Lossless Compression of Synchrophasor Measurement Unit Archives

R. Klump, H. Khurana, Z. Tate

TCIPG

Goals

- To take advantage of characteristics of power system data gathered by synchrophasor measurement unit devices (PMUs) to develop a compression algorithm for PMU data archives that will preserve important signal content but consume less disk space.
- To identify groups of streams of PMU data that exhibit similar trends and traits. This will help us identify related measurements in streams that lack metadata, which will be useful for subsequent system studies. It will also help us choose reference nodes for our compression algorithm in a more informed manner.
- To identify the noise component in PMU data to yield a more compressible data stream and to define the resolution with which such data should be stored.
- To characterize PMU data streams so that they may be generated more realistically in simulations.

Fundamental Questions/Challenges

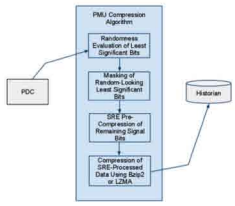
- As PMUs become more commonplace, and as their reporting rates increase, the amount of data reported by PMUs will grow, and so will PMU archives.
- Can the sizes of these archives be reduced without compromising the fidelity of the measurement?
- Are there characteristics unique to power systems that could be used to compress PMU data more effectively?
- Are all the bits of a PMU measurement actual data, or are some of the bits noise? And if some of the bits are noise, how beneficial to the compression process will filtering the noise bits be?

Research Plan

- Gather a diverse pool of PMU data sets.
- Identify characteristics of power system data that could enable pre-processing similar to what is done with images. For example, PNG image compression observes pixel-to-pixel color changes that are smooth and exhibit a certain degree of continuity. Can the same be said for data measured by PMUs? This work identified two such coherencies: temporal and spatial.
- To capture the temporal coherency, record the measurement-to-measurement changes in each data stream.
- To capture the spatial coherency, refer these measurement-to-measurement changes to corresponding changes in a reference stream.
- The resulting slack-referenced delta values yield data over a narrower and, therefore, more compressible range than the original data.
- Compressibility could be enhanced further if some of the bits prove to be noise rather than meaningful signal and can be filtered without impacting the fidelity of the measurement. The NIST test suite for randomness was applied to each bit position of each data stream to identify which bits might be noise and which might be signal. Bit positions that passed the NIST noise tests were filtered, and then the preprocessing and compression steps were performed. Predictably, removing the "noise" bits increased compression ratios significantly.

Research Results

- Here is a diagram of the algorithm for compressing PMU archives.



- Using this algorithm, compression ratios of up to 20 to 1 were seen with real event-related PMU data.

Quantity	Without Pre-Compression	With Pre-Compression	With Pre-Compression and 4-bit Noise Removal	With Pre-Compression and 8-bit Noise Removal
Voltage Mags	2.21	2.50	3.39	6.31
Phase Angles	2.47	3.94	6.92	20.81

Broader Impact

- As PMUs become increasingly used, efficient storage and retrieval of the data will increasingly become a challenge.
- This research effort will make more efficient storage of data possible.
- With smaller data sets, searching for and retrieving data should become more efficient.
- Peripheral aspects of this work that have focused on characterizing PMU data (e.g., signal vs. noise, coherencies) could provide a foundation for simulating PMU operation in integrated cyber-physical studies of the system.

Interaction with Other Projects

- Although it has not been investigated yet, this project could integrate with the "PMU Integration into Power Flow Software" project, as that would be an interesting way to determine whether filtering the noise bits impacts applications that use the data.

Future Efforts

- Test the algorithm with additional data sets and tweak it if necessary.
- Implement more intelligent techniques for selecting the reference node for modeling spatial coherency.
- Integrate the algorithm with an existing PMU data archive system, like openPDC.
- Code observations regarding the characteristics of PMU data as part of a simulation tool that models PMU data measurement and reporting.