Missing data recovery by exploiting low-dimensionality in synchrophasor measurements

Genevieve de Mijolla, Pengzhi Gao, Meng Wang & Joe Chow
(Rensselaer Polytechnic Institute),

Bruce Fardanesh, George Stefopoulos, Saman Babaei & Alan Ettlinger
(New York Power Authority),

Dan Iles & De Tran
(New York ISO)
Problem Motivation

- Phasor Measurement Units:
  - High sampling rate: 30/60 samples/second
  - Usually multi-channel: measure bus voltage phasors, line current phasors, and frequency.
- PMU data is considered a source of Big Data in power systems.
- Missing Data affect applications such as state estimation and disturbance identification.

Developing reliable and computationally efficient PMU data recovery methods.
Low-rank Property of PMU Data

- Existing missing data recovery: interpolation from measurements in the same channel.
- Our approach: Analyze PMU data of multiple time instants collectively from PMUs in electrically close regions and distinct control regions.
- Process spatial-temporal blocks of PMU data for tasks such as missing data recovery, data compression and storage, disturbance triggering, detection of data substitution attacks.
- Key feature: low-rankness of PMU data blocks. (also observed in Chen & Xie & Kumar 2013, Dahal & King & Madani 2012.)
Low-rank Property of PMU Data

- 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-rankness also used in Chen & Xie & Kumar 2013, Dahal & King & Madani 2012 for dimensionality reduction.

**PMUs in Central NY Power System**

**Current magnitudes of PMU data**

**Singular values of the PMU data matrix**
Our approach: leverage low-rankness of PMU data blocks.

Low-rank matrix completion Problem!

Quite a few recovery algorithms exist.
Low-rank Matrix Completion Methods

Nuclear norm minimization (Fazel 2002), recover the missing data by solving a convex program:

$$\min_X \|X\|_*$$

Singular Value Thresholding (Cai 2010), iteratively solve a modified version of the nuclear norm minimization problem:

$$\min_X \tau \|X\|_* + \frac{1}{2} \|X\|_F^2$$

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

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

$\|X\|_F^2 = \text{sum of the absolute squares of all entries of } X$
OLAP (Online Algorithm for PMU Data Processing)

OLAP can fill in the missing data in *real-time*.

OLAP continuously updates dominating singular values and singular vectors. The new data is viewed as a linear combination of existing singular vectors.

OLAP is adaptive to the dimension change of the subspace. Disturbances in the power system can cause the dimensionality of a PMU data matrix to vary rapidly with time.

Theoretical Analysis of PMU Data Recovery by Low-rank Matrix Completion Methods

The locations of missing PMU data are usually correlated.

- **Temporal** correlation: loss of consecutive measurements in one PMU channel.
- **Spatial** correlation: loss of measurements in multiple PMU channels simultaneously.

Theoretical guarantee of low-rank matrix completion when the locations of missing points are correlated.

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 $O(n^2 - \frac{1}{r+1} \frac{1}{rr+1} log^{r+1} n)$ entries are observed.

Data Recovery

- Tested: Ten 5-minute data segments of NYS PMU data during grid disturbances.

- Recovered:
  - Voltage magnitude (53 channels)
  - Voltage angle (53 channels)
  - Frequency (53 channels)
  - Current magnitude (264 channels)
  - Current angle (264 channels)
Varying the threshold parameter for the OLAP algorithm allows the submatrix rank to vary.

Computation time (2.60 GHz Intel Core i5 with 16 GB RAM):

<table>
<thead>
<tr>
<th></th>
<th>Voltage magnitude</th>
<th>Voltage angle</th>
<th>Frequency</th>
<th>Current magnitude</th>
<th>Current angle</th>
</tr>
</thead>
<tbody>
<tr>
<td>OLAP</td>
<td>1.305 secs</td>
<td>1.327 secs</td>
<td>1.239 secs</td>
<td>8.121 secs</td>
<td>9.113 secs</td>
</tr>
</tbody>
</table>
Data Recovery Results

Data recovery of a NYS disturbance for voltage magnitude, voltage angle, and frequency.
Data Recovery Results

Data recovery of a NYS disturbance for current magnitude and current angle.
Recovery Results – Consecutive data drops

OLAP and SVT both successfully recovered a capacitor switching event when filling in missing data.
Recovery Results – Numerous data drops

Illustration of data recovery results for OLAP and SVT when there are a large number of data drops.
Data Recovery Challenges

- Pre-processing:
  - Bad data check
  - Angle unwrapping

- Input parameters:
  - Especially hard to determine for SVT
Real-time OpenPDC Implementation

Other Applications
Ex: data visualization

Data Recovery Algorithm

Phasor State Estimator

ePDC

PMU data stream

Open PDC

PMU
Conclusions

- Missing data recovery by exploiting low-dimensional structures in PMU data.
- Both offline and online missing data recovery methods are reported.
- Missing PMU data was successfully recovered, even in cases of consecutive data drops or numerous data drops.
References


Questions?