Privacy Preservation in Industrial IoT via Fast Adaptive Correlation Matrix Completion

Aris S. Lalos, *Senior Member, IEEE*, Evangelos Vlachos, Kostas Berberidis, *Senior Member, IEEE*, Apostolos P. Fournaris, *Member, IEEE*, and Christos Koulamas, *Senior Member, IEEE*

***Abstract*—The Industrial Internet of Things (IIoT) is a key element of industry 4.0, bringing together modern sensor technology, fog and cloud computing platforms, and artificial intelligence to create smart, self-optimizing industrial equipment and facilities. Though, the scale and sensitivity degree of information continuously increases, giving rise to serious privacy concerns. The scope of this article is to provide efficient privacy preservation techniques, by tracking the correlation of multivariate streams recorded in a network of IIoT devices. The time-varying data covariance matrix is used to add noise that cannot be easily removed by filtering, generating obfuscated measurements and, thus, preventing unauthorized access to the original data. To improve communication efficiency between connected IoT devices, we exploit inherent properties of the correlation matrices, and track the essential correlations from a small subset of correlation values. Extensive simulation studies using constrained IIoT devices validate the robustness, efficiency, and effectiveness of our approach.**

***Index Terms*—Adaptive estimation, data processing, data privacy, distributed information systems.**

I. Introduction

INDUSTRIAL Internet of Things (IIoT) is a term used to describe the application of Internet of Things (IoT) in industrial environments, and more specifically the utilization of disruptive elements, such as sensors, actuators, control systems, machine-to-machine communication interfaces, and enhanced security mechanisms creating smart, self-optimizing industrial equipment and facilities [1]. The proliferation of IoT in industrial environments and value chains will allow companies, manufactures, and workers to operate in a more efficient manner and will have a great impact in several fields, including automation, industrial manufacturing, logistics, business processes, process management, and transportation [2], [3].

With the integration of a large number of computation components into industrial control systems, production systems, and factories, emerging mega trends, such as edge and cloud computing, big data analytics, and embedded artificial intelligence, are becoming important drivers of innovation in industry. Yet, industry has already realized that the true value of IoT is not on the physical interconnected devices per se, but on the massive datasets and crude, unrefined information they contain, and consequently how this hidden commodity can be efficiently processed in a fast and meaningful manner. Through IoT, industrial organizations will produce an unprecedented amount of raw information to evaluate accurately situational awareness, improving the intelligence, efficiency, and sustainability of several industrial systems [4]. This clearly violates user privacy, especially when considering that the user has willingly purchased the device that now may handle his personal data over to third-party data silos to be further processed. Consequently, it is essential for users to demand and legislative authorities to enforce a certain move toward data-centric privacy preserving schemes that will penalize paradox and improper data usage.

Conventional approaches to data protection are based on encryption and secure networks, encrypting at a gateway before transferring data to the Internet. This model is not the ideal one for cloud-based architectures since the data have to be decrypted on the cloud server for further processing, increasing the risk of experiencing a data breach. Alternatively, the data have to be brought back into the secure network for decryption. This approach has additional risks associated with the trust to users that have access to the network infrastructure (i.e., administrators). To the best of our knowledge, very few works focus on approaches that preserve privacy directly on streams of sensitive data, including defense, retail, performance, or even health care data. For example, consider two IoT networks deployed by two individual firms for monitoring several critical parameters, e.g., environmental, performance, health care, etc. Although, both networks could receive an advantage by a potential collaboration (e.g., to identify clusters over their real-time measurements), none of them is prepared for sharing the original monitored parameters. Therefore, there is an urgent need to design methods

Manuscript received September 30, 2019; accepted December 8, 2019. Date of publication December 17, 2019; date of current version September 18, 2020. This work was supported in part by the European Union's Horizon 2020 Research and Innovation Program CPSoSaware under Grant 873718, in part by the European Union's Horizon 2020 Research and Innovation Program CONCORDIA under Grant 830927, and in part by the project "I3T - Innovative Application of Industrial Internet of Things (IIoT) in Smart Environments" (MIS 5002434) implemented under the "Action for the Strategic Development on the Research and Technological Sector", funded by the Operational Programme "Competitiveness, Entrepreneurship and Innovation"(NSRF 2014-2020) and co-financed by Greece and the European Union (European Regional Development Fund). *(Corresponding author: Aris S. Lalos.)*

A. S. Lalos, E. Vlachos, A. P. Fournaris, and C. Koulamas are with the Industrial Systems Institute, Athena Research Center, 15125 Marousi, Greece (e-mail: lalos@isi.gr; evlachos@isi.gr; fournaris@isi.gr; koulamas@isi.gr).

K. Berberidis is with the Department of Computer Engineering and Informatics, University of Patras, 26500 Patras, Greece (e-mail: berberid@ceid.upatras.gr).

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TII.2019.2960275

that are able to utilize data, in an real-time fashion, without sacrificing privacy. One potential solution encompasses the so-called partially homomorphic encryption schemes [5]–[8] that, however, suffer from limitations attributed to their requirement of a trusted third party. Moreover, several secure multiparty computation techniques, presented in [9], become intractable either as the number of parties increases or in the presence of transmission errors, whereas the utility of the published data in different mining applications decreases with increasing level of privacy [10].

At this point, it should be also mentioned that several privacy preserving approaches suggest adding random perturbation, by projecting the noise to the principal components (PCs) of the data covariance increasing significantly different privacy preservation metrics, allowing at the same time the utilization of the data by different mining methods [11]–[14]. These approaches either work offline using stationary data streams [11] or they are capable of tracking correlations between time-evolving data streams, addressing challenges related to storage constraints or time evolving correlations. Despite the benefits offered by the online algorithms in evolving data streams, the communication overhead required for evaluating the PC analysis (PCA) subspace locally, increases significantly with the number of nodes, which are not resilient over intrinsic (e.g., power depletion of a node, link failures/packet loss) as well as to extrinsic failures (e.g., malicious nodes).

To address the aforementioned limitations, we provide a method for reconstructing the obfuscated data covariance matrix that have been captured by the nodes in an IIoT platform, when several entries of the matrix are missing, either due to intrinsic or extrinsic failures. To that end, we provide a fast correlation completion and tracking method, by solving at each step, a rank-one completion problem. The privacy of the exchanged data is properly preserved throughout the whole process.

Specifically, the contributions of this article can be summarized as follows.

1) We introduce a data obfuscation method that guarantees privacy and preserves the utilization of the data in a real-time fashion while time minimizes the communication and processing requirements. This method is capable of estimating the original data correlation subspace from a subset of obfuscated (i.e., privacy preserved) values, generated by a small number of IoT nodes. The data, before being exchanged, are properly obfuscated by adding noise, which is correlated with the recorded measurements.
2) To further minimize processing cost, we propose a low-complexity adaptive algorithm for tracking the evolving network-wide correlation subspace at each node. For the case of large-scale networks, the proposed algorithm requires only linear complexity over the number of nodes.
3) We provide both theoretical arguments and extensive evaluation studies in the Contiki Cooja wireless sensor network (WSN) simulator. More importantly, we validate the robustness and efficiency of our approach in terms of both computational complexity and privacy preservation on constrained IIoT devices.

