TrustData: Trustworthy and Secured Data Collection for Event Detection in Industrial Cyber-Physical System

Hai Tao†, Md Zakirul Alam Bhuiyan†, Md. Arafatur Rahman, Tian Wang, Jie Wu, Sinan Q. Salih, Yafeng Li, and Thaier Hayajneh

***Abstract*—Industrial Cyber-physical System (ICPS) is utilized for monitoring critical events such as structural equipment conditions in industrial environments. Such a system can easily be a point of attraction for the cyberattackers, in addition to system faults, severe resource constraints (e.g., bandwidth and energy), and environmental problems. This makes data collection in the ICPS untrustworthy, even the data are altered after the data forwarding. Without validating this before data aggregation, detection of an event through the aggregation in the ICPS can be difficult. This paper introduces `TrustData`, a scheme for high-quality data collection for event detection in the ICPS referred to as "`Trust`worthy and Secured `Data` Collection" scheme. It alleviates authentic data for accumulation at groups of sensor devices in the ICPS. Based on the application requirements, a reduced quantity of data is delivered to an upstream node, say, a cluster head. We consider this data might have sensitive information, which is vulnerable to being altered before/after transmission. The contribution of this paper is threefold. First, we provide the concept of `TrustData` to verify whether or not the acquired data is trustworthy (unaltered) before transmission, and whether or not the transmitted data is secured (data privacy is preserved) before aggregation. Second, we utilize a general measurement model that helps to verify acquired signal untrustworthy before transmitting towards upstream nodes. Finally, we provide an extensive performance analysis through real-world data set and our results prove the effectiveness of the `TrustData`.**

***Index Terms*—Industrial cyber-physical environments, industrial event monitoring, data collection, data trustworthiness, fault tolerance, privacy, security**

I. Introduction

THE Cyber-physical system (CPS) is ubiquitous today. The notion of cyber-physical systems (CPS) is to incorporate physical and engineering systems to monitor their operations with both discrete and dynamic behaviors, which are coordinated and controlled through the integration between the computing and communication core. Due to the potentials of permeative surveillance, CPS-based sensing has appealing real-world applications in numerous fields, e.g., crowd-sensing, social sensing, chemical explosions, mobile event detection, military intrusion tracking, and structural health monitoring (SHM) [1]–[3].

Particularly, industrial CPS (ICPS) is utilized for monitoring critical events such as structural equipment condition in industrial environments. These applications fall into the domain of structural health monitoring (SHM). Among all these aforementioned applications, the quality of the monitoring (QoM) or the quality of the data (QoD) and timely detection of an event are the prime concerning issues. In fact, an automated technique should be able to identify the acquired data faults through online and do the immediate recovery actions for avoiding meaningless monitoring operations and catastrophic situations due to a structural damage or fire.

Particularly in structural damage event detection, reliability is the most desired feature, since an alert of a structural event may play an important role in public safety and economic losses. There are a moderate set of works that suggest different techniques and protocols on reliable event detection and fusion [4]–[7]. A work on event-based data collection and fusion can be found, where heterogeneous sensors exchange data of events among each other [4], [5], [8], [9]. Though there are many work found similar to these, they mainly work data collection techniques, processing, and fusion rather data reliability.

The reliability of detection in the ICPS fully depends on the data trustworthiness in terms of QoD and QoM. However, the ICPS can easily be a point of attraction for the cyberattackers, in addition to system faults, severe resource constraints (e.g., bandwidth and energy), and environmental problems. This makes data collection in the ICPS untrustworthy. Even the data are compromised after the collected data forwarding. Without validating this before aggregation, detection of an event in the ICPS can be difficult in addition to system faults, severe resource constraints (e.g., bandwidth and energy), and environmental problems. This makes data collection in the ICPS is *untrustworthy*.

Even, after the collected data forwarding and/or before data aggregation, the collected data can be compromised, making detection of an event of interest (such as structural health event, mobile object) in the ICPS unreliable. Usually, the data

†Co-first Author

H. Tao and Y. Li is with the School of Computer Science, Baoji University of Art and Science, Shaanxi 721013, China.

M. Z. A. Bhuiyan and T. Hayajneh are with the Department of Computer and Information Sciences, Fordham University, New York, NY 10458, USA.

M. Rahman is with Faculty of Computing, and IBM Center of Excellence and Earth Resources and Sustainability Center, Universiti Malaysia Pahang, Kuantan 26300, Malaysia.

T. Wang is with College of Computer Science and Technology, Huaqiao University, Xiamen, China, 361021. (Corresponding author: wangtian@hqu.edu.cn

J. Wu is with the Department of Computer and Information Sciences, Temple University, Philadelphia, PA PA 19122, USA.

S. Salih is with the Institute of Research and Development, Duy Tan University, Da Nang 550000, Vietnam.

Manuscript received mm dd, yyyy; revised mm dd, yyyy.

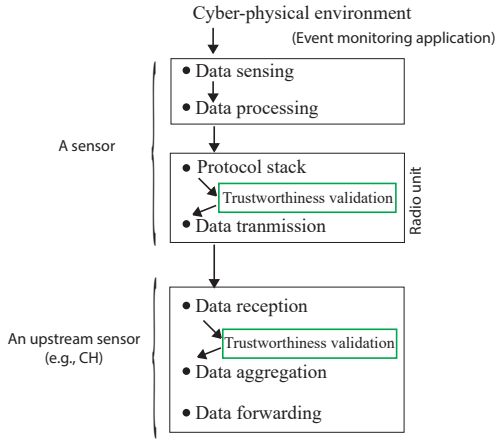


Fig. 1. Data trustworthiness validation in two levels in the ICPS: before data transmission and before data aggregation.

which is transmitted by sensors is not reliable because of several reasons, e.g., faults of sensor devices, limitation of sensor calibration, lack of observation, data modification caused by security attacks and background noise. If the aforesaid reasons are not being counteracted, the data cannot be considered as trustworthy while it is forwarded for aggregation to the upstream sensors, referred to as clusters. The QoM is affected severely because of such untrustworthy data.

Moreover, the energy of ICPS-based detecting may be unloosed due to the appropriate aggregation of untrustworthy information that acquired by different sensors and the submitted data maybe compromised (modified) before the data transmission. Whether the transmitted data is *trustworthy* or not, they can be further modified at some stage of transmission from sensor devices to cluster heads by the third party, making detection of an event of interest (such as structural health event, mobile object) in the ICPS unreliable. Some sensor devices uninterruptedly deliver reliable data whereas others may generate unreliable data, because of security attacks [10]–[12]. In fact, the cluster head should play a role of assessing the reliability of the data which is collected from the individual node. Thus, the collected data should be *secured* before aggregation. As a consequence, it is essential to characterize the reliability of received data before performing aggregation at a Cluster Head (CH). Overall, without addressing them, data collection tasks in the ICPS are *untrustworthy*.

