A Multi-Objective Indoor Localization Service for Smartphones

Andreas Konstantinidis
Department of Computer Science and Engineering, Frederick University
Nicosia, Cyprus
a.constantinides@frederick.ac.cy

Aphrodite Demetriades
Department of Computer Science and Engineering, Frederick University
Nicosia, Cyprus
st009195@stud.fit.ac.cy

Savvas Pericleous
Department of Computer Science and Engineering, Frederick University
Nicosia, Cyprus
s.pericleous@frederick.ac.cy

ABSTRACT

The bulk of indoor localization applications currently rely on either server-side, cloud-based services that raise critical data-disclosure concerns (e.g., reveal user's location to a central entity), or client-side services that introduce serious performance concerns (e.g., consuming precious smartphone battery and network bandwidth during content uploads). In this paper, we present a novel Multi-objective Indoor Localization Service (MILoS) that provides a fine-grained, energy-efficient indoor localization using only a subset of WiFi-based localization data on the client-side, maintaining user's privacy at the same time. MILoS follows a fingerprinting-based indoor localization model that concurrently optimizes several conflicting objectives (i.e., minimizes the smartphone's energy consumption and maximizes the area coverage induced by WiFi fingerprints importance), using a Multi-Objective Evolutionary Algorithm based on Decomposition (MOEA/D). To the best of our knowledge this is the first time that the WiFi fingerprinting approach is used in a Multi-Objective Optimization setting for indoor localization. We assess our proposed model using real datasets and realistic mobility scenarios.

KEYWORDS

indoor navigation, fingerprint-based localization, smartphones, multi-objective optimization, evolutionary

ACM Reference Format:
Andreas Konstantinidis, Aphrodite Demetriades, and Savvas Pericleous. 2019. A Multi-Objective Indoor Localization Service for Smartphones. In *The 34th ACM/SIGAPP Symposium on Applied Computing (SAC '19), April 8–12, 2019, Limassol, Cyprus.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3297280.3297395

1 INTRODUCTION AND RELATED WORK

Technological advances and the widespread use of smartphone devices led to a surge of location-aware applications and services for mobile phones [1]. Users can now, for example, employ their smartphones to post content to social media such as text, selfies/photos, or videos tagged with their location in real time (Google Latitude, Facebook Places, etc.). Indoor localization systems, in particular, have been gaining relevance as people tend to spend a considerable amount of their daily lives in indoors environments (shopping centres, universities, libraries, museums, hospitals, etc.), and the demand for various indoor location based services, such as, inventory management, in-building guidance and navigation has been growing [2, 3].

The satellite-based Global Positioning System (GPS), uses radio signals from satellites to offer super fine accuracy outdoors; moreover, the localization is performed directly on the smartphone, so location-sensitive information is not shared. But, it requires considerable energy, and its accuracy depends on additional factors, such as atmospheric conditions and signal blockage, making it unsuitable for indoor spaces [4].

Numerous other solutions for indoor localization, including Ultrasound, Bluetooth, LiFi and Infrared technologies can be found in the literature [5]. Although these can achieve location-accurate results, most of them carry a high cost in dedicated equipments and installation. The most common and cost-effective infrastructure that is already in place today in many buildings and can be used for indoor localization on users smartphones is Wi-Fi.

Various cloud-based *Indoor Positioning Services* (**IPS**), such as Skyhook, Google Indoor Maps, Navizon, Infsoft, IndoorAtlas, MazeMap, Indoo.rs and Anyplace [6] enable indoor location based applications by offering indoor models comprising of floors/buildings digital maps and *Points-Of-Interest* (**POI**s), along with cloud-based geolocation databases of radio signal intensities from mobile Cell Towers and WiFi Access Points (**AP**s).

The so called *WiFi Fingerprint*, is an array of values of the *Received Signal Strength* (**RSS**) of nearby WiFi APs at a specific location of the indoor area. The WiFi Fingerprints measured at a dense set of points of the indoor digital map are recorded and then, in offline mode, joint into a matrix, that is known as the WiFi *RadioMap*. Subsequently, a user that wishes to perform indoor localization can employ a smartphone to capture the RSS fingerprint of his/her current position and transfer it to the server, where it can be compared against the RadioMap, in order to find the best approximation, using, for example, the K-Nearest-Neighbour (**KNN**), or the *Weighted K-Nearest-Neighbour* (**WKNN**) [2] algorithms. The idea behind the KNN approach is to calculate the Euclidean distances between the user's currently observed fingerprint and all fingerprints in the RadioMap and select the K nearest ones. The user's location is then estimated as a convex combination of the locations of these K selected fingerprints. This method assigns the same importance to all K selected fingerprints regardless of their distance to the user's observed fingerprint. By assigning a different weight to each of the

SAC '19, April 8–12, 2019, Limassol, Cyprus

ACM ISBN 978-1-4503-5933-7/19/04…$15.00
https://doi.org/10.1145/3297280.3297395

K selected fingerprints based on this distance, the WKNN method obtains improved location accuracy results [6].

IPS often rely on one of the following RadioMap based approaches: *(i)* the WiFi RadioMap *Server-Side* approach, in which the smartphone user senses and uploads its current observed RSS fingerprint to the server, which is responsible to forward to the user the calculated location, and *(ii)* the WiFi RadioMap *Client-Side* approach, in which the whole RadioMap is downloaded to the smartphone and the localization takes place in-situ. In the former case, the user's smartphone requires minimal energy and network communication with the server, but privacy issues might arise, as the location of the user is revealed to the server. Moreover, the wireless network topology design of the building might lead to intermittent internet connectivity [7, 8], leading to communication breakdown between user and server. In the Client-Side case, although the user's location privacy is guaranteed, there is high battery consumption and network bandwidth overhead, as the user needs to download, albeit once, the whole WiFi RadioMap (which, in most cases, can be huge), and also performs all the localization calculations on the smartphone.

