Hessian Locally Linear Embedding of PMU Data for Efficient Fault Detection in Power Systems

Liu, Guohong; Li, Xiaomeng; Wang, Cong; Chen, Zhe; Chen, Ruonan; Qiu, Robert C.

Published in:
IEEE Transactions on Instrumentation and Measurement

Link to article, DOI:
10.1109/TIM.2022.3146905

Publication date:
2022

Document Version
Peer reviewed version

Hessian Locally Linear Embedding of PMU Data for Efficient Fault Detection in Power Systems

Guohong Liu, Xiaomeng Li, Cong Wang, *Member, IEEE*, Zhe Chen, *Student Member, IEEE*, Ruonan Chen, Robert C. Qiu, *Felllow, IEEE*

***Abstract*—This paper develops a computationally efficient fault-detection method for power systems, by exploiting phasor measurement unit (PMU) data and Hessian locally linear embedding (HLLE) technique. First, via HLLE technique, high-dimensional PMU data are transformed into low-dimensional embedding coordinates, which effectively captures the fluctuations of PMU data. Next, based on the feature space of embedding coordinates, T-squared statistic is employed for online fault detection. The method is evaluated on IEEE 39-bus system and a real-world system, exhibiting decent fault-detection performance as well as a considerably lower computational complexity when compared with existent methods.**

***Index Terms*—Power system, phasor measurement unit, fault detection, Hessian locally linear embedding, low complexity**

I. Introduction

PHASOR measurement units (PMUs) are widely deployed to provide high-accuracy measurement data in power systems [1]. With the PMU data, system operators can perform online fault detection. Up to today, a number of PMU-based online fault-detection methods have been proposed [2]–[7]. Among these methods, the majority, such as [4]–[7], are built upon principal component analysis (PCA). As they mathematically rely on linear dimensionality reduction, the fluctuations of PMU data cannot be effectively captured [8], [9]. In addition, they might suffer from performance degradation when PMU data get contaminated by strong noise or disturbance [8].

To effectively capture the fluctuations of PMU data, a method based on kernel principal component analysis (KPCA) is developed in [8]. By non-linearly mapping PMU data into some kernel space, the fault-detection performance is well improved. Furthermore, in order to avoid high-dimensional kernel-matrix constructions and decompositions, a Nystrom principal component analysis (NPCA) method is developed in [9]. Despite the technical merits of these methods, their high computational complexity is unfavorable for the online application.

This work was supported by the National Nature Science Foundation of China under Grant No. 62101210 and No. 61806085.

Guohong Liu, Xiaomeng Li, Cong Wang, and Ruonan Chen are with the College of Communication Engineering, Jilin University, Changchun, 130025, China (email: graceliu@jlu.edu.cn; lixm20@mails.jlu.edu.cn; wangcong2020@jlu.edu.cn; chenrn15@mails.jlu.edu.cn).

Zhe Chen is with Department of Electrical Engineering, Technical University of Denmark, 4000 Roskilde, Denmark (e-mail: zhech@elektro.dtu.dk).

Robert C. Qiu is with the Department of Electrical and Computer Engineering, Tennessee Technological University, Cookecille, 38505, USA (email: rcqiu@tntech.edu).

Corresponding author: Cong Wang.

Given the state of the art of PMU-based online fault-detection methods, this paper concentrates on low-complexity non-linear analysis of high-dimensional PMU data and develops an online fault-detection method by means of Hessian locally linear embedding (HLLE) technique [10]–[13]. In detail, by employing Euclidean distance to compute the k-nearest neighbors of every data sample, we first obtain the Hessian estimators for PMU data. Next, with reference to the feature space of embedding coordinates, the corresponding Hessian mapping matrix is determined. Finally, we derive the related T^2 statistic to identify whether there is an fault occurring in the power system. Compared with the recent KPCA-based and NPCA-based methods, our HLLE-based method has a considerably lower computational complexity while providing almost the same detection performance, and therefore enhances the practical applicability.

