

摘要

负荷预测作为智能电网中的一个重要模块,为电网、电源规划建设提供信息支撑,为电网调度和电力消费者的用电决策制定提供数据依据,是保障电力系统安全运作,节能减排,提高效益的重要工作之一。随着智能电网功能的进一步开发,负荷预测的对象不断细分,加上电网通讯系统的完善和信息采集设备的普及,预测对象从系统级负荷向住宅级负荷延伸。相对于系统级负荷,住宅级负荷具有更高的非线性,加之配电网中时变负荷需求的增加,使得原本复杂的住宅级负荷数据变得更加难以预测。如何提高住宅级负荷预测的精确度成为了众多学者研究的重点。

随着计算机计算能力的增长,一些原本很难实现的数据处理方式得以应用于负荷预测当中。近年来,深度学习的兴起为负荷预测提供了更多的方法,其强大的非线性处理能力使其逐渐取代传统的负荷预测模型,成为主流。本文针对住宅级负荷难预测的问题,提出了基于长短期记忆(Long-Short Term Memory, LSTM)网络的多目标优化组合模型。该模型以 LSTM 为基础,构建深度学习模型,辅之以多特征输入来提高网络的精度。使用基于分解的多目标优化算法(Multi-objective Evolutionary Algorithm Based on Decomposition, MOEA/D)对 LSTM 网络参数进行优化,得到多个优质的子模型。使用基于深度信念网络(Deep Belief Network, DBN)的结合策略,对多个子模型的输出进行集成学习,形成组合预测,以提高预测模型的泛化能力。具体工作内容如下:

(1) **总结住宅级负荷的特性。**以澳大利亚智能电网智能城市(Smart Grid Smart City, SGSC)项目提供的住宅用户负荷数据为基础,将住宅级负荷与系统级负荷进行对比,突显出住宅级负荷的特性,探索预测模型的重点优化方向。

(2) **构建基于多特征输入的数据预处理框架。**由于对住宅级负荷波动产生影响的因素较多,为了使预测模型充分挖掘数据中潜在的相关性,本文使用电能消耗分量、时间分量、周分量以及节假日分量组成多特征输入。希望通过增加输入数据的特征维度,来增加预测模型可以学习到的隐藏知识,以提高模型的预测精度。依托 SGSC 提供的数据,将多特征输入框架与 LSTM 网络相结合进行实验,与传统预测模型、深度学习模型进行比较,证明多特征输入框架能提高预测模型的整体精度。

(3) **提出基于 MOEA/D 的子模型生成策略。**为了提高网络的整体性能,使用 MOEA/D 算法对网络的参数进行优化。以 LSTM 网络的精确度和差异为目标空间,以 LSTM 网络的超参数为决策空间,对网络进行优化。以这种自适应的方

式，使得网络超参数自动适应负荷数据，避免了人为设置参数或者设置固定参数所造成的误差。同时，为了使 MOEA/D 算法与深度学习框架更加契合，本文对 MOEA/D 算法进行改进，并通过对比试验说明了改进的有效性。

(4)提出基于 DBN 的结合策略。使用 DBN 网络，将多个子模型的预测结果组合成最终预测结果，通过这种类似于集成学习的方法克服单个网络进行预测时的不稳定因素，以达到缓解数据过拟合的效果，提升模型的泛化能力。本文将基于长短期记忆网络的多目标优化组合预测模型与现有的多种智能预测模型进行实验对比，说明了本文提出模型的先进性。

关键词：用户侧负荷预测；多特征输入；多目标优化算法；长短期记忆网络；组合模型

Abstract

As an important part of smart grid, load forecasting provides information support for power grid and power supply construction planning, and provides data basis for power grid dispatching and power consumption decision-making of power consumers. It is one of the important tasks to ensure the safe operation of the power system, energy saving and emission reduction, and improve efficiency. With the further development of smart grid function, the object of load forecasting is continuously subdivided, coupled with the improvement of power grid communication system and the popularization of information acquisition equipment, the forecasting object extends from system level load to residential level load. Compared with system level load, residential level load has higher nonlinearity, and the increase of time-varying load demand in distribution network makes the complex residential level load data more difficult to predict. How to improve the accuracy of residential load forecasting has become the focus of many scholars.

With the growth of computer computing power, some data processing methods that are difficult to realize can be applied to load forecasting. In recent years, the rise of deep learning provides more methods for load forecasting. Its powerful nonlinear processing ability makes it gradually replace the traditional load forecasting model and become the mainstream. Aiming at the problem of high nonlinearity and difficult prediction of residential load, a multi-objective optimal combination model based on long short-term memory (LSTM) network is proposed in this paper. Based on LSTM, the model constructs a deep learning network, supplemented by multi feature input to improve the accuracy of the network. Multi objective evolutionary algorithm based on decomposition (MOEA/D) is used to optimize the parameters of LSTM network, and multiple high-quality sub models are obtained. The combined output strategy based on deep belief network (DBN) is used to combine the outputs of multiple sub models to form combined prediction, so as to improve the generalization ability of the prediction model. The specific work contents are as follows:

(1) Summarize the characteristics of residential load. Based on the residential user load data provided by the smart grid smart city (SGSC) project in Australia, this paper compares the residential load with the system load, highlights the characteristics of the residential load, and explores the key optimization direction of the prediction model.

(2) Build the data preprocessing framework based on multi feature input.

Because there are many factors that affect the change of residential load, in order to make the prediction model fully mine the potential correlation in the data, this paper uses the multi feature input composed of power consumption component, time component, week component and holiday component. We hope to increase the hidden knowledge that the prediction model can learn by adding the dimension of input data, so as to improve the prediction accuracy of the model. Based on the data provided by SGSC, the experiment is carried out by combining the multi feature input framework with LSTM network. Compared with the traditional prediction model and deep learning model, it is proved that the multi feature input framework can improve the overall accuracy of the prediction model.

(3) The network optimization strategy based on MOEA/D is proposed.

In order to improve the overall performance of the network, MOEA/D algorithm is used to optimize the parameters of the network. Taking the accuracy and difference of LSTM network as the target space and the super parameters of LSTM network as the decision space, the network is optimized. In this adaptive way, the network super parameters can automatically adapt to the load data, avoiding the error caused by artificially setting parameters or setting fixed parameters. At the same time, in order to make the MOEA/D algorithm more consistent with the deep learning framework, this paper improves the MOEA/D algorithm, and illustrates the effectiveness of the improvement through comparative experiments.

(4) The combined strategy based on DBN is proposed.

The DBN network is used to combine the prediction results of multiple sub models into the final prediction results. Through this method similar to ensemble learning, we can overcome the unstable factors in the prediction of a single network, so as to alleviate the effect of data over fitting and improve the generalization ability of the model. This paper compares the multi-objective optimization model based on long term and short-term memory network with the existing multiple intelligent prediction models to illustrates the advanced nature of the model.

Key Word: User side load forecasting; Multi feature input data; Multi-objective optimization algorithm; Short and long term memory network; Combination model

全文缩写与中文含义对照表

缩略语	英文全称	中文全称
LSTM	Long-Short Term Memory	长短期记忆网络
MOEA/D	Multi-objective Evolutionary Algorithm Based on Decomposition	基于分解的多目标优化算法
DBN	Deep Belief Network	深度信念网络
SGSC	Smart Grid Smart City	智能电网智能城市
AMI	Advanced Metering Infrastructure	高级计量基础设施
ToU	Time-of-daytariff	分时电价
EMS	Energy Management System	能量管理系统
RE	Renewable Energy	可再生能源
EV	Electric Vehicle	电动汽车
ARIMA	Autoregressive Integrated Moving Average model	差分整合移动平均自回归模型
ANN	Artificial Neural Network	人工神经网络
ELM	Extreme Learning Machine	极限学习机
EMD	Empirical Mode Decomposition	经验模态分解
PSO	Particle Swarm Optimization	粒子群优化
RBM	Restricted Boltzmann Machine	受限玻尔兹曼机
BPNN	Back Propagation Neural Network	反向传播神经网络
RNN	Recurrent Neural Network	循环神经网络
MOP	Multiobjective Optimization Problem	多目标优化问题
NSGA-II	Non-dominated Sorting Genetic Algorithm-II	非支配排序遗传算法-II
NSGA	Non-dominated Sorting Genetic Algorithm	非支配排序遗传算法
GA	Genetic Algorithms	遗传算法
PBI	Penalty-based Boundary Intersection	基于惩罚的边界交叉法
SVM	Support Vector Machine	支持向量机
RMSE	Root Mean Squared Error	均方根误差
MAPE	Mean Absolute Percentage Error	平均绝对百分比误差

全文符号与中文含义对照表

符号	中文含义
v	可视层向量
h	隐藏层向量
W	神经网络权重向量
L	损失函数
E_{θ}	能量函数
P_{θ}	DBN 单元激活概率
S_t	LSTM 隐藏层单元向量
f_t	LSTM 遗忘门向量
i_t	LSTM 输入门向量
g_t	LSTM 输入节点向量
o_t	LSTM 输出门向量
h_t	LSTM 中间输出向量
F_x	优化算法目标空间
Ω	优化算法决策空间
Y_j	拥挤度函数
λ	优化算法权重向量
ω_i	子模型权重向量
v_i	归一化因子
P	MOEA/D 种群大小
T	MOEA/D 领域大小
$iter$	MOEA/D 最大迭代次数
Z^*	MOEA/D 参考点

全文附表索引

表格	中文含义
表 1-1	负荷预测方法分类及其主要代表方法
表 3-1	缺失值处理方法及其特点
表 3-2	点 x^* 的相关信息
表 3-3	多特征输入框架有效性对比实验
表 3-4	重复 5 次实验运行中最大误差对比
表 3-5	多种方法的 MAPE 误差平均值
表 4-1	超参数择优范围
表 4-2	未经归一化目标函数值展示
表 4-3	有无第二阶段优化的模型精度对比
表 4-4	不同层数的 DBN 模型的误差
表 4-5	整体误差对比
表 4-6	个别误差对比

