High-dimensional Data Analysis in Power System Monitoring

Dr. Meng Wang

Rensselaer Polytechnic Institute





• SCADA (supervisory control and data acquisition) System provides power measurements every five seconds.

ヘロン ヘアン ヘビン ヘビン



• SCADA (supervisory control and data acquisition) System provides power measurements every five seconds.

ヘロン ヘアン ヘビン ヘビン





Example of PMU

http://www.macrodyneusa.com/
model_1690.htm

ヘロン ヘアン ヘビン ヘビン

npa-on-energy-issues/transmission/

 SCADA (supervisory control and data acquisition) System provides power measurements every five seconds.





Example of PMU

http://www.macrodyneusa.com/
model_1690.htm

```
npa-on-energy-issues/transmission/
```

- SCADA (supervisory control and data acquisition) System provides power measurements every five seconds.
- PMUs (Phasor Measurement Units) can provide synchronized phasor measurements of the power system at a sampling rate of 30 samples per second or more.

Phasor Measurement Units

- After the DOE smart grid investment program starting in 2009, 2000 PMUs have been installed in the North American power system.
- Multi-channel PMUs can measure bus voltage phasors, line current phasors, and frequency.





Current Installation of PMUs

http://www.riteh.uniri.hr/zav_katd_

PMU Data Analysis

• Applications:

- State estimation,
- Oscillation detection and electromechanical mode identification,
- Stability analysis and post-event analysis,
- Disturbance detection and location, dynamic security assessment,
- Task specific methods, developed when the coverage of PMUs on a power network was quite sparse.
- Mostly exploit the spatial and the temporal correlations in the measurements separately. Lack a common framework.

・ロ・ ・ 同・ ・ ヨ・ ・ ヨ・

Spatial-temporal Blocks of PMU data

- As the coverage of PMUs becomes denser, one can analyze PMU data collectively from PMUs located in electrically close regions and distinct control regions.
- Processing *spatial-temporal blocks* of PMU data for tasks such as missing data recovery, data compression and storage, disturbance triggering, and detection of cyber data attacks.

・ロト ・ 理 ト ・ ヨ ト ・

Spatial-temporal Blocks of PMU data

- As the coverage of PMUs becomes denser, one can analyze PMU data collectively from PMUs located in electrically close regions and distinct control regions.
- Processing *spatial-temporal blocks* of PMU data for tasks such as missing data recovery, data compression and storage, disturbance triggering, and detection of cyber data attacks.

How to process the high-dimensional PMU data blocks?

イロト 不得 トイヨト イヨト 二日





Pigh-dimensional Data Analysis Using Low-dimensional Models

▲ロト ▲団ト ▲ヨト ▲ヨト 三ヨー わらぐ





Detection of Cyber Data attacks







3 Missing PMU Data Recovery





Exploiting Low-dimensional Structures in High-dimensional Datasets

- It is relatively easy to acquire, store and process high-dimensional datasets if they have *low-dimensional* structures.
- Examples: sparse signals, low rank matrices.
- Wide existence of low-dimensional structures: images, rating matrices, network measurements,...
- Recent arts: compressed sensing theory, low rank matrix theory.

イロト 不得 とくほ とくほ とうほ

Wide Existence of Low-dimensional Structures

Sparse signals

- a vector in \mathbb{R}^n
- k nonzero coefficients under certain basis. (k ≪ n)
- Imaging,Communication, Biology



Compressed by 90%- 5.93 KB

Uncompressed- 57.5 KB

http://www.inkfarm.com/Image-File-Extensions

< < >> < </>

Wide Existence of Low-dimensional Structures

Sparse signals

- a vector in \mathbb{R}^n
- k nonzero coefficients under certain basis. (k ≪ n)
- Imaging,Communication, Biology

Low-rank matrices

- rank « dimension.
- Collaborative filtering (Netflix problem)
- System identification
- Internet traffic analysis



Compressed by 90%- 5.93 KB

Uncompressed- 57.5 KB

http://www.inkfarm.com/Image-File-Extensions



http://www.netflix.com

Movies



User ratings



Unknown sparse signal **x**, observation **y**, measurement matrix *A*. $\mathbf{y} = A\mathbf{x}, m \ll n.$

Given *A*, sparse x can be recovered from y!

