

基于 t-SNE 的电力系统电气距离可视化方法研究

史东宇¹, 胡文强¹, 李刚¹, 严剑峰¹, 臧德春²

(1. 中国电力科学研究院有限公司, 北京 100192; 2. 东南大学电气工程学院, 江苏 南京 211189)

摘要:我国已初步建成以特高压为骨干网架的交直流混联电网, 电网的运行方式更加复杂和多样, 这对调控运行的分析和展示提出了更高要求。文中从电力系统电气距离的角度出发, 利用基于 t 分布的随机近邻嵌入 (t-distribution stochastic neighbor embedding, t-SNE) 把设备之间高维、复杂的电气距离关系, 映射到二维或三维空间, 并结合可视化方法来展示设备间的紧密程度, 给运行和分析人员以更直观的印象。采用实际电网数据验证了方法的有效性, 本方法可以快速响应网架的变化, 满足在线分析和展示的要求。

关键词:电力系统; 电气距离; 可视化; t-SNE

中图分类号: TM734

文献标志码: A

文章编号: 2096-3203(2018)02-0078-05

0 引言

为了保障电能安全可靠的传输, 中国电网开展了西电东送、全国联网以及特高压输电等重大工程, 交直流混联的特大电网已经基本形成^[1]。随着电网规模的扩大, 电网安全稳定特性愈加多变。为提升调度人员掌控大电网运行的能力, 目前省级以上调度单位都已部署在线分析模块^[2-7], 而在线分析技术的应用也对可视化提出了更高的要求, 开展对运行电网全面细致的在线监视、分析和控制是调度自动化系统的迫切需求^[1]。

现有电网调度系统的可视化方法^[8-14]大多以表现设备间连接关系的二维图形为基础, 结合箭头、饼图、挂牌、动画、曲线、渲染图、3D 图等其他方式, 对复杂多样的电网运行状态进行监控, 地理图和单线图是最主要的 2 种可视化形式。这种图形通常是按照一定规则事先绘制或者自动生成的, 图上的厂站或设备位置只能表现固定的地理坐标或位置关系^[14]。而在电力系统分析中, 设备的“远近”是以电气距离为标准的, 它是电网稳定特性的重要参数, 与地理位置没有直接的关系。也就是说, 目前电网调度系统的可视化方法仅适用于电网监控, 无法直接表现出电网的稳定特性。

文中基于实际网架信息, 利用基于 t 分布的随机近邻嵌入 (t-distribution stochastic neighbor embedding, t-SNE) 算法^[15]把设备之间电气距离关系映射到二维或三维空间, 并结合可视化方法来展示不同设备间的紧密程度, 给运行分析人员以更直观的印象。

同时, 本方法还可进一步应用于同调机群分析、灵敏度分析、薄弱环节辨识等电网稳定的分析方法之中。

1 相关概念

1.1 电气距离

电气距离是指电力系统中任意 2 个节点之间的联系阻抗, 是决定电网稳定特征的重要参数, 阻抗值小代表这两点间的联系紧密, 相互影响作用明显。电气距离与地理距离之间没有必然的联系, 它主要取决于电压等级、两点之间直接或间接连接的支路数量以及支路阻抗大小等因素。

电力系统电气距离的求解方法已经较为成熟, 通常是基于节点阻抗矩阵 Z 来求解节点对的自阻抗, 而获取节点阻抗矩阵 Z 的方法一般有 2 种: 支路追加法和节点导纳矩阵 Y 求逆的方法。文中采用后者, 因此主要步骤包括:

(1) 遍历电网内全部支路(包括交流线和变压器), 形成节点导纳矩阵 Y ;

(2) 对 Y 矩阵进行求逆, 得到节点阻抗矩阵 Z ;

(3) 对任意节点对 i 和 j 的自阻抗: $Z_{ij,ij} = Z_{ii} + Z_{jj} - Z_{ij} - Z_{ji}$;

(4) 求取每个节点对自阻抗复数的幅值, 形成节点间电气距离矩阵。在不考虑移相器等因素的情况下, 该矩阵为对称阵, 且对角线数值为 0, 因此只需保存上三角矩阵的数值即可。

1.2 SNE 算法

电气距离存在于交流互联的任意两节点之间, 可以看作一个高维度的数据关系。高维数据不便于直接观察以发现其隐藏的特性, 随机近邻嵌入 (stochastic neighbor embedding, SNE) 算法^[16]采用非