TABLE I
SUMMARY OF NOTATIONS

$a, \mathbf{a}$ and $\mathbf{A}$	Scalar, vector and matrix variables
$\mathbf{A}^T$ and $\mathbf{A}^H$	Matrix transpose and Hermitian transpose
$\mathbf{A}^{-1}$ and $\mathbf{A}^\dagger$	Matrix inverse and pseudo-inverse
$[\mathbf{A}]_{i,j}$	Matrix element at the i-th row and j-th column
$\mathbf{M}^*$ and $\hat{\mathbf{M}}$	Raw and masked correlation matrix
$\mathbf{I}_N$	$N \times N$ identity matrix
$\mathbf{0}_{N\times K}$	$N \times K$ matrix with zeros
$\mathbf{I}_{N\times K}$	Concatenated matrix $[\mathbf{I}_N \ \mathbf{0}_{N\times K}]$
Ω	Set containing matrix positions of observed entries
$\mathcal{P}_\Omega(\mathbf{X})$	Operator that returns a Matrix consisting of $[X]_{ij}$ values if $(i,j) \in \Omega$ and zero otherwise.
$\|\cdot\|_*$, $\|\cdot\|_F$	Nuclear and Frobenius norms of matrix
$\times$	Scalar multiplication
$\circ$	Element-wise (Hadamard) matrix product
$\mathrm{diag}(\mathbf{a})$	Diagonal matrix with vector $\mathbf{a}$ on the diagonal
SVT	Singular Value Thresholding with threshold t
$\mathcal{L}(\mathbf{X}, \mathbf{Y})$	Lagrangian with primal ($\mathbf{X}$) and dual ($\mathbf{Y}$) variables
$\mathcal{E}\{\cdot\}$	statistical average value

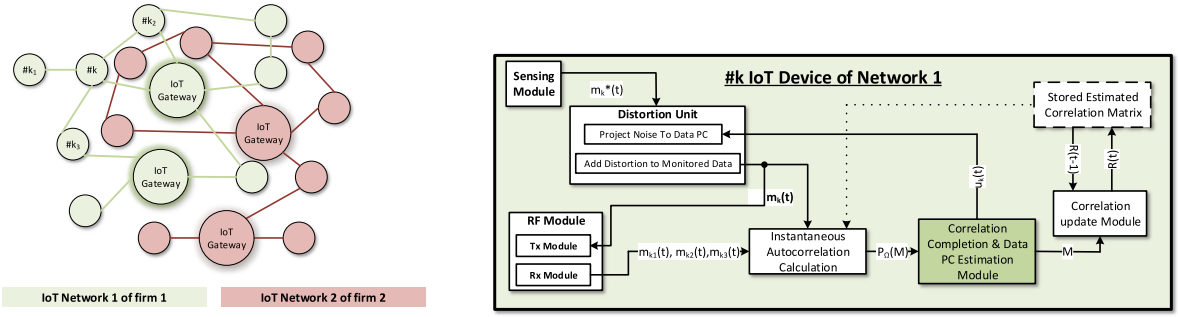
A summary of the notation used in this article is presented in Table I. The rest of this article is organized as follows. In Section II, we describe the privacy preservation scenarios that we focus and the architecture of the proposed solution. In Section III, we provide preliminaries, introducing the terminology and concepts of matrix completion (MC) and subspace tracking, and presents the proposed low-cost solution for reconstructing the network-wide correlation matrix from a subset of values. In Section IV, we present the simulation results on constrained IIoT devices. Finally, Section V concludes this article.

II. Data Model and Communication Scheme

Data obfuscation techniques perturb data values or attributes directly by additive and/or multiplicative noise. These techniques are based on the fact that the properly randomized samples could potentially preserve the underlying probabilistic properties. This property is very significant in our case, since we aim to compute the sample-based correlation matrix for all nodes, given that the true values of the data are unknown and cannot be revealed.

Let us consider two networks, shown in Fig. 1(a), deployed by two different firms/organizations for collaboratively monitoring environmental, performance, or even health care parameters over their real-time measurements. Though, none of them is prepared to share the original recordings. Our goal in this article is to provide a mechanism that could be adopted by any of the two networks that guarantee data privacy of data, preserving also utilization in an online fashion.

Without loss of generality, let us assume that Network 1 consist of K sensor nodes and each node k records the value $m_k^*(t) \in \mathbb{R}$ at time instant t. The conceptual architecture of the functional block designed to hide original data from IoT Network 2 is presented in Fig. 1(b). The original recorded data are forwarded to the *distortion unit*, which is responsible for generating obfuscated data using additive noise that is correlated with the recorded data from the other network IoT nodes. A



(a) (b)

Fig. 1. (a) Two IoT networks deployed by two individual firms for collaboratively monitoring environmental, performance parameters, or even health care parameters, though, they are not willing to provide the original measurements. (b) Schematic diagram of the functional blocks used by the proposed privacy preservation approach.

subset of network measurements [i.e., $\{k_1, k_2, k_3\}$ according to Fig. 1(b)] are used for forming the network autocorrelation matrix with various missing values, which is then completed in the *correlation completion and data PC estimation* module. The estimated network correlation matrix is then used to update the stored autocorrelation estimate using an exponentially decaying time window, whereas the estimated PC is used by the distortion unit to project independent and identically distributed (i.i.d.) noise to the PC of the instantaneous data autocorrelation matrix. A more detailed description of the process that is executed in each block is provided in the rest part of this section.

Consider that $m_k^*(t)$ is the measurement recorded by the sensing modules of the k_{th} node, at time instant t, which can be modeled as a discrete-time wide sense stationary stochastic process. A common method for data obfuscation is to simply use additive noise [15], however, it has been proved that this is not optimal for preserving privacy [11], [16]. This conclusion is justified by the fact that the original data can be reconstructed by using spectral or PCA based filtering approaches. To increase robustness to such an unauthorized reconstruction, the data owners should project random noise to the PCs of the data correlation matrix to prevent linear reconstruction methods from canceling additive noise.

To ensure privacy, each measurement is modified at the distortion unit by adding noise. All noisy sensor measurements form a vector $\mathbf{m}(t) = [m_1(t) \dots m_K(t)]^T \in \mathbb{R}^{K\times 1}$ and $m_k(t) = m_k^*(t) + n_k(t)$, where $k = 1, \dots, K$ and $\mathbf{n}(t) = [n_1(t) \dots n_K(t)]^T$ is the vector with the added noise for each node, with $\mathbf{n}(t) \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_n)$, where $\mathbf{C}_n \in \mathbb{R}^{K\times K}$ is the noise covariance matrix. In this article, we assume that the physical process that we monitor generates highly correlated samples. Thus, the original data covariance matrix $\mathbf{C} \in \mathbb{R}^{K\times K}$ is low rank.

Our goal is to add in an online fashion, random perturbations with statistical properties similar to those of the data streams. Let us assume that $\mathbf{C} = \mathcal{E}\{\mathbf{M}(t)\} \in \mathbb{R}^{K\times K}$, where $\mathbf{M}(t) = \mathbf{m}(t)\mathbf{m}^T(t)$, is the *correlation matrix* of the process $\mathbf{m}(t)$. Following the mean ergodic theorem, an exponentially decaying window can be used to estimate the correlation matrix $\mathbf{C}$

$$\mathbf{R}(t) = \frac{t-1}{t}\mathbf{R}(t-1) + \frac{1}{t}\mathbf{M}(t) = \frac{1}{t}\sum_{\tau=1}^{t}\mathbf{M}(\tau) \quad (1)$$

with $\lim_{t\to\infty} \mathbf{R}(t) = \mathbf{C}$, where $\text{rank}(\mathbf{C}) \ll K$, meaning that matrix $\mathbf{C}$ is a low-rank matrix.

