Better Late Than Never: Efficient Transmission of Wide Area Measurements in Smart Grids

Jingchao Bao(✉) and Husheng Li

University of Tennessee, Knoxville, TN 37996, USA
jcbao@utk.edu, husheng@eecs.utk.edu

Abstract. The persistent pursuit of reliability dates back to the birth of power system. In the era of smart grid, the harsh requirement extends to the whole system including communication infrastructure. The concentration of wide area synchronized measurements within large system is challenging. In this paper, we investigate the data aggregation issue of phasor measurement units (PMU) data stream in the synchrophasor network, where large latencies lead to unnecessary packet loss. We reduce the final packet loss rate by formulating the data aggregation problem as a multiple stopping time problem. Based on simulation, the success rate booms when compared with single optimal stopping time and multiple fixed-stopping time approaches. Our result could benefit the future development of protocol design, system state estimation and missing data recovery techniques.

1 Introduction

Nowadays, the synchronized PMUs based wide area measurement system (WAMS) accelerates the implementation of smart grid. Unlike the traditional power system, where measurements were gathered in supervisory control and data acquisition (SCADA) in an asynchronous fashion, all these synchronized measurements are marked with GPS time stamp and exchanged through the communication networks in real time to monitor, protect and control the dynamic operation of large area power system. The salient advantages of such system are the inborn time alignment and direct measurement of state instead of indirect system state estimation in the old time. The communication network becomes a critical component to build on. Just like a clot in the vein could cause severe damage to human brain, packet loss in a switch-based communication network for the smart grid system could blind the SCADA and lead to the catastrophic disasters.

However, most available communication infrastructures are built on the principle of probability and only promise to do the best under most circumstances. With no guarantee of the worst case packet loss and latency, the power community is reluctant to accept additional communication infrastructure although the potential benefits are huge. To reassure the doubt, more efforts should be made to mitigate and improve the system design considering the protocol, device and algorithm as a whole system.

L. Duan et al. (Eds.): GameNets 2017, LNICST 212, pp. 121–130, 2017.
DOI: 10.1007/978-3-319-67540-4_11

The concept of phasor network was proposed by 1990 s and has been implemented in the power system since the 21 century [1]. Within the rapid development of computing capability of modern computer and the boom of communication bandwidth in the past two decades, the once reasonable system design is worthy more consideration with new technology. Recently there are some debates on the existence of phasor data concentrator (PDC), mainly because of the additional latency introduced during the data transmission. Moreover, it appears that the time alignment function will magnify the packet loss problem in the large geographic system in two perspectives. The first factor is that current protocol will consider the messages arriving later than the deadline as lost, which converts part of the arrived packets as lost. The second negative factor is caused by the aggregated function in PDC. From the perspective of upper level receiver, a single packet loss from PDC means all the aggregated measurements from lower level are lost.

In [2], authors studied the missing data recovery using the matrix completion. However, it cannot recover the spike signal in the missing data. It is always better to attain the original measurements as much as possible when compared with possible post-recovery process. The authors in [3] discussed two scenarios where a dynamic waiting time is determined by the distributional information of all the latencies of different links. Then it becomes an optimal stopping time problem which could be solved by mathematical tools.

Unlike smart meter system, the phaser network are mostly constructed with wired network, especially fiber communication infrastructure. With the decreasing cost of bandwidth, it is preferable to trade bandwidth with reliability. Here we extend it to a multiple time data aggregation problem in one period with two simple observations. First, the communication bandwidth is considerably cheap compared with old days. We could watch 4 K videos stream on-line while the sample of PMUs are on the level of kbps. Today most of the synchrophasor networks are connected with optical fiber, where the bandwidth could be considered as huge pipe carrying a small stream, yet the reliability are not full optimized. Second, a second chance for packets arriving later than a conservative deadline will always improve the system packet loss rate. There are physical laws which we cannot break in any situation. However, current PDCs may be conservative to limit the deadline to be far ahead of this physical limit. More aggregated packets consisting of later arrived measurements will provide a more comprehensive vision for SCADA. The simulation result validates our assumption, which is given later in details.

The remainder of this paper is organized as follows. The system structure of synchrophasor network and the details of multiple aggregation in PDC are briefed in Sect. 2. Then in Sect. 3 we will give our system model and algorithm. Later we analyze the performance bound on our algorithm. Numerical simulation is compared with the original one in Sect. 4, and conclusions are drawn in Sect. 5.

