An Emulation Environment for Prototyping PMU Data Errors

Ikponmwosa Idehen, Zeyu Mao and Thomas Overbye
Electrical and Computer Engineering
University of Illinois
Urbana-Champaign, USA
{idehen2, zeyumao2, overbye}@illinois.edu

***Abstract*— High quality synchrophasor data ensure that mission-critical applications produce trustworthy results during grid monitoring. Hence, there is a need for risk assessment activities to screen incoming data for errors, and decide if erroneous portions of data be either rejected or passed on to the dependent-applications. Visualization of prototyped PMU data errors provides a means for better identifying and understanding error mechanisms in real-life PMU data. The objective of this work is to develop a framework for the simulation of prototype PMU data errors within an emulation environment. In this work, two errors have been simulated: error due to a time skew, and error due to mislabeled flag bits during a leap second transition.**

Keywords—phasor measurement unit (PMU), error emulation, risk assessment, prototype, noise

I. INTRODUCTION

Phasor measurement units (PMUs) provide precise, periodic and time-synchronized measurements from the power grid. These measurements, also known as synchrophasors, comprise voltage, current, frequency and rate of change of frequency (ROCOF). They are transmitted to control centers where they can be utilized by real-time, mission-critical applications for monitoring and control. Thus, high quality data enhances the grid operations. However, to provide assurance of high quality, risk assessment activities are employed to detect errors in these data prior to their usage by dependent applications. Error identification ensures that erroneous portions of data are either rejected (in order to prevent the application from processing wrong data) or accepted (with the assumption that acceptable performance limits of an application are not violated).

References [1], [2], define data quality to encompass aspects of accuracy, timeliness and availability. Timeliness is the attribute by which data is available when required by the dependent-application. It is accurate when it fully represents the physical phenomenon being measured. The process of risk assessment then involves the identification and categorization of those entities (errors) which affect any of the data quality aspects.

Several errors in synchrophasor data have been reported in the literature. References [3], [4], [5] mention the effects of a wrong time reference, harmonic interference and noise on the quality of data obtained from the PMU. Other errors include biases introduced by instrument transformers, algorithms used in phasor estimation, communication latency, bandwidth, data drop, wrong data alignment, limited capacity and operating speeds of application storage [6], [7], [8]. [9] also reports on how the increased deployment of endpoint phasor applications can reduce data quality.

Given these different sources of errors, the objective of this paper is to develop a framework that emulates different mechanisms of PMU data error. This is achieved by the specification of a prototype emulation environment where these errors are simulated. This aids activities in risk assessment which involve the identification of error patterns. PowerWorld Dynamic Studio (DS) [10] is used to run transient stability (TS) level simulations on an IEEE 118-bus system, from which C37.118 data is obtained. Considered perfect measurements, these data are exported into Matlab, where they are read by a client program. Data modification, in accordance with a user-specified error type, is then carried out to simulate the PMU data error.

This paper is organized as follows: synchrophasor error sources, based on data attribute categories, are discussed in Section II. Specification of the error emulation environment, and two error types are presented in Section III. Section IV discusses the aspect of noise injection in PMU data, and the conclusion and future work are stated in Section V.

II. ERRORS IN THE SYNCHROPHASOR NETWORK

The data flow in a typical synchrophasor network is illustrated in Fig. 1 [11].

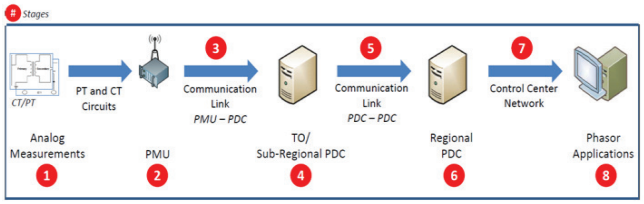

Fig. 1. The syncrophasor network data flow

From the instance of signal sampling by the instrument channels, through to the PMU which provides synchrophasor measurements, the phasor data concentrator (PDC) which aggregates and aligns these measurements, and finally to application usage, the synchrophasor network remains error-

978-1-5090-3270-9/16/$31.00 ©2016 IEEE

prone. Reference [1] categorizes PMU data using three levels of attributes: attributes of single data points, dataset and data stream availability. Attributes of single data point are concerned with the accuracy of the individual, time-stamped measurement. Inherent and extraneous measurement errors, synchronization issues, and positional metadata of PMUs are among the issues that affect data quality at this level. Table 1 shows error sources for different attributes of PMU data [12].

