Ad Hoc Networks 33 (2015) 55–70

Contents lists available at ScienceDirect

Ad Hoc Networks

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

# Sensor node tracking using semi-supervised Hidden Markov Models

Sudhir Kumar, Shriman Narayan Tiwari, Rajesh M. Hegde *

*Department of Electrical Engineering, Indian Institute of Technology, Kanpur, India*

ARTICLE INFO

*Article history:*
Received 30 August 2014
Received in revised form 16 April 2015
Accepted 17 April 2015
Available online 27 April 2015

*Keywords:*
Sensor node localization
Semi-supervised methods
Hidden Markov Model (HMM)
Sensor node tracking
Multi-sensor data fusion
Ad-hoc sensor networks

ABSTRACT

In this paper, a novel method for mobile sensor node tracking using semi-supervised Hidden Markov Models (HMM) is discussed. A new methodology to develop a combined attenuation model from data gathered from multiple sensors is also described. Observations emitted from the nodes are sparsely measured over the network area with beacons placed on the boundaries. HMMs are trained using observations measured at each grid point. The distances between a node passing through a specific grid point and beacons are estimated using likelihood maximization. The local location co-ordinates of the node positions are then computed by solving a constrained volume optimization problem. Quaternion rotation is used to finally obtain global coordinates of the node location. Several standard manoeuvres of mobile nodes are first simulated. Similar manoeuvres are also recorded from real field deployments. The experimental results are obtained for node localization and tracking from these experiments. Results indicate an improvement in the localization accuracy, when compared to the conventional localization methods.


## 1. Introduction

Recent advancements in wireless communication, MEMS technology, and sensing theory have accelerated the current research on wireless sensor networks (WSN) [1]. These technologies allow the fabrication of small size and low powered wireless sensor nodes. This has led to the development of novel algorithms for localization and related applications in the area of ad-hoc wireless sensor networks. Estimating the locations of the nodes over time is a challenging problem. This process of location estimation of a node is known as localization. Localization has found wide applications ranging from human tracking, animal tracking, habitat monitoring, military surveillance, disaster management, to medical applications. Various localization algorithms have been presented in the literature. The key idea behind all these algorithms is to use beacons whose absolute location is known either through global positioning system (GPS) or using any other location aware device. Localization of sensor nodes in a network is carried out with the help of location information of beacons. Although GPS has wide applications in outdoor environment, it has high energy consumption. Moreover, the location estimation error in GPS is high in indoor scenario. This is due to the non-availability of direct line of sight communications from the satellites. A large sensor network can have as many as thousands of sensor nodes. The use of GPS for location estimation in such a network increases both energy consumption and deployment cost. The high energy consumption of GPS devices and the large number of the sensor nodes make GPS unsuitable for WSN applications.

Several localization algorithms based on received signal strength (RSS) [2], direction of arrival (DOA) [3], time of

* Corresponding author.
*E-mail addresses:* sudhirkr@iitk.ac.in (S. Kumar), shriman.001@gmail.com (S.N. Tiwari), rhegde@iitk.ac.in (R.M. Hegde).

http://dx.doi.org/10.1016/j.adhoc.2015.04.004
1570-8705/

arrival (TOA) [4,5] and time difference of arrival (TDOA) [6,7] are employed to perform localization in wireless sensor networks. The RSS based method is cost effective and ubiquitous when compared to other measurement techniques. However, it performs poorly due to variations in a wireless channel. TOA based localization techniques require highly precise synchronization between nodes [8]. Special equipment such as antenna array are required to measure the angle of arrival of signal using DOA method of localization. When using Hidden Markov Model (HMM), the system is modeled as a Markov process with hidden or unobserved states [9]. In the context of localization, the hidden states refer to the range corresponding to the given observed sequence. This helps in estimating the distance between a node and a beacon from which the signal is received. This distance information obtained from HMM model is then further used for node localization. Another method, semi-definite programming (SDP) is described in [10], for noisy distance measurements as an approach to localization. However, this approach is less precise in estimating the location of static sensor nodes. In [11], an attempt to reduce the error in estimated location for static nodes is discussed. The algorithm based on the multidimensional scaling (MDS) [12,13], also computes the location of unknown nodes, given the set of distances between each pair of nodes. This method suffers from large localization errors.