2 Background of Synchrophasor Network and System Structure

Based on current standards of synchrophasor [4–6], the synchrophasor network consists of PMUs and PDCs in which data streams initiate from the lower level substations where PMUs are located, and then are sent to PDCs in a real time fashion. Typically, one PDC could aggregate these data streams from multiple PMUs in various key locations. Then these intermediate nodes could implement various sophisticated functions on the data streams for monitoring, control or protection.

The ownership of these PDCs usually belongs to different utilities or ISOs. Therefore, the network topology could be complicated as a directed acyclic graph or simple as a tree where measurements are gathered in one or multiple place. The system structure is shown in Fig. 1.

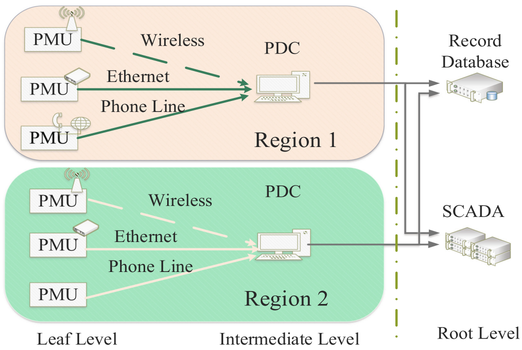


Fig. 1. Synchrophasor network

2.1 Data Aggregation in PDC

In the protocol [6], a PDC could perform different functions, such as data aggregation, data forwarding, data transfer protocols conversion, data latency calculation, redundant data handling etc., to relieve the burden of pre-process in the upper level control center. The specific configuration could be adjusted based on the need. This hierarchical structure usually offers great flexibility and scalability for a large distributed system.

Among all the functions, data aggregation and data forwarding are the most basic and core functions that PDC has. For data aggregation, it could be performed with or without time alignment. PDC should preserve data quality, time quality and time synchronization indication from each signal. For the case with time alignment, it refers to waiting for data with a given time stamp from all sources, placing that data in a packet, and forwarding it to next level. All the data coming to a PDC has been timestamped by the PMU with a time referenced to an absolute time. The PDC aligns received PMU/PDC data according to their timestamps, not their arrival order or arrival time, and transmits the combined data in one or more output data streams to other PDCs or applications such as archiving, visualization, or control.

However, unlike most traffic in commercial networks, the measurement streams in synchrophasor network is more time-critical. A large latency in switched network reduces the value of measurements for some time-stringent applications, especially protection or advanced control in the future. With data aggregation enabled, each level in phasor network would set a latency deadline, which inevitably introduces more latencies for the measurements arriving earlier than the deadline. Moreover, it will further lessen the time conservation of data transmission between this node and next level nodes, and the packets are prone to losses. We will provide detailed discussions in the following.

2.2 Packet Loss and Latency Thresholds

For synchrophasor networks, two metrics are used for measuring the performance. One is packet loss, while the other is latency. However, the problem is more troublesome in current situation.

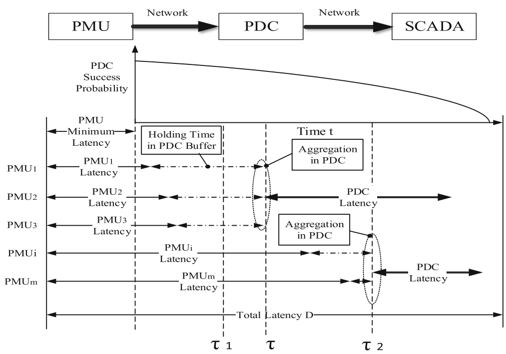


Fig. 2. Single deadline vs multiple grouping in PDC

Today, an empirical and conservative method is single fixed deadline policy, where the waiting time is determined by the empirical latency measurements and dedicated tuned with the best guess. The source of latency varies. To better demonstrate the problem, we compare two scenarios in Fig. 2. Here we denote the single waiting time by τ. With one time slot, there are m PMUs reporting to a common PDC, whose latency is noted by L_i. The total latency allowance D from PMUs to SCADA is determined by the specific application. The PDC will wait and aggregate whatever it receives before the deadline τ and then report to the SCADA in an integrated packet. From the view of SCADA, all these packets have a latency equaling τ instead of $L_i \leq \tau_1$. On the other hand, it further compresses the transmission time for PDC from $D - L_i$ to $D - T_D$, which translates to a higher packet loss rate since it increases the probability that this integrated packet arrives later than D. In addition to that, these packets arrived after τ are ignored by the PDC and considered as lost from the view of SCADA.

