Expert Systems with Applications 41 (2014) 6211–6223

Contents lists available at ScienceDirect

Expert Systems with Applications

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

Wi-Fi multi-floor indoor positioning considering architectural aspects and controlled computational complexity

Rafael Saraiva Campos [a,c], Lisandro Lovisolo [b,*], Marcello Luiz R. de Campos [c]

[a] CEFET-RJ, Brazil
[b] PROSAICO – PEL/DETEL – UERJ, Brazil
[c] PEE/COPPE/DEL/Poli, UFRJ, Brazil

ARTICLE INFO

Keywords:
Indoor positioning
WiFi networks
Received signal strength
Fingerprint techniques
Clustering
Kohonen layer
K-medians
Backpropagation

ABSTRACT

This paper focuses on the positioning of Wi-Fi nodes in multi-floor indoor environments. A target radio-frequency (RF) fingerprint – measured by the MS to be localized – is compared with georeferenced RF fingerprints, previously stored in a correlation database (CDB). Therefore, this strategy lies within the so-called Database Correlation Methods (DCM) used to locate mobile stations (MS) in wireless networks. To obtain best matches in terms of architectural structures such as floors, doors, aisles, among others, the authors apply two combined techniques that improve localization accuracy: unsupervised clustering (K-medians and Kohonen layer) and majority voting committees of backpropagation artificial neural networks (ANNs). The unsupervised clustering is employed to allow collected data (the fingerprints) to group freely in their natural space, without precluding – through the imposition of architectural constraints – any natural arrangement of the collected fingerprints. The proposed combined strategy improves floor identification accuracy, which in indoor multi-floor positioning must be high. The effects of the proposed solution on the DCM positioning accuracy are experimentally evaluated using actual measured data. In the trial the floor identification accuracy ranged from 91% to 97%, and the average 2D positioning error ranged from 4.5 to 1.7 m, depending on the size of the measurement window (from 1 to 25 samples).

1. Introduction

There is a growing number of Mobile Stations (MS) equipped with built-in Global Positioning System (GPS) receivers. In open areas, GPS yields the highest location precision, but is usually unavailable in indoor environments. In this scenario, Received Signal Strength (RSS) based location techniques are used both in cellular and WiFi networks (Bill, Cap, Kofahl, & Mundt, 2004). These techniques employ the measured received signal levels (their strengths) from different network nodes to estimate the MS position.

In cellular mobile telephony, indoor positioning might rely on the reception of signals from indoor micro or even pico-cells, or from outdoor cells with strong RSS by indoor mobile clients (Liu, Darabi, Banerjee, & Liu, 2007). However, the increasingly denser deployment and availability of WiFi access points (AP) – specially at corporate buildings, as exemplified in (International Meetings Review) – coupled with the fact that most mobile devices today – like smart-phones, tablets and PDAs – are WiFi enabled, makes the use of WiFi signals a preferable choice for indoor positioning, in comparison to cellular mobile telephony signals (Laoudias et al., 2012). Therefore, this paper focuses on RSS based DCM for WiFi multi-floor indoor positioning. DCM has already been established as a viable alternative for indoor WiFi positioning (King, Haenselmann, & Effelsberg, 2008; Mengual, Marbán, & Eibe, 2010; Mengual, Marbán, Eibe, & Menasalvas, 2013; Zhou et al., 2014). In fact, there is a wide variety of DCM solutions in the literature, but all share the same basic elements (Laitinen, Lahteenmaki, & Nordstrom, 2001; Campos & Lovisolo, 2011).

When implementing or using a location based system (LBS) for indoor positioning, one may not be troubled by small planar errors. For example, suppose that one is in a garage or mall, if the LBS estimates or indicates that an object or MS is within three or five meters from its actual location, this error may be acceptable. However, if the LBS reports that one is in the wrong floor this may not be acceptable at all, mostly due to the fact that buildings are built in a way that it is much easier to move within the same floor than among floors.

* Corresponding author. Tel.: +55 21 996565452.
E-mail addresses: rafael_saraiva@ig.com.br (R.S. Campos), lisandro@uerj.br (L. Lovisolo), campos@smt.ufrj.br (M.L.R. de Campos).

http://dx.doi.org/10.1016/j.eswa.2014.04.011
0957-4174/

Multi-floor indoor positioning is usually approached considering that RF fingerprints collected at the same floor are similar and will group together, when some sort of clustering is used. In such case, supervised training would be the obvious choice, as the targets (i.e., the floors) are known *a priori*. However, due to the inherent complexity of radio-wave propagation in indoor environments, such assumption is generally false. There is no guarantee that measurements of the RSS from a given Access Point (AP) collected at different locations inside a building will reflect the architectural aspects. For example, the RSS from the same AP in two different rooms may vary largely due to the small-fading characteristics of the WiFi channel in addition to the different obstacles that may be encountered along the radio-wave path (Lee & Buehrer, 2011). The same may occur among different floors. That is, in the RSS space, the similarities among data points may not reflect their proximity in the architectonic space.