TABLE I. PMU ERROR SOURCES

Categories	Error Sources
Single Data point	Accuracy, Noise, Phase-error, Harmonic Distortion, Estimation Algorithms, Asynchronous local behaviors (e.g. Time-skew), Instrument error
Dataset	Status code error, Improperly configured PMUs, Abnormal or loss of PDC configuration, Frequency calculation discrepancies, Mislabeling due to erroneous timestamps, CRC error, Invalid timestamp
Data stream	Network limitations - Data loss or drop-outs, network latency; Increase in endpoint applications

Attributes of dataset relate to features of aggregated data points usually at the PDCs. This corresponds to stages 4 and 6 in Fig. 1. At this level, trends in dataset are detected by examining status codes of PMUs and PDCs.

Attributes of data streams concern the ability of the network to deliver timely and reliable data streams to the dependent application. Quality is affected by the underlying communication link through measures such as the data drop-out and message rates.

III. THE ERROR-EMULATION ENVIRONMENT

In order to develop a prototype of PMU errors that have been introduced in Section II, an environment is proposed, and illustrated in Fig. 2. This environment is based on the IEEE C37.118 communication protocol [13], [14]. It is categorized into data acquisition and storage, error modification and visualization stages. The process for prototyping errors begins with the introduction of measurements we assume to be perfect.

We develop synthetic test cases using PowerWorld simulator, a power grid simulation software designed to simulate high voltage power system operations [11]. In order to capture streaming phasor measurements in C37.118 format, a TS simulation is run on a test grid. The TS engine runs an interactive power systems simulation, and can accept commands from other client programs in real time. The graphic user interface (GUI) block provides a medium for controlling the amount of C37.118 frames that are read given a defined time period. It also accepts user inputs, such as desired error to be simulated and error parameters. *CMD1* initiates C37.118 frame transition from PowerWorld DS. These data frames are stored in *Datastore1*, emulated as a cell array in Matlab. *Datastore1* provides the amount of data, as required by the error modifier block, during error emulation. Based on user's selection, the block, *Err* is activated in the *Error Modifier* where appropriate data modification and C37.118 field updates are carried out. Modified data are *re-packaged* to the binary coded decimal (BCD) format, and stored in *Datastore2*. For visualization, Matlab provides a platform to visualize both trends of the actual and erroneous data. Finally, in order to view global patterns of these signals, techniques such as data and cluster contouring are employed. An advanced display such as PowerWorld software can be used for this purpose.

The format of data flow is such that red and black arrows indicate paths for C37.118-formatted and decoded data frames respectively. User-issued commands are represented by the blue arrows.

A. Case Specification

The use of PowerWorld software ensures cases are run in TS mode by the simulation engine, and C37.118 data injected into the emulation environment by the DS. A dynamic case for a 138 kV, IEEE 118-bus system has been set up for a 20 second TS run. Table II shows some of the C37.118 options configured for the case.

