Anchor Node Placement for Localization in Wireless Sensor Networks

by

Benjamin Tatham

A Dissertation submitted to
the Faculty of Graduate Studies and Research
in partial fulfilment of
the requirements for the degree of
Master of Applied Science

Ottawa-Carleton Institute for
Electrical and Computer Engineering

Department of Systems and Computer Engineering
Carleton University
Ottawa, Ontario, Canada
November 2010

The undersigned recommend to

the Faculty of Graduate Studies and Research

acceptance of the Dissertation

Anchor Node Placement for Localization in Wireless Sensor Networks

Submitted by **Benjamin Tatham**

in partial fulfilment of the requirements for the degree of

Master of Applied Science

Dr. Thomas Kunz, Supervisor

Dr. H.M. Schwartz, Department Chair

Carleton University

2010

Abstract

Applications of wireless sensor network (WSN) often expect knowledge of the precise location of the nodes. One class of localization protocols patches together relative-coordinate, local maps into a global-coordinate map. These protocols require nodes that know their absolute coordinates, called anchor nodes. While many factors influence the node position errors, in this class of protocols, the placement of the anchor nodes significantly impacts the error. Through simulation, using the Curvilinear Component Analysis (CCA-MAP) protocol, we show the impact of anchor node placement and a set of rules to ensure the best possible outcome. Scientists are thus enabled to focus on the sensed data, and rely on a maximum node position error.

Dedicated to my wife and children who supported me through the long process of this research.

Acknowledgments

I would like to thank Professor Thomas Kunz for his guidance and encouragement that made this research possible. Further, I would like to thank Li Li whom, along with Professor Kunz, performed the initial CCA research and wrote the basis for the Matlab simulations used in this thesis. Finally, my colleagues Shafagh Alikhani and Ben Gardiner who provided a tireless ear to listen to ideas throughout the process.

Table of Contents

List of Figures

Chapter 1

Introduction

Scientists, engineers, and researchers use wireless sensor networks (WSN) for a wide array of applications. Many of these applications rely on knowledge on the precise position of each node. While some may only require relative coordinates within the network, most biological, geophysical, and other scientific applications require coordinates on a global coordinate system. Perhaps the obvious solution is for each node in the network to be equipped with GPS or other location positioning services. However, constraints on cost, power consumption, as well as visibility of satellites forbids this.