Based on the above discussion, the latency in communication system exacerbates packet loss. With the advancement of communication techniques, the cost of reasonable high bandwidth decreases dramatically yet quality of service(QoS) such as packet loss rate has not been improved equally. The idea comes naturally to trade bandwidth for better packet loss rate. Specifically, it means we could arrange PDC to send multiple combined packets to SCADA instead a single one. The benefits are manifold. The first group of measurement sent by $\tau_1 \leq \tau$ has a better chance to reach SCADA in time. On the other hand, the next few groups, such as τ_2, are not abandoned from PDC and could have a considerate probability to be successfully received by the destination. The last but not the least, the shooting time for single deadline approach is really conservative since you do not want to take a risk to choose shooting time with a very low success rate. However, you can choose a shooting time in a larger time period than the single deadline approach, because it could be accepted in the multiple-time aggregation framework.

With all the benefits mentioned, how to choose the shooting time delicately to maximize the benefits remains a unsolved problem. The balance between shooting times and bandwidth cost has not been studied before in the literature. We will attack the problem in the next section by modeling it as a multiple optimal stopping time issue.

3 Multiple Optimal Stopping Time Problem in Data Aggregation

The requirement of latency, denoted by D, varies based the time scale of applications. However, some applications could be more stringent than others. The communication system should and has to satisfy the most stringent application with top priority. Here we assume that D is pre-determined.

3.1 Problem Formulation

Without loss of generality, we assume that one PDC has m PMU data streams to gather. Each link has latency L_i while the latency from PDC to control center is $\hat{L}$. In our analysis, L_i and $\hat{L}$ follow an arbitrary distribution and mutually independent, yet not necessarily identically distributed. Since PDCs are equipped with latency calculation function, it can be safely assumed that we know the stochastic information of all latencies. Furthermore, we have $0 < j \leq n$ deadline τ_j as stopping time. Then, at any moment t, the total number of received measurements is given by

$$N(t) = \sum_{i=1}^{m} \mathbf{1}_{L_i < t} \tag{1}$$

where $\mathbf{1}_{\{\}}$ is the indicator function.

When $t = \tau_j$, we define the reward function as

$$R(\tau_j) = |N(\tau_j) - N(\tau_{j-1})| F_{\hat{L}}(D - \tau_j) - c. \tag{2}$$

In the above equation, $F_{\{\}}$ is the cumulative distribution function of random variable $\hat{L}$. It can be considered as a discount factor for the received yet unsent measurements. The later it sends, the less reward we will get. c is a constant cost for sending one combined packet from PDC to SCADA, which could force the PDC to send with patience and save the bandwidth. Here we ignored the process time for PDC since it can be considered as a fixed time for specific equipment and can be easily incorporated into the requirement of D.

If we have n shooting times, then the total reward is given as

$$R = \sum_{j=1}^{n} [|N(\tau_j) - N(\tau_{j-1})| F_{\hat{L}}(D - \tau_j) - c]. \tag{3}$$

To simplify the notation, we let $N(\tau_0) = 0$ and $\tau_0 = 0$. By maximizing reward R,

$$R^* = \arg \max_{0<\tau_1<\tau_2\cdots<\tau_n} R(\tau_1, \tau_2, \ldots, \tau_n) \tag{4}$$

provides the best aggregation strategy for PDC. Please notice that we neither assign nor limit the numbers of shooting times n in the PDC; however it should be automatically determined by the algorithm. The decision of these shooting times are similar to a sequential decision problem from the view of PDC over time. Therefore it is ready to be optimally solved by stochastic dynamic programming.

3.2 Stochastic Dynamic Programming

We consider a discrete time model for the PDC queue in which each time interval last t_s. The PDC will make a observation for the states. For simplicity, we let $K = \frac{D}{t_s}$ be an integer and time $t \in \{0, 1, \ldots, K\}$. The system has two states during the process. First one is the PMU measurements $N(t)$ received by time t. The other state $S(t)$ is the record of measurements that have been sent by PDC. We use $X(t)$ to represent the tuple $(N(t), S(t))$ concisely.

Action and Strategy: During each time slot, PDC need to make a decision of whether it transmits the messages received by then. We use

$$a(t) = u(X(t), t) = \begin{cases} 1 \text{ if PDC reach a stopping time.} \\ 0 \qquad\qquad \text{otherwise.} \end{cases}$$

as the action. Therefore, the number of shooting is determined by how many $a(t)$ is non-zero over one period.

