Crowd-Based Learning of Spatial Fields for the Internet of Things: From Harvesting of Data to Inference

Published Version:
Arias-De-Reyna, E., Closas, P., Dardari, D., Djuric, P.M. (2018). Crowd-Based Learning of Spatial Fields for the Internet of Things: From Harvesting of Data to Inference. IEEE SIGNAL PROCESSING MAGAZINE, 35(5), 130-139 [10.1109/MSP.2018.2840156].

Availability:
This version is available at: https://hdl.handle.net/11585/650801 since: 2018-11-24

Published:
DOI: http://doi.org/10.1109/MSP.2018.2840156

This is the post peer-review accepted manuscript of:

Arias-de-Reyna, E., Closas, P., Dardari, D., & Djuric, P. M. (2018). Crowd-Based Learning of Spatial Fields for the Internet of Things: From Harvesting of Data to Inference. *IEEE Signal Processing Magazine*, *35*(5), 130–139. https://doi.org/10.1109/MSP.2018.2840156

The published version is available online at: https://doi.org/10.1109/MSP.2018.2840156

Crowd-based Learning of Spatial Fields for the IoT

Eva Arias-de-Reyna, *Senior, IEEE*, Pau Closas, *Senior, IEEE*,

Davide Dardari, *Senior, IEEE*, and Petar M. Djurić, *Fellow, IEEE*

Index Terms

Crowdsensing, data fusion, Internet-of-Things, localization and mapping.

E. Arias-de-Reyna is with the Department of Signal Theory and Communications, Universidad de Sevilla, Camino de los Descubrimientos s/n, 41092 Sevilla (Spain). (e-mail: earias@us.es)

P. Closas is with the Department of Electrical and Computer Engineering at Northeastern University, 360 Huntington Avenue, Boston, MA 02115 (USA). (e-mail: closas@northeastern.edu)

D. Dardari is with the Department of Electrical, Electronic and Information Engineering (DEI) "Guglielmo Marconi" and CNIT, University of Bologna, Via Venezia 52, I-47521 Cesena (FC), Italy. (e-mail: davide.dardari@unibo.it)

P. M. Djurić is with the Department of Electrical and Computer Engineering at Stony Brook University, Stony Brook, NY 11794 (USA). (e-mail: petar.djuric@stonybrook.edu)

Crowd-based Learning of Spatial Fields for the IoT

I. Introduction

The knowledge of spatial distributions of physical quantities such as radio-frequency (RF) interference, pollution, geomagnetic field magnitude, temperature, humidity, audio and light intensity, will foster the development of new context-aware applications. For example, knowing the distribution of RF interference might significantly improve cognitive radio systems [1], [2]. Similarly, knowing the spatial variations of the geomagnetic field could support autonomous navigation of robots (including drones) in factories and/or hazardous scenarios [3]. Other examples are related to the estimation of temperature gradients, detection of sources of RF signals or percentages of certain chemical components. As a result, people could get personalized health-related information based on their exposure to sources of risks (e.g., chemical or pollution). We refer to these spatial distributions of physical quantities as *spatial fields*. All of the above examples have in common that learning the spatial fields requires a large number of sensors (agents) surveying the area [4], [5].

A common way to sense environmental variables is the deployment of dedicated wireless sensor networks (WSNs), which continue to stimulate fertile research activities in the scientific community. Typical WSN applications are oriented to sense specific physical quantities (e.g., temperature) in well-defined areas [6], [7]. Unfortunately WSNs are generally characterized by significant constraints in terms of deployment cost, energy limitation and need for maintenance. These constraints prevent them from becoming scalable and therefore from being the ultimate solution for automated and distributed sensing of the physical world.

The expected pervasive diffusion of Internet-of-things (IoT) devices (fixed and mobile) opens up a unique opportunity for a wide and massive sensing and mapping (that is, geo-referenciation of physical quantities). In fact, the IoT constitutes a paradigm where a multitude of heterogeneous devices is able to sense the environment, process data, and actuate, thus creating the necessary infrastructure for cyber-physical systems. This infrastructure galvanizes technologies such as smart grids, smart homes, smart cities, and intelligent transportation [8], [9].

From the wide variety of applications of the IoT, we are interested in those which will benefit from having a large spatial coverage of an area due to a large number of agents navigating through it. For instance, this can be the case when devices are carried by people or autonomous agents (e.g., vehicles, robots or drones) moving in outdoor and indoor populated environments like malls, stadiums, or crowded buildings. One can even imagine cities at large, if one considers much larger settings in size. Thanks to the widespread diffusion of IoT devices with heterogeneous sensors, the estimation of spatial physical fields is creating a new trend for next generation sensor networks, referred to as mobile crowd sensing networks [10]–[13]. This is basically a zero-effort approach to automatically collect and process data. Recently, as an example, this concept has been proposed for zero-effort automatic mapping of environmental features using sensors already embedded in smartphones, such as magnetometers and Wi-Fi [14]–[17]. In such settings, the contribution of the agent to the sensing process is as simple as while moving around, carrying the personal device in the pocket. Individuals are not even requested to be participatory,[1] as the sensing process could run in the background during the normal operation of the device. Thus, the sensing process is not an exclusive task, and it arises from the dynamic reality of humans or autonomous agents. The sensing process is a result of piggybacking on the capabilities of today's and future wireless personal devices. Including data generated by these devices will dramatically increase the amount of the data for sensing and mapping purposes, with obvious benefits in terms of the resulting accuracy.

In this context, the IoT is the technological enabler for crowdsensing and learning of spatial fields. Interestingly, IoT devices are in general able to communicate among themselves, either directly or through a fusion node that can potentially be in the cloud. Thanks to the communication capabilities, empirical data gathered by mobile agents (the crowd) can be collected and processed by learning algorithms located in the cloud. These algorithms exploit the correspondence between the position and the value of the physical quantity measured in that position to estimate the spatial field. As a consequence, positioning and spatial field estimation are intimately twinned as it will be illustrated later.

Crowdsourcing-based learning methods rely on the "experience" gained by previous agents. In principle, it is possible that with crowd-based learning one can perform optimal information fusion [10]. On the other hand, moving from the well controlled conditions of WSNs scenarios, where nodes are deployed in ad hoc known locations, to crowdsensing settings, where agents move around in an uncontrolled manner, entails a number of issues that need

[1]"Participatory" here means that the users are not requested to follow particular paths in order to make the learning process effective; however, they know that a sensing task is running in the background.

to be addressed. The methods rely on sharing through cloud mechanisms [18], but they can be of practical relevance in IoT applications only if their computational and memory requirements do not grow with the amount of collected data. Therefore, novel methodologies for multi-sensor data fusion and information processing are needed. They should guarantee efficient statistical representation of spatial fields and a computational complexity that does not depend on the number of measurements. Further, the algorithms need to be robust against irregular positioning and measurement errors.

