Joint Sensing Duty Cycle Scheduling for Heterogeneous Coverage Guarantee

Kin Sum Liu\textsuperscript{*}, Tyler Mayer\textsuperscript{†}, Hao Tsung Yang\textsuperscript{*}, Esther Arkin\textsuperscript{‡}, Jie Gao\textsuperscript{*}, Mayank Goswami\textsuperscript{†}, Matthew P. Johnson\textsuperscript{§}, Nirman Kumar\textsuperscript{¶}, Shan Lin\textsuperscript{∥}

\textsuperscript{*}Department of Computer Science, Stony Brook University. \{kiliu,haotyang,jgao\}@cs.stonybrook.edu
\textsuperscript{†}Dept. of Applied Math. & Stats., Stony Brook University. tyler.mayer@stonybrook.edu, \{estie,jsbm\}@ams.stonybrook.edu
\textsuperscript{‡}Department of Computer Science, Queens College, CUNY. mayank.goswami@qc.cuny.edu
\textsuperscript{§}Department of Computer Science, Lehman College, CUNY. mpjohnson@gmail.com
\textsuperscript{¶}Department of Computer Science, University of Memphis, nkumar8@memphis.edu
\textsuperscript{∥}Department of Electrical and Computer Engineering, Stony Brook University. shan.x.lin@stonybrook.edu

Abstract—In this paper we study the following problem: given a set of \( m \) sensors that collectively cover a set of \( n \) target points with heterogeneous coverage requirements (target \( j \) needs to be covered every \( f_j \) slots), how to schedule the sensor duty cycles such that all coverage requirements are satisfied and the maximum number of sensors turned on at any time slot is minimized. The problem models varied real-world applications in which sensing tasks exhibit high discrepancy in coverage requirements – critical locations often need to be covered much more frequently. We provide multiple algorithms with best approximation ratio of \( O(\log n + \log m) \) for the maximum number of sensors to turn on, and bi-criteria algorithm with \((\alpha, \beta)\)-approximation factors with high probability, where the number of sensors turned on is an \( \alpha = O\left(\frac{\log(n^{\alpha}) + \log(m^{\beta})}{\beta} \right) \)-approximation of the optimal (satisfying all requirements) and the coverage requirement is a \( \beta \)-approximation; \( \delta \) is the approximation ratio achievable in an appropriate instance of set multi-cover. When the sensor coverage exhibits extra geometric properties, the approximation ratios can be further improved. We also evaluated our algorithms via simulations and experiments on a camera tested. The performance improvement (energy saving) is substantial compared to turning on all sensors all the time, or a random scheduling baseline.

I. INTRODUCTION

Sensor deployment and scheduling has been a fundamental problem for sensor network applications. The complexity of the problem comes from the variety of sensing modalities and sensor coverage ranges (ranging from omnidirectional disk coverage [22] to line-of-sight visibility coverage [23], [31]), the application coverage requirement (full/blanket coverage [5], [33], full-view coverage [38], \( k \)-coverage [42], bar-rier coverage [32], etc), and networking requirement (sensors remain connected by wireless radio [5], [23]). A substantial amount of work has been invested on the many variations of this problem in the past few years [16], [36].

In this work, we assume that the sensor deployment problem has been solved and a near optimal set of sensors has been placed to provide full coverage of the target locations. We instead look at the optimal duty cycle scheduling of the sensors. We may need a large number of sensors to provide full coverage — but is it possible that many sensors may simultaneously cover the same target, exhibiting a high level of redundancy in the system. Many sensors, such as cameras, may consume a large amount of energy to operate, or require a large amount of bandwidth to upload data. Thus efficient scheduling of the operations of such sensors can result in significant energy/communication bandwidth savings. Second, in many monitoring applications the target locations may require different levels of quality of service. In a building monitoring scenario for safety applications, certain locations (high activity areas) such as entrances to the building typically require continuous or highly frequent monitoring while other locations such as individual offices may only need to be checked once in a while as activity levels are low. This heterogeneity in service quality also provides an opportunity to schedule the sensors in a resource efficient way without hurting the overall application. This combined with the joint sensing/coverage features can provide opportunities for a smart scheduling to save substantially in energy/operational costs. See Figure 1 for an example.

![Figure 1: An example of \( n + 1 \) target points (shown as crosses) and \( n \) sensor nodes (shown as solid disks) with line-of-sight coverage. \( n \) sensors are needed to ensure full coverage of the targets. Suppose that the \( n \) targets on top each need to be covered once in every \( n \) slots (with \( f = n \)) while the bottom target needs to be covered every slot (with \( f = 1 \)). Alternatively turning on each of the \( n \) sensors would suffice, with savings of \( n \) times the energy consumption compared to keeping all sensors on all the time.](image)

We formulate our problem in the following manner. We are given a set of \( m \) sensors and \( n \) targets, where each sensor \( i \) covers a subset of targets \( S_i \). Further, we assume that time is slotted and in each time slot one may turn on only a subset of sensors. Each target location \( j \) has a minimum coverage requirement — it must be covered at least once for every \( f_j \)
slots. These \( f_j \)'s can vary significantly. The problem asks to minimize the maximum number of sensors turned on at any time slot such that all the coverage requirements are satisfied.

In a closely related work by Liu et al. [27] a budget of \( k \) sensors per slot is given, and the goal is to find a duty cycle schedule so that all target points are monitored sufficiently frequently. They considered two optimization objectives. The first is to minimize the maximum amount of time that any target point remains unseen (called the min-max version), and the second is to minimize the average time that any target point remains unseen (called the min-average version). They show for the min-max version, that if the sensor set is a minimal cover for the domain then the round robin schedule with each time slot having \( \lceil m/k \rceil \) different sensors active is an optimal schedule, otherwise they can find an \( O(\log n) \) approximation.

