Efficient Node Localization in Energy-Harvesting Wireless Sensor Networks

Ahmad El Assaf*, Slim Zaidi*, Sofiène Affes*†, and Nahi Kandil*†
*INRS-EMT, Montreal, QC, Canada, Email: {elassaf,zaidi,affes}@emt.inrs.ca
†LRTCS-UQAT, Rouyn-Noranda, QC, Canada, Email: nahi.kandil@uqat.ca

***Abstract*—In this paper, we propose a novel localization algorithm tailored for energy-harvesting wireless sensor networks EH-WSNs where both the power budgets available at the EH sensors and hence their transmission capabilities are random. This characteristic, if not taken into account when designing the localization algorithm, may severely hinder its accuracy. Assuming a random transmission capability at each EH sensor, we develop a new approach to derive the expected hop progress (EHP). Exploiting the latter, we propose a localization algorithm that is able to accurately locate the EH sensors owing to a new implementation requiring no additional power consumption. Furthermore, we develop a correction mechanism which complies with the heterogeneous coverage nature of EH sensors to further improve localization accuracy without incurring any additional power cost. Simulation results show that the proposed algorithm, whether applied with or without correction, outperforms in accuracy the most representative WSN localization algorithms in EH powering contexts.**

***Index Terms*—Energy harvesting, wireless sensor networks, multi-hop, localization algorithm, low-cost localization.**

I. Introduction

Energy harvesting (EH) recently emerged as a promising solution aiming to power EH sensors using renewable energy sources, thereby securing their autonomy as well as the perpetuity of their target tasks without any human intervention. EH sensors are then able to harvest energy from for instance solar energy, vibration energy, thermoelectric energy, RF energy, etc. [1], convert it into an electrical energy and store it for future use. Since the harvested energy is limited, EH sensors usually transmit their gathered data to an access point (AP) in a multi-hop fashion. The latter allows in fact important energy saving especially if the AP is located in the far-field. However, the received data at the AP are often fully or partially meaningless if the location from where they have been measured is unknown, making the EH sensors' localization an essential task in multi-hop EH-WSNs (as is the case of any conventional multi-hop WSN). Owing to the low power consumption requirements of EH sensors and the randomness of their available harvested energy, unconventional paradigms for localization algorithm must yet be investigated. Many interesting solutions, originally developed for conventional WSN assuming very often homogeneous transmission coverage capabilities, exist in the literature. To properly localize each regular or position-unaware EH sensor, most of these algorithms require the distance between the latter and at least three position-aware EH sensors called hereafter anchors. Since it is very likely in multi-hop EH-WSNs that some regular sensors be unable to directly communicate with all anchors, the distance between each anchor-regular sensor pair is usually estimated using their shortest path. The latter distance is then obtained by summing the distances between any consecutive intermediate sensors located on the shortest path between the two sensors. Depending on the process used to estimate the latter distances, localization algorithms may fall into three categories: measurement-based, heuristic, and analytical [2]-[6].

Measurement-based algorithms exploit the measurements of the received signals' characteristics such as the received signal strength (RSS) or the time of arrival (TOA), etc. [2]. Using the RSS measurement, the distance between any sensors' pair could be obtained by converting the power loss due to propagation from a sensor to another based on some propagation laws. To obtain the power loss, a prior knowledge of the transmit power is, however, needed at the receiving sensor. Since this quantity is random due to the EH nature of sensors, it is quite impossible to guess its value beforehand and, hence, the distance cannot be estimated. Furthermore, even if the transmit power is known, the distance's estimate using the RSS would be far from being accurate due to the probable presence of noise and interference, thereby leading to unreliable localization accuracy. Using the TOA measurement, the EH sensors require high-resolution clocks and extremely accurate synchronization between them. While the first requirement may dramatically increase the cost and the size of the EH sensors, the second results in severe depletion of their power due to the additional overhead required by such a process. Additionally, in the presence of noise and/or multipath, the TOA measurement is severely affected thereby hindering EH sensors' localization accuracy. As far as heuristic algorithms are concerned, they also have a major drawback. Indeed, most of these algorithms are based on DV-HOP whose implementation in EH-WSNs requires a correction factor derived in a non-localized manner and broadcasted in the network by each anchor [3]. This causes an undesired prohibitive overhead and, hence, unnecessary high power consumption at each EH sensor.

Popular alternatives, more suitable for EH-WSNs, are the analytical algorithms [4]-[6] which evaluate theoretically the distance between any two consecutive intermediate sensors. The latter is in fact locally computable at each sensor, thereby avoiding unnecessary power consumption. In spite of their

Work supported by the CRD, DG, and CREATE PERSWADE <www.create-perswade.ca> Programs of NSERC and a Discovery Accelerator Supplement Award from NSERC.

valuable contributions, the approaches developed so far in [4]-[6] to derive that distance do not account for the randomness of the available harvested power at each sensor. Indeed, this phenomenon results in the randomization of the sensors' transmission capabilities, since the latters are closely related to the sensors' available powers. The transmission capabilities are then not only different from a sensor to another, but also unpredictable before EH-WSNs' deployment. As the approaches in [4]-[6] assume the same transmission capability throughout the network, their localization accuracy substantially deteriorates in EH-WSNs making them unsuitable for such networks. To the best of our knowledge, there is no analytical algorithm that accounts for the heterogeneous transmission coverage nature of EH sensors.