▶ < Ξ >



Unknown sparse signal \mathbf{x} , observation \mathbf{y} , measurement matrix A.

 $\mathbf{y} = A\mathbf{x}, m \ll n.$

Given *A*, sparse x can be recovered from y!

- significant reduction in the required number of measurements. *k*-sparse *n*-dimensional signals, $m = O(k \log \frac{n}{k})$.
- simple compression methods. *A* could be random matrices. Wide existence of good *A*'s.
- efficient recovery methods. e.g., ℓ_1 minimization.



 ℓ_0 -minimization

$$\min_{\mathbf{z}} \|\mathbf{z}\|_{\mathbf{0}} \qquad \text{s.t. } A\mathbf{z} = \mathbf{y},$$

where $\|\mathbf{z}\|_0 =$ number of non-zero entries of \mathbf{z} .

・ロット (雪) () () () ()



 ℓ_0 -minimization

 $\min_{\mathbf{z}} \|\mathbf{z}\|_{\mathbf{0}} \qquad \text{s.t. } A\mathbf{z} = \mathbf{y},$

where $\|\mathbf{z}\|_0 =$ number of non-zero entries of \mathbf{z} .

• bad news: computationally hard.

・ロット (雪) () () () ()



 ℓ_1 -minimization ℓ_1 -minimization

$$\min_{\mathbf{z}} \|\mathbf{z}\|_1 = \sum_i |z_i| \qquad \text{s.t. } A\mathbf{z} = \mathbf{y},$$

where $\|\mathbf{z}\|_1 = \sum_i |z_i|$.

The sparse signal \mathbf{x} is the solution to ℓ_1 problem!

- computationally efficient.
- theoretical guarantee. (Candès and Tao 2006, Donoho 2006)

Netflix Problem

Movies





イロト イポト イヨト イヨト

Netflix Problem

Movies





• • • • • • • • • • •

Low rank matrix completion problem.



Low-rank Matrix with Missing Entries

How shall we recover the missing entries of a low-rank matrix?

ヘロア ヘビア ヘビア・



Low-rank Matrix with Missing Entries

How shall we recover the missing entries of a low-rank matrix?

$$\min_{X} \operatorname{Rank}(X)$$

s.t. $X_{ij} = M_{ij}, \forall i \in \Omega$.

Ω: locations of the observed entries.

NP hard.

ヘロン ヘアン ヘビン ヘビン

Nuclear norm minimization (Fazel 2002, Candés & Recht 2009), recover the missing data by solving a convex program.

$$\min_{X} ||X||_{*} = \text{sum of the singular values of } X$$

s.t. $X_{ij} = M_{ij}, \forall i \in \Omega.$

- Convex relaxation of Rank minimization problem.
- computationally efficient. Can be solved by Semidefinite Programming.
- Theoretical guarantee. (Candés & Recht 2009)

イロト 不得 とくほ とくほ とうほ

Theorem (Candés & Recht 09, Gross 11, Recht 11)

All entries of a rank-r matrix $L \in \mathbb{C}^{n_1 \times n_2}$ can be corrected recovered, as long as $O(rn \log^2 n)(n = \max(n_1, n_2))$ randomly selected entries of *L* are observed.

Significant saving in the number of observations when r is small.

Theorem (Candés & Recht 09, Gross 11, Recht 11)

All entries of a rank-r matrix $L \in \mathbb{C}^{n_1 \times n_2}$ can be corrected recovered, as long as $O(rn \log^2 n)(n = \max(n_1, n_2))$ randomly selected entries of *L* are observed.

Significant saving in the number of observations when r is small.

Theorem (Candés & Tao 09)

If each entry is observed with prob. $p = \frac{m}{n_1 n_2}$, no method whatsoever can succeed with

 $m \leq Cnr\log n$, for some fixed constant C

・ロト ・回 ト ・ヨト ・ヨト … ヨ

Theorem (Candés & Recht 09, Gross 11, Recht 11)

All entries of a rank-r matrix $L \in \mathbb{C}^{n_1 \times n_2}$ can be corrected recovered, as long as $O(rn \log^2 n)(n = \max(n_1, n_2))$ randomly selected entries of *L* are observed.

Significant saving in the number of observations when r is small.