Algorithm 1: Reconstruction of the Network-Wide Correlation Matrix.

1: **for** $t = 1, 2, \dots$ each node **do**
2: Based on the available measurements of its own and the obfuscated ones received by the collaborating nodes, computes the corresponding correlation values.
3: Reconstructs the full correlation matrix from the known subset of the estimated values.
4: **end for**

Conventionally, to construct the entire correlation matrix in the instantaneous autocorrelation evaluation unit, each node has to receive, at each time instant t, K measurements $\mathbf{m}(t)$ and evaluate $(K-1)^2/2 + K$ multiplications. In this work, we assume that each node may have incomplete knowledge of the set of measurements [e.g., in Fig. 1(b)], node k receives only measurements from a subset of neighboring nodes $\{k_1, k_2, k_3\}$), since many data packets may be lost during network transmissions or due to energy depletion of the nodes or because there may not be direct links with all other nodes (which is actually the case in most networks). Hence, each node is required to reconstruct the entire correlation matrix $\mathbf{R}(t)$, at each time instance, based on incomplete measurements. Algorithm 1 presents the basic steps executed by the instantaneous autocorrelation evaluation unit for estimating $\mathcal{P}_\Omega(\mathbf{M})$ and by the correlation completion and data PC estimation module for estimating matrix $\mathbf{M}$ and its principal nonzero eigenvector $\mathbf{u}_k$, from a subset of obfuscated measurements received from other nodes, without having knowledge about the measurements of the entire network.

In the following part, we present the scheme for reconstructing the network-wide correlation matrix and tracking its principal eigenvector, with the aim of generating the correlated noise. This type of noise is hard to be filtered out by unauthorized nodes and does not change the statistical properties of the original data of nodes in the network. This scheme consists of two steps executed at each time instant t. In the first step, each

node computes the correlation between its own measurements and the raw obfuscated measurements that receives from the collaborating nodes, evaluating some of the entries of the kth row and column of the network correlation matrix $\mathcal{P}_\Omega(\mathbf{M})$. In the second step, a sparse matrix is being formed at each node. To obtain the missing entries, MC techniques can be successfully employed, since at each time update, the rank of the correlation updating term $\mathbf{M}(t)$ is equal to one. The completion steps of the network-wide correlation matrix are summarized in Algorithm 2 and correspond to the operations executed by the correlation completion and data PC estimation module. The output of this module is used: to update the stored network wide correlation matrix using (1) and to estimate the noise for generating the obfuscated measurements that have the same statistical properties with the original data.

The correlated noise $\mathbf{n}(t)$ is estimated at the distortion unit by projecting the generated vector $\mathbf{n}(t)$ to the estimated PC of the estimated data covariance $\mathbf{R}(t)$, i.e., if we assume that $\mathbf{R}(t) = \mathbf{U}\mathbf{\Sigma}\mathbf{U}^T$ is the eigenvalue decomposition of $\mathbf{R}(t)$, where $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_K] \in \mathbb{R}^{K\times K}$ are the K eigenvectors, whereas $\mathbf{\Sigma} = \text{diag}(\sigma_1, \ldots, \sigma_K)$ are the K eigenvalues in descending order. Then, the proposed projection can be expressed in matrix form

$$\mathbf{n}(t) = \mathbf{P}_{u_1}\mathbf{n}(t) \text{ with } \mathbf{P}_{u_1} = \mathbf{u}_1\mathbf{u}_1^T \tag{2}$$

where $\mathbf{u}_1$ is the principal eigenvector, which corresponds to the largest eigenvalue σ_1.

The noise $\mathbf{n}(t)$ is added to the original samples in order to generate the obfuscated data. More specifically, the obfuscated sample in node i at time instant t may be written as

$$m_k(t) = m_k^*(t) + u_{1,k}\mathbf{u}_1^T\mathbf{n}(t) \tag{3}$$

where $\mathbf{u}_{1,k}$ is the kth entry of the principal eigenvector $\mathbf{u}_1$.

As it is also verified in the simulation section, the PCs of the obfuscated streams $\mathbf{m}(t)$ corresponds to a good estimate of the PCs of data autocorrelation matrix of the process $\mathbf{m}^*(t)$. More importantly, this noise cannot be easily filtered out since it mirrors the dominant trends in the time series. To further justify this statement, let us assume that the data stream always has the same value. The right way to hide this value is to add the same noise sample through time, since any other noise can be filtered out by simple averaging. Similarly if the time series is a sine wave, again it should be hided by adding noise samples with the same frequency and potentially a different phase. The approach described earlier is a generalization of the aforementioned examples.

A. Unauthorized Reconstruction of the Original Measurements

Any node belonging to Network 2 that would like to recover the original data recorded by the Network 1 nodes could apply a linear filtering operation. The *linear reconstruction* of the original data $\mathbf{M}^*(t)$ can thus be expressed as $\tilde{\mathbf{M}}^*(t) = \mathbf{F}\mathbf{M}(t)$ where the matrix $\mathbf{F} \in \mathbb{R}^{K\times K}$ applies a low-pass filter to the perturbed data matrix. A principled approach that can be applied for this purpose is to project the data onto the signal (i.e., PC) subspace, such that most noise is removed while preserving the original data. More specifically, a PCA-based reconstruction scheme can be expressed as

$$\hat{\mathbf{M}}^*(t) \approx \underbrace{\mathbf{U}_k(t)\mathbf{U}_k^T(t)}_{\mathbf{F}}\mathbf{M}(t) \tag{4}$$

where $\mathbf{U}_k(t)$ corresponds to a $K \times k$ matrix with the k principal eigenvectors of the autocorrelation matrix $\mathbf{R}(t)$. However, when the noise is added exactly along the PC direction, the removed noise tends to zero, whereas additional projection error is included. In this case, the data obfuscation described earlier is robust toward this "unauthorized" reconstruction attempt in the sense that the discrepancy between the two versions of the data, defined as the normalized squared Frobenious norm, before $\|\mathbf{M}^*(t) - \hat{\mathbf{M}}(t)\|_F^2$ and after $\|\hat{\mathbf{M}}^*(t) - \hat{\mathbf{M}}(t)\|_F^2$ the reconstruction attempt will be the same.

III. Completion of the Network-Wide Correlation Matrix

In this section, we focus on the design of the correlation completion and data PC estimation module, minimizing the processing and communication cost at each node. This is achieved by a novel low-complexity fast adaptive algorithm for tracking the evolving network-wide correlation subspace. More specifically, we initially provide details on how we can reconstruct the unobserved instantaneous correlation values, by utilizing the MC framework. Then, motivated by the fact that the matrix we wish to recover at each time step is a rank one matrix, we utilize fast iterative subspace tracking approaches resulting at a low complexity scheme, capable of accurately estimating the obfuscated data covariance PC from a small subset of obfuscated data values.

A. Reconstruction of the Obfuscated Correlation Matrix

Before presenting the proposed approach, we review some concepts and terminology related to the MC theory. MC [17] theory provides an iterative technique to recover a low-rank matrix given a sampling of its entries. Typically, the completion of a matrix $\mathbf{X} \in \mathbb{R}^{K\times K}$ given a sampling of its entries, i.e., $\mathcal{P}_\Omega(\mathbf{C})$, can be expressed as

$$\min_{\mathbf{X}} \text{rank}(\mathbf{X}) \;\; \text{s. t. } \mathcal{P}_\Omega(\mathbf{X}) = \mathcal{P}_\Omega(\mathbf{C}) \tag{5}$$

where $\mathbf{C} \in \mathbb{R}^{K\times K}$ is the complete correlation matrix, Ω is the set with the matrix indices (i, j) of the nonzero entries, $\mathbf{X}$ is the unknown variable, and $\text{rank}(\mathbf{X})$ denotes the rank of the matrix $\mathbf{X}$. The problem (5) is computationally intractable (NP-hard), thus, it cannot be solved in polynomial time.

