



UBIQUITOUS (CB)RN(E) SENSOR NETWORKS

Analysis and Framework Study

9 May 2014

Kyle Grunden
Gina Guerra
William Leonard

Sponsor: Allen Harvey, TASC, INC.

Table of Contents

Abstract.....	5
Intro.....	6
The Sponsor	6
History of Security Concerns.....	6
Wireless Sensor Networks	6
Assumptions.....	8
The Model.....	9
Crowd Density.....	9
Source and Sensor Strength Measurement.....	9
Source Strength.....	9
Sensor Strength.....	10
See Distance.....	10
Sensor Creation.....	10
The Grid.....	10
The Parade Spectators	11
The Parade Participants	11
Source Movement.....	11
Source Localization	12
Source Emanation and Sensor Detection	12
Location Estimate.....	12
Outputs	13
GIS Model.....	13
Model Verification	14
Analysis and Results.....	14
Metric Relationships	14
Percent of Time Source is Seen.....	14
Average Error when Source is Seen	14
Maximum Error when Source is Seen	14
Last Position Error	15
Maximum Location Error	15
Base Case Analysis	15
Percent Seen	15

Average Error When Seen.....	15
Maximum Error When Seen.....	16
Last Position Error	16
Maximum Location Error	17
Excursion Variables	17
Negative Detections.....	18
Average Error When Seen.....	19
Random Source Movement	19
Crowd Movement	21
Triangular Vs. Uniform Vs. Base case.....	22
False Detections	23
Data Clusters	24
Framework.....	25
Sensors.....	25
Negative Detections.....	25
False Detections	25
Distributions.....	25
Final Note.....	26
Future Work.....	26
Multiple Target Tracking.....	26
Time Based Tracking	27
Fluid type dispersion.....	28
References	29
Appendix 1: Deriving See Distance	30
Appendix 2: Density Charts.....	31
Chart 1: Crowd Density of .25	31
Chart 2: Crowd Density of .925	31
Chart 3: Crowd Density of 1.6	31
Chart 4: Crowd Density of 2.275	31
Chart 5: Crowd Density of 2.95	32
Chart 6: Crowd Density of 3.625	32
Chart 7: Crowd Density of 4.3	32
Chart 8: Crowd Density of 4.975	32

Chart 9: Crowd Density of 5.65	32
Chart 10: Crowd Density of 6.325	33
Chart 11: Crowd Density of 7	33
Appendix 3: VBA Model	34
Figure 1: Excel Chart resulting from a run with a low density distribution of spectators and placement of participants	34
Figure 2: Excel Chart shows high spectator density distribution in blue, participant distribution in pink, source with a black x, and sensors that report source detection with a red highlight.	34
Appendix 4: QGIS Representation	35
Appendix 5: Data Charts	36
Appendix 5.1: Base Case Analysis Charts.....	36
Chart 12: Percent of Time the Source is Seen for Base Case	36
Chart 13: Average Error When the Source is Seen for Base Case.....	36
Chart 14: Maximum Error When the Source is Seen for Base Case	37
Chart 15: Last Position Error for Base Case (All Densities)	37
Chart 16: Last Position Error for Base Case (Low Densities).....	38
Chart 17: Last Position Error for Base Case (High Densities)	38
Chart 18: Maximum Location Error for Base Case.....	39
Appendix 5.2: Negative Detection Analysis Charts.....	40
Chart 19: Percent of Time the Source is Seen for Negative Detections (5%) – Density of .925	40
Chart 20: Percent of Time the Source is Seen for Negative Detections (5%) – Density of 2.95	40
Chart 21: Percent of Time the Source is Seen for Negative Detections (25%) – Density of .25	41
Chart 22: Percent of Time the Source is Seen for Negative Detections (25%) – Density of .925	41
Chart 23: Percent of Time the Source is Seen for Negative Detections (25%) – Density of 2.95	42
Chart 24: Percent of Time the Source is Seen for Negative Detections (45%) – Density of .25	42
Chart 25: Percent of Time the Source is Seen for Negative Detections (45%) – Density of .925	43
Chart 26: Percent of Time the Source is Seen for Negative Detections (45%) – Density of 4.3	43
Chart 27: Percent of Time the Source is Seen for Negative Detections (45%) – Density of 4.975	44
Chart 28: Average Error When the Source is Seen for Negative Detections (45%) – Density of .925	44
Chart 29: Last Position Error for Negative Detections (45%) – Density of .25.....	45
Chart 30: Last Position Error for Negative Detections (45%) – Density of .925.....	45
Chart 31: Maximum Location Error for Negative Detections (45%) – Density of .25	46
Appendix 5.3: Random Source Movement Analysis Charts.....	47
Chart 32: Percent of Time the Source is Seen for Random Source Movement – Density of .925.....	47

Chart 33: Percent of Time the Source is Seen for Random Source Movement – Density of 2.275....	47
Chart 34: Maximum Error When the Source is Seen Random Source Movement – Density of .925.	48
Appendix 5.4: Crowd Movement Analysis Charts.....	49
Chart 35: Last Position Error for Crowd Movement – Density of .25	49
Chart 36: Last Position Error for Crowd Movement – Density of 2.5	49
Chart 37: Last Position Error for Crowd Movement – Density of 4.3	50
Appendix 5.5: Distributions of Efficiencies Charts.....	51
Chart 38: Percent of Time the Source is Seen for Distribution of Efficiencies – Density of .925.....	51
Chart 39: Percent of Time the Source is Seen for Distribution of Efficiencies – Density of 1.6.....	51
Chart 40: Percent of Time the Source is Seen for Distribution of Efficiencies – Density of 4.3.....	52
Chart 41: Average Error When Source is Seen for Distribution of Efficiencies – Density of .925.....	52
Chart 42: Average Error When Source is Seen for Distribution of Efficiencies – Density of 2.275.....	53
Chart 43: Average Error When Source is Seen for Distribution of Efficiencies – Density of 3.625.....	53
Chart 44: Average Error When Source is Seen for Distribution of Efficiencies – Density of 7.....	54
Appendix 5.6: False Detection Analysis Charts.....	55
Chart 45: Percent of Time Source is Seen for False Detections – Density of .25.....	55
Chart 46: Average Error When Source is Seen for False Detections – Density of .25	55
Chart 47: Percent of Time Source is Seen for False Detections – Density of 2.....	56
Chart 48: Average Error When Source is Seen for False Detections – Density of .25	56
Chart 49: Percent of Time Source is Seen for False Detections – Density of 4.....	57
Chart 50: Average Error When Source is Seen for False Detections – Density of 4	57
Appendix 6: Earned Value Management Chart	58
Chart 51: Earned Value Management.....	58
Appendix 7: “See Distance” Intervals for Data	59
Chart 52: “See Distance” Intervals for Data.....	59
Appendix 8:Gantt Chart	60
Chart 53: Gantt Chart (1)	60
Chart 54: Gantt Chart (2)	61
Appendix 9: Reference of Key Terms.....	62

Abstract

As time goes on, our world continues to become more networked, and more advanced sensors become part of the devices we use daily. Our cell phones, watches, and cars will soon be able to connect regularly with one another, creating a ubiquitous network of communication. This network can be used in a variety of ways. One such application that we will discuss in this paper is the detection and surveillance of unregulated Chemical, Biological, Radiological, Nuclear, and Explosive (CBRNE) materials. These detections are crucial for safeguarding the nation from weapons of mass destruction. Ubiquitous sensors can be used to detect some of these materials. By communicating detections in time and location, illicit materials can be localized, tracked, and ultimately interdicted.

We created a model of a ubiquitous sensor network using Excel Visual Basic and modeled a network of people carrying sensors at the Macy's Thanksgiving Day Parade in New York City. The parameters of this event are discussed in the Scope Section of this report. The model generates a crowd (each with their own sensor) and several groups of parade participants. It then generates a source (with a radioactive/nuclear type of emanation) that enters the simulation at a fixed point and moves either directly or in a random manor to a fixed drop off location. The sensor network detects the movement of the source and estimates its position. The model outputs various metrics that are used to analyze the relationship of the various ubiquitous sensor networks' parameters. We analyze the effects of randomness of source movement, crowd movement, sensor density, "see distance", various distributions for sensor efficiency, and negative and false detections. From our results we have compiled a framework of desired parameters for a ubiquitous (CB)RN(E) sensor network to aid in a (CB)RN(E) event interdiction.

This project is being sponsored by TASC, Inc. with the purpose of introducing a framework and analysis for the detection of CBRNE materials through ubiquitous sensor networks.

Intro

The Sponsor

TASC, Inc. is a provider of systems engineering resulting in scientific, engineering, and technical services to federal, state, and local government agencies, as well as the military. Our sponsor, TASC, Inc., based out of Lorton, Virginia, is working to solve many of the most pressing national security and public safety challenges facing our nation and the world. In terms of defense, TASC, Inc. plays a key role in supporting the protections which prevent illicit agents from entering the United States. TASC, Inc. has requested an investigation of parameters and the introduction of a framework for further analysis. This study will be continued by future teams and eventually company computational scientists within a classified setting once more advanced sensor technologies exist.

History of Security Concerns

The United States' susceptibility to a CBRNE attack has increased over the years due to prior terrorist intentions and events. From the 1990's, the Department of Defense acknowledged both the creation of a growing number of CBRNE weapons of mass destruction and their impact on the United States. Since the terrorist attacks of September 11, 2001 and the Anthrax incident of 2001, there has been a heightened concern that nuclear weapons and chemical and biological materials pose a grave, future threat to the citizens of the United States. The Department of Homeland Security (DHS) was created to safeguard the country against terrorism and respond to any future attacks (Creation of the Department of Homeland Security). In addition to smuggling hazardous materials into the country, there are several chemical agents that are easy to obtain and produce that can be used as weapons or explosives (Nelson, 2012). Since the fear of chemical and radioactive agents getting into the wrong hands is so concerning and because of the detrimental effects of a CBRNE event or attack, it is a goal of the United States to focus their defense and security-related research toward the detection of explosive and chemical weapons that can be used as weapons of mass destruction.