To bridge this gap, we propose in this paper a novel analytical algorithm tailored for EH-WSNs where sensors have different and random transmission capabilities. Taking into account the EH nature of sensors, a new approach is developed to accurately derive the distances between any consecutive sensors. Using the so-obtained distances, the proposed algorithm is able to accurately locate the EH sensors owing to a new implementation avoiding any additional power consumption. Furthermore, we develop a correction mechanism which complies with the heterogeneous coverage nature of EH sensors to further improve localization accuracy without incurring any additional power cost. Simulations results show that the proposed algorithm, whether applied with or without correction, outperforms in accuracy the most representative WSN localization algorithms in EH powering contexts.

The rest of this paper is organized as follows: Section II describes the system model and discusses the motivation of this work. Section III derives the distance between consecutive sensors. A novel localization algorithm for EH-WSNs is proposed in section IV. Simulation results are discussed in Section VI and concluding remarks are made in section VII.

II. Network Model and Overview

The network model of our concern consists on N EH sensors deployed in a 2-D square area S. The transmission coverage of each sensor is assumed to be circular, i.e., the i-th sensor could directly communicate with any sensor located in $D\left(i, Tc_i\right)$, the disc having this sensor as a center and its transmission capability Tc_i as a radius. In a multi-hop transmission, note that the i-th sensor could also communicate with any sensor located outside its coverage area $D\left(i, Tc_i\right)$. Due to the EH nature of the sensors, their transmission capabilities $Tc_i,\ i = 1, \ldots, N$ are assumed to be independent but not necessary identically distributed random variables. It is also assumed that only a few sensors commonly known as anchors are aware of their positions. The other sensors, called hereafter position-unaware or regular sensors for the sake of simplicity, are oblivious to this information. Let N_a and $N_u = N - N_a$ denote the number of anchors and regular sensors, respectively. Without loss of generality, let $\left(x_i, y_i\right),\ i = 1, \ldots, N_a$ be the coordinates of the anchors and $\left(x_i, y_i\right),\ i = N_a + 1, \ldots, N$ those of the regular ones.

As a first step of any localization algorithm for multi-hop EH-WSNs aiming to estimate the regular sensors positions, the k-th anchor broadcasts through the network a message containing its position. If the $(i - N_a)$-th regular sensor (or the i-th sensor) is located outside the anchor coverage area, it receives this message through multi-hop transmission. For simplicity, let us assume that only one intermediate sensor j located over the shortest path between the k-th anchor and the i-th sensor is necessary (i.e., two-hop transmission). Assuming a high sensor density in the network, the distance d_{k-i} between the two sensors can be accurately approximated as [3]-[6]

$$d_{k-i} \simeq d_{k-j} + d_{j-i}, \tag{1}$$

where $d_{\star-*}$ is the effective distance between the $\star$-th and the $*$-th sensor. Two approaches have been so far developed to analytically estimate the distance d_{k-i} exploiting the aforementioned approximation [4]-[6]. In the first approach, the j-th sensor estimates the distance d_{k-j} using the number of common neighbors n_{kj} with the k-th sensor [5]. However, n_{kj} cannot be accurately obtained since some neighbors of the j-th and/or k-th sensor/s are not able to communicate with the latters due to their weaker capabilities and, therefore, this approach does not comply with the EH nature of sensors.

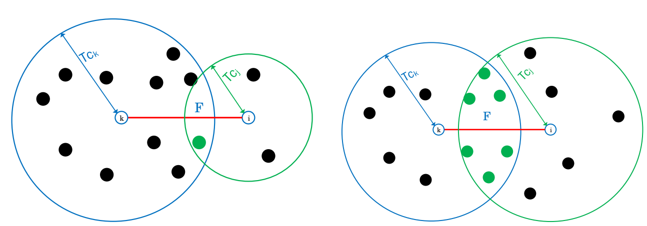
The second approach uses the fact that the minimum square error (MMSE) of the distance estimation is obtained if $\hat{d} = E\left(d\right)$ and, hence,

$$\hat{d}_{k-i} \simeq \bar{d}_{k-j} + \bar{d}_{j-i}, \tag{2}$$

where $\bar{d}_{k-j} = \mathrm{E}\left\{d_{k-i}\right\}$ is the expected hop progress (EHP) and $\bar{d}_{j-i} = \mathrm{E}\left\{d_{j-i}\right\}$ is the mean last hop (MLH). One of the well-known analytical expressions of EHP is the one developed in [4] as follows:

$$\bar{d}_{kj} = \sqrt{3}\lambda \int_0^{Tc_k} x^2 e^{-\lambda \frac{\pi}{3}\left(Tc_k^2 - x^2\right)} dx, \tag{3}$$

where λ is the sensor density, and x is the distance between the k-th and the i-th node. From (3), the EHP a priori depends only on the k-th sensor transmission capability Tc_k and, therefore, its computation does not supposedly require any knowledge of the j-th sensor transmission capability Tc_j. In what follows, and in contrast to (3), we will prove the EHP expression to be dependent on both Tc_k and Tc_j thereby revealing the expression derived in [4], as one example among too many others whose approaches are similar to the above but not discussed here for lack of space, to lack accuracy.



(a) Small Tc_j (b) Large Tc_j

Fig. 1. Effect of the intermediate sensor transmission capability.