全文插图索引

图	中文含义
图 1-1	我国 2011 年至 2020 年全社会用电量
图 1-2	智能电网调度体系框架
图 1-3	负荷预测分类及其主要作用
图 1-4	论文框架
图 2-1	RBM 结构示意图
图 2-2	DBN 结构示意图
图 2-3	RNN 结构示意图
图 2-4	LSTM 结构示意图
图 2-5	NSGA 流程图
图 2-6	NSGA-II流程图
图 2-7	MOEA/D 流程图
图 2-8	切比雪夫分解方法示意图
图 2-9	基于惩罚的边界交叉法示意图
图 3-1	SGSC 数据集文件图示
图 3-2	系统级负荷曲线与住宅级负荷曲线对比图
图 3-3	x^* 多特征处理后示意图
图 3-4	单层 MISO 结构示意图
图 3-5	LSTM 网络框架示意图
图 4-1	基于数分解的组合预测方法流程图
图 4-2	基于参数优化的组合预测方法流程图
图 4-3	基于长短期记忆网络的多目标优化组合预测方法流程图
图 4-4	全更新方法、基于最小适应值更新方法和基于最大适应差值的更新方法对比图
图 4-5	目标函数收敛趋势图
图 4-6	目标函数归一化对最终种群分布影响对比图
图 4-7	经过第二阶段优化后的种群分布
图 4-8	12 月负荷曲线拟合图
图 4-9	12 月负荷误差曲线
图 4-10	1 月负荷曲线拟合图
图 4-11	1 月负荷误差曲线
图 4-12	2 月负荷曲线拟合图
图 4-13	2 月负荷误差曲线
图 4-14	6 月负荷曲线拟合图
图 4-15	6 月负荷误差曲线

图 4-16	7 月负荷曲线拟合图
图 4-17	7 月负荷误差曲线
图 4-18	8 月负荷曲线拟合图
图 4-19	8 月负荷误差曲线

目录

摘要.....	I
Abstract.....	III
全文缩写与中文含义对照表.....	V
全文符号与中文含义对照表.....	VI
全文附表索引.....	VII
全文插图索引.....	VIII
第 1 章 绪论.....	1
1.1 研究背景及意义	1
1.2 负荷预测的研究现状	3
1.3 研究内容及论文章节安排	5
1.3.1 研究内容.....	5
1.3.2 论文章节安排.....	6
第 2 章 理论基础与相关工作.....	8
2.1 引言	8
2.2 深度学习方法	8
2.2.1 深度信念网络.....	8
2.2.2 长短期记忆网络.....	10
2.3 多目标优化算法	12
2.3.1 多目标优化问题.....	12
2.3.2 非支配排序遗传算法II.....	13
2.3.3 基于分解的多目标优化算法.....	15
2.4 本章小结	17
第 3 章 基于多特征输入的 LSTM 模型分析.....	19
3.1 引言	19
3.2 住宅级负荷数据特征分析	19
3.2.1 数据集介绍.....	19
3.2.2 数据缺失处理.....	20
3.2.3 探索性数据分析.....	21

3.3 基于多特征输入的 LSTM 模型介绍.....	23
3.3.1 负荷数据的多特征输入框架.....	23
3.3.2 基于 MISO 结构的 LSTM 网络.....	24
3.4 基于多特征输入的 LSTM 模型实验与分析.....	25
3.4.1 实验说明.....	25
3.4.2 预测性能评价指标.....	25
3.4.3 实验数据集的选择和参数设置.....	26
3.4.4 结果对比与分析.....	26
3.5 本章小结	28
第 4 章 基于长短期记忆网络的多目标优化组合预测模型.....	30
4.1 引言	30
4.2 组合预测方法	30
4.2.1 基于数据分解的组合预测方法.....	30
4.2.2 基于参数优化的组合预测方法.....	31
4.2.3 基于权重分配的组合预测方法.....	33
4.3 方法导论	33
4.4 基于 MOEA/D 的子模型生成策略	35
4.4.1 决策空间与目标空间的构成.....	35
4.4.2 基于最大适应度差值的更新方法.....	35
4.4.3 基于非支配排序的第二阶段优化策略.....	38
4.5 基于 DBN 的结合策略	39
4.6 对比实验与分析	40
4.6.1 参数设置.....	40
4.6.2 案例研究.....	40
4.7 基于 LSTM 组合预测模型的短期电力负荷预测实验与分析.....	44
4.7.1 整体误差对比.....	44
4.7.2 个别误差对比.....	44
4.8 本章小结	49
第 5 章 总结与展望.....	50
5.1 工作总结	50
5.2 工作展望	51
参考文献.....	52
致 谢.....	56

个人简介、攻读硕士学位期间的学术成果及科研项目.....57

第 1 章 绪论

1.1 研究背景及意义

19 世纪 60 年代发生的第二次工业革命将电能推向了时代潮流的第一线。作为一种清洁、高效的二次能源，电能具有运输安全性高、运输损耗少、能量形式转换灵活等优点^[1]。随着人类社会生产力的提升，对电能的需求也在逐步增加。

《中国能源大数据报告（2021）》对 2011 年至 2020 年我国全社会用电量进行了统计，统计结果如图 1-1 所示。根据该报告，我国“十二五”至“十三五”时期全社会用电量逐步增加，到 2020 年已经达到了 75110 亿千瓦时之巨。同时，该报告还指出，在 2021 年我国终端能源消费占比中，电能消费占比将达到 28% 左右。联系我国制定的“碳达峰”、“碳中和”的“双碳”目标，《中国 2060 年前碳中和研究报告》要求，在 2060 年的终端能源消费占比中，电能消费占比要求达到 70%。由此可见，电能已然成为了维持国民生产生活、促进人类发展进步的重要能源。

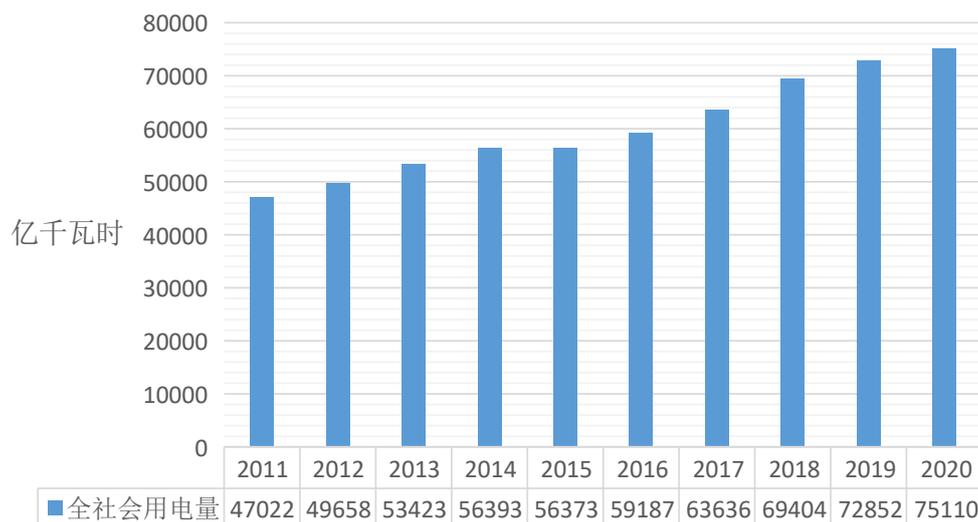


图 1-1 我国 2011 年至 2020 年全社会用电量

电力网在电力系统中，起到控制、调配以及输送电能的作用^[2]。传统电网的主要任务只停留在维持供电不间断，保证供电安全的层面。其运行模式多为被动，即只要系统没有发生故障就不会主动进行系统调整等操作。随着电力系统规模的增大，以及能源枯竭压力和环保压力的增加，传统电网的弊病也被暴露出来。第一是传统电网的脆弱性。传统电网无法对不正常的工作状态进行及时的自我调整，当故障发生时，其自愈功能不佳，几乎完全依靠人工操作，反应速度极慢。加之

传统电网的整体性强，区域性的故障时常会引发大规模的系统崩溃。第二是传统电网的局限性。传统电网内部信息共享能力低下，无法建立起反馈性的系统。同时，传统电网以集中式发电为主，无法很好的处理分布式能源并网的问题^[3]。由于以上的种种原因，传统电网被逐步淘汰，智能电网登上历史舞台成为电力系统发展的主流方向^[4]。智能电网具备强大的信息获取能力，其以实体电网为基础，配合以信息平台，对系统的各种实时信息变化做出快速反应。通过对信息的处理，快速制定调整策略，对电网运行进行准确调整，优化了电网的运行和管理，在节能减排、能源利用率最大化的同时保护了电网设备，延长了系统的整体运行寿命。

在图 1-2 中给出了智能电网的调度体系框架。智能电网通过强大的数据采集技术，收集大量的系统运行信息。通过对这些信息的处理，智能电网可以实现数据可视化、负荷预测、设备故障监测和自我修复等功能。电网端可以利用这些信息制定相应的发电计划、调度计划，减少电能损耗，提高电能利用率。

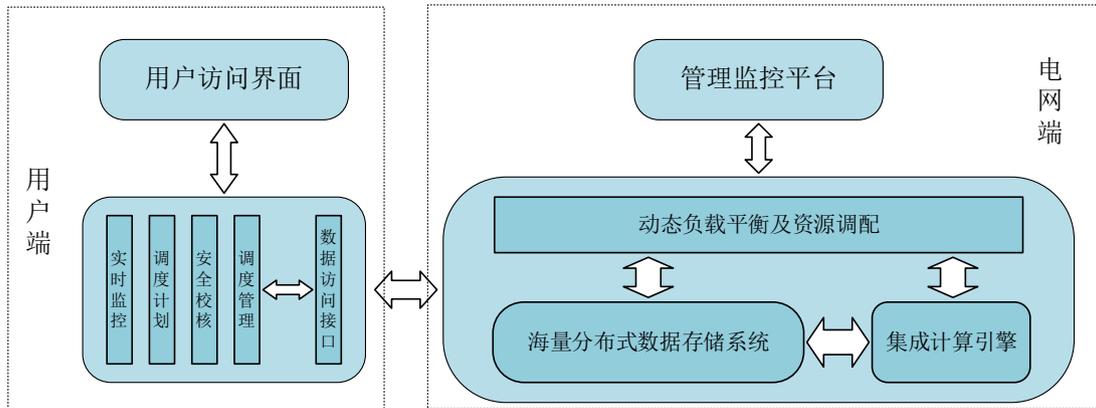


图 1-2 智能电网调度体系框架

负荷预测作为智能电网中重要的一环，为所有的计划制定提供数据支撑。可以根据所处理的负荷数据时间长短对负荷预测进行分类，其各种类负荷预测的作用及之间的关系如图 1-3 所示^[5]。本文所研究的对象是针对用户侧负荷的短期预测。以往的用户侧短期负荷预测主要是针对系统级的负荷进行预测。根据预测结果，可以对发电机组的启停进行控制，保持电能的产销平衡，减少不必要的发电设备的备用容量，为检修计划的制定提供信息支撑，在保障系统安全运行的同时，能大大减少运维成本，提高能源利用率^[6]。随着电网中高级计量基础设施^[7] (Advanced Metering Infrastructure, AMI) 的大量引入，海量的住宅级负荷数据被采集，使得对住宅级负荷数据的利用变为可能。住宅级负荷的短期预测在能量管理系统 (Energy Management System, EMS) 的建设中起到了决定性的作用^[8]。公用事业公司可以通过建立 EMS，配合能量存储设施和智能需求响应技术实现电力负荷“削峰填谷”，化解高峰电力短期问题，节约电源电网投资，达到高效、经济用电的目的。