For example, referring to Fig. 1, although the distances from the AP to the measurement points may be ordered as $d_B < d_{C_2} < d_A < d_{C_1}$, it may easily happen that the RSS are such that $\mathrm{RSS}_B > \mathrm{RSS}_A > \mathrm{RSS}_{C_2} > \mathrm{RSS}_{C_1}$. That is, the RSS order does not maps to the distance order. Moreover, although the measurement point C_2 is much closer to B than to C_1 it may occur that the measurement taken in C_2 is more similar to the ones taken in C_1 than to the ones in B, just to give another example. Obviously, the actual figures for these examples will depend on the geometry of the building, materials, transmission frequency, among other physical parameters.

In several localization strategies using WiFi RSS, there is an implicit assumption that similarities among data points in the RSS space will in some way correspond to similarities in the architectonic space, i.e., reflect the geometry of the building. However, this may not be case. In addition, for some indoor positioning applications, floor error might be very annoying or even unacceptable. For example, in the case of emergency call location, response teams (firemen, paramedics, policemen) need a highly accurate 2D position fix of the target terminal, and, particularly, a floor identification accuracy close to 100% (Moayeri, Mapar, Tompkins, & Pahlavan, 2011; Varshavsky, LaMarca, Hightower, & de Lara, 2007). The same applies for indoor positioning applications designed to help people with special needs, like children and elderly tracking (Moayeri et al., 2011), or to help the blind finding the floor of destination in multistory buildings (Bai, Jia, Zhang, Mao, & Sun, 2013).

Therefore, in this paper, a natural approach is explored for the multi-floor indoor positioning problem. By natural approach we mean that the data is considered in the space that it is collected without the imposition of any *a priori* restriction. In this sense, the RF fingerprints are grouped based on their similarity in their natural space (the RSS space) and not on the floor where they were collected. For that purpose, an unsupervised learning technique must be used, as no *a priori* assumptions are made about the

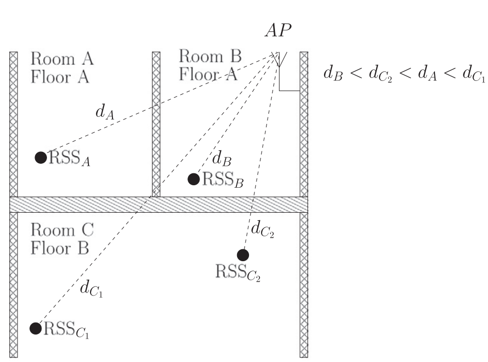

Fig. 1. Example of WiFi multi-floor environment.

clusters. Since the clusters of collected RSS levels scans obtained through the use of unsupervised techniques depend on the intrinsic similarities in data space and not on the output or response aimed for the data, one can, in each cluster, design more specialized classifiers for reporting the desired output (in the present work, this corresponds to the floor where the MS is located). For good understanding and coherence of reported locations one must, at some point of the positioning process, accommodate architectonic aspects, which requires supervised learning. However, using the presented approach one does not design classifiers that deal with all possible input data but with the data belonging to each cluster (previously identified in the first phase, using a unsupervised classifier). This diminishes the computational burden of classification both for the training phase and for its use.

Thus, to enhance accuracy in floor identification, this paper proposes a technique that combines both approaches: natural – data points grouped directly without considering building architectural data – and architectonic – data points grouped considering building architectural data. For the natural or RF approach, a single Kohonen layer with conscience is used. This tries to accommodate the similarities among data collected in the signal/natural (RSS) space using unsupervised clustering technique. In the sequence, a supervised classifier is employed for considering the architectonic aspects, i.e., floor classification. In this context a majority voting committee of backpropagation ANNs is used for each cluster for identifying the floor in which the MS is located from the RSS levels data. Outputs of joint voting classifiers are expected to have a lower variance than outputs of single binary classifiers (Leisch & Hornik, 1997), smoothing the stochastic component inherent to the ANN training and providing a more reliable and stable classification.

From the above, our system first associates the collected RSS levels at a given point to a cluster obtained by unsupervised learning, and subsequently the specific supervised classifier which depends on the cluster identity is applied to the collected RSS levels for identifying the floor. Sections 4.2 and 4.3 describe those topics in detail.

The proposed strategies are tried with actual data collected in a multi-floor building. Special care has been taken regarding data collection, obtaining a high number of measurement samples, so that the experimental results could be more reliable and representative of the type of environment under analysis. Most of the papers that approach the indoor positioning problem evaluate their proposals using a dataset spanning only one floor. Few articles in this field (Al-Ahmadi, Omer, Kamarudin, & Rahman, 2010; Al-Ahmadi, Rahman, Kamarudin, Jamaluddin, & Omer, 2011; Liu & Yang, 2011; Varshavsky et al., 2007) use a dataset that spans more than one floor. In this paper, tens of thousands of samples have been collected in 13 floors of a large building.

The remainder of this work is organized as follows. Section 2 introduces the basic elements of DCM; Section 2.5 provides a diagram of the proposed solution. Section 3 details the data collection campaign we have done. This is presented before the intelligent systems proposed in Section 4.1, since the collected data help illustrate why these are needed. More specifically, Section 4.2 describes the unsupervised clustering techniques and Section 4.3 describes the floor classification procedure. Section 5 details the experimental evaluation. Section 6 brings a brief conclusion.

2. Database correlation methods

DCM, also known as RF fingerprinting positioning, is a class of MS positioning methods that can be applied in any wireless network. The MS position is estimated by comparing a set of RF parameters (RF fingerprint) – measured by the MS or by its anchor