Distance and Time Based Node Selection for Probabilistic Coverage in People-Centric Sensing

Asaad Gad Elrab, Keiichi Yasumoto, Yukiko Yamauchi and Minoru Ito
Graduate School of Information Science, Nara Institute of Science and Technology (NAIST)
Email: {asaad-g, yasumoto, y-yamauchi, ito}@is.naist.jp

***Abstract*—Aiming to achieve sensing coverage for a given *Area of Interest* (AoI) in a *People-Centric Sensing* (PCS) manner, we propose a concept of (α, T)-coverage of the target field where each point in the field is sensed by at least one node with probability of at least α during the time period T. Our goal is to achieve (α, T)-coverage by a minimal set of mobile sensor nodes for a given AoI, coverage ratio α, and time period T. We model pedestrians as mobile sensor nodes moving according to a discrete Markov model. Based on this model, we propose two algorithms: the *inter-location* and *inter-meeting-time* algorithms, to meet a coverage ratio α in time period T. These algorithms estimate the expected coverage of the specified AoI for a set of selected nodes. The inter-location algorithm selects a minimal number of mobile sensor nodes from nodes inside the AoI taking into account the distance between them. The inter-meeting-time selects nodes taking into account the expected meeting time between the nodes. We conducted a simulation study to evaluate the performance of the proposed algorithms for various parameter setting including a realistic scenario on a specific city map. The simulation results show that our algorithms achieve (α, T)-coverage with good accuracy for various values of α, T, and AoI size.**

***Index Terms*—People centric sensing, Mobile sensors, Markov chain, Probabilistic coverage.**

I. Introduction

Nowadays, there is an increasing demand for realtime environmental information about specified regions in urban areas for various purposes such as surveillance, navigation, and event detection. The mobility of people opens up the possibility of using a set of mobile devices to cover a given *area of interest* (AoI) at low cost. Leveraging people as a part of the sensing infrastructure introduces a new sensing paradigm called *People-Centric Sensing* (PCS) [1]. In PCS, people with mobile devices play the role of mobile sensors to sense and collect information from their surroundings for the benefit of sensing applications and their users. Since the coverage in PCS relies on the uncontrollable mobility of people, we can only probabilistically guarantee the coverage of the target AoI.

An interesting motivating application comes from the real-time urban sensing scenario. For example, in a city sensing application, users like to know the information in a specific AoI such as crowded places, interesting spots, events on specific locations, and so on. In such an application, a user issues a query with a geographic area as the AoI, a required coverage ratio α, the required information (e.g., noise level), and a query interval (maximum allowable response time) T. After that, some people carrying mobile devices in the AoI, which satisfy the query requirements, will take part in the query responding process. However, we would like to select a minimal number of people with mobile devices that can provide the desired information. We refer to this problem as the *(α, T)-coverage problem*.

In this paper, we formally describe the (α, T)-coverage problem. Given an AoI, as a set of sensing points, a set of mobile nodes, and a query with a required coverage ratio α and a specified time interval T, the problem is to find the minimal set of mobile nodes such that each point in the AoI is visited and sensed by at least one node within T with a probability of at least α. To solve this problem, we need to be able to predict the future locations visited by each mobile node at any time depending on its initial location and its mobility. Thus, we model the mobility of the mobile nodes with a discrete Markov model. The solution of this problem depends critically on the number and the initial locations of mobile nodes inside and near the target AoI. Based on this, we propose two algorithms: the *inter-location* and *inter-meeting-time* algorithms, to meet a coverage ratio α in time period T. The inter-location algorithm estimates the probability of points in the AoI being visited by each mobile sensor node in T, and selects a minimal number of mobile nodes inside the AoI taking into account the distance between the nodes. The inter-meeting-time algorithm selects a minimal number of nodes taking into account the expected time until any two of the nodes will meet at a location.

We conducted a simulation study to evaluate the performance of the proposed algorithms for various parameter setting including a realistic scenario on a specific city map. As a result, we confirmed that the proposed algorithms achieve (α, T)-coverage with good accuracy for a variety of values of α, T, and AoI size, and the inter-meeting time algorithm selected smaller number of nodes without deteriorating coverage accuracy.

The rest of this paper is organized as follows. Section II reviews the related studies. Section III defines the (α, T)-coverage problem. Section IV describes the proposed algorithms based on the discrete Markov model. Section V shows the performance evaluation of the proposed algorithms in simulation-based experiments, and finally section VI concludes the paper.

II. Related Work

Many studies have proposed data gathering protocols to realize efficient communication between sensor nodes in wireless sensor networks (WSNs) [2]–[5]. Some studies also have proposed use of mobile sensor nodes in WSNs to improve coverage, lifetime, and/or fault-tolerance [6], [7].

Recently, information collection by pedestrians in PCS has received increasing attentions. PCS is different from existing WSNs because we cannot control the mobility of mobile nodes. There are several studies and research projects based on PCS [8]–[17].

SensorPlanet [8] is a platform that enables the collection of sensor data on a large and heterogeneous scale, and establishes a central repository for sharing the collected sensor data. Cartel [9] is a mobile communications infrastructure based on car-mounted communication platforms exploiting open WiFi access points in a city, and provides urban sensing information such as traffic conditions. CitySense [10] provides a static sensor mesh offering similar types of urban sensing data feeds. Bubble-sensing [12] is a sensor network that allows mobile phone users to create a binding between tasks and places of interest in the physical world. Mobile users are able to affix task bubbles at places of interest and then receive sensed data as it becomes available in a delay-tolerant fashion. PriSense [13] relies on data slicing and mixing and binary search to enable privacy-preserving queries, where each node slices its data into $(n+1)$ data slices, randomly chooses n other nodes, and sends a unique data slice to each of them. Finally, each node sends the sum of its own slice and the slices received from others to the aggregation server. GreenGPS [17] is a navigation service that uses participatory sensing data to map fuel consumption on city streets and find the most fuel-efficient route for vehicles between arbitrary endpoints.

Most of these approaches focus on information collection, but do not consider the probabilistic nature of coverage in PCS when the information collection period is restricted to a short time duration such as an on-demand query. They consider neither the difficulties of sensing coverage of a relatively wide area nor the time requirements of on-demand sensing by mobile users. However, these two criteria are very important in PCS. To meet these criteria, it is also very important to be able to estimate the area covered by each mobile node in a specified time interval. However, existing studies do not consider such a spatiotemporal coverage by mobile nodes.

