出的通过对用户主动询问方式对关键的评分位置进行填充，他设计一种名为Order&Extend的算法，该算法可以从评分矩阵中找出关键的评分位置，从而对用户主动询问关键位置的具体评分来达到其推荐算法更高的准确度。但是在我们的电力交易系统中，首先因为评分矩阵（用电企业对发电企业的评分）是稀疏的，其他数据是缺失的或者说是未知的。如果进行主动询问，询问量过与庞大。并且在我们的模式下是无法对企业进行直接询问的，因为直购电交易模式启动时间并不长，企业对许多其他发电企业并不了解，无法对一些发电企业进行具体的评分。为此我们针对这种情况下的稀疏评分矩阵提出基于用户（购电企业）相似度的矩阵填充算法，由于在试点地区用电企业对发电企业的选择是根据自身企业具体需求所做的决定，所以该评分具有很大的准确性，所以当用电企业的属性相似时，其对电厂的评分也是相似的。

### 3.2.1 用户相似度计算

根据文献的显示[11,12]，在基于用户的协同过滤推荐系统中，通常采用Pearon相关系数的方法确定相似用户集合，在不同的领域中，相似度度量方法还有向量余弦法、调整的向量余弦方法、改进的Pearon相关系数法、均方差或者Spearman秩相关系数法等。本文研究的电力交易推荐系统中，在用电企业对发电厂选择时，通常会根据自身的的一些具体情况来自主选择。如企业的规模，本企业的用电高峰期等特征来对电厂进行选择。所以本文将大用户影响对电厂选择的属性信息分为以下几类：1)大用户自身属性信息，其中包括：企业位置信息，企业规模，企业用电频率高峰期；2)大用户需求信息，其中包括电价需求和电量需求信息；3）大用户与发电企业交易属性，其中包括交易电价和交易电量信息。本文根据以上大用户属性信息，重新定义了大用户间不同属性相似度的度量方法，我们将以上所有的大用户的属性信息分为数值型属性与名称型属性。

根据各属性信息取值情况，其中用户数值型属性包括:电价（Eprice）和电量（Equatity），其中电价属性分为需求电价（Demand Eprice）和交易电价（Trading Eprice）。电量属性分为需求电量（Demand Equantity）和交易电量（Trading Equantity）。用户名称型属性包括：企业规模（Enterprise scale）,可以用大、中、小表示企业规模；大用户所处城市位置（Location），大用户用电频率高峰期（Peak season），用电高峰期通常分为工作日与周末、白天和黑夜等。每种类型的属性进行相应的距离度量计算，用户间的距离值越小，则表明他们之间的相似度越高，通过大用户的属性相似度并结合大用户对发电企业的评分形成最终的评分矩阵。根据以上对属性的分类，接下来具体对各个属性类型相似度的距离度量计算。

(1) 对于大用户数值型属性，为了方便对不同数值型属性的距离度量，我们首先对其数据进行去量纲处理，目的使用户的数值型属性的距离都映射在(0,1)这个区间内。

(3.1)

式4.1的表示用户的数值型属性值，U\_max表示在该数值型属性中的最大取值。为最终数值型数值去量纲后的属性值。那么不同用户之间的距离可以表示为：

(3.2)

其中，大用户的不同属性的数值绝对差值的最小值和最大值的差距区间，把这个最小值与最大值得差距分为n-1个小的区间：，本文主要设置了四个区间。针对每个分好的小区间定义一个数值表示属性的距离度量，当大用户间的数值型属性差值落在相应的小区间时，得到用户之间数值型的距离度量Dnum与simnum之间的映射关系。本文将大用户间的数值型属性定义不同距离区间：

如果，则；

如果，则;

如果，则；

如果，则；

针对本文不同的大用户之间的数值型属性的实际需求，将属性差值分为如下三个区间，并将距离度量与相似度之间产生一个映射关系，表示距离度量越小，相似度越大。

如果，即，则；

如果，即，则；

如果，即，则；

如果，即，则。

(2) 对于用电企业的名称型的属性中的企业规模属性（Enterprise scale）和用电高峰属性（Peak season），我们利用格雷编码进行距离度量，因为企业的规模与用电高峰期这两个属性对用户选择发电厂至关重要，尤其是用电高峰期属性，首先分别确定大用户每个属性的取值个数M，用确定最终的编码位数，将得到的m向上取整，确定需要编码的位数；然后，针对大用户的不同的名称型属性进行格雷编码，再将编码后的数值计算其海明距离，计算出不同的大用户之间的名称型属性的海明距离；最终用表示不同的大用户之间的海明距离。对于第一类名称型属性的距离度量，首先把不同的属性值进行编码，如大用户的企业规模属性：大、中、小分别编码为00、01、11，将大用户的名称型属性串联成二进制的形式组成一个二进制串。然后，组成的二进制串计算其海明距离，表示大用户间的名称型属性的距离度量。这样大用户和间第一类名称型属性值的距离度量：

(3.3)

(A属性的名称属性二进制串为01010000，用户B属性的名称型属性二进制串为11000101，则用户A与用户B之间的海明距离为4)。由于距离越近他们的相似性就越高，本文用其距离的倒数作为用户之间的相似度。

(3.4)

(3) 对于大用户的所处地理位置（Location）这一名称属性，由于电力交易模式改变后，发电厂与大用户的直接交易模式，使电网的盈利模式不再是赚取发电差价与输送电力所得利润之和，而是逐渐转换为通过大用户与发电企业的距离输送电能获取利润值，大用户之间如果所处位置相似，其中某些大用户所选择的交易对象距离其所处城市较近，那么相应的付出的输电成本就会减小，可能最终的撮合电价会相应降低，这对大用户来说百益无害。把这些离大用户最近的用户所选择的发电企业推荐给大用户，比较合理性。因此，针对大用户间所处地理位置的距离，我们用来表示用户地理位置数据的距离，并将其分为个等级，距离度量值越小，表示大用户间的距离相似度越越大。例如，当，表示两个大用户处在同一个城市；当时，表示两个大用户处在同一个省。以此类推，最终得到大用户间的地理位置距离度量。

然后，对于本文大用户的第二类名称属性距离度量，即大企业用户所处的地理位置，定义大用户和之间的地理位置距离D。

如果，即和都在同一个城市，则，其SimLocation=1；

如果，即和处于一个同一个省，但不处于同一个城市，则其SimLocation=2;

如果，即和处于一个同一个省，但不处于同一个城市，则其SimLocation=3。

最终，得到大用户和之间的需求属性距离度量,大用户间的需求属性相似度分别用以上的公式表示，每一部分所对应的权重的取值不同，相应得到的最终用户间的相似性也不同，可以根据本文实际的数据集实验得到，对大用户间的各个需求属性距离度量进行线性拟合，最终得到大企业用户间的相似性，若计算的值越小，说明大用户间越相似，影响力越强。如果计算得到的值越大，说明大用户间的相似性越低，其影响力相应也会减弱。以上数值型和名称型属性的距离度量代入式(3.5)，便会得到大用户和之间的属性相似度。对大用户和间的关于属性的总体相似度度量公式如下所示（其中是大用户每个属性相似度的权重，k表示用户的每一个属性,sim的值越大，表示ui和uj的总体相似度越大）：

(3.5)

综上所述计算用户间的属性值的相似最终得到用户的总体相似度，但是我们需要增加一个电厂企业选择相似性。原因之一是当用户同时选择了相同的发电企业，表明用户越相似；其次，增加这个影响因素也会避免在矩阵填充的过程中每次迭代是用户间的属性相似度相同的情况，从而提升矩阵填充的准确度。那么将电厂企业选择相似性定义为用户间每次“选择”发电企业中相同发电企业的个数，如式(3.6)所示，其中M表示每次评分矩阵填充时每个用户的前M个推荐列表，分子部分表示用户与选择相同的电厂的个数。

(3.6)

最终，电力交易中用户间的相似度由用户属性相似度和用户对发电厂选择相似性构成，用户的属性相似度与对发电企业的选择相似度对总体的影响是不同的，每一个相似度的权重值可以根据实际情况进行设置，将式(3.5)与(3.6)进行线性拟合为公式(3.7)所示，得到最终的用户相似度。的值越大，表明用户间的相似度越高，反之，则相似度越低。（其中，与分别表示两个子相似度的权重值，并满足）

(3.7)

### 3.2.3电力交易稀疏矩阵的填充

通过前文分析，本文将电力市场现有的大用户与发电企业交易数据以评分矩阵的形式展示出来，在实施直购电政策的大用户企业，如果该大用户选择了发电企业，就认为大用户满意该发电企业，对该电厂企业有评分。下面是对本文评分矩阵的具体数学描述。

设定存在一个大小为的矩阵T，矩阵T代表n个用户对m个发电企业的评分矩阵，在我们的评分矩阵中的每一行的评分数据，表示大用户对其选择发电企业的交易次数，由于现有实施直购电的企业，自主选择了认为合适的发电企业，对该企业有一次交易记录，将交易次数设为用户对发电企业的评分。但是由于电力改革启动时间并不长，该用电企业对其他发电企业并没有相应的交易记录，这便导致了我们这个评分矩阵的评分稀疏性。若用户Ui选择了电厂Ej，则在评分矩阵的第i行第j列有其评分Ri,j，没有交易选择的位置评分是缺失值，需要我们进行填充。为了得到用户对其他电厂企业的评分，本文提出了基于用户相似度的矩阵填充算法。