In order to address all the above issues, this paper introduces ‘`TrustData`, a scheme for high-quality data collection for event detection in the ICPS referred to as “`Trust`worthy and Secured `Data` Collection” scheme, as shown in Fig. 1. Based on the application requirements, a reduced quantity of data is delivered to an upstream node, say, cluster head (CH) (see Fig. 1). We consider this data might have event-sensitive information, which is vulnerable to being altered before/after transmission. The main concept of the `TrustData` is that a CH is able to verify two features before aggregating data in a cluster such as data trustworthiness and protection. The former one ensures that the data remains unchanged, which is estimated at the sensor device level. The later ensures that the data cannot be modified after transmission, which is done by CH. To ensure the trustworthiness of data, a measurement model extracted from Mutual

Information Independence (MII) has been utilized between two signals either from two different sensor nodes or the same one for evaluating results without considering the ground truth. The mutual statistical information may be utilized as an index to verify the trustworthiness of received data for structural damage event detection. The data is considered as trustworthy after successfully passing this check and subsequently, it will forward to the CH. To ensure the data protection, a Truth Status Value Finding approach has been introduced by having a goal to derive truthful facts from unreliable sensors. We have carried out extensive simulation for measuring the performance of `TrustData`. In the simulation, realistic data set is utilized, and it proves that the received data in `TrustData` is trustworthy and secured, which can make aggregation for event detection in particular applications trustworthy.

In summary, we make the following major contributions in this paper:

- We propose `TrustData`, a trustworthy and secured data collection scheme for the event detection in the ICPS. This can verify whether or not the acquired data is trustworthy (unaltered) and whether or not the transmitted data is secured (data privacy is preserved).
- We propose to utilize a general measurement model called MII that helps to verify whether the acquired signals are untrustworthy before transmitting towards upstream nodes.
- We present a signal validation algorithm using Truth Status Value Finding to validate signal trustworthiness before signal aggregation.
- We have carried an extensive performance analysis of `TrustData` exploiting real-world data set and our results prove the effectiveness of the `TrustData`.

The rest of the paper is summarized as follows. Section II shortly describes our scheme. In Section III, we illustrate the trustworthy data collection approach while Section IV gives truth status Finding approach. Section V presents the performance analysis of the proposed `TrustData` scheme. Finally, Section VI concludes the paper and recommends future direction of the work.

II. Related Work

Nowadays the researchers are focusing towards the data security and privacy research domains as they are attracted by the businesses, governments, individuals, and industrial networked environments like ICPS [9], [13], [14]. However, these existing protocols can be attacked by the attackers while they are handling data, what needs to be done then? This paper is a preliminary attempt to work on the data trustworthiness [9], [15].

ICPS society, healthcare providers and application users, have been working for offering protection to each device of the network [16], [17]. Having the data integrity feature, the device of the IoT/CPS network can be compromised during the data forwarding among the node. Recently, data

trustworthiness are taking a point of attraction from research, particularly, in patient healthcare. Some cloud based scheme for healthcare e-medical system is [8], [18], [19]. Patra et al. [8] suggested a cloud-based model that maintains patients privacy data. This model guarantee cost effectiveness and it is deployed for remote areas where implementation cost needs to be taken account. The patients can be served by the healthcare providers and professionals remotely exploiting cloud-based model. In [14], the authors propose a scheme, where data has been processed during collecting and delivering data. Zhang et al. [15] introduce cloud-enabled patient system, where three layers have been included, such as data-collection, data-management and data-service layer. In [20], the authors introduce blockchain technology with access control manager for health data in order to improve the interoperability of this system. In [9], the authors present a trustworthy data collections scheme for the cloud-enabled sensor systems. They consider three types of trust, which are described to assess the sensor and sink devices behaviors. However, it does not describe practically, how the data trustworthiness can be maintained either at the time of collection or transmission.

Traditionally, the voting scheme has been utilized to minimize conflicts for making decision analyzing the received data, i.e., making a trustworthy decision. This method is used for conducting majority-voting and the information related to the highest number of occurrences is regarded as the right answer. It is an assumption for designing a voting system that all the end device are reliable equally, consequently, the vote for various end devices weighted uniformly [21]. If one or two devices of the CPS get cyberattacks, it is tough to identify the device as the all avg or mean are used for all the devices. The rustworthiness of the estimation considering on the average or a maximum number of packets, weights, or votes cannot ponder the real facts in the CPS. There are various works on trust models that are used to determine to solve communication security problems [22].

Most of the work similar to the above considers trusted computing and trust communication, and so on. However, data trustworthiness at the time of collection, and before/after transmissions are not considered, a preliminary step of them are considered in `TrustData`.

III. TrustData: The Trustworthy and Secured Data Collection Scheme

The overview of the proposed trustworthy and secured data collection (`TrustData`) scheme will be exhibited in this section.

We consider a hierarchical ICPS having a set of sensor devices, which are deployed for a specific monitoring application. We focus on monitoring the health of civil structures (such as aircraft, building, bridge) as a representative application. A representative two dimensional (2D) model of a building is illustrated in Fig. 2a. In the figure, sensor devices (white circle) are employed in accordance with civil engineering-driven placement techniques [23]. Moreover, a distant application monitoring facility or a base station (BS) (colored circle) is at a distant location. The employed sensor devices are omnidirectional and self-organized into clusters utilizing some application-specific clustering algorithm [24]. Every CH collects the data, verifies, and then sends the ultimate decision or aggregated data towards the BS.

We consider that sensor devices will be assigned to do various classes of applications in the context related with SHM, such as temperature, pressure, damping, strain and sensing the vibration. Subsequently, the sensor devices will send their acquired information from the different application towards the intermediate nodes or a CH. For the convenience's sake, we take into account the vibration signals in this paper.

We assume a big number of sensor devices are employed to acquire data from the health monitoring area. To maintain the data collection security, each sensor device is given an ID and a paired cryptography key for encryption using some handshake procedure among nodes. In `TrustData`, sensor devices do not have any preshared secrets and they construct trust by producing numerous shared keys out of a channel. Particularly, every sensor device performs authentication with other devices. Furthermore, every sensor keeps a set of neighboring nodes information such as IDs, trust values and so on. The sensor devises with its CH maintain a relationship: if device A needs to transmit data to another device B, device A should have some information about the neighborhood and know some more information such as distance among them, trust value of device B and other. When node B is out of the communication range of A, the data are exchanged between them through multi-hop communication exploiting intermediate nodes. Every intermediate device should have its own decision for forwarding its information and look for the next hop to transmit the data from device A to device B.

Health event detection in the ICPS can easily be a point of attraction for the cyberattackers, in addition to system faults, and severe resource constraints (e.g., bandwidth and energy), and environmental problems. This makes data collection for event detection in the ICPS *untrustworthy*. The trustworthiness of health event detection depends on this trustworthy data collection. However, the data trustworthy is critically influenced by cyber threats/attacks in the ICPS. Ransomware, DDoS attacks, node replication, worm attacks, collusion attack, and so on are popular attacks in the ICPS [12]. Principally, these attacks convey new confronts, e.g., data can be compromised at the time of data collection or the data communication [25]. Even after the collected data forwarding or before aggregation, the data can be compromised, making detection of an event of interest (such as structural health event, mobile object) in the ICPS unreliable.

