一、学位论文开题报告

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 论 文 题 目 | 基于分钟级电表数据的非侵入式负荷监测方法的研究与实现 | | | | | | | | |
| 研 究  方 向 | 大数据分析与处理 | | | | | | | | |
| 题 目 来 源 | 国家 | 部委 | 省 | 市 | 厂、矿 | 自选 | 有无合同 | 经费数 | 备注 |
|  |  |  |  | √ |  |  |  |  |
| 题 目  类 型 | 理论  研究 | 应用  研究 | 工程  技术 | 跨学科  研 究 | 其他 |  | | | |
|  |  | √ |  |  |  | | | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **一、选题依据与工程应用价值**  **1、选题依据**  电力系统是世界上极为重要的工业化网络系统之一。为了提高电网的信息化程度，促进电网更加经济、稳定的运行，面向电力系统的智能感知技术得到快速发展，开展智能用电的关键在于获取电力用户用电负荷监测信息以及对用户用电行为的智能分析。居民用户具有用电行为多样、用户数量大等特点，是智能用电双向互动服务体系主要的服务群体。尽管我国智能电网建设日益完善，但是受技术等多方面因素的限制，我国并未大规模开展对居民用户侧负荷监测技术的研究，常规用电分析技术已经无法满足灵活互动的智能用电需求，居民用电负荷监测和智能分析技术亟待完善和创新。为了更深入了解居民用电行为，支撑智能电网建设，需要开展对负荷监测的相关研究。  居民用电设备负荷监测主要分为侵入式负荷监测和非侵入式负荷监测。侵入式负荷监测是指在用电负荷上安装硬件监测设备，实时获取每个设备负荷的运行状态与用电量，该方法虽然能精确实现负荷监测，但是会耗费大量的人力与财力，影响居民生活，且不易维护。非侵入式负荷监测（Non-intrusive load monitoring, NILM）是指直接在主电路上安装非侵入式负荷监测装置，获取总用电负荷的电流、电压与功率信息，根据这些信息识别居民家庭内每个用电设备的负荷运行状态。该方法不用深入居民家庭内部，仅在外部安装采集设备，相对简单方便，已成为近年来负荷监测研究的主流方向。但是，在实际应用中，现有常规非侵入式负荷监测方法通过改造或新增监测设备采集的高频数据量庞大，对数据中心和通信网络造成负担，面临投入较多、成本较高、涉及范围较广、推广难度较大的困难，同时智能分析的准确性和时效性还不能满足实际需求。  因此，为充分利用现有居民电表及数据采集环境，本课题将研究面向负荷监测与智能分析的分钟级居民电表数据采集方法，提出能够体现居民用电差异化的多维负荷特征提取方法，构建典型居民用电场景的负荷特征库；建立面向分钟级居民电表数据的负荷辨识智能计算模型，设计并实现居民用电负荷在线辨识系统，实现不同用电场景下居民用电设备负荷的快速准确识别，支撑用能策略优化、用电设备故障检测、用电定向营销等试点应用，拓展居民智能电表的非计量功能。  **2、工程应用价值**  通过非侵入式负荷监测获得的数据和信息对于家庭用户和电力公司都有很大的价值。目前大多数用户都对用电情况缺乏一定的了解，不知道电器具体的能耗。研究表明如果能为用户提供详细的用电消耗反馈，每个月用户的用电量会下降5%到15%[1]。普通用户可以通过这种方式了解各个电器的用电情况，掌握各个电器的耗电状态，从而理性安排电器的使用，或者购买更加节能的电器，最终做到合理用电，减少浪费，达到节约电费的效果。电力公司可以通过这种方式，在不增加投入和侵犯用户隐私的前提下，了解各个用户的电器使用情况，一方面可以检测用户对于违规电器的使用，及时处理，减少用电安全隐患，提高电网的安全性；另一方面可以对各个区域乃至各个家庭的用电行为进行分析，做到对每个用户都实现有差异的定向服务，提升用户满意度。电器制造商还可以通过负荷分解获得的电器信息，查看分析电器的实际使用情况和电器规格是否匹配，提升电器制造质量。因此，非侵入式负荷监测这一研究领域，融合了电力、计算机、制造业等各大行业，可以服务多方，实现共赢，是未来的一大发展趋势。作为电力技术和计算机技术的交叉方向，也同时兼备了理论研究意义和实际应用价值。  **二、国内外研究现状与发展动态**  非侵入式负荷监测这一概念是在20世纪80年代由MIT的Hart提出[2]，目的是以较低的硬件成本获得最多的信息。在Hart的开创性研究中使用了稳态有功功率和无功功率指标，以此改进具有类似功耗电器的NILM系统性能。