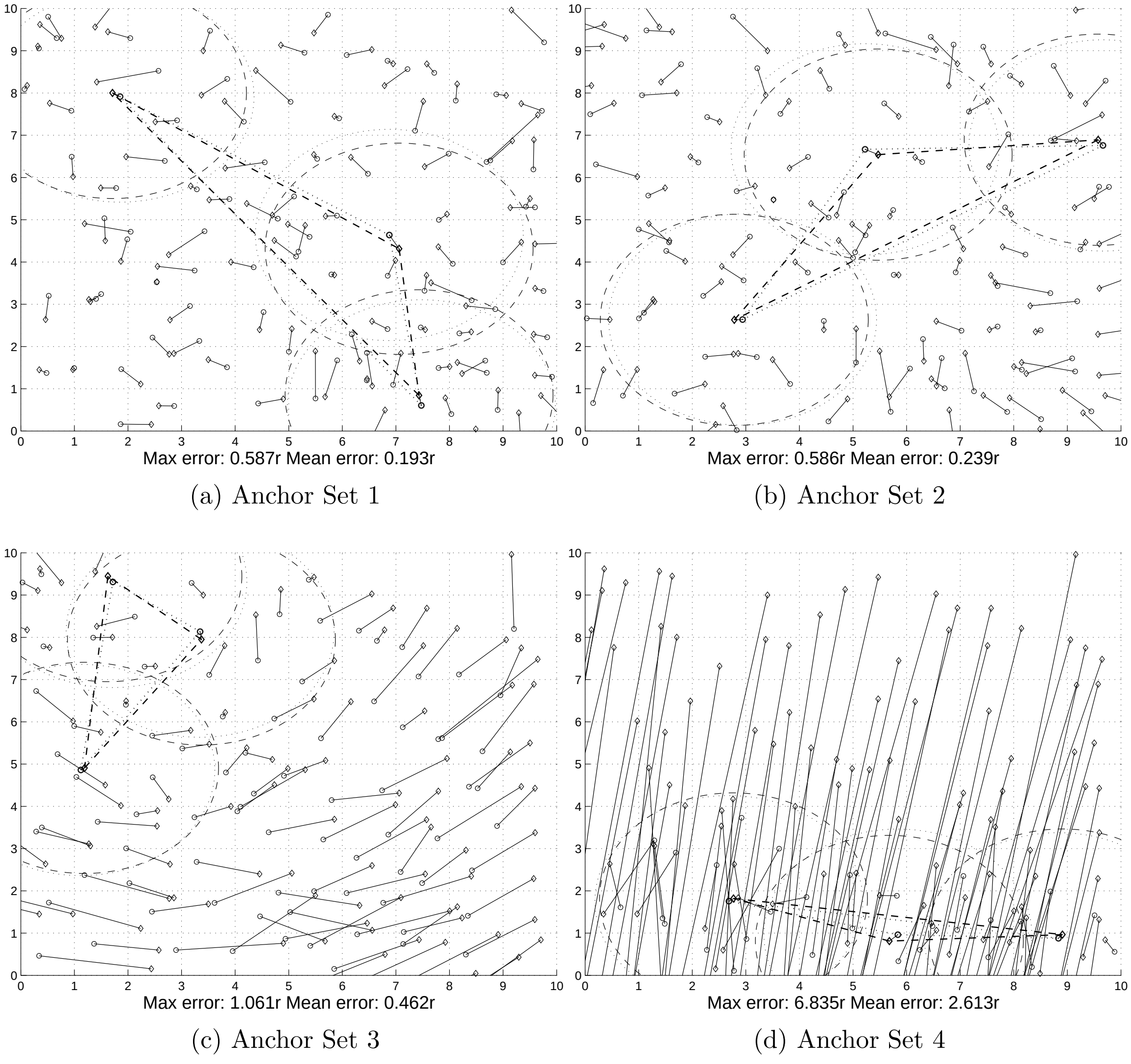
Many protocols have been proposed [1–3] to calculate relative positions amongst the nodes of a network. They vary in the required network functionality in terms of radio ranging or range-free. Radio ranging involves specialized hardware for to measure the distance between nodes based on physical data like signal strength or transmission delays. In order to convert from relative to global coordinates, some of the nodes do require a local source of global coordinates. This can be achieved by operators recording the global coordinates during network deployment, by embedding a GPS receiver in a subset of the nodes, or some other source. We call these enhanced nodes anchors. Here, we explore the effect of anchor node placement within the network on the overall localization errors, on a network-wide basis. This provides network planners with a set of general rules to minimize the number of anchor nodes required while avoiding poor node localization, allowing scientists to assume a maximum position error during their own research. Further, based on application requirements of location accuracy, planners can minimize the cost of the network associated with anchor nodes by using the minimum number and best position.

1.1 Motivation

During previous work designing localization protocols [4, p. 11] [2, p.2], authors often choose anchors at random within the network. Frequently, they simulate the network multiple times with different anchors in order to statistically exclude anchor node placement from their results.

Our initial investigations and simulations demonstrate that indeed the placement of anchor nodes in the network did have an often dramatic effect on the location error. The four plots shown in Figure 1.1 graphically establish that anchor node position does make a difference. Each plot is the same network with a different choice of three anchors. A line is drawn between the actual and calculated position of each node. The circles show the radio range of each anchor, and a triangle is drawn between the three anchors for clarity. While the first two choices have reasonable errors, the the third has almost two and a half times the mean error of the first. Further, the fourth choice has an extremely poor performance.

In practical applications, scientists and engineers do not have the luxury of running their localization protocols multiple times to determine the best location. In fact, this would defeat the purpose of the localization protocol in the first place. Therefore, it is critical to be able to assign a maximum location error to all the nodes in a network so that data processing that relies on location can effectively take the location error



(a) Anchor Set 1

(b) Anchor Set 2

(c) Anchor Set 3

(d) Anchor Set 4

Figure 1.1: Location error for 100 anchor set choices in the same network

into account.