There are two types of attacks that are considered in this paper: collusion attacks and trust-spoof attacks. In a trust-spoof attack, suspicious sensor devices deliberately push deceitful message to neighboring devices using some trust-value. For example, they normally offer lower-trust value for messages for transmission. Thus, the message under trust-spoof attacks cannot influence the real trustworthy data delivery route. Likewise, the trust-spoof attack may also deliberately offer higher-trust values for messages for evil sensor devices and help them obtain more opportunity to harm the ICPS. Another type of attacks in the ICPS we assume is the collusion attack. It

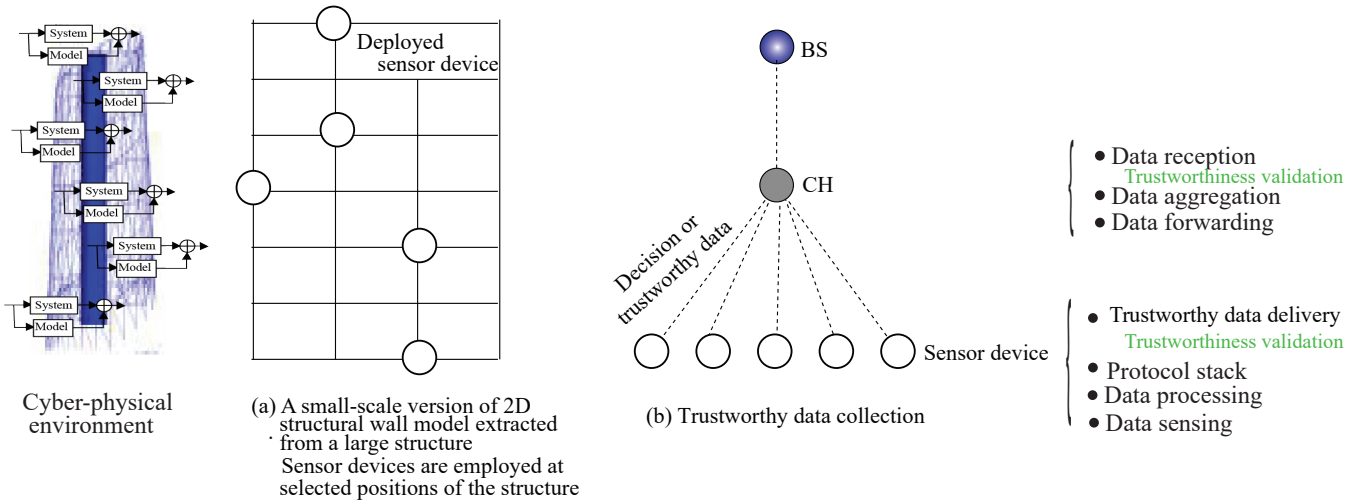


Fig. 2. The ICPS scheme for trustworthy data collection.

is a security attack where a device intentionally forms a hidden agreement with an evil device. The evil device might try to inject false or compromised data via one/more compromised nodes.

As shown in Fig. 2b, we propose trustworthiness validation in two-level: before data transmission and before data aggregation. We assume that, given the algorithms of this work, sensor devices A, B, and all collect data exploiting a state-space model [23]. The collected data has been examined locally in order to identify untrustworthy signals, and finally forward them. For better identifying trustworthy data, we also focus on a set of sensor faults by taking account of real-world wireless SHM system, such as sensor debonding fault (a wireless sensor slightly or completely debonds/detaches from the host structure), untrustworthy or faulty signals by precision degradation, etc., especially in vibration signal capturing, faults in offset, bias, and the amplification gain factor of signals. Sensor devices may produce ambiguous signals due to the security attacks.

Mutual Information Independence (MII) is being used in statistics as an indirect signal quality measurement. It helps to determine how much mutual information there is between two random variables. The hypothesis behind the information independence appears when mutual information is zero.

To achieve that, we consider a correlation model denoted by C, as discussed in [26]. This model is used to get a reference dataset. During the initialization of the ICPS, this dataset is automatically stored in the sensor local memory. To check the signal abnormality, a *MII* function of two signals of arbitrary sensor devices i and j at time t in a cluster is analyzed.

Basically, the security threats in a network (in a practical ICPS scenario) are coming from the sensor devices that are deployed in the network. The CH of a cluster might try to deduct the observation of each device. On the contrary, the sensor device might also try to derive the information of other intermediate nodes. Thus, it is of paramount importance to preserve sensor observation values (without alteration). We do not assume any modification of the received data at the CH regarding high-quality event detection. To deal with this, we propose to utilize a sensor status value truth finding technique, where providing true information of the sensors will be considered having truthful facts more often and the information that is supported by reliable sensor devices will be regarded as true facts.

IV. Acquired Signal Trustworthiness Validation before Transmission

In this section, we describe the trustworthy data collection.

To make reliable event detection in an application of the ICPS, we need trustworthy data collection in the application. In this case, we take SHM applications [23]. Data collection in SHM applications is usually used for structural health event detection within the structures, namely, damage. SHM schemes and their associated algorithms usually dominated by the civil or structural engineering domain execute to measure structural physical responses, which are due to ambient signals of vibration, strain, or forced excitation. A variety of ICPS sensor devices, including accelerometers, strain gauges, or displacements are integrated to acquire physical structural vibrations. There are numerous methods for data acquisition used by civil or structural engineering domains. We apply the *state space model*, which is commonly used by the engineering communities for signal acquisition. This can precisely acquire the structural physical dynamics [23].

Sensor devices in the ICPS are prone to generating faulty or untrustworthy signals collected form the physical structures. The structural signals of a sensor device are acquired by the vibration, which may be inaccurate compared to the signals of the neighboring devices, earlier signals, or reference sets of signals. First, we present data acquisition at each sensor device in the ICPS. A subset of sensor devices in a clustered neighborhood, denoted by D that is in a sensor device's minimum communication range, share the signals with each

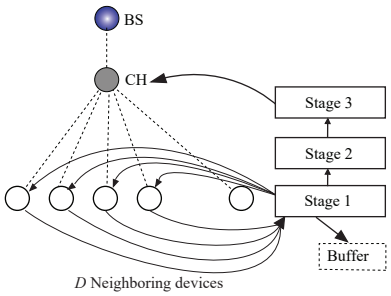


Fig. 3. Three stages of the data collection algorithm in `TrustData`.

other, and contribute to the detection of untrustworthy or faulty sensor device's signals.

We develop a data collection algorithm that simply shows the data acquisition techniques in a clustering neighborhood D. We assume that the clustering neighborhood D has the highest node degree, at most a constant. Although this process is comprised of multi-hop transmission theoretically, regarding the fact that for SHM application, the wireless radio transmission range of a sensor device surpasses in D where the sensor devices acquire physical structural signals. We maintain a limit on the number of sensor devices to transmit signals within one-hop neighboring sensor devices in D.

The trustworthy data collection algorithm has 3 stages, as shown in Fig. 3. These are described in the following:

Stage 1. Each sensor device collects signals acquired from the structural vibration responses (such as aircraft, nuclear plants), and puts them into buffers temporarily. It then transfers the signals, while it also receives acquired signals of the neighboring devices in D.

Stage 2. The sensor device verifies the received signals to identity whether there exist any faulty or untrustworthy signals.

Stage 3. In this stage, each sensor device carries out runs and computes another algorithm, which is called "decision-making on the acquired untrustworthy signals to identify whether acquired sensor device signals are untrustworthy or compromised.

When an outstanding change spot or signal modification happens in a sensor device's signal, there is a likelihood that a sensor devices signal is untrustworthy or altered, which can be found through Stage 2. We use the mentioned MII to identify the trustworthiness of the sensor device signals. We compute the numerical dependency amongst the acquired signals of any two sensor devices in D enumerated by the MII. Signal information of a sensor device is determined, which is denoted by ω and is distributed by another sensor in a set of acquired signals in the clustered neighborhood D. It is found that ω fluctuates once signal fault happens in a sensor device signal. Because this fluctuation is not found in the neighboring sensor device and there is no such fluctuation in the reference set of signals.

