IPSJ SIG Technical Report

Vol.2010-MBL-55 No.3
2010/9/2

# Probabilistic Methods for Spatio-Temporal Coverage in People-Centric Sensing

ASAAD AHMED , ,[†1] KEIICHI YASUMOTO ,,[†1]
YUKIKO YAMAUCHI ,[†1] and MINROU ITO [†1]

*People-Centric Sensing* (PCS) is a new paradigm for gathering environmental information of a specified region in the urban district. This paradigm relies on uncontrolled mobility of people to achieve sensing coverage for a given *areas of Interest* (AoI) over time at low cost. In this paper, we propose a concept of $(\alpha, T)$-coverage of the target field where each point in the field is sensed by at least one node at probability of at least $\alpha$ during time period $T$. Our goal is to achieve $(\alpha, T)$-coverage by as small number of mobile sensor nodes as possible for a given AoI, coverage ratio $\alpha$, and time period $T$. We model an urban sensing scenario with pedestrians as mobile sensor nodes moving according to a discrete Markov model. Based on this model, we propose two heuristic algorithms: *inter-location* and *inter-meeting-time* based algorithms, to meet a coverage ratio $\alpha$ and time period $T$. These algorithms estimate the expected coverage of the specified AoI for a set of selected nodes. The former algorithm selects some of the mobile sensor nodes inside the AoI taking into account the distance between them. The latter selects some of the nodes taking into account the expected meeting time between them. We conduct a simulation study to evaluate the performance of proposed algorithms for various parameter settings. Through simulation experiments, we confirmed that our algorithms achieve $(\alpha, T)$-coverage with good accuracy for various values of $\alpha$, $T$, and AoI size.

## 1. Introduction

Nowadays, there is an increasing demand for obtaining environmental information of a specified region in the urban district 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 a role of mobile sensors to sense and collect information from their surroundings for the benefits of some sensing applications and their users. Since the coverage in PCS relies on the uncontrolled mobility of people, we can guarantee the coverage of the target AoI only probabilistically. Hence, it is preferable to be able to measure the achievable coverage degree as the ratio. An interesting motivating application comes from the urban sensing scenario. For example, in the city sensing application, users like to know the information in a specific area of interest (AoI) such as crowds places, interested spots, events on the spots, and so on. In such an application, a user issues a query with a geographic area as AoI, required coverage ratio $\alpha$, required information (e.g. noise), and a time 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. We refer to this problem as the *$(\alpha, T)$-coverage problem*.

In this paper, we formally define the $(\alpha, T)$-coverage problem. Given an AoI, as a set of sensing points, a set of mobile nodes, and a query with a required coverage ratio $\alpha$ 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 $T$ at probability of at least $\alpha$. 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 mobile nodes' mobility by a discrete Markov model. The solution for this problem depends critically on the number and the initial locations of mobile nodes near the target AoI. Based on this fact, we propose two heuristic algorithms: *inter-location* and *inter-meeting-time* based algorithms, to meet a coverage ratio $\alpha$ and time period $T$. The former algorithm estimates the probability of points in the AoI visited by each mobile sensor node in $T$, and selects a minimal set of mobile nodes inside the AoI taking into account the distance between the nodes. The latter selects nodes taking into account the expected meeting time between the nodes instead of the distance.

Through simulation-based experiments, we confirmed that the proposed algorithms achieve $(\alpha, T)$-coverage with good accuracy for a variety of values of $\alpha$, $T$, and AoI size, and the inter-meeting time based algorithm selects smaller number of nodes without deteriorating coverage accuracy.

The rest of this paper is organized as follows. Section 2 overviews the related studies. Section 3 defines the $(\alpha, T)$-coverage problem. Section 4 describes the proposed algorithms based on the discrete Markov Model. Section 5 shows the performance evaluation of the proposed algorithms, and finally Section 6 concludes the paper.