This article addresses the challenges and solutions of learning spatial fields for the IoT whose multitude of connected devices sense the fields. In many real-world scenarios, one may have measurements acquired by thousands of people or autonomous mobile agents, interacting with each other and with things. The underlying idea is that each user takes advantage of the measurements acquired by previous users and, in turn, contributes to further improvement of the field estimates, which amounts to an indirect cooperative approach. More specifically, the article analyzes the main issues, techniques, and architectures for efficient crowd-based learning of spatial fields in the IoT. The nature of the problem suggests searching for solutions within the Bayesian framework, and this is what we have adopted in the paper. It is clear, however, that one may apply other methods including various types of data-driven or other non-Bayesian methods.

As already pointed out, an even more challenging application of this concept is the joint positioning and spatial field learning in indoor environments, where agents aim at self-localization and, at the same time, learn the position-dependent parameters of the underlying observation models (represented as spatial fields). We put particular emphasis on this topic to show the great potential of crowd-based learning approaches.

In particular, in Section 2, we provide details of inference methods that allow for learning spatial fields. Namely, in the absence of specific and accurate models for the fields, one approach is to assume that the field is a sample from a Gaussian process (GP). In this article we are interested in GPs and, particularly, in their representations through linear combinations of orthogonal basis functions. This approach enables the development of inference algorithms whose complexity does not grow with the number of observations, which is necessary in the context of crowd-based learning.

In Section 3, we discuss the issue of assuming perfect knowledge of the position where the data are sensed. This is almost never the case since the devices are typically positioned through the Global Navigation Satellite System (GNSS) or some other technology [19], and they provide position estimates with errors. It is therefore crucial to account for them in the process of learning spatial fields. Otherwise, the obtained results will be unreliable and/or inaccurate. Two examples of joint tracking and crowd-based learning methods are discussed.

Section 4 illustrates a use case related to the problem of indoor localization and tracking in the presence of biased ranging measurements caused by non-line-of-sight (NLOS) channel conditions, typical in time-based positioning systems, such as those based on the ultrawide-band (UWB) technology [20]. NLOS conditions might affect the position-dependent parameters of the observation model of the tracking algorithm and, thus, are treated as spatial fields to be estimated jointly with the agent position. Agents entering the area of interest take advantage of the knowledge acquired by previous agents so that the expected tracking performance improves as the number of agents participating in the crowdsourcing grows, as shown in our experiments.

An outlook of future directions of research is provided with the conclusions of the paper.

II. Inference Methods for Learning Spatial Fields

Suppose we want to estimate a static spatial field, which we denote with $f(\mathbf{x})$, where $\mathbf{x}$ contains location coordinates (the method can be modified to allow for the estimation of time-varying spatial fields). We make the initial estimate from T noisy observations $y_t = y(\mathbf{x}_t)$, acquired at known locations $\mathbf{x}_t$, $t = 1, 2, \ldots, T$. The set of these observations and locations, $\mathcal{D} = \{(y_t, \mathbf{x}_t) \mid t = 1, 2, \ldots, T\}$, represents a training set from which initial information about the spatial field $f(\mathbf{x})$ can be extracted. In particular, the objective is to make inference about the spatial field at locations without observations, i.e., not included in $\mathcal{D}$, and the corresponding uncertainty of the estimates. In the absence of specific and accurate models of the field, one approach is to assume that the field is a sample from a GP. We note that GPs represent a powerful and widely adopted machine learning methodology [21].

The resulting non-parametric regression problem has a well-known solution, which unfortunately suffers from a computational complexity, of the order of $O(T^3)$ [21]. It goes without saying that this solution is unaffordable in crowd sensing scenarios where T could grow to huge values. Several methods have been proposed to overcome this issue. For regular grids, fast solutions can be obtained through FFT-based approaches [22] or by approximatively

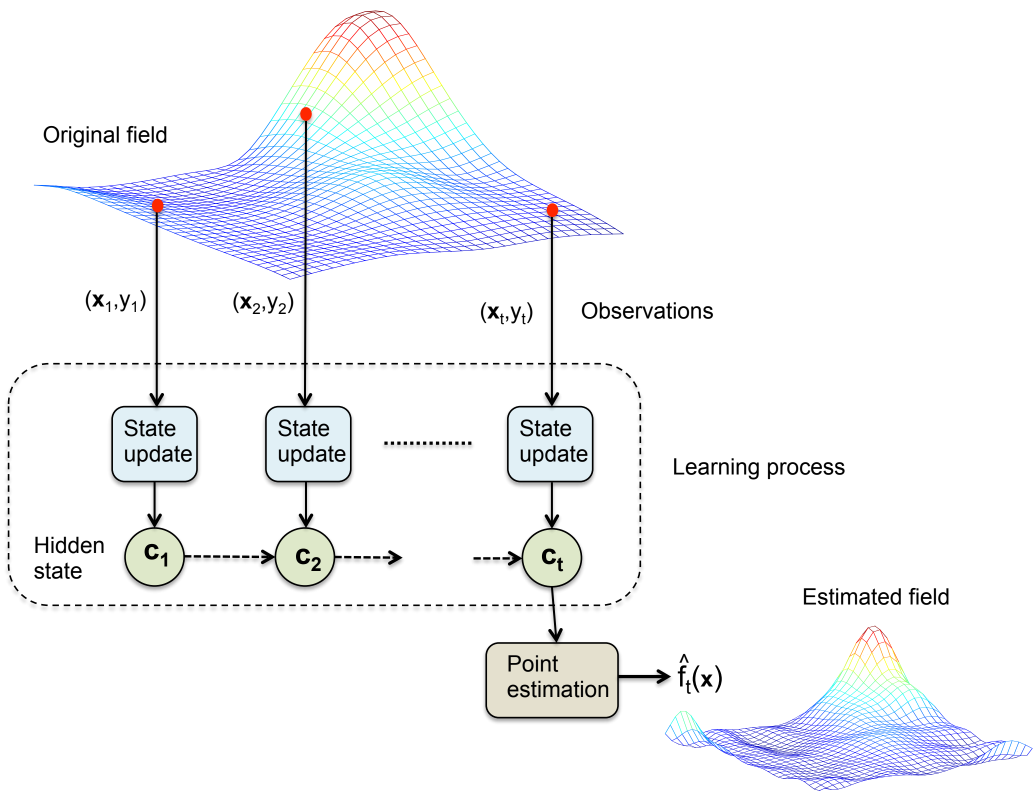

Fig. 1. Hidden Markov process applied to crowd-based learning. The parameters $\mathbf{c}_t$ capture information about the spatial field $f(\mathbf{x})$ at any location $\mathbf{x}$ available at time t, and $\hat{f}_t(\mathbf{x})$ is the estimate of the field at $\mathbf{x}$ after t measurements.

describing the GP through state-space models making the complexity independent of T under certain conditions [23]. Other methods are referenced in [24].

Most of the proposed solutions, however, are not computationally and memory efficient when applied to crowd-based learning, where all observations are not available simultaneously and might grow fast. Moreover, observations are usually obtained at random locations because the agents are generally not participatory, which entails that grid-based approaches cannot be applied in this context. On the other hand, we know that low-complexity incremental methods that update the field estimate once a new observation is acquired are preferable.

To tackle these issues, the authors in [25] and [26] propose a combined GP-state space method, whose complexity and memory requirements do not depend on the number of observations, in static and dynamic scenarios, respectively. This allows an efficient statistical characterization of the spatial field, which can be easily updated once new data become available, thus making it well-suited for crowd-based learning applications. The main idea is represented in Fig. 1 and summarized in the following.