Various privacy-enhancing localization techniques that appear in the literature are based on the following concepts: *(i)* sanitized locations [9, 10], where a set of fake locations (sanitized) of a user is also reported; *(ii)* spatial cloaking [11, 12], which tries to blur a user's exact location into a cloaked area that satisfies the user's privacy requirements; and *(iii)* space transformations [3, 13, 14], where the location of a user is transformed into another space in which his/her exact or approximate spatial relationships are maintained. These techniques try to mislead the server about the user's actual location by providing and requesting, so called, noisy data which guarantee more privacy, in the expense though, of increased resource consumption in terms of the smartphones battery power and network bandwidth.

In the literature, there are several techniques for dealing with intermittent connectivity [7]. One of the most popular is prefetching [8], that is, downloading and locally storing data, so that future requests for that data can be served in the event of a network failure. But again, the cost of prefetching in terms of increased energy and network bandwidth can be substantial.

From the user's point of view, it would be desirable to *(i)* maximize the accuracy of the indoor localization without *(ii)* deteriorating the resources of the smartphone device. The main idea put forward in this paper is to follow the paradigm of the Client-Side approach, but instead of the user downloading the whole RadioMap (as discussed above) from the IPS, the user just downloads a partial RadioMap *pRM* that is calculated offline through a fingerprint selection optimization process. The selected *pRM* is representative enough to "cover" the indoor area in order to provide the required localization accuracy, as well as minimize the energy consumption at the same time and inherently maintain the user privacy. These two objectives, however, are conflicting with each other and the respective problem must be therefore treated within the context of Multi-Objective Optimization.

A *Multi-Objective Optimization Problem* (**MOP**) can be mathematically formulated as

$$\min \ F(X) = (f_1(X), \dots, f_k(X)), \ \text{subject to } X \in \Omega, \qquad (1)$$

where Ω is the decision space and $X \in \Omega$ is a decision vector. $F(X)$ consists of k objective functions, and $\mathfrak{R}^k$ is the objective space. Improving on one objective may lead to deterioration of another, thus, no single solution exists that can optimize all objectives simultaneously. The best trade-off solutions, called the set of Pareto optimal (or non-dominated) solutions, is often required by a decision maker.

A vector $u = (u_1, \dots, u_k)$ is said to dominate another vector $v = (v_1, \dots, v_k)$, denoted as $u \prec v$, iff $\forall i \in \{1, \dots, k\}$, $u_i \leq v_i$ and $u \neq v$. A feasible solution $X \in \Omega$ of problem (1) is called *Pareto optimal solution*, iff $\nexists Y \in \Omega$ such that $F(Y) \prec F(X)$. The set of all Pareto optimal solutions is called the Pareto Set, denoted as

$$PS = \{X \in \Omega | \nexists Y \in \Omega, F(Y) \prec F(X)\}.$$

The image of the PS in the objective space is called the *Pareto Front* $PF = \{F(X) | X \in PS\}$.

Multi-Objective Evolutionary Algorithms (**MOEA**s) can obtain an approximate PF of a MOP in a single run by using various operators to iteratively generate a population of such solutions. The aim is to produce a diverse set of non-dominated solutions that is as close as possible to the real PF. Several techniques were proposed for improving their performance, such as nithcing techniques for improving diversity and/or local search methods for improving convergence [21]. For example, MOEAs, such as, NSGA-II [17] are based on Pareto Dominance, while *Multi-Objective Evolutionary Algorithm based on Decomposition* (**MOEA/D**) [18] is an example of a decompositional MOEA, that relies on conventional aggregation approaches to decompose a MOP into a number of scalar Single-Objective Optimization sub-problems, which are then solved simultaneously using neighborhood information, each time a new solution is generated. For recent surveys on the state of the art of MOEAs please refer to [19]. Finally, the decision maker chooses the single preferred solution from the PF, according to some external information relevant to the user's preferences, e.g., the user might be willing to sacrifice location accuracy in order to spend only a specified low percentage of battery power.

To the best of our knowledge there is no prior work that solves an indoor localization based on WiFi fingerprints problem in a Multi-Objective Optimization setting using MOEA.

In this paper, we define and formulate the *Multi-Objective Fingerprint Selection Optimization Problem* (**MO-FSOP**) for indoor localization that aims to optimize the area coverage and the smartphone's energy consumption, simultaneously, by calculating an offline selected partial RadioMap . We propose a novel *Multi-objective Indoor Localization Service* (**MILoS**) that adapts the well-known MOEA/D to solve the proposed MOP. We provide a series of experimental studies to evaluate the effectiveness of our proposal and in particular, to verify that as energy consumption / area coverage of the selected partial RadioMap increases, so does the localization accuracy calculated by the user's smartphone in-situ, given an observed fingerprint on the fly. Let us note, that once the user downloads from the server the selected partial RadioMap (corresponding to some chosen solution of the output PF of the MOEA/D), there is no further communication with the server. In particular, no information on the user's location is shared with the server.

The rest of the paper is organized as follows. The System Model and the definition and formulation of MO-FSOP, are presented next in Section 2. The main steps of the proposed MILoS approach are explained in Section 3. In Section 4, the performance of the proposed method is evaluated on real datasets with mobility scenarios and compared against NSGA-II, the state-of-the-art in MOEAs based on Pareto-dominance. Finally, Section 5 concludes the paper and discuss possible future research directions.

2 SYSTEM MODEL AND PROBLEM FORMULATION

In this section, we first outline the adopted system model and then formulate the MOP. The adopted symbol notations is summarized in Table 1.