图3.2 矩阵填充示意图

Fig. 3.2 Schematic diagram of matrix completion

如图3.2所示，在初始矩阵T中，我们只有一小部分观测到的矩阵评分数据，本文提出利用用户相似度以及初始矩阵T中的评分数据填充矩阵，最终生成评分数据丰满的矩阵，生成的评分矩阵。

基于用户相似度的矩阵填充算法其核心思想是通过计算各用户之间的相似度，通过加权累加相似用户对电厂企业的评分，来得到大用户企业对其未选择发电企业的评分。之所以可以利用这种加权累加相似性用户评价的形式来对我们的评价矩阵进行填充，是由于虽然大用户只对一个用电厂企业进行评分，但是该评分是经过企业综合考虑得到的，即该评分有很高的准确度。并且我们在上节内容很好的定义了影响大用户选择电厂的具体考虑因素，所以基于用户相似度的矩阵填充是适合本文评价矩阵的。

本文的矩阵填充过程示意图如图3.3所示：在初始矩阵T中大用户对发电企业、和的评分数据是缺失的。而大用户、和分别对发电企业、和有评分数据R。若要得到对发电企业、和的评分数据，首先计算大用户与、和的属性相似度，根据用户间的属性相似度以及相似用户对发电企业的评分数据最终得到大用户对发电企业、和的评分数据。由于相似的大用户选择的电厂是基本相同的，那么当只进行一次各用户之间相似性比较后通过累加评分来填充矩阵时，这个评分矩阵依然稀疏。为了能够得到更“丰满”我们的评价矩阵，为接下来的推荐提供更好的评价矩阵信息，我们会多次迭代这个累加过程，直到得到理想的评价矩阵信息。

图3.3 用户评分示意图

Fig. 3.3 Schematic diagram of user rating

本文的矩阵填充过程步骤如下：

(1) 获取大用户的自身属性值以及需求属性值，将所有属性分为数值型属性和名称型属性，通过3.2.2章节的计算方法得到大用户间的属性相似度。通过公式(3.7)计算大用户间的属性相似度。

(2) 通过大用户间的属性相似度以及用户对发电企业的评分进行具体评分矩阵填充。下列公式表示经过P次迭代后，大用户对发电企业的第p+1次迭代的评分结果。其中表示与大用户属性相似的用户集合，表示第P次迭代时对发电企业的评分，表示第P次迭代时计算得到的大用户对发电企业的评分，设定一个评分最小阈值，减小每次迭代计算评分时因评分太小而影响整个评分矩阵。当评分数据小于阈值时，将该评分剔除，不计入下一次评分矩阵的计算中。

(3.8)

(3) 将大用户对发电企业的评分归一化后，最后经过p+1次迭代的评分结果如下公式，并将最终评分结果填充矩阵，形成完整的用户-电厂评分矩阵：

(3.9)

## 3.3 基于协同过滤算法的发电企业推荐

基于用户的协同过滤推荐算法是最古老也是应用最广的推荐算法，首先根据用户与物品的评分矩阵矩阵，计算得出用户之间的相似度，找到与目标用户相似的用户集合，根据相似用户集合的推荐结果最后产生目标用户的推荐集合。本文基于用户需求属性相似度的协同过滤推荐算法首先分析了电力交易市场中大用户-发电企业的评分矩阵的特殊性，并提出矩阵填充算法对该特殊的稀疏矩阵进行填充，接下来主要针对系统中已经存在的大用户（即曾经与发电企业交易过的用户）和新用户，通过上述矩阵填充算法后形成相对完整的大用户-发电企业评分矩阵后，对存在用户及新用户产生推荐集合。

### 3.3.1 电力交易系统新增用户推荐集的产生

由于直购电交易模式的启动，电力市场交易系统中势必会加入新增大用户进行电力交易，而问题是在于现存的数据信息并没有这些用户交易信息，那么我们将新增用户分为新注册的大用户与匿名用户两种类型用户。由于新增用户并没有电力交易市场的评分数据，那么新用户加入到电力交易推荐系统中时，随即会产生冷启动问题，针对推荐系统中的冷启动问题，本文首先根据已有用户的交易信息建立的评分矩阵进行填充，使填充后的评分矩阵对新注册的大用户和匿名用户采用不同的推荐方法，使每一类新用户产生一个关于发电企业的推荐集。

对于在电力交易平台中新注册的大用户而言，当其注册该系统时，本系统就会为其产生一个唯一的ID，当新增用户注册为买方用户时，除了填写自身的属性信息以外，还要发布需求信息。本文根据其注册信息与发布的基本需求信息通过与已存在的交易用户进行相似度计算，选取相似度值较高的前1/3n个大用户为相似用户候选集，候选集中的每一个用户此时都有一个发电企业的评分列表集。最终，根据新增用户与用户候选集中用户的相似度及用户对推荐列表中发电企业的评分拟合，拟合的结果设定为新增用户对发电企业的评分列表，将被评分的发电企业定义为发电企业候选集，最终在该候选集中通过计算选出前1/3n个发电企业作为新增用户的关于发电企业的推荐列表，对于推荐列表中的发电企业，进行新增用户与发电企业的电量需求匹配。例如，选中的第一个发电企业同时推荐给其他用户，且被其他已有用户选择交易，则出现电力负荷供不应求的情况，需要新用户重新进行下一个选择。图3.4显示了对电力交易推荐系统中新注册用户的推荐过程。

图3.4 新增用户推荐过程

Fig. 3.4 New user recommended process

综合以上内容分析，以下的算法一表示了对电力交易系统中新增大用户产生推荐集的过程。

|  |
| --- |
| 算法一：&Users\_CF(，G，int R，int K)  //算法目标：为电力交易系统中已注册的新增大用户产生发电厂推荐集E。  输入：新增大用户U\_x，相似用户的前N个用户候选集G，矩阵填充后的评分数据R，产生的推荐个数K。 |
| Begin：   1. 通过公式(3.8)中的用户属性相似度计算方法计算出与相似的大用户集合，相似度排名前1/3的与相似的大用户集合G。集合定义如下： 2. 利用Top-k用户对其对自己选择前k个发电企业的评分以及的属性相似度，计算新增用户对相应发电企业的评分用表示，表示对的推荐发电企业： 3. 对于推荐列表中的发电企业，进行用户与电厂的电量需求匹配，剩余电力负荷满足新增用户的电量需求的发电企业作为最终的推荐集，不满足其电量需求的发电企业则被剔除推荐列表，产生最终的新增用户的发电企业推荐列表。   //算法输出：大用户，推荐集。  输出：新增注册大用户的发电企业推荐集。  End |

以上所述即对电力交易市场中新增用户的推荐过程，但是需要注意计算用户间相似性时由于新用户并没有自主选择过发电企业进行过交易，因此把用户间对发电企业的选择相似性设为0。

对于电力交易系统中的匿名用户，由于他们并没有注册自身属性信息，也没有发布需求，只是在网站上随意浏览。无法产生详细的推荐方法将发电企业推荐给匿名用户，本文秉着随机推荐比不推荐更有发展的原则，采用随机在大用户-发电企业存在大用户的推荐列表中Top\_1的发电企业，形成一个发电企业的推荐列表E推荐给匿名用户。

(3.10)

### 3.3.2 电力交易系统已交易用户的推荐集的产生

由于电力市场中直购电模式的初步启动，目前已有部分大用户企业直接与发电企业进行交易，这些已交易的用户企业必然是经过综合调研自身需求以及大量发电企业的情况，最终选择与发电企业进行撮合交易。也就是说，目前的交易行为表明当前用电的大用户企业对自己选择交易的发电企业是满意的。因此我们将双方企业之间的交易行为转化为企业用户与发电企业之间的初始评分矩阵，每一次交易行为都转换为评分数据，使初始矩阵中的评分为用户与发电企业的交易次数。这样，我们通过已交易用户之间的属性相似度以及对发电企业的评分数据进行矩阵填充有可行性较高。

在进行矩阵填充时，根据评分计算公式(3.8)可以看出用户与自己也要进行相似度比较，这是因为用户本身对自己的选择就是满意的结果，如果剔除自身相似性则对用户本身的选择产生不公平的结果，使用户对其选择的发电企业评分会因为迭代计算而越来越低。矩阵填充后，对已交易的大用户也会产生一个推荐列表，这个推荐列表可以在用户进行下一个季度交易时，当做一个反馈信息供其参考，如果下个季度进行交易是已入网的用户需求发生了变化，那么认为该用户为新增用户，推荐方法如3.3.1章节相同。