For the min-average variant they show how to produce a schedule which serves as a \((2 + \varepsilon)\alpha\)-approximation where \( \alpha = O(1) \) if \( k \geq \log m / \log \log m \) and \( \alpha = \log m / \log \log m \) otherwise.

The most significant difference in our problem formulation as compared to both the min-max and min-average version is that we strictly enforce the coverage quality requirement, which can be highly heterogeneous for different target locations. This is crucial for real-time applications with hard constraints on the coverage guarantee.

Our Results. We provide three algorithms for the joint duty cycle scheduling problem. First we show that restricting our solution to the family of periodic schedules can only suffer by a factor of two, and that finding the optimal periodic schedule is NP-hard. On the positive side, we have three different algorithms with different design ideas/approximation ratios.

- A combinatorial \( O(\log n + \log m) \)-approximation.
- An \( O\left(\gamma \log \left(\frac{f_{\text{max}}}{f_{\text{min}}}\right)\right) \) approximation in geometric instances where \( \gamma \) is the best approximation ratio achievable in the corresponding geometric set cover problem and \( f_{\text{max}}(f_{\text{min}}) \) is the largest (smallest) window size among all coverage requirements.
- An \( O(\gamma \log m) \)-approximation to the dual problem in geometric settings.
- An \( (\alpha = O(\frac{\delta \log n + \log m}{\beta}) \), \( \beta \) bi-criteria approximation algorithm where \( \alpha \) is an approximation to the optimal number of sensors required to maintain full coverage, and \( \beta \) is the approximation to the coverage requirement. In this case \( \delta \) is the best approximation achievable for the set multi-cover problem.

In simulations we observe that the three algorithms have their unique strength under different scenarios. If the frequency requirements follow a uniform distribution the combinatorial algorithm needs only 60% of sensors active compared to baseline random schedule. On the other hand, the geometric algorithm beats others and has only 36% of sensors active when the frequency requirements follow an exponential distribution. Experiments on a camera testbed verify the energy saving of our scheduling algorithms.

II. RELATED WORK

Our version of the duty cycle problem arises as a combination of a set multi-cover and a scheduling problem in sensor networks. In the set multi-cover problem one is given a universe \( U \) of \( n \) elements (with multiplicity) and a collection \( S \) of subsets of \( U \) and is asked to find a minimum cardinality sub-collection \( C \subseteq S \) which covers \( U \) (with, at least, the appropriate multiplicity for each element). This problem is clearly a generalization of set cover which has been shown NP-Hard and can’t be approximated to within a factor of \((1 - \varepsilon) \ln n \), \( \forall \varepsilon > 0 \), unless \( NP \subseteq DTIME(n^{O(\log \log n)}) \), by Feige [13]. Better approximations can be obtained in special cases. For example, [10] show a \( O(\log |OPT|) \) approximation when the set system has a constant VC dimension. The same approximation can be obtained for the set multi-cover problem [12]. For covering points by half-spaces in 3D, an \( O(1) \)-approximation is possible [12]. Bansal et al. [7] showed a randomized \( O(1) \)-approximation algorithm for the weighted geometric set multi-cover problem when the covering objects have linear union complexity. The set multi-cover problem using disks as sets has a constant factor approximation [8], [1], [9], as well as a QPTAS [6]. Altinel et al. [4] show that, for targets with uniform coverage requirements and sensor locations chosen from a regular grid with some density \( d \), an algorithm with \((6\pi, f(d))\) bi-criteria approximation which uses at most \( 6\pi \) times as many sensors from the regular grid coordinates as the optimal solution and covers at least \( f(d) \) percent of the total area where \( f(d) \) is a function based on the density chosen for the grid.

Given a fixed placement of sensors, the problem of scheduling these sensors to collaboratively cover the domain has been studied in the line of sight model [28], [19]. In visual sensor networks Munishwar et al. [30] study the problem of finding an angle of view for each camera to maximize the number of target points covered. They give an exact centralized algorithm which is suitable to run on small sized networks, as well as a decentralized heuristic for larger networks with a \( 1/2 \)-approximation guarantee.

The problem of near-optimal placement of sensors given a duty cycle schedule in order to obtain certain coverage requirements has been studied in [25], [23], [41]. Most of these heuristics are validated via numerical experiments.

One related paper by Feinberg et al. [14] studies a scheduling problem called the generalized pinwheel problem for a single server and a set of recurring jobs. Each job has a duration and a maximum allowable time between two consecutive runs. Authors show necessary conditions for the feasibility of the problem and present an exact algorithm (with exponential running time) based on dynamic programming as well as several heuristic approximations as the problem is NP-hard.

There is a huge literature on duty cycle scheduling in sensor networks. See [37] for a survey. The duty cycle problem has been studied in communication networks and energy management systems [34], [39], [18], [21]. Yoo et al. [40] study the problem of scheduling sensors with energy harvesting
capabilities based on the amount of remaining energy in the system, while Jaleel et al. [24] propose a probabilistic deployment algorithm based on a Poisson process and give upper bounds on the probability of coverage based on the energy decay function and spatial distribution of sensors over the network. These algorithms are largely non-collaborative, meaning that they don’t take into account the overlapping coverage ranges.

Several approaches to the duty cycle scheduling problem using integer linear programming (ILP) have also been explored. Meguerdichian et al. [29] solve the natural set cover ILP to find a minimum 0/1 set cover of the targets. They show how to find a minimum cover even for heterogenous coverage ranges.

