



(12) 发明专利

(10) 授权公告号 CN 109978222 B

(45) 授权公告日 2021.04.27

(21) 申请号 201910009520.1

(22) 申请日 2019.01.04

(65) 同一申请的已公布的文献号
申请公布号 CN 109978222 A

(43) 申请公布日 2019.07.05

(73) 专利权人 国网山东省电力公司电力科学研究
院

地址 250002 山东省济南市市中区望岳路
2000号

专利权人 山东大学
山东鲁能软件技术有限公司
国家电网有限公司

(72) 发明人 孙树敏 王士柏 赵岩 程艳
杨明 王楠 张兴友 王玥娇
滕玮 于芑 李广磊 魏大钧
王勃 赵元春 马嘉翼 王立峰
王尚斌 李洪海

(74) 专利代理机构 北京智绘未来专利代理事务
所(普通合伙) 11689

代理人 张红莲

(51) Int.Cl.
G06Q 10/04 (2012.01)
G06K 9/62 (2006.01)
G06Q 50/06 (2012.01)

(56) 对比文件
CN 103955779 A, 2014.07.30
CN 106933778 A, 2017.07.07
CN 104820868 A, 2015.08.05
CN 107846034 A, 2018.03.27
US 2012046917 A1, 2012.02.23
US 2015154504 A1, 2015.06.04
欧阳庭辉等. 风电功率爬坡事件预测时间
窗选取建模.《中国电机工程学报》.2015,(第13
期),全文.

审查员 张桂华

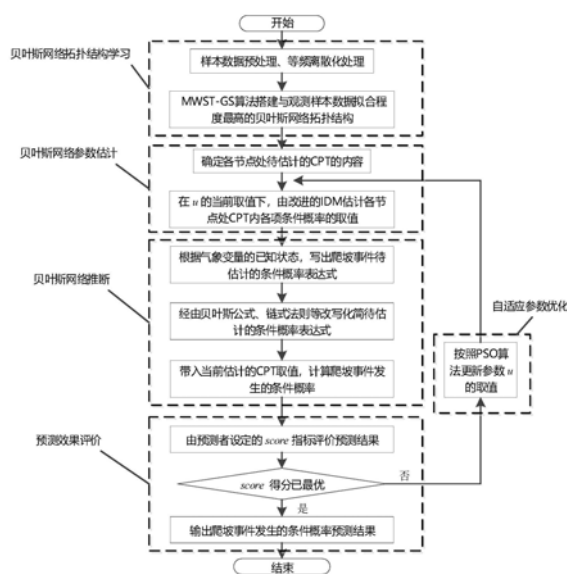
权利要求书4页 说明书14页 附图3页

(54) 发明名称

一种基于贝叶斯网络的风电爬坡事件概率
预测方法及系统

(57) 摘要

一种基于贝叶斯网络的风电爬坡事件概率
预测方法及系统,根据所掌握的观测样本数据,
挖掘风电爬坡事件与风速、风向、温度、气压、湿
度等相关气象影响因子间的相依关系,搭建与样
本数据拟合程度最高的贝叶斯网络拓扑结构;定
量描述爬坡事件与各气象因子间的条件相依关
系,估计贝叶斯网络各节点处的条件概率表内各
项条件概率的取值,与贝叶斯网络拓扑结构共同
组成风电爬坡事件预测的贝叶斯网络模型;由所
掌握的预测时刻的数值天气预报信息,推断爬坡
事件各状态发生的条件概率;自适应地调整各节
点处相应条件概率的取值,从而优化推断出的爬
坡事件各状态发生的条件概率结果,实现预测结
果可靠性与敏锐性的折中。



CN 109978222 B

1. 一种基于贝叶斯网络的风电爬坡事件概率预测方法,其特征在于,所述风电爬坡事件概率预测方法包括以下步骤:

步骤1:采集地区电网实测数据,由贝叶斯网络结构学习算法从历史观测样本数据中挖掘风电爬坡事件与风速、风向、温度、气压、湿度相关气象影响因子间的相依关系,搭建与观测样本数据拟合程度最高的贝叶斯网络拓扑结构,其中所述观测样本数据是指历史观测样本中的各气象影响因子-风电爬坡事件数据对;

步骤2:定量描述爬坡事件与各气象因子间的条件相依关系,采用改进的非精确狄利克雷模型IDM初步估计步骤1中所搭建的贝叶斯网络拓扑结构中各节点处的条件概率表CPT内各项条件概率也即贝叶斯网络的参数的取值,构建风电爬坡事件预测的贝叶斯网络模型;

其中,改进的非精确狄利克雷模型IDM估计条件概率 $P(A_1|B_1)$ 的取值范围的计算公式如下:

$$P(A_1|B_1) = \left[\frac{m_1}{M + u \times \log M}, \frac{m_1 + u \times \log M}{M + u \times \log M} \right]$$

其中,上式中的参数含义解释如下:节点变量A共有 n_a 个相互独立的取值,节点变量B共有 n_b 个相互独立的取值,通过统计观测样本数据可得,节点变量B取值 B_1 的样本个数为M,节点变量A取值 A_1, A_2, \dots, A_{n_a} 的样本个数分别为 m_1, m_2, \dots, m_{n_a} ,节点变量A取值 A_1, A_2, \dots, A_{n_a} 的概率分别为 r_1, r_2, \dots, r_{n_a} , u 为引入外生参数;

步骤3:依据步骤2所构建的贝叶斯网络模型,依次经由贝叶斯公式、链式法则、以及条件概率表达式,由所掌握的预测时刻的数值天气预报信息,推断爬坡事件各状态发生的条件概率;

步骤4:对预测的条件概率结果的可靠性与敏锐性进行评价,依据预先设定的可靠性与敏锐性赋予的权重,借助粒子群优化算法自适应地调整步骤2中初步估计的贝叶斯网络参数的取值,从而优化依步骤3所推断出的爬坡事件各状态发生的条件概率,实现预测结果可靠性与敏锐性。

2. 根据权利要求1所述的基于贝叶斯网络的风电爬坡事件概率预测方法,其特征在于:

在步骤1中,首先对地区电网实测数据进行预处理,剔除坏数据、补全缺失数据,对各实测数据实现等频离散划分,由连续量转化为状态量。

3. 根据权利要求2所述的基于贝叶斯网络的风电爬坡事件概率预测方法,其特征在于:

在步骤1中,结合最大权重生成树MWST算法与贪婪搜索GS算法搭建与观测样本数据拟合程度最高的贝叶斯网络拓扑结构。

4. 根据权利要求3所述的基于贝叶斯网络的风电爬坡事件概率预测方法,其特征在于:

在步骤1中,借助最大权重生成树算法构建与当前具体化样本数据集相匹配的树形结构,获取关于气象变量各要素与风电爬坡事件相依关系的启发性知识,借此初始化贪婪搜索算法,将初始拓扑结构限制在全局最优解的邻域内。

5. 根据权利要求1所述的基于贝叶斯网络的风电爬坡事件概率预测方法,其特征在于:

在步骤1中,具体包括以下内容:

1.1对电网实测数据进行预处理,剔除坏数据、补全缺失数据,并对风电爬坡事件的状态进行定义;

由下式对风电爬坡事件的各状态进行定义：

$$\begin{cases} -11\%P_R \leq P_t - P_{t-30\min} \leq 10\%P_R, & \text{不爬坡,} \\ P_t - P_{t-30\min} > 10\%P_R, & \text{上爬坡,} \\ P_t - P_{t-30\min} < -11\%P_R, & \text{下爬坡,} \end{cases}$$

其中, P_t 表示时刻 t 观测的风电功率, $P_{t-30\min}$ 表示前一观测时刻 $t-30\min$ 观测的风电功率, P_R 为装机容量；

1.2 经等频离散化过程分别将风速 (S)、风向 (D)、温度 (T)、湿度 (h)、风速变化量 (V) 以及爬坡事件即功率变化率 (H) 各划分为 3 个状态, 组成贝叶斯网络的六个节点变量；

1.3 借助最大权重生成树算法初步挖掘 1.2 中离散化所得的六个节点变量之间的相依关系, 并搭建与观测样本数据集相匹配的初始树形结构；

1.4 通过贪婪搜索算法对 1.3 得到的初始树形结构进行局部更新, 形成一系列候选模型, 然后由下式所示的 BIC 评分函数对一系列候选模型与观测样本数据之间的拟合程度进行逐一评价, 其中, 评价得分最优的候选模型成为“当前模型”；

$$g(G: D) = \sum_{i=1}^n \sum_{j=1}^{q_i} \sum_{k=1}^{r_i} N_{ijk} \log\left(\frac{N_{ijk}}{N_{ij}}\right) - \frac{1}{2} \log(N) \sum_{i=1}^n (r_i - 1) q_i$$

式中, n 为节点个数, 对于第 i 个节点 X_i , 假设其有 r_i 个状态 $(x_1, x_2, \dots, x_k, \dots, x_{r_i})$, 其父节点集写作 $\text{Pa}(X_i)$, 共有 q_i 种组合状态; N 为观测样本的总数, N_{ijk} 则代表观测样本中符合 $X_i = x_k$ 且 $\text{Pa}(X_i)$ 取第 j 种组合状态时的样本个数; N_{ij} 则表示符合 $\text{Pa}(X_i)$ 取第 j 种组合状态, 而 X_i 为任意取值的样本个数, 其中, 父节点是指: 对于网络拓扑结构中的两节点变量 A 与 B , 若它们之间的连接线由 A 指向 B , 则称 A 为 B 的父节点；

1.5 在后续的每一次迭代中, 均由贪婪搜索算法通过增边、减边、旋转边操作对“当前模型”局部更新得到一系列候选模型, 由 BIC 评分函数对每一候选模型进行评价；

