Data mining and playback of Hybrid Synchrophasor for Research and Education

Virgilio Centeno, PI
University: Virginia Polytechnic Institute and State University

Contact Information:
(540) 231 2045
vcenteno@vt.edu

Project Description

The aim of this project is to advance development in Synchrophasor data research and education in two fronts:

1) Facilitate Implementation of Data Mining Applications by
    a. developing and documenting procedures for applying data mining to Synchrophasor data
    b. proving and demonstrating the advantage of hybrid databases
    c. generating electrically equivalent system models untraceable to a utility and usable for educational and research applications
    d. generating a set of samples of Synchrophasor data of derived-events that are as realistic as recorded events and that will enhance the exposure of students to operation, limits and needs of power system

2) Provide a simulation tool for Synchrophasor devices that:
    a. is available as open source to academic institutions
    b. is capable of playing back data from recorded events and PSSE simulations
    c. allows simulation of multiple PMUs with a single computer
    d. allows the simultaneous use of multiple simulators to replicate the Synchrophasor outputs of any system size.

Data Mining:

Decision trees generated by classification and regression algorithms have the useful ability to select the variables to be used in forming the tree from a longer list. Classification and regression type of data mining software was developed for broad applicatio. Competing techniques like neural networks, for example, do not have the ability to select important variables. In attempting to locate Synchrophasors for a specific task in a large system classification and regression type of data mining has proven to be invaluable. The basic process followed is the creation of a database by running a simulation program like PSLF or PSSE and submitting the data labeled with the categories of interest to the analysis software (stable or unstable in the simplest power system case). The classification and regression software constructs a decision tree and uses an optimum selection of data to be used in forming the splitting nodes in the tree. The splitting nodes effectively give the location of the Synchrophasors to accomplish the task described by the database. This process has been used to provide Synchrophasor inputs to real-time, discrete-event control, predicting cascading events, voltage security, transient stability, detection of islanding, security assessment, and adaptive security dependability of relays.

The fact that major disturbances are rare makes it necessary to study some aspects of power system behavior with computer simulation. Particularly if model validation can be accomplished there are significant advantages to be gained by combining real Synchrophasor data with simulation results to find a hybrid database that is a candidate for data mining. In the recent past most decision trees have been found using only data from simulations. The approach for this project is to select simulation cases that will produce a balanced database (similar number of bad events and normal operating conditions). That is, although the event is rare in the real world a data base t

As the number of Synchrophasors increase the amount of archived Synchrophasor data will grow. The intent of this proposal is to merge the archived real data and the results of simulations to form a hybrid data base. In addition, the database will contain system-wide data at periodic intervals, and simulation data of a large number of possible events from properly validated system models. For example, after seeing a rare but interesting event rather than waiting for a reoccurrence of the event we can add a sufficient number of simulations of similar events so that data-mining can suggest a solution. Cascading outages are almost by definition rare events but they are exactly what we would like to have represented in a database. Dominion is interested in base lining the phase angle behavior in their territory. Data will be mined to extract mode shapes, damping ratios, and frequencies facilitating establishing warnings/alarm thresholds for operator action. In addition all of the Synchrophasor measurements will be correlated with the system performance measures for normal operating conditions and its variations over a period of time and during various limiting conditions like thermal limits, proximity to voltage instability or voltage collapse, transient stability, etc. This kind of analysis will facilitate establishing warnings/alarms thresholds for voltage phase angle measurements and determine site pairs of interest that are important in revealing systems stress and will recommend upper and lower limits for normal operation.

hat includes many extreme or rare events (from real events and simulations) is appropriate for creating decision trees.

Synchrophasor Data

The future training, research, and development on Synchrophasors will concentrate on the data provided by Synchrophasors and the problems associated with transmitting, filtering and using large amounts of time synchronized data. To enhance understanding these problems for teaching and research, educators need to have access to Synchrophasor type data. At the moment access to such data is only possible by signing restrictive NDA agreements with large utilities or by acquiring expensive real time data simulators that provide data output in the C37.118.2 Synchrophasor standard format. It is difficult for universities and specially for small research groups and for non-research schools to obtain access to real time Synchrophasor data from utilities. The minimum cost for an RTDS system capable of producing Synchrophasor data is over $200k and very few institution already have such a system.

Virginia Tech worked under agreement with PJM on the testing of PMUs and Phasor Data Concentrators (PDCs) used by the utilities in the PJM territory. As part of this project Virginia Tech developed a PMU testing system in collaboration with NIST. In addition to the PMU testing system Virginia Tech has developed a simple PMU simulator use to test loading on PDCs. This simulator was developed based on an existing Open Source iPDC PMU Simulator [23] Simulator. This simulator was design by iPDC to test

simple operation of their open source Phasor data concentrator. The software has been used and modified by Virginia Tech for the testing of Phasor Data concentrators as part of the PJM testing project.

The first task for the development of a simulator is to determine the time delay and time interval of the data output of commercial Synchrophasor units. Virginia Tech has 9 different commercial Synchrophasor units and has a Synchrophasor Testing system that can be modified to determine these required values. Since Synchrophasor output is limited to a maximum reporting rate of 60 messages per second it is expected that the synchronization requirements will not be as stringent for the simulator as they are for the Synchrophasor devices since the Simulator is not require to sample the signal but just to send the data within the observable delay window. For the proposed simulator is expected that the internal timing system of the computer could be synchronized to a 1pps of a GPS receiver and to available Network Timing Protocols (NTPs) with enough accuracy to emulate the typical time pattern of Synchrophasor devices. Preliminary tests at Virginia Tech has shown that the overhead of Microsoft Commercial operating systems generates large errors in synchronization but open access operating system bases on Linux produce very small errors in time Synchronization of the computer internal clock.

Another aspect of this subproject is the conversion of PSSE data to Synchrophasor data. In order for the simulation to be performed correctly the simulation point data (actual wave form) must be converted to a Phasor through a Synchrophasor algorithm that uses a data window to compute the Phasor as is the case on existing Synchrophasor devices. In addition, for correct operation a larger data window of simulation data must be used to compute the frequency of the signal similar to what a Synchrophasor unit has to do with a sampled signal. At the end of the project the Virginia Tech team will make the Synchrophasor simulator software and documentation available for academic use. This release will include Matlab data conversion software to translate PSEE simulation to the data format developed for the simulator.

Expected Research Outcome:

a) An example and a set of procedures on mining of Synchrophasor data with hybrid databases that will promote and enhance research in data mining in power systems.
b) An open source software that allows the ability to generate valid synchrophasor data for research and analysis on the interaction of Power Network and its communication network.

Expected Education Outcome:

a) A set of representative “derived” system events with their corresponding and sharable system model that allow instruction and research with realistic events without the requirements of a restrictive Non Discloser Agreement with a utility
b) An Open Access software and hardware requirements to implement a GPS Synchronized Synchrophasor simulator with a standard computer that can be used by academic institutions to instruct and research on the areas of real time protection and control of power systems.

Areas offered as opportunities for possible Collaboration with other projects

Other than collaboration in the PIs have no knowledge of the other projects and cannot offer advice on opportunities for possible collaborations.