Let F be the potential forwarding area wherein the intermediate sensor j could be located. Since this sensor should, at the same time, be located in the k-th sensor coverage area and communicate directly with the i-th sensor using its transmission capability Tc_j, F is given by

$$F = D(k, Tc_k) \cap D(i, Tc_j). \tag{4}$$

It is noteworthy that the EHP is nothing but the mean of all distances between the k-th sensor and all the potential intermediate sensors located in F and, hence, the EHP strongly depends on F. As can be observed from Fig. 3, if the intermediate sensor transmission capability Tc_j increases, the potential forwarding area F increases to include potential intermediate sensors closer to the k-th anchor, thereby decreasing the EHP. Likewise, if Tc_j decreases, F decreases to exclude potential intermediate sensors closer to the k-th anchor and, hence, the EHP increases. Consequently, the EHP depends not only on Tc_k, but also on Tc_j. Let us now turn our attention to the MLH. It is obvious that the transmission capability of the i-th sensor does not have any effects on the last hop size d_{j-i}. Therefore, in contrast with the EHP, the MLH depends only on the transmission capability of the transmitting sensor j. In the next section, novel approaches are developed to accurately derive the expressions of both the MLH and the EHP. These results will be exploited in Section IV to propose a localization algorithm that complies with the EH nature of sensors.

III. Analytical Evaluation of the MLH and EHP

In this section, the expression of the MLH as well as the EHP are accurately derived. To this end, we consider the same scenario described in Section II. For the sake of clarity, in what follows, we denote by X, Y, and Z the random variables that represent d_{k-i}, d_{j-i}, and d_{k-j}, respectively.

A. MLH derivation

Since the i-th regular sensor could be located anywhere in $D(j, Tc_j)$ (the j-th sensor coverage area) with the same probability, Y can be considered as a uniformly distributed random variable on $[0, Tc_j]$. Therefore, the MLH denoted hereafter by $h_{\text{last}}(Tc_j)$ is given by

$$h_{\text{last}}(Tc_j) = \int_0^{Tc_j} y f_Y(y) dy = \int_0^{Tc_j} \frac{y}{Tc_j} dy = \frac{Tc_j}{2}, \tag{5}$$

where $f_Y(y) = 1/Tc_j$ is the probability density function (pdf) of Y.

B. EHP derivation

In order to derive the EHP, one should first compute the conditional cumulative distribution function (CDF) $F_{Z|X}(z) = P(Z \leq z|x)$ of Z with respect to the random variable X. As can be shown from Fig. 2, $Z \leq z$ is guaranteed only if there are no sensors in the dashed area A. Therefore, the conditional CDF $F_{Z|X}(z)$ can be defined as

$$F_{Z|X}(z) = P(Z \leq z|x) = P(\mathrm{A}_0|F_1), \tag{6}$$

where $P(A_0|F_1)$ is the probability that the event $A_0 = \{$no sensors in the dashed area $A\}$ given $F_1 = \{$at least one sensor in the potential forwarding area $\}$ occurs. Since the sensors are uniformly deployed in S, the probability of having K sensors in F follows a Binomial distribution $\mathrm{Bin}(N, p)$ where $p = \frac{F}{S}$. For relatively large N and small p, it can be readily shown that $\mathrm{Bin}(N, p)$ can be accurately approximated by a Poisson distribution $\mathrm{Pois}(\lambda F)$ where $\lambda = N/S$ is the average sensor density in the network. Using the Bayes' theorem, $F_{Z|X}(z)$ could be rewritten as

$$F_{Z|X}(z) = \frac{P(\mathrm{F}_1|A_0)\,P(\mathrm{A}_0)}{P(F_1)}, \tag{7}$$

and, hence, for a large number of sensors N and small p, we have

$$F_{Z|X}(z) = \frac{e^{-\lambda A}\left(1 - e^{-\lambda B}\right)}{\left(1 - e^{-\lambda F}\right)}, \tag{8}$$

where $B = F - A$. In the equation above, note that we use the fact that $P(\mathrm{F}_1|A_0)$ is the probability that at least one sensor is in B. As can be observed from (8), when $z = a$, we have $B = 0$ and $A = F$ and, therefore, $F_{Z|X}(z) = 0$. This is expected since all potential intermediate sensors are located in the in the forwarding zone F where any sensor is at least at distance a from the k-th sensor (i.e. $P(Z \leq a) = 0$). Furthermore, if $z = Tc_k$, it holds that $B = F$ and $A = 0$ and, hence, $P(Z \leq Tc_k) = 1$. This is also expected since all potential intermediate sensors are located in the k-th sensor coverage area where any sensor is at most at distance Tc_k form the latter sensor (i.e. $P(Z \leq Tc_k) = 1$). It should noticed here that the latter properties are not satisfied by any previously developed expression CDF such as those in [4], [6]. Using some geometrical properties and trigonometric

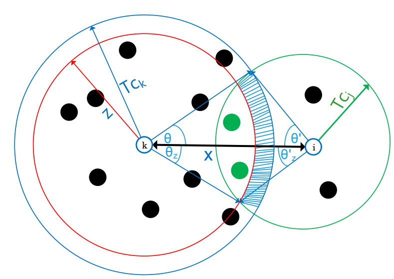


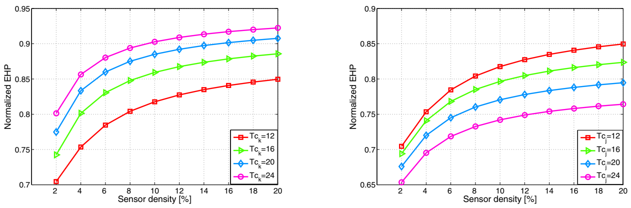
Fig. 2. Illustration of EHP analysis approach.