The contribution of this paper is the formulation of the (α, T)-coverage problem and the design and evaluation of probabilistic algorithms that consider on-demand sensing by mobile users, and probabilistic coverage in PCS based on the mobility of people.

III. The (α, T)-Coverage Problem

In this section, we first describe the assumptions and models for our target PCS application, then formulate the target problem to realize the application.

A. Assumptions and Models

1) System model: We assume an application such that when requested, some mobile users take part in a task to obtain the latest environmental information such as noise level, sunshine intensity, temperature, exhaust gas concentration, and so on, over a specified geographical area of the urban district in a PCS fashion. We assume that those participating users have some incentives to serve as mobile sensors such as electronic currency or coupons given by a service provider.

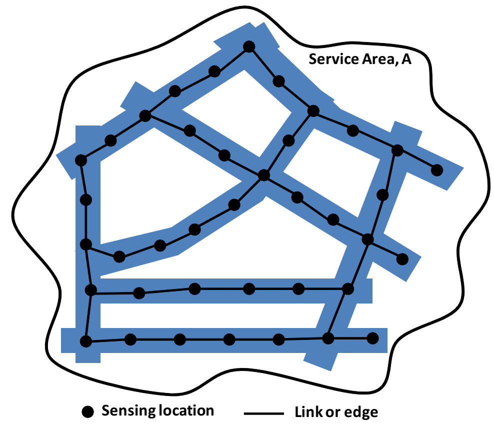


Fig. 1: Service area represented by a connected graph with sensing locations

We denote the whole service area by A. A road (street) network on which mobile users can move spans the area A. A service user wants to know the approximate condition of a specific area called the *Area of Interest (AoI)* produced by obtaining the environmental information about some locations in the AoI. Thus, we assume that there are multiple *sensing locations* with a uniform spacing Δ[1] (e.g., $\Delta = 50m$) on each road and that sensing coverage is achieved by obtaining the environmental information about all of the sensing locations in the specified AoI. We show an example road network with sensing locations in a service area in Fig. 1.

We represent the road network with sensing locations by a connected graph $G = (V, E)$, where V is the set of vertices corresponding to sensing locations and some of them are intersections and interesting spots and E is the set of edges corresponding to segments between neighboring sensing locations on roads.

Multiple service users of this application exist on the service area A and are moving on graph G. Each mobile user is equipped with a portable computing device such as a smart-phone capable of accessing the Internet via a cellular network (CDMA, GSM) from any place in A, measuring the current location, and sensing the nearby environmental information with its built-in sensors (camera, microphone, light-intensity sensor, etc). Hereafter, we refer to a service user with a mobile device simply as a *mobile node* or a *node*.

We assume that time progresses discretely as 0, 1, 2, and so on. Let U denote the set of mobile nodes on G at time 0. Each mobile node moves from one vertex to one of its neighboring vertices on G in a unit of time. Mobility of nodes is based

[1] We assume that each road can be divided into an integer number of segments with length Δ.

Sensing location	node u_1 $Prob(u_1, t, v_2, v)$			node u_2 $Prob(u_2, t, v_4, v)$			$SetProb(v, U', 2)$
v	$t = 0$	$t = 1$	$t = 2$	$t = 0$	$t = 1$	$t = 2$	$U' = \{u_1, u_2\}$
v_1	0	0.6	–	0	0	0	$1 - (1 - 0.6)(1 - 0) = 0.6$
v_2	1	–	–	0	0	0.2	$1 - (1 - 1)(1 - 0.2) = 1$
v_3	0	0.4	–	0	0.4	–	$1 - (1 - 0.4)(1 - 0.4) = 0.64$
v_4	0	0	0.2	1	–	–	$1 - (1 - 0.2)(1 - 1) = 1$

TABLE I: First time visiting and set coverage probabilities for the example in Fig. 2 with T = 2.

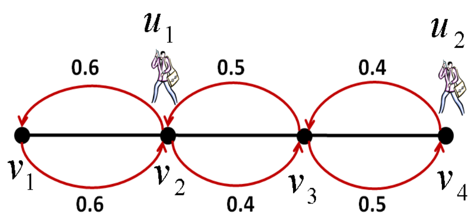


Fig. 2: Moving probability example for four sensing locations $\{v_1, v_2, v_3, v_4\}$ and initial locations of two mobile nodes $\{u_1, u_2\}$.

on a probabilistic model. Let $v_0^u \in V$ denote the initial (at time 0) location of node u. Let $Prob(u, t, v_0^u, v_t)$ denote the probability that each mobile node u with its location v_0^u at time 0 visits a vertex $v_t \in V$ at time t.

2) Service Model: We assume that our target application provides users with on-demand query service for sensing a specific AoI. We assume that there is a fixed server s in the Internet that can communicate with mobile nodes of U and executes required tasks.

We say that the AoI is α-*covered* if every sensing location in the AoI is visited (and thus the environmental information is sensed) at the location by at least one node with a probability of at least α. Here, we call α the *coverage ratio*. In our application, a mobile node sends s a query which asks for sensing a specified AoI with a specified coverage ratio α in a specified time interval T. We denote each query q by a quadruple $\langle AoI, S_{type}, \alpha, T\rangle$. Here, AoI is the area of interest in the service area specified by a set of sensing locations of V, and S_{type} specifies the type of environmental information to be sensed, such as temperature.

B. Problem Formulation

We call the probability of a set of mobile nodes $U'(\subseteq U)$ visiting a sensing location v $(\in V)$ in a time interval T, *the set coverage probability* denoted by $SetProb(v, U', T)$ and define it by the following equation.

$$SetProb(v, U', T) = 1 - \prod_{u \in U'} \prod_{t \in FVT_u} (1 - Prob(u, t, v_0^u, v)) \tag{1}$$

Here, FVT_u denotes the set of time steps no more than T at which u can visit v for the first time.