Dynamic of System States: $N(t)$ could be considered a stochastic process and it can be simplified as a Markov chain with transition probability $P(N(t+1)|N(t))$ under certain assumption. $S(t) = a(t)N(t) + (1 - a(t)) \times S(t-1)$. It will not change until an shooting action is carried out and $S(\tau_j)$ is updated as $N(\tau_j)$.

Benefit Function: Given all states and action of t, we define the gain function in each time slot.

$$C(X(t), a(t)) = a(t)[(N(t) - S(t))F_{PDC}(D - t) - c]$$

Bellman's Function: The key challenge of this scheduling algorithm is how to choose the shooting time given no knowledge of the evolution of states in the future. The action made before current time will have an impact on the expectation of future gain. Given above elements, we have following expectation form of Bell function.

$$J_t(X(t)) = \max_{a(t)\in\{0,1\}} (C(X(t), a(t)) + \mathbf{E}[J_{t+1}(X(t+1))|X(t)])$$

In the next section, we will analyze the performance in the simulation.

4 Numerical Results

Given the problem formulation, we conducted the numeral simulation to verify the performance. Without loss of generality, we assume one PDC between m PMUs and SCADA. We applied Monte Carlo method to generate the random latencies to calculate the average packet loss rates and average numbers of packet that have been sent in one period. We compared our method with other two strategies — single optimal stopping aggregation method proposed in [3] and the scheme of multiple fixed shooting times. The multiple fixed shooting times $T_{fixed}(i)$ are determined by number H of the average packets that have been sent. First we need to find the maximum positive integer $\lfloor H \rfloor$ in different scenarios, and then $T_{fixed}(i) = \frac{D}{\lfloor H \rfloor + 1} \times i$ where $0 < i \leq \lfloor H \rfloor$. All the latencies L_i and $\hat{L}$ are modeled as independent random variables following exponential distribution with parameter λ. The parameters we used in the simulation are listed in Table 1. We compare these approaches by varying one of these simulation parameters. Before we gave detailed results, some general results are summarized based on all cases.

4.1 General Result

As can be seen from all scenarios, our approach outperforms other approaches. However, in most cases, the scheme of multiple fixed shooting times outperforms the one optimal stopping time, if the average sending times are larger than 2. It shows that the packets arrived later than simple deadline should not be abandoned, as long as it does not reach the physical distance limitation. All these

Table 1. Parameters of simulation setup

Cost	0.6
λ_1	5
λ_2	5
PMU Number	15
Latency D	30 ms

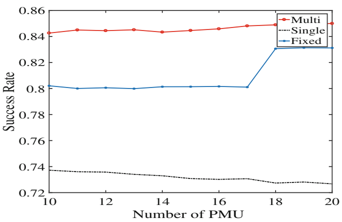
simulations demonstrate that there is still great room for improvement on packet loss in synchrophasor networks.

PMU Number: We varied PMU number m in the simulation scenario but the packet loss rate barely changed as showed in Fig. 3. However, from the average packet number in the figure, we learned that it increases with number of PMU data steams. Considering more PMUs could lead to more processing time in PDC node, it might be efficient to limit the number of PMUs within a small range to reduce the bandwidth cost and preserve more time for the net latency allowance.

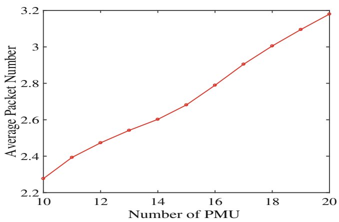
Cost per Packet: This coefficient put a penalty on the total shooting time. By adjusting the cost per packet c in the objective function, we could see that the packet success rate decreases with the cost, which fits our intuition. At the same time, the average shooting time also decreases, which reduces the bandwidth cost. Based on the simulation result, we could guarantee a better packet loss rate by reserving a higher bandwidth.

Latency Distribution: In this part, we change the parameter λ of exponential distribution and result is showed in Figs. 5 and 6. Since the total latency allowance D is fixed, the packet loss rate is negatively correlated with λ as shown in the figures. Since the expected latency increases with λ, we expect that the total budget becomes tighter and thus, more packets can not reach the sink in time. However, the average shooting time also increases with λ in general. We believe that PDC is trying to send more packets at the latter segment of the period. However, the success rate becomes lower due to the increasing expected latency between PDC and PMUs. In a nutshell, an over-tight total latency will waste both bandwidth and PMU measurements (Fig. 4).

Total Latency Allowance *D*: In Fig. 7, three methods finally reach the same level as we relax on the requirement of D. However, our method still beats others under more stringent situations in which total latency allowances are very limited.