目前，负荷分解的模型及算法可以分为基于组合优化的方法和基于模式识别的方法两大类。  基于组合优化的方法尝试将观察到的功率测量值与设备特征库中设备功率信号的可能组合进行匹配，以减少匹配误差作为优化策略[3]。这类方法原理直观，对于设备数量较少且设备特征已知的情况具有相对较好的分解效果。但是该方法本质上是NP完全问题，优化提升困难；其次，该方法假设各设备的特征满足叠加性，但是并不是所有设备的负荷特征都满足这一特性；最后，该方法需要一个完整的设备特征库，这一需求在实际应用中很难被满足[4]。  基于模式识别的方法直接从数据中学习设备用电特征模型，从而实现对负荷的分解。根据学习过程中是否利用标记信息，可以进一步划分为有监督学习和无监督学习两类。有监督学习方法除了需要总负荷数据以外，还需要使用单独监测每个设备的电力数据进行训练以获得模型。这类方法需要丰富的数据支持，直接从数据中学习设备运行特征和运行模式，实现负荷分解。这些数据在实际生活中需要为每台设备安装独立的检测装置获取。常用的有K最邻近分类算法[5]、决策树算法[6]、支持向量机算法[7]，以及最近较为流行的深度学习等[8]。而无监督学习方法不需要标签进行模型的训练，直接从数据中挖掘特征的相似性。例如通过K-means等方法进行负荷特征提取和设备辨识。无监督方法由于其预测准确率较低，实际应用较少。在众多负荷分解模型和方法中，隐马尔可夫模型、稀疏编码、深度学习比较受到人们的关注。  隐马尔可夫模型（Hidden Markov Model，HMM）被引入负荷分解领域之后就成为研究的热点之一[9]。马尔可夫链是概率论和数理统计中具有马尔可夫性质的一种随机过程，过程中t+1步的随机状态仅与它在t时刻的状态相关，与其它的历史状态无关。隐马尔可夫模型是马尔可夫链的一个拓展，初始状态概率向量π、状态转移模型、观测概率矩阵决定了一个马尔可夫模型。这一模型可以表示为：。其中，状态序列由和确定，确定观测序列。在非侵入式负荷监测问题中，观测序列为总电表采集的时间序列，隐含状态序列为每个电器的工作模式。基于隐马尔可夫模型的非侵入式负荷监测主要由模型学习部分和解码部分构成。其中，模型学习部分由电气数据估计训练参数，以最大化观测序列概率；解码部分由输入的观测序列及先前估计的模型，计算得出给定观测序列条件下概率最大的状态序列。隐马尔可夫模型自从被提出后，很多研究人员对它进行了相关研究。用HMM进行负荷分解的问题主要是模型所需的状态数量和运算量随着设备数量和设备状态数目的增加呈指数型增加。如果共有N个用电设备，每个用电设备有K个工作状态，则总的HMM状态数为,基于HMM的算法改进主要集中在如何降低算法的复杂度。Kolter等将因子隐马尔可夫模型（Factorial HMM, FHMM）引入了非侵入式负荷监测，该模型的隐含状态由多个独立的状态序列构成，每台用电设备对应一条独立的状态序列，观测序列的值由对应时刻的多个独立隐藏状态变量决定。但是FHMM中假设每台设备之间的状态是相互独立的，忽略了设备之间的相关性信息。此外，FHMM的模型复杂度仍然较高，并且要得出精确推理所需的状态数量仍然会呈指数型增加，这给实际应用带来了困难[10]。  稀疏编码也是较为受到关注的一种负荷分解方式。稀疏编码用来寻找一组“超完备”基向量来更高效地表示样本数据，使得能将输入向量表示为这些基向量的线性组合。这组基向量的好处是它们更能有效地找出隐含在输入数据内部的特征和模式。在负荷分解领域中，为每一台设备建立相应的负荷矩阵。训练时使用单独设备的负荷和总负荷去训练每台设备的模型，实际使用时将这些学习模型结合起来，仅使用它们的累计信号来分解从未见过的家庭中不同设备的功耗[11]。  基于深度学习的方法近些年来在各个领域都得到了运用且取得了不错的效果[12]。2015年，深度学习方法被引入非侵入式负荷监测领域，并被越来越多研究人员所接受并不断改进。其中，Mauch提出了采用深度神经网络和HMM相结合的方法[13]，经过在公开数据集REDD上测试后，该方法比之前的FHMM取得了更高的预测准确率。 Kelly提出了三种基于深度学习的非侵入式负荷监测方法[14]：基于长短期记忆网络（Long Short-Term Memory，LSTM）的方法[15]、基于降噪自动编码器（Denoising AutoEncoder，dAE）的方法[16]和深度神经网络方法。Kelly提出的方案属于Sequence to Sequence模型，即序列到序列模型。模型输入的是电器总功率序列，输出的序列是目标设备的功率，其采用了CNN[17]、RNN[18]等多种神经网络来学习电路总功率和单个电器功率之间的关系，构建非线性回归模型。