Given an appropriate set of 2D orthogonal basis functions (e.g., the 2D Fourier transform) in the area of interest, $\boldsymbol{\psi}(\mathbf{x}) = [\psi_1(\mathbf{x})\, \psi_2(\mathbf{x}), \ldots, \psi_J(\mathbf{x})]^\top$, the spatial field $f(\mathbf{x})$ is modeled by

$$f(\mathbf{x}) \quad = \quad \boldsymbol{\psi}^\top(\mathbf{x})\, \mathbf{c}, \tag{1}$$

where $\mathbf{c}$ is a $J \times 1$ Gaussian vector of coefficients with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. Thus, the GP is described by

$$f(\mathbf{x}) \quad \sim \quad \mathcal{GP}\left(m(\mathbf{x}), \kappa\left(\mathbf{x}, \mathbf{x}'\right)\right), \tag{2}$$

where $m(\mathbf{x}) = \boldsymbol{\psi}^\top(\mathbf{x})\boldsymbol{\mu}$ and $\kappa\left(\mathbf{x}, \mathbf{x}'\right) = \boldsymbol{\psi}^\top(\mathbf{x})\boldsymbol{\Sigma}\, \boldsymbol{\psi}(\mathbf{x}')$ are its mean and covariance, respectively.

The set of random coefficients $\mathbf{c}$ can be thought of as the state of a state-space model described by a hidden Markov process. The size of $\mathbf{c}$, J (dimension of the state-space), depends on the spatial variability of the physical field (spatial bandwidth). Thanks to (1), the problem of representing and estimating the spatial field $f(\mathbf{x})$ translates to characterizing the vector of coefficients $\mathbf{c}$, which does not depend on $\mathbf{x}$ and does not increase in size with the number of observations. Due to the Gaussian hypothesis, $\mathbf{c}$ is completely statistically characterized by the mean $\boldsymbol{\mu}$ and the covariance matrix $\boldsymbol{\Sigma}$.

The key to the crowd-based learning idea is to use all the observed data collected until time t by all past agents for updating $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. We denote the conditional vector of coefficients and its mean and covariance at time t by $\mathbf{c}_t$, $\boldsymbol{\mu}_t$ and $\boldsymbol{\Sigma}_t$, respectively. Note that at time $t = 0$, $\boldsymbol{\mu}_0$ and $\boldsymbol{\Sigma}_0$ represent the a priori statistical knowledge about the GP. Each time when a new noisy observation y_t is acquired, e.g., at position $\mathbf{x}_t$, it is used to update the characterization of the GP, $f(\mathbf{x})$, to $\hat{f}_t(\mathbf{x})$ given the observations $y_{1:t} = \{y_1, y_2, \ldots, y_t\}$ by using the model (1)-(2). This can be accomplished by properly updating $\boldsymbol{\mu}_t$ and $\boldsymbol{\Sigma}_t$ of $\mathbf{c}_t$ from the previous conditional mean and covariance $\boldsymbol{\mu}_{t-1}$ and $\boldsymbol{\Sigma}_{t-1}$. Such an iterative *learning process* can be expressed in general as

$$\left(\boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t\right) = \mathcal{L}\left[\left(\boldsymbol{\mu}_{t-1}, \boldsymbol{\Sigma}_{t-1}\right), y_t, \mathbf{x}_t\right], \tag{3}$$

where $\mathcal{L}[\cdot,\cdot,\cdot]$ represents the learning algorithm which updates $\left(\boldsymbol{\mu}_{t-1}, \boldsymbol{\Sigma}_{t-1}\right)$ to $\left(\boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t\right)$ by considering the latest

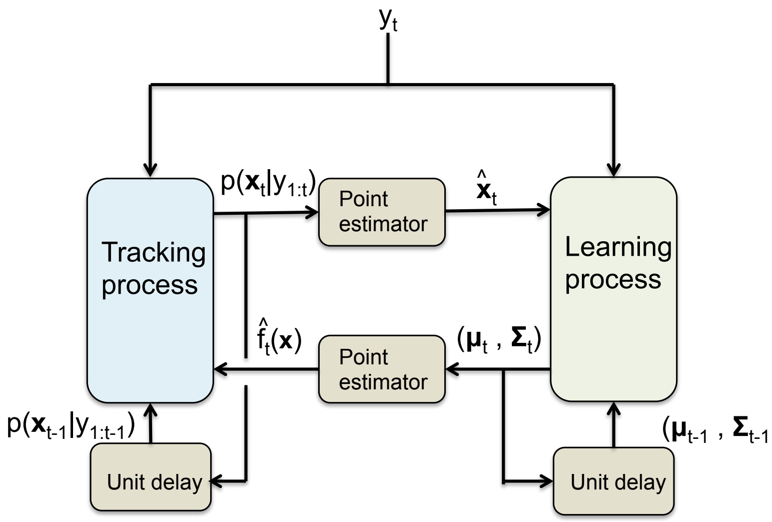

Fig. 2. A schematic illustration of the loose coupling method.

observation y_t and the position $\mathbf{x}_t$ at which it was collected. For instance, under the hypothesis of Gaussian observation noise and known $\mathbf{x}_t$, all the involved random variables (RVs) are Gaussian and the transformation (1) from the state vector to the spatial field is linear. Then the evolution of $\boldsymbol{\mu}_t$ and $\boldsymbol{\Sigma}_t$ in (3) can be computed efficiently through a Kalman filter algorithm with a complexity independent of the number of observations.

At any time t, one can compute a point estimate $\hat{f}_t(\mathbf{x})$ and/or the confidence interval of the GP at a certain position $\mathbf{x}$ conditioned on the history of observations by evaluating the maximum a posteriori (MAP) or minimum mean-square error (MMSE) estimate $\hat{\mathbf{c}}_t = \boldsymbol{\mu}_t$ of the coefficient vector $\mathbf{c}_t$ at time t and applying the transformation (1).[2] Therefore, the updated mean and covariance $(\boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t)$ represent the knowledge (sufficient statistics) about the GP acquired up to the current observation instant t. In this way it is not necessary to keep in the memory all the past observations whose number could grow to huge values. Instead, it suffices to store only $\boldsymbol{\mu}_t$ and $\boldsymbol{\Sigma}_t$ whose sizes depend on J only.

III. Sensing and Positioning

A common assumption in crowdsensing is that the agents sense the environment at perfectly known locations [5]. However, in the absence of this knowledge their positions are typically estimated – through GNSS or other technologies [27] – and thus, they are somewhat uncertain. Moreover, other sources of error might be present such as the relative location between the user's centroid and the mobile terminal. It is therefore crucial to account for

[2]We denote the estimate of the vector $\mathbf{c}_t$ with $\hat{\mathbf{c}}_t$ to convey that the estimate is made at time t using $y_{1:t}$.

