www.ietdl.org

Published in IET Communications
Received on 15th November 2013
Revised on 12th January 2014
Accepted on 23rd January 2014
doi: 10.1049/iet-com.2013.1019

ISSN 1751-8628

Adaptive multi-task compressive sensing for localisation in wireless local area networks

Rongpeng Li[1], Zhifeng Zhao[1], Yuan Zhang[1], Jacques Palicot[2], Honggang Zhang[1,3]

[1]Department of Information Science and Electronic Engineering, Zhejiang University, Zheda Road 38, Hangzhou 310027, People's Republic of China
[2]Supélec/IETR, Avenue de la Boulaie CS 47601, Cesson-Sévigné Cedex, France
[3]Université Européenne de Bretagne (UEB) & Supélec, Avenue de la Boulaie CS 47601, Cesson-Sévigné Cedex, France
E-mail: lirongpeng@zju.edu.cn

Abstract: The spatially distributed sparsity of the mobile devices (MDs) in indoor wireless local area networks (WLANs) makes compressive sensing (CS) based localisation algorithms feasible and desirable. In this Letter, the authors exploit the most recent developments in CS to efficiently perform localisation in WLANs and design an accurate indoor localisation scheme by taking advantage of the theory of multi-task Bayesian CS (MBCS). The proposed scheme assembles the strength measurements of signals from the MDs to distinct access points (APs) and jointly utilises them at a central unit or a specific AP to achieve localisation, thus being able to alleviate the burden of MDs while simultaneously giving a precise estimation of the locations. Afterwards, they give a deeper insight into the localisation problem in more practical scenarios with varying number of MDs and investigate two different adaptive algorithms to meet the satisfactory localisation error requirement. Compared with the conventional MBCS algorithms, simulation results validate that both adaptive algorithms could provide superior localisation accuracy and exhibit stronger resilience to the changes in the number of MDs.

1 Introduction

1.1 Motivation

Accompanying with the explosive advancement of multimedia-rich networking data services, location-based services (LBSs) have emerged in recent years [1, 2] and attracted considerable interest and wide applications in various fields, such as on-line social network [3], remote healthcare [4], e-commerce [5], personalised information delivery [6] and so on. Furthermore, location and mobility information is also being exploited to reduce the energy consumption in energy-constrained wireless communication [2, 7] and becomes essential for mobile user management in seamless and ubiquitous communications. As a result, researchers come up with numerous techniques to achieve localisation, with the aid of global position system (GPS) [8] and base stations [9] etc. Unfortunately, the cellular-based methods cannot provide adequate accuracy required in indoor applications [10]. On the other hand, although widely used in outdoor environment, GPS is often energy-consumable and not suitable for central urban or indoor area with heavy building shadowing [11]. In this light, owing to the ever-growing universal existence of wireless local area networks (WLANs), some novel localisation schemes have been proposed by relying on the received signal strength (RSS) from WLANs to estimate the locations of mobile devices (MDs). Notably, in spite of the possibility to adopt other positioning metrics like time of arrival and angle of arrival [8, 10], RSS is generally the featured choice in the context of localisation [9, 12–16]. However, because of the effect of noise and channel impairments, there exists a challenge to cope with the uncertainty in RSS measurements [16]. Therefore a Bayesian approach [17] could be exploited by firstly imposing a probabilistic model on the RSS measurements, and then trying to obtain the posterior distribution.

In this paper, we assume that the indoor positioning system can collect the strength measurements of signals from MDs to access points (APs) in the area of interest and assemble them at a network central unit (CU) or a specific AP to perform the localisation in WLANs. This methodology could provide several advantages. Firstly, the algorithms running on a CU could alleviate the burden of MDs, which usually have limited processing power, short battery lifetime and small memory [18]. Secondly, the location information could be more conveniently applied, since APs or LBS servers could directly utilise it to optimise the whole system. Besides, since the MDs are only located at few places of one large physical space, the locations of MDs can be regarded as sparse signals if we view them as a whole. Therefore the centralised localisation methodology could benefit from a plethora of existing compressive sensing (CS) algorithms [19–21], which explore the fact that a small collection of an originally sparse or compressible signal's linear projections contains sufficient information for signal recovery and thus require far fewer measurements than the Nyquist sampling theorem to accurately reconstruct the signal.

Motivated by the discussion above, we design a multi-task Bayesian CS (multi-task BCS, MBCS) [22] based localisation scheme. The term 'multi-task' implies that this scheme would take advantage of the intra- and inter-signal correlations of the RSS measurements at different APs. As a result, it could decrease the total amount of data required for accurate localisation. Besides, BCS algorithms are able to provide a criterion named 'error bars' [23] to gauge the accuracy of the reconstruction vector and make the adaptation of measurement number possible [24]. Following our previous works [25], we propose two adaptive MBCS algorithms to guarantee the effectiveness and accuracy when the number of MDs varies.

1.2 Related works

Thus far, there exists a substantial body of works towards the localisation problem in WLANs. In [12], Rizos *et al.* collected the RSS measurements and compared them with a pre-built radio environment map. The estimated position would be determined by the k-nearest neighbours (KNN) method, namely the point with the smallest Euclidean distance to the centroid of KNN in the map. In [10], Kushki *et al.* modified the KNN scheme and proposed a kernel-based technique to improve the positioning accuracy. As for CS-based localisation schemes, Zhang *et al.* [14] showed the corresponding feasibility theoretically. Meanwhile, researchers implemented their CS-based localisation schemes on experimental networks [15, 16]. For example, Au *et al.* [16] utilised CS to localise the position of one MD and exploited Kalman filtering to track the corresponding movement. However, these schemes [14–16] suffer from the uncertainty in RSS measurements and could not adapt to the varying number of MDs. Besides, they are still carried out on the MDs by averaging the RSS values, incurring indispensable yet energy inefficient pre-processing procedures such as AP selection [14] and orthogonalisation [15]. Nikitaki and Tsakalides [18] intended to bypass these procedures by addressing the localisation problem on the AP side. However, the authors did not consider the intra- and inter-signal correlations in the RSS measurements and provided no criterion to determine the sufficient number of measurements as well.

The rest of this paper is structured as follows. In Section 2, we give an introduction to the fundamentals of CS theory and describe the relevant MBCS framework. Section 3 presents the system model and formulates our compressed sensing based localisation scheme. Sections 4 and 5 provide two adaptive algorithms and present the corresponding simulation results. Finally, Section 6 concludes our works and offers our future research direction.

2 Basics of Bayesian compressive sensing

2.1 Compressive sampling process

The Nyquist sampling theorem demonstrates that signals, images, videos and other data can be exactly recovered from a set of uniformly spaced samples taken at the so-called Nyquist rate, which is twice the highest frequency in the signal of interest [26]. However, the resulting Nyquist rate in most situations is redundant such that we end up with far more necessary samples. The recently developed theory of CS states that if a real-world signal has a sparse representation in a certain transform bases, then it is possible to recover the signal with significantly fewer samples or measurements as required by the Nyquist rate. In this section, we give a brief review of the theory of CS.

Consider a signal $\boldsymbol{x} \in \mathbb{R}^{N\times 1}$, which can be represented in terms of $N \times N$ transform basis $\boldsymbol{\Psi}$ such that

$$\boldsymbol{x} = \sum_{i=1}^{N} s_i \psi_i = \boldsymbol{\Psi s} \tag{1}$$

The signal x is called K-sparse if its sparse representation $\boldsymbol{s}$ has at most K non-zero entries $\left(\|\boldsymbol{s}\|_{l_0} \le K\right)$, where $K \ll N$ and the l_0-norm $\|\boldsymbol{s}\|_{l_0}$ of the vector $\boldsymbol{s}$ is defined as the number of its non-zero components.

Consider a linear projection process that computes $M < N$ inner products between $\boldsymbol{x}$ and a set of $N \times 1$ vectors $\{\boldsymbol{\phi}_j\}_{j=1}^{M}$ as in $y_j = \phi_j^{\mathrm{T}}\boldsymbol{x}$, where $\{\cdot\}^{\mathrm{T}}$ represents the transpose operation. Collect the measurements y_j and form an $M \times 1$ vector $\boldsymbol{y}$, by arranging the projection vectors $\{\boldsymbol{\phi}_j^{\mathrm{T}}\}_{j=1}^{M}$ as rows of an $M \times N$ measurement matrix $\boldsymbol{\Phi}$. After substituting $\boldsymbol{x}$ with (1), the whole projection process can be represented as follows

$$\boldsymbol{y} = \boldsymbol{\Phi x} + \varepsilon = \boldsymbol{\Phi\Psi s} + \varepsilon = \boldsymbol{\Theta s} + \varepsilon \tag{2}$$

where $\boldsymbol{\Theta} = \boldsymbol{\Phi\Psi}$ is an $M \times N$ matrix and $\varepsilon = [\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_N]^{\mathrm{T}}$ represents the noisy environment effects with each entry ε_i being a zero-mean Gaussian random variable with variance σ^2. The CS theory aims to solve this under-determined problem and find an appropriate $\boldsymbol{x}$ satisfying (2).