实际运用中，因为统计的功率序列很长，同时作为训练数据和标签输入模型，会导致模型计算速度缓慢，难以使用。此外，因为较长的功率序列难以反映某个时间段内电器的功率变化曲线，训练效果也相对较差。为了应对这种问题，滑动窗口（Sliding Window）方法被应用进来。输入数据会被划分为一个个适当长度的滑动窗口，并根据不同的电器运行模型和功率曲线设置不同的滑动窗口大小，每次输入一个滑动窗口的数据对神经网络进行训练。出于提高运算速度和较小运算量的需求，实际应用中可以从全部可能的滑动窗口中选择部分窗口对模型训练。这种序列到序列模型对分解结果中序列的每个时间点都进行了多次预测，最终需要对每个时间点的多个预测结果取平均值，作为最终的结果。构建神经网络后，使用反向传播的方式对网络进行训练。训练神经网络的训练集使用了真实采集数据和混合数据进行训练，验证集和测试集则全部使用了真实数据。目前网络上有部分公开的电力数据集，Kelly使用了UK-DALE（Uk Domestic Appliance-Level Electricity）数据集对三种模型分别进行了测试[19]。在众多电器中，作者挑选了微波炉、冰箱、洗碗机、洗衣机和开水壶五种设备，使用Accuracy、F1、MSE等模型评价指标进行评估。通过测试结果不难发现，相比较传统的方式，使用上述三种神经网络均取得了更高的F1分数，并且模型的泛化性能也更为优秀。  目前，大多数对于非侵入式负荷监测的研究都集中于秒级或者毫秒级的数据，而现有居民电表采集数据的频率为分钟级，为了实现秒级或毫秒级数据的采集，需要增加专用的采集设备。实际应用中，考虑到硬件和软件成本，最现实的方案应当针对分钟级别的数据进行负荷分解。采集分钟级数据，只需要在软件层面对电表进行改造，成本较低。但是，与秒级和毫秒级这类相对高频数据相比，针对分钟级数据的非侵入式负荷监测存在着预测精准度较低、采集数据量的时间跨度较大等问题。因此，需要引入居民家庭的社会特征，以提高面向分钟级数据的负荷辨识精度。  **三、研究目标与研究内容**  **1、研究目标**  本课题的研究目标是，针对常规的非侵入式负荷监测成本较高、涉及范围较大、精确度不足等问题，基于深度学习等技术，研究数据采集和预处理方法，建立面向分钟级居民电表数据的负荷分解智能计算模型，设计并实现在线负荷识别系统，实现不同用电场景下居民用电设备负荷的快速准确识别，从而提供用户人机交互、数据可视化的能力，支撑用能策略优化、灾难预警等居民智能电表的非计量功能。  **2、研究内容**  **（1）数据采集和预处理**  为了解决当前非侵入式负荷监测方法存在的问题，研究基于居民智能电表的分钟级用电设备的电气特征指标及相应的采集与预处理方法；在此基础上，进一步研究能够反映家庭用户用电行为的社会特征指标及相应的采集与预处理方法，以弥补用电设备的分钟级低频电气数据刻画用户用电行为特征的不足。  **（2）基于混合数据的多输入负荷监测方法**  基于采集和预处理后的用电设备电气特征数据和居民用户的社会特征数据，建立可处理混合数据的多输入负荷监测模型，针对电气特征数据和社会特征数据分别采用不同的处理方式，对负荷监测模型进行学习训练，通过特征融合实现居民用户用电设备的负荷辨识。  **（3）非侵入式负荷监测系统的设计与实现**  设计并实现完整的非侵入式负荷监测系统，包括数据供给模块，负责从电表收集数据并保存数据至数据库；模型训练模块，负责设定训练参数，进行模型训练等；用户交互模块，负责向管理人员提供可视化信息如训练进度、预测与实际差值等。  **四、实施方案及可行性分析**  **1、实施方案**  **（1）数据采集和预处理**  数据采集包括采集居民用电设备的电气特征数据和居民家庭的社会特征数据两部分。所需电气数据直接从运行中的智能电表中读取，其收集主要通过采集器和集中器完成。采集器用来接收电子式电表用电产生的输出脉冲，记录电表用电情况并上传到集中器。集中器是远程集中抄表系统的中心管理设备和控制设备，负责定时读取终端数据、系统的命令传送、网络管理、事件记录等。目前正在使用的采集器和集中器一般是以15分钟到1小时的频率从智能电表中读取数据。该频率的数据捕捉不到大部分使用时间较短的电器的运行状态，如微波炉、烧水壶等，价值较低，一般只能应用于计算用户的用电总量，并不适用于用户用电行为分析等非侵入式负荷监测的应用。如果想监测更高频率的秒级数据，需要进行硬件升级，成本较高。由于不足1分钟的电器运行时间很短，监测意义不大，出于综合考虑，采用1分钟的频率进行数据采集。为此需要对集中器进行软件层面的改造，通过修改集中器中的采集程序，结合集中器——电表之间的通信能力，将数据采集频率从15分钟一次调整为1分钟一次，同时处理超时、容错等异常。  