Scene Analysis based Bluetooth Low Energy Indoor Positioning Methods
Master's Thesis Proposal v0.4

Srikanth Gadicherla

July 25, 2017

1 Goal of the thesis: Description

1. **Automatic Node Location Identification(ANOLI):**

   This task involves finding the locations of the beacons intelligently. Currently, Active Aheads, can read the RSSI values from all the other beacons, so, we need to employ an Machine Learning method for creating a topology of sensors. This task can also be named Self Organizing Sensor Network. The feasibility study is yet to be done but the main idea here is that combining the data from all the sensors and create a topology of sensors and get metadata out of it. This metadata would include to which cluster the particular sensor would belong to, clusters relative position to some landmark location, etc. [**?**]
   So, we will have 2 weeks after the commissioning of the luminaires.

2. **Develop an algorithm for Bluetooth Low Energy Indoor Positioning(BLE-IP):**

   In this task we want to find the position of a subject given the received signal strength indicator(RSSI) from the BLE beacons. This method would also make use of the information gained during the previous task. So, the idea here is that the metadata is loaded in to the mobile application, and when it starts reading the signal and it will relate the particular MAC(media access control) address to the topology and be used in the memory and non-memory methods. This metadata could also include the current sensor data, and the plan is to integrate as much data as possible for getting better accuracy. There is also conscious effort to make the application have less/no dependency on the internet.
   The first task would be to understand the different biases via experiments so as to build a broad model which could be used in variety of environments. Then the different methods are deployed to check the accuracy based on that.

   Challenges: Need to deal with the different sampling frequencies of different AP's(BLE, WIFI) or smart-phone sensors in the algorithm [Check balzer82 github page or check Torres-Sospedra et al, 2.3, Para 2]

2 Introduction

Indoor positioning(IP) is moving towards becoming ubiquitous and fulfilling the goal of Internet of Things(IoT) and Artificial Intelligence(AI) of Location Aware services. The explosion of usage of smartphones in the last decade has also enabled an exponential growth of this field. IP mainly leverages the sensor hub in smartphones like accelerometer, gyroscope, barometer, proximity, thermometer, pedometer, magnetic sensor, etc and wireless technologies like bluetooth and WiFi for accurately solving this problem. Signals of opportunities like magnetic field, pressure, light and sound intensity, Global Navigation Satellite System [Torres-Sospedra et al, A realistic evaluation of indoor positioning systems based on Wi-Fi fingerprinting, 2017], mobile network [**?**] and frequency modulation(FM)[**?**] signals have also been used which don't incur any additional infrastructure costs. IP also gets impetus from interrelated fields like geo-location, augmented reality, natural user interface, gaming, man-machine interaction, and 3D mapping.

[cite here]

The need for indoor positioning arose because the traditional Global Positioning System(GPS) fares badly indoors due to signals attenuation and scattering caused by roofs and walls. This leads to higher uncertainty in estimation of location which sometimes spans across multiple rooms sabotaging the whole positioning problem. Usage of indoor positioning(IP) can also be reasoned by following reasons condition, convenience and commercialization. Condition in the sense people tend to stay indoors 90% of the time according to Environment Protection Agency(EPA) [?]. Convenience as it can be used for locating rooms and tracking assets. Commercialization, in case of malls, groceries stores, can be used to send out discount coupons, personalized recommender engines. The positioning data could also be used to optimize the resource flow in hospitals or warehouse to minimize the operation cost[?].

put it to some other good place

change the adjective here, not quite intuitive

right word?

Formally, the estimation of location of mobile unit(MU) using wireless technologies is called location sensing, geo-location, position location or radio-location[?]. The conventional GPS suffer with problems like building occlusion, multipath effect [?] and signal attenuation [?] indoors, and hence many other technologies have been advocated. Even though the GPS technology can be improved using the GPS repeaters, the solution is not scalable and would require costly infrastructure. The new technologies for positioning include Assisted GPS, Bluetooth Low Energy(BLE), MEMS sensors, LED, WiFi, Magnetic anomalies or methods used in conjunction with each other [?].

cite classic text book

improve it

is it fine if I use abbreviations directly

IP has been solved using the triangulation techniques like lateration with time of arrival(TOA), time difference of arrival(TDOA), angle of arrival(AOA), received signal strength(RSS) or fingerprinting based scene analysis methods like decision based probabilistic method, k-Nearest Neighbour(k-NN), artificial neural networks(ANN) or proximity based methods[?] and Gaussian processes with latent variable models[?]. For the rest of the article, we use beacons and nodes interchangeably.

