This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/MCG.2022.3171506, IEEE Computer Graphics and Applications

PMUVis : A Large Scale Platform to Assist Power System Operators in a Smart Grid

Anjana Arunkumar
Arizona State University

Nitin Gupta
Arizona State University

Andrea Pinceti
Arizona State University

Lalitha Sankar
Arizona State University

Chris Bryan
Arizona State University

***Abstract*—Electric transmission power grids are being revamped with the widespread deployment of GPS-enabled Phasor Measurement Units (PMUs) for real-time wide-area monitoring and control via precise, time-synchronized measurements of voltage and current. Large, concurrently produced volumes of noisy data hinder PMU usability, particularly for the analysis of power oscillation and load fluctuation events in the grid. We examine visualization challenges for events in the electric power grid and develop PMUVis, a visualization platform that supports scalable analysis of grid network topology and anomalous events in near-time. PMUVis incorporates a novel FFT-based approach over raw and temporally aggregated data to examine oscillation event propagation through the grid network. We validate PMUVis with expert reviews and a case study, and discuss how visualization can be leveraged to enhance real-time, spatiotemporal grid analysis by advancing operator capabilities.**

Electric power grids are distributed, complex cyber-physical systems designed to provide reliable energy delivery. Due to congestion, atypical power flows, and increasing demand for renewable

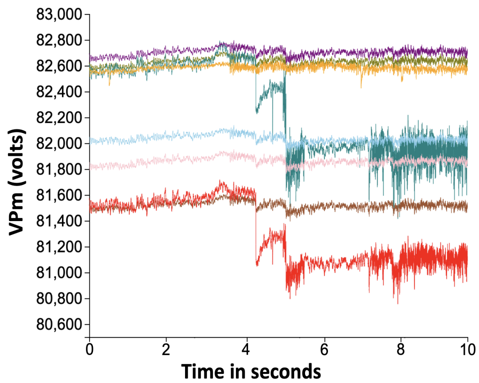


Figure 1: Voltage Magnitude (VPm) of eight PMUs from the North American power grid over a 10 minute time period. While basic trends can be seen with a small number of PMUs as shown here, raw data quickly becomes noisy as additional PMUs are added to the display.

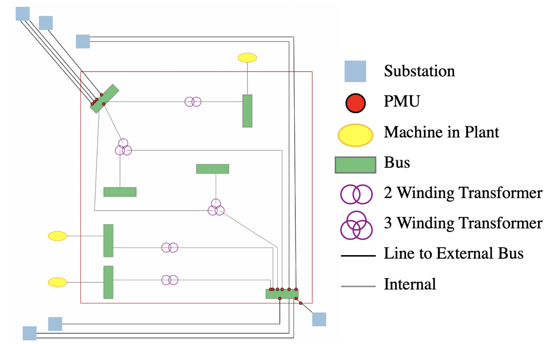


Figure 2: Schematic of a substation that depicts the interconnections between buses and transformers within the substation, as well as connections to other substations and external loads. PMUs monitoring voltage and current for certain lines are also present.

energy, power systems have been revamped as *smart grids* to combat heightened complexity and unpredictability, and enhance reliability and critical real-time decision-making.

Industry stakeholders are increasingly deploying GPS-synchronized phasor measurement units, or PMUs, for real-time wide-area monitoring and control. PMUs provide high speed and high resolution streaming data (usually 30–120 samples per second) of the phase/magnitude of voltages and currents [15]. Using a common time signal (in most cases provided by GPS), measurements are synchronized across wide geographical areas. This provides a clear and precise picture of the grid state in terms of power flows and power quality (voltage profile, frequency, etc.).

PMUs can hence capture dynamic subsecond behaviors that were previously unobservable using traditional supervisory control and data acquisition (SCADA) data. However, their usage comes with several challenges, including the massive scale of PMU data, and wideband noise [12] which impacts frequency and rate of change of frequency estimation in sensor measurements.

Unfortunately, current industry platforms provide limited flexibility and scalability for visually investigating grid events. For many operators, grid event analysis involves manually correlating network diagrams of the grid topology with changes in sensor values obtained by tediously inspecting line charts of PMU data streams (see Figure 1). This not only limits operator ability to conduct in-depth analyses for decision-making, but has also led to the slow and scattered adoption of PMU-equipped substations (Figure 2) by electric utilities.

In this paper, we focus on end-to-end, flexible, and scalable data management and visual analysis of PMU data. Our research contributions are as follows: (i) We survey a set of domain experts who work for U.S. power and electrical companies to understand current industry challenges and practices for handling and visualizing PMU data. (ii) We identify a set of 'focus' tasks required for effective visualization and analysis of large-scale PMU data to examine how events propagate through the grid network. (iii) Based on these requirements, we implement a software platform called PMUVis that supports scalable, interactive analysis of PMU data. (iv) To evaluate PMUVis, we conduct a case study and an in-depth expert review session with domain operators at a regional U.S. power and electrical company.

Related Work

As cyber-physical systems (CPS), smart grid power systems closely integrate computation, networking, and physical processes, and rely heavily on real-time monitoring, assessment, and decision making [10]. Within this context, visualization is an important technique for understanding the state of the grid.

Visualizing PMU data streams. Several industry and open-source software platforms support visualizing PMU and SCADA data streams. Real-Time Dynamics Monitoring System [4], or RT-DMS, is a platform for real-time grid monitoring that includes many "traditional" visualizations that are standard to domain operators. This includes geographic displays overlaid with the grid topology to show network state and connectivity and line charts of raw PMU data streams. Additionally, odometer-inspired radial dials are used to represent phase angle separation, voltage sensitivity, oscillation, and grid stress monitoring. Other wide area monitoring and control systems like Cozby et.al. [6] allow the capture and visualization of PMU data streams in real-time. Backend algorithms to identify anomalous behavior (i.e., events) such as oscillations or line faults are also incorporated. However, their visualization support is generally not tailored for event analysis. In contrast, while PMUVis currently supports historical data analysis, it is specifically designed to support in-depth event analysis, and can be adapted to work for real-time data.

