

Passive Network Synchronization Based on Concurrent Observations in Industrial IoT Systems

Pengyi Jia^{ID}, *Member, IEEE*, Xianbin Wang^{ID}, *Fellow, IEEE*, and Xuemin Shen^{ID}, *Fellow, IEEE*

Abstract—Accurate network synchronization is crucial to orchestrate distributed infrastructures in Industrial Internet of Things (IIoT) systems for accomplishing network-wide tight temporal collaboration. Traditional clock synchronization can be achieved with extensive exchanges of explicit timestamps for estimating clock offsets, which becomes impractical due to high overhead with the expansion of the network scale. The performance of conventional synchronization will also be dramatically deteriorated due to various uncertainties of IIoT networks. In this article, we propose a passive network synchronization scheme based on concurrent passive observations to calibrate the distributed clocks in IIoT systems while significantly reducing the explicit interactions and network resource consumption during synchronization. By processing the physical phenomena observed concurrently by a group of selected IIoT devices, the local clock offsets of the passive observing devices can be efficiently estimated according to the common time reference linked to the event observed. Multiple relay nodes are further coordinated by the cloud center to disseminate the reference time information throughout the IIoT system. Simulation results demonstrate that by utilizing a series of concurrent observations with efficient coordination, the proposed scheme can achieve accurate and reliable network synchronization for large-scale IIoT systems with significantly reduced network overhead.

Index Terms—Industrial Internet of Things (IIoT), passive network synchronization, principal component analysis (PCA), received signal strength, temporal correlation, timestamp-free.

I. INTRODUCTION

AS ONE of the indispensable enabling technologies for cohesive collaboration within distributed industrial systems, accurate clock synchronization is playing an increasingly important role in the design of Industrial Internet of Things (IIoT) systems to accommodate the stringent industrial requirements [1], [2]. A broad variety of advanced industrial subsystems in large-scale IIoT networks, including wireless sensor and actuator networks (WSAN), intelligent transportation system (ITS), and advanced manufacturing

system (AMS), are often designed to enhance the efficiency, productivity, and reliability of traditional industrial applications [3], [4]. However, these subsystems exclusively hinge on the precise temporal alignment of the data packets exchanged among the involved IIoT devices throughout the entire IIoT network for achieving tight distributed collaboration. Consequently, the network-wide synchronization becomes one of the prerequisites in fulfilling the advanced applications enabled by IIoT systems in terms of ubiquitous sensing and distributed cooperation [5].

Conventional point-to-point clock synchronization schemes, most of which evolved from network time protocol (NTP) [6], precision time protocol (PTP) [7], and flooding time synchronization protocol (FTSP) [8], rely heavily on the frequent exchange of timestamps between the master node (i.e., the node providing reference time) and the slave nodes with inaccurate clocks that require frequent clock calibration [9], [10]. Therefore, extensive interactions among a large number of industrial devices are considered indispensable in achieving accurate network synchronization, which poses daunting challenges in terms of network resource consumption for reliable timestamp provisioning throughout the large-scale IIoT systems. On the one hand, the excessive network overhead and accumulated latency are inevitably incurred by the frequent and explicit interactions during concurrent clock alignment processes of network synchronization, which become intolerable for IIoT systems with the increase of the network scale. The high synchronization overhead and the delayed scheduling of critical data due to the lengthy synchronization procedures for real-time industrial applications will lead to productivity deterioration and the potential risk of shutdown [11].

On the other hand, the underlying network dynamics and uncertainties among distributed IIoT devices will inevitably reduce the reliability of timestamps used during synchronization processes. Due to the dynamic communication conditions in the IIoT networks, including varying Internet link quality [12], random accessing contentions, and dynamic routing processes during end-to-end packet transmission, the timeliness of the timestamps exchanged will become unguaranteed. Meanwhile, the susceptibility of IIoT devices to internal malfunctions and external attacks will lead to unreliable timestamps during clock calibration [13], especially for resource-constrained IIoT devices [14], [15]. Consequently, designing network synchronization protocols without posing strict requirements on the network conditions is particularly useful for IIoT systems.

Manuscript received November 18, 2020; revised February 26, 2021; accepted March 17, 2021. Date of publication March 31, 2021; date of current version September 6, 2021. This work was supported in part by the Natural Sciences and Engineering Research Council of Canada Discovery Program under Grant RGPIN-2018-06254, and in part by the Idea to Innovation Program under Grant I2IPJ 538563-19. (Corresponding author: Xianbin Wang.)

Pengyi Jia and Xianbin Wang are with the Department of Electrical and Computer Engineering, Western University, London, ON N6A 5B9, Canada (e-mail: pjia7@uwo.ca; xianbin.wang@uwo.ca).

Xuemin Shen is with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: sshen@uwaterloo.ca).

Digital Object Identifier 10.1109/JIOT.2021.3070242

Motivated by these challenges, we are interested in developing new efficient network synchronization mechanisms by eliminating the synchronization overhead due to explicit timestamps exchanges. Inspired by the fact that neighboring network nodes, which are physically close to each other, can simultaneously observe the highly correlated physical phenomena around them, the concurrent observations of the same event by neighboring devices could be useful in analyzing the temporal correlation among a group of devices. Some typical applications of this include data fusion and reduction [16], [17], which are enabled by analyzing the temporal consistency of the different observations on the same target. Conversely, by processing concurrent observations of the same physical phenomenon from nearby nodes, the local clocks can be calibrated according to the obtained observation misalignment. As a result, Arao [18] proposed a control message-free synchronization scheme by utilizing the detected events in a wireless sensor network. The local detection time of the involved multiple sensors for the same event is regarded as identical so that their associated clocks can be corrected and synchronized accordingly. However, new protocol designs for global synchronization among distributed devices with different processing capabilities and locations are still needed to support large-scale IIoT systems.

Therefore, to take advantage of the concurrent observations while avoiding the associated deficiencies, a passive network synchronization based on concurrent observation (PANSO) scheme is proposed in this article, with the support of cloud-orchestrated reference time dissemination and PCA-assisted reliability enhancement. Specifically, by processing the concurrent observations of the same physical phenomena at a group of selected devices, a novel passive clock calibration mechanism with minimized explicit interactions is designed to reduce network overhead during the network synchronization process. According to the number of concurrent observations, the distributed nodes in large-scale networks are further orchestrated into isolated nodes and interactive nodes (INs), and reliable INs are selected as relays to disseminate the reference time throughout the network. Additionally, a principal component analysis (PCA) algorithm is designed in the cloud center to ensure reliability during network synchronization by analyzing the historical observation instants uploaded from the local devices while filtering unreliable nodes. The network synchronization performance in terms of achievable accuracy, cost efficiency, and reliability is significantly enhanced as the expected improvements.

The remainder of this article is organized as follows. The concurrent observations in a typical IIoT system are introduced in Section II. The proposed PANSO scheme is designed in Section III in detail, including the observation selection, observation processing, and passive offset estimation. Some proper mechanisms are further designed in Section IV to enhance the performance of PANSO for large-scale IIoT networks, including distance compensation, relay-enabled global synchronization, and PCA-assisted reliability enhancement. Simulation results are carried out in Section V to demonstrate the effectiveness of the proposed scheme in terms of the achievable

accuracy and reliability enhancement, followed by the conclusion in Section VI.

II. CONCURRENT OBSERVATIONS IN IIoT SYSTEMS

Concurrent observations of the same physical phenomenon at different neighboring IIoT devices are critical to provide the time reference to the group of observing devices, forming the cornerstone of the PANSO scheme. In a large-scale IIoT system, the available phenomena for concurrent observations are diverse, including the events during environment monitoring and the broadcasting of electromagnetic signals.

