Channel Model Identification in Wireless Sensor Networks Using a Fully Distributed Quantized Consensus Algorithm

Angelo Cenedese * and **Filippo Zanella** *

* *University of Padova, DEI - Department of Information Engineering, via G. Gradenigo 6/B, 35131 Padova, Italy (e-mails: angelo.cenedese@unipd.it , ing.filippo.zanella@gmail.com).*

Abstract: In this paper, we consider the problem of designing a distributed strategy to estimate the channel parameters for a generic Wireless Sensor-Actor Network (WSAN). To this aim, we present a distributed least-square algorithm that complies with the constraint of transmitting only integer data through the wireless communication, which often characterizes WSAN embedded architectures. In this respect, we propose a quantized consensus strategy that mitigates the effects of the rounding operations applied to the wireless exchanged floating data. Moreover, the approach is based on a symmetric random gossip strategy, making it suitable for the actual deployment in multiagent networks. Finally, the effectiveness of the proposed algorithm and of its implementation as an open-source application is assessed and the employment of the procedure is illustrated through the application to radio-frequency localization experiments in a real world testbed.

Keywords: Wireless sensor network; wireless channel modeling; consensus; quantized consensus; localization.

1. INTRODUCTION

In recent years, the academic community has devoted a huge research effort on distributed and network controlled systems, especially those employing mobile and wireless devices (see for example Akyildiz and Kasimoglu (2004), Hespanha et al. (2007),Yick et al. (2008)).

As a matter of fact, the idea of ad hoc networks to gather pervasive information and to interact with the surrounding environment through plug and play devices represents a sort of common framework inspiring this research. In particular, it appears how the embedding of ubiquitous Wireless Sensor Networks (WSNs)/Wireless Sensor-Actor Networks (WSANs) in the environment assumes a crucial rôle in bridging the gap between the theoretical study and the real world application (Oliveira and Rodrigues (2011)).

In this context, two main aspects need to be considered, one related to the fact that the topology of the network is usually not a well defined structure (mathematically represented by a graph), the other referring to the many sources of uncertainties and disturbances that characterize the environment and strongly affect the performance of the system (for example due to the presence of walls and obstacles, of people moving in the environment, of other interfering devices). In particular, when dealing with a radio-frequency (RF) wireless channel, it is of paramount importance to understand the properties of the wireless medium with respect to the specific network conditions. Beyond the main use of the wireless channel as a communication channel to transmit and share the information across the network, there are indeed many applications of localization and tracking where the distance among the nodes is inferred by the power level received by the same devices, avoiding in doing so the employment (and installation) of further localization circuitry and systems (see Cenedese et al. (2010) and references within).

Therefore, since the performance of any system based on RF wireless devices is strongly affected by the disturbances entering the communication channel, it is necessary to provide an accurate model of the transmission and to be able to detect the different contribution to path loss effects and power attenuation. With this scope in mind, in this work we aim at the design of a fully distributed algorithm that allows the network nodes to attain the identification of the channel model by their own without resorting to a central unit processing all the available information. In other words, we are seeking for an identification procedure that instead of being centralized (i.e the nodes gather data that are transmitted to the central unit to perform all the model computation) takes advantage of the computational grid provided by the network itself (i.e. the nodes gather data and through local sharing with neighbors they are able to infer the model parameters). Such approach inherits all the advantages common to the distributed algorithms, among which being more robust w.r.t. the centralized one and more adaptive to network topology changes. A special

⋆ The research leading to these results has received funding from the European Community's Seventh Framework Programme under agreements FP7-ICT-223866 FeedNetBack and FP7-ICT-257462 HYCON2 Network of excellence.

⋆⋆This technical report is an extended version of the paper presented at the 19^{th} IFAC World Congress 2014 (IFAC'14). In case of citation, please cite the IFAC paper (same authors and title).

attention is posed on the aspects of communication and quantization (Carli et al. (2010b)), which are particularly relevant in this scenario.

In summary, this work aims at providing a twofold contribution:

- on the theoretical side, a methodology for channel model identification that is completely distributed among the network nodes is studied with particular focus on the quantization effects;
- on the practical side, the experimental validation is presented and the implementation of the developed code is publicly available (Zanella (2013)).

The remainder of the paper is organized as follows: In Sec. 2 an overview of the state of the art on channel modeling and estimation is given, and in Sec. 3 the basic models of both the network and the transmission channel are described. Then, in Sec. 4 and Sec. 5 the main theoretical contribution are provided, respectively discussing the channel parameter identification and the proposed distributed approach to perform such procedure. In Sec. 6 the experimental results of the channel identification are shown and discussed, together with an application of localization and tracking. Finally, in Sec. 7 some conclusions are drawn.

2. RELATED WORK ON RF CHANNEL MODELING

In general, to understand the characteristics of radio transmission in a real world scenario, different approaches can be exploited, ranging from those exploiting a simulative approach through ray-tracing numerical algorithms, to others relying on a defined channel model whose parameters need to be identified either in real-time or exploiting previously acquired experimental data. Clearly, these approaches are strictly related to the application of interest and to the amount of a-priori information available on the installation characteristics.

The ray-tracing techniques have been used widely to predict radio propagation in indoor environments and can be used in combination with stochastic processes to model the temporal variability in the channel response due to people presence and movement in the environment (Valenzuela (1994), Kaya et al. (2009)). Although they provide accurate results if the geometry of the site is known, they are computationally intensive and require long execution time to calculate the propagation characteristics, especially when the indoor environment is large and complex.

On the other hand, the definition of an accurate model of radio channel, linking the received signal power to the distance between the emitter and the receiver, has been a hot research topic since the beginning of radio communication. In particular, the mathematical modeling of the radio transmission is well surveyed in Hashemi (1993), with reference to indoor environments. More accurate descriptions can also be derived, as in Hansen and Reitzner (2004) where both the log-normal transmission model and the power delay are considered, starting though from some a-priori knowledge of the environment such as the building maps and details on the building materials. In the present work, it is considered the log-distance path loss model (Rappaport (2002)), where the received power is linked to the transmission power through a log-normal model of path loss, and other contribution terms are added to take into account other attenuation effects, as better explained in Sec. 3. Despite the fact that the structure of the model is assumed, the values of its parameters are far than known, and characteristic of the installation environment (Wysocki and Zepernick (2000),Goldsmith (2005)), and thus, to employ the same model in different scenarios, a dedicated parameter tuning procedure can be implemented.

In this respect, one approach to parameter identification exploits centralized least square (LS) procedures (Söderström and Stoica (1988)) that start from a series of measurements and compute the parameter set of the whole system. An example of this method is presented in Durgin et al. (1998), where linear regression models are used together with LS methods to estimate the model parameters in different scenarios from a set of experimental campaigns. Such an approach will be considered as a reference for this work and referred hereafter as Centralized Channel Parameter Identification (CCPI). Similar techniques are reported also in Oestges et al. (2010), for different scenarios, and in Bardella et al. (2010), where the estimation of the channel parameters is obtained by applying statistical models to the data gathered through extensive experimental operation employing multichannel transmission.

