Selection of Sensors for Efficient Transmitter Localization

Arani Bhattacharya*, Caitao Zhan†, Himanshu Gupta†, Samir R. Das†, Petar M. Djuric†
*KTH Royal Institute of Technology, †Stony Brook University
aranib@kth.se, {cbzhan, hgupta, samir}@cs.stonybrook.edu, petar.djuric@stonybrook.edu

***Abstract*—We address the problem of localizing an (illegal) transmitter using a distributed set of sensors. Our focus is on developing techniques that perform the transmitter localization in an efficient manner, wherein the efficiency is defined in terms of the number of sensors used to localize. Localization of illegal transmitters is an important problem which arises in many important applications, e.g., in patrolling of shared spectrum systems for any unauthorized users. Localization of transmitters is generally done based on observations from a deployed set of sensors with limited resources, thus it is imperative to design techniques that minimize the sensors' energy resources.**

In this paper, we design greedy approximation algorithms for the optimization problem of selecting a given number of sensors in order to maximize an appropriately defined objective function of localization accuracy. The obvious greedy algorithm delivers a constant-factor approximation only for the special case of two hypotheses (potential locations). For the general case of multiple hypotheses, we design a greedy algorithm based on an appropriate auxiliary objective function—and show that it delivers a provably approximate solution for the general case. We develop techniques to significantly reduce the time complexity of the designed algorithms, by incorporating certain observations and reasonable assumptions. We evaluate our techniques over multiple simulation platforms, including an indoor as well as an outdoor testbed, and demonstrate the effectiveness of our designed techniques—our techniques easily outperform prior and other approaches by up to 50-60% in large-scale simulations.

I. Introduction

Wireless transmitter localization via analysis of the received signal from multiple receivers or sensors is an important problem. While the problem has been widely explored, the problem exposes new challenges in many emerging applications due to the constraints of the application. In this work, we are specifically interested in a distributed monitoring system where a set of distributed RF sensors are tasked to detect and localize transmitters. These transmitters could be of various type. For example, in certain spectrum allocation scenarios, unknown primary transmitters need to be detected/localized. Or, in spectrum patrolling scenarios, unauthorized transmitters need to be detected/localized [1]. Recent work has explored new approaches for such monitoring where the RF sensors are crowdsourced, perhaps using various low-cost spectrum sensing platforms [2], [3]. The crowdsourcing makes densely deployed, fine grain spectrum sensing practical by creating suitable incentive mechanisms [4], [2].

Crowdsourcing makes the sensing *cost-conscious*. The cost here could be incentivization cost, cost of power, backhaul bandwidth on the part of the spectrum owner or the opportunity cost – being low-cost platform, the sensors may be able to only sense smaller spectrum bands at a time. Thus, involving only a small number of sensors or sensors with low overall cost budget (for a suitable cost model) for sufficiently accurate localization performance is critical. Prior work [2] that discuss sensor selection in this context only presents heuristics without any performance guarantees.

We do not use geometric approaches which rely on hard-to-model mapping of received power to distance. Instead, we use a hypothesis-driven, Bayesian approach for localization [5]. We focus on the optimization problem of selecting a certain number of sensors from among the deployed sensors such that an appropriately defined objective of localization accuracy is maximized. This optimization problem can also be used to solve the dual problem of selecting a minimum number of sensors (or sensors with the minimum total cost budget) to ensure at least a given localization accuracy. We adopt the framework of a hypothesis-driven localization approach wherein each hypothesis represents a configuration (location, power, etc.) of the potential transmitters and then the localization is equivalent to determining the most-likely prevailing hypothesis. See Figure 1. The hypothesis-driven framework does not require an assumption of a propagation model, and works for arbitrary signal propagation characteristics. The framework does, however, require prior training to build joint probability distributions of observation vectors for each hypothesis.

Our Contributions. In the above hypothesis-based framework, we develop an overall approach that enables selection of sensors that are most relevant to localize transmitters. In particular, we develop algorithms that aim to maximize localization accuracy for a given budget of number of sensors to be used for localization. More specifically, we make the following contributions in the paper.

1) We design a greedy algorithm (GA) that selects sensors iteratively to maximize the objective function of localization accuracy, under the constraint of number of sensors selected. We prove that GA yields a constant-factor approximate solution for the special case of the problem wherein there are only two hypotheses.
2) For the general case of more than two hypotheses, we design an alternate greedy scheme (called AGA) based on maximizing an auxiliary objective function. We prove that

* Work done when the author was at Stony Brook University.

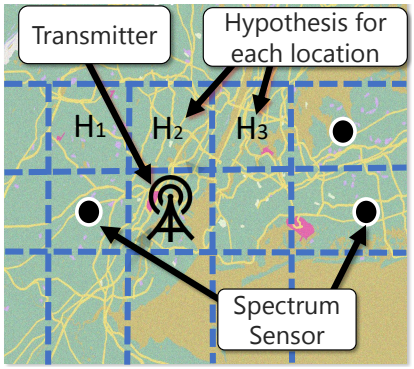


Fig. 1: Hypothesis-driven localization. The figure shows the simple case of localizing a single transmitter with fixed power; thus, there is a hypothesis created for each potential location. Observations from deployed sensors are analyzed to determine the most likely prevailing hypothesis (and thus, location).

AGA delivers a solution that has (i) an auxiliary objective value within a constant factor of the optimal auxiliary objective value, as well as (ii) a localization error within a certain factor of the optimal localization error.

3) We optimize the time complexity of our developed algorithms by a substantial factor, based on certain observations and reasonable assumptions. In addition, we generalize our techniques to more practical and useful settings.
4) We evaluate the performance of the developed algorithms over multiple evaluation platforms: (1) large-scale simulation using synthetically generated data using established signal propagation models, and (2) publicly available experimental data trace collected over an indoor WiFi network with 44 sensors, and (3) our own data collection using 18 outdoor software radio sensors in the 915 MHz band with a custom transmitter. Results show that our techniques outperform other state-of-the-art algorithm [2] substantially (up to a factor of 50-60%).

II. Background and Motivation

Problem Setting. The overall setting of the transmitter localization problem is as follows. Consider a geographic area, with a number of spectrum sensors deployed or available (if attached to mobile devices) at known locations. At any instant, one or more transmitters are allowed to transmit signals (on a common frequency). Each deployed/available spectrum sensor senses and processes the aggregate received signal, and reports appropriate metric (i.e., total received power or signal strength) to a central server which estimates the location of the transmitter(s) using the maximum-likelihood hypothesis algorithm as described below. The overall objective of our paper is to develop techniques to select an optimal subset of sensors in order to accurately localize any present transmitters. Though our developed techniques naturally extend to the case of multiple transmitters, for simplicity, we implicitly assume at most a single transmitter present at any instant. We consider the extension to multiple transmitters in §III-E. We start with defining basic notations used throughout the paper.

Hypotheses, Observations, and Inputs. We discretize the given space into locations $l_1, l_2, \ldots,$ and transmit power of a potential transmitter is similarly discretized into levels $p_1, p_2, \ldots$. We represent potential "configurations" of the possible transmitter by hypotheses $H_0, H_1, \ldots, H_m$, where each hypothesis H_i represents a configuration (l_i, p_i) of location l_i and transmit power p_i of a potential transmitter (see Figure 1). We use the convention that hypothesis H_0 corresponds to no transmitter being present. Localizing any potential transmitter is thus equivalent to determining the prevailing hypothesis. To do this, we use observations from a set of deployed sensors. We denote the observation vector of a subset of sensors $\mathbf{T}$ by $\mathbf{x_T}$ (we usually drop the subscript $\mathbf{T}$, as it is clear from the context).

