An Algorithm for Fast REM Construction

Sebastien Grimoud*, Berna Sayrac*, Sana Ben Jemaa* and Eric Moulines†
*Orange Labs, Issy-Les-Moulineaux, France
Email: {sebastien.grimoud, berna.sayrac, sana.benjemaa}@orange-ftgroup.com
†Telecom ParisTech, Paris, France
Email: eric.moulines@telecom-paristech.fr

***Abstract*—The Radio Environment Map (REM) stores radio environmental information that can be used to enhance cognitive radio resource management in wireless networks. In this paper, we propose an iterative REM building process based on Kriging interpolation technique that builds the REM using geolocated measurements performed by mobile terminals. As terminal measurements are costly in terms of signalling and battery consumption, we propose an algorithm that chooses the most appropriate measurements to be requested to the mobiles. We compare the performance of this algorithm with a random choice of measurements and show that our algorithm reduces the squared error of the power map by 16%. The proposed algorithm has also the merit of being fast enough to be implemented in an online fashion.**

I. Introduction

The Radio Environment Map [1], [2] is a promising concept for storing radio environmental information that can be used to enhance radio resource management in wireless networks. In cellular networks, the REM can be used to improve the network performances [3], or to minimize the operational costs by replacing or at least minimizing drive tests (MDT) [4], for troubleshooting for instance.

The REM information is built based on the terminals' measurement data, combined with location information and reported to a functional entity called REM manager. This entity exploits this information to build a complete map by interpolating the geolocalized measurements. Because measurement reporting is costly in terms of signalling overhead and battery consumption, the main challenge while building a REM is to find the optimal trade-off between the REM quality, i.e. the REM information accuracy and the measurements requested from the terminals.

In this paper[1], we propose an iterative REM building process based on Kriging interpolation technique. At each time a REM update is needed, this algorithm chooses the most appropriate candidate terminals for performing additional measurements in order to reach the target REM quality, with respect to the constraint of minimizing the number of measurements. The reminder of the paper is organised as follows. Section II gives an overview on the state of the art works related to REM building and sensor selection. Section III describes the studied scenario, section IV details the modelling assumptions and section V presents the mathematical derivation for sorting the candidate measurements. Simulation results are presented in section VI while section VII draws conclusions and gives some insight about further works.

II. Related Works

The concept of REMs has been first proposed by the Virginia Tech team [1]. They define REM as a database that contains information on the radio environment, including geographical features, available services, spectrum policies and regulations, location and activities of radio devices, past experiences etc. This database can be located anywhere in the network with different possible architectures: centralized, distributed or hybrid. Related with the architectural aspects, the amount of signalling overhead needed to disseminate the REM is of concern and treated in [5]. The REM proposed as such, has been mainly considered for IEEE 802.22 WRAN scenarios and applications [5] [6] [7] where the focus is on opportunistic spectrum access on TV whitespaces. We would like to underline that our focus is on a REM which stores incoming environmental data but also interpolates them to benefit from the spatial correlation that exists in the data. The concept of collecting geolocated information on the radio environment and constructing a map using this information has also been investigated and developped further by other research groups [2], [8], [9]. In these works, REMs have been handled in a more general Cognitive Radio (CR) context than TV whitespaces and it is considered as a mean to represent spatio-temporal characteristics of the radio environment by using concepts and tools from spatial statistics, like point processes, spatial random fields, pair correlation functions, point interaction models, spatial interpolation techniques, etc. Particularly, different from previsous work on REMs, [2] and [8] deal with construction aspects of REMs through spatial interpolation techniques like Kriging [10]. In this paper, we address an important practical issue associated with REM construction: the selection of the best additional measurement points needed to update the REM in order to achieve its target quality. Sensor selection has also been investigated in the context of cooperative spectrum sensing where the focus is on detecting the presence of the primary user. For example, in [11], the authors propose three sensor selection methods where sensors having the best detection performance with only hard (binary) local decisions are selected. On the other hand, [12] considers a cellular system peforming secondary spectrum access, and proposes sensor selection methods whose aim is to find the subset of sensors which experience uncorrelated shadow fading. For this purpose, different metrics are defined, which depend on shadowing correlation, on the positions of the sensors and on the distances between the sensors and the base stations. While there is this significant amount of literature on sensor selection in sensor networks and in cooperative spectrum sensing, there is no work that considers sensor selection in the context of REMs. Therefore, to the best of our knowledge, our work is the first to fill an important gap by providing a sensor selection mechanism for REM-update which can be implemented in an online fashion.