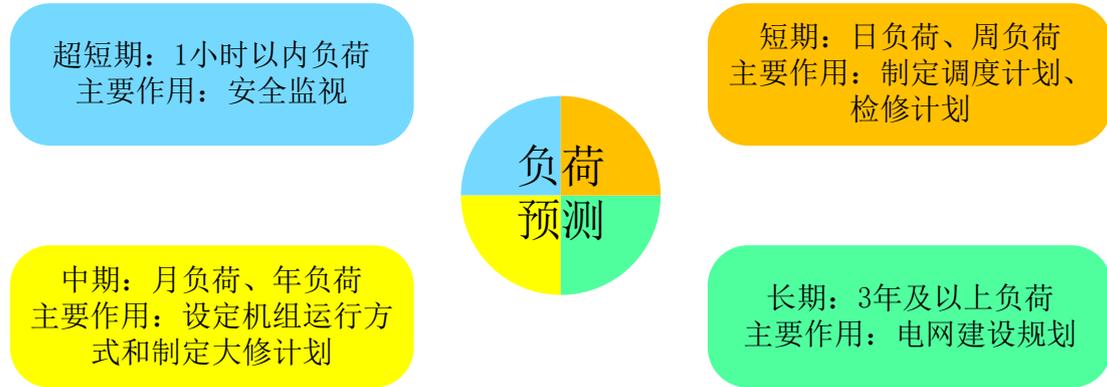


图 1-3 负荷预测分类及其主要作用

然而，相较于系统级负荷，住宅级负荷波动更加剧烈。住宅级负荷变化受到诸多因素的影响，其中包括人类活动和天气情况等，哪怕是细微的因素变化也会引起负荷的剧烈波动^[9]。并且，随着可再生能源(Renewable Energy, RE)、电动汽车(Battery Electric Vehicle, BEV)的加入，使得原本复杂的负荷数据变得更加难以预测。如何实现对住宅负荷的有效预测成为了当前的一个研究热点。

1.2 负荷预测的研究现状

20 世纪 70 年代，随着电网规模的不断扩大，负荷预测开始成为众多学者的研究焦点。因为工具的缺乏，最开始的负荷预测是由技术人员按照以往负荷曲线的变化趋势直接进行人工拟合，以来推测未来时间的负荷数据^[10]。这种完全依靠人类主观意愿而完成的负荷预测显而易见的会产生极大的误差，以至于在很长一段时间内，负荷预测根本无法用于工程实践。随着科技的进步，大量的高精尖技术以及功能强大的工具被运用到负荷预测上，特别是计算机计算能力的增长极大的促进了负荷预测方法的发展。目前，主要的负荷预测方法的分类及其代表方法如表 1-1 所示。

表 1-1 负荷预测方法分类及其主要代表方法

类别	代表方法
传统预测方法	时间序列方法、回归分析方法以及灰色预测方法等
现代智能预测方法	人工神经网络、模糊逻辑方法、深度学习以及组合预测方法等

电力负荷数据是一种明显的时间序列，时间序列方法就是对数据在很长一段时间内的变化趋势进行总结，并将其变化趋势外推到未来，使用过去的变化趋势来预测未来。文献^[11]采用最小二乘支持向量回归机对预处理过的电力负荷时间序

列进行建模,对电力负荷进行短期预测。在文献^[12]中作者将负荷和气温数据结合构建了基于差分整合移动平均自回归模型(Autoregressive Integrated Moving Average model, ARIMA)回归残差预测模型。时间序列方法的主要问题是其处理的对象只能是比较简单的数据。如果数据的复杂度增加,该方法的整体精度将会大幅下降。随着电力系统的不断细分,时间序列方法已经不再适用于如今复杂度很高的负荷数据。

回归分析方法是通过寻找主导负荷数据变化趋势的潜在影响因素,利用影响因素和数据变化之间的相关性对数据变化规律进行建模,通过搭建起来的模型对负荷进行预测。文献^[13]使用偏最小二乘回归分析方法对中长期电力负荷进行预测,具有计算速度快、精度高和实用性强的优点。文献^[14]首先使用随机森林对负荷输入数据进行预处理,再将处理过后的数据进行在线序列支持向量回归,构建回归模型。回归分析方法建模快,也就意味着模型结构简单,随着电力负荷数据的非线性加深,该方法也出现了预测精度逐渐降低的缺点。

对于一个系统而言,如果其内部特征存有部分已知、剩下部分未知的情况,就可以将该系统定义为灰色系统^[15]。灰色预测方法就是构建系统特征之间的相关性,通过生成微分方程对数据之间的规律进行描述。在文献^[16]中,作者首先使用灰色预测方法构建起多个预测模型,再使用神经网络对多个模型的预测结果进行整合,以提高最终预测结果的精度。

随着计算机的发展,一些需要大量计算量的数据处理方法得以实现。人工神经网络(Artificial Neural Network, ANN)的复兴为负荷预测提供了许多新的选择^{[17][18][19]}。文献^[20]使用极限学习机(Extreme Learning Machine, ELM)对电力系统短期负荷进行预测,该方法简单但对复杂数据进行预测时精度下降。文献^[21]将一个深度信念网络模型加入到一个前馈神经网络中,并使用该网络对马其顿地区的电力消耗进行了预测。文献^[22]使用 LSTM 网络对住宅级负荷进行预测。结果表明,相较于其他网络, LSTM 对住宅级负荷有着更好的拟合效果。

为了使负荷预测更加精确,学者们开始把深度学习网络与优化技术相结合,提出了组合模型的思想。文献^[23]提出了一种新的基于跟随领导概念的进化算法,再利用算法对网络参数进行适当调整,使网络性能接近最优化。文献^[24]通过经验模态分解(Empirical Mode Decomposition, EMD)对历史负荷数据进行处理,突出数据变化特点增加了模型的精确度。Hu 等人使用粒子群优化算法(Particle Swarm optimization, PSO)来优化深度学习网络的超参数^[25],得到了具有较高精度和稳定性的预测模型。但是优化算法只针对模型的预测精度进行了优化,使得模型对负荷数据产生了过拟合效应,降低了组合模型的泛化能力。以上所提到的方法,已经在系统级负荷预测上取得了成功,但当这些方法运用到预测难度更高的住宅级

负荷预测时，效果往往差强人意。在对住宅级负荷进行预测时，要求模型能对数据规律的突变做出及时有效的反应，能学习到所有与数据变换趋势相关联的隐藏信息，做到面面俱到。显然，单个模型已经不能到达这样的要求。为了弥补单个模型的不足，学者们开始把集成学习的思想引入到负荷预测当中。集成学习的核心是使用多个子模型对负荷数据进行拟合，再通过各种手段，把所有子模型结合起来。由于最终的预测结果结合了多个子模型对负荷数据的处理结果，使得预测精度相对单个模型进行预测要高得多。文献^[26]对多个不同的人工智能算法和机械学习算法进行集成学习，参考多个子模型的预测值来形成最终预测。但由于不同种类的人工智能算法和机械学习算法数量有限，限制了子模型的数量，降低了组合模型的最终性能。文献^[27]对多个相同的人工智能算法进行集成，相较于对不同种类的算法进行集成，该方法更加简单并不受子模型数量的限制。但该方法无法保证每个子模型的准确性和多样性，无法为组合模型提供最优质的子模型。文献^[28]在 ELM 的基础上引入了基于小波的结合策略，并利用偏最小二乘回归方法将个别子模型组合成集成预测。该方法相对于使用单个 ELM 模型进行预测，显著提高了模型的预测性能，但由于其结合策略过于依赖个别子模，而忽略了其他子模型的表现，无法保证最终组合模型的完整。

综上所述，随着智能电网功能的开发，负荷预测的对象不断细分，由最初的系统级负荷向住宅级负荷发展。预测对象的细分增加了负荷数据的波动性，使得负荷数据日益复杂。为了应对这一问题，负荷预测方法也不断进化，预测模型发生了由简单到复杂，由浅层到深层的演变。随着人类生活的电气化程度加深，负荷数据会变得更加复杂，如何提高预测模型的精度依然是当下学者们的研究重点。

1.3 研究内容及论文章节安排

1.3.1 研究内容

经过 1.2 节对现有方法的总结，可以发现，现有的负荷预测方法存在以下问题：以往简单的深度学习模型已经无法很好拟合现在复杂的住宅负荷数据；单目标优化算法在优化深度学习模型时，无法兼顾模型的多样性和准确性，降低了组合模型的泛化能力；现有的结合策略无法充分利用到所有子模的优点。

针对以上问题，为了得到具有强泛化能力和高精度的预测模型，实现对住宅级负荷的有效预测，本文提出了一种基于长短期记忆网络的多目标优化组合预测模型。该方法以 LSTM 为基础构建深度学习模型。用负荷数据、时间指数、周指数和假日指数组成多维输入提高了模型的精度；在组合学习方面，使用改进了的 MOEA/D 算法来生产和筛选子模型，并使用基于非支配排序的第二阶段优化对