Differently, a solution here is sought that is distributed and can be implemented in a multiagent setup of limited resource devices, so as to make unnecessary the presence of special units or coordinating nodes. In this sense, a first distributed solution is proposed in Bolognani et al. (2010), where the consensus theory is applied to solve the global LS optimization problem of parameter estimation. This solution is the base of the algorithm developed in this work (Distributed Channel Parameter Identification, DCPI).

3. SYSTEM MODEL

In general, we will use Roman capital letters to indicate matrices, bold fonts to indicate vectors, and plain italic fonts to indicate scalars. We use I to indicate the identity matrix, $\mathbb{1}$ for the unitary vector.

3.1 Network model

A sensor/actor network is commonly modeled as a graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$, where the ordered set $\mathcal{N} = \{1, \dots, N\}$ of N nodes communicate along the edges specified by the set $\mathcal{E}$. Referring to nodes and sensors/actors is equivalent, so we will use the two terms interchangeably. A graph $\mathcal{G}$ is called *undirected* when $(i,j) \in \mathcal{E} \Rightarrow (j,i) \in \mathcal{E}$. We denote with $\mathcal{V}(i) = \{j \,|\, (i,j) \in \mathcal{E}, i \neq j\}$ the set of neighbors of node i, with the degree $d(i) = |\mathcal{V}(i)|$ its cardinality. In a directed graph, a *walk* on a graph is an alternating series of nodes and edges, beginning and ending with a node, in which each edge is incident with the node immediately preceding it and the node immediately following it. A *path* is a walk in which all nodes are distinct. We say that a graph is *connected* if every pair of nodes (i,j) is connected by a path. In particular, a directed graph is called *strongly connected* if there is a path from each node in the graph to every other node.

The communication strategy among the nodes in the network can then fall into two categories: *broadcast* communication, where one node i transmits a message to all its neighbors $\mathcal{V}(i)$, and *gossip* communication, where a node i transmits a message to a specific node $j \in \mathcal{V}(i)$. The gossip communication can be either *symmetric* or *asymmetric*, respectively meaning that the transmitting node awaits for an answer from the receiver or not. In addition, the communication (broadcast or gossip) can be *synchronous*, in which all nodes communicate at the same time, or *asynchronous*, when nodes are triggered one at a time. In this case the resulting sequence of nodes can be *randomized* or *sequential*, w.r.t. the activation of a node to send messages to the neighbors. For more details, we refer the reader to the specialized literature (e.g. Mesbahi and Egerstedt (2010), Boyd et al. (2006)).

In our setup we consider a graph that is undirected, connected and not time-varying, in a randomized symmetric gossip type communication network.

3.2 Channel model

In the particular case of a WSN/WSAN, nodes communicate through a wireless channel, which is characterized in general by a non-zero packet loss probability. The communication reliability through the wireless medium is affected by both static and dynamic phenomena, the former basically related to the environment structure, the latter mainly due to interference and noise. Considering two nodes i and j, placed respectively at $\mathbf{z}_i \in \mathbb{R}^3$, $\mathbf{z}_j \in \mathbb{R}^3$ in an indoor environment, their distance being $d_{ij} := \|\mathbf{z}_i - \mathbf{z}_j\|_2$, a well agreed model is as follows, in terms of power P_{ij} received by node i when node j is transmitting with P_j^{tx}:

$$P_{ij} := P_j^{tx} + r_j + f_{pl}(d_{ij}) + f_{sf}(\mathbf{z}_i, \mathbf{z}_j) + f_a(\mathbf{z}_i, \mathbf{z}_j) + v_{ff}(t) + o_i, \quad (1)$$

where (all power levels are given in dBm):

- r_j and o_i are the transmitter and receiver offsets (w.r.t. the datasheet value). Being basically related to manufacturing mismatches, they are supposed to be constant in time;
- $f_{pl}(\cdot)$ is the *path loss* effect. This term is the power attenuation of the transmitted signal as the source to receiver distance increases, and it is given by a log-distance model Rappaport (2002):
$$f_{pl}(d_{ij}) := \beta - 10\gamma \log_{10}(d_{ij}) : \quad (2)$$
β is the receiver gain at the the nominal distance $d_{ij} = 1$ m and γ is the *loss factor*:wave reflection, diffraction and diffusion are some of the main phenomena that contribute to the path attenuation;
- $f_{sf}(\cdot)$ is the *slow fading*. This term is a slow varying component due to the effects of large obstacles along the propagation path (*shadowing*); it is supposed to be space dependent, symmetric, and modeled as a Gaussian variable (Gudmundson (1991));
- $f_a(\cdot)$ is the channel asymmetry factor. Related to signal non-symmetric reflections, it is spatially modeled as a zero-mean Gaussian random noise;
- $v_{ff}(\cdot)$ is the *fast fading*. This contribution is due to fast signal fluctuations, therefore it can be modeled as temporal zero-mean white noise.

4. CENTRALIZED PARAMETER IDENTIFICATION

The parameters of the Eq. (1) depend on the properties of both the environment where the network is installed and the specific hardware in use, and they need to be estimated on site, during a dedicated learning phase or continuously during runtime. In this perspective, in order to perform channel parameter identification, the model of Eq. (1) is simplified by taking the following assumptions:

- the transmission power of the sensors is set at the maximum level, $P_j^{tx} = 0$ dBm, $\forall j \in \mathcal{N}$, so that the transmitter offset is almost zero, $r_j \cong 0$ dBm, $\forall j \in \mathcal{N}$;
- the offset o_i is assumed to be zero, $\forall i \in \mathcal{N}$. Beside being an approximation, we want to notice that the offsets can be compensated exploiting a distributed strategy (Bolognani et al. (2010)) that is completely decoupled from the channel parameter estimation;
- the fast fading effect $v_{ff}(t)$ is removed, by averaging the received power over a set of $M_{ij} > 0$ consecutive measures: $\overline{P}_{ij} := \sum_{k=1}^{M_{ij}} P_{ij}^k$.

It follows that the average received power $\overline{P}_{ij}$ becomes:

$$\overline{P}_{ij} = \beta - 10\gamma \log_{10}(d_{ij}) + f_{sf}(\mathbf{z}_i, \mathbf{z}_j) + f_a(\mathbf{z}_i, \mathbf{z}_j). \quad (3)$$

Moreover, the components of slow fading and channel asymmetry are independent Gaussian random variables of variance σ_{sf}^2 and σ_a^2 respectively and can be combined into one zero-mean random variable w_{ij} with variance equal to $\sigma^2 = \sigma_a^2 + \sigma_{sf}^2$:

$$\overline{P}_{ij} = \beta - 10\gamma \log_{10}(d_{ij}) + w_{ij}. \quad (4)$$

4.1 Least-squares estimation

Such a channel model requires only the estimation of β and γ. Hence, we can write the model in linear form and employ a LS estimator: being a Markov estimator, this will coincide with the maximum likelihood estimator.

For each sensor $i \in \mathcal{N}$ we collect the self positions $\mathbf{z}_j$ communicated by neighbors $j \in \mathcal{V}(i)$, which allows the computation of the relative distance d_{ij}, and M_{ij} measurements of the received power P_{ij} to evaluate the average $\overline{P}_{ij} = \sum_{k=1}^{M_{ij}} P_{ij}^k$. We define $\mathcal{D}(i) := \{(\overline{P}_{ij}, d_{ij}) \,|\, j \in \mathcal{V}(i)\}$ as the actual data set available to each sensor i, and the communication flow is equal to $\sum_{i=1}^{N} \sum_{j \in \mathcal{V}(i)} M_{ij}$ to finally collect $M := \sum_{i=1}^{N} |\mathcal{D}(i)|$ total data pairs.