A. Identifying Faulty/Untrustworthy Signals

In this subsection, we analyze signals to identify whether or not the signals are faulty/trustworthy.

In order to identify in the collected signals, we apply a *correlation model* based on the multivariate Gaussian distribution. A correlation model is that it indicates the mutual relationship between two or more signals. This is to say, a correlation between two or more signals specify the quantity up to which the specified signals seem like another signals. Multivariate Gaussian distribution is widely applied to precisely model the correlation of different types of sensor devices signals in literature [27]. Each signal is broadcast among the sensor devices in the clustered neighborhood D. Let the ith sensor device signal be $y_i^t \in y_D^t$ and jth sensor device signal $y_j^t \in y_D^t$, $i, j \in D$. In order to simplify it, y_i^t as m and y_j^t as n are denoted hereafter.

Multivariate Gaussian distribution is widely applied to precisely model the correlation of different types of sensor devices signals [27]. Every collected signal is broadcast among the sensor devices in the clustered neighborhood D. Let the ith sensor device signal be $y_i^t \in y_D^t$ and jth sensor device signal $y_j^t \in y_D^t$, $i, j \in D$. In order to simplify it, y_i^t as m and y_j^t as n are denoted hereafter.

Therefore, it can be significant to consider how to obtain joint probability density between any two of the signals m and n. Signal Correlation is helpful for the reason that the signals can specify a predictive relationship, which is delayed in practice. The mutual information is referred to as statistical independence. The statistical dependency/independency between the two Gaussian distributed time signals m and n can be expressed in the form of the joint probability density $p(m, n)$ of signals. which is given as follows: The correlation coefficient between two signals m and n is denoted by ρ_{mn}. The correlation coefficient can also sometimes be used to decide if there is a statistical independency between two signals like m and n. In one hand, if $|\rho_{mn}| = 1$, there is a strong correlation between the two signals. On the other hand, if $|\rho_{mn}| = 0$, there is no correlation between the two signals. We need to note that a weak form of statistical dependency can appear in the correlation interpretation. According to [28], the statistical independence can be when there are two random variables, which are not shown correlated. This is one of the reason we assume statistical dependency or independency in this work. Let ρ_m and ρ_n be the product of the marginal densities of two signals m and n, respectively, which can be stated as follows:

$$p(m, n) = p(m)p(n) \tag{1}$$

The two signals are totally independent, if the expression in (1) is equivalent to the product of the marginal densities in (3). In order to compute the MII between the signals, there can be a possibility to calculate the statistical dependency between any of the two signals, which is as follows:

$$\omega(m, n, C) = \int \int p(m, n) \log \frac{p(m, n)}{p(m)p(n)} du\, dv \tag{2}$$

The basis of the logarithm determines the units by which information about statistical independency is determined. ω

can be zero, if m and n are independent according to (2). Finite bins and the number of sampled pair count can be achieved by a forward approach, which is to divide the range of m and n, i.r., $h_o = (u_o, v_o), o = 1, 2, \cdots, n$, falling into these finite bins. This count permits us to roughly decide on the probabilities. This leads to replacing (3) by the finite sum:

$$\omega_{bin}(m, n, C) = \sum_{a,b} p_{mn}(a, b) \log \frac{p_{m,n}(a, b)}{p_u(a)p_v(b)} \quad (3)$$

where $p_u(a) \approx n_u(a)/n$ and $p_u(b) \approx n_u(b)/n$ are the likelihood grounded on the number of points $n_u(a)$ and $n_v(b)$. This falls into the ath bin of m and the bth bin of n, respectively. We can have a joint probability based on the number of points $n(a, b)$ falling into box mumblers a and b, which is $p_{mn}(a, b) \approx n(a, b)/n$. Here, MII value is symmetric and non-negative and given as follows :

$$\omega(m, n, C) = \omega(n, m, C) \geq 0 \quad (4)$$

y_r and y_s are the MII values for all probable permutations of sensor device outputs (except $r = s$, where $i = 1, 2, \cdots, r, j = 1, 2, \cdots, s$). This leads to an ω-matrix for all permutations of r and s. The main idea is that when the presence of a signal compromise or fault f_r is there, the MII changes. Assume that it is in the rth index or channel:

$$\tilde{y}_r = y_r + f_r \quad (5)$$

This fault or untrustworthy situation happens only in the rth index. Therefore, we can suppose that all permutations with index r can demonstrate a decrease in ω. This enables us to pinpoint faulty or untrustworthy signals. One or more sensor devices' faulty or untrustworthy signals can be concurrently identified in the same way. Using the relative variation as a signal fault or untrustworthy indicator is to make a likelihood to picture the faulty or untrustworthy signals $\lambda^{\omega}_{y_r}$:

$$\lambda^{\omega}_{y_r} = \frac{|\omega_{y_r} - \omega_{ref}|}{\omega_{y_r}} \quad (6)$$

Here, the actual dataset is denoted by y_r and the reference data set is denoted by the lower index *ref*. Therefore, detecting sensor faults or untrustworthiness situations in different combinations of them is done by the method based on the MII.

B. Validation Decision

The untrustworthy or faulty signal detection can be executed in a distributed manner when each of the sensor devices can decide collected signal trustworthiness or faulty locally. This is because the distributed technique only needs the neighboring sensor devices' signals in a cluster to be synchronized. Moreover, the decision is nearly fast and online, since a sensor device does not need to delay for the signals from neighboring sensor devices at one or two hops away.

Through this algorithm, whether the collected signals are untrustworthy or faulty can be known from the local decision on the sensor device's collected signals, $\lambda^{\omega}_{y_r} > 0.5$. This implies that for the faulty or untrustworthy collected signals, the MII is high. It is worth noting that the MII is not dependent on a particular type of trustworthy or faulty situation. The algorithm relying on the MII is able to identify diverse types of signal faults (as discussed in Section III). By means of the algorithm, a sensor device can know whether or not its collected signals are trustworthy or fault, and forward the signals to the CH.

V. Signal Untrustworthiness Validation Before the Aggregation

In the previous section, the non-faulty or trustworthy signals are collected for event monitoring. Once a device transfers the trustworthy data, it may be compromised at the device level or intermediate sensor device before/after the device makes the data transmission. This is to say, a CH is most likely to receive compromised (or altered) data for the aggregation if the data is sent unprotected. We should guarantee that the acquired data is not compromised during the data exchange or transmission.

To identify an untrustworthy sensor device or its unprotected data at the time of data reception at the CH, we apply the Truth Status Value Finding in `TrustData`. Traditionally, truth value detection is applied in numerous fields for solving conflicts in compound noisy signals. The perception of the truth value identification algorithm is that it takes the first step with a random supposition of ground true facts, and iteratively performs the sensor device status value updates and truth updates until convergence [29]. As a result, algorithms can infer the sensor device trustworthiness or data trustworthiness of sensor device and the truth of data from each other's devices.