†1 Graduate School of Information Science, Nara Institute of Science and Technology

## 2. Related Work

Recently, information collection by pedestrians and vehicles in the name of People Centric Sensing (PCS) has received increasing attentions. There are several studies and research projects based on PCS [2–7].

SensorPlanet [2] 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 [3] 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 [4] provides a static sensor mesh offering similar types of urban sensing data feeds. Bubble-sensing [6] 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 [7] 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.

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

The contribution of this paper is the formulation of $(\alpha, T)$-coverage problem and the design and evaluation of the probabilistic algorithms that consider the sensing coverage of a relatively wide area, the on-demand sensing by mobile users, and the probabilistic coverage in PCS based on the uncontrolled mobility of people. Our goal is to achieve $\alpha$ coverage by the minimal set of mobile nodes for a given AoI within the specified time interval $T$. Since the mobility of nodes is uncontrollable, the coverage depends critically on the number and the initial locations of mobile nodes near the target AoI. Therefore, we develop in our algorithms mechanisms to probabilistically estimate the future coverage of the target AoI in time interval $T$ for a set of some selected nodes.

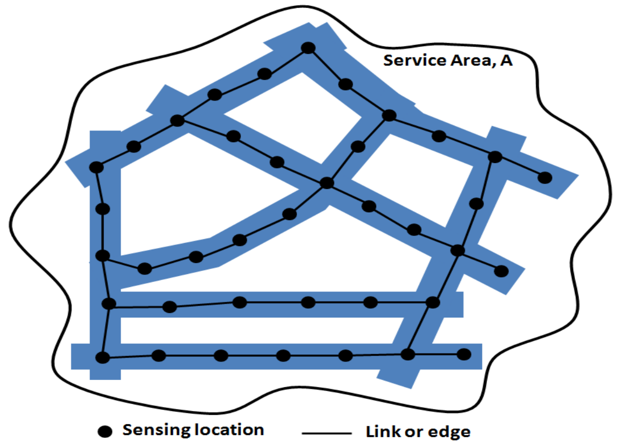


**Fig. 1** Service area represented by a connected graph with sensing locations

## 3. $(\alpha, T)$-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.

### 3.1 Assumptions and Models

#### 3.1.1 System model

We suppose an application such that as requested, some mobile users take part in a task for obtaining the latest environmental information such as a noise, sunshine intensity, temperature, exhaust gas concentration, and so on, over a specified geographical area in the urban district in a PCS fashion. We denote the whole service area by $A$. We assume that a road (street) network on which mobile users can move is spanned over the area $A$. We suppose that each service user wants to know the approximate condition of a specific area called *Area of Interest (AoI)* by obtaining the environmental information on some locations distributed over the AoI. Thus, we assume that there are multiple *sensing* *locations* with a uniform spacing $\Delta^{\star 1}$ (e.g., $\Delta = 50m$) on each road and that the sensing coverage is achieved by obtaining the environmental information on all of the sensing locations in the specified AoI. We show an example road network with sensing locations on 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 intersections, interested spots, and sensing locations and $E$ is the set of edges corresponding to road segments between neighboring sensing locations on roads.

We suppose that multiple service users of this application exist on the service area $A$ and are moving on graph $G$. We assume that each mobile user is equipped with a portable computing device like a smartphone or a small PC capable of accessing the Internet via a cellular network (WCDMA, GSM) from any place in $A$, measuring the current location (e.g., by its built-in GPS), and sensing the nearby environmental information with its built-in sensors (camera, microphone, light-intensity, etc). Hereafter, we refer to a service user with a mobile device simply by a *mobile node* or a *node*.

We assume that time progresses discretely like 0, 1, 2, and so on. Let $U$ denote the set of mobile nodes on $G$ at time $0$. We assume that all mobile nodes of $U$ move at a speed $\Delta$ per unit of time, which means that each mobile node moves from one vertex to one of its neighboring vertices on $G$ in a unit of time. We assume that each mobile node moves on graph $G$ according to a probabilistic mobility model and that we cannot control its moving route. 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 $t$.

