This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/LWC.2017.2785228

Target Localization Using Sensor Location Knowledge in Wireless Sensor Networks

Chun-Yi Wei, *Member*, *IEEE*, Hsuan-Yin Lin, *Member*, *IEEE*, Po-Ning Chen, *Senior Member*, *IEEE*, Yunghsiang S. Han, *Fellow*, *IEEE*, Pramod K. Varshney, *Fellow*, *IEEE*

***Abstract*—Incorporating error correcting coding techniques into target localization provides an immediate advantage that existing decoding algorithms can be used to determine which area the target is most likely located in. The important knowledge of exact sensor positions is, however, ignored in these decoding algorithms. This paper revisits the problem and shows that based on the weighted average of sensor positions with binary weightings from local decisions, a newly proposed decoding criterion can achieve a much better accuracy in target localization than the soft- and hard-decision rules proposed in [1], particularly when a certain number of sensors are under Byzantine attacks.**

***Index Terms*—Target Localization, Wireless Sensor Networks, Error Correcting Codes, Byzantines, Quantizer Design**

I. Introduction

Target localization in wireless sensor networks (WSNs) has become an important task due to the growing demand for surveillance of certain regions of interest (ROIs). Numerous localization techniques have been shown to perform well under various scenarios [1]–[5]. For example, time of arrival based techniques, e.g., [2] and [3], provide a reliable localization result but require highly synchronized sensors. Alternatively, localization techniques using analog measurements such as the received signal strength (RSS) [4], [5] are attractive due to their low energy consumption and cost-effective deployment but often lead to less reliable performance.

In 2014, an efficient iterative classification framework for target localization was proposed in [1]. With an elaborately devised error correcting code, it subdivides the ROI into multiple regions, each of which is associated with a codeword. The most likely region that the target lies in is then determined via a decoder at the Fusion Center (FC), and becomes the ROI in the next iteration. As such, the ROI, where the target is located in, is reduced in size iteratively. The center of the final ROI is regarded as the estimate of target location. It has been shown in [1] that an intuitive design of local thresholds for the quantization of local measurements ensures an asymptotically vanishing misclassification error (in number of sensors) and the proposed iterative classification framework can efficiently zoom to the target location by successfully excluding irrelevant candidate regions at each iteration. When a practical WSN with a finite number of sensors is considered, it was discovered in [6] that the intuitive threshold design is only sub-optimal and localization performance can be significantly improved by a better design that minimizes a misclassification error bound derived therein.

The decision rules in [1], [6] only consider the local observations from sensors. The important information of exact sensor positions that is actually known to the FC is not exploited. This leads to the contribution of this paper, where a new soft metric that incorporates the exact sensor positions into the estimate of target location is derived. The idea is to first provide an initial estimate of target location using the weighted average of sensor positions with binary weightings from local decisions. Then, the region that the target lies in can be determined based on the *a posteriori* probability given that the target was at the initially estimated location, which is optimal if the initial estimate is exact. Since the proposed soft metric is equivalent to compensating the decision rules in [1], [6] with multiplicative weightings and additive offsets, its implementation complexity can be considered comparable to those in [1], [6]. Simulation results confirm that the new soft metric can achieve a much more accurate estimate of target location than the legacy hard- and soft-decision rules in [1], [6], particularly in the presence of Byzantine attacks [7].

The rest of the paper is structured as follows. Section II introduces the system model and problem formulation. Section III derives the soft metric and Section IV presents the simulation results. Finally, Section V concludes the work.

II. Preliminaries

A. System model

Denote the target location in a 2-dimensional Cartesian plane by $\boldsymbol{\theta} \triangleq (\boldsymbol{x}, \boldsymbol{y})$, which is assumed to be uniformly distributed over the ROI.[1] Similarly, denote by $\theta_i = (x_i, y_i)$ the position of the ith sensor for $1 \leq i \leq N$, and assume that this information is known to the FC. There are a total of N sensors deployed in the ROI $\mathcal{R}$ of the WSN.