Theorem (Candés & Tao 09)

If each entry is observed with prob. $p = \frac{m}{n_1 n_2}$, no method whatsoever can succeed with

 $m \leq Cnr\log n$, for some fixed constant C

Near-optimal recovery via Nuclear norm minimization.

・ロン ・四 と ・ ヨン ・ ヨ

General setup Rank minimization

 $\min_{X} \operatorname{Rank}(X)$ s.t. $\langle A^{i}, X \rangle = \operatorname{Trace}((A^{i})^{T}X) = b^{i},$ $\forall i = 1, ..., m$

General setup Rank minimization

 $\min_{X} \operatorname{Rank}(X)$ s.t. $\langle A^{i}, X \rangle = \operatorname{Trace}((A^{i})^{T}X) = b^{i},$ $\forall i = 1, ..., m$ Convex relaxation.

$$\begin{split} \min_{X} \|X\|_{*} \\ \text{s.t.} \ \langle A^{i}, X \rangle &= \operatorname{Trace}((A^{i})^{T}X) = b^{i}, \\ \forall i = 1, ..., m \end{split}$$

イロト 不得 とくほ とくほ とうほ

General setup Rank minimization

Convex relaxation.

$$\begin{split} \min_{X} \operatorname{Rank}(X) & \min_{X} \|X\|_{*} \\ \text{s.t. } \langle A^{i}, X \rangle &= \operatorname{Trace}((A^{i})^{T}X) = b^{i}, \\ \forall i = 1, ..., m & \forall i = 1, ..., m \end{split}$$

Consider the special case $X = diag(\mathbf{x}), \mathbf{x} \in \mathbb{R}^n$, then

 $\operatorname{Rank}(X) = \|\mathbf{x}\|_0,$

 $||X||_* = ||\mathbf{x}||_1.$

・ロト ・ 同 ト ・ 三 ト ・ 三 ・ つへの

General setup Rank minimization

Convex relaxation.

 $\min_{X} \operatorname{Rank}(X) \qquad \min_{X} ||X||_{*}$ s.t. $\langle A^{i}, X \rangle = \operatorname{Trace}((A^{i})^{T}X) = b^{i}, \qquad \text{s.t. } \langle A^{i}, X \rangle = \operatorname{Trace}((A^{i})^{T}X) = b^{i},$ $\forall i = 1, ..., m \qquad \forall i = 1, ..., m$

Consider the special case $X = diag(\mathbf{x}), \mathbf{x} \in \mathbb{R}^n$, then

 $\operatorname{Rank}(X) = \|\mathbf{x}\|_0,$

$$||X||_* = ||\mathbf{x}||_1.$$

Reduce to compressed sensing!

・ロト ・ 同ト ・ ヨト ・ ヨー ・ つへの

Principle Component Analysis: Classical Method

- widely used for dimension reduction.
- convert observations to a set of values of linearly uncorrelated variables, called *principle components*.



Stack all the data points as column vectors of a matrix M, PCA seeks the best rank-k approximation by solving

 $\min_{L} \|M - L\|_2$
s.t. Rank $(L) \le k$

Principle Component Analysis: Classical Method

- widely used for dimension reduction.
- convert observations to a set of values of linearly uncorrelated variables, called *principle components*.



A B > 4
 B > 4
 B

Stack all the data points as column vectors of a matrix M, PCA seeks the best rank-k approximation by solving

```
\min_{L} \|M - L\|_2<br/>s.t. Rank(L) \le k
```

Not rubust to even a few outliers in M.

Robust PCA

The observation matrix is the sum of a low-rank matrix L₀ (actual data) and a sparse matrix (corruptions).

$$M = L_0 + S_0$$

• Decomposition of a low-rank matrix and a sparse matrix.

N

 Applications in computer vision, image processing, network traffic analysis, etc.

・ロト ・ 理 ト ・ ヨ ト ・

Robust PCA

(Hard) Optimization problem

 $\min_{L,C} \operatorname{Rank}(L) + \lambda ||C||_0$ s.t. L + C = M

- $||C||_0$: the number of nonzero entries in *C*.
- not tractable.