2.1 System Model

We assume that a weighted graph $G = (P, E)$ representing the connectivity of the building I is available on IPS s. The set of nodes P comprises of the POIs, which refer to rooms, offices, toilets, intersections, elevators, staircases, hallways etc., in the building as these have been provided by architects or crowdsourcers and the set of edges E comprises of the corridors, physical pathways, etc., aligned to floors inside the building, linking these POIs between them. The weight on each edge represents the distance of the physical transition between two nodes. The graph will allow us to compute and assign degrees of importance to each POI and also calculate shortest paths between any two of them. We note that paths between nodes are calculated using the graph-distance cost [15, 16], which reflects the topological constraints and physical entities of a building (e.g, elevators, corridors, walls, etc.), given that the Euclidean distance between two POIs may not always be appropriate to use.

We also assume that the indoor area I contains a finite set Q of N locations that are partially covered by a set $\{ap_1, ap_2, \cdots, ap_M\}$ of Wi-Fi APs. Each ap_i has a unique ID (i.e., MAC address) that is publicly broadcasted and passively received by anyone moving in I. The signal intensity at which the ID of ap_i is received at location $q \in Q$ is called the RSS of ap_i at q, with the value -110 indicating that ap_i is out of reach. The set of all offline measured RSS values and the AP-IDs captured at each point q is the fingerprint V_q. The $N \times M$ matrix of RadioMap (RM) comprises of all fingerprints V_q measured at all locations $q \in Q$ and is stored on IPS s.

We further assume that a pre-processing offline procedure is available so that any selected subset $X \subset P$ of POIs can be associated with a subset of rows of the RM matrix, i.e., a partial RadioMap (pRM). This can be achieved, for example, by linking each WiFi fingerprint V_q in RadioMap (RM) to its closest (in terms of Euclidean distance) POI. Moreover, the user u has to have installed on the smartphone a localization function *loc()*.

To perform localization at current position l in the indoor area I, user u employs a smartphone to capture the observed RSS fingerprint V_l and then calculates $loc()$ with input a partial RadioMap (pRM) that u has downloaded once from IPS s and the observed fingerprint V_l. We define the Localization Error $LocE_l$ for location l to be the distance between the outcome of the localization functions $loc(RM, V_l)$ and $loc(pRM, V_l)$.

2.2 MO-FSOP Formulation

Given the representation graph $G = (P, E)$ of a building I, the RadioMap RM from IPS s, and a procedure to associate subsets of P with partial RadioMaps, the *Multi-Objective Fingerprint Selection Optimization Problem* (**MO-FSOP**) can be stated as follows: select a subset $X \subset P$ of POIS with associated partial RadioMap pRM, such that

- pRM maximizes the area coverage of I, so that a smartphone user u located at any point $l \in I$, maximizes the obtained localization accuracy, i.e., minimizes the localization error $LocE_l$, and at the same time,
- minimizes the energy required to download the partial radiomap pRM associated with X from s, .

Table 1: Notation used throughout this work

Notation	Description
s, u	Indoor Positioning Service, smartphone user
P, X, n	set of all POIs in digital map, subset of P, $\lvert X\rvert$
E	set of physical pathways between neighbouring POIS
I, Q	Indoor area, pre-defined set of N points in I
ap_i, AP, M	Access Point i, set of all ap_i, $\lvert AP\rvert$
V_q	offline recorded fingerprint at point $q \in Q$ (MAC and RSS of its covering AP)
RM	$N \times M$ RadioMap matrix of all $V_q (q \in Q)$, stored on s
pRM	$N' \times M$ partial RM associated with X, selected to be downloaded by u
V_l	observed fingerprint by u at current location $l \in I$
$loc(*, V_l)$	localization output for location $l \in I$, for some $* \subseteq RM$
$LocE_l$	localization error for location $l \in I$, i.e., distance between points $loc(RM, V_l)$ and $loc(pRM, V_l)$

Given any two nodes r and t in P, let $D(r, t)$ denote the length of the shortest path between them, i.e., the distance between r and t in I. Denote by $maxD$, the maximum such length of the shortest path between any pair of nodes in P. Let X be the selected subset of n POIs in I. The *spread* of X within I is expressed as the normalized average distance between all possible pairs of nodes of X and is denoted by

$$Spread(X) = \frac{\overline{D}_X}{maxD} \tag{2}$$

where,

$$\overline{D}_X = \frac{\sum\limits_{r \neq t \in X} D(r, t)}{n(n-1)}.$$

Let $\sigma_{r,t}$ denote the total number of shortest paths from node r to node t and $\sigma_{r,t}(p)$ the number of those paths that pass through node p. We will use the Betweenness Centrality measure $B(p)$ to denote the importance of a POI p. This measures the extent to which p lies on paths between other POIs and is defined as the number of shortest paths in G from all vertices to all others that pass through p, i.e.,

$$B(p) = \sum_{r \neq p \neq t \in P} \frac{\sigma_{r,t}(p)}{\sigma_{r,t}}.$$

Denote by $maxB$ and $minB$, the maximum and minimum values respectively of the Betweenness function B over all nodes in P. The Normalized Betweenness Centrality value $NB(p)$ for node p is defined as

$$NB(p) = \frac{B(p) - minB}{maxB - minB}.$$

The *importance* of the selected subset X within I is expressed as the average of the Normalized Betweenness Centrality value $NB(p)$

over all nodes p in X and is denoted by

$$Importance(X) = \frac{1}{n} \sum_{p \in X} NB(p). \quad (3)$$

A representative **Area Coverage** measurement of the selected subset X of POIs with respect to the indoor space I, is expressed as a weighted combination, with weight $\gamma \in (0, 1)$, by

$$AreaCoverage(X) = \gamma\, Spread(X) + (1 - \gamma)\, Importance(X). \quad (4)$$

This Area Coverage definition aims to capture the most frequent user mobility patterns, by selecting the nodes that have higher probability to be visited (as represented by the use of the Betweenness function in Eq. 3), and at the same time, incorporate a dispersion factor (as represented by the Spread function used in Eq. 2 to account for more irregular mobility patterns.

