

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2020 年 7 月 16 日 (16.07.2020)



(10) 国际公布号
WO 2020/143253 A1

(51) 国际专利分类号:
G06Q 50/06 (2012.01) **G06N 3/063** (2006.01)

(21) 国际申请号: PCT/CN2019/108714

(22) 国际申请日: 2019 年 9 月 27 日 (27.09.2019)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201910016263.4 2019 年 1 月 8 日 (08.01.2019) CN

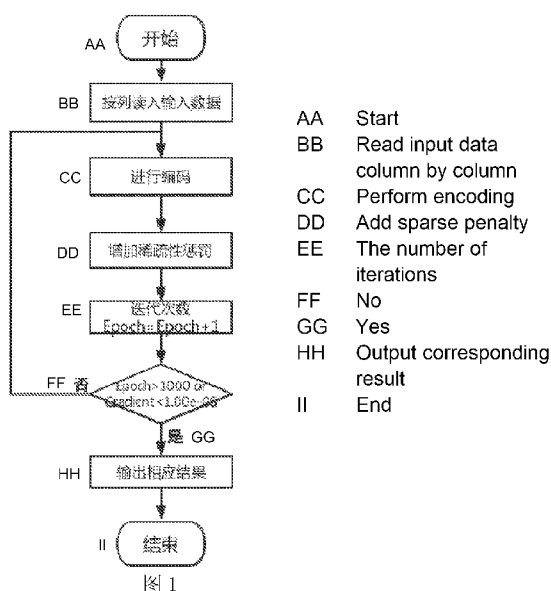
(71) 申请人: 西安交通大学 (XI' AN JIAOTONG UNIVERSITY) [CN/CN]; 中国陕西省西安市咸宁西路 28 号, Shaanxi 710049 (CN)。国网江苏省电力有限公司 (STATE GRID JIANGSU ELECTRIC POWER CO., LTD.) [CN/CN]; 中国江苏省南京市上海路 215 号, Jiangsu 210024

(CN)。全球能源互联网研究院有限公司 (GLOBAL ENERGY INTERNET RESEARCH INSTITUTE LTD.) [CN/CN]; 中国北京市昌平区未来科技城北区, Beijing 102200 (CN)。

(72) 发明人: 李更丰 (LI, Gengfeng); 中国陕西省西安市咸宁西路 28 号, Shaanxi 710049 (CN)。雷宇骁 (LEI, Yuxiao); 中国陕西省西安市咸宁西路 28 号, Shaanxi 710049 (CN)。徐春雷 (XU, Chunlei); 中国江苏省南京市上海路 215 号, Jiangsu 210024 (CN)。张啸虎 (ZHANG, Xiaohu); 中国北京市昌平区未来科技城北区, Beijing 102200 (CN)。史迪 (SHI, Di); 中国北京市昌平区未来科技城北区, Beijing 102200 (CN)。姜江枫 (JIANG, Jiangfeng); 中国陕西省西安市咸宁西路 28 号, Shaanxi 710049 (CN)。黄玉雄 (HUANG, Yuxiong); 中国陕西省西安市咸宁西路 28 号, Shaanxi 710049

(54) Title: METHOD EMPLOYING SPARSE AUTOENCODER TO CLUSTER POWER SYSTEM OPERATION MODES

(54) 发明名称: 一种基于 Sparse Autoencoder 的电力系统运行方式聚类方法



(57) Abstract: A method employing a sparse autoencoder to cluster power system operation modes. The method comprises: acquiring related data in a power system; setting a training parameter, the number of hidden layers and the number of neurons; using the related data to train an autoencoder model while simultaneously extracting a topological structure and a weight matrix from the model, and performing cluster analysis to obtain a number of typical scenarios; and decoding to obtain original data at centers of the respective scenarios. The method enables fast selection and dimensionality reduction to be performed with respect to feature vectors representing operation modes of a power system. The invention provides a novel idea and method for fast selection of a feature vector representing

(CN)。彭钰茗(PENG, Yuming); 中国陕西省西安市咸宁西路28号, Shaanxi 710049 (CN)。别朝红(BIE, Zhaohong); 中国陕西省西安市咸宁西路28号, Shaanxi 710049 (CN)。

(74) 代理人: 西安通大专利代理有限公司(XI' AN TONG DA PATENT AGENCY CO., LTD.); 中国陕西省西安市咸宁西路28号, Shaanxi 710049 (CN)。

(81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

(84) 指定国(除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告(条约第21条(3))。

an operation mode of a power system and generation of a typical operation scenario. The invention is useful in a novel application for neural networks.

