Community RF Sensing for Source Localization

Emmanouil Alimpertis, Nikos Fasarakis-Hilliard, and Aggelos Bletsas, Senior Member, IEEE

Abstract—This work experimentally demonstrates RF source location estimation error on the order of 50m, with received signal strength (RSS) measurements, from a network of smartphone users, within 800m (or more) from the source. This work incorporates careful modeling of the time-varying source transmission power, source antenna directionality (even with a simple 4-parameter model) and different path loss exponents among the various source-user links. More importantly, a vast number of RSS measurements is collected and exploited through the automated community smartphone network. The proposed methodology could be extended to other wireless scenarios.

Index Terms—Localization, RF Sources, RSS, community RF sensing, non-parametric estimation, particle filtering.

I. INTRODUCTION

T HIS work *experimentally* studies the problem of RF source location estimation, using a large collection of received signal strength (RSS) measurements, readily available from a community of mobile telephony, smartphone users. Specifically, approximately 1 measurement per second per smartphone user for a period of 7 months was automatically collected, utilizing a community of smartphone users as RF sensors of a common source (i.e. the base station).

Relevant literature in RSS-based RF source localization is quite rich. Localization error on the order of a few tens of meters is reported with Wi-Fi terminals and communication ranges on the order of tens of meters (e.g. work in [1] and references therein), as opposed to ranges on the order of hundreds of meters in this paper. Work in [2] solves a semidefinite programming problem (SDP), assuming a common and known path loss exponent (PLE) between RF source and measuring terminals, with unknown, time-constant RF source transmitted power. Work in [3] assumes that the PLE is unique across all measuring terminals and unknown, the RF source power is constant and known and offers a nonlinear least-square estimator. Work in [4] proposes a weighted least squares solution, considering both transmitted power and (the unique) PLE, constant and unknown. Moreover, work in [5] follows the same assumptions and exploits measurements collected by a moving sensor. Work in [6] assumes different PLE per measuring terminal and each grid point receives votes, based on a bounding-voting procedure; the grid point with the most votes is the estimated emitter location. All the aforementioned approaches are based on simulated data. Examples of experimental cell tower localization, based on



Fig. 1. Implemented community RF sensing: measurements from users' smartphones (e.g. RSS, GPS user location) are gathered at a web database. Mobile coverage maps and measurements (for 7 months, approx. 1 measurement per user per sec) are available (www.mysignals.gr/dataset.zip).

the strongest measurement or a weighted average of the *K*-strongest measurements [7], [8], [9], or user localization, based on modeling of the coverage areas with a student distribution [10], exploit *wardriving* (a.k.a. fingerprinting): RSS measurements and their respective GPS location are recorded, while moving in a specific area, before any processing is applied.

In sharp contrast to prior art, this work utilizes a vast collection of real world RSS measurements based on community RF sensing. Moreover, in sharp contrast to prior art, PLEs across different users are considered unequal and unknown, transmission power from source is considered time-varying, while cases between directional or isotropic source antenna are both considered. It is shown that the adoption of a directional antenna model, even a simple one, in conjunction with the rest of the aforementioned realistic assumptions, as well as the vast number of measurements - due to community RF sensing - can significantly reduce the absolute location error. Experimental results with absolute error on the order of 50m are reported, where the sources are real-world deployed mobile telephony base stations within extended ranges on the order of 800m from the measuring users-RF sensors.

II. LARGE-SCALE MEASUREMENT CAMPAIGN AND PROBLEM FORMULATION

A network of N mobile smartphone users record the RSS (in dBm) of specific base station cells, as well as their own location, through the GPS module of their smartphone (Fig. 1). Measurements and relevant information (e.g. cell ID, downlink carrier frequency and mobile user location) are uploaded periodically at a web server [11]. The known static location of user j is denoted as $\mathbf{x}_j \stackrel{\triangle}{=} \begin{bmatrix} x_j & y_j \end{bmatrix}^T$ and the location of the RF source (i.e. base station) to be estimated as $\mathbf{x}_{BS} = \begin{bmatrix} x_{BS} & y_{BS} \end{bmatrix}^T$. The collection of all user locations is denoted by $\mathbf{x} \stackrel{\triangle}{=} \begin{bmatrix} \mathbf{x}_1 \cdots \mathbf{x}_N \end{bmatrix}$. Each user j at a specific location records κ RSS measurements at time