收稿日期: 2017-11-15; 修回日期: 2017-12-19

基金项目: 国家电网有限公司科技项目 (XT71-15-047); 江苏省自然科学基金项目 (BK20160678)

线性降维的方法将高维的数据映射到三维或者二维空间,这样就能在立体图或者平面图中比较直观地展示数据的分布情况。

SNE 算法设计的原则是在映射的过程中保持数据点之间的距离关系,即在原始高维空间中距离较近(或较远)的点,映射到低维空间后对应点的距离也较近(或较远)。欧式距离是表示距离的一种常用方法,设 x_i, x_j 为原始高维空间中的任意两点, y_i, y_j 为其映射到低维空间中的点, $\|x_i - x_j\|^2$ 和 $\|y_i - y_j\|^2$ 分别表示高维和低维空间中的欧式距离。SNE 算法在欧式距离的基础上提出了一种条件概率来表示数据点的远近程度。在高维数据集 X 中该条件概率定义为:

$$p_{j|i} = \frac{\exp(-\|x_j - x_i\|^2/2\sigma_i^2)}{\sum_{k \neq i} \exp(-\|x_k - x_i\|^2/2\sigma_i^2)} \quad (1)$$

$p_{j|i}$ 表示点 x_j 出现在点 x_i 附近的条件概率,它采用高斯分布的形式, σ_i 为其对应的方差。 x_j 与 x_i 相隔越近, $p_{j|i}$ 越大,说明它们之间的相似度越高。设定 $p_{i|i} = 0$ 。

在低维数据集 Y 中也用类似的条件概率表示数据点的远近程度,其定义如下:

$$q_{j|i} = \frac{\exp(-\|y_j - y_i\|^2)}{\sum_{k \neq i} \exp(-\|y_k - y_i\|^2)} \quad (2)$$

$q_{j|i}$ 表示点 y_j 出现在点 y_i 附近的条件概率。其高斯分布的方差设定为 $1/\sqrt{2}$, 设定 $q_{i|i} = 0$ 。

这样对于任意的 i , 在高维空间和低维空间中分别形成一个条件概率分布 P_i 和 Q_i 。为了保持映射前后数据点之间的距离关系, P_i 和 Q_i 应尽量保持一致。KL 散度(Kullback-Leibler divergence)是衡量 2 个概率分布一致性的常用方法。所有的 i , P_i 和 Q_i 之间的 KL 散度为:

$$C = \sum_i D(P_i \| Q_i) = \sum_i \sum_j p_{j|i} \log \frac{p_{j|i}}{q_{j|i}} \quad (3)$$

C 即为 SNE 算法的代价函数,其值越小,分布越一致。可以通过梯度下降算法来求解 y_i 使得其值最小。 y_i 可以是二维或者三维坐标,求解出来后即可在用平面图或者立体图中标示出来。

1.3 t-SNE 算法

t-SNE 是 SNE 的改进算法,主要解决了 SNE 算法中的 2 个问题:不对称问题和拥挤问题。

(1) 不对称问题:SNE 算法定义的条件概率是不对称的,即 $p_{j|i} \neq p_{i|j}$ 和 $q_{j|i} \neq q_{i|j}$,按照常理,两点之间的距离或者相似度应该是唯一的,所以它们应该相等。为此,t-SNE 算法在式(1)和式(2)的基础

之上重新定义了 $p_{j|i}$ 和 $q_{j|i}$,并记为 p_{ij} 和 q_{ij} :

$$p_{ij} = \frac{p_{i|j} + p_{j|i}}{2n} \quad (4)$$

$$q_{ij} = \frac{\exp(-\|y_i - y_j\|^2)}{\sum_{k \neq i} \exp(-\|y_k - y_i\|^2)} \quad (5)$$