Fig. 2 shows an example for four sensing locations v_1, v_2, v_3, and v_4 and the moving probabilities between them. As shown in Fig. 2, there are two mobile nodes u_1 and u_2 at sensing locations v_2 and v_4, respectively. Table I shows the set coverage probabilities of v_1, v_2, v_3, and v_4 by $U' = \{u_1, u_2\}$ when T = 2.

Definition 1. ***(α, T)-coverage****: Given a graph $G = (V, E)$, an area specified by a set of sensing locations $AoI \subseteq V$, a set of mobile nodes $U' \subseteq U$, a required coverage ratio α, and a time interval T, the area AoI is called* ***(α, T)-covered*** *if the following condition holds.*

$$\forall v \in AoI, SetProb(v, U', T) \geq \alpha \tag{2}$$

We can now formally define the (α, T)-coverage problem as follows:

Definition 2. *Given the service area as a connected graph $G = (V, E)$, a set of mobile nodes U on G at time 0, and a query $q = \langle AoI, S_{type}, \alpha, T\rangle$, the* (α, T)***-coverage problem*** *is the problem of selecting a minimal set of mobile nodes $U' \subseteq U$ which achieves (α, T)-coverage of AoI.*

We define the objective function of this problem by the following equation.

$$\textbf{minimize } |U'| \tag{3}$$

$$\textbf{subject to } AoI \text{ is } (\alpha,\, T)\text{-}covered \tag{4}$$

This problem is NP-hard since it implies, as a special case, the *Minimum Set Covering Problem* (MSCP) which is known to be NP-hard [18].

IV. Algorithms

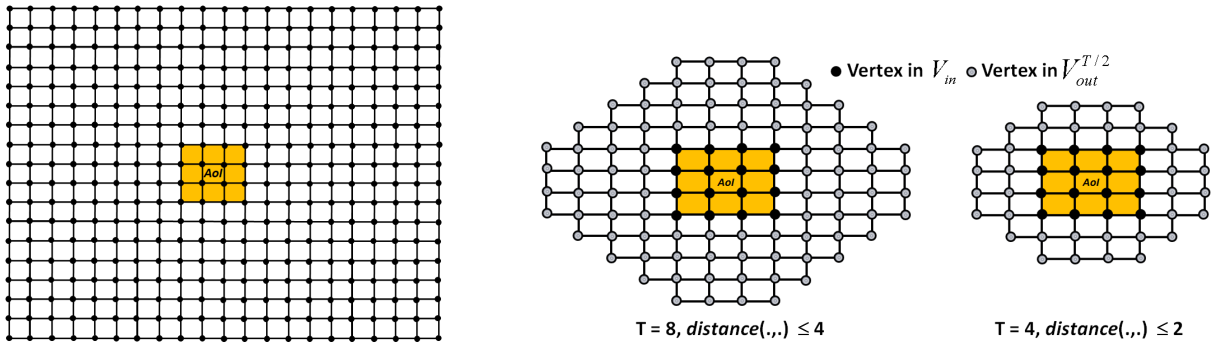
In this section, we propose two heuristic algorithms for the problem defined in Section III, named *Inter-Location Based (ILB)* and *Inter-Meeting Time Based (IMTB)* algorithms[2].

A. Preliminaries

Our basic idea is to select nodes that have higher probabilities of visiting distinct sensing locations in the specified AoI within a time interval T, prior to selecting other nodes.

The proposed algorithms depend on the probability $Prob(u, t, v_0^u, v_t)$ of each node u with initial location v_0^u visiting a location v_t at time t $(0 \leq t \leq T)$, that we call the *vertex-coverage probability*, hereafter. To simplify our

[2] We assume that all algorithms are executed by the server s in a centralized fashion.



(a) Original service area graph

(b) Reduced service area graph for T= 8, 4

Fig. 3: An example of a service area graph with AoI and its reduction for T= 8, 4

explanation, without loss of generality, we represent the graph $G = (V, E)$ for the service area by a grid of sensing locations (vertices) with a uniform spacing Δ between neighboring vertices and only vertical and horizontal edges (here, each edge is bi-directional), as shown in Fig. 3 and that mobile nodes move according to a discrete Markov model on this grid. Let N denote the number of vertices (i.e., $|V|$) and x_i denote the i-th vertex of $V(1 \leq i \leq N)$. We model the node movement on the grid as a Markov chain. For each node u, we define a vector with N states where the i-th state represents the probability that u is at vertex x_i.

Assuming that there are a sufficient number of nodes in the target area A, we select nodes only within the specified AoI. Let $U_0(\subseteq U)$ denote the set of nodes which are located in the target AoI at time 0.

1) Computation of vertex-coverage probability: Let $\boldsymbol{P}$ denote the probability matrix with size $N \times N$, where its i-th row and j-th column element represents the probability of a node at vertex x_i to move to vertex x_j by a unit of time. We define an initial state vector $\boldsymbol{v_0^u}$ representing that a node u is initially located at $x_i \in V$ by the following equation.

$$\boldsymbol{v_0^u} = (p_1, p_2, ..., p_N) \tag{5}$$

where

$$p_j = \begin{cases} 0 & (j \neq i) \\ 1 & (j = i) \end{cases} \tag{6}$$

Then, we can calculate the vertex-coverage probability of vertex $x_k \in V$ by node u at time t by the following equation.

$$Prob(u, t, v_0^u, x_k) = [\boldsymbol{v_0^u} \times \boldsymbol{P}^t]_k \tag{7}$$

Here, $[\,]_k$ denotes the k-th element in the resulted vector.

2) Reduction of probability matrix size: If the target service area contains many sensing locations, the probability matrix $\boldsymbol{P}$ will be large, resulting in a serious computational overhead in the server s. However, we only select nodes in the specified AoI and thus we do not need to consider the nodes which move more than $T/2$ away from the border of the AoI since such nodes never come inside the AoI again. This fact allows us to reduce the size of the probability matrix from $N \times N$ to $(M + L) \times (M + L)$, where M is the number of sensing locations included in the AoI and L is the number of sensing locations outside the AoI such that their shortest distances to the AoI border is less than or equal to $T/2$. Here, note that $N \gg M + L$ holds for typical scenarios where AoI and T are reasonably small and N is large.