(57) 摘要: 一种基于Sparse Autoencoder的电力系统运行方式聚类方法, 获取电力系统中的相关数据, 然后设置训练参数, 隐藏层层数和神经元个数, 对相关数据进行Autoencoder模型训练, 同时提取出模型的拓扑结构和权值矩阵, 进行聚类分析, 得出典型场景数目, 解码得出各场景中心的原始数据。本方法可以对表征电力系统运行方式的特征向量进行快速选取和降维, 为电力系统运行方式特征向量的选取并生成典型运行场景提供一种新思路和方法。同时为神经网络在这一方面的应用开创了一个先例。

一种基于 Sparse Autoencoder 的电力系统运行方式聚类方法

技术领域

本发明属于电力系统安全校验、规划运行技术领域，具体涉及一种基于 Sparse Autoencoder 的电力系统运行方式聚类方法。

背景技术

电力系统中使用典型运行方式对于电网的安全运行进行校验有着举足轻重的作用。在规划时期考虑到典型的运行方式并以此进行电力系统的运行校验，可以最大程度的防止电压越限、过负荷等事故的发生，保证电力系统对负荷和用户的持续供电能力。但是随着新能源的不断接入，电力系统的运行随机性大幅度提高，导致运行方式的特征也更加复杂，如何提取运行方式的特征向量生成典型场景变得尤为困难。然而利用传统的 PCA 方法无法准确地提取特征向量，时间复杂度过高，实用性也大幅度降低。

因此，为保证可靠地提取出表征电力系统运行方式的特征向量以进行典型场景分析，选取合理的特征向量提取方式是亟需认真考虑的问题。

针对上述问题，本发明提出了一种利用 Sparse Autoencoder 技术提取表征电力系统运行方式的特征向量的方法。

发明内容

本发明所要解决的技术问题在于针对上述现有技术中的不足，提供一种基于 Sparse Autoencoder 的电力系统运行方式聚类方法。

本发明采用以下技术方案：

一种基于 Sparse Autoencoder 的电力系统运行方式聚类方法，获取电力系统中的相关数据，然后设置训练参数，隐藏层层数和神经元个数，对相关数据进行 Autoencoder 模型训练，同时

提取出模型的拓扑结构和权值矩阵，进行聚类分析，得出典型场景数目，解码得出各场景中心的原始数据。

具体的，相关数据形成 n 行 m 列的输入矩阵 X_n^m ， n 为向量， m 为样本数量。

进一步的，相关数据包括电力系统内各个节点电压、电压幅值、各个节点的发电机有功功率和无功功率的数据以及电力系统在研究时间范围内的时序负荷数据。

具体的，设置训练参数，隐藏层层数和神经元个数如下：

设置相关参数为 α ， η 与最大迭代次数为初始化训练参数， α 是 L2 正则化方法的系数， η 是稀疏正则化的系数；设置隐藏层层数为单层，即 $l=1$ ；设置第 l 隐藏层神经元个数，即最终特征向量维数 $h_l=2$ 。

具体的，对相关数据进行 Autoencoder 模型训练步骤具体如下：

S201、以相关数据形成 n 行 m 列的输入矩阵 X_n^m 作为输入；

S202、输入可接受的误差 e 以及训练时间 t 进行可视化训练，观察误差和训练过程；

S203、提取最底层特征向量 $features_l$ ，并对 $features_l$ 进行聚类分析；

S204、找出 k 类场景中心，进行解码还原得出典型场景原始数据中心，同时还原出全部原始数据 $\hat{X}_n^A$ ；

S205、得出所需结果，循环结束。

进一步的，步骤 S202 中，若还原输入数据后与原始输入数据的欧氏距离大于 e ，增加迭代次数，重新训练模型；若训练模型时间大于 t ，即在迭代早期误差达到范围，减小迭代次数，重新训练模型。