式中: n 为原始数据点的总数,这样定义的 p_{ij} 和 q_{ij} 均满足对称性。

(2) 拥挤问题:数据降维可视化的理想效果是把相似或者相近的点聚在一簇,不同类别的簇尽量分隔开。SNE 算法中高维空间和低维空间的条件概率均采用高斯分布,假设求解后分布大概一致,由于空间维度减小,低维空间中不同类别的簇势必也会挨着很近,造成拥挤。为了解决该问题,t-SNE 算法将低维空间中的条件概率分布换成了 t-分布:

$$q_{ij} = \frac{(1 + \|y_i - y_j\|^2) - 1}{\sum_{k \neq i} (1 + \|y_k - y_i\|^2) - 1} \quad (6)$$

t-分布与高斯分布相比,中心部位偏低,尾部偏高偏长。将式(4)和式(5)中的 p_{ij} 和 q_{ij} 代替到式(3)中,得到新的代价函数为:

$$C = \sum_i D(P_i \| Q_i) = \sum_i \sum_j p_{ij} \log \frac{p_{ij}}{q_{ij}} \quad (7)$$

要使式(7)中代价函数 C 更小,则要求 p_{ij} 与 q_{ij} 分布更一致,那么在高维空间中相近的点映射到低维空间中后点的距离更近,高维空间中隔得较远的点映射到低维空间中点的距离更远。

2 算法描述

本算法可分为以下 3 个主要步骤:

(1) 求取自阻抗矩阵。基于电网拓扑和运行参数,形成导纳矩阵 Y ;求逆得到阻抗矩阵 Z ;根据公式 $Z_{i,j} = Z_{ii} + Z_{jj} - Z_{ij} - Z_{ji}$ 形成节点对的自阻抗矩阵。该矩阵即可反映任意两节点间的电气距离。

(2) t-SNE 算法进行降维。以最小化电气距离空间和低维空间内条件概率的 KL 散度为目标,将低维空间内每个节点坐标作为待求参数进行迭代优化计算。当 KL 散度足够小或迭代次数到达上限时,迭代计算结束,输出低维空间节点坐标。

(3) 根据降维结果进行绘图。根据第 2 步的节点坐标,在二维或三维空间中进行绘图,通过图形可直观地感受到各个节点间的远近程度。本步骤可根据用户需要进行选择性绘制,例如可以只绘制电厂节点展示电厂间的紧密程度,或者只绘制某一区域或电压等级的节点等。整个电气距离可视化方法流程如图 1 所示。

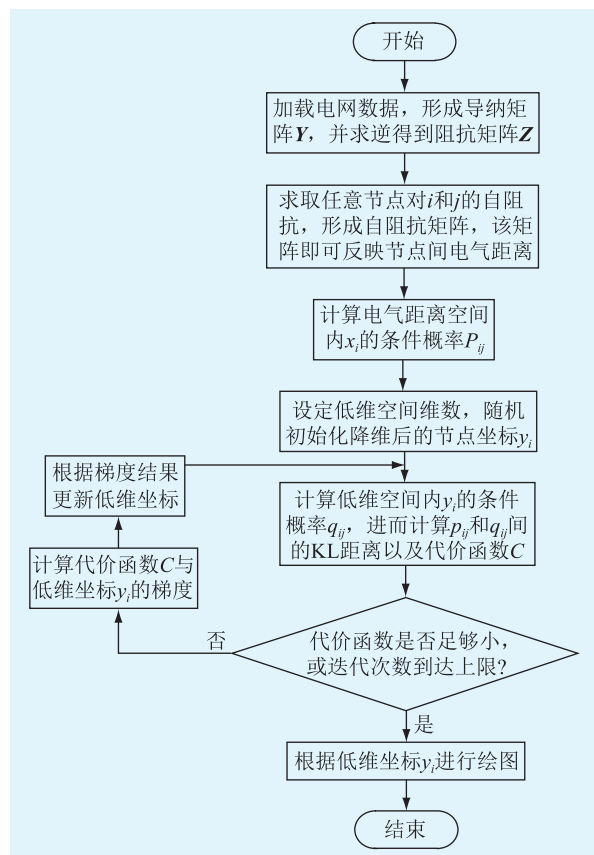


图1 基于 t-SNE 的电气距离可视化方法流程

Fig.1 Flow chart of visualization method for electrical distance based on t-SNE

3 算例分析

3.1 IEEE-118 算例

采用 IEEE-118 节点的标准测试系统进行算法验证,该系统含有 118 个节点,191 条交流线支路和 9 条变压器支路。假设系统处于全连接状态,首先可由 1.1 节所述方法得到节点间电气距离矩阵,此时节点 23 与 24 间电气距离为 0.046,远小于矩阵的平均值 0.217,说明相互间电气距离较近;之后经过 t-SNE 算法进行降维,可得到如图 2 所示的可视化结果,节点 23 与 24 均在红色方框之内,可见视觉距离与电气距离相符合;进一步采用 K 均值(K-means)聚类分析,可得到 7 个聚类子群,如表 1 所示,对应图 2 上所划分的 7 个子群,可见聚类结果与可视化效果也是基本一致的。图中的 X 轴和 Y 轴仅代表坐标,没有实际的物理含义。表 1 为原始 IEEE-118 系统的 K-means 分类结果。