This work was supported by the ERC-04-BLASE project, executed in the context of the "Education & Lifelong Learning" Operational Program of the National Strategic Reference Framework (NSRF), General Secretariat for Research & Technology (GSRT), funded through European Union-European Social Fund and Greek national funds. Authors are with School of ECE, Technical Univ. of Crete, Greece 73100. Email: {{ealimpetis,nfasarakis}@isc.tuc.gr, aggelos@telecom.tuc.gr}

instances $t_1, t_2, \ldots, t_{\kappa}$, for each specific cell ID $\mathbf{z}^{[j, \text{cID}]} \stackrel{\triangle}{=} \begin{bmatrix} P_{[j, \text{cID}]}^{(t_1)} \cdots P_{[j, \text{cID}]}^{(t_{\kappa})} \end{bmatrix}^T$ and the collection of all measurements is denoted by $\mathbf{z}^{[\text{cID}]} \stackrel{\triangle}{=} \begin{bmatrix} \mathbf{z}^{[1, \text{cID}]} \cdots \mathbf{z}^{[N, \text{cID}]} \end{bmatrix}$.

The RSS measurement (in dBm) for user j at a specific cell is modeled as a random variable distributed according to the log-normal distribution, with distinct PLE and variance:

$$P_{[j,\text{cID}]}^{(t)} = P_0^{(t)} - 10n_j \log_{10} \left(\frac{||\mathbf{x}_{\text{BS}} - \mathbf{x}_j||_2}{d_0} \right) + w_j^{(t)}, \quad (1)$$

where $P_0^{(t)}$ is the received power at reference distance d_0 and calculated at $d_0 = 1$ m in this work using the free-space path loss (Friis) transmission equation, with downlink frequency corresponding to the specific cell ID (typically at the 1800 MHz regime), user antenna gain equal to 0 dBi and base station antenna gain as explained below. Parameters n_j, w_j are the PLE and measurement noise, respectively, of the specific user *j*. Measurement noise is modeled as a zero-mean Gaussian with variance σ_j^2 (expressed in dB); noise r.v's are assumed independent across different users. It is noted that in contrast to this work, prior art has mainly focused on propagation models where the PLE is assumed equal across all users.

Eq. (1) implicitly assumes that the transmission power of the RF source is the same when the signal is received at reference distance d_0 or at distance $||\mathbf{x}_{BS} - \mathbf{x}_j||_2$. However, that may not be practically useful in experimental setups. For example, there is a specific downlink channel, used by GSM phones to measure the RSS, where two different transmission power levels are utilized from the base station-RF source [12].

In this work, the transmission power $P_{\text{TX}}^{(t)}$ of the RF source is modeled as a binary random variable with values P_{MIN} and P_{MAX} . The exact value of $P_{\text{TX}}^{(t)}$ is a priori unknown and must be estimated. A Markov chain is used to model the transition between $P_{\text{TX}}^{(t)}$ and $P_{\text{TX}}^{(t+1)}$ with transition probabilities that depend on the traffic load (Fig. 2); when traffic load increases, the probability for transmission at maximum power will also increase [11], [13].

A typical base station consists of three (or more) sectors (cells) typically served by directional antennas. A simplified 4-parameter model of a directional antenna is used in this work (Fig. 3) and will be contrasted to the isotropic antenna model. Specifically, a main antenna lobe at angle direction ϕ is assumed (in respect to the x-axis), with lobe opening of $2\phi_s$. The base station-RF source antenna gain at direction ϕ_j is given by (Fig. 3):

$$G_{H}(\phi_{j},\phi) = \begin{cases} G_{0}, & \phi - \phi_{s} \leq \phi_{j} \leq \phi + \phi_{s}, \\ G_{\text{SLL}}, & \phi + \phi_{s} < \phi_{j} \leq \phi + 2\phi_{s} \\ \text{or} & \phi - \phi_{s} > \phi_{j} \geq \phi - 2\phi_{s}, \\ G_{\text{BLL}}, & \text{elsewhere.} \end{cases}$$
(2)

For the numerical results section, $\phi_s = 30^\circ$, $G_0 = 0$ dB, $G_{SLL} = -3$ dB for the side lobe levels and $G_{BLL} = -23$ dB for the backside lobe level (BLL) (e.g. similar values for BLL can be found in [14]). The angle direction ϕ of base station-RF source antenna is unknown and must be estimated, while for the omnidirectional (isotropic) base station-RF source antenna,



Fig. 2. The two-level RF source transmission power Markov model for the power control at the broadcast control channel (BCCH) [12].