For most of the IIoT systems, a large number of sensor devices are typically deployed in achieving ubiquitous sensing and real-time monitoring of the target environment. Many external events could cause large variations of the environmental parameters monitored, which can be observed by the involved concurrent sensors. The observation instants can be recorded by different observers for further centralized correlation analysis in achieving network synchronization. However, the concurrent observation of the same environmental event requires common sensing capability, which could be hindered by the heterogeneity of the IIoT devices. Furthermore, the response time, processing capacity, limited sampling rate, and diverse locations of the observing sensors will inevitably lead to observation bias and recording inaccuracy, limiting the achievable synchronization accuracy.

In contrast, electromagnetic signals radiated from the nearby wireless transmitters are more accessible to IIoT devices for concurrent observations. Since neighboring network nodes can observe the signal from the same transmitter almost simultaneously, the receiving time at each IIoT device about the same signal can be recorded for further analyzing the observation difference. By sharing the temporal information on the recorded signals with neighboring nodes, synchronization among a group of concurrent observers can be achieved. Moreover, proper coordination for selecting the communication channel and target signals is critical to ensure observation accuracy and synchronization efficiency. Due to the ubiquity of electromagnetic signals among the heterogeneous IIoT devices, the rest of this article will choose the radio signals for a group of neighboring devices as the concurrent observation.

III. PASSIVE NETWORK SYNCHRONIZATION BASED ON CONCURRENT OBSERVATION

The proposed PANSO scheme is achieved by processing the commonly observed physical phenomena for multiple groups of industrial devices in a large-scale IIoT system with the orchestration from the cloud center, as demonstrated in Fig. 1. The overall process of PANSO consists of four successive phases, including concurrent observation selection, concurrent observation processing, reference time expansion, and synchronization performance enhancement. In this section, the local synchronization based on the concurrent observations among a group of nearby IIoT devices in a small network will be introduced in detail.

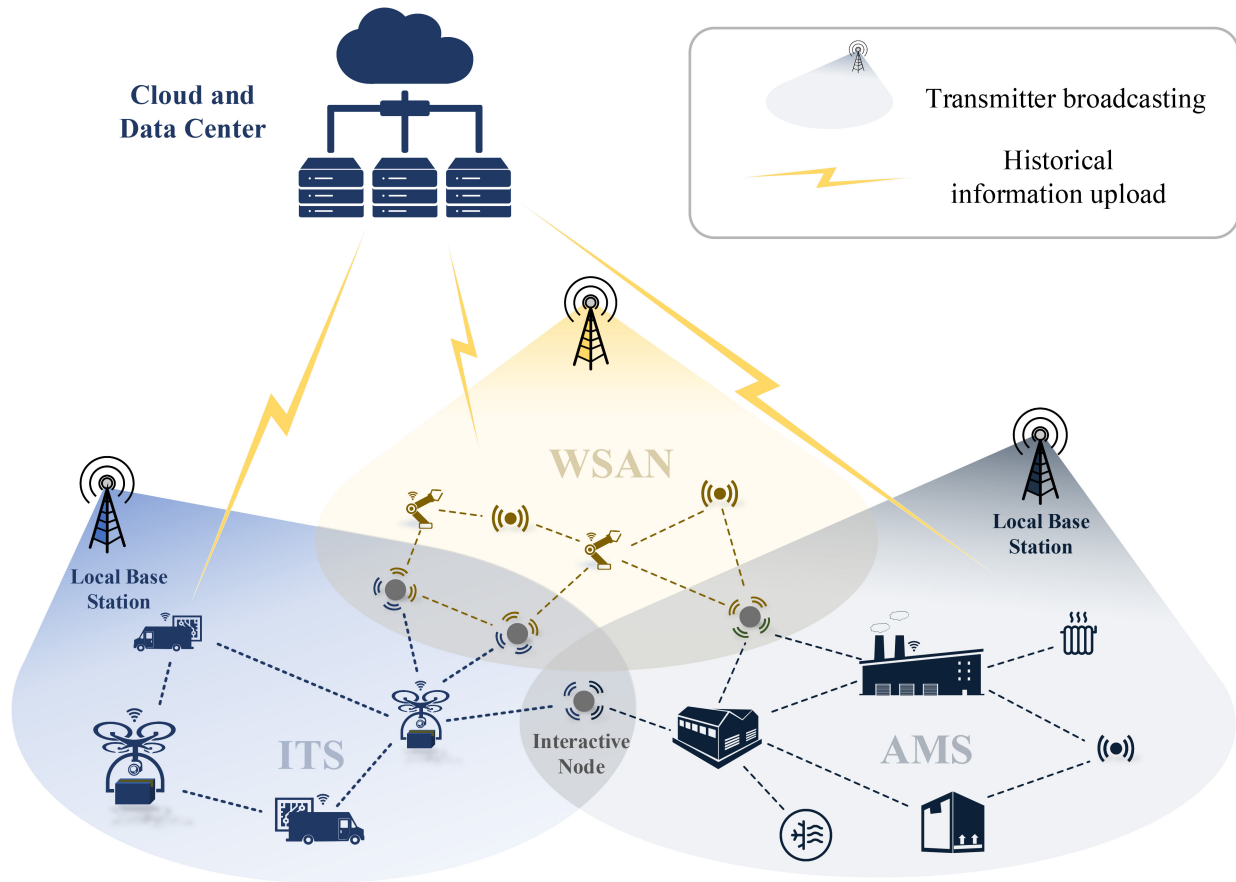


Fig. 1. Typical architecture of a large-scale IIoT system, which consists of multiple subsystems with heterogeneous devices and plants communicating via various protocols. INs are indispensable to share critical information among subsystems for distributed collaboration under the regulation of the remote cloud computing platform.

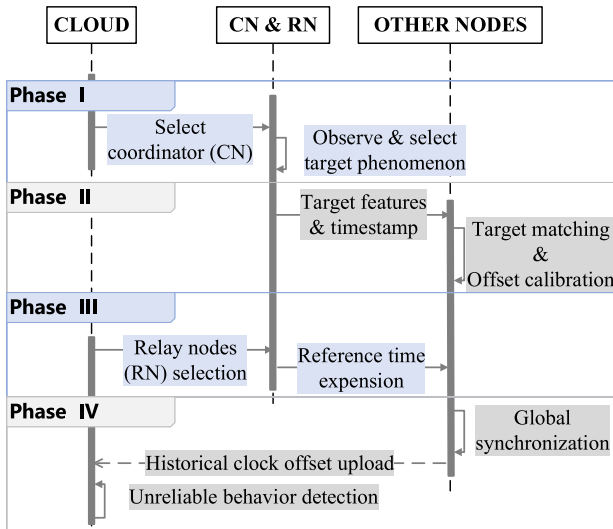


Fig. 2. Proposed PANSO scheme consists of four successive phases, namely, observation selection, observation processing, reference time expansion, and synchronization performance enhancement.

A. Concurrent Observation Selection

As shown in Fig. 2, the concurrent observation selection (i.e., *Phase I*) is initiated by selecting a proper coordinator node (CN) by the cloud center according to a series

of matrices, including the communication capability and clock quality. After the appropriate selection process, CN will be informed with the neighboring topology information and assigned two different tasks in guiding the subsequent synchronization procedures. Initially, CN should define the communication channel and target transmitter that its neighboring devices will listen to, which lays the foundation of the concurrent observation. As a direct result, it is assumed that all the involved IIoT devices are capable to capture the wireless signal from the same target. In the proposed PANSO scheme, target transmitters are defined as the IIoT device with strong communication capability, which can frequently broadcast signals widely to their surrounded devices. Moreover, since the target transmitters in IIoT systems are typically transmitting a large number of packets, CN should select the target signals in advance to enhance the observation efficiency. In practice, many different kinds of signals are unique and convenient to be observed, for instance, the boosting signal of a transmitter always changes dramatically from silence to a strong electromagnetic signal wave, which can be conveniently detected and recorded by its nearby devices. Furthermore, for the communication protocols commonly adopted in IIoT systems, the broadcasting signals are usually associated with the transmitting timestamps in its data packets, which can be further used to validate the observation accuracy.