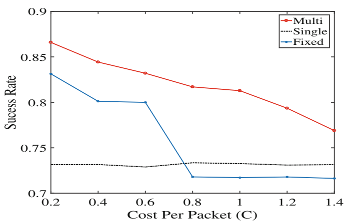


(a) Success Rate

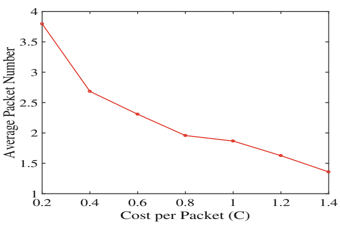


(b) Average Shooting Time

Fig. 3. The impact of number of PMUs

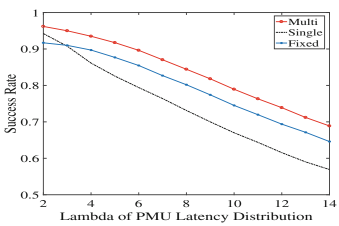


(a) Success Rate

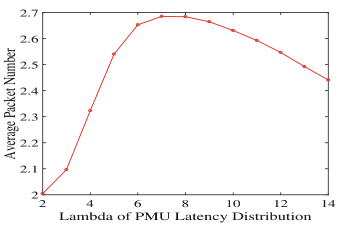


(b) Average Shooting Time

Fig. 4. The impact of cost per packet

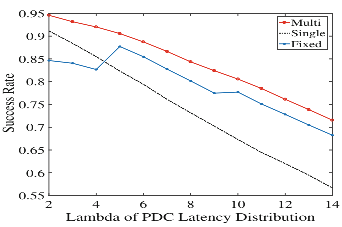


(a) Success Rate

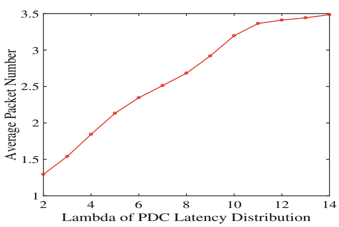


(b) Average Shooting Time

Fig. 5. The impact of distribution λ_1 of PMU latency

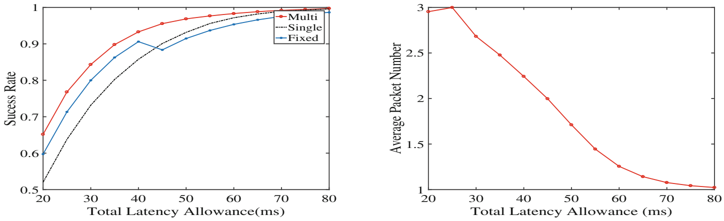


(a) Success Rate



(b) Average Shooting Time

Fig. 6. The impact of distribution λ_2 of PDC latency



(a) Suceess Rate (b) Average Shooting Time

Fig. 7. The impact of total latency D

5 Conclusion

The motivation of WAMS is to maintain a stable system state based on the measurements of geographically dispersed sensors. The latencies of measurements cannot be overlooked and need to be delicately treated to guarantee the QoS in such complex system. In this paper, we have discussed the packet loss issues caused by latency, and offered a simple yet effective approach to mitigate this problem. Without breaking the current protocol and system, the simulation results have shown that it outperforms existing methods. Beyond this result, we will further study more general cases with non-uniform distributions, where the number of states grows exponentially and some approximations are needed to avoid the curse of dimension.

Acknowledgment. The authors would like to thank the support of National Science Foundation under grants ECCS-1407679, CNS-1525226, CNS-1525418, CNS-1543830.

References

1. Phadke, A.G., Thorp, J.S.: Synchronized Phasor Measurements and Their Applications. Springer Science & Business Media, Boston (2008)
2. Gao, P., Wang, M., Ghiocel, S.G., Chow, J.H., Fardanesh, B., Stefopoulos, G.: Missing data recovery by exploiting low-dimensionality in power system synchrophasor measurements. IEEE Trans. Power Syst. **31**(2), 1006–1013 (2016)
3. He, M., Zhang, J.: Deadline-aware concentration of synchrophasor data: an optimal stopping approach. In: 2014 IEEE International Conference on Smart Grid Communications (SmartGridComm), pp. 296–301. IEEE (2014)
4. IEEE Standard for Synchrophasor Measurements for Power Systems. IEEE C37.118.1-2011 (2011)
5. IEEE Standard for Synchrophasor Data Traffic for Power Systems. IEEE C37.118.2-2011 (2011)
6. IEEE Guide for Phasor Data Concentrator Requirements for Power System Protection, Control, and Monitoring. IEEE C37.244 (2013)