Visualizing power grid topologies. A common approach for visualizing the grid topology is overlaying a node-link diagram atop a geographical map. Several industry platforms have extended this view to encode additional information. For example, contour maps are commonly utilized to show bus and line voltage/current profiles of regions either geographically or within bus systems [19]. Unfortunately, this approach interpolates virtual values in the "blank regions" of geographic space between disconnected network nodes. Another approach is showing grid nodes as 3D glyphs [20], which could lead to occlusion, as the PMUs are highly interconnected and clustered on subsets of substation lines. Our intention in noting these issues is not to say existing systems are wrong (to the contrary, many provide very robust analysis capabilities), but to emphasize that PMUVis strives to employ cohesive design practices. We use effective and expressive channel encodings, as well as employ well-known visualization techniques, such as heatmaps, in a novel application for anomaly detection.

Future visual analytic systems for the power grid can therefore benefit from collaboration with the visualization community. For example, industry applications could incorporate state-of-the-art dynamic network visualization techniques [2]. Additionally, techniques used in platforms that visualize a combination of cyber-physical sensor network data and live human reports, can be employed. For instance, [16] employ word streams and radial visualizations for the situational analysis of earthquakes. [17] incorporate climate sensor data into a pedestrian (tourist) navigation app using contour maps and path highlighting, to account for extreme weather conditions.

Visualizing grid events. Grid events refer to irregularities in power systems such as line/generation losses, voltage drops, oscillations, etc. Visual analysis of grid events generally combines backend processes with frontend techniques that show identified anomalies. For example, industry platforms such as BTRDB monitor PMU streams, and will alert operators when monitored values exceed thresholds, network connectivity drops between components, or anomalies are detected using ad-hoc algorithms [7]. However, current methods require high operator knowledge and overhead in processing. This regularly results in flagging of events further downstream where they can cause significantly more disruption (such as blackouts), necessitating efficient visualization methods to streamline event analysis.

When analyzing events, signal processing methods such as fast Fourier transform, matrix-pencil, and spectral analysis are a common approach for characterizing the time evolution of PMUs that are potentially relevant to an oscillation event. For example, Idehen et.al. [9] use a matrix-pencil technique to study large scale oscillation modes. The network's transient stability is visualized with line charts, network contour maps, and a "mode quality cost" function, which groups modes based on shape and frequency coherency. Similarly, PMUVis allows users to interactively apply signal processing to PMU data to analyze grid events, where selected PMU streams are decomposed using FFT and autocorrelation. In contrast to the above systems, we focus on egocentric visualization and analysis using a paired dataset of historical events, though existing event analysis techniques can be integrated into the system as future work to support real-time operations (see Discussion section).

Requirements for Scalable PMU Visualization

To better understand current practices and challenges for visualizing PMU data in the power grid, we surveyed three industry professionals who perform research and operations for U.S. power companies, and have extensive experience in the grid domain (13, 15, and 32 years). For survey responses, we note the number of practitioners who responded affirmatively in parentheses.

Challenges for visualizing PMU data	
(3) Retrieving data quickly and efficiently.	*
(2) Visualizing many PMUs at once.	*
(2) Linking events to visualized data.	*
(2) Tracing event propagation through the network.	*
(2) Isolating anomalous/outlier PMU behavior.	*
(1) Visualizing PMU data for a long duration.	*
(1) Identifying missing data.	*
(1) Streaming PMU data for real-time decision-making.	

Table 1: Our survey with power systems professionals identified the above items as salient challenges for visualizing power grid data. We note the number of users who referenced each challenge in parentheses. Items addressed by PMUVis are noted with an asterisk.

Each practitioner utilizes a variety of tools to visualize power grid data as a part of their job. Mentioned tools include software specifically developed for power systems (3: including RTMDS, SEL Synchrowave, ASPEN, PSS/E, DSATools, openECA, and OpenPDC), software developed in house (1), and general-purpose software/scripting tools/languages (3: mentioned software includes Excel, MATLAB, and Python). While platforms such as RTDMS are engineered to capture and display PMU data streams in real-time, and support backend modules for identifying anomalous behavior like islanding detection [3], they lack visualizations specifically tailored for event analysis.

Common data visualization techniques employed by the practitioners include line charts for showing raw PMU streams (3), node-link diagrams to show the grid network (3), 3D contour maps to show voltage information (2), line flow charts that superimpose scaled size arrows on lines in the network architecture to indicate the direction and magnitude of power flow (2), and bus schematic plots with overlaid heatmaps and voltage contour plots to identify an acceptable system state for the schema (2).

Table 1 lists the primary challenges the practitioners referenced when visualizing PMU data. Fast and customizable data retrieval was mentioned as a significant challenge by all three practitioners, though aspects of event analysis were also highlighted: tracing event propagation throughout the network (2), isolating anomalous PMU behavior (2), and linking events to PMUs (2).

Based on discussions and the identified challenges, many of which deal with event analysis, we outline a set of five design requirements (**DR1–DR5**) for visual analysis of the power grid using PMU data.

(DR1) Flexible and interactive data retrieval To enable interactive visualization, systems support real-time, customizable data retrieval. While the grid network is largely static—as PMUs primarily only go offline due to maintenance or severe unexpected events such as weather-related line faults—data storage and query can be primarily tailored around PMU data streams with different temporal granularities.

(DR2) Familiar techniques for familiar tasks For "common" operator visualization tasks, such as *show this PMU's data over time*, familiar techniques already exist. As operators are non-expert visualization users, techniques should be straightforward to minimize the learning curve.

(DR3) Scalable PMU visualization for trend and anomaly analysis While line charts can show PMU streams over time, this technique quickly becomes cluttered and noisy as either (i) more PMUs are added to the chart or (ii) the temporal duration of the plot increases. Alternative techniques that enable scalable visualization of hundreds (or thousands) of PMUs streams over long time durations will better support pattern, trend, and outlier analysis, including dropped or missing data.

(DR4) Link events to PMUs To support event analysis, especially an understanding of where events begin and how they evolve throughout the network, an event's origin should be sourced to a specific region within the network topology—such as to the PMU that is at (or adjacent to) an event's source. We refer to this PMU as the "event ego" PMU for the event.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/MCG.2022.3171506, IEEE Computer Graphics and Applications

(DR5) Support analysis of event propagation through the network Due to the high sampling frequency of PMUs, network imbalances and oscillations can be measured as they propagate between PMUs in the network using noise and voltage sequence techniques. Given the short duration of such events, propagation is difficult to track effectively when visualizing the raw data. Therefore, interactive visualization of signal processing techniques can be tailored specifically to analyze event evolution. For PMUVis, we compute the FFT and autocorrelation of the ego PMU as well as other user-selected PMUs, and compare the ego and selection against each other to see how an event affects the surrounding network.