To solve (5), the following approximation can be employed [18]:

$$\min_{\mathbf{X}} \kappa\|\mathbf{X}\|_* + \frac{1}{2}\|\mathbf{X}\|_F^2 \text{ subject to } \mathcal{P}_\Omega(\mathbf{X}) = \mathcal{P}_\Omega(\mathbf{C}) \tag{6}$$

where $\kappa \geq 0$ is a predefined weighting parameter. To proceed, let us express the Lagrangian of problem (6)

$$\mathcal{L}(\mathbf{X}, \mathbf{Y}) = \kappa\|\mathbf{X}\|_* + \frac{1}{2}\|\mathbf{X}\|_F^2 + \text{tr}(\mathbf{Y}^T\mathcal{P}_\Omega(\mathbf{C} - \mathbf{X})) \tag{7}$$

Algorithm 2: Completion of Low-Rank Correlation Matrix.

1: **for** $z = 1, \ldots, I_{\max}$ **do**
2: $\quad \mathbf{X}^{(z)} = \mathcal{D}_{\tau}(\mathbf{Y}^{(z-1)})$
3: $\quad \mathbf{Y}^{(z)} = \mathbf{Y}^{(z-1)} + \delta^{(z)}\mathcal{P}_{\Omega}(\mathbf{C} - \mathbf{X}^{(z)})$
4: **end for**

where $\mathbf{Y} \in \mathbb{R}^{K\times K}$ is the dual variable. Therefore, the solution of (6) can be obtained by obtaining the derivative of the Lagrangian function. First, with respect to $\mathbf{X}$, we have

$$\frac{\partial \mathcal{L}(\mathbf{X},\mathbf{Y})}{\partial \mathbf{X}} = \kappa \partial \|\mathbf{X}\|_* + \mathbf{X} - \mathcal{P}_{\Omega}(\mathbf{Y}) \tag{8}$$

which can be solved via the singular-value-thresholding (SVT) operator $\mathcal{D}_{\kappa}(\mathbf{Y})$ [18, Th. 2.1]. To be more specific, given that $\mathbf{Y} = \mathbf{U}\mathbf{\Sigma}\mathbf{U}^*$ is the EVD of a matrix $\mathbf{Y}$, then the SVT operator is defined as

$$\mathcal{D}_{\kappa}(\mathbf{Y}) = \mathbf{U}\text{diag}\big(\{([\mathbf{\Sigma}]_{ii} - \kappa)_+\}_{1\le i\le r}\big)\mathbf{U}^T$$

i.e., the singular values with $[\mathbf{\Sigma}]_{ii} < \kappa$ are replaced by zero. Thus, it provides the minimizer of (8), assuming a number of iterations

$$\mathbf{X}^{(z)} = \mathcal{D}_{\kappa}\left(\mathbf{Y}^{(z-1)}\right) \tag{9}$$

with $z = 1, \ldots, I_{\max}$. Considering now the minimization of the Lagrangian over $\mathbf{Y}$, we seek the minimizer of the expression

$$\frac{\partial \mathcal{L}(\mathbf{X},\mathbf{Y})}{\partial \mathbf{Y}} = \mathcal{P}_{\Omega}(\mathbf{C} - \mathbf{X}) \tag{10}$$

which is obtained by

$$\mathbf{Y}^{(z)} = \mathbf{Y}^{(z-1)} + \delta^{(z)}\mathcal{P}_{\Omega}\left(\mathbf{C} - \mathbf{X}^{(z)}\right) \tag{11}$$

where $\mathbf{Y}^{(z-1)}$ denotes the obtained matrix $\mathbf{Y}$ of the previous iteration, and $\delta^{(z)} > 0$ denotes the step-size parameter of the iterative procedure. In summary, the solution of (6) is provided by the iterative procedure of Algorithm 2.

It should be noted that in our case [presented in Fig. 1(b)], matrix $\mathbf{M}(t)$ that we wish to recover from a subset of obfuscated correlation values (i.e., $m_{k_i}(t), m_{k_j}(t)$, where $i, j = 1, 2, 3$), is a rank one matrix. Therefore, we are interested in estimating only the principal eigenvector [e.g., index $i = 1$ of the SVT operation in step presented in (9)], by deploying computational efficient approaches as the one presented below.

B. Completion for Rank-1 Correlation Matrix

To impose the rank-1 constraint, we replace (9) with (12)

$$\mathbf{X}^{(z)} = \lambda_{\max}\mathbf{u}_1^{(z)}\left(\mathbf{u}_1^{(z)}\right)^T \tag{12}$$

where $\lambda_{\max} = \sigma_1$. Subsequently, $\mathbf{Y}^{(z)}$ is updated as follows:

$$\mathbf{Y}^{(z)} = \mathbf{Y}^{(z-1)} + \delta\mathcal{P}_{\Omega}\left(\mathbf{M} - \mathbf{X}^{(z)}\right) \tag{13}$$

where we keep the same step size for all the executed iterations, i.e., $\delta^{(z)} = \delta$. Due to the symmetric property of matrices $\mathbf{M}$ and $\mathbf{X}^{(z)}$, matrix $\mathbf{Y}^{(z)}$ will also be symmetric, thus, the SVD operation collapses to eigenvalue decomposition (EVD).

Since (12) uses only the maximum eigenvalue, instead of executing SVT, we focus on solving a maximum eigenvalue estimation problem that can be written as follows:

$$\mathbf{u}_1^{(z)} = \arg\max_{\mathbf{u}} \frac{\mathbf{u}^T\mathbf{Y}^{(z-1)}\mathbf{u}}{\mathbf{u}^T\mathbf{u}}. \tag{14}$$

Note that the complexity of this problem is still high, of the same order of the SVT, i.e., $\mathcal{O}(K^3)$. To minimize this computational cost, several fast iterative approaches for updating the principal eigenvector $\mathbf{u}_1^{(z)}$ of the matrix $\mathbf{Y}^{(z-1)}$ can be deployed. A common fast approach is the one proposed in [19], that can be expressed as follows:

$$\mathbf{u}_1^{(z)} = \frac{\mathbf{u}_1^{(z-1)} + \alpha^{(z)}\mathbf{Y}^{(z-1)}\mathbf{u}_1^{(z-1)}}{\|\mathbf{u}_1^{(z-1)} + \alpha^{(z)}\mathbf{Y}^{(z-1)}\mathbf{u}_1^{(z-1)}\|} \tag{15}$$

where $\alpha^{(z)}$ is a step-size that affects convergence speed and the estimators variance. Therefore, by adopting the Oja rule, we derive Algorithm 3, representing the basic operations that are executed by the correlation completion and data PC estimation module in Fig. 1(b).

Proposition 1: Algorithm 3 converges to the maximum eigenvector of matrix $\mathcal{E}\{\mathbf{Y}^{(z)}\}$ after a number of iterations $I_{\max}$, for the sequence of matrices $\mathbf{Y}^{(z)} = \sum_{i=1}^{z}\mathcal{P}_{\Omega}(\mathbf{C} - \mathbf{X}^{(i)})$ with $z = 1, 2, \ldots, I_{\max}$.