[1]This work has been carried-out in the EU FP7 FARAMIR project framework

978-1-936968-19-0 © 2011 ICST
DOI 10.4108/icst.crowncom.2011.245788

III. Scenario under study

For the scenarios mentioned in section I, the REM must incorporate a power map taking into account each base station, and the measurement of interest is the signal power received on the pilot channel for one BS.

We consider that mobiles actively participate in building the REM as they are used as sensors, and act under the request of the REM manager. The algorithm we propose finds the locations where measurements would improve the REM. The request for measurements can be either a broadcast or a unicast message.Once a measurement is completed, the algorithm updates the measurements list or polls another mobile when needed. Deciding whether it is suitable to update the REM is based on two conditions:

- The current load of the network is not too high to handle extra signalling without disturbing the current calls
- A quality indicator of the REM is below a minimum threshold. As quality indicators we can think of metrics based on timestamps attached to measurements as “up-to-date” or “deprecated” indicators for the REM, as well as theoretical variance or hypothesis testing. However, choosing a quality metric for the REM is out of the scope of this paper.

Sorting potentially interesting locations is challenging as :

- This must be done with few prior information on what the radio environment is like,
- The decision must be taken fast enough so that the algorithm can be implemented online.

A simple alternative to this preferential sampling that fulfills those two conditions is to choose the location list randomly from the set of measurement free locations. However we will show that this approach is less efficient in terms of mean squared error of the predicted map, and this can even become critical when the traffic is not uniform. Indeed, the random strategy concentrates the measurements in the traffic hotspots and the accuracy of the interpolation becomes really poor outside those zones.

IV. Modelling assumptions

A. The wireless channel

There is an abundant amount of litterature on modelling the wireless channel. There are two main strategies, namely ray tracing and analytical models. Ray tracing requires a precise description of the environment and great computation capabilities. Analytical models vary greatly in complexity, capturing either the coarse effects of the environment, *e.g.* urban versus rural ones, or dealing with real environment data, *e.g.* accounting for walls. Analytical models are less demanding in terms of computing power but for accurate results, modelisation becomes a burden.

Another approach is followed here. It takes the pragmatic approach of using a simple analytical model in combination with a statistical description of what we are not able to model analytically. We decompose the wireless channel as a sum of a linear decay function of the logarithm of the distance and a shadowing term. When considering a homogeneous topology of the terrain, the factor of the linear decay can be considered as a constant. It is to be pointed out that this model does not prevent large deviations from this linear decay. The degree of these deviations is caracterised by the variance of the shadowing. Hence we make no assumption on the dominance of the linear decay or the shadowing over each other.

Because what causes a deviation from the analytical loss typically has significative dimensions compared to the area of interest, *e.g.* buildings and trees, the shadowing must be modelled with a spatially correlated grid. Obviously, the degree of correlation between two points is greater as the distance between them gets smaller. In the rest of this paper, the received power at location i is:

$$p_i = p_0 - 10\alpha \log_{10}(d_i) + s_i \tag{1}$$

where p_0 is the emitted power, α is the path loss coefficient, d_i and s_i are respectively the distance in meters and the shadowing in dB scale between the location of i and the BS. We do not include any temporal dependency as we consider that measurement duration is long enough to average out the fast fading effects. Discretizing the 2D space on a rectangular grid, (1) gives us the received power at any point. Once again, we stress out that fluctuations around the linear decay are within s_i while the linear decay reflects a trend. Not including the trend in the shadowing is of prime importance as it allows for specifying a stationary process for s_i. The shadowing is assumed to be a correlated Random Vector (RV) with log-normal distribution and a correlation distance exponentially decaying with the distance, as suggested in [13]:

$$\mathbb{E}(s_i s_j) = \sigma^2 \exp\left(-\frac{d_{ij}}{\phi}\right) \tag{2}$$

where d_{ij} represents the Euclidean distance between two locations indexed by i and j, σ^2 is the variance and ϕ the correlation distance of the shadowing.