PMUVis is designed to support these requirements; to help ensure this, during development we iteratively consulted our surveyed experts to assess the platform's design and user experience. In the following sections, we apply and illustrate how these principles **DR1–DR5** foster event-focused analysis of the power grid through the data management and frontend implementation of the PMUVis system.

Data Storage and Management

PMUVis is implemented using a historical PMU dataset from a U.S. gas and electric company consisting of ∼500 PMUs over a three year period. Each PMU records measurements for 18 attributes at a 30 Hz frequency. (At our institution, access to this PMU dataset is provided under non-disclosure agreement provisions, so we anonymize certain features such as PMU IDs and locations.) As this raw PMU data is very large and noisy (total raw dataset size is over 70 TB), we describe here the steps we take to aggregate, store, and retrieve it (**DR1**), as well as our use of signal processing for event analysis.

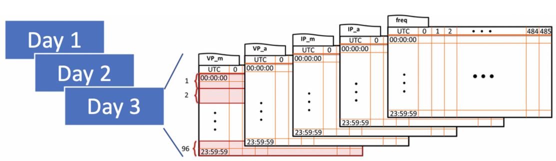
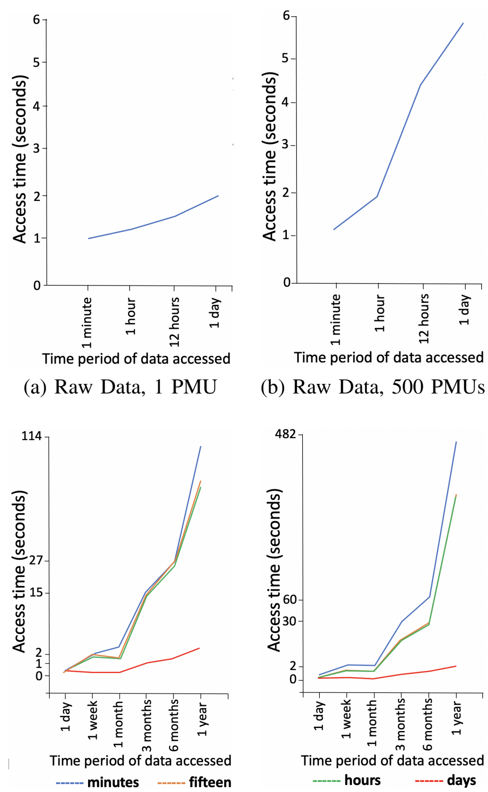


Figure 3: The storage format for PMU data. Raw data is saved into Parquet format (1 file equals 1 PMU attribute for 1 day), allowing fast column-wise reads for small time durations.

Storing and Accessing Raw PMU Data

Raw PMU data is saved on our compute cluster's servers using the Parquet file format. Parquet is a columnar storage scheme based on an algorithm for record shredding and assembly [14], that supports efficient and flexible compression and encoding. Blocks in a Parquet file are stored in the form of row groups, each of which in turn contains column chunks. In a column chunk, column values are stored in contiguous memory locations. As Parquet compression is transparent and can be adapted in a column-specific manner, column-wise queries can fetch values with much higher performance compared to row-wise queries.



(a) Raw Data, 1 PMU (b) Raw Data, 500 PMUs

(c) Aggr. Data, 1 PMU (d) Aggr. Data, 500 PMUs

Figure 4: Access times for PMU data stored in Parquet format. We plot access times for (a) one PMU and (b) all (∼500) PMUs when accessing two attributes (VPm and IPa). We plot access times of aggregated (aggr.) PMU data (each line represents a different aggregation granularity) for the same two attributes (VPm and IPa) over longer durations for (c) one PMU and (d) all PMUs.

In our dataset, each Parquet file stores one

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/MCG.2022.3171506, IEEE Computer Graphics and Applications

attribute's worth of data for all PMUs for one day (Figure 3). The files are subdivided into 96 fifteen-minutes groups, such that each group contains 27k rows. Each column corresponds to one PMU, and each row is one timestamp. As the PMUs record measurements at 30 Hz (raw values are 64-bit floating point values, null values are recorded during data dropout). Each Parquet file is ∼500 columns × 2.592M rows. Since each PMU records 18 attributes, each day contains 18 Parquet files.

Data stored in Parquet files can be compressed and decompressed on-the-fly using Gzip. Because of this, the compressed size of our Parquet files generally varies between attributes. For example, in compressed format, files for the magnitude attribute are generally between ∼1.5-2.5 GB. When reading data from a Parquet file, uncompressing a row group results in reading ∼20MB of data. Storage-wise, each day's set of 18 Parquet files is ∼50–70GB.

Figure 4(a–b) shows access times based on the number of PMUs and the time duration. Access times quickly become slow, prohibiting interactive visualization of raw PMU data over long timespans. Therefore, to enable interactive analysis of longer timespans, we utilize a hierarchical aggregation scheme for computed PMU statistics.

Aggregating PMU Statistics for Fast Querying

Our aggregation approach mimics the Parquet file structure for the raw data, only now storing aggregate computed statistics (Figure 5). This decision to maintain data storage and access via Parquet files was chosen based on the following reasons: (i) Aggregated data files can be stored in the same directory locations as raw data, making administration straightforward. (ii) APIs previously developed to access Parquet files storing raw PMU data can be tweaked to additionally access Parquet files storing aggregated statistics. This maintains a consistent data access API for both raw and aggregate data. (iii) Even when conducting subsecond analysis (as is done when examining grid events), PMU measurements are almost always considered over time ranges as opposed to a single measurement/timestep snapshot. Storing aggregated PMU statistics in column-wise format in Parquet files optimizes for the temporal queries. In contrast, SQL/NoSQL/time series databases would require implementing and maintaining indexes with additional overhead, thus slowing data retrieval. (iv) Besides raw data aggregation, we are able to compute aggregated statistics such as signal to noise ratio (SNR) over a sliding window and its standard deviation for all attributes, as well as positive and negative voltage sequences, simultaneously.