Proof: According to the performance analysis of [19], the convergence of the algorithm (15) is guaranteed if $\mathbf{Y}^{(I_{\max})} = \mathcal{E}\{\mathbf{Y}^{(I_{\max})}\}$. To prove this, let us express (11) as

$$\mathbf{Y}^{(z)} = \delta\sum_{i=1}^{z}\mathcal{P}_{\Omega}\left(\mathbf{C} - \mathbf{X}^{(i)}\right) \tag{16}$$

given that $\mathbf{Y}^{(0)} = \mathbf{0}$, i.e., matrix with all zeros. Then, capitalizing the properties of the projection operator $\mathcal{P}(\cdot)$, we have that

$$\mathbf{Y}^{(z)} = \delta z\mathcal{P}_{\Omega}(\mathbf{C}) - \delta\mathcal{P}_{\Omega}\left(\sum_{i=1}^{z}\mathbf{X}^{(i)}\right) \tag{17}$$

which for $z = I_{\max}$ is written as

$$\mathbf{Y}^{(I_{\max})} = \delta I_{\max}\mathcal{P}_{\Omega}(\mathbf{C}) - \delta\mathcal{P}_{\Omega}\left(\sum_{i=1}^{I_{\max}}\mathbf{X}^{(i)}\right). \tag{18}$$

Next, we take the expectation of $\mathbf{Y}_{I_{\max}}$

$$\mathcal{E}\{\mathbf{Y}^{(I_{\max})}\} = \delta I_{\max}\mathcal{P}_{\Omega}(\mathbf{C}) - \delta\mathcal{P}_{\Omega}\left(\sum_{i=1}^{I_{\max}}\mathcal{E}\{\mathbf{X}^{(i)}\}\right) \tag{19}$$

$$= \delta I_{\max}\mathcal{P}_{\Omega}(\mathbf{C}) - \delta I_{\max}\mathcal{P}_{\Omega}(\mathcal{E}\{\mathbf{X}\}). \tag{20}$$

Thus, assuming that the expected value of $\mathbf{X}^{(i)}$ is approximated by

$$\mathcal{E}\{\mathbf{X}\} \approx \frac{1}{I_{\max}}\sum_{i=1}^{I_{\max}}\mathbf{X}^{(i)} \tag{21}$$

then we have that $\mathbf{Y}^{(I_{\max})} = \mathcal{E}\{\mathbf{Y}^{(I_{\max})}\}$. ■

Complexity: The overall complexity of Algorithm 3 is $\mathcal{O}(KLI_{\max})$, where $I_{\max}$ is the maximum number of iterations. By increasing the number of iterations $I_{\max}$, we achieve a higher

Algorithm 3: Rank One Matrix Completion.

for $z = 1, \ldots, I_{\max}$ **do**
  $\mathbf{u}^{(z)} \leftarrow \mathbf{u}^{(z-1)} + \alpha^{(z)} \mathbf{Y}^{(z-1)} \mathbf{u}^{(z-1)}$
  $\mathbf{u}^{(z)} \leftarrow \frac{\mathbf{u}^{(z)}}{\|\mathbf{u}^{(z)}\|}$
  $\mathbf{X}^{(z)} = \lambda_{\max} \mathbf{u}^{(z)} (\mathbf{u}^{(z)})^T$
  $\mathbf{Y}^{(z)} = \mathbf{Y}^{(z-1)} + \delta \mathcal{P}_{\Omega}(\mathbf{M} - \mathbf{X}^{(z)})$
end for

estimation accuracy, though, the value of $I_{\max}$ seems to be independent of the number of nodes in the network K, whereas for a large-scale network, it could be $K \gg I_{\max}$. Thus, the complexity of the proposed algorithm increases *linearly* with the number of the nodes, given that $I_{\max} \ll K$ and $L \ll K$.

IV. Simulation Results

In this section, we initially evaluate the performance of the proposed technique based on Monte Carlo simulations and then we perform a practical evaluation to a realistic simulation environment of a resource constrained, IIoT system (6LoWPAN Contiki nodes and Cooja) and to real hardware nodes (Tmote Sky and Zolertia ReMote).

A. Simulation Setup and Algorithmic Evaluation

The efficiency of the proposed approach is evaluated by executing a large number of independent experiments, where a new scenario of a network is created with K sensor nodes. A number of $L = |\Omega|$ edges are randomly generated, whereas in all the executed scenarios, the topology graph is always a connected graph. We assume that the communication links between the sensor nodes are noiseless. The measurements $\mathbf{m}(t)$ are generated according to $\mathbf{m}(t) = \mathbf{C}\mathbf{d}(t)$, in order to impose correlations between the monitored parameters. The matrix $\mathbf{C} \in \mathbb{R}^{K \times K}$ is a fixed, symmetric matrix that represents the correlation among the sensor measurements. Its entries are drawn from a uniform distribution, i.e., $[\mathbf{C}]_{i,j} = [\mathbf{C}]_{j,i} \sim \mathcal{U}(0, 1)$. The vector $\mathbf{d}(t)$ represents the underlying random process, and its entries are normal random variables, i.e., $[d(t)]_i \sim \mathcal{N}(0, 1)$, for $i \in [1, \ldots, K]$.

Our goal is initially to evaluate privacy preservation of two different streaming scenarios where the obfuscated measurements are generated by adding : (i) i.i.d. noise and (ii) correlated additive noise. For the (ii) case, we evaluate the effect of both the number of missing entries in the correlation updating term and the executed iterations $I_{\max}$ in the privacy preservation. The goal of data obfuscation is to ensure that the introduced distortion cannot be filtered out with the PCA-based reconstruction approach. Thus, to measure the privacy, we have to consider the ability of an unauthorized node, in reconstructing the original data. More specifically, suppose that $\mathbf{M}_{\mathbf{s}}^* = [\mathbf{m}^*(1), \ldots, \mathbf{m}^*(t)] \in \mathbb{R}^{K \times t}$ and $\mathbf{M}_{\mathbf{s}} = [\mathbf{m}(1), \ldots, \mathbf{m}(t)] \in \mathbb{R}^{K \times t}$ are the original and obfuscated data generated at the nodes of Network 1 presented in Fig. 1(a), during the time instances $1, 2, \ldots, t$ and that $\hat{\mathbf{M}}_{\mathbf{s}}$ are the corresponding data reconstructed by any of the unauthorized nodes in Network 2, by following the approach described in Section II-A, e.g., $\hat{\mathbf{M}}_{\mathbf{s}} = \mathbf{U}_k(t)\mathbf{U}_k^T(t)\mathbf{M}_{\mathbf{s}}$. Then, privacy is the discrepancy (e.g., squared Frobenius norm) between the original and reconstructed streams, whereas privacy preservation is evaluated by the PD metric defined as the privacy normalized by the discrepancy between the original and obfuscated streams, e.g., $\text{PD} = ||\mathbf{M_s}^* - \hat{\mathbf{M}}_{\mathbf{s}}||_2^2 / ||\mathbf{M_s}^* - \mathbf{M_s}||_2^2$.