Let $V_{in}(\subseteq V)$ denote a set of vertices included in the AoI. Let $V_{out}(= V - V_{in})$ denote the set of vertices outside the AoI, but in the service area. Let $distance(x, y)$ denote the shortest distance from vertex x to vertex y in the service area. Let $V_{out}^{T/2}$ denote a set of vertices in V_{out} such that the shortest distance from any vertex of $V_{out}^{T/2}$ to at least one vertex of V_{in} is less than or equal to $T/2$. $V_{out}^{T/2}$ is defined by the following equation.

$$V_{out}^{T/2} = \{x \mid x \in V_{out} \land \exists y, distance(x, y) \leq T/2 \land y \in V_{in}\} \tag{8}$$

The vertices that belong to $V_{out}^{T/2}$ are illustrated in Fig. 3.

We can calculate the vertex-coverage probability of all vertices in V_{in} taking into account only the node moving probability at each vertex of $V_{in} \cup V_{out}^{T/2}$. Consequently, we define the new probability matrix $\boldsymbol{P'}$ for vertices of $V_{in} \cup V_{out}^{T/2}$.

We define the i-th row and j-th column element $p'_{i,j}$ of $\boldsymbol{P'}$ by the following equation.

$$p'_{i,j} = \begin{cases} p_{i,j} & (x_i, x_j \in V_{in} \cup V_{out}^{T/2} \land i \neq j) \\ \displaystyle\sum_{x \in Ngh(i)} p(i, x) & (x_i \in V_{out}^{T/2} - V_{out}^{T/2-1} \land i = j) \end{cases} \tag{9}$$

Here, $Ngh(i)$ is the set of neighboring vertices outside $V_{out}^{T/2}$ and $p_{i,j}$ is the probability of the corresponding edge in the original matrix $\boldsymbol{P}$. Equation (9) represents that the moving probability from x_i to x_j is the same as the original matrix $\boldsymbol{P}$ if i is not equal to j, but it is changed to the sum of probabilities of outgoing edges if x_i is on the border of $V_{out}^{T/2}$ and i is equal to j.

B. The Inter-Location Based Algorithm (ILB)

The ILB uses the distance between nodes as a metric to select a set of mobile nodes. We denote the distance between the initial locations of nodes u and u' in U_0 by $d_{u,u'}$ which is determined as the length of the shortest path between v_0^u and $v_0^{u'}$ on G. The ILB algorithm selects a minimal set of mobile nodes $U'(\subseteq U_0)$ such that the distance between any pair of nodes u and u' in U' is equal to or larger than a threshold d_{th}, and the specified AoI is (α, T)-covered. The above statement is defined as follows.

$$\textbf{minimize } |U'| \textbf{ subject to } (11)-(12) \tag{10}$$

$$d_{u,u'} \geq d_{th}, \forall u, u' \in U' \tag{11}$$

$$\text{AoI is } (\alpha,\, T)\text{-}\textit{covered} \tag{12}$$

The value of d_{th} should be dependent on three parameters: the total number of time steps T, the required coverage ratio α, and the maximum distance d_{max} that is the largest distance between the initial locations of two nodes in U_0. Intuitively, as T increases and/or α decreases, the number of selected nodes should decrease. On the contrary, as T decreases and/or α increases, the number of selected nodes must be increased to meet the (α, T)-coverage constraint. Thus, to minimize the number of selected nodes, we must choose an appropriate value for d_{th}. To reflect the above relationship among parameters, we define the distance threshold d_{th} by the following equation.

$$d_{th} = \min\left(\frac{T}{\alpha \cdot d_{max}}, d_{max}\right) \tag{13}$$

Algorithm 1 *The Inter-location based algorithm (ILB)*

Input: U, AoI, α, T, $G = (V, E)$
Output: U'

1: $U' \rightarrow \emptyset$
2: Compose V_{in}, V_{out}^T, U_0 from AoI and U
3: $\boldsymbol{P} \rightarrow$ ComputeProbMatrix(AoI,$V_{in} \cup V_{out}^T$)
4: **for** $\forall u \in U_0$ **do**
5: Compose u's initial state vector $\boldsymbol{v_0^u}$
6: **end for**
7: $d_{max} \rightarrow \max_{u,u' \in U_0} \{d_{u,u'}\}$
8: $d_{th} \rightarrow \min(\frac{T}{\alpha \cdot d_{max}}, d_{max})$
9: **while** $SetProp(v, U', T) < \alpha, \forall v \in V_{in}$ **do**
10: **if** $U_0 = \emptyset$ **then**
11: return $\emptyset$
12: **end if**
13: Select $u \in U_0$ at random
14: **if** $U' = \emptyset$ or $\min_{u' \in U'} \{d_{u,u'}\} \geq d_{th}$ **then**
15: $U' \rightarrow U' \cup \{u\}, U_0 \rightarrow U_0 - \{u\}$
16: **end if**
17: **end while**
18: **return** U'

Algorithm 1 shows the node selection process of ILB. The input parameters are the set of mobile nodes U, the area of interest AoI, the required coverage ratio α, the query interval time T, and the service area graph $G = (V, E)$. In line 1, the algorithm initializes U' to empty. In line 2, it composes the sets of vertices V_{in} and V_{out}^T, and set of nodes U_0. In line 3, it composes the probability matrix $\boldsymbol{P}$. In lines 4 to 6, it composes the initial state vector for each node $u \in U_0$. In lines 7 and 8, the algorithm determines the maximum distance d_{max} between nodes in U_0 and the distance threshold d_{th}, as defined in equation (13). In lines 9 to 18, the algorithm selects a set of nodes U' as follows: (i) while the AoI is not (α, T)-covered, the algorithm checks the state of U_0 and if U_0 is empty, the algorithm returns $\emptyset$ (i.e., the current U_0 is not sufficient to satisfy the required coverage α), as shown in lines 9 to 12, (ii) the algorithm selects a node $u \in U_0$ at random, as shown in line 13; and (iii) it adds the node u to the selected set of nodes U' if U' is empty or the distance between u and each node $u' \in U'$ is no less than the threshold d_{th}, as shown in lines 14 to 16. Finally, in line 18, the algorithm returns the selected set of nodes U'.

