SENSOR SELECTION FOR MODEL-FREE SOURCE LOCALIZATION: WHERE LESS IS MORE

Ehsan Tohidi[⋆] *Junting Chen*[†] *David Gesbert*[⋆]

[⋆] Communication Systems Department, EURECOM, Biot, France
[†] School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen, China

ABSTRACT

The ability for a wireless network to precisely localize the radio nodes composing it is a great tool towards system optimization and is increasingly seen as a basic service requirement. In the past, model-free algorithms such as weighted centroid localization (WCL) have proved popular, especially in the context of sensor networks, due to their simplicity and robustness to temporal changes in wireless propagation properties. However, WCL algorithms are biased since they implicitly require a uniform sensor distribution around the source in all directions. In this paper, we demonstrate that instead of employing all the sensors that result in a possibly unbalanced sensing pattern, it is better to reduce the number of sensors such that the subset of selected sensors symmetrically distributes around the source, which in principle would need to know the source location in advance. Here, we develop a sensor selection algorithm which manages that goal while blindly. Using less than half of the sensors, a 30% reduction in localization error is demonstrated from our numerical experiments.

***Index Terms*—** Sensor Selection, Source Localization, Model-Free Localization, Weighted Centroid.

1. INTRODUCTION

In wireless communications, exploiting the location information of transceivers often improves the capability of the network, including enhanced multi-cell radio resource management, dynamic location-aware routing, and better spatio-temporal sensing. In many cases, there is no handshake between a target transceiver and the infrastructure. For instance, in cellular communications, the network operator usually does not have access to a legacy user GPS data. Such a non-cooperative scenario hinders the use of conventional localization techniques such as those based on time-of-arrival (TOA) ranging [1, 2]. Here, the transceiver is treated as a signal source and the infrastructure is treated as a set of sensors, and the problem is referred to as a source localization problem.

Related works: The source localization problem has been studied extensively in different schemes. There are many works on range-based methods, e.g., ToA, [1, 3–5], range-free methods which require some parametric models, e.g., direction-of-arrival (DoA), [6–8]. Although model-based techniques offer better estimates, they require accurate information about the source in order to tune the parameters, that can be only achieved through a cooperation with the source or a precise knowledge of the environment [9]. In contrast, the model-free methods [10–14] have the strong advantage that they can operate with the sole received signal strength information, yielding some highly desired robustness properties with respect to model uncertainties. In particular, most existing model-free methods can be viewed as taking a weighted average of the sensor locations that measure the source's signal strength, where the intelligence of the algorithm then lies in the design of the weights. One example of these techniques is the WCL algorithm [9, 15]. WCL, in essence, indirectly estimates the source location via computing the centroid of the sensors. Unfortunately, in this paper, we point out that a significant bias exists in model-free approaches in some unfavorable source-sensor topologies, such as one where most sensors locate on one side of the source.

In the example illustrated in Fig. 1, we consider 10 sensors randomly distributed in a square region with a uniform distribution. A non-centered source is assumed, i.e., in the first quadrant. Fig. 1 presents the scenario and the results of performing WCL for source localization. The estimation error for this case is quite high because most of the sensors are located on one side of the source. An improved version of WCL to address the bias issue is the mean shift algorithm [16, 17], where only a subset of sensors is weighted. However the blind selection of a suitable subset of sensors in an open problem. The algorithm [17] iteratively evaluates different subsets till convergence. Generally, there also exists optimal trade-off in designing the sensor subset's size, between suppressing observation noise, i.e., a larger subset preferred, and reducing bias, i.e., a smaller subset preferred, the exploration of this trade-off is still an open problem. Further, a recent non-parametric matrix-based source localization algorithm is proposed in [18] which in contrast with the conventional model-free methods, exploits the unimodality and symmetry properties of the observation matrix to localize sources. However, the main drawbacks of the matrix-based algorithm