Since the target is generally mobile and a low frequency beacon signal could be used for a cost-economic system, the target-sensor links may be time-varying and also experience rapid RSS fluctuations over a travel distance of few wavelengths long. The so-called small-scale fading is accordingly considered in the target-sensor links in addition to the usual large-scale fading [8].[2] The signal captured by the ith sensor

C.-Y. Wei is with Dept. of Comm. Eng., Nat'l Taipei Univ., Taiwan.
H.-Y. Lin was with Inst. of Comm. Eng., Nat'l Chiao-Tung Univ., Taiwan, and is now with the Simula@UiB, Bergen, Norway.
P.-N. Chen is with Inst. of Comm. Eng., Nat'l Chiao-Tung Univ., Taiwan.
Y. S. Han is with School of Elec. Eng. & Intell., Dongguan Univ. of Tech., China, and College of Elec. Eng. & Comp. Sci., Nat'l Taipei Univ., Taiwan.
P. K. Varshney is with Dept. of Elec. Eng. & Comp. Sci., Syracuse Univ., USA.

[1] Throughout the paper, both $\boldsymbol{\theta}$ and $(\boldsymbol{x}, \boldsymbol{y})$ are used to denote the target location. As a convention, underlined letters denote vectors; boldface letters such as $\boldsymbol{u}_i$ and $\underline{\boldsymbol{u}} = (\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_N)$ respectively denote random variables and random vectors; and non-boldface letters such as u_i and $\underline{u} = (u_1, u_2, \ldots, u_N)$ are their realizations.

[2] In case the target is an acoustic source as assumed in [5] and [9], the small-scale fading effect can be neglected.

is therefore modeled by

$$\boldsymbol{s}_i = \boldsymbol{g}_i\boldsymbol{a}_i + \boldsymbol{n}_i \quad \text{for } i = 1, 2, \ldots, N,$$

where $\boldsymbol{g}_i$ denotes a unit-variance Rayleigh distributed fading attenuation, $\boldsymbol{a}_i$ is the beacon signal propagated from the target, and $\boldsymbol{n}_i$ is the noise experienced during the sensing operation and can be characterized as a zero-mean Gaussian random variable with variance σ^2. The signal amplitude $\boldsymbol{a}_i$ follows a power decay model as a function of the distance between the target and the sensor:

$$\boldsymbol{a}_i^2 = \boldsymbol{G}P_0\left(\frac{d_0}{d_{\mathrm{E}}(\boldsymbol{\theta}, \theta_i)}\right)^n \quad \text{for } i = 1, 2, \ldots, N, \tag{1}$$

where the log-normal distributed $\boldsymbol{G}$ reflects the random shadowing effect, $d_{\mathrm{E}}(\boldsymbol{\theta}, \theta_i) \triangleq \sqrt{(\boldsymbol{x} - x_i)^2 + (\boldsymbol{y} - y_i)^2}$ is the Euclidean distance between the target and the ith sensor, and P_0 is the power measured at reference distance $d_0 = 1$. Since the shadowing effect is imperceptible in a free space environment, $\boldsymbol{G} = 1$ and $n = 2$ are assumed in this work.

Due to limited resources, sensors are assumed to only send binary decisions to the FC. The binary decision made by the ith sensor is given by $\boldsymbol{u}_i = \mathbf{1}\{\boldsymbol{s}_i > \eta_i\}$, where $\mathbf{1}\{\cdot\}$ denotes the set indicator function, and η_i is the local threshold used by the ith sensor, for which a near-optimal design has been elaborated in [6].

The noisy wireless link between the ith sensor and the FC [10], [11] could be modeled as

$$\boldsymbol{v}_i = \boldsymbol{h}_i(-1)^{\boldsymbol{u}_i}\sqrt{\mathrm{E_b}} + \boldsymbol{w}_i, \tag{2}$$

where $\boldsymbol{h}_i$ denotes the unit-variance Rayleigh distributed fading attenuation of the wireless link, $\mathrm{E_b}$ is the energy required to transmit one sensor bit, and $\boldsymbol{w}_i$ is the additive Gaussian noise with mean zero and variance σ_{w}^2.