同时，每次新增用户加入电力交易推荐系统时，把他们对发电企业的评分矩阵也加入到我们的填充矩阵中，更新为一个新的评分矩阵，虽然评分矩阵越大推荐结果越准确，但是因为矩阵过大会造成计算复杂度过高且用户需求发生变化，我们需要每三个季度交易后将第一个季度的评分矩阵剔除，对于没有发生需求变化的用户来说，更新矩阵时保留其对发电企业的评分，对于发生需求变化的用户而言，重新计算其对发电企业的评分，将结果更新的评分矩阵中，达到整体评分矩阵的更新。

## 3.4 实验结果及分析

本章实验的结果在单机环境下运行而得到，运行环境的计算机处理器为Intel(R) Core(TM)2 Quad cpu Q9400 @ 2.66GHz，操作系统为64位Windows7，采用Java语言实现。实验数据集为2015年蒙东地区的大用户与发电企业交易的数据以及相应该地区其他大用户与发电企业的属性及需求数据。对于交易数据集中有2530个大用户与272个发电企业产生交易。对于交易数据集，抽取1/3作为测试集，2/3作为训练集。

### 3.4.1评分数据标准化及实验数据获取

对于获取大用户的自身属性及交易属性信息，计算大用户间的相似度。在本文中，推荐算法的大用户关于属性的信息来自于电力交易推荐系统移动端中用户的注册信息、发布需求的信息以及交易信息中获取得到的。利用大用户的所有属性值计算大用户间的数值型属性相似度和名称型属性相似度，最终拟合为大用户相似度。为了表示大用户的数值型属性、名称型属性和大用户属性，创建了四个数据字典，对于数据字典中的用户数值型属性、名称型属性和相似选择性分别建立属性的取值字典：数值型属性；名称型属性；用户的电厂选择属性则大用户的属性取值为，最终通过属性取值计算大用户间的相似度。

由于电力改革前的交易模式，使我们持有的电力交易数据显示每个大用户通过电网只能和唯一的一个发电企业进行交易，一个发电企业可以同时与多个大用户企业进行交易。本章的目的是通过电力交易的原始数据设置一个初始User-Elec的评分矩阵通过矩阵填充得到一个相对完整的评分矩阵，将交易行为转换为大用户对发电企业的评分。初始交易矩阵中将大用户对交易的发电企业评分数据之和定义为100分，矩阵填充后的矩阵评分考虑到大用户之间的相似性差异及与不同发电企业的交易，将评分范围设置为0到100之间，使大用户对唯一的发电企业的100分按照权重均分给其他发电企业。本文按公式（3.11）所示的方法将评分标准化为：

(3.11)

其中，分子部分表示大用户对发电企业第p+1次迭代的评分，分母部分表示大用户对所有发电企业的评分，评分标准化到0至100之间，减少大用户对发电企业最终评分的差异性，设定评分阈值为5分。

表3.1 部分数据集

Table 3.1 Partial data set

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 购方名称 | 售方名称 | 成交电量（MWh） | 购方电价（元/MWh） | 直接交易电价（元/MWh） | 发电类型 |
| 矿业集团 | 机组1#2# | 72200 | 355.6 | 206.5 | 火电 |
| 铜业集团 | 机组1#2# | 20000 | 360.55 | 211.45 | 火电 |
| 国电化工 | 发电厂1 | 20000 | 344.1 | 195 | 火电 |
| 钢铁集团 | 发电厂2 | 14700 | 334.66 | 211.56 | 火电 |
| 铝业集团 | 发电厂3 | 10000 | 289 | 165.9 | 火电 |
| 科技集团 | 热电公司 | 77800 | 288.6 | 165.5 | 火电 |
| 焦化企业 | 机组1#2# | 7500 | 324.05 | 174.95 | 火电 |
| 中网科技 | 兴安热电公司 | 77800 | 288.6 | 165.5 | 火电 |
| 得丰焦化 | 鄂温克1#2# | 7500 | 324.05 | 174.95 | 火电 |

如表3.1所示为部分大用户与发电企业的直接交易数据表，举例说明矩阵填充的过程，将该数据表的购买方（大用户）用表示，售方（发电企业）用表示，将交易数据转换为初始评分矩阵User\_Elec表示，初始评分矩阵中将大用户对自己交易的发电企业次数设置为评分，将每一行的总分设为100分，通过归一化后将初始评分映射在（0，100）区间内，最终通过大用户间的相似度以及相似用户对相应的发电企业的评分，且评分阈值设定，当每次评分填充时评分小于该阈值时，不计入下一次迭代计算中。最终根据公式(3.11)计算并迭代得到最终的大用户对发电企业的评分估计值且矩阵每一行中用户对发电企业的评分之和为100。

### 3.4.2 实验评价标准

评价标准是针对推荐算法的质量进行一种评估的方法，利用大用户与发电企业的交易数据集，把数据集分成测试集和训练集两个部分进行实验对比，把训练集的比例设为较大的数据集，放入本文的推荐算法中，通过多次迭代训练形成一个完整的评分矩阵，然后将产生的推荐列表与测试集中的数据进行对比，如果用户实际选择的发电企业在我们推荐的列表内，则表明我们的推荐方法产生的推荐结果是正确的，最终确定本文的推荐算法的准确性，从而引入了准确度(Accuracy)和召回率(Recall)的评价标准，根据本文的实际应用，分别定义如下公式所示。

(1) 准确度（Accuracy）：

(3.12)

我们考虑到推荐的Top\_k发电企业的评分，通过相应的评分来确定准确度，为了解决当真实选择的发电企业个数小于K时，在计算准确度时我们要对用户实际选择的发电企业进行扩展，扩展到K个。扩展方法是采用命中的发电企业中评分最小的进行扩展来计算推荐算法的准确度。其中T代表推荐算法为用户推荐的发电企业集合，代表对扩展后用户实际选择的发电企业集合，Rank代表用户对发电企业的评分。

(2) 召回率（Recall）：预测值与实际值的交集\实际值的个数。

(3.13)

N(R)表示企业用户实际选择的发电企业的个数，表示企业用户实际选择与我们推荐的发电企业的交集数目。

那么根据实验中的初始矩阵T以及矩阵T中的初始评分数据，我们要计算提出的推荐算法的性能，通过实验验证填充后的矩阵与初始矩阵T产生的推荐结果的准确率与召回率，使用定义在式3.12和3.13的准确率和召回率公式进行计算。

### 3.4.3 系统参数的确定

对用户相似度的权重系数的确定我们采用最小二乘法来确定,并满足。通过阅读参考文献[28]，充分考虑了用户属性相似度与用户选择相似度对用户相似性的影响，通过实验确定，使用户的相似度达到最大。

图3.5 准确度随迭代次数的变化

Fig. 3.5 Accuracy varies with the number of iterations

本文提出的矩阵填充算法中，是通过多次对用户间相似度计算之后得到一个较为“完整”的评分矩阵，然后根据这个通过多次迭代之后得到的评分矩阵，应用接下来提出的推荐算法为新增用户进行发电企业的推荐。图3.5显示了在推荐列表选择TOP-5时矩阵填充过程中随迭代次数增加，推荐系统准确率的变化曲线图。通过图3.5显示，矩阵填充过程中推荐系统的准确度随着填充算法的迭代次数而产生的变化，尤其是当迭代次数从0到10时，算法的准确度迅速提高，当迭代次数达到90次左右的时候，推荐系统的准确度趋于稳定。

在计算用户相似度时，其中影响总体相似度的因素中我们增加了用户对发电企业共同选择这一属性，通过对比用户间评分最高的M个发电企业的交集个数来得到用户间相似度，为了观察不同的M取值对系统精确度的影响，我们分别选取在推荐用户TOP-K时，K值的0.5,1,1.5,2倍作为M的取值。图3.6显示了在这些取值下系统精确度的示意图。从该图可以看出随着M值增加，精确度也是增加的特别是从0.5到1之间，而从1.5到2精确度增加并不会太明显，所以我们一般选TOP-K中K值的1.5倍作为在进行用户相似度计算中M的取值。

图3.6 准确度随M值的变化

Fig. 3.6 Accuracy varies with M value

为了验证在推荐系统中，向用户推荐不同个数发电企业后推荐算法的精确度，我们采用迭代90次之后评分矩阵作为推荐系统的矩阵，并且将M值的选择设置为TOP-K中K值的1.5倍。同时我们分别采用K值取1，3，5，7，10时对新用户推荐发电企业推荐算法的精确度，如图3.7所示，可以看出随着推荐电力企业数量的增多，推荐算法的精确度也随之增加，但是精确度从TOP-1到TOP-3，TOP-5提升十分明显，而超过推荐5个发电企业之后，推荐算法的精确程度提升并不会明显，这是由于用户企业会同时选择太多的发电企业原因造成的。

图3.7准确度随K值的变化

Fig.3.7Accuracy along with the change of K value

### 3.4.4 推荐系统准确率与召回率

针对本文提出的基于用户相似度协同过滤推荐算法，对矩阵填充前与填充后的准确度和召回率进行实验对比，选择矩阵填充90次的评分矩阵与填充前初始评分矩阵进行实验,同时，M值选择为K值的1.5倍。对于本文推荐算法的准确度与召回率计算过程利用公式（3.12）和（3.13）进行计算。