得到的多个优质子模型进行第二阶段筛选，保证了子模型的准确性和多样性；选用 DBN 作为最终的结合策略，通过 DBN 对多个子模型进行集成。本文的主要研究内容和创新如下所示：

(1) 以澳大利亚智能电网智能城市项目提供的住宅用户负荷数据为基础，对住宅级负荷进行了分析。以此为基准，提出了基于多特征输入的 LSTM 模型，为组合模型提供子模型基础。

(2) 为了实现模型网络参数的自适应，生成多个优质的子模型，本文提出了基于 MOEA/D 的子模型生成策略。使用 MOEA/D 算法对网络的超参数进行调整，并针对实际应用中出现的问题对 MOEA/D 算法进行了改进，使其更加契合实际操作。

(3) 为了能在(2)的基础上，对子模型进行进一步筛选，本文提出了基于非支配排序的第二阶段优化策略。

(4) 提出了基于 DBN 的结合策略。使用 DBN 强大的非线性拟合能力，构建起多个子模型与最终预测值之间的特征联系。充分利用每一个子模型的学习结果，提升了组合模型的精度和泛化能力。

1.3.2 论文章节安排

本文基于国家自然科学基金“基于自动需求响应的多能互补智能楼宇全直流微网动态优化”(61572416)项目需求，针对智能楼宇这一应用场景，对住宅级负荷短期预测展开研究。

本文的整体结构图如图 1-4 所示。

在第一章中，首先对电能、智能电网和负荷预测的背景进行了阐述，说明了对住宅级负荷进行准确预测的意义。然后对负荷预测方法的发展和现状进行了说明，逐一说明了现有方法的不足之处。最后针对这些不足，提出基于长短期记忆网络的多目标优化组合预测模型，并说明全文结构。

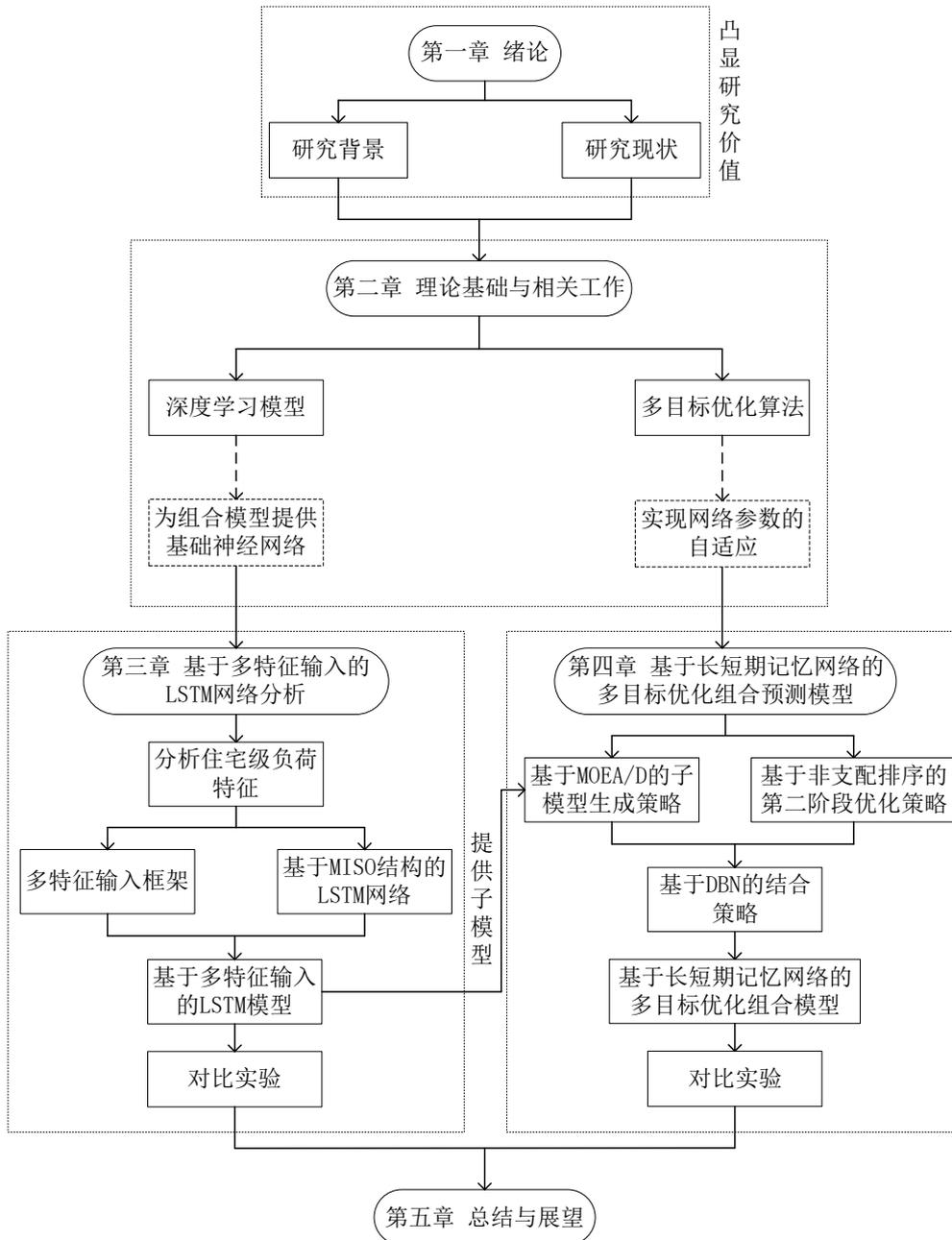
在第二章中，对本文所提出方法的全部基础性理论进行了说明，包括：深度学习模型，多目标优化算法。说明它们在本文提出的方法中各负责什么工作，起到什么作用，为后续的论证打下基础。

在第三章中，首先基于澳大利亚智能电网智能城市项目提供的住宅用户负荷数据对住宅负荷的特点进行探究。然后针对这些特点，提出多特征输入框架。将多特征输入框架与基于 MISO 结构的 LSTM 网络相结合，构成基于多特征输入的 LSTM 模型。最后，将引入现有的负荷预测方法，在多组数据集上进行对比实验，凸显所提出模型的优越性。

在第四章中，提出基于 MOEA/D 的子模型生成策略、基于非支配排序的第

二阶段优化策略和基于 DBN 的结合策略。延用第三章提出的模型，共同组成基于长短期记忆网络的多目标优化组合预测模型。对比实验分为两个部分。第一部分针对提出策略中的各种改进方法进行对比实验，说明改进方法的有效性。第二部分引入现有的多种预测模型以及组合预测方法，在多组数据集上进行对比实验，给出了多个图表结果，对所提出组合模型的优越性进行说明。

在第五章中，将给出本文的全部工作内容总结，总结所提出方法的优点和需要进一步改进的地方。同时也对本文提出的负荷预测组合模型未来的发展方向进行探讨。



第 2 章 理论基础与相关工作

2.1 引言

本文提出了基于长短期记忆网络的多目标优化组合预测模型,以提高对住宅级负荷进行短期预测的精度。该模型首先使用 LSTM 网络构建基础预测模型;然后使用改进了的 MOEA/D 算法对子网络的超参数进行优化,得到多个优质的子模型。在以上优化的基础上,继续使用基于非支配排序的第二阶段优化策略对子模型进行进一步择优。最后,使用基于 DBN 的结合策略对多个子模型的预测值进行组合输出,得到最终的预测值。由于该模型所涉及到的相关理论较多,为了方便之后的论述和实验比较,在本章对本文中所使用到的相关方法进行介绍,这些方法包括:在神经网络方面包括深度信念网络和长短期记忆网络;在多目标优化算法方面包括非支配排序遗传算法-II、基于分解的多目标进化算法;在组合预测方面包括基于网络参数优化、基于输入数据分解和基于相关联权重分配的组合预测方法。

2.2 深度学习方法

2.2.1 深度信念网络

G.E.Hinton 于 2006 年提出了名为深度信念网络的深度学习模型,并在神经计算(Neural Computation)期刊和科学(Science)上发表了相关论文^{[29][30]}。当时的学术界普遍认为深度神经网络模型太难优化,在处理实际问题时还不如直接使用浅层的机器学习模型,以至于深度神经网络的发展在很长一段时间内停滞不前。直到深度信念网络的提出,才再度掀起了深度神经网络的研究热潮,其历史地位不言而喻。

DBN 的基本神经元结构是受限玻尔兹曼机^[31] (Restricted Boltzmann Machine, RBM), 一个经典的 DBN 是由多个 RBM 共同组成的。RBM 的结构可以定义为如图 2-1 所示。图中, $v = (v_1, \dots, v_n)$ 表示可视层向量; $h = (h_1, \dots, h_m)$ 表示隐藏层向量; $a = (a_1, \dots, a_n)$ 表示可视层的偏置向量; $b = (b_1, \dots, b_m)$ 表示隐藏层的偏置; $W = (w_{ij})$ 表示可视层与隐藏层之间的权重向量。从图 2-1 中可知, 一个 RBM 包括可视层和隐藏层两个层结构。其中可视层包含了所有的输入数据, 一个可视层神经元用来表述输入数据的一个特征, 而隐藏层的意义无法明确, 一般将其定义为特征提取结果。可视层和隐藏层之间要求为全连接, 层内部要求没有连接, 这样的二分图结构使得当 RBM 给定了任意一个层神经元状态时, 另一层上的每一神