B. Problem Formulation

We first recapitulate the notations from [1]. The ROI $\mathcal{R}^{k-1}$ decided at the previous iteration is subdivided equally into L non-overlapping regions, denoted as $\mathcal{R}_1^k$, $\mathcal{R}_2^k$, ..., $\mathcal{R}_L^k$, where superscript $k = 1, 2, \ldots, K$ indexes the iteration. The number of sensors involved at the kth iteration is thus equal to $N_k \triangleq N/L^{k-1}$. A binary codeword $\underline{c}_\ell^k = (c_{\ell,1}^k, c_{\ell,2}^k, \ldots, c_{\ell,N_k}^k)$ that follows the design in [1] is associated with region $\mathcal{R}_\ell^k$ for $\ell = 1, 2, \cdots, L$, where $c_{\ell,j}^k = 1$ if $\theta_j \in \mathcal{R}_\ell^k$, and $c_{\ell,j}^k = 0$, otherwise. After receiving $\underline{v}^k = (v_1^k, v_2^k, \ldots, v_{N_k}^k)$ from the sensors, the FC calculates their corresponding log-likelihood ratios as $\xi_i^k = \log[f(\boldsymbol{v}_i^k = v_i^k|u_i^k = 0)/f(\boldsymbol{v}_i^k = v_i^k|u_i^k = 1)]$, $i = 1, 2, \ldots, N_k$, where $f(\boldsymbol{v}_i^k|u_i^k)$ is the probability density function (pdf) for the Rayleigh fading channel model in (2).

Based on the above framework, the soft-decision decoder adopted in [1], [6] decides the region, where the target is located in, according to

$$\hat{\ell}^k = \arg\max_{1\le\ell\le L}\Pr(\underline{v}^k|\theta \in \mathcal{R}_\ell^k) = \arg\min_{1\le\ell\le L} d_{\mathrm{E}}(\underline{\xi}^k, (-1)^{\underline{c}_\ell^k}). \tag{3}$$

Alternatively, the minimum Hamming distance fusion rule may be adopted to determine the most-likely ROI in terms of

$$\hat{\ell}^k = \arg\min_{1\le\ell\le L} d_{\mathrm{H}}(\underline{\hat{u}}^k, \underline{c}_\ell^k), \tag{4}$$

where $\hat{u}_i^k = \mathbf{1}\{\xi_i^k \le 0\}$ is the estimate of u_i^k, and $\underline{\hat{u}}^k = (\hat{u}_1^k, \hat{u}_2^k, \ldots, \hat{u}_{N_k}^k)$, and $d_{\mathrm{H}}(\cdot,\cdot)$ is the Hamming distance. It is obvious that the decision rules in both (3) and (4) are only functions of either $\underline{\xi}^k$ or $\underline{\hat{u}}^k$, and do not use the known information of sensor positions $\{\theta_\ell = (x_\ell, y_\ell)\}_{\ell=1}^N$.

A practical question that follows is whether further improvement of target localization accuracy can be obtained by exploiting $\{\theta_\ell = (x_\ell, y_\ell)\}_{\ell=1}^N$ in the fusion rule. An answer is given by the proposed new soft metric in the next section.

III. A New Soft metric for iterative classification

A. An initial estimate of target location θ

Noting that the closer the target is to sensor i, the higher the probability of $\hat{u}_i = 1$, one can infer that the target is more likely to reside in region $\mathcal{R}_J^k$, where $J \triangleq \arg\max_{1\le j\le L}\sum_{\ell\in\mathcal{S}_j^k}\hat{u}_\ell$ is the index of the region that yields the largest number of 1's in its respective $\hat{u}_\ell$-components, and $\mathcal{S}_j^k$ denotes the set of indices of sensors in region $\mathcal{R}_j^k$. A simple initial estimate of target location can, therefore, be given by:

$$\tilde{\theta} = (\tilde{x}, \tilde{y}) = \left(\frac{\sum_{\ell\in\mathcal{S}_J^k}\hat{u}_\ell\cdot x_\ell}{\sum_{\ell\in\mathcal{S}_J^k}\hat{u}_\ell}, \frac{\sum_{\ell\in\mathcal{S}_J^k}\hat{u}_\ell\cdot y_\ell}{\sum_{\ell\in\mathcal{S}_J^k}\hat{u}_\ell}\right). \tag{5}$$

B. A soft metric derived from probability-of-detection function

The probability of detection based on (4) is given by $\sum_{i=1}^L\int_{\theta\in\mathcal{R}_i^k} f_{\boldsymbol{\theta}}(\theta)\Pr(\underline{\hat{\boldsymbol{u}}}^k \in \mathcal{D}_i^k|\boldsymbol{\theta} = \theta)d\theta$, where $f_{\boldsymbol{\theta}}(\cdot)$ is the pdf of $\boldsymbol{\theta}$, and $\{\mathcal{D}_i^k\}_{i=1}^L$ are disjoint partitions on $\{0,1\}^{N_k}$ satisfying that for each $\underline{u}^k \in \mathcal{D}_i^k$,

$$d_{\mathrm{H}}(\underline{u}^k, \underline{c}_i^k) \le \min_{1\le j\le L,\ j\ne i} d_{\mathrm{H}}(\underline{u}^k, \underline{c}_j^k). \tag{6}$$