transformations, one can easily derive F and B as shown in [7]. Finally, the EHP $h(Tc_k, Tc_j)$ between the k-th and j-th sensors can be derived as

$$h(Tc_k, Tc_j) = \int_{Tc_k}^{Tc_k + Tc_j} \left(\alpha H_{Z|X}(\alpha) + \int_{\alpha}^{Tc_k} H_{Z|X}(z) dz \right) f_X(x)\, dx, \tag{9}$$

where $H(\star) = 1 - F_{Z|X}(\star)$ and $\alpha = x - Tc_j$ and $f_X(x)$ is the pdf of X. Note that the latter can be considered as a uniform random variable over $[Tc_k, Tc_k + Tc_j]$ and, hence, $f_X(x)$ can be substituted in the latter result by $1/Tc_j$. To the best of our knowledge, a closed-form expression for the

EHP in (9) does not exist. However, $h(Tc_k, Tc_j)$ can be easily implemented since it depends on finite integrals. As can be observed from (9), the proposed EHP depends on both Tc_k and Tc_j, in contrast to those such as in (3) proposed in all pervious works. It can be shown from Figs. 3(a) and 3(b) that



(a) Effect of Tc_k (b) Effect of Tc_j

Fig. 3. Effect of the transmission capabilities on the EHP.

the so-obtained EHP using decreases if Tc_j increases while it increases when Tc_k increases. This collaborates the discussion made above. Those figures also show that the proposed EHP increases with the sensor density. This is expected since it is very likely that the per-hop distance increases when the number of sensors located in F increases if, of course, both Tc_k and Tc_j are fixed.

IV. Proposed localization for EH-WSNs

In this section, based on the so-obtained EHP and MLH expressions, we propose a novel three-step localization algorithm for EH-WSNs.

A. Step 1: initialization

In this step, the k-th anchor starts by broadcasting through the network a packet which consists of a header followed by a data payload. The packet header contains the anchor position (x_k, y_k), while the data payload contains $(Tc_k, \hat{d}_k)$, where Tc_k is the transmission capability of the k-th anchor and $\hat{d}_k$ is the estimated distance initialized to zero. If the packet is successfully received by a sensor, the latter estimates the EHP using the above approach, adds it to $\hat{d}_k$, stores the resulting value in its database and then, rebroadcasts the resulting packet after substituting Tc_k by its own transmission capability. Once this packet is received by another sensor, its database information is checked. If the k-th anchor information exists and the stored estimated distance is larger than that of the received one, the sensor updates the k-th anchor's information, then broadcasts the resulting packet after substituting the received transmission capability by its own. Otherwise, the sensor discards the received packet. However, when the sensor is oblivious to the k-th anchor position, it adds this information to its database and forwards the received packet after substituting the received transmission capability by its own. This mechanism will continue until each regular sensor in the network becomes aware of each anchor position as well as the distance from the latter to the last intermediate sensor before reaching that sensor. Note that the implementation of the proposed algorithm requires that each sensor broadcasts the anchor information not only with its estimated distance but also its transmission capability to allow the EHP computation at the next receiving sensor. In contrast, the implementation of existing algorithms in EH-WSNs requires the broadcast of the anchor information and the estimated distance only. Yet we will prove next that the additional power consumption that could be incurred a priori when broadcasting the transmission capabilities can be easily avoided by the proposed algorithm.

B. Step 2: positions' computation

In this section, we will show how the so-received information can be exploited to get an initial guess of each regular sensor position. Using its available information, the $(i - N_a)$-th regular sensor (or the i-th sensor) computes an estimate of its distance to the k-th anchor as

$$\hat{d}_{k-i} = \hat{d}_k + h_{\text{last}}(Tc_{k+L}), \tag{10}$$

where $\hat{d}_k = \sum_{l=k}^{k+L-1} h\left(Tc_l, Tc_{l+1}\right)$ is the distance from the k-th anchor to the last intermediate sensor. In (10), we assume for simplicity, yet without loss of generality, that L intermediate sensors exist over the shortest path between the k-th anchor and the $(i - N_a)$-th regular sensor and that the l-th intermediate sensor is the $(k + l)$-th sensor. Using its estimated distances to the N_a anchors as well as the latters' coordinates, the position of the i-th sensor could be deduced by solving the following nonlinear equations system: $(x_k - \hat{x}_i)^2 + (y_k - \hat{y}_i)^2 = \hat{d}_{ik}^2$ for $k = 1, \ldots, N_a$. After some rearrangements that linearize the system above, we obtain

$$\mathbf{A}\hat{\mathbf{x}} = -\frac{1}{2}\mathbf{b}, \tag{11}$$

where $\hat{\mathbf{x}} = [\hat{x}_i, \hat{y}_i]$, $\mathbf{A}$ is a $(Na - 1) \times 2$ matrix with $[\mathbf{A}]_{k1} = x_k - x_{N_a}$ and $[\mathbf{A}]_{k2} = y_k - y_{N_a}$, and $\mathbf{b}$ is a $(Na - 1) \times 1$ vector with $[\mathbf{b}]_k = d_{ik}^2 - d_{iN_a}^2 - x_k^2 + x_{N_a}^2 - y_k^2 + y_{N_a}^2$. Since $\mathbf{A}$ is a non-invertible matrix, $\hat{\mathbf{X}}$ could be estimated with the pseudo-inverse of $\mathbf{A}$ as $\hat{\mathbf{x}} = -\frac{1}{2}\mathbf{A}^T\left(\mathbf{A}\mathbf{A}^T\right)^{-1}\mathbf{b}$. Therefore, the i-th regular sensor is able to obtain an initial guess of its coordinates as $\hat{x}_i = [\hat{\mathbf{x}}]_1$, and $\hat{y}_i = [\hat{\mathbf{x}}]_2$. It is also noteworthy from the definition of $\mathbf{A}$ and $\mathbf{b}$ that $\hat{x}_i$ and $\hat{y}_i$ are solely dependant on the information locally available at each sensor and, hence, their computation does not require any additional overhead (i.e., additional power cost). Unfortunately, errors are expected to occur when estimating the distance between each regular sensor-anchor pair, thereby hindering localization accuracy. As a third step of our proposed algorithm, we propose a correction mechanism aiming to reduce this error.