In our algorithm, we calculate sensor device status value to make a decision on the transmitted data, that is, trustworthy or not. In structural health event detection, the sensor devices need real values or facts for high-quality event detection. Note that usually the ground fact or truth of each structural health event is fixed. The main concept is that a high value is provided as a sensor device's status value, if the transmitted trustworthy data comes near to the assessed ground true facts. Typically, the sensor devices status values are calculated in the following:

$$S_k = \log\left(\frac{\sum_{k'=1}^{K}\sum_{m=1}^{M} d(x_m^{k'}, x_m^*)}{\sum_{m=1}^{M} d(x_m^k, x_m^*)}\right) \quad (7)$$

where $x_m^{k'}$ are the observation values and x_m^* are the estimated ground true facts. There is s distance function denoted by $d(.)$ used to determine the difference between sensor devices $x_m^{k'}$ and x_m^* [30].

The value of $d(.)$ is due to a specific sensing application situation. The presented scheme `TrustData` is envisioned to handle structural health event identification of SHM applications. In the case of SHM applications, where the sensor devices signals are continuous (e.g., structural vibration response), we adopt a standardized squared distance function, which is in the following:

$$d(x_m^k, x_m^*) = \frac{(x_m^k - x_m^*)^2}{std_m} \quad (8)$$

Here, to measure the amount of variation of the set of observation values of signals, std_m is used as the standard deviation of the observation values.

A. Truth Status Update.

Presume that each of the sensor device status values is static, which is not changed as long as some change happens. Then, the ground true facts for the m-th events can be estimated. Here, m is used as one or more than one change points can exist when there is a change in the structural health event. For example, when there is a health event, multiple sorts of sensing signals' information, including accelerometer, stain, displacement can be collected.

$$x_m^* \leftarrow \frac{\sum_{k=1}^{K} S_k x_m^k}{\sum_{k=1}^{K} S_k} \tag{9}$$

Since acceleration signals and strain signals are continuous steaming signals, x_m^* is used to get the estimated ground truth facts. The truth finding process is described by the Algorithm. The process of the algorithm commences with arbitrarily guessing the ground truth facts for a particular event, then iteratively updates sensor devices status values and approximate ground true facts until a convergence criterion is reached. Generally, the convergence criterion is taken with respect to the particular application requirements. It can be a threshold of the change points in the approximated ground truth facts in two successive iterations.

B. Truth Status Value Finding

In this subsection, we provide the details of the proposed truth status value finding through Algorithm 2. To achieve that, a security protocol has been used. A semantically secure (p, t)-threshold Paillier cryptography is undertaken, which is adopted from [31], [32]. The Paillier cryptosystem is an additive homomorphic cryptosystem, meaning that it can be used for evaluating some statistical information, such as the mean and the variance [32], which exactly matches our truth discovery in CH. The number of sensor devices in the D including both CH and sensor devices is denoted by p, and the smallest number of (cluster and sensor) devices required to calculate the decryption process is denoted by t. Therefore, the encryption key $p_k = (g, n)$ which is public is known at each of the sensor devices in `TrustData`, while the decryption key, which is private, is broken down and distributed to sensor devices in D (i.e., device i can have its private key share sk_i). At the level of data transmission, each of the sensor devices iteratively carries out two procedures, which are given as follows:

- *Status Value Update.* Every device calculates the distances between its trustworthy data (i.e., observation values) and the approximate ground true facts calculated by the CH with respect to the distance function $d(.)$. The sensor device then performs encryption of the distance information and sends the generated ciphertexts to the CH. After all the ciphertexts of all of the sensor devices are reached at the CH, the CH updates the status value with security, i.e., making in encrypted form for every sensor device. Finally, the updated status value of the ciphertext is transferred back to every corresponding sensor device.
- *Secure Truth Approximation.* Depending on the encrypted status value transferred by the CH, every sensor device calculates status value observation in the form of ciphertexts without the status value decryption. The sensor device then transfers them to the CH. As soon as the CH received all of the status value in the form of ciphertexts from every sensor device, the CH calculates the ground true facts for ensuring the trustworthy data.

Algorithm 2: Truth Status Value Finding

Input: Observation values of K sensor devices: $\{x_m^k\}_{m,k=1}^{M,K}$
Output: Generation of ground true facts for M signals: $\{x_m^*\}_{m=1}^{M}$
1. A CH arbitrarily sets the ground truth for each signal;
2. The CH transmits the calculated ground true facts ($\{x_m^*\}_{m=1}^{M}$);
3. Each device calculates $d(.)$ and and has the rounded values
   then encrypts them and transmits to the CH;
4. (upon the reception of ciphertexts of all the devices)
   i. The CH get the approximate x_m^* according to the *secure sum*;
   ii. Get updates of the encrypted weight of devices;
   iii. The updated encrypted weight is transmitted to every device;
5. When each device gets back weights in encrypted form from the CH,
   i. Do *secure sum* by the devices' weighted values;
   ii. Transmit these in form of ciphertexts towards the CH;
6. (after getting ciphertexts from the devices in the cluster)
   The CH computes the ground true facts;
7. Repeat steps 2 to 6 until a convergence criterion is met
8. Then get output ($\{x_m^*\}_{m=1}^{M}$);

In spite of this, every sensor devise data in form of plaintext must not be available to other party or sensor device regarding confidentially issues. We handle this issue through the secure sum protocol that is obtained by the threshold Paillier cryptosystem [33]. As shown in the Algorithm, such a secure sum protocol is able to help to compute the sum of sensor devices data in the form cyphertexts without revealing secret information to outsiders.

The two procedures above commence with an arbitrary initialization of the ground true fact for the trustworthy data. The two procedures are then iteratively carried out until a convergence is met. During the course of performing, the procedures are carried out fully on encrypted data. Thus, it can be confirmed that the observations of the status value of every device are recognizable simply to itself and any other device's do not know the sensor device correct status values.

In support of secure truth approximation and updated status value, we apply a secure sum protocol developed in [34] by which sensor devices are permitted to calculate the sum of the collected data while preserving their own data secret with added computation complexity to any suspicious party or the sensor devices of the system. In accordance with (9) and (11), a CH computes the sum of the data gathered from sensor devices so as to have updated status values and approximated ground true facts. In spite of this, every sensor device' data in form of plaintext must not be available to other party or sensor device regarding confidential issues. We handle these issues through the secure sum protocol that is obtained by the threshold Paillier cryptosystem [33]. As shown in the Algorithm 2, such a secure sum protocol is able to help

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2019.2950192, IEEE Transactions on Industrial Informatics

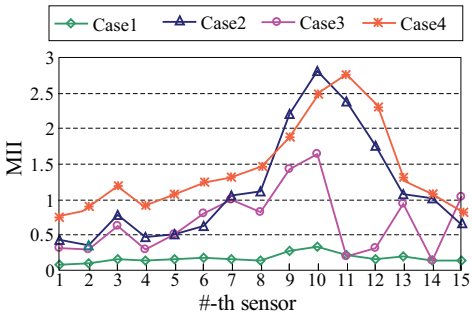


Fig. 4. Performance of TrustData in trustworthy data collection: achieved MII under sensor device signal faults.

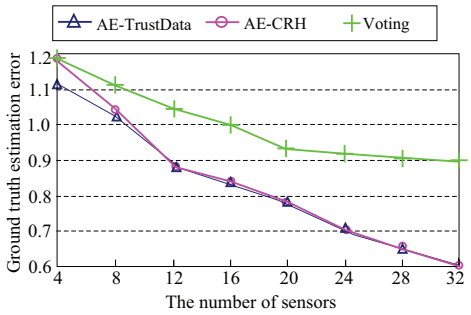


Fig. 5. Performance of data alternation and detection observing ground truth estimation.

compute the sum of sensor devices' data in the form of cyphertexts without revealing secret information to outsiders.