A continuing elaboration from Section III-A is that if the initial estimate $\tilde{\theta}$ is close to θ with high probability, $f_{\boldsymbol{\theta}}(\cdot)$ that is often assumed uniform over $\mathcal{R}^{k-1}$ could be replaced by one with high probability density near $\tilde{\theta}$. A soft metric can thus be proposed based on $\zeta_i^k \triangleq \Pr(\underline{\hat{\boldsymbol{u}}}^k \in \mathcal{D}_i^k|\boldsymbol{\theta} = \tilde{\theta})$.

Under $\tilde{\theta} \in \mathcal{R}_i^k$ and negligible sensing/channel imperfections, ζ_i^k is expected to be higher than $1/2$; hence, $\sum_{j=1}^L\zeta_j^k = 1$ implies $\zeta_i^k > \frac{1}{2} > \zeta_j^k\ \forall j \ne i$, which is equivalent to

$$\lambda_i^k \triangleq \log\left(\frac{\zeta_i^k}{1-\zeta_i^k}\right) > 0 > \lambda_j^k \triangleq \log\left(\frac{\zeta_j^k}{1-\zeta_j^k}\right)\ \forall j \ne i.$$

Subject to $\tau_i^k \triangleq d_{\mathrm{H}}(\underline{u}^k, \underline{c}_i^k)$ and $L > 2$, it can also be anticipated that $N_k - 2\tau_i^k > N_k - 2\tau_j^k > 0\ \forall j \ne i$ holds, based on which we obtain $(\tau_i^k - \frac{1}{2}N_k)\lambda_i^k < 0 < (\tau_j^k - \frac{1}{2}N_k)\lambda_j^k\ \forall j \ne i$. A new soft metric $(\tau_i^k - \frac{1}{2}N_k)\lambda_i^k$ for distributed classification is obtained, which gives a new decision rule as:

$$\hat{\ell}^k = \arg\min_{1\le\ell\le L}\left(\lambda_\ell^k d_{\mathrm{H}}(\underline{u}^k, \underline{c}_\ell^k) - \frac{1}{2}\lambda_\ell^k N_k\right). \tag{7}$$

We end this subsection by remarking that for given $\{\theta_\ell = (x_\ell, y_\ell)\}_{\ell=1}^N$, the initial estimate $\tilde{\theta}$ is only a function of the binary sequence $\{\hat{u}_\ell\}_{\ell\in\mathcal{S}^{k-1}}$, and so are $\{\zeta_i^k\}_{i=1}^L$ and $\{\lambda_i^k\}_{i=1}^L$. It is thus possible to establish a 2^{N_k}-by-L look-up table from $\{\hat{u}_\ell\}_{\ell\in\mathcal{S}^{k-1}}$ to $\{\lambda_i^k\}_{i=1}^L$ for the determination of the latter. As such, the computational effort of the new decision rule is mainly from the computation of (7) itself. In comparison with the legacy hard-decision fusion rule in (4), the newly proposed decision rule only requires additional L additions and $2L$ multiplications, yet simulations show that a much better performance in target localization can be achieved.

C. Approximation of ζ_i^k

$\zeta_i^k \triangleq \Pr(\underline{\hat{\boldsymbol{u}}}^k \in \mathcal{D}_i^k | \boldsymbol{\theta} = \tilde{\theta})$ depends on the statistics of sensing noises $\{\boldsymbol{n}_j\}_{j\in\mathcal{S}^{k-1}}$, propagation fading $\{\boldsymbol{g}_j\}_{j\in\mathcal{S}^{k-1}}$, channel fading $\{\boldsymbol{h}_j\}_{j\in\mathcal{S}^{k-1}}$ and channel noises $\{\boldsymbol{w}_j\}_{j\in\mathcal{S}^{k-1}}$, and hence, its derivation, although theoretically possible, could be a numerical burden. A simple approximation of ζ_i^k based on the union inequality is thus derived in this subsection. For simplicity, both superscript and subscript k that index the iteration will be omitted in the derivation below.