1.6 若本次迭代中最优候选模型的得分优于“当前模型”的得分, 则将此最优候选模型作为“当前模型”, 并返回步骤 1.5 继续搜索; 否则, 停止搜索, 且“当前模型”即为由贪婪搜索算法所构建的最优贝叶斯网络拓扑结构。

6. 根据权利要求 1 所述的基于贝叶斯网络的风电爬坡事件概率预测方法, 其特征在于:

在步骤 2 中, 采用改进的非精确狄利克雷模型 IDM 来初步估计贝叶斯网络的参数, 包括以下步骤:

2.1 依据步骤 1 所构建的与观测样本数据拟合程度最高的贝叶斯网络拓扑结构, 确定各节点变量的父节点集合, 确定各节点处待估计的条件概率表的内容;

2.2 使用改进的非精确狄利克雷模型估计各节点处待估计的条件概率表内各项条件概率的取值, 从而完成贝叶斯网络参数的初步估计。

7. 根据权利要求 1 所述的基于贝叶斯网络的风电爬坡事件概率预测方法, 其特征在于:

在步骤 3 中, 按照贝叶斯公式, 结合步骤 1 构建的贝叶斯网络拓扑结构与步骤 2 估计的贝叶斯网络参数, 推导爬坡事件各状态发生的条件概率的计算公式, 具体包括以下内容:

3.1 根据气象变量的已知状态, 写出爬坡事件待估计的条件概率表达式;

3.2 经由贝叶斯公式、链式法则化简待估计的条件概率表达式;

3.3 步骤 2 估计的贝叶斯网络参数带入化简后的待估计的条件概率表达式, 计算爬坡事

件发生的条件概率。

8. 根据权利要求1所述的基于贝叶斯网络的风电爬坡事件概率预测方法,其特征在于:

在步骤4中,根据预测者对爬坡事件各状态发生的概率预测结果的可靠性和敏锐性的风险态度,由粒子群优化算法自适应地调整步骤2所估计的贝叶斯网络中各节点处各项条件概率的取值,从而优化步骤3推断出的爬坡事件各状态的发生概率,满足预测者对结果可靠性与敏锐性的要求。

9. 根据权利要求8所述的基于贝叶斯网络的风电爬坡事件概率预测方法,其特征在于:步骤4具体包括以下内容:

4.1 使用指标 $score_1$ 对预测结果的可靠性进行评价;

预测结果的可靠性由指标 $score_1$ 来评价,表征预测结果对真实概率的覆盖能力, $score_1$ 得分的计算方式如下式所示:

$$score_1 = \begin{cases} 1, & \text{if } P^*(H_w | E_t) \in [P(H_w | E_t), \bar{P}(H_w | E_t)] \\ 0, & \text{otherwise} \end{cases}$$

其中, E_t 为由预测时刻的数值天气预报信息确定的各气象变量的所属状态, $[P(H_w | E_t), \bar{P}(H_w | E_t)]$ 为预测出的爬坡事件状态 H_w 发生的概率区间, $P^*(H_w | E_t)$ 为该项爬坡状态发生的真实概率;指标 $score_1$ 得分越高,则预测结果越可靠;

4.2 使用指标 $score_2$ 对预测结果的敏锐性进行评价;

预测结果的敏锐性由指标 $score_2$ 来评价,表征预测结果的不确定性的程度, $score_2$ 得分越低,预测结果的敏锐性越高:

$$score_2 = \bar{P}(H_w | E_t) - P(H_w | E_t)$$

4.3 预测者自由指定指标 $score_1$ 与指标 $score_2$ 在综合评价指标 $score$ 中所占的权重 wt_1 与 wt_2 ,以体现个人的风险态度,对步骤3推断出的爬坡事件各状态发生概率的预测结果进行综合评价;

预测结果的综合性能由下式所示的 $score_1$ 和 $score_2$ 的加权和来评估:

$$score = wt_1 \times score_1 - wt_2 \times score_2$$

$$\text{s. t. } wt_1, wt_2 > 0, wt_1 + wt_2 = 1$$

$score$ 指标得分越高,则模型的整体预测性能越优异;

4.4 使用粒子群优化算法PSO,以最大化综合评价指标 $score$ 的得分为目标,迭代优化步骤2中初步估计的贝叶斯网络参数的取值,从而得出与预测者风险态度相匹配的爬坡事件概率预测结果。

10. 一种利用权利要求1-9任一项权利要求所述风电爬坡事件概率预测方法的风电爬坡事件概率预测系统,包括贝叶斯网络拓扑结构学习模块、贝叶斯网络参数估计模块、爬坡事件的条件概率推断模块以及模型预测效果评价与自适应参数优化模块,其特征在于:

贝叶斯网络拓扑结构学习模块,其用于构建与所掌握的观测样本数据拟合程度最高的网络拓扑结构,以智能提取并抽象表达爬坡事件与多个气象变量之间的条件相依性关系,所述贝叶斯网络拓扑结构学习模块的输出信息分别上传至贝叶斯网络参数估计模块和爬坡事件的条件概率推断模块;

所述贝叶斯网络参数估计模块读入由贝叶斯网络拓扑结构学习模块传递的变量间父

子节点关系,用于由所掌握的观测样本数据估计贝叶斯网络各节点处条件概率表内各项条件概率的取值范围,以量化各变量间的条件相依性程度;所述贝叶斯网络参数估计模块的输出信息将上传至爬坡事件的条件概率推断模块;

爬坡事件的条件概率推断模块,其用于在获取相关气象变量证据后,对爬坡事件各状态发生的条件概率区间进行推导计算;

模型预测效果评价与自适应参数优化模块,其用于对所生成的爬坡事件概率预测结果的可靠性与敏锐性进行综合评价,并依据预测者的风险态度自适应地优化预测结果。

11. 根据权利要求10所述的风电爬坡事件概率预测系统,其特征在于:

所述贝叶斯网络拓扑结构学习模块采集包含历史风电出力以及风速、风向、温度、湿度、风速变化量气象信息在内的观测样本数据,对原始数据中的连续量信息进行离散化处理,经由最大权重生成树算法与贪婪搜索算法,输出与观测样本数据拟合程度最高的贝叶斯网络拓扑结构,以表征挖掘出的变量间的相依性关系。

12. 根据权利要求11所述的风电爬坡事件概率预测系统,其特征在于:

所述贝叶斯网络拓扑结构学习模块将输出信息上传至贝叶斯网络参数估计模块,提供变量间的父子节点关系,为确定各节点处需估计的条件概率的内容提供参考。

13. 根据权利要求11或12所述的风电爬坡事件概率预测系统,其特征在于:

所述贝叶斯网络拓扑结构学习模块还将输出信息上传至爬坡事件的条件概率推断模块,提供未知变量状态取值的概率推断所需的条件独立性信息。

14. 根据权利要求10所述的风电爬坡事件概率预测系统,其特征在于:

所述贝叶斯网络参数估计模块读入由贝叶斯网络拓扑结构学习模块传递的变量间父子节点关系,并采集观测样本数据用于对待估计的变量间各项条件概率的统计估计,输出网络中各节点处的条件概率表。

15. 根据权利要求10所述的风电爬坡事件概率预测系统,其特征在于:

所述爬坡事件的条件概率推断模块读入从贝叶斯网络拓扑结构学习模块传递的变量间条件独立性信息,以及从贝叶斯网络参数估计模块传递的网络中各节点处的条件概率表内的数值信息,化简爬坡事件条件概率表达式,输出爬坡事件各状态发生的条件概率预测结果。

一种基于贝叶斯网络的风电爬坡事件概率预测方法及系统

技术领域

[0001] 本发明属于风电功率预测领域,尤其涉及一种基于贝叶斯网络的风电爬坡事件概率预测方法及系统。

背景技术

[0002] 随着风电在电力系统中渗透率的持续增长,风电出力所固有的随机性、波动性及不确定性对电网安全稳定运行、经济调度和保护控制造成的影响日益严峻。风电场有功出力在短时间内发生的大幅度变化称作风电爬坡事件。我国并网风电具有大规模、高度集中的特点,当风电穿透功率超过一定值时,无预期的风电爬坡事件便会直接导致电力系统发电不平衡,易造成系统频率越限、恶化电能质量、威胁安全运行,甚至引起切负荷或大面积停电等事故,造成重大经济损失。实现风电爬坡事件的精确量化与准确预警,对辅助调度部门优化常规机组出力,合理配置旋转备用,减缓风/光功率波动对电网的冲击具有重要的理论价值和迫切的现实意义。

[0003] 在风电爬坡事件预测技术方面,目前使用较多的方法可大致分为间接预测法和直接预测法两类。间接预测法指先借助风电功率预测技术对风电功率时间序列进行预测,再就风电功率时间序列预测结果利用爬坡事件的定义来检测风电爬坡事件的发生,并提取相应爬坡特征量的预测信息,是当前风电爬坡事件预测的主流方法。然而,经典的风电功率预测模型往往有意忽略极端气象样本,并得到相对平滑的功率时序预测曲线以最小化整体预测误差。因此,依赖于功率预测结果的间接法或许会低估爬坡事件的严重程度,或导致爬坡预警的漏报。直接预测法利用历史样本数据挖掘相应爬坡特征量与区域气象信息间的相依关系,建立由气象信息到爬坡特征量的映射,无需进行风电功率预测。该类方法更加直观、预测精度高,但其模型的训练依赖于大量历史观测样本,样本的完备性和数据的准确性直接影响了该类方法的预测性能。此外,爬坡事件的罕见性导致了观测样本的有限性,缺乏足够的训练样本或许会给确定性爬坡事件预测带来难以避免的统计误差。所以,风电爬坡事件观测样本的稀缺性严重制约了现阶段确定性风电爬坡事件预测方法的应用与发展。在此背景下,结合贝叶斯网络理论与非精确概率统计技术,实现风电爬坡事件发生概率区间的可靠估计与爬坡状态的可靠预警,为解决有限样本条件下爬坡事件的有效预测提供了新的思路。