在实际运用中，使用普通的电表和经过特别定制的负荷监测信息采集装置进行数据采集。电表和集中器如图4-1所示，用电设备电气特征数据采集系统如图4-2所示。电表、采集器、集中器等设备依次通过通信通道相串联，最下层采集层级通过通信通道与智能终端相连接，最上层采集层级通过通信通道与边缘物联代理相连接，每个采集层级均运行周期采集任务，从而实现“分层采集、本地存储”的采集策略。    (a)电表 (b)集中器  图4-1 低频数据采集设备示意图    图4-2 用户电表分钟级数据采集示意图  根据电器设备不同的电气特征，可以将电器设备负荷特征分为两种：稳态特征和暂态特征。稳态特征是电器设备在保持稳定运行状态下体现出的电气特征，也是负荷监测的主要特征；暂态特征是电器设备在一个暂时的状态中，如启动、切换运行状态、关闭电器等的瞬间所体现出的电气特征。暂态特征由于其瞬时性的特性，分钟级数据采集难以捕捉到，作用有限，本课题主要采用稳态特征。  在社会特征方面，我们将电表测量的范围称为一个家庭，该家庭除了电表测量的功率特征外，还包括一系列社会特征，如家庭特征、地理特征、气象特征等，包括但不限于表4-1中的部分信息。对于人口数量较多的家庭，大型电器如洗衣机、热水器等使用频率可能相对较高；在温度较高的时期，空调的使用频率可能较高；对于地域靠北的家庭，空气加湿器等电器的使用频率可能相对较高。为此，我们对每个家庭进行社会特征空间构建，数据采集方式采用家庭走访和问卷调查的形式，对小区内每个家庭发放问卷，条件允许的话进行上门走访，最终完成社会特征数据的收集和整理。  表4-1 居民家庭社会特征指标   |  |  | | --- | --- | | 家庭 | 人口数量 | | 成员职业 | | 房屋面积 | | 房屋地域 | | 气候 | 气象温度 | | 气象湿度 | | 时间 | 是否是工作日 |   采集完数据集之后，需要对数据进行预处理。对于电气特征数据，由于电表采集缺失或者采集延迟等原因，并不是准确的每1分钟采集一次。实际电表在收集数据时，可能在当前时刻并未采集到数据或数据在网络传输过程中丢失，采集的具体时刻也有部分偏差，导致实际采集的数据并不完整。为了解决由于各类原因导致数据丢失的问题，减少对最终负荷辨识的影响，在训练过程中，对缺失数据采用取前后时刻均值填充的方式，形成一个较为平滑的数据曲线。而在在线负荷辨识过程中，如果当前时刻并未采集到数据，使用前填充的方式，取前一时刻的数据填充进来，这样可以在一定程度上弥补数据不完整或数据准确度不高的问题。当遇到大片数据缺失的情况，只有第一个缺失值会得到填充，防止大片填充相同或无意义的数据。最后，每个时间点都需要有总功率和该时刻相应电器的功率，当总功率或者相应电器功率二者任一为空时，即删除掉该记录，降低其对预测结果产生的影响。处理完后，使用标准化方式将数据转换为均值为0、标准差为1的分布。  社会特征数据方面，对于数值型这类具有连续特性的数据，如温度、人口数量等特征指标，采用最小-最大缩放，将数据归一化到[0, 1]区间，具体实现为。其中，为当前样本值，max为当前样本空间最大值，min为当前样本空间最小值。对于不同类型的分类型数据，采用one-hot编码的方式进行预处理，如工作日参数，以0和1的形式表示当前是工作日或不是工作日。  **（2）基于混合数据的多输入负荷监测方法**  本文综合居民用电设备的电气特征数据和居民家庭的社会特征数据，提出了基于混合数据的多输入负荷监测模型，如图4-3所示。    图4-3 基于混合数据的多输入负荷监测模型  电气特征数据处理部分，Kelly提出的Sequence to Sequence方法为每个电器设备单独建立一个模型，模型输出的是单个电器在该时间窗口内每个时间点的功率值，最后，取该时间点在多个时间窗口重叠部分的均值作为该电器设备负荷辨识的结果。考虑到每个用户家庭可能存在众多类型和同类型不同型号的电器设备，单独为每个电器训练模型会消耗大量的时间和算力。因此，本论文综合考虑了时间成本和电器设备之间的互相影响，对Sequence to Sequence方法进行了改进，输入是某个时间窗口大小的电气总功率数据，但输出是多个电器设备的功率，对同类型电器进行同时预测，极大的减少了模型训练的时间成本。  该模型使用了滑动窗口方法，假设中点元素可由输入总功率的非线性回归来表示，输入的训练数据为某个时间序列的总功率数据，输出为该时间窗口中点的对应电器的分解功率。对于功率变化，电器设备功率的状态与其前后时刻的总功率相关性较大，电器设备的功率可由总功率的变化曲线表现出来，总功率的变化情况是神经网络用来推断电器设备状态的重要特征之一。