进一步的，步骤 S203 中，选取 *K-means* 方法进行聚类，设聚类中心个数为 k ，设置初始值为 $k=1$ ，计算轮廓值 Sil_k^h ；给 $k=k+1$ ，计算轮廓值 Sil_k^h ，当 $k=h$ 时，退出循环；得出最大的轮廓值 Sil_k^h ，得出典型场景个数 k 。

更进一步的，若最大轮廓值 Sil_k^h 小于 0.85，当 $h_l < h_{l-1}$ 返回设置神经元个数， $h_l = h_l + 1$ ，重新训练模型；否则，返回设置隐藏层层数， $l=l+1$ ，重新训练模型。

进一步的，步骤 S204 中，计算矩阵 X_n^m 与 $\hat{X}_n^\Lambda$ 的欧氏距离 Φ_d ，若 $\Phi_d \leq \varepsilon$ 则接受。

进一步的，步骤 S204 中，若 $\Phi_d > \varepsilon$ ，若 $l > 1$ ，返回 $l=l-1$ ，重新训练模型；否则，返回 $h=h-1$ ，重新训练模型。

与现有技术相比，本发明至少具有以下有益效果：

本发明一种基于 Sparse Autoencoder 的电力系统运行方式聚类方法，将 Sparse Autoencoder 技术应用于电力系统特征向量的选取上，不需要复杂繁琐的人工标准化数据的过程，通过训练模型可以找到输入量之间的相关性，同时更重要的是可以降低特征向量的维数，确定聚类的初始场景数，同时大大减少了聚类时间上的复杂度。

进一步的，相关数据反应了电力系统运行的主要特征，将相关数据作为输入可以增加 Sparse Autoencoder 训练模型的速度和精度。

进一步的，根据不同电力系统模型聚类精度的要求，灵活设置初始的训练参数，隐藏层层数和神经元个数，便于针对不同的情况进行训练。

进一步的，通过进行 Autoencoder 模型训练，可以提高该模型的精度，准确提取出特征向量，为聚类分析提供了良好的条件。

进一步的，通过对 Autoencoder 模型训练得出的最底层特征向量 features，得出场景聚类轮廓值，用于判断模型的优劣以及修改模型。

进一步的，将训练所得最底层特征向量 features 进行还原，与输入向量进行比较，判断模型的还原度和误差，若满足要求则该模型可用。

进一步的，将训练所得最底层特征向量 features 进行还原，与输入向量进行比较，若误差过大，则修改参数重新训练模型。

综上所述，本发明可以对表征电力系统运行方式的特征向量进行快速选取和降维，为电力系统运行方式特征向量的选取并生成典型运行场景提供一种新思路和方法。同时为神经网络在这一方面的应用开创了一个先例。

下面通过附图和实施例，对本发明的技术方案做进一步的详细描述。

附图说明

图 1 为 Sparse Autoencoder 的程序流程图；

图 2 为 Sparse Autoencoder 的算法简图。

具体实施方式

由于 Sparse Autoencoder 技术可以避免复杂的电力系统数据标准化，训练出的特征向量作聚类分析误差小并且可以在解码后较好地还原电力系统原始数据，具有及其优良的特性，因此选择该技术。

Sparse Autoencoder 是一种无监督学习算法，它使用反向传播算法，并让目标值等于输入值，比如 $y^{(i)} = x^{(i)}$ ，该神经网络尝试学习一个 $h_{w,b}(x) \approx x$ 的函数。此时，当减少神经元的个数会强迫该神经网络去学习输入数据的压缩表示，这就实现了数据的降维过程。同时，由于这一算法有利于发现输入数据中的相关性，就更适用于电力系统。