Detection of explosive or chemical devices can be difficult due to the way they are concealed. Often, they are placed in crowded areas or main areas of travel so that it is difficult to be detected within the environment (Stankovic, 2006). More recently, explosive devices have been placed in overcrowded areas of high value and high visibility. In the United States, there are many public and transportation venues that are susceptible to a CBRNE incident. Such venues consist of, but are not limited to airports, sporting events, concerts, races, and speeches. Some examples from recent decades consist of the Boston Marathon tragedy, the World Trade Center bombings, and the New York City subway plot (Chakraborty, 2013). Thus, it is extremely important to provide an appropriate sensor network framework and architecture to protect the people present at these and other similar events.

Wireless Sensor Networks

A wireless sensor network (WSN) is a collection of sensor nodes organized into a cooperative network. The sensors will communicate wirelessly and self-organize when deployed, whether in an ad-hoc manner or not (Stankovic, 2006). To support the TASC, Inc. initiative to protect the United States from a CBRNE attack in particular, our group has created a model to simulate a network of sensors that detects and tracks a nuclear/radioactive source prior to its release or detonation.

Ubiquitous refers to the idea that anyone, anywhere, at any time, can be a part of a network (Park, 2005). Ubiquitous sensor networks allow for CBRNE detectors to continuously monitor the presence of illicit materials in any venue in order to prevent an incident from occurring, and just as importantly, to provide the proper authorities a timely warning (Nelson, 2012). The use of these detectors can provide intelligence and greatly reduce the time to discovery of a CBRNE source, as well as the interdiction of a CBRNE incident. The instantaneous feedback on the conditions of a venue's environment and the localization of a source are critical for a more rapid and well-organized response. If successful, an effective WSN will enable the quick detection and localization of an illicit source (Nelson, 2012).

The characteristics of a WSN are important for its efficacy. Firstly, the network must consist of a large number of sensor nodes ranging from ten to ten thousand dispersed throughout a venue (Park, 2005). These nodes can be arranged in an ad-hoc scenario or can be part of a ubiquitous and mobile network of sensors. The density and architecture of a network greatly influences its success. The more nodes that can hone in on a CBRNE source, the greater the area of detection overlay. Second, it is important that the sensors have a way to communicate, process, and transmit data among one another in order to cooperatively summarize observations of their environmental conditions (Park, 2005). Without this communication and data fusion, a network is nothing more than thousands of individual nodes acting alone, with no common purpose to track, detect, or monitor uncertain and illicit materials. What happens with this data that is collected is just as important as communication among one other. A network can have a gateway or a sink node that connects a network of sensors to other networks and the outside world. This gateway is a central location that combines the collection of information to be analyzed (Park, 2005). Lastly, a ubiquitous sensor network is useless unless the information is dealt with appropriately. Thus, the proper authorities must be alerted of any CBRNE incident as soon as the data is analyzed. In addition to the way the networks work together, their individual characteristics are significant in CBRNE source detection as well. In the case of detecting and preventing a CBRNE incident through a ubiquitous network, sensor nodes will consist of small, low cost, low-power devices that can communicate within short distances, sense environmental data, and perform minor data processing. The reduced size and reduced cost of these advanced sensors will not only be required but crucial for an omnipresent "surveillance." The reduced size would allow for the sensors to be used in numerous locations and allow them to be easily portable. The reduced cost would allow for the sensors to be used in a wide variety of devices without an increase in cost thus allowing for higher density networks.

One such sensor platform that is currently portable and used daily is the cell phone. Most cell phones have both a microphone and a camera, and can be connected to other phones via Bluetooth. Cell phones are actually part of their own network and have predictable power supplies and life expectancies. What makes them a significant entity within a large-scale ubiquitous network is the large number that exist and the vast spatial territory they encompass (Kansal, Goraczko, & Zhao, 2007). In the event that a CBRNE source is present at a public venue, cell phones that are equipped with the proper detectors can work together to collect and store location and detection data, which could be sent to a universal storage cloud. The network can localize the hazardous source and have authorities alerted. It is extremely likely that cell phones, combined with other electronic devices such as cars, watches, and cameras, can be used in unison to work together for the detection of harmful agents in the goal to protect the United States from CBRNE incidents in the near future.

Assumptions

For our scope, the team has decided to rule out topics that are not relevant to the specific goal of this project, which is to analyze the relationships between variables and metrics of wireless sensor networks. One significant assumption is that sensor implementation is feasible. The team will assume that these sensors could be implemented in order to analyze relationships between parameters of a sensor network that localizes a (CB)RN(E) source. This assumption also implies that the technology for sensors small enough to fit in a cell phone will be capable of detecting CBRNE materials in the future. This is not an unrealistic assumption. In fact, much work has already been done in this area (Cell-All: Super Smartphones Sniff Out Suspicious Substances) (Sutter, 2010).

The team assumes agnostic sensors; that is, the sensor in the network is not a specific make or model. The model assumes that there is one sensor per person. For the purpose of studying parametric relationships, the sensors' strength will remain constant for a single run or single data set of runs; varying from one set of data to the next throughout our study. In addition, the team will assume a single, agnostic source, classified by either of the following: radioactive or nuclear.

Due to the time available, we have decided that the model will only generate and track a single source. We created a preliminary algorithm to track multiple targets and gain better insight into its effects. It was concluded that multiple target tracking would add excessive time to the model creation and limit the time allotted for analysis. Simulating a single source provided the necessary data for a good understanding and analysis of the network parameters. It is to be noted that if the model generated more than one source, then multiple sources could be detected and properly tracked. However, to simulate more sources additional coding would have been required and thus exponentially expanding the required code. The biggest effect on the current model, resulting from not being able to track multiple targets, is the increased position errors when false detections are enabled. Finally, in the model, a generic source emanation function is used, which can be adjusted accordingly in order to allow a future team to conduct follow up analytics in anticipation of a particular source.

Detections by sensors will be modeled as a binary action. The source emanation will be consistent with the inverse square law. For our simulation, this emanation level will not have noise associated with it. There will be a threshold which will set off a sensor. If, at a particular range, the emanation is lower than the threshold a detection will not occur. If, at a particular range, the emanation is equal to or greater, then a detection will occur. The sensors will reset and show no detection once the emanation level decreases below the threshold level.

The variability of the source movement through the network has been limited. The source will enter the simulation area near the green square, as indicated in the diagram below, which will be fixed for all replications. It is a logical assumption for the source to enter into the area from a side street to start his movement through the simulation, as opposed to walking across the parade for example. The source will move through the simulation area either directly or with random movement, which will be discussed later. The simulation will end when the source reaches the drop off location indicated by the red square in the diagram below. We assume the source will be dropped off near the front of the sidewalk in order to cause significant damage. There are more people at the front of the sidewalk, and in addition, parade participants will be possible targets. This drop off location is fixed for all replications.

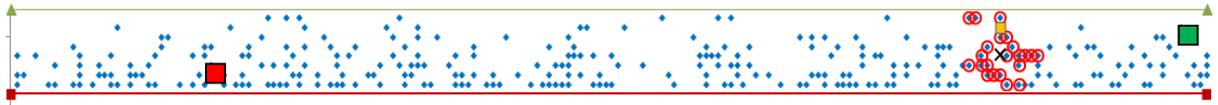


Diagram 1 - Source Path

Bandwidth and transmission speeds of the sensor network will be assumed adequate for the data requirements of the sensor networks in this study.

The team will not investigate the effects of weather on sensor detection or source emanation.

We will not be addressing the legalities of personal privacy.

The Model

Our team has created a model to simulate a sensor network in Excel Visual Basic that detects and localizes a single radioactive/nuclear source moving through an area. We created our model from scratch with input from Allen Harvey at TASC, Inc. The model has three main variables: density, source strength, and sensor strength. The model is initially run with a base-case to analyze the relationships between these variables. Excursions are then run to compile data on the efficiency of the sensor network. The model is broken down into sensor creation, source movement, sensor localization, and error calculation.

Crowd Density

At major public events such as parades, crowds of all forms and sizes can be seen depending on time and location. The model can step through these densities in as many steps as the user desires. In all of our simulation runs, the model steps from .25 people/m² to 7 people/m² by an interval of .625 people/m². The lower bound, .25 people/m², was chosen as a small enough density to emulate a very early point in the day of the parade. The team chose 7 people/m² as the maximum observed density to capture a very high density. See Appendix 2 for a visual example of each density.

Source and Sensor Strength Measurement

In this study, source and sensor strength measurement do not have units assigned to them. It is unknown what level of radiation will emanate from a source. There can be a large variance in radiological emanation. For example, an unexposed pound of uranium-238 would emanate .00015 curies, whereas a pound of cobalt-60 would emanate 518,000 curies. (Measurement: Activity: How Much Is Present?) Additionally there are different ways in which radiation can be measured. For example, a Geiger Counter can read out in roentgens per hour, milliroentgens per hour, rem per hour, millirem per hour, and counts per minute; another sensor such as a MicroR Meter can read out in microroentgens per hour and/or counts per minute (Radiation Basics: How Can You Detect Radiation?). There are other detectors as well. For our study, we will use sensing strength units and source strength units, which is simply a level of sensor or source strength.

Source Strength

The model can step through as many source strengths as the user desires. In all of our simulation runs, the model goes through 7 steps from .5 source strength units to 7 source strength units.