图3.8 矩阵填充前后精确度对比

Fig. 3.8 Comparison of Accuracy before and after matrix completion

图3.8和图3.9分别展示了随着K取值的增加，矩阵填充前后的推荐算法的准确度和召回率的柱状表，从整体看，推荐算法的准确度和召回率的总体趋势呈明显上升状态，当k值取到5时，准度度和召回率达到最大，表明我们的推荐算法的可行性；但是对于矩阵填充前的稀疏矩阵而言，填充后的评分矩阵对推荐算法的准确度和召回率明显优于填充前的推荐算法。表明我们对稀疏矩阵先进行评分填充这一过程是正确的。

图3.9 矩阵填充前后召回率对比

Fig. 3.9 Comparison of Recall before and after matrix completion

## 3.5 本章小结

本章主要针对电力市场交易模式的改变，如何对直购电用户推荐适合其自身需求的发电企业。首先，将现有已入网的直购电用户企业对自主选择交易的发电企业形成一个初始评分矩阵，用户与发电企业的交易次数经过归一化后设置为用户对发电企业的评分。其次，对形成的稀疏评分矩阵进行填充，填充方法是根据用户间的相似度以及用户对相应发电企业的评分拟合并迭代处理，最终形成完整的企业用户与发电企业评分矩阵。最后，根据填充后的评分矩阵对电力市场中新增的直购电用户进行推荐，将满足其需求的K个发电企业以推荐列表的形式推荐给企业用户。通过实验验证该推荐算法的准确性与可行性。

# 第4章 基于粒子群神经网络电力负荷预测

由于为用户推荐发电企业的时候需要考虑发电企业的电力负荷是否满足用户的额需求，因此需要对发电企业的下一个阶段的电力负荷进行预测，神经网络由于有很强的非线性拟合能力，可以映射任意复杂的非线性关系，同时它还具有很强的鲁棒性等优点，因此它非常适合对电力负荷进行预测[31]。于此同时在中长期负荷预测受到很多因素的影响如：天气变化，国家宏观调控，市场消耗的煤炭量等，这就需要我们建立好更合适的模型对电力负荷进行预测。优化BP神经网络算法以及对多种算法进行组合是算法改进的重要手段。本章根据电力负荷相关特点利用多种算法的组合对神经网络进行必要的优化。

本章首先将会介绍传统的神经网络算法在中长期电力负荷预测的应用所存在的相关问题，针对神经网络预测电力负荷存在的问题，采用粒子群优化算法改进并通过实验结果分析改进后的预测算法的准确性。

## 4.1 BP神经网络算法

### 4.1.1 BP神经网络模型

BP神经网络不仅是使用最广泛，也是最成熟和发达的神经网络技术，这种技术具有良好的非线性映射能力、自组织和自学习能力[9]。BP神经网络是一个上层网络通过全连通发送信号到下层网络的多层分层神经网络。当把学习样本提供给网络时，输入值从输入层通过隐藏层传播到输出层，可以通过神经网络输出层的神经元的响应得到与期望值的误差，将误差反向传播给隐藏层进行误差的反向传播，正确的训练BP网络往往会输出令人满意的答案。BP神经网络学习过程可分为两个部分：信号从输入层到输出层的正向传播和误差的反向传播过程。

(1) 正向传播

信号的正向传播过程是信息从输入层途经隐含层最后到输出层。在信号传播过程中网络中的权重是不变的，上一层的状态只会影响下一层的状态变化。如果输出的值满足不了期望输出的值，则转向误差的反向传播。

(2) 反向传播

误差值是指网络的实际输出和期望输出之间的差值，误差是由输出端逐渐的向后传播到输入端，这就是误差的反向传播。在误差向后传播的过程中，网络权值由误差反馈调节，通过调节网络中的权值使实际输出的值更加接近期望输出的值。

图4.1是一种常见的BP神经网络模型，其中输入层，隐含层，输出层以及之间的权重分别表示为(i=1,2,….,n)为网络中的输入值，(k=1,2,…,m)为网络中的输出值，为输入层到输出层的权重，为隐含层到输出层的连接权重。

图4.1 BP神经网络图

Fig. 4.1 The BP neural network diagram

### 4.1.2 BP神经网络算法

BP神经网络输入层的输入样例向量为，在隐藏层的输出向量为，在输出层的输出向量为,期望的输出向量被符号化为，从输入层到隐含层的权重矩阵定义为矩阵形式，如式(4.1)所示：

(4.1)

是指从输入层的第j个输入到隐藏层的第i个输出之间的权重，是指在隐藏层的第i个节点的权重向量，同样的，从隐藏层到输出层之间的权重矩阵也被定义为矩阵的形式，如式(3.2)所示：

(4.2)

是指从隐藏层的第k个输入到输出层的第j个输出之间的权重，是指在输出层第j个节点的权重向量。基于以上的这些定义，分别定义公式(4.3)和(4.4)：

( 4.3)

和

(4.4)

其中，非线性的和是传递函数，众所周知，BP神经网络的传递函数必须可微的和有界的。通过许多实践证明，Sigmoid函数是被广泛应用的，因此，和都可以表达为如下公式：

(4.5)

平方误差函数可以表示为如下公式：

(4.6)

因此，三层BP神经网络的权重调整公式可以被推导出如下公式：

 (4.7)

 (4.8)

其中，是学习速率。

针对本文电力负荷的预测，具体BP神经网络算法的步骤如下：

(1) 建立BP神经网络，设定输入节点，隐含层节点，输出节点个数，随机设定网络初始化权重和阈值，并设定学习步长。

(2) 输入一个电力数据的随机训练样本，数据包括电力负荷历史数据、影响因素（温度、工业用电量、煤炭价格、价格指数、GDP等）计算出向前各个隐含层输出值和输出层的值。

(3) 计算实际输出值和期望值之间的差值得出误差。

(4) 根据输出层和隐含层之间的权值连接反向传播的结果计算，并更新的值。

(5) 根据隐含层和输入层之间的权值连接反向传播结果计算出，并更新的值。

(6) 输入下一个样本训练，如果所有样本训练完毕，且训练后的误差值小于某个阈值则训练完毕，否则转向（2）。

## 4.2 PSO优化BP神经网络算法

粒子群（PSO）算法源于对鸟类觅食过程的研究[31,32]。研究者发现鸟类在觅食过程中经常突然改变，其运动轨迹总是不可预测，但总体上总是保持一致性。PSO基本思想就是根据此类过程总结而来：将所要求的变量值看成一个粒子，每个粒子在变量的维度以一定的速度进行搜索，通过适应度函数淘汰粒子，粒子根据自己最好的位置和其他粒子的位置，来调整飞行速度和移动位置，最终得出最优解。如下图为一个粒子群算法二维示例，显示粒子群算法搜索过程：算法开始时，初始化粒子群，输入每个粒子的初始位置和速度，评价每个粒子群的适应度，对于每一个粒子，把其当前适应值与其个体的历史最佳位置对应的适应值作比较，如果当前适应值比较高，则把当前适应值的位置更

新为历史最佳位置；对于每一个粒子，把它的当前适应值与全局最佳位置对应的适应值作比较，如果是当前适应值比较高，则把粒子的当前适应值更新为全局最佳位置。根据粒子群算法的公式更新粒子的速度与位置，通常算法停止的标志是迭代次数达到设定的最大值或者粒子的最佳适应值的增量小于设定的阈值。

图4.2 粒子群寻优过程

Fig. 4.2 Particle swarm optimization

### 4.2.1 BP神经网络算法的缺点

BP神经网络算法的基础是梯度下降法，梯度下降法所具有得线性收敛速度使网络在学习训练过程中次数过多最终导致训练时间过长。收敛速度过慢，这是BP神经网络的一个比较大的缺点之一。由于神经网络形成一个多维的空间，则很可能出现在求解最优值时出现局部最优的情况。从而可以总结出BP神经网络在电力负荷预测出现的问题[31,32]：

(1) BP神经网络学习算法在反馈调节权值和阈值过程中，是以负梯度下降方式来变化，所以当误差达到一定值时，这种调节就会停止。但是在预测电力负荷时，构建的神经网络计算最小误差时形成的最优函数，可能存在多个局部极小点的超曲面。这就使得BP神经网络在训练时只能收敛到某个局部最小值的情况。由于电力负荷可能受到多种因素影响，如天气，煤炭消耗，以及各种国家宏观调控等因素的影响。这就使得在进行构建BP网络时需要构建多个网络层和各层的多个节点。

(2) 在预测电力负荷时，运用BP神经网络进行迭代训练一般采用的增量训练法，这种方法需要将所有的样本的输入值放入网络中，训练后产生的全部误差再反馈到隐含层和输入层。此时再进行正向反馈，一直向这样迭代。在研究神经网络迭代过程中发现，如果想要网络稳定，需要大量的迭代次数，这就会出现这样的问题：网络稳定导致迭代次数多且向前传播过程中如果有多个隐含层时，导致迭代缓慢。