经元激活条件相互独立。

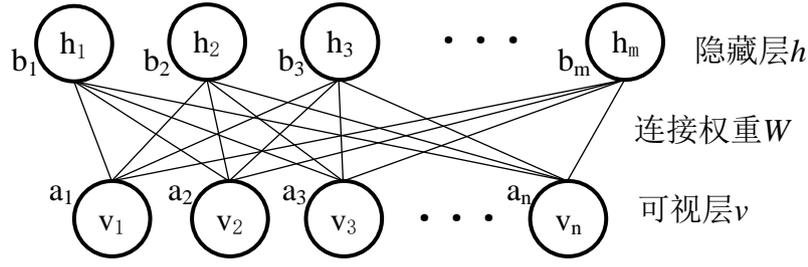


图 2-1 RBM 结构示意图

对于一个已经明确了各个神经元状态 (v, h) 的RBN来说,它的能量函数可以表示为式(2-1)。式中 $\theta = (a, b, W)$ 表示参数向量; n 和 m 表示可视层向量和隐藏层向量的长度; v_i 表示第 i 个可视层神经元对应的特征值; h_j 表示第 j 个隐藏层神经元对应的特征值; w_{ij} 表示 v_i 和 h_j 之间的连接权重值; a_i 表示 v_i 的偏置; b_j 表示 h_j 的偏置。

$$E_{\theta}(v, h) = -\sum_{i=1}^n a_i v_i - \sum_{j=1}^m b_j h_j - \sum_{i=1}^n \sum_{j=1}^m v_i w_{ij} h_j \quad (2-1)$$

从能量函数可以得到可视层与隐藏层之间的联合概率分布数学模型,如式(2-2)所示。因此,可以得出隐藏层神经元 h_j 被可视层 v 激活的概率为式(2-3)。由RBM的双向特性可知,可视层 v_i 也可以由隐藏层 h 激活,其概率可以定义为式(2-4)。

$$P_{\theta}(v, h) = \frac{e^{-E_{\theta}(v, h)}}{\sum_{v, h} e^{-E_{\theta}(v, h)}} \quad (2-2)$$

$$P_{\theta}(h_j = 1 | v) = \sigma \left(b_j + \sum_{i=1}^n w_{ij} v_i \right) \quad (2-3)$$

$$P_{\theta}(v_i = 1 | h) = \sigma \left(a_i + \sum_{j=1}^m w_{ij} h_j \right) \quad (2-4)$$

RBM的训练方法根据所处理任务的不同可分为监督学习^[32]以及无监督学习^[33]。在这里,只讨论在本文提出方法中会使用到的学习方式,无监督学习,其主要步骤如下所示:

- (1) 初始化参数向量 $\theta = (a, b, W)$;
- (2) 输入训练数据,记此时的可视层向量为 v ;
- (3) 由式(2-3)计算出所有隐藏层神经元的状态,记此时的隐藏层向量为 h_1 ;
- (4) 由式(2-4),通过 h_1 重构可视层 v_1 ;
- (5) 由损失函数 $L(v, v_1)$ 求得误差;

- (6) 由误差对参数向量 θ 进行微调；
 (7) 以 v_1 为可视层，重复(3)到(6)直至 RBM 收敛。

需要注意的是，在迭代的过程中，所有用来进行误差计算的基准皆为初始输入数据时所构建的可视层向量 v 。以上便是 RBM 的原理和训练方法，对多个 RBM 进行堆砌编组成了一个 DBN，DBN 结构如图 2-2 所示。

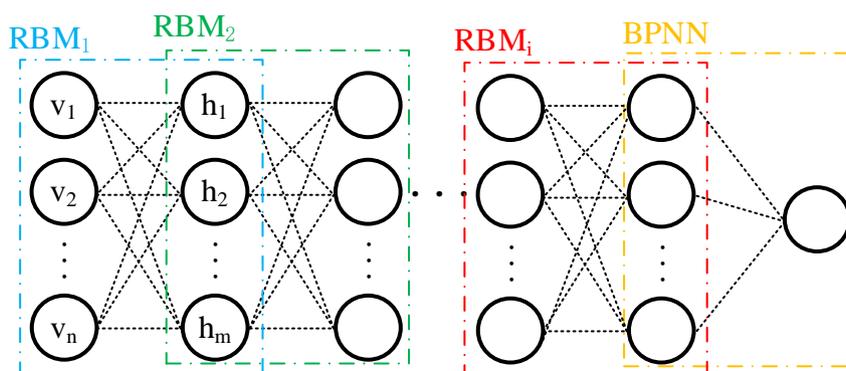


图 2-2 DBN 结构示意图

将 RBM_1 的隐藏层作为 RBM_2 的可视层便可实现两层 RBM 的堆叠。依照这种方式对 RBM 进行前后连接，一个 DBN 可以堆砌任意层数的 RBM。第一个 RBM 的可视层作为输入层，用来接收训练数据，往后的每一层隐藏层则负责去捕捉输入数据里的各种相关特性。在结构的最后，由一个 BP 神经网络(Back Propagation Neural Network, BPNN)来完成结尾，输出得到相应的结果。假设一个 DBN 包括了 i 个 RBM，则可以定义其联合概率为式(2-5)。

$$P_{\theta}(v, h_1, h_2, \dots, h_i) = P_{\theta}(v | h_1) P_{\theta}(h_1 | h_2) \cdots P_{\theta}(h_{i-2} | h_{i-1}) P_{\theta}(h_{i-1} | h_i) \quad (2-5)$$

本文希望借助 DBN 强大的非线性处理能力，构建起子模型预测值与最终预测值之间的联系。综合多个子模型的数据处理结果以来提高模型的泛化能力。

2.2.2 长短期记忆网络

长短期记忆网络是循环神经网络(Recurrent Neural Network, RNN)的一种衍生网络^[34]，在介绍长短期记忆网络之前，先对 RNN 进行简单介绍。循环神经网络^[35]早在二十世纪 80 年代就被提出，并在二十一世纪初期衍变为深度学习算法，其原始结构如图 2-3 所示。

图 2-3 的左边为一个基本的循环神经网络， x 表示输入向量； s 表示隐藏层神经元状态； y 表示输出向量； W_x 、 W_s 和 W_y 分别表示输入层、隐藏层和输出层的连接权重。图的右边为其平面展开图，下标 t 代表时间步。考虑到并不是每一

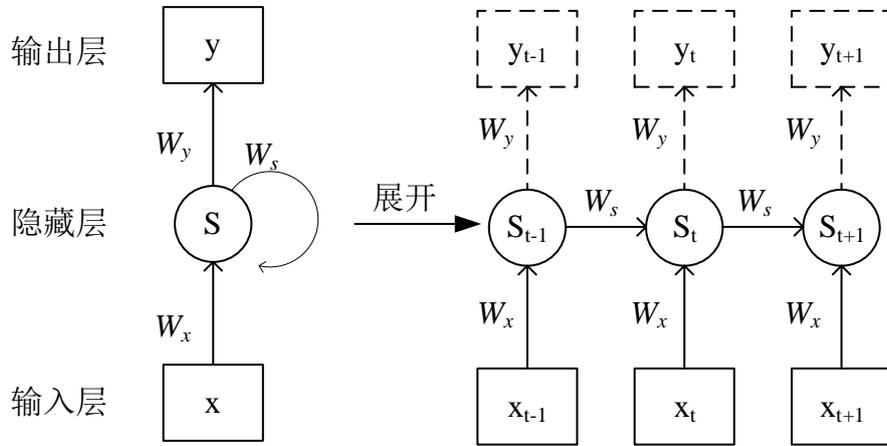


图 2-3 RNN 结构示意图

时间步都必须有输出层，而是根据不同的情况来确定需不需要设置输出层，所以在右边的结构中所有的输出层以虚线来表示。从图 2-3 中可以看出，在 RNN 中有着一条贯通始终的信息链。第 t 时间步的隐藏层神经元状态 S_t 是由当前输入 x_t 和前一时刻的状态 S_{t-1} 一同通过处理得到。同理， S_t 也会被用于计算第 $t+1$ 时间步的状态 S_{t+1} 。在训练时，通过不断的循环，对连接权重 W_x 、 W_s 和 W_y 进行调整，使得前面的信息可以影响到后面的信息，建立起信息间的时间关系。这样的特性使得 RNN 具备记忆功能，相对于传统的神经网络在学习序列型数据的非线性特征时有着明显的优势。其内部的数学模型归结如下：

$$S_t = f(W_x \cdot x_t + W_s \cdot S_{t-1} + b_s) \quad (2-6)$$

$$y_t = g(W_y \cdot S_t + b_y) \quad (2-7)$$

其中， f 表示隐藏层激活函数； g 代表输出层激活函数； b_s 和 b_y 表示隐藏层和输出层的偏置。

随着研究的深入，RNN 暴露出一个非常严重的问题——在对长序列数据研究时出现的梯度爆炸以及梯度消失问题^[36]。从式(2-6)可以看出，当对 RNN 进行反向传播时，梯度数值每经过一个时间步就需要对隐藏层激活函数进行一次处理。当序列达到一定长度时，梯度数值就会发散或者无限趋近于零，其发展方向取决于隐藏层激活函数的性质。为了克服这一问题，学者们提出了许多基于 RNN 的衍生结构，其中当数长短期记忆网络最为著名。LSTM 的最早出现可以追溯到 1997 年^[37]，经过不断的补全已然成为当前处理时序信号的主流方法。单个 LSTM 隐藏层神经单元如图 2-4 所示。

区别于传统的 RNN，LSTM 在每一个时间步的信息处理中，不再统一的对记忆向量 S_t 进行激活函数操作。这样的历史信息传递方式就从根本上缓解了网络在处理长期记忆时出现的梯度消失和梯度爆炸问题。同时，LSTM 单元还增设遗

