ARTICLE IN PRESS

Pervasive and Mobile Computing ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Pervasive and Mobile Computing

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

A novel approach towards utilizing Dempster Shafer fusion theory to enhance WiFi positioning system accuracy

Deepak Bhatt *, Swarna Ravindra Babu, Haresh S. Chudgar

Samsung Research India, Bangalore, India

ARTICLE INFO

Article history:
Received 25 May 2015
Received in revised form 6 April 2016
Accepted 25 June 2016
Available online xxxx

Keywords:
WPS
Support vector
Localization

ABSTRACT

WiFi positioning system (WPS) is the most opted technology to provide positioning information indoors. Among the two WPS approaches: Radio-based propagation model and Fingerprinting method, latter is the highly preferred approach for WPS wherein obtained RSSI signatures are matched statistically with the pre-calibrated radio map to localize the user. Existing approaches assume the RSSI distribution to be Gaussian and then utilize Bayesian approaches to derive the position information.

We present a novel approach wherein, firstly, RSSI distribution using kernel density estimation method is approximated via Support Vector Regression (SVR). Secondly, a method is presented to obtain the fused probability mass corresponding to different access points using Dempster Shafer (DS) theory. The fused probability mass thus obtained is used to identify the highly probable points or location of a user over pre-calibrated points. Experimental results demonstrated that the proposed two-step sequential approach is able to limit the 90th percentile localization error to be within meter level. Further the localization accuracy improved to about 88% using our SVR approach in comparison to Gaussian assumption.

1. Introduction

With the ever increasing smartphone devices penetration, a number of location-centric applications are developed such as targeted advertisement, locating places of interest, local offers, etc. Realization of such applications to the fullest relies on localizing the users in all kinds of environments. Locating users outdoors is made possible using GPS, a well known technology developed by Department of Defense which became operational in 1995 [1]. Though the state of the art positioning technology using GPS is accurate to about tens of meters [2], the major challenge lies in obtaining the navigation information indoors. Unlike GPS, there exists no mature technology capable of aiding navigation indoors uninterruptedly day and night. Limitations to the indoor environment arise due to the significant reduction in the line of sight radio signals received from the satellites located at about 20,200 km above the earth's surface.

To facilitate navigation indoors various research efforts have been carried out in order to utilize pre-existing or additional infrastructure such as infrared, ultrasonic, RF based, Wi-Fi, and optical light, etc. [3–7]. As discussed and described in [8,3] infrared based positioning system reads the unique code transmitted through ID card via infrared sensors installed on a ceiling. This information is then passed to the Base Station (BS) to identify a user's location. The Cricket system developed by the Massachusetts Institute of Technology (MIT) uses Time Difference of Arrival (TDOA) of ultrasonic and radio signals to compute the user's position. It is said to be accurate to about 2 cm and has a limited accuracy in Non Line of Sight environment

* Corresponding author.
E-mail address: deep_83bhatt@yahoo.co.in (D. Bhatt).

http://dx.doi.org/10.1016/j.pmcj.2016.06.014

(NLOS). Several RF based positioning systems have also been proposed such as RADAR, Spot-On, LandMarc, etc. and are accurate to about 3–5 m whereas the precision is defined to be 90% within 5.9 m for RADAR [9,10,5]. In [11] an indoor positioning system (IPS) based on visible light communication is proposed delivering an average estimation accuracy of 1.6 cm. Except Wi-Fi, majority of the IPS necessitates deploying special infrastructure such as LED's, RFID tags, or ultrasonic transmitter and a receiver, etc. As a result, WiFi based positioning system (WPS) has gained a lot of attraction in realizing IPS at no additional cost while making use of existing WiFi Access Points (APs) inside mall, office premises, residences, etc. Though the accuracy is still a major concern to the state of the art WPS which is about 3–30 m with an update range of few seconds [12].