1. AOA needs at least two known sources. 2. i-locate, open geo-data. http://www.i-locate.eu/pilot-sites/. 3. Before the advent of bluetooth low energy(BLE) beacons, the indoor positioning was ....

2.1 Setup

The BLE beacons are part intelligent luminaires are densely packed, the nodes can read the signal strengths from neighboring nodes. Generally, a location estimation task involves two hardware devices, one transmitting the signal and other receiving it[?]. Here, the transmitting device is bluetooth low energy chip(or WiFi router?) installed in Active Aheads. The other sensor data (there is no infrastructure for getting these data as of now), follow-strength values is planned to be downloaded once and updated biweekly or monthly. The Active Aheads are part of a mesh network.

3 Related Work

In this chapter, we discuss the existing methods and state of the art.

But it is essential to understand that the method and level of accuracy required is mostly application dependent.

add what each section does

3.1 Statistical Data analysis of RSSI values

The analysis of the received signal strength values could lead to better understanding of features and patterns for a particular fingerprint . Previously studies were based on RSSI and were mainly directed towards communication technologies rather than positioning systems, hence a clearer knowledge and analysis of this would lead to better modeling and design of positioning systems[?]. The RSSI values suffer attenuation due to reflection, refraction, multipath effect, number of people, indoor environment[Hightower et al, Location Sensing Techniques, A Technical report, 2001], orientation of phonesand quality of BLE chip in the MU's. One more [?]. There's always a trade-off between number of access points and signal interference, but even with the increase in nodes there is an irreducible error due to power measurement noise and quantization, hence can't be reduced below a certain value. Its also because of movement of people and MUs, there is time-correlated fluctuation of RSSI, which sometimes severely affect the performance of positioning solutions[?].

(using the compass), for experiments

The different ways of measuring the RSSI values, one, the measurement has standard deviation of 2.5 dBm while the phone was held still.The presence of human widely varied the signal strength values . But according to [**?**], the consideration phone orientation didn't improve the accuracy much.

add the rotating and cardinal directions.

how much and cite it

3.1.1 Preprocessing and Manipulation of RSSI values

From [**?**], for multiple measurements recorded at a single time step, generally the mean of all measurements is taken into consideration.

3.1.2 How hardware effects the measurement of RSSI values?

The effects of MU's hardware can be useful in estimating the measurement noise and modeling the same, it was seen that variation was 5dBm across different devices.[**?**]

3.1.3 Modeling based on variation in RSSI values

The probability distribution of RSSI is a function of location of time as the indoor setting is extremely convoluted. This leads to bimodality and asymmetric nature, hence modeling a better data model is daunting task. The skewness of the RSSI values depends on the distance and presence of obstacles leading to either left or right skewness, or sometimes it could also be symmetric. The left-skewness of the RSSI which is due to large variation of smaller values than information rich higher values values and these are modeled as log-normal distribution [Honkavirta et al, 2009]. It is also approximated using Gaussian distribution or exponential distribution. The limitation of smartphone in registering low signal values and quantization also exacerbates this problem.

3.2 RSS based Localization Algorithms

3.2.1 Triangulation

Triangulation includes methods like Lateration and Angulation, which are described. This method requires estimates from at least three different known sources.