#### 3.1.2 Service Model

We suppose 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 *$\alpha$-covered* if any sensing location in the AoI is visited and the environmental information is sensed at the location by at least one node at probability of at least $\alpha$. Here, we call $\alpha$ 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 $\alpha$ 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.

### 3.2 Problem Formulation

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

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

Here, $FVT_u$ denotes the set of time steps no more than $T$ where the probability of $u$ visiting $v$ for the first time at time $t(\in FVT_u))$ is more than 0.

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

**Definition 1.** *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 $q = \langle AoI, S_{type}, \alpha, T\rangle$, $(\alpha, T)$**-coverage problem** is the problem that selects a minimal set of mobile nodes $U' \subseteq U$ which guarantees $(\alpha, T)$-coverage of AoI.*

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

$$\textbf{minimize } |U'| \quad (2)$$

$$\textbf{subject to } \forall v \in AoI, SetProb(v, U', T) \geq \alpha \quad (3)$$

This problem is NP-hard since it is a typical combinatorial optimization problem and contains, as a special case, *Minimum Set Covering Problem* (MSCP) known to be NP-hard [8].

## 4. Algorithms

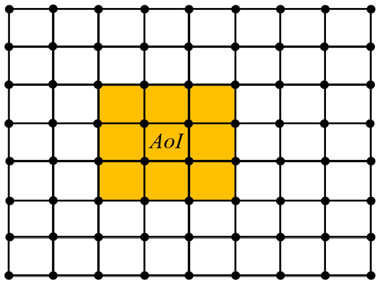
In this section, we propose two heuristic algorithms for the problem defined in Section 3, named *Inter-Location Based algorithm (ILB)* and *Inter-Meeting Time Based algorithm (IMTB)*$^{\star 2}$.

### 4.1 Preliminaries

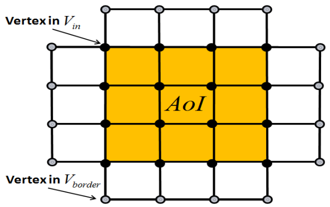
Our basic idea is selecting nodes which have higher probabilities to visit distinct sensing locations in the specified AoI within a time interval $T$, prior to other nodes. More specifically, we increase the expected probability of each sensing location in the AoI

⋆1 We assume that each road can be divided into integer number of segments with length $\Delta$.

⋆2 We suppose that all algorithms are executed by the server $s$ in a centralized fashion.




(a) Original service area graph




(b) Reduced service area graph

**Fig. 2** An example of service area graph with AoI and its reduction

visited by mobile nodes so as to exceed the required ratio $\alpha$ by incrementally selecting nodes with higher probability one by one.

The proposed algorithms depend on the probability $Prob(u, t, v_0^u, v_t)$ of each node $u$ with initial location $v_0^u$ to visit a location $v_t$ at time $t$ ($0 \leq t \leq T$), that we call the *vertex-coverage probability*, hereafter. In order to easily calculate the vertex-coverage probability, without loss of generality, we suppose that the graph $G = (V, E)$ for the service area is given by a grid of sensing locations (vertices) with a uniform spacing between neighboring vertices and only vertical and horizontal edges (here, each edge is bi-directional), as shown in Fig. 2(a) and that mobile nodes move according to a discrete Markov model on this grid where moving probability at each vertex to each of its four neighbors (up, down, left, and right) is the same (i.e., 0.25) ⋆1. 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 $i$-th state represents the probability that $u$ is at vertex $x_i$. Let $U_0 (\subseteq U)$ denote the set of nodes which are located in the target AoI at time 0.