Starting from these M measurements, consider the linear regression model $\mathbf{b} = \mathrm{S}\,\boldsymbol{\theta} + \mathbf{w}$, where $\boldsymbol{\theta} := [\beta, \gamma]^T$ is the parameter vector to be estimated, $\mathbf{w}$ is a M-dimensional zero-mean random vector with variance $\sigma^2\,\mathrm{I} = (\sigma_a^2 + \sigma_{sf}^2)\mathrm{I}$, and $\mathbf{b}$ and S are defined as

$$\mathbf{b} := \begin{bmatrix} \overline{P}_{1j_1^1} \\ \vdots \\ \overline{P}_{1j^1_{|\mathcal{D}(1)|}} \\ \hline \vdots \\ \overline{P}_{Nj_1^N} \\ \vdots \\ \overline{P}_{Nj^N_{|\mathcal{D}(N)|}} \end{bmatrix}, \; \mathrm{S} := \begin{bmatrix} 1 & -10\log_{10} d_{1j_1^1} \\ \vdots & \vdots \\ 1 & -10\log_{10} d_{1j^1_{|\mathcal{D}(1)|}} \\ \hline \vdots & \vdots \\ 1 & -10\log_{10} d_{Nj_1^N} \\ \vdots & \vdots \\ 1 & -10\log_{10} d_{Nj^N_{|\mathcal{D}(N)|}} \end{bmatrix}.$$

Under this assumptions it is well known (Rao et al. (1997)) that the LS identification of the parameter $\boldsymbol{\theta}$ is given by:

$$\hat{\boldsymbol{\theta}}_{CCPI} := \arg\min_{\boldsymbol{\theta}} \|\mathrm{S}\boldsymbol{\theta} - \mathbf{b}\| = (\mathrm{S}^T\mathrm{S})^{-1}\mathrm{S}^T\mathbf{b} \qquad (5)$$

with variance $N^{-1}\|\mathbf{b} - \mathrm{S}\hat{\boldsymbol{\theta}}_{CCPI}\|^2$, assuming that the matrix $\mathrm{S}^T\mathrm{S}$ is not singular.

This approach is referred to as Centralized Channel Parameter Identification (CCPI).

5. DISTRIBUTED PARAMETER IDENTIFICATION

A strategy is now presented to distributedly estimate the parameters β and γ of the communication channel in (4). First, we introduce some basic concepts of the consensus theory; then, we focus on the rôle of the quantization, which is of practical interest for our scopes. Lastly, we provide the theoretical tools to address the parameters identification problem in a distributed fashion through the Distributed Channel Parameter Identification (DCPI) algorithm.

5.1 Consensus

Consider a network of N nodes endowed with a state $x_i(t) : \mathbb{R} \to \mathbb{R}$, $\forall i \in \mathcal{N}$. The consensus problem can be summarized as the problem of allowing the nodes to reach an agreement regarding a quantity that is function of the state of all nodes. A consensus algorithm is an interaction rule of the form

$$\mathbf{x}(t+1) = \mathrm{Q}(t)\,\mathbf{x}(t), \qquad (6)$$

where $\mathbf{x}(0)$ is given, $\mathbf{x}(t) = [x_1(t), \ldots, x_N(t)]^T \in \mathbb{R}^N$ and $\mathrm{Q}(t) \in \mathbb{R}^{N\times N}$. Specifically, with reference to a randomized symmetric gossip communication, $\mathrm{Q}(t)$ is a random matrix extracted from an i.i.d. sequence $\{\mathrm{Q}(t)\}_{t\geq 0}$ of double stochastic matrices, defined as

$$\mathrm{Q}(t) = \mathrm{I} - \frac{1}{2}(\mathbf{e}_i - \mathbf{e}_j)(\mathbf{e}_i - \mathbf{e}_j)^T$$

where $\mathbf{e}_\ell = [0 \ldots 0\,1\,0 \ldots 0]^T \in \mathbb{R}^N$ is a null vector except for the ℓ-th component that is equal to 1; the couple (i, j) refers to the randomly selected edge at time instant t. Notice that when a link (i, j) fails in a symmetric gossip, there is no communication at all and then no update is performed ($\mathrm{Q}(t) = \mathrm{I}$ for some t). According to the previous statements, the system in (6) can be rewritten as

$$\begin{cases} x_i(t+1) = x_j(t+1) = \frac{1}{2}\left(x_i(t) + x_j(t)\right) \\ x_h(t+1) = x_h(t) \quad \text{if } h \neq i, j \end{cases} \qquad (7)$$

which highlights the structure of the symmetric gossip update.

From (Boyd et al. (2006)) we know that given an i.i.d. sequence $\{\mathrm{Q}(t)\}_{t\geq 0}$ of doubly stochastic matrix and a connected graph in which each edge is selected with a strictly positive probability, then, for every initial condition $\mathbf{x}(0)$, the sequence $\{\mathrm{Q}(t)\}_{t\geq 0}$ solves the *probabilistic (average) consensus problem*. That is, almost surely

$$\lim_{t\to\infty} x_i(t) = \frac{1}{N}\sum_{i=1}^{N} x_i(0) = \frac{1}{N}\mathbb{1}\mathbf{x}(0) \quad \forall i \in \mathcal{N}. \qquad (8)$$

Notice that the presented strategy and assumptions allow to accept link failures during the communication process. As an additional remark, the randomized symmetric gossip guarantees average consensus for all realizations, but it is quite expensive from a communication point of view. In fact, at least two acknowledged packets need to be exchanged at every step of the consensus iteration.

5.2 Quantized Consensus

The gossip algorithm in (7) is based on the crucial assumption that each node transmits to its neighbor the exact value of its state, but sometimes the quantization of the messages is practically imposed by communication constraints, application features or specific programming languages. This happens for example in TinyOS (Lewis (2006)), an open source operating system specially designed for WSNs, whose transmission protocol forces to use integer numbers (obtained by *round* operation). Because of this, nodes transmit measurements (i.e. their states) that differ from the actual values, and thus they have a knowledge about the neighbors' states that is not precise. This behavior can be assimilated to a quantization scheme and Carli et al. (2010a) have investigated the adaptation of the gossip algorithm to a network subject to quantized communication. In this scenario, our aim is to present and discuss what kind of quantization is convenient to adopt.