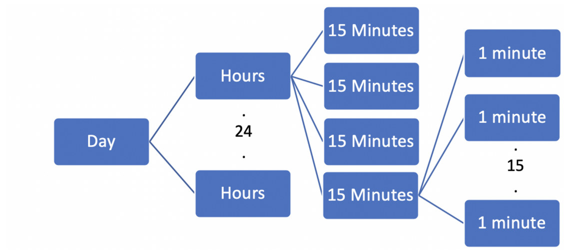


Figure 5: Storage of PMU data at different granularities. To enable long-duration data retrieval with fast access times, relevant statistics are computed at successive levels of granularity.

We consider four aggregation levels for PMU data (minute, fifteen minutes, hour, and day) by averaging over the granularity considered. Like before, we store each PMU's data to a single column, but each row now represents the respective time granularity considered. For example, by aggregating over minutes, each day's measurements are reduced to 1,440 rows. A week's worth of data can be stored to a single file using the minute granularity, with 10,080 rows and a size of 30MB. Like the raw data, each attribute is stored in a separate Parquet file. Figure 4(c–d) shows how access time increases based on the time duration and the aggregation level—we note that minute and day granularities are sufficient to efficiently conduct multi-year data retrievals (the hour and fifteen minute granularities performing equivalently well, with a less than 3 second difference than minute access time at maximum duration, i.e., 1 year). In this way, PMUVis uses data aggregation to support interactive querying similar to existing platforms for sensor data such as PingThings (https://www.pingthings.io) and BTrDB [1].

Despite this approach being straightforward, it proved successful both for facilitating fast data retrieval (**DR1**) as well as the retrieval of many PMUs for visualization (**DR3**).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/MCG.2022.3171506, IEEE Computer Graphics and Applications

Linking Event Reports and PMU Data

In addition to the PMU data, we also employ an auxiliary dataset consisting of historical event reports created by operators and technicians. These reports describe events of interest that occurred in the grid. They consist of unstructured text documents that do not provide any metadata linking the report to PMUs, substations, or locations in the grid (timestamps are the only included metadata). To link events to PMUs (**DR4**), these text reports must be associated to one or more PMUs in the network.

To do this, we first performed a regex matching between report text and PMU IDs, substation names (where the PMU is located), and PMU/substation locations, to identify specific associations between an event and PMU(s). If an association was found, the event was considered to originate from that PMU, or to originate such that the linked PMU was the closest PMU in the network. For example, if an event originated at a substation without any PMUs, the closest PMU would be the first to identify the event. Unmatched events were manually reviewed and either explicitly associated to a PMU or discarded (generally, these were cases where the event was irrelevant to PMU or grid analysis). In total, we generated a collection of 1,169 events that were matched to a PMU. While using PMUVis, when events were loaded and analyzed, if seemingly mis-linked event-PMUs were suspected, the data was reviewed and, if necessary, updated.

Interface

The PMUVis interface is shown in Figure 6. It consists of six linked and coordinated panels (A)–(F) to support the visual analysis of PMU data according to the design requirements **DR1**–**DR5**.

(A) The **control panel** allows the user to select events, query for PMU data (specifying time periods, manual/automatic selection of PMUs, levels of aggregated statistics, etc.), and update control settings for individual panels.

(B) The **network panel** visualizes the power grid network using a node-link diagram. Edges indicate power lines connecting substations (green rectangular nodes); red circular nodes attached to substations represent PMUs. Nodes and edges can be styled to indicate voltage/current levels and the presence/absence of external connections (buses, loads, etc.), as shown in Figure 7. Nodes can be positioned using either a force-directed layout or the substation's latitude and longitude coordinates (overlaid on a map layer). As node-link diagrams are found in existing grid monitoring systems, it is considered a "familiar technique" to domain operators (**DR2**).

The combination of control and network panels supports event and PMU selection. Identified events can be selected from a dropdown in the control panel; the corresponding ego PMU node is highlighted in the map. To expand the set of selected PMUs, the user can either: (i) individually select/deselect PMUs on the map, or (ii) trigger an "add hop" functionality, which selects all neighboring PMUs to any currently selected PMU. (Note that a PMU may perform multiple hops across substations or transformers without PMUs before encountering its closest neighbor PMU.) In this way, the map functions both as an overview visualization and as a mechanism to interactively select PMUs for further analysis.

(C) Individual PMUs can be plotted in the **line chart panel**. Similar to the node-link diagram, line charts are a familiar technique to domain operators (**DR2**). To avoid overplotting, one PMU attribute is shown in each chart; the user can scroll to see all charts.

(D) As a third familiar technique (**DR2**), the **substation panel** shows the schematic of an individual substation for reference and review. Substation components, including PMUs, buses, plants, loads, and two-winding/three-winding transformers, are rendered using a node-link diagram with component details displayed on hover.

(E) While line charts are suitable for inspecting individual PMUs, they suffer from several of the identified challenges listed in Table 1: visualizing too many PMUs at once and visualizing long durations has scalability issues, and it is difficult to isolate anomalous/outlier PMU behavior due to noise. Therefore, to support analysis of trends and anomalies (**DR3**), we utilize a **heatmap panel**.

