Les Cahiers du GERAD ISSN: 0711–2440

A new approach to reliability assessment of synchrophasor communications in smart grids

Y. Seyedi, H. Karimi,
C. Wette, B. Sansò

G–2020–09

January 2020

La collection *Les Cahiers du GERAD* est constituée des travaux de recherche menés par nos membres. La plupart de ces documents de travail a été soumis à des revues avec comité de révision. Lorsqu'un document est accepté et publié, le pdf original est retiré si c'est nécessaire et un lien vers l'article publié est ajouté.

Citation suggérée : Y. Seyedi, H. Karimi, C. Wette, B. Sansò (Janvier 2020). A new approach to reliability assessment of synchrophasor communications in smart grids, Rapport technique, Les Cahiers du GERAD G–2020–09, GERAD, HEC Montréal, Canada.

Avant de citer ce rapport technique, veuillez visiter notre site Web (https://www.gerad.ca/fr/papers/G-2020-09) afin de mettre à jour vos données de référence, s'il a été publié dans une revue scientifique.

The series *Les Cahiers du GERAD* consists of working papers carried out by our members. Most of these pre-prints have been submitted to peer-reviewed journals. When accepted and published, if necessary, the original pdf is removed and a link to the published article is added.

Suggested citation: Y. Seyedi, H. Karimi, C. Wette, B. Sansò (January 2020). A new approach to reliability assessment of synchrophasor communications in smart grids, Technical report, Les Cahiers du GERAD G–2020–09, GERAD, HEC Montréal, Canada.

Before citing this technical report, please visit our website (https://www.gerad.ca/en/papers/G-2020-09) to update your reference data, if it has been published in a scientific journal.

La publication de ces rapports de recherche est rendue possible grâce au soutien de HEC Montréal, Polytechnique Montréal, Université McGill, Université du Québec à Montréal, ainsi que du Fonds de recherche du Québec – Nature et technologies.

Dépôt légal – Bibliothèque et Archives nationales du Québec, 2020
– Bibliothèque et Archives Canada, 2020

The publication of these research reports is made possible thanks to the support of HEC Montréal, Polytechnique Montréal, McGill University, Université du Québec à Montréal, as well as the Fonds de recherche du Québec – Nature et technologies.

Legal deposit – Bibliothèque et Archives nationales du Québec, 2020
– Library and Archives Canada, 2020

GERAD HEC Montréal
3000, chemin de la Côte-Sainte-Catherine
Montréal (Québec) Canada H3T 2A7

Tél. : 514 340-6053
Téléc. : 514 340-5665
info@gerad.ca
www.gerad.ca

A new approach to reliability assessment of synchrophasor communications in smart grids

Younes Seyedi [a,b]

Houshang Karimi [b]

Constant Wette [c]

Brunilde Sansò [a,b]

[a] *GERAD, Montréal (Québec), Canada, H3T 2A7*

[b] *Department of Electrical Engineering, Polytechnique Montréal (Québec) Canada, H3C 3A7*

[c] *Ericsson, Montréal (Québec) Canada, H4S 0B6*

younes.seyedi@polymtl.ca
brunilde.sanso@polymtl.ca

January 2020
Les Cahiers du GERAD
G–2020–09

Copyright © 2020 GERAD, Seyedi, Karimi, Wette, Sansò, IEEE. This paper is a preprint (IEEE "submitted" status). Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Les textes publiés dans la série des rapports de recherche *Les Cahiers du GERAD* n'engagent que la responsabilité de leurs auteurs. Les auteurs conservent leur droit d'auteur et leurs droits moraux sur leurs publications et les utilisateurs s'engagent à reconnaître et respecter les exigences légales associées à ces droits. Ainsi, les utilisateurs:

- Peuvent télécharger et imprimer une copie de toute publication du portail public aux fins d'étude ou de recherche privée;
- Ne peuvent pas distribuer le matériel ou l'utiliser pour une activité à but lucratif ou pour un gain commercial;
- Peuvent distribuer gratuitement l'URL identifiant la publication.

Si vous pensez que ce document enfreint le droit d'auteur, contactez-nous en fournissant des détails. Nous supprimerons immédiatement l'accès au travail et enquêterons sur votre demande.

The authors are exclusively responsible for the content of their research papers published in the series *Les Cahiers du GERAD*. Copyright and moral rights for the publications are retained by the authors and the users must commit themselves to recognize and abide the legal requirements associated with these rights. Thus, users:

- May download and print one copy of any publication from the public portal for the purpose of private study or research;
- May not further distribute the material or use it for any profit-making activity or commercial gain;
- May freely distribute the URL identifying the publication.

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Abstract: Wireless communications can facilitate transfer of synchrophasor data between spatially separated phasor measurement units (PMUs) and phasor data concentrators (PDCs). However, such communication systems may impose random access delay and failure on PMU channels that lead to missing synchrophasor data frames at the output of the PDC. This paper presents a new approach for online reliability assessment and improvement of synchrophasor data communications. The proposed approach involves elaborate estimation and probabilistic prediction algorithms that trigger a prioritized handover mechanism in order to minimize the number of synchrophasor data frames that are missing over successive time-stamps. Extensive simulations based on LTE wireless communications confirm significant improvement in terms of data loss under both absolute and relative waiting time logics. This non-intrusive approach can be adopted by network operators to ensure reliable transmission of synchrophasor data to monitoring, control, and protection applications in smart grids.

Keywords: Phasor measurement unit, reliability, smart grid, synchrophasor, wireless communications

1 Introduction

As power distribution grids are shifting towards smart and complex systems, advanced metering and sensor technologies are being developed for data acquisition and grid intelligence purposes [1]. In particular, fast and accurate measurements provided by synchrophasor technology can significantly improve monitoring, control and protection applications in smart grids [2, 3]. Phasor measurement units (PMUs), phasor data concentrators (PDCs) and communication systems are the main components of synchrophasor networks [4]. PMUs provide synchrophasor data frames that carry the measured phasors, frequency, rate of change of frequency, real/reactive powers, and the unique time-stamps [5]. Communication systems deliver the data packets to the PDCs which collect and time align the data provided by geographically distributed PMUs [6, 7]. Communication systems directly affect data integrity and therefore play a crucial role in the performance of synchrophasor networks [8, 9].

Distribution grid operators may adopt wireless communications to facilitate transfer of synchrophasor data [10, 11]. In some distribution networks, wireless communications is the only feasible solution where Ethernet is not available. However, wireless systems may lead to unforeseen access failures and time-varying access delays. The access delays are essentially stochastic, and may impact PMU channels during different time intervals [12]. The extent and duration of the access failure/delay should be minimized to meet the requirements of time-critical control and protection applications in smart grids [13]. Access failures and access delays adversely intervene in the data collection mechanism at the PDC [14, 15]. Under such circumstances, incomplete synchrophasor datasets characterized by not-a-number (NaN) indicators appear at the PDC output. Incomplete datasets that bear multiple NaN indicators severely deteriorate the performance of control and protection applications [16, 17]. Smart grids may alleviate this problem and improve data integrity by using elaborate recovery methods [18, 19]. The data recovery methods require remarkable computations and archiving to replace the NaN indicators by estimates of the actual measurements. Moreover, the accuracy of the recovered data degrades with increasing the number of NaN indicators. Hence, data recovery methods still require reliable communication links to deliver high-quality synchrophasor datasets to the applications.