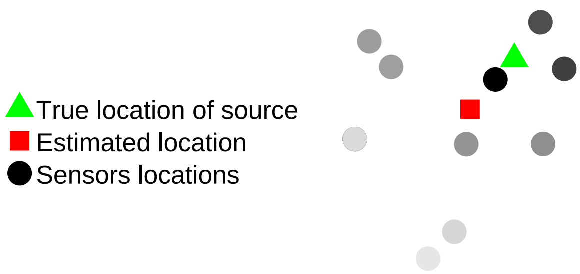


Fig. 1: A sample source localization scenario. Colors of sensors represent the level of their received power which is a function of distance (a darker color means a stronger power). The unbalanced density of sensor nodes around the source induces a bias in the estimated location.

are the required conditions on the sensor networks, e.g., a minimum distance among the sensors, and the high computational complexity as it needs to solve a matrix completion problem which make the method impractical.

Contributions: In this paper, introducing two desired properties of that selected sensors should retain, we propose a sensor selection algorithm that alleviates the problem of estimation bias. Specifically, the paper studies a model-free, range-free source localization technique using low-cost measurements from sensor networks. The data available to the network is a set of received signal strengths associated with their measurement locations. The algorithm extracts an optimized subset of sensors. The subset size is constrained in advanced below a certain threshold (e.g. based on maximum cost and/or energy considerations).

The contributions of this paper can be summarized as follows:

- We formulate two desired properties, namely closeness and evenness, to select the best subset of sensors and consequently reduce the bias of the WCL for source localization.
- We develop a sequential algorithm to iteratively improve the estimate of the source location from updating the best subset of sensors.
- We carry numerical experiments to demonstrate that using the proposed sensor selection strategy, the performance of the WCL can be improved with even fewer sensors.

2. SYSTEM MODEL

2.1. Topology and Signal Model

Consider a network consists of N sensors located in a square region with length D where the 2-D location of the nth sensor is $\mathbf{l}_n = [x_n, y_n]^T$. There is a source in the region of interest with location $\mathbf{s} = [x_s, y_s]^T$ and our aim is to obtain an estimation of its location $\hat{\mathbf{s}}$.

The following channel model for received signal strength is adopted where the received power of the nth sensor node from the source, P_n, is given by $P_n = h(d_n) + \xi_n$, where $d_n = ||\mathbf{l}_n - \mathbf{s}||_2$ is the distance between source and nth sensor with $||.||_2$ being l_2 norm, $h(.)$ represents the propagation loss versus distance, and ξ_n is the noise term for sensor n that might include thermal noise, shadowing effect, and error in sensor location. The collection of noise terms are denoted as $\boldsymbol{\xi} = [\xi_1, ..., \xi_N]^T$.

2.2. Conventional Model-Free Estimation Method

We aim to estimate the location of the source based on the measured powers and information about the sensor locations as $\hat{\mathbf{s}} = f(\mathbf{P}, \mathbf{L})$, where $f(.)$ is an estimation function, $\mathbf{P} = [P_1, ..., P_N]^T$ and $\mathbf{L} = [\mathbf{l}_1, ..., \mathbf{l}_N]^T$ are measured powers and location of sensors, respectively.

The WCL is a popular model-free estimation method that takes the form $\hat{\mathbf{s}} = \frac{\sum_{n=1}^N w_n \mathbf{l}_n}{\sum_{n=1}^N w_n}$, with $w_n = P_n$. Besides the advantages of the WCL algorithm such as being computationally simple and non-parametric, as explained in Fig. 1, it implicitly requires uniformly distributed sensors around the source in all directions. It is challenging to tackle such a restriction if we do not know the signal propagation model. One may want to put a significantly large weight to the sensor location that observes the highest signal strength so as to mitigate the bias due to non-uniform sensor distribution. For example, to set the weight as the squared of the received signal strength, the sensors with large measurement values will be significantly emphasized. However, such a highly selective weighting scheme may be vulnerable to noise, and therefore, one may want a mild weighting scheme, such as being a linear function of the signal strength, for noise tolerance. Yet, any logical design would require the parametric form of the signal as well as the statistics of the noise.