We define *deterministic quantizer* a map $q_d : \mathbb{R} \to \mathbb{Z}$ that converts a number $v \in \mathbb{R}$ into its nearest integer, i.e.

$$q_d(v) = n \in \mathbb{Z} \Leftrightarrow \begin{cases} v \in \left[n - \frac{1}{2}, n + \frac{1}{2}\right[& \text{if } v \geq 0 \\ v \in \left]n - \frac{1}{2}, n + \frac{1}{2}\right] & \text{if } v < 0. \end{cases} \qquad (9)$$

Instead, we call *standard probabilistic quantizer* the map $q_p : \mathbb{R} \to \mathbb{Z}$ that, for any $v \in \mathbb{R}$, is defined by

$$q_p(v) = \begin{cases} \lfloor v \rfloor & \text{with probability } \lceil v \rceil - v \\ \lceil v \rceil & \text{with probability } v - \lfloor v \rfloor. \end{cases} \qquad (10)$$

Now assume that (i, j) is the edge selected at time t. Considering the deterministic and standard probabilistic quantizer, sensors i and j, following the update rule in (7), apply the algorithms

$$x_i(t+1) = x_j(t+1) = \frac{1}{2}\left(q_d(x_i(t)) + q_d(x_j(t))\right) \qquad (11)$$

or

$$x_i(t+1) = x_j(t+1) = \frac{1}{2}\left(q_p(x_i(t)) + q_p(x_j(t))\right). \qquad (12)$$

It has been proved that the laws (11)-(12) drive the systems almost surely to consensus at an integer value, i.e. almost surely there exist $T \in \mathbb{Z}_{\geq 0}$ and $\bar{x} \in \mathbb{Z}$ such that $x_i(t) = \bar{x}$, for all $i \in \mathcal{N}$ and for all $t \geq T$. The previous statement highlights how these strategies don't preserve in general the initial average, but it can be proven that with the standard probabilistic quantizer it is at least

guaranteed that the average is preserved in expectation. We compare here the quantization strategies (11) and (12) in terms of the deviation

$$\tilde{x} = \mathbb{E}[|\bar{x} - N^{-1}\mathbb{1}\mathbf{x}(0)|]. \quad (13)$$

In Fig. 1(a), $\tilde{x}$ is plotted as a function of the number of sensors, i.e. generating complete graphs of increasing size, with random topology. Simulations show that using a q_p quantization the reached consensus is close to the average of the initial condition, although the average is preserved only in expectation. On the contrary, due to the accumulation of rounding errors, a q_d approach leads to a consensus point which is remarkably distant from the average of the initial condition. Moreover, this gap increases with N and depends on the initial condition.

In Fig. 1(b) we consider again the strategy

$$x_i(t+1) = x_j(t+1) = \frac{1}{2}\left(q_*(x_i(t)) + q_*(x_j(t))\right) \quad (14)$$

where q_* represents a generic quantization, as in (11)-(12). We compare different deterministic and probabilistic quantizers, in terms of the quantization error $\tilde{x}$ (13) as a function of the length of initial condition range $L = 2x_0$, where $\mathbf{x}(0) \sim U([-x_0 \;\; x_0])$, $x_0 \in [0 \;\; 100]$, $x_0 \in \mathbb{Z}$. The setup is given by a random graph built on a set of $N = 30$ sensors. To apply the strategy in (14), we select a deterministic quantizer q_d, a *ceiling deterministic quantizer* $q_c(v) := \lceil v \rceil$, a *flooring deterministic quantizer* $q_f(v) := \lfloor v \rfloor$, a *uniform probabilistic quantizer*

$$q_u(v) = \begin{cases} \lfloor v \rfloor & \text{if } \mathbb{P}_{[\mathcal{U}[0,1] > 0.5)]} \\ \lceil v \rceil & \text{otherwise} \end{cases}$$

and a *gaussian probabilistic quantizer*

$$q_g^{\lambda}(v) = \begin{cases} \lfloor v \rfloor & \text{if } \mathbb{P}_{[\mathcal{N}(0,1) > \lambda)]} \\ \lceil v \rceil & \text{otherwise} \end{cases}.$$

Simulations show that the quantizers lose their independence from initial conditions and they increase their $\tilde{x}$ error as they become more deterministic. It is clear from the plots reported in Fig. 1(b) that a rounding operation is better than using only a q_c and q_f quantizers (which are clearly comparable) since it switches between the ceiling and flooring of the state following the rule defined in (9).

A q_g^{λ} quantizer, which chooses to ceil or floor the state with a Gaussian probability becomes less deterministic as the λ-threshold moves from the mean of the Gaussian distribution. It is easy to see that the uniform probabilistic quantizer is best quantizer in terms of both independence from initial conditions and negligible $\tilde{x}$ error. The reason is that the q_u quantizer chooses with uniform probability to apply a ceiling or flooring rule, hence its randomness is higher than all the others strategies considered.

Considering our scenario, the equation (14) needs to be slightly modified into

$$x_i(t+1) = x_j(t+1) = q_*\left(\frac{1}{2}\left(x_i(t) + x_j(t)\right)\right). \quad (15)$$

This kind of modification does not affect substantially the behavior of the update in (14), in fact the deviation $\tilde{x}$ of the two approaches is comparable, as it is depicted in Fig. 2. Intuitively, (14) is better than (15) just because the former executes separately two quantizations at the same instant t for the pair of sensors (i, j), which means that (14) introduces more randomness in the system than (15), which executes a single instance of quantization. This is explanatory, since as we have seen before this additional randomness advantages the accuracy of the estimation of the real values of sensors' states.

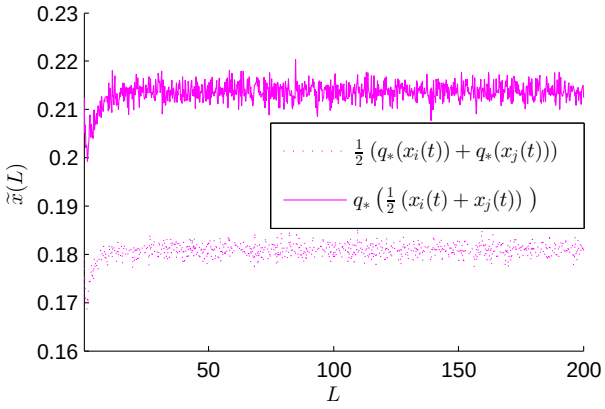

Fig. 2. Comparison in terms of $\tilde{x}$ as a function of the length of initial condition range $L = 2x_0$, where $\mathbf{x}(0) \sim U([-x_0 \;\; x_0])$, $x_0 \in [0 \;\; 10]$, between the update procedures in (14) and (15), for different initial conditions and complete graphs of size $N = 30$. The plotted values are the average of 1000 runs.

5.3 Least-squares estimation via consensus

The DCPI through a consensus approach is a procedure discussed in Bolognani et al. (2010). The LS estimation of the channel parameters can be computed as the solution of a distributed algorithm that does not require the knowledge of the total number of nodes N or the total number of data M available. Here we build on this result to design a formulation combined with the quantization procedure, which is suitable for our goal.