#### 4.1.1 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.

$$\boldsymbol{v_0^u} = (p_0, p_1, ..., p_N) \quad (4)$$

where

$$p_j = \begin{cases} 0 & (j \neq i) \\ 1 & (j = i) \end{cases} \quad (5)$$

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

$$Prob(u, t, v_0^u, x_k) = [\boldsymbol{v_0^u} \times \boldsymbol{P}^t]_k \quad (6)$$

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

#### 4.1.2 Reduction of probability matrix size

If the target service area contains a lot of sensing locations, the probability matrix $\boldsymbol{P}$ will be large, resulting in serious computational overhead in the server $s$.

Therefore, we adopt the following heuristic that reduces 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 but neighboring to it. Here, note that $N \gg M + L$ holds if the AoI is much smaller than the whole service area.

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 $V_{border}$ denote a set of vertices in $V_{out}$ that have nodes of $V_{in}$ as neighbors. $V_{border}$ is defined by the following equation.

$$V_{border} = \{x \mid x \in V_{out} \land \exists y, (x, y) \in E \land y \in Vin\} \quad (7)$$

The vertices that belong to $V_{border}$ is illustrated in Fig. 2(b).

Our idea of reducing the probability matrix is utilizing the fact that we are only interested in the vertex-coverage probability for each vertex of $V_{in}$. Thus, we modify the probability of movement at each vertex in $V_{border}$ so that 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 $V_{in} \cup V_{border}$. Consequently, we define the new probability matrix $\boldsymbol{P'}$ for vertices of $V_{in} \cup V_{border}$. Let denote vertices of $V_{in}$ by $y_1, y_2, ..., y_M$ and those of $V_{border}$ by $y_{M+1}, y_{M+2}, ..., y_{M+L}$. Then we define $i$-th row and $j$-th column

⋆1 We can treat any directed graph with arbitrary probability on each edge for representing a realistic urban district service area.

element $p'_{i,j}$ of $\boldsymbol{P'}$ by the following equation.

$$p'_{i,j} = \begin{cases} p_{i,j} \quad (i \leq M \vee j \leq M) \\ 1 - p_{i,k} \quad (\exists k(k \leq M \wedge p'_{i,k} > 0) \wedge M < i \wedge i = j) \\ 0 \quad (M < i \wedge M < j \wedge i \neq j) \end{cases} \tag{8}$$

Here, $p_{i,j}$ is the probability of the corresponding edge in the original matrix $\boldsymbol{P}$. By using $\boldsymbol{P'}$ instead of $\boldsymbol{P}$ in equation (6), we can calculate the vertex-coverage probability for each vertex in AoI with much lower overhead.

### 4.2 Inter-Location Based Algorithm (ILB)

ILB uses the distance between nodes as a metric to select a set of mobile nodes. We denote the distance between 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 ($\alpha$, $T$)-covered. The above statement is defined as follows.

$$\textbf{minimize } |U'| \textbf{ subject to } (10) - (11) \tag{9}$$

$$d_{u,u'} \geq d_{th}, \forall u, u' \in U' \tag{10}$$

$$\forall v \in AoI, SetProb(v, U', T) \geq \alpha \tag{11}$$

We determine the threshold $d_{th}$ as follows. The value of $d_{th}$ should be dependent on three parameters: total number of time steps $T$, required coverage ratio $\alpha$, and the maximum distance $d_{max}$ that is the largest distance between initial locations of two nodes in $U_0$. Intuitively, as $T$ increases and/or $\alpha$ decreases, the number of selected nodes should decrease. On the contrary, as $T$ decreases and/or $\alpha$ increases, the number of selected nodes must be increased to meet the $(\alpha, T)$-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.

$$d_{th} = \min(\frac{T}{\alpha \cdot d_{max}}, d_{max}) \tag{12}$$

where, $d_{max} = \max_{u,u' \in U_0} \{d_{u,u'}\}$.

ILB algorithm selects a node from $U_0$ whose distance to already selected nodes is no less than $d_{th}$, and repeats this process until the selected nodes $(\alpha, T)$-covers the AoI.