Inputs. For a given set of sensors deployed over an area, we assume the following available inputs, obtained via a priori training, data gathering and/or analysis[1]:

- Prior probabilities of the hypotheses, i.e. $P(H_i)$, for each hypothesis H_i.
- Joint probability distribution (JPD) of sensors' observations for each hypothesis. More formally, for each hypothesis H_j, we assume $P(\mathbf{x}_S|H_j)$ to be known for each observation $\mathbf{x}_S$ for the entire set $\mathbf{S}$ of deployed sensor. Note that this also gives us the JPD's of each subset $\mathbf{T} \subseteq \mathbf{S}$.

Maximum a Posteriori Localization (`MAP`) Algorithm. We use Bayes rule to compute the likelihood probability of each hypothesis, from a given observation vector $\mathbf{x_T}$ for a subset of sensors $\mathbf{T}$:

$$P(H_i|\mathbf{x_T}) = \frac{P(\mathbf{x_T}|H_i)P(H_i)}{\sum_{j=0}^{m} P(\mathbf{x_T}|H_j)P(H_j)} \quad (1)$$

We select the hypothesis that has the highest probability, for given observations of a set of sensors. Formally, the `MAP` algorithm returns the hypotheses based on the following equation:

$$\arg\max_{i=0}^{m} P(H_i|\mathbf{x_T}) \quad (2)$$

The above `MAP` algorithm to determine the prevailing hypothesis is known to be *optimal* [7], i.e., it yields minimum probability of (misclassification) error. The above hypothesis-based approach to localization works for arbitrary signal propagation characteristics, and in particular, obviates the need to assume a propagation model. However, it does incur a one-time training cost to obtain the JPDs, which can be optimized via independent techniques [8].

Selection of Sensors for Localization. As mentioned above, in a typical setting, spectrum sensors may be deployed at predetermined locations or available at certain locations (if part of mobile devices) to sense unauthorized signals and thus localize any unauthorized transmitters. Two immediate problems of interest in this context are: where to deploy given a number of sensors, and once deployed/available, which subset of sensors

[1] In our concurrent work [6], we discuss novel interpolation techniques to minimize such training cost.

to select for localization. The latter problem of selection of sensors is motivated by the fact that, in most realistic settings, the sensors (or their mobile devices) are not tethered to AC power outlets and hence have limited energy resources. Moreover, spectrum sensors also incur cost in transmitting sensing data to the fusion/cloud center [9]. Thus, it is critical to optimize resources and costs incurred in localization of unauthorized transmitters, e.g., via the selection of an optimal set of sensors. Note that the sensor-selection problem can also be used to effectively *deploy* a given number of sensor, by assuming sensors available at all potential locations.

III. Optimal Sensor Selection for Intruder Localization

In this section, we address the problem of sensor selection for transmitter localization; informally, the problem is to select an optimal set of B sensors such that the overall probability of error of localizing a transmitter is minimized, given appropriate JPDs as discussed in the previous section. We start with formulating the problem in the following subsection. In following subsection, we present a greedy algorithm for it and prove that it is guaranteed to deliver an approximation solution for the special case of two hypotheses. However, as shown, the greedy algorithm can perform arbitrarily bad for the general case of multiple hypotheses. Thus, we then modify our algorithm to use an "auxiliary" objective function and show that the modified algorithm delivers an approximation solution for the general case of multiple hypotheses albeit with a slightly worse approximation ratio. Finally, we discuss optimizing the computation complexity of the designed algorithms, certain extensions and other issues.

A. LSS Problem Formulation

We start with formally defining the optimization objective (probability of error or misclassification) for a given subset of sensors. Then, we formally define the sensor selection problem, hereto referred to as *Localization Sensor Selection (LSS)* problem. Throughout this section, we use hypotheses H_0 to represent the hypotheses with no transmitters present, and H_i to represent the hypotheses wherein a transmitter is present in i^{th} configuration.

Probability of Error ($\mathrm{P}_{\mathrm{err}}(\mathbf{T})$). Recall that, for a given observation vector, the MAP localization algorithm outputs the hypothesis that has the most likelihood among the given hypotheses. Thus, MAP can also be looked upon as a classification technique. Given a subset of sensors $\mathbf{T}$, we define the *probability of error* or misclassification as the probability of the MAP algorithm outputting a hypothesis different from the actual *ground truth* (i.e., prevailing hypothesis). The expected or overall probability of error is an expectation of the probability of error over all possible prevailing hypotheses and/or observation vectors $\mathbf{x_T}$ from $\mathbf{T}$. Our techniques generalize to the notion of distance-based localization error, as discussed in §III-E.

Formally, let $\mathtt{MAP}(\mathbf{x})$ be the output of the MAP algorithm on observation vector $\mathbf{x}$ from a given subset of sensors $\mathbf{T}$. Let $\delta_{\mathtt{MAP}(\mathbf{x})\neq i}$ be the binary predicate that denotes whether MAP algorithm outputs the hypothesis H_i or not; here, δ_p is the indicator function which is 1 if the predicate p is true and 0 otherwise. Given H_i as the ground truth and $\mathbf{x}$ as the observation vector, the probability of error $\mathrm{P}_{\mathrm{err}}(\mathbf{T}|H_i,\mathbf{x})$ can be written as:

$$\mathrm{P}_{\mathrm{err}}(\mathbf{T}|H_i,\mathbf{x}) = \delta_{\mathtt{MAP}(\mathbf{x})\neq i}. \tag{3}$$

If the observation vector $\mathbf{x}$ is not given, then the expected (or marginal) probability of error for a given ground truth H_i is just an expectation over the random variable $\mathbf{x}$. That is, $\mathrm{P}_{\mathrm{err}}(\mathbf{T}|H_i)$ can be written as:

$$\mathrm{P}_{\mathrm{err}}(\mathbf{T}|H_i) = \sum_{\mathbf{x}} \delta_{\mathtt{MAP}(\mathbf{x})\neq i} P(\mathbf{x}|H_i) = E_{\mathbf{x}|H_i}[\delta_{\mathtt{MAP}(\mathbf{x})\neq i}]$$

Since expectation of an indicator random variable is its probability, we can simplify the above equation as:

$$\mathrm{P}_{\mathrm{err}}(\mathbf{T}|H_i) = P(\mathtt{MAP}(\mathbf{x}) \neq i|H_i) \tag{4}$$

Above, the probability is over the random variable $\mathbf{x}$. Now, if the ground truth hypothesis is also not given, we can compute an expectation over all possible hypotheses. Thus, the (overall) *probability of error* for a given set of sensors $\mathbf{T}$ is given by:

$$\mathrm{P}_{\mathrm{err}}(\mathbf{T}) = \sum_i P(\mathtt{MAP}(\mathbf{x}) \neq i|H_i)P(H_i) \tag{5}$$

Localization Accuracy Function, $\mathrm{O}_{\mathrm{acc}}(\mathbf{T})$. To facilitate a greedy approximation solution, we formulate our sensor selection as a maximization problem—and thus, define a corresponding maximization objective. In particular, we define the *localization accuracy* $\mathrm{O}_{\mathrm{acc}}(\mathbf{T})$ as $1-\mathrm{P}_{\mathrm{err}}(\mathbf{T})$. Based on the above equation Eqn. 5, we get the expression for $\mathrm{O}_{\mathrm{acc}}(\mathbf{T})$ as:

$$\mathrm{O}_{\mathrm{acc}}(\mathbf{T}) = 1-\mathrm{P}_{\mathrm{err}}(\mathbf{T}) = \sum_i P(\mathtt{MAP}(\mathbf{x}) = i|H_i)P(H_i) \tag{6}$$

Localization Sensor Selection (LSS) Problem. Consider a geographic area with a set of sensors $\mathbf{S}$ deployed. Given a set of hypotheses and JPD's, as defined in previous section, the LSS problem is to select a subset $\mathbf{T} \subseteq \mathbf{S}$ of sensors with minimum probability of error $\mathrm{P}_{\mathrm{err}}(\mathbf{T})$ (or maximum localization accuracy $\mathrm{O}_{\mathrm{acc}}(\mathbf{T})$), under the constraint that $|\mathbf{T}|$ is at most a given budget B. The above formulation implicitly assumes a uniform cost for each sensor; we generalize our techniques to handle non-uniform sensor costs (see §III-E).