Let N sensors be equally distributed over regions $\mathcal{R}_1$, $\mathcal{R}_2$, …, $\mathcal{R}_L$, and assume $\bar{N} \triangleq N/L$ is an integer. From (6), $\beta_i \triangleq 1 - \zeta_i = \sum_{j=1,j\neq i}^{L} \zeta_j$ can be expressed as:

$$\begin{aligned}
\beta_i &= \Pr\left(\underline{\hat{\boldsymbol{u}}} \notin \mathcal{D}_i \middle| \boldsymbol{\theta} = \tilde{\theta}\right) \\
&= \Pr\left(\bigcup_{1\le j\le L, j\neq i} \left\{ d_{\mathrm{H}}(\underline{\hat{\boldsymbol{u}}}, \underline{c}_i) \ge d_{\mathrm{H}}(\underline{\hat{\boldsymbol{u}}}, \underline{c}_j) \right\} \middle| \boldsymbol{\theta} = \tilde{\theta} \right) \\
&= \Pr\left(\bigcup_{1\le j\le L, j\neq i} \left\{ \sum_{\ell=1}^{N} \hat{\boldsymbol{u}}_\ell \oplus c_{i,\ell} \ge \sum_{\ell'=1}^{N} \hat{\boldsymbol{u}}_{\ell'} \oplus c_{j,\ell'} \right\} \middle| \boldsymbol{\theta} = \tilde{\theta} \right),
\end{aligned}$$

where "$\bigcup$" denotes the OR logical inference for events and "$\oplus$" is the binary XOR operation. Since $c_{i,\ell} = c_{j,\ell} = 0$ for $\ell \notin \mathcal{S}_i \cup \mathcal{S}_j$, $c_{i,\ell} = 1 - c_{j,\ell} = 1$ for $\ell \in \mathcal{S}_i$, and $c_{j,\ell} = 1 - c_{i,\ell} = 1$ for $\ell \in \mathcal{S}_j$, we have

$$\sum_{\ell=1}^{N} \hat{\boldsymbol{u}}_\ell \oplus c_{i,\ell} \ge \sum_{\ell'=1}^{N} \hat{\boldsymbol{u}}_{\ell'} \oplus c_{j,\ell'} \Longleftrightarrow \sum_{\ell\in\mathcal{S}_i} \hat{\boldsymbol{u}}_\ell \le \sum_{\ell'\in\mathcal{S}_j} \hat{\boldsymbol{u}}_{\ell'}.$$

By noting that $\{\boldsymbol{X}_i \triangleq \sum_{\ell\in\mathcal{S}_i} \hat{\boldsymbol{u}}_\ell\}_{i=1}^{L}$ are independent given $\boldsymbol{\theta} = \tilde{\theta}$, an upper bound on β_i is derived as follows:

$$\begin{aligned}
\beta_i &= \Pr\left(\bigcup_{1\le j\le L, j\neq i} \{\boldsymbol{X}_i \le \boldsymbol{X}_j\} \middle| \boldsymbol{\theta} = \tilde{\theta} \right) \\
&\le \sum_{j=1,j\neq i}^{L} \Pr\left(\boldsymbol{X}_i \le \boldsymbol{X}_j \middle| \boldsymbol{\theta} = \tilde{\theta}\right) \\
&= \sum_{j=1,j\neq i}^{L} \sum_{m=0}^{\bar{N}} \Pr\left(\boldsymbol{X}_i = m \middle| \boldsymbol{\theta} = \tilde{\theta}\right) \Pr\left(\boldsymbol{X}_j \ge m \middle| \boldsymbol{\theta} = \tilde{\theta}\right) \\
&\le \sum_{j=1,j\neq i}^{L} \sum_{m=0}^{\bar{N}} \Pr\left(\boldsymbol{X}_i = m \middle| \boldsymbol{\theta} = \tilde{\theta}\right) \frac{\mathrm{E}[e^{t_j \boldsymbol{X}_j} | \boldsymbol{\theta} = \tilde{\theta}]}{e^{t_j m}} \qquad (8) \\
&= \sum_{j=1,j\neq i}^{L} \mathrm{E}[e^{-t_j \boldsymbol{X}_i} | \boldsymbol{\theta} = \tilde{\theta}]\ \mathrm{E}[e^{t_j \boldsymbol{X}_j} | \boldsymbol{\theta} = \tilde{\theta}] \\
&= \sum_{j=1,j\neq i}^{L} \prod_{\ell\in\mathcal{S}_i} \left(q_\ell + (1 - q_\ell)e^{-t_j}\right) \prod_{\ell'\in\mathcal{S}_j} \left(q_{\ell'} + (1 - q_{\ell'})e^{t_j}\right) \\
&\le \sum_{j=1,j\neq i}^{L} \left(\bar{q}_i + (1 - \bar{q}_i)e^{-t_j}\right)^{\bar{N}} \left(\bar{q}_j + (1 - \bar{q}_j)e^{t_j}\right)^{\bar{N}} \qquad (9)
\end{aligned}$$