进一步把节点 23 和节点 24 之间的支路断开,其余不变,此时节点 23 与 24 之间电气距离为 0.511,大于电气距离矩阵的平均值 0.241。经本算法分析可得到如图 3 的可视化结果,节点 23 与 24

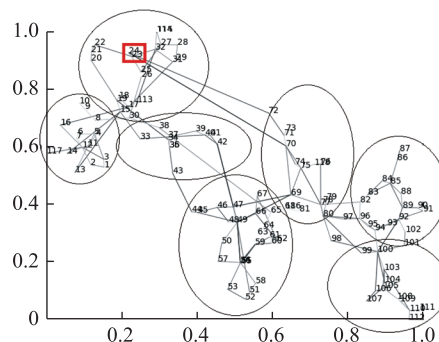


图2 原始 IEEE-118 系统可视化结果

Fig.2 Visualization result of original IEEE-118

表 1 原始 IEEE-118 系统的 K-means 分类结果

| 组号 | 所含母线编号 | 母线数量 |
|----|-------------------------|------|
| 1 | 1—14, 16, 117 | 16 |
| 2 | 15, 17—33, 113—115 | 21 |
| 3 | 34—43 | 10 |
| 4 | 44—64, 66, 67 | 23 |
| 5 | 65, 68—81, 98, 116, 118 | 18 |
| 6 | 82—97, 101, 102 | 18 |
| 7 | 99, 100, 103—112 | 12 |

分别在两个距离较远的红色方框内,对比图 2 可见,支路 23—24 运行时,两节点间电气距离较近,断开后则需要经过多条支路远距离互联,电气距离较远,两图的可视化结果很好地展示了这一变化。再次进行 K-means 分类可见,节点 23 和 24 分别被归入了不同的分组。此外,支路 23—24 断开后,电气距离矩阵平均值从 0.217 变为 0.241,说明系统总体的紧密程度有所下降。表 2 为断开 23—24 支路后的 K-means 分类结果。

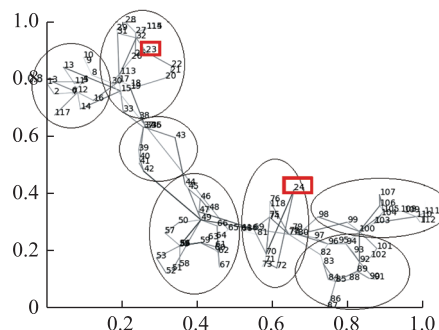


图3 断开 23—24 支路后 IEEE-118 系统可视化结果

Fig.3 Visualization result of IEEE-118 with the outage of line 23—24

3.2 两华电网算例

以国家电网公司某日在线计算数据为基础,验证本方法的有效性。当月华北—华中处于联网运

表2 断开 23—24 支路后的 K-means 分类结果

Tab. 2 Classification result of IEEE-118 with the outage of line 23—24 using K-means

| 组号 | 所含母线编号 | 母线数量 |
|----|-----------------------------|------|
| 1 | 1—14, 16, 117 | 16 |
| 2 | 15, 17—23, 25—33, 113—115 | 20 |
| 3 | 34—43 | 10 |
| 4 | 44—64, 66, 67 | 23 |
| 5 | 24, 65, 68—81, 98, 116, 118 | 19 |
| 6 | 82—97, 101, 102 | 18 |
| 7 | 99, 100, 103—112 | 12 |

行状态,因此在线数据中包含国调直调以及华北、华中所有 220 kV 以上的电网设备(注:所用在线数据中不包含蒙东电网数据)。图 4 展示了两华电网中所有厂站之间电气距离的可视化结果,明显呈现按省分群的特征,省内厂站间的电气距离通常小于省间,尤其是江西、重庆、湖南等群较为紧密;而四川、山东、京津冀的厂站数量较多,比较分散,占据了整个画面的较大的面积。上述可视化结果与两华电网特征相一致,验证了本算法的有效性。