In this study, it is found that integrating sensor selection with WCL may improve the localization performance under model-free scenarios as will be discussed in the following sections.

2.3. Why Sensor Selection is Essential?

2.3.1. Mitigating Estimation Bias

As discussed along with Fig. 1, if we have an initial guess that the source, without loss of generality, is in the center of the first quadrant of a bounded region, then a majority of measurements in the second to the fourth quadrant may only contribute to the estimation bias. In this case, even if we have collected all the measurements, it may sometimes be better not to use them to form the final estimation of the source location.

2.3.2. Communication and Energy Constraints

Communication and energy constraints in a sensor network further motivate to integrate sensor selection with source localization. In many sensor networks, a node operates, i.e., sensing, processing, and communication, on its limited amount of battery energy [19]. On the other hand, in scenarios with a significant number of sensors, if all the sensors are used for source localization, there would be a huge communication load between sensors and the fusion center [20].

Consequently, we can only select K sensors for source localization. Denoting sets of all and selected sensors by $\mathcal{N}$ and $\mathcal{K}$, respectively, the optimization problem to minimize the estimation error can be formulated in the following form

$$\begin{aligned} \min_{\hat{\mathbf{s}}, \mathcal{K} \subset \mathcal{N}} \quad & \mathbb{E}\{\|\mathbf{s} - \hat{\mathbf{s}}\|_2^2\} \\ \text{subject to} \quad & |\mathcal{K}| \leqslant K, \end{aligned} \tag{1}$$

where $\hat{\mathbf{s}}$ is the estimated location of source obtained by applying WCL over the set of sensors $\mathcal{K}$. Therefore, an essential question is how to select the sensors.

3. SENSOR SELECTION: LESS IS MORE

3.1. Property of the Desired Subset of Sensors

In the most ideal case, one may want to select sensors that surround the source in a *uniform* and *balanced* manner. In other words, the weighted center of those sensors is expected to coincide with the true source location. However, since the source location (what we want to estimate) is unavailable, we need to seek other properties for such a desire set of sensors.

The following two properties could serve as proxies to describe characteristics of the desired set of sensors:

Closeness: If all sensors of a subset have very high measurements, this could be a good indicator that these sensors are close to the source. In a counter-scenario, if some sensors observe high energy and some other sensors observe very low energy, then those low-energy-observed sensors may locate far away and contribute bias. Further, as the location of the source is unknown, the signal strength measure can be equivalently expressed as a closeness indicator. As a result, we introduce normalized received powers as the closeness indicator (the higher the closer):

$$P'_n = \frac{P_n}{\max\limits_{i \in \mathcal{N} \setminus \mathcal{K}} P_i}. \tag{2}$$

Evenness in all direction: In an ideal situation, it is expected to have a few sensors that locate in the center of the subset observing high values, and there are also sensors filling the boundary of the subset in all directions; the sensors at the boundary observe relatively smaller value. Thus, we define the evenness indicator in the following form:

$$\alpha_n = \max_{i \in \mathcal{K}} < \mathbf{r}_n, \mathbf{r}_i >, \tag{3}$$

where $< ., . >$ is inner product operator and $\mathbf{r}_n$ is the normalized location vector of sensor n given by

$$\mathbf{r}_n = \frac{\mathbf{l}_n - \hat{\mathbf{s}}}{\|\mathbf{l}_n - \hat{\mathbf{s}}\|}. \tag{4}$$

We normalize the evenness indicator to have the same scale as the closeness indicator, i.e., between 0 and 1, as follows:

$$\alpha'_n = \frac{\alpha_n + 1}{2}, \quad \forall n \in \mathcal{N} \setminus \mathcal{K}, \tag{5}$$

Thus, we have two indicators, namely evenness and closeness indicators. However, a scalar measure is required for the objective of the optimization problem. Since for the evenness indicator, the lower value is the better, while it is opposite for closeness indicator, the following unified sensor selection for source localization (S3L) scalar metric is introduced in a multiplicative form including the negation of the normalized evenness indicator $(1 - \alpha'_n)$ and the closeness indicator as

$$\beta_n = (1 - \alpha'_n) P'_n, \tag{6}$$

where the higher the β_n, the better the sensor n is (in terms of the introduced indicators).