C. Step 3: correction mechanism

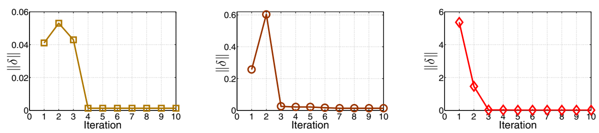
Let ϵ_{ik} denote the estimation error of the distance between the k-th anchor and the i-th regular sensor as

$$\epsilon_{ik} = \hat{d}_{k-i} - d_{k-i}, \tag{12}$$

where d_{k-i} is the true distance between the two sensors. As discussed above, this error hinders localization accuracy. As such, we have $x_i = \hat{x}_i + \delta_{x_i}$ and $y_i = \hat{y}_i + \delta_{y_i}$ where δ_{x_i} and δ_{y_i} are the location coordinates' errors to be determined. Retaining the first two terms of the Taylor series expansion of d_{k-i} and rewriting the result in a matrix form yields

$$\boldsymbol{\Gamma}_i\boldsymbol{\delta}_i = \boldsymbol{\zeta}_i - \boldsymbol{\epsilon}_i, \tag{13}$$

where $\mathbf{\Gamma}$ is a $Na \times 2$ matrix with $[\mathbf{\Gamma}]_{k1} = \frac{\hat{x}_i - x_k}{d^{\dagger}_{k-i}}$ and $[\mathbf{\Gamma}]_{k2} = \frac{\hat{y}_i - y_k}{d^{\dagger}_{k-i}}$, $\boldsymbol{\epsilon}_i = [\epsilon_{i1}, \epsilon_{i2}, \ldots, \epsilon_{iN_a}]^T$, $\boldsymbol{\zeta}_i = \left(\hat{d}_{i-1} - d^{\dagger}_{i-1}, \ldots \hat{d}_{i-N_a} - d^{\dagger}_{i-N_a}\right)^T$, $\boldsymbol{\delta}_i = [\delta_{x_i}, \delta_{y_i}]^T$, and $d^{\dagger}_{k-i} = \sqrt{(\hat{x}_i - x_k)^2 - (\hat{y}_i - y_k)^2}$. Many methods such as



(a) (b) (c)

Fig. 4. Convergence of $\|\boldsymbol{\delta}\|$ vs. the number of iterations.

the weighted least squares (WLS) might be used to properly derive $\boldsymbol{\delta}_i$. Using WLS, the solution of (13) is given by :

$$\boldsymbol{\delta}_i = \left(\mathbf{\Gamma}_i^T \mathbf{P}_i^{-1} \mathbf{\Gamma}_i\right)^{-1} \mathbf{\Gamma}_i^T \mathbf{P}_i^{-1} \boldsymbol{\zeta}_i, \tag{14}$$

where $\mathbf{P}_i$ is the covariance matrix of $\boldsymbol{\epsilon}_i$. Since ϵ_{ik} $k = 1, \ldots, N_a$ are independent random variables, $\mathbf{P}_i$ boils down to $\mathrm{diag}\left\{\sigma^2_{i1}, \ldots, \sigma^2_{iN_a}\right\}$ where σ^2_{ik} is the variance of ϵ_{ik}. A straightforward inspection of (14) reveals that $\boldsymbol{\delta}_i$ depends on all $\sigma^2_{ik}, k = 1, \ldots, N_a$, the entries of $\left(\mathbf{\Gamma}_i^T \mathbf{P}_i \mathbf{\Gamma}_i\right)^{-1}$ as well as some locally available information. Since $\mathbf{\Gamma}_i^T \mathbf{P}_i \mathbf{\Gamma}_i$ is a 2-by-2 matrix, the entries of its inverse can be analytically and easily derived using its locally available entries. Furthermore, it will be shown in Section V that the evaluation of $\sigma^2_{ik}, k = 1, \ldots, N_a$ requires a negligible additional power cost that could be easily avoided.

This proves that $\boldsymbol{\delta}_i$'s computation does not burden neither the implementation complexity of the proposed algorithm nor the overall cost of the EH-WSN. Once we get $\boldsymbol{\delta}_i$, the value of $(\hat{x}_i, \hat{y}_i)$ is updated as $\hat{x}_i$=$\hat{x}_i + \delta_{x_i}$ and $\hat{y}_i = \hat{y}_i + \delta_{y_i}$. The computations are repeated until $\|\boldsymbol{\delta}_i\|$ approaches zero. In such a case, we have $x_i \simeq \hat{x}_i$ and $y_i \simeq \hat{y}_i$ and, hence, more accurate localization is performed. As can be observed from Fig. 4, the latter iterations should be repeated at most 5 times to ensure the convergence of the proposed correction mechanism. Knowing that the required power to execute one instruction is in the range of 10^{-4} of the power consumed per transmitted bit [2], the power needed to execute the above mechanism is then very negligible with respect to the overall power consumed by each sensor. Consequently, the correction mechanism complies with EH-WSNs where the power is considered as a scarce resource. Note that the above discussion also holds for the position computation step and, hence, the proposed localization algorithm is perfectly tailored for EH-WSNs.