C. The Inter-Meeting Time Based Algorithm (IMTB)

The ILB algorithm is based on the distance between nodes. Hence, the selection process is location-dependent and does not take the query interval time T into consideration. To make the node selection more efficient taking into account the value of T, we propose an inter-meeting time based algorithm which uses the expected meeting time between nodes as a metric. This meeting time metric reflects the probability of nodes visiting distinct sensing locations of AoI and describes the expected first meeting time of any pair of nodes $u, u' \in U_0$. Intuitively, as the meeting time between nodes increases, the probability of visiting distinct sensing locations also increases[3] because those nodes explore different locations until they meet. We denote the expected meeting time between nodes u and u' in U_0 by $mt_{u,u'}$. The IMTB algorithm selects a minimal set of mobile nodes $U'(\subseteq U_0)$ such that the meeting time $mt_{u,u'}$ between any pair of nodes u and u' in U' is no less than a meeting time threshold mt_{th}, and the specified AoI is (α, T)-covered. The above statement is defined as follows.

$$\textbf{minimize } |U'| \textbf{ subject to } (15)-(16) \tag{14}$$

$$mt_{u,u'} \geq mt_{th}, \forall u, u' \in U' \tag{15}$$

$$\text{AoI is } (\alpha, T)\text{-}\textit{covered} \tag{16}$$

The values of $mt_{u,u'}$ and mt_{th} are calculated as follows.

The expected meeting time $mt_{u,u'}$ represents the earliest time when two nodes u and u' in U_0 may meet at some location $v_t \in V_{in}$ and is defined by the following equation.

$$mt_{u,u'} = \begin{cases} \min_{t \in MT_{u,u'}} \{t\} & (MT_{u,u'} \neq \emptyset) \\ T & (MT_{u,u'} = \emptyset) \end{cases} \tag{17}$$

where $MT_{u,u'}$ is a set of expected meeting time between u and u' during the time period T and is defined by the following

[3]This is not the case if the probability of a node staying the same location is high, but we suppose the environment where most of nodes near AoI move directly to their destinations.

equation.

$$MT_{u,u'} = \{t \mid Prob(u, t, v_0^u, v_t) > 0 \wedge Prob(u', t, v_0^{u'}, v_t) > 0, 0 \leq t \leq T, \exists\ v_t \in AoI\} \tag{18}$$

The meeting time threshold mt_{th} should be dependent on three parameters: the total number of time steps T, the required coverage ratio α, and the maximum estimated meeting time mt_{max} between pairs of nodes in U_0. Intuitively, as T increases and/or α decreases, the number of selected nodes will decrease. As a result, to minimize the selected number of nodes, the estimated meeting time threshold must be appropriately adjusted to meet the (α, T)-coverage constraint. To reflect the above relationship among parameters, we define the meeting time threshold mt_{th} as follows.

$$mt_{th} = \min(\frac{T}{\alpha \cdot mt_{max}}, mt_{max}) \tag{19}$$

Algorithm 2 *The Iner-meeting time based algorithm (IMTB)*

Input: U, AoI, α, T, $G = (V, E)$
Output: U'

1: $U' \rightarrow \emptyset$
2: Compose V_{in}, V_{out}^T, U_0 from AoI and U
3: $\boldsymbol{P} \rightarrow$ ComputeProbMatrix(AoI,$V_{in} \cup V_{out}^T$)
4: **for** $\forall u \in U_0$ **do**
5: compose u's initial state vector $\boldsymbol{v_0^u}$
6: **end for**
7: $mt_{max} \leftarrow \max_{u,u' \in U_0}\{mt_{u,u'} : mt_{u,u'} \neq \infty\}$
8: $mt_{th} \leftarrow \min(\frac{T}{\alpha \cdot mt_{max}}, mt_{max})$
9: **while** $SetProp(v, U', T) < \alpha, \forall v \in V_{in}$ **do**
10: **if** $U_0 = \emptyset$ **then**
11: return $\emptyset$
12: **end if**
13: Select $u \in U_0$ at random
14: **if** $U' = \emptyset$ or $\min_{u' \in U'}\{mt_{u,u'}\} \geq mt_{th}$ **then**
15: $U' \rightarrow U' \cup \{u\}, U_0 \rightarrow U_0 - \{u\}$
16: **end if**
17: **end while**
18: **return** U'

Algorithm 2 shows the node selection process of IMTB. The input parameters are the set of mobile nodes U, the area of interest AoI, the required coverage ratio α, the query interval time T, and the service area graph $G = (V, E)$. In lines 1 to 6, the algorithm does the same steps as lines 1 to 6 in algorithm 1. In lines 7 and 8, the algorithm determines the maximum estimated meeting time mt_{max} between nodes in U_0 and the threshold mt_{th}, as defined in equation (19). In lines 9 to 18, the algorithm selects a set of nodes U' as follows: (i) while the AoI is not (α, T)-covered, the algorithm checks the state of U_0 and if U_0 is empty, the algorithm returns $\emptyset$ (i.e., the current U_0 is not sufficient to satisfy the required coverage α), as shown in lines 9 to 12, (ii) the algorithm selects a node $u \in U_0$ at random, as shown in line 13; (iii) it adds the node u to the selected set of nodes U' if U' is empty or the estimating meeting time between u and each node $u' \in U'$ is no less than the threshold mt_{th}, as shown in lines 14 to 16. Finally, in line 18, the algorithm returns the selected set of nodes U'.

V. Performance Evaluation

In this section, we report the results of simulations that we examine the coverage performance of the proposed algorithms in terms of the number of mobile sensor nodes selected and the accuracy of achieved coverage ratio.

TABLE II: Configuration Parameters.