2.2 Recovery solutions in compressive sensing

In this section, we detail the recovery solutions in CS. Actually, traditional techniques like least-square algorithms [27] could not readily handle the under-determined problem in (2), since the knowledge of the non-zero element positions of $\boldsymbol{x}$ is usually unknown beforehand. Hence, rather than requiring the knowledge of zeros element positions, the recovery algorithms in CS solve this under-determined problem by the means of greedy algorithms [28, 29].

Tsaig and Donoho [30] state that if $M \ge 2K$, the recovery problem, under specific conditions, could be treated as follows

$$\hat{s} = \arg\min_{\boldsymbol{s}} \|\boldsymbol{s}\|_{l_0} \quad \text{s.t. } \left\|\boldsymbol{y} - \boldsymbol{\Phi\Psi s}\right\|_{l_2} \le \epsilon \tag{3}$$

However, (3) requires to solve a non-convex problem which is usually 'NP-hard'. In [31], it has been proved that we can approximate the aforementioned l_0 optimisation problem by its l_1 relaxation [For an integer real number $p \ge 1$ and a vector $\boldsymbol{z}$, the corresponding l_p-norm $\|\boldsymbol{z}\|_{l_p}$ equals $\left(\sum_i |z_i|^p\right)^{(1/p)}$.] with a bounded error only if the matrix $\boldsymbol{\Theta} = \boldsymbol{\Phi\Psi}$ satisfies the restricted isometry property (RIP)

$$\hat{s} = \arg\min_{\boldsymbol{s}} \|\boldsymbol{s}\|_{l_1} \quad \text{s.t. } \left\|\boldsymbol{y} - \boldsymbol{\Phi\Psi s}\right\|_{l_2} \le \epsilon \tag{4}$$

Here, a matrix $\boldsymbol{\Theta}$ is said to conform to the RIP for order K with constant $\delta_K \in (0, 1)$ if

$$(1 - \delta_K)\,\|\boldsymbol{c}\|_{l_2}^2 \le \|\boldsymbol{\Theta c}\|_{l_2}^2 \le (1 + \delta_K)\,\|\boldsymbol{c}\|_{l_2}^2 \tag{5}$$

for any N-length vector $\boldsymbol{c}$ satisfying $\|\boldsymbol{c}\|_{l_0} \le K$. Although designing and evaluating a suitable measurement matrix $\boldsymbol{\Phi}$

to ensure the RIP for $\boldsymbol{\Theta} = \boldsymbol{\Phi}\boldsymbol{\Psi}$ is challenging, there fortunately exists a large set of matrices that have proven to obey the RIP. Generally, if $\boldsymbol{\Phi}$ is a random matrix with identical independently distributed (i.i.d.) Gaussian or Bernoulli entries or a matrix made up of randomly selected rows of an orthogonal matrix [e.g. the discrete Fourier transform (DFT)], the RIP would be satisfied [32]. Typically, conditioned on the RIP, the recovery problem in (4) could be solved by techniques such as basis pursuit (BP) [29], orthogonal matching pursuit (OMP) [28] and OMP's variants [14]. Although recovery through l_1 optimisation proves to provide highly accurate solutions, the algorithms above often incur high computational complexity as well.

On the other hand, the problem could be solved from a Bayesian prospective as well. Compared with the aforementioned greedy algorithms, Bayesian compressed sensing (BCS) [23] recovers the signal based on the posterior probability instead of a single value and could yield better performance in reconstruction with noisy measurements y in terms of l_0-norm [33]. Therefore we apply the BCS algorithm to our localisation problem. Moreover, in order to combat the negative fading and shadowing effect on measurements and achieve even better recovery performance, we take into account the intra- and inter-signal correlations and adopt an MBCS framework [22]. Specifically, let us assume there are P sets of CS measurements $\{\boldsymbol{y}^i\}_{i=1}^P$, projected from P sets of original compressive signals $\{\boldsymbol{x}^i\}_{i=1}^P$, namely

$$\boldsymbol{y}^i = \boldsymbol{\Phi}^i\boldsymbol{x}^i + \varepsilon^i = \boldsymbol{\Phi}^i\boldsymbol{\Psi}^i\boldsymbol{s}^i + \varepsilon^i = \boldsymbol{\Theta}^i\boldsymbol{s}^i + \varepsilon^i, \quad \forall i \in \{1, \ldots, P\} \tag{6}$$

where each $\boldsymbol{x}^i \in \mathbb{R}^N$ exploits a disparate random measurement matrix $\boldsymbol{\Phi}^i \in \mathbb{R}^{M^i \times N}$ to derive $\boldsymbol{y}^i \in \mathbb{R}^{M^i}$, with every sparse vector in $\{\boldsymbol{s}^i\}_{i=1}^P$ similar or equal to each other. $\varepsilon^i \in \mathbb{R}^{M^i}$ denotes Gaussian noise and if $P = 1$, (6) would be simplified into (2), namely a single-task case. Obtained from repeated P experiments on similar scenarios or the same type of tasks, MBCS algorithm could collect a subset of highly correlated measurements, which is exactly suitable for the case of RSS-gathering process in our localisation scheme and contributes to the information-sharing between tasks.

Commonly, $\boldsymbol{\varepsilon}^i$ can be modelled as an M^i i.i.d. zero-mean multivariate Gaussian random variable with variance σ_0^2. As a result, conditioned on s^i and $\lambda_0 = 1/\sigma_0^2$, the likelihood function for $\boldsymbol{y}^i$ in (6) could be formulated as

$$p(\boldsymbol{y}^i|\boldsymbol{s}^i, \lambda_0) = (2\pi/\lambda_0)^{-M^i/2} \exp\left(-\frac{\lambda_0}{2}\left\|\boldsymbol{y}^i - \boldsymbol{\Theta}^i\boldsymbol{s}^i\right\|_{l_2}^2\right), \quad \forall i \in \{1, \ldots, P\} \tag{7}$$

The difficulty towards applying BCS lies in the typically intractable computation of the posterior probability function [34]. Hence, we are forced to accept some approximations and establish a hierarchical prior to the sparse vectors [23]. In MBCS, we just assign a common prior with common hyperparameters $\boldsymbol{\lambda} = [\lambda_1, \ldots, \lambda_N]$ to $\boldsymbol{s}^i$ and connect several recovery tasks together [23]

$$p(\boldsymbol{s}^i|\boldsymbol{\lambda}) = \prod_{j=1}^N \mathcal{N}(s_j^i|0, \lambda_j^{-1}), \quad \forall i \in \{1, \ldots, P\} \tag{8}$$

where $\mathcal{N}(\cdot|0, \lambda_j^{-1})$ represents the Gaussian distribution with zero mean and variance λ_j^{-1}. By the Bayes' rule, the posterior probability function for $\{\boldsymbol{s}^i\}_{i=1}^P$ can be derived by using (7) and (8), namely

$$p(\boldsymbol{s}^i|\boldsymbol{y}^i, \boldsymbol{\lambda}, \lambda_0) = \frac{p(\boldsymbol{y}^i|\boldsymbol{s}^i, \lambda_0)p(\boldsymbol{s}^i|\boldsymbol{\lambda})}{p(\boldsymbol{y}^i|\boldsymbol{\lambda}, \lambda_0)} = (2\pi)^{-(N+1/2)}\left|\Sigma^i\right|^{-(1/2)} \exp\left(-\frac{1}{2}(\boldsymbol{s}^i - \boldsymbol{\mu}^i)^{\mathrm{T}}(\Sigma^i)^{-1}(\boldsymbol{s}^i - \boldsymbol{\mu}^i)\right) \tag{9}$$

with mean and covariance given by

$$\boldsymbol{\mu}^i = \lambda_0\Sigma^i(\boldsymbol{\Theta}^i)^{\mathrm{T}}\boldsymbol{y}^i, \quad \boldsymbol{\Sigma}^i = (\lambda_0(\boldsymbol{\Theta}^i)^{\mathrm{T}}(\boldsymbol{\Theta}^i) + \Lambda)^{-1}, \quad \Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_N) \tag{10}$$