Chakrabarty et al. [11] show how to use ILP to find the minimum cover of sensors from a regular grid of positions for coverage as well as target localization in the network. Gu et al. [17] show how a column generation approach to solving the associated ILP can be used to find a minimum cover even for heterogeneous coverage requirements of the targets. Although ILP methods can solve these problems exactly, they are often only feasible for modest problem sizes, as the class of problems are still NP-Complete.

III. JOINT DUTY CYCLE SCHEDULING

Given a set of \( m \) sensors \( S \) and a set of \( n \) targets \( U \) to monitor, sensor \( i \) covers a subset of targets \( S_i \subseteq U \). For each target \( j \), we are given the coverage requirement as the window size \( f_j \), such that for any time window of length \( f_j \) target \( j \) is covered at least once. Time is slotted and at any time slot we wish to only turn on a subset of the sensors. The problem is to design a duty cycle schedule such that all targets meet their coverage frequency requirement and the maximum number of sensors turned on at any time slot (which is referred to by \( k \)) is minimized.

This problem is clearly NP-hard as set cover is a special case when \( f_i = 1 \) for all \( i \). Before we describe approximation algorithms for this problem, we first present a few important observations on periodicity of the optimal schedule.

It is not clear that an optimal schedule will be periodic, but here we show that if we restrict ourselves to periodic schedules with a sufficiently large period, such schedule is within a constant factor 2 of the optimal schedule. Moreover, we may force the period to be no greater than \( m \), if we turn on only one extra sensor in every time slot on top of any periodic schedule.

We can model a schedule \( \mathcal{F} \) as a set of computable functions \( \mathcal{F} = \{ F_1, \ldots, F_m \} \) where each \( F_i : \mathbb{N} \rightarrow \{ \text{true}, \text{false} \} \) specifies if sensor \( i \) is to be turned on at time \( t \), i.e., \( F_i(t) = \text{true} \) iff sensor \( i \) is turned on at time \( t \) for any \( t \in \mathbb{N} \). Such a schedule is periodic with period \( T \), if each \( F_i \) is a periodic function of \( t \) with period \( T \), i.e., \( F_i(t+T) = F_i(t) \) for each \( i = 1, \ldots, m \) and any \( t \in \mathbb{N} \).

Let \( \mathcal{F}_{opt} \) denote the optimal schedule and \( \mathcal{F}_{T \_opt} \) denote the optimal schedule with period \( T \). For any schedule \( \mathcal{F} \) let \( k(\mathcal{F}) \) denote the maximum number of sensors turned on at any time slot, called the cost of the schedule:

\[
k(\mathcal{F}) = \max_{t \in \mathbb{N}} \{|i : F_i(t) = \text{true}|\}
\]

Let \( f_{max} = \max_{j=1,...,n} f_j \) be the largest window size. We have the following lemmas. Omitted proofs of this section can be found in Appendix A.

**Lemma 3.1.** For any \( T \geq f_{max} \), \( k(\mathcal{F}_{T \_opt}) \leq 2k(\mathcal{F}_{opt}) \).

From the above lemma, we can focus on finding a periodic schedule with period \( T = f_{max} \) such that we only lose an approximation factor of two in the cost. But \( f_{max} \) can still be fairly large, e.g., exponential in \( m \) or \( n \). The lemma below shows we can assume the period is \( m \) (the number of sensors), at the cost of turning on only one more sensor in each slot, on top of any periodic schedule of period of \( f_{max} \).

**Lemma 3.2.** \( k(\mathcal{F}_{opt}^m) \leq 2k(\mathcal{F}_{opt}) + 1 \).

**Proof:** Separate the targets \( U \) based on their minimum window size into two sets \( U', U'' \) where \( U' = \{ j : f_j \leq m \} \) and \( U'' = \{ j : f_j > m \} \). Let \( \mathcal{F}_{opt} \) be the optimal schedule for \( U' \) and \( \mathcal{F}_{opt}^m \) be the optimal schedule for \( U'' \). Obviously, \( k(\mathcal{F}_{opt}^m) \leq k(\mathcal{F}_{opt}) \). From Lemma 3.1, we know that \( k(\mathcal{F}_{opt}^m) \leq 2k(\mathcal{F}_{opt}) \).

To cover \( U'' \), we can turn on one extra sensor in each time slot by simply doing round robin on the \( m \) sensors. This suffices to cover all targets in \( U'' \) with window size greater than \( m \). In fact, these targets are covered more frequently.

Put together, the new schedule has period of \( m \) and cost at most \( 2k(\mathcal{F}_{opt}) + 1 \). \( \square \)

Therefore, in the rest of the paper we will assume without loss of generality that \( f_{max} \leq m \) and will focus on periodic schedules with period at most \( m \). This restriction only induces a constant factor.

(LEY: prove the periodic schedule can be optimal)

1) **New lemma:**

**Lemma 3.3.** For any optimal schedule \( \mathcal{F}_{opt} \), there exists a periodic schedule with finite length \( W \) which is also optimal. That is, \( k(\mathcal{F}_{opt}^W) = k(\mathcal{F}_{opt}) \).

**Proof:** To help for the following proof, we firstly show a way to express a schedule into a \( T \times m \) matrix with its schedule length \( T \). For example, \( \mathcal{F}^T \) can be expressed as a matrix \( M \) with the following:

\[
M = \begin{bmatrix}
F_1(1) & F_2(1) & F_3(1) & \ldots & F_m(1) \\
F_1(2) & F_2(2) & F_3(2) & \ldots & F_m(2) \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
F_1(T) & F_2(T) & F_3(T) & \ldots & F_m(T)
\end{bmatrix}
\]
We also define $M[a,b]$ is the submatrix by including from row $a$ to row $b$ of $M$.