In this model, the parameters p_0, α, σ^2 and ϕ are unknown. However, we do have prior information on what their values might be. For instance, we know that the radiated power must be something around the power at the feeder, α is about 3.5 in urban areas [14] and σ typically ranges between 8 and 11 dB for typical outdoor Above RoofTop to Below RoofTop scenarios [13].

B. Bayesian Estimator

In this paper, we use Kriging for interpolating the power map. A good introduction to spatial statistics is given in [15] where Kriging is explained in details. Kriging applies when the underlying stochastic process can be modelled as having a normal probability density function (pdf). In our case, the received power in dBm at each point in space is normally distributed. Since those values are correlated, it must be described with the RV $\mathbf{p}$ which has the following distribution:

$$\mathbf{p} \equiv \mathcal{N}\left(\mathbf{D}\begin{pmatrix} p_0 \\ \alpha \end{pmatrix}, \sigma^2 \mathbf{R}(\phi)\right) \tag{3}$$

where D is a matrix of regressors:

$$\mathbf{D} = \begin{pmatrix} 1 & -\log_{10}(d_1) \\ \vdots & \vdots \\ 1 & -\log_{10}(d_N) \end{pmatrix} \tag{4}$$

and $\sigma^2 \mathbf{R}(\phi)$ is the correlation matrix whose $(i,j)^{\text{th}}$ term is given by (2). In the following, we use subscripts 1 and 2 to denote quantities related to the points for which we do not and do have measurement values respectively. For correlation matrices, R_{12} denotes the correlation of $\mathbf{p_1}$ with $\mathbf{p_2}$.

The parameters p_0,α,σ^2 and ϕ are assumed to be random values for which we are able to specify prior pdfs. Hence, we are interested in the following pdf, expected value of which gives the best map estimator with respect to the mean squared error criteria:

$$p(\mathbf{p_1}|\mathbf{p_2}) = \int p(\mathbf{p_1}|p_0,\alpha,\sigma^2,\phi,\mathbf{p_2})p(p_0,\alpha,\sigma^2,\phi|\mathbf{p_2})\mathrm{d}p_0\mathrm{d}\alpha\mathrm{d}\sigma^2\mathrm{d}\phi \quad (5)$$

The first term in the integral is the gaussian pdf, for which we are able to find conjugate priors for the unknown parameters $\boldsymbol{\beta} = \begin{pmatrix} p_0 \\ \alpha \end{pmatrix}$ and σ^2:

$$\boldsymbol{\beta} \equiv \mathcal{N}(\mathbf{m_b}, \sigma^2\mathbf{V_b}) \quad (6)$$

$$p(\sigma^2) = (\sigma^2)^{-\frac{n_\sigma}{2}+1}\exp(-\frac{n_\sigma S_\sigma^2}{2\sigma^2}) \quad (7)$$

where $\mathbf{m_b}$, S_σ^2 and $\mathbf{V_b}$, n_σ are respectively the mean and variance parameters of the priors. The prior for ϕ is chosen to be a discrete distribution for convenience purposes. The resulting pdf is [15]:

$$p(\mathbf{p_1}|\mathbf{p_2}) \propto \sum_\phi p(\phi|\mathbf{p_2})t_{n_\sigma+n}(\boldsymbol{\mu}^*(\phi), S^2(\phi)\mathbf{R}^*(\phi)) \quad (8)$$

where $t_{n_\sigma+N_2}$ is the multivariate Student pdf with $n_\sigma+N_2$ degrees of freedom, having the following parameters:

$$\begin{aligned}\boldsymbol{\mu}^*(\phi) &= (\mathbf{D_1}-\mathbf{R_{12}}\mathbf{R_{22}}^{-1}\mathbf{D_2})\mathbf{V}_{\beta^*}\mathbf{V_b}^{-1}\mathbf{m_b} \\ &+ [\mathbf{R_{12}} + (\mathbf{D_1}-\mathbf{R_{12}}\mathbf{R_{22}}^{-1}\mathbf{D_2})\mathbf{V}_{\beta^*}\mathbf{D_2}^T)]\mathbf{R_{22}}^{-1}\mathbf{p_2}\end{aligned} \quad (9)$$