3.2. Sensor Selection Algorithm

As the indicator (6) requires a prior estimate of the source location, we need a sequential approach to first make an initial guess, and then, step-by-step improve the estimation. Specifically, the proposed algorithm consists of two stages:

- Initialization stage: Pick the center of the region of interest as the initial estimate $\hat{\mathrm{s}}^{(0)}$ of the source location.
- Sequential estimation stage: Sequentially localize the source and improve the estimation accuracy as follows. At the tth step, starting with an empty set of selected sensors, i.e., $\mathcal{K} = \varnothing$, select K sensors iteratively based on the S3L metric (7) below, where $\hat{\mathrm{s}}^{(t-1)}$ is used to replace $\hat{\mathrm{s}}$ in (4).

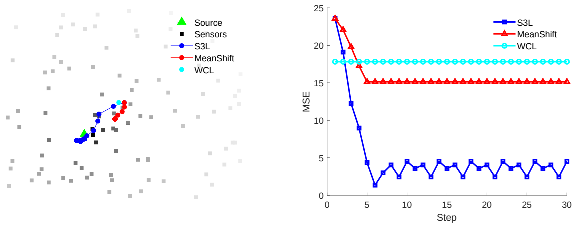
$$\begin{aligned} n^* &= \arg\max_{n \in \mathcal{N} \setminus \mathcal{K}} \beta_n, \\ \mathcal{K} &\leftarrow \mathcal{K} \cup \{n^*\}. \end{aligned} \tag{7}$$

  We continue the procedure for T steps and update the location estimation $\hat{\mathrm{s}}^{(t)}$ at the end of each step.

The computational complexity of the proposed S3L algorithm is $O(TNK^2)$ which is a linear function of the total number of sensors N.

4. SIMULATIONS

In this section, we compare the mean shift based algorithm in [17] and the conventional WCL algorithm with our proposed algorithm. Since the WCL algorithm is not an iterative



(a) (b)

Fig. 2: A sample localization problem (a) trajectory of the location estimation for different steps, (b) MSE versus step number.

method, its estimation result is the same for all the steps. In the following simulations, we consider a square region of interest with dimension $D = 100m$ and locations of the source and sensors are selected randomly with a uniform distribution. Moreover, $N = 100$ is assumed. Further, we consider noise with a Gaussian distribution, zero-mean, and variance σ_s^2.

Fig. 2 depicts the results for a sample sensor network where $K = 10$ sensors are selected. In Fig. 2(a), the estimated locations of the source in different steps are presented. In Fig. 2(b), estimation MSE of the three algorithms versus the step number is plotted. Both S3L and mean shift algorithms get the final estimation in less than 10 steps.

For the second scenario, the simulations are implemented for two different noise levels and the results are averaged over 1000 iterations to obtain plots in Fig. 3. In this scenario, $K = 20$ sensors are selected. As shown in Fig. 3, the MSE of the S3L increases drastically by increasing the noise level. Furthermore, comparing the result of the mean shift and WCL algorithms, it is clear that selecting a subset of closest sensors does not necessarily improve the estimation accuracy, particularly, when the selected subset of sensors is not well distributed around the source.

