



(12)发明专利申请

(10)申请公布号 CN 109446474 A

(43)申请公布日 2019.03.08

(21)申请号 201811029062.X

(51)Int.Cl.

(22)申请日 2018.09.05

G06F 17/16(2006.01)

G06F 17/18(2006.01)

(71)申请人 中国电力科学研究院有限公司

G06F 17/24(2006.01)

G06K 9/62(2006.01)

地址 100192 北京市海淀区清河小营东路
15号

申请人 国家电网有限公司

中国科学院沈阳自动化研究所

(72)发明人 张立春 刘彬 程永锋 姬昆鹏

汉京善 费香泽 唐凤珍 斯白露

司佳钧 杨加伦 展雪萍 赵彬

李鹏 李丹煜

(74)专利代理机构 北京安博达知识产权代理有
限公司 11271

代理人 徐国文

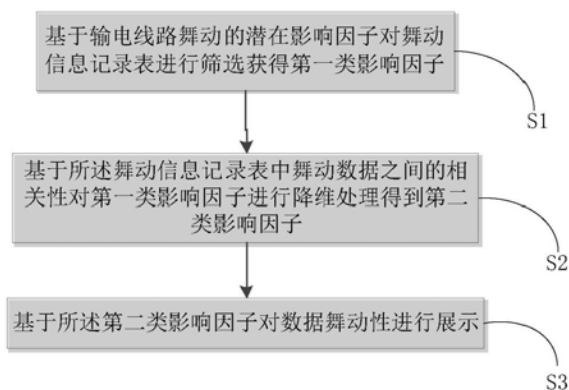
权利要求书3页 说明书12页 附图2页

(54)发明名称

一种数据舞动性的展示方法及系统

(57)摘要

一种数据舞动性的展示方法及系统,包括:基于输电线路舞动的潜在影响因子对舞动信息记录表进行筛选获得第一类影响因子;基于所述舞动信息记录表中舞动数据之间的相关性对第一类影响因子进行降维处理得到第二类影响因子;基于所述第二类影响因子对数据舞动性进行展示。本发明提供的技术方案可以直观地展示数据正负样本的分布情况,便于了解舞动数据特点,为通过舞动数据预测输电线路舞动提供了依据。



1. 一种数据舞动性的展示方法,其特征在于,包括:

基于输电线路舞动的潜在影响因子对舞动信息记录表进行筛选获得第一类影响因子;

基于所述舞动信息记录表中舞动数据之间的相关性对第一类影响因子进行降维处理得到第二类影响因子;

基于所述第二类影响因子对数据舞动性进行展示。

2. 如权利要求1所述的展示方法,其特征在于,所述基于所述舞动信息记录表中舞动数据之间的相关性对第一类影响因子进行降维处理得到第二类影响因子,包括:

从所述舞动信息记录表中选择第一影响因子及所述第一影响因子涉及的输电线路相关数据,生成分析数据表;

基于所述分析数据表利用主成分分析法对第一类影响因子进行降维处理得到第二类影响因子。

3. 如权利要求2所述的展示方法,其特征在于,所述基于所述分析数据表利用主成分分析法对第一类影响因子进行降维处理得到第二类影响因子,包括:

基于所述分析数据表中舞动数据之间的相关系数构建相关系数矩阵;

基于雅可比法得到所述相关系数矩阵的特征值和特征向量;

基于所述特征值和第一类影响因子的数量,计算所述第一类影响因子的累计贡献率;

基于所述累计贡献率和预设范围的关系,选择满足条件的所述累计贡献率对应的特征值;

基于所述对应的特征值和所述特征向量计算主成分载荷;

基于所述主成分载荷和第一类影响因子按设置的公式计算得到第二类影响因子。

4. 如权利要求3所述的展示方法,其特征在于,所述基于所述分析数据表中舞动数据之间的相关系数构建相关系数矩阵,如下式所示:

$$R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1j} & \cdots & r_{1p} \\ r_{21} & r_{22} & \cdots & r_{2j} & \cdots & r_{2p} \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ r_{i1} & r_{i2} & \cdots & r_{ij} & \cdots & r_{ip} \\ \vdots & \vdots & & \vdots & \ddots & \vdots \\ r_{p1} & r_{p2} & \cdots & r_{pj} & \cdots & r_{pp} \end{bmatrix}$$

式中:R:相关系数矩阵; r_{ij} :第一类影响因子中第i个影响因子 x_i 与第j个影响因子 x_j 的相关系数;p:第一类影响因子的数量。

5. 如权利要求4所述的展示方法,其特征在于,所述第一类影响因子中第i个影响因子 x_i 与第j个影响因子 x_j 的相关系数 r_{ij} ,按下式计算:

$$r_{ij} = \frac{\sum_{k=1}^n (x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j)}{\sqrt{\sum_{k=1}^n (x_{ki} - \bar{x}_i)^2 \sum_{k=1}^n (x_{kj} - \bar{x}_j)^2}}$$