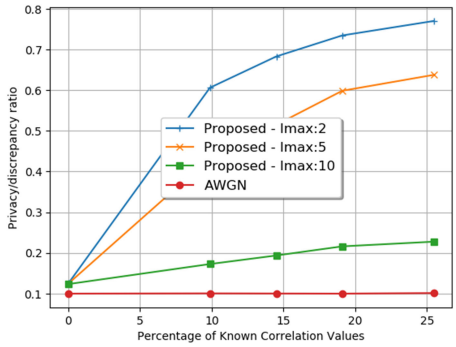


Fig. 2. Privacy evaluation with respect to the known entries of the correlation matrix. The number of nodes is $K = 20$. Discrepancy is 0.1.

The step size α_k in (15) is selected to be equal to $\alpha_k = 1/k$, ensuring convergence according to the analysis that is provided in [19]. Moreover, the step size δ in (13) is selected to be fixed across iterations and equal to one, since according to [18, Th. 4.2], the convergence for the reconstruction scheme is guaranteed when $0 < \delta < 2$. In the following part of this section, we study the impact of the Algorithm iterations, the missing entries, and the number of network nodes in privacy preservation, as well as the utility of the obfuscated data. Considering the obfuscated versus the original data, the larger the variance of the added noise, the larger the distortion of the data. In other words, a larger distortion hides the original data, though we are interested in investigating whether it hides also information about the relations of the data. To this end, we define as utility, the discrepancy between the covariance of the original and the obfuscated data.

1) Impact of Algorithm Iterations and Missing Entries in Privacy Preservation: In Fig. 2, we provide the evolution of privacy versus discrepancy ratio with respect to the known correlation values that occur in various WSN network setups with $K = 20$ nodes. Each curve corresponds to the average of 100 realizations, executing the same number of $I_{\max}$ iterations for completing the missing entries. From the figure, it is obvious that the privacy preservation is significantly increased with the number of iterations executed during the reconstruction of the rank-1 matrix. More importantly, by inspecting this figure, it is obvious that a small number of iterations (e.g., 5) affects significantly the privacy preservation metric as compared to the conventional AWGN case (e.g., 400% increase assuming 10% of known correlation values).

2) Impact of the Number of Nodes: In Fig. 3, we present the evolution of the privacy versus discrepancy ratio versus the number of nodes in various WSN setups. The discrepancy is fixed to 0.3. The number of connections in the network corresponds on average to 25% of connections of a fully connected network and the number of executed iterations $I_{\max}$ of Algorithm 2 is set to 25. The privacy metric for the correlated additive case remains almost unaffected with the number of nodes in the network K,

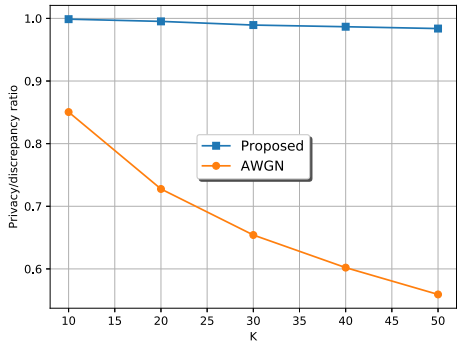


Fig. 3. Privacy evaluation with respect to the number of the IoT nodes. The discrepancy is 0.3.

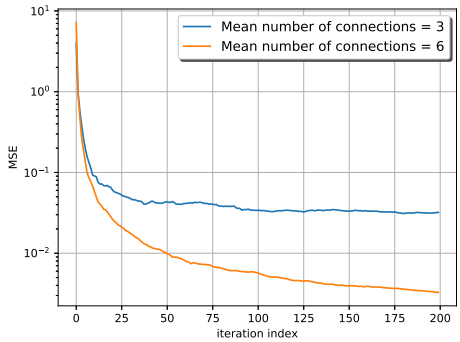


Fig. 4. NMSE with respect to the iteration index of Algorithm 2. The number of nodes is $K = 20$ and the discrepancy is 0.1.

whereas in the AWGN case, the privacy preservation reduces linearly with K.

3) Obfuscated Versus Original Stream Covariance: To evaluate the effect of using the obfuscated measurements for the estimation of the data correlation matrix, we make use of the normalized mean square error (NMSE), which is defined as $\text{NMSE} = 1/I_{\max} \sum_{r=1}^{I_{\max}} \|\mathbf{C}^* - \mathbf{R}(r)\|_F^2 / \|\mathbf{C}^*\|_F^2$ where the supper r denotes the MC iteration index, with $r = 1, 2, \ldots, I_{\max}$. $\mathbf{C}^* = 1/T \sum_{t=1}^{I_{\max}} \mathbf{m}^*(t)\mathbf{m}^*(t)^T$ is the data covariance matrix and $\mathbf{R}(r)$ is the estimation of the obfuscated data covariance at iteration r. Fig. 4 presents this NMSE with respect to the time index assuming WSNs with 20 nodes, with different number of mean connections per node. The number of executed iterations is fixed to 20. By inspecting this figure, it can be seen that the covariance estimation using obfuscated data is accurate even in very sparse networks.

B. Practical Evaluation in Real IoT Nodes

For the feasibility and efficiency assessment of the proposed method, an automated simulation and testing platform have been developed, in order to seamlessly transfer the same parametric environment from the algorithmic development and testing tools (Julia and Juno IDE [20]) to a realistic environment on a resource constrained, IIoT system that uses the Contiki 3.0 OS [21] (6LoWPAN Contiki nodes and Cooja). The tests on the IoT nodes consisted of two different approaches, a simulation-based approach using the Contiki Cooja simulator on a TI MSP430 based embedded system for measurements related to the energy consumption of the proposed solution and a real hardware based approach using two different hardware nodes (Tmote Sky and Zolertia ReMote) for evaluating the time delay of the MC operation in comparison to the number of involved nodes. Thus, we can evaluate the solution in a WSN node in terms of energy consumption for motes that focus explicitly on low power consumption (Tmote Sky) and also to evaluate how the solution behaves in terms of time delay (that is related to the practicality of the solution) as the number of connected nodes increases, for typical higher performance but still constrained motes (such as the Zolertia ReMote).

The Contiki Cooja simulator provides a very good simulation solution since it supports assembly level system simulation for the TI MSP430 processor that was adopted in our experiments. More specifically, we studied a Contiki 3 OS implementation of the Tmote Sky embedded mote featuring a TI MSP430 16 b CPU with 48 KB of Flash memory, 10 KB of RAM, and a TI CC2420 2.4 GHz IEEE 802.15.4 RF interface. On this processor, we were interested in evaluating the power consumed: at the ON and LPM CPU states during the MC operations, as well as at the CC2420 wireless communication subsystem during transmit (Tx), receive (Rx), and idle states. The Energest module is supported for the specific platform and it was selected to be used in order to collect energy consumption measurements during Cooja simulations. By using current consumption data from the manufacturer datasheet of each of the system on chips (SoCs) under test for every state, we collected measurements through Energest on the time spent in each state and were able to calculate the energy spent on each state using the formula $E(j) = t_{\text{state}} \cdot V_{\text{in}} \cdot I_{\text{state}}$, where t_{state} represents the time spent in a state and V_{in} and I_{state} are the supply voltage and current, respectively, which are selected as reported in each SoC's datasheet.

The networking testing platform that was used is generated through a number of initializing Julia functions that create different WSN topologies with varying characteristics, exporting these concurrently to the appropriate graph objects in Julia, as well as to automatically generated Cooja simulation (.csc) and C header (.h) files. These files are compiled and linked with the C code implementation of the proposed method and the networking code for the data exchange in the WSN.