$$\begin{aligned}\mathbf{R}^*(\phi) &= \mathbf{R_{11}} - \mathbf{R_{12}}\mathbf{R_{22}}^{-1}\mathbf{R_{12}}^T + (\mathbf{D_1}-\mathbf{R_{12}}\mathbf{R_{22}}^{-1}\mathbf{D_2}) \\ &\times (\mathbf{V_b}^{-1}+\mathbf{V}_{\beta^*}^{-1})(\mathbf{D_1}-\mathbf{R_{12}}\mathbf{R_{22}}^{-1}\mathbf{D_2})^T\end{aligned} \quad (10)$$

$$\begin{aligned}S^2(\phi) &= \frac{1}{n_\sigma+N_2}(n_\sigma S_\sigma^2 \\ &+ \mathbf{m_b}^T\mathbf{V_b}^{-1}\mathbf{m_b} + \mathbf{p_2}^T\mathbf{R_{22}}^{-1}\mathbf{p_2} - \boldsymbol{\beta^*}^T\mathbf{V}_{\beta^*}^{-1}\boldsymbol{\beta^*})\end{aligned} \quad (11)$$

where N_2 is the size of $\mathbf{p_2}$ and:

$$\mathbf{V}_{\beta^*} = (\mathbf{V_b}^{-1} + \mathbf{D_2}^T\mathbf{R_{22}}^{-1}\mathbf{D_2})^{-1} \quad (12)$$

$$(13)$$

$$\boldsymbol{\beta^*} = \mathbf{V}_{\beta^*}(\mathbf{V_b}^{-1}\mathbf{m_b} + \mathbf{D_2}^T\mathbf{R_{22}}^{-1}\mathbf{p_2}) \quad (14)$$

$$p(\phi|\mathbf{p_2}) = p(\phi)\sqrt{|\mathbf{V}_{\beta^*}||\mathbf{R_{22}}^{-1}|}(S^2)^{-\frac{N_2+n_\sigma}{2}} \quad (15)$$

For notational convenience we omitted the ϕ dependency in the correlation matrices.

V. Measurement Requesting

We choose to collect measurements in a way that minimizes the mean squared error of the predicted map:

$$MSE = (\hat{\mathbf{p}}-\mathbf{p})^T(\hat{\mathbf{p}}-\mathbf{p}) \quad (16)$$

where $\hat{\mathbf{p}}$ and $\mathbf{p}$ denotes respectively the predicted and the real world power maps. Obviously, this quantity can not be derived as $\mathbf{p}$ is unknown. However we can estimate this value by taking the expected value, conditioning on the measurements:

$$\begin{aligned}\mathbb{E}(MSE|\mathbf{p_2}) &= \mathbb{E}((\hat{\mathbf{p}}-\mathbf{p})^T(\hat{\mathbf{p}}-\mathbf{p})|\mathbf{p_2}) \\ &= \mathbb{E}\left((\mathbb{E}(\mathbf{p_1}|\mathbf{p_2})-\mathbf{p_1})^T(\mathbb{E}(\mathbf{p_1}|\mathbf{p_2})-\mathbf{p_1})|\mathbf{p_2}\right) \\ &= \mathrm{trace}\left(\mathrm{cov}\left(\mathbf{p_1}|\mathbf{p_2}\right)\right)\end{aligned} \quad (17)$$

For each candidate measurement, we assign a score which is closely related to (17), where the candidate measurement is added to the second measurement set. This score is given by (8). However, since we do not know the value of the candidate measurement yet, its value in $\mathbf{p_2}$ is replaced by its predicted value taken from the most recent prediction, $\mathbb{E}(\mathbf{p_1}|\mathbf{p_2})$.