Finally, Fig. 4 presents the MSE vs. the number of selected sensors K where plots are displayed for three noise levels. All the plots are for step 10 of both algorithms. The WCL algorithm is not plotted in this figure because the result of the WCL is exactly equal to the results of the other two algorithms for $K = 100$, i.e., selecting all the sensors. Based on the performance of the proposed S3L algorithm in Fig. 4, for the noise-less case, increasing the number of selected sensors K increases the MSE. This is due to the fact that with a lower number of selected sensors, it is easier to reach a better evenness indicator. However, increasing the noise level, there would be a trade-off between the evenness and correctness of the measurements. For instance, for the case with $\sigma_s = 5$, increasing K, until some point (e.g., $K = 40$ sensors) the MSE is reduced (the more sensors, the result is more robust against the nuisance term), while after that, the MSE is increasing. This trade-off is less evident for the case with $\sigma_s = 10$ because the noise term is too strong and we need more sensors to obtain a good location estimation.

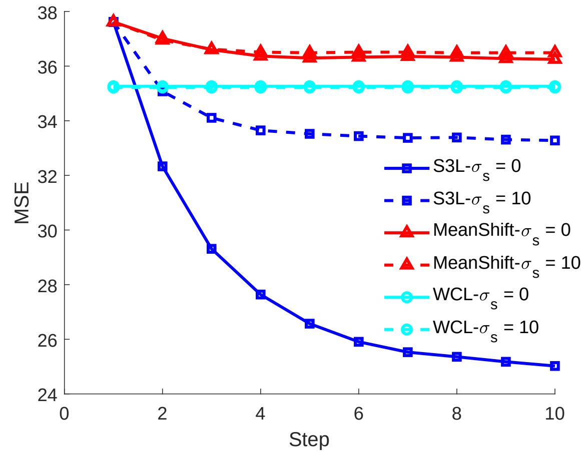


Fig. 3: MSE versus step number for noise levels of $\sigma_s = 0$ and $\sigma_s = 10$.

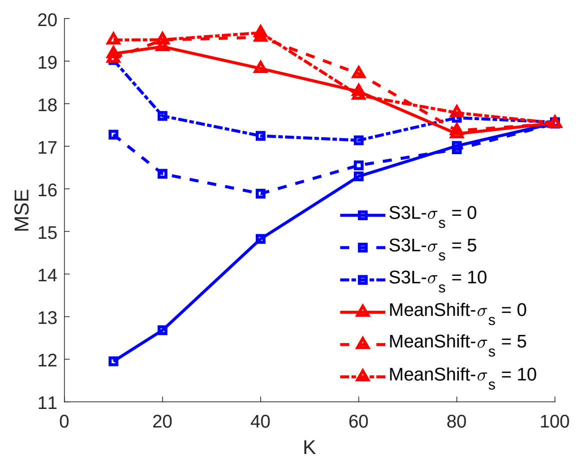


Fig. 4: MSE error versus number of selected sensors K for three different noise levels σ_s.

5. CONCLUSION

In the context of model-free source localization, we formulated two desired indicators to select the best of sensors and consequently improve the localization accuracy of the WCL algorithm. We proposed a sequential algorithm with a computational complexity that grows linearly with the number of sensors and iteratively improves the estimate of the source location form updating the best subset of sensors. Through numerical experiments demonstrated the 30% reduction in localization error with less than half of the sensors.

6. REFERENCES

[1] Baoqi Huang, Lihua Xie, and Zai Yang, "TDOA-based source localization with distance-dependent noises," *IEEE Transactions on Wireless Communications*, vol. 14, no. 1, pp. 468–480, 2014.

[2] X. Qu, L. Xie, and W. Tan, "Iterative constrained weighted least squares source localization using TDOA and FDOA measurements," *IEEE Transactions on Signal Processing*, vol. 65, no. 15, pp. 3990–4003, Aug 2017.

[3] Y. Zou, Q. Wan, and H. Liu, "Toa source node self-positioning with unknown clock skew in wireless sensor networks," in *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2019, pp. 4385–4389.

[4] Y. Sun, K. C. Ho, and Q. Wan, "Solution and analysis of TDOA localization of a near or distant source in closed form," *IEEE Transactions on Signal Processing*, vol. 67, no. 2, pp. 320–335, Jan 2019.