V. Variance evaluation

This section aims to derive the expression of the variances σ^2_{ik}, $k = 1, \ldots, N_a$ which are required for the proposed algorithm's implementation. As such, two different methods, analytical and non-parametric , are proposed.

A. Analytical method

Assuming a high sensor density in the network, the distance d_{k-i} between two sensors can be rewritten as

$$d_{ik} \simeq \sum_{l=k}^{k+L} d_{l,l+1}, \tag{15}$$

where L is the number of intermediate sensors over the shortest path and $d_{l,l+1}$ is the distance between the l-th and $(l+1)$-th intermediate sensor. It follows from (10) and (15) that $\varepsilon_{ki} \simeq \sum_{l=k}^{k+L-1} e_l + e_{\text{last}}$ with $e_l = h\left(Tc_l, Tc_{l+1}\right) - d_{l-(l+1)}$ is the distance estimation error incurred during the $(l-k+1)$-th hop and $e_{\text{last}} = h_{\text{last}}(Tc_{k+L}) - d_{(k+L)-i}$ is the error incurred at the last hop. It can be readily shown that $\sigma^2_{ik} = \sum_{l=k}^{k+L-1} \sigma^2_l + {\sigma_{\text{last}}}^2$ where σ^2_l and ${\sigma_{\text{last}}}^2$ are the variances of e_l and e_{last}, respectively. Using the results developed in Section III, we obtain ${\sigma_{\text{last}}}^2 = \frac{Tc^2_{k+L}}{12}$ and

$$\begin{aligned}{\sigma_l}^2 = &\int_{Tc_l}^{Tc_l+Tc_{l+1}} \left(\alpha^2 H(\alpha) + 2\int_{\alpha}^{Tc_l} zH(z)dz\right) f_X(x)dx - \\ &\left(\int_{Tc_l}^{Tc_l+Tc_{l+1}} \left(\alpha H(\alpha) + \int_{\alpha}^{Tc_l} H(z)dz\right) f_X(x)dx\right)^2.\end{aligned} \tag{16}$$

Note that ${\sigma_l}^2$ is obtained using any the CDF developed in Section III-B. As can be observed from the latter results, ${\sigma_{\text{last}}}^2$ is locally computable by the i-th sensor while ${\sigma_l}^2$ should be computed at the $(l+1)$-th intermediate sensor, added to the term $\sum_{m=k}^{l-1} \sigma^2_m$, then forwarded to the next intermediate sensor. This results in an additional few bits that must be transmitted by each sensor in the network and, hence in an additional power cost.

Let p_l be the available power at the l-th intermediate sensor, b_l be the length in bits of the original packet (i.e, when ${\sigma_l}^2$ is not included), and a_l be the cost in bits if ${\sigma_l}^2$ is added to the packet. If the power p_l allows the l-th intermediate sensor to transmit b_l bits over a Tc_l coverage distance, this power will also allow the latter sensor to transmit $b_l + a_l$ bits but over a coverage distance $\bar{Tc_l} < Tc_l$, where $\bar{Tc_l}$ is the new transmission capability of the l-th intermediate sensor. Since no matter are the transmission capabilities of the l-th and $(l+1)$-th intermediate sensors, the latter is always able to compute the $(l+k-1)$-th EHP, the fact that Tc_l decreases to $\bar{Tc_l}$ does not affect the performance of the proposed localization algorithm. Therefore, σ^2_{ik}, $k = 1, \ldots, N_a$ can be obtained at the i-th sensor without any additional power cost and, further, without any performance loss. This further proves that the proposed algorithm complies with the nature of EH sensors. Such extremely important and crucial features to EH-WSN localization no longer hold true for the previously proposed localization algorithms [4], [6] where only the effect of the sender sensor's transmission capability is taken into account when deriving the EHP. Note that the above discussion also holds for the additional bits (with respect to former algorithms) incurred when broadcasting the transmission capabilities.

B. Non-parametric method

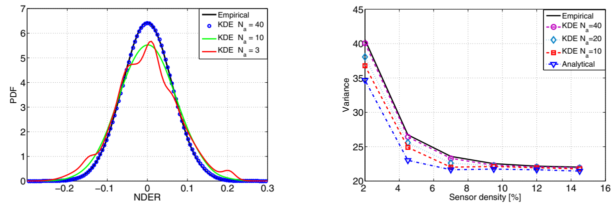
In the previous section, the analytical expression of σ_{ik}^2 was derived using the approximation in (15) which holds only for highly dense networks. However, if the latter assumption is not satisfied (i.e., lowly dense network), (15) would no longer be valid and, hence, σ_{ik}^2's expression would no longer be accurate enough. In such a case, to properly derive σ_{ik}^2, we propose to exploit the PDF of the distance estimation error ε_{ki} denoted by $f(\varepsilon)$. Unfortunately, to the best of our knowledge, there is no closed form solution for such a PDF. In this work, we propose to use a non-parametric technique to estimate it owing to some potential observations available at anchors. So far, many non-parametric techniques have been proposed in the literature such as the histogram and the well-known kernel density estimation (KDE) techniques [8]. In this paper, we are only concerned by the latter which can estimate an arbitrary distribution without much observations. Such observations can in fact be easily obtained at the k-th anchor. Indeed, since this anchor is aware of all other anchor positions, it is able to derive the actual distances between it and the latters. Using (10), the k-th anchor could also obtain the estimated distances between it and the other anchors and, therefore, derive ε_{ki}. Hence, if N_a anchors exist in the network, the total number of available observations is $n_o = N_a(N_a - 1)$. Let $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{n_o}$ denote such observations. Using the KDE technique, $f(\varepsilon)$ can be then approximated by