Configuration parameter	**Value in simulation**
# nodes	25 to 200
Node speed	1 m/s
Field size	$500m \times 500m$
Required coverage, α	0.2, 0.4, 0.5, 0.6, 0.8, 0.9
Total # sensing locations in A	121
AoI-Ratio (# sensing locations)	0.01 (4), 0.25 (36), 0.45 (56), 0.5 (66), 0.65 (77), 0.85 (99)
Δ	50 m
Total # steps (time period), T	2, 3, 4,..., 20

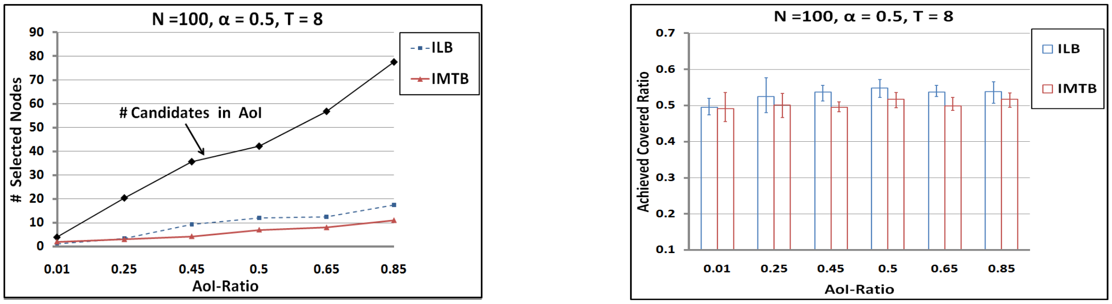
A. Simulation Environment

The QualNet [19] simulator was used with the input parameters listed in table II, such as service area size, number of nodes, node speed, etc. In addition, the node mobility was based on a discrete Markov model as described in section IV. The service area was represented as a grid of sensing locations arranged with uniform spacing, 50 *meters*. We selected the AoI as a rectangular region where its position was selected at random within the service area and the ratio of its size to the service area size, called *AoI-Ratio*, was selected from {0.01, 0.25, 0.45, 0.5, 0.65, 0.85} and the corresponding number of sensing locations in each AoI was {4, 36, 56, 66, 77, 99}. The initial node location was selected at random among sensing locations in the service area. The probability of a node at a location moving to one of its neighboring locations was set uniformly to 0.25. We repeated every simulation experiment 5 times with different initial node distributions, then averaged the results.

We measured the performance of the proposed algorithms in terms of the number of selected nodes and the achieved coverage ratio, by changing the number of nodes, the AoI-Ratio, the total number of time steps (query interval time), and the required coverage ratio. Here, we define the achieved coverage ratio as the ratio of the number of sensing locations visited by at least one node to the total number of sensing locations in AoI. We say that the algorithms satisfy the required coverage ratio if the average achieved coverage ratio of several simulation runs is no less than the required ratio.

B. Simulation Results

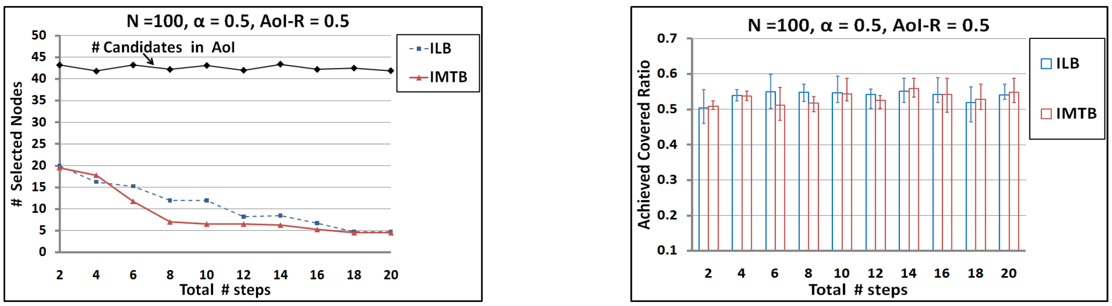
We show the simulation results in Fig. 4, Fig. 5, Fig. 6, and Fig. 7 (The black line with solid diamonds in Fig. 4 (a),



(a) Number of selected nodes vs. AoI-Ratio

(b) Achieved coverage ratio vs. AoI-Ratio

Fig. 4: Performance for different AoI-Ratios



(a) Number of selected nodes vs. Total number of steps

(b) Achieved coverage ratio vs. Total number of steps

Fig. 5: Performance for different time steps

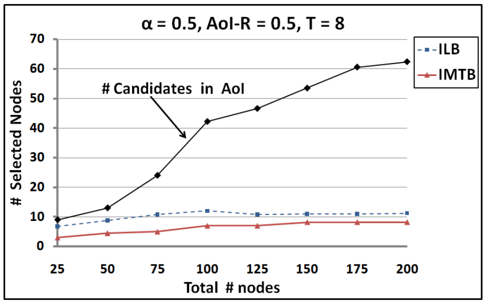
Fig. 5 (a), Fig. 6 (a), and Fig. 7 (a) represents the the number of candidate nodes in the AoI).

Fig. 4 shows the performance for different AoI-Ratio in case of medium required coverage ratio and medium number of time steps. The number of nodes was 100, the required coverage α was 0.5, and the total number of steps was 8. In Fig. 4 (a), the number of selected nodes increased as AoI-Ratio increased. This is because, when the AoI-Ratio increased, we need more nodes to satisfy the required coverage ratio. The number of selected nodes for the proposed algorithms is much smaller than the number of candidate nodes in the AoI. The number of selected nodes for ILB was lower than IMTB when the AoI-Ratio was 0.01 (very small). For a higher AoI-Ratio, the number of selected nodes for IMTB was much lower than ILB. In Fig. 4 (b), the required coverage is almost satisfied by both algorithms. When the AoI-Ratio was 0.01, the variance of ILB was smaller than IMTB. For a higher AoI-Ratio, the variance of IMTB was smaller than ILB. As a result, the performance of IMTB is the best, since IMTB selected lowest number of nodes and satisfied the required coverage with small value of variance for the higher values of AoI-Ratio.

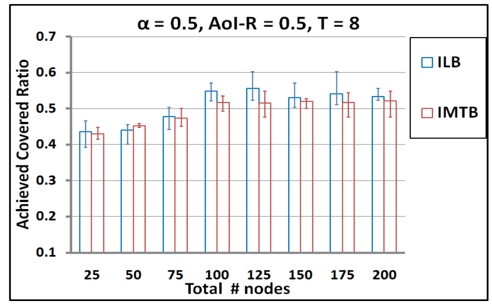
Fig. 5 shows the performance for different numbers of time steps with a medium size AoI and a medium required coverage ratio. The number of nodes was 100, the required coverage α was 0.5, and the AoI-Ratio was 0.5. In Fig. 5 (a), the number of selected nodes decreased as the total number of steps increased. This is because the distance and meeting time threshold increases in proportion to the total number of steps. The number of selected nodes for IMTB was lower than ILB since the required coverage is medium and the meeting time increased when the AoI-Ratio is medium. In Fig. 5 (b), both algorithms satisfied the required coverage and the variance of IMTB was smaller than ILB.

