Power System Data Quality and Privacy Enhancement

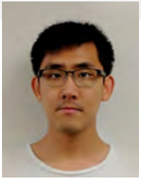
Meng Wang

Assistant Professor
Department of Electrical, Computer & Systems Engineering
Rensselaer Polytechnic Institute (RPI)
Troy, NY, USA

2019 Future Energy Systems Technology Conference
April 10, 2019

Acknowledgment

Dr. Yingshuai Hao



Dr. Pengzhi Gao

Shuai Zhang

Ren Wang

Prof. Joe H. Chow

Big Data in Power Systems

- Phasor Measurement Units (PMUs)
  - PMUs provide synchronized phasor measurements at a sampling rate of 30 or 60 samples per second.
  - Multi-channel PMUs can measure bus voltage phasors, line current phasors, and frequency. 2000+ PMUs in the North America.
  - Data availability and quality issues, e.g., data losses due to communication congestions.
  - Limited incorporation into the real-time operations.
- Smart Meters
  - 90% of power outages and disturbances are rooted in distribution networks. SCADA measurements are available only at the substation level.
  - Smart meters provide fine-grained measurements of power consumptions of customers and enhance the distribution system visibility.

Low Dimensionality of PMU data

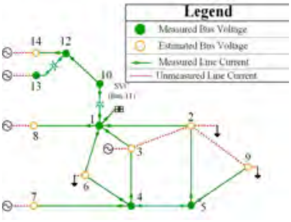


Figure: PMUs in Central NY Power Systems

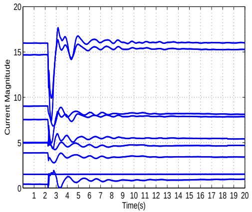


Figure: Current magnitudes of PMU data

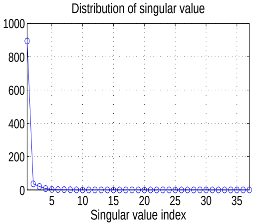


Figure: Singular values of the PMU data matrix

- 6 PMUs measure 37 voltage/current phasors. 30 samples/second for 20 seconds.
- Singular values decay significantly. Mostly close to zero. Singular values can be approximated by a sparse vector.
- Low-dimensionality also used in Chen, Xie, Kumar 2013, Dahal, King, Madani 2012 for dimensionality reduction.

Data-driven Approaches by Exploiting Low-Dimensional Structures

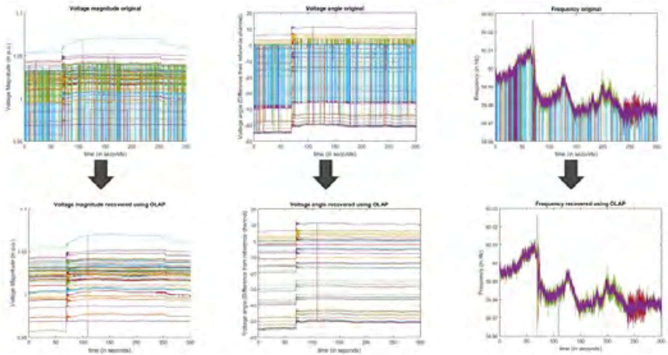
Objective: Develop computationally efficient data-driven methods for power system situational awareness.

- PMU data quality improvement: missing data recovery, bad data correction, and detection of cyber data attacks.
- Data clustering and pattern extraction from privacy-preserving measurements.

Outline

PMU Data Quality Issues

- Data losses and errors resulting from communication congestions and device malfunction.
- California Independent System Operator reported that 10%-17% of data in 2011 had availability and quality issues.
- Reliable data needed for real-time situational awareness and control.



Simultaneous and Consecutive Data Losses

A recorded PMU dataset: consecutive data losses on three phases of line for an hour.