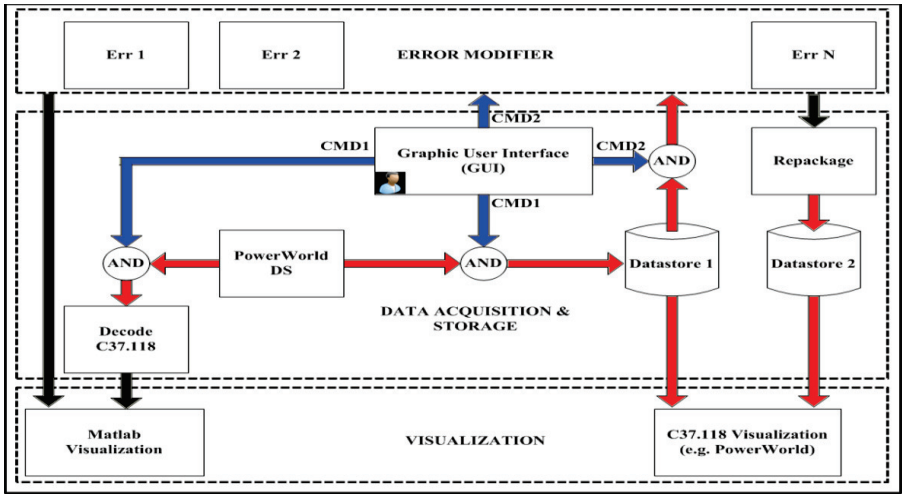

Fig. 2. The PMU error emulation environment

TABLE II. THE C37.118 OPTIONS

Option	Value
C37.118 Protocol Port	4712
IDCODE	1
PHNMR	118
AHNMR	0
Configuration Frame 1 & 2 Size	2414 bytes
Time Step (Sampling period)	0.5 cycles (0.0083333 sec)
Data Rate	120 frames/sec

Only one data stream (IDCODE = 1) is identified for the case, thus, the values of all estimated 118 phasors are located in this stream. However, no analog value is reported. The simulator provides an option to save results at every *n* time step during the TS run. For the emulation environment, *n* has been chosen to equal 1, thus ensuring that all 120 frames are saved during TS simulation. To provide more flexibility in the environment, an option that allows a user to select a preferred C37.118 reporting rate, even after the TS run, has been incorporated. Thus, reduced reporting rates of 5, 10, 20, 30 and 60 frames/sec are made available to the user. The process of decoding the values of phasors, frequency and rate-of-change-of-frequency in the data frame is carried out with the aid of the configuration data obtained from the PMU [14].

B. Frequency and Rate-of-Change-of-Frequency (ROCOF)

An assumption for this work is that the TS simulation provides error-free sets of phasor measurements. Thereafter, modification of phasor magnitudes and angles are carried out to simulate PMU errors. [13] does not specify a particular method by which device manufacturers should estimate frequency and ROCOF. As a result, an approach based on [15] for balanced sets of three-phase inputs has been utilized for this work.

Let $\omega(t)$, ω_o, $\Delta\omega$ and ω' denote the instantaneous, nominal, deviation values and rate of change of angular frequencies respectively; and φ_o, $\varphi(t)$ denote the values of the initial and instantaneous phase angles respectively. It follows that:

$$\omega(t) = \omega_o + \Delta\omega + t\omega' \tag{1}$$

$$\phi(t) = \int \omega(t) = \phi_o + t\omega_o + t\Delta\omega + \frac{1}{2}t^2\omega' \tag{2}$$

Since the nominal angular velocity is uniform for all phase angles, this term can be neglected. As a result, we have:

$$\phi(t) = \phi_o + t\Delta\omega + \frac{1}{2}t^2\omega' \tag{3}$$

(3) is a quadratic expression in *t*, and may be expressed as:

$$\phi(t) = a + bt + ct^2 \tag{4}$$

where *a*, *b* and *c* correspond to ω_o, $\Delta\omega$ and (1/2) ω' respectively. Solving for *a*, *b* and *c* using least-squares methods, the frequency deviation and ROCOF are evaluated as follows:

$$\Delta f = \frac{b}{2\pi}(\text{Hz}), f' = \frac{c}{\pi}(Hz/s) \tag{5}$$

The actual frequency, f_{act} is then computed as:

$$f_{act} = f_o + \Delta f \tag{6}$$

C. Simulating PMU Errors

Based on Table I, the mechanisms for time-skew error and mislabeled bits which affect attributes of single data point and dataset respectively are presented.