- **Lateration** Lateration[**?**] accomplishes the position estimation problem by measuring the distances from known reference points, hence it is also called **range measurement technique**. Most of the times distance is indirectly measured via either received signal strength(RSS), time of arrival(TOA) or time difference of arrival(TDOA). TOA works on the principle that time and distance positively correlate with each other and given one other can be estimated. But care has to be taken that transmitter and receiver are properly synchronized. TDOA works on time difference of time taken by different measurement devices for receiving the signal. Signal attenuation based models(RSS) try to learn the path loss component due to propagation as the signal in indoor setting is degraded due to multipath fading. Other methods include roundtrip-time-of-flight(RTOF) and received signal phase method. These methods could be used in conjunction with each other to better the accuracy.

  Drawbacks in Lateration: Multi path and lack of line of sight complicates the problem in TOA, TDOA and Signal attenuation based methods. The location of nodes should be known beforehand.

- **Angulation** : The location estimate is found out using finding angle of the target along some fixed reference with at least two known nodes.

  Drawbacks in Angulation: Need expensive, large and complicated hardware equipment. This method also suffers from multipath reflection.

3.2.2 Static Scene Analysis(Fingerprinting)

Scene Analysis involves inferring based on features already recorded for a particular location and static scene analysis looks up into a pre-recorded dataset which maps locations with it features[**?**]. Hence, it is also called Fingerprinting as it aims at creating unique feature map based on location dependent measurements. This method involves matching the measured signal with location-dependent ground truth signal. To obtain the ground truth signal, some pre-defined points called

reference points are selected and RSSI measurements are recorded from all the available nodes. These ground truth measurements/pre-recorded dataset with location tag is called **reference table**. The reference table can be used to create a radio-map unlike [?].This can be done either with a simple pathloss interpolation or Gaussian process regression. This reference table in conjunction with test measurements can be used with deterministic or probabilistic methods for obtaining the location estimate[?]. [?] calls this method as template matching algorithm, with reference table and test measurements data called as offline and online data respectively. [?] calls the offline phase as calibration phase. These data can also be used in sequential Monte Carlo methods like Bayesian Filters (Grid Filter, Kalman Filter, Extended Kalman Filter, Unscented Kalman Filter, Particle Filter) in conjunction with Gaussian Processes both simple and Latent Variable Models or k-Nearest Neighbors.

my defintion of radio-map

check this out,

For the construction of radio map, most of the authors have recorded the measurements keeping their MU's still but [?] have rotated at the given position while recording the measurements to minimize the orientation bias and at the same time they have increased the calibration time to maximize the reliability of the measured data[?]. These have also been recorded in the four different cardinal directions[?].This goes in-line with [?] which advocates more the samples better the fingerprint during calibration phase. [?] splits the floor into cells and assumes the likelihood to be constant inside the cell and in addition also uses methods like histogram and histogram comparison.

Advantages of scene analysis over other methods is that it doesn't require any additional information about the devices in contrast to triangulation methods like time-of-flight, time-difference of flight where extra synchronization is essential and in conjunction with GP's it is scalable[?].

Few observation about Calibration phase: Calibration time more than 10 secs doesn't help. And also increasing in AP's doesn't increase accuracy beyond some point. The Grid size for calibration depends on the floor plan and location of the access points.
In simple Gaussian process(GP) model, we use simple GP regression to predict the RSSI values at the other locations and then use a simple Gaussian model for evaluating the likelihood model. One major advantage of using semi-parametric model is that with enough evidence the non-parametric counterpart can override the parametric one[?].

what is the learning when we are using k-NN?

What methods have been used? [Mazzullah Khan et al,2017] have used pathloss measurement model in Unscented Kalman Filter in conjunction with weighted kNN(WKNN).

One of the drawbacks for Gaussian process latent variable models(GP-LVM) is that it requires unique and information-rich measurements, hence limits it's applicability

add online and offline method

which paper?

Drawbacks of Fingerprinting: When the indoor setting is restructured or locations of AP's is changed then the current radio-map would become redundant, needing a new radio-map[Yiu et al, 2016].