The communication aspects of synchrophasor networks have been addressed by several papers [20]–[23]. The authors in [20] propose a hybrid infrastructure based on power-line and wireless communications to constrain the packet delays while maintaining the bandwidth at low levels. In [21], it is shown that synchrophasor networks can benefit from compressive sampling technique to minimize delays and reduce the communication bandwidth. Castello et al. propose a method that evaluates the delay statistics in individual PMU links [22]. Their adaptive method dynamically changes waiting time and data aggregation at the PDC to meet the application requirements. The impact of communication constraints on optimal PMU placement is studied in [23]. The authors present an optimization procedure that accounts for data rate and routing to achieve fully observable power grids [23].

The previous methods do not provide a systematic procedure for online reliability assessment of synchrophasor communications. They require computer simulations and detailed knowledge of the network parameters to deal with the communication constraints. Moreover, extra computations can be imposed on the PMU/PDC. In many practical scenarios, however, the communication operators do not have access to the power utility premises, and prior knowledge of the PMU/PDC parameters is not available. To overcome such challenges, this paper presents a new approach for reliability assessment and improvement of data communications which can be applied to any synchrophasor network with arbitrary topology/parameters in real-time. The reliability is evaluated in terms of NaN sequences (NaNS) which represent a series of missing data frames over successive time-stamps at the PDC output. Parameter estimation, probabilistic prediction, and prioritized handover are the main steps involved in the proposed approach. Parameter estimation deals with extracting the reporting rate of PMUs and the delay thresholds by analyzing the time of arrival (TOA) and time of departure (TOD) of synchrophasor packets. This non-intrusive approach does not require prior knowledge of the PMU/PDC parameters. The extracted parameters are used by prediction algorithms that evaluate the emergence of NaNS under absolute and relative waiting time strategies. The developed algorithms can predict latent unreliable communication conditions which are not detectable by the previous methods in the literature. Moreover, early prediction of NaNS triggers prioritized handover for unreliable PMU channels and improves the reliability of synchrophasor communications in a timely manner.

2 Proposed approach

Figure 1 demonstrates the general schematic of a distribution-level synchrophasor network based on the wireless communications. In this architecture, an uplink node (UN) is the ingress to the communication system which directly receives synchrophasor packets from PMUs. A downlink node (DN) is the egress from the communication system which sends synchrophasor packets to the local PDC. The UNs and the DN, which are typically base stations, are indirectly interconnected by means of a core network that is responsible for packet routing in accordance with the Internet protocol. The core/Internet is a complex network composed of multitude of routers and heterogeneous links.

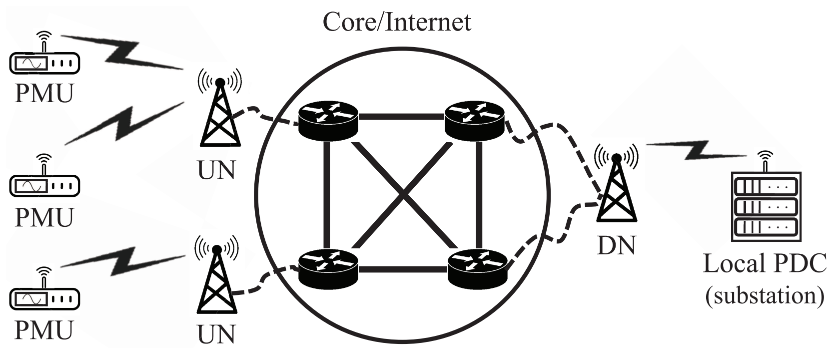

Figure 1: A distribution-level synchrophasor network that involves PMUs, local PDC, and wireless communications

The alignment of the synchrophasor data is carried out based on the coordinated universal time (UTC) and the time-stamp of the synchrophasor data frames. If a synchrophasor packet arrives at the local PDC later than a pre-determined threshold, then the corresponding entry of the dataset is filled by a NaN indicator [5]. Data pushing with time alignment can be performed based on either absolute or relative waiting time strategies [6]. Figure 2 shows example datasets at the PDC output with 4 PMU channels. The NaNS in Figure 2(a) indicate unreliable communications, i.e., missing data frames for a period of 5 time-stamps. In contrast, the datasets depicted in Figure 2(b) exhibit scattered data losses which can be estimated by the data recovery methods [18, 19], hence, Figure 2(b) is not considered as a case of unreliable communications.

The block diagram depicted in Figure 3 shows the main steps and the important parameters involved in the proposed approach. Parameter estimation, prediction algorithms, and prioritized handover are the constituent elements of the proposed approach. The basic premise is that the UNs and the DN are able to record the TOA and TOD of synchrophasor packets.

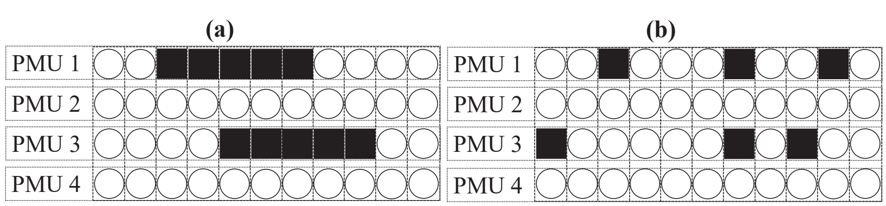

Figure 2: Illustrative examples of incomplete synchrophasor datasets at the PDC output. Empty circles and filled squares are measurements and NaN indicators, respectively. (a): Unreliable communications yield NaNS in PMU channels 1 and 3 (b): Sparse NaN indicators in PMU channels 1 and 3

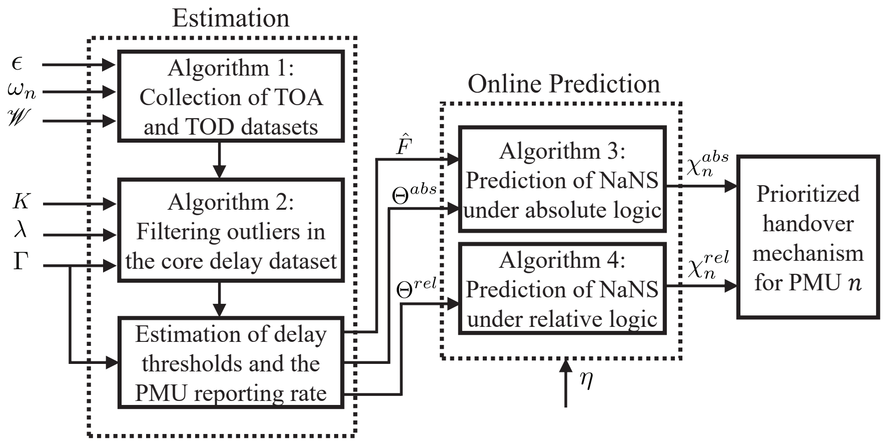

Figure 3: The main functional elements of the proposed approach

2.1 Construction of time-aligned delay datasets

Suppose that the synchrophasor network has N spatially separated PMUs, and their reporting rate is equal to F frames-per-second (fps). The first step towards communications reliability assessment is to estimate the important parameters of the synchrophasor data streams. To this aim, a time window is required for measuring the TOA and TOD data.