与Sequence to Sequence模型相比，该模型定义的神经网络将输入总功率的滑动窗口映射到对应滑动窗口对应的中点。该模型可表示为  其中，为电器在时间窗口中点的功率，为时间窗口长度，为当前时刻，为一个时间窗口内的总功率，为W维的高斯噪声。将时间窗口映射到一个该时间窗口中点。而用来训练网络的损失函数可表示为  其中，为网络的参数，为时间序列总长度。这样做的好处是对于每个，都有一个单一的预测，而不是每个窗口都进行一次预测后取平均值。当需要辨识的电器设备规模为N个时，与Sequence to Sequence模型相比，本文所提方法的输出总量从减少为。其中，是输入滑动窗口的长度，是输入时间序列的总长度。这一特性使得神经网络主要注重于预测窗口的中点，而不是预测较为困难的窗口边缘数据。电气特征数据部分用一维卷积层对输入的数据特征进行自动提取，双向LSTM层用来对整体的时间序列进行预测，最后，使用全连接层输出数据至联合层。经验证该模型相比Sequence to Sequence方案能产生更加精准的预测结果。  社会特征输入数据为该时刻的温度、湿度、家庭人口构成、家庭地区等多种社会特征，每种特征数据都会进行单独的预处理，如缩放、标准化等。社会特征处理模型部分，采用一维卷积层进行特征提取，连接多层全连接层，最后输出数据至联合层。联合层用于特征联合，将社会特征与电气特征进行融合。最终通过全连接层输出负荷辨识结果，即为每台电器设备在该时间窗口中点的功率，最终输出的功率数量为同时进行辨识的电器设备数量。  **（3）非侵入式负荷监测系统的设计与实现**  经过多次验证和模型结构优化后，需要设计并实现一个完整的非侵入式负荷监测系统。该系统由以下功能模块构成，图4-4展现了具体的系统功能模块图。    图4-4 非侵入式负荷监测系统模块图  1. 数据处理模块：该模块可以用于收集实时电力数据，并将数据保存至数据库。在训练时向模型输出训练数据，辨识时提供总功率数据来检测模型辨识效果。数据预处理部分用于进行缺失数据填充，错误数据处理，标准化，归一化等。  2. 模型训练模块：该模块用于设定模型训练的参数和训练模型。设定参数，如一次性加载的数据量大小、输入窗口大小、训练周期、采样时间、场景选择等，在精度要求不同、时间及时性需求不同、电器组合不同等场景下，可设定不同参数，以达到最优效果。该模块为每个电器训练一个模型，并保存在磁盘上，用于实际分解时加载使用。  3. 在线辨识模块：该模块主要用于负荷在线辨识。该模块读取某个时间段内的总电表数据，并进行负荷辨识并将辨识的结果以图形的方式展现。同时，该模块会输出辨识的准确率，精确率，召回率等信息。  数据部分采用MySQL数据库进行存储。每台电表对应一张数据表，记录原始数据。每张表记录信息包括采集时间戳、电表id、功率、电压、电流、是否为工作日、温度、家庭人数等，如表4-2所示。  表4-2 电表采集信息表   |  |  |  |  | | --- | --- | --- | --- | | 字段名 | 类型 | Key | 备注 | | id | INT | PRI | 主键id | | meter | INT | - | 电表id | | KW | INT | - | 功率 | | V | INT | - | 电压 | | A | INT | - | 电流 | | timestamp | BIGINT | - | 采集时间 | | isworkday | VARCHAR(10) | - | 是否为工作日 | | centigrade | INT | **-** | 温度 | | people | INT | **-** | 家庭人数 |   对于社会特征信息建立数据表，包括各类社会特征，如表4-3所示。  表4-3 社会特征信息表   |  |  |  |  | | --- | --- | --- | --- | | 字段名 | 类型 | Key | 备注 | | id | INT | PRI | 主键id | | people | INT | **-** | 家庭人数 | | occupation | INT | - | 成员职业 | | space | INT | - | 房屋面积 | | location | INT | - | 房屋地域 | | centigrade | INT | **-** | 温度 | | humidity | INT | - | 湿度 | | isworkday | VARCHAR(10) | - | 是否为工作日 | | timestamp | BIGINT | - | 采集时间 |   电表分为总电表和分电表，记录接在某台总电表下的分电表，使用电表层次信息表存储，如表4-4所示。  