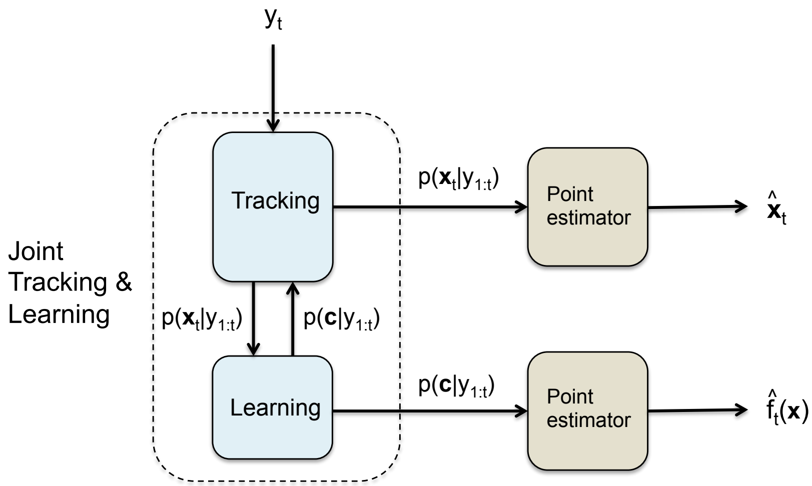

Fig. 3. A schematic illustration of the tight coupling method.

this uncertainty in the process of learning the spatial fields; otherwise the obtained results will be unreliable and/or inaccurate. Localization techniques such as multi-lateration, fingerprinting, sources of opportunity, etc., rely on the availability of position-dependent physical measurements (e.g., time-of-arrival (TOA), received signal strength indicator (RSSI)) from which observations are derived. Regardless of the adopted localization technology, the main performance limitation is often imposed by model mismatches, i.e., the discrepancy between the reality and the applied models for characterizing the observations used in the localization process. Even if a model is accurate, the model parameters might depend on the agents' positions and hence they can be treated as spatial fields to be estimated. In turn, estimates of the locations depend on the model parameters, i.e., on the (unknown) spatial fields. Such a "chicken and egg" problem is challenging and can be tackled only through joint crowd-based learning and localization methods.

In the following we illustrate two possible approaches to estimate jointly the position and the field(s) characterizing the observation model parameter(s). For both approaches we suppose that a central unit, or more generally the cloud, keeps receiving the observations y_t (e.g., distance estimates) from all the agents in the space of interest. Based on these observations, this central unit has to estimate the position $\mathbf{x}_t$ of each user and to update the estimate of the coefficients $\mathbf{c}_t$ for the spatial field of interest.

A. Loose coupling approach

The first approach, which we refer to as "loose coupling," was originally proposed in [28] and extended in [29]. According to the approach, each of the unknowns $\mathbf{x}_t$ and $\mathbf{c}_t$ is estimated by its own method. The two methods communicate with each other by exchanging sequentially their respective estimates (see Fig. 2).

Specifically, at time t the estimate of an agent's position $\mathbf{x}_t$ is derived starting from the incoming position-dependent observation of that agent, y_t, as well as all the previous measurements through a *Tracking algorithm:*[3]

$$p\left(\mathbf{x}_t|y_{1:t}\right) = \mathcal{P}\left[p\left(\mathbf{x}_{t-1}|y_{1:t-1}\right), y_t, \hat{f}_{t-1}(\cdot)\right] \;, \tag{4}$$

where $p\left(\mathbf{x}_t|y_{1:t}\right)$ is the a posteriori probability density function (pdf) of $\mathbf{x}_t$, $\mathcal{P}[\cdot,\cdot,\cdot]$ denotes any iterative positioning/tracking algorithm able to provide the a posteriori pdf at time t as a function of the a posteriori pdf at time $t-1$, the new observations y_t and the estimate of the spatial field(s) $\hat{f}_{t-1}(\cdot)$,[4] where the spatial field is a part of the observation model. Such algorithms can be derived using well-known Bayesian filtering tools that also allow for mobility models. From $p\left(\mathbf{x}_t|y_{1:t}\right)$, an estimate $\hat{\mathbf{x}}_t$ of the position can easily be obtained, e.g., MAP, MMSE, or any other point estimate. More details can be found in [27].

The updated position estimate is then used as the input of the *Learning algorithm*:

$$\left(\boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t\right) = \mathcal{L}\left[\left(\boldsymbol{\mu}_{t-1}, \boldsymbol{\Sigma}_{t-1}\right), y_t, \hat{\mathbf{x}}_t\right], \tag{5}$$

where $\mathcal{L}[\cdot,\cdot,\cdot]$, as in (3), updates $\left(\boldsymbol{\mu}_{t-1}, \boldsymbol{\Sigma}_{t-1}\right)$ of $\mathbf{c}_t$ to $\left(\boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t\right)$ by considering the latest observation y_t. Note that here, differently from (3), the estimated position $\hat{\mathbf{x}}_t$ is used instead of the true one, $\mathbf{x}_t$, which is not available. In addition, as it will be detailed in Sec. IV, the observation used to train the Learning process might not be directly y_t, but a function of it and the estimated position $\hat{\mathbf{x}}_t$.

In the following step, the estimated field $\hat{f}_t(\cdot)$, obtained from $\left(\boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t\right)$, is sent to the Tracking algorithm, which obtains the estimate $\hat{\mathbf{x}}_{t+1}$. The algorithms continue to update and exchange estimates as new measurements keep coming.

The main advantage of this method is its relatively easy implementation because the tracking and learning tasks

[3]The symbol y_t represents a scalar, but in general it can be a vector that contains multiple measurements.

[4]Actually, there may be more than one spatial field of interest, as it will be illustrated in Section IV.

are separated. The learning algorithm amounts to a standard recursive least-squares-type method. For tracking, one can use any particular algorithm (particle filtering, extended Kalman filtering, etc.). The choice of filter would depend on various factors including the adopted mobility and observations models (typically non-linear) and the allowed computational complexity of the tracking process.

B. Tight coupling approach

In the "tight coupling" approach, $\mathbf{x}_t$ and $\mathbf{c}$ are estimated by way of integrating out $\mathbf{c}$ while estimating $\mathbf{x}_t$, with $\mathbf{c}$ still being estimated (see Fig. 3). At each recursion, the joint posterior pdf of the unknowns $\mathbf{x}_t$, $p\left(\mathbf{x}_t|y_{1:t}\right)$, also known as filtering pdf, is calculated from the previous joint posterior and the new measurement y_t, or

$$p\left(\mathbf{x}_t|y_{1:t}\right) = \mathcal{T}\left[p\left(\mathbf{x}_{t-1}|y_{1:t-1}\right), y_t\right], \tag{6}$$

where $\mathcal{T}[\cdot,\cdot]$ is an algorithm that performs the computation of the a posteriori pdf of $\mathbf{x}_t$. Unlike in (4), it appears that there is no use of $\mathbf{c}$ in the algorithm. However, this is not the case. More specifically, if we denote the posterior of $\mathbf{c}$ by $p(\mathbf{c}|y_{1:t})$ and according to the usual Markovian assumption for the underlying state-space model [27], the equations that describe how $p\left(\mathbf{x}_t|y_{1:t}\right)$ and $p(\mathbf{c}|y_{1:t})$ are updated are expressed as follows:

$$p\left(\mathbf{x}_t|y_{1:t}\right) \propto \int p(y_t|\mathbf{x}_t, y_{1:t-1})\; p(\mathbf{x}_t|\mathbf{x}_{t-1})p(\mathbf{x}_{t-1}|y_{1:t-1})\,\mathrm{d}\mathbf{x}_{t-1}, \tag{7}$$

$$p(\mathbf{c}|y_{1:t}) \propto \int p(y_t|\mathbf{x}_t,\mathbf{c})p(\mathbf{x}_t|y_{1:t-1},\mathbf{c})p(\mathbf{c}|y_{1:t-1})\,\mathrm{d}\mathbf{x}_t, \tag{8}$$

where $p(\mathbf{x}_t|\mathbf{x}_{t-1})$ is the mobility model. These equations use two pdfs, $p(y_t|\mathbf{x}_t, y_{1:t-1})$ and $p\left(\mathbf{x}_t|y_{1:t-1},\mathbf{c}\right)$, which are defined by

$$p(y_t|\mathbf{x}_t, y_{1:t-1}) = \int p(y_t|\mathbf{x}_t,\mathbf{c})\,p(\mathbf{c}|y_{1:t-1})\,\mathrm{d}\mathbf{c}, \tag{9}$$

$$p\left(\mathbf{x}_t|y_{1:t-1},\mathbf{c}\right) = \int p(\mathbf{x}_t|\mathbf{x}_{t-1})\,p(\mathbf{x}_{t-1}|y_{1:t-1},\mathbf{c})\,\mathrm{d}\mathbf{x}_{t-1}. \tag{10}$$

Finally, the pdf $p(\mathbf{x}_{t-1}|y_{1:t-1},\mathbf{c})$ that appears in (10) is obtained from

$$p(\mathbf{x}_{t-1}|y_{1:t-1},\mathbf{c}) \propto p(y_{t-1}|\mathbf{x}_{t-1},\mathbf{c})p(\mathbf{c}|y_{1:t-2})\int p(\mathbf{x}_{t-1}|\mathbf{x}_{t-2})p(\mathbf{x}_{t-2}|y_{1:t-2})\mathrm{d}\mathbf{x}_{t-2}, \tag{11}$$

where all the necessary pdfs in (11) are known from previous recursions.

Point estimates $\hat{\mathbf{x}}_t$ of the position or estimates of the predicted spatial field $\hat{f}_t(\cdot)$ can easily be obtained from $p\left(\mathbf{x}_t|y_{1:t}\right)$ and $p(\mathbf{c}|y_{1:t-1})$ (and (1)-(2)). As evident from the previous equations, the design of the algorithm $\mathcal{T}[\cdot,\cdot]$ is more complex than that of the loose coupling approach.

IV. Use Case

We discuss a relevant use case in the context of joint crowd-based learning and localization. We consider agents navigating in a certain area whose behavior is modeled as a random walk, which is quite common in the literature in the absence of any other information on user behavior. Each user is recording TOA physical measurements with respect to fixed reference nodes. We call these nodes anchor nodes because we know their locations, and we compute distance estimates with respect to them using TOA measurements. Such estimates could be subjected to unknown bias due to NLOS conditions that might characterize the channel between the agent and the ith anchor node according to the following observation model [20], [27]:

$$y^{[i]} = \left\|\mathbf{x} - \mathbf{x}_{\mathrm{A}}^{[i]}\right\| + f^{[i]}(\mathbf{x}) + \nu^{[i]} \ , \tag{12}$$

where $\mathbf{x}_{\mathrm{A}}^{[i]}$ denotes the (known) position of the ith anchor, $\left\|\mathbf{x} - \mathbf{x}_{\mathrm{A}}^{[i]}\right\|$ is the true distance between anchor i and the agent, $\nu^{[i]}$ is a zero mean Gaussian perturbation, and $f^{[i]}(\mathbf{x})$ represents the spatial field characterizing the spatial behavior of the bias induced by the NLOS condition. Obviously we have one spatial field for each anchor. It is worth highlighting that in loose coupling, the distance observation $y^{[i]}$ cannot be used directly by the Learning process because the latter needs bias (i.e., spatial field) observations that are not available. A possible solution is to derive a 'virtual' bias observation $\tilde{y}^{[i]}$ from (12) using the estimated position $\hat{\mathbf{x}}$ instead of the true one, i.e., $\tilde{y}^{[i]} = y^{[i]} - \left\|\hat{\mathbf{x}} - \mathbf{x}_{\mathrm{A}}^{[i]}\right\|$.

While an agent crosses the area of interest, it takes advantage of the available estimated field obtained from measurements acquired by previous agents. In turn, the estimate of this field is updated by the measurements of this agent. Thereby, subsequent agents can also benefit by using the field for their own localization. In the introduction, we referred to this sharing of information among the users as indirect cooperation.

Now we present some numerical results for the NLOS use case validating the crowd-approach discussed in the paper. This set of representative experiments highlights the benefits of leveraging the crowd to learn the NLOS-induced bias field, ultimately improving the knowledge of the field at a reduced cost and without calibration requirements. We point out that in the case of cooperative calibration with some users navigating along prescribed trajectories, performance could be improved.

We based the simulation of the observation values on real measurements taken in a typical office indoor environment with walls made of concrete, with floor plan shown in Fig. 4. We had four anchors, denoted in the figure as tx_i, $i = \{1, 2, 3, 4\}$ (note that the figure also shows a fifth anchor, but its measurements were not used). The figure shows also a set of 20 test locations where $1,500$ range measurements were taken for each anchor using a commercial UWB radio operating in the 3.2 - 7.4 GHz band. The complete description of the measurement campaign can be found in [20].

For each anchor the following procedure was carried out. For each of the 20 test locations we computed the mean and the standard deviation σ of the range measurements. Then we evaluated the bias as the difference between the mean range observation and the true distance. The value used in the simulations as true bias for any other location was obtained by interpolation from the set of bias values.

It can be observed that most test locations were in NLOS with respect to the anchors, thus making the localization and fields estimation processes quite challenging.

In the simulations, the users were entering in succession and moved randomly within a square area of side $L = 9$ meters in Fig. 4. Each user had a total of 50 measurements during the sojourn in the area, taken at intervals of 1 s. The simulated trajectories followed a random-walk model, with noise covariance matrix equal to $\rho^2 \mathbf{I}$ with $\rho = 0.1$ and where $\mathbf{I}$ is the identity matrix, and the ranging measurements had errors with a standard deviation of $\sigma = 0.1\,\mathrm{m}$, in addition to the bias when in NLOS. Notice that if another technology different from UWB is used to collect $y^{[i]}$, the standard deviation of the measurements should be changed accordingly.