[0004] 贝叶斯网络是一种表达不确定性知识和进行因果推理的非精确概率图模型,其将图模型与非精确概率理论结合,具有表达更一般不确定性问题和进行不确定性推理的能力。面对电力调度系统获取可靠风电出力预测信息的要求,将贝叶斯网络理论用于风电爬坡事件的预测是非常有利的,尤其在历史爬坡事件观测样本有限的条件下,贝叶斯网络统计技术可以智能挖掘爬坡事件与多气象变量间的隐含的相依关系,并以非精确条件概率的形式量化相依性程度,在获取相关气象证据的条件下实现爬坡事件发生概率的可靠推理,从而提高爬坡事件发生概率的预测精度。

发明内容

[0005] 为了解决现有风电爬坡事件预测技术受风电功率预测精度的影响较大,对训练样本非完备性和数据量测非准确性预测情景的适应性差的问题,本申请提供了一种基于贝叶斯网络的风电爬坡事件概率预测方法以及一种基于贝叶斯网络的风电爬坡事件概率预测系统。

[0006] 为实现上述发明目的,本申请具体采用以下技术方案。

[0007] 一种基于贝叶斯网络的风电爬坡事件概率预测方法,其特征在于,所述风电爬坡事件概率预测方法包括以下步骤:

[0008] 步骤1:根据所掌握的地区电网实测数据,由贝叶斯网络结构学习算法从历史观测样本数据中挖掘风电爬坡事件与风速、风向、温度、气压、湿度等相关气象影响因子间的相依关系,搭建与观测样本数据拟合程度最高的贝叶斯网络拓扑结构,其中所述观测样本数据是指历史观测样本中的各气象影响因子-风电爬坡事件数据对;

[0009] 步骤2:定量描述爬坡事件与各气象因子间的条件相依关系,从而初步估计步骤1中所搭建的贝叶斯网络拓扑结构中各节点处的条件概率表(CPT)内各项条件概率也即贝叶斯网络的参数的取值,构建风电爬坡事件预测的贝叶斯网络模型;

[0010] 步骤3:依据步骤2所构建的贝叶斯网络模型,借助贝叶斯网络推断算法,由所掌握的预测时刻的数值天气预报信息,推断爬坡事件各状态发生的条件概率;

[0011] 步骤4:对预测的条件概率结果的可靠性与敏锐性进行评价,并根据预测者对预测结果的可靠性与敏锐性的风险态度,借助粒子群优化算法自适应地调整步骤2中初步估计的贝叶斯网络参数的取值,从而优化依步骤3所推断出的爬坡事件各状态发生的条件概率,实现预测结果可靠性与敏锐性的折中。

[0012] 所述基于贝叶斯网络的风电爬坡事件概率预测方法进一步包括以下优选方案。

[0013] 在步骤1中,首先对地区电网实测数据进行预处理,剔除坏数据、补全缺失数据,对各实测数据实现等频离散划分,由连续量转化为状态量。

[0014] 在步骤1中,结合最大权重生成树(MWST)算法与贪婪搜索(GS)算法搭建与观测样本数据拟合程度最高的贝叶斯网络拓扑结构。

[0015] 其中,借助最大权重生成树算法构建与当前具体化样本数据集相匹配的树形结构,获取关于气象变量各要素与风电爬坡事件相依关系的启发性知识,借此初始化贪婪搜索算法,将初始拓扑结构限制在全局最优解的邻域内。

[0016] 在步骤1中,包括以下内容:

[0017] 1.1对电网实测数据进行预处理,剔除坏数据、补全缺失数据,并对风电爬坡事件的状态进行定义;

[0018] 由下式对风电爬坡事件各状态进行定义:

$$[0019] \begin{cases} -11\%P_R \leq P_t - P_{t-30\min} \leq 10\%P_R, & \text{不爬坡,} \\ P_t - P_{t-30\min} > 10\%P_R, & \text{上爬坡,} \\ P_t - P_{t-30\min} < -11\%P_R, & \text{下爬坡,} \end{cases} \quad (1)$$

[0020] 其中, P_t 表示时刻 t 观测的风电功率, $P_{t-30\min}$ 表示前一观测时刻 $t-30\min$ 观测的风电功率, P_R 为装机容量。

[0021] 1.2经等频离散化过程分别将风速(S)、风向(D)、温度(T)、湿度(h)、风速变化量

(V) 以及爬坡事件即功率变化率(H) 各划分为3个状态,组成贝叶斯网络的六个节点变量;

[0022] 1.3借助最大权重生成树算法初步挖掘1.2中离散化所得的六个节点变量之间的相依关系,并搭建与观测样本数据集相匹配的初始树形结构;

[0023] 1.4通过贪婪搜索算法对1.3得到的初始树形结构进行局部更新,形成一系列候选模型,然后由式(2)所示的BIC评分函数对一系列候选模型与观测样本数据之间的拟合程度进行逐一评价,其中,评价得分最优的候选模型成为“当前模型”;

$$[0024] \quad g(G:D) = \sum_{i=1}^n \sum_{j=1}^{q_i} \sum_{k=1}^{r_i} N_{ijk} \log\left(\frac{N_{ijk}}{N_{ij}}\right) - \frac{1}{2} \log(N) \sum_{i=1}^n (r_i - 1) q_i \quad (2)$$

[0025] 式(2)中,n为节点个数,对于第i个节点 X_i ,假设其有 r_i 个状态($x_1, x_2, \dots, x_k, \dots, x_{r_i}$),其父节点集写作 $\text{Pa}(X_i)$,共有 q_i 种组合状态; N 为观测样本的总数, N_{ijk} 则代表观测样本中符合 $X_i = x_k$ 且 $\text{Pa}(X_i)$ 取第j种组合状态时的样本个数; N_{ij} 则表示符合 $\text{Pa}(X_i)$ 取第j种组合状态,而 X_i 任意取值的样本个数,其中,父节点是指:对于网络拓扑结构中的两节点变量A与B,若它们之间的连接线由A指向B,则称A为B的父节点。

[0026] 1.5在后续的每一次迭代中,均由贪婪搜索算法通过增边、减边、旋转边等操作对“当前模型”局部更新得到一系列候选模型,由式(2)的BIC指标对每一候选模型进行评价;

[0027] 1.6若本次迭代中最优候选模型的得分优于“当前模型”的得分,则将此最优候选模型作为“当前模型”,并返回步骤1.5继续搜索;否则,停止搜索,且“当前模型”即为由贪婪搜索算法所构建的最优贝叶斯网络拓扑结构。

[0028] 在步骤2中,采用改进的非精确狄利克雷模型(IDM)来初步估计贝叶斯网络的参数,包括以下步骤:

[0029] 2.1依据步骤1所构建的与观测样本数据拟合程度最高的贝叶斯网络拓扑结构,确定各节点变量的父节点集合,确定各节点处待估计的条件概率表的内容;

[0030] 2.2使用改进的非精确狄利克雷模型估计各节点处待估计的条件概率表内各项条件概率的取值,从而完成贝叶斯网络参数的初步估计;

[0031] 基于改进的非精确狄利克雷模型估计条件概率 $P(A_1|B_1)$ 的取值范围的计算公式如下:

$$[0032] \quad P(A_1|B_1) = \left[\frac{m_1}{M + u \times \log M}, \frac{m_1 + u \times \log M}{M + u \times \log M} \right]$$

[0033] 其中,上式中的参数含义解释如下:节点变量A共有 n_a 个相互独立的取值,节点变量B共有 n_b 个相互独立的取值,通过统计观测样本数据可得,节点变量B取值 B_1 的样本个数为M,节点变量A取值 A_1, A_2, \dots, A_{n_a} 的样本个数分别为 m_1, m_2, \dots, m_{n_a} ,节点变量A取值 A_1, A_2, \dots, A_{n_a} 的概率分别为 r_1, r_2, \dots, r_{n_a} ,u为引入外生参数。

[0034] 在步骤3中,按照贝叶斯公式,结合步骤1构建的贝叶斯网络拓扑结构与步骤2估计的贝叶斯网络参数,推导爬坡事件各状态发生的条件概率的计算公式,具体包括以下内容:

[0035] 3.1根据气象变量的已知状态,写出爬坡事件待估计的条件概率表达式;

[0036] 3.2经由贝叶斯公式、链式法则化简待估计的条件概率表达式;

[0037] 3.3步骤2估计的贝叶斯网络参数带入化简后的待估计的条件概率表达式,计算爬坡事件发生的条件概率。

[0038] 在步骤4中,根据预测者对爬坡事件各状态发生的概率预测结果的可靠性和敏锐性的风险态度,由粒子群优化算法自适应地调整步骤2所估计的贝叶斯网络中各节点处各项条件概率的取值,从而优化步骤3推断出的爬坡事件各状态的发生概率,满足预测者对结果可靠性与敏锐性的要求。

[0039] 具体包括以下内容:

[0040] 4.1使用指标 $score_1$ 对预测结果的可靠性进行评价;

[0041] 预测结果的可靠性由指标 $score_1$ 来评价,表征预测结果对真实概率的覆盖能力, $score_1$ 得分的计算方式如式(10)所示:

$$[0042] \quad score_1 = \begin{cases} 1, & \text{if } P^*(H_w | E_l) \in [P(H_w | E_l), \bar{P}(H_w | E_l)] \\ 0, & \text{otherwise} \end{cases} \quad (10)$$