3.3 Challenges with RSS based methods

1. Multipath and fading.
2. Movement of user causes fluctuation called small scale fading[?].
3. Heavy interference due to 2.4 GHz license-free frequency which is also used by cordless phones, bluetooth devices, WiFi signals and microwaves. [?]
4. Signal attenuation based on presence of number of people. This is because human bodies are bags of waters which can adsorb signals. [?]
5. Absence of ping from any BLE device due to malfunction can effect the positioning accuracy. Usually, the unheard devices are set to their minimum power device sensitivity level[?]and similarly even the test data is given the same treatment.
6. Inaccuracies due to multipath fading can be mitigated using spatial and temporal patterns [?]. This method will find the correlation between positional data and walking trajectory, internal building structure and nodes location (ANOLI).

3.4 Signal Technologies used for Indoor Positioning

In general location estimation scenario, Global Positioning system(GPS) is widely used technology for outdoors. For indoors, technologies like Infra-red, ultrasonic, radio-frequency, optical systems[**?**], visual light communication.

3.5 Applications of Indoor Positioning

1. **Asset tracking**: Tracking an asset based on reference signal sent by the asset.
2. **Personalized recommender engines**: Based on the location data from a store, personal recommendations can be sent to users.

3.6 Performance Metric for Localization Systems

Similar to [Honkavirta et al, 2009], the mean of the norms(ME), median. RMSE, maximum and 95% confidence interval.

4 Work done

1. Naive implementation of k-NN and GP measurement model for BLE based indoor positioning[**?**].
2. Naive implementation of asset tracking using linear Pathloss model.
3. An smart-phone application for getting the RSSI values from the nodes.

5 Methods and Materials

5.1 Methodology in the thesis

The methods used in the thesis would be based on methods and models. Models could be either based on either Gaussian Processes(GP) or Pathloss model. Gaussian processes can be used both during the fingerprinting phase or as the measurement likelihood(or data) model. Pathloss model could be inaccurate as the signal suffers from multiple attenuations indoors so inclusion of prior knowledge is crucial here. So the pathloss model is used in conjunction with either GP as the mean function [**?**], some intelligent mapping between physical state space and signal space[**?**] or interpolation for improvement in the accuracy [**?**].

5.1.1 Different kinds of Methods

1. **Non-memory based methods**: These methods are deterministic or intelligent methods which try to relate the spatial space with the signal space. These methods are Trilateration/ Triangulation, k-Nearest Neighbors, Artificial Neural Networks etc. These methods estimate directly based on the RSSI values with no knowledge about the previous state or previous measurements, hence the name Non-memory based methods.
2. **Memory based methods (or Sequential Monte Carlo)**: By adding a dynamic model or the prior to the non-memory based methods we arrive at Memory based methods which are also called Sequential Monte Carlo based methods. The non-memory based methods are used as the measurement model, hence combining with the prior knowledge we arrive at the posterior or state estimate. There are different dynamic models used like stationary state model and constant velocity model(Bearing only Tracking; BOT) [**?**], the Augmented Coordinated Turn model could also be used to include the heading. Then, the sensor data like compass have to be include to get a better of this.

   Particle filters can be initialized using the location of node which gives highest RSSI [**?**] value in conjunction with motion detection data using PIR sensors. And, also the particle filter could only be advanced when an event from the accelerometer is detected rather than updating it always [**?**]. In other case, the particle states included the position, step length

and heading offset with constant number of particles, so the individual particles were moved according to estimated step length and heading angle [**?**]. Calibrated Orientation sensor or compass for direction in which the person is moving[**?**] shows that gyroscope and compass together can be used for this purpose.

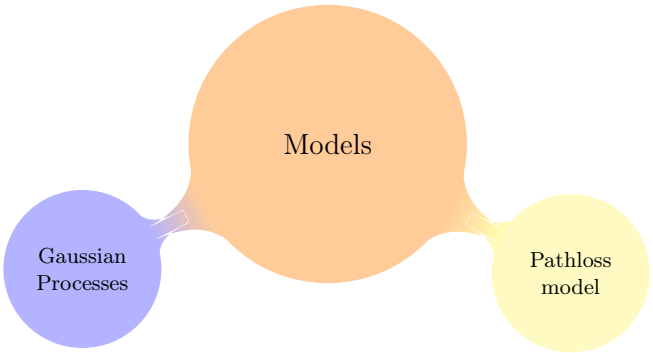

Figure 1: The two models used in the thesis.