The challenge when deriving (17) is that it has to be done fast enough for every candidate so that the REM can be updated within a reasonable amount of time. This actually prevents us from deriving (17) as it is, and we use the following equivalent expression instead:

$$\begin{aligned}\mathrm{trace}\left(\mathrm{cov}\left(\mathbf{p_1}|\mathbf{p_2}\right)\right) &= \sum_\phi p(\phi|\mathbf{p_2})S^2(\phi)\Big[N_1 - \mathbf{1}^T\mathbf{R_{12}}^2\circ\mathbf{R_{22}}^{-1}\mathbf{1} \\ &+ \mathbf{1}^T(\mathbf{D_1}-\mathbf{R_{12}}\mathbf{R_{22}}^{-1}\mathbf{D_2})^2\circ(\mathbf{V_b}^{-1}+\mathbf{V}_{\beta^*}^{-1})^{-1}\mathbf{1}\Big] \\ &+ \sum_j\Big[\sum_\phi p(\phi|\mathbf{p_2})\mu_j^*(\phi) - (\sum_\phi p(\phi|\mathbf{p_2})\mu_j^*(\phi))^2\Big]\end{aligned} \quad (18)$$

where $\mathbf{A}^2 = \mathbf{A}^T\mathbf{A}$, $\circ$ is the Hadamard product, $\mathbf{1}$ a vector of ones and N_1 the size of $\mathbf{p_1}$. Because we add a candidate measurement to the former measurement set, $\mathbf{R_{22}}^{-1}$, $|\mathbf{R_{22}}^{-1}|$ and $\mathbf{R_{12}}^T\mathbf{R_{12}}$ can be iteratively calculated from the result of the most recent interpolation. In the following equations, we index with "cdt" the new quantities to calculate, and we split the covariance matrices in blocks wrt the row and column related to the candidate measurement.

$$\begin{aligned}\mathbf{R}_{\mathbf{22},\mathrm{cdt}}^{-1} &= \begin{pmatrix}\mathbf{R_{22}^{11}} & \mathbf{c_{22}^1} & \mathbf{R_{22}^{12}} \\ \mathbf{c_{22}^1}^T & 1 & \mathbf{c_{22}^2}^T \\ \mathbf{R_{22}^{21}} & \mathbf{c_{22}^2} & \mathbf{R_{22}^{22}}\end{pmatrix}^{-1} \\ &= \mathbf{P}\begin{pmatrix}\mathbf{R_{22}^{-1}} + \frac{1}{k}\mathbf{R_{22}^{-1}}\mathbf{c_{22}}\mathbf{c_{22}}^T\mathbf{R_{22}^{-1}} & -\frac{1}{k}\mathbf{R_{22}^{-1}}\mathbf{c_{22}} \\ -\frac{1}{k}\mathbf{c_{22}}^T\mathbf{R_{22}^{-1}} & \frac{1}{k}\end{pmatrix}\mathbf{P}^T\end{aligned} \quad (19)$$

$$k = 1 - \mathbf{c_{22}}^T\mathbf{R_{22}^{-1}}\mathbf{c_{22}}$$

where $\mathbf{P}$ is a transposition matrix that places the row corresponding to the candidate position in R_{22} to the last position.

$$|\mathbf{R_{22}^{-1}}|_{\mathrm{cdt}} = \frac{1}{k}|\mathbf{R_{22}^{-1}}| \quad (20)$$

Defining:

$$\mathbf{R_{12}} = \begin{pmatrix}\mathbf{R_{12}^{11}} & \mathbf{R_{12}^{12}} \\ \mathbf{r_{12}^1} & \mathbf{r_{12}^2} \\ \mathbf{R_{12}^{21}} & \mathbf{R_{12}^{22}}\end{pmatrix}$$

$$\mathbf{R_{12}}_{\mathrm{cdt}} = \begin{pmatrix}\mathbf{R_{12}^{11}} & \mathbf{c_{12}^1} & \mathbf{R_{12}^{12}} \\ \mathbf{R_{12}^{21}} & \mathbf{c_{12}^2} & \mathbf{R_{12}^{22}}\end{pmatrix}$$

$$\mathbf{R_{12}}^T\mathbf{R_{12}} = \mathbf{R} + \mathbf{r}^T\mathbf{r}$$

we obtain:

$$\mathbf{R_{12}}^T\mathbf{R_{12}}_{\mathrm{cdt}} = \begin{pmatrix}\mathbf{R^{11}} & \mathbf{u^1} & \mathbf{R^{12}} \\ \mathbf{u^1}^T & \mathbf{c}^T\mathbf{c} & \mathbf{u^2}^T \\ \mathbf{R^{21}} & \mathbf{u^2} & \mathbf{R^{22}}\end{pmatrix} \quad (21)$$