The main contribution of the paper are as follows. One major contribution of this work lies in the development of the individual and combined attenuation model from sparse data collected from multiple sensors. Additionally, it is a semi-supervised method that can perform well in limited data conditions. A novel mapping technique is also proposed to obtain correspondence between received observation sequence and range using a semi-supervised HMM. Field sensor deployments in both outdoor and indoor environment indicate a significantly improved performance due to these advantages. The $O(N)$ complexity of the localization algorithm also makes it computationally efficient.

The rest of the paper is organized as follows. Section 2 describes the related work. The basic framework and algorithm for mobile node tracking is discussed in Section 3. The performance evaluation of the proposed algorithm on node tracking has been discussed in Section 4 followed by conclusion in Section 5.

## 2. Related work

Various localization algorithms have been proposed in literature. The algorithms proposed in [4–7,14,2] use a beacon that is fixed and aware of its location. All these methods assume that the node is unaware of its location. The node estimates its location from the signal received from the beacon. Localization based on the received signal strength is also widely used in the literature. In [2], stochastic RSS-map model and RSS channel model are discussed. However, this method does not consider the issue of robustness to environmental changes. Because of time varying nature of wireless channel, RSS [2] method performs badly. RSS fluctuates due to wall, furniture, and people movements in NLOS conditions. Obstacles such as building, forest are the major components of this cause. In this context, Bayesian network called semi-supervised Hidden Markov Model has the advantage. Semi-supervised Laplacian regularized least squares method and HMM based RSS-EKF (Extended Kalman Filter) method using RSS are described in [15,16] respectively. The performance of these algorithms are illustrated using simulated data. In contrast to these methods, the proposed algorithm utilizes low cost multi-sensor data such as radio, acoustic, visible light and fused signal acquired from real field deployment. Another popular method in literature is based on angle of arrival (AOA) [14] of the beacon signal impinging on the node. The AOA method requires an antenna array to measure angle of arrival of the beacon signal.

The method proposed in [4,5] describes the node localization scheme based on time of arrival (TOA) of signal. Time of arrival (TOA) method performs localization using the information about time of arrival of signal from different beacons. But TOA requires a reference time stamp to synchronize the local clock of the node. To overcome this issue of reference offset, a time difference of arrival (TDOA) method has been proposed in [6,7]. TDOA is based on the principle of time difference of arrival of the beacon signals at a pair of nodes. In this method, one of the beacons is generally taken as a reference sensor. The TDOA method has high localization accuracy in low noise environments. However, this method is highly susceptible to errors from non-line of sight (NLOS) measurements. In [4,5], TOA-based localization under NLOS scenario is discussed. Additionally, these methods do not assume the statistics of NLOS distribution. The performance of all the aforementioned methods like RSS, TDOA, TOA and AOA methods degrade under harsh weather conditions. On the other hand, machine learning techniques such as neural networks, and Hidden Markov Model requires large training data. But they have added advantage of performing well under noisy and limited data conditions.

In this paper, a semi-supervised method for mobile node tracking using Semi-Supervised Hidden Markov Model (SS-HMM) over a scalable ad-hoc sensor network is described. The observed sequence is employed to perform the task of mobile node tracking. This is because, the received signal measurement is cost effective as most of the experimental nodes have this capability. As the node can move in any direction and can have varying velocities, a method that is independent of geometry is most suitable in the context. A SS-HMM can be utilized to ensure that independence. Random motion of node is considered to account for both varying direction and speed of a node. Hence, this random mobility model provides all possible motion scenarios. The observed sequences are used to obtain initial estimate of node distance from beacons using Hidden Markov Models. Coarse location estimates of the node are then computed using these distances. A refinement technique is subsequently applied to compute the final node location. The proposed algorithm tracks the mobile node precisely and efficiently. Also, it is robust even if the distance measurement between a node and beacons is slightly inaccurate. This makes the proposed algorithm