Subsequently, we have to estimate the hyperparameters $\boldsymbol{\lambda}$ and λ_0 in order to obtain the posterior probability function. The estimates for $\boldsymbol{\lambda}$ and λ_0 can be derived by maximising the logarithm marginal likelihood

$$\begin{aligned} L(\boldsymbol{\lambda}, \lambda_0) &= \sum_{i=1}^P \log p(\boldsymbol{y}^i|\boldsymbol{\lambda}, \lambda_0) \\ &= \sum_{i=1}^P \log \int p(\boldsymbol{y}^i|\boldsymbol{s}^i, \lambda_0)p(\boldsymbol{s}^i|\boldsymbol{\lambda})\,\mathrm{d}\boldsymbol{s}^i \\ &= -\frac{1}{2}\sum_{i=1}^P\left[M^i \log 2\pi + \log\left|\boldsymbol{C}^i\right| + (\boldsymbol{s}^i)^{\mathrm{T}}(\boldsymbol{C}^i)^{-1}\boldsymbol{s}^i\right] \end{aligned} \tag{11}$$

with the unit matrix $\boldsymbol{I}$ and

$$\boldsymbol{C}^i = \lambda_0^{-1}\boldsymbol{I} + \boldsymbol{\Theta}^i\Lambda^{-1}(\boldsymbol{\Theta}^i)^{\mathrm{T}}, \quad \forall i \in \{1, \ldots, P\} \tag{12}$$

Afterwards, differentiate (11) with respect to $\boldsymbol{\lambda}$ and λ_0, respectively, and then set the results to zero, yielding

$$\lambda_j^{\mathrm{new}} = \frac{P - \lambda_j\sum_{i=1}^P \boldsymbol{\Sigma}_{(j,j)}^i}{\sum_{i=1}^P\left(\boldsymbol{\Sigma}_j^i\right)^2}, \quad \forall j \in \{1, 2, \ldots, N\} \tag{13}$$

and

$$\lambda_0^{\mathrm{new}} = \frac{\sum_{i=1}^P (M^i - N + \sum_{j=1}^N \lambda_j \sum_{(j,j)}^i)}{\sum_{i=1}^P \left\|\boldsymbol{y}^i - \boldsymbol{\Theta}^i\boldsymbol{\mu}^i\right\|_{l_2}^2} \tag{14}$$

Note that $\boldsymbol{\lambda}^{\mathrm{new}}$ and λ_0^{new} are functions of $\{\boldsymbol{\mu}^i\}_{i=1}^P$ and $\{\boldsymbol{\Sigma}^i\}_{i=1}^P$, whereas $\{\boldsymbol{\mu}^i\}_{i=1}^P$ and $\{\boldsymbol{\Sigma}^i\}_{i=1}^P$ are pertinent to $\boldsymbol{\lambda}$ and λ_0 in (10). So we can alternatively iterate between (10), (13) and (14) until convergence. In this process, Wipf *et al.* [34] notes that in order to maximise the likelihood in (11), some entries of $\boldsymbol{\lambda}$ tend to infinity, which in turn means the corresponding entries in $\boldsymbol{\mu}^i$ tightly approach zero. In other words, the algorithm finally finds sparse vectors $\{\boldsymbol{s}^i\}_{i=1}^P$ and could estimate the location vector in terms of the average of $\{\boldsymbol{\mu}^i\}_{i=1}^P$. Besides, since the update of $\boldsymbol{\lambda}^{\mathrm{new}}$ requires the knowledge of the whole set $\{\boldsymbol{\mu}^i\}_{i=1}^P$ and $\{\boldsymbol{\Sigma}^i\}_{i=1}^P$, MBCS

IET Commun., 2014, Vol. 8, Iss. 10, pp. 1736–1744
doi: 10.1049/iet-com.2013.1019

algorithm implicitly exploits the intra- and inter-correlations in the RSS signal measurements.

3 BCS-based localisation scheme in WLANs

3.1 System model and problem description

In this paper, we primarily focus on a localisation scenario (Fig. 1) with a set of N non-overlapping grids. Meanwhile, the area of interest is covered by P APs and K MDs equipped with WLAN adapters. As mentioned in Section 1, the locations of the MDs can be estimated by comparing the current RSS measurements with a pre-set radio environment map of this area by the CU in the 'backbone' network. Here, the radio environment map, which is assumed to be readily generated, refers to a table of pre-measured RSS readings of a similar device corresponding to every grid of the area.

For clarity of representation, let an N-length vector $[0,\ 0,\ 1\cdots]^{\mathrm{T}} \in \mathbb{R}^{N\times 1}$ denote the logical location of an MD. To be specific, if the MD is located in the ith $[(i \in \{1,\ \ldots,\ N\})]$ grid, the corresponding ith element in the location vector would equal '1'. Therefore, if there exist $K \ll N$ MDs in the area of interest, the location vector $\boldsymbol{s} \in \mathbb{R}^{N\times 1}$, which needs to be estimated, contains K entries equalling 1. Thus, the location estimation problem could be reformulated into a sparse signal recovery problem, due to the facts that an MD can be reasonably located in exactly one of the grids within the whole area of interest and the number of MDs is very comparatively smaller than that of the grids in aggregate.

Although a number of available methods can be utilised to carry out the comparison between the RSS measurements and the radio environment map, we choose the BCS-based scheme to incorporate the inherent sparsity feature of the localisation problem and seek for the best matching, which will be verified to be more accurate and more efficient, lately.

3.2 BCS based localisation

Given the property of sparseness in the localisation scenario, the CS-based algorithms or their variants can be applied here. In this section, we would demonstrate how our BCS-based localisation scheme could accurately locate the MDs in a noisy environment with an optimal number of measurements.

In our BCS-based scheme, instead of the common transform bases such as DFT basis and discrete wavelet transform basis, we finally decide to choose the radio environment map matrices $\{\boldsymbol{\Psi}^i\}_{i=1}^{P}$ as the projection matrices, after taking into account the obvious sparse structure of the corresponding location vector $\boldsymbol{s}$. For each column vector of the radio environment map matrix $\boldsymbol{\Psi}^i$, $\psi_j^i \in \mathbb{R}^N$ denote the vector of collected RSS samples at grid j from the ith AP, wherein N indicates the sampling length and entries of ψ_j^i should be different because of the fading and shadowing effect. By means of concatenating the vectors for all the N grids, a CU will construct a single matrix $\boldsymbol{\Psi}^i \in \mathbb{R}^{N\times N}$ for the ith AP. Therefore the actual signals at the AP side can be expressed as a linear combination of several columns of $\boldsymbol{\Psi}^i$. In other words, the received signals $\boldsymbol{x}^i$ are the multiplication product of the radio map matrix $\boldsymbol{\Psi}^i$ and the $N\times 1$ location vector $\boldsymbol{s}$, which means that $\{\boldsymbol{s}^i\}_{i=1}^{P}$ in MBCS are the same and equal $\boldsymbol{s}$.

Now, assume that every AP has collected a train of RSS measurements from MDs at unknown locations. Then, considering the tremendous amount of measurements and the corresponding inherent sparsity property, each AP, say AP i would sample M^i measurements out of $\boldsymbol{x}^i \in \mathbb{R}^N$ from every MD according to the measurement matrix $\boldsymbol{\Phi}^i \in \mathbb{R}^{M^i\times N}$ and feedback them to the CU via backhaul link. When the incoherent measurement process completes, the overall RSS measurements for MDs associated with the ith AP can be described using (6) at the CU side, where P is the number of APs in this area and $\boldsymbol{\varepsilon}^i$ represents the Gaussian noise. $\boldsymbol{\Phi}^i$ takes the standard Gaussian matrix with its columns normalised to unit norm, thus guaranteeing the required incoherence between $\boldsymbol{\Phi}^i$ and $\boldsymbol{\Psi}^i$ to meet the RIP.

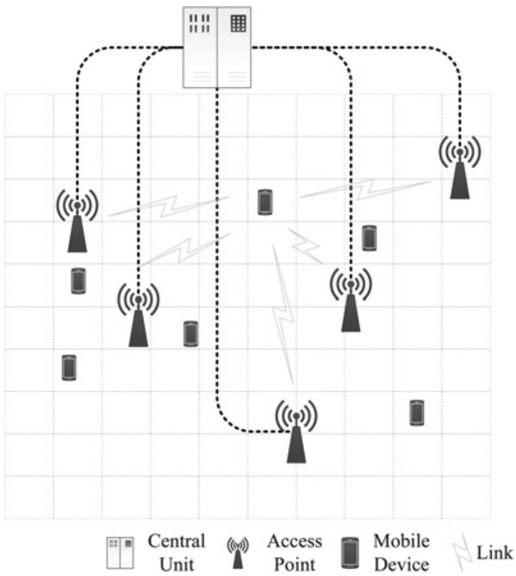

Fig. 1 *Localisation scenario in wireless local area networking environment*

Finally, the P sets of measurements $\boldsymbol{y}^i$ can be used jointly to recover the location vector. Therefore we can simply take advantage of MBCS in Section 2 to solve this reconstruction problem. After completing reconstruction process, the locations of MDs can be precisely determined based on the average of $\{\boldsymbol{\mu}^i\}_{i=1}^{P}$.