式中:n:样本个数; x_{ki} :第k个样本中的第一类影响因子中第i个影响因子 x_i ; \bar{x}_i :第一类影响因子中第i个影响因子 x_i 的均值; x_{kj} :第k个样本中的第一类影响因子中第j个影响因子

x_j ; \bar{x}_j : 第一类影响因子中第j个影响因子 x_j 的均值;

其中, 所述第一类影响因子中第i个影响因子 x_i 的均值 \bar{x}_i 和所述第一类影响因子中第j个影响因子 x_j 的均值 \bar{x}_j , 按下式计算:

$$\begin{cases} \bar{x}_i = \frac{\sum_{k=1}^n x_{ki}}{n} \\ \bar{x}_j = \frac{\sum_{k=1}^n x_{kj}}{n} \end{cases}。$$

6. 如权利要求3所述的展示方法, 其特征在于, 所述基于所述特征值和第一类影响因子的数量, 计算所述第一类影响因子的累计贡献率, 按下式计算:

$$F_m = \frac{\sum_{i=1}^m \lambda_i}{\sum_{i=1}^p \lambda_i} \quad (i=1, 2, \dots, p, m < p)$$

式中: F_m : 第一类影响因子中m个影响因子的累计贡献率; λ_i : 第i个影响因子的特征值; p: 第一类影响因子的数量。

7. 如权利要求6所述的展示方法, 其特征在于, 所述基于所述累计贡献率和预设范围的关系, 选择满足条件的所述累计贡献率对应的特征值, 包括:

当所述第一类影响因子中m个影响因子的累计贡献率 F_m 有且只有一个在预设范围内时, 选择所述累计贡献率 F_m 对应的特征值;

当所述第一类影响因子中m个影响因子的累计贡献率 F_m 有且至少有两个在预设范围内时, 选择最小的m对应的累计贡献率 F_m , 并获取所述累计贡献率 F_m 对应的特征值。

8. 如权利要求3所述的展示方法, 其特征在于, 所述基于所述对应的特征值和所述特征向量计算主成分载荷, 按下式计算:

$$l_{ij} = \sqrt{\lambda_i} e_{ij} \quad (i=1, 2, \dots, m, j=1, 2, \dots, p, \text{且} m < p)$$

式中: l_{ij} : 第一类影响因子中第i个影响因子与第j个影响因子的主成分载荷; λ_i : 第i个影响因子的特征值; e_{ij} : 第一类影响因子中第i个影响因子与第j个影响因子的特征向量; p: 第一类影响因子的数量。

9. 如权利要求3所述的展示方法, 其特征在于, 所述基于所述主成分载荷和第一类影响因子按设置的公式计算得到第二类影响因子, 按下式计算:

$$Z = \begin{cases} z_1 = l_{11}x_1 + l_{12}x_2 + \dots + l_{1p}x_p \\ z_2 = l_{21}x_1 + l_{22}x_2 + \dots + l_{2p}x_p \\ \dots \\ z_m = l_{m1}x_1 + l_{m2}x_2 + \dots + l_{mp}x_p \end{cases} \quad (m < p)$$

式中: Z: 第二类影响因子; z_m : 第二类影响因子中的第m个主成分; l_{ij} : 第一类影响因子中第i个影响因子与第j个影响因子的主成分载荷; x_p : 第一类影响因子中的第p个影响因子。

10. 如权利要求1所述的展示方法, 其特征在于, 所述基于输电线路舞动的潜在影响因

子对舞动信息记录表进行筛选获得第一类影响因子,包括:

基于输电线路舞动的潜在影响因子获取所述潜在影响因子涉及的输电线路相关数据;

基于潜在影响因子和所述潜在影响因子涉及的输电线路相关数据构建舞动信息记录表;

从所述舞动信息记录表筛选出待分析的影响因子组成第一类影响因子;

其中,所述输电线路舞动的潜在影响因子,包括:风速、风向、线路走向与风向夹角、温度、湿度、降水形式、导线覆冰厚度、地形、海拔高度、电压等级、杆塔回路数、绞线截面积、导线排列方式和导线分裂数。

11. 如权利要求10所述的展示方法,其特征在于,所述基于输电线路舞动的潜在影响因子对舞动信息记录表进行筛选获得第一类影响因子,还包括:

将第一类影响因子中通过文字描述的影响因子按设置的转换原则转化为数值形式。

12. 如权利要求3所述的展示方法,其特征在于,所述基于所述第二类影响因子对数据舞动性进行展示,包括:

从第二类影响因子中选择前两个主成分,并将分析数据表中所述主成分对应的输电线路相关数据绘制在直角坐标系中,通过直角坐标系展示数据舞动性。

13. 一种数据舞动性的展示系统,其特征在于,包括:

筛选模块,用于基于输电线路舞动的潜在影响因子对舞动信息记录表进行筛选获得第一类影响因子;

降维模块,用于基于所述舞动信息记录表中舞动数据之间的相关性对第一类影响因子进行降维处理得到第二类影响因子;

