# Improved Area Estimates for Localization in Wireless Sensor Networks

[1, 2]Venkatesh A. Pillai, [3]Winston K.G. Seah, [1]Yong Huat Chew

[1]Institute for Infocomm Research – Agency for Science, Technology and Research, Singapore
[2]Department of Electrical Engineering – National University of Singapore, Singapore
[3]Engineering and Computer Science – Victoria University of Wellington, New Zealand

*Abstract* — **To accurately determine the location of every sensor in large wireless sensor networks is normally very computation intensive and hence, is not cost-effective. On the other hand, a coarse estimate of the sensors' location is usually sufficient for many applications. The Area Localization Scheme (ALS) has been proposed as a scheme that estimates the position of every sensor within a certain region rather than identifying its exact location. Experimental results have shown that ALS is a promising technique with an accuracy of over 80% of sensors successfully localized. However, the accuracy of the scheme is dependent of the location of the sensors, as the regions into which the sensors are localized are not of uniform sizes. If a sensor is predicted to lie in a small region, then there is a higher probability that the prediction is inaccurate, and vice versa. In this paper, we aim to correct such discrepancy in defining positioning error. We propose to aggregate smaller regions to form more uniform regions within which nodes are localized. This effort improves the overall scheme's accuracy according to some predefined threshold.**

***Keywords - Wireless sensor networks; Positioning errors; Area localization;***

## I. INTRODUCTION

Localization is the process of determining the location of a sensor or tracking of a mobile. It is a challenging process as the reliance on technologies such as GPS is infeasible due to either the physical location, cost or energy constraints. There are many applications where localization is important, for example, to deploy low cost wireless sensors to detect the occurrence of natural disasters such as tsunamis and wildfires, ecosystem monitoring, health monitoring, and even military surveillance. The data that these sensors monitor and store are often interpreted with reference to the sensors' location. For some applications, wireless sensor networks may also involve small-scale sensor nodes which have limited computational power and memory.

Various localization techniques proposed take into account of factors such as the network topology, device capabilities, signal propagation models and energy requirements. Most localization schemes require the location of some nodes in the network to be known, often referred to as anchor nodes or reference nodes (or simply anchors). The localization schemes that use reference nodes can be broadly classified into two categories: range-based schemes and range-free schemes. In the range-based localization [3][4], specialized hardware is used to obtain point-to-point distance or angle estimates to compute location. In the range-free localization [5][6][7], no such hardware is required and instead, only connectivity information is utilized. There also exist schemes that do not require such reference locations in the network [8].

Many parameters can be used to compute the location of a sensor or mobile, namely, the use of angle-of-arrivals of the paths, delay times of received signals, maximum likelihood estimation, etc. Generally, estimation accuracy is a tradeoff between computation complexities and costs. However, finding the exact location of the sensor is not always necessary, instead, a coarse estimate of a sensor's position within a certain region is sometimes sufficient. The Area Localization Scheme (ALS) [1] has been proposed for this purpose and has been shown to successfully localize over 80% of sensors in a network [2]. However, due to the non-uniform regions that different sensors are localized into, the positional errors are overestimated – the accuracy of the scheme becomes dependent on the size of the region we referred to. Typically, the localization of a sensor node into a small region is more likely to be less accurate and vice versa. As a result, the likelihoods of having errors in ALS differ with different regions.

In this paper, we address this problem by aggregating much smaller regions to finally obtain regions with areas more evenly distributed. We felt in ALS, it is sufficient to just keep the tolerable positioning errors at each region to within some predefined threshold area. Unnecessarily making small tolerable positioning errors at some regions give no additional advantage – these regions are more likely to have positioning errors even if the channels are slightly deviated. In the following sections, we first provide an overview of ALS and related schemes, and then highlight the key issues before presenting the algorithm for aggregating regions. By defining new regions, there is more consistency in obtaining the performance metric as long as all the regions have an area less than the predefined threshold. Performance improvements in terms of probability of area positioning errors achieved by the approach are discussed before concluding the paper.