A、稀疏性的定义：

神经元的平均输出激活测度定义为：

$$\hat{\rho}_j = \frac{1}{m} \sum_{i=1}^m [a_j^{(1)}(x^{(i)})]$$

$a_j^{(1)}$ 表示隐藏神经元 j 的激活度，我们将使用 $a_j^{(1)}$ 来表示在给定输入为 x 情况下，自编码神经网络隐藏神经元 j 的激活度。同时为了增加模型的稀疏性，加入稀疏性约束：

$$\hat{\rho}_j = \rho$$

其中， ρ 是稀疏性参数，通常是一个接近于 0 的较小的值（比如 $\rho = 0.03$ ）。同时为了实

现这一限制，我们将会在我们的优化目标函数中加入一个额外的惩罚因子，而这一惩罚因子将惩罚那些 $\hat{\rho}_j$ 和 ρ 有显著不同的情况从而使得隐藏神经元的平均活跃度保持在较小范围内。

惩罚因子的具体形式有很多种合理的选择，我们将会选择以下这一种：

$$\Phi_{sparsity} = \sum_{j=1}^{s_1} \rho \log \frac{\rho}{\hat{\rho}_j} + (1-\rho) \log \frac{1-\rho}{1-\hat{\rho}_j}.$$

s_1 是隐藏层中隐藏神经元的数量，而索引 j 依次代表隐藏层中的每一个神经元。

B、L2 正则化方法：

正则化是机器学习中一个防止过拟合的重要手段，因为实际的模型可能并没有那么复杂，同时学习的模型拓扑以及学到的权值矩阵只是在训练数据上表现的结果可能会比较好。所用的特征过多，样本较少时就很容易陷入过拟合。那么我们就需要把它转化成为更为简单的模型。

本发明使用了 L2 正则化方法。

定义以下公式：

$$\Phi_{weights} = \frac{1}{2} \sum_l^L \sum_j^n \sum_i^k (w_{ji}^{(l)})^2$$

L 是隐藏层的层数， n 是观测值的数量， k 是训练集中变量的个数。

C、成本函数：

$$E_{cost} = \frac{1}{N} \sum_{n=1}^N \sum_{k=1}^K (x_{kn} - x_{kn}^{\Lambda})^2 + \alpha * \Phi_{weights} + \eta * \Phi_{sparsity}$$

α 是 L2 正则化方法的系数， η 是稀疏正则化的系数，可以通过 L2WeightRegularization 和 SparsityRegularization 函数去分别修改。

本发明提供了一种基于 Sparse Autoencoder 的电力系统运行方式聚类方法，获取电力系统中的相关数据，如：电力系统中节点电压、电压幅值、节点负荷、发电机有功和无功出力等；然后设置训练参数，隐藏层层数和神经元个数，训练相关模型，同时提取出模型的拓扑结构

和权值矩阵，然后进行聚类分析；最后得出典型场景数目，并解码得出各场景中心的原始数据。利用本发明方法可以对表征电力系统运行方式的特征向量进行快速选取和降维，为电力系统运行方式特征向量的选取并生成典型运行场景提供一种新思路和方法。。

请参阅图 1 和图 2，本发明一种基于 Sparse Autoencoder 的电力系统运行方式聚类方法的步骤如下：

S1、数据简单初始化；

粗略进行电力系统运行的数据筛选，例如：获取该系统内各个节点电压、各个节点的发电机有功功率和无功功率的数据以及系统在研究时间范围内的时序负荷数据。这些数据构成 n 维，即 n 行向量。同时样本数量共有 m 个，即形成 n 行 m 列输入矩阵，记为 X_n^m 。

S2、对步骤 S1 得到的数据矩阵进行 Autoencoder 模型训练，提取最底层特征向量进行聚类，确定典型场景个数，解码还原全部数据 $\hat{X}_n^A$ ；

设置相关参数 α ， η 与最大迭代次数， α 是 L2 正则化方法的系数， η 是稀疏正则化的系数，即初始化训练参数；第 l 隐藏层神经元个数，即最终特征向量维数，记为 $h_l=2$ ；隐藏层数，默认为单层，记为 $l=1$ ；