1) Time-Skew Effect

Time synchronization, based on the coordinate universal time (UTC), offers numerous advantages to PMUs over the conventional SCADA measurements. The time-skew effect is a result of mis-synchronization between the global positioning clock (GPS) and the sampling clock of the PMU. It can be caused by any of the devices.

Assuming a reporting rate of *n* frames/sec, Fig. 3 depicts two sets of phasor measurements being time-tagged [4]. Time-reference is provided for the first measurement (shown in the upper portion of the figure) by an accurate PMU sampling clock. A constantly drifting clock, with a time delay ε (in microseconds), provides a time reference for the second measurement (shown in the lower part of the figure). At the beginning of a second, a 1 pulse per sample (1-pps) signal is received, and the PMU is calibrated – the timestamp of the first sample in a reporting frame is thus accurate. An internal clock then allocates this signal among the other phasors, defined by the reporting rate.

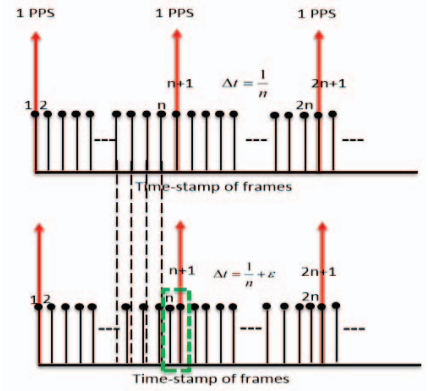

Fig. 3. Two phasor measurements for an accurate and erroroneous PMU clock

This is accurately done for the clock of the first measurement, with period Δt. However, a constant error, ε is introduced in the clock of the second measurement, which then accumulates until the last phasor is reached. A synchronizing 1-pps signal in the next reporting window results in a large time offset ((*n* -1) × ε). The time error introduced at any i^{th} data point in a reporting window is given as,

$$\Delta t_e^{\,i} = (i-1)\varepsilon \tag{7}$$

Let f_o be the nominal frequency, then the bias phasor angle introduced due to the clock error is obtained from (8).

$$\phi_e = 2\pi f_o \varepsilon + 2\pi \Delta f \varepsilon \text{ (Radians)} \quad (8)$$

$$\phi_e^\circ = \frac{\phi_e}{\pi} \times 180^\circ \text{ (Degrees)} \quad (9)$$

In steady state systems, the second term in (8) is often negligible so that $\Delta f = 0$.

Hence,

$$\phi_e^\circ = \frac{2\pi f_o \varepsilon}{\pi} \times 180 = 360 f_o \varepsilon^\circ \quad (10)$$

Considering a time error of ε sec, the phase error introduced at i^{th} data point is determined in (11)

$$\delta_e^{\,i} = (i-1) \times \phi_e^\circ \quad (11)$$

A similar analysis can be carried out for cases with intermittent GPS pulses and non-linear drifting of PMU clocks. Algorithm 1 below illustrates the procedure used to simulate the error in the emulation environment.

The right display in Fig. 4 illustrates the time-skew effect obtained for a 5 µsec sampling clock time delay. The signal duration and reporting rate are 10 sec and 30 frames/sec respectively. The dotted black circle indicates the data points at which a large offset in the phase angle, and spikes in both frequency and ROCOF signals are observed. Since the constant error is associated with the sampling clock, this phenomenon is observed periodically at every 1-pps resynchronization time. Loose and long cable connections to the GPS receiver, intermittent GPS pulses, inaccurate or degraded PMU sampling clocks GPS shields have been reported as causes of time synchronization issues [16]. Other global causes include the effects of solar flares, and a recent case of GPS time anomaly due to a malfunctioning satellite [17].

2) Mislabeling of Flagged Bits

The status of phasor measurements are conveyed through flagged bits in the IEEE C37.118.2-2011 protocol. Some of these flag bits are found in the FRACSEC field of the data frame [14]. Quality of reported time and leap second indication are provided by flagged bits of the message time quality. To ensure UTC time remains in sync with the irregular rotation of the earth, the addition or deletion of a leap second is implemented. Bits 7-4 of the first byte in the FRACSEC field have been set as indicators for this purpose. Table III shows each flag bit and definition.