Fig. 6 shows the performance for different numbers of nodes with a medium size AoI, a medium required coverage ratio, and a medium number of time steps. The number of nodes was 100, the required coverage α was 0.5, and the total number of steps was 8. In Fig. 6 (a), when the number of nodes was 25 to 125, the number of selected nodes increased as the number of nodes increased. This is because, as the number of nodes increases, the number of selected nodes increases. However, when the number of nodes was higher than 125, the number of selected nodes was fixed since the number of selected nodes is bounded by the number of nodes needed to satisfy the required coverage. In Fig. 6 (b), the required coverage was not satisfied by ILB and IMTB algorithms when the total number of nodes was 25 to 75. For a higher number of nodes, both algorithms satisfied the required coverage. The variance of IMTB is smaller than ILB.

Fig. 7 shows the performance for different required coverage ratio with medium size AoI and medium number of time steps. The number of nodes was 100, the AoI-Ratio was 0.5, and the total number of steps was 8. In Fig. 7 (a), the number of selected nodes increased as required coverage ratio increased. This is because, as the required coverage ratio increases, we need more nodes to satisfy it. The number of selected nodes

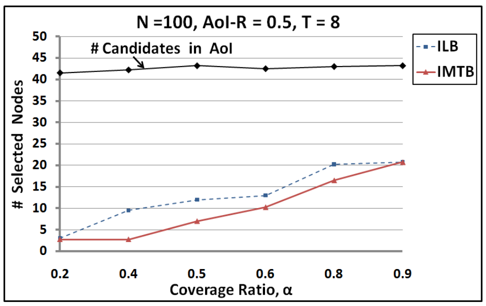


(a) Number of selected nodes vs. Total number of nodes

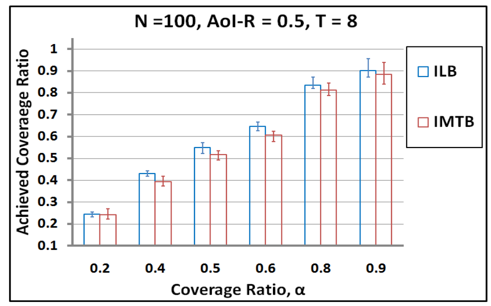


(b) Achieved coverage ratio vs. Total number of nodes

Fig. 6: Performance for different number of nodes



(a) Number of selected nodes vs. Required coverage ratio



(b) Achieved coverage ratio vs. Required coverage ratio

Fig. 7: Performance for different required coverage ratios

for IMTB was lower than ILB. In Fig. 7 (b), the required coverage was satisfied by both algorithms and the variance of IMTB is smaller than ILB.

Here, we summarize the simulation results as follows.

- Both ILB and IMTB algorithms reduce the number of selected nodes to a great extent for (α, T)-coverage compared to the number of candidates nodes in the AoI.
- For a small AoI, ILB can select a smaller number of nodes to meet the required coverage with smaller variance than IMTB.
- For medium and large AoI, IMTB can select smaller number of nodes to meet the required coverage with smaller variance than ILB.
- When only a small number of nodes are initially located in the AoI, the ILB and IMTB can not meet the required coverage.

C. Realistic Scenario Evaluation

To show the performance of the proposed algorithms on a realistic scenario, we conducted a simulation on a specific city map near Osaka station in Japan. Fig. 8 shows the city map with its road network and its sensing locations with a uniform spacing which was 50 meters. Here, the number of nodes was 100, the required coverage α was 0.5, and the AoI-Ratio was 0.5. The blue rectangle in Fig. 8 represents the selected AoI area. We used MobiREAL simulator [20] to generate a realistic mobility trace for mobile nodes from the actually observed number of nodes on each street. Based on the city map and the generated trace, the moving probability of mobile nodes was determined. We show the simulation results in Fig. 9 (The black line with solid diamonds in Fig. 9 (a) represents the the number of candidate nodes in the AoI).

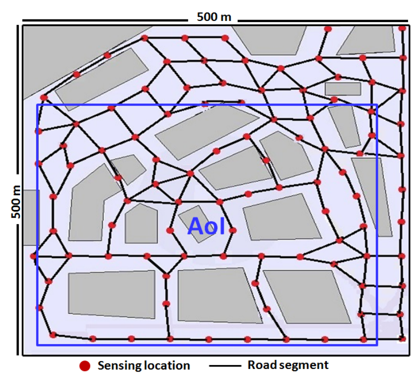
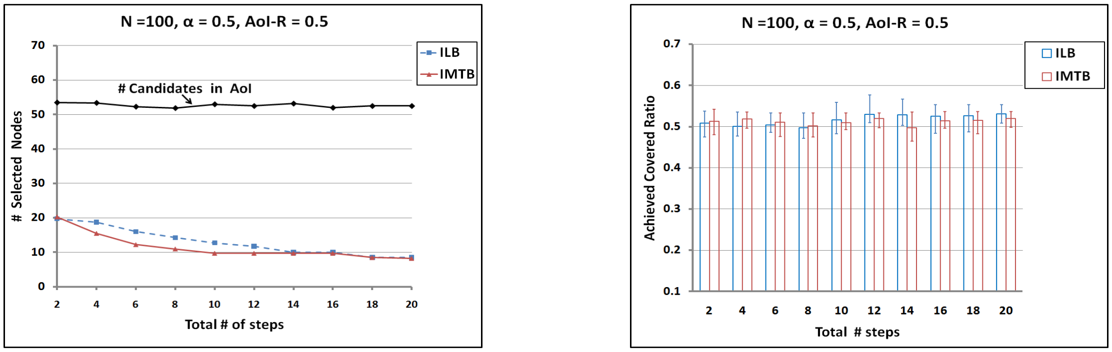


Fig. 8: A city map with roads, sensing locations, and AoI

In Fig. 9 (a), the trend on the number of selected nodes is almost similar to the case of Fig. 5, but more nodes was selected when the number of steps was bigger. This is because, the probability matrix in this experiment is not uniform and there are a smaller number of nodes moving towards some sensing locations in AoI. In Fig. 9 (b), ILB and IMTB satisfied the required coverage and the variance of IMTB is in most cases smaller than ILB.