(3) 在运用BP神经网络预测电力负荷时，预测出来的值为连续值，而非常见的用于分类的离散值。在很多的文献中处理这种连续值时采用的是输出层只有一个节点的方法，但是这将很大程度上降低网络的复杂度，从而对预测结果的精确度有很大的影响。鉴于此中情况本文将在实验部分进行对连续值进行相关处理，让BP神经网络预测连续值时也能存在多个输出节点，增加网络的复杂度，从而使拟合的效果更佳好。

### 4.2.2 粒子群基本算法

粒子群算法(pso)受到鸟类和鱼类觅食的影响而产生[32,33,34]，它与不同于其他著名的进化算法（EA）的解决方案是通过探测搜索空间找到潜在的解决方案，但不包括操作符，通过进化算法的启发，来生成一个新的有前途的关于种群的解决方案。相反，粒子群算法中，人口（称为群）的每个个体（称为粒子）通过调整自己以及其通过拓扑领域的任何成员的历史最佳位置（称为Pbest），粒子之间分享不同种类的信息，在粒子群算法的全局区域内，整个种群被认为是领域，因此，在寻找有前途的位置信息时粒子受益于其他全局信息所有同伴的经验及共享信息。另外，有些区域的变体算法中粒子只利用自己的信息和相邻最好的邻居的信息。

在粒子群算法（PSO）中，每一个粒子有两个特征：它的位置和它的速度。假设在d维空间上第i个粒子，它当前的搜索空间的位置和速度可以分别被定义为和，第i个粒子的历史最佳位置被定义

为。

有不同类型的PSO算法，包括有惯性权重（GWPSO）的全局PSO算法、惯性权重（LWPSO）的局部PSO算法、收缩因子（GCPSO）的全局PSO算法和收缩因子（LCPSO）的局部PSO算法。

在GWPSO中，搜索者很受欢迎，有两种方法更新每个粒子的位置和速度。在整个群体中更新第K次迭代寻找最佳位置使用第一种方法，通过采用当前位置整个群体更新当前最佳位置使用第二种方法。

在第一种方法中，粒子i的第k次迭代的位置和速度更新公式如下：

 (4.9)

 (4.10)

在公式(4.10)中，w为惯性权重。和是正常数，分别代表认知度和社会参数。和是在区间[0,1]内随机选择的随机数，常数和代表每个粒子寻找当前最佳位置和全局最佳位置的的随机加速项的权重，通常设置为。

第二种方法中，gbest被代替，稍后显示，混合变量中的数值型例子中或者只有离散变量的问题中，相比于gbest更适合应用，即的成功率高于gbest。原因一是因为gbest的快速收敛，原因二是粒子无力逃脱gbest的局部最小值。换言之，因为离散变量快速收敛，一个有限的连续变量必须搜索特定的区域，但这个区域可能不是最佳搜索区域。

惯性权重w在粒子群算法的收敛行为中扮演重要的角色，惯性权重被用来控制当前粒子的速度受之前历史速度的影响，权重w使群体在全局搜索和局部搜索达到平衡的能力。在PSO中，当搜索持续时，惯性权重的下降线性公式为：

 (4.11)

其中和分别代表惯性权重的最大值和最小值。代表最大的迭代次数。有时随着粒子的振动搜索广泛，系统可能会面临奔溃的倾向，为了阻止这种倾向的产生简单定义了一个参数来控制在每一个d维空间的每个粒子i的速度，防止超过设定的速度。如果速度超过预定速度，则被修改为：

, then, (4.12)

(4.13)

### 4.2.3 粒子群优化神经网络算法步骤

电力负荷预测中，运用传统的BP神经网络减少预测误差时通常使用梯度下降法来调节权重和阈值，但是，BP神经网络存在许多不足之处，如收敛速度较慢且容易陷入局部极小值，网络的内外部结果难以确定等，针对电力负荷预测，前一节已经具体描述了应用BP神经网络算法的缺点。对于BP神经网络的缺点，几年来许多学者提出了各种改进的方法，比如采用自适应的调参，运用启发式的算法优化BP神经网络等。

粒子群优化算法实质上是一种并行的全局性随机搜索算法，它比遗传算法更加简单的同时搜索能力更加快速等特点。而对神经网络的优化实质上是对误差函数的寻优的优化，粒子群优化算法的优越性和特点如下：

(1) PSO算法简单，只是内部调节粒子的位置和速度，并不需要调节很多的参数，所以更简单易行。

(2) PSO算法本身收敛速度快，搜索能力具有很好的并行性和全局性，而且不需要目标函数是连续的，所以，PSO优化算法可以快速接近最优解。

(3) PSO算法符合自然规律，且其并不需要借助算法的本身的特征信息，比如梯度。

(4) PSO算法同样具有标准BP算法的自学习能力、快速寻找问题的最优解以及联想存储的功能。

这就为两者之间的结合提供了一个很好的思路。粒子群和神经网络的结合主要有以下两种方式：

(1) 利用粒子群的全局搜索能力优化神经网络拓扑结构，连接权重，粒子群算法结合神经网络算法提高网络的学习能力以及提高神经网络的全局优化能力和局部搜索能力，加速整体搜索能力。

(2) 神经网络嵌入在粒子群优化（pso）算法中，神经网络学习能力可以优化粒子群算法的性能，从而提高了粒子群的收敛速度，同时还可以减少计算量。

因此，PSO优化算法改进神经网络并不会因为两个算法的特点而冲突，通过结合可以改进BP网络的性能，本文的电力负荷预测采用粒子群优化算法改进BP神经网络。

PSO是一种优化算法，利用PSO算法进行网络的优化实质上就是一个寻优过程，更加具体来说就是进行向前网络的计算，向后网络反馈时进行优化。这其中涉及到了两个主要问题就是粒子与网络中的权值的关系以及粒子群中如何构造适应度的问题。

PSO优化BP神经网络的关键两点如下：

(1) 首先，将PSO粒子个体与BP神经网络连接权值之间产生一种产生映射关，即粒子群中每个个体怎样表示成求解BP神经网络优化函数的权值。

(2) 其次，将神经网络中实际值与预测值的均方误差作为粒子群中的适应函数，通过PSO算法优化过程求解全局位置。

在三层BP神经网络结构图中，网络结构的输入层节点个数为M,假设隐含层节点个数为N，输出层节点个数为Q。粒子群对神经网络的优化一方面体现在同时优化输入层到隐含层，隐含层到输出层之间的权重，他们分别为和，从而可以使用相应的编码规则对粒子群个体与权重更新产生映射。则需要的每个粒子维度D为

(4.13)

其中表示为输入层中存在的权值+阈值个数，同样表示为隐含层到输出层中的权值+阈值个数。

对于适应度函数的设计过程，粒子群优化神经网络中对于权值的更新是一个反复迭代的过程。通常为保证实现的算法的评估准确程度，往往将数据集拆分成两个部分，训练集和测试集。评估准确性即个体在群体中的优劣程度就需要适应度。在PSO优化神经网络的适应度如下：

(4.14)

其中n为训练样例的个数，为第样例的实际值，对应的预测值。从表达式上看，适应度越小则说明预测出来的误差就越小。

图4.3 PSO优化BP神经网络过程

Fig. 4.3 PSO optimization BP neural network process

## 4.3 电力负荷预测实验结果分析

根据前文对电力负荷预测方法BP神经网络和粒子群优化算法的概述，本节内容主要是根据粒子群优化算法改进BP神经网络应用到电力负荷预测中。

### 4.3.1 数据集选取

本文主要针对中长期的电力负荷进行预测，使电力负荷预测到的结果应用到接下来的推荐算法中。电力负荷是一个不稳定的动态变化过程，除了受用户的用电量影响，还受到其他多种外界影响因素的影响。本章的电力负荷预测中历史电力负荷数据及其他影响因素来自项目“电力交易信息采集与数据挖掘技术研究与开发”（批准号：2014-0-1-16489），关于影响因素中天气的数据是利用爬虫技术在中国天气网等网络中获取的。统计了2013年到2016年的数据，将2013到2015年的数据作为训练集，2015到2016年的数据作为测试集。下表为数据样本表。

表4.1 某地区中长期的电力负荷样本表

Table 4.1 Long term electric load sample table in a certain area

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 年份 | 价格指数  （分/度） | 汽车增长（万辆） | 工业用电量（103Mwh） | 高低温天气（天数） | 钢材消耗（万吨） | 对外经济（亿） | 电力负菏(Mwh) |
| 2013.01 | 102.4 | 7.6 | 86.43 | 20 | 267.45 | 27.1 | 23469 |
| 2013.02 | 103.6 | 20.21 | 66.95 | 24 | 258.36 | 19.6 | 21444 |
| 2013.03 | 102.5 | 14.49 | 79.91 | 4 | 244.51 | 29.3 | 19220 |
| 2013.04 | 102.5 | 18.32 | 83.45 | 5 | 253.12 | 23.96 | 18464 |
| 2013.05 | 102.5 | 21.54 | 90.76 | 5 | 286.23 | 28.84 | 17667 |
| 2013.06 | 103 | 13.21 | 99.5 | 12 | 315.21 | 28 | 19100 |
| 2013.07 | 102.9 | 21.96 | 117.44 | 21 | 349.24 | 37.18 | 26989 |
| 2013.08 | 102.9 | 15.23 | 101.23 | 23 | 321.12 | 36.73 | 26561 |
| 2013.09 | 103.1 | 20.84 | 75.12 | 3 | 316.29 | 34.51 | 25407 |
| 2013.10 | 103.1 | 24.63 | 92.54 | 1 | 295.62 | 30.27 | 17551 |
| 2013.11 | 102.9 | 16.48 | 95.33 | 7 | 302.15 | 34.61 | 19712 |
| 2013.12 | 102.5 | 18.19 | 117.34 | 26 | 298.45 | 34.25 | 23915 |
| … | ... | … | … | … | … | …. | … |