Now, suppose $J_{opt}$ is non-periodic and its matrix expression is $A$, which has infinite rows. Denote $H$ is a schedule with period $f_{max}$ that each function is exactly the same from time 1 to time $f_{max}$. That is, $\forall t \leq t \leq f_{max}, H(t) = F(t)$, and $H = \{H_1, H_2, \ldots, H_m\}$. $H$ can be expressed by a submatrix of $A$, which is exactly $A[1, f_{max}]$. In addition, since $m$ is finite and each entry in $A$ only has two values (true, false), there are only finite number of $f_{max} \times m$ matrices, which means there must be another submatrix $A[x,x+f_{max}-1]$ for some integer $x$ such that $A[1, f_{max}] = A[x, x + f_{max} - 1]$. Denote $S$ is a periodic schedule which is expressed by a matrix $B$ with $B = A[1, x-1]$. For any target $j$ in the schedule $S$, consider a window starting at time $t$ and having size $f_j$. Since $f_j \leq f_{max} < x-1$, the corresponding schedule of $S$ during this period from $t$ to $t+f_j-1$ is also a part of schedule in $J_{opt}$ for any $t$. Thus, the coverage requirement are also satisfied in $S$, which means $S$ is also optimal.

\[ \square \]

IV. COMBINATORIAL ALGORITHM

In this section, we construct a periodic schedule $F^m$ of period length $m$. The algorithm is for the purely combinatorial setting, i.e., the set of targets covered by a sensor can be arbitrary. We do not make any assumptions on sensing range and the result applies in the most general setting.

The main challenge lies in how to extend the concept of coverage to the temporal domain. For each target $j$, we partition one cycle of $m$ slots into $m/\lfloor f_j/2 \rfloor$ intervals of $\lfloor f_j/2 \rfloor$ time slots each. For each interval $b$, introduce a new element $e_{j,b}$ which represents the target in the time frame $b$. The total number of elements is $\sum_{j=1}^{n} \lfloor f_j/2 \rfloor = O(nm)$. Let $E$ be the universe of these new elements.

Sensors are placed in the $m$ slots to create a periodic schedule $F^m$. If a sensor $i$ is placed in slot of time frame $b$, and $i$ covers target $j$, we say the element $e_{j,b}$ is covered. Then, if our schedule covers all target $j$'s corresponding elements $\{e_{j,b} : b \in \{1, \ldots, m/\lfloor f_j/2 \rfloor\}\}$, it is clear that $j$ is covered at least once in every window of size $f_j$ – since every such window will fully contain at least one interval.

Now we are ready to formulate a new set cover problem. We work with the family $V$ of vectors of dimension $m$. Each vector $V \in V$, $V = \langle i_1, i_2, \ldots, i_m \rangle$, corresponds to sensor $i$, being turned on at time slot $t$, for $1 \leq t \leq m$. We choose from this family $V$ a vector and use it as one layer of the periodic schedule. Eventually we choose $k$ vectors of $V$, overlay them along the temporal domain, to create a schedule with cost of $k$.

We want to minimize $k$ such that all elements $\{e_{j,b} : \forall j, \forall b\}$ are covered.

Formally, in this set cover $\{S(V), V \in V\}, E$ problem, each set in the collection $V$ is a subset of elements that are covered by the vector $V$, denoted by $S(V) \subseteq E$, $V \in V$. We wish to choose a minimum number of such vectors, such that the elements they collectively cover is $E$.

Lemma 4.1. $F^m$ constructed from the optimal solution $C^*$ of the set cover Problem $\{S(V), V \in V\}, E$ is a $2$-approximation to $F^m_{opt}$.

Proof: $F^m$ is the best $m$-period schedule to cover every element of $E$. Clearly any schedule that covers all elements in $E$ satisfies all the coverage window requirements. This establishes the correctness of the solution $F^m$.

For the approximation factor of $F^m$, note that the optimal solution $F^m_{opt}$ that satisfies all the coverage window requirements, may not cover all elements of $E$. The latter is a stronger, sufficient but not necessary condition of the former.

To establish the approximation factor, we take $F^m_{opt}$ and make it a new schedule $F^{m/2}$ by merging the sensors in the $(2t-1)$-th and $2t$-th time slot as the sensors turned on in the $t$-th time slot in $F^m_{opt}$, for each $1 \leq t \leq m/2$. We have two observations of $F^{m/2}$:

- $k(F^{m/2}) = 2k(F^m_{opt})$.
- $F^{m/2}$ will have at least one sensor that covers target $j$ in each window of size $\lfloor f_j/2 \rfloor$, for each $j$. Therefore, $F^{m/2}$ covers all elements in $E$. By the optimality of $F^m$, we know $k(F^m) \leq k(F^{m/2})$.

Combining the two we have $k(F^m) \leq 2k(F^m_{opt})$. \[ \square \]

Running the classical greedy algorithm for set cover on this new instance $\{S(V), V \in V\}, E$ immediately gives an $O(\log n + \log m)$-approximation. The issue is that there are exponentially many sets. $|V| = m^m$, since in each time slot we may have $m$ choices of sensors. Therefore at each step, we cannot afford to enumerate over all remaining vectors in $V$ to select the one that covers the maximum number of new elements of $E$. However, it turns out that this greedy selection does not need to be done optimally. [20] shows that if we use a pseudo-greedy implementation that picks a $\beta$-approximate best set in each step (i.e., we choose a set that covers at least $\beta h$ new elements if the best set can cover $h$ new elements), such pseudo-greedy cover yields an $O(\frac{1}{\beta} \ln |E|)$-approximation.

Below we show how to do this pseudo-greedy step. In particular, to find the next vector $V \in V$, we run another greedy algorithm to gradually fill in the sensors for all the $m$ slots in vector $V$. The next sensor we turn on is the one that covers the maximum number of new elements in $E$. We continue this iteration until all slots of $V$ are filled up. The entire algorithm is summarized in the pseudo code below.