II. Main Method

A. Nonlinear Dimensionality Reduction via HLLE

We assume that there are M PMUs in the power system. These PMUs are placed according to practical applications and might not ensure full observability of the system.[1]

In a period of steady-state operation, we synchronously collect N fault-free points from all the PMUs and form a matrix $\mathbf{Y} = [\mathbf{y}_1, \cdots, \mathbf{y}_N]$. For this fault-free PMU data matrix, $\mathbf{y}_i \in \mathbb{R}^M$, $i = 1, \cdots, N$ represents the time-synchronized PMU sample of the system at one instant.

For each sample $\mathbf{y}_i$, we compute the k-nearest neighbors $\mathbf{z}_j^i, j = 1, \cdots, k$ from $\mathbf{y}_{i'}, i' \neq i$, by means of Euclidean distance. Then, form a re-centered matrix $\mathbf{M}^i$ as follows, which can mitigate the effect of PMU offset to certain extent.

$$\mathbf{M}^i = \left[\mathbf{z}_1^i - \frac{1}{k}\sum_{q=1}^{k}\mathbf{z}_q^i\ ,\ \cdots\ ,\ \mathbf{z}_k^i - \frac{1}{k}\sum_{q=1}^{k}\mathbf{z}_q^i\right]^{\mathbf{T}} \in \mathbb{R}^{k\times M} \tag{1}$$

After a singular-value-decomposition (SVD) on $\mathbf{M}^i$, we have

$$\mathbf{M}^i = \mathbf{U}\mathbf{D}\mathbf{V}^{\mathbf{T}} \tag{2}$$

where $\mathbf{U}$ and $\mathbf{V}$ respectively denote the left and right singular-vector matrices, and diagonal matrix $\mathbf{D}$ collects the singular values in decreasing order.

[1] A power system is fully observable if all its buses are observable. To make a bus observable, one PMU should be placed on either this bus or at least one of its neighboring buses [14].

By extracting the first d columns out of $\mathbf{U}$, we can construct a matrix $\mathbf{X}^i \in \mathbb{R}^{k\times(1+d+d(d+1)/2)}$ as follows

$$\mathbf{X}^i = \Big[\ \mathbf{1} \quad \mathbf{U} \quad \mathbf{U}(:,a)\circ\mathbf{U}(:,b)\ _{\{1\le a\le b\le d\}}\Big] \tag{3}$$

where $\circ$ is the Hadamard product. Further, through Gram-Schmidt orthogonalization, $\mathbf{X}^i$ is transformed into a new matrix $\bar{\mathbf{X}}^i$ of the same size, which gives an Hessian estimator as

$$\mathbf{H}^i = \Big(\bar{\mathbf{X}}^i(:, d+2:1+d+d(d+1)/2)\Big)^{\mathbf{T}}. \tag{4}$$

Once the Hessian estimators $\mathbf{H}^1,\cdots,\mathbf{H}^N$ have been computed for all $\mathbf{y}_1,\cdots\mathbf{y}_N$, we establish a quadratic form $\varpi_{i,j}$ for $i,j\in\{1,\cdots,k\}$ as follows

$$\varpi_{i,j} = \sum_{l=1}^{N}\sum_{r=1}^{d(d+1)/2}(\mathbf{H}^l)_{r,i}(\mathbf{H}^l)_{r,j}\ . \tag{5}$$

Finally, after wrapping up $\varpi_{i,j}$ into square matrix $\boldsymbol{\varpi}$, we perform an eigen-value computation $\boldsymbol{\varpi}\boldsymbol{\varphi} = \lambda\boldsymbol{\varphi}$ and arrange the eigen-values in descending order, which yields the low-dimensional embedding coordinates as shown below,

$$\mathbf{S} = \Big([\boldsymbol{\varphi}_1,\cdots,\boldsymbol{\varphi}_d]\mathbf{R}^{-1/2}\Big)^{\mathbf{T}} \in \mathbb{R}^{d\times k} \tag{6}$$

where $\mathbf{R}\in\mathbb{R}^{d\times d}$ with $(\mathbf{R})_{r,s} = \boldsymbol{\varphi}_r^{\mathbf{T}}\boldsymbol{\varphi}_s$. So far, the fault-free PMU data $\mathbf{Y}$ is non-linearly transformed into $\mathbf{S}$.