It can be shown (proof omitted here) that the above LSS problem is NP-hard, via reduction from the well-known maximum-coverage problem. Thus, we develop approximation algorithms below; in particular, our focus is on developing greedy approximation algorithms. The key challenge lies in showing that the objective function satisfies certain desired properties that ensure the approximability of the algorithm.

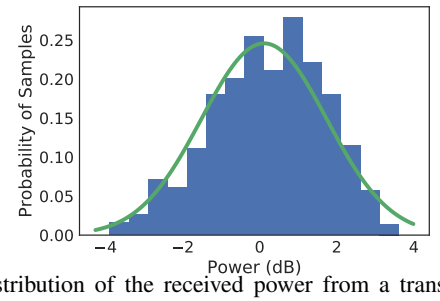


Fig. 2: Distribution of the received power from a transmitter at an RTL-SDR sensor, and the Gaussian fit (green line) of the observed distribution.

B. Greedy Algorithm (GA)

In this subsection, we analyze a simple greedy approach and show that it delivers a constant-factor approximate solution for the special case of two hypotheses and Gaussian JPD's. In the next subsection, we present a modified greedy algorithm for the general case of more than two hypotheses.

Greedy Algorithm (GA): A straightforward algorithm for the LSS problem is a greedy approach wherein we iteratively select a single sensor at each stage. At each stage, we select the sensor that improves the localization accuracy $\mathrm{O}_{\mathrm{acc}}(\mathbf{T})$ the most. The algorithm iterates until the given budget B is reached. We call this algorithm Greedy Algorithm (GA); see Algorithm 1 for the pseudo-code.

Constant-Factor Approximation for 2 Hypotheses. We observe that when the spectrum sensors are deployed outdoors, the joint probability distribution (JPD) of the observation vectors is approximately Gaussian. See Figure 2, which shows the distribution obtained by a single RTL-SDR [10] based spectrum sensor and a USRP-based transmitter. This assumption of Gaussian JPDs allows us to derive closed-form expressions for the objective functions, at least for the case of 2 hypotheses, and thus prove a performance guarantee of 63%. The result is stated in Theorem 1 below.

Theorem 1. *For the special case of two hypotheses and Gaussian JPDs, GA gives a subset* $\mathbf{T}$ *of sensors whose localization accuracy is at least 63% of the optimal.* □

We defer the proof of the above theorem to Appendix A, but the performance guarantee of the greedy approach holds because the localization accuracy function $\mathrm{O}_{\mathrm{acc}}()$ can be shown to be *"monotone" and "submodular"* for the above special case. The function $\mathrm{O}_{\mathrm{acc}}()$ being *monotone* signifies that for a given $\mathbf{T}$ and a sensor $s \notin \mathbf{T}$, $\mathrm{O}_{\mathrm{acc}}(\mathbf{T} \cup \{s\}) \geq \mathrm{O}_{\mathrm{acc}}(\mathbf{T})$. Intuitively, the monotone property means that adding a sensor to a set of already selected sensors can never decrease the localization accuracy. Also, $\mathrm{O}_{\mathrm{acc}}()$ being *submodular* signifies that for any subsets $\mathbf{T}_1$ and $\mathbf{T}_2$ such that $\mathbf{T}_1 \subseteq \mathbf{T}_2$, we can show that for any sensor $s \notin \mathbf{T}_1, \mathrm{O}_{\mathrm{acc}}(\mathbf{T}_1 \cup \{s\}) - \mathrm{O}_{\mathrm{acc}}(\mathbf{T}_1) \geq \mathrm{O}_{\mathrm{acc}}(\mathbf{T}_2 \cup \{s\}) - \mathrm{O}_{\mathrm{acc}}(\mathbf{T}_2)$. Intuitively, the submodular property means that the "benefit" of adding a sensor s decreases over GA's iterations, i.e., as the selected set of sensor grows (from $\mathbf{T}_1$ to $\mathbf{T}_2$, here). It is well known that if an objective function is both monotone and submodular, then a greedy approach that iteratively maximizes the objective function will return a constant-factor approximate solution [11].

Algorithm 1 Greedy Algorithm (GA).

INPUT: Set of available sensors $\mathbf{S}$, budget B, objective $\mathrm{O}_{\mathrm{acc}}$
OUTPUT: Subset of sensors $\mathbf{T}$

$\mathbf{T} \leftarrow \phi$
while $|\mathbf{T}| \leq B$ **do**
  $L \leftarrow \mathrm{O}_{\mathrm{acc}}(\mathbf{T})$
  $\max \leftarrow 0$
  for all $s \in \mathbf{S} \setminus \mathbf{T}$ **do**
    $M = \mathrm{O}_{\mathrm{acc}}(\mathbf{T} \cup \{s\}) - L$
    if $M > \max$ **then**
      $\max \leftarrow M$
      $r \leftarrow s$
    end if
  end for
  $\mathbf{T} \leftarrow \mathbf{T} \cup \{r\}$
end while
return $\mathbf{T}$

Performance of GA for more than two Hypotheses. For the case of more than two hypotheses, GA no longer provides a constant-factor approximation. In fact, we can show via a counter-example that the $\mathrm{O}_{\mathrm{acc}}()$ is not submodular for more than 2 hypotheses, even if the given JPDs are Gaussian; we omit the details for space constraints.

C. Auxiliary Greedy Algorithm (AGA)

In the section, we design an approximation algorithm for the general case of multiple hypotheses based on an auxiliary objective function. To do so, we first analyze the proof of Theorem 1 and see why it doesn't generalize if the number of hypotheses is greater than 2. This insight helps in defining an "auxiliary" objective function that is the key to designing the approximation algorithm for the general case.

Auxiliary Function. Let us consider a special case of MAP algorithm, viz., $\mathtt{MAP}_{ij}$ which compares the likelihood of only two hypothesis H_i and H_j and returns the one with a higher likelihood. It is easy to formulate the objective function $\mathrm{O}_{\mathrm{acc}}$ in terms of $\mathtt{MAP}_{ij}$ too. From Equation 6, we easily get:

$$\mathrm{O}_{\mathrm{acc}}(\mathbf{T}) = \sum_{i=0}^{m} P(\bigcap_{j \neq i} \mathtt{MAP}_{ij}(\mathbf{x}) = i | H_i) P(H_i) \tag{7}$$

$$\mathrm{O}_{\mathrm{acc}}(\mathbf{T}) = \sum_{i=0}^{m} [1 - P(\bigcup_{j \neq i} \mathtt{MAP}_{ij}(\mathbf{x}) = j | H_i)] P(H_i) \tag{8}$$

Above, $\mathbf{x}$ represents the observation vector for the set of sensors $\mathbf{T}$. For the case of two hypothesis, the above expression is just $\sum_{i=0}^{1} [1 - P(\mathtt{MAP}_{ij}(\mathbf{x}) = j | H_i)] P(H_i)$ where j is 1 if i is 0 and vice-versa; Theorem 1 essential shows that the term $P(\mathtt{MAP}_{ij}(\mathbf{x}) = i | H_i)$ is submodular. However, for the case of multiple hypothesis, computing the probability for a union of events involves product (and sum) of appropriate probability terms. Note that product of submodular functions need not be submodular, while sum of submodular functions is

submodular. Thus, we *approximate* the above $\mathrm{O}_{\mathrm{acc}}()$ expression as follows, so that it is a sum of submodular terms. In effect, in defining the auxiliary objective $\mathrm{O}_{\mathrm{aux}}()$, we estimate the probability of union of events in the above equation by just taking a summation of the probability of events, i.e., we ignore the other terms involving subsets of events. Formally, we define the auxiliary objective $\mathrm{O}_{\mathrm{aux}}()$ for a set of sensors $\mathbf{T}$ as:

$$\mathrm{O}_{\mathrm{aux}}(\mathbf{T}) = 1 - \sum_{i=0}^{m} \sum_{j \neq i} P(\texttt{MAP}_{ij}(\mathbf{x}) = j | H_i) P(H_i) \quad (9)$$

The above auxiliary objection function is submodular if the JPDs are Gaussian, as it is a sum of submodular functions ($P(\texttt{MAP}_{ij}(\mathbf{x}) = i | H_i)$ is submodular, as per Theorem 1's proof). Note that, for a competitive algorithm for the original LSS problem, we also need to show that maximizing $\mathrm{O}_{\mathrm{aux}}()$ also maximizes the original objective function $\mathrm{O}_{\mathrm{acc}}()$.