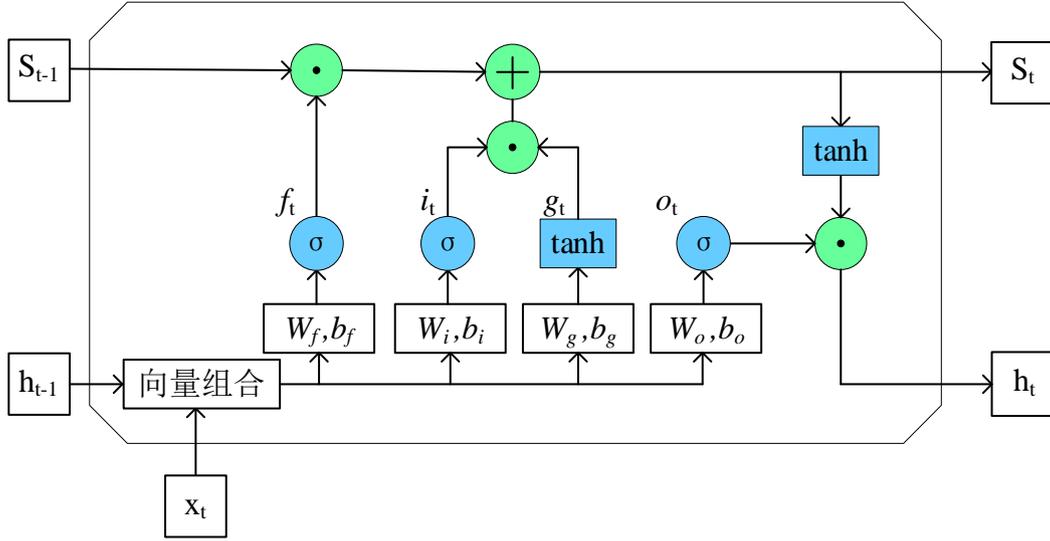


图 2-4 LSTM 结构示意图

忘门 f_t 、输入门 i_t 、输入节点 g_t 和输出门 o_t 来对输入 x_t 和历史信息进行处理，以对记忆向量 S_t 进行更新和形成中间输出 h_t [38]。单元内部的数学模型如下所示，其中， W 和 b 表示权重向量和偏置向量。

$$f_t = \sigma(W_f [h_{t-1}, X_t] + b_f) \quad (2-8)$$

$$i_t = \sigma(W_i [h_{t-1}, X_t] + b_i) \quad (2-9)$$

$$g_t = \tanh(W_g [h_{t-1}, X_t] + b_g) \quad (2-10)$$

$$o_t = \sigma(W_o [h_{t-1}, X_t] + b_o) \quad (2-11)$$

$$S_t = S_{t-1} \cdot f_t + i_t \cdot g_t \quad (2-12)$$

$$h_t = \tanh(S_t) \cdot o_t \quad (2-13)$$

本文选用 LSTM 为组合模型的基础网络，希望依托其强大的时间序列处理能力，提高最终模型对住宅级负荷的预测精度。

2.3 多目标优化算法

2.3.1 多目标优化问题

我们在实际生活中所遇到的优化问题多数是由多个子问题组成，这些子问题相互冲突，而平衡这些子问题，使它们在某种层面上达到同步最优，便是对这个优化问题的求解。这种能被分解成多个子问题的优化问题被统称为多目标

优化问题^[39](Multiobjective Optimization Problem, MOP)。一个 MOP 的数学定义如式(2-14)所示。

$$\begin{aligned} & \text{maximize } F(x) = (f_1(x), \dots, f_m(x)) \\ & \text{subject to: } x \in \Omega \end{aligned} \quad (2-14)$$

式(2-14)中, $F(x)$ 表示目标空间, Ω 表示决策空间。目标空间包括了 m 个子问题 $f_i(x) (i=1, \dots, m)$, 这些子问题由决策变量 x 通过对应的目标函数映射得到。目前, 多目标优化算法的求解方法有两种。一种是通过对于子问题分配权重, 将多目标优化问题展开成多个单目标问题并逐一对其进行求解, 另一种是利用帕累托最优求出一系列互不支配的解。第一种方法要求预先对所有子问题分配权重, 并且权重的分配方式对结果有着很大的影响。而在实际问题的求解中, 权重往往是无法预知的, 大多都是随着人的偏好所设置, 严重影响了结果的客观性。所以目前普遍使用的是第二种方法。

帕累托最优^[40](Pareto Optimality)是指在保证整体不退化的情况下, 使得局部更优化的一种优化方式, 由 Vilfredo Pareto 于 1896 年提出。该方法规定, 对于两个决策空间的点 x_1 和 x_2 , 如果对于所有子问题有 $f_i(x_1) \leq f_i(x_2) (i=1, \dots, m)$ 则称 x_1 弱支配 x_2 , 如果至少存在一个 $j \in \{1, \dots, m\}$, 使得 $f_j(x_1) < f_j(x_2)$, 则称 x_1 支配 x_2 。对于决策空间中, 那些不被任何其他点所支配的点, 称为 Pareto 最优解, 这些解的集合就称为 Pareto 最优解集, 其所构成的曲面称为 Pareto 前沿。对于一个多目标优化问题, 只要设法求得其 Pareto 最优解集, 再从中按要求挑选出一个最满意的解即可。

2.3.2 非支配排序遗传算法 II

非支配排序遗传算法-II (Non-dominated Sorting Genetic Algorithm-II, NSGA-II)是在非支配排序遗传算法(Non-dominated Sorting Genetic Algorithm, NSGA)的基础以上衍变而来^[41]。1995年, Srinivas 以及 Deb 两位学者以遗传算法^[42] (Genetic Algorithms, GA)为基础提出了 NSGA, 其算法流程图如图 2-5 所示。

如图 2-5 所示, NSGA 对 GA 的主要改进是在进行选择、交叉、变异等操作之前, 根据个体之间的支配关系对种群进行了非支配排序, 并对同一级别个体进行了适应度共享, 其具体操作步骤如下所示:

- (1) 挑出当前种群中处于非支配地位的所有个体, 组成第一非支配等级;
- (2) 对剩下的个体重复步骤(1), 直到全部个体都被分配到一个非支配等级。
- (3) 对每个非支配等级重新指定一个虚拟适应度, 非支配等级越靠前, 虚拟适应度越大。

经过上述操作, 将适应度值规范化, 使得优秀的个体能有更多的机会参与繁

殖下一代种群，防止了超级个体对种群多样性的影响，克服了 GA 容易陷于局部

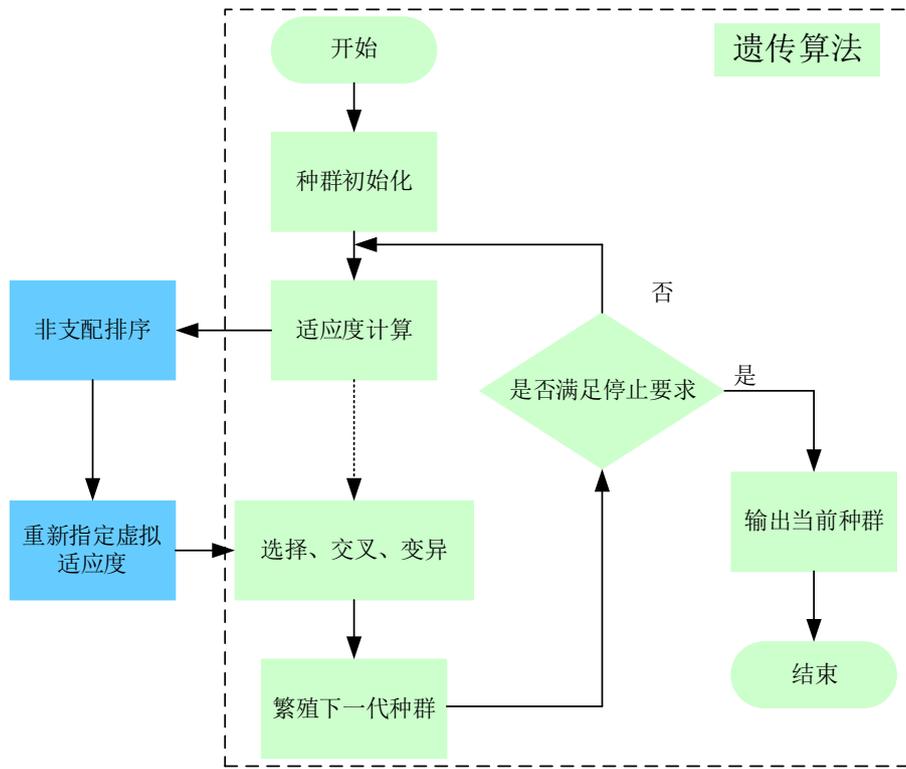


图 2-5 NSGA 流程图

最优的缺点。2000 年，Deb 又针对 NSGA 中的缺点提出了 NSGA-II，其算法流程图如图 2-6 所示。

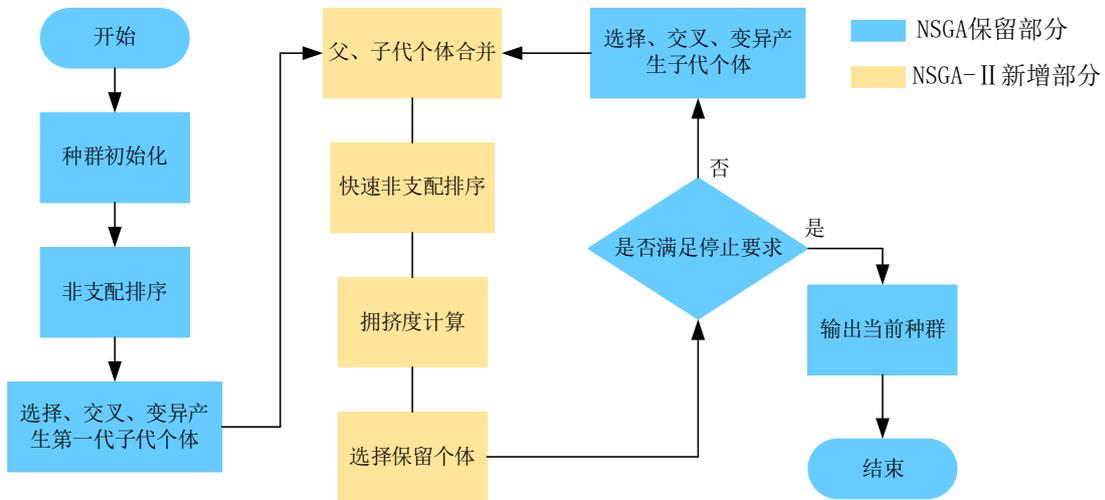


图 2-6 NSGA-II流程图

如图 2-6 所示，NSGA-II 为了弥补 NSGA 中的一些潜在问题又提出了精英策略、拥挤度算子和快速非支配排序法三个改进方法。其中，精英策略要求每一次迭代时父代和子代都要参与到竞争当中，这样就拓宽了采样空间，使得父代中优秀的个体有机会保留下来，加快了种群的进化速度。快速非支配排序法主要作