S201、将 X_n^m 作为输入，在 Matlab 中进行 Autoencoder 模型训练；

S202、可视化训练过程，观察误差和训练过程，输入可接受的误差 e 以及训练时间 t 进行可视化训练，若还原输入数据后与原始输入数据的欧氏距离大于 e ，增加迭代次数，重新训练模型；若训练模型时间大于 t ，即在迭代早期误差可达到范围，减小迭代次数，重新训练模型；

S203、提取最底层特征向量，记为 $features_l$ ，并对 $features_l$ 进行聚类分析；

选取 $K-means$ 方法进行聚类，聚类中心个数设为 k ，初始值设置为 $k=1$ ，计算轮廓值大小，记作 Sil_k^h ；不断给 $k=k+1$ ，计算轮廓值大小，记作 Sil_k^h ，当 $k=h$ 时，退出循环；得出最大的轮

廓值 Sil_k^h ，得出 k 值，即为典型场景个数；若认为最大轮廓值 Sil_k^h 小于 0.85，当 $h_l < h_{l-1}$ 返回设置神经元个数， $h_l = h_l + 1$ ，重新训练模型；否则，返回设置隐藏层层数， $l = l + 1$ ，重新训练模型；

S204、找出 k 类场景中心，进行解码还原得出典型场景原始数据中心；同时还原出全部原始数据，记为 $\hat{X}_n$ ；

计算矩阵 X_n^m 与 $\hat{X}_n$ 的欧氏距离，记为 Φ_d ，若 $\Phi_d \leq \varepsilon$ 则接受上述模型和结果；

若 $\Phi_d > \varepsilon$ ：

若 $l > 1$ ，返回 $l = l - 1$ ，重新训练模型；

否则，返回 $h = h - 1$ ，重新训练模型；

S205、得出所需结果，循环结束。

S3、提取出模型拓扑以及学到的权值矩阵，根据需要分析变量的相关性。

提取 S2 中的最佳 k 值，即为典型场景个数，并提取出对应的场景中心原始数。

为使本发明实施例的目的、技术方案和优点更加清楚，下面将结合本发明实施例中的附图，对本发明实施例中的技术方案进行清楚、完整地描述，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。通常在此处附图中的描述和所示的本发明实施例的组件可以通过各种不同的配置来布置和设计。因此，以下对在附图中提供的本发明的实施例的详细描述并非旨在限制要求保护的本发明的范围，而是仅仅表示本发明的选定实施例。基于本发明中的实施例，本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例，都属于本发明保护的范围。

下面结合附图和运用 IEEE-14 节点系统例对本发明做详细说明。

初步选取的输入量如表 1，共有 30000 个样本数据，各具有 53 个特征向量记为 X_{53}^{30000} ：

表一 输入数据

特征向量/组号	a(1,2,3...10000)	b(1,2,3...10000)	c(1,2,3...10000)
Ua_1	-3.7,-3.8,-3.7...-3.6	-5.0,-4.9,-5.0...-5.0	-5.9,-6.0,-5.9...-6.0
...	...	...	...
Ua_{14}	-12.1,-12.3,-12.5...-12.2	-16.0,-16.0,-16.0...-16.0	-18.8,-18.9,-19.2...-19.0
Um_1	1.06,1.06,1.06...1.06	1.06,1.06,1.06...1.06	1.06,1.06,1.06...1.06
...	...	...	...
Um_{14}	1.04,1.04,1.04...1.04	1.04,1.04,1.04...1.04	1.04,1.04,1.04...1.04
Pd_1	17.1,16.6,17.3...17.8	21.7,21.7,21.7...21.7	25.5,24.9,25.6...25.8
...	...	...	...
Pd_{14}	11.6,11.2,11.7...11.5	14.9,14.9,14.9...14.9	17.5,17.3,17.2...16.9
Pg_1	175.8,173.2,179.4...180.2	232.4,232.4,232.4...232.4	275.1,279.8,283.5...277.5
...	...	...	...
Pg_5	30.3,30.8,31.1...31.8	40,40,40...40	47.9,48.1,46.8...47.9
Qg_1	-7.9,-7.4,-8.8...-7.6	-16.5,-16.5,-16.5...-16.5	-22.7,-22.1,-21.8...-21.9
...	...	...	...
Qg_5	15.5,15.6,15.2...14.9	17.6,17.6,17.6...17.6	19.4,18.7,19.6...19.3