As we explained in this section, 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 more efficient node selection taking into account the value of $T$, we propose an inter-meeting time based algorithm which uses expected meeting time between nodes as a metric in the next subsection.

### 4.3 Inter-Meeting Time Based Algorithm (IMTB)

IMTB uses the meeting time between nodes as a metric to select a set of mobile nodes. This meeting time metric reflects the probability of nodes to visit 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 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 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 ($\alpha$, $T$)-covered. The above statement is defined as follows.

$$\textbf{minimize } |U'| \textbf{ subject to } (14) - (15) \tag{13}$$

$$mt_{u,u'} \geq mt_{th}, \forall u, u' \in U' \tag{14}$$

$$\forall v \in AoI, SetProb(v, U', T) \geq \alpha \tag{15}$$

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.

$$mt_{u,u'} = \begin{cases} \min_{t \in MT_{u,u'}} \{t\} \quad (MT_{u,u'} \neq \emptyset) \\ \infty \quad (MT_{u,u'} = \emptyset) \end{cases} \tag{16}$$

where $MT_{u,u'}$ is a set of expected meeting time between $u$ and $u'$ during time period $T$ in the case that they meet in the AoI and is defined by the following equation.

$$MT_{u,u'} = \{t \mid Prob(u, t, v_0^u, v_t) > 0 \ \wedge \ Prob(u', t, v_0^{u'}, v_t) > 0, \\ 0 \leq t \leq T, v_t \in AoI\} \tag{17}$$

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

$$mt_{th} = \min(\frac{T}{\alpha \cdot mt_{max}}, mt_{max}) \tag{18}$$

where, $mt_{max} = \max_{u,u' \in U_0} \{mt_{u,u'} : mt_{u,u'} \neq \infty\}$.

IMTB algorithm selects a node from $U_0$ whose meeting time with already selected nodes is no less than $mt_{th}$, and repeats this process until the selected nodes $(\alpha, T)$-covers the AoI.

### 4.4 Extended versions of IBL and IMTB Algorithms

As we described in the previous two subsections, ILB and IMTB algorithms apply the selection process only to a set of nodes located inside AoI at time 0, $U_0$, and do not consider the nodes outside AoI. The number of nodes inside AoI at time 0 may not be sufficient to guarantee the $\alpha$-coverage of AoI in time period $T$, if it is too small. To cope with this situation, we extend the algorithms to add more nodes located outside AoI in the selection process. The addition of a node located outside AoI should be dependent on its initial location and the time period $T$. In other words, it should be dependent on shortest distance from the added node to at least one vertex in $V_{border}$. Intuitively, if this distance of a new added node is more than $T$, then the node will not visit any locations in AoI within the time period $T$. So, the distance must be less than or equal to $T$. Avoiding the number of added nodes to be very large, we add only nodes if the shortest distance to locations of $V_{border}$ is less than or equal to $\lfloor \frac{T}{3} \rfloor$. This means that each added node has a probability to visit $2\lfloor \frac{T}{3} \rfloor$ locations within the time period $T$. We denote the extended versions of ILB and IMTB by *eILB* and *eIMTB*, respectively.

## 5. Performance Evaluation

### 5.1 Simulation Environment

The QualNet [9] simulator is used with input parameters as listed in table 1, such as service area size, number of nodes, node speed, etc[⋆1]. In addition, the node mobility is based on discrete Markov model which was described in Section 4 where the time step is 50 seconds. The service area is represented as a grid of sensing locations arranged with uniform spacing. In our simulation, the spacing is 50 *meters*. We selected the AoI as a rectangle region where its position is selected at random within the service area and its ratio to the service area called *AoI-Ratio* is selected from $\{0.01, 0.5\}$. The probability of a node at a location move to one of its neighboring locations is set uniformly to 0.25.

⋆1 We did not simulate communications between mobile nodes and the server $s$, but we used QualNet just for reproducing mobility of nodes.