用是降低了算法的计算复杂度,在处理种群规模庞大的问题时大大缩短了算法运行时间,由于本文所提出的模型没有使用这一方法,遂不在这里对其进行详细说明。拥挤度算子用式(2-15)来计算所有个体的拥挤度。

$$Y_j = \sum_{i=1}^s \frac{f_i(j+1) - f_i(j-1)}{f_i(max) - f_i(min)} (j=1, \dots, P) \quad (2-15)$$

式(2-15)中,用 Y_j 来表示第 j 个个体的拥挤度, s 表示目标函数的维度, P 表示种群大小, $f_i(j+1)$ 为第 $j+1$ 个个体在目标函数 i 上的解, $f_i(max)$ 为所有个体在目标函数 i 上的解的最大值。其具体含义为以第 $j+1$ 和第 $j-1$ 个个体为顶点作包含第 j 个个体但不包含其他个体的长方形,求这个长方形的边长之和,用来反应区域中个体的密度。在 NSGA-II 中,使用拥挤度算子取代了 NSGA 中的适应度共享策略,使之不需要额外设置共享半径这一参数,并使用拥挤度作为“选择保留个体”操作时的一项指标。其规定,当“选择保留个体”时,优先考虑非支配等级考前的个体,当两个个体处于同一非支配等级时,则选择拥挤度较大的个体。通过比较拥挤度实现了保持种群多样性的效果。

本文在 MOEA/D 对子网络进行优化之后,使用包含了拥挤度的非支配排序对种群进行进一步筛选,剔除不良个体的同时进一步优化种群,为组合输出提供最优质的子网络。

2.3.3 基于分解的多目标优化算法

2007年, Qingfu Zhang 发表了名为《MOEA/D: A Multiobjective Evolutionary Algorithm Based on Decomposition》的论文^[43],正式提出了 MOEA/D 这一基于 Pareto 最优概念的多目标优化算法。这个算法的核心思想是用一组彼此分布均匀的权重向量对目标空间进行分割,将目标框架分解成多个子问题,依托权重向量对每个子问题同步优化,最终使种群逼近 Pareto 最优面。该算法的整体流程图如图 2-7 所示。

图 2-7 中,在初始化阶段,每个个体对应一个权重向量,并且每一个权重向量可以代表 1 个子问题。权重向量的维度就等于目标函数个数,具体的权重值代表在当前子问题中各个目标函数所占的比例大小。这样的处理将子问题个数固定为权重向量个数,避免了不必要的计算资源浪费。在种群更新阶段,每个父代个体只能和其领域内的个体进行繁殖,产生子代个体。进行比较更新时,也只有领域内的点才能进行竞争。这使算法在优化子问题时,将采样空间缩小到了该子问题的领域当中而不再是全局搜索。一方面减少了算法的计算复杂度,另一方面,在种群进化的过程中,个体紧紧围绕着其子问题所对应的权重向量,向着 Pareto 前沿逼近。也就是说只要选择好适当的分解方法,以及设置好子问题的个数(生

成一组合适的权重向量)就能保证种群的多样性。

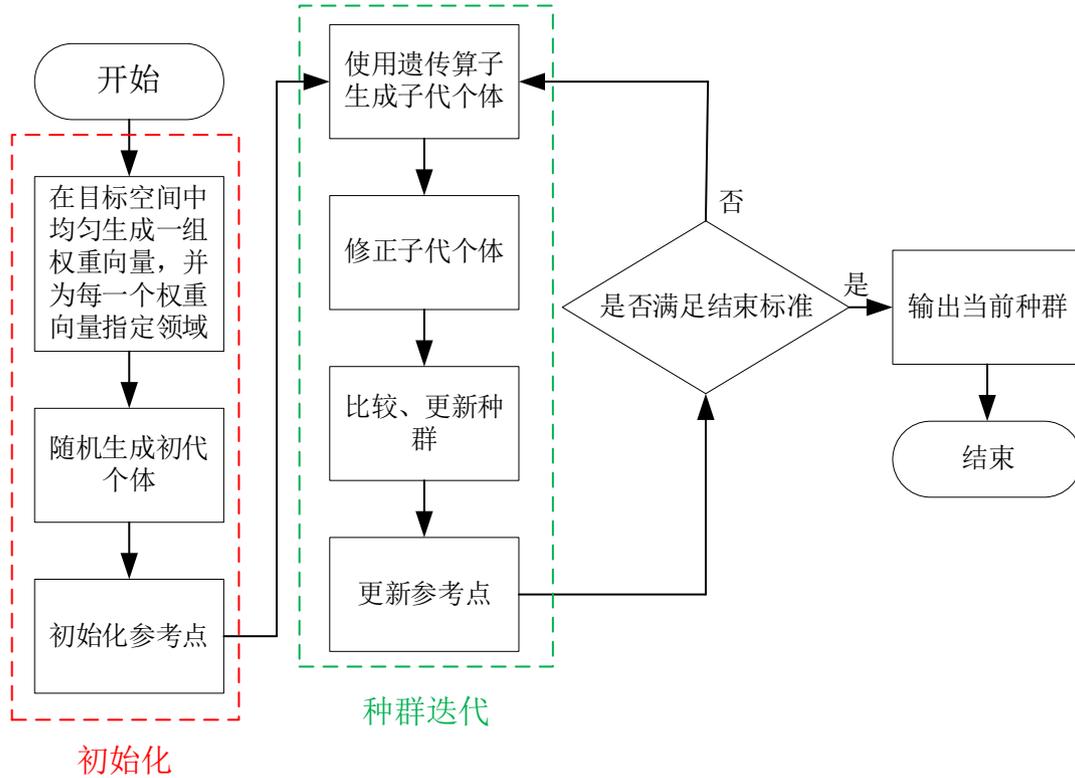


图 2-7 MOEA/D 流程图

目前主要的分解方法包括切比雪夫方法(Tchebycheff)、基于惩罚的边界交叉法^[44](Penalty-based Boundary Intersection, PBI)。切比雪夫方法的聚合函数定义为:

$$\min_{x \in \Omega} g(x | \lambda, Z^*) = \max_{1 \leq i \leq m} \lambda_i |f_i(x) - z_i^*| \quad (2-16)$$

式(2-16)中 Ω 代表决策空间, x 代表个体, m 代表目标函数的个数, $f_i(x)$ 代表个体 x 在第 i 目标函数上的表现, $\lambda = \lambda_1, \dots, \lambda_m$ 表示权重向量, $Z^* = z_1^*, \dots, z_m^*$ 表示参考点(一般选用当前最优个体作为参考点)。从式(2-16)中可以看出, 当 $\lambda_i |f_i(x) - z_i^*|$ 取最大时, 意味着对于个体 x , 在第 i 个目标函数上与参考点相差最大。在优化的过程中, 不断改变 x , 是这个差值不断减小, 即可使得个体越来越趋近最优解, 完成对整个种群的优化。假设 $m=2$, 则可用图 2-8 来模拟这个过程。

如图 2-8 所示, $\lambda = \lambda_1, \lambda_2$ 将目标空间分割成两个部分。在 λ 的上方, 有 $\lambda_1 |f_1(x) - z_1^*| < \lambda_2 |f_2(x) - z_2^*|$, 则对于这个区域的个体, 其主要优化任务为 $\min \lambda_2 |f_2(x) - z_2^*|$ 。以此内推, 在 λ 的下方有 $\lambda_1 |f_1(x) - z_1^*| > \lambda_2 |f_2(x) - z_2^*|$, 其主要优化任务为 $\min \lambda_1 |f_1(x) - z_1^*|$ 。每一次优化都对最需要优化的目标函数进行优化, 能够确保最终个体在各个目标上都取得最优。

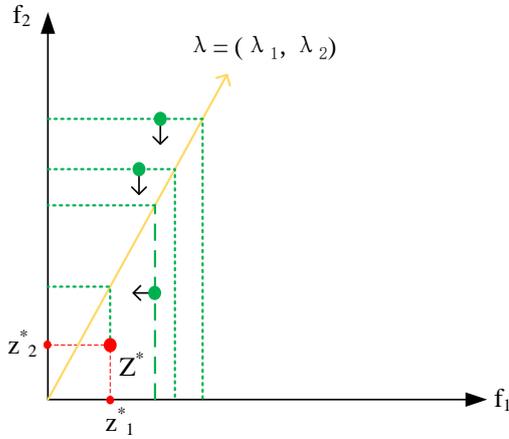


图 2-8 切比雪夫分解方法示意图

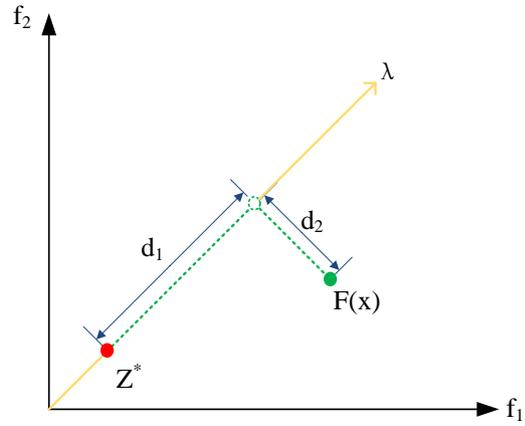


图 2-9 基于惩罚的边界交叉法示意图

基于惩罚的边界交叉法的聚合函数可以定义为式(2-17)，其种群优化过程示意图如图 2-9 所示。

$$\begin{aligned} \min g(x) | \lambda, Z^* &= d_1 + \theta d_2 \\ x &\in \Omega \end{aligned} \quad (2-17)$$

$$d_1 = \frac{\|Z^* - F(x)^T \lambda\|}{\|\lambda\|}, \quad d_2 = \|F(x) - Z^* - d_1 \lambda\|$$

在式(2-17)中， d_2 表示 $F(x)$ 在 λ 上的投影， d_1 表示该投影到参考点 Z^* 的距离， θ 即为惩罚系数。该方法是通过缩短 $d_1 + \theta d_2$ ，实现对个体的优化，使得个体越来越逼近参考点。惩罚系数 θ 是需要进行人为设置的参数，用以约束种群向权重向量方向分布。已有文献^[45]证明在 MOEA/D 框架下，当目标空间维度数大于 2 时，PBI 所求出的最优解能够取得更加均匀分布。但是 PBI 需要设置惩罚系数，如果设置不当将会严重影响算法性能。本文所需要优化的多目标问题维度数为 2，并且 PBI 的参数设置需求，势必会降低组合模型的泛化能力。综合考虑，在本文所提出的方法中，使用切比雪夫方法作为 MOEA/D 的分解方法。

