

Sensor Selection from Independence Graphs using Submodularity

Thomas Powers
Dept of Electrical Engineering
University of Washington
Seattle, WA 98195-2500
Email: tcpowers@uw.edu

David W. Krout
Applied Physics Laboratory
University of Washington
Seattle, WA 98105-6698
Email: dkrou@apl.washington.edu

Les Atlas
Dept of Electrical Engineering
University of Washington
Seattle, WA 98195-2500
Email: atlas@uw.edu

Abstract—In this paper we develop a framework to select a subset of sensors from a field in which the sensors have an ingrained independence structure. Given an arbitrary independence pattern, we construct a graph that denotes pairwise independence between sensors, which means those sensors can operate simultaneously. The set of all fully-connected subgraphs (cliques) of this independence graph can form a set of matroid constraints over which we maximize a submodular objective function. Since we choose the objective function to be submodular, the algorithm returns a near-optimal solution with approximation guarantees. We also argue that this framework generalizes to any network with a defined independence structure between sensors, and intuitively models problems where the goal is to gather information in a complex environment. We apply this framework to ping sequence optimization for active multistatic sonar arrays by maximizing sensor coverage and not only achieve significant performance gains compared to conventional round-robin sensor selection, but approach optimal performance as well.

I. INTRODUCTION

Sensor selection problems, or more generally subset selection problems, are important for many applications in areas such as machine learning and wireless communications [1]–[4]. This paper addresses general sensor selection, with a specific focus on sensor networks with interfering sensors, and we demonstrate improved performance on active multistatic sonar arrays [5]. As is the case with most optimization problems, it would be advantageous for the problem to be convex. However, framing sensor selection problems as convex has two main problems. First, sensor selection has inherently discrete optimization variables since selecting a sensor is an absolute choice. One cannot “half-choose” a sensor, and convex optimization requires continuous variables. Second, convex optimization is unable to handle dependence constraints between variables. There is no way to enforce dependent values between variables in a convex framework, which means if a pair of sensors interfere, convex optimization provides no guarantee that the two sensors will not be present in the solution together.

Submodular function optimization (SFO) provides a more intuitive framework for handling these two problems, since it inherently uses set functions and can be constrained to optimize over matroids. Matroids are a structure that generalize the notion of linear independence from vector spaces to set

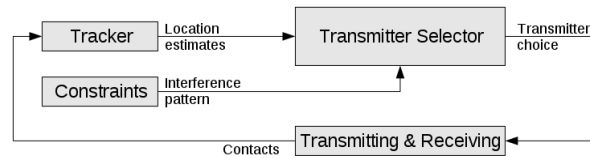


Fig. 1: General system diagram for sensor selection.

systems, and can be used to form constraints in SFO. They will be addressed further in Section II.

The main focus of this paper is modeling sensor networks as graphs and then using the graph structure to form matroid constraints in submodular function optimization. SFO can handle constraints that make problems nonconvex or non-polynomial (NP) hard and find polynomial time solutions that are provably optimal or are near-optimal with performance guarantees [6]. As we will show in Section II, modeling sensor network interference patterns as an independence graph can be folded directly into SFO as a series of matroid constraints [7], [8]. We apply this approach to active multistatic deep-water sonar arrays, or ping sequence optimization (PSO), in which we repeatedly optimize to select a subset of buoys that maximize a probabilistic coverage metric. We detail this application in Section III. In order to demonstrate the superior utility of the proposed approach, called SFO-Greedy, we compare the performance of SFO with matroid constraints to the standard round-robin sensor selection, single buoy, and exhaustive search approaches in Section IV.

II. SUBMODULARITY AND INDEPENDENCE GRAPHS

The binary nature of sensor selection makes optimization difficult. Typically, one represents the sensor nodes in an indicator vector with a selected sensor node as ones and unselected sensor nodes as zeros. These independence constraints make optimization problems nonconvex. One of the main contributions of this paper is modeling independence constraints on the sensor networks.