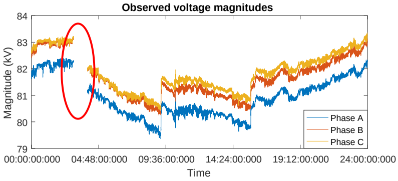


Figure: Measured voltage phasor magnitudes

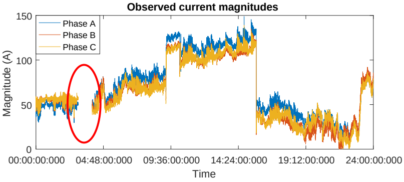


Figure: Measured current phasor magnitudes

Our Contribution

Our developed data-driven data recovery and error correction methods

- Recover/correct data losses/errors including simultaneous and consecutive data losses/errors.
- Differentiate bad data from system events.
- No modeling of power system dynamics is needed.
- First-order algorithms to solve nonconvex optimization problems with provable global optimality.

Zhang, Hao, Wang, Chow. *IEEE Journal of Selected Topics on Signal Processing*, 2018.

Hao, Wang, Chow, Farantatos, Patel, *IEEE Transactions on Power Systems*, 2018.

Zhang, Wang. *IEEE International Symposium on Information Theory*, 2018.

Zhang, Wang. *under revision with IEEE Transactions on Signal Processing*, 2018.

Low-rank Hankel Structure of PMU Data

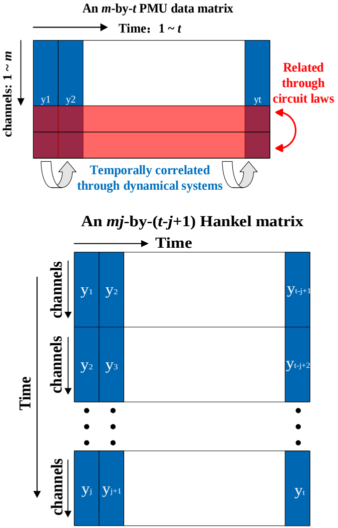
Observation matrix:

$$\mathbf{Y} = [\mathbf{y}_1, \ \mathbf{y}_2, \ \cdots, \ \mathbf{y}_n] \in \mathbb{C}^{n_c \times n}$$

Hankel structure:

$$\mathcal{H}_\kappa(\mathbf{Y}) = \begin{bmatrix} \mathbf{y}_1 & \mathbf{y}_2 & \cdots & \mathbf{y}_{n-\kappa+1} \\ \mathbf{y}_2 & \mathbf{y}_3 & \cdots & \mathbf{y}_{n-\kappa+2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{y}_\kappa & \mathbf{y}_{\kappa+1} & \cdots & \mathbf{y}_n \end{bmatrix}.$$

$\mathcal{H}_\kappa(\mathbf{Y}) \in \mathbb{C}^{\kappa n_c \times (n-\kappa+1)}$ can still be approximated by a low-rank matrix.



The low-rank Hankel property results from the reduced-order dynamical system.

Low-rank Hankel Structure of PMU Data

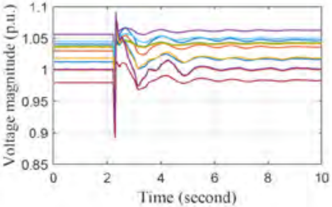


Figure: Measurements that contain a disturbance

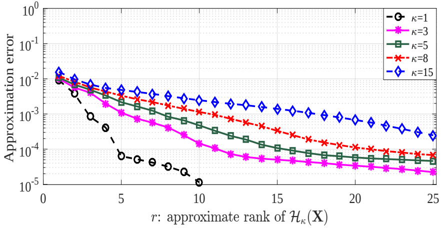


Figure: The low-rank approximation errors to $\mathcal{H}_\kappa(\mathbf{Y})$

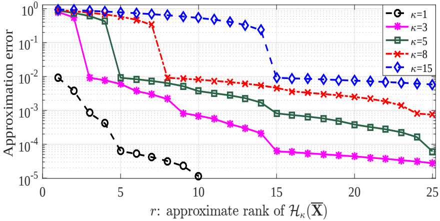


Figure: The low-rank approximation errors to $\mathcal{H}_\kappa(\overline{\mathbf{Y}})$, where $\overline{\mathbf{Y}}$ is a column permutation of $\mathbf{Y}$.