4 Adaptive multi-task CS algorithm for localisation in WLANs

The previous sections have addressed the means of BCS-based localisation scheme in WLANs. Yet, the varying number of MDs in mobility oriented networks makes it challenging to accurately determine the required measurement number in prior and thus suffering to directly adopt the BCS-based localisation scheme. Meanwhile, the criterion, which measures the localisation error and balances whether the current measurement is sufficiently precise, is worthy to know. However, it is still unavailable in the previous localisation researches. In this section, we take into the aforementioned considerations and aim to tackle how to achieve the optimal measurement number for location reconstruction, by deriving an adaptive multi-task BCS algorithm, or the AMBCS algorithm.

Recall that, the location vector $\boldsymbol{s}$ is finally estimated as the average of $\{\boldsymbol{\mu}^i\}_{i=1}^{P}$ or the mean of a multivariate Gaussian distribution in (9). Indeed, the distribution also provides a by-product, namely the covariance $\boldsymbol{\Sigma}^i$ in (10), which can be interpreted as 'error bars' in [23, 35]. As the name suggests, the 'error bars' could be applied as a metric on the accuracy of the recovered location vector $\hat{\boldsymbol{s}}$ and furthermore as a criterion to decide how many measurements are necessary. Intuitively, if the value of 'error bars' is larger than anticipation, it implies that the recovered vector still remains quite uncertain and it needs more RSS measurements to meet the accuracy requirement. On the opposite side, if the 'error bars' are smaller than expected, which means the measurements are already sufficient, then we could attempt to decrease the number to reduce the possibly unnecessary redundant measurements.

These insights inspire us to understand the impact of measurement number on the uncertainty for localisations or the 'error bars'. Beforehand, we need to define the formal definition of 'error bars'.

Definition 1: For each task $i \in \{1, \ldots, P\}$ in an MBCS algorithm, the 'error bars' is defined as the determination of the corresponding covariance $\boldsymbol{\Sigma}_v^i$ in Lemma 2.
Indeed, we observe that the increase in measurement number could theoretically contribute to reduce the 'error bars'.

Theorem 1: Given a set of hyperparameters $\boldsymbol{\lambda}$ and λ_0, a radio map matrix $\{\boldsymbol{\Psi}^i\}_{i=1}^P$ and M^i measurements $\{\boldsymbol{y}^i\}_{i=1}^P$ generated based on a current measurement matrix $\{\boldsymbol{\Phi}^i\}_{i=1}^P$, for any $i \in \{1, \ldots, P\}$, if another $(M^i + 1)$th measurement $y_{M^i+1}^i$ is sampled out of $\boldsymbol{x}^i$ according to a new measurement vector $\{\boldsymbol{r}^i\}^{\mathrm{T}}$, the uncertainty for recovering the location vector $\boldsymbol{s}$, which is measured by the 'error bars' $\det(\Sigma_v^i)$, would certainly decrease.
We leave the proof of Theorem 1 in Appendix. Furthermore, the corresponding proof offers the following corollary.

Corollary 1: To decrease the 'error bars' to the possibly least, for any $i \in \{1, \ldots, P\}$, the newly added measurement vector $\boldsymbol{r}^i$, on which a new sample is based, should be $\boldsymbol{r}_o^i = arg\,\underset{\boldsymbol{r}}{max}\, \boldsymbol{r}^{\mathrm{T}} \boldsymbol{\Psi}^i (\boldsymbol{\Sigma}^i)(\boldsymbol{\Psi}^i)^{\mathrm{T}} \boldsymbol{r}$.

Remark: Theorem 1 tells that according to the 'error bars', it is feasible to exploit the AMBCS algorithm to obtain a satisfactory localisation accuracy by dynamically increasing or decreasing the number of measurements. Besides, based on Corollary 1, the AMBCS algorithm could be further modified to decrease the 'error bars' at a faster pace. Specifically, in order to achieve the maximum of $(\boldsymbol{r}^i)^{\mathrm{T}} \boldsymbol{\Psi}^i (\boldsymbol{\Sigma}^i)(\boldsymbol{\Psi}^i)^{\mathrm{T}} \boldsymbol{r}$, we can design the new measurement vector $\boldsymbol{r}^{\mathrm{T}}$, by performing an eigen-decomposition of the matrix $\boldsymbol{Q} = \boldsymbol{\Psi}^i (\boldsymbol{\Sigma}^i)(\boldsymbol{\Psi}^i)^{\mathrm{T}}$ and assigning the eigenvector with the largest eigenvalue to $\boldsymbol{r}$. Hence, the rate, at which the uncertainty of the recovered location vector diminishes, could be optimised. By merging this adaptive idea to the localisation problem, we could design a greedily adaptive MBCS (GAMBCS) algorithm, which could save the computational cost of repetitively adding the measurements and performing the MBCS operations.
Finally, we summarise the AMBCS algorithm for localisation in Algorithm 1.

5 Simulation analysis and numerical results

In this section, we verify the localisation performance of our proposed BCS-based algorithm and its variants, by simulating in an area of $20\,\mathrm{m} \times 20\,\mathrm{m}$ with discretised grid size of $1.0\,\mathrm{m} \times 1.0\,\mathrm{m}$. Correspondingly, the length of the location vector would be $N = 400$. Moreover, we assume there exist $P = 6$ APs altogether in this Gaussian noise-corrupted environment. For ease of comparison, we assume the number of CS measurements M^i for each AP i (each task i) is equivalent. Meanwhile, we adopt the indoor signal propagation model in [36] to calculate the collected RSS at the AP side. Besides, all the simulation results presented hereafter are an average of 100 independent Monte Carlo runs.

Algorithm 1

Initialisation: For $i \in \{1, 2, \cdots, P\}$, initialise the initial measurement number $M^i = M_0^i$ and generate the radio map matrix by $\boldsymbol{\Psi}^i$ by collecting RSS measurements from MDs at the CU side. Initialise the hyperparameters $\boldsymbol{\lambda}$ and λ_0.

Repeat:

1) APs are assigned a $\boldsymbol{\Phi}^i \in \mathbb{R}^{M^i \times N}$, for $\forall i \in \{1, 2, \cdots, P\}$;
2) APs obtain the measurements vector $\{\boldsymbol{y}^i\}_{i=1}^P$ according to $\boldsymbol{\Phi}^i$ for $\forall i \in \{1, 2, \cdots, P\}$ and feed them back to the CU side via the backhaul link;
3) For $\forall i \in \{1, 2, \cdots, P\}$, CU calculates the mean $\boldsymbol{\mu}^i$ and variance $\boldsymbol{\Sigma}^i$ of posterior probability function $\boldsymbol{s}^i$ by Eq. (9) and Eq. (10) and updates the hyperparameters $\boldsymbol{\lambda}$ and λ_0 by Eq. (13) and Eq. (14);
4) CU notifies AP $i \in \{1, 2, \cdots, P\}$ to adaptively adjust the measurement number:
   a) If the "error bars" $\det(\boldsymbol{\Sigma}_v^i)$ exceeds a predefined accuracy requirement E_b, CU notifies AP i to perform a new measurement. Return to Step 2.
   b) Otherwise, CU notifies AP to decrease the measurement number by 1 and remove the last row vector from $\boldsymbol{\Phi}^i$. CU calculates the average of $\{\boldsymbol{\mu}^i\}_{i=1}^P$ and assigns it as the recovered location vector $\hat{\boldsymbol{s}}$.

End Repeat

Output: the recovered location vector $\hat{\boldsymbol{s}}$.