Submodularity is a property that describes set functions similar to how convexity describes functions in a continuous space. For ping sequence optimization, submodular functions can be used to find optimal subsets of buoys to achieve

objectives like maximizing coverage of non-interfering buoys, or maximizing probability of target detection in a target tracking scenario. Rather than exhaustively searching over all combinations of subsets, submodular functions provide a fast and tractable framework to compute a solution [6], [9], [10].

Let the set of available objects, known as the ground set, be denoted as V . A submodular function f maps a set of objects denoted by a binary indicator vector of length V to a real number. The binary indicator vector is represented by the expression 2^V since the variable can take two values and has length V . As mentioned previously, a value of 1 or 0 for the i^{th} element of the indicator vector denotes the inclusion or exclusion of the i^{th} element of the ground set V . Therefore, we can define a submodular function f with the following inequality.

Definition: A function $f : 2^V \rightarrow \mathbb{R}$ is submodular if for any $A, B \subseteq V$

$$f(A) + f(B) \geq f(A \cup B) + f(A \cap B) \quad (1)$$

Note that inequality is very similar to the definition of a convex function, and in fact, submodularity can be viewed as a discrete analog to convexity [9].

More intuitively, submodularity can be expressed by the notion of diminishing returns. This means that the incremental value of the objective function shrinks as more elements of the ground set are added. Drawn out for a particular sequence of elements, the objective function looks like either a concave or convex function sampled at equal intervals. An alternate but equivalent definition is as follows.

A function $f : 2^V \rightarrow \mathbb{R}$ is submodular if for any $A \subseteq B \subseteq V$ and $v \in V \setminus B$

$$f(A \cup \{v\}) - f(A) \geq f(B \cup \{v\}) - f(B) \quad (2)$$

This is the notion of diminishing returns.

Submodularity is very closely tied to structures known as matroids, which generalize the notion of linear independence in vector spaces [7]. One can think of matroids as a generalization of matrices, which extend the definition of rank beyond column vectors to arbitrary independent subsets of a ground set. More importantly, SFO allows for matroid constraints on the problem, which means you can encode complicated variable dependence patterns into the problem and obtain polynomial time solutions. The pair $(V, \mathcal{I})$ is a matroid if the family of sets $\mathcal{I}$ satisfies the following three properties:

- 1) $\emptyset \in \mathcal{I}$
- 2) $I_1 \subseteq I_2 \in \mathcal{I}$
- 3) $I_1, I_2 \in \mathcal{I}, |I_1| < |I_2| \implies \exists v \in I_2 \setminus I_1 : I_1 \cup v \in \mathcal{I}$

This leads us to the independence graphs, where nodes on the graph represent sensors and edges denote pairwise independence between sensors. An edge between two sensors, in other words, means both sensors can be used at the same time. For this setup, the nodes in any fully connected subgraph (clique) is an allowable subset. The set of all cliques from this independence graph G can form a partition matroid if the stable sets of the complimentary graph $\bar{G}$ form a partition [11].

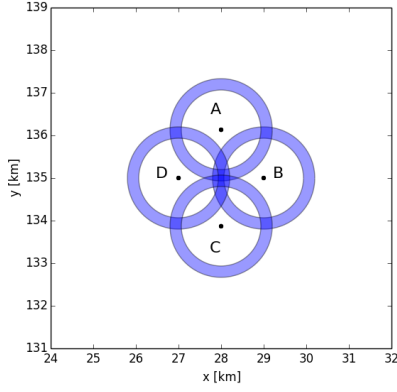
Unfortunately, this is not true in general for the independence graphs generated from the interference patterns in a sensor field. In these cases, the set of all cliques can be represented by a set of n matroids, where n equals the number of maximum size cliques of the graph G . To prove this, consider a maximum size clique C in G . Let the independent set be $\mathcal{I} = \{I | I \subseteq C\}$ and the ground set be C . $\mathcal{I}$ satisfies the first property because the empty set is a subset of every set. $\mathcal{I}$ satisfies the second property because every subset of C is included in the definition. It is easy to see that $\mathcal{I}$ satisfies the third property, which is also referred to as the basis exchange property. Consider two subsets $C_1, C_2 \subseteq C$ such that $|C_1| < |C_2|$. Adding any element $i \in C_2 \setminus C_1$ to C_1 will necessarily be a subset of C since $i \in C$ and $C_1 \subseteq C$. Therefore, $(C, \mathcal{I})$ is a matroid, and n maximum size cliques of a graph can be converted into n matroids. An example of the independence graph is in Fig. 2d. By turning the interference pattern of a sensor field into a set of matroid constraints, we can guarantee that two interfering sensors will not be chosen in the solution.