Robust Data Recovery

Let $\mathbf{M} = \mathbf{Y} + \mathbf{S}$ denote the partially corrupted measurements, where $\mathbf{S}$ denotes the sparse errors.
The robust data recovery problem is formulated as

$$\begin{aligned} \min_{\mathbf{X},\mathbf{S}\in\mathbb{C}^{n_c\times n}} & \|\mathcal{P}_\Omega(\mathbf{X}+\mathbf{S}-\mathbf{M})\|_F^2 \\ \text{subject to } & \text{rank}(\mathcal{H}_\kappa(\mathbf{X})) = r, \|\mathbf{S}\|_0 \leq s. \end{aligned} \tag{1}$$

Our proposed alternating projection algorithm

Initialization: $\mathbf{X}_0 = \mathbf{0}$, thresholding ε_0;
Two stages of iterations:

- In the k-th outer iteration:
  - Increase the desired rank k from 1 to r gradually;
- In the l-th inner iteration:
  - Update $\mathbf{S}_l$ based on the current estimated thresholding ξ_l;
  - Update $\mathbf{X}_l$ along the gradient descent direction $\mathcal{P}_\Omega(\mathbf{X}_l + \mathbf{S}_l - \mathbf{M})$;
  - Project the Hankel matrix $\mathcal{H}_\kappa \mathbf{X}_l$ into the rank-k matrix set;
  - Obtain $\mathbf{X}_{l+1}$ from the matrix after projection;
  - Update ξ_{l+1} based on $\mathbf{X}_{l+1}$.

Theoretical results

- Required number of observations: $\mathcal{O}(r^3 \log^2(n))$, less than the bound $\mathcal{O}(nr \log^2(n))$ of recovery with convex relaxation approach;
- Fraction of corruptions it can correct: $\mathcal{O}\left(\frac{1}{r}\right)$ in each row;
- Low computational complexity: $\mathcal{O}(rn_c n \log n \log(1/\epsilon))$, in comparison, a convex alternative takes $\mathcal{O}(n_c n^3/\epsilon)$
- Recovery guarantees on simultaneous data losses and corruptions across all channels.

Zhang, Wang. *IEEE International Symposium on Information Theory* 2018.

Zhang, Wang. *To appear in IEEE Transactions on Signal Processing*, 2019.

Numerical experiments

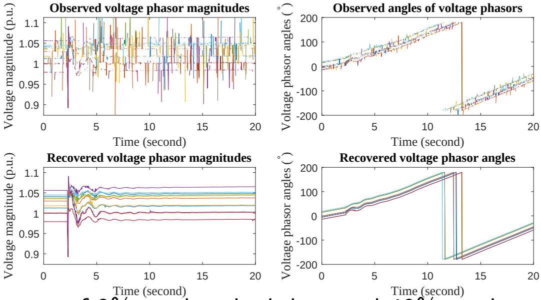


Figure: One case of 8% random bad data and 40% random missing data

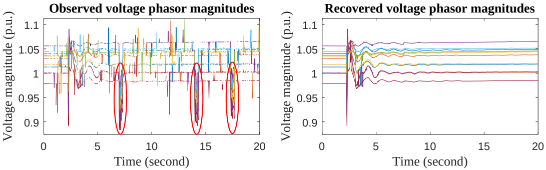


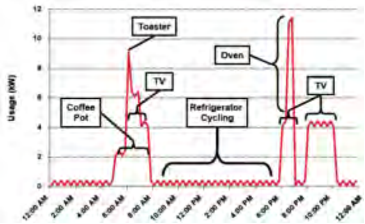
Figure: Consecutive bad data, 3% random bad data and 20% missing data

Hao, Wang, et al. IEEE Trans. Power Systems, 2018

Outline

Privacy Concerns

- Smart meter data
  - Non-intrusive load monitoring approaches can identify individual appliances from the household total demand [Hart et al., 1992].
  - User behaviors and habits can be extracted [Lisovich et al., 2010].
- The operator clusters customers with similar load patterns to enhance the load forecasting accuracy, design incentives for demand response, and identify abnormal user patterns.