Definition 1 *The estimation window, $\mathcal{W}$, is the time interval during which the streaming of synchrophasor packets is analyzed for parameter estimation.*

The length of the estimation window should be much larger than the maximum reporting time of PMUs to allow for processing of a large number of packets, and to increase the accuracy of the parameter estimation. The communication system has prior knowledge of the basic parameters of the links between the PMUs and the UNs including capacity and bandwidth of the channels, packet sizes, etc. Hence, the network can determine the normal uplink delay, i.e., the time that a packet requires to be successfully received by a UN when the uplink is in the normal operation condition. Let the parameter ω_n denote the normal uplink delay for PMU n. The upper-bound of the uplink delay in the normal state is

$$\pi_n = \omega_n + \epsilon \tag{1}$$

where the parameter ϵ represents the largest variation in the packet transmission time (due to PMU processing). This *a priori* known parameter satisfies the inequality $0 < \epsilon \ll 1/F_{\max}$ where $F_{\max}$ shows the maximum PMU reporting rate that can be supported by the wireless communication system. The role of the parameter ϵ in (1) is to prevent false alarms in the prediction of NaNS. Suppose that $t_{n,k}$ shows the time instant that the k^{th} synchrophasor packet is generated by PMU $n \in \mathcal{N}$, $\mathcal{N} = \{1, 2, ..., N\}$. Accordingly, the following parameters are defined:

Definition 2 *The uplink TOA, $t^u_{n,k}$, is the time instant at which the k^{th} data packet from the n^{th} PMU arrives at a UN.*

Definition 3 *The downlink TOA, $t^d_{n,k}$, is the time instant at which the k^{th} data packet from the n^{th} PMU arrives at the DN that serves the PDC.*

Definition 4 *The downlink TOD, $t'_{n,k}$, is the time instant at which the k^{th} data packet from the n^{th} PMU is successfully transmitted by the DN that serves the PDC.*

Based on the above definitions, the end-to-end communication delay, denoted by $\tau_{n,k}$, is given by $\tau_{n,k} = t'_{n,k} - t_{n,k}$, which is essentially a random process. Algorithm 1 shows the procedure for collection of TOA and TOD data over the estimation window. This algorithm consists of three subroutines and overcomes the packet reordering problem by using a hash code matching technique. The outputs of Algorithm 1 are; the uplink TOA datasets $\mathcal{T}_n^u$, the core delay datasets $\mathcal{T}_n^c$, and the downlink delay datasets $\mathcal{T}_n^r$. Subroutine **A** is carried out by the UNs which extract the uplink TOAs and the hash codes. Subroutine **B** is carried out by the DN which finds the downlink TOAs, TODs, and the hash codes. The UNs and the DN can calculate the hash codes by discarding the IP and TCP/UDP headers of the synchrophasor packets followed by applying SHA-1 algorithm [24].

At the end of the estimation window, Subroutine **C** is carried out which first aligns the recorded TOAs/TODs for different PMUs, and then constructs sorted datasets by comparing the hash codes. The alignment task is necessary for maintaining synchronization between different PMU channels such that the elements with identical indices refer to identical time-stamps. It should be noted that, the elements of $\mathcal{T}_n^u$ are intrinsically sorted in ascending order, i.e., $\mathcal{T}_n^u(i) < \mathcal{T}_n^u(j) \iff i < j$.

Algorithm 1 Collection of TOAs and TODs over $\mathcal{W}$

Inputs: ω_n, ϵ
Initialize: $i_n = j_n = 0$. Execute **A** and **B** simultaneously
A: Uplink Probing: At every UN serving a PMU:
1: **if** $t \in \mathcal{W}$, **then**
2: Upon arrival of a new packet from PMU n go to **A**-6
3: **else**
4: Proceed to **C**
5: **end if**
6: Set $i_n = i_n + 1$ and record the uplink TOA t^u_{n,i_n}
7: Add the hash code of the uplink payload to the set $\mathcal{H}_n^u$
8: Go to **A**-1
B: Downink Probing: At every DN serving the PDC:
1: **if** $t \in \mathcal{W}$, **then**
2: Upon arrival of a new packet from PMU n go to **B**-6
3: **else**
4: Proceed to **C**
5: **end if**
6: Set $j_n = j_n + 1$ and record the downlink TOA and TOD t^d_{n,j_n} and t'_{n,j_n}, respectively
7: Add the hash code of the downlink payload to the set $\mathcal{H}_n^d$
8: Go to **B**-1
C: Alignment & Sorting: Set $m = \arg\max_{n\in\mathcal{N}} t^u_{n,1}$
1: **for** $n = 1$ to N **do**:
2: **if** $t^u_{m,1} - t^u_{n,1} \le |\omega_m - \omega_n| + \epsilon$, **then**
3: Set $l_n = 1$
4: **else**
5: Set $l_n = 2$
6: **end if**
7: **for** $i = l_n$ to $\| \mathcal{H}_n^u \|$ **do**:
8: **if** $\mathcal{H}_n^u(i) = \mathcal{H}_n^d(j)$ for some $j \ge 1$, **then**
9: Add $t^u_{n,i}$ to $\mathcal{T}_n^u$
10: Add $t^d_{n,j} - t^u_{n,i}$ to $\mathcal{T}_n^c$
11: Add $t'_{n,j} - t^d_{n,j}$ to $\mathcal{T}_n^r$
12: **end if**
13: **end for**
14: **end for**
Outputs: $\mathcal{T}_n^u$, $\mathcal{T}_n^c$ and $\mathcal{T}_n^r$

2.2 Filtering outliers in the core delay datasets

It is known that the IP-based core networks may impose relatively huge delays on the synchrophasor packets [10]. Such abnormal core delays can span a large number of synchrophasor time-stamps and degrade decision making if not detected. Therefore, it is crucial to find the outliers in the core delay datasets and discard them before parameter estimation. Algorithm 2 explains the procedure of finding the outliers based on the local outlier probabilities (LOPs) of the measured core delays [25]. In this algorithm, the vectors $\mathbf{x}_i$ represent N-dimensional data samples and $\alpha(\mathbf{x}_i|K,\lambda,\mathcal{D})$ denotes the LOP of x_i with respect to the dataset $\mathcal{D}$. The LOPs are calculated based on the K^{th} nearest neighbor search and the parameter λ is the significance level. The threshold Γ in Algorithm 2 is the confidence level in the delay analysis, and is typically very close to the unity, e.g., $\Gamma = 0.99$. The operator erf(.) denotes the error function. According to Algorithm 2 if the LOP of a sample exceeds the threshold Γ, then that sample is labeled as an outlier and is ignored. The dataset of samples is updated once a sample with LOP less than Γ is found in the recorded core delays. The outputs of Algorithm 2 are the filtered (outlier free) delay datasets denoted by $\mathcal{C}_n$.