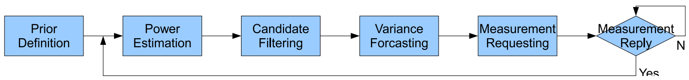


Fig. 1. Algorithm block diagram

where

$$\begin{pmatrix} \mathbf{u^1} \\ \mathbf{u^2} \end{pmatrix} = \begin{pmatrix} \mathbf{R_{12}^{11}}^T \mathbf{c_{12}^1} + \mathbf{R_{12}^{21}}^T \mathbf{c_{12}^2} \\ \mathbf{R_{12}^{12}}^T \mathbf{c_{12}^1} + \mathbf{R_{12}^{22}}^T \mathbf{c_{12}^2} \end{pmatrix} \tag{22}$$

Thanks to equations (19), (20) and (21) it is possible to estimate the variance of the future map for any candidate. It is to be noted that the above equations rely on quantities derived upon reception of the last measurements. Hence, there is no issue related to propagation of quantization noise as we iterate only once.

VI. Simulation results

A. Actors and roles

For simulation purposes we consider that the REM manager creates a list of measurements of interest, refered to as the measurement request. The measurement request is then broadcated or sent via unicast messages to the mobile terminals.The mobile terminals reply back, provided they are located in a location of interest. Upon reception of a measurement, the REM manager updates the power map prediction and subsequently derives the next measurement request.

B. Implemented algorithm

The algorithm implemented by the REM manager is summarized in Figure 1.

- The Prior Definition block is the definition of the parameters of the prior pdfs. The values are provided by the network engineers.
- The Power Estimation block derives $\mathbb{E}(\mathbf{p_1}|\mathbf{p_2})$, as well as $\mathbf{R_{22}}^{-1}$, $|\mathbf{R_{22}}^{-1}|$ and $\mathbf{R_{12}}^T\mathbf{R_{12}}$.
- Candidate Filtering embeds heuristics for reducing the number of measurement candidates. This is for further decreasing the computation time of “Variance Forcasting”.
- The Variance Forcasting block is the key contribution of this paper. It is in charge of estimating how a candidate measurement would reduce the variance of the power map estimator. The output is a score assigned to each candidate. The derivation is done for any candidate as specified by the Candidate Filtering block.
- The Measurement Requesting block takes decisions on how the measurement request is built, based on the score of the candidates. It can typically be the K top candidates.

C. Main result

Table I summarizes the environment parameters that have been used throughout the simulations. The maps are squares of 41×41 grid points and gridwidth is $25\,\mathrm{m}$, hence covering $1\,\mathrm{km}^2$ which is typically greater than the coverage of a cell in urban area. The traffic distribution is modeled as a normally distributed random vector, normalized so as to have a probablity one over the zone of interest.

Parameter	Value
Radio Environment	
Radiated Power p_0	$10\log_{10} 60\,\mathrm{dBW}$
Path loss Exponent α	3.5
Shadowing Std σ	$8\,\mathrm{dB}$
Shadowing correlation distance ϕ	$150\,\mathrm{m}$
Traffic distribution	
Mean	$6\,10^{-4}$
Variance	$2.6\,10^{-8}$
Correlation Distance	$500\,\mathrm{m}$

TABLE I
ENVIRONMENT PARAMETERS

	β		σ^2		ϕ
Simu	mean	variance	S_σ^2	n_σ	
1	$\begin{pmatrix}10\log_{10}60\\3.5\end{pmatrix}$	$\begin{pmatrix}3&0\\0&3\end{pmatrix}$	64	20	$\frac{1}{6}[75:25:200]$
2	$\begin{pmatrix}10\log_{10}63\\3.4\end{pmatrix}$	$\begin{pmatrix}12&0\\0&12\end{pmatrix}$	60	5	$\frac{1}{6}[75:25:200]$
3	$\begin{pmatrix}10\log_{10}60\\3.5\end{pmatrix}$	$\begin{pmatrix}3&0\\0&3\end{pmatrix}$	64	20	$\frac{1}{6}[75:25:200]$