1.2 Thesis Contribution

We establish that by avoiding certain poor anchor node placements, extremely poor location errors can be avoided. Further, we demonstrate that other than these edge cases, the location error falls within a statistically insignificant range, where the range is a function of network connectivity and topology. Specifically, we show that the probability of extremely high location error result from anchor nodes being roughly in a geographically straight line. As the anchor nodes are spread out from a straight line, the probability of high errors decreases, leaving network designers a relatively simple chore when choosing anchor nodes locations.

While some papers have touched upon anchor node placement, we have yet to come across a comprehensive study of the optimal anchor node placement. This paper provides a comprehensive study of possible anchor node placements and their differing effects on overall network localization accuracy.

1.3 Methodology

Many localization protocols and algorithms provide a set of relative coordinates that are then transformed into global coordinates. For the purpose of this research, we chose CCA-MAP [3, 4] as the algorithm to provide simulation results. A Matlab ©simulation of this algorithm already existed from Li Li [3], and was modified to provide the necessary output statistics presented here. CCA-MAP is described in more detail in 2.2.3 on page 9.

1.4 Thesis Organization

A brief background of Wireless Sensor Networks and localization protocols in general are presented in Chapter 2. Chapter 3 presents the limited related work in anchor node placement. Chapter 4 contains the various anchor node placements and summaries of how they perform. The cause of the extreme edge cases comes to light in Chapter 5, along with a discussion of the Procrustes algorithm. Chapter 6 presents conclusions and future work.

Chapter 2

Overview of Wireless Sensor Networks and Localization

2.1 Wireless Sensor Networks

A wireless sensor network (WSN) consists of a set of nodes tasked with sensing environmental phenomenon at or near each node. Nodes communicate via radios to send their data back to a central acquisition system. Nodes are typically small, cheap devices and are designed with power efficiency in mind to prolong the lifetime of the network's ability to collect data. Nodes are often distributed in the field of interest randomly, sometimes even by dropping them from the air, as on a military battlefield. Other times, they are placed in specific, but unknown a priori, locations, as in placing them in bird nests [5]. Or, they may be rolled into a transportation tunnel to give firefighters and emergency crews current information about heat and oxygen levels [6]. The list of applications goes on and on.

A number of issues arise when designing a WSN. Each node must be able to communicate with each other and send data to a central collection site. Each node must know what time it is, for purposes of data sampling, and often for routing protocols as well. Further, each node must know where it is so spatial data can be properly correlated. Location can also be useful for geographic routing protocols. This thesis focuses on determining the location of each node as accurately as possible.

2.2 Localization Protocols

There are two general classes of localization protocols: ranging and range-free.

Ranging protocols rely on information from the radio. With this information, a fairly accurate network topology can be built. Ranging techniques can use a variety of metrics to build the network topology. These include Time-of-Arrival (TOA), like GPS [7], Time-Differential-of-Arrival (TDOA) [8], Angle-of-Arrival (AOA) [9], or Received-Signal-Strength-Indicator (RSSI) [10].

However, the special hardware and power requirement to perform these ranging techniques is counter to the goal of low-cost, low-power nodes, and thus we exclude ranging protocols from our study. Regardless, if a ranging protocol does build a relative map, and then does a post processing step by mapping this relative map to a global map based on a subset of anchor nodes, the results of this thesis apply to ranging as well as range-free protocols.

Range-free protocols do not rely on any specialized hardware for additional information. Rather, they rely solely on network connectivity, specifically knowledge of their direct neighbors. Often, a node will collect information about their direct neighbors' neighbors as well, known as two-hop information. Knowledge of each further node requires more information to be shared and there transmitted between nodes, thus requiring more power for radio transmission. It is for this reason that only one-hop or possibly two-hop knowledge is preferred.

2.2.1 Ad Hoc Positioning System

Niculesu, et al. propose a distributed localization algorithm known as Ad Hoc Positioning System (APS) [1]. It is similar to GPS in that is uses triangulation to determine node position. In APS, each node maintains a table of distances to each anchor. The distance can be represented as a hop count, estimated distance using RSSI, or Euclidean distance. As a distributed algorithm, each node determines its own position based on the distances to the anchor nodes. Thus, APS does not perform well in anisotropic network, that is networks with holes or "C" shapes in the topology, because the communication distance can be far greater than the geometric distance between two nodes.