Algorithm 2 Filtering outliers in the core delay datasets

Inputs: $\mathcal{T}_n^c$, K, λ, Γ
Initialize:
1: Set $M = \min\limits_{n\in\mathcal{N}} \parallel \mathcal{T}_n^c \parallel$
2: Construct the sample vectors:
$\mathbf{x}_i = \left[\mathcal{T}_1^c(i)\ \mathcal{T}_2^c(i)\ ...\ \mathcal{T}_N^c(i)\right]^T,\ i = 1,2,...,M$
3: Construct the initial dataset as $\mathcal{D} = \{\mathbf{x}_1,\ \mathbf{x}_2, ...,\ \mathbf{x}_K\}$
A: Filter & Update:
1: **for** $i = K+1$ to M, **do**
2: Execute **B** and determine the LOP $\alpha(\mathbf{x}_i|K,\lambda,\mathcal{D})$
3: **if** $\alpha(\mathbf{x}_i) < \Gamma$, **then**
4: **for** $n = 1$ to N, **do**
5: Add $\mathcal{T}_n^c(i)$ to $\mathcal{C}_n$
6: **end for**
7: Append sample data $\mathbf{x}_i$ to the dataset $\mathcal{D}$
8: **end if**
9: **end for**
B: LOP Calculation: For given K, λ, and $\mathcal{D}$:
1: Calculate the distances between the samples:
$d(\mathbf{x}_j,\mathbf{x}_l) = \sqrt{(\mathbf{x}_j-\mathbf{x}_l)^T(\mathbf{x}_j-\mathbf{x}_l)}, j \neq l, \forall \mathbf{x}_j,\mathbf{x}_l \in \mathcal{D}$
2: Set $\mathcal{L}(\mathbf{x}_j)$ as the context set by finding the K nearest neighbors to the sample $\mathbf{x}_j$
3: Calculate the standard distance for the sample $\mathbf{x}_j$:
$$\bar{d}(\mathbf{x}_j) = \sqrt{\tfrac{1}{K}\textstyle\sum_{\mathbf{x}_l\in\mathcal{L}(\mathbf{x}_j)} d^2(\mathbf{x}_j,\mathbf{x}_l)}$$
4: Calculate the probabilistic local outlier factor for $\mathbf{x}_j$:
$$p(\mathbf{x}_j) = \frac{K\bar{d}(\mathbf{x}_j)}{\sum_{\mathbf{x}_l\in\mathcal{L}(\mathbf{x}_j)} \bar{d}(\mathbf{x}_l)} - 1$$
5: Calculate the standard probabilistic local outlier factor:
$$s = \sqrt{\tfrac{1}{\|\mathcal{D}\|}\textstyle\sum_{\mathbf{x}_j\in\mathcal{D}} p^2(\mathbf{x}_j)}$$
6: Calculate the LOP of $\mathbf{x}_i$ as:
$$\alpha(\mathbf{x}_i) = \max\left\{0, \operatorname{erf}\left(\tfrac{p(\mathbf{x}_i)}{\sqrt{2}\lambda s}\right)\right\}$$
7: Proceed to **A**-3
Outputs: $\mathcal{C}_n$

2.3 Estimation of delay thresholds

When the PDC uses the absolute waiting time strategy, the reference time for data pushing is the time-stamp or equivalently the packet generation time. In this case, the following delay threshold is used for predicting the data loss:

$$\Theta^{abs} = \max_{n\in\mathcal{N}} \theta_n^a \tag{2}$$

$$\theta_n^a = \pi_n + \min\left\{\tau \middle| \Phi_n^c(\tau) \geq \Gamma\right\} + \nu_n \tag{3}$$

$$\Phi_n^c = \mathbb{C}(\mathcal{C}_n) \tag{4}$$

$$\nu_n = \arg\max_a h_n^r(a) \tag{5}$$

$$h_n^r = \mathbb{H}(\mathcal{T}_n^r) \tag{6}$$

Operators $\mathbb{C}(.)$ and $\mathbb{H}(.)$ show respectively the empirical cumulative distribution function (CDF) and histogram.

For the relative waiting time strategy, the arrival of the first packet with a given time-stamp is the reference for the data pushing. In this case, the following delay threshold is used which accounts for the range of the measured delays:

$$\Theta^{rel} = \epsilon + \min\Big\{\theta \Big| \Phi^r(\theta) \geq \Gamma\Big\} \tag{7}$$

$$\Phi^r = \mathbb{C}\Big(\big\{\theta_1^r, \theta_2^r, ...\big\}\Big) \tag{8}$$

$$\theta_i^r = \max_{n\in\mathcal{N}} \theta_{n,i} - \min_{l\in\mathcal{N}} \theta_{l,i} \tag{9}$$

$$\theta_{n,i} = \pi_n + \mathcal{C}_n(i) + \nu_n \tag{10}$$

where the parameter ν_n is given by Equation (5) and $\mathcal{C}_n(i)$ represents the i^{th} element of the filtered core delay dataset.

2.4 Estimation of the PMU reporting rate

The uplink TOA datasets can be used to estimate the reporting rate of the PMUs in the synchrophasor network. The estimate of the PMU reporting rate is crucial in prediction of the data loss under the absolute waiting time logic. First, the uplink inter-packet times are calculated as:

$$\delta_n(i) = \mathcal{T}_n^u(i+1) - \mathcal{T}_n^u(i), i \in \{1, 2, ..., \parallel \mathcal{T}_n^u \parallel -1\} \tag{11}$$

Then, the peak analysis of the empirical histogram of the uplink inter-packet times provides the estimate of F as

$$\hat{F} = \arg\min_{f\in\mathcal{F}} \left| f - \frac{1}{\arg\max_a h_\delta(a)} \right| \tag{12}$$

$$h_\delta = \mathbb{H}\Big(\{\delta_1\}, \{\delta_2\}, ..., \{\delta_N\}\Big) \tag{13}$$

$$\mathcal{F} = \Big\{1, 10, 12, 15, 20, 30, 60, 120\Big\} \tag{14}$$

It is worth mentioning that $\mathcal{F}$ contains the permissible PMU reporting rates tailored to power systems that operate at 60 Hz frequency [5].

2.5 Prediction of NaNS

Once the delay thresholds and the PMU reporting rate are estimated, unreliable communications can be detected through prediction of NaNS in one or more PMU channels. The prediction of NaNS under the absolute and relative strategies is based on Algorithms 3 and 4, respectively. These event-based algorithms trigger an operation either upon arrival of a synchrophasor packet at a UN, or at a pre-calculated time instant. The parameter η shows the minimum number of successive NaN indicators that are considered as a NaNS. These prediction algorithms ignore a group of successive NaN indicators if their number is less than η. The parameter μ_n is calculated as

$$\mu_n = \mathbb{E}\big(\mathcal{C}_n\big) \tag{15}$$

where the operator $\mathbb{E}(.)$ denotes the arithmetic mean function. The choice of expected value for μ_n is justified by recent studies that report bi-modal distributions for the core delays [10]. Therefore, compared to the mode of distribution the mean value produces more accurate predictions. The outputs of Algorithms 3 and 4 are the sequence signals χ_n^{abs} and χ_n^{rel}, respectively. As a convention, $\chi_n = 0$ shows that the channel of PMU n is not likely to produce NaNS, whereas $\chi_n = 1$ indicates that NaNS are very likely and thus unreliable communications for PMU channel n is detected.

Probabilistic prediction of NaNS under the absolute waiting time logic has three main steps, as shown in Algorithm 3. Step I is carried out when a packet arrives at a UN. Steps II and III are carried out at pre-determined time instants. Decision-making about NaNS is performed at either Step I or Step III whichever comes first. It should be noted that, the sets $\mathcal{A}_n$ contain the packet indices that are not received by UNs in a timely manner. Step II updates the sets $\mathcal{A}_n$. The parameters ρ_n in Step I denote the estimated values of the uplink delay for the packets that are not received in time, i.e., their indices are found in $\mathcal{A}_n$. For such delayed packets, the expected values of the total communication delay are given by $\rho_n + \mu_n + \nu_n$ which set the binary indicators as $\beta_n = 1$ if go beyond the threshold Θ^{abs}. If delayed packets are not received by the time specified in Step III, then the likelihood that their late arrivals do not meet the threshold Θ^{abs} is high, and consequently the indicators β_n are set as unity. The sequence signals χ_n^{abs} are set to unity when β_n change form 0 to 1 for the recent η indices.