III. APPLICATION TO PING SEQUENCE OPTIMIZATION

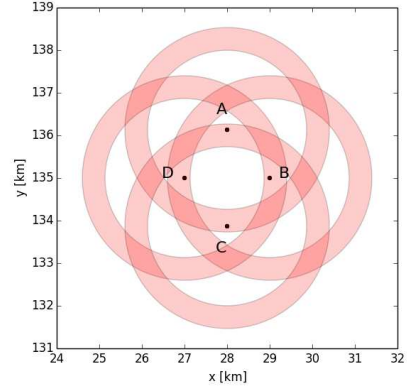
We apply this sensor selection framework to active sonar arrays, where each buoy has a co-located transmitter and receiver that operates monostatically. However, since SFO allows for multiple buoys to be selected, the array functions multistatically in that multiple receivers are operating simultaneously and at potentially overlapping regions. An example of a spatial buoy arrangement where some of the buoys interfere can be found in Figs. 2a, 2b, and 2c. The four buoys are arranged in a diamond pattern with locations represented by black dots. In Fig. 2a, the blue rings denote the coverage regions for each buoy and the red rings in Fig. 2b denote the regions where another buoy will interfere with a given buoy. Coverage is defined by the probability of target detection for a buoy. If two interfering buoys transmit simultaneously, the direct path signal from the first will arrive at the second when the second buoy's reflections would arrive. The relationship between the coverage and interference regions for the buoys can be found in Fig. 2c. In this arrangement, the buoys across from each other, i.e. the top and bottom pair and left and right pair, will interfere with each other, since the buoys in each pair are in the red interference region of the other buoy. However, any other pair of buoys can ping simultaneously [5].

We go beyond the conventional approach for buoy selection by allowing for simultaneous pinging. Specifically, we can select buoys based on target state that significantly improve the system's ability to track existing targets and search for new targets. We demonstrate the increase in performance compared to a conventional round-robin approach and optimized single buoy selection. In addition, our approach allows for simultaneous search and track objectives within the system.

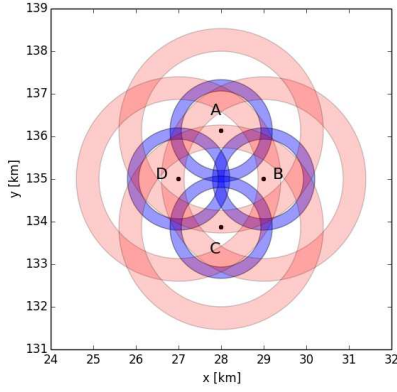
In order to find out the maximum number of buoys that can ping simultaneously, the largest set of nodes is picked such that all the nodes in the set are connected to every node in the set. Note that self-loops are implied, since a



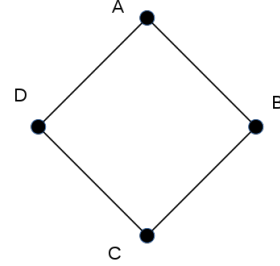
(a) Four buoys (black dots) with coverage regions (blue rings).



(b) Four buoys (black dots) with interference regions (red rings).



(c) Four buoys (black dots) with coverage regions (blue rings) and interference regions (red rings).



(d) Independence graph for four buoys as nodes with edges that signify pairwise independence.