Lemma 4.2. The pseudo-greedy step gives a $\frac{1}{2}$-approximation, i.e., the number of elements covered is at least half of the maximum number of elements that can be covered by the optimal vector $V$.

Proof: We observe that our problem of finding the best vector is an instance of Problem (1.5) in Fisher’s analysis [15] with $P = 1$. Specifically, our objective can be re-stated as a matroid-constrained sub-modular maximization:

$$\max_{X \subseteq E} \{z(X) : X \in I, M = (E, I) \text{ a matroid}, z(X) \text{ sub-modular and non-decreasing}\}.$$
Algorithm 1 Combinatorial Algorithm

Inputs: Universe $\mathcal{E}$
Output: Periodic Schedule $\mathcal{F}^m$

1: $\mathcal{C} = \emptyset$
2: while not all $e_{j,b} \in \mathcal{E}$ is covered do
3: $k \leftarrow k + 1$
4: while not all $m$ time slots are used do
5: choose $(r, t)$ pair that covers the most remaining elements, where sensor $r$ is turned on at an empty time slot $t$
6: $i_t \leftarrow r$
7: end while
8: $\mathcal{C} = \mathcal{C} \cup S(< i_1, ..., i_m >)$
9: end while
10: construct $\mathcal{F}^m$ from $\mathcal{C}$

where

$E = \{(i, t) : i \in \mathcal{S}, t \in \mathbb{N}_{\leq m}\}$

$I = \{A : A \subseteq \{(i_1, 1), ..., (i_m, m) : i_1, ..., i_m \in \mathcal{S}\}\}$

$z(X) = z(I(\{(x_1, 1), ..., (x_m, m)\})) = |S(< x_1, ..., x_m >)|$

Here $E$ is the set of all possible allocations of individual time slots to a sensor and $I$ is the collection of all possible (partial) allocations of $m$ time slots. $\mathcal{M}$ is a matroid because $z(\cdot)$ is a cardinality function which is clearly sub-modular and non-decreasing. For the hereditary property, if $X \subseteq Y$ and $Y \in I$, then $X \in I$ from our construction of the independent sets $I$. For the independent set exchange property, if $X, Y \in I$, and $|Y| > |X|$ (so $X$ must be a partial allocation), then we must be able to find a time slot $t_e$ allocated in $Y$ but not in $X ((\cdot, t_e) \notin X)$ and assigning that to $X$ is still feasible. So we have $\exists (i_e, t_e) \in Y \setminus X, X \cup (i_e, t_e) \in I$.

From this observation, we know that our algorithm is equivalent to using greedy heuristic for Problem (1.5) in [15] and we can directly apply Theorem (2.1) to get an approximation factor of $\frac{3}{2}$. □

Theorem 4.3. The algorithm has an approximation factor $O(\log n + \log m)$.

From Lemma 4.2, our algorithm picks a $\beta$-approximate best vector with $\beta = \frac{1}{2}$. Plugging in $\beta = \frac{1}{2}$ into Corollary 2 of [20], Algorithm 1 is a $O(2 \log |\mathcal{E}|)$-approximation, i.e., $O(\log n + \log m)$-approximation.

V. GEOMETRIC ALGORITHM

The algorithm we discussed so far does not make any assumptions on the coverage range of the sensors. In many scenarios the coverage ranges have some geometric properties which can be leveraged to obtain better approximation ratios.

For example, consider the following problem: among the sensors in $\mathcal{S}$, choose a minimum number of them to cover all targets in $\mathcal{U}$. As mentioned, this problem in general cannot be approximated better than $(1 - \varepsilon) \log n$, $\forall \varepsilon > 0$. However, if the sensor coverage ranges have constant VC-dimension, the approximation ratio can be improved to $O(\log |OPT|)$ where $|OPT|$ is the number of sets in the optimal solution [10]. If the sensor coverage ranges are axis-parallel rectangles or disks (of possibly different radii), the set cover problem can be solved with approximation of $O(\log \log |OPT|)$ or $O(1)$ respectively [3]. If the sensor coverage ranges are unit disks, then a simple greedy algorithm gives a constant approximation [2] and a PTAS exists [26]. In the following we denote by $\gamma$ the best approximation ratio one can obtain for the particular instance of geometric set cover.

In this section, we will present a different algorithm that can benefit from results in geometric set cover problems.

Uniform Window Size. As an appetizer we look at the special case when all the targets have the same window size in coverage requirement, i.e., $f_j = f$, for all $j$. Define the optimal set cover solution as $C^* \subseteq \mathcal{S}$ that covers all targets. We can find a geometric set cover solution $C$ with size $|C| \leq \gamma|C^*|$ for the universe $\mathcal{U}$ with $n$ targets and $\gamma$ is the approximation ratio.

Now we have the following scheduling algorithm: Do round robin on $C$ with $k = \lceil |C|/f \rceil$ sensors. The period of this schedule is at most $f$, which satisfies the coverage frequency requirement.

We argue that the optimal schedule has to use $k^* \geq \lceil |C^*|/f \rceil$. This is because all the targets must be covered at least once in one full period, thus the set of sensors that appear in one full period is a set cover solution, with size at least $|C^*|$. Thus in this simple setting we find a sensor scheduling with approximation ratio of $\gamma$.

General Setting. When the targets have different coverage window sizes $f_j$, we take $f_{\min} = \min_j f_j$ and $f_{\max} = \max_j f_j$. Clearly, turning on $\lceil |C|/f_{\min} \rceil$ at each time slot is a feasible solution and a lower bound is to turn on $\lceil |C^*|/f_{\max} \rceil$ sensors.