[0043] 其中, E_l 为由预测时刻的数值天气预报信息确定的各气象变量的所属状态, $[P(H_w | E_l), \bar{P}(H_w | E_l)]$ 为预测出的爬坡事件状态 H_w 发生的概率区间, $P^*(H_w | E_l)$ 为该项爬坡状态发生的真实概率;指标 $score_1$ 得分越高,则预测结果越可靠。

[0044] 4.2使用指标 $score_2$ 对预测结果的敏锐性进行评价;

[0045] 预测结果的敏锐性由指标 $score_2$ 来评价,表征预测结果的不确定性的程度, $score_2$ 得分越低,预测结果的敏锐性越高:

$$[0046] \quad score_2 = \bar{P}(H_w | E_l) - P(H_w | E_l) \quad (11)$$

[0047] 4.3预测者自由指定指标 $score_1$ 与指标 $score_2$ 在综合评价指标 $score$ 中所占的权重 wt_1 与 wt_2 ,以体现个人的风险态度,对步骤3推断出的爬坡事件各状态发生概率的预测结果进行综合评价;

[0048] 预测结果的综合性能由式(12)所示的 $score_1$ 和 $score_2$ 的加权和来评估:

$$[0049] \quad \begin{aligned} score &= wt_1 \times score_1 - wt_2 \times score_2 \\ \text{s.t.} \quad &wt_1, wt_2 > 0, \quad wt_1 + wt_2 = 1 \end{aligned} \quad (12)$$

[0050] $score$ 指标得分越高,则模型的整体预测性能越优异。

[0051] 4.4使用粒子群优化算法(PSO),以最大化综合评价指标 $score$ 的得分为目标,迭代优化步骤2中改进的非精确狄利克雷模型引进的超参数 u 的取值,从而得出与预测者风险态度相匹配的爬坡事件概率预测结果。

[0052] 在粒子群优化算法的每一步迭代中,首先在超参数 u 的当前取值下由步骤2中改进的非精确狄利克雷模型估计贝叶斯网络各节点处条件概率表内各项条件概率的取值范围;再由步骤3推导出的爬坡事件条件概率计算公式,在已知相关气象变量状态取值的条件下,推断未知的爬坡事件变量各状态发生的概率区间;接着,由步骤3.4中设计的综合评价指标 $score$ 对预测结果的可靠性与敏锐性进行综合评价;若此次迭代所得的评价指标 $score$ 的得分优于上次迭代中 $score$ 的得分,则将 u 的当前取值作为最优取值,计算参数 u 下一步的搜索方向与更新步长,跳转到下一步迭代中继续寻优;若此次迭代所得的评价指标 $score$ 的得分未能优于上次迭代中 $score$ 的得分,则停止迭代过程,输出当前 u 的取值为最优取值,且此次迭代中计算所得的爬坡事件各状态发生的条件概率区间即为优化后的预测结果。

[0053] 本申请还提供了一种利用前述风电爬坡事件概率预测方法的风电爬坡事件概率预测系统,技术方案如下。

[0054] 一种利用前述风电爬坡事件概率预测方法的风电爬坡事件概率预测系统,包括贝叶斯网络拓扑结构学习模块、贝叶斯网络参数估计模块、爬坡事件的条件概率推断模块以及模型预测效果评价与自适应参数优化模块,其特征在于:

[0055] 贝叶斯网络拓扑结构学习模块,其用于构建与所掌握的观测样本数据拟合程度最高的网络拓扑结构,以智能提取并抽象表达爬坡事件与多个气象变量之间的条件相依性关系,所述贝叶斯网络拓扑结构学习模块的输出信息分别上传至贝叶斯网络参数估计模块和爬坡事件的条件概率推断模块;

[0056] 所述贝叶斯网络参数估计模块读入由贝叶斯网络拓扑结构学习模块传递的变量间父子节点关系,用于由所掌握的观测样本数据估计贝叶斯网络各节点处条件概率表内各项条件概率的取值范围,以量化各变量间的条件相依性程度;所述贝叶斯网络参数估计模块的输出信息将上传至爬坡事件的条件概率推断模块;

[0057] 爬坡事件的条件概率推断模块,其用于在获取相关气象变量证据后,对爬坡事件各状态发生的条件概率区间进行推导计算;

[0058] 模型预测效果评价与自适应参数优化模块,其用于对所生成的爬坡事件概率预测结果的可靠性与敏锐性进行综合评价,并依据预测者的风险态度自适应地优化预测结果。

[0059] 所述风电爬坡事件概率预测系统进一步包括以下优选方案:

[0060] 所述贝叶斯网络拓扑结构学习模块采集包含历史风电出力以及风速、风向、温度、湿度、风速变化量气象信息在内的观测样本数据,对原始数据中的连续量信息进行离散化处理,经由最大权重生成树算法与贪婪搜索算法,输出与观测样本数据拟合程度最高的贝叶斯网络拓扑结构,以表征挖掘出的变量间的相依性关系。

[0061] 所述贝叶斯网络拓扑结构学习模块将输出信息上传至贝叶斯网络参数估计模块,提供变量间的父子节点关系,为确定各节点处需估计的条件概率的内容提供参考。

[0062] 所述贝叶斯网络拓扑结构学习模块还将输出信息上传至爬坡事件的条件概率推断模块,提供未知变量状态取值的概率推断所需的条件独立性信息。

[0063] 所述贝叶斯网络参数估计模块读入由贝叶斯网络拓扑结构学习模块传递的变量间父子节点关系,并采集观测样本数据用于对待估计的变量间各项条件概率的统计估计,输出网络中各节点处的条件概率表。

[0064] 所述爬坡事件的条件概率推断模块读入从贝叶斯网络拓扑结构学习模块传递的变量间条件独立性信息,以及从贝叶斯网络参数估计模块传递的网络中各节点处的条件概率表内的数值信息,化简爬坡事件条件概率表达式,输出爬坡事件各状态发生的条件概率预测结果。

[0065] 该模块需读入从爬坡事件的条件概率推断模块传递的爬坡事件条件概率预测信息,并采集观测样本数据以对爬坡事件各状态发生的真实概率进行统计,计算综合评价指标的得分,并以最大化该评分指标的取值为目标对超参数 u 进行迭代寻优。在每一步的迭代过程中,该模块输出对当前参数 u 取值下贝叶斯网络模型预测结果的评价得分,并计算下一步迭代前对参数 u 的取值所做的更改。若该模块的计算结果并没有满足参数寻优的收敛判据,则该模块输出的关于参数 u 的更新信息将回传至贝叶斯网络参数估计模块,更新所搭建的贝叶斯网络模型的参数;若收敛判据得以满足,则该模块的输出信息将不再回传,迭代寻优过程结束,超参数 u 的当前取值即为与预测者风险态度相适应的最优取值。

[0066] 与现有技术相比,本发明的有益效果是:

[0067] (1)经典的风电功率预测方法有意忽略极端样本,并获得相对平滑的功率时序预测曲线以最小化整体预测误差。因此,传统的从预测的风电功率序列中检测爬坡事件发生的预测方法可能会低估爬坡事件发生的概率。相比之下,本发明直接探索了不同气象条件下风电爬坡事件各状态的概率分布,因此可以避免在风电功率预测环节引入的累积误差。

[0068] (2)本发明将风电爬坡事件视为多状态随机变量,其状态根据爬坡幅度划分。由非精确概率质量函数表示的预测结果可以根据预测时刻的气象条件有效地估计每个可能的爬坡状态的概率区间,并量化预测中的不确定性,从而为调度运行人员提供更全面的预测信息。

[0069] (3)大多数基于场景分类的预测方法均是根据气象条件对观测样本进行划分,以找出易触发爬坡事件的极端气象条件。然而,对应于极端气象的样本量通常很少,或将导致预测结果不可靠。本发明通过提取爬坡事件与相关气象变量之间的相依关系,可以增加用于条件概率预测的有效样本容量(针对某一气象条件进行爬坡概率预测时,并不局限于此气象条件下的观测样本),使得在有限样本条件下也得以提高爬坡事件预测的可靠性。

附图说明

[0070] 构成本申请的一部分的说明书附图用来提供对本申请的进一步理解,本申请的示意性实施例及其说明用于解释本申请,并不构成对本申请的不当限定。

[0071] 图1为本申请预测理论中用于展示贝叶斯网络推理过程的简单三节点网络拓扑结构示意图;

[0072] 图2为本申请的基于贝叶斯网络的风电爬坡事件概率预测方法流程示意图;

[0073] 图3为本申请的基于MWST-GS算法的贝叶斯网络拓扑结构学习算法处理流程示意图;

[0074] 图4为本申请实例验证中由MWST算法搭建的树形结构及由GS算法获取的最优拓扑结构示意图;

[0075] 图5为本申请实例验证中在8种气象条件下预测的爬坡事件各状态的非精确条件概率结果示意图;

[0076] 图6是本申请的基于贝叶斯网络的风电爬坡事件概率预测系统的结构示意图。

具体实施方式

[0077] 应该指出,以下详细说明都是例示性的,旨在对本申请提供进一步的说明。除非另有指明,本申请中使用的所有技术和科学术语具有与本申请所属技术领域的普通技术人员通常理解的含义。

[0078] 需要注意的是,这里所使用的术语仅是为了描述具体实施方式,而非意图限制根据本申请的示例性实施方式。如在这里所使用的,除非上下文另外明确指出,否则单数形式也意图包括复数形式,此外,还应当理解的是,当在本说明书中使用术语“包含”和/或“包括”时,其指明存在特征、步骤、操作、器件、组件和/或它们的组合。

[0079] 本申请实施例对宁夏地区的某风电场进行了爬坡事件的概率预测,下面结合附图与实施例对本申请作进一步说明。

[0080] 如图2所示,基于贝叶斯网络的风电爬坡事件概率预测方法,主要包括如下步骤:

[0081] 步骤1:根据所掌握的观测样本数据,由贝叶斯网络结构学习算法挖掘风电爬坡事件与风速、风向、温度、气压、湿度等相关气象影响因子间的相依关系,搭建与样本数据拟合程度最高的贝叶斯网络拓扑结构。

[0082] 按照图3示出的由最大权重生成树算法初始化的贪婪搜索算法搭建贝叶斯网络拓扑结构的处理流程,本实施例在搭建最优贝叶斯网络拓扑结构时的主要实施方式如下:

[0083] 1.1对宁夏某风电场的量测数据(包括输出功率、风速、风向、温度、湿度)进行预处理,剔除坏数据、补全缺失数据;

[0084] 1.2经等频离散划分将各变量的取值由连续量转化为状态量;

[0085] 按照式(1)定义风电爬坡事件的各个状态。此外,5种气象变量,也即风速(S)、风向(D)、温度(T)、湿度(h)、风速变化量(V)均由等频离散化过程划分为3个状态,如表1所示。

[0086] 表1贝叶斯网络中节点变量的状态

节点变量	状态 1	状态 2	状态 3
风速变化量(V)/(m·s ⁻¹)	V ₁ :[-11.2,-0.8)	V ₂ :[-0.8,0.8)	V ₃ :[0.8,16.7]
风速(S)/(m·s ⁻¹)	S ₁ :[0,3.6)	S ₂ :[3.6,5.3)	S ₃ :[5.3,23.2]
[0087] 风向(D)/(°)	D ₁ :[1,120)	D ₂ :[120,240)	D ₃ :[240,360]
温度(T)/(°C)	T ₁ :[-21.7,5.7)	T ₂ :[5.7,16.8)	T ₃ :[16.8,36.2]
湿度(h)/(%)	h ₁ :[9.5,36.5)	h ₂ :[36.5,58.5)	h ₃ :[58.5,98.5)
爬坡事件(H)(功率变化量)/(MW)	H ₁ :[-4.0,3.6)	H ₂ :(3.6,36.0]	H ₃ :[-36.0,-4.0)

[0088] 1.3由最大权重生成树算法初步挖掘六变量间的相依关系,搭建与观测样本数据集相匹配的初始树形结构;

[0089] 在六节点的网络结构空间中搜索最优拓扑结构是一个NP困难的问题,故本申请拟采用最基本的启发式搜索算法—贪婪搜索(GS)算法以搜寻与样本数据拟合程度最高的网络拓扑结构。

[0090] 贪婪搜索算法中,初始网络结构可以是空图、随即图或利用专家知识构造的先验结构,然而,不合理的初始结构或导致搜索迭代过程复杂化甚至陷入局部最优解。故在本申请中,首先借助最大权重生成树(MWST)算法搭建与具体化样本数据集相匹配的树形结构,获取关于各变量相依关系的启发性知识,借此初始化贪婪搜索算法,将初始拓扑结构限制在全局最优解的邻域内。

[0091] 最大权重生成树算法是一种基于约束测试的结构学习算法,首先计算两两变量间的联合概率分布,使用互信息函数(式13)度量节点间的相依性程度,将此作为各节点间连接边的权重。

$$[0092] \quad I(X, Y) = \sum_{x, y} P(X = x, Y = y) \log \frac{P(X = x, Y = y)}{P(X = x) P(Y = y)} \quad (13)$$

[0093] 依各边权重从大到小排序,并从类节点出发添加权重最大的一条边,然后按照有向无环图原则逐步添加剩余边中权重最大的边,形成的n-1边生成树即是针对当前历史样本集由最大权重生成树算法构建的最大权重生成树。本申请实施例针对宁夏某风电场的观测样本数据构建的最大权重生成树结构如图4(a)所示。

[0094] 1.4应用贪婪搜索算法处理1.3得到的最大权重生成树结构,生成一系列候选模

型,并逐一评价候选模型与观测样本数据之间的拟合程度;

[0095] 贪婪搜索算法由一个初始模型出发开始搜索,每一步搜索中首先利用搜索算子(增边、减边、反转边)对当前模型进行局部更新,然后由式(2)所示的BIC评分指标评价所得的一系列候选模型,量化候选模型与观测样本数据之间的拟合程度。其中,评价得分最优的候选模型成为“当前模型”。

[0096] 1.5在贪婪搜索算法后续的每一次迭代中,均由BIC指标对本次迭代中局部更新生成的候选模型进行评价;

[0097] 1.6在每一次迭代中,若评价后选出的最优候选模型的得分优于“当前模型”的得分,则将此最优候选模型作为“当前模型”,并返回步骤1.5继续搜索;否则,停止搜索,并输出“当前模型”作为由贪婪搜索算法构建的最优贝叶斯网络拓扑结构。本申请实施例针对宁夏某风电场的观测样本数据构建的最优贝叶斯网络拓扑结构如图4(b)所示。

[0098] 步骤2:采用改进的非精确狄利克雷模型,初步估计步骤1所搭建的网络拓扑结构中各节点处的相关条件概率。

[0099] 2.1确定步骤1所搭建的贝叶斯网络拓扑结构中各节点的父节点集合,明确各节点处待估计的条件概率的内容;

[0100] 根据图4(b)所示的最优拓扑结构,可知,节点V的父节点集合为{节点H},节点S的父节点集合为{节点H与节点V},节点D的父节点集合为{节点H},节点T的父节点集合为{节点H与节点S},节点h的父节点集合为{节点D与节点T},节点H无父节点。

[0101] 由此,节点V处待估计的条件概率的内容包括: $P(V_y | H_w)$,其中, $y=1,2,3;w=1,2,3$ 。

[0102] 节点S处待估计的条件概率的内容包括: $P(S_r | H_w), P(S_r | V_y)$,其中, $y=1,2,3;r=1,2,3;w=1,2,3$ 。

[0103] 节点D处待估计的条件概率的内容包括: $P(D_q | H_w)$,其中, $q=1,2,3;w=1,2,3$ 。

[0104] 节点T处待估计的条件概率的内容包括: $P(T_k | H_w), P(T_k | S_r)$,其中, $r=1,2,3;k=1,2,3;w=1,2,3$ 。

[0105] 节点h处待估计的条件概率的内容包括: $P(h_d | D_q), P(h_d | T_k)$,其中, $q=1,2,3;k=1,2,3;d=1,2,3$ 。

[0106] 2.2使用改进的非精确狄利克雷模型估计2.1列出的待估计的条件概率的取值,完成贝叶斯网络参数的初步估计。

[0107] 在估计2.1中列出的各项条件概率时,通过对观测样本数据的统计估计,便可得出式(4)中参数 m_1 与M的相应取值,随后计算式(4)便得出2.1中各项条件概率的取值范围,也即实现了贝叶斯网络参数的初步估计。

[0108] 下面以条件概率 $P(A_1 | B_1)$ 为例,介绍非精确狄利克雷模型对此条件概率进行统计估计的理论,并详细给出本申请对非精确狄利克雷模型所做的改进。

[0109] 假设节点变量A共有 n_a 个相互独立的取值,节点变量B共有 n_b 个相互独立的取值,通过统计观测样本数据可得,节点变量B取值 B_1 的样本个数为M,节点变量A取值 A_1, A_2, \dots, A_{n_a} 的样本个数分别为 m_1, m_2, \dots, m_{n_a} ,节点变量A取值 A_1, A_2, \dots, A_{n_a} 的概率分别为 r_1, r_2, \dots, r_{n_a} ,则根据非精确狄利克雷模型估计理论,条件概率 $P(A_1 | B_1)$ 的取值范围可写作:

$$[0110] \quad P(A_1 | B_1) = \left[\frac{m_1}{s+M}, \frac{m_1+s}{s+M} \right] \quad (3)$$

[0111] 其中,参数s为非精确狄利克雷模型为解决有限样本估计的非精确性而引入的超参数,其取值大小反应了非精确狄利克雷模型对条件概率估计所引入的不确定性的程度。

[0112] 然而,本申请在非精确狄利克雷模型的实际应用中发现,对于传统经典模型所设置的常值参数s,当满足父节点取值条件要求的观测样本数量较多时(即M取值较大时),在条件概率估计中所引入的不确定性s/(M+s)衰减过快,使得步骤3中所推断的风电爬坡事件各状态发生的非精确条件概率结果的可靠性难以满足要求。因此,考虑将恒定的参数s替换为满足父节点取值条件要求的观测样本数量M的函数,即 $s=f(\log M)$,以延缓非精确条件概率估计的不确定性的衰减速度,而这一延缓作用的强度则通过引入外生参数u来控制,即令 $s=u \times \log M$ 。参数u的取值将在步骤4中对爬坡事件各状态发生概率的预测结果的评价中由粒子群优化算法自适应地优化,以实现预测结果的可靠性与敏锐性的折中。

[0113] 基于改进的非精确狄利克雷模型估计条件概率 $P(A_1 | B_1)$ 的取值范围的计算公式如下:

$$[0114] \quad P(A_1 | B_1) = \left[\frac{m_1}{M+u \times \log M}, \frac{m_1+u \times \log M}{M+u \times \log M} \right] \quad (4)$$

[0115] 凭借本申请在改进的非精确狄利克雷模型中所引入的可调参数u,在步骤4中便可依据预测者的风险态度自适应地调整步骤2中估计的各节点处条件概率表内各项条件概率的取值,从而实现在步骤3中对爬坡事件各状态发生概率的预测结果的自适应优化。