Consider

$$\mathbf{b}_i := \begin{bmatrix} \overline{P}_{ij_1^i} \\ \vdots \\ \overline{P}_{ij_{|\mathcal{D}(i)|}^i} \end{bmatrix}, \; \mathrm{S}_i := \begin{bmatrix} 1 & -10\log_{10} d_{ij_1^i} \\ \vdots & \vdots \\ 1 & -10\log_{10} d_{ij_{|\mathcal{D}(i)|}^i} \end{bmatrix},$$

as the blocks of size $|\mathcal{D}(i)|$ of $\mathbf{b}$ and those of size $|\mathcal{D}(i)| \times 2$ of S, introduced in Sec. 4, and compute

$$\widetilde{\mathrm{S}}_i := {\mathrm{S}_i}^T \mathrm{S}_i = \begin{bmatrix} |\mathcal{D}(i)| & -10 \sum_{k=1}^{|\mathcal{D}(i)|} \log d_{ij_k^i} \\ -10 \sum_{k=1}^{|\mathcal{D}(i)|} \log d_{ij_k^i} & 100 \sum_{k=1}^{|\mathcal{D}(i)|} (\log d_{ij_k^i})^2 \end{bmatrix} \quad (16)$$

and

$$\widetilde{\mathbf{b}}_i := {\mathrm{S}_i}^T \mathbf{b}_i = \begin{bmatrix} \sum_{k=1}^{|\mathcal{D}(i)|} \overline{P}_{ij_k^i} \\ \sum_{k=1}^{|\mathcal{D}(i)|} \overline{P}_{ij_k^i} (-10\log d_{ij_k^i}) \end{bmatrix}. \quad (17)$$

In this way, each component of the matrices $\widetilde{\mathrm{S}}_i$ and vectors $\widetilde{\mathbf{b}}_i$ can be computed independently, i.e. six different consensus procedures can be applied at the same time to

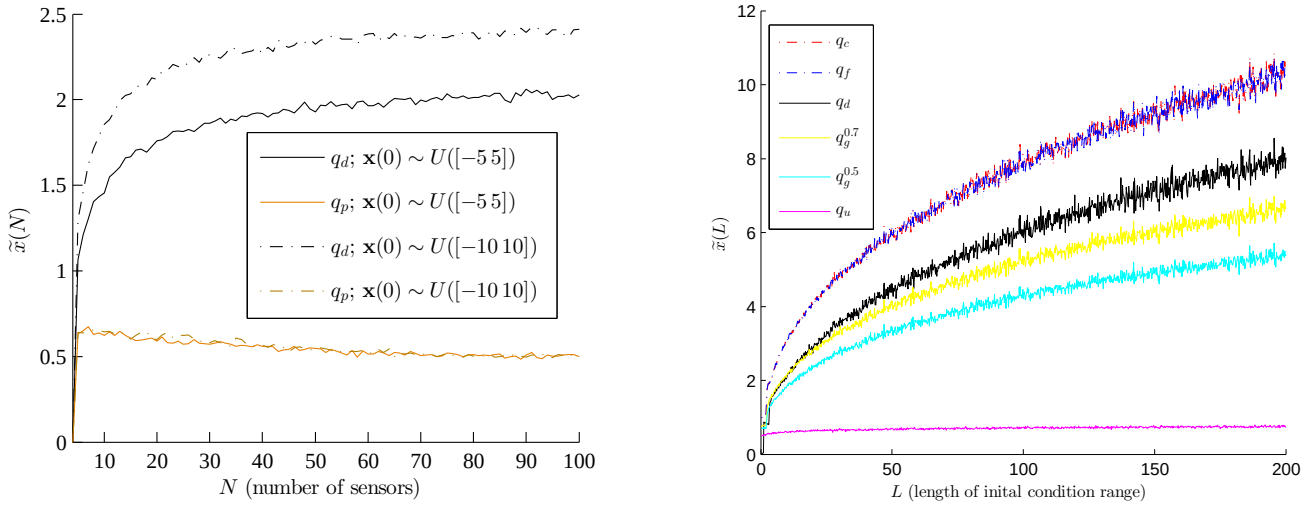

Fig. 1. Deterministic and probabilistic quantizers. (a) Comparison in terms of the deviation $\tilde{x}$ between q_d and q_p, for complete graphs of N sensors. The plotted values are averaged over 1000 runs. The initial conditions are chosen from a uniform distribution on the given intervals. (b) Comparison in terms of $\tilde{x}$ as a function of the length of initial condition range $L = 2x_0$, where $\mathbf{x}(0) \sim U([-x_0\ x_0])$, $x_0 \in [0\ 100]$, for different initial conditions and complete graphs ($N = 30$). The plotted values are the average of 10000 runs.

distributedly compute the matrices components that yield the identification of the channel parameters. In fact:

$$\hat{\boldsymbol{\theta}}_{CCPI} = (\mathrm{S}^T\mathrm{S})^{-1}\mathrm{S}^T\mathbf{b} = \left(\frac{1}{N}(\mathrm{S}^T\mathrm{S})\right)^{-1}\frac{1}{N}\left(\mathrm{S}^T\mathbf{b}\right)$$
$$= \left(\frac{1}{N}\sum_{i=1}^{N}\widetilde{\mathrm{S}}_i\right)^{-1}\frac{1}{N}\sum_{i=1}^{N}\widetilde{\mathbf{b}}_i =: \left(\overline{\mathrm{S}}\right)^{-1}\overline{\mathbf{b}},$$

where $\overline{\mathrm{S}}$ and $\overline{\mathbf{b}}$ have been introduced to simplify the notation. By introducing state X_i as the unknown channel parameters $\mathrm{X}_i := \left[\widetilde{\mathrm{S}}_i \,|\, \widetilde{\mathbf{b}}_i\right]$, the consensus strategy with uniform quantization can be applied to the node pair (i, j)

$$\begin{cases} \mathrm{X}_i(t+1) = \mathrm{X}_j(t+1) = q_u\left(\frac{1}{2}\left(\mathrm{X}_i(t) + \mathrm{X}_j(t)\right)\right) \\ \mathrm{X}_h(0) = \left[\widetilde{\mathrm{S}}_h \,|\, \widetilde{\mathbf{b}}_h\right] \quad \forall h = i, j \end{cases}. \tag{18}$$

After a finite time T the nodes reach consensus over the elements $\overline{\mathrm{S}}$ and $\overline{\mathbf{b}}$, which are needed for the distributed channel identification, i.e.

$$\mathrm{X}_h(t \geq T) = \left[\overline{\mathrm{S}} + \mathrm{E} \,|\, \overline{\mathbf{b}} + \mathbf{e}\right]$$

$\forall h = 1, \ldots, N$, where E and $\mathbf{e}$ are the (bounded) errors introduced by the quantization performed at each iteration. Hence, the DCPI is given by

$$\hat{\boldsymbol{\theta}}_{DCPI} = \left(\overline{\mathrm{S}} + \mathrm{E}\right)^{-1}\left(\overline{\mathbf{b}} + \mathbf{e}\right). \tag{19}$$

By the Matrix Inversion Lemma (Tylavsky and Sohie (1986)), assuming that E is non-singular, the identity

$$(\overline{\mathrm{S}} + \mathrm{E})^{-1} = (\overline{\mathrm{S}})^{-1} - (\overline{\mathrm{S}})^{-1}((\mathrm{E})^{-1} + (\overline{\mathrm{S}})^{-1})^{-1}(\overline{\mathrm{S}})^{-1}$$

is used to rewrite (19) as

$$\hat{\boldsymbol{\theta}}_{DCPI} = \left(\overline{\mathrm{S}}\right)^{-1}\overline{\mathbf{b}} + \tilde{\mathbf{e}} = \hat{\boldsymbol{\theta}}_{CCPI} + \tilde{\mathbf{e}}$$

meaning that the distributed estimation corresponds to its centralized version corrupted by a bounded additive error that depends on the quantization errors on E and $\mathbf{e}$:

$$\tilde{\mathbf{e}} = (\overline{\mathrm{S}})^{-1}\left[\mathbf{e} - \left((\mathrm{I} - ((\mathrm{E})^{-1} + (\overline{\mathrm{S}})^{-1})^{-1}(\overline{\mathrm{S}})^{-1}\right)(\overline{\mathbf{b}} + \mathbf{e})\right].$$