Algorithm 1: PMU Linear Drifting Clock

Input: ϵ (in µsec), PMU, Stream of C37.118 data frames, Start Transmission Rqst, DATA_ RATE, Conversion factor (*k)* in config frame, Time duration in sec, T_d

Output: Time-skewed phasor angles

Compute: $\Delta\delta = \epsilon \times f \times 360$ % Estimate phase error

TOTALFRAMES = $T_d \times$ DATA_RATE

Stream of Incoming Data Frame

```
for i = 1: TOTALFRAMES
    set j = i modulo DATA_RATE
  Extract: 4-byte phasor for PMU (in PHASORS field)
    Convert: BCD phasor → Decimal → Polar form
    Extract: δj
    if j = 0 Do nothing
    else
        Update: δj = δj + (j-1) * Δδ
                Compute new frequency and ROCOF
    end
    Re-package: Polar Form → Decimal → Modified BCD phasor
    Insert: Modified BCD into PHASORS field
end
```

TABLE III. C37.118 FLAGGED BITS FOR LEAP SECOND TRANSITION

Bit #	Definition
7	Reserved
6	Leap Second Definition – 0 for add, 1 for delete
5	Leap Second Occurred – set in the first second after the leap second occurs and remains set for 24 h
4	Leap Second Pending – shall be set not more than 60 s nor less than 1 s before a leap second occurs, and cleared in the second after the leap second occurs

Setting the bits during leap second transitions is essential for time and reporting accuracy. However, mislabeling of these bits leads to an error that affects a complete data set. The response of eight PMUs to leap second was investigated in [18].

We consider the case of a PMU having a report rate of 60 frames/sec, and whose flagged bits were never set all through the leap second transition. Table IV shows the expected and actual (in parenthesis) bit settings during a leap second transition. Assuming accurate bit settings, the SOC should be repeated for the leap second (i.e. 8799). However, no flag bit was set and a new SOC (i.e. 8800) was reported during the leap second transition. As a result, there was no phasor data reported for that additional second.

TABLE IV. LEAP SECOND TRANSITION

	SOC	FS	Time of Day	Bit 6	Bit 5	Bit 4
	8799	0.9833	23:59:59:9833	0	0	1(0)
LS	8799 (8800)	0.0000	23:59:60:0000 (00:00:00:0000)	0	0	1(0)
LS	8799 (8800)	0.0167	23:59:60:0167 (00:00:00:0167)	0	0	0

LS = Leap second, SOC = Seconds of Century (Only the last 4 digits have been shown), FS = FRACSEC

Fig. 5 shows the number of reports with respect to the time-of-day (TOD). The leap second is at TOD 0, where no data was reported. Time synchronization issues resulted for subsequent samples with timestamps that were 1 sec ahead of UTC. Synchronism was restored when SOC repeated at 8816 (17 sec later), but at the expense of reporting two sets of data. An effect of this labeling error is the time alignment issues that arise during data aggregation at the PDC level. It is best observed when other PMUs report data of accurate timestamps. Other cases of wrongly-timed, leap second bit settings resulting in erroneous PMU data are documented in [18].