表4-4 电表层次信息表   |  |  |  |  | | --- | --- | --- | --- | | 字段名 | 类型 | Key | 备注 | | id | INT | PRI | 主键id | | parent | INT | - | 总电表id | | child | INT | - | 分电表id |   对于电器设备信息建立表，包括电器名、平均功率、最大功率、电器类型等信息，如表4-5所示。  表4-5 电器设备信息表   |  |  |  |  | | --- | --- | --- | --- | | 字段名 | 类型 | Key | 备注 | | id | INT | PRI | 主键id | | meter | INT | - | 电器名 | | average | INT | - | 平均功率 | | max | INT | - | 最大功率 | | type | VARCHAR(30) |  | 电器类型 |   数据传输流程如下：  （1）电表采集电器数据并保存至数据库，存储到对应的电表数据表。数据进行去噪等预处理工作后，存储到该电表对应的预处理数据表；  （2）负荷监测模型从数据库中读取预处理后的数据，进行模型训练；  （3）读取某时间段的总电表数据，进行在线辨识后，将数据以图表和数字的方式展示给用户。  系统开发环境为ubuntu18.04，运行环境版本为jdk1.8、python3.7、tensorflow2.0。前端采用Thymeleaf进行页面展示，提供参数设定、设备信息管理、分解信息查看、对比等功能。后台采用SpringBoot框架，主要使用Java进行系统开发，Python实现数据处理和负荷分解模型的具体实现，并提供接口供调用。  **2、可行性分析**  （1）技术可行性  目前，使用深度学习的相关方法解决非侵入式负荷监测问题已经越来越成熟。本论文着重于在基于混合数据的多输入模型上，综合电气特征和社会特征对用电负荷进行分解，并且提高分解效率和分解精准度。在需要对用户用电行为进行分析时，能够提供较为准确的分解结果，足够支撑开展基于现有电表数据的用户用能模式、用电设备故障检测，用能策略优化等非侵入式负荷监测应用研究。并且已经在公开数据集上进行了验证，将采样频率降低到1分钟，取得了不错的分解结果。  （2）个人能力可行性  本人在本科阶段就学习计算机科学与技术专业，学习过数据结构等计算机基础知识。并且研究生入学后，不但注重必修课程的学习，还利用课外时间，认真学习机器学习、深度学习等相关知识，仔细阅读了很多优秀的书籍和负荷分解相关的论文，并从中汲取了不少理论知识，学习了他人切入问题的角度和解决问题的能力等。此外，我也不断锻炼自己的开发和编码能力。在学校时期，学习了主要用于深度学习的python语言和常用的深度学习框架tensorflow等，具备了运用神经网络等相关知识独立解决问题的能力。  （3）企业支持可行性  本论文选题来自国家电网正在投入研究的科研项目，整体方案经过多次仔细评估和可行性分析，公司对此很重视并投入了不少资源，可以与公司同事互相探讨研究项目过程中遇到的阻碍，迅速解决问题。因此，在企业的全方位支持下，能够完成此次论文的研究工作。  （4）导师指导可行性  本论文从研究的方向以及选题的确定，都和校内外导师进行了长时间的沟通和交流。其中，校内导师多年从事计算机教学和学生指导工作，并在电力方向有过项目经验，专业知识丰富，能够在学习道路上提供很多学习方法，在小问题上提供思路和理论指导，大方向上考虑项目的整体设计，把控研究方向。企业导师也关注于非侵入式负荷分解相关的研究，对项目中存在的阻碍和难点，能够及时提供相关的技术指导和详细的解决方案。  **五、参考文献**   1. DARBY S, et al. The effectiveness of feedback on enery consumption[J]. A Review for DEFRA of the Literature on the Literature on Metering, Billing and direct Displays, 2006, 486(2006): 26 2. Hart G W. Nonintrusive appliance load monitoring[J]. Proceedings of the IEEE, 1992, 80(12):1870-1891 3. Du Y, Du L, Lu B, Harley R, Habetler T. A review of identiﬁcation and monitoring methods for electric loads in commercial and residential buildings[J]. Proceedings of 2010 IEEE, Energy Conversion Congress and Exposition. Atlanta, USA: IEEE, 2010. 