Algorithm 3 Prediction of NaNS under the absolute logic

Inputs: Θ^{abs}, $\hat{F}$, π_n, μ_n, ν_n, ϵ, η

Initialize: Set the reference times for event detection:

```
Set m = arg max_{n∈N} t^u_{n,1}
for n = 1 to N do:
    if t^u_{m,1} − t^u_{n,1} ≤ |ω_m − ω_n| + ϵ, then
        Set t^b_n = t^u_{n,1}, b_n = 0
    else
        Set t^b_n = t^u_{n,2}, b_n = 1
    end if
end for
Set L_n = 1, i_n = j_n = 2, A_n = ∅, χ^abs_n = 0, ∀n ∈ N
```

Detection & Prediction: $\forall t > \max_{n \in \mathcal{N}} t_{r,n}$

I: Upon arrival of a packet at the UN that serves PMU n:

```
Increase counter L_n = L_n + 1
if L_n ∈ A_n, then
    Calculate ρ_n = t^u_{n,L_n+b_n} − t^b_n − (L_n − 1)/F̂ + π_n
    if ρ_n + μ_n + ν_n > Θ^abs then
        Set β_n(L_n) = 1
        if β_n(L_n − p) = 1, ∀p ∈ {0, 1, η − 1}, then
            Set χ_n = 1
        end if
    else
        Set β_n(L_n) = 0
    end if
    Remove L_n from A_n
else
    if L_n ≥ j_n, then
        Set β_n(L_n) = 0
    end if
end if
```

II: At $t = t_n^b + (i_n - 1)/\hat{F} + \epsilon$:

```
if L_n ≠ i_n, then
    Add i_n to the set A_n
end if
Set i_n = i_n + 1
```

III: At $t = t_n^b + (j_n - 1)/\hat{F} + \Theta^{abs} - (\pi_n + \mu_n + \nu_n) + \epsilon$:

```
if j_n ∈ A_n, then
    Set β_n(j_n) = 1
    if β_n(j_n − p) = 1, ∀p ∈ {0, 1, η − 1}, then
        Set χ^abs_n = 1
    end if
    Remove j_n from A_n
end if
Set j_n = j_n + 1
```

Outputs: χ_n^{abs}

As shown in Algorithm 4, the prediction of NaNS under the relative waiting time logic involves two steps: Step I which is executed upon arrival of packets at UNs, and Step II which is carried out at pre-calculated times. The objective of Step I is to identify the packets which are expected to arrive at the PDC earlier than their counterparts. Once the earliest packet is identified, the parameter φ is updated to properly adjust the decision-making time for Step II. The value of φ is calculate based on the expected value of the interval between packet arrival at the UNs and packet arrivals at the PDC. The objective of Step II is to find the packets whose relative delay (with respect to the delay of the earliest packet) exceeds the threshold Θ^{rel} with a high probability. Any delayed packet in PMU channel n that is not received by the specified time yields a positive indicator β_n. Finally, the sequence signals χ_n^{rel} are set to unity when the instantaneous indicators β_n change form 0 to 1 for the recent η indices.

Algorithm 4 Prediction of NaNS under the relative logic

Inputs: Θ^{rel}, μ_n, ν_n, ϵ, η

Initialize: Set the reference times for event detection:

Set $m = \arg\max_{n\in\mathcal{N}} t_{n,1}^u$

for $n = 1$ to N **do**:

  if $t_{m,1}^u - t_{n,1}^u \leq |\omega_m - \omega_n| + \epsilon$, **then**

    Set $t_n^b = t_{n,1}^u, b_n = 0$

  else

    Set $t_n^b = t_{n,2}^u, b_n = 1$

  end if

end for

Set $L_n = 1, i = 1, q = 0, \mathcal{B}_n = \varnothing$, $\chi_n^{rel} = 0, \forall n \in \mathcal{N}$

Detection & Prediction: $\forall t > \max_{n\in\mathcal{N}} t_{r,n}$

I: Upon arrival of a packet at the UN that serves PMU n:

Increase counter $L_n = L_n + 1$

if $L_n \in \mathcal{B}_j$, for some $j \in \mathcal{N}$, **then**

  Set $\zeta_n = t_{n,L_n+b_n}^u - t_{\mathcal{E}(q),L_n+b_{\mathcal{E}(q)}}^u + \mu_n + \nu_n$

  if $\zeta_n \leq \mathcal{S}(q)$, **then**

    Set $\mathcal{S}(q) = \zeta_n$

    Set $\mathcal{E}(q) = n$

  end if

else

  if $L_n > i$, **then**

    Set $q = q + 1$

    Set $\mathcal{S}(q) = \mu_n + \nu_n$

    Set $\mathcal{E}(q) = n$

  end if

end if

if $q > 0$, **then**

  Add L_n to $\mathcal{B}_n$

end if

Set $\varphi = \mu_{\mathcal{E}(1)} + \nu_{\mathcal{E}(1)} - \min_{n\in\mathcal{N}}(\mu_n + \nu_n)$

II: At $t = t_{\mathcal{E}(1),\mathcal{B}_{\mathcal{E}(1)}(1)+b_{\mathcal{E}(1)}}^u + \Theta^{rel} + \varphi + \epsilon$:

Set $i = i + 1$

if $\mathcal{B}_n(1) = \mathcal{B}_{\mathcal{E}(1)}(1)$, **then**

  Set $\beta_n(i) = 0$

else

  Set $\beta_n(i) = 1$

  if $\beta_n(j_n - p) = 1,\ \forall p \in \{0, 1, \eta - 1\}$, **then**

    Set $\chi_n^{rel} = 1$

  end if

end if

Remove the first elements of $\mathcal{E}$, $\mathcal{S}$, and $\mathcal{B}_n$

Set $q = q - 1$

Outputs: χ_n^{rel}

2.6 Prioritized handover: Improving reliability

A promising use case of the developed algorithms is to improve the reliability of synchrophasor data communications over shared wireless systems. As soon as the sequence signals χ_n^{abs} or χ_n^{rel} rise from 0 to 1, the communication system searches for another UN that can best serve PMU n. Upon detection of the substitute UN, a prioritized handover mechanism is initiated to accelerate data transmission and prevent further data losses. In this regime, early prediction of NaNS combined with prioritized handover can significantly reduce the number of missing data frames, and improve the integrity of synchrophasor datasets at the PDC output. The resulting improvement is of remarkable importance for data recovery methods as well as control/protection applications in smart distribution systems.

3 Simulation results

The performance and effectiveness of the developed algorithms have been verified by extensive simulations of a distribution-level synchrophasor network based on LTE wireless communications. In the case of LTE, the UNs and DN are called evolved Node B (eNodeB), and the core is called evolved packet core (EPC) [26]. The test synchrophasor network has 10 PMUs that are distributed within a neighborhood in the city of Montreal, QC Canada. Table 1 gives the values of the simulation parameters, and Figure 4 shows the locations of the PMUs, the local PDC and the neighboring eNodeBs. An elaborate LTE simulator [27] is used which reads realistic eNodeB information, such as location, antenna type and antenna height, available from a public database, and handles the data packets that are generated by the PMUs. Base station distances to the PMUs and the PDC, as well as coverage, based on wireless propagation, are taken into account by the LTE simulator. Also considered is how LTE handles the access, the latency inserted by the EPC and the effect of other devices competing for the same LTE resources. The outlier-free EPC delays follow a bi-modal probability distribution function (PDF) as depicted in Figure 5(a). The downlink transmission delays for all synchrophasor packets are modeled as Gaussian random processes with the PDF shown in Figure 5(b). Moreover, the EPC delay and the downlink delay in different PMU channels are statistically independent. More description of the LTE simulator and its use with smart city communications are found in [27] and [28], respectively.