VI. Performance Evaluation

A. Simulation Methods

In order to measure the performance and prove the effectiveness of the proposed scheme (i.e., TrustData), we have done comprehensive simulations in MATLAB by considering real data sets that are collected by the SHM system applied on the Guangzhou National TV Tower [35], [36] and a SHM system toolsuite [37]. The considered dataset collects data from 800 sensor nodes and 100 sensor cases are utilized in the simulations. The ICPS deployment has been performed using our ICPS-based deployment scheme as presented in [38] The network size of the simulation environment is a 450×50 sensing field for the structural environment, e.g., aircraft, building.

For convenience, we take into account the vibration signals in this simulation, which are influenced by the 100 sensor locations in the field. Moreover, we add random Gaussian noise in all the data, with zero mean and 10% standard deviation of real signals. There are two parts in the data sets: i) reference data for training the joint distribution; and ii) another set is used for testing, where noise is added in both cases so that it is reflected in the trained correlation model. The sensor node updates its MII once a decision has been received.

We have chosen an existing truth finding technique in the simulation as a baseline technique for status-value-truth-finding scheme [30], which does not counteract the sensor security during the whole procedure. The status value truth finding is implemented by considering the Paillier Threshold Encryption Toolbox, where we have set a $(p, \lfloor \frac{p}{2} \rfloor)$-threshold Paillier cryptosystem (http://cs.utdallas.edu/dspl/cgi-bin/pailliertoolbox/).

Again, a Voting scheme has been considered for performing comparison. It is used to reduce conflicts during decision-making based on received data. This method is used for conducting majority-voting and the information related with the highest number of occurrences is regarded as the right answer. Usually, an assumption for a voting system is that all the end devices are reliable equally and, consequently, the vote for various end devices is weighted. We considered an existing network-based-voting algorithm, as presented in [21].

B. Performance measures

In order to measure the performance of the TrustData, we consider two metrics i.e., Error Rate (ER) and mean of absolute error (AE) and the continuous data. The first metric i.e., ER is the percentage of the approach's output that are different from the trustworthy data and ground true facts. The second one. i.e. AE, is to calculate the mean of absolute distance between the ground true facts and estimated results. Each simulation is running 50 times.

C. Results

We have conducted three sets of simulations. First, we executed the proposed TrustData scheme for trustworthy data acquisition and secured data transmission. We inject compromised (untrustworthy) signal information into the sensor devices. Here, we regard compromised signals as the faulty signals. This is done via arbitrarily modifying some sensor devices' signals in the data sets. We have arbitrarily chosen a portion of the sensor devices from the network of the ICPS and the faulty or compromised signals are injected into their acquisition modules and make them as untrustworthy or faulty sensor devices. In the simulations, the amount of the faulty or untrustworthy signals varies from 15% to 25%. Each device broadcasts its signals to the sensor devices in the cluster neighborhood. Every faulty or untrustworthy signal is substituted by an arbitrary amount independently obtained through a uniform distribution in the sensor device deployment environment (0, 450). The reason for selecting this type of fault model is that it results in a set of uncorrelated data in the equal magnitude according to the gathered signals in practice.

The MII obtained in the first five successful simulation cases, with diverse sensor untrustworthy or fault injection are shown in Fig. 4. There is no any untrustworthy/fault signal injection in Case 1. This implies that the collected data is not distorted almost in all of the sensor devices by any signal faults or security compromised. In Case 2. it is observed that the high MII value at some sensor devices includes sensor 9 and sensor 10. Their signals are untrustworthy or faulty or partly altered, which is apparently detected. When the rate of untrustworthy or faulty signal injection increases, i.e., in case 4, we can see that the values of the MII at sensor devices in the neighborhood becomes the maximum. In Case 3, changes are seen in only in one or two points (such as sensor 10). This is also due to the

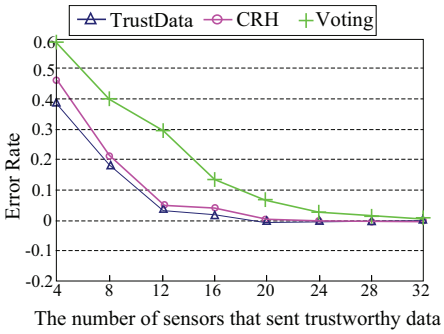


Fig. 6. Performance on the error rate under both trustworthy data and altered data situations in all the schemes.

influence of untrustworthy signal. As a result, data collected from these sensor devices cannot be trustworthy. This validates the precision of the untrustworthy signal detection. Every time a CH gets such data, the CH may drop the collected data before aggregation or a signal reestablishment process might be used for the portion of untrustworthy data.

In the second set of simulations, we make a comparison between `TrustData` and CRH in terms of the correctness of ground true facts. The approximation errors of `TrustData` are presented by arbitrarily predicting the ground true facts, since we arbitrarily introduce the approximate ground true facts, and apply a threshold of the change point in the approximate ground true facts in two successive iterations until the convergence criterion. The ground truth approximation errors of `TrustData`, CRH, and Voting under different random values are shown in Fig. 5. AE is used to measure the error. It can be seen in the figure that the approximation errors in `TrustData` are almost the same as the CRH while the amount of sensor devices is changing.

Even `TrustData` demonstrates better performance than the performance of CRH in some cases; specifically, when the amount of sensor devices in the ICPS is not large. In addition, it is seen that, with the increase of the amount of the sensor devices, the approximation errors decrease. Another scheme, Voting, depicts the poor performance when compared to both `TrustData` and CRH. One of the probable causes is that trustworthiness approximation based on the highest amount of data packets or votes may not imitate the true facts in the ICPS.

Nonetheless, the widely applied method to remove the conflicts in a decision on a faulty or untrustworthy sensor is to carry out majority voting so that the accurate response come through information fusion from the average or maximum amount of occurrences. From our results, it is evident that the concern with such Averaging or Voting schemes is that all of them presume all the data packets from all the sensor devices are equivalently trustworthy, and thus the votes from diverse sensor devices are equally weighted. However, the alteration of packets or votes is not included in these schemes.

In the last set of simulations, we attempt to study the error rates in different schemes. As shown in Fig. 6, we demonstrate the performance of the schemes in terms of error rates on the SHM data set. It can be seen that `TrustData` achieves a lower error rate on the SHM data set compared to that of the CRH and Voting. We can see that the CRH takes all of acquired data where some of the signals are untrustworthy or faulty. Untrustworthy data is discovered in both the CRH and Voting. In the Voting scheme, sensor device signals are considered equally trustworthy and/or the sensor device is trustworthy which is with the highest votes. Sensor devices in `TrustData` verify signal trustworthiness and include the trustworthy signals. When a fraction of the trustworthy signals again are compromised during the signal transfer, this signal is not included in the aggregation. Therefore, the error rates in `TrustData` become lower than the error rates in the CRH.