Auxiliary Greedy Algorithm (AGA). We now modify our Greedy Algorithm (Algorithm 1) to iteratively maximize the auxiliary objective $\mathrm{O}_{\mathrm{aux}}()$ instead of the original objective $\mathrm{O}_{\mathrm{acc}}()$. We call this algorithm as Auxiliary Greedy Algorithm (AGA). From the submodularity of the $\mathrm{O}_{\mathrm{aux}}()$ for Gaussian JPDs, it is easy to see that AGA delivers a solution $\mathbf{T}$ s.t. $\mathrm{O}_{\mathrm{aux}}(\mathbf{T})$ is within 63% of the optimal $\mathrm{O}_{\mathrm{aux}}()$ possible. The following lemma states that maximizing $\mathrm{O}_{\mathrm{aux}}$ also maximizes $\mathrm{O}_{\mathrm{acc}}$. See Appendix B for a proof.

Lemma 1. *Let $\mathbf{T}$ be a subset of sensors already selected by AGA at some iteration. We claim that $\mathrm{O}_{\mathrm{aux}}(\mathbf{T}) \leq \mathrm{O}_{\mathrm{acc}}(\mathbf{T}) \leq 1 - \frac{1}{k}(1 - \mathrm{O}_{\mathrm{aux}}(\mathbf{T}))$, where k is a value less than m that decreases as $\mathbf{T}$ grows (i.e., over AGA's iterations).* □

We empirically evaluate the value of k defined above in §IV. The above lemma easily yields the below result.

Theorem 2. *For Gaussian JPDs, AGA delivers a subset $\mathbf{T}$ of sensors such that*

$$P_{err}(\mathbf{T}) \leq 0.37 + 0.63 k P_{err}(\texttt{OPT}),$$

where k is as defined in the above Lemma and OPT is the optimal solution. □

D. Optimizing AGA's Computation Cost

In a straightforward implementation of AGA (akin to Algorithm 1 for GA), $\mathrm{O}_{\mathrm{aux}}$ function is computed (using Eqn. (9)) Bn number of times where n is the total number of sensors. Eqn. (9) requires m^2 computations of the expectation $P(\texttt{MAP}_{ij}(\mathbf{x}) = j | H_i)$, which, for Gaussian distributions, effectively requires computing the Eqn. 10 and thus takes $O(B^2)$ time as it involves matrix multiplication of the observation vector of dimension B with the covariance matrix of dimension $B \times B$. Thus, the overall time complexity of a straightforward implementation of AGA is $O(m^2 n B^3)$. As mentioned before (and in §II), the number of hypotheses m can be large due to the large number of potential transmitter locations and power values; however, we can reduce the time complexity to $O(Bn)$ as discussed below, based on some observations and optimizations.

Reducing Number of Comparisons. Consider a sensor s whose benefit is to be computed in the `for` loop of Algorithm 1. Below, we show that effectively we only need to consider a constant number of (H_i, H_j) pairs in Eqn. (9) when computing s's benefit, and thus removing the m^2 factor from the time complexity. We implicitly assume a single transmitter in the below discussion, and later extend our argument to multiple transmitters. Let us use R to denote the maximum transmission "range" of the transmitter; formally, R is such that, beyond R, the probability distribution of the signal received from the transmitter is similar to the signal received when there is no transmitter present. We stipulate that any observation x_s at s, $P(x_s|H_{i1}) = P(x_s|H_{i2})$ for any two hypotheses H_{i1} and H_{i2} whose corresponding locations l_{i1} and l_{i2} are more than R distance away from s. The implication of the above observation is that, for a given sensor s, we can group all the hypotheses H_i with corresponding location l_i more than R distance away from s into one single "super" hypothesis H_s. Then, if the total number of hypotheses with corresponding locations within a distance of R from s is say G_R, then the total number of comparisons between pairs of hypotheses in Eqn. (9) is effectively only $(G_R+2)^2$, involving G_R hypotheses, H_0, and H_s. The above brings down the overall time complexity of AGA to $O(G_R^2 n B^3)$. Note that G_R is essentially equal to the number of grid locations within a circle of radius R times the total number of power levels, and thus, can be considered as constant (does not vary across problem instances)—which reduces AGA's time complexity to $O(nB^3)$.

Independent Sensor Observations. If we assume that the observations across sensors are conditionally independent, then the JPDs can be instead represented by independent probability distributions at each sensor location. In this case, the covariance matrix is purely diagonal, which allows us to "incrementally" compute the benefit of a sensor from one AGA iteration to another and thus reduce AGA's time complexity by an additional factor of B^2—and thus to $O(nB)$. See Appendix C for details.

E. Generalizations

Weighted (Distance-Based) Objective Function. The probability of error $\mathrm{P}_{\mathrm{err}}$ function penalizes *uniformly* for each misclassification. However, in general, it would be useful to assign different penalties or weights for different misclassifications. E.g., Eqn (9) should be generalized to:

$$\mathrm{O}'_{\mathrm{aux}}(\mathbf{T}) = 1 - \sum_{i=0}^{m} \sum_{j \neq i} w_{ij} P(\texttt{MAP}_{ij}(\mathbf{x}) = j | H_i) P(H_i)$$

Above, w_{ij} is the weight function. We note that our techniques and proofs of performance guarantees generalize easily to the above generalized function, irrespective of the weight function. In particular, weight w_{ij} can be the Euclidean distance between the locations l_i and l_j corresponding to the hypotheses

H_i and H_j. For the general case of multiple transmitters, where H_i and H_j may represent configuration of multiple transmitters, a minimum-cost matching based objective can be used to define the weight w_{ij}; if the number of transmitters in H_i and H_j are different, then an appropriately penalty for misses or false-alarms can be added to the weight.

Non-Uniform Sensor Cost. Another generalization of interest is to allow non-uniform cost for sensors, e.g., to prefer sensors with more (remaining) battery resources. Here, each sensor s may have a different cost $c(s)$, and the LSS problem constraint becomes: total cost of the selected set of sensors must be less than a given cost budget. For this version of the LSS problem, our algorithms need to be slightly modified in that we should pick the sensor that offers the highest improvement in the objective function per unit cost. To ensure a theoretical performance guarantee, we also need to use the "knapsack trick," i.e., to pick better of the two solutions: one returned by the modified algorithm, and the other the best one-sensor solution [12]. It can be shown that the overall algorithm still offers a theoretical performance guarantee for submodular functions, but the performance ratio is reduced by a multiplicative factor of 2. The above model is useful when designing a "load-balanced" strategy to maximize network lifetime of a system—therein, the sensor-selection algorithm must be run periodically based on the remaining battery resources.

Multiple Transmitters. Till now, we have implicitly assumed that a single transmitter was present. Our selection[2] techniques naturally generalize to the case of multiple transmitters, if we represent each *combination* of configurations of present transmitters by a separate hypothesis. Since the GA and AGA algorithms are formulated in terms of hypotheses, they generalize naturally to localization of multiple transmitters. However, the key challenge arises due to the large number of hypotheses—exponential in the number of potential transmitters— and thus, the high time complexity of AGA. Fortunately, the techniques discussed in previous subsection can be extended for the case of multiple transmitters as follows.