Table 1: The parameters and their values in the simulations

Parameter	**Value**
Number of PMUs: N	10
PMU reporting rate: F	60 fps
PMU packet size (UDP/IP)	116 bytes
Normal uplink delay: ω_n	10 msec
Maximum delay in PMU packet transmission: ϵ	3 msec
Number of nearest neighbors: K	20
Significance level: λ	3
Confidence level: Γ	99.8%
Estimation window: $\mathcal{W}$	1 hour

At any simulation time instant, the eNodeBs that serve the PMUs and the PDC are determined based on a set of criteria including the channel quality, antenna coverage, available resource blocks, etc. Unreliable communications are created by failure of all antennas on specified eNodeBs. After the failure instant, the LTE simulator finds the best available eNodeB and executes a handover mechanism for the affected PMU channels.

Table 2 reports the estimated delay thresholds without filtering the core delay outliers that exist for 10 seconds of the estimation window. In this scenario, the core delay outliers are created by superimposing excessive delays on the outlier-free delays with the PDF shown in Figure 5(a). The values of the superimposed excessive delays are indicated in first column of Table 2. If the core delay outliers are filtered out by Algorithm 2, then the delay thresholds are estimated as $\Theta^{abs} = 99$ msec and $\Theta^{rel} = 36$ msec. It should be noted that, the experiments based a real-world LTE system report 100 msec and 40 msec for the absolute and relative waiting times, respectively [10]. This validates the performance of Algorithms 1 and 2 as the estimated delay thresholds are very close to the realistic PDC waiting times reported in [10].

Table 2 reveals that the estimated delay thresholds can be highly inaccurate when the EPC delay outilers are not removed from the core delay datasets. For instance, if only one PMU channel incurs 100 msec excessive delay, then the estimation errors under absolute and relative logics are 61 msec and 70 msec, respectively. Table 2 also indicates that the estimation error aggravates with increasing the excessive core delays. Under such circumstances, a large number of missing data frames cannot be predicted.

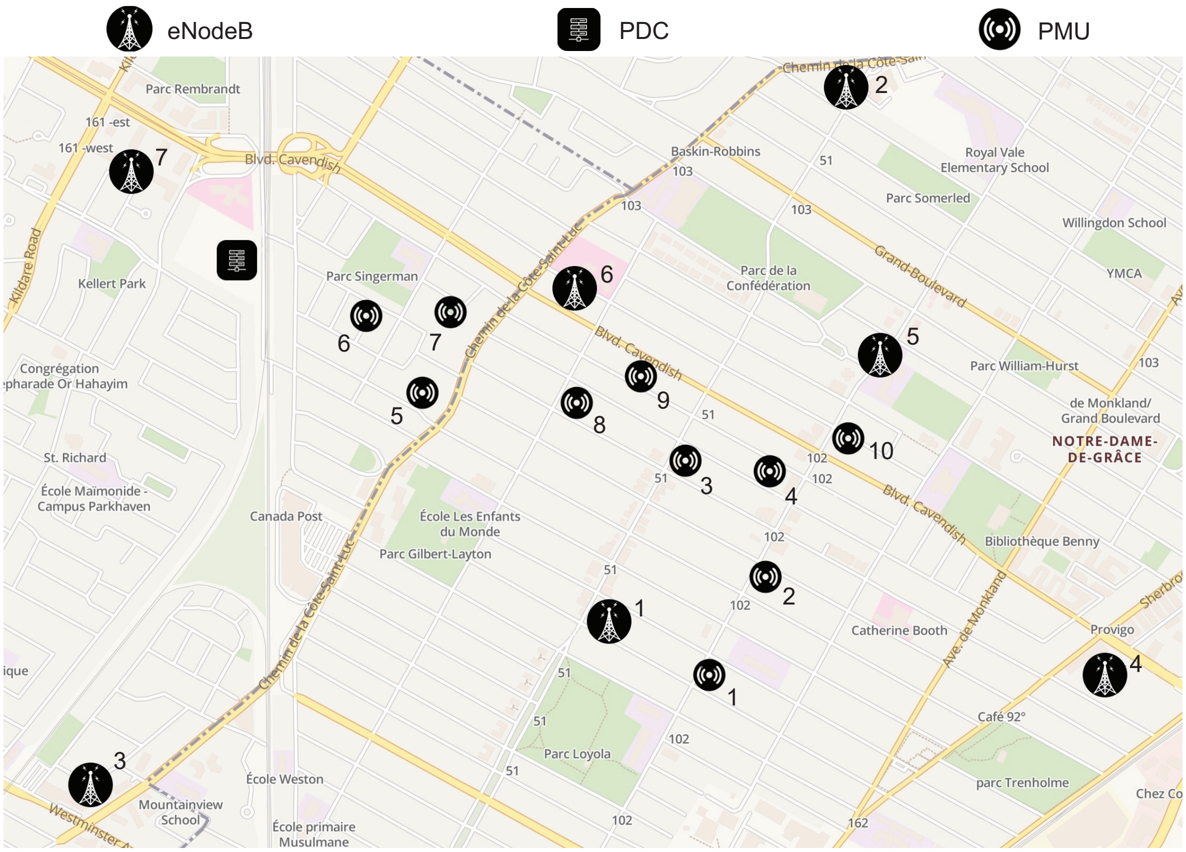

Figure 4: The map of the test synchrophasor network

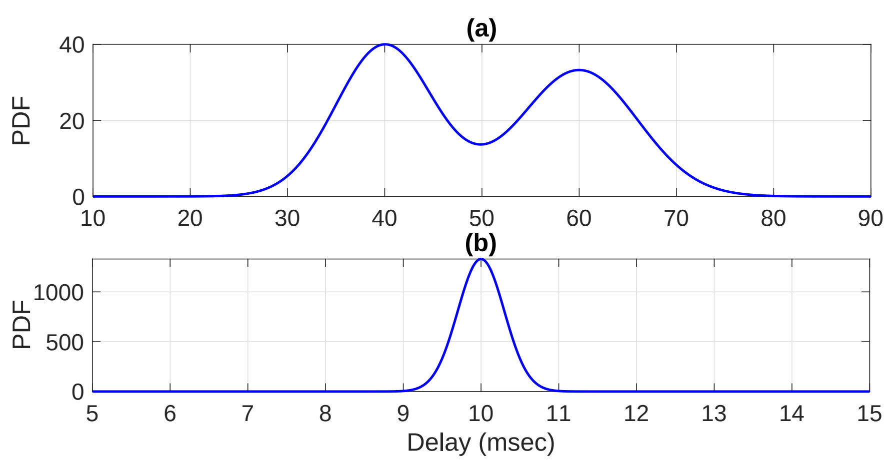

Figure 5: The distribution of random delays in the simulations (a) The PDF of outlier-free EPC delays (b) The PDF of the downlink transmission delays

Figures 6(a)–(d) illustrate the prediction of NaNS when eNodeB 1 which serves PMU 10 fails at $t = 2$ sec. After the instant of eNodeB failure, the subsequent packet delays result in NaNS for both absolute and relative logics. Specifically, three successive NaN indicators are predicted at $t = 2.072$ sec and $t = 2.082$ sec for absolute and relative logics, respectively. It should be noted that, the mean value of the end-to-end communication delay (excluding the core delay outliers) in this synchrophasor network is around 70 msec.