B. Fault-Detection Processes

In what follows, we develop an online fault-detection method for power systems. Specifically, when a new PMU sample $\hat{\mathbf{y}}$ arrives, we find its k-nearest neighbors $\hat{\mathbf{z}}_1,\cdots,\hat{\mathbf{z}}_k$ from $\mathbf{y}_1,\cdots,\mathbf{y}_N$ and construct a matrix $\hat{\mathbf{M}}$ as follows.

$$\hat{\mathbf{M}} = \left[\hat{\mathbf{z}}_1 - \tfrac{1}{k}\sum_{q=1}^{k}\hat{\mathbf{z}}_q\ ,\ \cdots\ ,\ \hat{\mathbf{z}}_k - \tfrac{1}{k}\sum_{q=1}^{k}\hat{\mathbf{z}}_q\right]^{\mathbf{T}} \in \mathbb{R}^{k\times M} \tag{7}$$

Then, our goal is to approximate $\mathbf{S}$ by $\Big(\hat{\mathbf{M}}[\mathbf{a}_1,\cdots,\mathbf{a}_d]\Big)^{\mathbf{T}}$ with $\mathbf{a}_i\in\mathbb{R}^M$ for all i. In the sense of least squares,we have

$$\mathbf{a}_i = \arg\min_{\mathbf{a}}\sum_{j=1}^{k}\left(\mathbf{a}^{\mathbf{T}}\left(\hat{\mathbf{z}}_j - \frac{1}{k}\sum_{q=1}^{k}\hat{\mathbf{z}}_q\right) - (\mathbf{S})_{i,j}\right)^2 \tag{8}$$

Using $\mathbf{a}_1,\cdots,\mathbf{a}_d$, the low-dimensional embedding coordinate $\hat{\mathbf{s}}$ of sample $\hat{\mathbf{y}}$ is expressed as

$$\hat{\mathbf{s}} = [\mathbf{a}_1,\cdots,\mathbf{a}_d]^{\mathbf{T}}\left(\hat{\mathbf{y}} - \frac{1}{k}\sum_{q=1}^{k}\hat{\mathbf{z}}_q\right) \tag{9}$$

and the T^2 statistic is derived as follows

$$T^2 = \hat{\mathbf{s}}^{\mathbf{T}}{\mathbf{R}_{\mathbf{S}}}^{-1}\hat{\mathbf{s}} \tag{10}$$

where $\mathbf{R}_{\mathbf{S}} = \mathbf{S}\mathbf{S}^{\mathbf{T}}/(k-1)$.

According to [15], the confidence limit of T^2 is given by

$$T_\alpha^2 = \frac{d(k-1)}{k-d}F(d,k-d,\alpha) \tag{11}$$

where F is F-distribution and α is a significance level. Therefore, for the power system, we report a fault when $T^2 > T_\alpha^2$.

C. Discussions for Practical Deployment

1) Computational Complexity: The computational complexity of our method comes from two major aspects: (i) offline transformation of fault-free PMU data $\mathbf{Y}$ into the embedding coordinates $\mathbf{S}$, and (ii) online calculation of the T^2 statistic that is associated with the new sample $\hat{\mathbf{y}}$.

For both aspects, the complexity is caused by a series of matrix multiplications, decompositions and inversions. As the matrix dimensions are determined by parameters d and k, the complexity is upper bounded by a cubic polynomial in d,k. As can be seen, by appropriately controlling the parameters d and k, the total computational complexity of our method can be lower than that of the previous KPCA-based and NPCA-based methods.