[0116] 步骤3:根据所构建的贝叶斯网络,在获取相关气象证据的条件下,实现爬坡事件各状态发生的非精确条件概率估计。

[0117] 为详细展示由贝叶斯网络拓扑结构和贝叶斯网络参数推断爬坡事件发生的条件概率的计算方法,下面将借助图1所示的简单三节点网络拓扑结构,依次经由贝叶斯公式、链式法则、以及拓扑结构中隐含的条件独立性对待估计的条件概率表达式进行变形、拆分以及化简,推导根据贝叶斯网络模型,在掌握已知变量取值的证据条件下,推断未知变量取值的算法。

[0118] 3.1根据已知变量的取值,确定已知变量的所属状态,写出未知变量与已知变量之间待估计的条件概率表达式;

[0119] 在图1所示的简单三节点网络拓扑结构中,假设A,C为两状态节点,B为三状态节点,也即 $A = \{A_i | i=1,2\}$, $B = \{B_k | k=1,2,3\}$, $C = \{C_d | d=1,2\}$ 。在已掌握的证据条件 $\{B_k, C_d\}$ 下,未知变量A取值 A_1 的条件概率可表示为 $P(A_1 | B_k, C_d)$,即为本示例部分待估计的条件概率。

[0120] 3.2根据贝叶斯公式将3.1中的条件概率表达式改写:

[0121] 根据贝叶斯公式,此条件概率表达式可写作:

$$[0122] \quad P(A_1 | B_k, C_d) = \frac{P(A_1) \cdot P(B_k, C_d | A_1)}{\sum_{i=1}^2 P(A_i) \cdot P(B_k, C_d | A_i)} \quad (5)$$

[0123] 根据链式法则将3.2中变形后的条件概率表达式拆分;

[0124] 根据链式法则,概率 $P(B_k, C_d | A_1)$ 可拆分为:

[0125] $P(B_k, C_d | A_i) = P(B_k | A_i) \cdot P(C_d | A_i, B_k)$ (6)

[0126] 因此,式(5)可改写为:

[0127]
$$P(A_1 | B_k, C_d) = \frac{P(A_1) \cdot P(B_k | A_1) \cdot P(C_d | A_1, B_k)}{\sum_{i=1}^2 P(A_i) \cdot P(B_k | A_i) \cdot P(C_d | A_i, B_k)}$$
 (7)

[0128] 按照贝叶斯网络拓扑结构中变量间的条件独立性关系将3.3中拆分后的条件概率表达式化简,完成对贝叶斯推断算法的推导;

[0129] 由图1所示贝叶斯网络拓扑结构可知,节点B的父节点为A,节点C的父节点为A,节点C的取值与节点B的取值之间不存在直接相关关系,即在已知变量A取值状态条件下,变量B与变量C之间条件独立,故式(7)中有:

[0130] $P(C_d | A_i, B_k) = P(C_d | A_i)$ (8)

[0131] 由此,拆分后的条件概率表达式可进一步化简为:

[0132]
$$P(A_1 | B_k, C_d) = \frac{P(A_1) \cdot P(B_k | A_1) \cdot P(C_d | A_1)}{\sum_{i=1}^2 P(A_i) \cdot P(B_k | A_i) \cdot P(C_d | A_i)}$$
 (9)

[0133] 3.3将步骤2估计的贝叶斯网络参数带入3.4所推导出的待估计条件概率的计算公式,由所掌握的预测时刻的数值天气预报信息,推断爬坡事件各状态发生的条件概率。

[0134] 式(9)中所含的条件概率子式 $P(A_i)$ 、 $P(B_k | A_i)$ 以及 $P(C_d | A_i)$ 的取值范围均可由步骤2中改进的非精确狄利克雷模型对观测样本数据的统计估计计算得出,即在超参数u的某一取值下,通过式(4)估计各项条件概率子式的取值范围,再由式(9)计算出待估计条件概率 $P(A_1 | B_k, C_d)$ 的上限和下限,便可得出由贝叶斯网络推断算法估计的未知变量在已知证据下取值的概率分布信息。

[0135] 在本申请实施例中,3.1写出爬坡事件与五项气象变量之间待估计的条件概率表达式;

[0136] 根据表1所确定的六变量的可取状态,预测时刻共存在243种不同的气象证据,表示为 $E_1 = \{V_y, S_r, D_q, T_k, h_d\}$, $y, r, q, k, d \in \{1, 2, 3\}$ 。则实施例中待估计的爬坡事件条件概率表达式应写作: $P(H_w | E_1)$ 。本申请所涉及的爬坡事件概率预测方法可同时量化估计中的不确定性,故预测结果为非精确概率的形式,即 $P_{im}(H_w | E_1) = [P(H_w | E_1), \bar{P}(H_w | E_1)]$ 。

[0137] 3.2根据贝叶斯公式改写3.1中的条件概率表达式:

[0138]
$$\begin{cases} \underline{P}(H_w | E_1) = \min \frac{P(H_w) \times P_{im}(V_y, S_r, D_q, T_k, h_d | H_w)}{\sum_{w=1}^3 P(H_w) \times P_{im}(V_y, S_r, D_q, T_k, h_d | H_w)} \\ \bar{P}(H_w | E_1) = \max \frac{P(H_w) \times P_{im}(V_y, S_r, D_q, T_k, h_d | H_w)}{\sum_{w=1}^3 P(H_w) \times P_{im}(V_y, S_r, D_q, T_k, h_d | H_w)} \end{cases}$$
 (14)

[0139] 3.3根据链式法则拆分3.2中变形后的条件概率表达式;

[0140] 根据链式法则,概率 $P_{im}(V_y, S_r, D_q, T_k, h_d | H_w)$ 可拆分为:

[0141]
$$\begin{aligned} & P_{im}(V_y, S_r, D_q, T_k, h_d | H_w) \\ & = P_{im}(V_y | H_w) \cdot P_{im}(S_r | H_w, V_y) \cdot P_{im}(D_q | H_w, V_y, S_r) \\ & \quad \cdot P_{im}(T_k | H_w, V_y, S_r, D_q) \cdot P_{im}(h_d | H_w, V_y, S_r, D_q, T_k) \end{aligned}$$
 (15)

[0142] 因此,式(14)可进一步改写为:

$$\begin{cases}
 \underline{P}(H_w | E_l) = \min \frac{P(H_w) \times P_{im}(V_y | H_w) \cdot P_{im}(S_r | H_w, V_y) \cdot P_{im}(D_q | H_w, V_y, S_r) \cdot P_{im}(T_k | H_w, V_y, S_r, D_q) \cdot P_{im}(h_d | H_w, V_y, S_r, D_q, T_k)}{\sum_{w=1}^3 P(H_w) \times P_{im}(V_y | H_w) \cdot P_{im}(S_r | H_w, V_y) \cdot P_{im}(D_q | H_w, V_y, S_r) \cdot P_{im}(T_k | H_w, V_y, S_r, D_q) \cdot P_{im}(h_d | H_w, V_y, S_r, D_q, T_k)} \\
 \overline{P}(H_w | E_l) = \max \frac{P(H_w) \times P_{im}(V_y | H_w) \cdot P_{im}(S_r | H_w, V_y) \cdot P_{im}(D_q | H_w, V_y, S_r) \cdot P_{im}(T_k | H_w, V_y, S_r, D_q) \cdot P_{im}(h_d | H_w, V_y, S_r, D_q, T_k)}{\sum_{w=1}^3 P(H_w) \times P_{im}(V_y | H_w) \cdot P_{im}(S_r | H_w, V_y) \cdot P_{im}(D_q | H_w, V_y, S_r) \cdot P_{im}(T_k | H_w, V_y, S_r, D_q) \cdot P_{im}(h_d | H_w, V_y, S_r, D_q, T_k)}
 \end{cases} \quad (16)$$

[0144] 3.4根据2.1中列出的六节点之间的父子节点关系,易知六变量间的条件独立性关系,从而待估计的爬坡事件条件概率取值的上下界 $\overline{P}(H_w | E_l)$ 与 $\underline{P}(H_w | E_l)$ 可简化为:

$$\begin{cases}
 \underline{P}(H_w | E_l) = \min \frac{P(H_w) \times P_{im}(V_y | H_w) \times P_{im}(S_r | H_w, V_y) \times P_{im}(D_q | H_w) \times P_{im}(T_k | H_w, S_r)}{\sum_{w=1}^3 P(H_w) \times P_{im}(V_y | H_w) \times P_{im}(S_r | H_w, V_y) \times P_{im}(D_q | H_w) \times P_{im}(T_k | H_w, S_r)} \\
 \overline{P}(H_w | E_l) = \max \frac{P(H_w) \times P_{im}(V_y | H_w) \times P_{im}(S_r | H_w, V_y) \times P_{im}(D_q | H_w) \times P_{im}(T_k | H_w, S_r)}{\sum_{w=1}^3 P(H_w) \times P_{im}(V_y | H_w) \times P_{im}(S_r | H_w, V_y) \times P_{im}(D_q | H_w) \times P_{im}(T_k | H_w, S_r)}
 \end{cases} \quad (17)$$

[0146] 在式(17)中,先验概率 $P(H_w)$ 可由状态 H_w 在观测样本数据中出现的频率来估计,其他所涉及的非精确条件概率均为直接从2.2所构建的条件概率表中读取。

[0147] 步骤4:评估模型的预测效果,并将模型预测性能与对照模型作比较。依据预测者为预测结果的可靠性与敏锐性赋予的权重自适应地优化2.2中超参数 u 的取值。

[0148] 本发明实施例所选取的对照模型为基于中心极限定理的非精确概率估计模型。

[0149] 中心极限定理(CLT)是估计统计样本数据平均分布的常用方法。假设 μ 和 σ^2 分别代表样本数据集的均值和方差。当由 M 所表示的数据集样本量足够大时,根据CLT理论,样本均值可以近似通过正态分布 $N(\mu, \sigma^2/M)$ 来拟合。从而,在设定的置信水平下,可进一步获取均值的置信区间。