Sensor Strength

The model can step through as many sensor strengths as the user desires. In all of our simulation runs, the model goes through 7 steps from .5 sensing strength units to 7 sensing strength units. The lower bound of the sensor strength is based on the upper bound of the source strength. The lower bound of sensor strength is equal to $1/(\text{upper bound of source strength})$. Using this ratio in our simulation runs, the weakest sensor is able to detect the strongest source of at least a distance of one grid unit. If the lower bound of the sensor strength is stronger than this, data points would be excluded. If the lower bound of the sensor strength is weaker than this, then no sensor will be able to see the strongest source. The upper bound of sensor strength should be strong enough to detect the weakest sensor: $1/(\text{upper bound of sensor strength})$, which will at most be the lower bound of the source strength. The upper bounds for source and sensor strength are chosen to give a max “see distance” of 7 grid units (“see distance” will be discussed in the next section). The upper bound of 7 grid units is chosen because this is approximately the width of the sidewalk from the edge of the street to the side of the building, so this ensures nearly everyone along this dimension will see the source.

See Distance

“See distance” is the relationship between source strength and sensor strength. A nuclear/radiological emanation is determined by the inverse square law (Nave, 2014). The relationship is given by:

$$\text{See Distance} = \sqrt{\text{sensor efficiency} * \text{sensor strength} * \text{source strength}}$$

Therefore, for a detection to occur, the distance from source to sensor must be less than or equal to this value. In addition, the “see distance” of a sensor must be at least 1 in order for a detection to occur. The reason for this is that we are using $1/R^2$ for the decay rate of the source. If the “see distance” is less than 1 grid unit, then the decay would be greater than the initial source strength. The derivation of this is provided in the Appendix 1.

Sensor Creation

The first part of the model generates sensors on a grid that overlay the area of focus within the Macy’s Thanksgiving Day Parade. There are two sensor groups created; the crowd and the parade participants. These groups can be seen in Figure 1 in Appendix 3. The top portion represents the crowd and the bottom portion represents the parade participants. These sensors are assumed to be located on an electronic device carried by either a person at the parade or a person in the crowd; each person has exactly one sensor.

The Grid

The grid is determined by the parade dimensions and the maximum crowd density; both of which are predetermined. The average person has a minimum space requirement of .085 square meters (Oberhagemann, 2012). This space requirement gives a density of 11.8 people/m². However, taking into account clothing and actual observed densities, maximum crowd density is between seven to nine people/m² (Oberhagemann, 2012). As a result, the team has decided that 8 people/m² is an appropriate value for maximum density. Therefore, each person must be at least $1/\sqrt{8}$ meters from another person. This value is used as the unit of length in our grid (one grid unit = $1/\sqrt{8}$ meters). As for parade dimensions, our grid takes into account the outline of the parade. The vertices of the rectangular sidewalk and street area used in our model have latitude and longitude as follows:

Sidewalk Grid Area		Street Grid Area	
40.77816° N	73.97472° W	40.77807° N	73.97449° W
40.77815° N	73.97469° W	40.77815° N	73.97469° W
40.77763° N	73.97511° W	40.77762° N	73.97508° W
40.77762° N	73.97508° W	40.77754° N	73.97486° W

The parade starts at 77th Avenue, but viewing is not permitted until 75th Avenue, and only on the west side of the street (Macy’s Thanksgiving Day Parade, 2011). The dimensions above cover the west side sidewalk along Central Park West from 75th Avenue to 74th Avenue.

The Parade Spectators

Parades are made up of packs of people. The way people arrive and form these crowds is not an exact science. Crowds rarely pack in regular formation, but rather with some random distribution (Still, 2014). In an attempt to portray the spectators of the parade, they are given randomly generated grid coordinates. The coordinates of each crowd member is then stored to prevent more than one person from being in the same grid cell. The x-coordinate is drawn from a discrete uniform distribution from 1 to 191 with 191 being the length of the simulated area in grid units. The team has chosen a uniform distribution because we have assumed it is equally probable for a spectator to line the parade route at each location. The y-coordinate is drawn from a geometric distribution with parameter $p = .3$. We assume that in a parade setting a person has an inclination to be closer to the parade. A geometric distribution with $p=.3$ appropriately distributes the spectators so that a majority of the spectators line up in the front rows of the crowd. In the model, as an input to our base case sensor network, the parade spectators are stationary. As an excursion, however, a crowd movement option can be selected. If crowd movement is selected, the model allows for the input of the uniform probability that each crowd member moves. This chance ranges from 0% (no chance for movement) to 100% (crowd member must move). The spectators will move in a random direction to one of the eight adjacent grid locations but will never share the same position as another spectator on the grid; for example, one spectator could move one block away as another one enters its previous location but if all adjacent grid locations are occupied, the spectator would not move and remain in the same location.

The Parade Participants

The parade participants are mobile and are not randomly generated; they have predetermined coordinates. The participants form three separate entities: a band, a balloon, and a float. These three entities will continually loop through the area of focus until the end of each run.

Source Movement

The model moves a source through the parade grid from a designated entrance point to the point where the source is abandoned. The source has a generic walking speed, which is modeled as a series of discrete events. In the base-case, the source moves through the parade grid in a shortest path (least amount of steps), moving directly from the designated entry point to the source abandonment point. As an excursion, the model can be altered by enabling the “Random Walk” mode, which adds another stochastic input to the model. When enabled, for each step the source moves, the source has a chance to randomly move from its current location to any grid location one cell away or remain in the same spot. The randomness of the source movement is variable and ranges from 0% (no deviation from the

shortest path) to 100% (completely random movement, in which case the simulation will only end if the source happens to randomly make it to the drop off location).

Source Localization

Source Emanation and Sensor Detection

The source and each sensor will have a strength associated with it. The maximum distance that a sensor can detect a source is calculated using the Inverse Square Law from these values (Nave, 2014). The team refers to this relationship as the “see distance,” which is the radius around the source.

Location Estimate

As the source moves through the parade area, sensor detection turns “on” and “off” as the source radius overlaps the sensor or distances itself from the sensor, respectively. A false positive and/or negative parameter is used to add realism to the model.

For each step the source takes towards the drop off location, the model will determine an estimated position for the source. Using the source strength and sensor strength, the model will determine which sensor should have detections. If there are no negative detections, all detections will be added to the detection list. If negative detections are on, then for each detection, the model will generate a random uniform number and compare it against the negative recur rate for that sensor. If a negative detection is generated, then that detection is not added to the detection list. Next, if false detections are enabled, the model will determine if any sensor(s) generate false detections which is based on their false detection recur rate in a similar manner to negative detections. If a false detection occurs, it will be added to the detection list. When false detections occur, the model uses additional parameters, called attenuation parameters, to eliminate these false detections (and possibly true detections). These parameters include a range and number. The attenuation range gives a radius (in terms of grid units) in which to communicate with other sensors who may also detect the source. The attenuation number gives the number of sensors required to detect a source within the attenuation range for that sensor to be a true detection. An example of attenuation range and number can be seen in the diagram below.

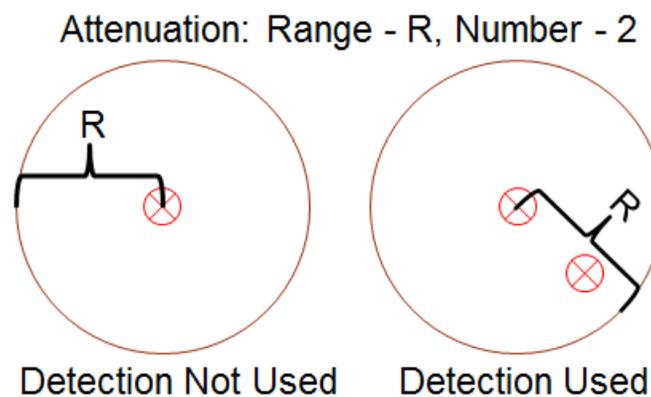


Diagram 2 – Attenuation Parameters

Each detection in the list of detections is checked for these parameters. If a sensor has a detection and the required number of detections are present in that range, then a detection remains on the detection

list and is used for position estimation. If the number of detections is not met, the detection is eliminated. While this can eliminate false detections, it also has the chance to eliminate true detections when densities are low.

The model calculates an estimation of the source location by calculating the center of mass of all detections that remain on the detection list. The equations used are as follows:

$$Est. x = \frac{1}{\sum_{i=1}^n \left(\frac{100}{Eff[i] * Str[i]} \right)} \sum_{i=1}^n x_i \left(\frac{100}{Eff[i] * Str[i]} \right)$$

$$Est. y = \frac{1}{\sum_{i=1}^n \left(\frac{100}{Eff[i] * Str[i]} \right)} \sum_{i=1}^n y_i \left(\frac{100}{Eff[i] * Str[i]} \right)$$

Where:

Eff[i]=efficiency of sensor I;

Str[i]=strength of sensor I;

n=number of sensors in detection list; and

x_i, y_i are the x and y positions of the i^{th} sensor respectively.

The model allows the following variables to be altered before each run: crowd density, source strength, sensor strength, minimum sensor efficiency, chance of false detection and negative detection, and an attenuation distance and factor. In Appendix 3, Figure 2 shows a source being detected as a source moves through the parade area. As previously stated, randomness of source movement, option for crowd movement, and parade movement can be adjusted as well.

Outputs

The model outputs consist of the error of actual source location and estimated source location, which includes the calculation of average error, minimum error, maximum error, the error of the distance from the source drop off location to the last estimated position when the source is seen, and the error of the distance from the last position the source is seen to the next time the source is seen. Other results calculated consist of total steps, total time, percentage of time that the source is seen within the network, and average number of sensors used to localize the source.

The model also outputs the locations of the crowd spectators, parade participants, source location, source estimate, false detections, and detecting sensors at each step that the source moves. The number of outputs can be altered as needed. The outputs are in Cartesian coordinate notation, but can be converted to a latitude and longitude to create a visual of the model.

GIS Model

The GIS model is a visualization of the team's localization results. It aids viewers in understanding how the model works. The model shows the parade spectators on the west-side sidewalk from the corner of Central Park West and 75th avenue to the corner of Central Park West and 74th avenue, as well as the

parade participants in the street from the corner of Central Park West and 75th avenue to the corner of Central Park West and 74th avenue. The localization model writes an output of sensor detections, source position, and source estimation that is plotted over a Google Maps overlay within Quantum GIS. These areas are shown in Appendix 4, Figure 3. Using the Quantum GIS experimental plug-in, Time Manager, the model visualizes and mobilizes the location of the above outputs over 30 steps (or seconds) of time, creating the effect that the source, source detects, and parade are mobile. The model is set to estimate localization at the Macy's Thanksgiving Day Parade around 10:04 AM, over 30 seconds, assuming that one step is tantamount to one second of time. From this model, the team created a video of this time period for visual understanding of the VBA model and base-case parameters.