1、a, b 和 c 组分别代表 IEEE14 节点系统三种不同负荷水平所得到的数据作为输入集合。

将 X_{53}^{30000} 作为输入，进行 2) 中操作；

2、进行模型训练：设置最大迭代次数为 1000， $\alpha=0.01$ ， $\eta=4$ ；设置初始 $h_l=2$ 以及 $l=1$ 等按照 2) 中方法不断进行循环找出最佳结果；

3、提取出模型拓扑以及学到的权值矩阵，根据需要分析变量的相关性。提取 2) 中的最

佳 k 值，即为典型场景个数，并提取出对应的场景中心原始数据。

将计算所得轮廓值展示如下表：

表二 轮廓值计算值

$\begin{matrix} \text{h} \\ \text{k} \end{matrix}$	2	3	4	5
2	0.78	0.98	0.86	0.73
3	0.77	0.97	0.84	0.72
4	0.80	0.96	0.83	0.70
5	0.73	0.96	0.81	0.67
6	0.76	0.96	0.80	0.67
7	0.76	0.97	0.82	0.68
8	0.74	0.96	0.81	0.67

由表 2 可知当典型场景数目为三类时，计算所得的轮廓值最大为 0.96 左右，得出在该输入数据进行训练的情况下，最佳的聚类水平应该分为三类。其聚类结果符合预期随着负荷水平分类的三种情况，具有极显著的特性。

同时在训练的场景数目不变的情况下，聚类的时间和参与聚类的特征向量的维数几乎成线性关系，即特征向量维数越高，聚类的时间越长。体现了使用 Sparse Autoencoder 对于典型场景的分类，在聚类效果几乎不变的情况下，当对特征向量进行降维时，其极大减少消耗的时间，这满足了电力系统在计算时的快速性。同时由结果可见，若电网的规模更大即节点数，特征向量维数更高，通过 Sparse Autoencoder 降低特征向量维数将对于聚类效果的提升有着更加显著的效果，对实际计算有着极大的帮助。

以上内容仅为说明本发明的技术思想，不能以此限定本发明的保护范围，凡是按照本发

明提出的技术思想，在技术方案基础上所做的任何改动，均落入本发明权利要求书的保护范围之内。

权利要求书

1、一种基于 Sparse Autoencoder 的电力系统运行方式聚类方法，其特征在于，获取电力系统中的相关数据，然后设置训练参数，隐藏层层数和神经元个数，对相关数据进行 Autoencoder 模型训练，同时提取出模型的拓扑结构和权值矩阵，进行聚类分析，得出典型场景数目，解码得出各场景中心的原始数据。

2、根据权利要求 1 所述的基于 Sparse Autoencoder 的电力系统运行方式聚类方法，其特征在于，相关数据形成 n 行 m 列的输入矩阵 X_n^m ， n 为向量， m 为样本数量。

3、根据权利要求 1 或 2 所述的基于 Sparse Autoencoder 的电力系统运行方式聚类方法，其特征在于，相关数据包括电力系统内各个节点电压、电压幅值、各个节点的发电机有功功率和无功功率的数据以及电力系统在研究时间范围内的时序负荷数据。

4、根据权利要求 1 所述的基于 Sparse Autoencoder 的电力系统运行方式聚类方法，其特征在于，设置训练参数，隐藏层层数和神经元个数如下：

设置相关参数为 α ， η 与最大迭代次数为初始化训练参数， α 是 L2 正则化方法的系数， η 是稀疏正则化的系数；设置隐藏层层数为单层，即 $l=1$ ；设置第 l 隐藏层神经元个数，即最终特征向量维数 $h_l=2$ 。

5、根据权利要求 1 所述的基于 Sparse Autoencoder 的电力系统运行方式聚类方法，其特征在于，对相关数据进行 Autoencoder 模型训练步骤具体如下：