 $G_H(\phi_j, \phi) = 0$ dB, $\forall (\phi_j, \phi)$. Thus, the RF source radiated power at angle ϕ_j can be expressed as:

$$P_{\text{TX}}^{(t)}(\phi_j, \phi) = P_{\text{TX}}^{(t)} + G_H(\phi_j, \phi) \text{ (dBm)}.$$
 (3)

The goal is to estimate vector $\theta^{(t)}$, with all the unknowns (including RF source coordinates), from *real-world measurements* $\mathbf{z}^{[cID]}$, provided by the community of N smartphones-RF sensors. Measurements for 7 months were collected (approx. 1 measurement per smartphone user per sec):

$$\boldsymbol{\theta}^{(t)} \stackrel{\triangle}{=} \begin{bmatrix} x_{\text{BS}} & y_{\text{BS}} & \phi & P_{\text{TX}}^{(t)} & n_1 \cdots n_N \end{bmatrix}^T.$$
(4)

Thereinafter, $\mathcal{U}[a, b]$ denotes the uniform distribution in [a, b] and $\mathcal{N}(a, b)$ denotes the Gaussian distribution with expected value a and variance b.

III. LOCALIZATION CASE STUDY

It was experimentally found that an initial estimate of the PLEs (to be refined in a latter step) reduced the localization error, compared to the case where PLEs were randomly initialized. That can be attributed to the fact that a small variance in PLE significantly alters RSS, according to Eq. (1). *Initial Estimation of PLEs:* Assuming availability of an initial estimate $\hat{\mathbf{x}}_{BS}^{init}$ and $d_0 = 1$ m, an initial estimate of \hat{n}_j results from Eq. (1):

$$\widehat{n}_{j} = \frac{\overline{P}_{0_{\text{MAX}}} - (1/\kappa) \sum_{i=1}^{\kappa} P_{[j,\text{cID}]}^{(t_{i})}}{10 \log_{10} \|\mathbf{x}_{j} - \widehat{\mathbf{x}}_{\text{BS}}^{\text{init}}\|_{2}}, \forall j,$$
(5)

where $\overline{P}_{0_{MAX}}$ is the received power at $d_0 = 1$ m given by the Friis transmission equation (free space path-loss), assuming that the base station is transmitting at P_{MAX} and κ RSS



Fig. 3. Simple 4-parameter RF source directional antenna gain modeling. Parameters $G_0, \phi_s, G_{SLL}, G_{BLL}$ are needed.



Fig. 4. Results for cell-ID 60562 located at the city of Chania city center. K = 2000 different PF runs were performed. T = 60 min.

measurements were randomly selected for each user j. An initial estimate of $\hat{\mathbf{x}}_{BS}^{init}$ is given by the voting algorithm in [6], where PLEs are considered unknown and bounded.

Particle Filtering: The probability density $p(\theta | \mathbf{z}^{[\text{cID}]})$ is modeled and estimated with non-parametric procedures based on particle filtering [15]. Specifically, M particles at time t are considered, with corresponding weights $\{w_t^{[m]}\}$:

$$\left\{ \boldsymbol{\theta}_{t}^{[m]} \!=\! \begin{bmatrix} x_{\text{BS}}^{[m]} & y_{\text{BS}}^{[m]} & \phi^{[m]} & P_{\text{TX}}^{(t),[m]} & n_{1}^{[m]} \cdots n_{N}^{[m]} \end{bmatrix}^{T} \right\}.$$

The RF source is assumed static (i.e. immobile) where (for the case of RF source directional antenna), each cell is illuminated by the corresponding directional antenna at a fixed direction (unchanged for the measurement duration). Moreover, for the time window of utilized experiments, the PLEs are considered constant.

Thus, the *update* operation of particle filtering is restricted to the transmission power of the base station and is given by:

$$P_{\text{TX}}^{(t),[m]} \sim \Pr\left(P_{\text{TX}}^{(t),[m]} | P_{\text{TX}}^{(t-1),[m]}\right),\tag{6}$$

where the conditional probability is described by the Markov chain of Fig. 2.