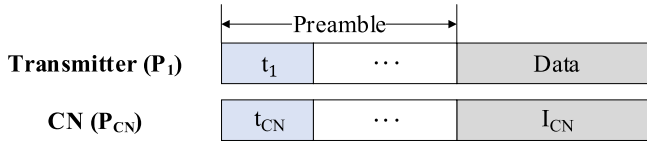


Fig. 3. Example of the PHY packets of CN with regard to the signal transmitted from the target transmitter adopting IEEE 802.11ac protocol. The synchronization-related information is contained in the data field as I_{CN} .

After recording a sufficient number of signals transmitted from the target transmitter, CN will select one of the signals, s_1^{CN} , as the target signal, which is associated with a broadcast address in the MAC layer and obvious features. The transmitting time of s_1^{CN} from the target transmitter (i.e., t_1) enclosed in the corresponding PHY packet P_1 will be recorded for further verification during the successive signal matching process. Meanwhile, the local reception time $\hat{t}_1^{CN}$ and the receive signal strength (RSS) value RSS_1^{CN} associated with the signal s_1^{CN} will also be recorded. Thereby, CN will transmit the relevant information to all the other neighboring devices that are associated with the target transmitter in the data field of its packet P_{CN} . An example that CN adopting IEEE 802.11ac during concurrent observation is shown in Fig. 3, where the shared critical information I_{CN} in the data field of the packet P_{CN} is closely related to the subsequent processing and passive synchronization, given by

$$I_{CN} = (s_1^{CN}, t_1, \hat{t}_1^{CN}, RSS_1^{CN}). \quad (1)$$

B. Concurrent Observation Matching

Since both P_1 and P_{CN} will be received by the neighboring devices of CN, the concurrent observation can be matched for further clock calibration. In Phase II, once the information I_{CN} is received at the local device i , it will compare the target signal s_1^{CN} with its previously recorded local signals s^i that delivered from the same transmitter. To ensure the matching accuracy, both the features of the signals and the corresponding transmission timestamps must be identical. The matched signal is given by

$$s_1^i = \{s | s = s_1^{CN} \cap t_1^i = t_1, \forall s \in s^i\} \quad (2)$$

where t_1^i is the timestamp associated with the signal s^i transmitted from the target transmitter and recorded at the device i . The verification condition $t_1^i = t_1$ can guarantee that the matched signal s_1^i is identical to the target signal defined by CN. Meanwhile, the local device i will also record the reception time $\hat{t}_1^i$ and the associated RSS value RSS_1^i of the received signal for the following processing procedures. Consequently, the local information recorded at the device i after successful signal matching regarding the target transmitter can be written as

$$I_1^i = (\hat{t}_1^i, \hat{t}_1^{CN}, RSS_1^i, RSS_1^{CN}). \quad (3)$$

Since the group of neighboring devices will receive the same signal via dynamic communication environments, the multipath propagation effect will pose a noticeable influence on the target signal analysis. In this case, each local device

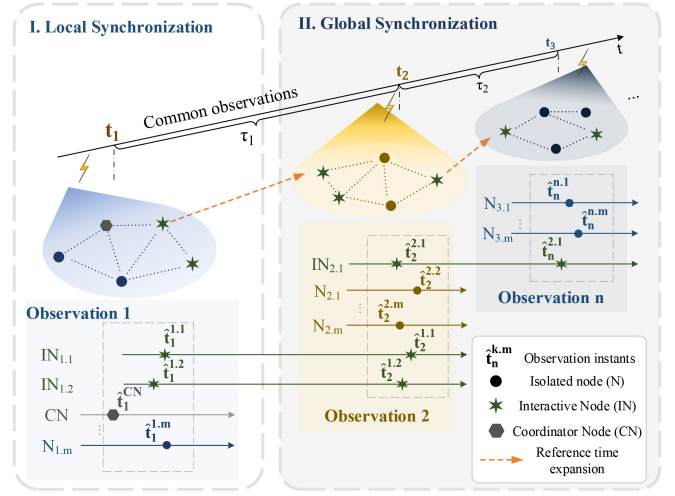


Fig. 4. Time information for every device in a network corresponding to each concurrent observation. The INs are aware of more common observations compared to isolated nodes. The observation instants $\hat{t}_n^{k,m}$ refer to when the n th concurrent observation is recorded at a device, which is the m th member of the group k .

might receive the same signal multiple times continuously with similar data information at slightly different time instants, leading to confusion and inaccuracy of the signal matching. The most straightforward approach in addressing this issue is to sort all matched signals based on the reception time at the local devices and select the earliest one while ignoring the rest. Although it cannot guarantee the elimination of the multipath effect, the observation bias and estimation error induced can be minimized. The final recorded matching time at each device i can be further written as

$$\hat{t}_1^i = \min\{\hat{t}_1^i(k)\} \quad (4)$$

where k denotes a series of observation instants that the target signal is matched at the local device.

C. Passive Offset Estimation

Passive network synchronization of each device is achieved by eliminating the estimated clock offset passively according to the concurrent observation difference between its local information and CN. In achieving passive offset estimation, all devices only need to stay listening to the broadcasting signals transmitted from the surrounded devices. Only CN will be responsible for actively coordinating and transmitting temporal information as the references for other nodes.

A group of devices nearby the target transmitter can have slightly different observation instants on the same signal, shown as the group 1 in Fig. 4. Each device will be associated with a unique record due to its clock offset and the observation bias on the target signal. The observation bias at different devices can be induced by various issues (e.g., the distance between each device and the target transmitter). However, for small-scale networks, it is rational to assume that the same signal can be received at each device almost simultaneously with a negligible error.

The generation time at transmitter 1 and the local observation time at each device regarding the target signal are listed

TABLE I
TIMESTAMPS FOR SIGNAL GENERATED BY THE FIRST TARGET
TRANSMITTER RECORDED AT DISTRIBUTED DEVICES

Node	Transmitter 1	CN	N_1	N_2	...	N_m
Timestamp	t_1	$\hat{t}_1^{CN}$	$\hat{t}_1^{1,1}$	$\hat{t}_1^{1,2}$	...	$\hat{t}_1^{1,m}$

in Table I. Inspired by the fact that the local clock offset will dominant the recorded observation instants, CN will transmit its observation time $\hat{t}_1^{CN}$ on the target signal to its neighboring devices for offset estimation. Based on the information, device i can compare its local observation timestamp $\hat{t}_1^i$ in discovering the clock offset, given by

$$o_i = \hat{t}_1^i - \hat{t}_1^{CN} \quad (5)$$

which will be used to calibrate its local clock. Since only CN is required to disseminate synchronization-related message throughout the network, explicit interactions from the local IIoT devices are eliminated. Therefore, passive clock calibration can be achieved for each local device with reduced explicit resource consumption.

However, in the case that all nodes in one area are distantly distributed, the difference of the signal propagation time will lead to large observation bias, meaning that even each device with an accurate clock will have a distinctive observation time for the target signal. By simply assuming all devices can simultaneously observe the target signal will cause synchronization error as an inevitable consequence. Affected by the difference of the propagation distance, the updated offset of device i compared to CN can be estimated as

$$\hat{o}_i = \hat{t}_i^n - \hat{t}_{CN}^n - \frac{d_{ci}}{c} \quad (6)$$

where c is the transmission speed of an electromagnetic signal, while d_{ci} is the relative difference between the distance of CN and device i with respect to their common target transmitter.

IV. GLOBAL SYNCHRONIZATION AND PERFORMANCE ENHANCEMENT

With the expansion of the IIoT network scale, local passive synchronization for small networks is insufficient to support the overall performance of the clock behavior, due to the long propagation distance and limited communication range. As a consequence, proper mechanisms should be designed to guarantee the global synchronization performance.