**Table 1** Configuration Parameters.

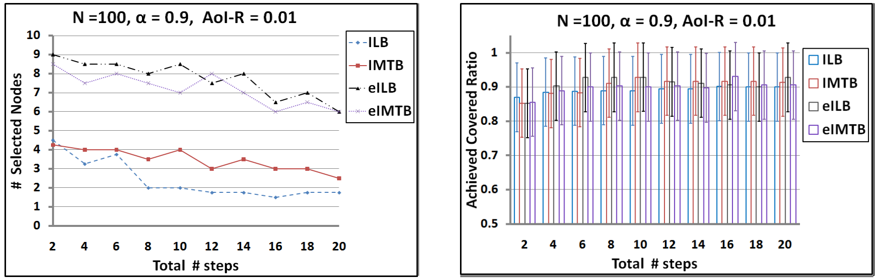
| Configuration parameter | Value in simulation |
|---|---|
| # nodes | 25 to 200 |
| Node speed | 1 m/s |
| Service area size | $500m \times 500m$ |
| Required coverage, $\alpha$ | 0.2, 0.4, 0.5, 0.6, 0.8, 0.9 |
| AoI-Ratio | 0.01, 0.5 |
| $\Delta$ | 50 m |
| Total # steps (time period), T | 2, 3, 4,..., 20 |

We measured the performance of ILB and IMTB and the extended algorithms eILB and eIMTB 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, because we cannot measure the probability of each sensing location to be actually visited by nodes for each simulation run. We regard 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. We repeated every simulation experiment 5 times with different node distribution, then averaged the results.

### 5.2 Simulation Results

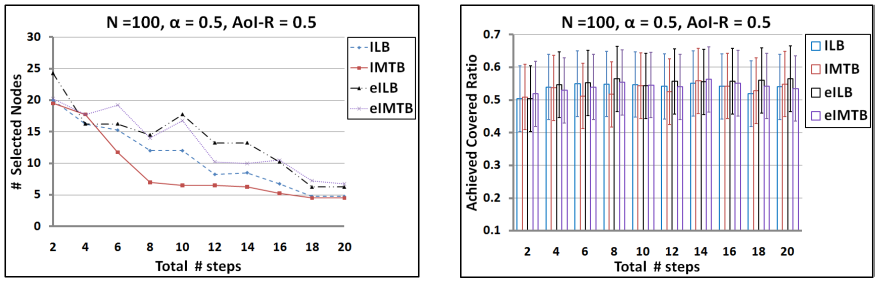
We show the simulation results in Fig. 3, Fig. 4, Fig. 5, and Fig. 6.

Fig. 3 shows the performance for different number of time steps with very small size AoI and high required coverage ratio. Here, the number of nodes was 100, the required coverage $\alpha$ was 0.9, and the AoI-Ratio was 0.01. In Fig. 3(a), the number of selected nodes decreased as the total number of steps increased. This is because, as the total number of steps increases, the distance and meeting time threshold increases in proportion the total number of steps. The number of selected nodes for ILB is lower than other algorithms. This is because, the AoI-Ratio is very small and the number of selected




(a) Number of selected nodes vs. Total number of steps
(b) Achieved coverage Ratio vs. Total number of steps

**Fig. 3** Performance for small AoI and high required coverage ratio

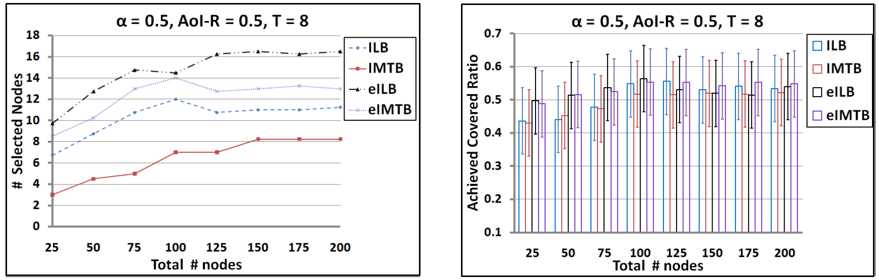



(a) Number of selected nodes vs. Total number of steps
(b) Achieved coverage Ratio vs. Total number of steps