In its simplest form, APS uses a propagation technique called DV-HOP to determine distances between nodes. DV-HOP is based on classical distance vector exchange from general network protocols like TCP/IP. Each node maintains a table of hop counts between all known nodes. Each node exchanges this table only with its direct neighbors. When an anchor has discovered a hop count to another anchor, the anchor estimates the average distance for each hop since it knows the absolute location of itself and the other anchor. This correction factor is sent to the entire network. DV-HOP thus minimizes the amount of data that must be transmitted in the network.

Further, APS can employ a propagation technique called DV-distance. DV-distance is similar to DV-hop except that it uses RSSI to determine each hop distance and sends this distance instead of hop count. This difference allows DV-distance to effectively detect holes and curves in the network as each anchor can see that the transmission path between them is larger than the Euclidean distance.

2.2.2 MDS-MAP

Shang, et al. attempt to correct the errors introduced by APS and other distributed algorithms through a centralized localization algorithm called MDS-MAP(C) [2]. MDS-MAP(C) is divided into three phases. In phase one, shortest path distances or hop counts are exchanged via distance vector exchange, similar to APS. This provides a rough estimate of the distance between each pair of sensors. In phase two, multi-dimensional scaling (MDS) is applied, resulting in a relative map. MDS is a general data analysis tool originating from psychophysics to transform data from many to few dimensions. In simple terms, MDS takes a set of distances between points and creates a structure that fits those distances. Often, it is used for general data visualization. In this case, the relative map conforms closely to the pair-wise distances provided. In phase three of MDS-MAP(C), the relative map is transformed into a global coordinate system using at least three anchors.

The authors provided a modified, distributed version, MDS-MAP(P) [11]. This variation simply divides the network into smaller, more manageable sections to the algorithm can be performed locally, with the limited node resources available. Each local map is then merged together, although this part of the algorithm is not distributed. Local map merging begins at a randomly selected node's local map, and chooses the local map with the most overlapping nodes. A linear transformation of translation, reflection, rotation and scaling is used to combine the two relative coordinates systems into one. The process continues until all the local maps are merged together.

2.2.3 CCA-MAP

Li, et al, propose a similar style algorithm to MDS-MAP called CCA-MAP [3, 4]. It is similar in that it generates relative, local maps of sections of the network and then patches them together into a global coordinate system. CCA-MAP improves on MDS-MAP in that the algorithm is more efficient. MDS is a non-linear reduction algorithm and has a computational cost of $O(n^3)$. CCA [12] on the other hand, is a self-organized neural network performing quantization and non-linear projection. CCA-MAP has a total computational cost of $O(n^2)$. CCA runs in a series of iterations, where each iteration has a computational cost of $O(n)$.

CCA-MAP has four phases. In the first phase, each node builds a local map of nodes within R hops. For that local map, the shortest distance matrix is accumulated, as in APS and MDS-MAP. The second phase involves performing the CCA algorithm itself on each local map, generating relative coordinates for each node in the local map. In phase three, the local maps are merged together, as in MDS-MAP(P), and finally, in phase four, the relative coordinates are transformed into absolute coordinates based on the known coordinates of the anchor nodes. Phase four can only be performed with a minimum of three anchors for *2D* space or four anchors for *3D* space.

CCA-MAP is flexible as to where computations can be performed. Local map calculations can be performed at the nodes themselves, if computing resources allow, or outsourced to more powerful gateway nodes or a central server. Further, local map merging can be performed in parallel at selected nodes in the network, or again at a central server. Further, if in any sub-map sufficient anchors are found, then absolute coordinates can be calculated.

Chapter 3

Related Work on Anchor Node Placement

While much attention has been paid to localization accuracy and computational effort, anchor node placement is often recognized, but dismissed as future study.

3.1 Empirical Evidence

Often, authors will come across anchor placement by accident and discuss it based on their own empirical evidence. Shang, et al. [2, p. 964] and Li, et al. [3, p. 11] both choose anchors at random within the network. Although, Shang does mention that a co-linear set of anchors chosen in one example ”represents a rather unlucky selection”, without supporting evidence of why this is unlucky.

Earlier work by Doherty, et al. [13] requires anchor nodes to be placed at the edges, and ideally at the corners of the network. In this case, however, the algorithm is a simple constraint problem. One constraint requires that all the unknown nodes be places within the convex hull of the anchors, and therefore, better results are obtained when anchors are at the corners.