2) Fast Search for k-Nearest Neighbors: In high-dimensional big-data applications, there exist fast algorithms (e.g., [16], [17]) for finding the exact k-nearest neighbors. Technically, in our case, these algorithms first perform an offline clustering over the fault-free PMU data $\mathbf{Y}$. Then, online searches for the k-nearest neighbors are conducted via triangle inequality in only the clusters that are nearest to the specific sample. Here, the specific sample can be either a fault-free $\mathbf{y}_i$ in $\mathbf{Y}$ or the new sample $\hat{\mathbf{y}}$.

Compared to the traditional exhaustive search, these algorithms compute the Euclidean distance for only a small portion of samples, leading to considerably lower complexity.

3) PMU Missing Data Recovery: The PMU data loss can occur in practice, as a result of communication congestion, hardware failure, and transmission delay. If the power system is in steady state, then its PMU data are ambient, and hence can be effectively recovered by interpolation or low-rank matrix method (e.g., [18], [19]). In the cases where the power system is disturbed, the missing PMU can only be approximately recovered, via methods such as singular value decomposition [20].

III. Performance Evaluation

In this section, we numerically evaluate our HLLE-based method on IEEE 39-bus system and a real-world system. For comparison, the recent KPCA-based method [8] and NPCA-based method [9] are also evaluated. All implementations are based on a desktop with Intel(R) Core(TM) $\mathbf{i7-8550U}$ CPU@1.8 GHz and 16GB RAM.

A. Evaluation on IEEE 39-Bus System

Scenario 1. We simulate the system operation for a period of 160s, during which the voltage-phasor data are collected at 50Hz by PMUs. Upon the time instant of 140s, a single-phase-to-ground fault occurs to bus 8 and lasts for 100ms. To render bus 8 unobservable, we remove the PMU data with respect to buses 5, 7, 8, 9, leading to $M = 35$.

In this scenario, we take the following parameters. For HLLE-based method, the number of the nearest neighbors is $k = 16$, and the dimension of the embedding coordinates is $d = 4$. For KPCA-based method, the cumulative variance is 80%, and the order of polynomial kernel is 18. For NPCA-based method, the Nystrom approximation is performed on a

subset of 70 samples, and the cumulative variance together with kernel function are the same as those in KPCA-based method. Moreover, we use $N = 2000$ fault-free samples to construct the $\mathbf{Y}$ matrix.

Given the above parameters, we first evaluate the false-alarm rates under different confidence limits. To this end, white Gaussian noise (WGN) and colored Gaussian noise (CGN) are respectively added to the PMU data, mimicking the real case of industry-grade PMUs. For these noises, the signal-to-noise ratio (SNR) is set to 92dB [5]. As the confidence limit varies from 90% to 99% with an incremental step size of 1%, the variation of false-alarm rates is shown in Fig. 1. Clearly, both HLLE-based and KPCA-based methods achieve false-alarm rates lower than NPCA-based method. Apart from the false-alarm rates, the runtime for HLLE-based method is 1.63s, which is 83.62% less than KPCA-based method and 73.99% less than NPCA-based method.

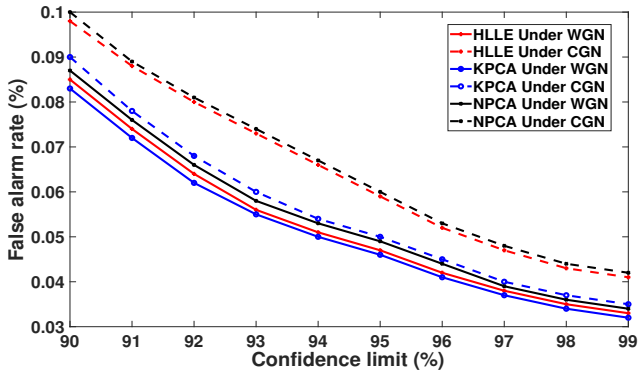

Fig. 1. False-alarm rates under different confidence limits.