$$\hat{f}(\varepsilon) = \frac{1}{p s_\varepsilon} \sum_{i=1}^{n_o} K\left(\frac{\varepsilon - \varepsilon_i}{s_\varepsilon}\right), \tag{17}$$

where s_ε is a smoothing parameter determined using the method in [9] and $K(\varepsilon) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}\varepsilon^2)$ is the Gaussian kernel. As can be noticed from (17), the estimated PDF is computed by averaging the Gaussian density over all observations. Substituting $K(\varepsilon)$ in (17) and using the resulting PDF to compute σ_{ik}^2 yields

$$\sigma_{ik}^2 = \frac{\sum_{i=1}^{n_o}\left(X_i G_i - Y_i^2\right)}{\sum_{i=1}^{n_o} G_i^2}, \tag{18}$$

where $X_i = \left(s_\varepsilon^2 + \varepsilon_i^2\right) G_i - s_\varepsilon^2 \left((\varepsilon_i + 1)\, e^{-\frac{(1-\varepsilon_i)^2}{2 s_\varepsilon^2}} + (\varepsilon_i - 1)\, e^{-\frac{(1+\varepsilon_i)^2}{2 s_\varepsilon^2}} \right)$, $G_i = s_\varepsilon \sqrt{2\pi}\left(Q\left(\frac{\varepsilon_i - 1}{s_\varepsilon}\right) - Q\left(\frac{\varepsilon_i + 1}{s_\varepsilon}\right)\right)$, and $Y_i = \varepsilon_i G_i - s_\varepsilon^2 \left(e^{-\frac{(1-\varepsilon_i)^2}{2 s_\varepsilon^2}} - e^{-\frac{(1+\varepsilon_i)^2}{2 s_\varepsilon^2}} \right)$ with $Q(x)$ being the Q-function. Fig. 5(a) plots the empirical $f(\varepsilon)$ as well as $\hat{f}(\varepsilon)$ for different numbers of anchors. We see there that only a few anchors (i.e., few observations) are required to accurately estimate the localization errors' PDF. Furthermore, from Fig. 5(a), the estimated PDF approaches the empirical one, as N_a increases. This gives a sanity check for the proposed nonparametric method. Fig. 5(b) plots the error variance for different sensor densities. It shows, as expected, that the variance decreases when the sensor density increases. Beyond a sensor density threshold of less than 10%, both the analytical and the non-parametric methods start to yield about the same variance as the empirical one. Furthermore, when N_a increases, more so at large enough values, the efficiency of the non-parametric method increases even at low sensor densities. Note that increasing the number of anchors N_a does not only result in a more accurate variance, but also in a more reliable localization [3].



(a) NDER's PDF using KDE. (b) Global NDER's variance.

Fig. 5. Normalized distance estimation error (NDER).

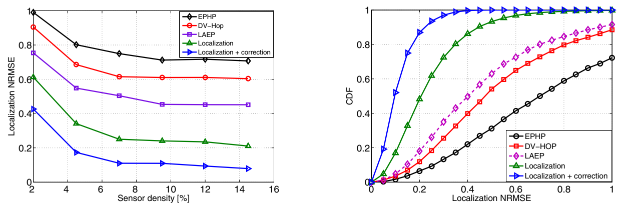
Nevertheless, in order to derive σ_{ik}^2 using this approach, each regular sensor needs to be aware of all observations. If this is not properly done, it will be very expensive in terms of power consumption, since each anchor would recur to a second broadcast to share its observations with the regular sensors. In order to circumvent this problem, we propose in what follows a power-efficient observation sharing protocol where anchors periodically broadcast their information. In fact, during the first time slot, only the first anchor should broadcast its own information while the $(N_a - 1)$ other anchors only execute the tasks described in Section IV-A. At the second time slot, the second anchor derives an estimation error observation using the information received from the first anchor, adds it to its packet and broadcasts the resulting packet in the network. Upon reception of this information, the rest of anchors derive and store a second observation. Two observations are then available at the third anchor which also broadcasts them in the network. This process will continue until each regular sensor becomes aware of a sufficient number of observations. Note that if N_a is large enough so that $N_a - 1$ observations are sufficient to accurately derive the PDF, only two time slots are required. Indeed, after the first time slot, $(Na-1)$ observations are available and can be simultaneously broadcasted by the $(Na - 1)$ anchors in the network. It is noteworthy that each anchor could transmit few observations without incurring any power cost and, hence, the proposed non-parametric method also complies with the nature of EH-WSNs.

VI. Simulations results

In this section, we evaluate the performance of the proposed algorithm in terms of localization accuracy by simulations using Matlab. These simulations are conducted to compare, under the same network settings, the proposed algorithm with some of the best representative localization algorithms currently available in the literature, i.e., DV-Hop [3], LAEP [4] and EPHP [6]. All simulation results are obtained by averaging over 100 trials. In the simulations, sensors are uniformly deployed in a 2-D square area $S = 100 \times 100\ m^2$. We assume that all sensors' transmission coverage capabilities are uniformly distributed between 5 and 30 m. We also assume that the number of anchors is set to be 20 in all simulations. As a performance metric, we propose to the normalized root mean square error (NRMSE) defined as

$$\mathrm{NRMSE} = \frac{1}{N_u} \sum_{i=1}^{N_u} \frac{\sqrt{(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2}}{Tc_i}. \quad (19)$$

Fig. 6(a) plots the localization NRMSE achieved by DV-Hop,



(a) NRMSE vs. sensor density.