Table 2: Delay thresholds with unfiltered EPC delay outliers

Excessive Core Delay (msec)	Θ^{abs} (msec)	Θ^{rel} (msec)
100	160	106
200	260	206
300	360	306
400	460	406
500	560	506
600	660	606
700	760	706
800	860	806
900	960	906
1000	1060	1006

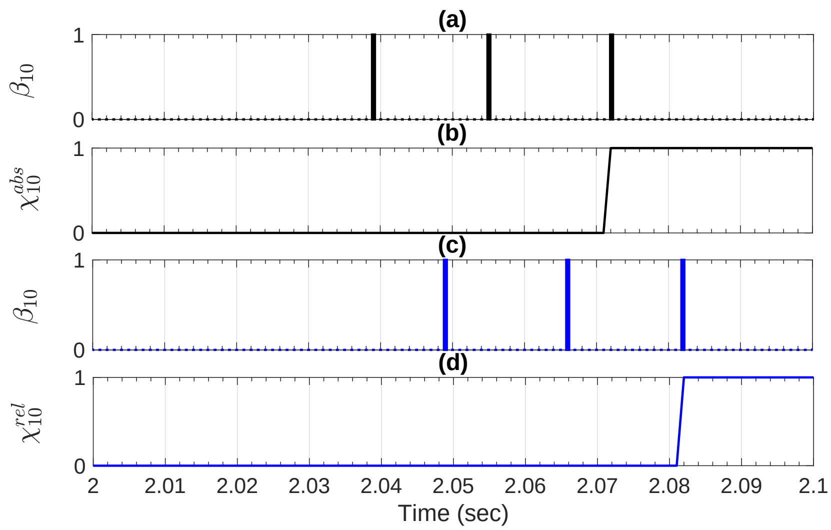

Figure 6: The prediction of NaNS in the channel of PMU 10 with $\eta = 3$: (a) Successive indicators under absolute logic (b) Sequence signal under absolute logic (c) Successive indicators under relative logic (d) Sequence signal under relative logic

The average time required for prediction of unreliable communications is demonstrated in Figure 7. It is seen that, for both absolute and relative logics, the average prediction time linearly increases with increasing the value of parameter η. The fastest prediction is achieved under the absolute logic with $\eta = 2$ which has an average prediction time equal to 56 msec. The average prediction time under the relative logic with the same value of η is 10 msec higher. It is concluded that a proper choice of η is critical for effective improvement of the synchrophasor communications. Smaller η (e.g., less than 5) should be used if the synchrophasor network feeds protection applications, otherwise the prolonged data losses deteriorate the performance of the application.

Figures 8(a) and (b) illustrate the number of missing data frames at the PDC output when eNodeB 3 fails at $t = 2$ sec. The horizontal axis represents the time-stamp of the synchrophasors, and the vertical axis represents the total number of missing data frames for the given time-stamp. In the absence of the prediction algorithms, only one regular handover is carried out that transfers PMU 5 to eNodeB 6. Reconnection of PMU 5 adversely affects the link of PMU 3 such that incremental access delays bring about several missing data frames at the PDC output, as shown in Figure 8(a). This case shows a latent unreliable communication condition for PMU PMU 3 which is not detectable by the previous methods in the literature. The results based on Algorithm 3 with $\eta = 2$ combined with prioritized handover are shown in Figure 8(b). In this case, two prioritized handovers are carried out. The first prioritized handover establishes the link of PMU 5 with eNodeB 6, and the second prioritized handover connects PMU 3 to eNodeB 5. It is seen that the proposed approach can prevent detrimental loss of synchrophasor data such that only 2 incomplete datasets are observed at the PDC output.

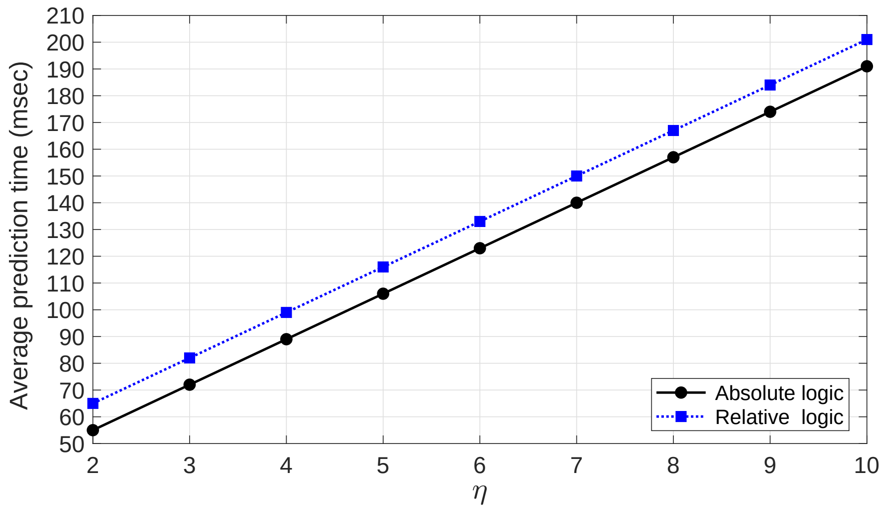

Figure 7: The average prediction time vs. the minimum length of NaNS when the communication link between PMU 5 **and eNodeB** 3 **fails**

Figures 9(a) and (b) demonstrate the average number of missing data frames per PMU when different eNodeBs fail. The results shown in Figure 9(a) indicate that failure of eNodeBs can lead to significant number of missing data frames in the absence of the prediction algorithms. For example, failure of eNodeB 3 entails 112 and 62 missing data frames per PMU under absolute and relative logics, respectively. In this network, failure of the rest of eNodeBs may result in 10–13 missing data frames on average. Figure 9(b) reveals that the average number of missing data frames per PMU can be significantly reduced by using the proposed prediction and prioritized handover mechanism. In particular, the number of missing data frames per PMU are around 1 and 2 for the absolute and relative logics, respectively. The simulation results corroborate the effectiveness of the proposed approach for real-time reliability assessment and improvement of the synchrophasor communications.

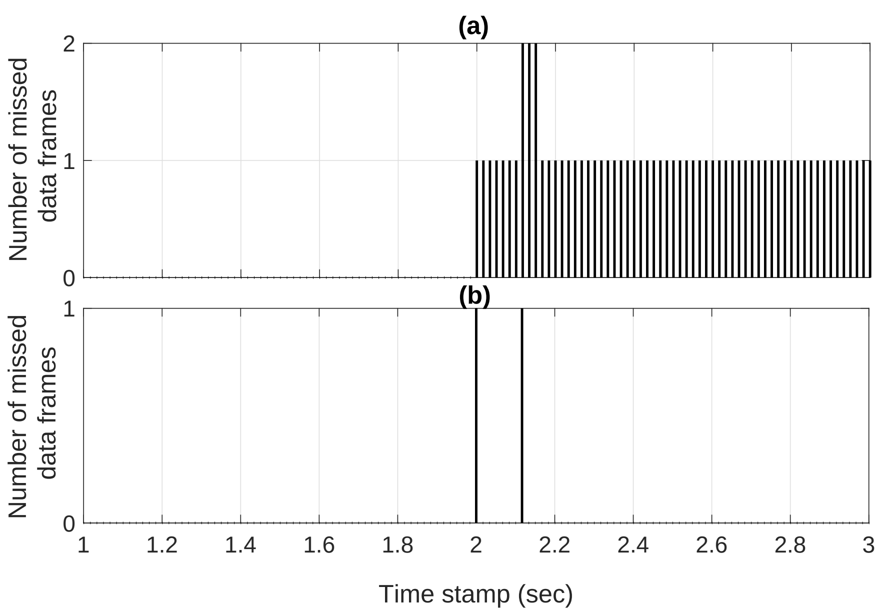

Figure 8: A snapshot of the total number of missed data frames at the PDC output with absolute waiting time 100 **msec (a) Without prediction (b) With the proposed prediction and prioritized handover mechanism**