Figure 5(a) shows the corresponding spatial field $f^{[2]}(\mathbf{x})$ for anchor number two. The rest of the illustrations in Fig. 5 correspond to results obtained by the loose coupling approach described in the previous section. As 2D orthogonal basis functions $\boldsymbol{\psi}(\mathbf{x})$ in the Learning process, we used the exponentials of the 2D Fourier series

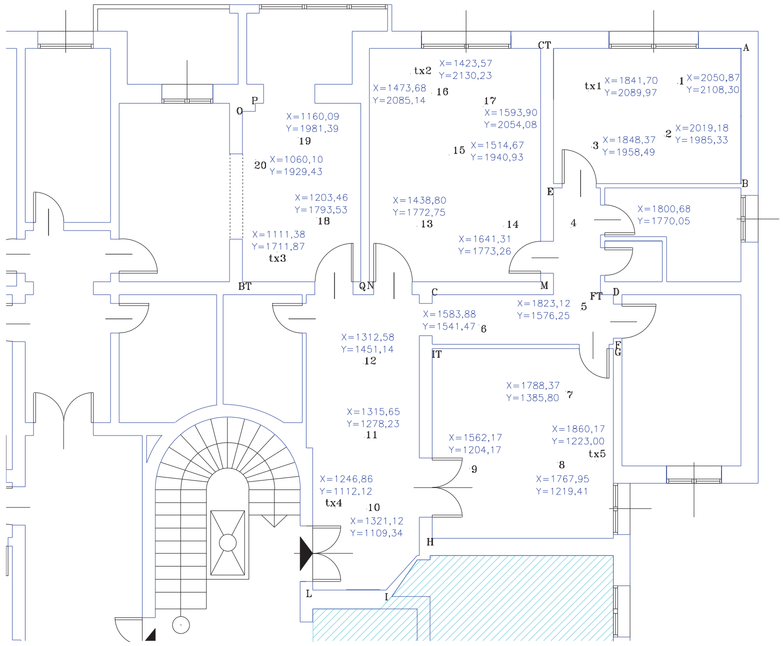

Fig. 4. Floor plan of the office environment considered in the case study, based on the measurement campaign in [20]. There are four anchors of fixed known location and 20 test locations.

expansion of the periodical repetition of $f^{[i]}(\mathbf{x})$ with a period L in each dimension. They have been mapped into the corresponding real and imaginary components and truncated to $J = 578$ terms. Here the 2D Fourier exponentials have been taken as an example. In general the choice of the basis functions is strictly dependent on the application and deserves further research. The Tracking process was carried out by means of particle filtering (PF) with 500 particles.

Figures 5(b)-(d) show the estimated spatial field of anchor two, after 20, 50 and 200 users, respectively. More specifically, the figures show the mean $m(\mathbf{x})$ of the GP representing the spatial field $f^{[2]}(\mathbf{x})$ at these instants. A progressive improvement of the field knowledge can be observed, which is visually apparent from the closer resemblance of the estimated fields to the true field in Fig. 5(a) as the number of users grow.

Figure 6 shows the localization error, namely the distance between the true and estimated locations, as a function of the time step, for the 200th user of the same simulation of Fig. 5. Besides the described technique with crowd-based learning, two more algorithms based on the same PF are included for comparison; one of them assumes always line-of-sight (LOS) condition while the other has perfect knowledge of the range bias value for each agent-anchor pair at each time instant. The latter is unrealistic and used as a benchmark, as it has the best performance,

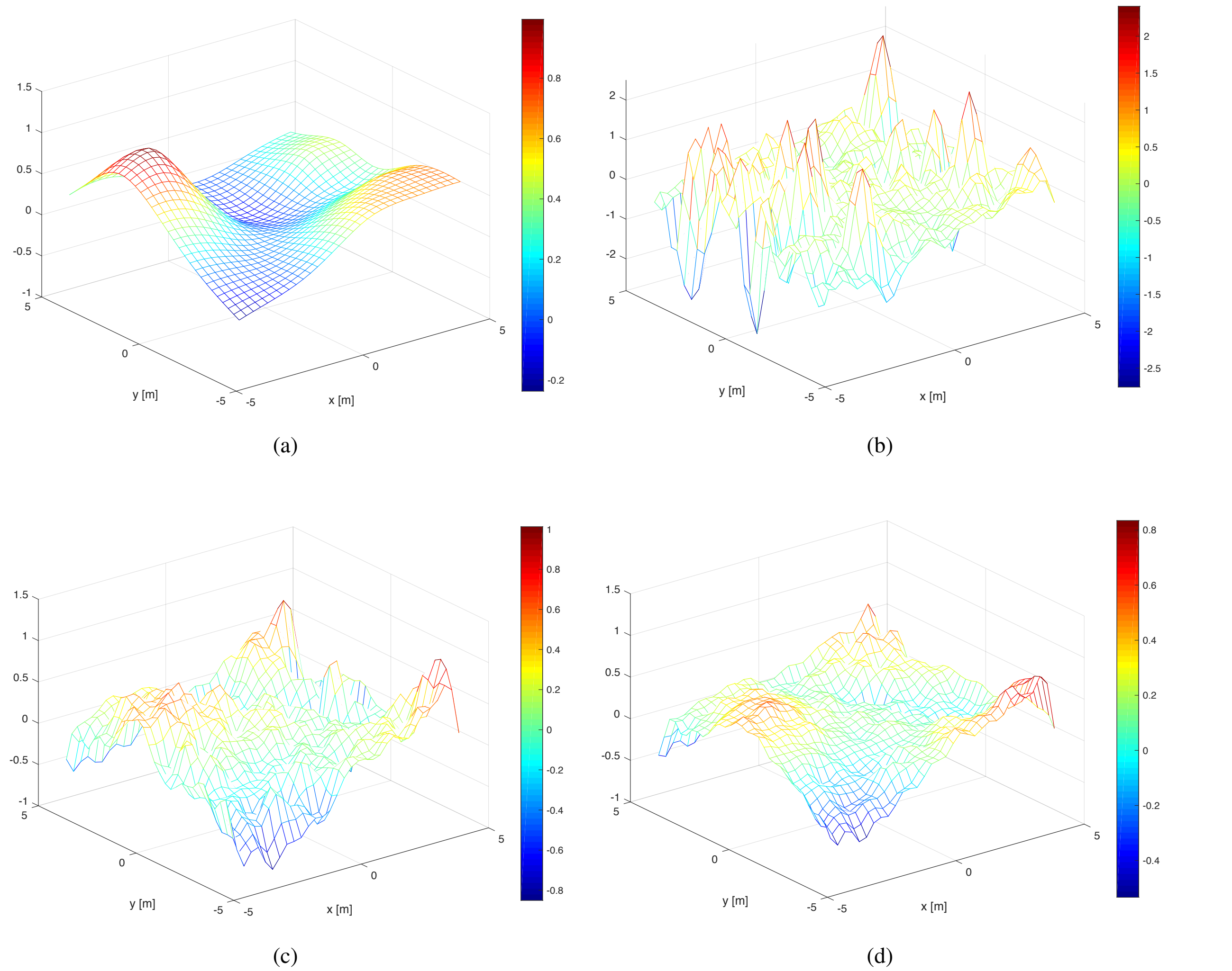

Fig. 5. Radiofrequency surveying can be achieved by crowd-based approaches, where knowledge is gained with agents navigating an area. For the scenario described and for anchor number 2, (a) the true bias field, its estimation after (b) 20 users, (c) 50 users, and (d) 200 users.

according to Fig. 6. It can also be observed that the naive method assuming always LOS has the worst performance. The method based on crowd-based learning achieves an improvement with respect to the method assuming LOS.