Model Verification

The model has been verified. Due to the fact that the model creates visual outputs, the team was able to ensure that the model was working the way it was intended. Through debugging the model and watching visual runs, it was verified that the model is working and calculating what it is supposed to calculate.

Analysis and Results

Metric Relationships

In order for the team to analyze the relationships between the input variables of the sensor network model, the team graphed and did minor statistical analysis of the outputs of the following metrics:

1. Percent of Time Source is Seen
2. Average Error When Source is Seen
3. Maximum Error When Source is Seen
4. Last Position Error
5. Maximum Location Error

Percent of Time Source is Seen

It is difficult for one to deem a particular percentage of time the source is seen as good or bad. Instead, the team decided to analyze the overall trend of this metric in relation to different input variables of the model. The percent of time the source is seen is described as the quotient of the number of steps a source is seen and the total steps in the model.

Average Error when Source is Seen

The average error when the source is seen is described as the average of the distance from the estimated source position to the actual position of the source. The team analyzed the overall trend of this metric in relation to different input variables of the model.

Maximum Error when Source is Seen

The maximum error when the source is seen is described as the maximum of all the distances from the estimated source position to the actual position of the source. The team analyzed the overall trend of this metric in relation to different input variables of the model.

Last Position Error

The last position error is described as the difference of the last estimated source position and the source abandon location. The team analyzed the overall trend of this metric in relation to different input variables of the model.

Maximum Location Error

The maximum location error is described as the distance the source is able to move away from the point that the network last detected it. This metric closely relates to the Percent of Source Seen. The team analyzed the overall trend of this metric in relation to different input variables of the model.

Base Case Analysis

After running the base case, it is clear that “see distance” has a significant effect on the percent of time the source is seen, the maximum and average position error, and the last position error.

For the base case, each data point is an average of each metric over 200 simulation runs for a specific “see distance” and density.

Percent Seen

Refer to Chart 12 in Appendix 5.1. Holding see distance constant, as density increases, the percent of time the source is seen increases, but at a decreasing rate. This relationship is one that was expected because the more sensors one adds to the network the higher chance the source will be seen.

Holding density constant, as “see distance” increases, the percent of time the source is seen increases, but at a decreasing rate. Each density curve concaves downward and is equivalent to one at some “see distance” and greater. The reason for this concavity for a given density is as “see distance” increases, the “see distance” of each sensor overlaps with the “see distance” of the other sensors. For example, imagine two sensors and a given source location between them. At a low “see distance,” neither sensor detects the source. At a higher “see distance,” one sensor can see it, but the other cannot, so there is a detection. At an even higher “see distance,” both sensors can now see the source; but since there is overlap in their “see distances,” there is no increase in percent seen because a sensor closer to the source already sees it. Therefore, an increase in “see distance” becomes less important as it gets higher. The same type of argument holds for why when holding “see distance” constant, as density increases, the percent of the time the source is seen increases as well, but at a decreasing rate. When adding more sensors to the network (increase in density), there again will be overlap. So a sensor that was just added may not contribute anything if its “see distance” is overlapped completely by other sensors in the network. An interesting observation is that all networks with a “see distance” of 3 grid units or greater and density greater than or equal to .925 people/m² can see the source at least 90% of the time.

Average Error When Seen

Refer to Chart 13 in Appendix 5.1. Holding see distance constant, as density increases, the average error when the source is seen decreases. As more sensors are added, more detections are likely at each step; therefore, the source is pinpointed more accurately.

Holding density constant, as see distance increases, the average error when the source is seen increases. Increasing the “see distance” increases the area of detection which allows for more possible error. For example, a single sensor with a “see distance” of one also has a possible maximum error (when the

source is detected) of one. Increasing the “see distance” to ten would increase the maximum error to ten.

As “see distance” increases to the point of approaching the length of the crowded region, the source estimate becomes less useful for that dimension. For example, at a “see distance” of 7 grid units (about the length of the grid in the y direction as seen below) the y-coordinate of the estimate is no longer beneficial because it will only change due to the position of the sensors rather than the position of the source. As seen below, sensors activated from the building to the street.

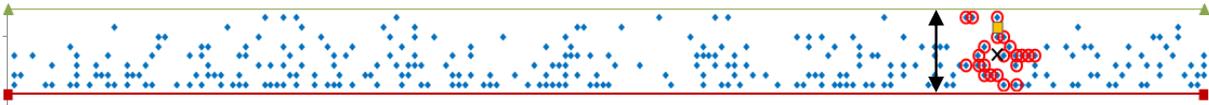


Diagram 3 – Vertical Estimate of Source

As the source moves up and down there are no more sensors that can turn “on” or “off” to give a better y estimate. Alternatively, the x-coordinate still provides a good estimate for the fact that as the source moves from side to side sensors turn “on” and “off,” moving the estimate based on the path of the source. All densities greater than or equal to $.925 \text{ people/m}^2$ have an average error when seen less than 2.14 grid units. Note that all the density curves except for a density of $.25 \text{ people/m}^2$ are nearly identical in shape (with the exception of a vertical shift of values).

Maximum Error When Seen

Refer to Chart 14 in Appendix 5.1. Similar results are experienced with maximum error as with average error. The only significant difference is that the maximum error has a higher magnitude of results.

Last Position Error

Refer to Charts 15 through 17 in Appendix 5.1. For a density of $.25 \text{ people/m}^2$, as “see distance” increases, the last position error decreases. This density has the worst (highest) last position error regardless of “see distance.” There are too few sensors, and so, any increase in “see distance” increases the chance that a sensor will see the source closer to its end point. The following relationships discussed exclude a density of $.25 \text{ people/m}^2$.

In general, from a “see distance” of 1 to 2 grid units, as density increases, last position error decreases. Refer to Chart 17 in Appendix 5.1. Densities of 4.975 people/m^2 , 5.65 people/m^2 , 6.325 people/m^2 , and 7 people/m^2 start out with the lowest last position error and end up with the highest as “see distance” increases. Densities of 2.95 people/m^2 and higher start out lower and then increase, with the higher densities increasing at a quicker rate therefore having a higher last position error at a “see distance” of 7 grid units.

Refer to Chart 16 in Appendix 5.1. Lastly, densities of 2.275 people/m^2 and below start out high, then decrease until about a “see distance” of 3 grid units, and finally start increasing again either immediately or around a “see distance” of 4 grid units. The lower densities have the higher last position error regardless of the “see distance.” The higher densities seem to increase quicker, but never get to the point where they go above a lower density, which is the case for densities higher than 2.275 people/m^2 . Densities of 2.275 people/m^2 and below start decreasing since increasing “see distance” increases the chance of having a sensor closer to the end point seeing the source. Another effect of increasing “see distance” is that it increases the number of sensors involved in detecting the source. This higher density

decreases detection error. However, as soon as see distance gets to about 3 grid units, this distance is now large enough to offset the increase in sensor density near the source drop off point. More sensors in the middle and back are pulling the estimated position from the front where the source really is. Therefore, error starts to increase. Questions may arise as to why this trend of decreasing then increasing does not occur for average error when seen at low densities. The reason lies in the fact that error is higher at the back of the viewing area (fewer sensors) than at the front of the viewing area (more sensors). At the back, an increase in “see distance” is less likely to increase the number of sensors detecting the source, thereby only exacerbating the error.

Maximum Location Error

Refer to Chart 18 in Appendix 5.1. Holding “see distance” constant, as density increases, maximum location error decreases. Meanwhile, holding density constant, we see a decrease in maximum location error and then an increase (for lower densities this change occurs at a higher “see distance”). The reason for the decrease is because when increasing “see distance” the source is seen more often, and so it can’t step as far without being seen. The reason for this increase is that when you get to a certain “see distance,” the network is seeing the source most of the time. Therefore, the curve will start to look like maximum error when source seen.

Significant findings: Base Case

In terms of percent of time the source is seen, it is preferable for a network to have higher densities with higher “see distances.” However, percent seen becomes marginally important at some “see distance”, for each density. In terms of average error when the source is seen it is preferable to have higher densities with lower see distances. In terms of last position error, it is preferable for a network to have higher densities and low “see distances” or medium to low densities with lower “see distances.” In terms of maximum location error it is preferable for a network to have higher densities and lower “see distances.” Overall, for each density it is necessary to have a certain “see distance” to obtain a good percent of time the source is seen. Once this “see distance” is obtained, increasing it more will only increase error.

Excursion Variables

In order to analyze the results of our sensor network model, the team explored the relationships between the model’s variables by altering the ones deemed to be significant. Once the team determined the feasible region of sensor networks and ran the base-case of the model varying crowd density, sensor and source strength, the team conducted additional excursions for the following variables:

1. 50% Random Walk
2. Negative Detections
3. False Detections
4. Crowd Movement
5. Varying Distributions for Sensor Efficiencies

The model does 200 runs for each excursion and records the outputs as data and visual graphs for analysis (with the exception of Random Source Movement with a source strength, sensor strength, and density of 7 people/m², where 50 simulations were run). Each excursion was run with the same densities, source strengths, and sensor, strengths unless otherwise noted.

It is important to take note that when nothing is seen, the model will not average a zero value for a metric like maximum error. The difference between the base case and excursion along with the half widths were used to compare the excursion to the base case.

$$Halfwidth = t_{95\%,n-1} \sqrt{\frac{Standard\ Deviation^2}{n}}$$

If the difference was large enough to affect the network’s performance, a quick investigation as to whether it was statistically significant was conducted. Since some intervals have more than one data point, the average of the difference in each interval is taken and then all these values are averaged to obtain the overall average for the density

Negative Detections

A negative detection is defined as a sensor’s failure to see the source when it should see it. Keeping the base-case variables as is, the model is set so that 100% of the sensors have a chance to false detect. For each sensor, three different probabilities were run giving each sensor a chance of giving a negative detection. The three probabilities run were 5%, 25%, and 45%.