## II. AREA LOCALIZATION SCHEME

ALS can be classified as a coarse range-based localization technique. In principle, we are still using the received power strength to determine the appropriate region in which the sensor lies. This can be achieved by monitoring the beacon signals which a sensor can correctly receive from the anchors in the implementation – all anchors are assigned to transmit periodically at a few pre-determined transmit power levels which define the boundaries according to the desired area accuracy. In this way, this method can also be considered as range-free approach since the sensors simply only need to interpret the received beacon packets to make decision, with no real ranging needed.

ALS works on the principle that the received power is inversely proportional to the distance between the transmitter and the receiver. The knowledge of the path loss is assumed to be known in order to calibrate the system for positioning. For propagation over free space, the received power at a distance $d$ from the transmitter is given by:

$$P_r = P_t G_t G_r \left( \frac{\lambda}{4\pi d} \right)^2 \qquad (1)$$

Practical measurement shows that in an open space, the path loss is decreasing more than just -20dB/decade which is predicted by (1). The two-ray ground reflection model considers both the direct path and a ground reflection path, and gives a better prediction on the propagation path loss in an open area. The received power at a distance $d$ is given by:

$$P_r = \frac{h_r^2 h_t^2 G_r G_t P_t}{d^4} \quad \text{for} \quad d > \frac{4\pi h_t h_r}{\lambda} \qquad (2)$$

where $P_r$ is the received power, $P_t$ is the transmitted power, $d$ is the distance between the transmitter and receiver, $\lambda$ is the wavelength and, $h_t$ and $h_r$ are the heights of the transmitter and receiver respectively; $G_t$ and $G_r$ represent the transmitter and receiver gains respectively.

Using (1) or (2) and the threshold power that each sensor can receive, anchors can calculate the power required to reach different distances. Each reference node then broadcasts a set of beacon packets with increasing power levels, which need not be synchronized but have to be scheduled to avoid collisions. The transmitted set of power levels need not be the same for all the anchors. Neither is it necessary for the anchors to know each other's position, nor the set of power levels transmitted by one another.

Each anchor sends out beacon signals $N_r$ times at all the power levels in the set $\mathcal{P}$. The sensor nodes use a simple signal coordinate representation to indicate their location information to the sinks. Power contour lines can be obtained based on the set of beacon signal power levels $\mathcal{P}$ transmitted by each anchor as shown in Figure 1. The contour lines represent the furthest distances that the beacon signals at each power level can travel. Let us assume there are $n$ anchors referred to as $R_1$, $R_2$, … and $R_n$. For a sensor in the area, let the lowest transmitted power levels it receives from the $n$ anchors be $S_1$, $S_2$ … and $S_n$ respectively. $S_1$, $S_2$ … and $S_n$ are simple integer numbers associated with the different power levels rather than the actual signal strengths. Only the anchors and sinks know the mappings between numbers and actual power values. The signal coordinate is defined as the $n$-tuple $<S_1, S_2, \ldots, S_n>$ such that each $S_i$ ($i^{th}$ coordinate) is the lowest power level received from $R_i$. Assuming that each of the four anchors located at the corners transmit at three power levels, represented by {1,2,3}, a sensor located in the shaded region will have a signal coordinate <3,3,3,2> as this sensor can only receive packets that is sent at the highest power level from anchors 1, 2 and 3, whereas it can receive a packet at a lower power level from anchor 4.

In real conditions, fading and shadowing can cause the received power level to vary erratically from the expected signal strength predicted by the path loss model. Hence, the lowest signal power level received by a sensor from a reference node need not be the same for all the rounds 1 to $N_r$. Hence, a threshold value *CONF_LVL* is used which represents the confidence level with which the values $S_1$, $S_2$ .., $S_n$ can be estimated. This and other techniques to improve the accuracy of the signal coordinate information are discussed in [1] and [2].