where (8) holds for arbitrary positive $\{t_j\}_{j=1}^{L}$ because of Markov's inequality, $\bar{q}_i \triangleq (1/\bar{N}) \sum_{\ell\in\mathcal{S}_i} q_\ell$ with $q_\ell \triangleq \Pr\left(\boldsymbol{u}_\ell = 0 | \boldsymbol{\theta} = \tilde{\theta}\right)$ for $1 \le \ell \le N$, and (9) follows from Jensen's inequality. When $\bar{q}_i < \bar{q}_j$, we obtain that

$$\begin{aligned}
&\min_{t_j>0} \left(\bar{q}_i + (1 - \bar{q}_i)e^{-t_j}\right)\left(\bar{q}_j + (1 - \bar{q}_j)e^{t_j}\right) \\
&= \left(\sqrt{\bar{q}_i \bar{q}_j} + \sqrt{(1 - \bar{q}_i)(1 - \bar{q}_j)}\right)^2, \qquad (10)
\end{aligned}$$

where the minimum is achieved by taking $t_j = (1/2)\log[(1 - \bar{q}_i)\bar{q}_j]/[(1 - \bar{q}_j)\bar{q}_i]$. However, when $\bar{q}_i \ge \bar{q}_j$ for some $j \neq i$, $\min_{t_j>0} \left(\bar{q}_j + (1 - \bar{q}_j)e^{-t_j}\right)\left(\bar{q}_i + (1 - \bar{q}_i)e^{t_j}\right)$ is equal to 1 and therefore substituting (10) into (9) can no longer form an upper bound to β_i. Nevertheless, regardless of the order of $\bar{q}_i$ and $\bar{q}_j$, numerical examinations indicate that at medium-to-high signal-to-noise ratios, the desired ζ_i and its approximation

$$\tilde{\zeta}_i \triangleq \min\left\{ \max\left(\frac{1}{(L-1)} \sum_{j=1}^{L} \tilde{\beta}_j - \tilde{\beta}_i, 0 \right), 1 \right\} \qquad (11)$$

can lead to mostly the same decision in (7),[3] when taking

$$\tilde{\beta}_i \triangleq \sum_{j=1,j\neq i}^{L} \left(\sqrt{\bar{q}_i \bar{q}_j} + \sqrt{(1 - \bar{q}_i)(1 - \bar{q}_j)}\right)^{2\bar{N}}.$$

Therefore, $\tilde{\zeta}_i$ can be used as a replacement of ζ_i to save the computational effort. Note that unlike ζ_i that needs to accumulate the probabilities of $\underline{\hat{\boldsymbol{u}}}$ over an area $\mathcal{D}_i$ of irregular shape, $\tilde{\zeta}_i$ is simply the sum of each individual $\tilde{\beta}_j$, which in turn can be obtained from the average of $q_\ell = \Pr\left(\hat{\boldsymbol{u}}_\ell = 0 | \boldsymbol{\theta} = \tilde{\theta}\right)$. By [6, Eq. (14)], we can further derive

$$\begin{aligned}
q_\ell &= \Pr(\hat{\boldsymbol{u}}_\ell = 0 | \boldsymbol{\theta} = \tilde{\theta}) \\
&= \Pr\left(\boldsymbol{g}_\ell \tilde{a}_\ell + \boldsymbol{n}_\ell < \eta_\ell\right) \approx \Phi\left(\frac{\eta_\ell - \tilde{a}_\ell}{\sigma}\right), \qquad (12)
\end{aligned}$$

where $\tilde{a}_\ell = \sqrt{P_0}(d_0/d_{\mathrm{E}}(\tilde{\theta}, \theta_\ell))$, the approximation in (12) follows from the simplification of $\boldsymbol{g}_\ell \approx 1$, and $\Phi(\cdot)$ is the cumulative distribution function (cdf) of the standard normal distribution. Identification of the most likely ROI via (7) can accordingly be conducted more efficiently.