展示模块,用于基于所述第二类影响因子对数据舞动性进行展示。

14. 如权利要求13所述的展示系统,其特征在于,所述降维模块,包括:

生成子模块,用于从所述舞动信息记录表中选择第一影响因子及所述第一影响因子涉及的输电线路相关数据,生成分析数据表;

降维子模块,用于基于所述分析数据表利用主成分分析法对第一类影响因子进行降维处理得到第二类影响因子。

一种数据舞动性的展示方法及系统

技术领域

[0001] 本发明涉及输电线路防灾减灾领域,具体涉及一种数据舞动性的展示方法及系统。

背景技术

[0002] 输电线路舞动事故是威胁电网安全稳定运行的重要灾害形式之一,近年来,随着人工智能技术的快速发展,其在电网防灾减灾领域的应用日渐深入,人们将人工智能技术引入输电线路舞动预测预警工作中,利用积累的大量舞动相关数据,对舞动事故的发生可能性进行预测预警,从而使得线路运维部门能够提前预判灾害的发生,制定应对措施,降低舞动事故对电网产生的损害。然而,舞动相关数据形式多样,如何对海量复杂的数据中进行直观展示,便于直观的反应舞动研究中最具价值的信息,需要进行深入研究。

发明内容

[0003] 为了解决现有技术中所存在的舞动预测准确率低的问题,本发明提供一种数据舞动性的展示方法及系统。

[0004] 本发明提供的技术方案是:一种数据舞动性的展示方法,包括:

[0005] 基于输电线路舞动的潜在影响因子对舞动信息记录表进行筛选获得第一类影响因子;

[0006] 基于所述舞动信息记录表中舞动数据之间的相关性对第一类影响因子进行降维处理得到第二类影响因子;

[0007] 基于所述第二类影响因子对数据舞动性进行展示。

[0008] 优选的,所述基于所述舞动信息记录表中舞动数据之间的相关性对第一类影响因子进行降维处理得到第二类影响因子,包括:

[0009] 从所述舞动信息记录表中选择第一影响因子及所述第一影响因子涉及的输电线路相关数据,生成分析数据表;

[0010] 基于所述分析数据表利用主成分分析法对第一类影响因子进行降维处理得到第二类影响因子。

[0011] 优选的,所述基于所述分析数据表利用主成分分析法对第一类影响因子进行降维处理得到第二类影响因子,包括:

[0012] 基于所述分析数据表中舞动数据之间的相关系数构建相关系数矩阵;

[0013] 基于雅可比法得到所述相关系数矩阵的特征值和特征向量;

[0014] 基于所述特征值和第一类影响因子的数量,计算所述第一类影响因子的累计贡献率;

[0015] 基于所述累计贡献率和预设范围的关系,选择满足条件的所述累计贡献率对应的特征值;

[0016] 基于所述对应的特征值和所述特征向量计算主成分载荷;

[0017] 基于所述主成分载荷和第一类影响因子按设置的公式计算得到第二类影响因子。

[0018] 优选的,所述基于所述分析数据表中舞动数据之间的相关系数构建相关系数矩阵,如下式所示:

$$[0019] \quad R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1j} & \cdots & r_{1p} \\ r_{21} & r_{22} & \cdots & r_{2j} & \cdots & r_{2p} \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ r_{i1} & r_{i2} & \cdots & r_{ij} & \cdots & r_{ip} \\ \vdots & \vdots & & \vdots & \ddots & \vdots \\ r_{p1} & r_{p2} & \cdots & r_{pj} & \cdots & r_{pp} \end{bmatrix}$$

[0020] 式中:R:相关系数矩阵; r_{ij} :第一类影响因子中第i个影响因子 x_i 与第j个影响因子 x_j 的相关系数;p:第一类影响因子的数量。

[0021] 优选的,所述第一类影响因子中第i个影响因子 x_i 与第j个影响因子 x_j 的相关系数 r_{ij} ,按下式计算:

$$[0022] \quad r_{ij} = \frac{\sum_{k=1}^n (x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j)}{\sqrt{\sum_{k=1}^n (x_{ki} - \bar{x}_i)^2 \sum_{k=1}^n (x_{kj} - \bar{x}_j)^2}}$$

[0023] 式中:n:样本个数; x_{ki} :第k个样本中的第一类影响因子中第i个影响因子 x_i ; \bar{x}_i :第一类影响因子中第i个影响因子 x_i 的均值; x_{kj} :第k个样本中的第一类影响因子中第j个影响因子 x_j ; \bar{x}_j :第一类影响因子中第j个影响因子 x_j 的均值;

[0024] 其中,所述第一类影响因子中第i个影响因子 x_i 的均值 \bar{x}_i 和所述第一类影响因子中第j个影响因子 x_j 的均值 \bar{x}_j ,按下式计算:

$$[0025] \quad \begin{cases} \bar{x}_i = \frac{\sum_{k=1}^n x_{ki}}{n} \\ \bar{x}_j = \frac{\sum_{k=1}^n x_{kj}}{n} \end{cases} .$$