A. Propagation Effect Compensation

According to (6), the distribution of the devices in an IIoT system will lead to inaccuracy during offset estimation. Depending on the network scale, this inaccuracy can be tiny as submicrosecond levels with a distance of tens of meters, or a few milliseconds if the distance increases to kilometers. Therefore, distance compensation is necessary to enhance observation accuracy. The location awareness for the IIoT devices is different due to their heterogeneity in terms of mobility, cost, and functionality. For instance, in the smart transportation subsystem, every node is typically

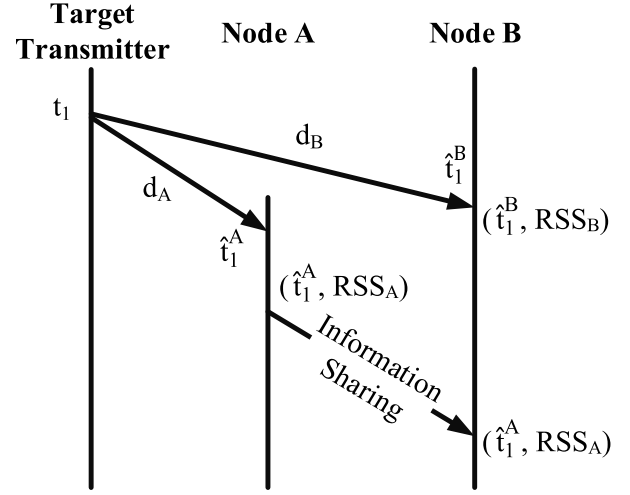


Fig. 5. Critical information transmission in terms of observation instants and RSS values between the common target transmitter and two devices, where the difference of the distance will induce synchronization error.

equipped with a location-aware unit (e.g., GPS) to accurately estimate the distance between each device and its selected target transmitter. However, for devices without costly GPS units or deployed in indoor and underground environments, other methods for calculating the relative distance are necessary. Since these IIoT devices are typically assigned to perform regular tasks with steady objectives, schedules, and places, it is reasonable to assume that their locations are fixed in most of their life cycles.

Since the target signals are always transmitted with the corresponding RSS values during information exchange, the RSS values can be used to compensate for the effect of distance [19] according to the fact that each device can calculate the approximate distance to the target transmitter based on the RSS value. To achieve RSS-based relative distance estimation, CN will deliver the RSS value for the received signal from the target transmitter together with the time information to its subordinates, as shown in Fig. 5. By assuming the fixed devices follow the free space propagation model during their communication, the relation between the distance and the corresponding RSS value can be obtained, while the distance between the target transmitter and CN can be calculated by

$$R_c = A_0 - 10n \lg(d_c) \quad (7)$$

and the distance for the device i can be similarly written as

$$R_i = A_0 - 10n \lg(d_i) \quad (8)$$

where n is the path-loss factor affected by the environment, and A_0 is the transmission power from the target transmitter. d_c and d_i are the distances from the transmitter to CN and local devices, respectively. According to (7) and (8), the difference between the two distances can be obtained as

$$d_{ci} = d_c - d_i \quad (9)$$

so that the offset estimation can be further updated from (6). Due to different issues, including the inaccuracy of propagation models and the uncertainty of RSS values, slight errors

will be inevitably induced to the distance estimated by (9). However, the proposed RSS-based distance compensation can still bring substantial improvement to the achievable synchronization accuracy, especially for large-scale IIoT networks.

B. Reference Time Expansion

By calibrating the estimated offset $\hat{o}_i$ for each local device, the nodes around CN can be synchronized accurately. However, for a large-scale network with massive nodes that cannot receive the same signal from a common target transmitter, global synchronization is unachievable with a single CN. Therefore, multiple nodes should be selected in *Phase III* of the proposed scheme to efficiently disseminate the reference time throughout the entire network without reducing the clock accuracy for the distantly distributed devices. Since IN can receive signals from multiple target transmitters in different areas, some of them will be assigned as relay nodes (RNs), according to Algorithm 1.

Specifically, all devices in the network will first listen to the broadcasting signal from the surrounded transmitters and upload the received transmitter ID to the cloud center. If more than one ID of the target transmitters are recorded at the cloud center for a single device, it will be marked as an IN, and one of the INs will be randomly selected as the RN to expand the time reference if multiple INs exist between any two subsystems. Due to the randomness of the initial RN election, successive RN selection and filtering mechanisms are necessary to guarantee the accomplishment of global synchronization. Finally, all other devices will only need to passively synchronize to the received time reference without expending the limited network resources on explicit synchronization processes.

After the selection process, RNs will be responsible for defining succeeding target signals and deliver the associated temporal information to the neighboring areas. As shown in Fig. 4, the selected RNs are aware of the information from multiple target transmitters, while isolated nodes can only sense partial information due to lack of connectivity. Consequently, RNs can expand the reference time gradually in achieving global synchronization. For the isolated node i within the j th group, its local clock offset can be calculated as

$$o_i = \hat{t}_j^{j,i} - \hat{t}_j^{j-1} = \hat{t}_j^{j,i} - \hat{t}_1^1 - \sum_{k=1}^{j-1} (o_k^{\text{RN}} + \tau_k) \quad (10)$$

where o_k^{RN} is the local clock offset for each RN compared to the clock reference and τ_k is the interval between each two common signal observations, which can be accurately calculated by each RN.

C. PCA-Assisted Reliability Enhancement

Another critical issue in *Phase IV* is the robust synchronization provisioning under various security and reliability issues during synchronization in hostile industrial environments. Lacking security-related mechanisms can lead to untrustworthy clock information expansion and potential inconsistent cooperation among the connected devices. As PCA is powerful

Algorithm 1 Cloud-Assisted Initial Node Arrangement

```

1: Cloud select the coordinator node (CN)
2: for Every target transmitter  $m$  in its subsystem do
3:   Broadcast its identification  $ID_m$  to surrounding devices

4:   for Each of its belonging devices do
5:     Record  $ID_m$  and upload to cloud center
6:   end for
7: end for
8: for Each device in the system  $i$  do
9:   Cloud center calculate its uploading times  $k_i$ 
10:  if  $k_i > 1$  then
11:    Mark it as an interactive node (IN)
12:    Record uploaded transmitters as neighboring ones
13:  else
14:    Mark it as an isolated node
15:  end if
16: end for
17: for Each neighboring transmitter do
18:  if Only one IN exists then
19:    Mark it as the relaying node (RN)
20:  else
21:    Randomly select one IN as RN
22:  end if
23: end for

```

in investigating abnormal behaviors [20], [21], in this section, a PCA-assisted scheme is designed to enhance the synchronization reliability among the involved devices. Moreover, since the performance of PCA algorithms is affected by the network issues severely, it will be beneficial to consider a series of historical concurrent observations, which is not significantly affected by the network conditions, as the core of the successive analysis. Specifically, the cloud center will collect the observation timestamp at the k th physical phenomenon from each node i as $\hat{t}_k^i$ and its corresponding relay timestamp as $\hat{t}_k^1$, so that the relative clock offset can be calculated as

$$o_k^i = \hat{t}_k^i - \hat{t}_k^1 \quad (11)$$

while the relative skew between node i and its RN can be estimated by

$$\hat{\alpha}_k^i = \frac{\hat{t}_k^i - \hat{t}_{k-1}^i}{\hat{t}_k^1 - \hat{t}_{k-1}^1}. \quad (12)$$

Due to the uniqueness of each clock skew and the effect of external operating temperature on the clock oscillation frequency as validated in [22], the estimated clock skew can be used as critical evidence for detecting abnormality of the clocks. Together with the local temperature for the node denoted by T_k^i , a $4 \times k$ matrix can be formulated as X_i for the node i with four physical observations and k samples, i.e.,

$$X_i = \begin{bmatrix} \hat{t}_k^i \\ o_k^i \\ \hat{\alpha}_k^i \\ T_k^i \end{bmatrix} = \begin{bmatrix} \hat{t}_1^i & \hat{t}_2^i & \dots & \hat{t}_k^i \\ o_1^i & o_2^i & \dots & o_k^i \\ \hat{\alpha}_1^i & \hat{\alpha}_2^i & \dots & \hat{\alpha}_k^i \\ T_1^i & T_2^i & \dots & T_k^i \end{bmatrix}. \quad (13)$$

Thereby, the raw data matrix X_i should be transformed into a new coordinate subsystem by

$$W_i = Q_i X_i \quad (14)$$

where Q_i is the transformation matrix and W_i is the corresponding projection referred to as score matrix. Since the time information and temperature information have different scales, data normalization should be adopted to mitigate the effect. Therefore, a zero-mean and unit-variance matrix $\hat{X}_i$ for each node is obtained, which can be used to derive the covariance matrix by

$$C_{X_i} = \frac{1}{n-1} \hat{X}_i^* \hat{X}_i \quad (15)$$

where $*$ is the conjugate transpose operator.