IV. Experimental Results

In this section, simulations results for the newly proposed soft metric based localization are provided. The setting that we employ is given below. There are $N = 1024$ sensors aligned in a square grid, where the distance between neighboring sensors is equal to 1 meter. Three iterations are performed. At each iteration, the size of the ROI is reduced by a factor of $L = 4$; thus, after three iterations, iterative classification at the FC will report which 4×4 region the target is most likely located in.

According to (1), the sensing signal-to-noise ratio (SNR) for the radiating target is equal to $\gamma_{\text{rad}} \triangleq P_0/\sigma^2$. Likewise, it can be obtained from (2) that the SNR for the Rayleigh faded wireless links between the sensors and the FC is given by $\gamma_{\text{w}} \triangleq$

[3] Approximation (11) can be justified by $\zeta_i = 1 - \beta_i = \frac{1}{(L-1)} \sum_{j=1}^{L} \beta_j - \beta_i$. Since $\frac{1}{(L-1)} \sum_{j=1}^{L} \tilde{\beta}_j - \tilde{\beta}_i$ may be either negative or larger than one when sensing and channel imperfection is serious, a simple clipping operation is adopted to simplify the computations.

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/LWC.2017.2785228

E_b/σ_w^2, which is fixed as 5 dB in our simulations. Without loss of generality, both P_0 and E_b are set to be 1. The local thresholds $\{\eta_i\}_{i=1}^N$ are taken from [6]. Also investigated are the performances of the new soft metric under Byzantine attacks, where α fraction of sensors are randomly picked to send faulty data to the FC by flipping their local binary decisions [7].

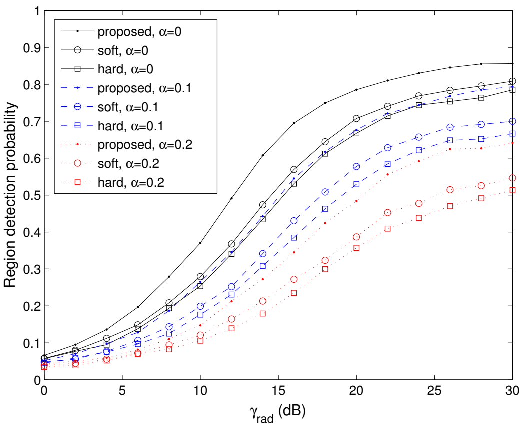
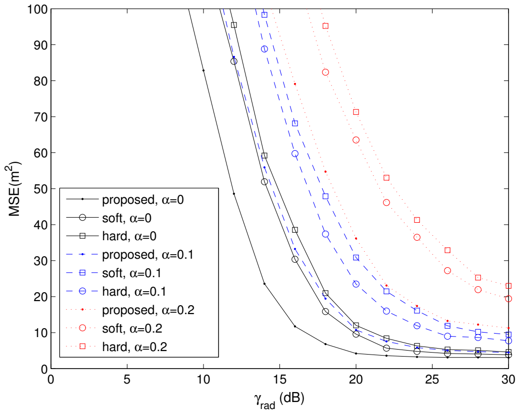
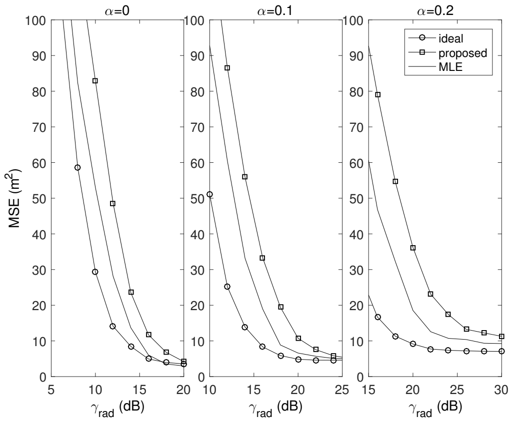


Fig. 1. Region detection probabilities for iterative classification



(a) MSEs under Byzantine attacks with $\boldsymbol{\theta} = \tilde{\theta}$



(b) MSEs for taking $\boldsymbol{\theta} = \tilde{\theta}$ (proposed) and $\boldsymbol{\theta} = \theta$ (ideal)