6. EXPERIMENTAL RESULTS

The strategy described in the previous sections have been assessed and validated by implementing the algorithms on a WSN installed at the ground floor of the Department of Information Engineering (DEI) of the University of Padova. This area, shown in Fig. 3, has an approximate size of $15 \times 36m^2$, is partially unstructured, with laboratory/office furniture and equipment, and is subject to a reasonable level of interference and electromagnetic noise. The sensor network testbed employs TMOTE$^{\text{TM}}$ SKY nodes connected via USB (serial) hubs that provide power supply and allow to collect log data for debugging purpose. The sensors are also connected to embedded computers that act as gateways and connect via Ethernet to a central server from which to monitor, manage and check the entire WSN.

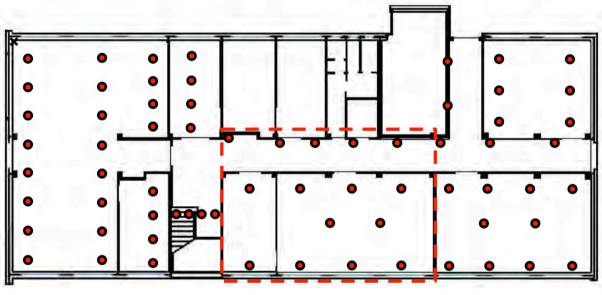

Fig. 3. Planimetry of the testbed (WSN layout).

6.1 Quantization effects

Before dealing with the channel identification procedure, the efficacy of the symmetric gossip communication is assessed by testing the deterministic version described by (9). A subset $\mathcal{N} = \{s_1, \ldots, s_N\}$ of fixed sensors is considered, placed in arbitrary positions but in a range

such that each sensor $s_i \in \mathcal{N}$ can communicate with at least another $s_j \in \mathcal{N}$. A state value x_i, $i = 1, \ldots, N$, is then assigned to each of these sensors, and simulations are performed gradually increasing N and assigning suitable values to the timings and to the desired percentage of collision probability, to converge much faster without compromising the final outcome. The simulations confirm the convergence towards a common value that does not coincide exactly with the expected value $\overline{x}$ because of the behavior of the deterministic quantizer, as stated in Sec. 5. Fig. 4 shows the evolution of the state for one of these experiments in which $N = 7$ sensors are used (randomly taken node subset in the dashed red box highlighted in Fig. 3) to reach the consensus on $\overline{x} = \sum_{i=1}^{N} x_i$.

The detail of the magnification box of Fig. 4 highlights how all node states converge to a common value that does not coincide exactly with the expected value $\overline{x}$ because of the behavior of the deterministic quantizer, as stated in Sec. 5.

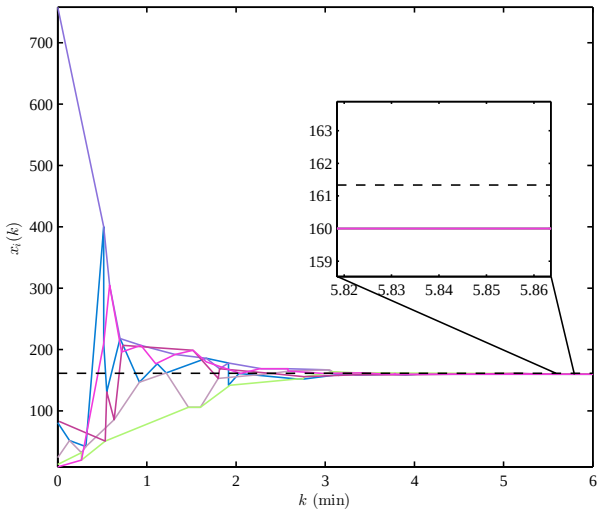

Fig. 4. Convergence of x_i using symmetric gossip with a deterministic quantizer. The dashed line is the value $\overline{x}$.

Again, with a configuration example of $N = 7$ nodes in a random connected graph (taken in the dashed red box highlighted in Fig. 3), it can be assessed that a symmetric gossip with probabilistic quantizers is generally better than a deterministic one, by comparing q_d and q_u quantizers for different initial conditions $x_i(0)$ of the scalar state $x_i(t)$, $i = 1, \ldots, N$. Fig. 5 shows the consensus errors $\tilde{x}_d := |\overline{x}_d - N^{-1}\mathbb{1}\mathbf{x}(0)|$ and $\tilde{x}_u := |\overline{x}_u - N^{-1}\mathbb{1}\mathbf{x}(0)|$, for different lengths of the initial condition range $L = 2x_0$, where $\mathbf{x}(0) \sim U([-x_0 \; x_0])$, and $\overline{x}_d, \overline{x}_u$ are the consensus values reached by the two quantizers. It is easy to see that in the considered cases $\tilde{x}_u$ is almost always lower than $\tilde{x}_d$.

6.2 DCPI algorithm

Then, the distributed identification of the β and γ parameters of the channel model is carried out by means of the symmetric gossip algorithm of Sec. 5.

It is important to observe that the procedure requires the node synchronization because the algorithm consists of

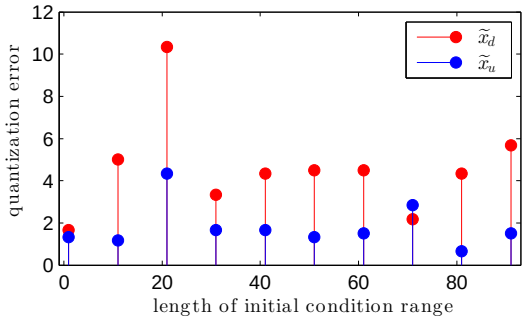

Fig. 5. Quantization error $\tilde{x} = \mathbb{E}[|\bar{x} - N^{-1}\mathbb{1}\mathbf{x}(0)|]$ for $N = 7$ sensors considering a deterministic ($\tilde{x}_d$) and a probabilistic ($\tilde{x}_u$) quantizer of uniform distribution for different values of the initial condition range $L = 2x_0$, where $\mathbf{x}(0) \sim U([-x_0 \; x_0])$, $x_0 \in \{1/2, 11/2, \ldots, 91/2\}$.

a unique phase in which the sensors perform the power measurements needed to build consensus matrices (16) and (17). Alternatively, the synchronization requirement can be removed by modifying the algorithm in such a way that each node can make Received Signal Strength (RSS) requests to the WSN even when other nodes have already started the consensus iterations. Consider the matrix $\mathrm{X}_i(t)$ of the consensus strategy (18):

$$\mathrm{X}_i(t) = \left[\widetilde{\mathrm{S}}_i(t) \,|\, \widetilde{\mathbf{b}}_i(t)\right] = \left[\begin{array}{cc|c} x_i^{(1,1)}(t) & x_i^{(1,2)}(t) & x_i^{(1,3)}(t) \\ x_i^{(2,1)}(t) & x_i^{(2,2)}(t) & x_i^{(2,3)}(t) \end{array}\right] :$$

the trajectories of the entries $x_i^{(h,k)}(t)$ converge, some of these showing an oscillatory behavior for a limited period of time, due to the quantization effect that causes the node states to randomly oscillate around nearby values. A convergence example is shown in Fig. 6, where the trajectories of the entry $x_i^{(1,2)}(t)$ are plotted for a subset of nodes. In this example, the oscillations occur between values -94 and -95 in an interval of about 10 minutes.