The **Energy consumption** of downloading from s, the partial radiomap pRM associated with selected set X of POIs, is defined as

$$Energy(X) = \frac{N'}{N} \quad (5)$$

where N' is the number of Wifi Fingerprints in pRM and N the total number of registered fingerprints in RM.

The fingerprint selection optimization process then aims to:

$$minimize \quad F(X) = (f_1(X), f_2(X)), \quad \text{subject to } X \subseteq P, \quad (6)$$

with **objective functions**

$$f_1 = Energy(X), \qquad f_2 = -AreaCoverage(X), \quad (7)$$

defined above in Equations 5 and 4, respectively.

3 PROPOSED APPROACH

In this section we present MILoS, our proposed Multi-Objective Indoor Localization Service. First we explain how to utilize MOEA/D to solve the Multi-Objective Fingerprint Selection Optimization Problem (MO-FSOP) formulated in Section 2.2 above.

3.1 Multi-Objective Optimization Module

MOEA/D accepts as input a representation graph $G = (P, E)$ of an indoor area I, the registered RadioMap RM, and a procedure of associating subsets of P with partial RadioMaps pRM. It outputs a set of trade-off candidate solutions, i.e., points of the Pareto Front PF, that concurrently optimize the problems objectives (Energy and Area Coverage). Each solution X is a subset of the set P of POIs and is associated to a partial Radiomap pRM.

MOEA/D requires first some pre-processing procedures at **Step 0**, before initiating the main part of the algorithm. The main steps are briefly summarized and discussed next.

Encoding Representation: A solution X of MO-FSOP is represented by a a binary vector of size equal to the number $|P|$ of POIs, whose components signify whether a POI is included in X or not.

Decomposition: Initially, the MO-FSOP is decomposed into a number of N scalar subproblems using the Tchebycheff approach as originally proposed in [18]. Given the objective vector $F(X) = (f_1(X), f_2(X))$ of Equation 6, weight vector λ^i, $(1 \leq i \leq N)$, that remains fixed for each subproblem for the whole evolution and a reference point $z^* = (z_1, z_2)$, which is a vector with all the best values z_k found so far for each objective f_k, the objective function of a subproblem i is stated as:

$$g(X|\lambda^i, z^*) = \sum_{k=1}^{2} |\lambda^i_k f_k(X) - z_k|. \quad (8)$$

Algorithm 1 Solving MO-FSOP using MOEA/D

Input:
- an instance of MO-FSOP (see Section 2.2);
- the number N of decomposed subproblems = population size;
- uniformly spread weight vectors $\{\lambda^1, \dots, \lambda^N\}$;
- the size of the neighbourhood T of each subproblem;
- tournament size t, crossover rate c_r and mutation rate m_r;
- a termination criterion: max number of generations = gen^m;

Output: a set of non-dominated solutions PF.

Step 0 - Pre-processing:
Decomposition: into a set of N single-objective subproblems having weights $\{\lambda^1, \dots, \lambda^N\}$ respectively;
Neighborhoods: Define B^i for the i^{th} subproblem to include the T closest weight vectors of λ^i.
Setup: Set $PF := \emptyset$; $gen := 0$; $IP_{gen} := \emptyset$;

Step 1 - Initialization: Set Pareto Front $PF = \emptyset$ and reward vectors $R^i = 0$. For each subproblem, umniformly randomly generate and evaluate an initial internal population $IP_0 = \{X^1, \cdots, X^N\}$. Set $gen = 1$.

Step 2: For $i = 1, \dots N$ **do**
Step 2.1 - Genetic Operators: For i^{th} subproblem, generate new solution Y^i using the genetic operators.
Step 2.2 - Update: Update reference point z^* and use Y^i to update IP_{gen}, PF and the neighborhood B^i of the T closest neighbor solutions of Y^i.

Step 3 - Stopping criterion: If stopping criterion is satisfied, i.e., $gen = gen^m$, **then** stop and output PF, **otherwise** $gen = gen + 1$, go to Step 2.

Neighbourhood: A neighbourhood (or subpopulation) B^i is maintained for each of the N subproblems associated with weight vector λ^i, composed of the indices of the subproblems whose associated weight vectors are the T closest (in terms of Euclidean distance) to λ^i. One expects optimal solutions in neighbouring sub-problems to be close to each other in the search space, so the exchange of genetic information should be helpful.

Step 1 - Initialization: The algorithm commences by creating an initial population $IP_0 = \{X^1, ..., X^N\}$ of solutions one for each subproblem, named Internal Population (IP) of generation $gen = 0$. The initial solutions are randomly generated and each individual is evaluated as described earlier. Set $gen = 1$;

Step 2.1 - Genetic Operation: At each generation gen, for each subproblem i with objective function $g(X^i|\lambda^i, z^*)$, the population IP_{gen} is evolved by generating a new solution Y^i, known as offspring using conventional genetic operators (i.e., Selection, Crossover and Mutation as in [18]). In particular, two parent solutions are selected from the neighbourhood B^i of subproblem i using the well-known tournament selection approach with a tournament size t. The two parent solutions are recombined using a two-point crossover to produce a new solution - the offspring - with a probability r_c. The offspring is then modified with a random mutation operator with a probability r_m. Evaluate the new solution Y^i using Eq. (8).

Step 2.2 - Update: Use solution Y^i to update the reference point $z*$, the internal population IP, the set of non-dominated solutions PF found so far and neighbourhood B^i of the sub-problem i B^i. If $i < N$ then $i = i + 1$ and goto Step 2.1. The same process is followed for all N sub-problems.
Step 3 - Stopping Criteria: If $gen = gen^m$ then terminate the algorithm and output the PF, otherwise goto Step 2.1.