S201、以相关数据形成 n 行 m 列的输入矩阵 X_n^m 作为输入；

S202、输入可接受的误差 e 以及训练时间 t 进行可视化训练，观察误差和训练过程；

S203、提取最底层特征向量 $features_l$ ，并对 $features_l$ 进行聚类分析；

S204、找出 k 类场景中心，进行解码还原得出典型场景原始数据中心，同时还原出全部原始数据 $\hat{X}_n^A$ ；

S205、得出所需结果，循环结束。

6、根据权利要求5所述的基于 Sparse Autoencoder 的电力系统运行方式聚类方法，其特征在于，步骤 S202 中，若还原输入数据后与原始输入数据的欧氏距离大于 e ，增加迭代次数，重新训练模型；若训练模型时间大于 t ，即在迭代早期误差达到范围，减小迭代次数，重新训练模型。

7、根据权利要求5所述的基于 Sparse Autoencoder 的电力系统运行方式聚类方法，其特征在于，步骤 S203 中，选取 *K-means* 方法进行聚类，设聚类中心个数为 k ，设置初始值为 $k=1$ ，计算轮廓值 Sil_k^h ；给 $k=k+1$ ，计算轮廓值 Sil_k^h ，当 $k=h$ 时，退出循环；得出最大的轮廓值 Sil_k^h ，得出典型场景个数 k 。

8、根据权利要求7所述的基于 Sparse Autoencoder 的电力系统运行方式聚类方法，其特征在于，若最大轮廓值 Sil_k^h 小于 0.85，当 $h_l < h_{l-1}$ 返回设置神经元个数， $h_l = h_l + 1$ ，重新训练模型；否则，返回设置隐藏层层数， $l=l+1$ ，重新训练模型。

9、根据权利要求5所述的基于 Sparse Autoencoder 的电力系统运行方式聚类方法，其特征在于，步骤 S204 中，计算矩阵 X_n^m 与 $\hat{X}_n^A$ 的欧氏距离 Φ_d ，若 $\Phi_d \leq \varepsilon$ 则接受。

10、根据权利要求5所述的基于 Sparse Autoencoder 的电力系统运行方式聚类方法，其特征在于，步骤 S204 中，若 $\Phi_d > \varepsilon$ ，若 $l > 1$ ，返回 $l=l-1$ ，重新训练模型；否则，返回 $h=h-1$ ，重新训练模型。

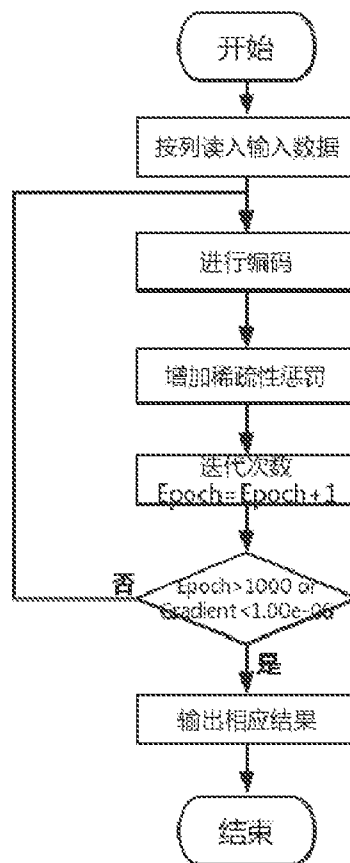


图 1

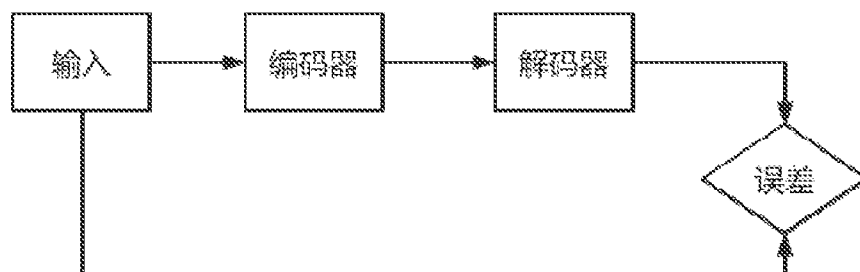


图 2

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/108714

A. CLASSIFICATION OF SUBJECT MATTER

G06Q 50/06(2012.01)i; G06N 3/063(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06Q; G06N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WPI, EPODOC, CNPAT, CNKI: 自编码, 自动编码, 无监督, 稀疏, 电力, 聚类, 分类, 神经网络, sparse, autoencoder, power, cluster+, training