TABLE II
PRIOR PARAMETERS

For the first simulation, the priors correspond to a well informed case, where pdf modes match the real parameter values, as shown in table II. The Candidates Filtering block selects candidates on a sub-sampled grid with gridwidth $50\,\mathrm{m}$ and the measurement request is only the best candidate ($K = 1$).

Figure 2 shows the mean squared error of the power map as measurements are added, while the blue curve of Figure 3 compares this strategy against a random choice of the measurements. It can be seen that the proposed algorithm gives an error reduction of $16\,\%$ on the average. The simulated power map, interpolated power map and mean square error map are shown in figure 4 for 200 measurements points.

D. Influence of the prior choice

In this section we show that the method is robust even when we have little information on the parameters. This is reflected by inaccurate and diffuse priors, as those chosen in the second line of table I. The black curve of Figure 3 shows the gain over the random selection strategy. When compared with the well informed case, we can see that the gain tends to be higher. This is because the proposed algorithm takes full advantage of the Bayesian approach which embeds the variance of the parameters for building the measurement request.

E. Influence of the request size

Here we focus on the measurement requesting block. The last simulation forms the request list as the top 25 candidate measurements ($K = 25$). Allowing for more measurements to be retrieved enables faster REM construction as it is more likely that a mobile is located at a location of interest. On the other hand, the sampling becomes dependent of the traffic distribution. For instance a very large K would cause the measurements to be located mainly on traffic hotspots. This is clearly not suitable as this would lead to the random selection case. The red curve of figure 3 shows that we

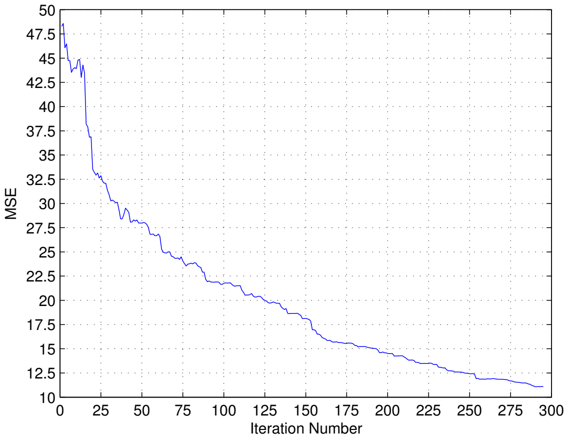


Fig. 2. Mean Square Error as measurements are added

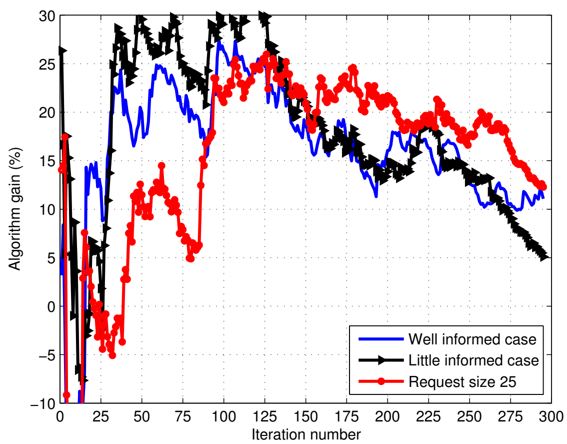


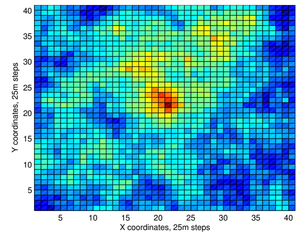
Fig. 3. Algorithm gain compared to the random choice of measurements

still have good improvement with a request size 25, and the mean gain is $15\,\%$.