Overall, by including negative detections, there is no change in trends, only change in magnitudes. In other words, the curves have similar shapes to the base case.

Percent Seen

Refer to Charts 19 through 27 in Appendix 5.2. By including negative detections, the percent seen is lower than the base case. With 5% chance of negative detection, the average difference in the percent seen between the excursion and base case for a density of .25 people/m² is about 1.4%, and the difference gets smaller as density increases. Therefore, there is no significant effect on the percent of time seen for 5% (see Charts 19 and 20 in Appendix 5.2).

Refer to Charts 21 through 23 in Appendix 5.2. With a 25% chance of negative detections, the average difference of percent seen for a density of .25 people/m² is about 7.83% and gets smaller as density increases. A large part of this difference occurs at lower “see distances.” Therefore, at low densities or low “see distances,” there is a moderate effect (11% difference is the absolute largest that occurs at a density of .25 people/m²). The average difference from 1 grid unit to 1.1412 grid units remains at about 10% for densities up to 2.95 people/m². However, as density increases, the difference at lower “see distances” becomes less pronounced. All networks with a “see distance” of 3 grid units or greater and density greater than or equal to .925 people/m² can see the source at least 84% of the time.

With a 45% chance of negative detection, the average percent seen difference for a density of .25 people/m² is about 16.6% (see Chart 24 in Appendix 5.2), and the difference gets smaller as density increases. For densities greater than .25 people/m², a large part of this difference occurs at lower “see distances.” At a density of .925 people/m², for see distances between 1 grid units and 4 grid units, the average is 15% with a maximum 19.9% and minimum at 9.6% (see Chart 25 in Appendix 5.2). As density increases this difference lessens. However, it is still fairly substantial. At a density of 4.3 people/m² and “see distance” between 1 unit and 2 units, the average difference is about 12% (see Chart 26 in Appendix 5.2). At a density of 4.975 people/m² and “see distance” above 2 grid units, the differences are less than 2% (see Chart 27 in Appendix 5.2). All networks with a “see distance” of 3.6 grid units or greater and density greater than or equal to .925 people/m² can see the source at least 87% of the time.

Average Error When Seen

Including negative detections generally will result in higher average error in estimated position because there are fewer sensors to pinpoint the source. With a 5% chance, the results are nearly identical to the base case. The average difference from the base case is .014 grid units. With a 25% and a 45% chance, the average difference is .074 grid units and .17 grid units respectively. At low densities, higher “see distances” have larger differences. As density gets higher, the difference becomes more pronounced at lower “see distances” (especially between 1 grid units and 2 grid units). However, the largest difference (high density, low “see distance”) is about .18 grid units (25%) and .32 grid units (45%). These differences are small enough to not have a major effect on the network. At a 45% chance all densities greater than or equal to .925 people/m² have an average error when seen less than 2.5 grid units (see Chart 28 in Appendix 5.2).

Maximum Error

Including negative detects generally results in a higher maximum error in estimated position than the base case. With a 5% chance, the results are nearly identical to the base case. With a 25% and a 45% chance, the average difference is .19 grid units and .38 grid units respectively. At low densities, higher “see distances” have larger differences, and as density increases, the difference becomes more pronounced at lower “see distances.” However, the largest difference (high density, low “see distance”) is about .4 grid units (25%) and .76 grid units (45%). These differences are small enough to not have a major effect on the network.

Last Position Error

The following numbers were taken from a 45% chance which shows that a smaller chance will be better. All differences are less than 3.7 grid units for a density of .25 people/m² with the overall average being .89 grid units, and less than .7 grid units for a density of .925 people/m² (see Charts 29 and 30 in Appendix 5.2). As density increases the overall difference lessens.

Maximum Location Error

The following numbers were taken from a 45% chance which shows that a smaller chance will be better. All differences are less than 7.24 grid units for a density of .25 people/m² with the overall average being 1.98 grid units (see Chart 31 in Appendix 5.2). As density increases the overall difference lessens.

Significant Findings: Negative Detections

From our analysis, we have concluded that negative detections take a density and lower it a certain degree. At 5% there is no significant effect regardless of density and “see distance.” At higher probabilities the only major effect negative detection has on the network is with percent seen and maximum location error, and it is only at low densities and/or low “see distances” where this effect may have cause for concern.

Random Source Movement

Keeping the base-case variables as is, the random source movement variable is set to 50%. This means the source has a 50% chance that in any given step, that the source will move one cell in a random direction and possibly deviate from the shortest path from the source starting point to the source abandon point.

Overall, by including random source movement, there is no change in trends, only change in magnitudes. In other words, the curves have similar shapes to the base case.

Percent Seen

When analyzing the percent time that a source is seen, we found the results to be nearly identical to the base case. For a majority of the see distances, the difference is negligible (see Charts 32 and 33 in Appendix 5.3).

Average Error

The average error when source is seen shows the resulting curves to have a close resemblance in shape to the base case. The only differences are that the random source movement generally gives a higher average error. As density increases, the difference becomes more pronounced and the higher the “see distance,” the more pronounced the difference. The results are nearly identical to base case at a density of 2.275 people/m² and lower (average errors are less than .05 grid units). The reasoning for this is that at higher densities and “see distances,” nearly every sensor from back to front can see the source. Therefore, the estimated source position will be in the vertical middle of the grid regardless of the source movement. With random source movement, the source is more able and likely to go to the extremes of the grid (very top or very bottom) thereby creating a larger average error. However, this larger error is not substantial with a density of 7 people/m² having the largest overall average difference of error of .178 grid units.

Maximum Error

Maximum error when the source is seen shows similar results to the base case. As density increases, the difference becomes more pronounced and the difference starts to appear at lower “see distances.” The maximum error occurs when the source is “outside” the majority of sensors that detect it. With lower densities, this can occur horizontally and vertically in the grid. As densities increase, this will generally occur at the top of the grid because fewer sensors are present. This could also occur at the bottom of the grid when both density and “see distance” are high. When random source movement is added, there is more chance that the source is traveling at the very top of the grid. This allows the source to be at the extremes of the detection ranges, which will result in a larger maximum error. Without the random source movement, the source starts at the line below the very top and never reaches the very top or the very bottom. The largest overall average difference is .38 grid units for a density of 7 people/m² (see Chart 34 in Appendix 5.3).

Last Position Error

The last position error shares the same trend as the base case, making it nearly identical for this metric as well. The reason for this is that since it measures last position, it should be almost exactly the same, especially at higher densities.

Maximum Location Error

Maximum location error is a little higher with random source movement, but still follows the same trend as the base case. This reasoning is similar to the argument about maximum error. The source is more likely to roam around sensor-less spaces or be at the very top of the grid where there are no/fewer sensors, thereby increasing the chance for a higher maximum location error.

Significant Findings: Random Source Movement

Adding in random source movement does little to nothing to affect the network.

Crowd Movement

In reality, people do not stand absolutely still. For the crowd movement excursion, keeping the base-case variables as is, the model is set so that each parade spectator has a 50% chance of moving. This means that each step for the simulation, each crowd member has a 50% chance of moving in a random unoccupied adjacent grid location. If there are no unoccupied spaces the crowd member will remain in their current location. The excursion is run with three densities: .25 people/m², 2.5 people/m², and 4.3 people/m². We assumed that at the highest of densities, crowd movement would not matter as much.

It appears that over time the crowd becomes more uniform over the vertical axis as opposed to geometric.

Percent Seen

The percent of time a source is seen shows a similar trend to that of the base case. For a density of .25 people/m², results are nearly identical to the base case. For a density of 2.5 people/m² and “see distances” between 1 and 2 grid units, the average difference is 3.7%. For a density of 4.3 people/m² and “see distances” between 1 and 2 grid units, there is an average difference 1.4%. As “see distance” increases, the difference decreases as well.

Average Error

Crowd movement also shares a similar trend to the base case in terms of average error when a source is seen. The only difference is that crowd movement results in lower average errors. This is believed to be a function of the crowd density change that occurs over time. The density from the front of the sidewalk is denser than the back of the sidewalk. The original geometric distribution of crowd shifts towards a more uniform distribution and this seems to account for the lower average errors. For a density of .25 people/m², results are nearly identical to the base case. The maximum difference occurs at higher “see distances” where the overall average difference is .04 grid units. For a density of 2.5 people/m², the average difference is .12 grid units and for a density of 4.3 people/m², the average difference is .1 grid units.

Maximum Error

The maximum error when a source is seen has similar results to the base case as well, although crowd movement results in lower maximum errors when seen. For a density of .25 people/m², results are nearly identical to the base case (the average difference is .0032 grid units). For a density of 2.5 people/m², the overall average difference is .1374 grid units, and the maximum difference is .2511 grid units. For a density of 4.3 people/m², the overall average difference is .089 grid units, and the maximum difference is .17 grid units.

Last Position Error

Crowd movement has higher last position error than the base case, but follows the same trend. This makes sense because the vertical is now more uniform as opposed to geometric and thus fewer people than before are in front of the crowd which is where the source ends. For a density of .25 people/m², the overall average difference is 1.5 grid units and the maximum difference is 5.4 grid units (R=1) (refer to Chart 35 in Appendix 5.4). For a density of 2.5 people/m², the overall average difference is .552 grid units and the maximum difference is 1.03 grid units (R=7) (refer to Chart 36 in Appendix 5.4). For a density of 4.3 people/m², the overall average difference is .4236 grid units and the maximum difference is .8383 grid units (R=7) (refer to Chart 37 in Appendix 5.4).