[0026] 优选的,所述基于所述特征值和第一类影响因子的数量,计算所述第一类影响因子的累计贡献率,按下式计算:

$$[0027] \quad F_m = \frac{\sum_{i=1}^m \lambda_i}{\sum_{i=1}^p \lambda_i} \quad (i=1,2,\dots,p, m < p)$$

[0028] 式中: F_m :第一类影响因子中m个影响因子的累计贡献率; λ_i :第i个影响因子的特征值;p:第一类影响因子的数量。

[0029] 优选的,所述基于所述累计贡献率和预设范围的关系,选择满足条件的所述累计贡献率对应的特征值,包括:

[0030] 当所述第一类影响因子中m个影响因子的累计贡献率 F_m 有且只有一个在预设范围

内时,选择所述累计贡献率 F_m 对应的特征值;

[0031] 当所述第一类影响因子中 m 个影响因子的累计贡献率 F_m 有且至少有两个在预设范围内时,选择最小的 m 对应的累计贡献率 F_m ,并获取所述累计贡献率 F_m 对应的特征值。

[0032] 优选的,所述基于所述对应的特征值和所述特征向量计算主成分载荷,按下式计算:

$$[0033] \quad l_{ij} = \sqrt{\lambda_i} e_{ij} (i=1,2\dots m, j=1,2\dots p, \text{且 } m < p)$$

[0034] 式中: l_{ij} :第一类影响因子中第 i 个影响因子与第 j 个影响因子的主成分载荷; λ_i :第 i 个影响因子的特征值; e_{ij} :第一类影响因子中第 i 个影响因子与第 j 个影响因子的特征向量; p :第一类影响因子的数量。

[0035] 优选的,所述基于所述主成分载荷和第一类影响因子按设置的公式计算得到第二类影响因子,按下式计算:

$$[0036] \quad Z = \begin{cases} z_1 = l_{11}x_1 + l_{12}x_2 + \dots + l_{1p}x_p \\ z_2 = l_{21}x_1 + l_{22}x_2 + \dots + l_{2p}x_p \\ \dots \\ z_m = l_{m1}x_1 + l_{m2}x_2 + \dots + l_{mp}x_p \end{cases} (m < p)$$

[0037] 式中: Z :第二类影响因子; z_m :第二类影响因子中的第 m 个主成分; l_{ij} :第一类影响因子中第 i 个影响因子与第 j 个影响因子的主成分载荷; x_p :第一类影响因子中的第 p 个影响因子。

[0038] 优选的,所述基于输电线路舞动的潜在影响因子对舞动信息记录表进行筛选获得第一类影响因子,包括:

[0039] 基于输电线路舞动的潜在影响因子获取所述潜在影响因子涉及的输电线路相关数据;

[0040] 基于潜在影响因子和所述潜在影响因子涉及的输电线路相关数据构建舞动信息记录表;

[0041] 从所述舞动信息记录表筛选出待分析的影响因子组成第一类影响因子;

[0042] 其中,所述输电线路舞动的潜在影响因子,包括:风速、风向、线路走向与风向夹角、温度、湿度、降水形式、导线覆冰厚度、地形、海拔高度、电压等级、杆塔回路数、绞线截面积、导线排列方式和导线分裂数。

[0043] 优选的,所述基于输电线路舞动的潜在影响因子对舞动信息记录表进行筛选获得第一类影响因子,还包括:

[0044] 将第一类影响因子中通过文字描述的影响因子按设置的转换原则转化为数值形式。

[0045] 优选的,所述基于所述第二类影响因子对数据舞动性进行展示,包括:

[0046] 从第二类影响因子中选择前两个主成分,并将分析数据表中所述主成分对应的输电线路相关数据绘制在直角坐标系中,通过直角坐标系展示数据舞动性。

[0047] 基于同一发明构思,本发明还提供了一种数据舞动性的展示系统,包括:

[0048] 筛选模块,用于基于输电线路舞动的潜在影响因子对舞动信息记录表进行筛选获得第一类影响因子;

[0049] 降维模块,用于基于所述舞动信息记录表中舞动数据之间的相关性对第一类影响因子进行降维处理得到第二类影响因子;

[0050] 展示模块,用于基于所述第二类影响因子对数据舞动性进行展示。

[0051] 优选的,所述降维模块,包括:

[0052] 生成子模块,用于从所述舞动信息记录表中选择第一影响因子及所述第一影响因子涉及的输电线路相关数据,生成分析数据表;

[0053] 降维子模块,用于基于所述分析数据表利用主成分分析法对第一类影响因子进行降维处理得到第二类影响因子。

[0054] 与现有技术相比,本发明的有益效果为:

[0055] 本发明提供的技术方案,基于输电线路舞动的潜在影响因子对舞动信息记录表进行筛选获得第一类影响因子;基于所述舞动信息记录表中舞动数据之间的相关性对第一类影响因子进行降维处理得到第二类影响因子;基于所述第二类影响因子对数据舞动性进行展示,可以直观地展示数据正负样本的分布情况,了解舞动数据特点,为通过舞动数据预测输电线路舞动提供了依据。