Each heatmap row corresponds to one PMU with the horizontal axis showing time. Each cell represents an aggregated timestep (minute, hour, day, etc.), using color to show attribute value. (PMUVis supports several color palettes.) As rows can be shrunk to only a few pixels tall and wide, they provide excellent scalability. PMUVis

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/MCG.2022.3171506, IEEE Computer Graphics and Applications

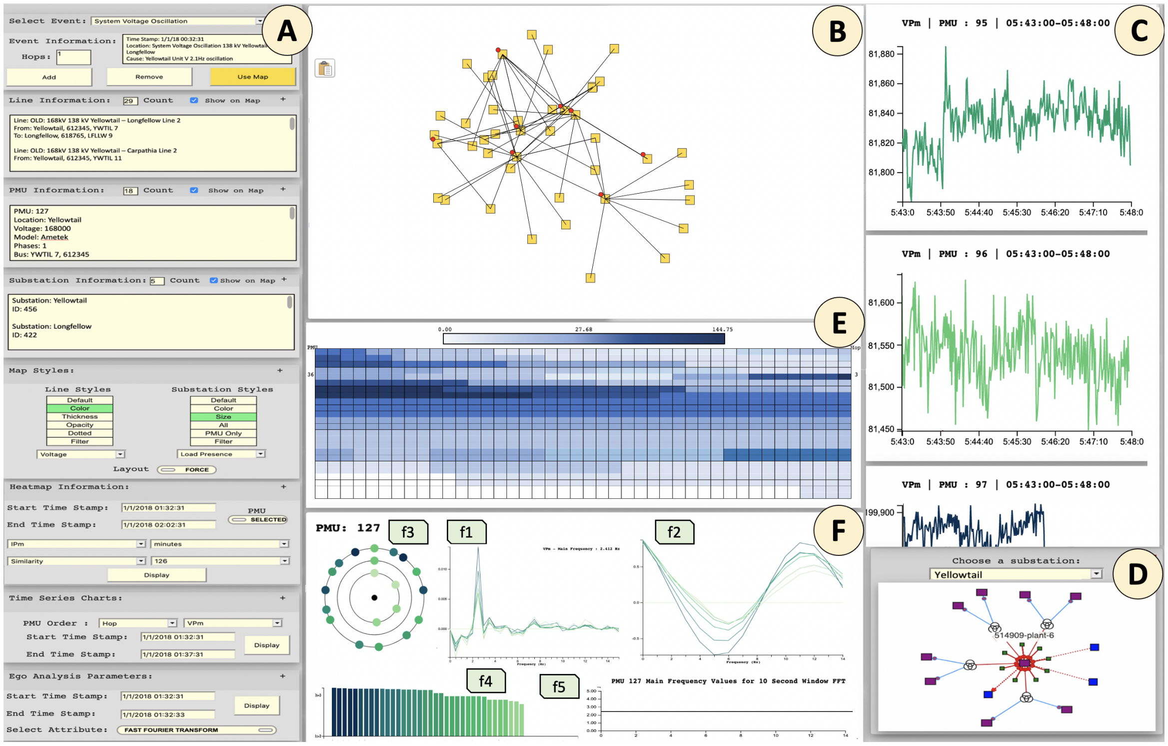


Figure 6: The PMUVis interface for visualizing PMU data and analyzing grid events. (A) The control panel supports event selection and interface settings. (B) The graph panel displays the power grid network. (C) The heatmap panel displays aggregate PMUs over long timespans. (D) The raw time series of PMUs are shown in the line chart panel, while (E) the ego analysis panel supports egocentric event analysis based on an "origin" PMU. (F) Finally, the architecture of substations can be reviewed.

supports ordering PMU rows in multiple ways: (i) by ID, (ii) by hop distance to an ego PMU, and (iii) based on pairwise PMU similarity, calculated as cosine similarity of a PMU row with the ego PMU. This approach is similar to, for example, visualization of the parameter spaces for large simulation ensembles [11] for scalable trend and anomaly analysis.

(F) Finally, the **event panel** supports propogation analysis of an event (**DR5**). It consists of five visualizations that juxtapose the event's ego PMUs with other selected PMUs.

(f1) The **FFT chart** shows the computed FFT values based on a user-selectable 2/5/10 second interval over the voltage magnitudes of the current PMU selection. For each PMU, we compute its similarity to the ego PMU based on the Euclidean distance of its FFT values. Line color is based on these values: the ego PMU (with a distance of zero to itself) is dark blue, which shades to green as PMUs become less similar. These computed similarities and color values are also used both by the other charts in this panel, and by the other panels (e.g., in Figure 6(C), the PMU in each line chart is colored this way).

(f2) The **autocorrelation chart** shows computed autocorrelation values using the same 2/5/10 intervals as the FFT chart, with line colors based on FFT similarity.

(f3) The **ego radar chart** shows several concentric rings. The ego PMU is shown as a dark blue circle at the center, and other selected PMUs placed on the surrounding rings based on their 'hop distance' from the ego (the path distance between the PMUs in the network). Position within the ring is determined based on both the FFT similarity of PMUs and physical proximity to other PMUs in the ring (i.e., if PMUs are located on the same line or substation).

(f4) Below the ego radar plot, the **similarity bar chart** shows all selected PMUs ordered and sized by their FFT similarity to the ego PMU.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/MCG.2022.3171506, IEEE Computer Graphics and Applications

(f5) Finally, the **main frequency plot** shows the main frequency, i.e., the frequency at which the maximum FFT value occurs for every 2/5/10 second window considered—as previously selected—over a five-minute interval for the ego PMU. This view helps in understanding at a macro-level how system oscillations increase, decrease, or remain consistent over long durations.

Case Studies

We now demonstrate how PMUVis can be used to analyze oscillatory power grid events via two case studies. Low-frequency oscillations are always present in large, interconnected transmission grids [18]. While usually harmless, certain fault conditions can progressively exacerbate oscillations such that they grow to a magnitude disruptive enough to lead to partial or total power system breakdown. Wide-area system monitoring using synchrophasor data from PMUs is often used to control these oscillations using anomaly detection techniques [5]. System operators must monitor the system's ability to damp such oscillations, and reduce power transfer if required, while maintaining awareness of relevant events and information on other parts of the grid not directly under their control.

Case Study: Forced Oscillation

We first analyze a historical long duration forced oscillation event. Specific steps are shown in Figure 7.