Fig. 2 *Adaptive multi-task compressive sensing (AMBCS) algorithm for localisation in WLANs*

On the other hand, we evaluate the performance based on the mean localisation error (MLE), which is defined between the recovered and corresponding original location vectors (e.g. $\{\hat{\boldsymbol{s}}^i\}_{i=1}^P$ and $\{\boldsymbol{s}^i\}_{i=1}^P$) as follows

$$\mathrm{MLE} = \frac{1}{P} \sum_{i=1}^{P} \frac{\|\hat{\boldsymbol{s}}^i - \boldsymbol{s}^i\|_{l_2}}{\|\boldsymbol{s}^i\|_{l_2}} \tag{15}$$

Meanwhile, we calculate the 'normalised error bars' (NEB) for MBCS algorithm by (Fig. 2):

Equation (16) to take into account the localisation variance of recovered location vectors

$$\mathrm{NEB} = \frac{1}{P} \sum_{i=1}^{P} \left(\det(\boldsymbol{\Sigma}_v^i)\right)^{\left(1/\|\hat{\boldsymbol{s}}^i\|_{l_0}\right)} \tag{16}$$

Here, the $\left(1/\|\hat{\boldsymbol{s}}^i\|_{l_0}\right)$ order of $\det(\boldsymbol{\Sigma}^i)$ makes it possible to avoid the negative influence incurred by the varying number of MDs.

Firstly, Fig. 3 presents an illustration of localisation performance for the MBCS algorithm and AMBCS algorithm. Therein, every AP collects 60 RSS measurements interfered with zero-mean Gaussian noise power ($\sigma_0 = 0.005$) and feedbacks them to the CU. It can be observed that the MBCS algorithm could provide almost precise results. However, because of the increase in localisation variance (heavier colour) at the static number of measurements in the MBCS algorithm, the localisation confidence would be impaired when more MDs emerge. However, the proposed AMBCS algorithm could offer the result with nearly the same confidence when the number of MDs varies.

doi: 10.1049/iet-com.2013.1019

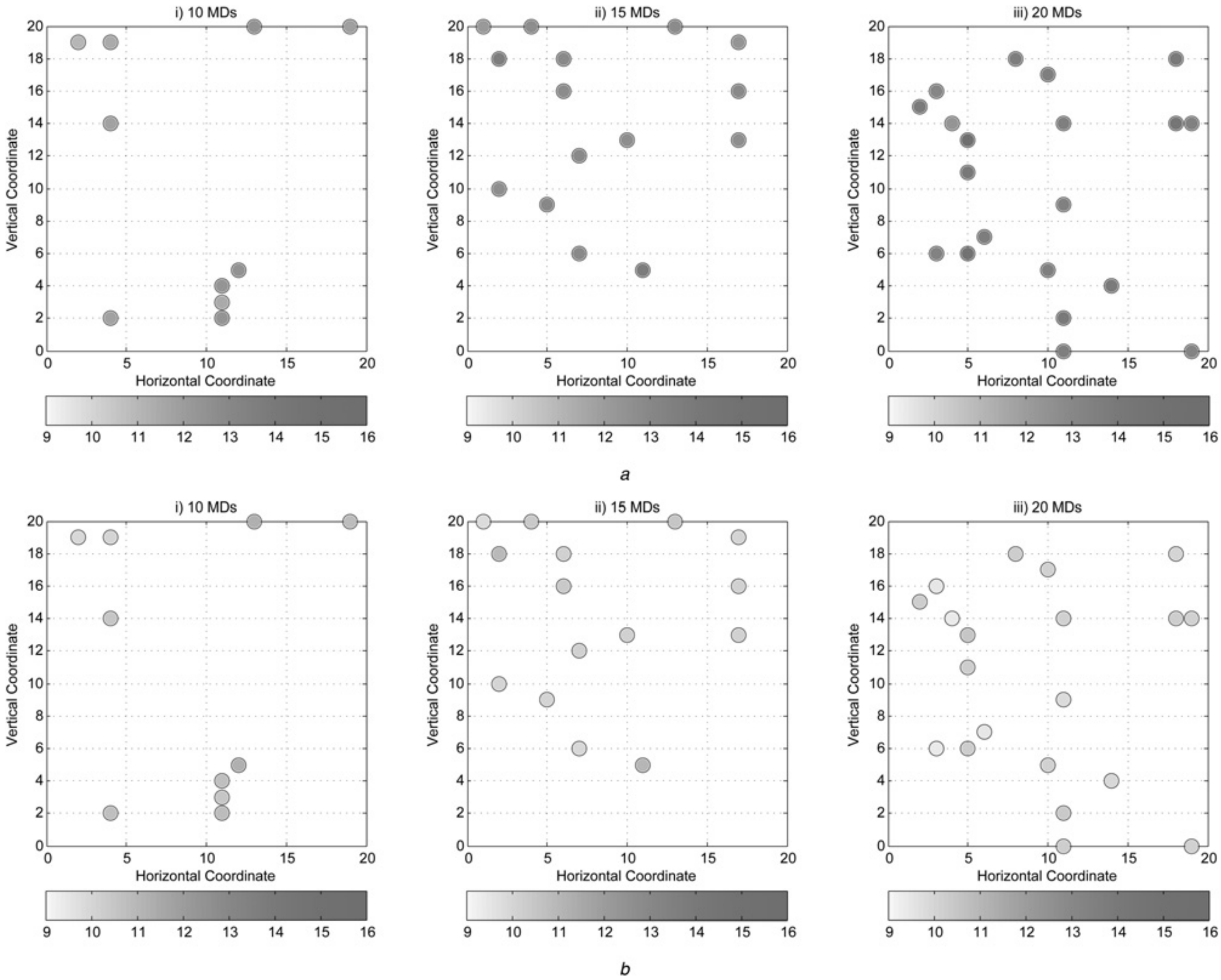

Fig. 3 *Localisation performance in three scenarios for the MBCS algorithm (a) and the AMBCS algorithm (b)*

Circles represent the real locations of MDs, whereas the dotted ones represent the estimated locations and the corresponding colour reflects the value of localisation variance by (16)
a MBCS algorithm
b AMBCS algorithm

Fig. 4 compares the MLE metric-based localisation performance among algorithms such as MBCS, AMBCS, GAMBCS, single-task BCS (SBCS), single-task BP (SBP) and single-task OMP (SOMP). Notably, it is worthwhile to note here that the single-task algorithm means that every AP aims to separately recover the location vectors, depending on its own measurements, and the algorithm finally obtains the final estimation result by combining and averaging them. Fig. 4 shows the localisation error decreases by collecting more measurements. Besides, MBCS algorithm could yield a significantly better performance using the same number of measurements by taking advantage of the shared knowledge in inter-AP measurements. Moreover, we also examine the localisation performance of both adaptive algorithms when the initial number of measurements is varying and the threshold for NEB is set to be 70. We can find when the number of measurements is small, the adaptive algorithms could provide superior performance than MBCS by dynamically adding more measurements.

Fig. 5 continues the evaluation under the same scenario assumptions using MBCS algorithm and shows that

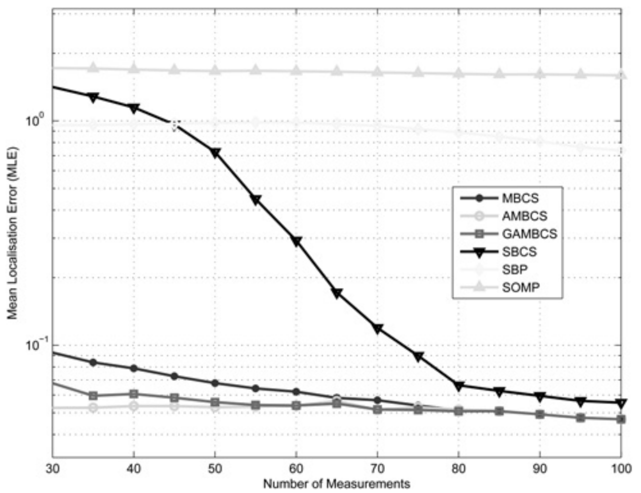

Fig. 4 *Localisation performance against varying number of measurements for the algorithms when there exist 10 MDs and noise power $\sigma_0 = 0.005$*

For AMBCS and GAMBCS, the initial number of measurements is varying and the NEB is set to be 70

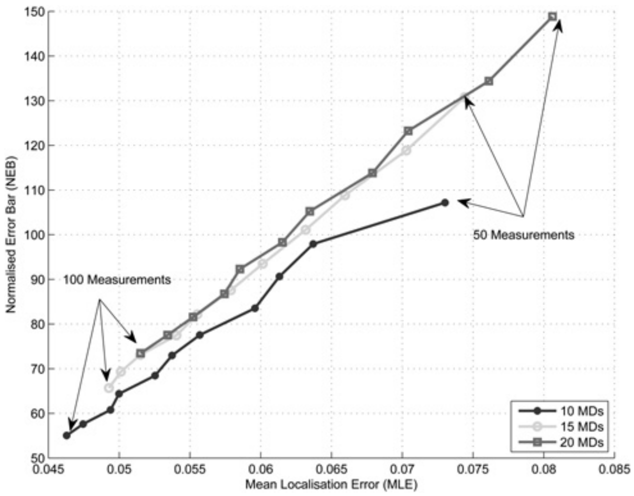

Fig. 5 *Relationship between MLE and 'NEBs' when the MBCS algorithm is used and noise power $\sigma_0 = 0.005$*

the MLE performance monotonically varies along with the NEB performance. Besides, it can be observed that the curves under different number of MDs are tightly close while exhibiting the same varying trend. Consequently, given that MLE could not be obtained in practical scenarios, NEB could be a strong alternative and help to determine the necessary number of measurements for a reliable localisation performance.