Maximum Location Error

Crowd movement has smaller maximum location error as compared to the base case. For a density of $.25 \text{ people/m}^2$, the overall average difference is 1.9 grid units and the maximum difference is 7.96 grid units ($R=1.71$). This difference seems large, but at the “see distance” for which this occurs, both crowd movement and base case have very large maximum location errors which is considered poor. For a density of 2.5 people/m^2 , the overall average difference is $.4597$ grid units and the maximum difference is 2.3 grid units ($R=1.71$). For a density of 4.3 people/m^2 , the overall average difference is $.1645$ grid units and the maximum is $.7866$ grid units ($R=1.71$).

Significant Findings: Crowd Movement

Crowd movement changes the geometric distribution of the y-axis into a uniform distribution over time. This distribution change in the crowd has a small effect on the network’s performance. There is a very minor effect on percent seen at the lowest “see distances” for a density of 2.5 people/m^2 . Also, there is a moderate effect on last position error for the density of $.25 \text{ people/m}^2$.

Triangular Vs. Uniform Vs. Base case

The minimum sensor efficiency percent is set to 0, and sensors get assigned efficiencies based on a distribution. The two distributions run were Uniform[0, 100] and Triangular with mean 75.

Percent Seen

The uniform distribution has the lowest percent seen of all three with the base case being the highest. A “see distance” in the interval [1,1.33] grid units regardless of density has a large average difference between the base case and the distributions due to the fact that nearly no sensor has a high enough “see distance” to detect the source. Excluding this “see distance” interval, as density increases, the average differences decrease and follow the same trend as the base case, just at a lower magnitude. At a density of 1.6 people/m^2 and “see distance” of 2.5 grid units, the uniform and triangular distribution networks are seeing the source about 80% of the time (refer to Charts 38-40 in Appendix 5.5).

Average Error

See Charts 41-44 in Appendix 5.5. The uniform and triangular distributions initially have higher average error than the base case, but at a certain “see distance” (depending on the density) they begin to have a lower average error. We have seen already that smaller “see distances” have smaller average errors in the base case. Therefore, it makes sense for the triangular and uniform to have smaller average errors than the base case as “see distance” increases. This is because the distributed efficiencies cause a reduction in “see distance” for all the sensors. As density increases, the base case benefits from having more sensors to detect the source (reducing the average error); however, with the uniform and triangular excursions, an increase in density at a low “see distance” does little to increase the actual number of sensors that can detect the source, hence, the reason we see the difference between the base case and distributions grow as density increases. The reason for the decrease then increase in the curves is a similar reasoning as previously mentioned for last position error of the base case. An increase in the “see distance” recruits more sensors to help pinpoint the source; however, the “see distance” eventually becomes large enough that the reduction in error from the additional sensors recruited is outweighed by the increase in error from having a larger region the source can be in. Despite the excursion having a different shape than the base case, it is still not too much different with the maximum difference being around $.8$ at the higher densities with low “see distances.”

Maximum Error

Overall results are similar to the base case in magnitude with a difference no larger than .9 grid units.

Last Position Error

The last position error of the excursion has a similar trend as the average error. Comparing it to the base case, the largest differences occur at low “see distances” (less than or equal to 1.33 grid units). The difference lessens for 1.33 grid units as density increases. At a density of 1.6 people/m² and “see distances” greater than 1.33 grid units, the difference is less than 1 grid unit. At “see distances” greater than around 3-4 grid units the last position error of the distributions is less than the base case for densities greater than .25 people/m². Though, the largest difference here is about .77 grid units.

Maximum Location Error

The maximum location error of the excursion has a similar trend as the base case. The curve starts out decreasing and then it increases. The uniform and triangular excursions have a much steeper decrease than the base case. The error is also larger than the base case during the decrease. For the distribution of efficiencies, increasing the “see distance” at low values substantially increases the number of sensors likely to detect the source. . Eventually the curve becomes nearly identical to the max error curve because the “see distance” reaches a point where the source is seen all the time.

Significant Findings: Distribution of Sensor Efficiency

By giving a distribution to the efficiencies, there is a substantial impact at “see distances” less than 1.41 grid units for percent seen. A see distance of 1 grid unit is nearly useless since it is highly unlikely that a sensor will be able to detect the source. Changing the distribution of the sensor efficiency (triangular vs. uniform) does not have a significant impact on the effectiveness of the sensor network. We conjecture that a normal distribution would have similar results as the triangular. By having a distributed sensor efficiency the only major effect on the network is with percent seen and maximum location error, and it is only at low densities and/or low “see distances” where this effect may have cause for concern. “See distances” in the interval [2, 4] grid units seem to give the most effective network with a “see distance” of 4 grid units being used for the lower densities, and as densities increase, lowering the “see distance” to 2 grid units for a lower average error.

False Detections

A false detection is defined as a sensor showing a detection when there is no source present within its “see distance” range. With this excursion we introduce attenuation parameters: range and number. For a sensor detection to be considered in the estimated position, there must be a certain number (attenuation input parameter) of sensor detections within the range (attenuation input parameter) of the sensor in question. This will be referenced as attenuation (number of sensors required for detection to occur, range of detection in grid units). Keeping the base-case variables as is, the model is set so that 100% of the sensors can be bad sensors. From the base-case, for each sensor, the probabilities were run giving each sensor a 5% chance of getting a false detection. The three densities run were .25 people/m², 2 people/m², and 4 people/m². At each density, attenuation parameters are varied.

For a density of 4 people/m², at all see distances, in terms of percent of time the source is seen and average error, the attenuation parameter that gave the best results was an attenuation range of 2 grid units and a sensor detection requirement of 5 sensors: attenuation (2,5). If we specifically look at “see distances” higher than 2.5 grid units, attenuation (3, 5) is best (see Charts 49 and 50 in Appendix 5.6).

Using this trend, looking at a density of 2 people/m², we found that the best attenuation parameters for percent of time the source is seen and average error are attenuations (3, 4) and (4, 5) (see Charts 47 and 48 in Appendix 5.6).

Lastly, with a density of .25 people/m², there were no acceptable ratios because the density is too low and the sensors from the parade were confounding the network (see Charts 45 and 46 in Appendix 5.6).

Data Clusters

There are times the data from the simulation runs occur in clusters. The reason for these clusters is due to the grid structure of our model. Recall that for a detection to occur the distance from source to sensor must be less than or equal to the “see distance” for that sensor. Sources and sensors can only be located on integer coordinates; therefore, this creates intervals for the “see distances.” For example, in the chart below the sensor is located at (0, 0). The distance from this sensor to source location 1 is 1 grid unit. If the “see distance” of the sensor is 1 grid unit, it can see the source at location 1, but not at locations 2 and 3. If the sensor has a “see distance” of 1.2 grid units, it has the same effect on the network as if it had a “see distance” of 1 grid unit.

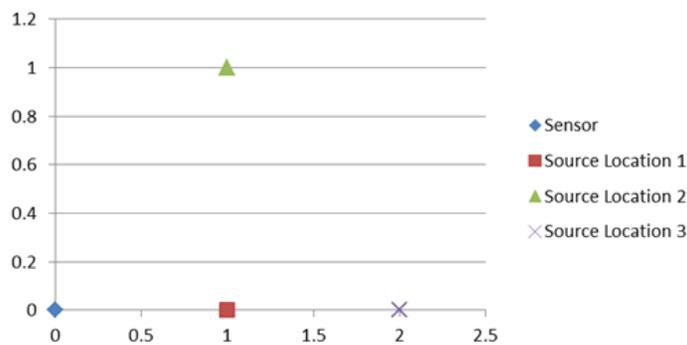


Diagram 4 - Clustering

Not until the “see distance” of the sensor becomes 1.41 grid units will it be able to see a source further away than 1 grid unit. Diagram 5, on the next page, shows an example of clustering. The clusters are more pronounced at the lower left of the chart, but can be seen throughout the data series.

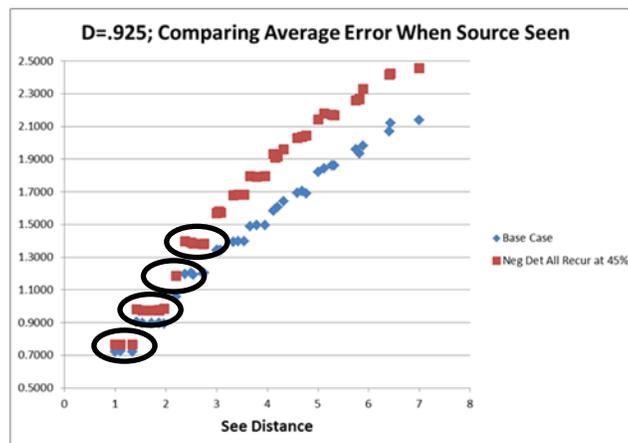


Diagram 5 - Clustering Example

Framework

From our analysis, we have introduced a framework for a (CB)RN(E) event interdiction using a ubiquitous sensor network. The framework takes into consider the variables that can be controlled, rather than those like crowd density (which can only be controlled to an extent) or source strength.

Sensors

If we can estimate source strength, we can estimate what “see distance” is necessary and thus can adjust the strength of a sensor, since “see distance” is based on a source and sensor ratio.

A beneficial characteristic of a ubiquitous sensor network is having sensors that are strong and can be dampened for a given density. For example, if a sensor sees a source, its strength can be dampened to better localize the source.

Additionally, it would be beneficial for a sensor to be able to detect levels of source emanation. This would allow for the localization algorithm to better approximate source location and reduce estimate errors.

Negative Detections

Limiting the percentage chance of a negative detection occurring would be beneficial to a ubiquitous sensor network. This can be done by creating a manufacturing standard for all sensors that are placed within cell phones and other electronics. However, if negative detections cannot be reduced, we recommend having a sensor strength that gives a “see distance” of at least 3 grid units. This allows for a high percent seen for all but the smallest of densities.

False Detections

There are three ways that false detections should be managed. Firstly, given a known density about the network, an attenuation range and number should be chosen in order to reduce false detections. (See False Detection discussion for some initial suggestions). Secondly, based on our analysis, false detections have a much greater effect on the network than negative detections.