The key observation is that, for a given hypothesis H_i, the probability distribution of observations at a sensor s depends only on the configuration of transmitters in H_i that within a distance of R of s. I.e., for any observation x_s at a sensor s, $P(x_s|H_{i1}) = P(x_s|H_{i2})$ for any two hypotheses H_{i1} and H_{i2} that have the same configuration (locations and powers) for transmitters that are within a distance of R of s. The implication of the above observation(s) is that, for a given s, we can group the given hypotheses into *equivalence classes* based on the configuration of transmitters close of s, and to compute the benefit of a sensor s with AGA's iteration, we only need to compare pairs of equivalence classes (rather than the original hypotheses). The number of such equivalence classes is easily seen to be equal to G_R^T where G_R is the number of locations (grid cells) within R times the number of power levels, and T is the maximum number of transmitters possible/allowed within a range R of s (or any location). Thus, computation of benefit of s requires consideration of G_R^{2T} pairs of equivalence classes. If we assume T to be a small constant, then the overall time complexity of AGA reduces to $O(nB^3)$ as before, and to $O(nB)$ if we assume conditional independence of sensor observations.

[2]In our concurrent work [6], we optimize the MAP algorithm itself for the case of localizing multiple transmitters.

IV. Evaluation

In this section, we evaluate the performance of our algorithms developed in the previous sections. We first describe the evaluation platforms used in our experiments.

A. Evaluation Platforms

We use the following three evaluation platforms with varying fidelity of signal propagation characteristics, to demonstrate the performance of our techniques.

- *Simulation based on synthetic data.* To demonstrate the scalability of our techniques and the sensitivity of our algorithms to changes in settings, we consider a large geographic area of 4km by 4km, with signal path-loss values generated using the SPLAT! [13] application for the Longley-Rice [14] model. We use the noise in the sensor measurements (measured independently) to generate the required JPDs. We assume observations to be conditionally independent, thus representing the JPDs as set of probability distributions, one for each sensor and intruder configuration pair. Unless otherwise stated, for this large-scale platform, we use 100m x 100m grid cells giving 1600 potential locations, randomly deploy a transmitter at the height of 30m at a random power between 27-33 dBm which corresponds to roughly 250-750m of transmission range. We randomly deploy 100 spectrum sensors in the area.
- *Indoor Data.* We use publicly available data [15], which deploys transmitters and receivers at 44 locations in an indoor building of an area of $14m \times 14m$, with the transmitters sending signals at 2.443 GHz. Here, we use 1m x 1m grid cells, thus giving us a total of 196 potential locations and hypotheses.
- *Outdoor Testbed.* Finally, to evaluate our techniques in a more practical outdoor setting, we deploy our own testbed in a parking area of dimension $10m \times 10m$. Each grid cell has size of 1m x 1m. We place a total of 18 sensors on the ground. The sensors consist of single-board computers such as Raspberry Pi's and Odroid-C2's connected to an RTL-SDR dongle. The RTL-SDRs use dipole antennas. We collect raw Inphase-Quadrature (I/Q) samples from the RTL-SDRs [10], while transmitting data using a USRP-based transmitter from each grid cell at a height of 1.5m. We perform FFT on the I/Q and update the evaluation dataset description samples with a bin size of 256 samples to get the signal power values, and then utilize the mean and standard deviation of the power reported for each of the sensors.

Metrics We evaluate the performance of a localization strategy in terms of two key metrics: (i) Localization accuracy, i.e., $\mathrm{O}_{\mathrm{acc}}(\mathbf{T})$, and (ii) Weighted localization error, which weights

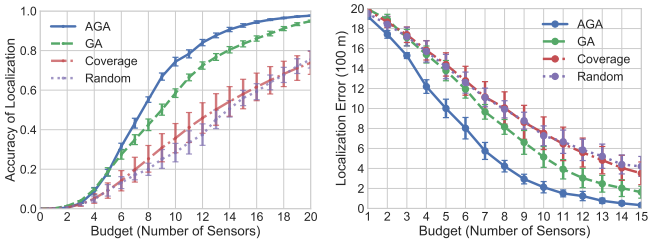


Fig. 3: Comparison of various techniques for (i) Localization accuracy (O_{acc} ()), and (ii) Weighted localization error, for varying available budget (number of sensors).

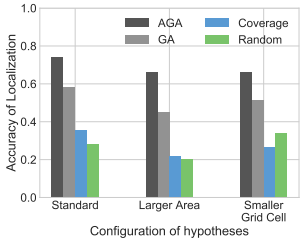


Fig. 4: Comparison for configurations with different number of hypotheses, with a fixed budget of 10 sensors.

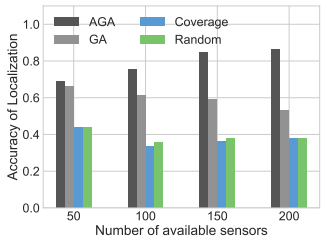


Fig. 5: Comparison for varying number of available sensors, with a fixed budget of 10 sensors.

the misclassification error by the Euclidean distance between the actual and the predicted location (§III-E).

Compared Algorithms. We implement both of our designed algorithms, AGA and GA. We also implement two other techniques for comparison purposes. The first technique, called `Coverage`, is the selection heuristic from the recent work [2], which essentially tries to maximize the "coverage" of the sensors in a greedy manner.[3] We also implement a `Random` algorithm which selects the required sensors randomly. We implement these algorithms in python, with extensive use of *numpy* library for vectorized operations. For our algorithms, we also leverage the data-level parallelism inherently present in computation of O_{acc} and O_{aux} by utilizing a GPU using numba library [16]. With the above, GA and AGA run in less than a minute, over a 3.3GHz i9-7900X CPU with 20 cores. To ensure that our results are statistically significant, we run each of the algorithms 100 times; the range of values is plotted in each of the figures.

B. Simulation Based on Synthetic Data

Varying Budget. Figure 3 shows the performance of our techniques for budgets varying from 1 to 20 sensors. We observe that AGA and GA easily outperform other two algorithms in terms of both metrics, with AGA outperforming even GA quite significantly. For example, AGA outperform `Coverage` by up to 39% and 56% for localization accuracy and error respectively, while outperforming GA by 15% and 50% for the two metrics respectively.

Varying Number of Hypotheses. We now show the performance of our algorithms in terms of localization accuracy, for varying number of hypotheses. In Figure 4, we plot three different cases: (i) the default configuration of 100m by 100m grid cells, (ii) a larger area of 6km by 6km with 100m by 100m grid cells giving 3600 potential locations, and finally (iii) a configuration with default 4km by 4km area, but smaller 62.5m by 62.5m grid cells. First, we observe that AGA continues to outperform other techniques significantly across different cases, with the performance gap between AGA and others increasing with increase in number of hypotheses. Also, as expected, with increase in area and thus number of hypotheses, the accuracy of each of the algorithms falls, but deterioration in AGA's accuracy is much less compared to other techniques.

[3]Their approach *Metropolis* performs worse than their greedy approach in open areas [2], and hence, not compared.

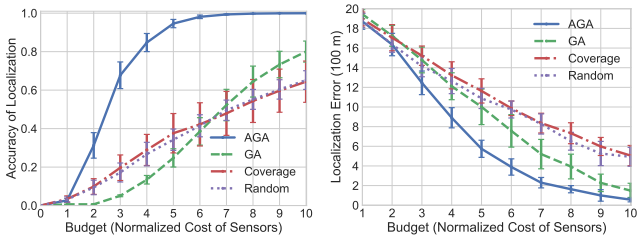


Fig. 6: Comparison of various techniques, for sensors with non-uniform cost.

Varying Sensor Density, and Non-Uniform Sensor Costs. Figure 5 shows the accuracy of localization for varying sensor density (i.e., number of available sensors) with a fixed budget of 10 sensors. We observe that AGA continues to outperform other techniques, with localization accuracy of AGA significantly improving with increase in number of sensors. Surprisingly, however, the performance of other techniques reduces slightly with increase in number of sensors, across these specific set of experiments. We also evaluate performance of techniques under the setting of sensors with non-uniform cost. See Figure 6. We observe that AGA continues to outperform the other techniques in both metrics.