Previous simulations demonstrate the superior localisation performance of AMBCS and GAMBCS in the scenarios where the number of MDs is steady. Here, we examine the response of our adaptive algorithms to varying number of MDs. We set the initial number of measurement and NEB to be 50 and 70, respectively. Hence, according to Fig. 5, the default value (e.g., 70) for 'error bars' implies that if we adopt the MBCS algorithm the MLE will be below 0.055 and the final number of measurements will require 70–90, depending on the exact number of MDs. During the simulation process, we assume another five MDs emerge at the sixth slot and disappear at the eleventh slot. The corresponding localisation performance is plotted with both adaptive algorithms in Fig. 6. It can be observed in Fig. 6, the adaptivity of both algorithms could be well accustomed to the sudden appearance of MDs without deteriorating the performance. Specifically speaking, the NEB in the localisation process would exceed the threshold when more MDs emerge. As a result, every AP will automatically sample more measurements to the CU according the method inside. For example, the AMBCS algorithm requires another 16 measurements to meet the NEB requirement (e.g., 70) and gradually reduce the measurements once they become unnecessary. Meanwhile, because of the greed choice of added observation vector, GAMBCS algorithm only requires another 2 measurements, but it offers slightly inferior performance.

In Fig. 7, we examine the impact of working APs on the performance of the AMBCS algorithm. We can find the loss of APs only causes limited negative impact on the corresponding performance. Moreover, AMBCS with fewer APs could still offer superior performance than MBCS, benefiting from the adaptation inside. Fig. 8 presents the

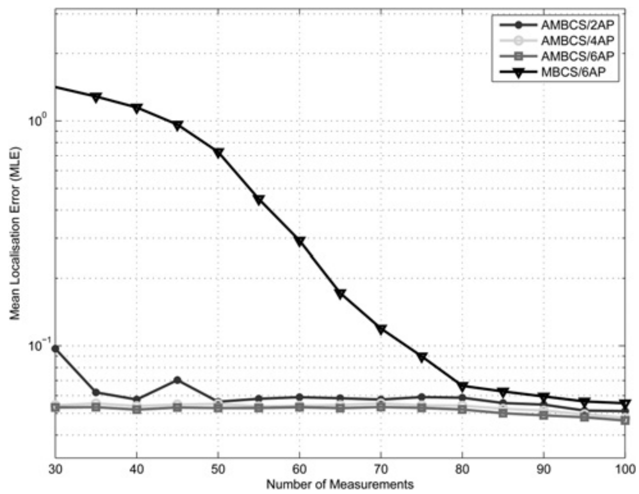

Fig. 7 *Localisation performance of the AMBCS algorithm against the number of working APs when there exist 10 MDs and 60 measurements are sampled for the CU*

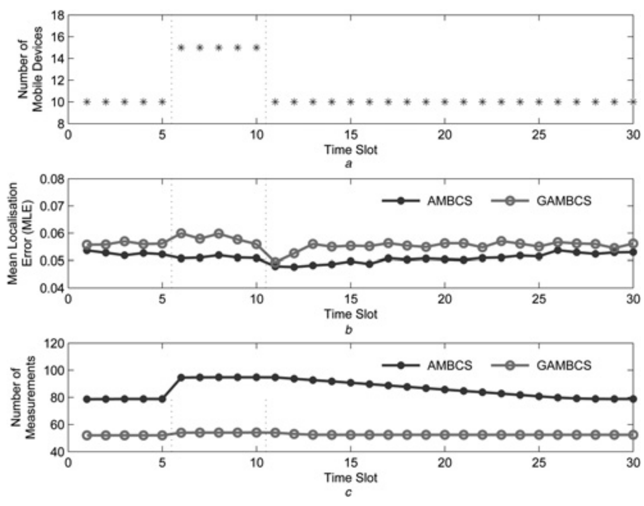

Fig. 6 *Localisation performance as a function of varying number of MDs for AMBCS and GAMBCS when initial number of measurements is set to be 50, NEB is set to be 70 and noise power $\sigma_0 = 0.005$*

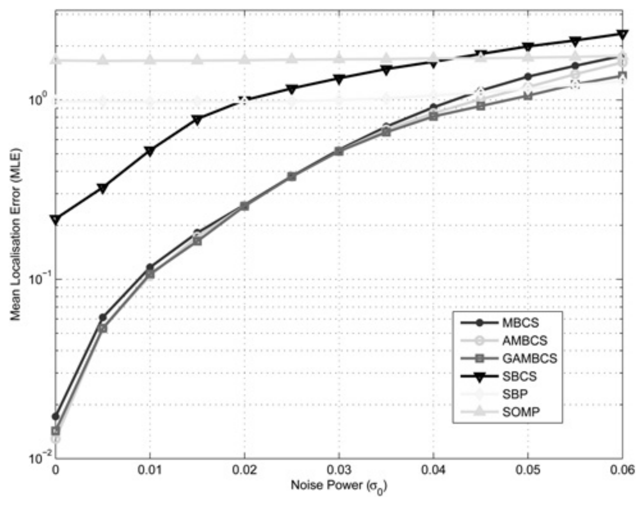

Fig. 8 *Localisation performance as a function of increasing noise power (σ_0) for MBCS, AMBCS, GAMBCS, SBCS, SBP and SOMP when there exist 10 MDs and 60 measurements are sampled at the CU side*

doi: 10.1049/iet-com.2013.1019

localisation error under different signal-to-noise ratio (SNR) levels. Each measurement is corrupted by Gaussian noise. The proposed MBCS algorithm outperforms much better than the single-task algorithms, especially in the low SNR environment. Moreover, both adaptive algorithms achieve even better results.

6 Conclusion and future works

In this paper, we designed an MBCS localisation scheme to estimate the locations of MDs more accurately. Specifically, we explored the intra- and inter-signal correlations of the RSS measurements and exploited the common structures between different tasks to achieve more precise localisation. Moreover, we proposed two adaptive algorithms to dynamically determine the necessary number of RSS measurements, by virtue of the theoretically validated criterion 'error bars'. The simulation results further verified that both the proposed algorithms could achieve superior localisation performance, compared with conventional MBCS algorithms.

Although our BCS-based scheme has showed the accuracy and effectiveness by simulations, it is still worthwhile to validate it in practical experimental networks. Therefore we are dedicated to handle the related meaningful works in the future.

7 Acknowledgment

This paper is supported by the National Basic Research Program of China (973Green, no. 2012CB316000), the Key (Key grant) Project of Chinese Ministry of Education (no. 313053), the Key Technologies R&D Program of China (no. 2012BAH75F01) and the grant of 'Investing for the Future' program of France ANR to the CominLabs excellence laboratory (ANR-10-LABX-07-01).

