Pervasive and Mobile Computing 23 (2015) 59–79

Contents lists available at ScienceDirect

Pervasive and Mobile Computing

journal homepage: www.elsevier.com/locate/pmc

Determining the best vector distance measure for use in location fingerprinting

Vahideh Moghtadaiee *, Andrew G. Dempster

School of Electrical Engineering and Telecommunications, University of New South Wales, Sydney, Australia

A R T I C L E I N F O

Article history:
Received 7 April 2014
Received in revised form 21 October 2014
Accepted 3 November 2014
Available online 11 November 2014

Keywords:
Indoor positioning
Fingerprinting
Distance measures

A B S T R A C T

Location fingerprinting is a technique that records vectors of received signal strength (RSS) from several transmitters at some reference points (RPs) into a database, and later matches these recorded vectors to a new measurement to position the user. In this work we deal with deterministic fingerprinting algorithms based on the nearest neighbor algorithm (NN). Distance measures between the recorded RSS values and the new measurements are then necessary, which can be calculated using Minkowski distance. The most popular cases of Minkowski distance, Manhattan, Euclidean and Chebyshev, are widely used in various studies in location fingerprinting. However, their positioning performance has not been analytically discussed yet. This causes unknown performance degradation when these signal distances are utilized, which can affect the positioning procedure. In this paper, the positioning performance using Manhattan, Euclidean and Chebyshev distance in terms of the probability of error (POE) are compared analytically and then the results are confirmed by simulations and real experiments. The relationship between the POE and distance error is also analyzed. The results show that using the NN method and a RSS Gaussian distribution assumption, Euclidean distance is the optimum distance which provides the lowest POE and mean distance error (MDE) for indoor fingerprinting.

1. Introduction

Indoor positioning has become highly important because of the difficulties that satellite-based technologies such as the Global Positioning System (GPS) experience operating in such areas, e.g. low received signal power and low visibility of satellites. Non-satellite-based technologies, therefore, are important for indoor localization. Utilizing signals of opportunity is an alternative to GPS due to much higher power levels and wider coverage in indoor environments [1,2].

Many studies have effectively employed wireless networks for indoor localization based on the Received Signal Strength (RSS)-based location fingerprinting technique [3]. Fingerprinting is capable of alleviating some of the problems caused by multipath and Non-Line of Sight (NLOS) propagation in an indoor environment [4]. Fingerprinting requires a survey of Radio Frequency (RF) signal strength vectors to be made ahead of the system's use for localization. Fingerprinting has two stages: "training" and "positioning". It stores the location-dependent characteristics of a signal recorded at Reference Points (RPs) in a database in the training stage, and in the positioning stage, pattern matching algorithms are applied to find the best match between the fingerprint of the user and the database, and eventually estimates the position of the user based on good matches.

The matching methods can be implemented in various ways from a mathematical point of view. They are based on deterministic [5] and probabilistic [6] algorithms both of which have been used in Wi-Fi [3], FM radio [7,8], and mobile

* Corresponding author.
E-mail addresses: v.moghtadaiee@unsw.edu.au (V. Moghtadaiee), a.dempster@unsw.edu.au (A.G. Dempster).

http://dx.doi.org/10.1016/j.pmcj.2014.11.002
1574-1192/

phone [9] networks. The measurements of the RSS values at one location can vary considerably, but in deterministic location fingerprinting the average value is stored for the post processing and position determination stage. Nearest Neighbor (NN), K-Nearest Neighbor (KNN), and K-Weighted Nearest Neighbor (KWNN) methods [5] are the most popular matching methods based on deterministic analysis of fingerprinting.

In this work we deal with the simple and popular NN deterministic method which has been broadly used in fingerprinting. In this algorithm, the position of the nearest RP is considered as the estimated position of the user. In order to select this nearest RP, the signal distance measures between the recorded RSS fingerprints at the test point (TP) and the RPs (in the database) are necessary, which are mostly calculated using some sort of Minkowski distance. Manhattan, Euclidean, and Chebyshev are special and popular cases of the Minkowski distance. Most authors dealing with fingerprinting algorithms use Euclidean distance [10–13]. Manhattan distance is mostly employed when lower computational load is necessary [14]. Chebyshev distance is also easy to use and has less complexity in the calculations [15]. However, there is a lack of research on reasons for choosing one distance measure over the others and no analytical criteria have been proposed for choosing the best. Therefore, without having proper information about the effect of using a distance measure, unknown performance errors can be resulted in the positioning systems.

This paper develops an analytical framework for indoor positioning systems that enables us to analyze the accuracy performance of fingerprinting. By employing this framework, we fully investigate the most important cases of Minkowski distance and compare Manhattan, Euclidean, and Chebyshev distances both mathematically and by simulation so that we can identify which provides least distance error. We will show that results from the simulation agree with the analytical methods and are verified by real experiments results.

The rest of the paper is organized as follows. Section 2 presents the related work on the analytical model for indoor fingerprinting systems and on utilizing different cases of Minkowski distance between fingerprints. An overview of the indoor location fingerprinting system model and the Minkowski distance are described in Section 3. In Section 4 the probability of returning the wrong position using the NN method is explained. The probability of error (POE) in positioning for the special cases of Minkowski distance is investigated mathematically in Section 5. In Section 6, the analytical model of indoor location fingerprinting is employed and Monte Carlo simulation is also carried out to obtain numerical results. The relationship between the POE in positioning and mean distance error (MDE) is also analyzed. Section 7 shows the results of a real experiment. Finally, the conclusions of the work are discussed in Section 8.

2. Related work

There are limited studies that partly analyze and estimate the accuracy and precision performance of indoor positioning techniques based on location fingerprinting [10,16]. The authors in [10] introduce a model for predicting performance of an indoor positioning system based on the probability of correct positioning which is the chance of the distance between the RSS at TP and the correct location fingerprint being less than the distance between the RSS at TP and the wrong location fingerprint. However, their work is based on Euclidean distance only. They also do not discuss the relationship between the probability of correct positioning and distance error.

Various distance measures have been introduced [17,18]. However, most researchers dealing with fingerprinting algorithms limit their work to a single distance measure between TP measurements and RP fingerprints stored in a database, such as Manhattan distance [3,19,20], Euclidean distance [10–13], and Chebyshev distance [15]. Furthermore, although the significant part of fingerprinting is the matching technique, no criteria have been proposed for choosing the optimal distance measures. Regarding Manhattan and Euclidean distances, authors in [18] compare two scenarios of the same and different indoor environment conditions for training and positioning stages of fingerprinting. Their simulation results show that the lowest error is provided by Manhattan and Euclidean in the same and different environment conditions scenarios respectively, but the difference is not significant.

This work, however, examines the Minkowski distance family and highlights a framework that helps utilize the most suitable distance. It analytically investigates the most popular and widely used cases of Minkowski: Manhattan, Euclidean, and Chebyshev distance in order to find the best for fingerprinting. To the best of our knowledge this is the first time mentioned distances analytically compared in fingerprinting positioning. The simulation and experimental results are also discussed.

3. Indoor location fingerprinting

In this section we describe the modified preliminary model of indoor location fingerprinting described in [10] for Minkowski distance.

3.1. System model

It is assumed there are P APs and R RPs with known locations in a 2D area. The whole area here is divided into R small regions where the center of each region is a RP. Fingerprinting is carried out for each RP, and a fingerprint vector which consists of the statistical mean of the RSS from each AP is saved in a database. We denote the mean RSS vector of the rth RP