・ロト ・ 同ト ・ ヨト ・ ヨト … ヨ

Robust PCA

(Hard) Optimization problem

 $\min_{L,C} \operatorname{Rank}(L) + \lambda \|C\|_0$ s.t. L + C = M

||C||₀: the number of nonzero entries in C.
not tractable.

Convex relaxation

$$\min_{L,C} ||L||_* + \lambda ||C||_1$$

s.t. $L + C = M$

- $||C||_1 = \sum_{ij} |C_{ij}|$
- Theoretical guarantee:

Candés & Li & Ma & Wright 11, Chandrasekaran & Sanghavi & Parrilo & Willsky 11, Recht & Fazel & Parrilo 10.
Does low-dimensionality exist in PMU measurements?

<ロ> <四> <四> <四> <三</td>

Low-rank Property of PMU data



PMUs in Central NY Power Current magnitudes of Systems PMU data

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 observed in Chen & Xie & Kumar 2013, Dahal & King & Madani 2012)

How could we use the low dimensionality in PMU data management for power system monitoring?

イロト イポト イヨト イヨト 三日





2) High-dimensional Data Analysis Using Low-dimensional Models







Missing Data Recovery



We observe partial entries of the complex matrix (rectangular form of voltage and current phasors). How shall we recover the missing points for offline applications like system identification and past event analysis?

Missing Data Recovery



We observe partial entries of the complex matrix (rectangular form of voltage and current phasors). How shall we recover the missing points for offline applications like system identification and past event analysis?

Low-rank Matrix Completion Problem!

21/44

- Wide applications in collaborative filtering, computer vision, machine learning, remote sensing, and system identification.
- Problem formulation: given part of the entries of a matrix, recover the remaining entries.
- the rank of the matrix is much less than its dimension.





Low-rank Matrix with Missing Entries

ヘロン 人間 とくほど くほとう



Low-rank Matrix with Missing Entries

Nuclear norm minimization

 $\min_{X} ||X||_* = \text{sum of singular values of } X$ s.t. X is consistent with the observed entries,

・ロト ・ 理 ト ・ ヨ ト ・



Low-rank Matrix with Missing Entries

Nuclear norm minimization

 $\min_X \|X\|_* = \text{sum of singular values of } X$

s.t. X is consistent with the observed entries,

 Quite a few recovery algorithms exist, e.g., singular value thresholding (SVT) (Cai *et al.* 2010), information cascading matrix completion (ICMC) (Meka *et al.* 2009).

Theorem (Candés & Recht 09, Gross 11, Recht 11)

All entries of a rank-r matrix $L \in \mathbb{C}^{n_1 \times n_2}$ can be corrected recovered, as long as $O(rn \log^2 n)(n = \max(n_1, n_2))$ randomly selected entries of *L* are observed.

프 에 에 프 어

Theorem (Candés & Recht 09, Gross 11, Recht 11)

All entries of a rank-r matrix $L \in \mathbb{C}^{n_1 \times n_2}$ can be corrected recovered, as long as $O(rn \log^2 n)(n = \max(n_1, n_2))$ randomly selected entries of *L* are observed.

• Significant saving in the number of observations when *r* is small.

・ 同 ト ・ ヨ ト ・ ヨ ト

Theorem (Candés & Recht 09, Gross 11, Recht 11)

All entries of a rank-r matrix $L \in \mathbb{C}^{n_1 \times n_2}$ can be corrected recovered, as long as $O(rn \log^2 n)(n = \max(n_1, n_2))$ randomly selected entries of *L* are observed.

- Significant saving in the number of observations when *r* is small.
- Existing analysis assumes that the locations of missing points are selected randomly.

・ロ・ ・ 同・ ・ ヨ・ ・ ヨ・

Our Results

• The locations of missing PMU data are usually correlated.

- temporal correlation: loss of consecutive measurements in one PMU channel.
- channel correlation: loss of measurements in multiple PMU channels simultaneously.

ヘロン ヘアン ヘビン ヘビン

Our Results

• The locations of missing PMU data are usually correlated.

- temporal correlation: loss of consecutive measurements in one PMU channel.
- channel correlation: loss of measurements in multiple PMU channels simultaneously.