8 References

1 Schiller, J., Voisard, A.: 'Location-based services' (Morgan Kaufmann, 2004). Available at http://www.amazon.com/Location-Based-Services-Kaufmann-Management-Systems/dp/1558609296
2 Kpper, A.: 'Location-based services: fundamentals and operation' (John Wiley & Sons, Ltd., 2005). Available at http://www.amazon.com/Location-Based-Services-Fundamentals-Axel-pper/dp/0470092319
3 Scellato, S., Noulas, A., Lambiotte, R., Mascolo, C.: 'Socio-spatial properties of online location-based social networks'. Proc. ICWSM, Barcelona, Spain, July 2011
4 Boulos, M.N., Rocha, A., Martins, A., *et al.*: 'CAALYX: a new generation of location-based services in healthcare', *Int. J. Health Geogr.*, 2007, **6**, (1), p. 9
5 Tsalgatidou, A., Veijalainen, J., Markkula, J., Katasonov, A., Hadjiefthymiades, S.: 'Mobile e-commerce and location based services: technology and requirements'. Proc. ScanGIS, Espoo, Finland, April 2003
6 Harroud, H., Ahmed, M., Karmouch, A.: 'Policy-driven personalized multimedia services for mobile users', *IEEE Trans. Mob. Comput.*, 2003, **2**, (1), pp. 16–24
7 Gustafsson, F., Gunnarsson, F.: 'Mobile positioning using wireless networks: possibilities and fundamental limitations based on available wireless network measurements', *IEEE Signal Process. Mag.*, 2005, **22**, (4), pp. 41–53
8 Patwari, N., Ash, J.N., Kyperountas, S., Hero III, A.O., Moses, R.L., Correal, N.S.: 'Locating the nodes: cooperative localization in wireless sensor networks', *IEEE Signal Process. Mag.*, 2005, **22**, (4), pp. 54–69
9 Bahl, P., Padmanabhan, V.: 'RADAR: an in-building RF-based user location and tracking system'. Proc. IEEE INFOCOM, Tel Aviv, Israel, March 2000
10 Kushki, A., Plataniotis, K., Venetsanopoulos, A.: 'Kernel-based positioning in wireless local area networks', *IEEE Trans. Mob. Comput.*, 2007, **6**, (6), pp. 689–705
11 Patterson, C.A., Muntz, R.R., Pancake, C.M.: 'Challenges in location-aware computing', *IEEE Pervasive Comput.*, 2003, **2**, (2), pp. 80–89
12 Rizos, C., Dempster, A.G., Li, B., Salter, J.: 'Indoor positioning techniques based on wireless LAN'. Auswireless Conf., Sydney, Australia, March 2006
13 He, T., Huang, C., Blum, B.M., Stankovic, J.A., Abdelzaher, T.: 'Range-free localization schemes for large scale sensor networks'. Proc. ACM Mobicom, San Diego, CA, USA, September 2003
14 Zhang, B., Cheng, X., Zhang, N., Cui, Y., Li, Y., Liang, Q.: 'Sparse target counting and localization in sensor networks based on compressive sensing'. Proc. IEEE INFOCOM, Shanghai, China, April 2011
15 Feng, C., Au, W.S.A., Valaee, S., Tan, Z.: 'Compressive sensing based positioning using RSS of WLAN access points'. Proc. IEEE INFOCOM, San Diego, CA, USA, March 2010
16 Au, A., Feng, C., Valaee, S., *et al.*: 'Indoor tracking and navigation using received signal strength and compressive sensing on a mobile device', *IEEE Trans. Mob. Comput.*, 2013, **12**, (10), pp. 2050–2062
17 Roos, T., Myllymki, P., Tirri, H., Misikangas, P., Sievnen, J.: 'A probabilistic approach to WLAN user location estimation', *Int. J. Wirel. Inf. Netw.*, 2002, **9**, (3), pp. 155–164
18 Nikitaki, S., Tsakalides, P.: 'Localization in wireless networks via spatial sparsity'. Conf. Record of the Forty Fourth Asilomar Conf. Signals, Systems and Computers (ASILOMAR) 2010, Monterey, CA, USA, November 2010
19 Donoho, D.: 'Compressed sensing', *IEEE Trans. Inf. Theory*, 2006, **52**, (4), pp. 4036–4048
20 Candes, E.J., Tao, T.: 'Near-optimal signal recovery from random projections: universal encoding strategies?', *IEEE Trans. Inf. Theory*, 2006, **52**, (12), pp. 5406–5425
21 Cands, E.J., Romberg, J., Tao, T.: 'Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information', *IEEE Trans. Inf. Theory*, 2006, **52**, (2), pp. 489–509
22 Ji, S., Dunson, D., Carin, L.: 'Multitask compressive sensing', *IEEE Trans. Signal Process.*, 2009, **57**, (1), pp. 92–106
23 Ji, S., Xue, Y., Carin, L.: 'Bayesian compressive sensing', *IEEE Trans. Signal Process.*, 2008, **56**, (6), pp. 2346–2356
24 Li, X., Hong, S., Han, Z., Wu, Z.: 'Bayesian compressed sensing based dynamic joint spectrum sensing and primary user localization for dynamic spectrum access'. Proc. IEEE Globecom, Houston, TX, USA, December 2011
25 Zhang, Y., Zhao, Z., Zhang, H.: 'Adaptive bayesian compressed sensing based localization in wireless networks'. Proc. Chinacom, Kunming, China, August 2012
26 Romberg, J., Wakin, M.: 'Compressed sensing: a tutorial'. Proc. IEEE SSP Workshop, Madison, WI, USA, August 2007. Available at http://www.people.ee.duke.edu/~willett/SSP//Tutorials/ssp07-cs-tutorial.pdf
27 Lawson, C.L., Hanson, R.J.: 'Solving least squares problems' (SIAM, 1974), vol. 161. Available at http://www.epubs.siam.org/doi/pdf/10.1137/1.9781611971217.fm
28 Tropp, J.A., Gilbert, A.C.: 'Signal recovery from random measurements via orthogonal matching pursuit', *IEEE Trans. Inf. Theory*, 2007, **53**, (12), pp. 4655–4666
29 Chen, S.S., Donoho, D.L., Saunders, M.A.: 'Atomic decomposition by basis pursuit', *SIAM J. Sci. Comput.*, 1998, **20**, (1), pp. 33–61
30 Tsaig, Y., Donoho, D.L.: 'Extensions of compressed sensing', *Signal Process.*, 2006, **86**, (3), pp. 549–571
31 Cands, E.J.: 'The restricted isometry property and its implications for compressed sensing', *C. R. Math.*, 2008, **346**, (9), pp. 589–592
32 Berger, C., Wang, Z., Huang, J., Zhou, S.: 'Application of compressive sensing to sparse channel estimation', *IEEE Commun. Mag.*, 2010, **48**, (11), pp. 164–174
33 Wipf, D., Rao, B.: 'Comparing the effects of different weight distributions on finding sparse representations'. Proc. NIPS, Vancouver, Canada, December 2006
34 Wipf, D., Palmer, J., Rao, B.: 'Perspectives on sparse bayesian learning', *Adv. Neural Inf. Process. Syst.*, 2004, **16**, pp. 249–256
35 Tipping, M.E.: 'Sparse bayesian learning and the relevance vector machine', *J. Mach. Learn. Res.*, 2001, **1**, (1), pp. 211–244
36 Seidel, S., Rappaport, T.: '914 MHz path loss prediction models for indoor wireless communications in multifloored buildings', *IEEE Trans. Antennas Propag.*, 1992, **40**, (2), pp. 207–217
37 Cover, T.M., Thomas, J.A.: 'Differential entropy', in: 'Elements of information theory' (Wiley-interscience, 2012)
38 Horn, R.A., Johnson, C.R.: 'Matrix analysis' (Cambridge University Press, 1990). Available at http://www.amazon.com/Matrix-Analysis-Roger-A-Horn/dp/0521386322

9 Appendix

9.1 Proof of Theorem 1 and Corollary 1

Firstly, we give a lemma in the field of information theory, which bridges the link between covariance and differential entropy [37] and implies why the 'error bars' could offer the informational uncertainty of our localisation problem.

Lemma 1: For a multivariate Gaussian distribution $\mathcal{N}(\boldsymbol{f})$ with mean value $\boldsymbol{\mu}_f$ and covariance Σ_f, the differential entropy for a continuous distribution $p(\boldsymbol{f})$, which is defined as $H(\boldsymbol{f}) = -\int p(\boldsymbol{f}) \log p(\boldsymbol{f}) d\boldsymbol{f}$, would equal $H(\boldsymbol{f}) = (1/2)\log|\Sigma_f| + C$. Here, C denotes a constant independent with the distribution $\mathcal{N}(\boldsymbol{f})$.

Moreover, as detailed in the following lemma, $\{\boldsymbol{\Sigma}^i\}_{i=1}^{P}$ in (10) could turn into a positive definitive matrix $\{\boldsymbol{\Sigma}_v^i\}_{i=1}^{P}$ after limited number of elementary algebraic transforms [38].