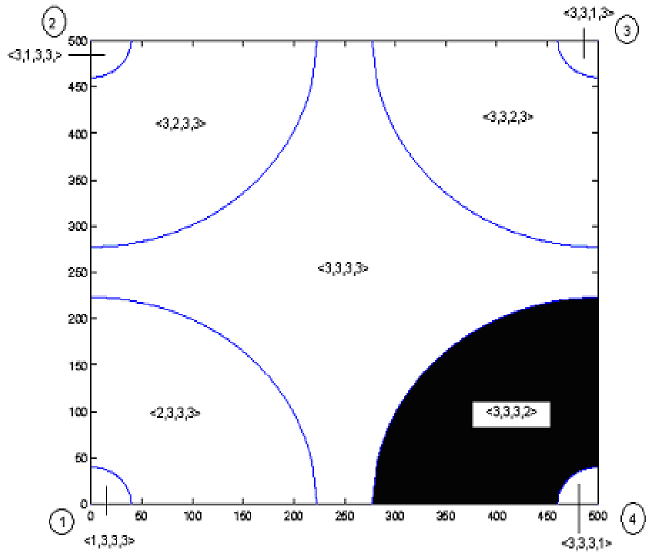


Figure 1 Example of ALS under Ideal Isotropic Conditions

## III. Non-Uniform Regions Problem

In the proposed ALS [1], the regions are obtained by overlapping regions given by the final signal coordinates of the sensor giving rise to regions of various sizes, varying from tiny regions (less than 1% of the network area) to as much as 5% of the network area. A sensor is said to be successfully localized if its actual location falls within the region obtained by ALS. Hence there is a higher probability of localizing a sensor node if its final region has a larger area. However, when the final region is very small, it is difficult to localize the node accurately in practice, due to deviation in the propagation path loss predicted by (1) in the presence of shadowing, fast fading and path loss constant. Thus, it is meaningless to localize a node within such regions which are very sensitive to positioning errors, because this degrades the overall positioning accuracy. Ideally we will like all regions to be equal in area, which is just at the sufficient size.

In other words, although the average granularity of the system is small (less than 4% of the overall area), there are regions with very high granularity and more regions with very low granularity. Consequently, this makes the accuracy dependent on the placement of the sensors. In order to standardize or quantitatively analyze the results of the ALS and compare it to other localization schemes, it is necessary that the individual regions demarked by unique signal coordinates are of similar sizes. In the next section, we present an approximation technique to address this problem.

## IV. Region Aggregation

The technique works on the principle that smaller regions, which surround a region of bigger area, are merged into the bigger region and thus, approximated to have the same signal coordinate as the bigger region. As a result, the total number of regions estimated by ALS decreases and variation in area sizes also reduces.

A neighboring region of region *A* is defined as the region whose signal coordinate differs from region *A*'s signal coordinate by at most one value. For example, regions with signal coordinates <1,2,3,2> and <1,2,4,2> are neighbors because they are different only at the $3^{rd}$ coordinate value and the $3^{rd}$ coordinates' values differ by 1. Another possible neighbor to <1,2,3,2> is <1,2,2,2>. However, <1,2,3,2> and <1,3,2,2> are not neighbors because both these regions differ at two coordinates, namely, the $2^{nd}$ and $3^{rd}$ coordinates (even if their values differ by 1).

Given a region's signal coordinate with *n* coordinate values, there are *2n* possible neighbors to the region. This is obtained by incrementing and decrementing each coordinate value. For instance, a region whose signal coordinate is <3,2> has four neighbors, namely, <2,2>, <4,2>, <3,1> and <3,3>. However, not all four of these possible signal coordinates might map to a region defined by ALS. A signal coordinate is said to be valid if and only if there exists a unique and distinct region defined by the signal coordinate. In other words, the region defined by the signal coordinate should be formed by the overlap of regions from all the anchors.

### A. *Algorithm*

The following algorithm performs the region aggregation on a user-defined network space where ALS is deployed to localize the sensors:

Step (i) Find all valid signal coordinates and the areas of the corresponding regions within the network.

Step (ii) Start with an available region that occupies the maximum area. Let that region be *A*. Define the maximum approximated region area of the network, *t*, to be 110% of the area of *A*.