Fig. 2: Relationship between the independence graph, coverage regions and interference regions for a four buoy arrangement.

buoy does not interfere with itself. The problem of finding the largest subset of fully connected nodes is a well known problem in computer science [12]. Exact methods for solving this problem run in exponential time, but for reasonable graph sizes (a hundred vertices), the algorithm runs fairly quickly. For example, if the graph meets certain conditions, i.e. if the graph is “planar” or “perfect,” finding the largest clique can be solved in polynomial time [13]. For the arrangement in Fig. 2c, there is a four-way tie for largest clique, which are the adjacent pairs (top and left buoys, left and bottom buoys, bottom and right buoys, and right and top buoys). The independence graph for this arrangement is depicted in Fig. 2d. In a real scenario, the detection regions will not be perfect rings, so one of the pairs might have better coverage than the others. A more complicated interference pattern will emerge as the number of buoys is increased, which is demonstrated in Fig. 3.

Our objective function is a variant of probabilistic coverage. It utilizes target state estimates to help determine which buoys are selected. Let V be the set of N buoys $b_i, i = 1 \dots N$. Let $B \subseteq V$ such that B is a clique of G , where G is the

independence graph determined by the interference pattern of all the buoys b_i in V . Let the set of all sets of sensors that form cliques on the graph be a partition matroid $\mathcal{I}$. Coverage is a positive, non-decreasing objective, so the goal is to maximize the objective function. Then the optimal set of buoys is given by

$$B^* = \operatorname{argmax}_{B \in \mathcal{I}} \frac{1}{M} \sum_{\phi=1}^M f_{\phi}(B) \quad (3)$$

where $\phi = 1, \dots, M$ corresponds to the predicted target locations and M is the number of targets. The functions $f_{\phi} : 2^V \rightarrow \mathbb{R}$ are given by the equation

$$f_{\phi}(B) = 1 - \prod_{b_i \in B} (1 - P_{\phi, b_i}) \quad (4)$$

where P_{ϕ, b_i} is the probability of detection of buoy b_i at location ϕ determined by a table look-up for pre-computed probability of detection maps for each buoy.

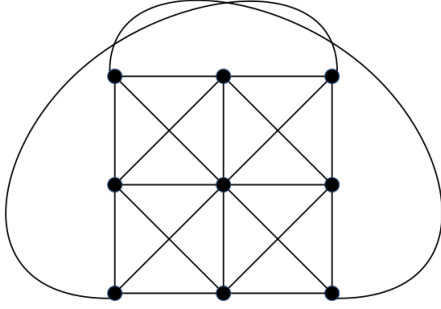


Fig. 3: Independence graph for the nine buoys in a 3x3 grid.

We use the greedy algorithm from the SFO toolbox to solve the above optimization problem [10]. This submodular objective function is monotonic non-decreasing and subject to a matroid constraint. Submodular maximization for functions of this form have been studied and have certain performance guarantees. For a monotonic non-decreasing objective function subject to a matroid constraint, the solution has a worst case performance bound of $\frac{1}{2}$, and the bound scales with the number of matroids, i.e. for k matroid constraints the worst case performance bound is $\frac{1}{k+1}$ [14].

Our approach is tracking-centric in that the objective prioritizes covering areas where known targets are located, but it provides good coverage as well. After the algorithm addresses coverage of the known targets, it adds as many non-interfering buoys as are available, and thus provides an effective simultaneous track and search framework.

IV. EXPERIMENTAL RESULTS

In this section we compare the performance of the proposed ping sequence optimization algorithm, SFO-Greedy, to the standard round-robin approach, optimized single buoy selection, and exhaustive search in a Monte Carlo simulation. For the experiment, we have nine buoys in a grid pattern, with each buoy 60 km away from its neighbors. The independence graph for the buoys can be found in Fig. 3, and the interference pattern in Fig. 4. We assume the buoys have a probability of detection of $P = 0.8$ in the coverage region and $P = 0$ everywhere else. We assume here that there is no sensor drift during the experiment. Two targets with random initial location, constrained to be within the buoy array's detection area, and constant velocities are present for each trial. The experiment consisted of ten thousand trials, and each trial lasted thirty time-steps or until a target moved out of the array's detection area.