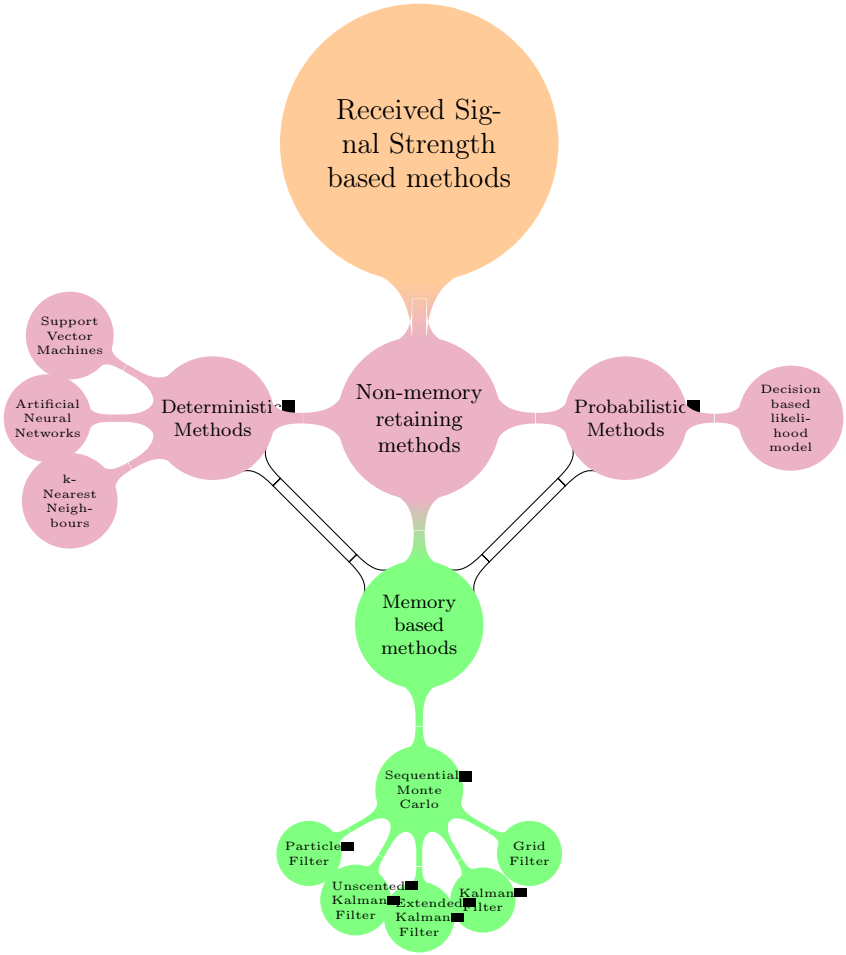

Figure 2: The colorless edges denote that memory based methods can be used in conjunction with non-memory based methods.

6 Plan

6.1 Experiments:

6.1.1 Data Analysis of RSSI values.

1. Different Biases in RSSI measurements
   - User's presence and without user's presence.
   - Different phones.
   - Different orientation of phones, angles, height?
   - Different material of luminaires.(metal, plastic, metal+plastic). Its an added bonus if we can get the different luminaire material information from the BLE signal.
   - Different orientation of BT chip (cardinal directions and upwards and downwards)
   - Different calibration time for fingerprinting.
   - Different calibration points.
   - Fingerprint measuring while being still, only in 4 cardinal positions and while rotating.
2. With nearest node not working.
3. Experiment with different orientation of the mobile device. Hypothesis is that the there is 5 dBm difference for the change in orientation [**?**].

6.1.2 Methods

Next, based on the results of the experiments above implement the methods described in the **??**.