Theorem (Gao & Wang & Ghiocel & Chow 14)

Although the locations of the missing entries of a rank-*r* matrix are temporally or spatially correlated, all missing entries can be correctly recovered as long as $O(n^{2-\frac{1}{r+1}}r^{\frac{1}{r+1}}\log^{\frac{1}{r+1}}n)$ entries are observed.

イロト 不得 とくほ とくほ とうほ

Our Results

• The locations of missing PMU data are usually correlated.

- temporal correlation: loss of consecutive measurements in one PMU channel.
- channel correlation: loss of measurements in multiple PMU channels simultaneously.

Theorem (Gao & Wang & Ghiocel & Chow 14)

Although the locations of the missing entries of a rank-*r* matrix are temporally or spatially correlated, all missing entries can be correctly recovered as long as $O(n^{2-\frac{1}{r+1}}r^{\frac{1}{r+1}}\log^{\frac{1}{r+1}}n)$ entries are observed.

• The first theoretical guarantee of low-rank matrix completion when the locations of missing entries are correlated.

イロン 不良 とくほう 不良 とうほ

Simulation



Current magnitudes in four datasets.

э

・ロト ・ ア・ ・ ヨト ・ ヨト

Simulation Results

• Recover missing data points that are temporally correlated.

the first 5-second PMU data of dataset #1.







Relative recovery error of the ICMC algorithm

Simulation Results

• Recover missing data points that are spatially correlated.

the first second PMU data of dataset #1.



Relative recovery error of the SVT algorithm



Relative recovery error of the ICMC algorithm

Online Missing Data Recovery

 We also develop an online missing data recovery algorithm that could fill in the missing data instantaneously and detect system disturbance.

Table : Relative recovery error of OLAP on four datasets

p_{avg}	Dataset #1	Dataset #2	Dataset #3	Dataset #4
0.05	0.0082	0.0026	0.0006	0.0186
0.10	0.0089	0.0039	0.0009	0.0235
0.15	0.0119	0.0053	0.0011	0.0442
0.20	0.0145	0.0060	0.0013	0.0494
0.25	0.0153	0.0066	0.0014	0.0726
0.30	0.0191	0.0075	0.0017	0.0779
0.35	0.0204	0.0082	0.0018	0.1137
0.40	0.0227	0.0091	0.0019	0.1189



Missing data recovery by OLAP algorithm

Gao, Wang, Ghiocel, Chow, IEEE Trans. Power Systems 2016.





2 High-dimensional Data Analysis Using Low-dimensional Models





Detection of Cyber Data attacks



・ロ・・聞・・用・・用・ ひゃぐ

• Measurements V_1 , V_2 , I_{12} , and I_{13} . Estimate V_3 .

A B + A B +
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A



► < E >



- Measurements V_1 , V_2 , I_{12} , and I_{13} . Estimate V_3 .
- Redundancy in measurements can be used to detect bad data.



- Measurements V_1 , V_2 , I_{12} , and I_{13} . Estimate V_3 .
- Redundancy in measurements can be used to detect bad data.
- Cyber data attack: manipulate *I*₁₂ and *I*₁₃ simultaneously.



- Measurements V_1 , V_2 , I_{12} , and I_{13} . Estimate V_3 .
- Redundancy in measurements can be used to detect bad data.
- Cyber data attack: manipulate *I*₁₂ and *I*₁₃ simultaneously.
- Can result in significant error in *V*₃.



- Measurements V_1 , V_2 , I_{12} , and I_{13} . Estimate V_3 .
- Redundancy in measurements can be used to detect bad data.
- Cyber data attack: manipulate *I*₁₂ and *I*₁₃ simultaneously.
- Can result in significant error in *V*₃.
- The removal of attacked measurements will make *V*₃ unobservable.

Cyber Data Attacks

• The worst interacting bad data. (Liu & Ning & Reiter 11).

Cyber Data Attacks

- The worst interacting bad data. (Liu & Ning & Reiter 11).
- An intruder with the system topology information can simultaneously manipulate multiple measurements so that these attacks cannot be detected by any bad data detector.
- Cyber data attacks can potentially lead to significant errors to the outcome of state estimation.

・ロト ・ 理 ト ・ ヨ ト ・

Cyber Data Attacks

- The worst interacting bad data. (Liu & Ning & Reiter 11).
- An intruder with the system topology information can simultaneously manipulate multiple measurements so that these attacks cannot be detected by any bad data detector.
- Cyber data attacks can potentially lead to significant errors to the outcome of state estimation.
- Also termed as "unobservable attacks", because the removal of affected measurements would make the system unobservable. (Kosut & Jia & Thomas & Tong 10).