Data Access and Privacy Issues Related to Smart Grid Technologies, by Megan Hertzler

Tradeoff Between Privacy and Accuracy

- Data privacy preserving approaches for smart meter data:
  - Aggregating the data of co-located customers [Li et al., 2011].
  - Adding noise to the measurement through signal processing approaches [Pedro et al., 2014].
  - Physically adding rechargeable batteries to the households [Stephen et al., 2011].

Privacy-preserving Measurements ⇒ Inaccurate Information for the Operator

Tradeoff Between Privacy and Accuracy

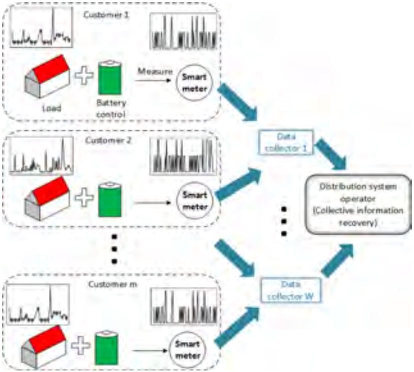
- Data privacy preserving approaches for smart meter data:
  - Aggregating the data of co-located customers [Li et al., 2011].
  - Adding noise to the measurement through signal processing approaches [Pedro et al., 2014].
  - Physically adding rechargeable batteries to the households [Stephen et al., 2011].

Privacy-preserving Measurements ⇒ Inaccurate Information for the Operator

We can achieve enhanced data privacy, reduced data communication, and accurate information recovery simultaneously!

Our Approach: Simultaneous Achievement of Data Privacy and Information Accuracy

- Quantize the power consumption to one of a few levels to hide information using a probability distribution depending on the actual power consumption.



Our Approach: Simultaneous Achievement of Data Privacy and Information Accuracy

Question: how can the operator recover the data from the quantized measurements and cluster them into the right group?

Our Approach: Simultaneous Achievement of Data Privacy and Information Accuracy

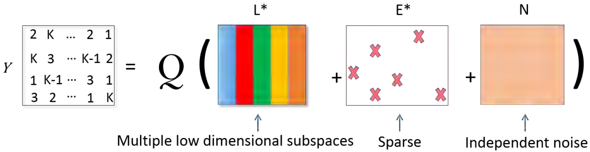
Question: how can the operator recover the data from the quantized measurements and cluster them into the right group?

- We propose a data recovery and clustering method for the operator.
- Our approach provides accurate results with a sufficient number of measurements.→ The operator has the correct information, but a cyber intruder with partial measurements does not.

Gao, Wang, Wang, and Chow, *IEEE Transactions on Signal Processing*, 2018

Wang, Wang, and Xiong, *IEEE Journal of Sel. Topics in Signal Process., Special Issue on Robust Subspace Learning and Tracking: Theory, Algorithms, and Appl.*, 2018.

Problem Formulation



- Each subspace has the dimension d. The rank of L^* is r $(r \leq pd)$.
- $E^* \in \mathbb{R}^{m \times n}$: At most s nonzero entries.
- $N \in \mathbb{R}^{m \times n}$: i.i.d. noise with known cdf $\Phi(z)$.
- $\|L^*\|_\infty \leq \alpha_1$ and $\|E^*\|_\infty \leq \alpha_2$ for some constants α_1, α_2.

$$Y_{ij} = \mathcal{Q}(L^*_{ij} + E^*_{ij} + N_{ij}), \ \forall (i,j). \tag{2}$$

Given K-level quantization boundaries $\omega_0 < \omega_1 < \ldots < \omega_K$,

$$\mathcal{Q}(x) = l \ \text{ if } \omega_{l-1} < x \leq \omega_l, \ \ l \in [K]. \tag{3}$$

Problem Formulation

One can check that

$$Y_{ij} = l \text{ with probability } f_l(X^*_{ij}),\ \forall (i,j),\ X^*_{ij} = L^*_{ij} + E^*_{ij} \quad (4)$$

where $\sum_{l=1}^{K} f_l(X^*_{ij}) = 1$, and

$$f_l(X^*_{ij}) = P(Y_{ij} = l | X^*_{ij}) = \Phi(\omega_l - X^*_{ij}) - \Phi(\omega_{l-1} - X^*_{ij}). \quad (5)$$

How can we estimate L^*, E^*, and C^* given Y and Φ?