**Fig. 4** Performance for medium AoI and medium required coverage ratio

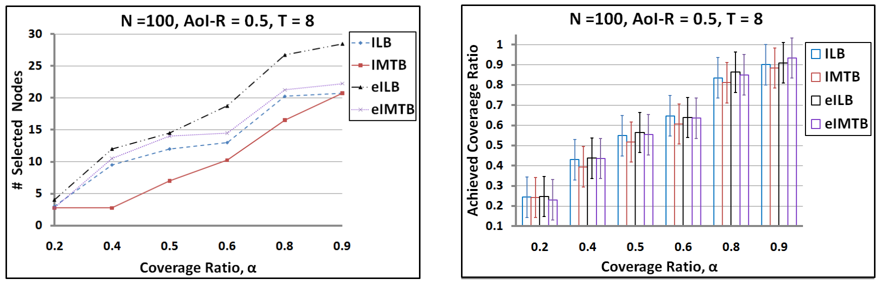
nodes reduces as the distance between nodes reduces in ILB. In Fig. 3(b), the required coverage is satisfied by neither algorithms when the total number of steps was 2 and 4. While for higher number of steps, all algorithms satisfied the required coverage. The variance of ILB is the smallest among all.

Fig. 4 shows the performance for different number of time steps with medium size AoI and medium required coverage ratio. The number of nodes was 100, the required coverage $\alpha$ was 0.5, and the AoI-Ratio was 0.5. In Fig. 4(a), the number of selected nodes decreases as the total number of steps increases. This is because, as the total number of steps increases, the distance and meeting time threshold increases proportionally




(a) Number of selected nodes vs. Total number of nodes
(b) Achieved Coverage Ratio vs. Total number of nodes

**Fig. 5** Number of selected nodes and achieved coverage ratio vs. Total number of nodes




(a) Number of selected nodes vs. Required coverage ratio
(b) Achieved Coverage Ratio vs. Required coverage ratio

**Fig. 6** Number of selected nodes and achieved coverage ratio vs. Required coverage ratio

to the total number of steps. The number of selected nodes for IMTB is lower than other algorithms. This is because, the required coverage is medium and the meeting time increased when AoI-Ratio is medium. In Fig. 4(b), the required coverage is satisfied by all algorithms and the variance of IMTB is the smallest among all.

Fig. 5 shows the performance for different number of nodes with medium size AoI, medium required coverage ratio, and medium number of time steps. The number of nodes was 100, the required coverage $\alpha$ was 0.5, and the total number of steps was 8. In Fig. 5(a), when number of nodes was 25 to 125, the number of selected nodes increased as number of nodes increased. This is because, as the number of nodes increases, the number of the selected nodes increases when the number of nodes was higher than 125, the number of selected nodes was fixed. This is because, the number of selected nodes is bounded by the number of nodes needed to satisfy the required coverage. In Fig. 5(b), the required coverage is not satisfied by ILB and IMTB algorithms when the total number of nodes was between 25 to 75. For higher number of nodes, all algorithms satisfied the required coverage. The variance of eIMTB is the smallest among all when the total number of nodes was 25 to 75 while the variance for IMTB is the smallest among all when the total number of nodes was higher than 75.

Fig. 6 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. 6(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 for IMTB is lower than other algorithms. In Fig. 6(b), the required coverage is satisfied by all algorithms and the variance of IMTB is the smallest among all.

We summarize the simulation results as follows.