・ロ・ ・ 同・ ・ ヨ・ ・ ヨ・

Existing Approaches

Cyber attacks in SCADA system. A protection perspective.

 Usually protect key PMUs to avoid these attacks.(Kosut & Jia & Thomas & Tong 10, Kim & Poor 11, Bobba *et al.* 10, Dán & Sandberg 10)

・ロト ・ 理 ト ・ ヨ ト ・

Existing Approaches

Cyber attacks in SCADA system. A protection perspective.

- Usually protect key PMUs to avoid these attacks.(Kosut & Jia & Thomas & Tong 10, Kim & Poor 11, Bobba *et al.* 10, Dán & Sandberg 10)
- Sedghi & Jonckheere 13: Detection of cyber data attacks in SCADA system. Assume the measurements at different time instants are i.i.d. distributed.

・ロ・ ・ 同・ ・ ヨ・ ・ ヨ・

Our Contributions

A new detection method of cyber data attacks in PMU measurements.

- no other assumptions expect for the low-rankness of the PMU data.
- The detection method can identify the attacks even when the system is under disturbance.
- The PMU channels under attack can be identified by solving a convex optimization problem.

・ロト ・ 理 ト ・ ヨ ト ・

System Model

• π model of a transmission line.



 π model of a transmission line

• Current I^{ij} from bus *i* to bus *j* is related to bus voltage V^i and V^j by

$$I^{ij} = \frac{V^i - V^j}{Z^{ij}} + V^i \frac{Y^{ij}}{2}.$$

 Voltage and current phasor measurements can be represented by linear functions of state variables.

< ∃→

A B + A B +
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A

Attack Model

- The intruder can only attack a small number of PMUs continuously.
- The intruder injects an unobservable attack at each time instant.
- An unobservable attack is a linear combination of errors in state variables.

ヘロン ヘアン ヘビン ヘビン

Mathematical Formulation

- *t*: total number of sampling time instants.
- *p*: total number of PMU channels.
- *n*: total number of buses. *n* < *p*
- $\overline{L} \in \mathbb{C}^{t \times p}$: the actual voltage and current phasors in *t* instants.
- $\overline{D} \in \mathbb{C}^{t \times n}$: the additive error in state variables.
- $N \in \mathbb{C}^{t \times p}$: the noise.
- $M \in \mathbb{C}^{t \times p}$: the obtained measurements that are under attack.
- $W \in \mathbb{C}^{t \times p}$: relates state variables with phasors.

$$M = \bar{L} + \bar{C}W^T + N$$
Assumptions



- \overline{L} : low-rank. From correlations in measurements.
- \bar{C} : column sparse. The intruder has limited access to the system.
- N: $||N||_F \leq \varepsilon$.

Given *M* and *W*, how could we identify \overline{L} and \overline{C} ? Decomposition of a low-rank matrix and a transformed column-sparse matrix.

Matrix Decomposition



- Decomposition of a low-rank matrix and a sparse matrix.
- Through convex optimization.

$$\min_{L \in \mathbb{C}^{t \times p}, C \in \mathbb{C}^{t \times n}} \|L\|_* + \lambda \sum_{i,j} |C_{ij}| \quad \text{s.t.} \quad \|L + C - M\|_F \le \varepsilon$$
(1)

- Theoretical guarantee: Candés & Li & Ma & Wright 11, Chandrasekaran & Sanghavi & Parrilo & Willsky 11, Recht & Fazel & Parrilo 10.
- Applications: Internet traffic analysis, image processing, medical imaging, etc.

Connection to Related Work

• Xu & Caramanis & Sanghavi 12: Decomposition of a low-rank matrix and a column-sparse matrix.



Our methods and proofs are built upon those in Xu & Caramanis & Sanghavi 12. Extension to general cases.



Connection to Related Work

 Mardani & Mateos & Giannakis 13: Decomposition of a low-rank matrix plus a compressed sparse matrix. Internet traffic anomaly detection.



Our focus: column-sparse matrices, *W* is arbitrary.