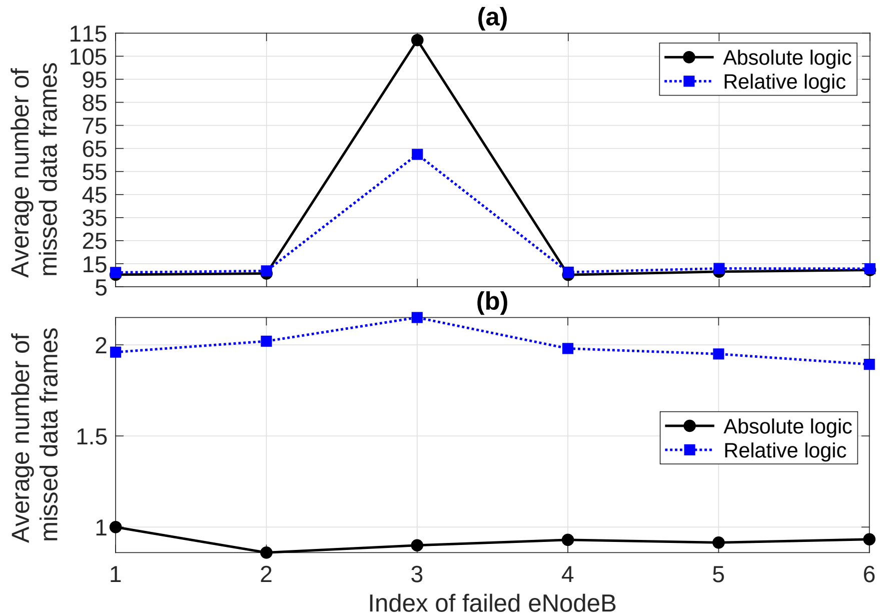

Figure 9: The average number of missing data frames per PMU channel (a) Without prediction (b) With the proposed prediction and prioritized handover mechanism

4 Conclusion

A new non-intrusive approach for online reliability assessment and improvement of synchrophasor communications is proposed. The proposed approach involves three main steps: 1) parameter estimation that deals with the PMU reporting rate and delay thresholds by analyzing measured TOA and TOD datasets. 2) Prediction of NaNS under absolute and relative waiting time logics. 3) Prioritized handover mechanism carried out for unreliable PMU channels. The simulation results based on LTE communication systems indicate significant reduction of the number of missing data frames and corroborate effectiveness of the developed algorithms.

References

[1] J. Romero Aguero, et al., Managing the new grid: delivering sustainable electrical energy, IEEE Power Energy Mag., 17(4):75–84, July-Aug. 2019.

[2] W. Koong Chai, et al., An information-centric communication infrastructure for real-time state estimation of active distribution networks, IEEE Trans. Smart Grid, 6(4):2134–2146, Jul. 2015.

[3] H. Gharavi, and B. Hu, Synchrophasor sensor networks for grid communication and protection, Proceedings of the IEEE, 105(7):1408–1428, Jul. 2017.

[4] H. Gharavi, and B. Hu, Scalable synchrophasors communication network design and implementation for real-time distributed generation grid, IEEE Trans. Smart Grid, 6(5):2539–2550, Sep. 2015.

[5] IEEE standard for synchrophasor measurements for power systems, Std C37.118.1-2011, Dec. 28, 2011.

[6] IEEE guide for phasor data concentrator requirements for power system protection, control, and monitoring, Std C37.244-2013, May 10, 2013.

[7] IEEE standard for synchrophasor data transfer for power systems, Std C37.118.2-2011, Dec. 28, 2011.

[8] B. Yang, et al., Cost-efficient low latency communication infrastructure for synchrophasor applications in smart grids, IEEE Systems Journal, 12(1):948–958, Mar. 2018.

[9] C. F. M. Danielson, et al., Analysis of communication network challenges for synchrophasor-based wide-area applications, in Proc. IREP Symposium Bulk Power System Dynamics and Control - IX Optimization, Security and Control of the Emerging Power Grid, Aug. 2013.

[10] A. Derviskadic, et al., Architecture and experimental validation of a low-latency phasor data concentrator, IEEE Trans. Smart Grid, 9(4):2885–2893, Jul. 2018.

[11] R. H. Khan, and J. Y. Khan, Wide area PMU communication over a WiMAX network in the smart grid, in Proc. Int. Conf. on Smart Grid Communications, Nov. 2012.

[12] M. Chenine, and L. Nordstrom, Investigation of communication delays and data incompleteness in multi-PMU wide area monitoring and control systems, in Proc. Int. Conf. on Electric Power and Energy Conversion Systems, Nov. 2009.

[13] A. Armenia, and J. H. Chow, A flexible phasor data concentrator design leveraging existing software technologies, IEEE Trans. Smart Grid, 1(1):73–81, Jun. 2010.

[14] R. Khan, et al., Analysis of IEEE C37.118 and IEC 61850-90-5 synchrophasor communication frameworks, in Proc. Power and Energy Society General Meeting, Jul. 2016.

[15] M. Kezunovic, et al., Signal processing, communication, and networking requirements for synchrophasor systems, in Proc. Int. Workshop on Signal Processing Advances in Wireless Communications, Jun. 2012.

[16] Y. Seyedi, H. Karimi, and J. M. Guerrero, Centralized disturbance detection in smart microgrids with noisy and intermittent synchrophasor data, IEEE Trans. Smart Grid, 8(6):2775–2783, Nov. 2017.

[17] C. Lackner, et al., A tool to characterize delays and packet losses in power systems with synchrophasor data, IEEE Power Energy Technol. Syst. J., 5(4):117–128, Dec. 2018.

[18] C. Genes, et al., Robust Recovery of missing data in electricity distribution systems, IEEE Trans. Smart Grid, Early Access, DOI:10.1109/TSG.2018.2848935.

[19] Y. Hao, et al., Modelless data quality improvement of streaming synchrophasor measurements by exploiting the low-rank Hankel structure, IEEE Trans. Power Syst., 33(6):6966–6977, Nov. 2018.

[20] K. V. Katsaros, et al., Low latency communication infrastructure for synchrophasor applications in distribution networks, in Proc. Int. Conf. on Smart Grid Communications, Nov. 2014.

[21] S. Das, and T. S. Sidhu, Application of compressive sampling in synchrophasor data communication in WAMS, IEEE Trans. Ind. Inf., 10(1):450–460, Feb. 2014.

[22] P. Castello, et al., Adaptive management of synchrophasor latency for an active phasor data concentrator, in Proc. Int. Instrumentation and Measurement Technol. Conf., May 2017.

[23] X. Zhu, et al., Optimal PMU-communication link placement for smart grid wide-area measurement systems, IEEE Trans. Smart Grid, Early Access, DOI:10.1109/TSG.2018.2860622.

[24] D. Eastlake 3rd and P. Jones, US secure hash algorithm 1 (SHA1), RFC 3174, DOI:10.17487/RFC3174, Sep. 2001.

[25] H. P. Kriegel, et al., LoOP: local outlier probabilities, in 18th ACM Conference on Information and Knowledge Management, Nov 2009, pp. 1649–1652.

[26] LTE; Feasibility study for further advancements for E-UTRA (LTE-Advanced), 3GPP TR 36.912 v. 14.0.0, R. 14, 2017.

[27] F. Malandra, et al., Traffic characterization and LTE performance analysis for M2M communications in smart cities," Pervasive and Mobile Computing,48:59–68, 2018. DOI:10.1016/j.pmcj.2018.05.006.

[28] F. Malandra, et al., Impact of PMU and Smart Meter Applications on the Performance of LTE-based Smart City Communications,"IEEE PIMRC, pp. 1–6, Sept. 2018, Bologna.