The *correction* operation for each particle weight at time t is given by:

$$w_t^{[m]} = p\left(\mathbf{z}^{(t)} | \boldsymbol{\theta}_t^{[m]}\right) = \prod_{j=1}^N p\left(P_{[j]}^{(t)} | \boldsymbol{\theta}_t^{[m]}\right), \tag{7}$$

where independence of RSS measurements across different users has been exploited. The conditional p.d.f. value of each measurement is based on the log-normal p.d.f. (Eq. (1)):

$$p\left(P_{[j]}^{(t)}|\boldsymbol{\theta}_{t}^{[m]}\right) = \frac{1}{\sigma_{j}\sqrt{2\pi}}\exp\left(-\left(P_{[j]}^{(t)}-\overline{P_{[m,j]}^{(t)}}\right)^{2} / 2\sigma_{j}^{2}\right).$$
(8)

Parameter $\overline{P_{[m,j]}^{(t)}} \stackrel{\triangle}{=} \mathbb{E}\left[P_{[j]}^{(t)} | \boldsymbol{\theta}_t^{[m]}\right]$ is the expected value of RSS at location \mathbf{x}_i , given $\boldsymbol{\theta}_t^{[m]}$:

$$\overline{P_{[m,j]}^{(t)}} = P_0^{(t),[m]} - 10n_j^{[m]} \log_{10} \left(\|\mathbf{x}_{BS}^{[m]} - \mathbf{x}_j\|_2 / d_0 \right), \quad (9)$$

where $P_0^{(t),[m]}$ is the received signal strength at reference distance $d_0 = 1$ m calculated by the Friis transmission equation

with base station transmission power equal to $P_{\text{TX}}^{(t),[m]}$ and base station cell antenna gain given by Eq. (2). Angle ϕ_j for user j given $x_{\text{BS}}^{[m]}, y_{\text{BS}}^{[m]}$ is easily calculated from Fig. 3:

$$\phi_j = \arctan\left((y_j - y_{\text{BS}}^{[m]})/(x_j - x_{\text{BS}}^{[m]})\right).$$
 (10)

The standard deviation σ_j is estimated from the whole measurement data-set for user j at a specific location, exploiting standard techniques for parameter estimation with Gaussian random variables (the RSS in dBm values are assumed Gaussian random variables according to Eq. (1)) [16].

Finally, the low-variance sampling procedure of the particles was utilized [15]. The pseudo-code of the PF follows.

PF Algorithm: Particle Filtering for estimating $p(\theta|\mathbf{z}^s)$

- (1): Initialization of Variables:
- (2): Set M (number of Particles), (3): X, Y (dimensions of area of interest), T (time window). (4): Initialization of Particles, $\forall m = 1 : M$ (4): Initialization of Particles, $\forall m = 1 : M$ (5): $x_{t=0}^{[m]} \sim \mathcal{U}[0, X], y_{t=0}^{[m]} \sim \mathcal{U}[0, Y], \phi_{t=0}^{[m]} \sim \mathcal{U}[0, 2\pi].$ (6): $P_{\text{TX}}^{(t=0),[m]} = b, b \in \{P_{\text{MIN}}, P_{\text{MAX}}\}, \text{Pr}(b) = 0.5.$ (7): $n_{j,t=0}^{[m]} \sim \mathcal{N}(\hat{n}_{j}, \sigma_{n}^{2}), \text{ with } \hat{n}_{j} \text{ from (5).}$ (8): $\boldsymbol{\theta}_{t=0}^{[m]} = \left[x_{\text{BS}}^{[m]} \ y_{\text{BS}}^{[m]} \ \phi^{[m]} \ P_{\text{TX}}^{(t),[m]} \ n_{1}^{[m]} \cdots n_{N}^{[m]} \right]^{T}.$ (9): $\mathbf{z}^s \leftarrow \mathbf{Select} \ T$ consecutive RSS measurements from $\mathbf{z}^{[\text{cID}]}$ (10): for t = 1 : T do {*time-step*} for m = 1 : M do {particle index} (11): Sample $P_{\text{TX}}^{(t),[m]} \sim \Pr\left(P_{\text{TX}}^{(t),[m]}|P_{\text{TX}}^{(t-1),[m]}\right)$. Calc. $\phi_j, \forall j$, from (10). Calc. $P_{\text{TX}}^{(t)}(\phi_j, \phi^{[m]}), \forall j$ from (3) given $\boldsymbol{\theta}_t^{[m]}$. $w_t^{[m]} = p\left(\mathbf{z}^{(t),s}|\boldsymbol{\theta}_t^{[m]}\right)$, from (7). (12): (13): (14): (15): (16): end for Normalize weights $\boldsymbol{w}_t^{[1:M]}$. (17) $\boldsymbol{\theta}_{t}^{[1:M]} \leftarrow \text{Low-VarianceSampler}\left(\boldsymbol{\theta}_{t}^{[1:M]}, \boldsymbol{w}_{t}^{[1:M]}\right).$ (18)(19): end for