### 4.3.2 数据规范化

数据规范化的目的是所有数据的度量单位都会对数据分析产生影响，从而可能导致的结果完全不同。比如，数值属性具有较大的值域可能是由较小的单位表示该数值属性而导致的，那么结果就是这样的属性会趋向于具有较大的“权重”，使其对结果的影响力较大。为了避免这种情况，应该对数据进行标准化或规范化，涉及到变换数据，使处理的数据落入较小的共同区间内。

对于本文需要处理的天气影响因素而言，由于高温和低温天气对电力负荷的影响较大，而温度适应（10-26°）的天气则影响相对较小。因此对天气高低温天数的处理方法是分别统计每个月的高温与低温天气数目X和Y，总天数为N，最终结果D映射到[0,1]区间内。公式如下：

 (4.15)

对于上表中其他数据的规范化处理是对原始的数据进行一个线性变换，设和分别表示属性A的最大值和最小值，最终将属性A的值x映射到中的。其数据最大值与最小值规范化公式为：

 (4.16)

### 4.3.3 BP神经网络的结构设计与参数设置

对于BP神经网络的网络结构设计并没有固定的可遵循的步骤，许多的参数也是靠经验获取，并通过大量实验进行比较。一般而言，BP神经网络通常采用三层网络，其可以任意的逼近一个非线性函数。由于BP神经网络的简单可行、计算量小等特点，许多多层网络的应用实例仍采用该网络结构进行训练，在本文的电力负荷预测中，采用三层神经网络模型对电力负荷进行预测。

神经网络中的输入层和输出层的神经元个数是由所解决问题的本身而决定的。在本章关于电力负荷预测的实验中，输入了6个输入属性，目标数据的维度为3。对电力负荷预测，BP神经网络的输出层节点为1，对神经网络进一步优化过程中将输出层节点个数设置为3，这样做的目的在于增加网络的复杂度，增强预测能力,网络结构的每一层采用的激活函数为Sigmoid函数。而确定隐含层的神经元数目相对而言较为复杂，其数目的确定与拟解决问题的输入和输出神经元多少有直接的关系。如果隐含层的神经元数目较多，则全局最小点越容易收敛获取且局部的极小值越少，同时使神经网络的映射能力越强，但是隐含层的神经元数目过多，会使神经网络因为训练时间过长导致神经网络的容错性能降低；如果隐含层的神经元数目较少，则使算法在训练过程中不能达到收敛。通过实验证明下列公式可以确定隐含层的神经元数目H，其中I表示输入层神经元数目，O表示输出层神经元数目，为[1,10]区间的常数。

 (4.15)

表4.2 隐含层个数对误差值的影响

Table 4.2 The influence of the number of hidden layers on the error value

|  |  |
| --- | --- |
| 隐含层神经元个数 | 训练后的达到的误差 |
| 4 | 0.009 |
| 6 | 0.0043 |
| 8  10  12 | 0.0035  0.0012  0.023 |

通过实验证明，当隐含层神经元数目为10时，电力负荷的预测值与实际值误差最小。

粒子群中参数有种群规模M,惯性因子，学习因子和和区间速度，在本章中种群规模M=100。粒子维度值在本文设置的。通常情况下，学习因子和一般取相同的值2。实验表明，如果取得恰当的惯性因子，学习因子和，它的收敛的速度将同使用相差不大。因此本文选取的子，。算法终止迭代次数设置的尽可能高，目的是为了实验中对个算法的性能进行有效地测试。

### 4.3.4 PSO-BP预测电力负荷实验结果

本节内容主要是通过改进的PSO优化BP神经网络算法以及没有优化的BP神经网络算法对经过处理的实验数据进行实验。

图4.4 2015年度电力负荷预测值与实际值

Fig. 4.4 2015 Annual electricity load forecast and actual value

图4.4所示为通过对比PSO-BP算法和BP神经网络算法，选取了2015年度的电力负荷实际值与预测值进行对比。两种算法的预测电力负荷曲线与真实值曲线总体趋势上是相适应的，但是通过两种算法的结果对比，PSO优化BP神经网络算法效果远远好于BP算法，基本上是顺应实际值的趋势。预测的电力负荷值经过归一化后，与实际值误差在最小可达到0.008；而BP神经网络算法预测值与实际值误差最小为0.1026且根据图示显示，在预测过程中BP神经网络算法预测的电力负荷值与实际值之间波动较大。

图4.5 PSO-BP与BP预测电力负荷的误差曲线

Fig. 4.5 PSO-BP and BP forecast power load of the error curve figure

从图4.5的展示结果表明，在对训练集的学习过程中，两种算法都进行了不断寻找最优值的过程。而从始至终，从图示中可以看出，PSO优化BP神经网络时在迭代34次时误差值有明显的下降趋势，在迭代90次时，PSO优化算法就已经使误差值降到0.01以下并一直降低直到误差值达到平稳，说明了PSO优化BP神经网络算法无论在寻优速度还是在误差变化中，都比BP神经网络算法性能更优。

## 4.4 本章小结

本章内容主要是针对电力交易推荐系统中大用户与发电企业电量匹配中关于电力负荷的预测，利用BP神经网络算法进行电力负荷预测，但是由于BP神经网络算法容易陷入局部最优以及收敛速度过慢而导致预测结果并不理想，因此，采用粒子群优化算法对BP神经网络模型进行优化，通过实验证明，优化的BP神经网络算法对电力负荷的预测结果更加精确。

# 第5章 电力交易推荐系统的移动端实现

现如今，随着移动网络的普及，国内4G网络正式启动且不断发展是移动网络速度甚至超越了宽带技术。截止到2012年，智能手机的单位出货量已经超越了台式计算机和笔记本电脑出货量的总数，互联网产品也逐渐将PC端应用网络发展为移动端应用网络的开发[35,36]。基于此，本文设计了基于Andriod平台实现了电力交易推荐系统。使用户与企业可以通过接入移动网络及无线网络随时随地进行交易，不必受限于PC端应用的网络与地理位置的局限性。电力交易推荐系统主要是通过针对大用户与发电企业的交易形式与交易数据，产生一个“特殊”的评分矩阵。对于发电企业而言，分析发电企业的历史电力负荷数据及影响电力负荷的因素，预测下一个阶段的短期电力负荷，将预测得到的电力负荷应用到推荐系统中；对于大用户而言，分析大用户之间的自身属性相似度和需求属性相似度，并结合其他大用户对发电企业的评分，从而给目标大用户提供一个关于发电企业的推荐列表，使大用户有一个良好的交易体验。良好的推荐结果可以增加该电力交易系统中大用户的数目，同时也会为日后的发电企业带来更多商业价值。总体来说，一个完整的电力交易推荐系统的移动端的实现通常由三个部分组成：客户端的页面显示、预测电力负荷算法与推荐算法实现模块和后台的数据库模块。

## 5.1 系统架构设计

本文的电力交易推荐系统主要包括移动端界面展示、推荐模块、数据存储[37]三个方面。用户的界面展示部分主要给大用户提供对“商品”的体验以及将推荐列表显示给大用户。推荐模块主要包括PSO优化神经网络算法预测电力负荷和基于大用户属性相似性协同过滤推荐算法组成，属于系统的离线处理部分。数据存储部分主要存储系统中大用户与发电企业的信息，以及记录交易信息，同时将推荐模块处理的推荐结果存储在数据库中。最后，还有一个在线处理部分，主要功能是及时对大用户产生更好地推荐结果。电力交易推荐的移动系统总体系统架构如图5.1所示。

图5.1 电力交易推荐系统架构设计图

Fig. 5.1 The architecture design of electricity trading recommendation system

整个电力交易推荐系统的架构流程如图所示，下面针对每个流程进行详细的讲解：

(1) 大用户打开该系统的移动端APP时，用户界面会根据其曾经的交易情况为其产生一个关于发电企业的推荐列表，大用户可以根据推荐列表中的信息对发电企业进行浏览查看发电企业的各项信息，并选择与选中的发电企业进行线下撮合交易。进而产生了交易记录。

(2) 根据大用户与发电企业的交易记录，在线处理模块开始工作，计算大用户间的需求属性和自身属性相似度，并将交易模式转换为用户与发电企业的评分数据，调整基于用户的推荐算法产生的推荐列表。同时将交易记录保存到数据库中。