Proposed Approach

Simultaneously recover and cluster the data by solving a constrained maximum likelihood problem

- We estimate (L^*, E^*, C^*) by $(\hat{L}, \hat{E}, \hat{C})$, where

$$(\hat{L}, \hat{E}, \hat{C}) = \arg\min_{L,E,C} -\sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{l=1}^{K} \mathbf{1}_{[Y_{ij}=l]} \log(f_l(L_{ij} + E_{ij})), \quad \text{s.t. } (L, E, C) \in \mathcal{S}_f, \tag{6}$$

and the feasible set $\mathcal{S}_f$ is defined as

$$\mathcal{S}_f = \{(L, E, C) : L = LC, \operatorname{rank}(L) \leq r, \|L\|_\infty \leq \alpha_1, \|E\|_\infty \leq \alpha_2, \|E\|_0 \leq s, \|c_i\|_0 \leq d, C_{ii} = 0, \forall i \in [n]\}. \tag{7}$$

- Apply spectral clustering on $\hat{C}$.
- Problem (6) is nonconvex due to the nonconvexity of $\mathcal{S}_f$.

Recovery and Clustering Results for Multiple Subspaces

- Theorem 1[1]: If columns of $\hat{L}$ belong to p subspaces, each of which has dimension smaller or equal to d, then

$$\frac{\|(\hat{L}+\hat{E})-(L^*+E^*)\|_F}{\sqrt{mn}} \leq O(\sqrt{\frac{d}{m}}) \text{ and } \frac{\|\hat{L}-L^*\|_F}{\sqrt{mn}} \leq O(\sqrt{\frac{d}{m}}), \quad (8)$$

- Theorem 2: Consider any algorithm that, for any $L+E \in \mathcal{S}_f$, takes $Y = L+E+N$ as the input and returns $\hat{L}+\hat{E}$. Then there exists $L+E \in \mathcal{S}_f$ such that with probability at least $\frac{3}{4}$,

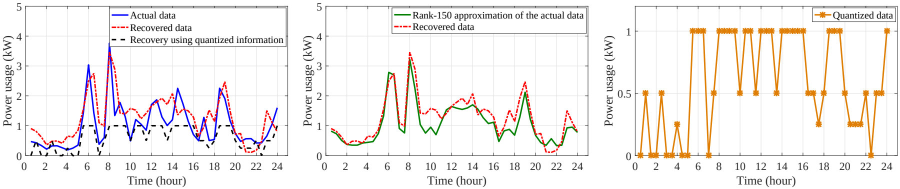
$$\frac{\|(\hat{L}+\hat{E})-(L^*+E^*)\|_F}{\sqrt{mn}} \geq O(\sqrt{\frac{d}{m}}) \text{ and } \frac{\|\hat{L}-L^*\|_F}{\sqrt{mn}} \geq O(\sqrt{\frac{d}{m}}), \quad (9)$$

- Theorem 3: The global minimizer $\hat{C}$ of (6) has subspace-preserving property of $\hat{L}$.

[1] Wang, Wang, and Xiong, *IEEE Journal of Sel. Topics in Signal Process., Special Issue on Robust Subspace Learning and Tracking: Theory, Algorithms, and Appl.*, 2018.