According to the calculated covariance matrix, its eigenvalue and eigenvector can be obtained by

$$V_i = U_i^{-1} C_{X_i} U_i \quad (16)$$

where V_i is the diagonal matrix with eigenvalues and U_i is the matrix of the eigenvectors of C_{X_i} , respectively. By selecting a few largest columns in U_i as the principal components, the original eigenvalue matrix V_i and eigenvector matrix U_i can be reduced to $\hat{V}_i$ and $\hat{U}_i$, with the principal components (i.e., the most dominant parts). By selecting the principal parts of the eigenvector matrix as the transformation matrix Q_i , the projection of the raw data can be described as

$$\hat{W}_i = Q_i \hat{X}_i = \hat{U}_i^T \hat{U}_i. \quad (17)$$

Consequently, the remaining value after extracting the principal components, referred to as the residual value, can be obtained as

$$e_{X_i} = \hat{X}_i - Q_i^T \hat{W}_i \quad (18)$$

based on which an SPE score for each local device can be calculated as the criteria for anomalous timestamps detection, given by

$$SPE_i = \sqrt{e_{X_i}^T e_{X_i}}. \quad (19)$$

Since SPE_i score after principal extraction is typically very small, any induction of anomalous timestamps will result in a huge fluctuation, based on which the corresponding local device can be diagnosed as unreliable. The real-time SPE_i value will be calculated and compared to the historical score $\overline{SPE}_i$ so that any newly collected timestamps meet the corresponding criteria will be reported by the cloud center as unreliable, which is given by

$$SPE_i(t) - \overline{SPE}_i > \rho_i \sigma_{SPE} \quad (20)$$

where σ_{SPE} is the standard deviation of the SPE score for the historically collected data, while ρ_i is the threshold for determining the abnormality.

In terms of any unreliable nodes are detected, synchronization will be conducted by filtering the abnormal devices and their associated timestamps to enhance the overall synchronization reliability. Meanwhile, RN should be reselected from the trustworthy devices to avoid unexpected deterioration of the synchronization performance.

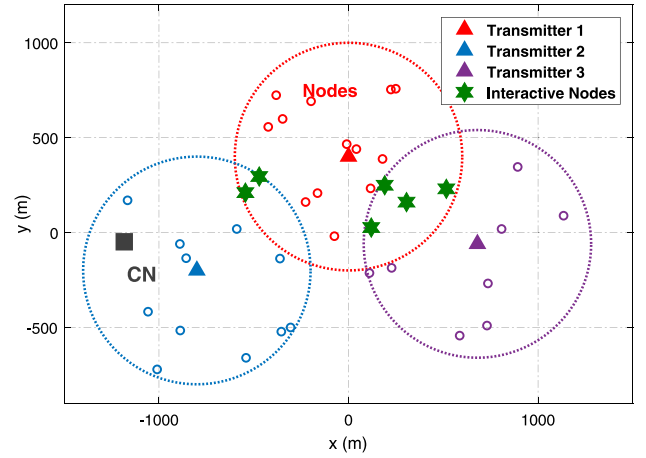


Fig. 6. Distribution of the devices with three target transmitters in the proposed system. One CN and multiple INs exist in the network.

V. PERFORMANCE EVALUATION

In this section, a series of simulations are conducted to evaluate the performance of the proposed PANSO scheme from several aspects, including the distance-compensated clock offset estimation, reference time expansion, and reliability enhancement with abnormal timestamps.

A. Simulation Settings

In this simulation, a total number of 40 nodes are randomly deployed in a large industrial environment of $3000 \text{ m} \times 2000 \text{ m}$, where the distance between two nodes in one group can range from 30 m to 1000 m. Surrounded by the 40 nodes, there are three target transmitters that can transmit and broadcast wireless signals to their neighboring nodes occasionally. The distribution of the devices is shown in Fig. 6, where CN is selected by the cloud center prior to the simulation. Meanwhile, there are multiple INs between every two groups, which can receive the broadcast information from more than one target transmitter. Two RN will be further selected from the available INs during the reference time expansion phase. The simulations will be initiated by the CN in terms of the concurrent observation selection within its local group of devices, which surrounds the same target transmitter.

Moreover, the effect of unreliable issues in terms of internal malfunctions of the IIoT devices and external malicious attacks are considered during performance evaluation. Three kinds of different abnormalities that may lead to anomalous timestamps are considered, including inner malfunction, delay attacks, and tampering attacks, as listed in Table II for definition and mathematical expressions. More specifically, external attacks are contrary to internal malfunctions, where the faulty timestamps are caused by the physical damage or interference of the enclosed crystal oscillator. The corresponding consequence of the internal malfunctions is the abnormal sensitivity of the enclosed oscillator to the environmental variations, where η_i is the temperature sensitivity factor for node i , and γ_i reflects the degree of abnormality. Moreover, two types of external attacks are, respectively, considered, while each of them will

TABLE II
THREE KINDS OF ABNORMAL BEHAVIORS CONSIDERED IN THE
SIMULATION EVALUATION (SECTION V-D)

Types of Unreliability	Descriptions	Quantitative Expression
<i>Internal malfunctions</i>	Unexpected skew variation with environmental changes	$\eta_i = \gamma\eta_i$
<i>Delay attacks</i>	Delayed timestamps are delivered by the malicious relay nodes	$t_i^k = t_i^k - d_m$
<i>Tampering attacks</i>	Fake timestamps are transmitted by the abnormal devices	$t_i^k = t_m^k$

deteriorate the trustworthiness of the timestamps delivered. On the one hand, malicious RN will take advantage of delay attacks to transmit outdated timestamps to the local devices that are waiting for reference time information. By synchronizing with delayed timestamps, the clock accuracy cannot be improved as one of the direct consequences. Meanwhile, since the delayed timestamps are generated by the same clock, some of the existing mechanisms that rely on accurate clock modeling (e.g., [23]) may become less effective. On the other hand, tampering attacks will be harmful to all devices since their observation instants recorded can be manipulated so that the offset estimation and succeeding synchronization will be inaccurate. The synchronization performance under tampering attacks will be random and unpredictable due to the injection of unreliable timestamps.

B. Distance-Compensated Offset Estimation

The most fundamental part of synchronization is the estimation of the clock offset for the distributed devices within the network compared to the reference time, as well as the timely elimination of the clock offset. Typically, one of the critical criteria for determining synchronization performance is the averaged clock error among the involved devices throughout the network operation. In the proposed PANSO scheme, the clock offset is estimated by comparing the timestamps of the locally matched observation with the received target signal at each local device. Meanwhile, as analyzed in Section V-B, the effect of propagation delays will lead to inevitable observation bias if the location of each device is unknown. As a consequence, an RSS-based method is adopted in the proposed scheme to address this issue, where the relative distances from the target transmitter to its neighboring devices are calculated according to the RSS values when the target signals are locally received, in accordance with (7)–(9).