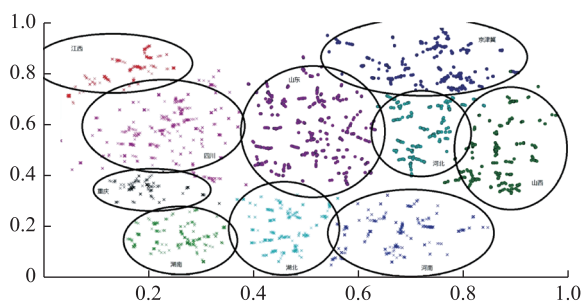


图 4 某日两华电网可视化结果

Fig.4 Visualization of north China and central China power grid

本算法虽然在分群上取得了较好的效果,但对于群与群之间的关系处理尚不完善,例如山东电网不应处于两华电网的中心地带。主要原因在于:低维空间坐标 y_i 刚开始是随机初始化的,虽然迭代过程中不断收缩变化,但最终的群所在位置受到了初始化结果的影响。可结合连接关系或设备区域等信息来对初始化过程进行改进,以取得更好的效果。

4 结语

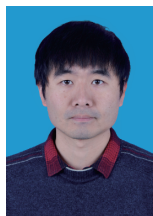
文中提出了一种基于 t-SNE 算法的电力系统电气距离可视化方法,并通过实际算例验证了方法的有效性。本方法可以生成以母线或厂站为单位的可视化图形,同时也可以以此为底图,叠加更多的电网运行或稳定分析数据,这样形成的图形会更加具有说服力。

参考文献:

- [1] 刘振亚. 特高压交直流电网[M]. 北京:中国电力出版社,2013.
LIU Zhenya. Ultra-high voltage AC & DC grid[M]. Beijing: China Electric Power Press, 2013.
- [2] 严剑峰,周孝信. 电力系统在线动态安全监测与预警技术[M]. 北京:中国电力出版社,2015.
YAN Jianfeng, ZHOU Xiaoxin. Power system online dynamic security monitoring and early-warning technology[M]. Beijing: China Electric Power Press, 2015.
- [3] 汤涌,王英涛,田芳,等. 大电网安全分析、预警及控制系统的研发[J]. 电网技术, 2012, 36(7):1-11.
TANG Yong, WANG Yingtao, TIAN Fang, et al. Research and development of stability analysis, early-warning and control system for huge power grids[J]. Power System Technology, 2012, 36(7):1-11.
- [4] 孙树明,谢昶,吕颖,等. 电力系统在线安全稳定分析应用模式[J]. 电网技术, 2015, 39(10):2875-2881.
SUN Shuming, XIE Chang, LYU Ying, et al. Power system online security and stability analysis application modes[J]. Power System Technology, 2015, 39(10):2875-2881.
- [5] 严剑峰,于之虹,田芳,等. 电力系统在线动态安全评估和预警系统[J]. 中国电机工程学报, 2008, 28(34):87-93.
YAN Jianfeng, YU Zhihong, TIAN Fang, et al. Dynamic security assessment & early warning system of power system[J]. Proceedings of the CSEE, 2008, 28(34):87-93.
- [6] 王大江,江叶峰,仇晨光,等. 江苏电网在线动态安全评估系统及应用研究[J]. 电力工程技术, 2017, 36(2):1-6.
WANG Dajiang, JIANG Yefeng, QIU Chenguang, et al. Research on the on-line dynamic security assessment system and application of Jiangsu power grid [J]. Electric Power Engineering Technology, 2017, 36(2):1-6.
- [7] 忽浩然,沈凤杰,汤伟,等. 新一代安控装置集中管理系统设计及工程应用[J]. 电力工程技术, 2017, 36(2):42-48.
HU Haoran, SHEN Fengjie, TANG Wei, et al. Design and application of new security and stability control management system[J]. Electric Power Engineering Technology, 2017, 36(2):42-48.
- [8] 李大勇,马冬雪,王晓宁,等. 电网信息可视化应用研究[J]. 电力系统保护与控制, 2009, 37(23):156-158.
LI Dayong, MA Dongxue, Wang Xiaoning, et al. Application of network information visualization[J]. Power System Protection and Control, 2009, 37(23):156-158.
- [9] 张菁,蒋愈勇,王化龙. 基于科学计算可视化技术的电网调度系统应用[J]. 计算机系统应用, 2010, 19(4):123-126.
ZHANG Jing, JIANG Yuyong, WANG Hualong. Application of power grid dispatching system based on visualization in scientific computing[J]. Computer Systems & Applications, 2010, 19(4):123-126.
- [10] 沈国辉,余东香,孙湃,等. 电力系统可视化技术研究及