3.2 Explicit Studies of Anchor Node Placement

While few, there have been a some explicit studies of anchor node placement. Hara, et al. [14] propose a method of choosing anchor node locations to achieve a specific accuracy target. The proposal, however, only applies to rectangular network areas and that anchor nodes must be placed at the center of a sub-rectangle of the original rectangle when divided into equal sized rectangles. Further, it assumes simple RSSI-based localization.

Ash, et al. [15] provide analytical proof that placing anchor nodes uniformly around the perimeter of a network proves the best results, in the absence of any other information about the sensor node positions. However, again this assumes a rectangular network, and more importantly a simple localization algorithm like [13] or other multi-lateration techniques. When using all inter-node distances at once, as in MDS-MAP and CCA-MAP, this analysis breaks down.

Karl and Willig dedicate an, albeit short, sub-chapter to the *Impact of anchor placement* in their book [16, p. 247-248]. Referencing [13] and [17], again they defer to perimeter anchor placement as the optimal choice. Unfortunately, the technique proposed involves adaptive deployment, whereby a mobile node with absolute positioning available, like GPS, wanders through the network and attempts to determine the optimal anchor placements as it travels. For the purposes of a priori planning, this technique is not feasible.

Cheng, et all. [18] present a novel technique to handle the effects of adverse anchor placement, specifically in clumps. The algorithm, *HyBloc*, is a hybrid of MDS and proximity-distance map (PDM) [19]. It works by using MDS to add artificial, secondary anchor nodes using MDS to improve the performance of PDM, since on its own, MDS can adequately and efficiently determine the position of a few nodes based on arbitrarily place primary anchor nodes. PDM is shown to a be good choice for anisotropic networks.

Other studies also focus on the effect of indoor conditions and anchor placement as it relates to RSSI and other radio propagation measurements [20].

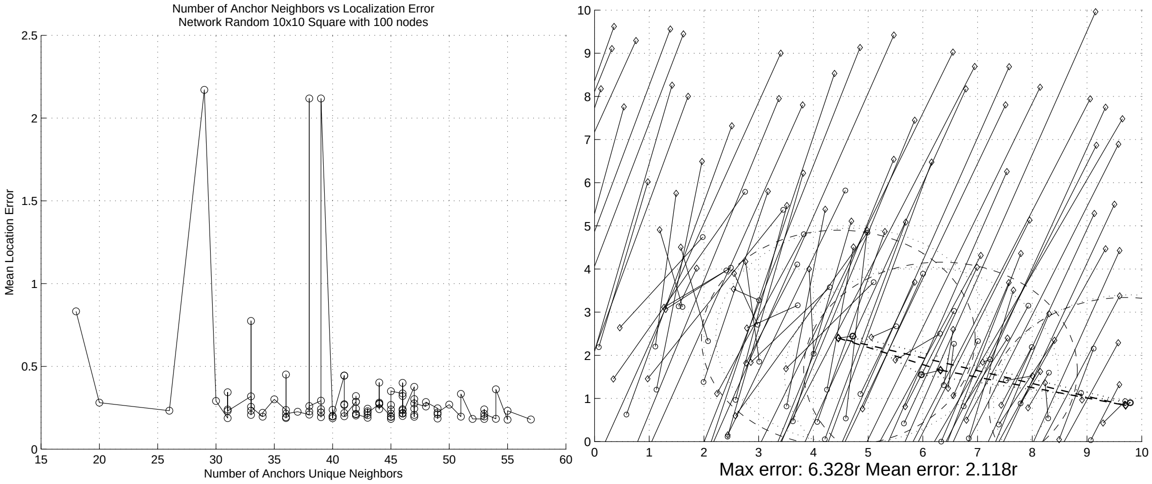
Chapter 4

Survey of Anchor Node Placements

4.1 Measuring Location Error

Before searching for the best anchor node placement we must first define what *best* means, in terms of location error. First of all, location error is measured as a factor of radio radius (or range). Since this study addresses range-free networks, and thus relies solely on network connectivity, the actual units of distance do not matter for general study. What is important is how many other nodes in the network fall within the radio range of a given node. The average numbers of nodes within range of a each node is known as connectivity.

