Simplicial flat-norm based structural validation of power distribution networks

Anonymous submission

Abstract

We explore the problem of comparing a pair of two dimensional geometries in this paper. A geometry can be a combination of multiple points or line segments on a Euclidean two dimensional plane. The usage of such comparison method can be understood through the following example - one of the geometries can be actual *ground truth*, while the other might be its synthesized *digital duplicate*. In this context, the comparison would serve as a metric of quality of the digital duplicate. This problem can be extended to comparing a pair of graphs embedded in a metric space and is a problem of interest in transportation network analysis.

Introduction

The increased adoption of rooftop solar photovoltaics (PVs) and electric vehicles (EVs) augmented with residential charging units has altered the energy consumption profile of an average consumer. The power distribution system, which was once considered as a passive entity in power grid planning and operation, is gaining attention for the challenges it poses in the present day. The requirement of an extensive dataset pertaining to power distribution network and residential consumer demand is vital for public policy researchers and power system engineers alike. Some examples of frameworks which synthesize extensive set of realistic power distribution networks include, but are not limited to (Li et al. 2020; Schweitzer et al. 2017; Mateo et al. 2020; Meyur et al. 2020).

An important aspect of creating such frameworks is validating the created synthetic power distribution to their actual physical counterpart. We require well-defined metrics to rank the frameworks which generate synthetic distribution networks, and thereby judge the strength and weakness of each framework. Though the prior works validate the statistical and power engineering attributes of the created networks, the structural aspect often gets neglected. In this context, (Meyur et al. 2022) has attempted to compare the network geometries using Hausdorff distance after partitioning the geographic region into small rectangular grids followed by comparing the geometries for each grid. However, the comparison is not robust to different resolution of the rectangular grids. Similar approach has been undertaken in (Brovelli et al. 2017) to compare a pair of road networks in a given geographic region, and suffers from the same drawback. This necessitates a well-defined distance metric between networks with geospatial embedding (Ahmed, Fasy, and Wenk 2014).

Approach

The terms *precision* and *recall* has been widely used in machine learning literature as performance metrics for classification and clustering problems. In general, *precision* denotes the fraction of estimated data which is relevant; while *recall* denotes the fraction of relevant data which is estimated. The term 'relevance' of estimated data can be defined based on its closeness to the actual ground truth. In our context, we can define as follows.

Definition 1 (Relevance)**.** *Two embedded graphs* $\mathcal{G}_0\left(\mathcal{V}_0, \mathcal{E}_0\right)$ *and* $\mathcal{G}_1\left(\mathcal{V}_1, \mathcal{E}_1\right)$ *are said to be 'δ-relevant' if* $d\left(\mathcal{G}_0, \mathcal{G}_1\right) \leq \delta$.

We define *precision* and *recall* in the context of synthetic distribution networks as follows. Consider actual distribution network $\mathcal{G}_{\text{act}}$ to be the ground truth and $\mathcal{G}_{\text{syn}}$ is a synthetic distribution network created by a framework.

Definition 2 (Precision)**.** *Given* n *networks sampled from the* ϵ*-neighborhood of* $\mathcal{G}_{syn}$*, and* k *of them being* δ*-relevant to* $\mathcal{G}_{act}$*, the precision is computed as* $\mathsf{pre}_{\mathsf{act,syn}} = \frac{k}{n}$.

Definition 3 (Recall)**.** *Given* m *networks sampled from the* ϵ*-neighborhood of* $\mathcal{G}_{act}$*, and* k *of them being* δ*-relevant to* $\mathcal{G}_{syn}$*, the recall is computed as* $\mathsf{rec}_{\mathsf{act,syn}} = \frac{k}{m}$.

Sampling from ϵ-neighborhood. Each edge in the graph is denoted by a straight line segment connecting a pair of nodes. We consider n vertices sampled from ϵ-neighborhood (radius) of each node in the graph. Each set of vertices can be joined as in the original graph to construct n network samples in the ϵ-neighborhood of the original graph. An example of sampling from ϵ-neighborhood of graph for $\epsilon = 0.1$ is shown in Fig. 1.

Measuring distance between networks. Need to define a metric to measure the area between a pair of network geometries. Simplicial flat-norm might be a way of addressing this problem.

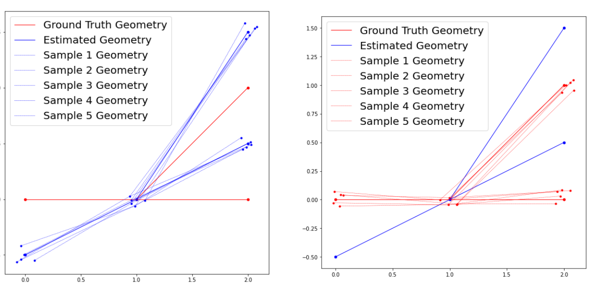

Figure 1: Sampling networks within ϵ-neighborhood of a graph for $\epsilon = 0.1$. Five network samples within ϵ-neighborhood of synthetic network are shown in the left plot. The right plot shows five network samples within the ϵ-neighborhood of the ground truth network.

References

Ahmed, M.; Fasy, B. T.; and Wenk, C. 2014. Local Persistent Homology Based Distance between Maps. In *Proceedings of the 22nd ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, SIGSPATIAL '14, 43–52. New York, NY, USA: Association for Computing Machinery. ISBN 9781450331319.

Brovelli, M. A.; Minghini, M.; Molinari, M.; and Mooney, P. 2017. Towards an Automated Comparison of OpenStreetMap with Authoritative Road Datasets. *Transactions in GIS*, 21(2): 191–206.

Li, H.; Wert, J.; Birchfield, A.; Overbye, T.; Roman, T.; Domingo, C. M.; Marcos, F.; Martinez, P. D.; Elgindy, T.; and Palmintier, B. 2020. Building Highly Detailed Synthetic Electric Grid Data Sets for Combined Transmission and Distribution Systems. *IEEE Open Access Journal of Power and Energy*, 7: 478–488.

Mateo, C.; Postigo, F.; de Cuadra, F.; Roman, T. G. S.; Elgindy, T.; Dueñas, P.; Hodge, B.; Krishnan, V.; and Palmintier, B. 2020. Building Large-Scale U.S. Synthetic Electric Distribution System Models. *IEEE Transactions on Smart Grid*, 11(6): 5301–5313.

Meyur, R.; Marathe, M.; Vullikanti, A.; Swarup, S.; Mortveit, H.; Centeno, V.; and Phadke, A. 2020. Creating Realistic Power Distribution Networks using Interdependent Road Infrastructure. In *IEEE Big Data Conference*.

Meyur, R.; Vullikanti, A.; Swarup, S.; Mortveit, H.; Centeno, V.; Phadke, A.; Poor, H. V.; and Marathe, M. 2022. Ensembles of Realistic Power Distribution Networks. Accepted with revisions.

Schweitzer, E.; Scaglione, A.; Monti, A.; and Pagani, G. A. 2017. Automated Generation Algorithm for Synthetic Medium Voltage Radial Distribution Systems. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, 7(2): 271–284.