2.4 本章小结

本章对所提出组合预测模型中的重要相关理论进行了阐述，包括：深度学习模型和多目标优化算法。主要针对它们的原理，优越性和在本文所提出方法中的作用进行了说明。

本章首先对 DBN 和 LSTM 两种深度学习模型进行了说明。LSTM 模型因其独特的结构，在拟合时间序列上有着独特的优势。在本文中，将会使用 LSTM 作为所提出组合模型的基础网络。DBN 则着重于探究神经元之间的激活概率，在本文中，使用 DBN 来对多个子模型进行组合输出。

其次对多目标优化算法进行了阐述。在本文所提出的组合模型中，使用多目标优化算法对基础网络的参数进行优化。这样做的目的是实现预测网络参数自适应，保证子模型对不同数据集进行拟合时，都能设置最优的相关参数，以提升整体的泛化能力。

第 3 章 基于多特征输入的 LSTM 模型分析

3.1 引言

作为智能电网的重要组成部分,智能电表以频繁的间隔获取用户的消费信息,并通过通信系统及时反馈给电网。在此基础上,对住宅级负荷进行准确的预测,将极大地促进电力系统的运行稳定、经济。然而,相较于系统级负荷,住宅级负荷波动更加剧烈。住宅级负荷变化受到诸多因素的影响,其中包括人类活动和天气情况等,哪怕是细微的因素变化也会引起负荷的剧烈波动。并且,随着可再生能源、电动汽车和配电网中时变负荷需求的加入,使得原本复杂的负荷数据变得更加难以预测。为了提高对住宅级负荷的预测精度,本章将住宅级负荷与系统级负荷进行对比,分析住宅级负荷数据的相关特性,并在此基础上设计了一种基于多特征输入的 LSTM 模型。为了验证模型的优越性,将该模型与诸多现有的预测模型进行对比实验。实验结果表明,本章所提出的模型对住宅级负荷有着更好的拟合效果,为之后提出的组合模型打下基础。

3.2 住宅级负荷数据特征分析

3.2.1 数据集介绍

在本文中,所使用到的数据集是由澳大利亚智能电网智能城市项目(Smart Grid Smart City, SGSC)提供的住宅用电数据^[46]。SGSC 是在 2010 年启动的澳大利亚首个商业规模的智能电网项目,在四年的项目开发期内,为新南威尔士州约一万名不同用户提供了智能电表,用于收集住宅用电数据。

该数据集目前处于完成公开的状态,一共包括两个 Excel 文件,都可以在指定网站中直接下载。其中,“sgsc-ct_customer-household-data-revised”文件记录了每一个用户的相关信息,如图 3-1 (a)所示。表中的“A”列表示用户编号,每一个编号对应一个用户,其他列则表示用户的用电基本情况和其他消费信息,主要包括有关电器的使用、居住位置的气候、零售和分销商品报价等。由于 SGSC 数据集过于庞大,对所有用户进行采样实验是不可能的,于是在本文的实验中,只挑选一组子数据集进行有关实验,而这些用户相关信息则为挑选实验对象提供了依据。文件“CD_INTERVAL_READING_ALL_NO_QUOTES”则记录了每一个用户 2010 年 1 月 1 日至 2014 年 12 月 31 日的全部半小时负荷数据(及一天共 48 个负荷数据采样点),如图 3-1 (b)所示。表中“A”列表示用户编号,“B”列表示数据采集时间,“E”列表示负荷数据。在本文中,之后的实验皆是以该数据

集为基础展开的。

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	CUSTOMER_KEY	TRIAL_CU	CONTR	TARIFF_PI	FEEDBAC	FEEDBAC	LIFESTYLE	INFERRE	VERIFIED	SERVICE		CONTR	NET_SOL	GROS
2	8459115	Retail	N					14		Domestic	1	0	0	
3	10015988	Retail	N					60		Domestic	1	0	0	
4	8922450	Retail	N					28		Domestic	1	0	0	
5	8922462	Network	N					28		Domestic	1	0	0	
6	8922472	Network	N	NETDPRPE				64	64	Domestic	1	0	0	
7	11189125	Retail	N					1		Domestic	1	0	0	
8	11189127	Retail	N					37		Domestic	1	0	0	
9	11189133	Retail	N					73		Domestic	1	0	0	
10	11189135	Retail	N					37		Domestic	1	0	0	

(a) “sgsc-ct_customer-household-data-revised”文件

	A	B	C	D	E	F	G	H	I	J
1	CUSTOMER_ID	READING_DATETIME	CALENDA	EVENT_K	GENERAL_SUPPLY_KWH	CONTRC	GROSS_C	NET_GEN	OTHER_KWH	
2	10006414	2012/2/10 8:00	233632	0	0.141	0	0	0	0	
3	10006414	2012/2/10 8:30	233635	0	0.088	0	0	0	0	
4	10006414	2012/2/10 9:00	233638	0	0.078	0	0	0	0	
5	10006414	2012/2/10 9:30	233641	0	0.151	0	0	0	0	
6	10006414	2012/2/10 10:00	233644	0	0.146	0	0	0	0	
7	10006414	2012/2/10 10:30	233647	0	0.077	0	0	0	0	
8	10006414	2012/2/10 11:00	233650	0	0.052	0	0	0	0	
9	10006414	2012/2/10 11:30	233653	0	0.055	0	0	0	0	
10	10006414	2012/2/10 12:00	233656	0	0.055	0	0	0	0	

(b) “CD_INTERVAL_READING_ALL_NO_QUOTES”文件

图 3-1 SGSC 数据集文件图示

3.2.2 数据缺失处理

对于任何种类的数据样本，数据缺失都是不可避免的。导致数据缺失发生的原因主要有两种，分别为机械因素和人为因素。其中，机械因素是由于数据采集设备的故障所导致的数据收集失败，或者存储设备的故障所导致的数据存储失败和丢失。人为因素是指由于操作者的不当行为，导致数据的采集、传输、存储不及时，或对存储数据的误删等原因所产生的数据丢失。由于负荷数据属于时序信号，前后的样本之间有着很强的相关性。如若发生数据缺失，将会严重影响学习机对数据相关性的学习。因此，如何处理数据缺失是负荷预测中一个至关重要的问题。

表 3-1 中，记录了当下常用的几种缺失值处理方法以及其特点。由于 SGSC

表 3-1 缺失值处理方法及其特点

方法	特点
删除法	直接删除缺失值所对应的样本，适用于大规模的数据集。
固定值法	以固定的数值对缺失值进行填充。
均值法	采用样本平均值对缺失值进行填补，不适用于存在大规模缺失值的样本。
插值法	通过对样本分析，计算出插入值，插入缺失值位置。比较常用的计算方法包括拉格朗日插值和牛顿插值。
回归方法	对样本进行建模，通过回归模型计算缺失值。该方法复杂、费时。

所提供的数据集非常的庞大，所以在本文中选用删除法来处理数据缺失的问题。

3.2.3 探索性数据分析

与系统层面的负荷不同，单个住宅用户负荷缺乏明显的一致性，使得对其进行准确的预测变得十分困难。一般来说，系统级负荷主要包括地区电网的总负荷、大型工业园区的总负荷等，这些负荷数据聚合了多个用户的负荷使得整体的负荷曲线更加具有一致性，且抗干扰能力强。图 3-2 中记录了新威尔士州区域电网和 10017466、10006572 两用户 2013 年 1 月 7 日至 2013 年 1 月 13 日连续七天的负荷曲线，正好对应周一到周日一个周期。从图中可以明显的看到，相比较于单个用户的负荷曲线，系统级负荷曲线更加平滑。文献^[47]指出，在单个家庭中，居民的日常生活习惯和生活方式以及拥有的主要电器类型都可能会对短期负荷曲线产生更直接的影响。例如，一些家庭在使用洗衣机后可能会继续使用烘干机对衣物进行烘干处理，这就意味着在随后的半个小时内会消耗大量电力。但是，如果当日的天气条件允许，或者住宅主人不急于使用清洗过的衣物，便很有可能不使用烘干机而选择晾晒。由于烘干机属于大功率电器，上述两种不同的选择就会造成负荷曲线的波动。

图 3-2 (b)记录的是用户 10017466 连续 7 天的负荷曲线，从图中可以明显看出，负荷曲线分为两种类型。第一种如横坐标 49 至 96(周二)、145 至 192(周四)、241 至 288(周六)和 289 至 336(周日)所示。这几天的负荷曲线表现出白天负荷大，晚上负荷小的特征，且峰值统一出现在上午，这符合人类日出而作日落而息的特点。第二种如横坐标 1 至 48(周一)、97 至 144(周三)和 193 至 240(周五)所示。这几天的负荷一直保持较低的数值，且每日负荷曲线之间的一致性较大。由此估计这几天住宅内无人人类活动，仅有常供电电器(例如冰箱)运行。图 3-2 (c)记录的是用户 10006572 连续 7 天的负荷曲线。与用户 10017466 不同，该用户连续 7 日的负荷值都很低，且保持着每间隔大约 8 小时就出现一次极大值的规律。类比于用户 10017466 的第二种负荷曲线，可以估计在这 7 天中，住宅内无人人类活动。

通过分析，可以得出，住宅负荷数据在时间上具有一定的连续性。如用户 10017466 的负荷曲线就是以天为单位，夹杂着少许的波动进行往复。由此，在本文的实验中将 LSTM 网络的时间步长设置为 48，以增加数据之间的相关性，方便模型对其进行学习。同时，也可以发现日期对负荷曲线的变化起着一定的影响。对于用户 10017466，可以猜测其住户是周一、周三、周五外出，而周二、周四、周末双休日在家，从而导致了其 7 天负荷呈现两种趋势。对于用户 10006572，可以猜测在 1 月 7 日至 1 月 13 日期间，其住户长期外出，才导致负荷曲线呈现长期无人回家的趋势。由此，将时刻、日期等特征加入到 LSTM 网络的输入中是

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/077131021052006031>