3.2 Multi-objective Indoor Localization Service - MILoS

Our proposed MILoS approach for smartphone users proceeds in three steps as described next.
Graph Generation Module: We use the free and open source IPS *Anyplace* [6] for data collection. Given as input a series of queries on the Building ID and Floor number, the Anyplace system returns the Points of Interest (POIs) of the building/floor combination, and the connections between them in JSON file as well as its RadioMap RM as a text file. All this data is then reconstructed by our own system to form the representation graph $G = (P, E)$ of area I. Moreover, we define a procedure for associating subsets of POIs with partial RadioMap: each WiFi fingerprint is linked to its closest POI.
Multi-Objective Optimization Module: The reconstructed graph $G = (P, E)$ and the association map between subsets of POIs and partial RadioMaps are fed into a MOEA/D (Algorithm 1) which solves MO-FSOP offline. The MOEA/D obtains a set of non-dominated solutions (PF) that concurrently optimizes the problem's objectives (Energy and Area Coverage).
Localization Module: Given some specified user's criteria, a solution X is selected from the PF by the decision maker and its associated partial RadioMap pRM is downloaded once on the smartphone. In order to perform localization at current position l in I the user employs a smartphone to initially capture the observed RSS fingerprint V_l and then uses the WKNN localization method (with input the partial RadioMap pRM and the observed fingerprint V_l) in-situ.

Algorithm 2 MILoS - A Multi-Objective Indoor Localization Service

Step 1 - Graph Generation Module:
Input 1.1: Data from IPS Anyplace.
Output 1.1: Reconstructed representation graph $G = (P, E)$.
Output 1.2: Association between POIs and partial Radiomaps by linking each WiFi fingerprint to its closest POI.
Step 2 - Multi-Objective Optimization Module (Algorithm 1):
Input 2.1: Output 1.1 and 1.2.
Output 2.1: Pareto Front (PF) set of non-dominated solutions.
Step 3 - Localization Module:
Input 3.1: Any solution X (set of POIs) from PF of Output 2.1.
Input 3.2: Partial RadioMap pRM associated to X from Output 1.2.
Input 3.3: User's observed fingerprint V_l at location $i \in I$.
Input 3.4: A localization function $loc()$ - we use WKNN method.
Output 3.1: Calculated location $loc(pRM, V_l)$.

4 EXPERIMENTAL EVALUATION

In this section, we describe the details of our experimental methodology composed of our datasets, algorithms, algorithmic parameters, evaluation metrics and some realistic mobility scenarios. We then present the results of our MOEA/D performance evaluation and the validation of the obtained near-optimal solutions with respect to indoor localization accuracy on the mobility scenarios.

4.1 Datasets

To carry out our trace-driven experimentation, we used the following real data:

CSUCY Data: Data is collected in a typical building at the Computer Science (CS) department of the University of Cyprus using three Android devices. In particular, it consists of 45,000 reference fingerprints taken from ~120 Wi-Fi APs installed in the three floors of the *CS* and neighboring buildings. On average, 10.6 APs are detected per location. We collected our data by walking over a path that consists of 2,900 locations. The CSUCY data has a size of ~2.6 MBs. The three floors of the CS UCY building can be further characterized as follows:

- Floor 1: Normalized # of fingerprints: 715; # of POIs: 121;
- Floor 2: Normalized # of fingerprints: 686; # of POIs: 115;
- Floor 3: Normalized # of fingerprints: 752; # of POIs: 57;

4.2 Multi-Objective Evolutionary Algorithms

The proposed MOEA/D is compared with the state-of-the-art in MOEAs based on Pareto-dominance NSGA-II. NSGA-II maintains a population IP of size N at each generation gen, for gen^m generations. NSGA-II adopts the same evolutionary operators (i.e. selection, crossover and mutation) for offspring reproduction as MOEA/D. The key characteristic of NSGA-II is that it uses a fast non-dominated sorting and a crowded distance estimation for comparing the quality of different solutions during selection and to update the IP and the PF. We refer interested readers to [17] for details.

4.3 MOEA Parameters

The algorithmic parameters in the following experiments are set as follows: termination criterion gen^m=200, population size and number of subproblems N=500, crossover rate r_c=0.9, mutation rate r_m=0.05, neighbourhood size $T = 10$ and tournament size $t = 5$. Note that in our experimental studies we have used the same number of function evaluations for all methods, for fairness, and each algorithm is executed 20 times in each study. The value of $\gamma = 0.5$ was used for Equation 4. All algorithms were coded in Java programming language and run on an Intel(R) Core(M) i5 CPU 2.4GHz Windows 7 server with 4 GB RAM.

4.4 MOEA Performance Metrics

It is desirable that the obtained non-dominated set of a MOEA is of high quality, that is as close to the true Pareto Front as possible, and distributed as diversely and uniformly as possible. In the literature, there is no single metric that can reflect both of these aspects and thus a number of metrics are often used [20, 21]. In this study, we have used the following metrics to evaluate our proposed approach:

- **Coverage** (C)**:** commonly used for comparing two sets of non-dominated solutions A and B, the $C(A, B)$ metric calculates the ratio of the non-dominated solutions in B dominated by the non-dominated solutions in A, divided by the total number