We conducted a Monte Carlo simulation by varying the set of user trajectories. Specifically, the simulation consisted of 500 Monte Carlo realizations where, at each realization, K users entered the area in sequence and followed a random path. The users were tracked and the field estimates of the different anchors were updated. At the beginning of each realization, the crowd-based learning method started without any knowledge about the true bias fields of the different anchors, i.e., the learning process was reset. The results in terms of (empirical) cumulative distribution function (CDF) of the localization error are shown in Fig. 7. The curves encompass all the values of localization error for each of the 50 measurements per user and each of the 500 realizations. For

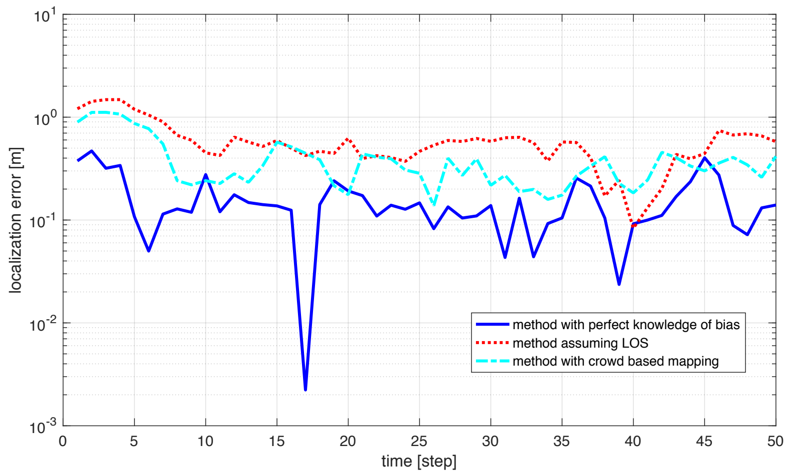

Fig. 6. Positioning error as a function of time for the 200th user.

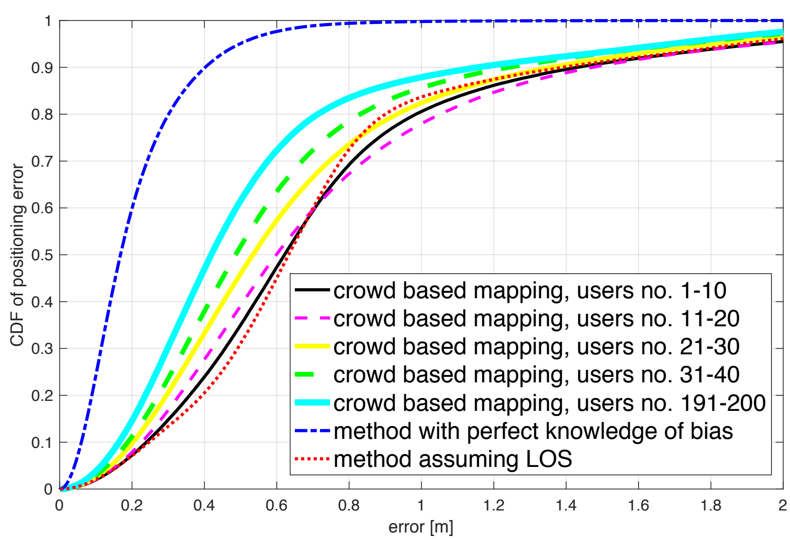

Fig. 7. CDF of localization error for the three methods considered.

the method with crowd-based learning, several curves are displayed, each of them encompassing the results for a specific set of users. It can be observed that the result improves as the number of users grow, first very quickly and later slowing down. Even in this challenging situation with typical NLOS range bias values prevailing, and in the absence of any calibration, the crowd enhanced method achieves a notable improvement with respect to the simple method assuming always LOS.

As an example, if one sets a target performance at 0.5 m, with the described loose approach, after 20 users having navigated in the area, 40% of the locations meet the target performance (covered) with respect to 30% assuming always LOS. After 200 users, more than 60% of the area has the same target performance. The root mean square error (RMSE) for the crowd-based method, computed from the error values corresponding to the curve for users 191 to 200 is 0.72 m, an intermediate value between those of the method assuming LOS (0.88 m) and the benchmark with perfect knowledge (0.26 m).

V. Conclusions

We addressed the problem of estimating the spatial distribution of physical quantities (spatial fields), by taking an advantage of the pervasive diffusion of IoT mobile devices equipped with sensors collecting measurements related to the spatial field at different locations. This crowd-based learning scenario, where the knowledge about the field is refined as new agents enter the area, poses several challenges mainly caused by the random and uncertain position of the agents and the unbounded growth of the amount of collected data. We have shown how these issues can be tackled within the framework of signal processing by illustrating a couple of methods for efficient joint learning and positioning (in terms of memory and computational burden). To demonstrate the potential of these methods, one of them has been applied to the problem in indoor positioning in the presence of NLOS using real measurements in which the range bias value has been modeled as a spatial field to be jointly estimated with the location of the agent.

Several other topics deserve further investigations such as energy consumption minimization, privacy-preserving schemes, and the incentivization of users to participate in the crowd sensing process [30]. The choice of basis functions in representing the spatial fields and the dimension of the spanned space is also important.

Information fusion of multiple agents accessing the cloud in an asynchronous manner as well as the possibility to perform distributed learning are potentially fertile directions for future research investigations. For example, we will commonly have two or more users that will be tracked simultaneously, and there is more than one way to fuse the information extracted from the measurements about the spatial field. In the case of loose coupling, the fusion is less challenging because the system needs to keep track basically of the mean and covariance of $\mathbf{c}$ for each spatial field. The update of these statistics can take place after a new measurement is received (from any user), and the tracking algorithm will always be fed with the latest statistics. Further, the updates can be asynchronous. By contrast, the tight coupling approach provides interesting challenges because implementations of fusion after every received measurement are not easy. The alternative is to fuse the information after the user leaves the area.

Yet, another set of questions about fusion arise when the users do not transmit their measurements to the central unit and instead, given the estimates of the spatial fields from the central unit, they track themselves and at the same time continue to improve the estimates of the spatial fields. At some point in time, before leaving the area, they report their estimates of all the spatial fields to the central unit, which now has to fuse them with the existing information.

Last but no least, the amount of computations needed to implement the tracking of the users and the update of the spatial fields can also be an issue. When most of the computations take place away from the users, this is not so critical. However, when the users employ apps for self-tracking as suggested above, then it is essential that the required need for computing power is minimized. So it is expected that there will be research in methods that minimize computational costs while maintaining “guaranteed” accuracy of self-localization.

VI. Acknowledgments

EA has been supported by the Spanish Ministry of Economy and Competitiveness and the Spanish National Research Agency under grant TEC2016-78434-C3-2-R (AEI/FEDER, EU), by the European Union (FEDER) and by Junta de Andalucía (TIC-155). DD has been supported by the European H2020 project XCycle (Grant 635975). PMD has been supported by NSF under Awards CCF-1618999 and CNS-1642965.

References

[1] R. Di Taranto, S. Muppirisetty, R. Raulefs, D. Slock, T. Svensson, and H. Wymeersch, “Location-aware communications for 5G networks: How location information can improve scalability, latency, and robustness of 5G,” *IEEE Signal Processing Mag.*, vol. 31, no. 6, pp. 102–112, Nov 2014.