Every network has its own application requirements, and thus there are many options for what statistics to examine for accessing the quality of locations. The simplest criteria are to look at the mean and maximum location error across all nodes in the network. For the most part, this is the basis for the results in this study. However, this assumes that all nodes in the network must be used in the final results. If the network designers know which nodes have poor locations, they may wish to exclude these nodes from the final results. Therefore, if the economics of the network deployment allow, it may be beneficial to look at the best, for example, 80% of nodes in the network. In practice, the designers do not know which nodes to exclude, so this



(a) The Data

(b) The Network

Figure 4.1: Number of Anchor Neighbors vs Location Error

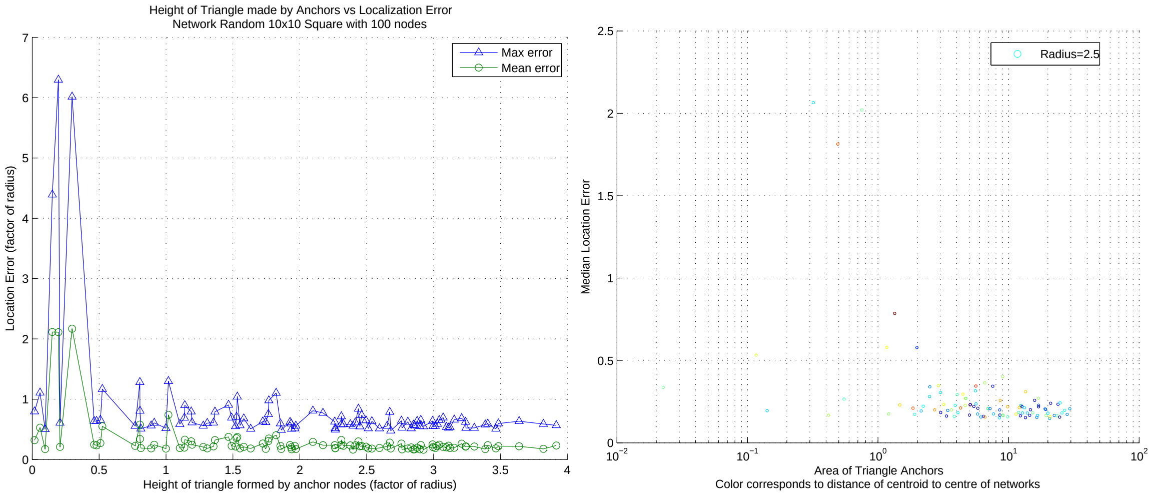
study also attempts to identify the worst areas of a network for location accuracy, depending on the anchor node placement.

4.2 Insignificant Factors

A number of hypotheses of anchor node placement were attempted, but showed no significant correlation to location error. The following is a summary of what some may assume matter, but in reality does not.

4.2.1 Number of Neighbors

One theory is that the nodes that are one-hop away from an anchor would have better location accuracy. As shown in Figure 4.1, the effect of the number of neighbors is random. Figure **??** shows the best of the chosen anchor sets, in terms of mean node location error and the difference between the actual and calculated locations. Further, the three nodes with the worst locations are within one hop of an anchor. In the end, having more neighbors means that the error of the anchor node itself has



(a) Triangle Height

(b) Triangle Area

Figure 4.2: Anchor Triangle Height and Area

a better calculated location accuracy as show in TODO, but this does not translate into better network-wide accuracy.

4.2.2 Anchors Farther Apart

Another seemingly obvious optimal anchor placement would be to have the anchors as far apart as possible. This should ensure that the resulting linear transformation generated takes into account as much of the network variation as possible. TODO...

4.3 Significant Factors

Some other network layout metrics do have a stark impact on the overall location accuracy.

4.3.1 Triangle Formed by Anchors

The triangle formed by a set of three anchors turns out to be a good indicator of the worst case of anchor placement as shown in 4.2.

List of References

[1] D. Nicelescu and B. Nath, “Ad Hoc Positioning System (APS),” in *Global Telecommunications Conference (GLOBECOM 2001)*, vol. 5, (San Antonio, Texas), pp. 2926–2931, November 2001.

[2] Y. Shang, W. Rumi, Y. Zhang, and M. Fromherz, “Localization from connectivity in sensor networks,” *Parallel and Distributed Systems, IEEE Transactions on*, vol. 15, pp. 961 – 974, nov. 2004.