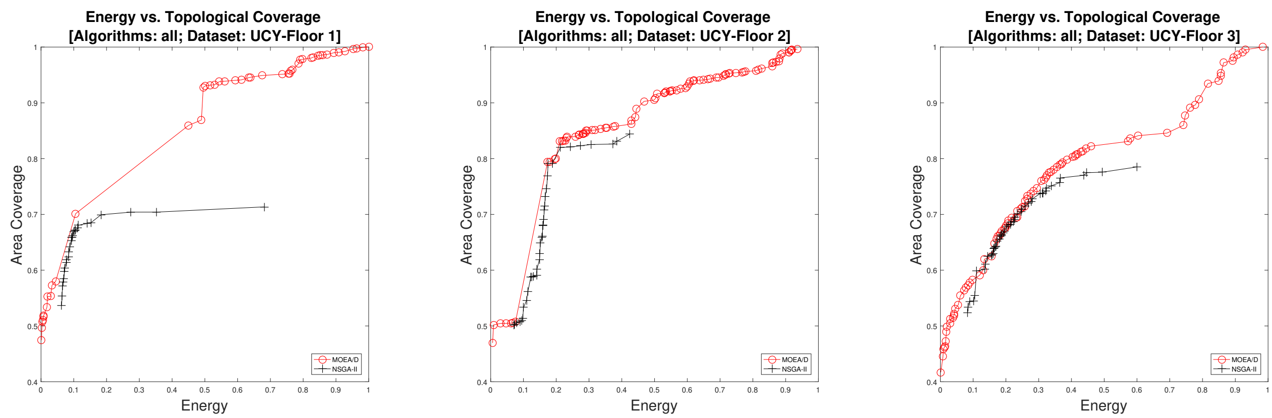

Figure 1: Experimental Series 1 - Comparison between MOEA/D and NSGA-II on the three floors of the UCY dataset.

Table 2: Experimental Series 1 - Comparison between MOEA/D and NSGA-II in terms of the performance metrics I_D, I_H, NDS and C. The best results of each test instance are denoted in bold.

Alg:	MOEA/D			NSGA-II				
Metric:	I_D	I_H	NDS	I_D	I_H	NDS	C(M,N)	C(N,M)
f1:	0.12	**1.00**	**46.00**	0.12	0.94	27.00	**0.39**	0.00
f2:	**0.08**	**0.99**	**87.00**	0.10	0.92	36.00	**0.34**	0.00
f3:	**0.08**	**1.00**	**89.00**	0.09	0.92	49.00	**0.56**	0.08
mean:	**0.0882**	**0.9974**	**77.7500**	0.1023	0.9237	40.2500	**0.4649**	0.0408
std:	0.0215	0.0026	21.1877	0.0138	0.0101	10.7510	0.1134	0.0471

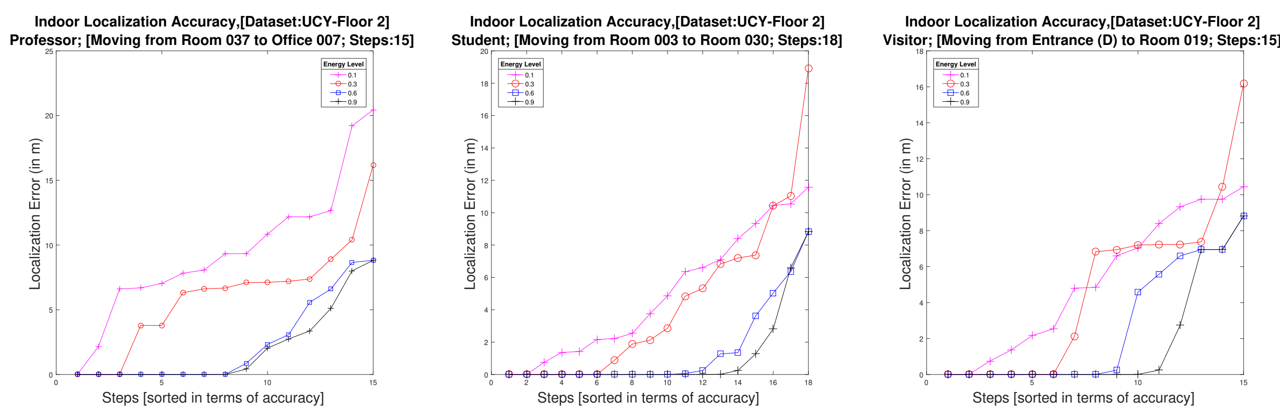

Figure 2: Experimental Series 2 - Indoor Localization Accuracy during navigation for various mobility scenarios.

of non-dominated solutions in B; a higher value for $C(A, B)$ is an indication of higher quality of solutions in A than in B.

- **Distance from reference set (I_D):** shows the average distance from a solution in the reference set R to the closest solution in A. The smaller the value of I_D the closer the set A is to R indicating better convergence. In the absence of the real reference set R, the average distance of each single point to the nadir point is used.
- **Hypervolume (I_H):** indicates the area dominated by at least one solution in the obtained non-dominated set A. Therefore high I_H indicates better diversity.
- **Number of Non-Dominated Solutions (NDS):** a straightforward metric proposed by Weicker et al. in [22] that is usually considered in cases of real-life discrete optimization problems showing the cardinality or the number of Non-Dominated Solutions in set A, i.e. $NDS(A) = |A|$. In these cases, it is more

desirable to obtain a high number of $NDS(A)$ in order to provide an adequate number of Pareto optimal choices. It is usually desirable to have a high number of NDS when the solutions is of high quality (i.e. low C-metric) and spread (i.e. low I_D-metric) in the objective space.

4.5 Mobility Scenarios

For the validation of the obtained PF solutions with respect to indoor localization accuracy, we constructed realistic mobility patterns of three different users (professor, student and visitor) navigating within the CSUCY building. Their navigation is composed of around 15-20 localization steps, where at each step the WKNN method is used (Step 3 of Algorithm 2), as discussed in Section 1.

The three scenarios are as follows:

- A professor moving from a lecture room to his/her office.
- A student moving from a lecture room to a computer lab.
- A visitor navigating from the department's main entrance to a professor's office.

4.6 Experimental Series 1: MOEAs Performance Evaluation

Experimental Series 1 aims at evaluating the performance of the MOEA/D approach against the NSGA-II approach, in all UCY datasets described in subsection 4.1 and with respect to all the performance metrics of subsection 4.4.

Figure 1 shows that the proposed MOEA/D approach outperforms the NSGA-II in all datasets in terms of both diversity and convergence. Particularly, the Pareto Front (PF) obtained by MOEA/D dominates all the solutions obtained by NSGA-II, since it is much closer to the zenith point of the objective space (i.e., upper left corner). Moreover, MOEA/D's PF is also wider than NSGA-II's PF and therefore provides more near-optimal solution choices to the decision maker. This is due to the fact that NSGA-II is trapped to local optima and mainly fails to improve the Area Coverage objective in all 3 cases.

