Computers and Electrical Engineering 000 (2017) 1–12

Contents lists available at ScienceDirect

Computers and Electrical Engineering

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

Transmitter source location estimation using crowd data☆

Arif Selcuk Ogrenci*, Taner Arsan

Kadir Has University, Faculty of Engineering and Natural Sciences, Istanbul, 34083 Turkey

ARTICLE INFO

Article history:
Received 15 April 2017
Revised 23 September 2017
Accepted 25 September 2017
Available online xxx

Keywords:
Source localization
Neural networks
Learning
Received signal strength
Nonlinear regression

ABSTRACT

The problem of transmitter source localization in a dense urban area has been investigated where a supervised learning approach utilizing neural networks has been adopted. The cellular phone network cells and signals have been used as the test bed where data are collected by means of a smart phone. Location and signal strength data are obtained by random navigation and this information is used to develop a learning system for cells with known base station location. The model is applied to data collected in other cells to predict their base station locations. Results are consistent and indicating a potential for effective use of this methodology. The performance increases by increasing the training set size. Several shortcomings and future research topics are discussed.

1. Introduction

Massive use of mobile phones utilizing cell-based technologies created huge opportunities for the society and changed the life style in a fundamental way. As almost all of the smart mobile devices are equipped with GPS (Global Positioning System) support, the mobile users can easily locate themselves provided that the satellite based GPS signals are available. In outdoor environments, this has allowed service providers and third party developers to supply numerous location based services to the customers spanning fields of navigation support, active marketing, and social networking [1]. A serious and everlasting problem has been the estimation of the location of the mobile user when GPS is not available, e.g. in indoor environments or when GPS signals are not available because of the buildings and other environmental conditions. There are several solutions to the problem with varying levels of accuracy and cost. A simple and inaccurate solution can be to use the location of the base station as broadcasted by the base station itself provided that the service provider allows such a transmission. Assisted GPS is a more accurate solution where the GPS module of the mobile phone and the service provider's servers work together in different modes for better positioning. Assisted GPS is an enhanced GPS technology commonly used on smartphones. In the Assisted GPS system, an assistance server in the mobile network provides accurate time stamps and accurate GPS satellite orbital information. By using this information, it is possible to obtain initial position information in seconds. The overall accuracy is between 5 m and 10 m in Assisted GPS. Just like GPS, it has several pros and cons. The main advantage of Assisted GPS is its high accuracy. Moreover Assisted GPS can be helpful in densely populated areas where clear GPS signals may not be available. The quick location collection, which is useful for location-based services, is another advantage of Assisted GPS. One of the main disadvantages is the dependency on the mobile network provider. Assisted GPS can only work where mobile network reception is possible [1].

☆ Reviews processed and recommended for publication to the Editor-in-Chief by Associate Editor Dr. Z. Arnavut.

* Corresponding author.
E-mail addresses: ogrenci@khas.edu.tr (A.S. Ogrenci), arsan@khas.edu.tr (T. Arsan).

https://doi.org/10.1016/j.compeleceng.2017.09.026

Please cite this article as: A.S. Ogrenci, T. Arsan, Transmitter source location estimation using crowd data, Computers and Electrical Engineering (2017), https://doi.org/10.1016/j.compeleceng.2017.09.026