Now we partition the targets into $\log R$ sets, where $R = f_{\max}/f_{\min}$, such that in each set $\mathcal{G}_i$, the frequencies of the targets do not differ by more than a factor of 2. Take $h_i = \min_{j \in \mathcal{G}_i} f_j$ and $H_i = \max_{j \in \mathcal{G}_i} f_j$, $H_i \leq 2h_i$. To cover the set of targets in $\mathcal{G}_i$, let $C_i$ be the set cover solution and $C^*_i$ be the optimal set cover solution. $|C_i| \leq \gamma|C^*_i|$, where $\gamma$ is the approximation factor in the geometric set cover problem.

For each set $\mathcal{G}_i$, we will use round robin and turn on $k_i = \lceil |C_i|/h_i \rceil$ sensors at each time slot, which is sufficient to guarantee that all targets are covered frequently enough. The total number of sensors we need to turn on is $k = \sum_i k_i$.

Theorem 5.1. The algorithm has an approximation factor $O(\gamma \log R)$, where $R = f_{\max}/f_{\min}$.

Proof: Just to cover the group of sensors $\mathcal{G}_i$, we have a lower bound $k^* \geq \lceil |C^*_i|/H_i \rceil$. Thus

$k_i = \lceil |C_i|/h_i \rceil \leq \gamma|C^*_i|/h_i \leq 2\gamma|C^*_i|/H_i \leq 2k^* \gamma$.

There are $\log R$ groups. This finishes the proof. □

Discussion. The approximation ratio of this algorithm depends on $R$, the ratio of the largest window size and the smallest
window size. Recall that we can assume that $f_{\text{max}}$ is at most $m$ with only a sacrifice of a factor of 2 in Lemma 4.1, thus we may assume that $\log R = O(\log m)$. In the case when the coverage ranges are disks this approximation ratio is no worse than that of the combinatorial algorithm. But when $R$ is much smaller, i.e., this algorithm can be better than the combinatorial algorithm. We tested a variety of geometric setting in our simulations and indeed it shows that in the geometric setting this algorithm starts to outperform the combinatorial algorithm.

VI. DUAL AND BICRITERIA ALGORITHMS

In this section, we study the tradeoff between coverage requirement and resource constraints. In the original problem we fix the coverage frequency requirement and minimize $k$ (the maximum number of sensors to turn on in a slot). Here we would like to study two other variations:

- The dual problem: we fix $k$ (the number of sensors one can turn on in each slot) and minimize the factor $\beta$, by which we stretch the window size of the coverage requirement. That is, each target $i$ is covered at least once in every window of size $\beta f_j$, with $\beta \geq 1$. We denote this as the $\beta$-relaxed coverage requirement. We wish to minimize $\beta$.
- $(\alpha, \beta)$-approximation: we turn on at most $\alpha k^*$ sensors in each slot such that each target $i$ is covered at least once in every window of size $\beta f_j$, ($\beta \geq 1$), where $k^*$ is the optimal solution for the original problem (when the coverage window size is $f_j$ for target $j$).

A. The Dual Problem

In the dual problem we look at the case when $k$ is given and we may not be able to meet the requirements for all targets. Thus we consider stretching each window size $f_j$ by a factor $\beta$ and we hope to minimize $\beta$.

We use the same partition of targets into groups $G_i$ as in Section V. Now for a total quota of $k$ sensors to turn on at each time slot, we give a quota of $k/\log R$ to each group $G_i$. The stretch $\beta_i$ beyond the requirement is bounded by

$$\beta_i \leq \frac{k_i}{k/\log R} = \frac{k_i \log R}{k}.$$  

Let the stretch factor in the optimal solution be $\beta^*$. Consider only group $i$, neglecting other targets, and allocate the entire quota of $k$ sensors for monitoring targets in $G_i$. This solution over $G_i$ is no larger than the optimal solution. Thus a lower bound on the stretch factor over group $G_i$ is a lower bound for $\beta^*$. By this argument we have

$$\beta^* \geq \frac{|C_i^*|}{h_i k}.$$  

Re-arrange, we have

$$\beta \leq \frac{k_i \log R}{k} \leq \frac{|C_i^*| O(\gamma \log R)}{h_i k} \leq \beta^* O(\gamma \log R).$$

Thus, we have the following theorem.

Theorem 6.1. The dual problem, can be approximated within a factor of $O(\gamma \log R) = O(\gamma \log m)$, where $\gamma$ is the approximation ratio of geometric set cover and $R = f_{\text{max}}/f_{\text{min}}$ is the aspect ratio of the window size.

B. Bicriteria Approximation

For an $(\alpha, \beta)$-bi-criteria approximation we relate the duty cycle scheduling problem to the set multi-cover problem. The set multi-cover problem is an extension of the set cover problem. Given a set of targets $U$, we want to choose a multi-set $C$ (i.e., a sensor can appear multiple times) with sensors from $S$ with minimum number of sensors such that target $j$ must be covered at least $r_j$ times. $r_j$ is called the coverage multiplicity. In our case, we choose $r_j = f_{\text{max}}/f_j$.

The optimal solution to this set multi-cover problem is denoted as $C_{\text{SMC}}^*$. We can run essentially the same greedy algorithm to find a solution to the set multi-cover problem denoted as $C_{\text{SMC}}$: choose the next sensor $i$ that maximizes the utility, defined as the number of targets covered by $i$ among those targets whose coverage multiplicity has not been met. By [35], we know that $|C_{\text{SMC}}|/|C_{\text{SMC}}^*| \leq \delta$, where $\delta$ is the best approximation ratio achievable for a particular instance of set multi-cover. $\delta = O(\log m)$ in general, and in the geometric case [12] show one can obtain a $O(\log (|OPT|))$ for set systems with bounded VC-Dimension, and an $O(1)$-approximation when covering points in 3D with half-spaces, and an $O(\log \log |OPT|)$-approximation for covering points by triangles in the plane.

