**一种基于强化学习的新能源场站数据处理方法**

**技术领域**

本发明涉及新能源场站数据处理领域，设计一种基于强化学习的新能源场站数据清洗方法。

**背景技术**

随着全球经济的发展，能源需求逐年递增，由于不可再生能源的枯竭和环境污染问题，新能源以其安全性、可再生性、环保性受到各国高度重视和广泛应用，并呈现出良好的发展前景，其中新能源发电以光伏发电和风电发电为主。随着电网中风电、光伏占比的不断升高，特高压直流输电逐步投运，电网运行与结构愈加复杂，系统功率平衡及调频难度不断加大。为提升新能源消纳比例，电源结构发生较大变化，电网可用的快速频率响应资源逐步减少。对于电源输送端大电网来说，电网频率控制特性的结构性困境日趋明显，发生电网频率安全问题的风险进一步加大。因此对新能源发电并网运行监测数据进行研究就显得十分重要。

然而，由于电场运行时机组弃风、弃光、检修、极端天气情况、外界电磁干扰或设备故障等原因，导致原始数据中存在大量不完整的数据和异常的数据。这些不良数据严重影响预测模型的参数估计，造成预测精度低、预测偏差大等后果。因此，有必要进行有效的数据处理。

在文献《数据清洗在新能源功率预测中的研究综述和展望》中，作者从异常值剔除和缺失值重构两个方面，详细地阐述、分析和归纳了新能源异常数据的类别和处理方法。然而传统的数据插补法虽然简单快速，但是如果数据波动较大，插值不准确，会丢失一些数据的特征信息，还可能会引入偏差，而且没有考虑相邻数据间的时序信息。

**发明内容**

本发明的目的是设计一种基于强化学习的新能源场站数据处理方法，通过对采集到的监测数据进行特征学习，实现数据的有效扩增，一定程度解决采集过程中的数据缺失等噪声问题，在填补缺失数据的基础上，不会丢失时序数据的前后依赖关系。处理后的数据可为后续的新能源场站相关研究提供数据支撑。

本发明根据已有数据样本建立深度Q网络强化学习模型（Deep Q Network, DQN）。DQN 可以处理具有时间序列特性的问题，模型以缺失点附近的数据作为特征，预测缺失点的数据，通过海量的历史数据中挖掘相似的变化，从而进行更精准的数据填充。其填充数据不仅只关注自身的时序演变，同时关注相邻时序，特别是有影响关系的时序指标之间的影响，例如，对于风速数据的缺失，可以通过 DQN 学习历史风速数据的变化趋势和模式，来预测缺失的风速值。

本发明的有益效益是，能够解决采集新能源场站数据过程中的数据缺失等噪声问题，处理后的数据可为后续的新能源场站相关研究提供数据支撑。

**附图说明**

图1是常见的数据缺失处理方法及优缺点；

图2是该发明的强化学习模型与环境交互；

图3是该发明的流程图；

**具体实施方案**

1.明确数据缺失情况

分析新能源场站数据缺失的类型，如电压和电流的基波、谐波和间谐波分量的幅值、相角、频率和频率变化率等，确定哪些数据对场站的运行和决策至关重要，以下假设新能源场站数据的频率缺失。

2．定义强化学习环境

状态空间：包括当前已有的新能源场站数据（未缺失的数据）、时间信息、数据缺失的位置标识等，用向量

来表示状态空间，其中表示电压和电流的基波，表示其谐波，分别表示间谐波分量的幅值，相角和频率变化率。

动作空间：可以是预测模型或数据插补方法，如回归模型预测，时间序列模型（如自回归积分滑动平均模型）预测，多重插补法插补等，用向量

其中分别表示回归模型，时间序列模型和多重插补法的填充值。

奖励函数：设计奖励函数以衡量算法的性能，根据数据恢复的准确性、对场站运行决策的帮助程度等因素来确定奖励。本发明通过计算恢复后的数据与真实数据之间的误差来确定奖励，误差越小奖励越高。恢复后的数据与真实数据之间的误差由均方误差方法计算：

其中，是数据点的数量，是真实数据值，是恢复后的数据值。

奖励函数设计如下：

其中为很小且不为0的常数，以避免除数为0的情况。

3.构建DQN模型

神经网络结构设计：使用卷积神经网络来提取缺失点附近数据的特征，Batch Normalzation(BN)层提高模型训练效率，池化层减少网络参数，必要时可在数据输入时使用dropout操作，减少过拟合，输入层接受状态表示，最后全连接层作为输出层对应每个动作的Q值估计。

经验回放与目标网络：设置经验缓冲区，存储模型与环境交互产生的经验：

其中，分别为时刻模型所处的状态，选择的动作，获得的奖励，为时刻模型的状态。

在训练过程中，随机从缓冲区中抽取小批量的经验进行学习，打破数据的相关性。

同时，引入目标网络，目标网络在 DQN 中的主要作用是提供一个相对稳定的目标值，用于计算损失函数和更新主网络。目标网络的结构通常与主网络（即当前用于决策的网络）相同。这意味着它具有相同的输入层、隐藏层和输出层。通过定期更新目标网络的参数，使得训练过程更加稳定，减少训练中的波动，提高算法的收敛速度和性能。使用软更新策略更新目标网络：

其中，是目标网络的参数，是主网络的参数，是一个小于1的更新比例参数。较小的值会使目标网络更新更缓慢，更稳定；较大的值会使目标网络更新更快速，但可能会导致不稳定，通过实验确定合适的软更新比例参数。