For each trial, we initialize the target location and velocity and pass the initial state estimates into the SFO algorithm. Based on the objective function output for each target, we sample the probability that each target has a successful detection at the next time step and pass in the updated state estimates for the detected targets. For the round-robin algorithm, we simply choose the next buoy in the sequence and calculate its objective function value, and for the optimized single buoy

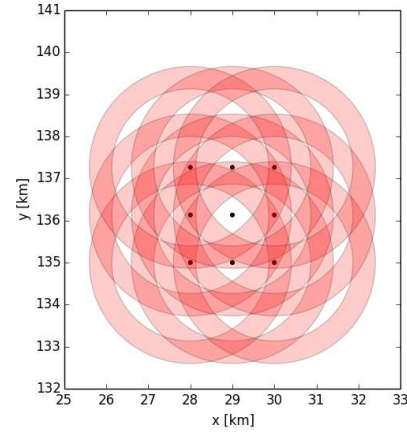


Fig. 4: Interference pattern for the nine buoys in a 3x3 grid.

selection, we greedily choose the buoy that provides the best coverage according to the objective function in (3). We also compare the proposed algorithm to an exhaustive search over all possible combinations of non-interfering buoys to see how close to optimal the greedy approximation is.

Over the course of the trial, we accumulate the objective function values which form a cumulative probability of detection (CPD) score for the two algorithms. The mean probability of detection (PD) scores for each algorithm can be found in Table I, and a plot of the PD scores over the different trials is found in Fig. 5.

The results in Table I and Fig. 5 demonstrate the practical utility of the proposed method. A round-robin approach to sensor selection detected the targets about 7% of the time on average, whereas the proposed SFO-Greedy approach detected the targets over 30% of the time. In addition, the greedy algorithm performed nearly as well as an exhaustive search over all sets of non-interfering buoys. The results also demonstrate the advantage of using multiple buoys. By allowing simultaneous pinging, we gained 50% better coverage over optimized single buoy selection. The size of the error-bars in Fig. 5 can be attributed to the fact that nine buoys have gaps in their overall coverage, so many simulated targets were impossible to detect given the arrangement. However, in no trial did the round-robin approach beat the proposed algorithm. The worst case bound for the greedy algorithm provides a nice floor for worst-case behavior, but in practice the algorithm is as good as exhaustively searching over exponentially growing sets of non-interfering buoys.

TABLE I

Method	Single Buoy		Multiple Buoys	
	Round-robin	SFO-Greedy (Proposed)	Exhaustive search	SFO-Greedy (Proposed)
Mean PD	0.068	0.210	0.320	0.315

Probability of detection (PD) results for Monte Carlo simulation.

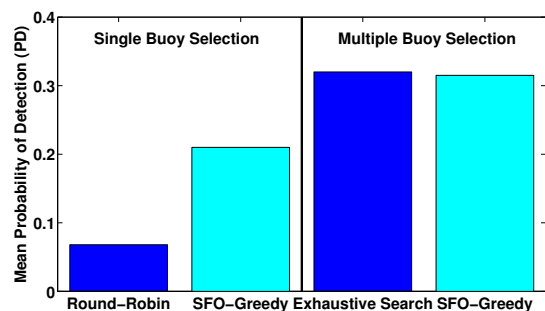


Fig. 5: Probability of detection (PD) results for Monte Carlo simulation comparing the proposed SFO-Greedy method to a round-robin strategy for single buoy selection, and SFO-Greedy to exhaustive search for multiple buoy selection.

V. CONCLUSION

In applying independence graphs to a sensor selection problem, we demonstrate the utility of submodular function optimization (SFO) to the problem domain. Specifically for ping sequence optimization (PSO), SFO allows us go beyond the standard approach for buoy selection by allowing for simultaneous pinging. By posing the PSO as a submodular optimization problem, we are able to derive near-optimal solutions for combinatorial problems. We can select buoys based on target state that significantly improve the probability of detecting targets over a standard approach and achieve equivalent performance to an optimal exhaustive search approach. Moreover, our approach allows for simultaneous search and track objectives within the system.