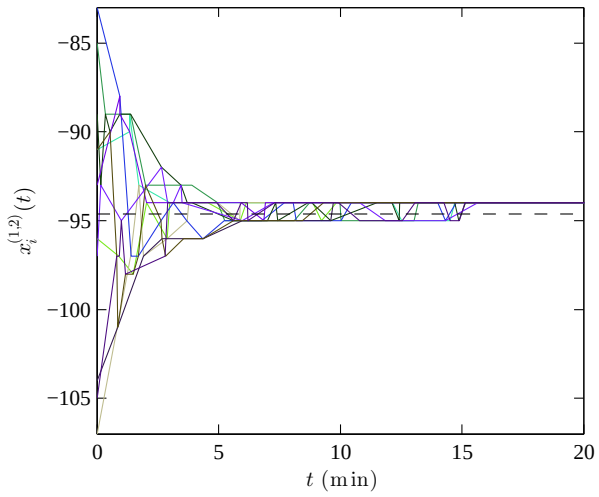

Fig. 6. DCPI convergence of the estimated $x_i^{(1,2)}(t)$ for a node subset. The dashed line is the expected value.

In addition, Fig. 7 shows the trajectories of the estimates for $x_i^{(2,2)}(t)$, $x_i^{(1,3)}(t)$, $x_i^{(2,3)}(t)$ of the sensors $i = 1, \ldots, N$ in the whole testbed environment. The estimates

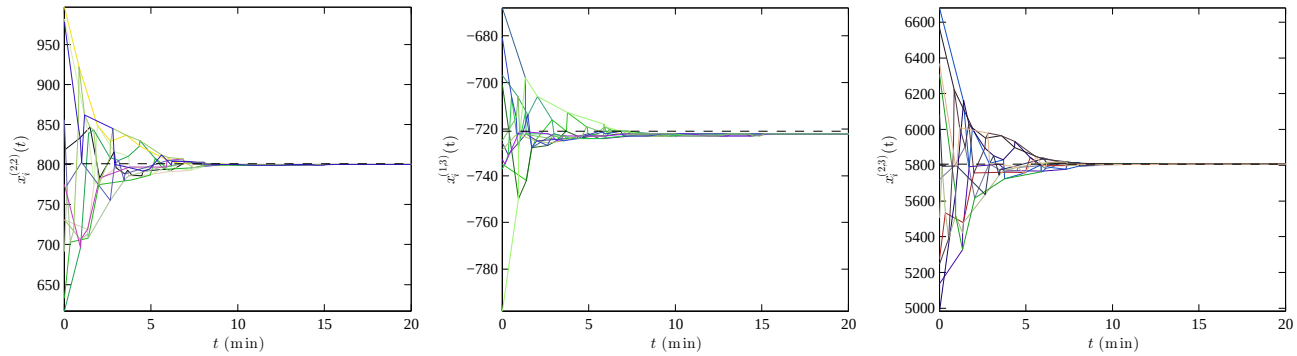

Fig. 7. DCPI convergence of the estimated $x_i^{(2,2)}(t)$, $x_i^{(1,3)}(t)$, $x_i^{(2,3)}(t)$ for all the sensors $i \in \mathcal{N}$ of the testbed. The dashed lines are the expected values $\overline{x}^{(2,2)}$, $\overline{x}^{(1,3)}$, $\overline{x}^{(2,3)}$.

of $x_i^{(1,1)}(t)$ and $x_i^{(2,1)}(t)$ are not shown since the former is identically constant for all sensors and the latter is equal to $x_i^{(1,2)}(t)$ by the symmetry of $\widetilde{\mathrm{S}}_i$ (see (16)-(17)).

To compare the performance of the DCPI with that of the CCPI, a set of 20 experimental RSS measurements for a subset of nodes has been collected by a centralized server communicating through wired connection with the testbed nodes (thus the data in this case are not quantized). The $\hat{\boldsymbol{\theta}}_{CCPI}$ parameters are computed through (5) for the testbed environment, obtaining the estimates listed in Tab. 1, which are compared with the $\hat{\boldsymbol{\theta}}_{DCPI}$ estimates affected by the uniform quantization in the gossip messages of the DCPI approach.

Parameter	DCPI	CCPI
$\hat{\beta}$	-40.5	-35.13
$\hat{\gamma}$	1.94	2.62

Table 1. Channel model identification using the CCPI and DCPI methods.

Fig. 8 shows the two channel models, described analytically by $P^{rx}(d) = \hat{\beta} - 10\hat{\gamma}\log_{10}(d)$ as a function of the distance d. Although the parameters identified through the two procedures differ in the absolute values, interestingly the two models are very close in the range of higher interest ($d \in [3, 9]$). This consideration justifies the possible employment of the DCPI in practical WSN applications, where the distributed technique may result more suitable for the less demanding requirements (in terms of transmission policy, communication flow, computational burden) with respect to the centralized approach. In particular, the advantage of using a distributed approach is clear when considering the communication burden, which affects also the energy consumption: both these issues are of paramount importance in networked systems composed by limited resource embedded devices.

6.3 Localization experiments

Actually, as a final experiment, we evaluate the performance of the DCPI technique and the goodness of the derived channel model by studying a typical application, that of localization in a WSN, to understand how the discrepancy possibly induced by the identification procedure may affect the localization result.

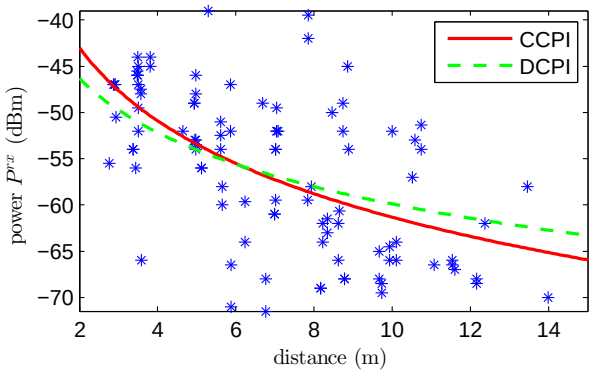

Fig. 8. Channel model identification with the CCPI (solid red line) and the DCPI (dashed green line) algorithm. The blue stars are the points of the RSS data set.

In this sense, the localization problem is solved here using the algorithm described in Bertinato et al. (2008) and based on the channel model considered in this paper. In short, this algorithm applies a maximum-likelihood method to estimate the distances of a mobile node in a network of fixed nodes (where the neighbors of the former dynamically change during the motion) by basically inverting the relation of Eq. 4. Then, the algorithm computes a LS estimation of the position of the mobile node by exploiting the estimated distances.