WLAN, also known as WiFi operates in the 2.4 GHz and 5 GHz based on the IEEE 802.11 standards and is the most dominant and hugely deployed local wireless area networking standard. It offers a gross bit rate of 11, 54, or 108 Mbps and a range of 50–100 m [12]. Due to the long coverage of WLAN, these signals are additionally studied and utilized to facilitate navigation indoors. In [9], an in-building user location and tracking system, known as RADAR, is proposed. Two approaches for location estimation are described wherein the first utilizes nearest neighbor in signal space (NNSS) using signal strength data collected offline whereas the second approach utilizes trilateration method based on signal propagation model. RADAR system demonstrated 50th percentile accuracy of 2.37–2.65 m. Horus system as described in [13,14] uses probabilistic approach to identify the closest match with the obtained real-time RSSI measurements and pre-calibrated points in order to locate the user also known as fingerprinting method. The system is accurate to 90% within 2.1 m. An extensive research has been carried out to demonstrate the feasibility of WPS. For a detailed review of the existing approaches, the readers are referred to [12,8].

The research work presented here focuses on WLAN based positioning system predominantly using the fingerprinting method. The main contributions of this paper are as follows:

- A novel method to approximate the RSSI distribution obtained using kernel density estimation method is proposed using Support Vector Regression (SVR), a form of Support Vector Machines (SVMs). The proposed method obtains improved estimate of probability mass as opposed to method wherein RSSI distribution is assumed to be Gaussian [15].
- Further, the probability masses for each RSSI measurements thus obtained are fused using the Dempster Shafer Theory to find the highest probable location of a user over a pre-calibrated radio map. Finally, the positioning information is estimated using weighted sum of highly probable points location information and is accurate within a meter level 90% of the time.

The rest of the paper is organized as follows: Section 2 gives an overview of Support Vector Machines and its application to approximate kernel density estimates of RSSI measurements. Section 3 describes how the evidences as derived from received signal strengths in an online phase are combined to identify the highest probable location of a user. Finally, Section 4 discusses and presents the results obtained using the proposed approach leading to conclusion in Section 5.

2. Density approximation using SVR

2.1. Introduction to SVM

Support Vector Regression (SVR), a form of SVM has been widely applied in various tasks of regression due to its ability to avoid local minima and improved generalization. As a result, it approximates a given input–output relationship effectively and is able to estimate output or forecast for an unseen input. Given a set of input–output sample pairs as in (1), the primal objective of SVR is to approximate a function of the form given by (2) such that closeness and flatness are maintained in order to avoid overfit and possess least estimation error.

$$\{x_1, y_1\}, \{x_2, y_2\}, \ldots, \{x_n, y_n\} \tag{1}$$

$$f(x) = \langle w, x \rangle + b. \tag{2}$$

In (1) x_i represents the input feature vector and y_i represents the corresponding target value. In the current context, x_i represents the recorded RSSI measurement corresponding to WiFi AP and y_i is the probability mass. To meet the objectives of closeness and flatness, the primal objective of the problem reduces to:

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*) \\ \text{subject to} \quad & y_i - \langle w, x_i \rangle - b \le \epsilon + \xi_i \\ & \langle w, x_i \rangle + b - y_i \le \epsilon + \xi_i^* \\ & \xi_i, \xi_i^* \ge 0. \end{aligned} \tag{3}$$

In Eqs. (2) and (3), w represents the weight vector normal to the separating hyperplane, b corresponds to the dot product of w and a point on the plane whereas ϵ is a deviation of a function $f(x)$ from its actual value. ξ_i and ξ_i^* are the additional slack variables which determine that deviations of magnitude ξ above ϵ are tolerated. The constant C, known as the regularization parameter, determines the tradeoff between the flatness of f and tolerance of error above ϵ.

Please cite this article in press as: D. Bhatt, et al., A novel approach towards utilizing Dempster Shafer fusion theory to enhance WiFi positioning system accuracy, Pervasive and Mobile Computing (2016), http://dx.doi.org/10.1016/j.pmcj.2016.06.014