[0056] 本发明提供的技术方案,可以提取出与舞动发生率密切相关的影响因子,降低舞动预测模型中输入向量的维度,提高舞动预测预警的准确率。

附图说明

[0057] 图1为本发明提供的展示方法流程图;

[0058] 图2为本发明实施例提供的舞动信息记录表示意图;

[0059] 图3为本发明实施例提供的风向和线路走向数据格式转换的坐标变换图;

[0060] 图4为本发明实施例中的主成分分析示意图;

[0061] 图5为本发明实施例中舞动数据中正负样本分布示意图。

具体实施方式

[0062] 为了更好地理解本发明,下面结合说明书附图和实例对本发明的内容做进一步的说明。

[0063] 实施例1:

[0064] 图1为本发明提供的展示方法流程图,包括:

[0065] S1、基于输电线路舞动的潜在影响因子对舞动信息记录表进行筛选获得第一类影响因子;

[0066] S2、基于所述舞动信息记录表中舞动数据之间的相关性对第一类影响因子进行降维处理得到第二类影响因子;

[0067] S3、基于所述第二类影响因子对数据舞动性进行展示。

[0068] 本实施例中采用的文本数据为“舞动信息记录表”为例,通过本发明提供的展示方法进行展示,其中部分数据如图2所示。

[0069] S1、基于输电线路舞动的潜在影响因子对舞动信息记录表进行筛选获得第一类影响因子,具体包括:

[0070] 1.1舞动数据描述

[0071] 1.1.1原始数据及其格式化

[0072] 输电线路舞动相关数据庞杂,形式多元,包括图像、文本、视频等多种类型,本发明只针对文本数据进行研究。

[0073] 基于输电线路舞动的潜在影响因子获取潜在影响因子涉及的输电线路相关数据;

[0074] 基于潜在影响因子和潜在影响因子涉及的输电线路相关数据构建舞动信息记录表。

[0075] 图2中标记为舞动的记录记为正样本,标记为非舞动的记录记为负样本,显然正负样本分布不均衡,比例约为1:7。

[0076] 由于本实施例旨在分析输电线路舞动的主要影响因素,因此,只选取与舞动成因相关的数据进行分析。而输电线路舞动的潜在影响因子可以归为三大类,一是气象因素,二是地理因素,三是输电线路自身的结构特征。

[0077] 1) 气象因素:风速、风向、线路走向与风向夹角、温度、湿度、降水形式、导线覆冰厚度。

[0078] 2) 地理因素:地形、海拔高度。

[0079] 3) 输电线路自身的结构特征:电压等级、杆塔回路数、绞线截面积、导线排列方式、导线分裂数。

[0080] 根据舞动发生机理以及舞动规律的理论分析结果,筛选以下12个影响因子组成第一类影响因子进行分析:电压等级、海拔高度、线路走向、风速、风向、风向与线路夹角、温度、湿度、降水形式、导线覆冰厚度、外绞线面积与导线分裂数。

[0081] 从舞动信息记录表中选择第一影响因子及第一影响因子涉及的输电线路相关数据,生成分析数据表。

[0082] 将第一类影响因子中通过文字描述的影响因子按设置的转换原则转化为数值形式。

[0083] 由图2可知,线路走向、风向、降水形式3个影响因子的值为文字描述,其余9个影响因子均为数值型信息,为了能够被数据挖掘工具所识别,需要将文字描述转换为数值型信息,其中线路走向和风向按照图3所示的坐标进行转换。

[0084] 具体的格式转换原则如表1.1至表1.3所示。

[0085] 表1.1线路走向格式转换原则

线路走向	对应数值 (°)
南北/北南	90
东西/西东	0
东北-西南/西南-东北	45
西北-东南/东南-西北	135
东西偏北 N° 度	N
东西偏南 N 度	180-N
南北偏东 N 度	90-N
南北偏西 N 度	90+N
注*: N 为整数, 且 $0 < N < 180$	

[0087] 表1.2风向格式转换

[0088]

风向	对应数值
北(偏北)	90
东北	45
西北	135

[0089] 表1.3降水形式格式转换

[0090]

降水形式	对应数值
雾凇	1
雨凇	2
湿雪	3

[0091] 1.1.2数据直观展示

[0092] 为了较为直观地展示舞动样本(正样本)与非舞动样本(负样本)的分布情况,首先比较正负样本各影响因子的均值和标准差的区别,如表1.4和表1.5所示。

[0093] 表1.4正负样本各影响因子的均值比较

[0094]

影响因子	舞动（正样本）	非舞动（负样本）	相差
电压等级	194.1402	505.2507	311.1105
海拔高度	145.3601	297.4079	152.0478
线路走向	84.2568	33.6358	50.6209
风速	5.5299	3.9333	1.5966
风向	102.3253	101.7933	0.5320
风向与线路夹角	69.0740	57.0856	11.9883
温度	-2.3427	-0.1376	2.2051
湿度	83.1602	86.4076	3.2474
降水形式	2.3875	3	0.6125