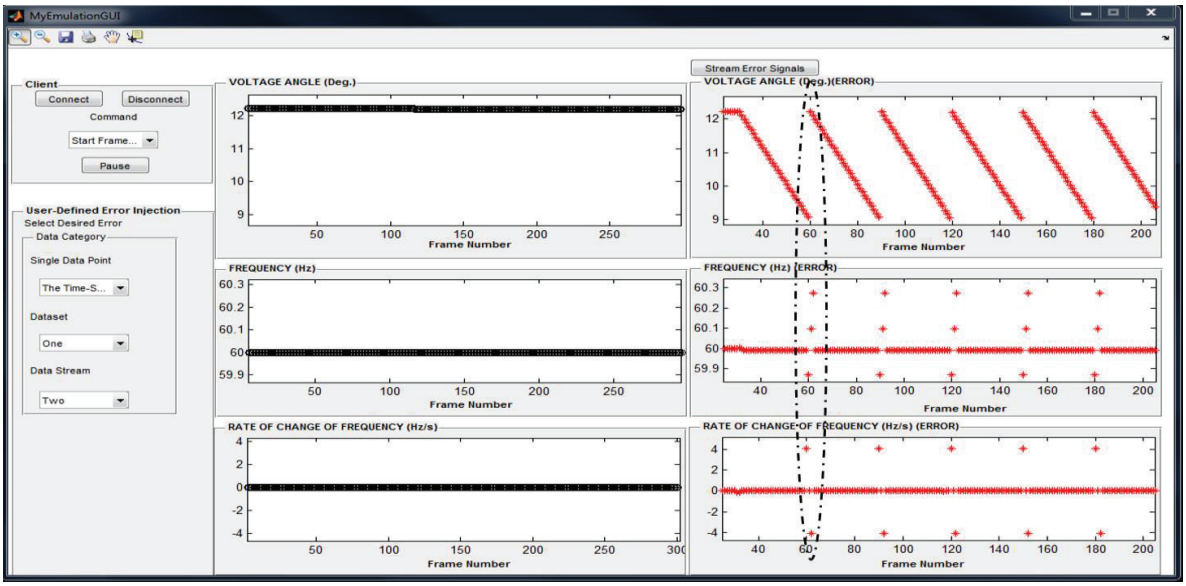

Fig. 4. The time-skew effect for a 5μs sampling clock time delay

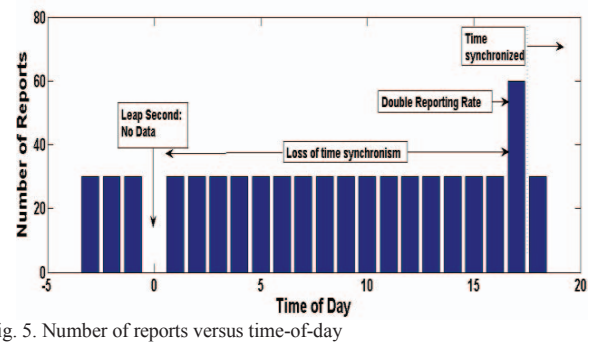

Fig. 5. Number of reports versus time-of-day

The display in Fig. 6 shows the stream of voltage measurements for three different PMUs simulated for 3 secs in the prototype environment. The report rate is set at 20 frames/sec. Leap second transition is assumed to occur immediately after the 2nd sec (i.e. 41st frame). No bit flag is set for the blue PMU during the leap second transition (LS, black arrow). It fails to recognize the addition of the second, hence reports its data at a time stamp 1 sec ahead of others. The red arrow indicates the loss of synchronization with other PMUs. Also, another effect of the loss of synchronism is observed in the reported phase, as shown in (12).

$$\phi_e = 2\pi\Delta f\,\varepsilon \text{ (radians)} \tag{12}$$

However, since the problem does not arise from the sampling clock, only the 2nd term in (8) contributes to the phase error. This error is noticeable when the actual system frequency is different from the nominal frequency.

IV. Noise in PMU Data

Noise is unwanted electrical energy that interferes with the actual measured signal of interest. Not considered actual errors, noise in PMU data can be attributed to several sources, which include the measurement process of instrument channels and interference due to external electric fields. The literature models this noise as an additive, Gaussian distribution parameterized by a zero-mean and finite variance (σ^2) [19], [20]. The corresponding power level (in decibels) is obtained from the SNR (Signal-Noise Relationship) and standard deviation (σ) relationship:

$$SNR = 20 \bullet \log(\tfrac{1}{\sigma}) \tag{13}$$

Thus,
$$\sigma = 10^{-(SNR/20)} \tag{14}$$

In order to provide more realistic PMU signals within the prototype environment, white Gaussian noise is added to the measurements obtained from the simulator. Algorithm 2 shows the steps that were used to incorporate noise into the magnitude of all voltage measurements obtained in every frame. User-specified input is the noise mean and deviation used to create a normal distribution (N) for every received frame. The multiplier noise factor (n), used to update these voltage magnitudes is random selected from N. The result is a set of noisy measurement with the same noise characteristics across the different phasors. Noise statistics can be further improved by considering the shape of the normal distribution. Properties such as the skew and kurtosis provide more statistics of the noise levels in actual signals which enable the prototyping of real-life PMU measurements.

V. Conclusion and Future Work

In this work, a framework for prototyping PMU errors in an emulation environment has been presented. It utilizes a C37.118 data source to provide grid data, after which modification of the relevant fields are done to emulate actual field errors. The mechanisms of time-skew error, and mislabeled bits during leap second transition have been discussed.

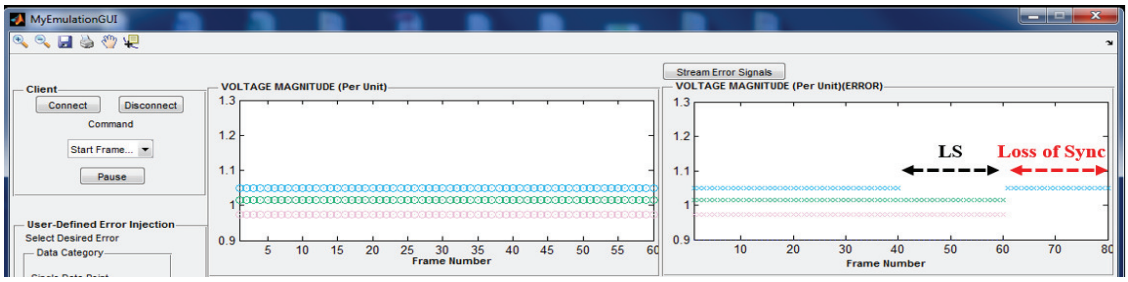

Fig. 6. Inaccurate data reporting and time mis-synchronization due to mislabeled leap transition bits

Algorithm 2: Process for Noise Injection

Input: Stream of C37.118 data frames, Start Transmission Rqst, Mean (μ), Noise level (in dB), Number of phasors (PHNMR)

Output: Noisy voltage magnitude

Stream of Incoming Data Frame

```
for i = 1: TOTALFRAMES
    Evaluate: Deviation, σ
    Generate: Normal distribution, N(μ,σ)
    Random Selection: Noise factor, n
    for j = 1: PHNMR
        Convert: BCD phasor → Decimal → Polar form
        Extract: Voltage magnitude |V|
        Increment |V| by |V|× n
        Convert: BCD phasor → Decimal → Polar form
        Re-package: Modified BCD phasor
        Insert: Modified BCD into PHASORS field
    end
end
```

In addition, prototype errors for each of the cases has been implemented, and visualized in a Matlab environment. In order to provide a better emulation of real-life PMU measurements, the addition of noise was also discussed.

In future work, other mechanisms of PMU data errors will be incorporated, thus expanding on the current environment. Also, an overall framework for the analysis and visualization of PMU data errors is being developed. The methods being proposed to achieve this goal will include the computation of statistical correlations and clustering of data to quantify errors in prototyped and real-life PMU data.

ACKNOWLEDGMENT

This research was supported by the Power Systems Engineering Research Center (PSERC) under project T-57.

REFERENCES

[1] NASPI and PNNL. (2016). NASPI-2016-TR-002 Synchrophasor Data Quality Attributes and a Methodology for Examining Data Quality Impacts upon Synchrophasor Applications. [Online]. Available: https://www.naspi.org/FIle.aspx?fileID=1689

[2] Qiang Zhang, Xiaochuan Luo, David Bertagnolli, Sergey Maslennikov, and Brock Nubile. PMU data validation at ISO new england. In *Power and Energy Society General Meeting (PES), 2013 IEEE*, pages 1–5. IEEE, 2013.