Step (iii) Find all the valid available neighbors of *A* and their respective areas.

Step (iv) Find the neighbor with the least area and if the cumulative area of *A* and this neighbor does not exceed the threshold *t*, merge the neighbor's area into *A*. Mark this neighbor as unavailable.

Step (v) Repeat step (iv) for all valid available neighbors of *A* until the area of *A* exceeds the threshold area, *t*. Then, mark the region *A* as unavailable.

Step (vi) Repeat steps (ii) to (v) with the next largest region until all the regions become unavailable.

The algorithm is represented in the flowchart as shown in Figure 2.

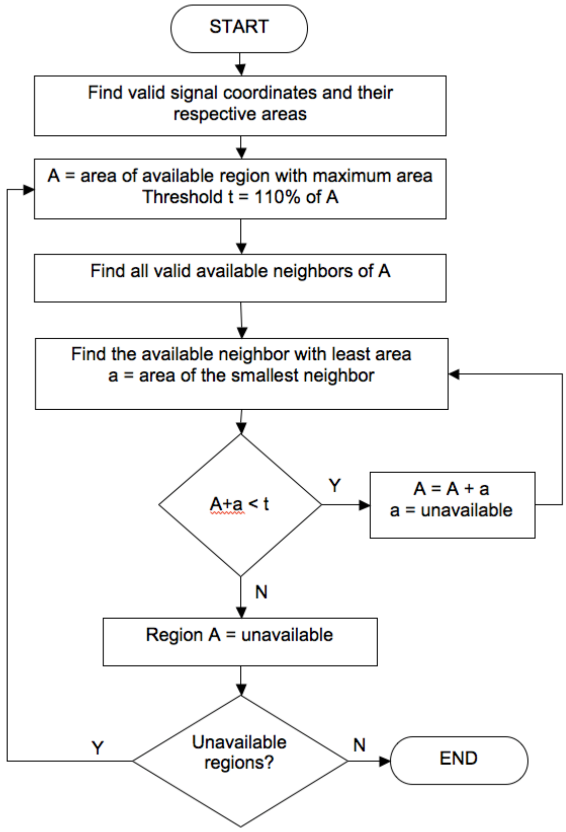


Figure 2 Flowchart of region aggregation algorithm

### B. *Application Example*

We now apply the algorithm on a case study involving a 30m × 30m network with eight anchors placed at the four corners and midpoints of the four sides of the network. These eight anchors send packets at eight different power levels: "1" being the lowest and "8" being the highest power level. When ALS is applied to this configuration, the 30m × 30m space is divided into regions as shown in Figure 3.

All valid signal coordinates and their corresponding regions are first obtained. From the list of signal coordinates, it can be seen that the signal coordinate <3,4,5,4,2,4,4,2> occupies the largest area (about 1.607% of the network area; cf: TABLE I. ). Therefore, *A* = <3,4,5,4,2,4,4,2>, shaded region in Figure 3.

Next, the list of all possible neighbors of A and their areas are gathered. Since there are eight anchors used in this setup, there are eight component values in a signal coordinate and thus, there are 2 × 8 = 16 possible neighbors. Table II lists all these neighbors and their areas, if they are valid regions.

The merge procedure (step (iv)) continues till the algorithm reaches the signal coordinate <3,4,5,4,3,4,4,2> whose area is 15. The cumulative area of A, so far, is 181, which is right at the threshold value. Thus, the approximation stops for the region *A* and goes to the next available region with the largest area. Figure 4 shows the final approximated region for the signal coordinate <3,4,5,4,2,4,4,2>.