(a) Number of selected nodes vs. Total number of steps

(b) Achieved coverage Ratio vs. Total number of steps

Fig. 9: Performance in realistic scenario

VI. Conclusion

In this paper, we tackled the (α, T)-coverage problem in people centric sensing with a motivating application scenario. We formulated this problem as an optimization problem with the objective of minimizing the number of selected nodes to meet the demanded coverage ratio α within a query interval time T. To resolve this problem, we proposed two heuristic algorithms: the inter-location and inter-meeting-time algorithms. The inter-location algorithm that selects a minimal number of mobile sensor nodes from nodes inside and near AoI taking into account the distance between them. The inter-meeting-time algorithm that selects nodes taking into account the expected meeting time between the nodes.

Our simulation results showed that the proposed algorithms achieved (α, T)-coverage with good accuracy for a variety of values of α, T, and AoI size, and that the inter-meeting time based algorithm selects a smaller number of nodes without deteriorating coverage accuracy.

In this paper, the proposed algorithms are based only on the initial locations of mobile nodes. If we react the location of nodes in and near the AoI during period T, we may be able to achieve more accurate coverage with lower cost by removing useless nodes and adding some extra nodes that more contribute coverage. The implementation and evaluation of the update mechanism is part of our future work. Also, we considered only the case where a single query is issued at a time. In future work, we will try to make the proposed algorithms adaptive to the case of multiple simultaneous queries.

References

[1] A. Campbell, S. Eisenman, N. Lane, E. Miluzzo, R. Peterson, H. Lu, X. Zheng, M. Musolesi, K. Fodor, and G. Ahn, "The rise of people-centric sensing," *IEEE Internet Computing Special Issue on Mesh Networks*, vol. 12, no. 4, pp. 12–21, Jul./Aug. 2008.

[2] J. N. Al-Karaki and A. E. Kamal, "Routing techniques in wireless sensor networks: A survey," *IEEE Wireless Communications*, vol. 11, no. 6, pp. 6–28, Dec. 2004.

[3] K. Akkaya and M. Younis, "A survey on routing protocols for wireless sensor networks," *IEEE Wireless Communications*, vol. 3, no. 3, pp. 325–349, May 2005.

[4] L. Lima and J. Barros, "Random walks on sensor networks," in *the 5th International Syposium on Modeling and Optimization in Mobile, Ad hoc, and Wireless Networks*, Apr. 2007.

[5] R. Shah, S. Roy, S. Jain, and W. Brunette, "Data mules: Modeling a three-tier architecture for sparse sensor networks," in *IEEE Workshop on Sensor Network Protocols and Applications*, May 2003.

[6] R. Katsuma, Y. Murata, N. Shibata, K. Yasumoto, and M. Ito, "Extending k-coverage lifetime of wireless sensor networks using mobile sensor nodes," in *the 5th IEEE International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob'09)*, Oct. 2009.

[7] W. Wang, V. Sirinivasan, and K. Chua, "Trade-offs between mobility and density for coverage in wireless sensor networks," in *the 5th IEEE International Conference on Mobile Computing and Networking (MobiCom'07))*, Sep. 2007.

[8] V. Tuulos, J. Scheible, and H. Nyholm, "Combining web, mobile phones and public displays in large-scale: Manhattan story mashup," in *the 5th International Conference on Pervasive Computing*, May 2007.

[9] B. Hull, V. Bychkovsky, Y. Zhang, K. Chen, M. Goraczko, A. Miu, E. Shih, H. Balakrishnan, and S. Madden, "Cartel: A distributed mobile sensor computing system," in *the 4th International Conference on Embedded Networked Sensor Systems*, Nov. 2006, pp. 125–138.

[10] R. Murty, A. Gosain, M. Tierney, A. Brody, A. Fahad, J. Bers, , and M. Welsh, "Citysense: A vision for an urban-scale wireless networking testbed," in *IEEE International Conference on Technologies for Homeland Security*, May 2008.

[11] S. Eisenman, N. Lane, E. Miluzzo, R. Peterson, G. Ahn, and A. Campbell, "Metrosense project: People-centric sensing at scale," in *Workshop on World-Sensor-Web*, Oct. 2006, pp. 6–11.

[12] H. Lua, N. Lane, S. Eisenmanb, and A. Campbella, "Bubble-sensing: Binding sensing tasks to the physical world," *Journal of Pervasive and Mobile Computing*, vol. 6, no. 1, pp. 58–71, Feb. 2010.

[13] J. Shi, R. Zhang, Y. Liu, and Y. Zhang, "Prisense: privacy-preserving data aggregation in people-centric urban sensing systems," in *IEEE INFOCOM'10*, Mar. 2010.

[14] A. Campbell, S. Eisenman, N. Lane, E. Miluzzo, and R. Peterson, "People-centric urban sensing," in *WICON06*, Aug. 2006.

[15] C. Cornelius, A. Kapadia, D. Kotz, D. Peebles, M. Shin, and N. Triandopoulos, "Anonysense: privacy-aware people-centric sensing," in *MobiSys08*, Jun. 2008.

[16] T. Abdelzaher *et al.*, "Mobiscopes for human spaces," *IEEE Pervasive Computing*, vol. 6, no. 2, pp. 20–29, Apr. 2007.

[17] R. K. Ganti, N. Pham, H. Ahmadi, S. Nangia, and T. F. Abdelzaher, "Greengps: a participatory sensing fuel-efficient maps application," in *MobiSys10*, Jun. 2010.

[18] C. Vercellis, "A probabilistic analysis of the set covering problem," *Annals of Operations Research*, vol. 1, no. 3, pp. 255–271, Oct. 1984.

[19] *Qualnet 4.0 network simulator*, Scalable Network Technologies Inc., 2007.

[20] K. Maeda, A. Uchiyama, T. Umedu, H. Yamaguchi, K. Yasumoto, and T. Higashino, "Urban pedestrian mobility for mobile wireless network simulation," *Journal on Ad hoc Networks*, vol. 7, no. 1, pp. 153–170, Jan. 2009.