Lemma 6.2 (Lower Bound on $\text{opt}$). If we require each target to be covered at least once in every window of size $f_j$, in some time slot we need to turn on at least $[C_{\text{SMC}}^*/f_{\text{max}}]$ sensors. That is, $k^* \geq [C_{\text{SMC}}^*/f_{\text{max}}]$.

Proof: Consider the optimal schedule that satisfies the coverage frequency requirement and consider all the sensors turned on in a contiguous window of $f_{\text{max}}$ slots, denoted by a multi-set $Y^*$. We know that $Y^*$ will cover each target at least $r_j = f_{\text{max}}/f_j$ times. Thus $Y^*$ is a solution to the set multi-cover problem. This says that $|Y^*| \geq |C_{\text{SMC}}^*|$. By pigeon hole principle, there is at least one slot in which $|Y^*|/f_{\text{max}}$ sensors are turned on. This gives a lower bound on $k^*$:

$$k^* \geq \frac{|Y^*|}{f_{\text{max}}} \geq \frac{|C_{\text{SMC}}^*|}{f_{\text{max}}}.$$  

Now we are ready for the scheduling algorithm that makes use of the set multi-cover solution $C_{\text{SMC}}$. Define $z = |C_{\text{SMC}}|$. We take $q$ random permutations $\pi_i$ of $C_{\text{SMC}}$, for $i = 1, \cdots, q$. For each of the permutation $\pi_i$, we stretch it (or shrink it) to fit in a schedule of period $f_{\text{max}}$:

- If $z \leq f_{\text{max}}$, we spread the elements of $\pi_i$ uniformly on $f_{\text{max}}$ slots. The first element in $\pi_i$ is turned on in each of the first $f_{\text{max}}/z$ slots, the second element in $\pi_i$ is turned on in each of the second $f_{\text{max}}/z$ slots and so on.
- If $z > f_{\text{max}}$, we turn on multiple sensors in one slot. The first $z/f_{\text{max}}$ elements in $\pi_i$ are turned on in the first slot, the second $z/f_{\text{max}}$ elements in $\pi_i$ are turned on in the second slot, and so on.
A. Simulations

VII. Simulations & Experiments

A. Simulations

We compare our algorithms to two baseline algorithms, Base_random, and Base_heuristic in geometric simulation. The baseline algorithms can be divided into two steps. First, for each target \( j \), generate a periodic schedule with length \( f_j \) that satisfies the target frequency requirement. The sensor chosen for each schedule is different: in Base_random, we randomly pick a sensor that covers this target, whereas Base_heuristic, selects the sensor which covers the most uncovered targets including this one. To avoid congestion in specific time slots, the schedule of each target randomly starts between 1 to \( f_j \). Second, merge all schedules to create an overall schedule that satisfies frequency requirements of all targets.

We randomly generate 1000 targets and 64 sensors for the simulation. Each target is covered by at least one of the sensors. Each sensor is located at a point randomly generated in the plane. A sensor covers a target if the distance between them is at most one. Each target \( j \) gets randomly assigned to a particular sensor \( i \) and the location of the target is then generated randomly within the unit disk centered at sensor \( i \). We introduce a parameter \( \delta \) such that the frequency requirement of each target is chosen between 1 and \( 2^\delta \). The values of window sizes follow two different frequency distributions: uniform or exponential. In the uniform setting we sample the window size from 1 to \( 2^\delta \) with uniform probability; in the exponential setting the probability of choosing a larger window size is exponentially higher than others. That is, for a specific target \( j \), \( g \in \mathbb{N}_{\leq \delta} \) and

\[
\text{Prob}\{2^{g-1} \leq f_j \leq 2^g\} = 2^{-g} \cdot (2^{g-1} - 2^g).
\]

Figure 2 shows a comparison between the algorithms. The \( y \)-axis represents the average of 50 simulations of the maximum number of sensors turned on in any time slot (\( k \)) in each setting. We run this experiment for \( \delta = \{2, 3, ..., 6\} \). For fairness we set \( \beta = 1 \) and \( q = 8 \) to ensure coverage with high probability (where \( q \) is the number of random permutations we overlay) in the bi-criteria algorithm.

Our results show that the combinatorial algorithm performs the best under a uniform frequency distribution in Figure 2a. However, if the frequency distribution is exponential, the geometric algorithm improves the most when the window size is more skewed to larger values. The reason is that the geometric algorithm has a smaller hidden constant factor than the combinatorial one and the effective grouping benefits (more targets with larger window size are grouped into the same cluster).

B. Experiment

We have implemented a testbed to evaluate the algorithms. This testbed has four wireless camera nodes, and each of them is built based on the off-the-shelf BeagleBone low power development board. A 3.1 Megapixel Aptina CMOS digital image sensor MT9T111 and a USB Edimax EW-7811Un Wi-Fi adapter are plugged into the board. Events are emulated by an LED. The LED is controlled by a Raspberry Pi 2 development board.

For this experiment we set each time slot to be five seconds. Our schedule decides which cameras take a picture at each five seconds. Each LED \( j \) is a target and is randomly turned on for \( f_j \) slots multiple times. We check whether the cameras can capture every event. Our testbed is shown in Figure 3. Camera 1 and 3 cover three targets each, \( \{a, b, c\} \), and \( \{e, f, g\} \)
In this paper we propose three algorithms for duty cycle scheduling of sensors with different design principles. They leverage properties and structures of the domain to construct efficient schedules with theoretical bounds while remain simple enough for implementation. The flexible formulation of coverage makes it appealing for many different applications in wireless sensor networks. In future work, we will explore the extension of this formulation to connectivity and load balancing.