TABLE I. LIST OF ALL POSSIBLE NEIGHBORS OF A, AND THEIR VALIDITY

| Signal Coordinates | Area* | Signal Coordinates | Area |
|---|---|---|---|
| $A$ = <3,4,5,4,2,4,4,2> | 164 | | |
| <2,4,5,4,2,4,4,2> | 42 | <3,4,5,4,1,4,4,2> | Invalid |
| <4,4,5,4,2,4,4,2> | 3 | <3,4,5,4,3,4,4,2> | 15 |
| <3,3,5,4,2,4,4,2> | Invalid | <3,4,5,4,2,3,4,2> | 2 |
| <3,5,5,4,2,4,4,2> | 5 | <3,4,5,4,2,5,4,2> | Invalid |
| <3,4,4,4,2,4,4,2> | Invalid | <3,4,5,4,2,4,5,2> | Invalid |
| <3,4,6,4,2,4,4,2> | Invalid | <3,4,5,4,2,4,3,2> | 2 |
| <3,4,5,3,2,4,4,2> | Invalid | <3,4,5,4,2,4,4,1> | Invalid |
| <3,4,5,5,2,4,4,2> | 5 | <3,4,5,4,2,4,4,3> | 15 |

*The area of a region is computed using a coarse method of marking the entire network area with points (in total, 3000 x 3000) and then the "area" of the region is represented by the number of dots within it.

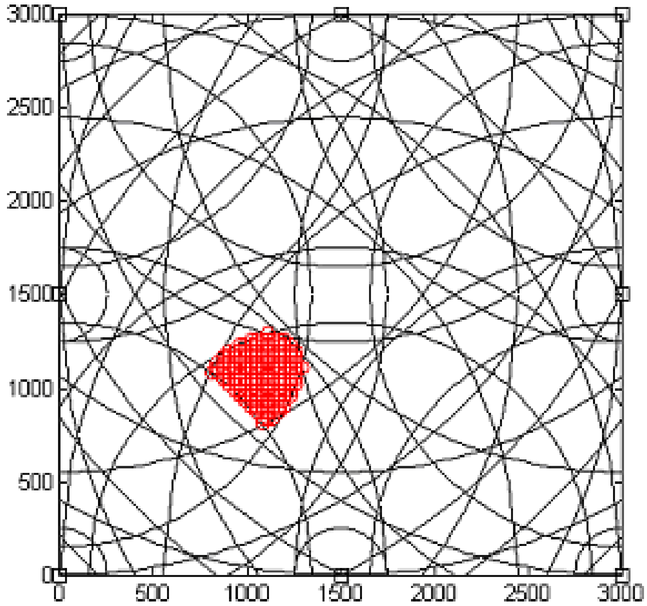


Figure 3 Shaded region represents the unique signal coordinate <3,4,5,4,2,4,4,2>

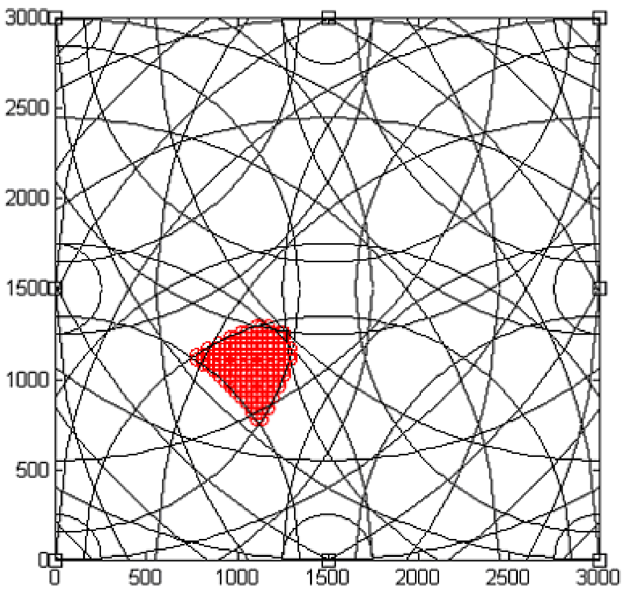


Figure 4 Final approximated region for <3,4,5,4,2,4,4,2>. The neighbouring smaller regions has been merged.