[0095]

导线覆冰厚度	10.2933	4.1242	6.1691
外绞线面积	282.9317	537.8688	254.9371
导线分裂数	1.8822	4.0138	2.1316

[0096] 表1.5正负样本各影响因子的标准差比较

[0097]

影响因子	舞动（正样本）	非舞动（负样本）	相差
电压等级	155.1839	34.7129	120.4710
海拔高度	304.4459	78.1421	226.3038
线路走向	26.4801	27.6218	1.1417
风速	1.8422	1.4086	0.4336
风向	139.0564	109.3122	29.7443
风向与线路夹角	22.6357	59.5522	36.9168
温度	3.3518	4.0985	0.7446
湿度	14.3700	6.2863	8.0837
降水形式	1.3439	0	1.3439
导线覆冰厚度	10.0645	6.5918	3.4727
外绞线面积	106.9778	50.7202	56.2576
导线分裂数	1.2534	0.2200	1.0334

[0098] 由表1.4和表1.5可知,正负样本各个影响因子的均值和标准差的差别较大,表明舞动样本与非舞动样本的分布有明显的区别。

[0099] S2、基于所述舞动信息记录表中舞动数据之间的相关性对第一类影响因子进行降维处理得到第二类影响因子,包括:

[0100] 为了更为直观地展示正负样本的分布情况,绘制样本分布图是最佳选择,但是舞动数据具有12个维度,属于高维数据,若想在二位或三维空间中进行展示,需对数据进行降维。

[0101] 本发明采用主成分分析方法对舞动数据进行降维,主成分分析(Principal Component Analysis, PCA)是一种对于数据进行压缩常用的方法,该方法是将原来的变量通过线性变换,建立尽可能少的新变量,使得这些新变量是两两不相关的,而且这些新变量尽可能保持原有的信息。主成分分析的实质是对原始坐标系进行平移和旋转,使得新坐标的原点与数据群点的重心重合,新坐标系的第一轴与数据变异的最大方向对应,新坐标的第二轴与第一轴标准正交,并且对应于数据变异的第二大方向对应,依此类推。

[0102] 如图4所示,这些新轴可以称为第一主轴 e_1 ,第二主轴 e_2 等等,经舍弃少量信息后,主轴 e_1, e_2, \dots, e_m ($m < p$)能够有效的表示原数据的变异情况,由原来 p 维空间降至 m 维生成新的空间 $L(e_1, e_2, \dots, e_m)$ 称为 m 维主超平面,样本原点在主超平面上的投影近似表达原群点。

[0103] 原群点在主超平面第一轴上的投影构成新数据表的第一变量 $z_1 \in R^p$,称为第一主成分,一般地, z_h 被称为第 h 主成分 $h=1, 2, \dots, m$ 。

$$[0120] \quad r_{ij} = \frac{\sum_{k=1}^n (x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j)}{\sqrt{\sum_{k=1}^n (x_{ki} - \bar{x}_i)^2 \sum_{k=1}^n (x_{kj} - \bar{x}_j)^2}} \quad (\text{式 4})$$

[0121] 其中, \bar{x}_i 、 \bar{x}_j ($i, j=1, 2, \dots, p$) 分别表示 x_i 和 x_j 的均值, 并按下式计算:

$$[0122] \quad \begin{cases} \bar{x}_i = \frac{\sum_{k=1}^n x_{ki}}{n} \\ \bar{x}_j = \frac{\sum_{k=1}^n x_{kj}}{n} \end{cases}。$$

[0123] (2) 计算特征值与特征向量

[0124] 解特征方程 $|\lambda I - R| = 0$, 常用雅可比法 (Jacobi) 求出特征值 λ_i ;

[0125] 分别求出对应于特征值 λ_i 的特征向量 e_i ($i=1, 2, \dots, p$), 要求 $\|e_i\| = 1$, 即 $\sum_{j=1}^p e_{ij}^2 = 1$ 其

中 e_{ij} 表示第一类影响因子中第 i 个影响因子与第 j 个影响因子的特征向量。

[0126] (3) 按下式计算主成分的累计贡献率

$$[0127] \quad F_m = \frac{\sum_{i=1}^m \lambda_i}{\sum_{i=1}^p \lambda_i} \quad (i=1, 2, \dots, p, m < p)$$

[0128] 当累积贡献率大于 85% 时, 就认为能足够反映原来变量的信息了, 对应的 m 就是抽取的前 m 个主成分。

[0129] (4) 计算主成分载荷

$$[0130] \quad l_{ij} = p(z_i, x_j) = \sqrt{\lambda_i} e_{ij} \quad (i=1, 2, \dots, m, j=1, 2, \dots, p) \quad (\text{式 7})$$

[0131] 根据 (式 7) 计算得到的主成分载荷和第一类影响因子通过 (式 2) 得到新变量

[0132] 降维后选择第一、第二主成分来描述正负样本的分布情况, 如图 5 所示, 其中舞动样本即正样本, 非舞动样本, 即负样本; 显然图 5 可以明确地看出正负样本具有较为明显的区别。