(3) 在线处理后的信息及时保存到数据库中，同时根据不同比例从数据库中获取推荐列表中基于用户相似度推荐的前K个发电企业供大用户浏览其信息。

(4) 后台离线方式产生推荐列表，针对发电企业采用PSO优化BP神经网络算法对其下一阶段的电力负荷进行精确预测，以用于接下来对大用户推荐时的电量匹配。针对大用户需要基于大用户之间的相似性以及大用户对发电企业的评分产生推荐结果。这些计算的结果利用数据库中的发电企业与大用户的注册信息与需求信息计算获得。计算得到的结果再次存储到数据库中。这一流程的目的是将存在数据库中的推荐列表推送到用户界面，展示给电力交易系统移动端的用户，供其浏览选择。

以上内容是整个电力交易推荐系统的移动端的整体数据流架构方案。

## 5.2 电力推荐系统的移动端的具体实现

本文的电力交易推荐系统的移动端的具体实现过程中，系统的服务器端使用目前较为流行的Web应用程序开源框架ssh（struts+spring+hibernate）框架，采用Javaweb语言实现整体框架；客户端使用Andriod开源框架实现；数据库使用Mysql实现。该框架使用的是MVC模型，即分别为模型层（Model）、视图层（View）和控制层（Cotroller[36,37]。

### 5.2.1 系统架构层次设计与实现

电力交易推荐系统每个层次的设计和实现如图5.2所示，使用了MVC的经典设计模式Struts，通过采用JavaServlet技术实现了基于Java EEWeb应用的MVC设计模式应用框架。同时还使用了Hibernate开源对象关系映射框架，将数据库进行非常轻量级的对象封装，可以使用对象封装的思想操纵数据库。本系统采用将Struts和Hibernate两个框架结合起来进行MVC架构的管理协调。下面分别详细介绍MVC的每一层：

* 视图层

视图层是对用户显示数据的一个展示层，这一层中显示用户的注册，登录等操作功能界面，同时用户可以发布需求，浏览发电企业，查看推荐列表，选择满足自身需求的发电企业进行交易等功能。针对本系统的移动端视图，使用的视图框架为Andriod自身的andriod.View包来实现整个系统视图的界面展示。

* 控制层

控制层是调用业务层的相关代码来实现与数据库的交互及其他相关功能，以捕获用户请求并控制请求转发。本系统在控制层采用了Struts，struts提供了一个控制器组件ActionServlet，它继承自HttpServlet，并重载了HttpServlet的doGet（）和doPost（）方法，可以接受HTTP的响应，并进行转发，同时还提供了使用XML进行转发Mapping的功能。

* 模型层

数据模型层主要是负责业务逻辑以及数据库的交互。使用hibernate和Mysql数据库结合的技术，hibernate可以将对象映射到数据库中的数据表，并对数据表进行操作，实现封装对象的思想来操作数据库。选用Mysql数据库的原因是它是一个关系型的数据库，这种关联数据库可以将数据保存在不同的表中并建立关系，而不是把所有的数据都要放到一个大仓库中，提高了数据库的灵活性并增加了速度。

图5.2 MVC模型图

Fig. 5.2 MVCmodle

### 5.2.2 系统数据库的设计与实现

选用Mysql数据库作为本系统的数据仓库，Mysql属于关系型数据库，把这种“关系型”理解为表格的概念。那么一个关系型的数据库便是由一个或者多个表格组成。表头（header）表示每一列的名称；列（row）表示相同的数据类型的数据集合；行（col）用来描述用户或者事物的详细信息；值（value）指的是行的具体信息，且每一个值必须与该列的数据类型相同；键（key）是数据表中用来识别一个特定的用户或者事物的方法，在当前列具有唯一性。图5.3是本系统的数据库逻辑设计图。

图5.3 数据库整体逻辑图

Fig. 5.3 Database integrated logic diagram

用户表中记录了大用户的注册信息，包括其所处地理位置，企业规模，用电峰值等，其中大用户的ID是交易记录表和推荐表的外键；发电企业表中记录了发电企业的信息，其ID也是其他两个表的外键。推荐列表中存放了为大用户提供的推荐发电企业的ID及发电企业的ID列表。交易记录中主要记录大用户与发电企业的交易行为，为离线的推荐技术提供数据源。

### 5.2.3 系统功能模块的设计与实现

电力交易推荐系统移动端的各个功能的设计和实现过程其实就是控制层（Controller）的设计和实现。每一个模块的功能都由一个Controller类表示，实现每个类的过程就是实现每个模块的过程。针对本文的电力交易推荐系统的移动端的功能模块分为：大用户模块、发电企业预测模块和推荐模块。

图5.4 功能模块设计图

Fig. 5.4 The function of module design

用户模块主要处理该系统中的两种用户的相关功能操作：大用户和发电企业。

如实现这两类用户的登陆和注册等功能。针对发电企业预测模块，主要是对发电企业的通过pso优化神经网络算法实现电力负荷的预测，并将发电企业的信息展示出来。推荐模块主要是在线处理用户的推荐列表，实时计算大用户间的属性和需求相似度，并根据计算结果适时调整评分矩阵中的评分信息，将评分较高的前K个发电企业推荐给大用户。而基于用户属性相似度的协同过滤算法得到的推荐结果依旧按照之前的排序推荐给大用户，变化的只是基于用户属性推荐的列表内容。

### 5.2.4系统界面设计与实现

该系统的界面的设计与实现即实现MVC中的视图层（View），电力交易推荐系统主要是给大用户推荐发电企业，所以对于移动端的界面设计并不需要很复杂，但是需要体现出电力交易推荐系统这一平台的功能。该系统界面采用Andriod自带的界面框架，整个界面的框架结构使用ViewGroup类别组件来实现，为实现过程提供了丰富的控件、样式等，使系统的界面设计方便快捷，无需浪费太多的时间。如图5.5所示为该移动端的部分界面，电力交易系统的主页要突出该系统的核心目的，通过综合考虑用户的需求以及掌握电力交易的关键业务，将电力交易系统移动应用的功能主要分为电力大用户与发电企业交易、针对大用户的推荐服务、个人中心等三方面功能。通过设置tab标签作为导航，并点击标签来实现每个功能的切换。将电力交易业务作为默认的移动端应用首页，包括市场成员管理（大用户与发电企业）、信息发布管理、交易管理和结算管理；对于推荐模块，将前期通过推荐算法处理的结果以列表的形式展示给用户企业，供其浏览选择，与选中的发电企业进行交易数据申报，提交申报数据后与相应的发电企业进行线下撮合交易；对于个人中心模块主要展示用户的注册登录等功能。

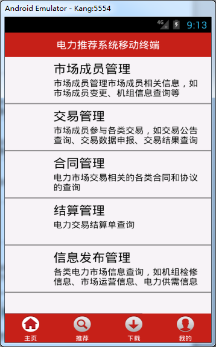


图5.5 部分显示界面

Fig. 5.5 Partial display interface

## 5.3 本章小结

本章主要是设计了基于Andriod平台的电力交易推荐系统。使用户与企业可以通过接入移动网络及无线网络随时随地与发电企业进行交易，不必受限于PC端应用的网络与地理位置的局限性。该移动端的整体架构主要包括移动端界面展示、推荐模块、数据存储三个方面。移动端的具体实现中，服务器端使用Web应用程序开源框架ssh（struts+spring+hibernate）框架，采用Javaweb语言实现整体框架；客户端使用Andriod开源框架实现；数据库使用Mysql实现。整体框架使用的是MVC框架模型，即分别为模型层（Model）、视图层（View）和控制层（Cotroller），分别对每一层具体实现。最终实现电力交易推荐系统的移动端。

# 第6章 总结与展望

## 6.1 总结

本文主要围绕电力改革中电力交易方式模式的改变为背景，分析了国内外电力市场的改革过程，并描述了推荐系统的相关发展历程和推荐算法的相关技术，同时，介绍了电力负荷预测的相关方法。接下来，整个论文围绕背景内容设计一个电力交易推荐系统移动端，以供企业用户选择直接交易的发电企业，实现双方互惠互利。其中，移动端的设计是为了用户可以随时随地接入移动网络或者无线网络进行查看与交易，体现了移动设备的灵活性以及普适性。在该移动平台上，实现电力负荷预测算法，采用粒子群优化神经网络的算法进行预测发电企业的中长期电力负荷，将预测到的电力负荷应用到接下来的发电企业推荐中。同时，在该移动平台上实现对电力市场中具有直购电规模的用户企业推荐适合其需求的发电企业的功能，主要是基于用户的协同过滤推荐方法，针对交易模式的改变将交易数据初始化为评分矩阵，矩阵中的评分数据是企业用户自主选择的交易次数归一化产生，由于电力改革启动时间较短，并没有丰富的数据供我们使用，导致了评分矩阵的数据稀疏化，本文提出了矩阵填充方法，填充后的完整评分矩阵进行入网用户的推荐，矩阵填充方法是基于用户相似度与用户选择相似性拟合并迭代处理，最终利用填充后的矩阵进行推荐系统的实现。

## 6.2 展望

虽然本文在电力交易推荐平台上实现了推荐，预测及移动端等功能，但是由于本人才疏学浅，对相关知识了解的不够透彻，使整个研究还不够完善，对于本文的进一步有待完善的内容为：电力交易模式的改变，势必会导致电力数据的增多，导致海量数据的产生，针对这种海量数据的现象，下一步需要应用分布式计算模型实现电力负荷的预测及电力用户间的推荐；同时，移动端的设计也要考虑用户数量爆发后的承载能力，进一步优化该系统；最后，对于推荐算法的应用，需要考虑到不仅可以应用到电力交易推荐中，也要应用到其他领域。

# 参考文献

[1] Yang X, Guo Y, Liu Y, et al. A survey of collaborative filtering based social recommender systems[J]. Computer Communications, 2014, 41(5):1-10.