(b) NRMSE's CDF.

Fig. 6. Localization NRMSE.

EPHP, LAEP and the proposed algorithm for different sensor densities λ in EH-WSNs. From this figure, the proposed algorithm, with or without localization correction, always outperforms its counterparts. Indeed, our proposed algorithm turns out to be until about two, three and four times more accurate than LAEP, DV-Hop, and EPHP, respectively. Furthermore, as can be observed from Fig. 6(a), the NRMSE achieved by the proposed algorithm significantly decreases when the sensor density λ increases while those achieved by its counterparts slightly decreases then quickly saturate. This is expected since two conflicting phenomena arise when λ grows large. The first is that the approximation in (15) becomes more realistic and, hence, more accurate localization is performed. The second is the increase of the number of different transmission capabilities due to the heterogeneous-coverage nature of EH-WSNs when the sensor density increases. Since more different are the transmission capabilities in the network, worse is the accuracy of the former algorithms. This explains why their performance quickly saturates when the sensor density increases. The proposed algorithm's accuracy, in contrast, increases with λ since it takes into account the difference between the transmission capabilities that is typical of EH-WSN. This further proves the efficiency and suitability of the proposed localization algorithm to the heterogeneous-coverage nature of EH-WSNs. Fig. 6(b) illustrates the localization NRMSE's CDF. Using the proposed algorithm, 90% of the regular sensors could estimate their position within almost the fifth of their transmission capabilities. In contrast, 20% of the sensors achieve the same accuracy with LAEP, about 14% with DV-Hop, and only 9% with EPHP. This further proves the efficiency of the proposed algorithm. Fig. 7 shows the NRMSE standard deviation achieved by all localization algorithms. As can be observed from this figure, the NRMSE standard deviation achieved by the proposed algorithm substantially decreases when the sensor density increases while those achieved by the other algorithms slightly decrease. This is due once again to the fact that the proposed algorithm complies with the heterogeneous-coverage nature of the EH-WSN when the former algorithms

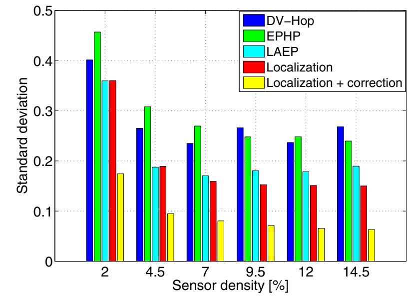


Fig. 7. Standard deviation vs. sensor density.

do not. Furthermore, the NRMSE standard deviation achieved by the proposed algorithm using approaches zero. This means that implementing our algorithm in EH-WSNs results in a high accurate localization guaranteed for any given realization. This result is very interesting in terms of implementation strategy, since it proves that the result in Fig. 6(a) becomes more and more meaningful as λ grows large.

VII. Conclusion

In this paper, a low-cost localization algorithm which accounts for the heterogeneous transmission coverage nature of EH sensors was proposed. A new approache were developed to accurately derive the expected hop progress (EHP). Using the latter the proposed algorithm is able to accurately locate the EH sensors owing to a new implementation that avoids any additional power consumption. Furthermore, we developed a correction mechanism which complies with the heterogeneous coverage nature of EH sensors to further improve localization accuracy without incurring any additional power cost. We showed that the proposed algorithm, whether applied with or without correction, outperforms in accuracy the most representative WSN localization algorithms in EH powering contexts.

References

[1] Y. Luo, J. Zhang, and K. B. Letaief, "Optimal Scheduling and Power Allocation for Two-Hop Energy Harvesting Communication Systems,“ *IEEE Trans. Wireless Commun.,* vol. 12, no. 9, pp. 4729-4741, Sep. 2013.

[2] N. Patwari *et al.*, "Locating the nodes: cooperative localization in wireless sensor networks,“ *IEEE Signal Process. Mag.,* vol. 22, no. 4, pp. 54-69, Jul. 2005.

[3] D. Niculescu and B. Nath, "Ad hoc Positioning System (APS),“ *Proc. IEEE GlobeCom'2001,* Nov. 25-29, 2001.

[4] Y. Wang *et al.*, "Range-Free Localization Using Expected Hop Progress in Wireless Sensor Networks,“ *IEEE Trans. Parallel Distrib. Syst.,* vol. 25, no. 10, pp. 1540-1552, Oct. 2009.

[5] D. Ma, M.J. Er, and B. Wang , "Analysis of Hop-Count-Based Source-to-Destination Distance Estimation in Wireless Sensor Networks With Applications in Localization,“ *IEEE Trans. Veh. Technol.,* vol. 59, no. 6, pp. 2998-3011, Jul. 2010.

[6] L. Kleinrock and J. Silvester, "Optimum Transmission Radii for Packet Radio Networks or Why Six is a Magic Number,“ *Proc. IEEE NTC '1978,* Dec. 4-6, 1978.

[7] A. EL Assaf, S. Zaidi, S. Affes, and N. Kandil, "Low-cost localization for multi-hop heterogeneous wireless sensor networks". Accepted for publication in *IEEE Trans. Wireless. commun.*, August 2015.

[8] A.R. Webb and K.D. Copsey, *Statistical Pattern Recognition,* 3^{rd} edition Wiley, USA, 2011.

[9] A.W. Bowman and A. Azzalini, "Applied Smoothing Techniques for Data Analysis: The Kernel Approach with S-Plus," *Proc. IEEE GlobeCom'2001,* Nov. 25-29, 2001.