Month	**Task**	**Remarks**
February	REST API	Completed
Meeting 1	Slides:	`https://goo.gl/Up9ejx`
March	Literature Review	1. Made extensive notes of the reading 2. Will move it to thesis 3. Made plan.
	Phase 1	
April & May	Get better model:	Influence of various biases; Experiment in **??**
June: week 23, 24	Buffer period for 2 weeks	End of Phase 1
	Phase 2	
June: week 25, 26 July: week 27	Methods	Non-memory based methods. As discussed in **??**:
July: rest August	Methods	Memory based methods
	Phase 3	
September	Writing	Compile results, write it to thesis, and survey paper.
October Novemeber	Writing	Polish the thesis,comments from instructor and supervisors.

Table 1: Plan for the thesis

References

[1] Yeqing Fang, Zhongliang Deng, Chen Xue, Jichao Jiao, Hui Zeng, Ruoyu Zheng, Shunbao Lu, "*Application of an Improved K Nearest Neighbor Algorithm in WiFi Indoor Positioning*,"

China Satellite Navigation Conference(CSNC), 2015 Proceedings: Volume III2015342, Berlin, Germany, Springer 517524, Lecture Notes in Electrical Engineering, 10.1007/978-3-662-46632-2_45, doi:10.1155/2014/649276

[2] Xianghui Yuan, Feng Lian, and Chongzhao Han, "*Models and Algorithms for Tracking Target with Coordinated Turn Motion*," Mathematical Problems in Engineering, vol. 2014, Article ID 649276, 10 pages, 2014. doi:10.1155/2014/649276

[3] H. Liu, H. Darabi, P. Banerjee and J. Liu, "*Survey of Wireless Indoor Positioning Techniques and Systems*," IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), vol. 37, number 6, 1067-1080 pages, Nov 2007, doi:10.1109/TSMCC.2007.905750

[4] Yiu, H. Darabi, P. Banerjee and J. Liu, "*Locating User Equipment and Access Points using RSSI fingerprints*," IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), vol. 37, number 6, 1067-1080 pages, Nov 2007, doi:10.1109/TSMCC.2007.905750

[5] Arjun. P. Kamath,"*Indoor Location-Based Services In The Telecommunications Network*", Master Thesis, Department of Communications and Networking, School of Electrical Engineering, Aalto University, 2014.

[6] B. Ferris, D. Fox, and N. Lawrence. WiFi SLAM Using Gaussian Process Latent Variable Model. In JCAI, January 2007.

[7] Beacon, link: https://store.kontakt.io/our-products/30-double-battery-beacon.html.

[8] Discussion with Max Björkgren on 9th March.

[9] Martin Sauter (2010). "3.7.1 Mobility Management in the Cell-DCH State". From GSM to LTE: An Introduction to Mobile Networks and Mobile Broadband (eBook). John Wiley & Sons. p. 160. ISBN 9780470978221.

[10] http://www.geospatialworld.net/wp-content/uploads/magazine/Geospatial-World-August-2014.pdf.

[11] Discussion with Tommi Raivisto on 16th March.

[12] Project done for the course CS-E4030 - Seminar on Gaussian processes, School of Science, Aalto University.

[13] Suining He, S.-H. Gary Chan. *"Wi-Fi Fingerprint-based Indoor Positioning:Recent Advances and Comparisons"*, IEEE Communications Surveys and Tutorials, 2015.

[14] Kaemarungsi et al, Analysis of WLAN's received signal strength indication for indoor location fingerprinting,2011

[15] Philipp Richter, Revisiting Gaussian Process Regression Modeling for Localization in Wireless Sensor Networks, 2015.

[16] Honkarvirta et al, 2009; Prasithsangaree et al, 2002

[17] De Luca et al, Performance evaluation of Indoor Localization Techniques Based on RF Power measurements from Active or Passive Devices.

[18] Hightower et al, Location Sensing Techniques, A Technical report, 2001

[19] Mazzullah Khan et al,2017

[20] Umar Ahmad, iBeacon Localization, Master Thesis.

[21] Smailagic et al, 2002

[22] Prasithsangaree et al, 2002

[23] Torres-Sospedra et al, A realistic evaluation of indoor positioning systems based on Wi-Fi fingerprinting, 2017

[24] Nurminen et al, Particle Filter and smoother for Indoor localization,2013