[5] K. Yang, G. Wang, and Z. Luo, "Efficient convex relaxation methods for robust target localization by a sensor network using time differences of arrivals," *IEEE Transactions on Signal Processing*, vol. 57, no. 7, pp. 2775–2784, July 2009.

[6] Bin Zhao, Xinping Guan, Lihua Xie, and Wendong Xiao, "Sensor selection for received signal strength-based source localization in wireless sensor networks," *Journal of Control Theory and Applications*, vol. 9, no. 1, pp. 51–57, 2011.

[7] Michael Rabbat, Robert Nowak, and James Bucklew, "Robust decentralized source localization via averaging," in *Proceedings.(ICASSP'05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005.* IEEE, 2005, vol. 5, pp. v–1057.

[8] Chen Meng, Zhi Ding, and Soura Dasgupta, "A semidefinite programming approach to source localization in wireless sensor networks," *IEEE signal processing letters*, vol. 15, pp. 253–256, 2008.

[9] J. Wang, P. Urriza, Y. Han, and D. Cabric, "Weighted centroid localization algorithm: Theoretical analysis and distributed implementation," *IEEE Transactions on Wireless Communications*, vol. 10, no. 10, pp. 3403–3413, October 2011.

[10] N. Bulusu, J. Heidemann, and D. Estrin, "Gps-less low-cost outdoor localization for very small devices," *IEEE Personal Communications*, vol. 7, no. 5, pp. 28–34, Oct 2000.

[11] Z. Ma, W. Chen, K. B. Letaief, and Z. Cao, "A semi range-based iterative localization algorithm for cognitive radio networks," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 2, pp. 704–717, Feb 2010.

[12] H. Chen, Q. Shi, R. Tan, H. V. Poor, and K. Sezaki, "Mobile element assisted cooperative localization for wireless sensor networks with obstacles," *IEEE Transactions on Wireless Communications*, vol. 9, no. 3, pp. 956–963, March 2010.

[13] Tian He, Chengdu Huang, Brian M Blum, John A Stankovic, and Tarek Abdelzaher, "Range-free localization schemes for large scale sensor networks," in *Proceedings of the 9th annual international conference on Mobile computing and networking*. ACM, 2003, pp. 81–95.

[14] Hongyang Chen, Yiu-Tong Chan, H Vincent Poor, and Kaoru Sezaki, "Range-free localization with the radical line," in *2010 IEEE International Conference on Communications*. IEEE, 2010, pp. 1–5.

[15] Christine Laurendeau and Michel Barbeau, "Centroid localization of uncooperative nodes in wireless networks using a relative span weighting method," *EURASIP Journal on Wireless Communications and Networking*, vol. 2010, pp. 6, 2010.

[16] Yanjun Chen, Quan Pan, Yan Liang, and Zhentao Hu, "Awcl: Adaptive weighted centroid target localization algorithm based on rssi in wsn," in *2010 3rd International Conference on Computer Science and Information Technology*, July 2010, vol. 9, pp. 331–336.

[17] Q. Zhou, X. Li, and Y. Xu, "Mean shift based collaborative localization with dynamically clustering for wireless sensor networks," in *2009 WRI International Conference on Communications and Mobile Computing*, Jan 2009, vol. 2, pp. 66–70.

[18] J. Chen and U. Mitra, "Unimodality-constrained matrix factorization for non-parametric source localization," *IEEE Transactions on Signal Processing*, vol. 67, no. 9, pp. 2371–2386, May 2019.

[19] C. Zhan, Y. Zeng, and R. Zhang, "Energy-efficient data collection in UAV enabled wireless sensor network," *IEEE Wireless Communications Letters*, vol. 7, no. 3, pp. 328–331, June 2018.

[20] E. Tohidi, M. Coutino, S. P. Chepuri, H. Behroozi, M. M. Nayebi, and G. Leus, "Sparse antenna and pulse placement for colocated MIMO radar," *IEEE Transactions on Signal Processing*, vol. 67, no. 3, pp. 579–593, Feb 2019.