A use case example of such a WSN topology is presented in Fig. 5, where the generated graph plot in Julia at the left has been transferred to Cooja, with a one-to-one mapping of node coordinates, relationships, and initial conditions. Nodes exchange obfuscated data packets, according to the graph links, representing application level relationships, and broadcast their limited correlation knowledge under the constraints of a noisy radio environment that allows only 1% of data to reach the furthest nodes. This results into a sparse WSN with few connections. Thus, there are several missing entries for the correlation matrix that is formulated at each node.

Regarding the Cooja simulation itself per mote, in terms of CPU and radio power consumption, Fig. 3 depicts these values in a specific mote (with 16 collaborating WSN nodes) as they evolve in time, during the operation of the network. In the figure, there are observable peak values of energy consumption for radio

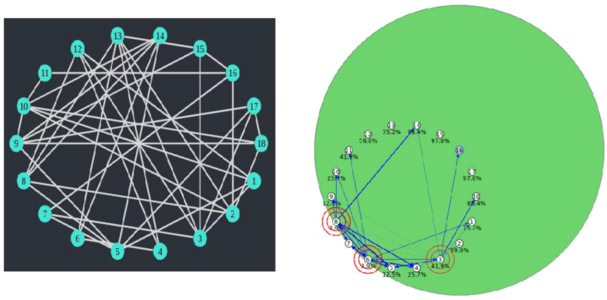


Fig. 5. WSN topology example. The subfigure on the right presents a simulation example, where each percentage assigned to the corresponding node represents the successful packet reception rate from a node that is selected by the user. In this case node, 16 located at the center of the green circle is the sender, whereas the green circle highlights the Tx range for this specific transmitting node.

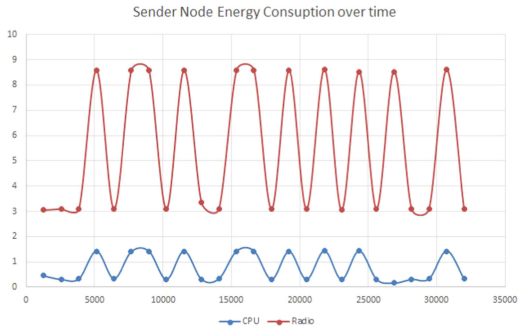


Fig. 6. CPU and radio power consumption evolution (mW) in time (ms).

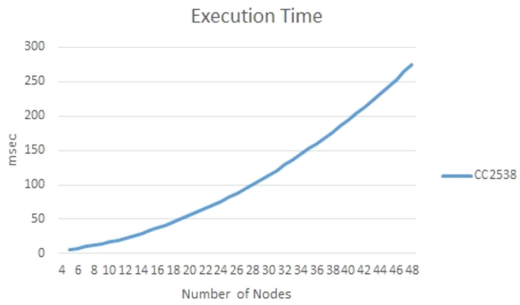


Fig. 7. Processing requirements in real constraint hardware (TI's CC2538 in Zolertia ReMote).

transmission and CPU energy consumption, whenever there is a receipt of a new obfuscated measurement by a collaborating node followed by a trigger of a new MC function and then a broadcast transmission of a new obfuscated measurement (collected from the mote's sensors). It can be remarked that the energy consumption footprint of the radio transactions is considerably higher from the CPU energy consumption due to the MC operation.

In the isolated, core processing time delay evaluation (that complements the CPU-only related power consumption of Fig. 6), Fig. 7 presents the CPU time measurements of the main MC routine, as a function of the nodes number, for a representative constrained node CPU, namely the Texas Instruments CC2538,[1] as this is utilized by the Zolertia ReMote. As it can be observed from the figure, the execution time of the MC function falls under practically affordable values for typically expected node populations.

[1][Online]. Available: http://www.ti.com/product/CC2538

V. Conclusion

In this article, we considered the problem of privacy preservation in IIoT platforms where the raw data measurements are obfuscated with additive noise with statistical properties similar to the statistical properties of the data. To achieve that, the measurement correlation matrix has been decomposed into a time-sequence of rank-one matrices, which is partially known at each node. Thus, a rank one correlation completion problem is solved at each node via a novel low-complexity technique. The proposed approach converges in a number of iterations to the full-rank data correlation matrix, which is later used for data obfuscation. The complexity cost of the proposed privacy preserving algorithm (ideally suited for correlated data streams) can be linear over the number of the sensor nodes as the density of the network increases. Extensive evaluation studies in the Contiki Cooja WSN simulator validate the robustness and efficiency of our approach in terms of both computational complexity and privacy preservation on constrained IIoT devices.

References

[1] E. Sisinni, A. Saifullah, S. Han, U. Jennehag, and M. Gidlund, "Industrial Internet of Things: Challenges, opportunities, and directions," *IEEE Trans. Ind. Informat.*, vol. 14, no. 11, pp. 4724–4734, Nov. 2018.

[2] R. Schmidt, M. Möhring, R.-C. Härting, C. Reichstein, P. Neumaier, and P. Jozinović, "Industry 4.0-potentials for creating smart products: Empirical research results," in *Proc. Int. Conf. Bus. Inf. Syst.*, 2015, pp. 16–27.

[3] Z. Sheng, C. Mahapatra, C. Zhu, and V. C. Leung, "Recent advances in industrial wireless sensor networks toward efficient management in IoT," *IEEE Access*, vol. 3, pp. 622–637, 2015.

[4] A. Z.-k. K. Q. Fung and C. Hardwick, "Smart cities and sustainability initiative," Amer. Planning Assoc., Chicago, IL, USA, Tech. Rep., 2015. [Online]. Available: https://planning-org-uploaded-media.s3.amazonaws.com/legacy_resources/leadership/agendas/2015/spr/pdf/SmartCitiesSustainabilityFinal.pdf

[5] P. Paillier, "Public-key cryptosystems based on composite degree residuosity classes," in *Proc. Int. Conf. Theory Appl. Cryptographic Techn.*, 1999, pp. 223–238.

[6] T. Graepel, K. Lauter, and M. Naehrig, "ML confidential: Machine learning on encrypted data," in *Proc. Int. Conf. Inf. Secur. Cryptology*, 2012, pp. 1–21.

[7] L. J. Aslett, P. M. Esperança, and C. C. Holmes, "Encrypted statistical machine learning: New privacy preserving methods," 2015, *arXiv:1508.06845*.

[8] V. Nikolaenko, U. Weinsberg, S. Ioannidis, M. Joye, D. Boneh, and N. Taft, "Privacy-preserving ridge regression on hundreds of millions of records," in *Proc. IEEE Symp. Secur. Privacy*, 2013, pp. 334–348.

[9] Y. Lindell, "Secure multiparty computation for privacy preserving data mining," in *Proc. Encyclopedia Data Warehousing Mining*, 2005, pp. 1005–1009.

[10] D. Kifer and J. Gehrke, "Injecting utility into anonymized datasets," in *Proc. ACM SIGMOD Int. Conf. Manag. Data*, 2006, pp. 217–228.

[11] Z. Huang, W. Du, and B. Chen, "Deriving private information from randomized data," in *Proc. ACM SIGMOD Int. Conf. Manag. Data*, 2005, pp. 37–48.

[12] F. Li, J. Sun, S. Papadimitriou, G. A. Mihaila, and I. Stanoi, "Hiding in the crowd: Privacy preservation on evolving streams through correlation tracking," in *Proc. IEEE 23rd Int. Conf. Data Eng.*, 2007, pp. 686–695.