(1) Selecting the event in PMUVis' control panel highlights its corresponding ego PMU #122 in the network panel. Per the operator report, a ∼2 hour-long oscillation was noticed at the Yearling substation (where PMU #122 is located) with a beginning timestamp of 20:44. (2) To create a subgraph suitable for analyzing the oscillation's propagation, additional PMUs are selected via the add hop functionality and manual selection/deselection. (3) In total, 15 PMUs are selected within 1–3 hops of PMU #122. In Figure 7, the graph's display settings are toggled to highlight

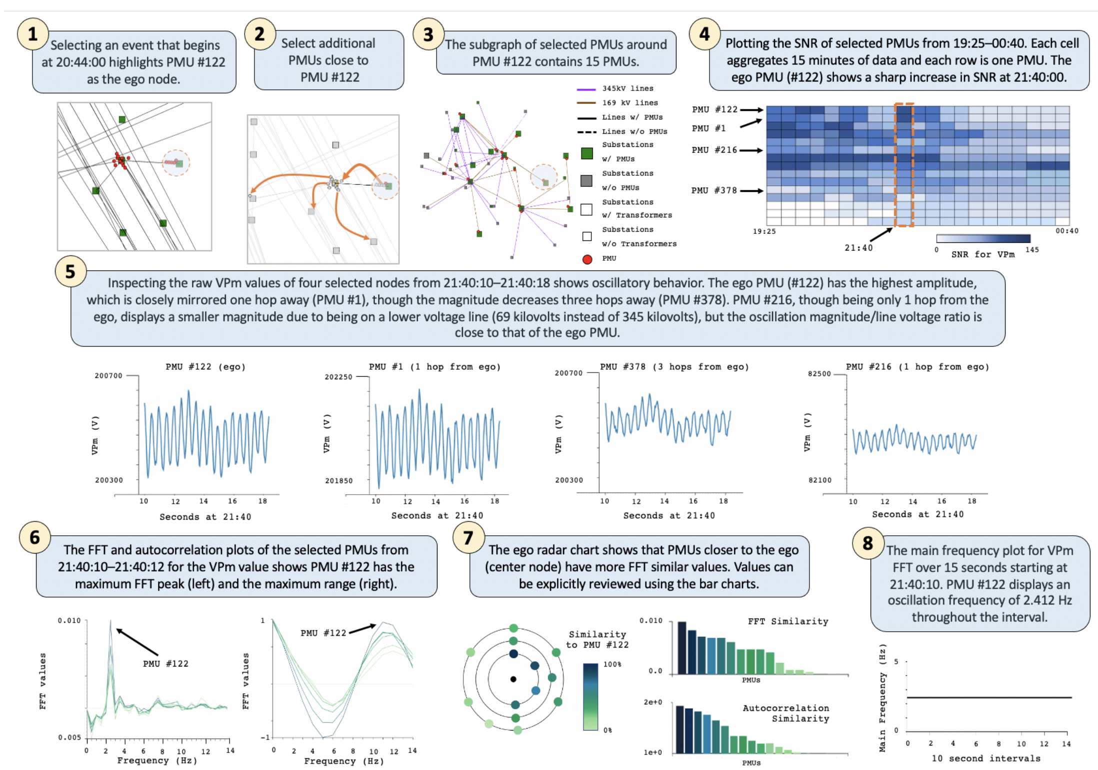


Figure 7: In the Forced Oscillation case study, (1) PMU #122 is selected as the event's ego PMU. (2–3) After selecting additional PMUs and (4) identifying 21:40 as a timestep for subsequent analysis, (5–8) additional charts show how the oscillation propagates out from the ego PMU to nearby PMUs. V: volts, VPm: positive sequence voltage magnitude.

the subgraph's equipment specifications (types of line, loads, etc.).

(4) For an overview of the oscillation, the heatmap loads a ∼4.5-hour time period starting at the event's beginning, displaying the signal-to-noise ratio (SNR) for the VPm attribute (positive sequence voltage magnitude) of selected PMUs aggregated into 15-minute blocks. Rows are ordered based on PMU hop distance and similarity to the ego PMU #122 (the top row). Color fluctuations across a row indicates a PMU is experiencing varying SNR values. Several PMUs close to PMU #122 have increased fluctuation in the left half of heatmap (during the ∼2 hours that the oscillation occurred). We select timestep 21:40, which shows a sharp increase in SNR for PMU #122, as a place for further investigation.

(5) To examine the oscillation with non-aggregated PMU data, we select four PMUs from the heatmap and display the raw VPm values in the line chart panel. Oscillatory behavior is visible in the ego PMU #122 and in PMU #1 (one hop away). PMU #378 (three hops away) shows a reduced oscillation with smaller magnitude. As PMU #216 is on a lower voltage line, despite being one hop from PMU #122, it also shows a smaller magnitude oscillation; however, the ratio between its oscillation magnitude and voltage rating is close to that of the ego PMU.

(6) The 15 selected PMUs are loaded into the ego panel. The FFT chart confirms that PMU #122 has the highest FFT and autocorrelation peaks, helping to confirm it is at the oscillation's source.

(7) Comparing the ego PMU to other selected PMUs using the ego radar chart and similarity bar charts shows that, generally speaking, PMUs closer to the ego (i.e., with fewer hops) tend to be more similar. This importantly demonstrates the oscillation's effects lessen as they propagate from the ego PMU throughout the network. (8) Finally, the main frequency plot shows a steady main frequency of ∼2.4Hz, indicating the oscillation at the Yearling substation is indeed a constant, long-duration event.

Case Study: Damped Transitory Oscillation

The first case study considered a long-duration, sustained oscillation. Here, we consider a short, damped transitory oscillation event that was not included in our dataset of historical operator reports, but was instead discovered while analyzing the previous case study. Specific steps are shown in Figure 8.