At the end of particle filtering algorithm the final particles $\theta_T^{[1:M]}$ produce an estimate of the conditional density function $p(\theta|\mathbf{z}^s)$ using the histogram approach [15]. With marginalization, an estimate of the conditional density function $p(\mathbf{x}_{BS}|\mathbf{z}^s)$

is readily available. The estimate of base station location $\widehat{\mathbf{x}}_{BS}$ is offered by the conditional mean:

$$\widehat{\mathbf{x}}_{BS} = [\widehat{x}_{BS} \ \widehat{y}_{BS}]^T = \mathbb{E}[\mathbf{x}_{BS} | \mathbf{z}^s].$$
(11)

The procedure could be repeated for K different z^s and the average value of all estimates produces the final outcome:

$$\widehat{\mathbf{x}}_{\mathrm{BS}}^{mean} = \sum_{k=1}^{K} \frac{1}{K} [\widehat{x}_{\mathrm{BS}}^{k} \ \widehat{y}_{\mathrm{BS}}^{k}]^{T}.$$
(12)

IV. NUMERICAL RESULTS

Numerical results are reported for real-world measurements from similar sensitivity smartphones and comparison with prior art is limited to techniques applicable to real-world, nonsimulated data. Number of particles M is set to M = 140000, X = Y = 1000 m and T = 60 min (line (2) of the algorithm), with approximately 1 measurement per minute per user. In line (7) of the algorithm, $\sigma_n = 0.3$ has been selected heuristically, since a smaller value resulted in particle depletion [15], while a larger value also required a larger number of particles; values $P_{MAX} = 43.5$ dBm, $P_{MIN} = 41.5$ dBm were provided by the mobile telephony provider specifications. Location of the N users-smartphones was apriori known (in order to avoid smartphone GPS inaccuracies) and RSS measurements for the GSM modem were utilized (3G RSS measurements were also readily available).

Fig. 4-(a) offers the final estimated base station-RF source location, when K = 2000 different datasets of measurements were used, for N = 4 users and a specific cell (60562). In other words, $2000 \times 60 \times 4 = 480000$ independent RSS measurements were utilized. The RF source antenna was assumed isotropic (no directionality). The reported estimated value is based on Eq. (12) and the true location is also depicted. An additional user is added in Fig. 4-(b) (i.e. N = 5) and the estimation is repeated without significant improvement. When the antenna directionality model is included (Fig. 4-(c)), specific areas of interest are excluded and the final estimate is significantly improved. In other words, the rich measurement dataset, due to the community RF sensing infrastructure, was beneficial when the (simple) base station-RF source antenna directionality model was incorporated.

Table 1 shows that the absolute location error in cell 60562 with base station-RF source directionality modeling is approximately three times smaller (scenario 6) compared to the case with isotropic antennas (scenarios 4, 5). It is emphasized that the reported localization error on the order of 50m is achieved even when communication distances between smartphone users and RF source can exceed 800m. Performance

Absolute Localization Error(m) 5 3 4 6 60561 516.9126.793.6 382.484.2 48.260562 277.7 419.2 158.3145.7149.658.3

TABLE I

1. TOP-K RSSI, 2. Strongest RSSI, 3. Grid Voting. 4., 5. PF-Isotropic Antenna (N = 4 and N = 5 respectively), 6. PF-Directional Antenna.