Next, we change the SNRs of WGN and CGN and test whether the methods are sensitive to noise. For this purpose, we fix the confidence limit to 99%, and let the SNR vary from 75dB to 110dB with an incremental step size of 5dB. As shown in Fig. 2, HLLE-based method provides a fault-detection performance that is (i) better than NPCA-based method under all SNRs, and (ii) almost the same as KPCA-based method for SNR larger than 90dB.

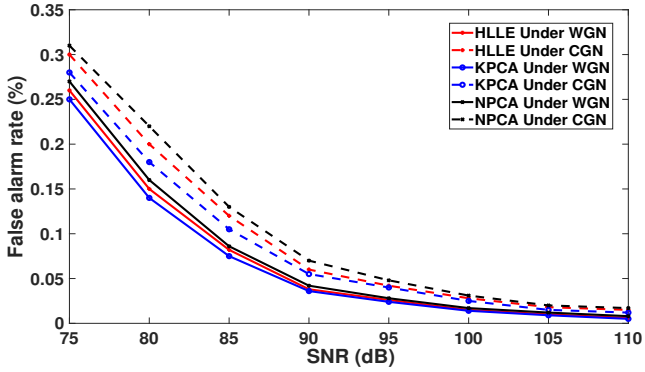

Fig. 2. False-alarm rates under different SNRs.

Scenario 2. Now, let us evaluate the methods by two faults that are specified in Table I. Physically, fault 1 is caused by three-phase short circuit, whereas fault 2 is incurred by two-phase short circuit and grounding.

In this scenario, parameters remain the same as before, except that (i) the order of polynomial kernel is 15 for KPCA-based and NPCA-based methods, (ii) the Nystrom approximation is performed on a subset of 80 samples for NPCA-based method, and that (iii) a constant confidence limit of 99% is adopted for fault detection.

TABLE I
DETAILS OF THE TWO FAULTS

	Fault 1	Fault 2
Instant of fault occurrence (s)	140	140
Instant of fault clearance (s)	140.12	140.10
Bus of fault	26	17
Buses without PMU	25, 26, 27, 28, 29	16, 17, 18, 27

With the faults in Table I, we evaluate the methods and show the results in Table II. As can be seen,

- the three methods all present zero detection-time deviation and hence correctly report the occurrence of the faults[2],
- the false-alarm rates are comparable for the three methods (though HLLE-based and KPCA-based methods are slightly better than NPCA-based method).

However, when observe the runtime, we find that HLLE-based method is 81.40% less than KPCA-based method and 71.36% less than NPCA-based method, which implies that HLLE-based method can be deployed for online applications.

TABLE II
COMPARISON OF THE THREE METHODS WITH THE TWO FAULTS

	HLLE	KPCA	NPCA
False-alarm rate for fault 1 (%)	0.032	0.032	0.034
Detection-time deviation of fault 1 (s)	0	0	0
False-alarm rate for fault 2 (%)	0.034	0.033	0.036
Detection-time deviation of fault 2 (s)	0	0	0
Average runtime (s)	1.67	8.98	5.83
Reduction of runtime (%)	N/A	81.40	71.36

B. Evaluation on a Real-World Power System

We consider a real-world power system that is controlled by the State Grid Cooperation of China [9]. In this system, we

- collect voltage-phasor data at 50Hz from $M = 14$ PMUs, rendering a large portion of the system unobservable;
- assume no prior knowledge of the system topology for confidentiality of the interconnected areas.