We suggest that in order to counter and reduce false detects, sensors have an internal capability of turning “off” or going into a “non-detect mode.” Once a sensor realizes that it is detecting when there is no threat, the sensor will cease to detect and become a negative detector. This will stop unnecessary data overload. A sensor realizes there is no threat by communicating with surrounding sensors. No surrounding detects or detection patterns should signify there is no real threat. Setting an attenuation range and number should aid in this process. Often, one sensor going off in a field of sensors will not be a determining factor of a threat being present.

Lastly, as a look into the future, looking at the effects of other localization methods (i.e. time based) and multi targeting algorithms are significant things that will aid in lowering error. This aspect is left for a future team to complete and analyze the model using multiple targets and exploring other localization models.

Distributions

Distribution of sensor efficiency cannot be changed, however, randomized distribution can be created within an AD HOC sensor network. Sensors of varying efficiencies are better and dampening sensors will help create such a network. Adding larger sensors within the network is another way to make the

sensor AD HOC with different strengths of sensors. With this dampening, random distribution creation can be achieved.

Final Note

When considering characteristics for the framework, random sensor movement and crowd movement for example, are variables in real life that there will be little to no control over. Characteristics such as these where not be considered for the framework.

Future Work

Due to the existing research and time constraints, the team does not intend to model nor include analysis for some variables and metrics. However, we have compiled work that would advance the research of ubiquitous sensor networks for future findings.

Since the model only localizes a Radiological and Nuclear source, an extension of the model is a possibility for future groups. There is potential to insert code for chemical, biological and explosive sources that the current team did not have time to complete. Some initial research and modeling has been conducted for a future team to continue and complete. Other areas for future research include multiple target tracking, time based tracking, and fluid type dispersion.

Multiple Target Tracking

There are two areas in which multiple target tracking will be of importance; first, there can be multiple sources, and second, in determination of false positives. A rough coding idea was considered but the schedule did not allow for implementation. Below is an example of the code where there are four targets originating from the same location. Target one will move up at one grid per chart. Target two will move right one grid and up one half a grid per chart.

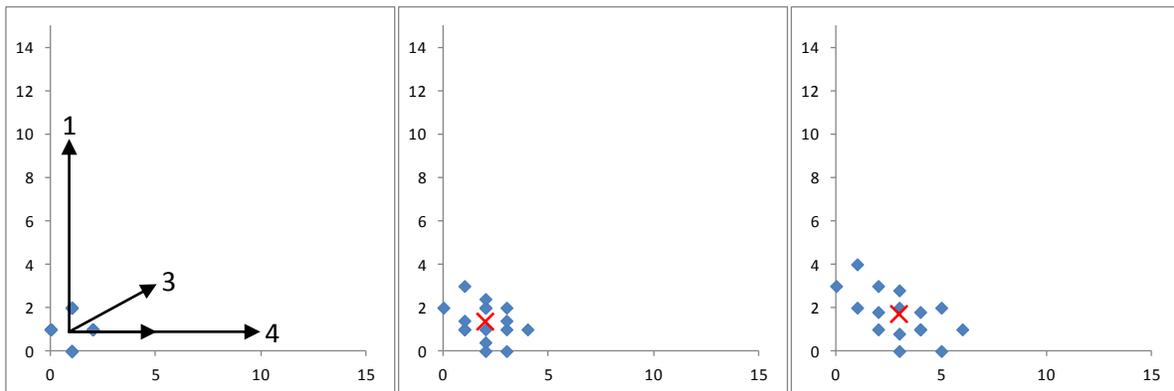


Diagram 6: Direction

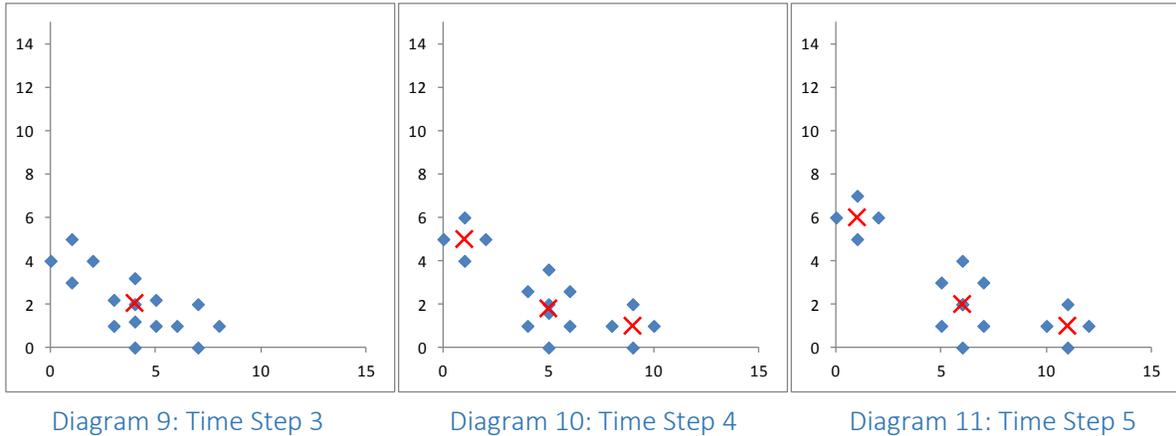
Diagram 7: Time Step 1

Diagram 8: Time Step 2

Target three moves right one grid per chart and target four moves right at two grids per chart. The actual paths are shown in Diagram 6 above by the arrows. The “blue diamond” icons represent detections and the “red X” icons represent estimated positions.

The first and most important characteristic for separating multiple targets in order to be considered a new target is the distance between detections (or resolution). If this distance is too small each detection would be considered a target. If the distance is too large, many false targets are then included

in the target location estimation which will result in larger location errors. The six charts representing a multiple sources uses 1.5 grid as a required resolution. In Diagram 7, the detections are grouped close together and only a single source is identifiable. In Diagram 8, visually, it appears as if there are several targets, but the resolution is insufficient to break them out.



In Diagram 9, it is apparent that there is more than one target, but again due to the chosen resolution, only a single target is seen. If a smaller resolution were chosen, two targets could be broken out in Diagram 9. With the chosen resolution, the model sees three target groups in Diagram 10 and can track them as separate targets until the distance between them falls below the chosen resolution. The last chart in Diagram 11, shows that the two center targets do not have sufficient space between them to differentiate them. The above example is just one method of multiple target tracking and a simplistic one at that.

Adding multiple targeting ability will increase the performance of the sensor grid. The first benefit is the ability to track multiple sources. The second and closely related is that the model would be able to differentiate between groups of detections. These groups could fall into one of three different categories. The first, as mentioned earlier, is an actual source. The second is a group of false detections (there happen to be several erroneous detections in a given area even though nothing of significance in the area). The third is a group of detections that were set off by something that wasn't actually a source, but something in the area and it is being detected as one. Using time in conjunction with multiple target tracking will enhance the models ability to differential between the three categories of detections.

Time Based Tracking

As mentioned above time based tracking will help to differential types of targets when multiple targets are seen. The first category is the group of false detections, in which there was a random group of detections in the same area. It is likely these will not persist for a long duration and not have a movement associated with them. By using a timeframe requirement for the group to exist, these false targets can be eliminated and improve estimated position. The other two groups will be harder to differential. An actual source and a false source could have very similar characteristics. It is possible this

may only be eliminated through sensor design, unless there is a possible behavior different associated that could be programmed to separate the two.

Time based tracking will also allow for better positioning though future estimation of position. With multiple positions and times, speed and direction can now be estimated and future positions can be estimated. With map overlays, additional information could be gleaned. One parameter studied was distance from last estimated position. This position was a stationary position until the source was seen again. With time based tracking, new estimates could be produced and this error could be reduced as well as adding new metric to fine tune future estimations. Multiple target and time based tracking open up new areas to study.

Fluid type dispersion

The current model only allows for the tracking of radiological/nuclear type device due to the emanation type. Once multiple targets can be tracked and sensed over time, chemical/biological devices can be tracked. This type of dispersion, which is fluid in the nature of its dispersion, requires time to model effectively. The multiple target tracking is important in estimating the source position. As a chemical/biological device moves through an area, it will have characteristics of multiple targets trailing behind the actual source location. Looking at these targets behind the source as discrete points, each will start as a concentrated mass with small diameter. As time continues, it will increase in diameter and decrease in strength. Eventually, it will disperse to be a large area with little strength, and at some point in time, the sensor will no longer be able to sense the substance. This brings several new parameters into the problem. Not only is source strength important, but also its dispersion characteristics. How fast will it disperse? What effects will crowd density have on its dispersion? Weather would have an impact and could be modeled. The research that can take place in this type of sensor network is vast. The major problem facing the next team to look at this is the same one we faced, how to scope the problem from the many details that could be included and studied.

References

Cell-All: Super Smartphones Sniff Out Suspicious Substances. (n.d.). Retrieved from Homeland Security: <https://www.dhs.gov/cell-all-super-smartphones-sniff-out-suspicious-substances>

Creation of the Department of Homeland Security. (n.d.). Retrieved from Homeland Security: <http://www.dhs.gov/creation-department-homeland-security>

Kansal, A., Goraczko, M., & Zhao, F. (2007). *Building a Sensor Network of Mobile Phones.* Cambridge, MA.: Microsoft.

Knoll, G. (2010). *Radiation Detection and Measurement.* John Wiley & Sons.

Macy's Thanksgiving Day Parade. (2011). Retrieved from about.com: gonyc.about.com/od/thanksgiving/tp/2011-Macys-Thanksgiving-Day-Parade

Measurement: Activity: How Much Is Present? (n.d.). Retrieved from Guidance for Radiation Accident Management: <http://orise.orau.gov/reacts/guide/measure.htm>

Nave, C. R. (2014). *Inverse Square Law, Radiation.* Retrieved from HyperPhysics: <http://hyperphysics.phy-astr.gsu.edu/hbase/forces/isq.html>

Nelson, R. (2012). *Concept of Operations for Wireless Sensor Networks.* Naval Postgraduate School.