This is also evident from the analytical results of Table 2 in which the best performance value of each metric and for each dataset is shown in bold. MOEA/D outperforms NSGA-II in all three datasets for the I_H, NDS and C-metric and provides better or equal performance for the I_D metric. Particularly, MOEA/D provides around 20% and 10% better performance, on average, for the I_D and I_H metrics. It obtains about 37 more non-dominated solutions in its PF than NSGA-II and the PF obtained by MOEA/D dominates 46%, on average, the PF obtained by NSGA-II.

4.7 Experimental Series 2: Indoor Localization Accuracy on Mobility Scenarios

This experimental series aims at *(i)* validating the performance of the PF solutions obtained by MOEA/D and *(ii)* demonstrating the relation of both the coverage and energy objectives with respect to the indoor localization accuracy on various mobility scenarios described in Subsection 4.5.

Figure 2 shows the indoor localization accuracy per localization step in all three scenarios, while varying the energy level. Here it is important to notice that different energy levels means different solutions from the PF and therefore different partial RadioMaps (*pRM*) used during the localization process. For example, an energy level of 0.1 means that the decision maker selects the PF solution that is closer to an energy objective value equal to 0.1 and its solution in the decision space (i.e., the associated *pRM*) is then used throughout the navigation of the user in our mobility scenarios. The results show that the higher the energy consumption of the selected solution is, the better the accuracy (and therefore the less localization error) is achieved by the user during navigation.

This is also evident from the results of Figure 3 that shows the average localization error of all localization steps per scenario. Here it is important to notice the declining trend of the bar plots of top row as the energy increases and therefore the contradiction between energy and accuracy. Moreover, the high quality of the obtained solutions is also evident, since for a 90% energy consumption, which means that 90% of the whole *RM* is downloaded to the smartphone, the average localization error is close to 1m, which is acceptable for an indoor environment.

Finally, the 3 bar plots of bottom row of Figure 3 show the relation between the area coverage objective and the indoor localization accuracy. The results suggest that the proposed area coverage objective is a good representation of the localization accuracy, since in all three scenarios, as its objective value increases, the localization error decreases.

5 CONCLUSIONS AND FUTURE WORK

In the paper, we propose a novel Multi-objective Indoor Localization Service (MILoS) that provides a fine-grained, energy-efficient localization on the client-side, using only a partial RadioMap of WiFi fingerprints.

MILoS proceeds in three steps. Firstly, it uses Anyplace IPS [25] to reconstruct building digital maps, marking Points-of-Interest (POIs), and pathways joining them. Also associates with any subset X of POIs, a partial RadioMap *pRM* by linking each WiFi fingerprint to its closest POI. The reconstructed graph is fed into a MOEA/D to produce offline a set of non-dominated solutions, that concurrently optimize several conflicting objectives (i.e., minimize the smartphone's Energy Consumption and maximize the Area Coverage). Finally, under some specified user's criteria a solution is selected by the decision maker and the associated partial RadioMap *pRM* is downloaded once on the user's smartphone. No further communication with the server is required and no information on the user's location is ever shared with the server, thus maintaining user privacy. To perform localization at current indoor position, the user employs a smartphone to capture the observed RSS fingerprint and then uses the WKNN localization method (with input the *pRM* and the observed fingerprint) in-situ.

The performance of the proposed method is evaluated on real datasets over mobility scenarios and compared against NSGA-II, the state-of-the-art in MOEAs based on Pareto-dominance. In particular, we experimentally verify that the two objectives are conflicting with each other and that as the energy / area coverage of the selected solution increases, so does the localization accuracy calculated by the user.

Directions for future work include the following: *(i)* Allow the user in advance to indicate preferences (in terms of probabilities)

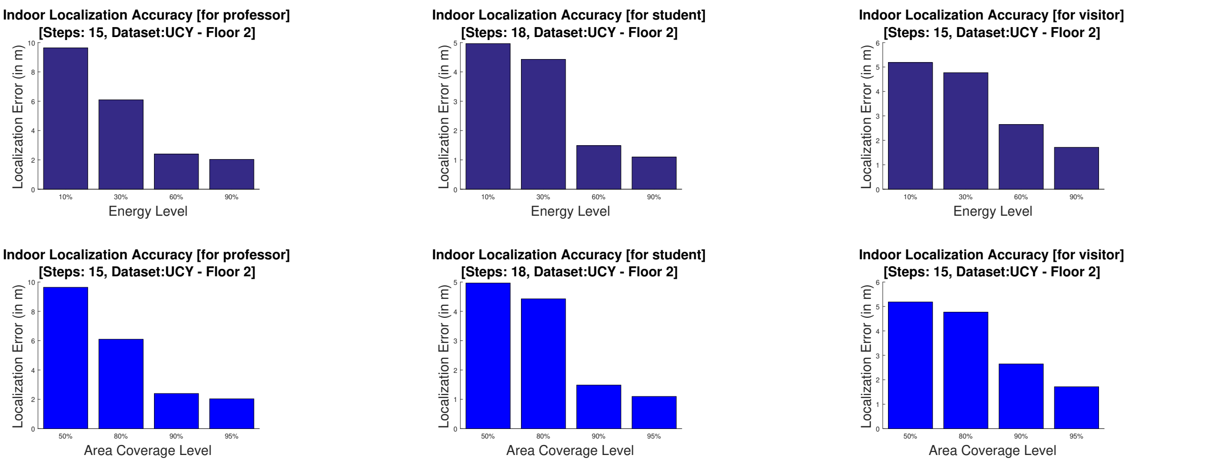

Figure 3: Experimental Series 2 - Relation between Indoor Localization Accuracy and Energy objective on various mobility scenarios.

of visiting POIs, instead of a blind visit, *(ii)* Improve data preprocessing, by introducing a more dynamic method for associating POIs subsets with partial RadioMaps, *(iii)* Hybridize the MOEA/D with Local Search Heuristics for further improving its performance, *(iv)* Design a Multi-Objective Indoor Navigation Service that suggests to the user a path between two POIs, that maximizes localization accuracy. *(v)* Quantify the notion of Privacy (conditional on the required localization accuracy error).