4.模型的训练与优化

初始化模型：基于高斯分布，随机初始化 DQN 的当前网络和目标网络的参数，设置学习率、折扣因子等超参数。

与环境交互进行训练：模型根据当前环境选择一个动作，执行动作后观察奖励和下一个状态，将这个经验存储到回放缓冲区中。从回放缓冲区中随机抽取小批量经验，计算目标 Q 值并更新当前网络的参数，使得当前 Q 网络对动作在状态下的 Q 值估计接近目标 Q 值。定期更新目标网络的参数，例如每隔一定的训练步数复制当前网络的参数到目标网络。

计算损失函数：对于一个状态和动作，主网络预测的Q值为:，目标网络预测的目标值为

其中是当前的奖励，是折扣因子，损失函数计算为：

更新主网络：使用Adam优化算法来最小化损失函数，从而更新主网络的参数。

优化策略：使用优先经验回放，根据经验的重要性进行抽样，给予重要的经验更高的被抽取概率，从而使模型更快地学习到关键的信息。在新能源场站数据缺失问题中，重要的经验是那些导致较大误差或对数据恢复有重大影响的状态转移。使用TD误差（Temporal Difference error）来计算优先级，TD误差是当前估计的Q值与目标Q值之间的差异。TD误差：

将TD误差的绝对值作为经验的优先级。

5.模型的评估与应用

模型评估：使用测试集数据对训练好的 DQN 模型进行评估。计算数据恢复的准确性、对频率预测等场站决策任务的改善程度等指标。比较恢复后的数据与真实数据的误差，评估模型在不同缺失模式下的性能。

实际应用：将训练好的 DQN 模型应用于新能源场站的实际数据缺失问题中。当出现数据缺失时，模型根据当前状态选择最优的数据插补方法进行数据恢复。持续监测模型的性能，根据实际情况进行调整和优化。

**说 明 书 附 图**

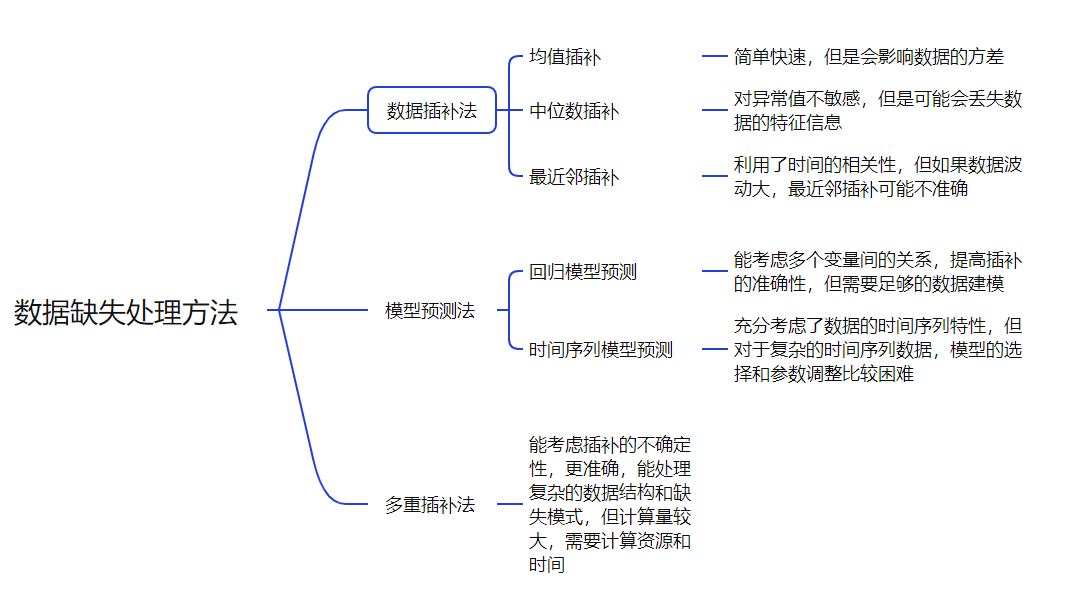


图1

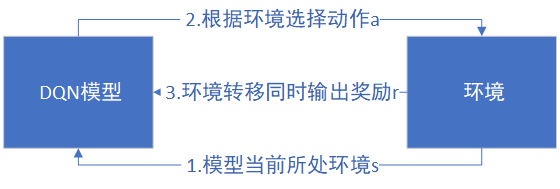


图2

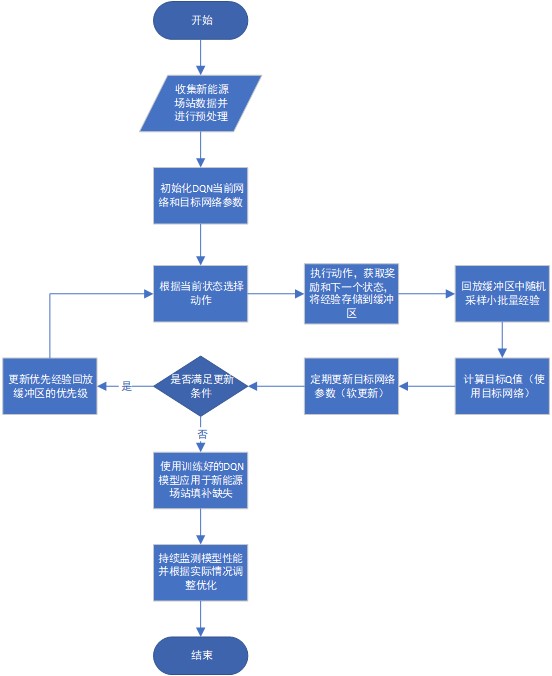


图3