Fig. 2. MSE performances for the proposed soft metric

Observe from Figs. 1 and 2(a) that the newly proposed soft metric considerably improves the performance of soft-decision (cf. (3)) and hard-decision (cf. (4)) iterative classifications proposed in [1] and [6]. For example, at $\gamma_{\text{rad}} = 20$ dB, the newly proposed soft metric has nearly 10% improvement in region detection probability (RDP) in comparison with the soft-decision fusion rule under $\alpha = 0$, 10% and 20%. More improvement can be observed when the newly proposed soft metric based and the hard-decision based localization rules are compared. Additionally, by using the center of the final 4×4 region as the final estimated target location, the newly proposed soft metric can achieve 5 and 36 (m^2) mean square error (MSE) improvement when it is compared with the soft-decision fusion rule under no Byzantine attack and with $\alpha = 20\%$ Byzantine sensors, respectively.

Lastly, we demonstrate the impact of $\tilde{\theta}$ (equiv. $\tilde{a}_\ell$ in (12)) on localization accuracy in Fig. 2(b), along with the MSE of a maximum-likelihood estimator (MLE) [1, Eq. (4)] as a benchmark. Our results suggest that under $\alpha = 0$, the simple initial estimate in Section III-A performs close to the computationally intensive MLE when sensing SNR≥ 20.

V. Conclusion

In this paper, a new soft metric that incorporates the important information of exact sensor positions was proposed. With a manageable incremental complexity, a much better performance than that of the legacy soft- and hard-decision rules that do not employ the sensor location knowledge can be obtained, particularly when a certain number of sensors suffer Byzantine attacks. Our results show the significant impact of the sensor location knowledge on the performance of target localization even when a certain number of unreliable local measurements exist.

References

[1] A. Vempaty, Y. S. Han, and P. K. Varshney, "Target localization in wireless sensor networks using error correcting codes," *IEEE Trans. Inf. Theory*, vol. 60, no. 1, pp. 697–712, Jan. 2014.

[2] E. Xu, Z. Ding, and S. Dasgupta, "Wireless source localization based on time of arrival measurement," in *IEEE Int. Conf. Acoustics, Speech, Signal Proc. (ICASSP)*, Mar. 2010, pp. 2842–2845.

[3] K. Yang, J. An, X. Bu, and G. Sun, "Constrained total least-squares location algorithm using time-difference-of-arrival measurements," *IEEE Trans. Veh. Technol.*, vol. 59, no. 3, pp. 1558–1562, Mar. 2010.

[4] D. Li and Y. H. Hu, "Energy-based collaborative source localization using acoustic microsensor array," *EURASIP J. App. Signal Proc.*, no. 4, pp. 321–337, 2003.

[5] X. Sheng and Y. H. Hu, "Maximum likelihood multiple-source localization using acoustic energy measurements with wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 53, no. 1, pp. 44–53, Jan. 2005.

[6] C. Y. Wei, P. N. Chen, Y. Han, and P. Varshney, "Local threshold design for target localization using error correcting codes in wireless sensor networks in the presence of byzantine attacks," *IEEE Trans. Inf. Forens. & Secur.*, vol. 12, no. 7, pp. 1571–1584, Jul. 2017.

[7] A. Vempaty, L. Tong, and P. K. Varshney, "Distributed inference with byzantine data: State-of-the-art review on data falsification attacks," *IEEE Signal Proc. Mag.*, vol. 30, no. 5, pp. 65–75, Sept. 2013.

[8] T. S. Rappaport, *Wireless Communications: Principles and Practice*, 2nd ed. Prentice Hall, 2002.

[9] O. Ozdemir, R. Niu, and P. K. Varshney, "Channel aware target localization in wireless sensor networks," in *Int. Conf. Info. Fusion*, Quebec City, Canada, Jul. 2007.

[10] I. Bahceci, G. Al-Regib, and Y. Altunbasak, "Parallel distributed detection for wireless sensor networks: performance analysis and design," in *IEEE Global Commun. Conf.*, vol. 4, Dec. 2005.

[11] R. Jiang and B. Chen, "Fusion of censored decisions in wireless sensor networks," *IEEE Trans. Wireless Comm.*, vol. 4, no. 6, pp. 2668–2673, Nov. 2005.