In this work, we have investigated the inverse problem of locating the transmitting station based on the data collected by a crowd of devices. The problem of determining transmitter location based on sensors is a research area that is well developed for certain techniques. Common techniques utilized for this purpose are based on TOA (Time Of Arrival), TDOA (Time Difference Of Arrivals), AOA (Angle Of Arrival), RSS (Received Signal Strength), and the hybrid use of those mentioned. In all of those methods the major requirement is that the location of the receiving sensors are known. The computations of the transmitter location can be based on geometrical lateration techniques which may be converted to the solution of nonlinear set of equations employing methods of least squares or other optimization techniques. The reader may refer to the surveys for a comprehensive list of applications [2–8]. The major advantages and disadvantages, or requirements of these methods are as follows. Both TOA and TDOA based methods offer a high accuracy provided that LOS (Line Of Sight) conditions are met. Moreover, TOA requires time synchronization of both the transmitter and the receiver sensors whereas TDOA requires time synchronization for the received signal among the receivers only. AOA methods offer also high accuracy for LOS conditions where special smart antenna arrays are needed to determine the angle of the signal at its highest intensity whereas time synchronization is not required. RSS based methods also do not require synchronization however an accurate signal propagation model is necessary for accuracy. A major concern in all those methods is that fading in signals would change the accuracy drastically. The use of hybrid approaches for non LOS conditions tries to remove this problem to a certain extent. The problems associated with the transmitter location estimation get complicated in heterogeneous environments such as in dense urban environments where the propagation models cannot be determined accurately. The errors arising in the computations of sensor locations and received signal strengths also cause difficulties. The lack of necessary hardware capabilities in the sensor nodes to perform the required complex calculations is another obstacle hindering effective use of those methods [3,6,8].

A major and novel problem is the use of information collected by a crowd to determine the transmitter location even if there is no synchronization among those receiver sensors. The active research area for wireless sensor networks can be extended to this field where TOA, TDOA, and AOA methods cannot be utilized. In the literature, there are papers using the RSS and sensor location information to predict the source location. Among them, [9] describes an iterative method that estimates source location and the other unknown parameters of power and path loss coefficients based on each other where RSS values are used. The simulations assuming zero mean Gaussian noise, indicate convergence to the theoretical lower bounds (Cramer–Rao Lower Bound), however, there is no real life data. Similarly, in [10], the authors employ a novel weighted least squares algorithm for source localization where the simulation results are given for an area of 40 m by 60 m. The results indicate that the prediction accuracy will be enhanced if the signal to noise ratio of the transmitter is increased. The recent work of [11] employs a Bayesian theory for data fusion which is simulated for an evenly distributed set of sensors within an area of 1 km by 1 km. This approach does not estimate the source location directly; it is used to predict the source in the closest proximity. The experimental work in [12] fits histograms of RSS measurements to spline functions in order to predict the source location in a 150 m^2 area where the mean error comes out to be between 10 m and 34 m depending on the position of the transmitter above the ground. The authors conclude that an “alternative approach to solving the positioning equations that takes into account the variations in the path-loss” has to be developed. A field work paper compares different propagation models in an urban area using real life measurements of GSM signals [13]. As the focus of the paper is not source localization, there is no performance data about this, however, the comparison yields that the SUI model ([14]) is the closest one for the dense urban environment. The deviation between the closest model and the experimental data is more than 10 dB in distances less than 1 km and this reduces the applicability of such propagation models in source localization as the location error becomes 80% for typical values of propagation coefficients. Lastly, the work in [15] deals with the problem of estimating the theoretical bounds of source localization. The methods are based on using the RSS measurements, however, the paper states that non-Bayesian methods have to be deployed if the propagation model is not known. Review of those papers mentioned above indicates that hard computing methods would not suffice in dealing with our problem where an accurate propagation model cannot be obtained from the collected data only. However, research indicates that soft computing methods such as neural networks can be efficiently used for source localization in a setting where the nonlinear propagation model is highly unpredictable [16]. A recent review of RF based indoor positioning methods also indicates that there are soft computing methods utilized for proper source localization [17]. Based on those work in the literature, we have investigated the problem of transmitter location estimation using machine learning from data supplied by a set of heterogeneous sensors which only supply RSS and location data. To the best of our knowledge, there is no work that utilizes GSM signals in outdoor for source localization. As the case study, we have performed real life tests using the GSM signals of base stations received by mobile phones in a dense urban area. In the remaining of the paper, Section-2 describes the problem settings and the methodology employed. Section-3 includes the experimental data collected and processed. Finally Section-4 concludes the paper by discussing the results, shortcomings and future research topics.

2. Problem and methodology

2.1. Problem statement

In its most general form, the research problem is to determine the location of a transmitter source using the data collected by a crowd of sensors where the data includes the “id” of the transmitter, the strength of the received signal (RSS)