Simulation on Smart Meter Data

- Irish smart meter trial consists more than 5000 residential customers [2]. The power usage was measured in kW in every 30 minutes.
- $m = 1440$ (July 14 - August 12, 2009), $n = 1448$, $r = 150$, $\|L^*\|_\infty = 6$
- The entries of the noise matrix N are drawn i.i.d. from $\mathcal{N}(0, 0.3^2)$.
- Sparse matrix: $E^* \in R^{m \times n}$ (Nonzero entries have random locations and are uniformly selected from $[-0.5, -6]$ and $[0.5, 6]$). Average corruption rate is 5%
- Level-K=5: $\omega_0 = -\infty$, $\omega_1 = 0.25$kW, $\omega_2 = 0.5$kW, $\omega_3 = 1$kW, $\omega_4 = 3$kW, and $\omega_5 = \infty$



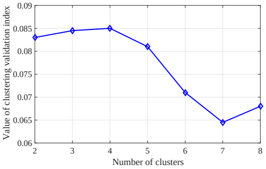
[2] Commission for Energy Regulation Smart Metering Project

Simulation on Smart Meter Data (Multiple Classes)

- Clustering index: Let a_j denote the angle of the data point x_j ($j \in [N]$) to the subspace of its own group. Let b_j be the minimum angle of x_j to the subspaces of other groups.

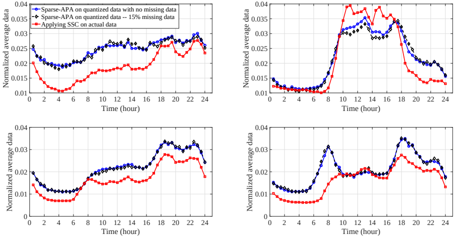
$$s_j = \frac{b_j - a_j}{\max(a_j, b_j)}, Index = \frac{1}{N}\sum_{j=1}^{N} s_j$$

- Clustering Validation
  - Use data from August 17 to August 23 in 2009 as validation dataset.
  - Compare the clustering validation index under different numbers of clusters using Sparse Subspace Clustering (SSC).



Simulation on Smart Meter Data (Multiple Classes)

- $m = 1440$ (July 14 - August 12, 2009), $n = 4780$, $d = 50$, and $k = 4$. All other parameters are set to be the same with one class case.
- Mean daily profiles are obtained by first normalizing data ($\|L_i\|_2 = 1$), and then averaging in the same group.



Mean daily profiles by (a) using our method with no missing data (b) using our method with 15% missing data (c) applying SSC on actual data

	Quantized Clustering	SSC on original data	SSC on recovered data	SSC on quantized data	Random selection
Clustering Index	0.082	0.085	0.073	0.06	0.051

Conclusions

- A framework of power system data analytics by exploiting the low-dimensional structure of spatial-temporal data blocks.
- Data quality improvement with analytical guarantees. (Missing data recovery, detection of cyber data attacks.)
- A new approach to enhance the data privacy and reduce the communication burden without too much information loss.
- Other work: real-time event identification approach using a small number of recorded single events for training.

Q & A

References

Candes, Emmanuel, Benjamin Recht (2012)
Exact Matrix Completion via Convex Optimization
Communications of the ACM, 111 – 119.

Fazel M (2002)
Matrix Rank Minimization with Applications
PhD thesis, Stanford University

Mateos G, Giannakis G B (2013)
Load Curve Data Cleansing and Imputation via Sparsity and Low Rank
IEEE Transactions on Smart Grid, 2347 – 2355.

Kekatos V, Zhang Y, Giannakis G B (2014)
Electricity Market Forecasting via Low-rank Multi-kernel Learning
IEEE Journal of Selected Topics in Signal Processing, 1182 – 1193.

Liu Y, Ning P, Reiter M K (2011)
False Data Injection Attacks Against State Estimation in Electric Power Grids
ACM Transactions on Information and System Security (TISSEC), 13.

References

Xie L, Mo Y, Sinopoli B (2010)
False Data Injection Attacks in Electricity Markets
Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on, 226 – 231.

Kosut O, Jia L, Thomas R J, et al (2010)
Malicious Data Attacks on Smart Grid State Estimation: Attack Strategies and Countermeasures
Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on, 220 – 225.

Sedghi H, Jonckheere E (2013)
Statistical Structure Learning of Smart Grid for Detection of False Data Injection
Power and Energy Society General Meeting (PES), 1 – 5.

Liu L, Esmalifalak M, Ding Q, et al (2014)
Detecting False Data Injection Attacks on Power Grid by Sparse Optimization
IEEE Transactions on Smart Grid, 612 – 621.