[0133] 实施例 2

[0134] 基于同一发明构思, 本实施例还提供了一种数据舞动性的展示系统, 包括:

[0135] 筛选模块, 用于基于输电线路舞动的潜在影响因子对舞动信息记录表进行筛选获得第一类影响因子;

[0136] 降维模块, 用于基于所述舞动信息记录表中舞动数据之间的相关性对第一类影响因子进行降维处理得到第二类影响因子;

[0137] 展示模块, 用于基于所述第二类影响因子对数据舞动性进行展示。

[0138] 实施例中, 所述降维模块, 包括:

[0139] 生子模块, 用于从所述舞动信息记录表中选择第一影响因子及所述第一影响因子涉及的输电线路相关数据, 生成分析数据表;

[0140] 降维子模块, 用于基于所述分析数据表利用主成分分析法对第一类影响因子进行

降维处理得到第二类影响因子。

[0141] 实施例中,所述降维子模块,包括:

[0142] 构建单元,用于基于所述分析数据表中舞动数据之间的相关系数构建相关系数矩阵;

[0143] 第一计算单元,用于基于雅可比法得到所述相关系数矩阵的特征值和特征向量;

[0144] 计算累计贡献率单元,用于基于所述特征值和第一类影响因子的数量,计算所述第一类影响因子的累计贡献率;

[0145] 比较单元,用于基于所述累计贡献率和预设范围的关系,选择满足条件的所述累计贡献率对应的特征值;

[0146] 第二计算单元,用于基于所述对应的特征值和所述特征向量计算主成分载荷;

[0147] 结果单元,用于基于所述主成分载荷和第一类影响因子按设置的公式计算得到第二类影响因子。

[0148] 实施例中,所述筛选模块,包括:

[0149] 获取单元,用于基于输电线路舞动的潜在影响因子获取所述潜在影响因子涉及的输电线路相关数据;

[0150] 构建记录表单元,用于基于潜在影响因子和所述潜在影响因子涉及的输电线路相关数据构建舞动信息记录表;

[0151] 重组单元,用于从所述舞动信息记录表筛选出待分析的影响因子组成第一类影响因子;

[0152] 其中,所述输电线路舞动的潜在影响因子,包括:风速、风向、线路走向与风向夹角、温度、湿度、降水形式、导线覆冰厚度、地形、海拔高度、电压等级、杆塔回路数、绞线截面积、导线排列方式和导线分裂数。

[0153] 显然,所描述的实施例是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其它实施例,都属于本发明保护的范围。

[0154] 本领域内的技术人员应明白,本申请的实施例可提供为方法、系统、或计算机程序产品。因此,本申请可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且,本申请可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质(包括但不限于磁盘存储器、CD-ROM、光学存储器等)上实施的计算机程序产品的形式。

[0155] 本申请是参照根据本申请实施例的方法、设备(系统)、和计算机程序产品的流程图和/或方框图来描述的。应理解可由计算机程序指令实现流程图和/或方框图中的每一流程和/或方框、以及流程图和/或方框图中的流程和/或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理设备的处理器以产生一个机器,使得通过计算机或其他可编程数据处理设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的装置。

[0156] 这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理设备以特定方式工作的计算机可读存储器中,使得存储在该计算机可读存储器中的指令产生包括指令装置的制造品,该指令装置实现在流程图一个流程或多个流程和/或方框图一个方框或

多个方框中指定的功能。

[0157] 这些计算机程序指令也可装载到计算机或其他可编程数据处理设备上,使得在计算机或其他可编程设备上执行一系列操作步骤以产生计算机实现的处理,从而在计算机或其他可编程设备上执行的指令提供用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的步骤。

[0158] 以上仅为本发明的实施例而已,并不用于限制本发明,凡在本发明的精神和原则之内,所做的任何修改、等同替换、改进等,均包含在申请待批的本发明的权利要求范围之内。