Our Approach

• Find (*L*^{*}, *C*^{*}), the optimum solution to the following optimization problem

$$\min_{L \in \mathbb{C}^{t \times p}, C \in \mathbb{C}^{t \times n}} \|L\|_* + \lambda \|C\|_{1,2} \quad \text{s.t.} \quad \|L + CW^T - M\|_F \le \varepsilon$$
(2)

- Compute the SVD of $L^* = U^* \Sigma^* V^{*\dagger}$.
- Find column support of $D^* = C^* W^T$, denoted by \mathscr{I}^* .
- Return $L^*_{\mathscr{I}^{*c}}$, U^* and \mathscr{I}^* .

(2) is convex and can be solved efficiently.

イロン 不良 とくほう 不良 とうほ

Theoretical Guarantee

Theorem (Noiseless measurements, N = 0)

With a properly chosen λ , the solution returned by our method

- identifies the PMU channels under attack.
- identifies the measurements that are not attacked.
- recovers the correct subspace spanned by actual phasors.

글 에 에 글 어 !!

Theoretical Guarantee

Theorem (Noiseless measurements, N = 0)

With a properly chosen λ , the solution returned by our method

- identifies the PMU channels under attack.
- 2 identifies the measurements that are not attacked.
- recovers the correct subspace spanned by actual phasors.

Theorem (Noisy measurements, $N \neq 0$)

With a properly chosen λ , the solution returned by our method is sufficiently close (with distance depending on the noise level) to a solution such that 1-3 hold.

イロト 不得 とくほ とくほ とうほ

Numerical Results



- Simulate the case that the intruder alters the PMU channels that measure I^{12} , I^{52} , I^{13} and I^{43} .
- The voltage phasor estimates of Buses 2 and 3 are corrupted



Actual values and corrupted values

Recovered error in each channel

20 30

Column index

(b) Disturbance data

norm of the column

2

Gao, Wang, Chow, Ghiocel, Fardanesh, Stefopoulos, and Razanousky, to appear in IEEE Trans.

Conclusion

- Leverage the low-dimensional structures in high-dimensional data to address the challenges in data acquisition, storage, and information extraction.
- Connection of low-rank methods with power system monitoring. Analysis of spatial-temporal blocks of PMU measurements.
- Missing data recovery: theoretical guarantee of successful recovery when the locations of the missing points are correlated.
- Detection of cyber data attacks: decomposition of a low-rank and transferred column-sparse matrix.

・ロト ・ 理 ト ・ ヨ ト ・



◆□▶ ◆□▶ ◆臣▶ ◆臣▶ 臣 のへで

Compressed sensing attempts to recover a k-sparse (only contains k nonzero entries) vector $\hat{\mathbf{x}}$ in \mathbb{R}^n ($k \ll n$) from *m*-dimensional measurements $\mathbf{y} = A\hat{\mathbf{x}}$ (m < n) by solving the following ℓ_1 minimization problem,

$$\min_{\mathbf{x}\in\mathcal{R}^n} \|\mathbf{x}\|_1 \quad \text{subject to } A\mathbf{x} = \mathbf{y}. \tag{1}$$

Under mild assumptions, the signal of interest $\hat{\mathbf{x}}$ is indeed the solution to (2).

A general low-rank matrix completion problem attempts to recover a rank-r matrix M in $\mathbb{R}^{n \times n}$ from m linear measurements $b^i = \langle A^i, M \rangle, i = 1, ..., m$, by solving the Nuclear-norm minimization problem,

$$\min_{X \in \mathcal{R}^{n \times n}} \|X\|_* \quad \text{subject to } \langle A^i, X \rangle = b^i, \quad \forall i = 1, ..., m$$
(2)

where

- $||X||_*$ is the sum of the singular values of X,
- A^i is in $\mathbb{R}^{n \times n}$ for every i,
- b^i is in \mathbb{R} for every i,
- $\langle A, B \rangle$ is the inner product of matrices A and B in $\mathbb{R}^{n \times n}$. $\langle A, B \rangle = \sum_{k=1}^{n} \sum_{j=1}^{n} A_{kj} B_{kj}$.

Show that ℓ_1 minimization problem can be viewed as a special case of Nuclear-norm minimization problem.