V. CONCLUSION

A large number of RSS measurements, through an automated *community* user network, can significantly improve the location estimate of a RF source, as this work experimentally demonstrated, even when distances between RF source and users exceed 800m. The proposed methodology is a concrete example of a crowd-sensing application [17] and could be extended to other RF source transmission power/antenna directionality models and wireless propagation scenarios.

REFERENCES

- Y. Cheng, Y. Chawathe, A. LaMarca, and J. Krumm, "Accuracy characterization for metropolitan-scale Wi-Fi localization," in ACM Int. Conf. on Mobile Systems, Applications and Services (MobiSys), Seattle WA, USA, 2005, pp. 233–245.
- [2] R. M. Vaghefi, M. R. Gholami, and E. G. Strom, "RSS-based sensor localization with unknown transmit power," in *IEEE Int. Conf. Acoustics*, *Speech, and Signal Processing (ICASSP)*, Prague, Czech Republic, May 2011, pp. 2480–2483.
- [3] X. Li, "RSS-based location estimation with unknown pathloss model," *IEEE Trans. Wireless Commun.*, vol. 5, no. 12, pp. 3626–3633, Dec. 2006.
- [4] G. Wang, H. Chen, Y. Li, and M. Jin, "On received-signal-strength based localization with unknown transmit power and path loss exponent," *IEEE Wireless Commun. Lett.*, vol. 1, no. 5, pp. 536–539, Oct 2012.
- [5] A. A. Gorji and B. D. O. Anderson, "Emitter localization using receivedstrength-signal data," *Elsevier Signal Process.*, vol. 93, no. 5, pp. 996– 1012, May 2013.
- [6] J. Shirahama and T. Ohtsuki, "RSS-based localization in environments with different path loss exponent for each link," in *IEEE Vehicular Technology Conf. (VTC).*, Singapore, May 2008, pp. 1509–1513.
- [7] J. Yang, A. Varshavsky, H. Liu, Y. Chen, and M. Gruteser, "Accuracy characterization of cell tower localization," in ACM Int. Conf. on Ubiquitous Computing (Ubicomp), Copenhagen, Denmark, Sep 2010, pp. 223–226.
- [8] A. Varshavsky, D. Pankratov, J. Krumm, and Eyal De Lara, "Calibree: Calibration-free localization using relative distance estimations," in *Springer Pervasive Computing*, pp. 146–161. May 2008.
- [9] M. Y. Chen, T. Sohn, D. Chmelev, D. Haehnel, J. Hightower, J. Hughes, A. LaMarca, F. Potter, I. Smith, and A. Varshavsky, "Practical metropolitan-scale positioning for GSM phones," in *Int. Conf. on Ubiquitous Computing (Ubicomp)*, pp. 225–242. Newport Beach, California, Sep 2006.
- [10] M. Dashti, S. Ali-Löytty, L. Wirola, P. Müller, H. Nurminen, and R. Piché, "Robust Kalman filter for positioning with wireless BS coverage areas," in 9th Workshop on Positioning Navigation and Communication (WPNC), Dresden, Germany, March 2012, pp. 83–88.
- [11] E. Alimpertis, "Cell Tower Discovery via Particle Filtering: Progress Report," Tech. Rep., School of ECE, Technical Univ. of Crete, Crete, July 2013, [Online at http://tinyurl.com/techReportPFs].
- [12] M. Sailly, G. Sebire, and E. Riddington, GSM/EDGE Evolution and Performance, Wiley, Chichester, UK, 2011, ch. 13.5.3, Energy Saving Mode on BCCH transceiver.
- [13] S. Bregni, R. Cioffi, and M. Decina, "An empirical study on statistical properties of GSM telephone call arrivals," in *IEEE Global Commun. Conf. (Globecom)*, San Francisco, CA, Nov 2006, pp. 1–5.
- [14] Specification sheet, 900 MHz Sector Antenna, REN 68715 SN-90, CSG Networks.
- [15] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*, The MIT Press, Cambridge, MA, 2006.
- [16] B. C. Levy, Principles of Signal Detection and Parameter Estimation, Springer, New York, USA, 2008.
- [17] W. Guo and S. Wang, "Mobile crowd-sensing wireless activity with measured interference power," *IEEE Wireless Commun. Lett.*, vol. 2, no. 5, pp. 539–542, Oct 2013.