The distance estimation accuracy before clock calibration is shown in Table III, where the distance from each local device to its belonging target transmitter is different due to the random deployment, leading to an averaged distance in each group ranging from around 100 m to 400 m. According to the RSS values obtained from the target signals, the estimated averaged distances are very close to the ground truth, with only a few meters as the estimation bias. As a consequence,

TABLE III
RECEIVED-SIGNAL-STRENGTH-BASED DISTANCE ESTIMATION
ACCURACY FOR THREE GROUPS OF DEVICES

	Real Avg. distance (m)	EST Avg. distance (m)	EST bias (m)	EST bias (μ s)
Group 1	250.40	260.95	10.55	0.035
Group 2	365.29	368.92	3.62	0.012
Group 3	136.90	132.17	4.73	0.016

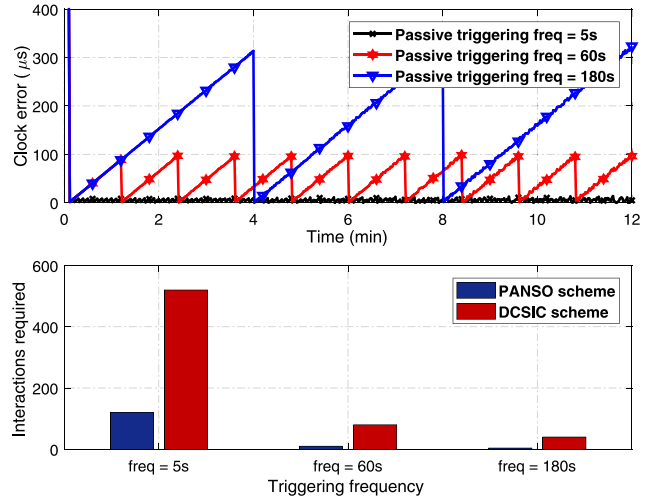


Fig. 7. Evolution of clock errors after adopting PANSO with three different triggering frequencies. Compared to the conventional active synchronization methods, the passive one only requires a very small number of interactions for various triggering frequencies.

the estimation bias in terms of the propagation time is much smaller than 1μ s, meaning that the effect of estimation bias on the synchronization performance can be negligible for further clock calibration.

After adopting the proposed RSS-based distance compensation method, the clock calibration can be achieved among the first group of devices based on the reference time provided by CN. As shown in Fig. 7, clock errors can be accurately calibrated and almost eliminated after triggering the passive synchronization actions. Meanwhile, a higher triggering frequency can result in a better synchronization performance in terms of the achievable accuracy in long-term operations. For example, with a triggering frequency of 5 s, submicrosecond synchronization performance is attainable for a group of devices, while a synchronization frequency of 180 s will show an averaged clock error that exceeding 150μ s. Therefore, application-oriented synchronization can be provisioned with controllable clock accuracy to avoid overwhelming synchronization processes. Moreover, compared to the conventional synchronization methods that hinge on frequent explicit interactions, e.g., DCSIC scheme [10], PANSO can achieve accurate clock calibration with significantly reduced interactions for all IIoT devices. As a direct consequence, the uncertain network conditions will have a less significant effect on the synchronization performance, while more

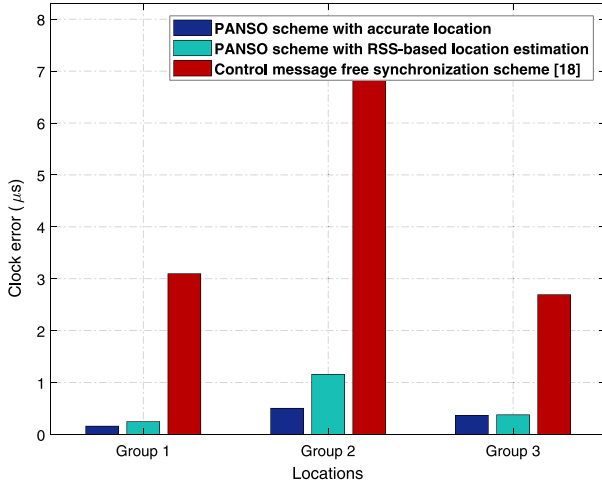


Fig. 8. Comparison for the averaged achievable clock accuracy among each group of devices in three situations considering different distance-related strategies. RSS-based method can greatly enhance the clock accuracy.

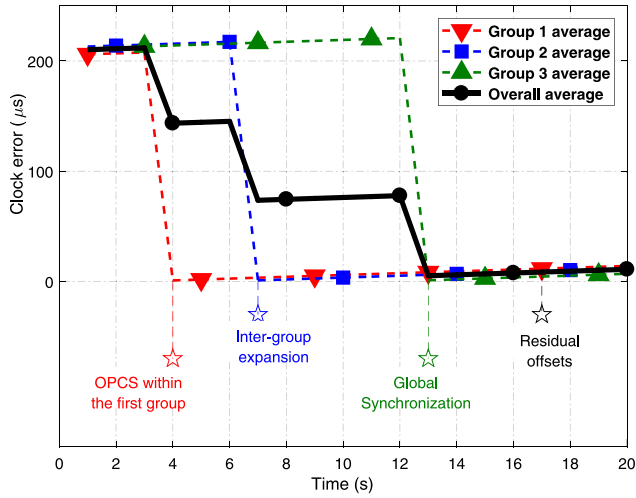


Fig. 9. Expansion of time reference for three areas assisted by the RNs. Global synchronization is achieved gradually with minor residual offset after synchronization.

network resources can be saved for other critical industrial applications.

C. Reference Time Expansion

After achieving a consensual clock for a group of devices, global synchronization should be subsequently performed with the assistance of RNs. For each group of devices, the proposed RSS-based distance compensation scheme should be adopted in advance to enhance the offset estimation accuracy. As shown in Fig. 8, a simulation is conducted to evaluate the performance of the proposed PANSO scheme in three different types of location-related situations, namely, aware of the accurate location of each device, RSS-based distance compensation method, and the control message free scheme [18], which disregarded the distribution nature of the nodes. It can be observed that for all three groups of devices, the clock errors for the case with accurate location information is the smallest, while PANSO adopting an RSS-based distance compensation

approach can reduce the clock error by a few microseconds, which is a significant improvement as compared to the existing study. Moreover, it indicates that with more hospitable communication environments or GPS-embedded devices in outdoor situations, a higher clock accuracy is expected since the distance estimation accuracy will be enhanced.

With the distance information estimated, global synchronization can be achieved accordingly, which is illustrated in Fig. 9. Since there are three neighboring areas required to be synchronized, the network synchronization is achieved gradually with the reference time expansion. Initially, the PANSO is conducted in the first group where CN is selected, while the averaged clock error within the first group will dramatically drop after clock calibration. Then, the intergroup time reference expansion is conducted by the selected RNs, and other groups of devices will be synchronized accordingly. Therefore, the averaged clock errors for the three groups will reduce in different time instants, and the overall averaged clock error will appear a decrement three times. The accomplishment of global synchronization is achieved at the end, where all devices are accurately synchronized based on the local clock offsets calculated. After global synchronization, there will be a minor clock error remaining, which is due to the calculation error and the effect of the clock skew. This residual issue will contribute to the clock errors in the long-term operation so that periodic clock calibrations are necessary to limit the residual offset within the application-oriented expectations.

D. Unreliable Node Detection

The reliability and security during synchronization can be affected by a wide range of issues, as illustrated in Table II. Both the internal and external issues will lead to anomalous timestamps during information exchange, causing unexpected synchronization errors, especially when unreliable RNs are involved during reference time expansion.