Davenport M A, Plan Y, van den Berg E, et al (2014)
1-Bit Matrix Completion

References

Pengzhi Gao, Ren Wang, Meng Wang, and Joe H. Chow (2016)
Low-rank Matrix Recovery from Quantized and Erroneous Measurements: Accuracy-preserved Data Privatization in Power Grids
Asilomar Conference on Signals, Systems and Computers, 374 – 378.

Pengzhi Gao, Meng Wang, Scott G. Ghiocel, Joe H. Chow, Bruce Fardanesh, and George Stefopoulos (2016)
Missing Data Recovery by Exploiting Low-dimensionality in Power System Synchrophasor Measurements
IEEE Trans. Power Systems 31(2), 1006 – 1013.

Pengzhi Gao, Meng Wang, Joe H. Chow, Scott G. Ghiocel, Bruce Fardanesh, George Stefopoulos, and Michael P. Razanousky (2016)
Identification of Successive "Unobservable" Cyber Data Attacks in Power Systems
IEEE Trans. Signal Processing 64(21), 5557 – 5570.

References

Le Xie, Yang Chen, and P. R. Kumar (2014)
Dimensionality Reduction of Synchrophasor Data for Early Event Detection: Linearized Analysis
IEEE Trans. Power Systems 29(6), 2784 – 2794.

Mark A. Davenport, Yaniv Plan, Ewout van den Berg, and Mary Wootters (2014)
1-bit Matrix Completion
Information and Inference 3(3), 189 – 223.

Sonia A. Bhaskar (2016)
Probabilistic Low-Rank Matrix Completion from Quantized Measurements
Journal of Machine Learning Research 17(60), 1 – 34.

Akshay Soni, Swayambhoo Jain, Jarvis Haupt, and Stefano Gonella (2016)
Noisy Matrix Completion under Sparse Factor Models
IEEE Trans. Information Theory 62(6), 3636 – 3661.

References

Tony Cai, and Wen-Xin Zhou (2013)
A Max-Norm Constrained Minimization Approach to 1-Bit Matrix Completion
Journal of Machine Learning Research 14(1), 3619 – 3647.

Olga Klopp, Jean Lafond, Eric Moulines, and Joseph Salmon (2015)
Adaptive Multinomial Matrix Completion
Electronic Journal of Statistics 9(2), 2950 – 2975.

Andrew S. Lan, Christoph Studer, and Richard G. Baraniuk (2014)
Matrix Recovery from Quantized and Corrupted Measurements
IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 4973 – 4977.

Pengzhi Gao, Meng Wang, Joe H. Chow, Scott G. Ghiocel, Bruce Fardanesh, George Stefopoulos, and Michael P. Razanousky (2016)
Identification of Successive "Unobservable" Cyber Data Attacks in Power Systems.
IEEE Transactions on Signal Processing, 64 (21): 5557-5570.

References

H. Farhangi (2010)
The Path of the Smart Grid
IEEE Power and Energy Magazine, 8(1), 18 – 28.

Quilumba Franklin L., Lee Wei-Jen, Huang Heng, Wang David Y., and Szabados Robert L. (2015)
Using Smart Meter Data to Improve the Accuracy of Intraday Load Forecasting Considering Customer Behavior Similarities
IEEE Transactions on Smart Grid, 6(2), 911 – 918.

S. Haben, C. Singleton, and P. Grindrod (2016)
Analysis and Clustering of Residential Customers Energy Behavioral Demand Using Smart Meter Data
IEEE Transactions on Smart Grid, 7(1), 136 – 144.

G. W. Hart (1992)
Nonintrusive appliance load monitoring
Proceedings of the IEEE, 80(12), 1870 – 1891.

References

M. A. Lisovich, D. K. Mulligan, and S. B. Wicker (2010)
Inferring Personal Information from Demand-Response Systems
IEEE Security Privacy, 8(1), 11 – 20.

Li Fengjun, Luo Bo, and Liu Peng (2011)
Secure and Privacy-Preserving Information Aggregation for Smart Grids
International Journal of Security and Networks, 6(1), 28 – 39.