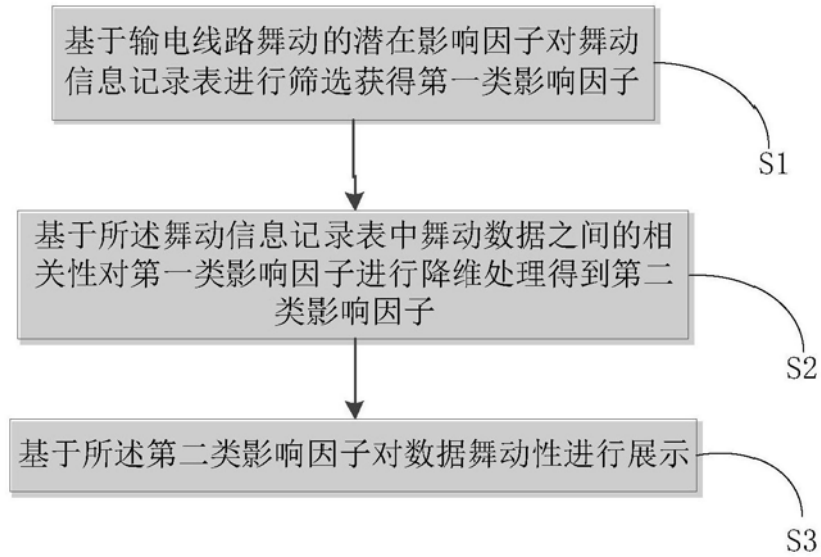


图1

序号	单位	线路名称	电压等级 (kV)	导线舞动区地形 (山区、丘陵、平原)	海拔高度 (m)	舞动段线路走向	风速 (m/s)	风向	风向与舞动段线路走向夹角 (°)	温度 (°C)	湿度	降水形式 (雨淞、雾淞、湿雪)	导线覆冰厚度 (mm)	杆塔回路数	导线型号	导线排列方式 (垂直、水平、三角)	导线分裂数	是否舞动
1	长春超高压局	丰徐2号线	500	丘陵	353	东西	6	西北	26	15	98%	雨淞	5	1	400/35	水平	4	是
2	锦州超高压局	董辽一线	500	山区	200	东西	12m-18m/s	北风	90	2~-7	50%-70%	雨淞	15	单回	LGJQ-300	水平	4	是
3	锦州超高压局	董辽一线	500	山区	250	东西	4m-6m/s	北风	90	0~-8	50%-70%	雨淞	5-8	单回	LGJQ-300	水平	4	否
4	锦州超高压局	董辽二线	500	山区	250	东西	4m-6m/s	北风	90	0~-8	50%-70%	雨淞	5-8	单回	LGJ-300/40	水平	4	否
5	锦州超高压局	沙董一线	500	平原、丘陵	100	东西	4m-6m/s	北风	90	0~-8	50%-70%	雨淞	5-8	单回	LGJ-400/35	水平	4	否

图2

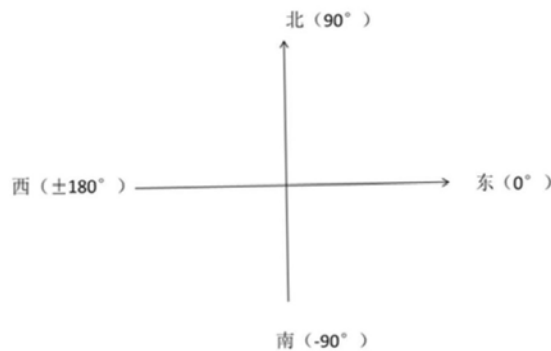


图3

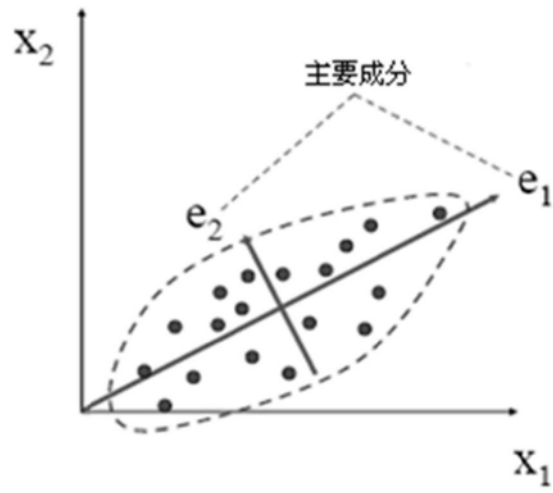


图4

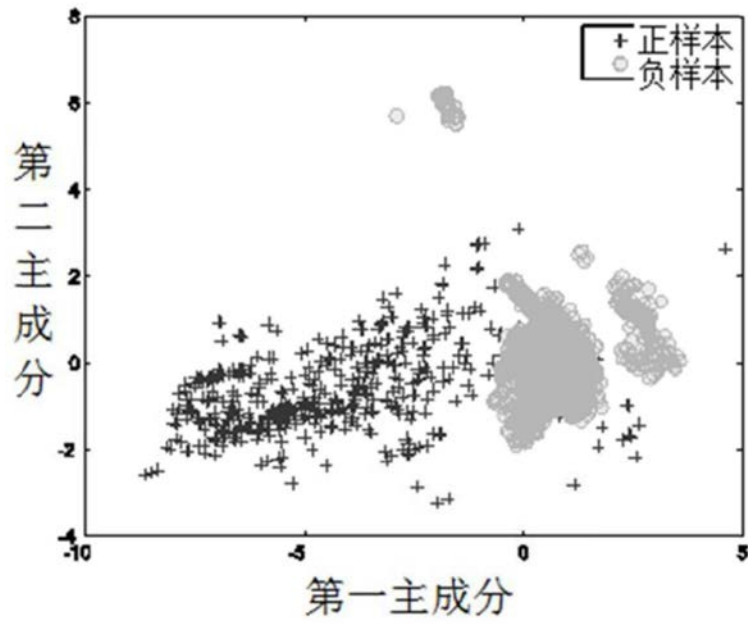


图5