The entire algorithm continues for all the regions in the defined space until there are no more available regions to be merged. Finally, after the algorithm has been applied to all the regions for this particular network example, the approximated regions are as shown in Figure 5 . Each colored region represents a distinct final region after approximation. All the regions are of approximately the same area. Each unique region is surrounded by two or more regions that form the 'within-one-neighborhood' regions. It can, however, be observed that there are certain regions which cannot be approximated. This is due to the inherent condition set in the algorithm. A neighboring signal coordinate is defined to be one that differs from the actual signal coordinate by at most one value. However, there are certain 'neighbors' whose signal coordinate values differ by more than one value and these regions are therefore neglected. However, based on the performance of this algorithm, the definition of 'neighboring signal coordinate' can be modified accordingly.

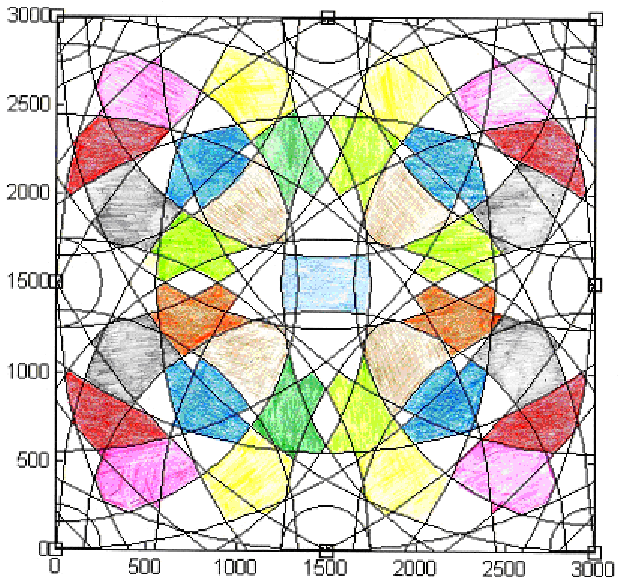


Figure 5 Final approximated regions for a 30m x 30m space using 8 anchors with 8 power levels

## V. ALS WITH REGION AGGREGATION

Using the aggregation procedure with a threshold of 1.8% of the network area, it can be seen that for a 30m × 30m space, with eight anchors sending out packets at eight power levels, the number of regions derived from ALS has decreased and, the granularity of the area has also increased with the removal of many small regions. Thus, accuracy is much less dependent on the placement of the sensors, if at all. Random distribution of the sensors achieves a consistent high accuracy, i.e. higher % of sensors localized successfully.

In order to validate the effectiveness of the proposed approximation technique, we apply the aggregation algorithm on the results obtained from the experimental setup of ALS in an outdoor environment with obstacles [2] (i.e. a square plot in a park with dimensions 30m × 30m; 8 anchors transmitting at 8 power levels). In that experiment, out of the total thirty sensors that were placed within the square plot, sixteen sensors were localized by ALS within the predicted region, nine sensors could be localized within one-hop and the remaining five sensors (shown in Figure 6) could not be localized. Through the application of the region aggregation algorithm, positions of twenty-three out of the thirty sensors are correctly detection, and six are within one-hop. We show that we are able to reduce that number of unlocalized sensors by providing a more consistent size for the localized area.

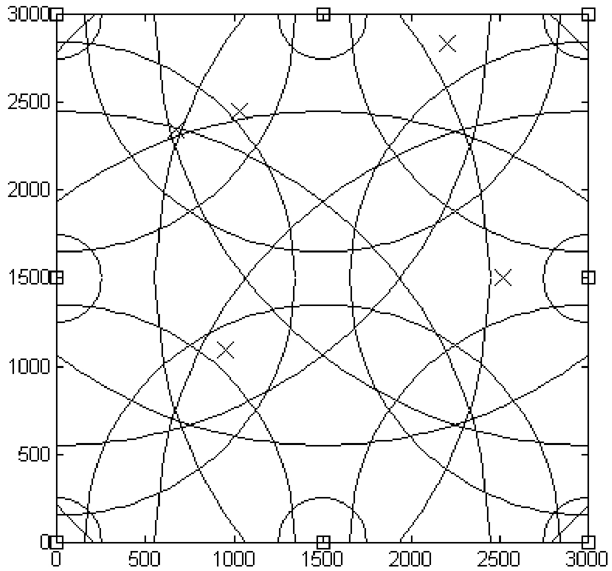


Figure 6 Location of sensors (denoted by 'X') that were unable to be localized by ALS during the outdoor experiment (open space with obstacles).