Experiments with different paths in the environment have been studied (an example is reported in Fig. 9) and the mean localization error is computed as

$$\frac{1}{K}\sum_{k=1}^{K} ||\hat{\mathbf{z}}_{\mathbf{0}}(k) - \mathbf{z}_{\mathbf{0}}(k)||,$$

where $\mathbf{z}_{\mathbf{0}}(k) = [x_0(k), y_0(k)]$ is the real position in the $\mathbb{R}^2$ plane of the mobile sensor and $\hat{\mathbf{z}}_{\mathbf{0}}(k)$ is the estimated position of the mobile sensor returned by the localization process in each $k = 1, \ldots, K$ time steps.

The results, with mean localization error in the range $[1.2-2.8m]$ according to the different path, are well in agreement with those from similar experiments in Bertinato et al. (2008), Cenedese et al. (2010).

This observation also hints at the local nature of the channel model and promotes the employment of a local identification procedure apt for a distributed approach.

Fig. 9. Example of one path chosen for the localization experiments.

Indeed, the localization experiments support the idea that this distributed approach can be adopted in similar applications to get the parameters describing communication channel, thus avoiding the use of a centralized method.

7. CONCLUSIONS

In this work, a distributed procedure to identify the parameters of a RF-channel model is proposed. The approach resorts to a quantized consensus algorithm, where different kind of quantizers are compared and discussed.

A set of validation experiments are then conducted, both in simulation and in a real experimental setup, to assess the quantized distributed procedure. In particular, for the real scenario experiments, the channel identification procedure is applied to the RF-based localization problem, whose performance is strongly affected by a correct modeling of the channel.

The results of these experiments prove the validity of the distributed approach for channel modeling as proposed in the paper and show consistency with the centralized identification, while being beneficial in terms of resource requirements with respect to the centralized approach.

Future works will focus on a more comprehensive and quantitative comparison between the CCPI an DCPI approaches, and a deeper study on the deviation error bounds for each quantization method.

REFERENCES

Akyildiz, I.F. and Kasimoglu, I.H. (2004). Wireless sensor and actor networks: research challenges. *Ad Hoc Networks*, 2(4), 351–367.

Bardella, A., Bui, N., Zanella, A., and Zorzi, M. (2010). An experimental study on IEEE 802.15.4 multichannel transmission to improve RSSI-based service performance. In *Proc. of the 4th Int. Conf. on Real-world wireless sensor networks (REALWSN10)*, 154–161.

Bertinato, M., Ortolan, G., Maran, F., Marcon, R., Marcassa, A., Zanella, F., Zambotto, P., Schenato, L., and Cenedese, A. (2008). RF localization and tracking of mobile nodes in wireless sensors networks: Architectures, algorithms and experiments. Technical report, University of Padova. Available online: `http://paduaresearch.cab.unipd.it/1046/` (accessed: Mar. 2014).

Bolognani, S., Del Favero, S., Schenato, L., and Varagnolo, D. (2010). Consensus-based distributed sensor calibration and least-square parameter identification in WSNs. *Int. J. of Robust and Nonlinear Control*, 20(2), 176–193.

Boyd, S., Ghosh, A., Prabhakar, B., and Shah, D. (2006). Randomized gossip algorithms. *IEEE/ACM Trans. Netw.*, 14, 2508–2530.

Carli, R., Frasca, P., Fagnani, F., and Zampieri, S. (2010a). Gossip consensus algorithms via quantized communication. *Automatica*, 46, 70–80.

Carli, R., Bullo, F., and Zampieri, S. (2010b). Quantized average consensus via dynamic coding/decoding schemes. *Int. J. of Robust and Nonlinear Control*, 20(2), 156–175.

Cenedese, A., Ortolan, G., and Bertinato, M. (2010). Low density wireless sensors networks for localization and tracking in critical environments. *IEEE Trans. on Vehicular Technology*, 59(6), 2951–2962.

Durgin, G., Rappaport, T.S., and Xu, H. (1998). Measurements and models for radio path loss and penetration loss in and around homes and trees at 5.85 GHz. *IEEE Trans. on Communications*, 46(11), 1484–1496.

Goldsmith, A. (2005). *Wireless Communications*. Cambridge University Press, New York, NY, USA.

Gudmundson, M. (1991). Correlation model for shadow fading in mobile radio systems. *Electronics Letters*, 27(23), 2145–2146.

Hansen, J. and Reitzner, M. (2004). Efficient indoor radio channel modeling based on integral geometry. *IEEE Trans. on Antennas and Propagation*, 52(9), 2456–2463.

Hashemi, H. (1993). The indoor radio propagation channel. *Proc. of the IEEE*, 81(7), 943–968.

Hespanha, J.P., Naghshtabrizi, P., and Xu, Y. (2007). A survey of recent results in networked control systems. *Proc. of IEEE* Special Issue on Technology of Networked Control Systems, 95(1), 138–162.

Kaya, A., Greenstein, L., and Trappe, W. (2009). Characterizing indoor wireless channels via ray tracing combined with stochastic modeling. *IEEE Trans. on Wireless Communications*, 8(8), 4165–4175.

Lewis, P. (2006). TinyOS programming. Available online: `http://www.tinyos.net/tinyos-2.x/doc/pdf/tinyos-programming.pdf` (accessed: Mar. 2014).

Mesbahi, M. and Egerstedt, M. (2010). *Graph Theoretic Methods for Multiagent Networks*. Princeton University Press, Princeton, NJ, USA.

Oestges, C., Czink, N., Bandemer, B., Castiglione, P., Kaltenberger, F., and Paulraj, A. (2010). Experimental Characterization and Modeling of Outdoor-to-Indoor and Indoor-to-Indoor Distributed Channels. *IEEE Trans. on Vehicular Technology*, 59(5), 2253–2265.

Oliveira, L.M. and Rodrigues, J.J. (2011). Wireless sensor networks: a survey on environmental monitoring. *J. of Communications*, 6(2), 143–151.

Rao, C.R., Rao, R., and Toutenburg, H. (1997). *Linear Models: Least Squares and Alternatives*. Springer.

Rappaport, T.S. (2002). *Wireless Communications: Principles and Practice*. Prentice Hall.

Söderström, T. and Stoica, P. (1988). *System identification*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA.

Tylavsky, D. and Sohie, G. (1986). Generalization of the matrix inversion lemma. *Proc. of the IEEE*, 74(7), 1050–1052.

Valenzuela, R.A. (1994). Ray tracing prediction of indoor radio propagation. In *IEEE Int. Symp. on Personal,*

Indoor and Mobile Radio Communications, volume 1, 140–144.

Wysocki, T. and Zepernick, H.J. (2000). Characterisation of the indoor radio propagation channel at 2.4 GHz. *J. of Telecommunications and Information Theory*, 1(3–4), 84–90.

Yick, J., Mukherjee, B., and Ghosal, D. (2008). Wireless sensor network survey. *Comput. Netw.*, 52(12), 2292–2330.

Zanella, F. (2013). Zophar – a TinyOS application for least-square channel identification using consensus algorithm. Available online: `https://github.com/r4m/zophar` (accessed: Mar. 2014).