[0150] 表2分析了在243种气象证据条件下基于CLT模型的爬坡事件预测结果和基于贝叶斯网络(BN)模型的爬坡事件预测结果的平均预测性能(CLT模型的置信度选为90%)。在此表中,综合评价指标score中的权重 wt_1 被依次选作0.3,0.5,0.7以体现本申请的预测方法对预测者的不同风险态度的普适性。

[0151] 表2预测结果分析

指标	$wt_1 = 0.3$		$wt_1 = 0.5$		$wt_1 = 0.7$	
	BN	CLT	BN	CLT	BN	CLT
$score_1$	570	410	608	410	646	410
$score_2$	88.7	107.1	105.2	107.1	133.2	107.1
$score$	108.9	48.0	251.4	151.4	412.2	254.9
平均覆盖率	78.2%	56.2%	83.4%	56.2%	88.6%	56.2%
平均区间宽度	0.122	0.147	0.144	0.147	0.183	0.147
区间宽度 ≥ 0.1	60.5%	54.3%	70.3%	54.3%	84.0%	54.3%
区间宽度 ≥ 0.2	8.2%	32.0%	16.8%	32.0%	38.3%	32.0%
区间宽度 ≥ 0.3	0.8%	16.1%	2.4%	16.1%	8.2%	16.1%

[0152] 风险寻求者或许会指定 $wt_1=0.3, wt_2=0.7$ 以反映他们对预测结果敏锐性的关注。在这种情况下, BN模型中所应用的PSO参数优化方法为超参数 u 的取值寻得的最优结果为 $u=11.7$, BN模型预测的概率区间的平均宽度仅为0.122。超过90%的预测概率区间窄于0.2, 只有极少数的区间宽于0.3, 这比CLT模型的预测结果要好得多。此外, 在此项测试中, 78.2%的预测概率区间可以覆盖真实概率, 这一比例高于CLT模型的结果, 体现了BN模型在预测可靠性方面的优异性能。

[0154] 相反, 风险规避者或许会选择 $wt_1=0.7, wt_2=0.3$ 来反映他们对预测结果可靠性的关注。在这种情况下, BN模型中所应用的PSO参数优化方法为超参数 u 的取值寻得的最优结果为 $u=18.2$, BN模型预测结果所实现的覆盖率接近90%, 体现了BN模型在可靠性方面卓越的预测性能。然而, 作为达到较高可靠性的代价, 此情况下预测区间的平均宽度显著增加, 当 wt_1 从0.3增加到0.7时, 预测区间的平均宽度增加了50%。

[0155] 有趣的是, 当 $wt_1=wt_2=0.5$ 时, BN模型的预测结果与CLT模型的预测结果恰巧实现了几乎相同的平均区间宽度。在这种情况下, BN模型中所应用的PSO参数优化方法为超参数 u 的取值寻得的最优结果为 $u=13.5$, BN模型预测结果的更高的覆盖率清晰地反映了其在可靠性与敏锐性的综合评价方面相对CLT模型的优势。

[0156] 总的来说, 从表2展示的BN模型和CLT模型的测试结果中可以得出以下结论:

[0157] 1. 在测试中, 综合评价指标 $score$ 中的权重 wt_1 和 wt_2 反映了预测者个人的风险态度。权重 wt_1 越大, 所得的预测结果的可靠性越强, 同时, 作为高可靠性的代价, 预测的概率区间将相对较宽。

[0158] 2. 对于不同的风险态度, 可以相应地优化预测模型中的超参数 u , 以使模型获取更好的预测性能。从表2的预测结果分析中可以清楚地观察到, 本申请的爬坡事件预测模型预测的概率区间可以根据预测者个人的风险态度进行自适应地调整, 体现了该模型的灵活性。

[0159] 3. 从表2的分析中可见, 无论预测者对可靠性与敏锐性的风险态度如何, 本申请的风电爬坡事件预测模型总是表现出比CLT模型更优异的预测性能。

[0160] 图5形象展示了当 $wt_1=wt_2=0.5$ 时, 在8种不同的气象条件下, 本申请的BN模型和作为对照的CLT模型对爬坡事件各状态发生的非精确概率预测结果。

[0161] 在图5中可以观察到, 在气象条件 E_1, E_2 和 E_3 下, 通过CLT模型获得的预测区间过宽, 虽覆盖住了真实概率, 但过宽的预测区间也模糊了所揭示的统计规律, 弱化了预测的有效性。在这三种气象条件下, BN模型在保证覆盖真实概率的基础上, 总是可以获得更狭窄的预

测区间,体现了本申请的爬坡事件概率预测方法优异的敏锐性。

[0162] 在气象条件 E_4 和 E_5 下,由于观测样本数量有限,由训练集和验证集所统计的真实概率分布存在很大差异。在这两种预测情景中,CLT模型就样本充裕量方面的应用要求难以满足。在图5中可以观察到,CLT模型估计的概率区间并不能覆盖住真实概率,存在估计偏差。然而,本申请所提出的BN模型在这两种气象条件下均实现了对真实概率的覆盖,体现了该方法在有限样本条件下对揭示真实统计概率的优异性能。

[0163] 此外,对于气象条件 E_6 和 E_7 ,训练集中均无爬坡事件发生,因此,依据CLT模型的预测理论仅可得到单值预测结果 $P(H_1|E_1) = 1$ 且 $P(H_2|E_1) = P(H_3|E_1) = 0$,带来无法容忍的预测误差。与此相反,在图5中可以观察到,本申请所提出的BN模型在这两种气象条件下仍可保证良好的预测性能,实现了对统计真实概率的覆盖。

[0164] 在最不利的预测条件 E_8 下,由于严重缺乏观测样本,训练集和验证集所统计的真实概率之间的偏差十分显著,使得BN模型和CLT模型在可靠性方面的预测性能均被劣化。然而,从图5中可以看出,BN模型预测的概率区间相对验证集统计出的真实概率的偏差更小些,体现了其与CLT模型相比,在这种非常不利的预测情景下也能更准确地揭示潜在分布规律。

[0165] 图6是本申请的基于贝叶斯网络的风电爬坡事件概率预测系统的结构示意图。

[0166] 如图6所示,本申请的一种基于贝叶斯网络的风电爬坡事件概率预测系统,包括:

[0167] (1) 贝叶斯网络拓扑结构学习模块,其用于构建与所掌握的观测样本数据拟合程度最高的网络拓扑结构,以智能提取并抽象表达爬坡事件与多个气象变量之间的条件相依性关系。

[0168] 该模块需采集包含历史风电出力以及风速、风向、温度、湿度、风速变化量等气象信息在内的观测样本数据,对原始数据中的连续量信息进行离散化处理,经由最大权重生成树算法与贪婪搜索算法,输出与观测样本数据拟合程度最高的贝叶斯网络拓扑结构以表征挖掘出的变量间的相依性关系。该模块的输出信息将上传至贝叶斯网络参数估计模块,提供变量间的父子节点关系,为确定各节点处需估计的条件概率的内容提供参考;该模块的输出信息还将上传至爬坡事件的条件概率推断模块,提供未知变量状态取值的概率推断所需的条件独立性信息。(2) 贝叶斯网络参数估计模块,其用于由所掌握的观测样本数据估计贝叶斯网络中各节点处条件概率表内各项条件概率的取值范围,以量化各变量间的条件相依性程度。

[0169] 该模块需读入由贝叶斯网络拓扑结构学习模块传递的变量间父子节点关系,并采集观测样本数据用于模块内改进的非精确狄利克雷模型对待估计的变量间各项条件概率的统计估计,输出网络中各节点处的条件概率表。该模块的输出信息将上传至爬坡事件的条件概率推断模块,提供爬坡事件发生的条件概率的计算过程中所需的相关数量信息。

[0170] (3) 爬坡事件的条件概率推断模块,其用于在获取相关气象变量证据后,对爬坡事件各状态发生的条件概率区间进行推导计算。

[0171] 该模块需读入从贝叶斯网络拓扑结构学习模块传递的变量间条件独立性信息,以及从贝叶斯网络参数估计模块传递的网络中各节点处的条件概率表内的数值信息,经由贝叶斯公式、链式法则等对待估计的爬坡事件条件概率表达式进行改写、拆分与化简,输出爬坡事件各状态发生的条件概率预测结果。该模块的输出信息将上传至模型预测效果评价与

自适应参数优化模块,提供在超参数 u 的当前取值下,依所搭建的贝叶斯网络模型对爬坡事件发生的概率区间的预测结果。

[0172] (4) 模型预测效果评价与自适应参数优化模块,其用于对所生成的爬坡事件概率预测结果的可靠性与敏锐性进行综合评价,并依据预测者的风险态度自适应地优化预测结果。

[0173] 该模块需读入从爬坡事件的条件概率推断模块传递的爬坡事件条件概率预测信息,并采集观测样本数据以对爬坡事件各状态发生的真实概率进行统计,计算综合评价指标的得分,并以最大化该评分指标的取值为目标对超参数 u 进行迭代选优。在每一步的迭代过程中,该模块输出对当前参数 u 取值下贝叶斯网络模型预测结果的评价得分,并计算下一步迭代前对参数 u 的取值所做的更改。若该模块的计算结果并没有达到参数寻优的收敛判据,则该模块输出的关于参数 u 的更新信息将回传至贝叶斯网络参数估计模块,更新所搭建的贝叶斯网络模型的参数;若收敛判据得以满足,则该模块的输出信息将不再回传,迭代寻优过程结束。