Fig. 2: Performance of algorithms in geometric domain against upper bound of maximum window size ($\delta$). Since every target must be covered once for every window, a larger window size allows it to be covered less frequently so the maximum number of active sensors in each time slot is reduced. All algorithms benefit when the targets are required to be covered less frequently.

Fig. 3: Experiment testbed of covering targets using a camera network. There are 4 cameras (rectangles) with line-of-sight coverage and 7 point targets (represented by red crosses).

respectively, and camera 2 covers \{b, c, d, e, f\}. Camera 4 only covers three targets \{b, c, d\}.

Generally, our three algorithms generate schedules with similar $k$. However, if we set targets with small window size ($f_g$ equals to 1 and others equal to 2) the bi-criteria algorithm obtains $k = 2$, while others obtain $k = 3$. The reason is that the geometric algorithm clusters targets with window size = 1, 2 in the same group. Also, the combinatorial algorithm requests each target $j$ to be covered within each $f_j/2$ time slots, which is also every slot. Therefore, they both cover all targets for each time slot which needs three sensors. On the other side, bi-criteria algorithm avoid this situation by taking directly the set multi-cover solution. For this simple setting, our scheduling algorithms reduce the number of active sensors in each time slot thus saving energy in the network.

Fig. 4: Transforming a computable schedule to a periodic one. The example has $T = 4$. The disk, cross etc. represent different sensors, and their appearance represents the time slot they are scheduled. The interval [11, 14] corresponds to a target covered by a sensor shown as a empty box. The corresponding interval [3, 6] is also covered.

**APPENDIX A**

**PROOFS FOR SECTION III**

**Lemma 3.1.** For any $T \geq f_{\text{max}}$, $k(F^T_{\text{opt}}) \leq 2k(F_{\text{opt}})$.

**Proof:** Let $F_{\text{opt}}$ denote the optimal schedule, and $F_{1,\text{opt}}, \ldots, F_{n,\text{opt}}$ denote the corresponding functions. We construct functions $F_1, \ldots, F_n$ each with period $T$ such that the cost of the corresponding schedule is at most 2 times that of the optimal schedule specified by $F_{\text{opt}}$. Our definition is as follows: For $t = 1, \ldots, T$ we define $F_i(t) = F_{i,\text{opt}}(t) \vee F_{i,\text{opt}}(t+T)$. Let $F^T$ denote the corresponding periodic schedule i.e., where we have extended these functions $F_i$ over the natural numbers periodically, i.e., for $t \in [\gamma T + 1, (\gamma + 1)T]$ we have $F_i(t) = F_i(t - \gamma T)$ for $\gamma \geq 1$. As we show later this schedule is valid (i.e., it satisfies the required coverage of the targets). For now, we show that its cost is at most 2 times the cost of optimal. See Figure 4 for an illustration of how to construct the periodic schedule from a given schedule.
For any \( t \) in \( \{1, \ldots, T\} \), sensor \( i \) is turned on iff it is turned on in \( F_{opt} \) in either of times \( t, t+T \). Thus the number of sensors turned on at any point in time is at most the sum of the number of sensors turned on at time \( t, t+T \) which is at most \( 2k(F_{opt}) \).

Consider now any target \( j \). Let \( S_j \subseteq \mathbb{N}_{\leq m} \) be the indices of the sensors covering target \( j \). Consider an interval of length \( f_j \). Since \( T \geq f_{max} \geq f_j \) such an interval can intersect at most two of the intervals \( \{1, T\}, \{T+1, 2T\}, \ldots \). If such an interval lies inside \( [\gamma T + 1, (\gamma + 1)T] \) then suppose it extends from \( [\gamma T + x, \gamma T + x + f_j - 1] \) where \( x \geq 1 \). Now consider the corresponding interval \([x, x + f_j - 1]\). Since this interval is satisfied for target \( j \) in the optimal schedule we have that at least one of the sensors with index in \( S_j \) is turned on during the times \([x, x + f_j - 1]\). By definition of \( F \) such a sensor will also be turned on once during the time interval \( [\gamma T + x, \gamma T + x + f_j - 1] \). Now, consider an interval that intersects at least two periods, see Figure 4 for an example. Suppose it extends from \( [\gamma T - f_j + x, \gamma T + x] \). In this case consider the interval \([\gamma T - f_j + x, \gamma T + x + f_j - 1] \) in the optimal schedule. A sensor in \( S_j \) is turned on at least once during this interval. Suppose this was sensor \( i \) turned on during some time in \([T - f_j + x, T]\) then as per the definition of \( F \) it is turned on during the corresponding interval in \( F \) as well. On the other hand, if it was turned on during \([T + 1, T + x]\) then, it will be turned on during \([1, x]\) in \( F \) (since \( F_i(t) = F_i(t) \vee F(t) \) for \( t \) in \( \{1, \ldots, T\} \), and by periodicity also during \([T + 1, T + x]\). So a sensor in \( S_j \) will be turned on at least once during \([T - f_j + x, T]\) By periodicity, that sensor will be turned on during \([\gamma T - f_j + x, \gamma T + x]\). This proves the validity of our periodic schedule. \( \square \)

Acknowledgement: E. Arkin would like to acknowledge NSF support through CCF-1526406. K. Liu, H. Yang and J. Gao would like to acknowledge the support through NSF DMS-1418255, CCF-1535990, CNS-1618391 and AFOSR FA9550-14-1-0193. K. Liu, H. Yang and S. Lin are grateful for support through DE-EE0007682, NSF CNS 1553273, CNS 1536086, CNS 1463722, and IIS 1460370.

REFERENCES