Barbosa Pedro, Brito Andrey, Almeida Hyggo, and Clau Sebastian (2014)
Lightweight Privacy for Smart Metering Data by Adding Noise
Proceedings of the 29th Annual ACM Symposium on Applied Computing, SAC '14, 531 – 538.

McLaughlin Stephen, McDaniel Patrick, and Aiello William (2011)
Protecting Consumer Privacy from Electric Load Monitoring
Proceedings of the 18th ACM Conference on Computer and Communications Security, 12, 87 – 98.

References

N. Mahmoudi-Kohan, M. Parsa Moghaddam, M.K. Sheikh-El-Eslami, and E. Shayesteh (2010)

A Three-Stage Strategy for Optimal Price Offering by a Retailer Based on Clustering Techniques

International Journal of Electrical Power & Energy Systems, 32(10), 1135 – 1142.

C. Leon, F. Biscarri, I. Monedero, J. I. Guerrero, J. Biscarri, and R. Millan (2011)

Variability and Trend-Based Generalized Rule Induction Model to NTL Detection in Power Companies

IEEE Transactions on Power Systems, 26(4), 1798 – 1807.

Figueiredo Vera, Rodrigues Fatima, Vale Zita, and Gouveia Joaquim Borges (2005)

An Electric Energy Consumer Characterization Framework Based on Data Mining Techniques

IEEE Transactions on Power Systems, 20(2), 596 – 602.

Saeed Aghabozorgi, Seyed Shirkhorshidi Ali, and Ying Wah Teh (2015)

Time-Series Clustering–A Decade Review

Information Systems, 53, 16 – 38.

References

Piao, Minghao and et al (2014)
Subspace Projection Method Based Clustering Analysis in Load Profiling
IEEE Transactions on Power Systems, 29(6), 2628 – 2635.

Chen Yi, Nasrabadi Nasser M, and Tran Trac D (2011)
Hyperspectral Image Classification Using Dictionary-Based Sparse Representation
IEEE Transactions on Geoscience and Remote Sensing, 49(10), 3973 – 3985.

Pengzhi Gao, Meng Wang, Joe H. Chow, Matthew Berger, and Lee M. Seversky (2017)
Missing Data Recovery for High-dimensional Signals with Nonlinear Low-dimensional Structures
IEEE Transactions on Signal Processing, 65(20), 5421 – 5436.

Sonia A. Bhaskar (2016)
Probabilistic Low-Rank Matrix Completion from Quantized Measurements
Journal of Machine Learning Research 17(60), 1 – 34.

References

Ng Andrew, Jordan Michael, and Weiss Yair (2002)
On Spectral Clustering: Analysis and an Algorithm
Advances in neural information processing systems, 2, 849 – 856.

Elhamifar Ehsan, and Vidal Rene (2013)
Sparse Subspace Clustering: Algorithm, Theory, and Applications
IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(11), 2765 – 2781.

Liu Guangcan, Lin Zhouchen, Yan Shuicheng, Sun Ju, Yu Yong, and Ma Yi (2013)
Robust Recovery of Subspace Structures by Low-Rank Representation
IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(1), 171 – 184.

Rahmani Mostafa, and George K. Atia (2017)
Innovation Pursuit: A New Approach to Subspace Clustering
IEEE Transactions on Signal Processing, 65(23), 6276 – 6291.

References

Tseng Paul (2000)
Nearest Q-Flat to m Points
Journal of Optimization Theory and Applications 105(1), 249 – 252.

Pengzhi Gao, Ren Wang, Meng Wang, and Joe H. Chow (2018)
Low-rank Matrix Recovery from Noisy, Quantized and Erroneous Measurements
IEEE Transactions on Signal Processing, 60(11), 2918 – 2932.

Andrew S. Lan, Christoph Studer, and Richard G. Baraniuk (2014)
Matrix Recovery from Quantized and Corrupted Measurements
IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 4973 – 4977.

Mark A. Davenport, Yaniv Plan, Ewout van den Berg, and Mary Wootters (2014)
1-bit Matrix Completion
Information and Inference 3(3), 189 – 223.