VII. Conclusion

This paper has introduced a scheme for event detection in the ICPS referred to as Trustworthy and Secured Data Collection Scheme (`TrustData`). The scheme ensures authentic data collection for aggregation inside a cluster of the ICPS. In order to achieve the trustworthiness of data in the ICPS, an algorithm has been proposed, which ensures the integrity of transmitted data. In order to have secured data in the ICPS, we have improved the Truth Status Value Finding to derive trusted facts from untrusted sensor data. Finally, we have carried out an extensive performance analysis of `TrustData` exploiting real-world data set and the results exhibit that the collected data exploiting the proposed scheme is trustworthy and secured, which may ensure a reliable decision-making in event detection in the ICPS. The future direction of this work is to provide security and authenticity of data transmission by exploiting the sensor status value finding approach.

References

[1] M. Z. A. Bhuiyan, J. Wu, G. Wang, , and J. Cao, "Sensing and decision-making in cyber-physical systems: The case of structural health monitoring," *IEEE Transactions on Industrial Informatics*, vol. 12, no. 6, pp. 2103–2114, 2016.

[2] Y. Wang, "Trust quantification for networked cyber-physical systems," *IEEE Internet of Things Journal*, vol. 5, no. 3, pp. 2055–2070, 2018.

[3] L.-A. Tang, X. Yu, S. Kim, Q. Gu, J. Han, A. Leung, and T. L. Porta, "Trustworthiness analysis of sensor data in cyber-physical systems," *Journal of Computer and System Sciences*, vol. 79, no. 3, pp. 383 – 401, 2013.

[4] A. P. Plageras, K. E. Psannis, C. Stergiou, H. Wang, and B. Gupta.

[5] C. Esposito, A. Castiglione, F. Palmieri, M. Ficco, C. Dobre, G. V. Iordache, and F. Pop, "Event-based sensor data exchange and fusion in the internet of things environments," *Journal of Parallel and Distributed Computing*, vol. 118, pp. 328 – 343, 2018.

[6] C. Cobb, S. Sudar, N. Reiter, R. Anderson, F. Roesner, and T. Kohno, "Computer security for data collection technologies," *Development Engineering*, vol. 3, pp. 1 – 11, 2018.

[7] R. Yadav, A. K. Pradhan, and I. Kamwa, "Real-time multiple event detection and classification in power system using signal energy transformations," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 3, pp. 1521–1531, 2019.

[8] M. R. Patra, R. K. Das, and R. P. Padhy, "Crhis: cloud based rural healthcare information system," in *Proc. of the ACM 6th International Conference on Theory and Practice of Electronic Governance*, 2012, pp. 402–405.

[9] T. Wang, Y. Li, W. Fang, W. Xu, J. Liang, Y. Chen, and X. Liu, "A comprehensive trustworthy data collection approach in sensor-cloud system," *IEEE Transactions on Big Data*, pp. 1–14, 2018.

[10] M. Rezvani, A. Ignjatovic, E. Bertino, and S. Jha, "Secure data aggregation technique for wireless sensor networks in the presence of collusion attacks," *IEEE Transactions on Dependable and Secure Computing*, vol. 12, no. 1, pp. 98–110, 2015.

[11] D. Jiao, M. Li, Y. Yu, and J. Ou, "Self-healing key-distribution scheme with collusion attack resistance based on one-way key chains and secret sharing in wireless sensor networks," *International Journal of Distributed Sensor Networks*, vol. 2015, pp. 1–7, 2015.
[12] H. Tao, M. Z. A. Bhuiyan, A. Abdalla, M. Hassan, J. Jain, and T. Hayajneh, "Secured data collection with hardware-based ciphers for iot-based healthcare," *IEEE Internet of Things Journal (IEEE IoT-J)*, pp. 1–10, 2018, https://doi.org/10.1109/JIOT.2018.2854714.
[13] M. Pouryazdan, B. Kantarci, T. Soyata, L. Foschini, and H. Song, "Quantifying user reputation scores, data trustworthiness, and user incentives in mobile crowd-sensing," *IEEE Access*, vol. 5, pp. 1382–1397, 2017.
[14] C. Rolim, F. Koch, J. W. C. Westphall and, A. Fracalossi, and G. Salvador, "A cloud computing solution for patients data collection in health care institutions," in *Proc. of the Second International Conference oneHealth, Telemedicine, and Social Medicine (ETELEMED'10)*, 2010, pp. 95–99.
[15] Y. Zhang, M. Qiu, C.-W. Tsai, M. M. Hassan, and A. Alamri, "Healthcps: Healthcare cyber-physical system assisted by cloud and big data," *IEEE Systems Journal*, vol. 11, no. 1, pp. 88–95, 2017.
[16] E. Luo, Q. Liu, and G. Wang, "Hierarchical multi-authority and attribute-based encryption friend discovery scheme in mobile social networks," *IEEE Communications Letters*, vol. 20, no. 9, pp. 1772–1775, 2016.
[17] J. Li, Y. Zhang, X. Chen, and Y. Xiang, "Secure attribute-based data sharing for resource-limited users in cloud computing," *Computers & Security*, vol. 72, pp. 1–12, 2018.
[18] E. Hendrick, B. Schooley, and C. Gao, "Cloudhealth: Developing a reliable cloud platform for healthcare applications," in *2013 IEEE 10th Consumer Communications and Networking Conference (CCNC)*, 2013, pp. 887–891.
[19] O. Gul, M. Al-Qutayri, C. Y. Yeun, and Q. H. Vu, "Framework of a national level electronic health record system," in *Proc. of 2012 International Conference on Cloud Computing Technologies, Applications and Management (ICCCTAM),*, 2012, pp. 60–65.
[20] L. A. Linn and M. B. Koo, "Blockchain for health data and its potential use in health it and health care related research," in *ONC/NIST Use of Blockchain for Healthcare and Research Workshop. Gaithersburg, Maryland, United States: ONC/NIST*, 2016. [Online]. Available: https://www.healthit.gov/sites/default/files/11-74-ablockchainforhealthcare.pdf
[21] H.-T. Pai, , and Y. S. Han, "Power-Efficient Direct-Voting Assurance for Data Fusion in Wireless Sensor Networks," *IEEE Transactions on Computers*, vol. 57, no. 2, pp. 261–273, 2008.
[22] M. Henze, R. Hummen, R. Matzutt, and K. Wehrle, *A Trust Point-based Security Architecture for Sensor Data in the Cloud*. Cham: Springer International Publishing, 2014, pp. 77–106.
[23] M. Z. A. Bhuiyan, G. Wang, J. Wu, J. Cao, X. Liu, and T. Wang, "Dependable structural health monitoring using wireless sensor networks," *IEEE Transactions on Dependable and Secure Computing*, vol. 14, no. 4, pp. 363–376, 2017.
[24] X. Liu, J. Cao, S. Lai, C. Yang, H. Wu, and Y. Xu, "Energy efficient clustering for WSN-based structural health monitoring," in *Proc. of IEEE INFOCOM*, 2011, pp. 2768–2776.
[25] S. H. Islam, K. Arijit, G. Biswas, M. Zakirul, P. Vijayakumar, , and M. Karuppiah, "Provably secure identity-based signcryption scheme for crowdsourced industrial internet of things environments," *IEEE IoT Journal*, vol. 0, pp. 1–10, 2017.
[26] P. Zhuang, D. Wang, and Y. Shang, "Distributed faulty sensor detection," in *Proc. of IEEE GLOBECOM* , 2009, pp. 1–6.
[27] P. Schaffer and I. Vajda, "CORA: Correlation-based resilient aggregation in sensor networks," *Ad Hoc Network*, vol. 7, no. 6, pp. 1035–1050, 2009.
[28] T. Clouqueur, K. K. Saluja, and P. Ramanathan, "Fault tolerance in collaborative sensor networks for target detection," *ACM Transactions on Computers*, vol. 53, no. 3, pp. 320–333, 2004.
[29] Y. Li, J. Gao, C. Meng, Q. Li, L. Su, B. Zhao, W. Fan, , and J. Han, "A Survey on Truth Discovery," *ACM SIGKDD Explorations Newsletter*, vol. 17, no. 2, pp. 1–16, 2015.
[30] Q. Li, Y. Li, J. Gao, B. Zhao, W. Fan, and J. Han, "Resolving conflicts in heterogeneous data by truth discovery and source reliability estimation," in *Proc. of ACM SIGMOD*, 2014, pp. 1–12.
[31] I. Damgard and M. Jurik, "Generalisation, a simplification and some applications of paillier's probabilistic public-key system," in *Proc. of PKC*, 2001, pp. 119–136.
[32] T.-Y. Youn, N.-S. Jho, and K.-Y. Chang, "Design of additive homomorphic encryption with multiple message spaces for secure and practical storage services over encrypted data," *The Journal of Supercomputing*, vol. 74, no. 8, pp. 3620–3638, 2018.
[33] R. Cramer, I. Damgard, and J. B. Nielsen, "Multiparty computation from threshold homomorphic encryption," in *Proc. of EUROCRYPT*, 2001, pp. 280–300.
[34] S. Mehnaz, G. Bellala, and E. Bertino, "A secure sum protocol and its application to privacy-preserving multi-party analytics," in *Proceedings of the 22Nd ACM on Symposium on Access Control Models and Technologies*. New York, NY, USA: ACM, 2017, pp. 219–230.
[35] Y. Ni, Y. Xia, W. Liao, and J. Ko, "Technology innovation in developing the structural health monitoring system for Guangzhou New TV Tower," *Struct. Cont. and Health Monit.*, vol. 16, no. 1, pp. 73–98, 2009.
[36] B. Li, D. Wang, F. Wang, and Y. Q. Ni, "High quality sensor placement for SHM systems: Refocusing on application demands," in *Proc. of IEEE INFOCOM*, 2010, pp. 1–9.
[37] ISHMP Toolsuite. [Online]. Available: http://shm.cs.uiuc.edu/
[38] S. Wei, Y. Meng, and C. Jean-Pierre, "Resilient secure localization and detection of colluding attackers in WSNs," in *Proc. of Ad-hoc, Mobile, and Wireless Networks*, 2012, pp. 181–192.