4527−4533 4. Cheng Xiang, Li Lin-Zhi, Wu Hao, Ding Yi, Song Yong-Hua, Sun Wei-Zhen. A survey of the research on nonintrusive load monitoring and disaggregation[J]. Power System Technology, 2016, 40(10): 3108−3117 5. Yang C C, Soh C S, Yap V V. A systematic approach in appliance disaggregation using k-nearest neighbours and naive Bayes classiﬁers for energy eﬃciency[J]. Energy Eﬃciency, 2017, 11(1): 239−259 6. Buddhahai B, Wongseree W , Rakkwamsuk P . A Non-Intrusive Load Monitoring System Using Multi-Label Classification Approach[J]. Sustainable Cities and Society, 2018, 39 7. Suykens J A K , Vandewalle J . Least Squares Support Vector Machine Classifiers[J]. Neural Processing Letters, 1999, 9(3):293-300 8. Yao X . Evolving artificial neural networks[J]. Proceedings of the IEEE, 1999, 87(9):1423-1447 9. Makonin S, Popowich F , Baji I V, et al. Exploiting HMM Sparsity to Perform Online Real-Time Nonintrusive Load Monitoring[J]. Smart Grid, IEEE Transactions on, 2016, 7(6):2575-2585 10. Kolter J Z, Jaakkola T. Approximate inference in additive factorial HMMs with application to energy disaggregation[J]. Proceedings of the 5th International Conference on Artiﬁcial Intelligence and Statistics. Canary Islands, Spain:JMLR, 2012. 1472−1482 11. Kolter J Z, Batra S, Ng A Y. Energy Disaggregation via Discriminative Sparse Coding[C]. International Conference on Neural Information Processing Systems. Curran Associates Inc. 2010. 69-73 12. Bengio Y. Learning Deep Architectures for AI[J]. Foundations & Trends in Machine Learning, 2009, 2(1):1-127 13. Mauch L, Yang Bin. A novel dnn-hmm-based approach for extracting single loads from aggregate power signals[J]. Proceedings of the IEEE 2016 Acoustics, Speech and Signal Processing International Conference (ICASSP). Shanghai, China: IEEE, 2016. 