[13] R. Mendes and J. P. Vilela, "Privacy-preserving data mining: Methods, metrics, and applications," *IEEE Access*, vol. 5, pp. 10562–10582, 2017.

[14] Z. Xiao, X. Fu, and R. S. M. Goh, "Data privacy-preserving automation architecture for industrial data exchange in smart cities," *IEEE Trans. Ind. Informat.*, vol. 14, no. 6, pp. 2780–2791, Jun. 2018.

[15] R. Agrawal and R. Srikant, "Privacy-preserving data mining," *SIGMOD Rec.*, vol. 29, no. 2, pp. 439–450, May 2000. [Online]. Available: http://doi.acm.org/10.1145/335191.335438
[16] H. Kargupta, S. Datta, Q. Wang, and K. Sivakumar, "On the privacy preserving properties of random data perturbation techniques," in *Proc. 3rd IEEE Int. Conf. Data Mining*, 2003, p. 99. [Online]. Available: http://dl.acm.org/citation.cfm?id=951949.952160
[17] E. Candès and B. Recht, "Exact matrix completion via convex optimization," *Found. Comput. Math.*, vol. 9, no. 6, pp. 717–772, 2009.
[18] J.-F. Cai, E. J. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM J. Optim.*, vol. 20, no. 4, pp. 1956–1982, 2010.
[19] E. Oja and J. Karhunen, "On stochastic approximation of the eigenvectors and eigenvalues of the expectation of a random matrix," *J. Math. Anal. Appl.*, vol. 106, pp. 69–84, 1985.
[20] J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah, "Julia: A fresh approach to numerical computing," *SIAM Rev.*, vol. 59, no. 1, pp. 65–98, 2017. [Online]. Available: https://doi.org/10.1137/141000671
[21] A. Dunkels, B. Gronvall, and T. Voigt, "Contiki—A lightweight and flexible operating system for tiny networked sensors," in *Proc. 29th Annu. Int. Conf. Local Comput. Netw.*, 2004, pp. 455–462. [Online]. Available: https://doi.org/10.1109/LCN.2004.38

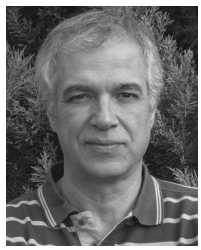
Aris S. Lalos (M'19) received the Diploma, M.A.Sc., and Ph.D. degrees in signal processing and communications from the Computer Engineering and Informatics Department (CEID), School of Engineering (SE), University of Patras (UoP), Rio Patras, Greece, in 2003, 2005, and 2010, respectively.

He was a Research Fellow with the Signal Processing and Communications Laboratory, CEID, SE, UoP, from 2005 to 2010; the Signal Theory and Communications Department, Technical University of Catalonia, Barcelona, Spain, from October 2012 to December 2014; and with the Visualization and Virtual Reality Group from January 2015 to 2018. From October 2011 to October 2012, he was a Telecommunication Research Engineer with Analogies S.A, Patras, Greece, an early stage start up. In May 2018, he was an elected Principal Researcher (Associate Research Professor Level with tenure) with the Industrial Systems Institute, "ATHENA" Research Center. His general research interest include digital communications, adaptive filtering algorithms, geometry processing, wireless body area networks, and biomedical signal processing. He is an author of 94 research papers in international journals (33), conferences (53), and book chapters (4). He received the best demo award in IEEE CAMAD 2014, the Best Paper Award in IEEE ISSPIT 2015, the World's FIRST 10K Best Paper Award in IEEE ICME 2017 in January 2015, he was nominated as Exemplary Reviewer for the IEEE Communications Letters.

Evangelos Vlachos received the Diploma degree in computer engineering and informatics, the M.Sc. degree in signal processing and telecommunications, and the Ph.D. degree in signal processing for wireless communications from the University of Patras (UoP), Patras, Greece, in 2005, 2009, and 2015, respectively.

From 2015 to 2016, he was a Postdoctoral Researcher with the Laboratory of Signal Processing and Telecommunications, UoP, in Computer Engineering and Informatics, working on distributed signal processing over networks. During 2016, he was a Postdoctoral Researcher with the Visualization and Virtual Reality Group, UoP, on graph signal processing. From 2017 to 2019, he was a Postdoctoral Researcher in Signal Processing for Communications. He is currently with the Institute for Digital Communications, The University of Edinburgh, Edinburgh, U.K. In 2019, he was an elected Research Associate with the Industrial Systems Institute, "ATHENA" Research Centre, Marousi, Greece. His general research interest include wireless communications, machine learning and optimization, adaptive control, and filtering algorithms.

Dr. Vlachos was the recipient of the Best Paper Award from the IEEE International Conference on Multimedia and Expo in 2017.



Kostas Berberidis (S'87–M'90–SM'07) received the Diploma degree in electrical engineering from the Democritus University of Thrace, Komotini, Greece, in 1985, and the Ph.D. degree in signal processing and communications from the University of Patras, Patras, Greece, in 1990.

During 1991, he was with the Signal Processing Laboratory, National Defense Research Center. From 1992 to 1994 and from 1996 to 1997, he was a Researcher with the Computer Technology Institute, Patras. From 1994 to 1995, he was a Postdoctoral Fellow with the Centre Commun d'Études de Télévision et Télécommunications, Centre National d'Études des Télécommunications, Rennes, France. Since December 1997, he has been with the Computer Engineering and Informatics Department, University of Patras, where he is currently a Professor and Head of the Signal Processing and Communications Laboratory. Also, since 2008, he has been the Director of the Signal Processing & Communications Research Unit, Computer Technology Institute and Press "Diophantus," Patras. His research interests include distributed signal and information processing and learning, adaptive signal processing, signal processing for communications, wireless communications and sensor networks, array signal processing, smart grid, etc.

Apostolos P. Fournaris (M'09) received the Ph.D. degree in public key cryptography systems design from the Electrical and Computer Engineering Department, University of Patras, Patras, Greece, in 2008.

He was with the Sophia Antipolis Hitachi Europe SAS R-D Centre for two years as a Researcher. He is currently a Principal Researcher (Research Associate Professor) with the Industrial Systems Institute, Research Center ATHENA, Marousi, Greece. He was also a Visiting Sessional Lecturer with Monash University, Melbourne, VIC, Australia, where he lectured courses on cryptography, computer, and network security. He is the author of more than 80 research papers. His research interests include asymmetric cryptography, side channel attacks and analysis, hardware attack resistance, security–privacy, and trust in IoT systems.

Dr. Fournaris is a Member of the International Association for Cryptologic Research, IEEE Computer Society, and IEEE Circuits and System Society.

Christos Koulamas (M'06–SM'18) received the Engineering Degree in computer engineering and informatics and the Ph.D. degree in electrical and computer engineering from the University of Patras, Patras, Greece, in 1994 and 2004, respectively.

He is currently a Research Director with the Industrial Systems Institute, ATHENA Research and Innovation Center, Marousi, Greece. He has more than 25 years of experience in R&D, in the areas of real-time distributed embedded systems, industrial networks, wireless sensor networks, and their applications in various industry sectors. He is the author or coauthor of more than 70 scientific publications in international journals, conferences, and books, and more than 110 publicly available or confidential technical reports. His current research interests include cyber–physical systems and industrial Internet of Things technologies, and their applications in smart environments.

Dr. Koulamas has served as a Guest Editor for the MDPI *Electronics* and *Sensors* Journals, as a PC/TPC member in many IEEE and other conferences and workshops, and as a member of the board of European Research Consortium for Informatics and Mathematics.