- 应用[J]. 电网技术, 2009(17):31-36.
SHEN Guohui, SHE Dongxiang, SUN Pai, et al. Research and application of power system visualization technology[J]. Power System Technology, 2009(17):31-36.
- [11] 沈国辉, 刘金波, 陈光, 等. 特高压调度运行支持系统关键技术[J]. 电网技术, 2009(20):33-37.
SHEN Guohui, LIU Jinbo, CHEN Guang, et al. Key technologies for UHV dispatching operation support system[J]. Power System Technology, 2009(20):33-37.
- [12] 沈国辉, 李立新, 狄方春, 等. 特高压电网调度自动化系统的数据集成和可视化展示[J]. 电力系统自动化, 2009, 33(23):94-97.
SHEN Guohui, LI Lixin, DI Fangchun, et al. Data integration and visualization display for UHV power dispatching automation system[J]. Automations of Electric Power Systems, 2009, 33(23):94-97.
- [13] 赖晓文, 陈启鑫, 夏清, 等. 基于SVG技术的电力系统可视化平台集成与方法库开发[J]. 电力系统自动化, 2012, 36(16):76-81.
LAI Xiaowen, CHEN Qixin, XIA Qing, et al. Development of power system visualization platform and methods library based on SVG technology[J]. Automations of Electric Power Systems, 2012, 36(16):76-81.
- [14] 马韬韬, 郭创新, 曹一家, 等. 电网智能调度自动化系统研究现状及发展趋势[J]. 电力系统自动化, 2010, 34(9):7-11.
MA Taotao, GUO Chuangxin, CAO Yijia, et al. Current status and development trend of intelligent dispatching automation system of power system[J]. Automations of Electric Power Systems, 2010, 34(9):7-11.
- [15] LAURENS V D M, HINTON G. Visualizing data using t-SNE[J]. Journal of Machine Learning Research, 2008, 9(2605):2579-2605.
- [16] HINTON G, ROWEIS S. Stochastic neighbor embedding[J]. Advances in Neural Information Processing Systems, 2002, 41(4):833-840.
- [17] TANG Jian, LIU Jingzhou, ZHANG Ming, et al. Visualizing large-scale and high-dimensional data[C] // Proceedings of the 25th International Conference on World Wide Web, 2016:287-297.

作者简介:



史东宇

史东宇(1980—),男,硕士,高级工程师,从事电力系统在线安全稳定分析工作(E-mail:sdy@epri.sgcc.com.cn);

胡文强(1990—),男,硕士,工程师,从事计算机技术应用工作;

李刚(1984—),男,本科,工程师,从事计算机技术应用工作。

Study on Visualization Method of Electrical Distance in Power System Using t-SNE

SHI Dongyu¹, HU Wenqiang¹, LI Gang¹, YAN Jianfeng¹, ZANG Dechun²

(1. China Electric Power Research Institute, Beijing 100192, China;

2. Southeast University School of Electrical Engineering, Nanjing 211189, China)

Abstract: The AC / DC hybrid power grid has already been built in China, which uses Ultra-High Voltage (UHV) grid as its backbone. Operation modes of power grid become more complex and diverse. The requirement for online analysis and visualization of dispatching operation will be much higher. From the view of electrical distance, this paper proposes a new method to represent the closeness degree of power equipment, which maps the high-dimensional and complicated relations of electrical distances into 2D images by using t-SNE (t-distribution stochastic neighbor embedding) algorithm, and gives dispatcher more intuitive impressions. The validation is verified by taking the actual data of power grid as an example, indicating the method can represent the variety of grid rapidly and meet online application.

Key words: power system; electrical distance; visualization method; t-distribution stochastic neighbor embedding

(编辑 钱悦)