[3] Qing Zhang, V Vittal, G Heydt, Y Chakhchoukh, Naim Logic, and Steve Sturgill. The time skew problem in PMU measurements. In *Power and Energy Society General Meeting, 2012 IEEE*, pages 1–6. IEEE, 2012.

[4] Jiecheng Zhao, Lingwei Zhan, Yilu Liu, Hairong Qi, Jose R Garcia, and Paul D Ewing. Measurement accuracy limitation analysis on synchrophasors. In *Power & Energy Society General Meeting, 2015 IEEE*, pages 1–5. IEEE, 2015.

[5] Qiang Frankie Zhang and Vaithianathan Mani Venkatasubramanian. Synchrophasor time skew: Formulation, detection and correction. In *North American Power Symposium (NAPS), 2014*, pages 1–6. IEEE, 2014.

[6] Prashant Kansal and Anjan Bose. Bandwidth and latency requirements for smart transmission grid applications. *Smart Grid, IEEE Transactions on*, 3(3):1344–1352, 2012.

[7] Markos Asprou and Elias Kyriakides. The effect of time-delayed measurements on a PMU-based state estimator. In *PowerTech, 2015 IEEE Eindhoven*, pages 1–6. IEEE, 2015.

[8] Biju Naduvathuparambil, Matthew Valenti, Ali Feliachi, et al. Communication delays in wide-area measurement systems. In *Southeastern Symposium on System Theory*, volume 34, pages 118–122. Citeseer, 2002.

[9] J.D. Taft. (2016). Grid Architecture 2. [Online]. Available: http://gridarchitecture.pnnl.gov/media/white-papers/GridArchitecture2final.pdf

[10] PowerWorld, "PowerWorld Simulator Overview," available at http://www.powerworld.com/products/simulator/overview.

[11] NASPI. (2011). Synchrophasor data quality. [Online]. Aavailable: https://www.naspi.org/Badger/content/File/FileService.aspx?fileID=34

[12] Kenta Kirihara, Karl E Reinhard, Andy K Yoon, and Peter W Sauer. Investigating synchrophasor data quality issues. In *Power and Energy Conference at Illinois (PECI), 2014*, pages 1–4. IEEE, 2014.

[13] IEEE Standard for Synchrophasor Measurements for Power Systems, IEEE Standard C37.118.1-2011.

[14] IEEE Standard for Synchrophasor Data Transfer for Power Systems, IEEE Standard C37.118.2-2011.

[15] Arun G Phadke and John Samuel Thorp. *Synchronized phasor measurements and their applications*. Springer Science & Business Media, 2008.

[16] Ken Fodero, Chris Huntley, and Dave Whitehead. Secure, wide-area time synchronization. In *proceedings of the 12th Annual Western Power Delivery Automation Conference, Spokane, WA*, 2010.

[17] Merja Tornikoski Joni Tammi Ari Mujunen, Juha Aatrokoski. GPS Time Disruptions on 26-Jan-2016. [Online]. Available: https://aaltodoc.aalto.fi/handle/123456789/19833

[18] Ya-Shian Li-Baboud Allen Goldstein, Dhananjay Anand. (2015). Nistir 8077: Investigation of PMU Response to Leap Second: 2015. [Online]. Available: http://dx.doi.org/10.6028/NIST.IR.8077

[19] Ning Zhou, Da Meng, Zhenyu Huang, and Greg Welch. Dynamic state estimation of a synchronous machine using PMU data: A comparative study. *Smart Grid, IEEE Transactions on*, 6(1):450–460, 2015.

[20] Di Shi, Daniel J Tylavsky, and Naim Logic. An adaptive method for detection and correction of errors in PMU measurements. *Smart Grid, IEEE Transactions on*, 3(4):1575–1583, 2012.