[0174] 上述虽然结合附图对本发明的具体实施方式进行了描述,但并非对本发明保护范围的限制,所属领域技术人员应该明白,在本发明的技术方案的基础上,本领域技术人员不需要付出创造性劳动即可做出的各种修改或变形仍在本发明的保护范围以内。

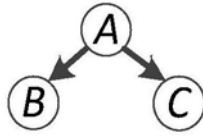


图1

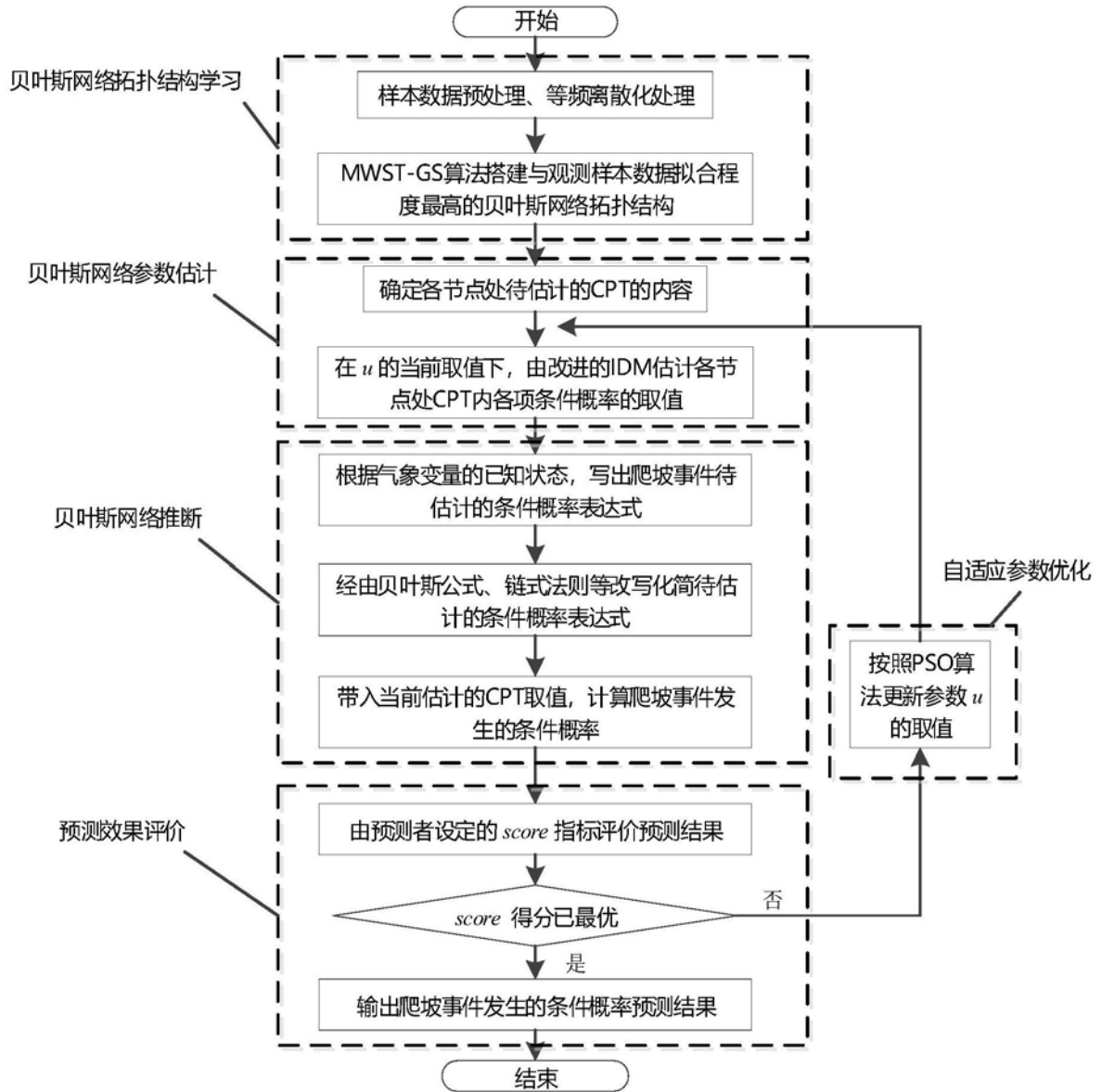


图2

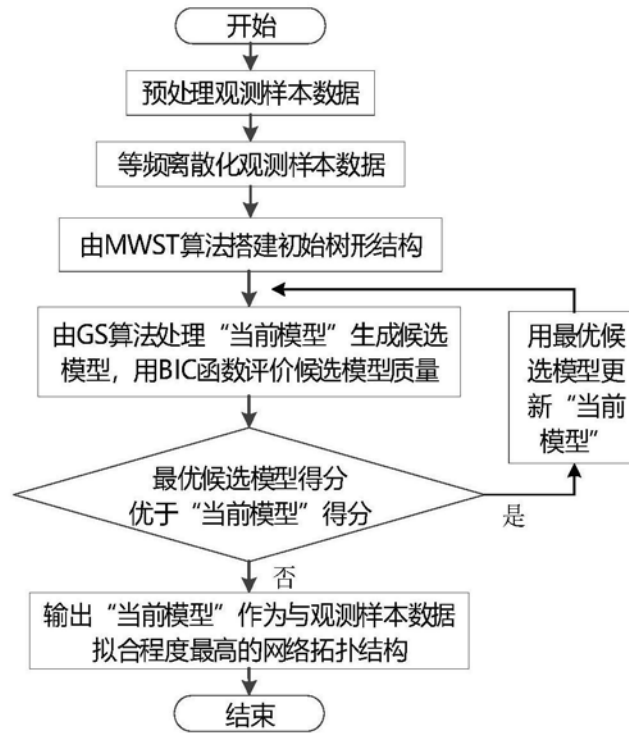


图3

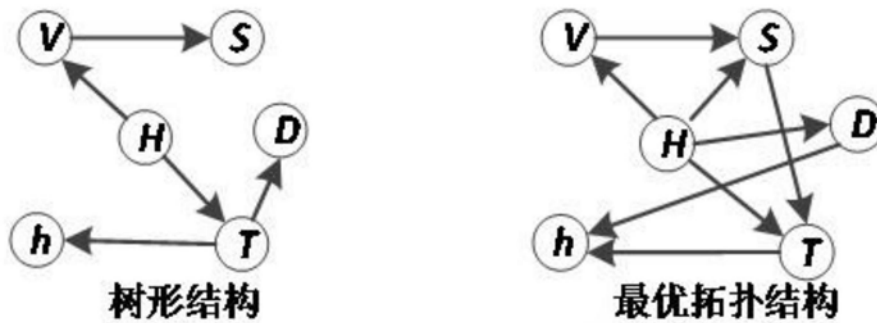


图 4a

图 4b

图4

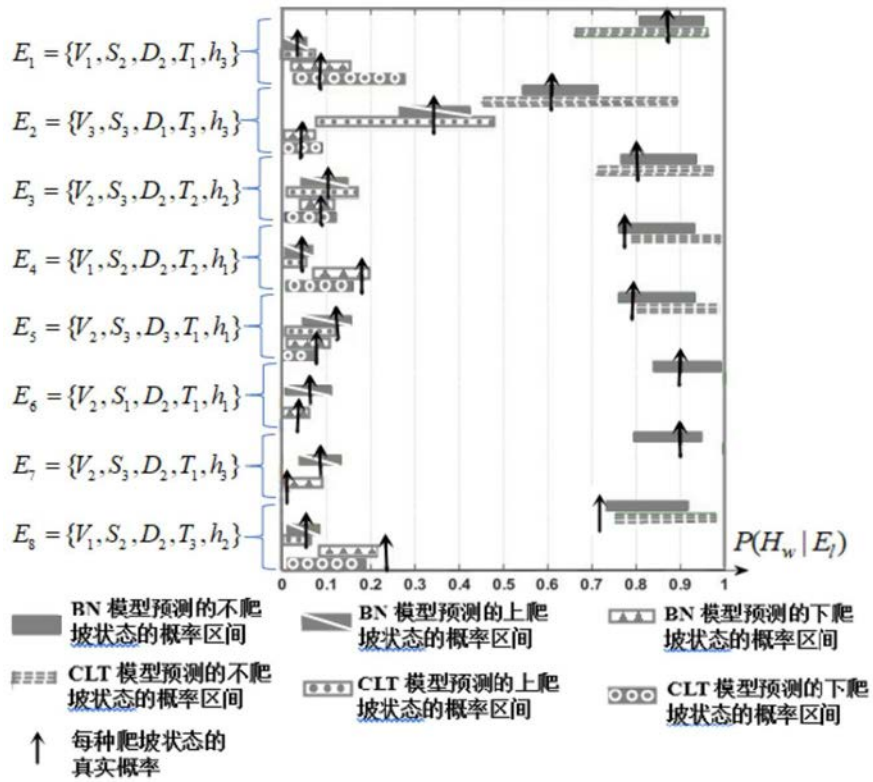


图5

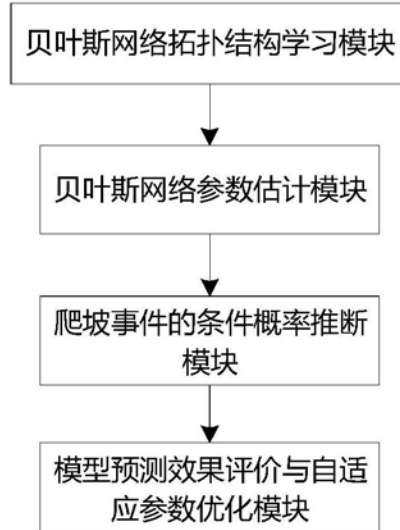


图6