[2] Milicevic A K, Nanopoulos A, Ivanovic M. Social tagging in recommender systems: a survey of the state-of-the-art and possible extensions[J]. Artificial Intelligence Review, 2010, 33(3):187-209.

[3] Jahrer M, Töscher A, Legenstein R. Combining predictions for accurate recommender systems[C].ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Washington, Dc, Usa, July. 2010:693-702.

[4] 新闻中心，<http://www.hbhygnfd.com/new_detail/id/133.html>

[5] 罗朝春.二级电力市场交易模式与方法研究[D]. 华中科技大学, 2005.

[6] 刘树杰, 罗圣华. 电力市场设计:深化电力体制改革的前提[J]. 宏观经济管理, 2014(12):37-38.

[7] 文安, 刘年, 黄维芳,等. 英国电力市场的价格机制分析[J]. 南方电网技术, 2015, 9(1):1-6.

[8] 崔强. 国外电力工业体制改革[M]. 中国电力出版社, 2001.

[9] 赵菁, 许克明. 神经网络和模糊理论在短期负荷预测中的应用[J]. 电力系统及其自动化学报, 2010, 22(3):129-133.

[10]陈哲, 冯天瑾. 小波分析与神经网络结合的研究进展[J]. 电子与信息学报, 2000,

[11]项亮，推荐系统实践[M].北京：人民邮电出版社，2012.1-30

[12]Dietmar Jannach，Markus Zanker(著)，蒋凡(译)，推荐系统[M]. 北京：人民邮电出版社，2013,9-20

[13]Yang X, Steck H, Liu Y. Circle-based recommendation in online social networks[C]. ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2012:1267-1275.

[14]Yu L, Pan R, Li Z. Adaptive social similarities for recommender systems[C]. ACM Conference on Recommender Systems. ACM, 2011:257-260.

[15]Jamali M, Ester M. A matrix factorization technique with trust propagation for recommendation in social networks[C]. ACM Conference on Recommender Systems, Recsys 2010, Barcelona, Spain, September. 2010:1055-1066.

[16]MaH, King I, Lyu M. Learning to recommend with social trust ensemble[C]. International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2009, Boston, Ma, Usa, July. 2009:203-210.

[17]孟祥武, 胡勋, 王立才,等. 移动推荐系统及其应用[J]. 软件学报, 2013, 24(1):91-108.

[18]Tintarev N, Flores A, Amatriain X. Off the beaten track: a mobile field study exploring the long tail of tourist recommendations[C]. Conferenceon Human-Computer Interaction with Mobile Devices and Services, Mobile Hci 2010, Lisbon, Portugal, September. 2010:209-218.

[19]Gavalas D, Kenteris M. A web-based pervasive recommendation system for mobile tourist guides[J]. Personal and Ubiquitous Computing, 2011, 15(7):759-770.

[20]Kantor P B, Rokach L, Ricci F, et al. Recommender Systems Handbook[J]. 2011.

[21]张立毅, 刘婷, 孙云山,等. 遗传算法优化神经网络权值盲均衡算法的研究[J]. 计算机工程与应用, 2009, 45(11):162-164.

[22]刘章, 蒋杰, 罗日成,等. 基于小波神经网络的绝缘子泄漏电流监测装置的设计[J]. 电气工程应用, 2012(2):29-33.

[23]谭健雄. 基于社交网络用户特征的协同过滤推荐算法[J]. 科技致富向导, 2015(6):32-32.

[24]火生旭. 协同过滤推荐算法在社交网络中的研究与应用[D]. 湖南大学, 2013.

[25]Kong W, Liu Q, Yang Z, et al. Collaborative filtering algorithm incorporated with cluster-based expert selection[J]. 2012, 9(12):3421-3429.

[26]Kong W, Liu Q, Wang S, et al. Relation-based collaborative filtering algorithm[J]. Journal of Computational Information Systems, 2012, 8(15):6257-6265.

[27]Ruchansky N, Crovella M, Terzi E. Matrix Completion with Queries[C]. ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2015:1025-1034.

[28]Keshavan R H, Oh S, Montanari A. Matrix completion from a few entries[J]. IEEE Transactions on Information Theory, 2009, 56(6):2980 - 2998.

[29]Király F J, Theran L, Tomioka R. The Algebraic Combinatorial Approach for Low-Rank Matrix Completion[J]. Journal of Machine Learning Research, 2012, 62(2):299-321.

[30]Wen Z, Yin W, Zhang Y. Solving a low-rank factorization model for matrix completion by anonlinear successive over-relaxation algorithm[J]. Mathematical Programming Computation, 2012, 4(4):333-361.

[31]Cheng W, Feng P. Network traffic prediction algorithm research based on PSO-BP neural network[C]. International Conference on Intelligent Systems Research and Mechatronics Engineering. 2015.

[32]Gudise V G, Venayagamoorthy G K. Comparison of particle swarm optimization and backpropagation as training algorithms for neural networks[C]. Swarm Intelligence Symposium, 2003. Sis '03.Proceedings of the.2003:110--117.

[33]Shi X H, Liang Y C, Lee H P, et al. Particle swarm optimization-based algorithms for TSP and generalized TSP[J]. Information Processing Letters, 2007, 103(5):169-176.

[34]Chen J, Qin Z, Liu Y, et al. Particle Swarm Optimization with Local Search[C]. International Conference on Neural Networks and Brain, 2005. Icnn&b. 2005:481-484.

[35]张舒. 基于Android的虚拟交友社区研究与实现[D]. 北京邮电大学, 2010.

[36]孙晓宇. Android手机界面管理系统的设计与实现[D]. 北京邮电大学, 2009

[37]李恒新. 基于组合推荐技术的音乐推荐引擎研究与实现[D]. 广东工业大学, 2015.

.

# 致 谢

两年半的研究生生涯马上结束了，在这期间，我收获了很多：敬爱的老师，友爱的同学，浓烈的学习氛围。东北大学曾经就像是一个遥远的梦，从来没有想过我会来到这所高校学习，让我体会到了一所好的大学所带来的好的资源。虽然只有短短的两年半的时间，却也让我获益匪浅，让我对东北大学以及计算机科学与工程学院有着特殊的感情，将会为我未来的道路添加最精彩的一笔。

首先，感谢我的研究生导师焦明海副教授，焦老师为人和蔼亲和，对学生关怀有加。在两年半的研究生生涯中，焦老师对我学习上帮助很多，经常帮我解答我表示困惑的地方，尤其是在毕业设计中，焦老师努力与我一起讨论，让我有更明确的目标；感谢张岩峰老师在毕业设计中和每次的组会中对我的指导和帮助，虽然我每次对论文的理解都达不到老师的要求，但张老师依旧很耐心的和我交流和讨论。经过两年半的组会讨论，我也学到了很多知识。张老师无论对学术的严谨还是生活的热心都有他独特的魅力，身边的每位同学对张老师都充满了崇拜之情；同时感谢于戈老师，有机会认识于老师是我的荣幸，在于老师身上学到凡事都要悉心学习，无论在何种场合遇到于老师，都会与我们聊上几句，从生活到学习都会关怀备至，同学们对于老师完全敬佩和尊重。

其次，感谢我的实验室同学李帅，付笑晗，李玉国，王祝兵，他们是我生活中的好友，学习上的老师，对我有很大的帮助，每次遇到困难都会帮我一起想办法解决。在他们身上，我看到的更多的是阳光向上的态度和积极进取的决心，遇到他们是我研究生生涯中最值得的事情，希望他们有一个美好的未来。同时，还要感谢我的室友，虽然她们说我是她们的开心果，但是如果没有她们的帮助与鼓励，我也不会那么开朗向上。

最后，感谢我的父母，感谢他们将我养育成人，不论生活多么艰难，父母都坚持供我完成学业，二十多年的学生生涯即将结束，父母也因为供我和弟弟读书而老去，感谢他们的伟大与无私，我会继续努力，成为他们的骄傲。

# 攻读硕士学位期间的论文和项目情况

1.2014.12-2017.12，教育部留学回国人员科研启动基金项目：基于服务的云数据资源优化关键理论与算法研究(项目编号：49-1)。

2.2015.01-2016.12，企事业单位委托科技项目：电力交易信息采集与数据挖掘技术研究与开发，(项目编号： 2014-0-1-16489)。