[3] L. Li and T. Kunz, “Localization applying an efficient neural network mapping,” in *Proceedings of the 1st international conference on Autonomic computing and communication systems (Autonomics 2007)*, vol. 302, (Rome, Italy), 2007.

[4] L. Li and T. Kunz, “Cooperative node localization using nonlinear data projection,” *ACM Transactions on Sensor Networks (TOSN)*, vol. 5, February 2009.

[5] J. Kumagai, “The secret life of birds,” *IEEE Spectrum*, vol. 41, pp. 42–48, April 2004. This is the article that originally piqued my interest in WSN, and inspired me to pursue thesis research in this area.

[6] “RUNES (Reconfigurable Ubiquitous Networked Embedded Systems) IST Project.” http://www.ist-runes.org, 2004-2006.

[7] B. Wellenhoff, H. Lictenegger, and J. Collins, *Global Positioning System: Theory and Practice*. Berlin, Germany: Springer, 4th ed., 1997.

[8] A. Savvides, C. Han, and M. Strivastava, “Dynamic finegrained localization in ad-hoc networks of sensors,” in *Proceedings of the 7th Annual International Conference on Mobile Computing and Networking (MobiCom’01*, (Rome, Italy), pp. 166–169, July 2001.

[9] D. Nicelescu and B. Nath, “Ad hoc positioning system (APS) using AOA,” in

Twenty-Second Annual Joint Conference of the IEEE Computer and Communications (INFOCOM 2003), vol. 3, (San Francisco, California), pp. 1734–1743, April 2003.

[10] N. Patwari, A. Hero, M. Perkins, N. Correal, and R. O'Dea, "Relative location estimation in wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 51, no. 8, pp. 2137–2148, 2003.

[11] Y. Shang and W. Ruml, "Improved mds-based localization," in *INFOCOM 2004. Twenty-third AnnualJoint Conference of the IEEE Computer and Communications Societies*, vol. 4, pp. 2640 – 2651 vol.4, 7-11 2004.

[12] P. Demartines and J. Herault, "Curvilinear component analysis: a self-organizing neural network for nonlinear mapping of data sets," *Neural Networks, IEEE Transactions on*, vol. 8, pp. 148 –154, jan 1997.

[13] L. Doherty, K. Pister, and L. El Ghaoui, "Convex position estimation in wireless sensor networks," in *INFOCOM 2001. Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, vol. 3, pp. 1655 –1663 vol.3, 2001.

[14] S. Hara and T. Fukumura, "Determination of the placement of anchor nodes satisfying a required localization accuracy," in *Wireless Communication Systems. 2008. ISWCS '08. IEEE International Symposium on*, pp. 128 –132, 21-24 2008.

[15] J. Ash and R. Moses, "On optimal anchor node placement in sensor localization by optimization of subspace principal angles," in *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*, pp. 2289 –2292, march 2008.

[16] H. Karl and A. Willig, *Protocols and Architectures for Wireless Sensor Networks.* John Wiley & Sons, Ltd, 2005.

[17] C. Savarese, J. M. Rabaey, and K. Langendoen, "Robust positioning algorithms for distributed ad-hoc wireless sensor networks," in *ATEC '02: Proceedings of the General Track of the annual conference on USENIX Annual Technical Conference*, (Berkeley, CA, USA), pp. 317–327, USENIX Association, 2002.

[18] K.-Y. Cheng, K.-S. Lui, and V. Tam, "Hybloc: Localization in sensor networks with adverse anchor placement," *Sensors*, vol. 9, no. 1, pp. 253–280, 2009.

[19] H. Lim and J. Hou, "Localization for anisotropic sensor networks," vol. 1, pp. 138 – 149 vol. 1, mar. 2005.

[20] R. Zemek, M. Takashima, S. Hara, K. Yanagihara, K. Fukui, S. Fukunaga, and K. Kitayama, "An effect of anchor nodes placement on a target location estimation performance," in *TENCON 2006. 2006 IEEE Region 10 Conference*, pp. 1 –4, 14-17 2006.

Appendix A

Matlab©Simulation Code

All source code used to simulate CCA-MAP and generate resulting data can be found in Google©Code. The code was originally written by Li Li and modified for the purposes of readability and different output data and plots.

See `http://code.google.com/p/sim4j/source/browse/#svn/trunk/thesis/matlab` for details.