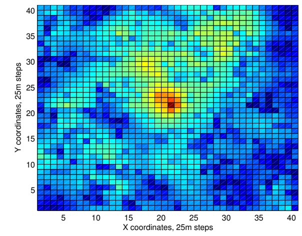
VII. Conclusions and Future Work

In this paper we propose an algorithm for collecting power measurements performed by mobile terminals. Measurements are recorded into the REM which is a powerful device for cognitive management and optimization of cellular networks. We have derived equations that enables fast REM construction. In the last section we highlight the benefit of using our algorithm comparing to the random measurement selection strategy. The gain in terms of mean squared error is around $16\,\%$, this value depends on the number of measurements we are willing to retrieve.

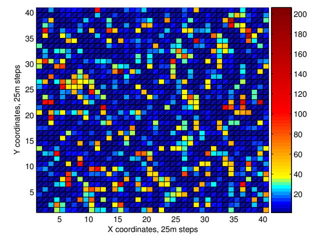
We believe that this paper constitutes a major step toward implementation of the REM. Interesting challenges still exist for adressing the design of both the Candidate Filtering and Measurement Resquesting blocks. In the near future, we will focus on adapting the algorithm so that the REM manager can detect and react to changes in the radio environment.



(a) Simulated map



(b) Interpolated map



(c) Square error map

Fig. 4. Simulation results

References

[1] B. Fette, *Cognitive Radio Technology (Communications Engineering)*. Newnes, 2006, ch. 11.

[2] B. H. Alaya-Feki, B. Sayrac, S. B. Jemaa, and E. Moulines, "Interference cartography for hierarchical dynamic spectrum access," *DySPAN*, 2008.

[3] S. Grimoud, S. Ben Jemaa, B. Sayrac, and E. Moulines, "A rem enabled soft frequency reuse scheme," *GLOBECOM 2010. BWA Workshop*, December 2010.

[4] *3GPP; TSG RAN; UTRA and E-UTRA; Radio measurement collection for Minimization of Drive Tests (MDT); Overall description; Stage 2 (Release 10)*, 3GPP Std. 37.320, Rev. 1.0.0. [Online]. Available: http://www.3gpp.org/ftp/tsg_ran/WG2_RL2/TSGR2_71/Docs/R2-105238.zip

[5] Y. Zhao, J. Reed, S. Mao, and K. Bae, "Overhead analysis for radio environment map-enabled cognitive radio networks," *1st IEEE Workshop on Networking Technologies for Software Defined Radio Networks, pp.18-25*, 2006.

[6] Y. Zhao *et al.*, "Development of radio environment map enabled case- and knowledge-based learning algorithms for ieee 802.22 wran cognitive engines," *CROWNCOM, pp.44-49*, 2007.

[7] Y. Zhao and al., "Applying radio environment maps to cognitive wireless regional area networks," *DySPAN, pp.115-118*, 2007.

[8] J. Riihijarvi, P. Mahonen, M. Wellens, and M. Gordziel, "Characterization and modelling of spectrum for dynamic spectrum access with spatial statistics and random fields," *PIMRC 2008*.

[9] J. Riihijarvi, P. Mahonen, M. Petrova, and V. Kolar, "Enhancing cognitive radios with spatial statistics: From radio environment maps to topology engine," *CROWNCOM*, 2009.

[10] N. A. Cressie, *Statistics for Spatial Data*. Wiley Series in Probability and Statistics, 1993.

[11] Z. Khan, J. Lehtomki, K. Umebayashi, and J. Vartiainen, "On the selection of the best detection performance sensors for cognitive radio networks," *IEEE Signal Processing Letters, vol.17, no.4*, Apr. 2010.

[12] Y. Selen, H. Tullberg, and J. Kronander, "Sensor selection for cooperative spectrum sensing," *DySPAN*, 2008.

[13] G. Senarath and al, "Multi-hop relay system evaluation methodology," IEEE 802.16 Broadband Wireless Access Working Group, 2007.

[14] H.-S. Jo and J.-G. Yook, "Path loss characteristics for imt-advanced systems in residential and street environments," *Antennas and Wireless Propagation Letters, IEEE*, pp. 867 –871, 2010.

[15] P. J. Diggle and P. J. Ribeiro, *Model-based Geostatistics*. Springer, 2007.