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 109711483 A (XI'AN JIAOTONG UNIVERSITY et al.) 03 May 2019 (2019-05-03) claims 1-10	1-10
Y	CN 108846410 A (XIANGTAN UNIVERSITY) 20 November 2018 (2018-11-20) description, paragraphs [0039]-[0065]	1-10
Y	付晓 等 (FU, Xiao et al.). "基于特征聚类的稀疏自编码快速算法 (An Optimized Sparse Auto-encoder Network Based on Feature Clustering)" <i>Chinese Journal of Electronics</i> , Vol. 46, No. (5), 31 May 2018 (2018-05-31), section 3	1-10
A	CN 108985330 A (HUAZHONG UNIVERSITY OF SCIENCE AND TECHNOLOGY) 11 December 2018 (2018-12-11) entire document	1-10
A	CN 106447039 A (XI'AN JIAOTONG UNIVERSITY) 22 February 2017 (2017-02-22) entire document	1-10
A	CN 108491859 A (GUO, Jingqiu et al.) 04 September 2018 (2018-09-04) entire document	1-10



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

16 December 2019

Date of mailing of the international search report

27 December 2019

Name and mailing address of the ISA/CN

China National Intellectual Property Administration
No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing
100088
China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2019/108714

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN	109711483	A	03 May 2019	None	
CN	108846410	A	20 November 2018	None	
CN	108985330	A	11 December 2018	None	
CN	106447039	A	22 February 2017	None	
CN	108491859	A	04 September 2018	None	

国际检索报告

国际申请号

PCT/CN2019/108714

A. 主题的分类

G06Q 50/06(2012.01)i; G06N 3/063(2006.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

G06Q; G06N

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))

WPI, EPDOC, CNPAT, CNKI: 自编码, 自动编码, 无监督, 稀疏, 电力, 聚类, 分类, 神经网络, sparse, autoencoder, power, cluster+, training

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
PX	CN 109711483 A (西安交通大学 等) 2019年 5月 3日 (2019 - 05 - 03) 权利要求1-10	1-10
Y	CN 108846410 A (湘潭大学) 2018年 11月 20日 (2018 - 11 - 20) 说明书第[0039]-[0065]段	1-10
Y	付晓 等. “基于特征聚类的稀疏自编码快速算法” 电子学报, 第46卷, 第5期, 2018年 5月 31日 (2018 - 05 - 31), 第3节	1-10
A	CN 108985330 A (华中科技大学) 2018年 12月 11日 (2018 - 12 - 11) 全文	1-10
A	CN 106447039 A (西安交通大学) 2017年 2月 22日 (2017 - 02 - 22) 全文	1-10
A	CN 108491859 A (郭静秋 等) 2018年 9月 4日 (2018 - 09 - 04) 全文	1-10

☐ 其余文件在C栏的续页中列出。☒ 见同族专利附件。

* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件	“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件
--	---

国际检索实际完成的日期

2019年 12月 16日

国际检索报告邮寄日期

2019年 12月 27日

ISA/CN的名称和邮寄地址

中国国家知识产权局(ISA/CN)
中国北京市海淀区蓟门桥西土城路6号 100088

传真号 (86-10)62019451

受权官员

丁文勃

电话号码 86-(10)-53961330

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2019/108714

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN	109711483	A	2019年 5月 3日	无	
CN	108846410	A	2018年 11月 20日	无	
CN	108985330	A	2018年 12月 11日	无	
CN	106447039	A	2017年 2月 22日	无	
CN	108491859	A	2018年 9月 4日	无	