During an afternoon, two typical chain-reaction faults occur at 13:49:55 and 13:49:56, as shown in Fig. 3. To detect these faults, the parameters are determined as follows. For HLLE-based method, the number of nearest neighbors is $k = 12$

[2]Note that the detection-time deviation is reported as a relative value around the instant of fault occurrence. A zero detection-time deviation means that the method correctly reports an occurrence of the fault.

and the dimension of the embedding coordinates is $d = 3$. For KPCA-based method, the cumulative variance is set to 80%, and the order of polynomial kernel is 15. For NPCA-based method, the Nystrom approximation is performed on a subset of 60 samples, and the cumulative variance together with kernel function are the same as those in KPCA-based method. Moreover, we use $N = 2500$ fault-free samples to construct the $\mathbf{Y}$ matrix and set the confidence limit to 99%.

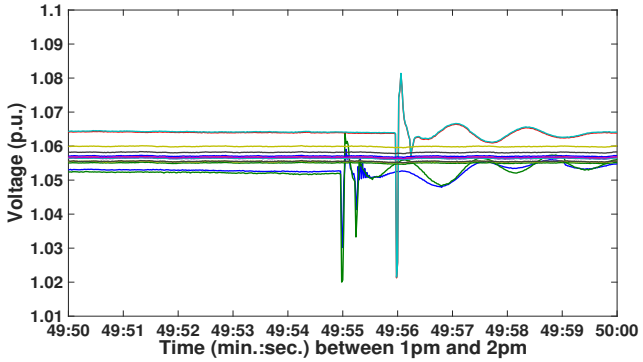

Fig. 3. The collected PMU data that cover two typical chain-reaction faults.

In Table III, we present the fault-detection performance. By observation, the three methods achieve the same false-alarm rate and zero detection-time deviation for fault 1. However, for fault 2, only our HLLE-based method provides zero detection-time deviation[3]. With respect to the runtime, the HLLE-based method achieves reduction of 80.28% and 69.74%, when compared with its two competitors.

TABLE III
COMPARISON OF THE THREE METHODS FOR REAL-WORLD CHAIN-REACTION FAULTS

	HLLE	KPCA	NPCA
False-alarm rate (%)	0.002	0.002	0.002
Detection-time deviation of fault 1 (s)	0	0	0
Detection-time deviation of fault 2 (s)	0	-0.02	-0.02
Runtime (s)	3.15	15.97	10.41
Reduction of runtime (%)	N/A	80.28	69.74

IV. CONCLUSION

This paper has exploited HLLE technique to achieve efficient dimensionality reduction of the PMU data, and has further developed a method for low-complexity fault detection in power systems. The method has been evaluated via both simulated and real-world PMU data. When compared with the typical KPCA-based and NPCA-based methods, the proposed HLLE-based method exhibits a considerably lower computational complexity while providing almost the same detection performance. The future studies refer to realizing the fault localization, by mining the relation between the T^2 statistic and PMU placement.

[3]A negative detection-time deviation means that the method incorrectly reports a later occurrence of the fault.

REFERENCES

[1] L. Chen, W. Zhao, F. Wang, and S. Huang, "Harmonic phasor estimator for p-class phasor measurement units," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 4, pp. 1556–1565, 2020.

[2] M. Jamei, R. Ramakrishna, T. Tesfay, R. Gentz, C. Roberts, A. Scaglione, and S. Peisert, "Phasor measurement units optimal placement and performance limits for fault localization," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 1, pp. 180–192, 2020.

[3] P. Tehrani and M. Levorato, "Frequency-based multi task learning with attention mechanism for fault detection in power systems," in *IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids*, 2020, doi: 10.1109/SmartGridComm47815.2020.9302968.

[4] L. Souto, J. Melendez, and S. Herraiz, "Comparison of principal component analysis techniques for pmu data event detection," in *IEEE Power and Energy Society General Meeting*, 2020, doi: 10.1109/PESGM41954.2020.9281512.

[5] L. Xie, Y. Chen, and P. R. Kumar, "Dimensionality reduction of synchrophasor data for early event detection: Linearized analysis," *IEEE Transactions on Power Systems*, vol. 29, no. 6, pp. 2784–2794, 2014.