REFERENCES

[1] A. Konstantinidis, G. Chatzimilioudis, D. Zeinalipour-Yazti, P. Mpeis, N. Pelekis and Y. Theodoridis, "Privacy-Preserving Indoor Localization on Smartphones," *IEEE TKDE*, vol. 27, iss. 11, pp. 3042–3055, 2015.

[2] B. Li, J. Salter, A. G. Dempster, and C. Rizos. Indoor positioning techniques based on wireless lan. In *Proceedings of the 1st Intl. Conference on Wireless Broadband and Ultra Wideband Communications*, pages 13–16, June 2006.

[3] A. Khoshgozaran and C. Shahabi. Blind evaluation of nearest neighbor queries using space transformation to preserve location privacy. In *Proceedings of the 10th Intl. Conference on Advances in Spatial and Temporal Databases*, SSTD'07, pages 239–257, 2007.

[4] J. Xiao, Z. Zhou, Y. Yi and L.M. Ni., "A Survey on Wireless Indoor Localization from the Device Perspective," *ACM Comput. Surv.*, vol. 49, iss. 2, pp. 25:1–25:31, 2016.

[5] H. Liu, H. Darabi, P. Banerjee, and J. Liu, "Survey of wireless indoor positioning techniques and systems," *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, vol. 37, no. 6, pp. 1067–1080, Nov 2007.

[6] D. Zeinalipour-Yazti and C. Laoudias, "The Anatomy of the Anyplace Indoor Navigation Service," *ACM SIGSPATIAL Special*, ACM Press, vol. 9, pp. 3–10, 2017.

[7] Y. Xia and C. K. Yeo. Mobile internet access over intermittent network connectivity. *Journal of Network and Computer Applications*, 40(0):126 – 138, 2014.

[8] A. Konstantinidis and P. Irakleous and Z. Georgiou and D. Zeinalipour-Yazti and P. K. Chrysanthis, IoT Data Prefetching in Indoor Navigation SOAs. *ACM Transactions on Internet Technology (TOIT'18)*, , 20 pages, 2018.

[9] H. Kido, Y. Yanagisawa, and T. Satoh. An anonymous communication technique using dummies for location-based services. In *Proceedings of the Intl. Conference on Pervasive Services*, ICPS '05, pages 88–97, 2005.

[10] M. L. Yiu, C. S. Jensen, X. Huang, and H. Lu. Spacetwist: Managing the tradeoffs among location privacy, query performance, and query accuracy in mobile services. In *Proceedings of the 24th IEEE Intl. Conference on Data Engineering*, ICDE '08, pages 366–375, 2008.

[11] C.-Y. Chow, M. F. Mokbel, and X. Liu. A peer-to-peer spatial cloaking algorithm for anonymous location-based service. In *Proceedings of the 14th Annual ACM Intl. Symposium on Advances in Geographic Information Systems*, GIS '06, pages 171–178, 2006.

[12] M. Gruteser and D. Grunwald. Anonymous usage of location-based services through spatial and temporal cloaking. In *Proceedings of the 1st Intl. Conference on Mobile Systems, Applications and Services*, MobiSys '03, pages 31–42, 2003.

[13] G. Ghinita, P. Kalnis, A. Khoshgozaran, C. Shahabi, and K.-L. Tan. Private queries in location based services: Anonymizers are not necessary. In *Proceedings of the ACM SIGMOD Intl. Conference on Management of Data*, SIGMOD '08, pages 121–132, 2008.

[14] M. L. Yiu, G. Ghinita, C. S. Jensen, and P. Kalnis. Outsourcing search services on private spatial data. In *Proceedings of the 25th IEEE Intl. Conference on Data Engineering*, ICDE '09, pages 1140–1143, 2009.

[15] C. Becker and F. Dürr "On location models for ubiquitous computing," *Personal and Ubiquitous Computing*, vol. 9, iss. 1, pp. 20–31, 2005.

[16] H. Lu, C. Guo, B. Yang and C.S. Jensen: "Finding Frequently Visited Indoor POIs Using Symbolic Indoor Tracking Data," *EDBT*, 2016.

[17] K. Deb, A. Pratap, S. Agarwal, T. Meyarivan, A fast and elitist multiobjective genetic algorithm: NSGA II, IEEE Transactions on Evolutionary Computation 6 (2) (2002) 182–197.

[18] Q. Zhang, H. Li, MOEA/D: A multi-objective evolutionary algorithm based on decomposition, IEEE Transactions on Evolutionary Computation 11 (6) (2007) 712–731.

[19] C. A. C. Coello, D. A. V. Veldhuizen, G. B. Lamont, Evolutionary Algorithms for Solving Multi-Objective Problems, Vol. 5, Kluwer Academic Publishers, 2002.

[20] E. Zitzler, L. Thiele, M. Laumanns, C. M. Fonseca, V. G. da Fonseca, Performance assessment of multiobjective optimizers: An analysis and review, IEEE Transactions on Evolutionary Computation 7 (2) (2003) 117–132.

[21] C. A. C. Coello, D. A. V. Veldhuizen, and G. B. Lamont, *Evolutionary Algorithms for Solving Multi-Objective Problems*. Springer, 2007.

[22] N. Weicker, G. Szabo, K. Weicker, P. Widmayer, Evolutionary multiobjective optimization for base station transmitter placement with frequency assignment, IEEE Trans. on Evolutionary Computation 7 (2) (2003) 189–203.