PCA-assisted clock information analysis is achieved in the cloud center for both the abnormal nodes detection and reliable RNs selection. Any nodes with uncommon time information will be regarded as unreliable, while only reliable INs will be selected as RNs for reference time expansion. The performance of the proposed PCA-assisted scheme is demonstrated in Fig. 10 based on the temporal information uploaded from the local devices. Specifically, the effectiveness of the detection scheme is denoted by the true positive rate (TPR), which is the ratio of the correct detection r_c and the overall anomalous instants M_p . TPR for each local device can be calculated by

$$\text{TPR} = \frac{r_c}{M_p} = \frac{r_c}{\sum_{i=1}^{N_p} \sum_{j=1}^{S_p} C_i(t_j)} \quad (21)$$

where $C_i(t_j)$ is the clock generation at the instant j of the device i . N_p and S_p are the total number of positive cases (i.e., anomalous devices and instants), respectively. It can be observed from Fig. 10 that with the increment of the detection threshold ρ , the detection precision is enhanced accordingly. In other words, more unreliable devices can be successfully discovered by setting a stricter condition when the PCA-assisted

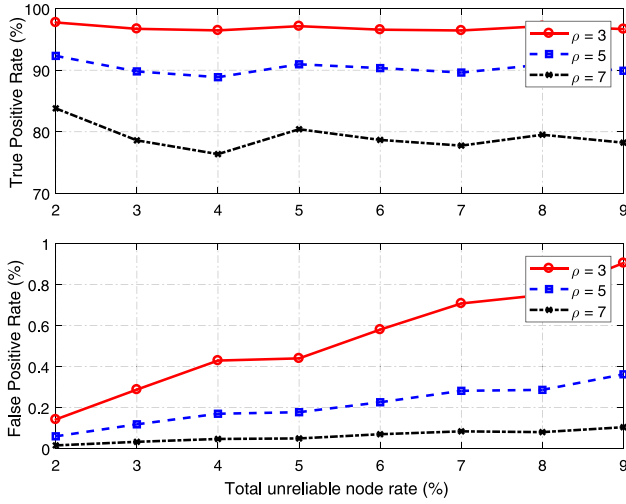


Fig. 10. Performance of the proposed PCA-assisted unreliable node detection. With more unreliable nodes and higher detection threshold, the detection precision is increased with slightly increased false positive cases.

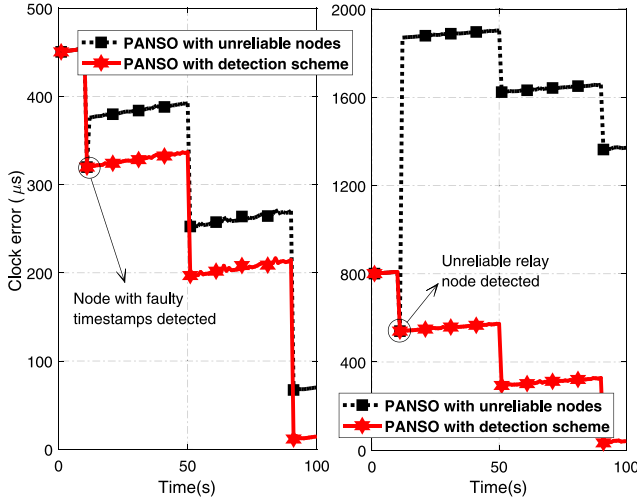


Fig. 11. Improvement of clock accuracy after adopting the PCA-assisted unreliable node detection. The enhancement can be more significant if unreliable RNs are successfully detected.

scheme is adopted. However, a stricter condition will lead to false positive (FP) cases (i.e., normal devices are incorrectly detected as anomalies). The false-positive rate (FPR) in the proposed study is defined as the ratio of the FP reports r_e and the total number of normal instants M_n during the simulation period S_t , given by

$$\text{FPR} = \frac{r_e}{M_n} = \frac{r_t - r_c}{N_t S_t - M_p} \quad (22)$$

where r_t and N_t are the total number of anomaly reports and the total number of devices in the network, respectively. It can be observed that the FPR is slightly increased with a stricter detection threshold, while the largest FPR is still lower than 0.5%, indicating that a reliable detection accuracy without overwhelming false alarms is achieved. Therefore, in this scheme, a stricter detection is preferred so that more unreliable devices can be successfully detected and the synchronization performance can be enhanced accordingly.

By utilizing the PCA-assisted detection scheme according to the historically recorded observation instants, several kinds of unreliable devices can be filtered so that future synchronization can be achieved with trustworthy nodes. The improvement of the synchronization performance is shown in Fig. 11, where it can be seen that an isolated node with an error message will lead to the increment of the clock inaccuracy. Moreover, unreliable RNs will cause significant clock error due to the failed expansion of accurate reference time to the neighboring groups. However, by adopting the proposed anomaly detection scheme, both unreliable local devices and RNs can be effectively detected so that the clock errors are dramatically reduced. It is worth noting that if any of the initially assigned RNs are reported as unreliable after performing the PCA-assisted detection scheme in the cloud center, other reliable INs should be further selected as RNs to ensure the accomplishment of time information expansion and network-wide synchronization.

VI. CONCLUSION

In this article, PANSO has been designed to achieve synchronization without dedicated resource consumption during timestamps information exchange. Specifically, by processing the commonly observed physical phenomenon in a group of IIoT devices, passive clock calibration can be achieved at each local IIoT device with the alignment of the observation instants. RSS-based distance estimation method has been adopted to compensate for the effect of propagation latency. Additionally, a PCA-assisted anomaly detection scheme has been designed at the cloud center to investigate unreliable devices based on the historically uploaded observation instants to substantially enhance the synchronization performance under various uncertainties. Extensive simulation results have demonstrated that the proposed PANSO scheme can achieve passive clock calibrations and accurate network-wide synchronization under various kinds of abnormalities. In the future, we will establish an experimental platform to validate the practical performance of the proposed PANSO scheme under different communication conditions.

REFERENCES

- [1] M. Lévesque and D. Tipper, "A survey of clock synchronization over packet-switched networks," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 4, pp. 2926–2947, 4th Quart., 2016.
- [2] Z. Ma, M. Xiao, Y. Xiao, Z. Pang, H. V. Poor, and B. Vucetic, "High-reliability and low-latency wireless communication for Internet of Things: Challenges, fundamentals, and enabling technologies," *IEEE Internet Things J.*, vol. 6, no. 5, pp. 7946–7970, Oct. 2019.
- [3] M. Wollschlaeger, T. Sauter, and J. Jasperneite, "The future of industrial communication: Automation networks in the era of the Internet of Things and industry 4.0," *IEEE Ind. Electron. Mag.*, vol. 11, no. 1, pp. 17–27, Mar. 2017.
- [4] Y. Liao, E. de Freitas Rocha Loures, and F. Deschamps, "Industrial Internet of Things: A systematic literature review and insights," *IEEE Internet Things J.*, vol. 5, no. 6, pp. 4515–4525, Dec. 2018.
- [5] A. Mahmood, M. I. Ashraf, M. Gidlund, J. Torsner, and J. Sachs, "Time synchronization in 5G wireless edge: Requirements and solutions for critical-MTC," *IEEE Commun. Mag.*, vol. 57, no. 12, pp. 45–51, Dec. 2019.
- [6] D. Mills, J. Martin, J. Burbank, and W. Kasch, "Network Time Protocol Version 4: Protocol and Algorithms Specification," ITU, RFC 5905, Jun. 2010. [Online]. Available: <http://www.rfc-editor.org/rfc/rfc5905.txt>