Empirical Evaluation of k Value. We now evaluate the k value as defined in Lemma 1. In particular, the performance guarantee of AGA depends on the value of k, with better performance guarantee for lower k (ideally, k should be equal to 1). Figure 7 shows the value of k for varying budget. We observe that at low budgets, the value of k is large, but it reduces with increases in budget. In particular, for budgets of 10 and 15 sensors, k equals 1.78 and 1.19 respectively. This shows that AGA's performance guarantee as per Theorem 2 reaches its near-best for a moderately small budget.

Comparison with Optimal in Small Instances. To confirm AGA's performance with respect to optimal, we consider small instances of the problem and compare AGA with an optimal algorithm (exhaustive search). See Figure 8. We observe that AGA and optimal perform near-identically, with the optimal algorithm yielding at most 0.7% higher localization accuracy than AGA.

C. Evaluation in Indoor and Outdoor Testbeds

Indoor Data. We now evaluate our techniques over a publicly available data-trace taken in an indoor environment, as

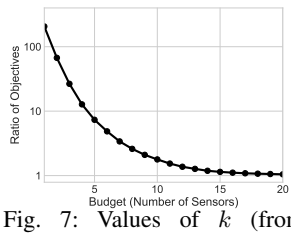


Fig. 7: Values of k (from Lemma 1) for varying budget.

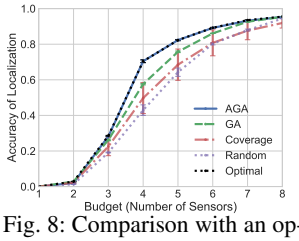
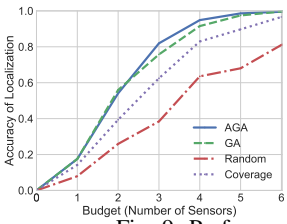


Fig. 8: Comparison with an optimal algorithm.



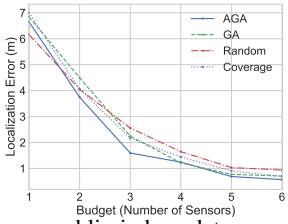


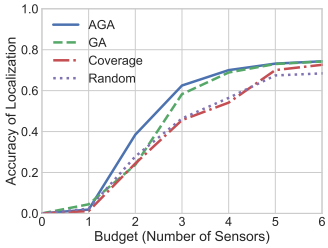
Fig. 9: Performance over public indoor data.

described in the previous subsection. See Figure 9. We again observe similar performance trends as in previous experiments, for both the performance metrics. The relatively smaller performance gap between AGA and GA is likely due to smaller a number of hypotheses.

Outdoor Testbed Figure 10 shows the performance of various algorithms over our outdoor testbed described in the previous subsection. Observe that AGA again performs the best among all techniques in both the metrics. As in the indoor testbed, the performance gap between the AGA and GA is less compared to the large-scale simulations due to a small number of hypotheses. Note that because of the noise in the dataset, the localization accuracy reaches a maximum of only 75% even with all the 18 sensors.

V. Related Work

Sensor Selection for Transmitter Localization. A large number of works have developed techniques for detecting and localizing transmitters or intruders that emit radio signals [17], [18]. Note that the transmitter localization problem is slightly different from another well-studied problem of indoor localization [19], wherein a user is localized based on signal *received* from multiple transmitters; herein, the issue of selecting transmitters to localize the user has no incentive, and hence not been addressed before. To the best of our knowledge, none of these prior works on transmitter localization either have addressed the optimization problem addressed in the paper. The closest related works are [1] and [2] as discussed next. The work [1] focuses on *detection* of unauthorized transmitters using low-cost sensors in the context of shared spectrum systems; they consider the problem of selection of sensors in this context, and propose a heuristic with no performance guarantees. The key difference of our work from theirs is that they focus on detection of transmitters, which is a much simpler problem than localization of transmitters. In addition, [2] considers selection of sensors for transmitter localization, but with a objective of maximizing the "coverage" of the region by the sensors. They present heuristics without



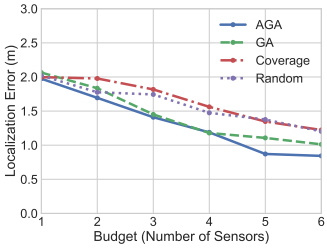


Fig. 10: Performance over outdoor testbed data.

any performance guarantees. Nevertheless, we implement their approach and compare with our techniques (§IV).

Sensor Selection in Sensor Networks. Sensor selection is a natural problem to address in the context of wireless sensor networks deployed to detect and/or localize an event or phenomenon (see [20] for a survey). Many of these works have leverage the submodularity property to develop greedy approximation algorithms. The closest work among these is that of [21] which shows approximability of the greedy approach for the problem of minimizing uncertainty in estimating a spatial phenomenon (e.g., temperature). However, in general, the key difference of our work with these works is our desired objective function ($\mathrm{O_{acc}}$ or $\mathrm{P_{err}}$)—and thus, the making the proof of monotonicity and/or submodularity of the objective function very different. In our case, we had to even circumvent the non-submodularity of the objective function $\mathrm{O_{acc}}$ by considering an appropriate auxiliary objective function.

Online Selection of Sensors. An alternate formulation of our sensor selection problem could be to select sensors adaptively *based* on the observations of previously selected sensors. This online problem is similar to the adaptive stochastic optimization problem addressed in other contexts [22], [23], [24], [25]. However, in online selection, a sensor is selected based on analysis (which will incur non-trivial latency) of observations of previous sensors. This makes localization based on *near-simultaneous* sensor observations, required to localize intermittent transmitters, infeasible. Also, note that online selection needs to be done anew for each localization, which may be performed very frequently (e.g., every second or fraction of a second) in many applications, e.g., spectrum patrolling. Thus, our focus is on offline selection.

VI. Conclusion

In this work, we consider the hypothesis-driven approach for localization of transmitters, and develop techniques to optimize the localization accuracy under a constraint of limited resources. Developed techniques are shown to yield provably approximate solutions. Our work can be instrumental in maximizing the network lifetime of a spectrum monitoring and/or patrolling system. Our future work focusses on improving our theoretical performance guarantee results, and developing similar sensor selection approximation algorithms for other localization approaches that are not hypothesis-driven.

Appendix A: Proof of Theorem 1

Let $\mathbf{T}$ be a given subset of sensors. For simplicity and without any loss of generality, let use assume that (i) the

JPD for H_0 has a zero mean, and (ii) the variance of the JPDs for both H_0 and H_1 is same (=Σ). T hus, the JPD for H_0 is $N(0, \Sigma)$ and for H_1 is $N(\mathbf{p}, \Sigma)$, where $N(\mu, \sigma)$ is a normal distribution with mean μ and variance σ, p is the vector (one dimension for each sensor in $\mathbf{T}$) of means, and Σ is the covariance matrix. To prove the theorem, we will show the following for a given set of sensors $\mathbf{T}$:

1) $\mathrm{P}_{\mathrm{err}}(\mathbf{T}) = Q(\frac{1}{2}\sqrt{\mathbf{p}^T\Sigma^{-1}\mathbf{p}})$, where $\mathbf{p}^T$ is the *transpose* of the p vector and $Q()$ is the tail function.
2) $\mathrm{O}_{\mathrm{acc}}(\mathbf{T}) = 1 - \mathrm{P}_{\mathrm{err}}(\mathbf{T})$ is montone and submodular.

The theorem follows easily from the above, as it is well known that greedy algorithms for a monotone and submodular objective function yield a 63% approximation [11].