[2] J. Lunden, V. Koivunen, and H. V. Poor, “Spectrum exploration and exploitation for cognitive radio: Recent advances,” *IEEE Signal Process. Mag.*, vol. 32, no. 3, pp. 123–140, May 2015.

[3] W. Storms, J. Shockley, and J. Raquet, “Magnetic field navigation in an indoor environment,” in *Proc. of the IEEE Conf. on Ubiquitous Positioning Indoor Navigation and Location Based Service (UPINLBS)*, 2010.

[4] I. Nevat, G. W. Peters, and I. B. Collings, “Random field reconstruction with quantization in wireless sensor networks,” *IEEE Transactions on Signal Processing*, vol. 61, no. 23, pp. 6020–6033, Dec 2013.

[5] J. Unnikrishnan and M. Vetterli, “Sampling and reconstruction of spatial fields using mobile sensors,” *IEEE Transactions on Signal Processing*, vol. 61, no. 9, pp. 2328–2340, May 2013.

[6] R. Verdone, D. Dardari, G. Mazzini, and A. Conti, ," in *Wireless Sensor and Actuator Networks: technologies, analysis and design*. 2008, Elsevier Ltd, London.

[7] D. Dardari, A. Conti, C. Buratti, and R. Verdone, "Mathematical evaluation of environmental monitoring estimation error through energy-efficient wireless sensor networks," *IEEE Trans. Mobile Comput.*, vol. 6, no. 7, pp. 790–802, jul 2007.

[8] J. Shi, J. Wan, H. Yan, and H. Suo, "A survey of cyber-physical systems," in *2011 International Conference on Wireless Communications and Signal Processing (WCSP)*, Nov 2011, pp. 1–6.

[9] C. X. Mavromoustakis, G. Mastorakis, and J.M. Batalla, *Internet of Things (IoT) in 5G Mobile Technologies*, Modeling and Optimization in Science and Technologies. Springer International Publishing, 2016.

[10] H. Ma, D. Zhao, and P. Yuan, "Opportunities in mobile crowd sensing," *IEEE Communications Magazine*, vol. 52, no. 8, pp. 29–35, 2014.

[11] B. Guo, Z. Wang, Z. Yu, Y. Wang, N. Y. Yen, R. Huang, and X. Zhou, "Mobile crowd sensing and computing: The review of an emerging human-powered sensing paradigm," *ACM Comput. Surv.*, vol. 48, no. 1, pp. 7:1–7:31, Aug. 2015.

[12] M. Zappatore, A. Longo, and M. A. Bochicchio, "Crowd-sensing our smart cities: a platform for noise monitoring and acoustic urban planning," *Journal of communications software and systems*, vol. 13, no. 2, pp. 53–67, June 2017.

[13] J. Liu, H. Shen, and X. Zhang, "A survey of mobile crowdsensing techniques: A critical component for the internet of things," in *2016 25th International Conference on Computer Communication and Networks (ICCCN)*, Aug 2016, pp. 1–6.

[14] A. Rai, K. K. Chintalapudi, V. N. Padmanabhan, and R. Sen, "Zee: Zero-effort crowdsourcing for indoor localization," in *Proceedings of the 18th Annual International Conference on Mobile Computing and Networking*, New York, NY, USA, 2012, Mobicom '12, pp. 293–304, ACM.

[15] M. Alzantot and M. Youssef, "Crowdinside: Automatic construction of indoor floorplans," in *Proceedings of the 20th International Conference on Advances in Geographic Information Systems*, New York, NY, USA, 2012, SIGSPATIAL '12, pp. 99–108, ACM.

[16] T. Higuchi, H. Yamaguchi, and T. Higashino, "Context-supported local crowd mapping via collaborative sensing with mobile phones," *Pervasive and Mobile Computing*, vol. 13, pp. 26–51, 2014.

[17] I. Koukoutsidis, "Estimating spatial averages of environmental parameters based on mobile crowdsensing," *ACM Trans. Sen. Netw.*, vol. 14, no. 1, pp. 2:1–2:26, Dec. 2017.

[18] E. S. Lohan, J. Torres-Sospedra, H. Leppäkoski, P. Richter, Z. Peng, and J. Huerta, "Wi-Fi Crowdsourced Fingerprinting Dataset for Indoor Positioning," *Data*, vol. 2, no. 4, pp. 32, 2017.

[19] P. Closas, M. Luise, J. Ávila-Rodriguez, C. Hegarty, and J. Lee, "Advances in Signal Processing for GNSSs [From the Guest Editors]," *IEEE Signal Processing Magazine*, vol. 34, no. 5, pp. 12–15, 2017.

[20] D. Dardari, A. Conti, J. Lien, and M. Z. Win, "The effect of cooperation on localization systems using UWB experimental data," *EURASIP Journal on Advances in Signal Processing, Special Issue on Cooperative Localization in Wireless Ad Hoc and Sensor Networks*, vol. 2008, 2008, Article ID 513873, 12 pages.

[21] C. E. Rasmussen, *Gaussian Processes for Machine Learning*, MIT Press, 2006.

[22] J. Fritz, I. Neuweiler, and W. Nowak, “Application of FFT-based algorithms for large-scale universal Kriging problems,” *Mathematical Geosciences*, vol. 41, no. 5, pp. 509–533, 2009.

[23] S. Särkkä, A. Solin, and J. Hartikainen, “Spatio-temporal learning via infinite-dimensional Bayesian filtering and smoothing,” *IEEE Signal Processing Mag.*, vol. 30, no. 4, pp. 51–61, 2013.

[24] A. Solin and S. Särkkä, “Hilbert space methods for reduced-rank Gaussian process regression,” *arXiv preprint arXiv:1401.5508*, 2014.

[25] D. Dardari, A. Arpino, F. Guidi, and R. Naldi, “A combined GP-State space method for efficient crowd mapping,” in *IEEE ICC 2015 - Workshop on Advances in Network Localization and Navigation*, London, United Kingdom, June 2015.

[26] D. Dardari, G. Pasolini, and F. Zabini, “An efficient method for physical fields mapping through crowdsensing,” *Pervasive and Mobile Computing Journal, Elsevier*, 2017, submitted.

[27] D. Dardari, P. Closas, and P. M. Djurić, “Indoor tracking: Theory, methods, and technologies,” *IEEE Trans. Veh. Technol.*, vol. 64, no. 4, pp. 1263–1278, Apr. 2015.

[28] E. Arias-de-Reyna, D. Dardari, P. Closas, and P. M. Djurić, “Enhanced indoor localization through crowd sensing,” in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, March 2017, pp. 2487–2491.

[29] E. Arias-de-Reyna, D. Dardari, P. Closas, and P. M. Djurić, “Estimation of spatial fields of NLOS/LOS conditions for improved localization in indoor environments,” in *2018 IEEE Statistical Signal Processing Workshop (SSP)*, June 2018.

[30] L. G. Jaimes, I. J. Vergara-Laurens, and A. Raij, “A survey of incentive techniques for mobile crowd sensing,” *IEEE Internet of Things Journal*, vol. 2, no. 5, pp. 370–380, Oct 2015.