We also demonstrate the theoretical advantages of SFO over convex optimization for sensor selection. Carefully framing sensor selection problems as convex has two main problems: the inability to handle discrete optimization variables and independence constraints. Submodular function optimization provides a more intuitive framework for handling these two problems. It inherently uses set functions and can be constrained to optimize over matroids, which can be used to encode complex independence patterns between sensors. Not only can our proposed use of submodular function optimization handle complex constraints, but it provides guaranteed near-optimal polynomial time solutions to combinatorial problems.

ACKNOWLEDGMENT

This work was funded by the U.S. Office of Naval Research, Contract Number N00014-12-G-0078, Delivery Order #24, and the ARO grant number W911NF1210277. We acknowledge the helpful discussions with Scott Wisdom and Rishabh Iyer of the University of Washington Department of Electrical Engineering.

REFERENCES

- [1] R.K. Iyer and J.A. Bilmes, "Submodular optimization with submodular cover and submodular knapsack constraints," in *Advances in Neural*

- Information Processing Systems* 26, pp. 2436–2444. Curran Associates, Inc., 2013.
- [2] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *Communications magazine, IEEE*, vol. 40, no. 8, pp. 102–114, 2002.
- [3] A. Krause, B. McMahan, Guestrin C., and Gupta A., "Robust submodular observation selection," *Journal of Machine Learning Research (JMLR)*, vol. 9, pp. 2761–2801, December 2008.
- [4] S. Joshi and S. Boyd, "Sensor selection via convex optimization," *Signal Processing, IEEE Transactions on*, vol. 57, no. 2, pp. 451–462, Feb 2009.
- [5] D.W. Krout and T. Powers, "Sensor management for multistatics," in *Information Fusion (FUSION), 2014 17th International Conference on*, July 2014, pp. 1–6.
- [6] S. Fujishige, *Submodular functions and optimization*, vol. 58, Elsevier, 2005.
- [7] J. Edmonds, "Matroids and the greedy algorithm," *Mathematical Programming*, vol. 1, no. 1, pp. 127–136, 1971.
- [8] G. Calinescu, C. Chekuri, M. Pál, and J. Vondrák, "Maximizing a submodular set function subject to a matroid constraint," in *Integer programming and combinatorial optimization*, pp. 182–196. Springer, 2007.
- [9] L. Lovsz, "Submodular functions and convexity," in *Mathematical Programming The State of the Art*, Achim Bachem, Bernhard Korte, and Martin Grtschel, Eds., pp. 235–257. Springer Berlin Heidelberg, 1983.
- [10] A. Krause, "SFO: A toolbox for submodular function optimization," *J. Mach. Learn. Res.*, vol. 11, pp. 1141–1144, Mar. 2010.
- [11] K. Kashiwabara, Y. Okamoto, and T. Uno, "Matroid representation of clique complexes," *Discrete Applied Mathematics*, vol. 155, no. 15, pp. 1910–1929, 2007.
- [12] C. Bron and J. Kerbosch, "Algorithm 457: finding all cliques of an undirected graph," *Communications of the ACM*, vol. 16, no. 9, pp. 575–577, 1973.
- [13] M.R. Garey, D.S. Johnson, and L. Stockmeyer, "Some simplified np-complete graph problems," *Theoretical computer science*, vol. 1, no. 3, pp. 237–267, 1976.
- [14] M.L. Fisher, G.L. Nemhauser, and L.A. Wolsey, "An analysis of approximations for maximizing submodular set functions-ii," in *Polyhedral Combinatorics*, M.L. Balinski and A.J. Hoffman, Eds., vol. 8 of *Mathematical Programming Studies*, pp. 73–87. Springer Berlin Heidelberg, 1978.