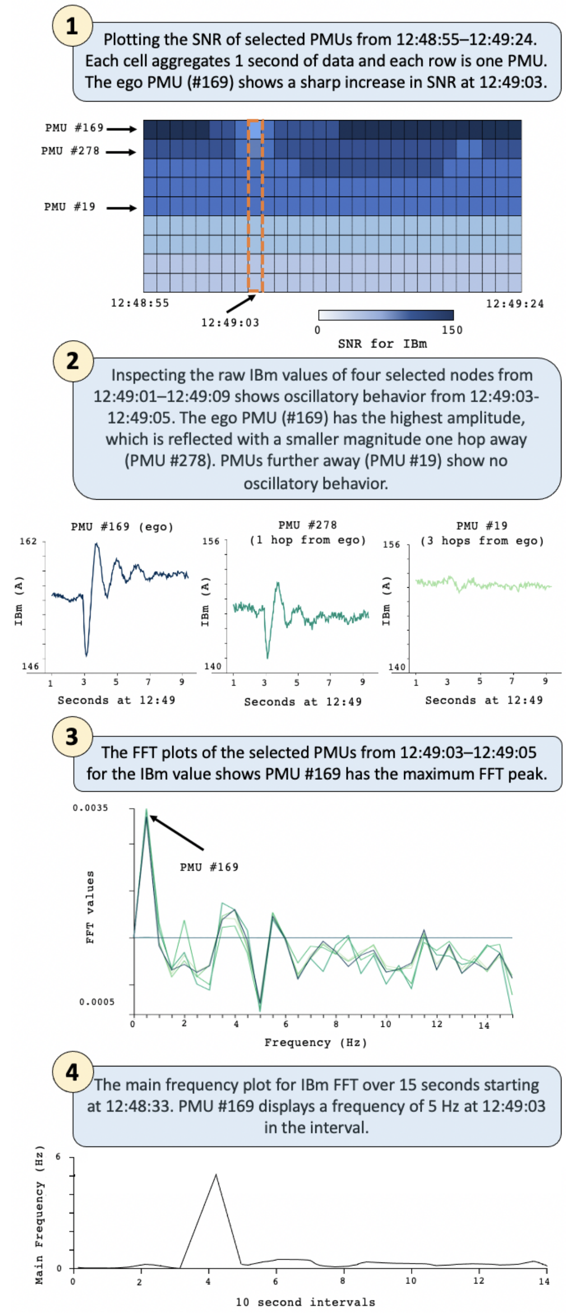


Figure 8: In the Damped Transitory Oscillation case study, (1) PMU #169 (top row of the heatmap) experiences a transitory (sub-second) oscillation, which can be seen in the (2) line charts, (3) the FFT plot, and (4) the main frequency chart. A: amperes, IBm: current magnitude of phase B.

While exploring different configurations with the heatmap, we noticed that PMU #169 had an unexplained color fluctuation. (1) Selecting this PMU as the new ego with a selection of 8 additional PMUs and reloading the heatmap at a granularity of one-second-per-cell for the IBm attribute (current magnitude of phase B) shows a fluctuation for PMU #169 that also seems to affect PMU #278 (one hop away). (2) This is also seen in the line charts panel, where the oscillation is visible in PMU #169 and with a smaller magnitude in PMU #278. The oscillation is difficult to see in PMU #19, located three hops away.

Examining the oscillation in the event panel shows this is a well-damped transitory current oscillation—one that lasts for approximately 2 seconds. (3) High overlap of FFT values is seen for immediate neighbors of the ego PMU; on further analysis, we note that neighboring PMUs are located on low voltage lines (69kV), and are connected to step-down transformers, accounting for the highly localized oscillation (which is also symmetrical over IAm and ICm). (4) The main frequency oscillation values for the ego PMU, calculated over 10 second time windows (for the time interval chosen to construct raw IBm time charts), show that the system is almost completely at rest, with only a momentary sharp increase in the main frequency of the ego PMU; the event quickly dies out as the oscillation damps, with the system returning to an "at rest" state comprised of very slow oscillations that can be filtered out using a high-pass filter. Current signals contain decaying DC components, and exhibit such damped, transitory power oscillations when transient events such as "earth faults" occur [13]. These comprise of line/generator tripping or high load fluctuation.

Feedback from Domain Experts

To further evaluate PMUVis, we conducted a two-hour demo session with two industry engineers (30+ and 1+ years of experience). These practitioners regularly conduct data-driven analyses on the same regional power grid network that constitutes our dataset; thus, they were familiar with the grid's topology and could use the raw (non-anonymized) interface.

While feedback was generally positive, we focus here on highlighting how PMUVis differs from many standard approaches used in industry. For example, PMU data storage and processing is considered a significant industry challenge, which (despite being relatively straightforward in PMUVis) is efficiently addressed via aggregated Parquet storage.

One engineer remarked that plotting PMUs directly onto substations in the network panel was appreciated: "*A lot of programs, there can be several PMUs at a substation… they don't even display them properly. Knowing where the PMU is located, having a visual is really nice. These little things can make the analysis so much easier and faster.*" Likewise, the "add hop" functionality was appreciated: "*Let's say there's a line trip. You can pick out where the PMU closest to the event typically is: there's a strong correlation between that signal energy and the source. [Then,] I'm trying to figure out: what lines tripped? It's difficult going back and forth between map and data, map and data. I really liked seeing that [hop interaction]. You can start with an ego, and then just hop around and help you find where the actual source is.*"

Interestingly, though heatmaps are well-known in the visual analytics community, its usage was considered quite innovative: "*The heatmap is a really interesting plot. …The consolidated, aggregated [aspect] seems really helpful. …[This type of view is] not really supported by tools out there today.*

The engineers were particularly excited about the ego panel as a way to characterize the impact and evolution of oscillations on both the ego and surrounding PMUs. "*Different locations have different phasing. [Showing that] is really important. …Autocorrelation, that's really helpful. You can tell the phasing between different locations.*" Using FFT values to rank oscillation impacts across PMUs was considered useful, and an apt target for future work integrating spectral analysis. "*You could do a sliding FFT signal on that one signal. Maybe even do a waterfall diagram of the spectral analysis.*" "*Spectral analysis comparing multiple events would be interesting. …Say you had 0.01 Hz sampling data superimposed in your FFT plot for different events for one PMU, now that would be useful.*"

Discussion and Conclusion

Based on feedback from our expert review sessions, PMUVis effectively supports **DR1–DR5**. In contextualizing PMUVis to the domain challenges outlined in Table 1, the primary challenge it does not address is handling *streaming PMU data for real-time decision-making*, which is a significant industry hurdle and one we are pursuing in future research. Currently, PMUVis supports ego-centric analysis of events based on historical data from PMUs, and can be used to develop new indicators and protocols by experienced operators to flag events further upstream based on historical event analysis procedures. Supporting real-time event analysis requires: (i) algorithmic integration, such as a modal or spectral analysis that can flag anomalous behavior in real time, and (ii) backend support for scalable, on-the-fly data processing, such as platforms like BTrDB [1].