- ILB can select the smaller number of nodes to meet the required coverage with small variance for small AoI, than others.
- IMTB can select the smaller number of nodes to meet the required coverage with small variance for medium and large AoI, than others.
- When only small number of nodes are initially located in the AoI, the extended algorithm eIMTB can meet the required coverage with small variance.

## 6. Conclusion

In this paper, we presented the $(\alpha, T)$-coverage problem in people centric sensing with motivating application scenarios. We formulated this problem as an optimization problem with the objective of minimizing the number of selected nodes to meet the demanded coverage ratio $\alpha$ within a query interval time $T$. To effectively estimate the probability that each mobile node visits, we used a discrete Markov model to model the mobility of pedestrians with sensors. Based on this estimation mechanism, we proposed two heuristic algorithms. First, inspired by the fact that each node in the specified area of interest (AoI) will explore the proximity sensing locations with higher probability, we proposed the inter-location based algorithm that selects nodes with their inter-distance more than a threshold. Second, taking into account the query interval time $T$, we proposed the inter-meeting-time based algorithm that selects nodes whose possible future meeting time is more than a threshold within $T$. Last, we extended the algorithms to select more nodes so as to cope with the case that AoI does not have sufficient number of nodes. Our simulation results showed that the proposed algorithms achieve $(\alpha, T)$-coverage with good accuracy for a variety of values of $\alpha$, $T$, and AoI size, and that the inter-meeting time based algorithm selects smaller number of nodes without deteriorating coverage accuracy for medium and large AoI.

In this paper, we considered only the case that a single query is issued at one time. In the future work, we will try to make the proposed algorithms adaptive to multiple queries case. In addition, we will apply our algorithms to more realistic urban sensing scenarios with more realistic mobility models.

## References

1) Campbell, A., Eisenman, S., Lane, N., Miluzzo, E., Peterson, R., Lu, H., Zheng, X., Musolesi, M., Fodor, K. and Ahn, G.: The Rise of People-Centric Sensing, *IEEE Internet Computing Special Issue on Mesh Networks*, Vol.12, No.4, pp.12–21 (2008).
2) Tuulos, V., Scheible, J. and Nyholm, H.: Combining Web, Mobile Phones and Public Displays in Large-Scale: Manhattan Story Mashup, *Proc. of the 5th Int. Conf. on Pervasive Computing* (2007).
3) Hull, B., Bychkovsky, V., Zhang, Y., Chen, K., Goraczko, M., Miu, A., Shih, E., Balakrishnan, H. and Madden, S.: CarTel: A Distributed Mobile Sensor Computing System, *Proc. of the 4th Int. Conf. on Embedded Networked Sensor Systems*, pp.125–138 (2006).
4) Murty, R., Gosain, A., Tierney, M., Brody, A., Fahad, A., Bers, J., and Welsh, M.: CitySense: A Vision for an Urban-Scale Wireless Networking Testbed, *Proc. of IEEE Int. Conf. on Technologies for Homeland Security* (2008).
5) Eisenman, S., Lane, N., Miluzzo, E., Peterson, R., Ahn, G. and Campbell, A.: MetroSense Project: People-Centric Sensing at Scale, *Workshop on World-Sensor-Web*, pp.6–11 (2006).
6) Lua, H., Lane, N., Eisenmanb, S. and Campbella, A.: Bubble-sensing: Binding Sensing Tasks to The Physical World, *Journal of Pervasive and Mobile Computing*, Vol.6, No.1, pp. 58–71 (2010).
7) Shi, J., Zhang, R., Liu, Y. and Zhang, Y.: PriSense: privacy-preserving data aggregation in people-centric urban sensing systems, *Proc. of IEEE INFOCOM2010* (2010).
8) Vercellis, C.: A Probabilistic Analysis of The Set Covering Problem, *Annals of Operations Research*, Vol.1, No.3, pp.255–271 (1984).
9) Scalable Network Technologies Inc.: *Qualnet 4.0 network simulator* (2007).