[6] M. Rafferty, X. Liu, D. M. Laverty, and S. McLoone, "Real-time multiple event detection and classification using moving window PCA," *IEEE Transactions on Smart Grid*, vol. 7, no. 5, pp. 2537–2548, 2016.

[7] M. Grbovic, W. Li, P. Xu, A. K. Usadi, L. Song, and S. Vucetic, "Decentralized fault detection and diagnosis via sparse PCA based decomposition and maximum entropy decision fusion," *Journal of Process Control*, vol. 22, no. 4, pp. 738–750, 2012.

[8] X. Liu, J. M. Kennedy, D. M. Laverty, D. J. Morrow, and S. McLoone, "Wide-area phase-angle measurements for islanding detection: An adaptive nonlinear approach," *IEEE Transactions on Power Delivery*, vol. 31, no. 4, pp. 1901–1911, 2016.

[9] G. Liu, H. Chen, X. Sun, N. Quan, L. Wan, and R. Chen, "Low-complexity nonlinear analysis of synchrophasor measurements for events detection and localization," *IEEE Access*, vol. 6, pp. 4982–4993, 2018.

[10] S. Zhang, Z. Ma, and H. Tan, "On the equivalence of HLLE and LTSA," *IEEE Transactions on Cybernetics*, vol. 48, no. 2, pp. 742–753, 2018.

[11] H. Rajaguru and S. K. Prabhakar, "Performance analysis of local linear embedding (LLE) and hessian LLE with hybrid ABC-PSO for epilepsy classification from EEG signals," in *IEEE International Conference on Inventive Research in Computing Applications*, 2018, pp. 1084–1088.

[12] D. L. Donoho and C. Grimes, "Hessian eigenmaps: Locally linear embedding techniques for high-dimensional data," *Proceedings of the National Academy of Sciences*, vol. 100, no. 10, pp. 5591–5596, 2003.

[13] X. Lu, H. Wang, R. Zhou, and B. Ge, "Using hessian locally linear embedding for autonomic failure prediction," in *2009 World Congress on Nature & Biologically Inspired Computing (NaBIC)*. IEEE, 2009, pp. 772–776.

[14] E. Abiri, F. Rashidi, and T. Niknam, "An optimal PMU placement method for power system observability under various contingencies," *International Transactions on Electrical Energy Systems*, vol. 25, no. 4, pp. 589–606, 2015.

[15] H. Hotelling, "The generalization of Student's ratio," *The Annals of Mathematical Statistics*, vol. 2, no. 3, pp. 360–378, 1931.

[16] A. Almalawi, A. Fahad, Z. Tari, M. Cheema, and I. Khalil, "kNNVWC: An efficient k-nearest neighbors approach based on various-widths clustering," *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 1, pp. 68–81, 2016.

[17] Y. Pan, Z. Pan, Y. Wang, and W. Wang, "A new fast search algorithm for exact k-nearest neighbors based on optimal triangle-inequality-based check strategy," *Journal of Knowledge-Based Systems*, vol. 189, no. 1, pp. 1–11, 2020.

[18] K. Jones, A. Pal, and J. Thorp, "Methodology for performing synchrophasor data conditioning and validation," *IEEE Transactions on Power Systems*, vol. 30, no. 3, pp. 1121–1130, 2015.

[19] Y. Hao, M. Wang, J. Chow, and M. Patel, "Model-less data quality improvement of streaming synchrophasor measurements by exploiting the low-rank hankel structure," *IEEE Transactions on Power Systems*, vol. 33, no. 6, pp. 6966–6977, 2018.

[20] Z. Yang, H. Liu, T. Bi, Z. Li, and Q. Yang, "An adaptive PMU missing data recovery method," *Journal of Electrical Power and Energy Systems*, vol. 116, no. 1, pp. 1–14, 2020.