- [7] *IEEE Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems*, IEEE Std 1588–2008 (Revision of IEEE Std 1588–2002), pp. 1–300, 2008.
- [8] K. S. Yildirim and A. Kantarci, “Time synchronization based on slow-flooding in wireless sensor networks,” *IEEE Trans. Parallel Distrib. Syst.*, vol. 25, no. 1, pp. 244–253, Jan. 2014.
- [9] H. Wang, L. Shao, M. Li, and P. Wang, “Estimation of frequency offset for time synchronization with immediate clock adjustment in multihop wireless sensor networks,” *IEEE Internet Things J.*, vol. 4, no. 6, pp. 2239–2246, Dec. 2017.
- [10] P. Jia, X. Wang, and K. Zheng, “Distributed clock synchronization based on intelligent clustering in local area industrial IoT systems,” *IEEE Trans. Ind. Informat.*, vol. 16, no. 6, pp. 3697–3707, Jun. 2020.
- [11] L. Chettri and R. Bera, “A comprehensive survey on Internet of Things (IoT) toward 5G wireless systems,” *IEEE Internet Things J.*, vol. 7, no. 1, pp. 16–32, Jan. 2020.
- [12] F. Naem, M. Tariq, and H. V. Poor, “SDN-enabled energy-efficient routing optimization framework for industrial Internet of Things,” *IEEE Trans. Ind. Informat.*, early access, Jul. 3, 2020, doi: [10.1109/TII.2020.3006885](https://doi.org/10.1109/TII.2020.3006885).
- [13] K. Fan *et al.*, “Blockchain-based secure time protection scheme in IoT,” *IEEE Internet Things J.*, vol. 6, no. 3, pp. 4671–4679, Jun. 2019.
- [14] H. Fang, A. Qi, and X. Wang, “Fast authentication and progressive authorization in large-scale IoT: How to leverage AI for security enhancement,” *IEEE Netw.*, vol. 34, no. 3, pp. 24–29, May/Jun. 2020.
- [15] L. Xing, “Cascading failures in Internet of Things: Review and perspectives on reliability and resilience,” *IEEE Internet Things J.*, vol. 8, no. 1, pp. 44–64, Jan. 2021.
- [16] T. Yu, X. Wang, and A. Shami, “A novel fog computing enabled temporal data reduction scheme in IoT systems,” in *Proc. IEEE Glob. Commun. Conf. (GLOBECOM)*, 2017, pp. 1–5.
- [17] F. H. Bijarbooneh, W. Du, E. C. H. Ngai, X. Fu, and J. Liu, “Cloud-assisted data fusion and sensor selection for Internet of Things,” *IEEE Internet Things J.*, vol. 3, no. 3, pp. 257–268, Jun. 2016.
- [18] A. Arao and H. Higaki, “Clock synchronization algorithm between wireless sensor nodes without additional control message exchanges,” *WSEAS Trans. Commun.*, vol. 18, pp. 8–16, Jan. 2020.
- [19] Y. Hu and G. Leus, “Robust differential received signal strength-based localization,” *IEEE Trans. Signal Process.*, vol. 65, no. 12, pp. 3261–3276, Jun. 2017.
- [20] S. C. Chan, H. C. Wu, and K. M. Tsui, “Robust recursive eigendecomposition and subspace-based algorithms with application to fault detection in wireless sensor networks,” *IEEE Trans. Instrum. Meas.*, vol. 61, no. 6, pp. 1703–1718, Jun. 2012.
- [21] V. Chatzigiannakis and S. Papavassiliou, “Diagnosing anomalies and identifying faulty nodes in sensor networks,” *IEEE Sensors J.*, vol. 7, no. 5, pp. 637–645, May 2007.
- [22] P. Jia, X. Wang, and X. Shen, “Digital-twin-enabled intelligent distributed clock synchronization in industrial IoT systems,” *IEEE Internet Things J.*, vol. 8, no. 6, pp. 4548–4559, Mar. 2021.
- [23] D. Huang, W. Teng, C. Wang, H. Huang, and J. M. Hellerstein, “Clock skew based node identification in wireless sensor networks,” in *Proc. IEEE GLOBECOM IEEE Glob. Telecommun. Conf.*, 2008, pp. 1877–1881.



Pengyi Jia (Member, IEEE) received the M.Eng. and Ph.D. degrees in electrical and computer engineering from Western University, London, ON, Canada, in 2016 and 2020, respectively.

He is currently a Postdoctoral Associate with Western University, London, ON, Canada. His research interests include distributed clock synchronization, intelligent coordination of Internet of Things systems, digital twin, and networked control systems.



Xianbin Wang (Fellow, IEEE) received the Ph.D. degree in electrical and computer engineering from the National University of Singapore, Singapore, in 2001.

He was with the Communications Research Centre Canada (CRC) as a Research Scientist/Senior Research Scientist from July 2002 to December 2007. He is a Professor and a Tier 1 Canada Research Chair with Western University, Canada. From January 2001 to July 2002, he was a System Designer with STMicroelectronics, Geneva, Switzerland. His current research interests include 5G and beyond, Internet of Things, communications security, machine learning, and intelligent communications. He has over 400 peer-reviewed journal and conference papers, in addition to 30 granted and pending patents and several standard contributions.

Dr. Wang has received many awards and recognitions, including the Canada Research Chair, the CRC Presidents Excellence Award, the Canadian Federal Government Public Service Award, the Ontario Early Researcher Award, and six IEEE Best Paper Awards. He currently serves as an Editor/Associate Editor for IEEE TRANSACTIONS ON COMMUNICATIONS, IEEE TRANSACTIONS ON BROADCASTING, and IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY. He was also an Associate Editor for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS from 2007 to 2011 and IEEE WIRELESS COMMUNICATIONS LETTERS from 2011 to 2016. He was involved in many IEEE conferences, including GLOBECOM, ICC, VTC, PIMRC, WCNC, and CWIT, in different roles, such as the Symposium Chair, the Tutorial Instructor, the Track Chair, the Session Chair, and the TPC Co-Chair. He is currently serving as the Vice Chair of IEEE London Section and the Chair of ComSoc Signal Processing and Computing for Communications Technical Committee. He is a Fellow of the Canadian Academy of Engineering and the Engineering Institute of Canada, and an IEEE Distinguished Lecturer.



Xuemin (Sherman) Shen (Fellow, IEEE) received the Ph.D. degree in electrical engineering from Rutgers University, New Brunswick, NJ, USA, in 1990.

He is a University Professor with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada. His research focuses on network resource management, wireless network security, Internet of Things, 5G and beyond, and vehicular ad hoc and sensor networks.

Dr. Shen received the R.A. Fessenden Award in 2019 from IEEE, Canada, the Award of Merit from the Federation of Chinese Canadian Professionals, Ontario, in 2019, the James Evans Avant Garde Award in 2018 from the IEEE Vehicular Technology Society, the Joseph LoCicero Award in 2015 and Education Award in 2017 from the IEEE Communications Society, and the Technical Recognition Award from Wireless Communications Technical Committee in 2019 and AHSN Technical Committee in 2013. He has also received the Excellent Graduate Supervision Award in 2006 from the University of Waterloo and the Premiers Research Excellence Award in 2003 from the Province of Ontario, Canada. He is the President Elect of the IEEE Communications Society. He was the Vice President for Technical & Educational Activities, the Vice President for Publications, a Member-at-Large on the Board of Governors, the Chair of the Distinguished Lecturer Selection Committee, and a member of IEEE Fellow Selection Committee of the ComSoc. He served as the Technical Program Committee Chair/Co-Chair for IEEE Globecom16, IEEE Infocom14, IEEE VTC10 Fall, and IEEE Globecom07, and the Chair for the IEEE Communications Society Technical Committee on Wireless Communications. He served as the Editor-in-Chief for the IEEE INTERNET OF THINGS JOURNAL, IEEE NETWORK, and *IET Communications*. He is a registered Professional Engineer of Ontario, Canada, a Fellow of the Engineering Institute of Canada, Canadian Academy of Engineering, and Royal Society of Canada, a Foreign Member of the Chinese Academy of Engineering, and a Distinguished Lecturer of the IEEE Vehicular Technology Society and Communications Society.