2384−2388 14. Kelly J, Knottenbelt W J. Neural nilm: deep neural networks applied to energy disaggregation[J]. Proceedings of the 2nd ACM International Conference on Embedded Systems for Energy-Eﬃcient Built Environments. Seoul, Korea: ACM, 2015. 55−64 15. Hochreiter S, Schmidhuber J . Long Short-Term Memory[J]. Neural Computation, 1997, 9(8):1735-1780 16. Vincent P, Larochelle H, Bengio Y, et al. Extracting and Composing Robust Features with Denoising Autoencoders[C]. Machine Learning, Proceedings of the Twenty-Fifth International Conference (ICML 2008), Helsinki, Finland, June 5-9, 2008. 2008 17. Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks[J]. Advances in neural information processing systems, 2012, 25(2) 18. Mauch L, Yang B. A new approach for supervised power disaggregation by using a deep recurrent LSTM network[J]. Proceedings of IEEE Global Conference on Signal and Information Processing (GlobalSIP). Orlando, USA:IEEE, 2015. 63−67 19. Kelly J, Knottenbelt W. The uk-dale dataset, domestic appliance-level electricity demand and whole-house demand from ﬁve uk homes[J]. Scientiﬁc Data, 2015, 3   研究生签名  年 月 日  （注：本页不够可附页） |

二、学位论文工作实施计划

（一）论文的理论分析与硬件要求及其预期达到的水平与结果

|  |
| --- |
| 理论分析：   1. 深入理解非侵入式负荷监测相关技术 2. 深入理解应用深度学习方法进行负荷分解 3. 深入理解模型优化策略   硬件要求：   1. 实时监测电表和经过软件优化的采集器 2. 具有高效运算速度的GPU   预期达到的水平与结果：  采用综合电气特征和社会特征的方案，在只对电表进行软件层面改造的基础上，对电器设备负荷特征数据和居民家庭社会特征数据进行采集、预处理和训练，提出一个完善的非侵入式负荷监测模型，能够较为精确地将总功率分解为单个电器的功率。最后设计并实现非侵入式负荷监测系统，提供完整的数据处理和训练流程，以及人性化的系统界面，方便管理人员使用。 |

（二）论文工作进度与安排

|  |  |  |  |
| --- | --- | --- | --- |
| 起讫  日期 | | 工 作 内 容 和 要 求 | 备 注 |
| 2020.9 -2020.12 | | 熟悉课题方向的背景和理论知识，查找文献论文，确定论文研究内容，阅读相关资料 |  |
| 2021.1-2021.3 | | 研究数据采集和预处理方法 |  |
| 2021.4-2021.8 | | 设计并优化负荷分解模型 |  |
| 2021.9-2021.11 | | 设计并实现非侵入式负荷监测系统 |  |
| 2021.12-2022.1 | | 系统测试与优化 |  |
| 2022.2-2022.3 | | 撰写和修改学位论文 |  |
| 学校指导教师对开题报告的综合意见 | 指导教师（签字）  年 月 日 | | |
| 校外指导  教师对开题报告的综合意见 | 指导教师（签字）  年 月 日 | | |