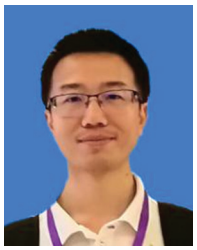
ACKNOWLEDGMENT

This work is supported by the Fordham University Faculty Grant, in part by the National Natural Science Foundation of China (No. 61971005), by the Natural Science Basic Research Plan in Shaanxi Province of China (No. 2018JM6043), and by the doctoral scientific research initial funding project of Baoji University of Arts and Sciences (ZK2018062), and in part by the UMP Grant no. RDU192202.

Tao Hai , PhD, is currently an Associate Professor in Baoji University of Arts and Sciences. Prior to that, he was a Senior Lecturer of Faculty of Computer Systems and Software Engineering of University Malaysia Pahang. He got his B.Sc in School of Computer and Information Science from Northwest University of Nationalities in 2004. He got his M.S in School of Mathematics and Statistics from Lanzhou University in 2009. Finally, he obtained his PhD in Fauclty of Computer System and Software Engineering from University Malaysia Pahang in 2012. His current research interests include machine learning, the Internet of things and optimization computation.

Md Zakirul Alam Bhuiyan , PhD (M'09-SM'17), is currently an Assistant Professor of the Department of Computer and Information Sciences at the Fordham University, NY, USA, the Founding Director of Fordham Dependable and Secure System Lab (DependSys). Earlier, he worked as an Assistant Professor at the Temple University. His research focuses on dependability, cybersecurity, big data, and IoT/CPS Applications. His work (including 40+ JCR Q1 papers) in these areas published in top-tier venues. Several research work of Dr. Bhuiyan have got recognition of ESI Highly Cited Papers. He has received numerous awards, including the IEEE TCSC Early Career Researcher, the IEEE Outstanding Leadership Award, and IEEE Service Award. He is a senior member of IEEE and a member of ACM.

Md arafatur Rahman , is currently an assistant professor in the Faculty of Computing, Universiti Malaysia Pahang. He got his Ph.D. degree from the University of Naples Federico II, Italy, and his Masters Degree from the International Islamic University Malaysia. His research interests include wireless communication, cognitive radio networks, security and Big Data. He is an IEEE Senior Member. He is also a research fellow of IBM CoE and ERAS at UMP, Malaysia. He is a co-author of more than 70 international journals and conferences.



Tian Wang received his BSc and MSc degrees in Computer Science from the Central South University in 2004 and 2007, respectively. He received his PhD degree in City University of Hong Kong in 2011. Currently, he is a professor at the College of Computer Science and Technology, Huaqiao University, China. His research interests include internet of things, edge computing and mobile computing.

Jie Wu is the Director of the Center for Networked Computing and Laura H. Carnell professor at Temple University. He also serves as the Director of International Affairs at College of Science and Technology. His research interests include mobile computing and wireless networks, routing protocols, cloud computing, and network trust and security. Dr. Wu publishes in scholarly journals, conference proceedings, and books. He serves on several editorial boards. Dr. Wu was general co-chair for MASS 2006, IPDPS 2008, ICDCS 2013, MobiHoc 2014, ICPP 2016, and CNS 2016, as well as program chair for INFOCOM 2011 and CNCC 2013. Dr. Wu is a Fellow of the IEEE.

Sinan Q. Salih received the B.Sc. degree in information systems from the University of Anbar, Anbar, Iraq, in 2010, the M.Sc. degree in computer sciences from the Universiti Tenaga National (UNITEN), Malaysia, in 2012, and the Ph.D. degree in soft modeling and intelligent systems from the Universiti Malaysia Pahang (UMP). His current research interests include optimization algorithms, nature-inspired metaheuristics, machine learning, and feature selection problem for real-world problems.

Yafeng Li received the Ph.D. degree from Xidian University, Xi'an, China, in 2011. He is currently a Professor in the School of Computer Science and Technology, Baoji University of Arts and Science. His research is focused on variation, sparse representation, optimization and their applications in pattern recognition and image processing.

Thaier Hayajneh Ph.D. is the founder and director of Fordham Center for Cybersecurity, University Professor, and director of the cybersecurity programs at Fordham University, New York. Prior to joining Fordham, he was a full-time faculty of computer science at New York Institute of Technology and founding director of NYIT Center of Excellence in Cyber Security (2014-2016). He received his Ph.D. and M.S. degrees in Information Sciences with specialization in cybersecurity and Networking from the University of Pittsburgh, PA. He also received his MS and BS in Electrical and Computer Engineering from Jordan University of Science and Technology, Irbid, Jordan, in 1999 and 1997, respectively.