Figure 7 to Figure 9 show the various scenarios involved in localizing some of the five sensors which could not be localized earlier. In Figure 7, the shaded region represents the approximated region after region aggregation and the "×" represents the actual location of the sensor. The presence of "×" within the shaded region implies a successful localization. In Figure 8, while the sensor does not lie within the predicted region, it is within one of the adjacent regions and can be localized as a result of region aggregation.

In Figure 9, the sensor neither lies within the aggregated region nor the in one region neighborhood, hence, we consider the sensor cannot be localized (even after region aggregation) and its location is considered to be unknown or undefined. The inability to localize the sensor can be attributed to (non) line-of-sight and/or other fading or shadowing problems, where the sensor was unable to receive beacon packets from the anchor nodes properly. Thus, the signal coordinates of this sensor may be corrupted.

Applying region aggregation on the experimental setup presented in [2], we are able to localize four out of the five sensors within the one region neighborhood. The accuracy of ALS has improved with aggregation, as demonstrated by the higher number of successful localization of sensors. The performance of ALS with region aggregation has also been evaluated using simulations and the results showed a significant improvement in successful sensor localization.

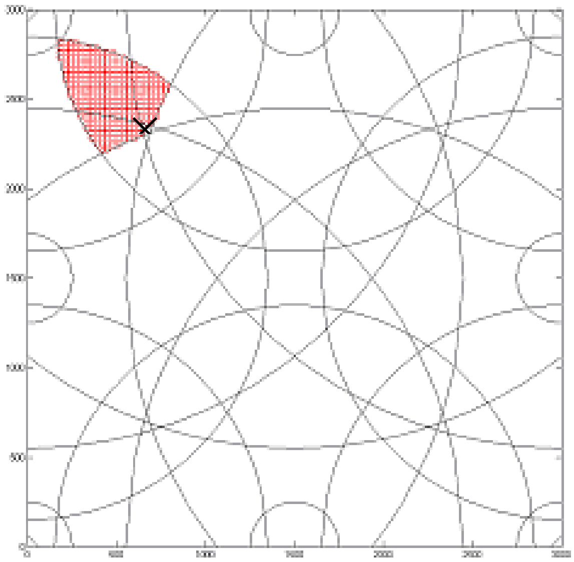

Figure 7 Sensor that is successfully localized

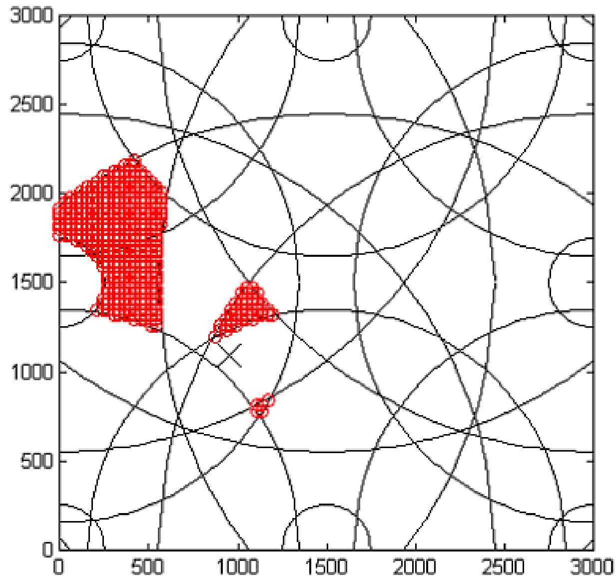


Figure 8 Sensor that is localized within one neighbourhood region

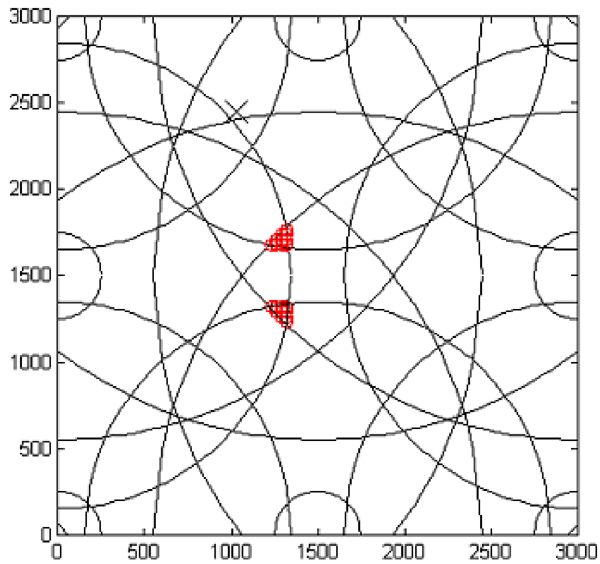


Figure 9 Unsuccessful localization

## VI. Conclusion and Future Work

The ALS is a simple range-free localization technique that provides a coarse estimate of the location of the sensor nodes within certain regions of a network. The sensors simply record information based on the beacon packets received from the anchors, which locates the sensors with pre-calibrated data. Experimental results have shown that up to 80% of the sensors can be localized within one neighboring region (or one region neighborhood.)

With the region aggregation algorithm, many small neighboring regions are merged into bigger regions in order to create regions of similar size, thus increasing the granularity of the regions as well. The results of ALS can now be easily compared against other localization schemes, as the accuracy of ALS is constant regardless of the placement of sensors within the network. Moreover, region aggregation has also increased the accuracy of the ALS with nearly all the sensors localized within the aggregated region or within one neighboring region.