Lemma 2: After limited number of elementary algebraic transforms, for any $i \in \{1, \ldots, P\}$, Σ^i in (10) could be reformulated as

$$\Sigma^i \to \begin{bmatrix} \boldsymbol{\Sigma}_v^i & 0 \\ 0 & 0 \end{bmatrix} \tag{17}$$

Proof: Without loss of generality, we omit the superindex i in $\boldsymbol{\Sigma}^i$ during the proof. As discussed in the end of Section 2, certain elements in λ tend to be infinity [34]. Consequently, after limited elementary transforms $\mathcal{T}$, Λ in (10) could be reformulated as

$$\Lambda \to \begin{bmatrix} \Lambda_v & 0 \\ 0 & \Lambda_{\text{inf}} \end{bmatrix} \tag{18}$$

Meanwhile, there only exists a set $\boldsymbol{s}_v$ of non-zero recovered entries in $\boldsymbol{s}$. Thus, we could only consider the $|\boldsymbol{s}_v|$ column vectors of transform basis $\boldsymbol{\Psi}$, one of which exactly corresponds to the RSS measurements from one estimated MD location to the AP. By the same elementary transforms $\mathcal{T}$, $\boldsymbol{\Psi}$ can be represented as

$$\boldsymbol{\Psi} \to \begin{bmatrix} \boldsymbol{\Psi}_v & \boldsymbol{\Psi}_{\text{inf}} \end{bmatrix} \tag{19}$$

Therefore, after the same transformation $\mathcal{T}$, Σ can turn into

$$\begin{aligned} \Sigma &\to \left(\lambda_0 \begin{bmatrix} \boldsymbol{\Psi}_v^{\text{T}} \\ \boldsymbol{\Psi}_{\text{inf}}^{\text{T}} \end{bmatrix} \boldsymbol{\Phi}^{\text{T}}\boldsymbol{\Phi}\begin{bmatrix} \boldsymbol{\Psi}_v & \boldsymbol{\Psi}_{\text{inf}} \end{bmatrix} + \begin{bmatrix} \Lambda_v & 0 \\ 0 & \Lambda_{\text{inf}} \end{bmatrix}\right)^{-1} \\ &= \left(\begin{bmatrix} A \triangleq \lambda_0 \boldsymbol{\Psi}_v^{\text{T}}\boldsymbol{\Phi}^{\text{T}}\boldsymbol{\Phi}\boldsymbol{\Psi}_v + \Lambda_v & B \triangleq \lambda_0 \boldsymbol{\Psi}_v^{\text{T}}\boldsymbol{\Phi}^{\text{T}}\boldsymbol{\Phi}\boldsymbol{\Psi}_{\text{inf}} \\ C \triangleq \lambda_0 \boldsymbol{\Psi}_{\text{inf}}^{\text{T}}\boldsymbol{\Phi}^{\text{T}}\boldsymbol{\Phi}\boldsymbol{\Psi}_v & \Lambda_{\text{inf}} \end{bmatrix}\right)^{-1} \\ &\overset{\text{(a)}}{=} \begin{bmatrix} (A - B\Lambda_{\text{inf}}^{-1}C)^{-1} & A^{-1}B(CA^{-1}B - \Lambda_{\text{inf}})^{-1} \\ (CA^{-1}B - \Lambda_{\text{inf}})^{-1}CA^{-1} & (\Lambda_{\text{inf}} - CA^{-1}B)^{-1} \end{bmatrix} \\ &= \begin{bmatrix} A^{-1} & 0 \\ 0 & 0 \end{bmatrix} \end{aligned} \tag{20}$$

where the equality (a) comes from formula 0.7.3 in [38]. Moreover, $\boldsymbol{A}$ is a positive definitive matrix, since $\boldsymbol{A}$ is a combination of a Hermitian matrix and a positive diagonal matrix. Hence, $\boldsymbol{\Sigma}_v = \boldsymbol{A}^{-1}$ is a positive definitive matrix. Now, the lemma comes. □

Then, we can give the proof of Theorem 1.

Proof: Without loss of generality, we omit the superindex i in $\boldsymbol{x}^i$, $\boldsymbol{y}^i$, $\boldsymbol{s}^i$, $\boldsymbol{\Phi}^i$ and $\boldsymbol{\Psi}^i$ during the proof. Lemma 2 implies that only the non-zero recovered entries $\boldsymbol{s}_v$ and corresponding transform basis $\boldsymbol{\Psi}_v$ take effect. By Lemma 1, the differential entropy for $\boldsymbol{s}$ can be directly derived based on the corresponding posterior probability in (9), namely

$$\begin{aligned} &H(p(s_v|\boldsymbol{y}, \boldsymbol{\lambda}, \lambda_0)) \\ &= -\int p(s_v|\boldsymbol{y}, \boldsymbol{\lambda}, \lambda_0) \log p(s_v|\boldsymbol{y}, \boldsymbol{\lambda}, \lambda_0)\, \mathrm{d}p \\ &= \frac{1}{2}\log|\boldsymbol{\Sigma}_v| + C = -\frac{1}{2}\log|\lambda_0 \boldsymbol{\Theta}_v^{\text{T}}\boldsymbol{\Theta}_v + \Lambda_v| + C \end{aligned} \tag{21}$$

where $\boldsymbol{\Theta}_v = \boldsymbol{\Phi}\boldsymbol{\Psi}_v$ and Λ_v is defined in Lemma 2. Moreover, the dependence of differential entropy on the measurement $\boldsymbol{y}$ is reflected by the related covariance and hyperparameters. If we add another measurement according to the new measurement vector $\boldsymbol{r}^{\text{T}}$, we could obtain a new measurement matrix $\boldsymbol{\Phi}_{\text{new}}^i = \begin{bmatrix} \boldsymbol{\Phi}^{\text{T}} & \boldsymbol{r} \end{bmatrix}^{\text{T}}$ and a new $\boldsymbol{\Theta}_{\text{new}} = \boldsymbol{\Phi}_{\text{new}}\boldsymbol{\Psi}$. Meanwhile, (21) would be updated as

$$\begin{aligned} &H(p(\boldsymbol{s}_v|\boldsymbol{y}_{\text{new}}, \boldsymbol{\lambda}, \lambda_0)) \\ &= -\frac{1}{2}\log|\lambda_0(\boldsymbol{\Theta}_{\text{new}})^{\text{T}}(\boldsymbol{\Theta}_{\text{new}}) + \boldsymbol{\Lambda}| + C \\ &= -\frac{1}{2}\log\left|\lambda_0\left(\boldsymbol{\Psi}_v^{\text{T}}\boldsymbol{\Phi}^{\text{T}}\boldsymbol{\Phi}\boldsymbol{\Psi}_v + \boldsymbol{\Psi}_v^{\text{T}}\boldsymbol{r}\boldsymbol{r}^{\text{T}}\boldsymbol{\Psi}_v\right) + \boldsymbol{\Lambda}_v\right| + C \\ &= -\frac{1}{2}\log\left|\boldsymbol{\Sigma}_v^{-1} + \lambda_0\boldsymbol{\Psi}_v^{\text{T}}\boldsymbol{r}\boldsymbol{r}^{\text{T}}\boldsymbol{\Psi}_v\right| + C \end{aligned}$$

$$\begin{aligned} &\overset{\text{(a)}}{=} -\frac{1}{2}\log\left(|\boldsymbol{\Sigma}_v^{-1}||1 + \lambda_0\boldsymbol{r}^{\text{T}}\boldsymbol{\Psi}_v\boldsymbol{\Sigma}_v\boldsymbol{\Psi}_v^{\text{T}}\boldsymbol{r}|\right) + C \\ &= H(p(\boldsymbol{s}_v|\boldsymbol{y}, \boldsymbol{\lambda}, \lambda_0)) - \frac{1}{2}\log\left|1 + \lambda_0\boldsymbol{r}^{\text{T}}\boldsymbol{\Psi}_v\boldsymbol{\Sigma}_v\boldsymbol{\Psi}_v^{\text{T}}\boldsymbol{r}\right| \end{aligned} \tag{22}$$

Here, the equality (a) follows from the Schur complements and determinantal formula [38]. By Lemma 2, the last term $\frac{1}{2}\log\left|1 + \lambda_0\boldsymbol{r}^{\text{T}}\boldsymbol{\Psi}_v\boldsymbol{\Sigma}_v\Psi_v^{\text{T}}\boldsymbol{r}\right|$ of (22) is positive. Consequently, a newly added measurement could contribute to decrease the 'error bars', since the measurement vector could be selected intentionally. The claim follows. □

Next, Corollary 1 can be easily proved, after taking into account that the equation $\boldsymbol{\Psi}^i(\boldsymbol{\Sigma}^i)(\boldsymbol{\Psi}^i)^{\text{T}} = \boldsymbol{\Psi}_v^i(\boldsymbol{\Sigma}_v^i)(\boldsymbol{\Psi}_v^i)^{\text{T}}$ holds for any $i \in \{1, \ldots, P\}$.

IET Commun., 2014, Vol. 8, Iss. 10, pp. 1736–1744
doi: 10.1049/iet-com.2013.1019