Expression for $\mathrm{P}_{\mathrm{err}}(\mathbf{T})$. We start with computing $\mathrm{P}_{\mathrm{err}}(\mathbf{T}|H_0)$, i.e., the probability that MAP picks H_1 (i.e., $P(H_1|\mathbf{x}) > P(H_0|\mathbf{x})$) when the prevailing hypothesis in H_0, based on an observation vector $\mathbf{x}$ from $\mathbf{T}$. We get:

$$\begin{aligned} P(H_1|x) &= \frac{1}{\sqrt{2\pi\Sigma|}}\exp[-\frac{1}{2}(\mathbf{x}-\mathbf{p})^T\Sigma^{-1}(\mathbf{x}-\mathbf{p})] \\ P(H_0|x) &= \frac{1}{\sqrt{2\pi|\Sigma|}}\exp[-\frac{1}{2}\mathbf{x}^T\Sigma^{-1}\mathbf{x}] \end{aligned}$$

Now, the expression $P(H_1|\mathbf{x}) > P(H_0|\mathbf{x})$ is equivalent to $P(H_1|\mathbf{x})/P(H_0|\mathbf{x}) > 1$ which simplifies to:

$$\mathbf{x}^T\Sigma^{-1}\mathbf{p} > \frac{1}{2}\mathbf{p}^T\Sigma^{-1}\mathbf{p}. \tag{10}$$

We are interested in computing the probability of above expression being true, given H_0. In essence, given H_0, we want to compute:

$$\mathrm{P}_{\mathrm{err}}(\mathbf{T}|H_0) = P(\mathbf{x}^T\Sigma^{-1}\mathbf{p} > \frac{1}{2}\mathbf{p}^T\Sigma^{-1}\mathbf{p})$$
$$= P(\frac{\mathbf{x}^T\Sigma^{-1}\mathbf{p}}{\sqrt{\mathbf{p}^T\Sigma^{-1}\mathbf{p}}} > \frac{1}{2}\sqrt{\mathbf{p}^T\Sigma^{-1}\mathbf{p}}) = Q(\frac{1}{2}\sqrt{\mathbf{p}^T\Sigma^{-1}\mathbf{p}}),$$

where Q is the tail function of the standard Gaussian distribution. The last equation follows, since in H_0, the expression $\frac{\mathbf{x}^T\Sigma^{-1}\mathbf{p}}{\sqrt{\mathbf{p}^T\Sigma^{-1}\mathbf{p}}}$ has a mean of 0 and a standard deviation of 1. Similarly, we can show that $\mathrm{P}_{\mathrm{err}}(\mathbb{T}|H_1) = Q(\frac{1}{2}\sqrt{\mathbf{p}^T\Sigma^{-1}\mathbf{p}})$, since variance of the JPDs in H_0 and H_1 is same (Σ). Thus, $\mathrm{P}_{\mathrm{err}}(\mathbf{T}) = Q(\frac{1}{2}\sqrt{\mathbf{p}^T\Sigma^{-1}\mathbf{p}})$, and the localization accuracy $\mathrm{O}_{\mathrm{acc}}(\mathbf{T})$ is:

$$\mathrm{O}_{\mathrm{acc}}(\mathbf{T}) = 1 - \mathrm{P}_{\mathrm{err}}(\mathbf{T}) = 1 - Q(\frac{1}{2}\sqrt{\mathbf{p}^T\Sigma^{-1}\mathbf{p}}) \tag{11}$$

$\mathrm{O}_{\mathrm{acc}}$ () is Monotone and Submodular. First, we note that the value of $\mathbf{p}^T\Sigma^{-1}\mathbf{p}$ increases monotonically with "growth" (more dimensions) in $\mathbf{p}$. Now, since $Q(z)$ is a monotonically decreasing function, we get $\mathrm{O}_{\mathrm{acc}}(\mathbf{T}) > \mathrm{O}_{\mathrm{acc}}(\mathbf{T} \cup \{s\})$. For submodularity, we note that $Q(z)$ is continuous, differentiable, with $\frac{d^2Q(z)}{dz^2} > 0, \forall z > 0$. Thus, the rate of reduction of $Q(z)$ reduces with an increase in z. Thus, $\mathrm{O}_{\mathrm{acc}}$ is submodular.

Appendix B: Proof of Lemma 1

We prove the lemma in three parts.

$\mathrm{O}_{\mathrm{aux}}(\mathbf{T}) \leq \mathrm{O}_{\mathrm{acc}}(\mathbf{T})$. This directly follows from an application of Boole's inequality [26] which states that the probability of a union of events is never greater than the sum of the probabilities of individual events. In particular, by Boole's inequality, we have for all i:

$$P(\cup_{j\neq i}\mathtt{MAP}_{ij} = j|H_i) \leq \sum_{j\neq i} P(\mathtt{MAP}_{ij} = j|H_i) \tag{12}$$

Then, by multiplying each by $P(H_i)$, summing over all i, subtracting each side from 1, and noting that $\sum_i P(H_i) = 1$, we get $\mathrm{O}_{\mathrm{aux}}(\mathbf{T}) \leq \mathrm{O}_{\mathrm{acc}}(\mathbf{T})$ using Eq (8) and Eq (9).

$\mathrm{O}_{\mathrm{acc}}(\mathbf{T}) \leq 1 - \frac{1}{k}(1 - \mathrm{O}_{\mathrm{aux}}(\mathbf{T}))$. To get this, we utilize the fact that the probability of a union of events is more than the probability of each of the individual events. Thus,

$$P(\cup_{j\neq i}\mathtt{MAP}_{ij}(\mathbf{x}) = j|H_i) \geq \max_{j\neq i}\{P(\mathtt{MAP}_{ij}(\mathbf{x}) = j|H_i)\} \ \forall i.$$

We also have the below, as maximum is greater than mean:

$$\max_{j\neq i}\{P(\mathtt{MAP}_{ij}(\mathbf{x}) = j|H_i)\} \geq \frac{1}{m}\sum_{j\neq i} P(\mathtt{MAP}_{ij}(\mathbf{x}) = j|H_i) \quad \forall i,$$

where $0 \leq i \leq m$. Now, using Eq (8) and the above two equations, we get:

$$\begin{aligned} \mathrm{O}_{\mathrm{acc}}(\mathbf{T}) &\leq 1 - \frac{1}{m}\sum_{i=0}^{m}\sum_{j\neq i} P(\mathtt{MAP}_{ij}(\mathbf{x}) = j|H_i)P(H_i) \\ &= 1 - \frac{1}{m}(1 - \mathrm{O}_{\mathrm{aux}}(\mathbf{T})). \end{aligned}$$

The lemma now follows from the following fact, whose proof we omit for lack of space.

$$\frac{1 - \mathrm{O}_{\mathrm{aux}}(T')}{1 - \mathrm{O}_{\mathrm{acc}}(T')} \leq \frac{1 - \mathrm{O}_{\mathrm{aux}}(T)}{1 - \mathrm{O}_{\mathrm{acc}}(T)}, \quad \text{for any} \quad T' \supseteq T$$

Appendix C: Independent Sensor Observations

From Theorem 1's proof and notations therein, note that Eq (9) can be written as:

$$\mathrm{O}_{\mathrm{aux}}(\mathbf{T}) = 1 - \sum_i\sum_{j\neq i} Q((\mathbf{p}_j - \mathbf{p}_i)\Sigma^{-1}(\mathbf{p}_j - \mathbf{p}_i)^T)P(H_i), \tag{13}$$

where $Q(x)$ denotes the Gaussian Q-function [27]. Now, suppose we wish to compute $\mathrm{O}_{\mathrm{aux}}(\mathbf{T} \cup \{s_k\})$ for a sensor s_k whose observations have a mean of p_{ki} for hypothesis H_i and a variance is σ_k^2. We denote the argument of $Q()$ in Eq (9) by $\mathbf{q}_{ij}(\mathbf{T})$. Then, we have the following recurrence relation:

$$\mathrm{O}_{\mathrm{aux}}(\mathbf{T}\cup\{s_k\}) = 1 - \sum_i\sum_{j\neq i} Q(\mathbf{q}_{ij}(\mathbf{T}\cup\{s_k\}))P(H_i) \tag{14}$$

We note that computing $\mathbf{q}_{ij}(\mathbf{T})$ directly using Eq (13) takes $O(B^2)$ time. However, we can reduce the time complexity by computing $\mathbf{q}_{ij}(\mathbf{T})$ incrementally in constant time by using:

$$\mathbf{q}_{ij}(\mathbf{T} \cup \{s_k\}) = \mathbf{q}_{ij}(\mathbf{T}) + (p_{ki} - p_{kj})/\sigma_k^2.$$

As computing the Q-function takes constant time, the above reduced the time complexity by a factor of $O(B^2)$.