Moving forward, routing protocols can be designed to exploit the localization information provided by ALS for directional and/or location based routing. Similarly, navigation algorithms can, for instance, direct a mobile robot to search a region within the network based on the information provided by a sensor. If the sensor is not present in the specified region, the robot could extend its search to all the neighboring regions to locate the sensor and investigate any phenomenon detected by the sensor. Since, the region aggregation technique provides a higher accuracy of localization, it makes ALS more promising for implementation and deployment in real-life conditions.

## References


[1] Q. Yao, *et al.*, "An Area Localization Scheme for large Wireless Sensor Networks", *Proceedings of the IEEE 61st Semiannual Vehicular Technology Conference (VTC2005-Spring)*, May 30-Jun 1, 2005, Stockholm, Sweden.

[2] V. Chandrasekhar, *et al.*, "Experimental Analysis of Area Localization Scheme for Sensor Networks", *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC)*, Mar 11-15, 2007, Hong Kong.

[3] S. Gezici, *et al.*, "Localization via Ultra Wide Band Radios", *IEEE Signal Processing Magazine*, Vol. 22, No. 4, Jul 2005, pp. 70-84.

[4] N.B. Priyantha, A. Chakraborty and H. Balakrishnan. "The Cricket Location-Support system", *Proceedings of the 6th ACM International Conference on Mobile Computing and Networking (Mobicom)*, Aug 6-11, 2000, Boston, MA, USA.

[5] T. He, *et al.*, "Range-Free Localization Schemes for Large Scale Sensor Networks", *Proceedings of the 9th ACM International Conference on Mobile Computing and Networking (Mobicom 2003)*, Sep 14-19 2003, San Diego, CA, USA.

[6] D. Niculescu and B. Nath. "DV Based Positioning in Ad Hoc Networks", *Telecommunication Systems,* Vol. 22, No. 1-4, pp 268-280, 2003.

[7] S.Y. Wong, *et al.*, "Density-aware Hop-count Localization (DHL) in wireless sensor networks with variable density", *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC)*, Mar 13-17, 2005, New Orleans, LA, USA.

[8] L. Doherty, K. Pister and L. Ghaoui, "Convex Position Estimation in Wireless Sensor Networks", *Proceedings of the 20th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, Apr 22-26, 2001, Anchorage, AK, USA.