A significant factor in analyzing real-time grid data is that events will not be linked to PMUs *a priori*. In this case, visual analysis of events can begin from a "non-egocentric" perspective, where identifying the ego PMU is a first step. While anomaly detection algorithms have been proposed to identify ego PMUs, recent research has indicated such approaches are ripe for misclassification [8]. Human-in-the-loop visual interfaces provide a viable strategy in this scenario, as users can employ domain expertise to review algorithm recommendations when determining the source of an event. The scalability of such an approach is partly validated in PMUVis by the usage of the heatmap panel to minimize overhead in identifying PMUs and time periods of interest during events. Additional affordances, panels, and interactions can be tailored for real-time analyses to quickly identify ego PMUs.

While real-time analytics of power grid data is a pressing concern, future visual analytics efforts can also address critical emerging research themes for the power grid, including *ante-mortem event analysis* (prediction, classification, and mitigation) and *cybersecurity forensics and countermeasuring*. Such scenarios presume streaming grid data to simulate real-time event scenarios that require time-critical and situational decision making.

Acknowledgment

This research was supported by the U.S. National Science Foundation through grant OAC-1934766.

References

1. M. P. Andersen and D. E. Culler. Btrdb: Optimizing storage system design for timeseries processing. In *14th {USENIX} Conference on File and Storage Technologies ({FAST} 16)*, pages 39–52, 2016.
2. F. Beck, M. Burch, S. Diehl, and D. Weiskopf. A taxonomy and survey of dynamic graph visualization. In *Computer Graphics Forum*, volume 36, pages 133–159. Wiley Online Library, 2017.
3. H. Chen, K. Martin, B. Bhargava, V. Budhraja, and J. Ballance. On-line islanding detection application in the realtime dynamics monitoring system. In *2014 IEEE PES T D Conference and Exposition*, pages 1–4, 2014.
4. H. Chen, J. Mo, U. Kothapa, B. Bhargava, and J. Dyer. Development of an on-line synchrophasor wide area dynamics monitoring platform. In *2014 IEEE PES T D Conference and Exposition*, pages 1–5, 2014.
5. Y. Chen. *Early anomaly detection and classification with streaming synchrophasor data in electric energy systems*. PhD thesis, 2015.
6. R. Cozby, D. Bushmitch, and M. Badger. Visualization of a performance metric against a performance threshold, Sept. 10 2019. US Patent 10,410,134.
7. M. Farajollahi, A. Shahsavari, E. M. Stewart, and H. Mohsenian-Rad. Locating the source of events in power distribution systems using micro-PMU data. *IEEE Transactions on Power Systems*, 33(6):6343–6354, 2018.
8. T. Huang, N. M. Freris, P. R. Kumar, and L. Xie. A synchrophasor data-driven method for forced oscillation localization under resonance conditions. *IEEE Transactions on Power Systems*, pages 1–1, 2020.
9. I. Idehen, B. Wang, K. Shetye, T. Overbye, and J. Weber. Visualization of large-scale electric grid oscillation modes. In *2018 North American Power Symposium (NAPS)*, pages 1–6, 2018.
10. S. K. Khaitan and J. D. McCalley. Cyber physical system approach for design of power grids: A survey. In *2013 IEEE Power Energy Society General Meeting*, pages 1–5, 2013.
11. M. Luboschik, S. Rybacki, F. Haack, and H.-J. Schulz. Supporting the integrated visual analysis of input parameters and simulation trajectories. *Computers & Graphics*, 39:37–47, 2014.

12. D. Macii, D. Fontanelli, G. Barchi, and D. Petri. Impact of acquisition wideband noise on synchrophasor measurements: A design perspective. *IEEE Transactions on Instrumentation and Measurement*, 65(10):2244–2253, 2016.
13. R. K. Mai, L. Fu, Z. Y. Dong, B. Kirby, and Z. Q. Bo. An adaptive dynamic phasor estimator considering DC offset for PMU applications. *IEEE Transactions on Power Delivery*, 26(3):1744–1754, 2011.
14. S. Melnik, A. Gubarev, J. J. Long, G. Romer, S. Shivakumar, M. Tolton, and T. Vassilakis. Dremel: Interactive analysis of web-scale datasets. In *Proc. of the 36th Int'l Conf on Very Large Data Bases*, pages 330–339, 2010.
15. A. Monti, C. Muscas, and F. Ponci. *Phasor measurement units and wide area monitoring systems*. Academic Press, 2016.
16. H. N. Nguyen and T. Dang. Eqsa: Earthquake situational analytics from social media. In *2019 IEEE Conference on Visual Analytics Science and Technology (VAST)*, pages 142–143, 2019.
17. T. Opach, C. Navarra, J. K. Rød, and T.-S. Neset. Towards a route planner supporting pedestrian navigation in hazard exposed urban areas. In *Proceedings of the 17th International Conference on Information Systems for Crisis Response and Management, ISCRAM*, 2020.
18. K. Prasertwong, N. Mithulananthan, and D. Thakur. Understanding low-frequency oscillation in power systems. *International Journal of Electrical Engineering Education*, 47(3):248–262, 2010.
19. M.-A. Sanchez-Hidalgo and M.-D. Cano. A survey on visual data representation for smart grids control and monitoring. *Sustainable Energy, Grids and Networks*, 16:351–369, 2018.
20. S. Thakur and A. Chakrabortty. Multi-dimensional wide-area visualization of power system dynamics using synchrophasors. In *2013 IEEE Power Energy Society General Meeting*, pages 1–5, 2013.

Anjana Arunkumar is a Ph.D. student at Arizona State University. Her research areas include information visualization, natural language processing, supply chain management, and visual cognition. Contact: aarunku5@asu.edu

Nitin Gupta received the Masters degree from Arizona State University. Contact: ngupta68@asu.edu

Andrea Pinceti received his Ph.D. in Electrical Engineering from Arizona State University in 2021. Currently, he is working on cyber-security and data analytics related to power systems. Contact: Andrea.Pinceti@asu.edu

Lalitha Sankar is an Associate Professor in the School of Electrical, Computer, and Energy Engineering at Arizona State University. Her research interests include the development and application of information sciences to complex networks, data privacy, and algorithmic fairness. Contact: lsankar@asu.edu

Chris Bryan is an Assistant Professor in the School of Computing and Augmented Intelligence at Arizona State University, where he directs the Sonoran Visualization Laboratory (SVL@ASU). His research areas include information visualization, human-computer interaction, and virtual reality. Contact: cbryan16@asu.edu