ACKNOWLEDGMENTS

This work is supported by the NSF grant CNS-164296 and CNS-1815306.

REFERENCES

[1] A. Bhattacharya, A. Chakraborty, S. R. Das, H. Gupta, and P. M. Djurić, "Spectrum patrolling with crowdsourced spectrum sensors," *IEEE Transactions on Cognitive Communications and Networking*, (accepted) 2019. doi: 10.1109/TCCN.2019.2939793

[2] M. Khaledi, M. Khaledi, S. Sarkar, S. Kasera, N. Patwari, K. Derr, and S. Ramirez, "Simultaneous power-based localization of transmitters for crowdsourced spectrum monitoring," in *Proceedings of the 23rd Annual International Conference on Mobile Computing and Networking - MobiCom*, 2017. doi: 10.1145/3117811.3117845 p. 235–247.

[3] A. Nika, Z. Li, Y. Zhu, Y. Zhu, B. Y. Zhao, X. Zhou, and H. Zheng, "Empirical validation of commodity spectrum monitoring," in *Proceedings of the 14th ACM Conference on Embedded Network Sensor Systems - SenSys*, 2016. doi: 10.1145/2994551.2994557 p. 96–108.

[4] A. Chakraborty, M. S. Rahman, H. Gupta, and S. R. Das, "SpecSense: Crowdsensing for efficient querying of spectrum occupancy," in *IEEE INFOCOM - IEEE Conference on Computer Communications*, may 2017. doi: 10.1109/infocom.2017.8057113 pp. 1–9.

[5] V. Cevher, P. Boufounos, R. G. Baraniuk, A. C. Gilbert, and M. J. Strauss, "Near-optimal bayesian localization via incoherence and sparsity," in *International Conference on Information Processing in Sensor Networks - IPSN*, Apr 2009, pp. 205–216.

[6] C. Zhan, H. Gupta, A. Bhattacharya, and M. Ghaderibaneh, "Efficient localization of multiple intruders for shared spectrum system," in *International Conference on Information Processing in Sensor Networks - IPSN*, Apr 2020.

[7] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern classification*. John Wiley & Sons, 2012.

[8] P. K. Penumarthi, A. Q. Li, J. Banfi, N. Basilico, F. Amigoni, J. O'Kane, I. Rekleitis, and S. Nelakuditi, "Multirobot exploration for building communication maps with prior from communication models," in *International Symposium on Multi-Robot and Multi-Agent Systems (MRS)*, Dec 2017. doi: 10.1109/mrs.2017.8250936 pp. 90–96.

[9] "NSF workshop on spectrum measurements infrastructure," p. 22–23, 2016.

[10] "RTL-SDR (RTL2832U) and software defined radio news and projects," https://www.rtl-sdr.com/, 2018, accessed on Oct 18, 2018.

[11] G. L. Nemhauser, L. A. Wolsey, and M. L. Fisher, "An analysis of approximations for maximizing submodular set functions—i," *Mathematical programming*, vol. 14, no. 1, 1978. doi: 10.1007/BF01588971

[12] S. Khuller, A. Moss, and J. S. Naor, "The budgeted maximum coverage problem," *Information Processing Letters*, vol. 70, no. 1, pp. 39 – 45, 1999. doi: 10.1016/S0020-0190(99)00031-9

[13] J. A. Magliacane, "SPLAT! a terrestrial rf path analysis application for linux/unix," Downloaded from https://www.qsl.net/kd2bd/splat.html, 2008.

[14] K. Chamberlin and R. Luebbers, "An evaluation of Longley-Rice and GTD propagation models," *IEEE Transactions on Antennas and Propagation*, vol. 30, no. 6, pp. 1093 – 1098, 1982. doi: 10.1109/TAP.1982.1142958

[15] N. Patwari, "CRAWDAD dataset utah/cir (v. 2007-09-10)," Downloaded from https://crawdad.org/utah/CIR/20070910, Sep. 2007.

[16] S. K. Lam, A. Pitrou, and S. Seibert, "Numba: A LLVM-based Python JIT Compiler," in *Proceedings of the Second Workshop on the LLVM Compiler Infrastructure in HPC*, 2015. doi: 10.1145/2833157.2833162 pp. 1 – 6.

[17] R. K. Martin and R. Thomas, "Algorithms and bounds for estimating location, directionality, and environmental parameters of primary spectrum users," *IEEE Transactions on Wireless Communications*, vol. 8, no. 11, pp. 5692–5701, 2009. doi: 10.1109/TWC.2009.090494

[18] F. Zafari, A. Gkelias, and K. K. Leung, "A survey of indoor localization systems and technologies," *IEEE Communications Surveys Tutorials*, vol. 21, no. 3, pp. 2568–2599, thirdquarter 2019. doi: 10.1109/COMST.2019.2911558

[19] A. Yassin, Y. Nasser, M. Awad, A. Al-Dubai, R. Liu, C. Yuen, R. Raulefs, and E. Aboutanios, "Recent advances in indoor localization: A survey on theoretical approaches and applications," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 2, pp. 1327–1346, 2017. doi: 10.1109/comst.2016.2632427

[20] H. Rowaihy, S. Eswaran, M. Johnson, D. Verma, A. Bar-Noy, T. Brown, and T. La Porta, "A survey of sensor selection schemes in wireless sensor networks," in *Unattended Ground, Sea, and Air Sensor Technologies and Applications IX*, vol. 6562. International Society for Optics and Photonics, 2007.

[21] A. Krause, A. Singh, and C. Guestrin, "Near-optimal sensor placements in gaussian processes: Theory, efficient algorithms and empirical studies," *Journal of Machine Learning Research*, vol. 9, 2008.

[22] H. Wang, K. Yao, G. Pottie, and D. Estrin, "Entropy-based sensor selection heuristic for target localization," in *International Conference on Information Processing in Sensor Networks - IPSN*, 2004. doi: 10.1145/984622.984628 pp. 36–45.

[23] Y. Chen, S. H. Hassani, A. Karbasi, and A. Krause, "Sequential information maximization: When is greedy near-optimal?" in *Proceedings of The 28th Conference on Learning Theory*, vol. 40, 03–06 Jul 2015, pp. 338–363. [Online]. Available: http://proceedings.mlr.press/v40/Chen15b.html

[24] A. Krause and C. Guestrin, "Near-optimal nonmyopic value of information in graphical models," in *Proceedings of the Twenty-First Conference on Uncertainty in Artificial Intelligence*, ser. UAI'05, Arlington, Virginia, USA, 2005, p. 324–331.

[25] D. Golovin and A. Krause, "Adaptive submodularity: Theory and applications in active learning and stochastic optimization," *Journal of Artificial Intelligence Research*, vol. 42, no. 1, p. 427–486, Sep. 2011.

[26] E. Seneta, "On the history of the strong law of large numbers and boole's inequality," *Historia Mathematica*, vol. 19, no. 1, pp. 24 – 39, 1992. doi: 10.1016/0315-0860(92)90053-E

[27] M. Abramowitz, I. A. Stegun *et al.*, "Handbook of mathematical functions with formulas, graphs, and mathematical tables," 1972.