Oberhagemann, D. (2012). *Static and Dynamic Crowd Densities at Major Public Events.*

Park, J. (2005). *Management of Ubiquitous Sensor Networks.* Kyungpook National University.

Radiation Basics: How Can You Detect Radiation? (n.d.). Retrieved from Health Physics Society: <http://hps.org/publicinformation/ate/faqs/radiation.html>

Stankovic, J. (2006). *Wireless Sensor Networks.* University of Virginia.

Still, K. G. (2014). *Crowd Risk Analysis.* Retrieved from GKStill.com: <http://gkstill.com/>

Sutter, J. (2010). *Smart Dust Aims to Monitor Everything.* Retrieved from CNN: <http://www.cnn.com/2010/TECH/05/03/smart.dust.sensors/>

Appendix 1: Deriving See Distance

Let

R = distance from source to sensor (grid units)

So = source strength (source units)

Se = sensor strength

A sensor has the ability to detect $\frac{1}{Se^2} \frac{\text{source units}}{\text{grid units}^2}$ and greater

The source degrades using the equation: $\frac{So}{R^2}$

Therefore, for a detection to occur, $\frac{So}{R^2} \geq \frac{1}{Se}$

This simplifies to $R \leq \sqrt{So \times Se}$

We define $\sqrt{So * Se}$ as the see distance of the sensor.

Appendix 2: Density Charts

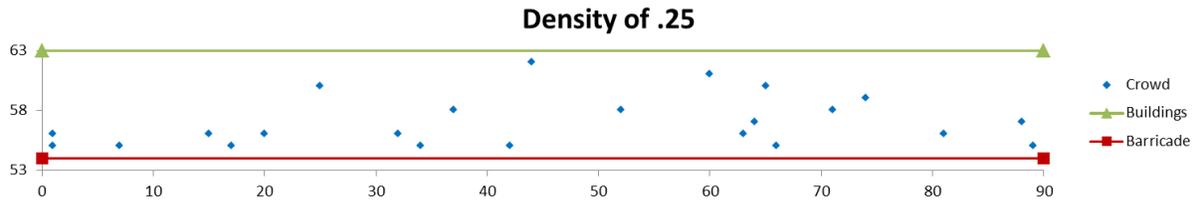


Chart 1: Crowd Density of .25

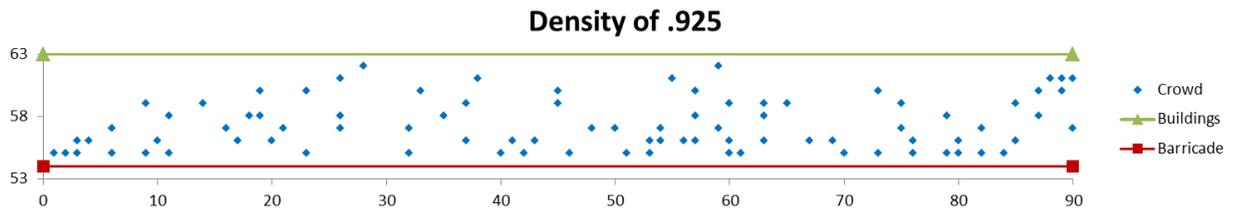


Chart 2: Crowd Density of .925

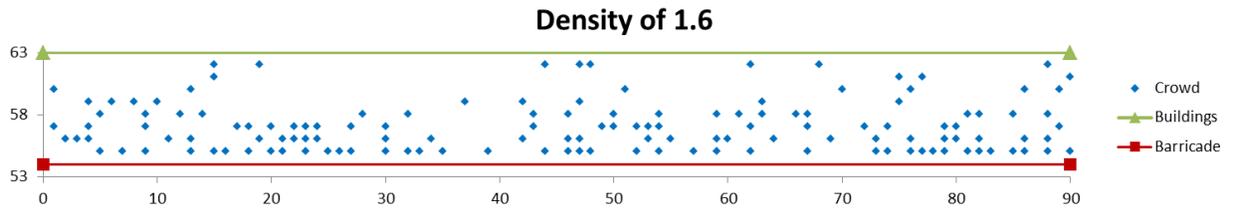


Chart 3: Crowd Density of 1.6

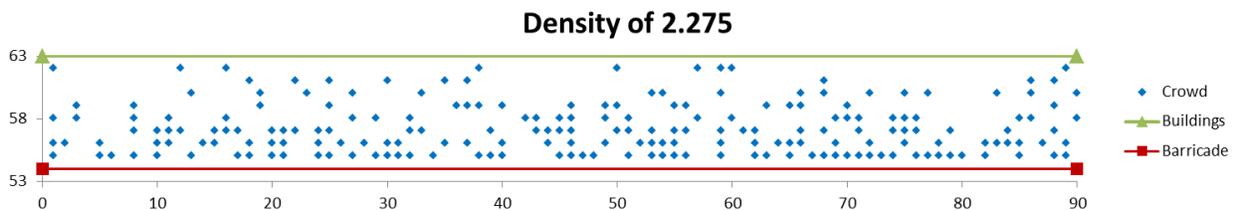


Chart 4: Crowd Density of 2.275

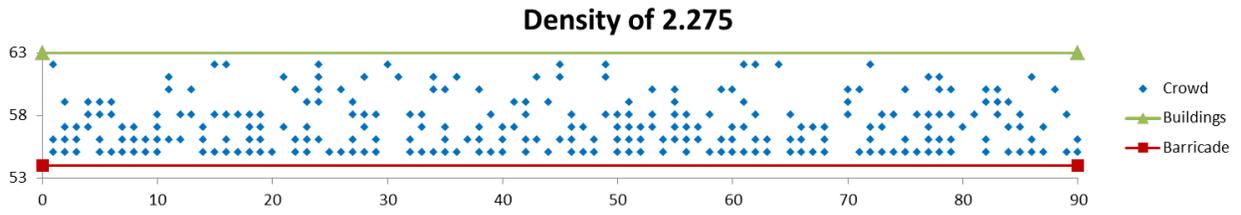


Chart 5: Crowd Density of 2.95

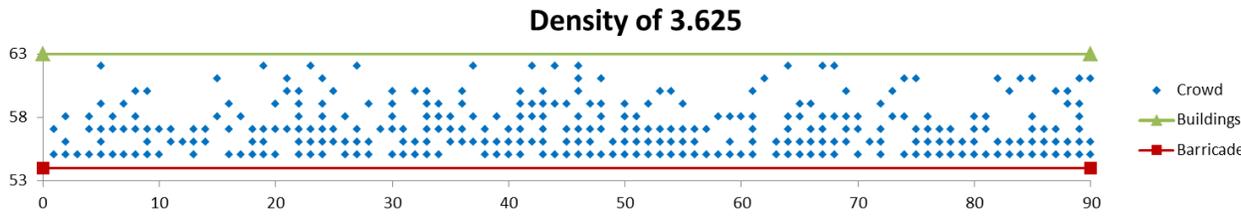


Chart 6: Crowd Density of 3.625

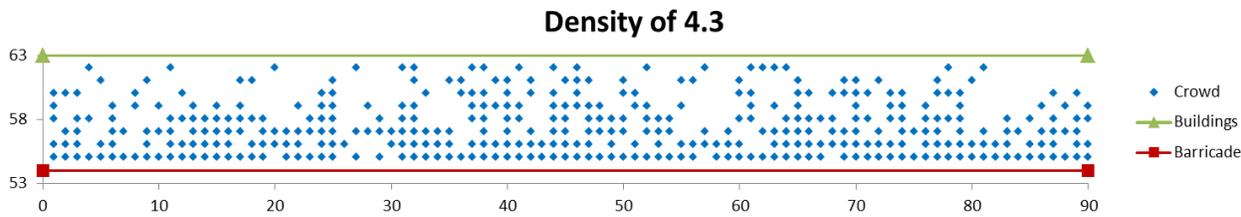


Chart 7: Crowd Density of 4.3

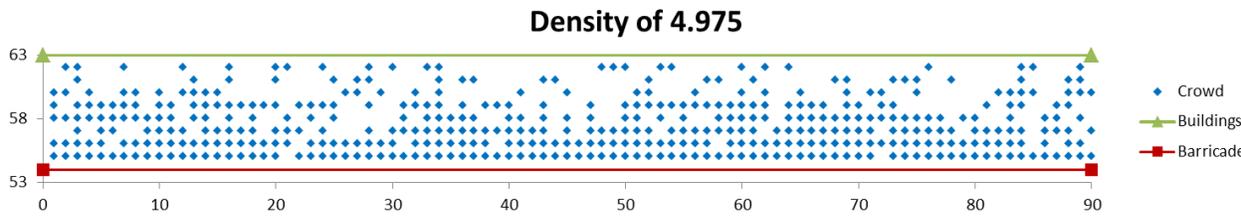


Chart 8: Crowd Density of 4.975

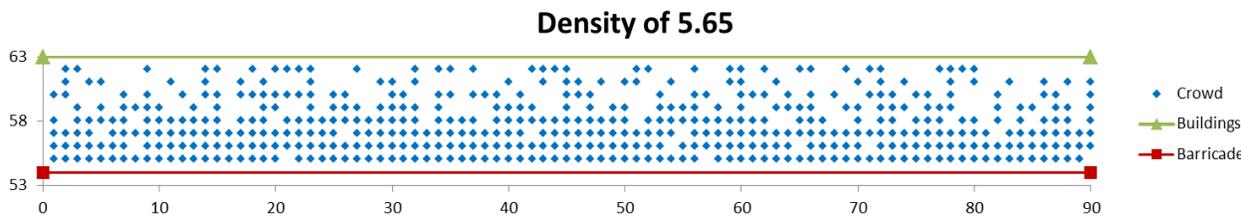


Chart 9: Crowd Density of 5.65

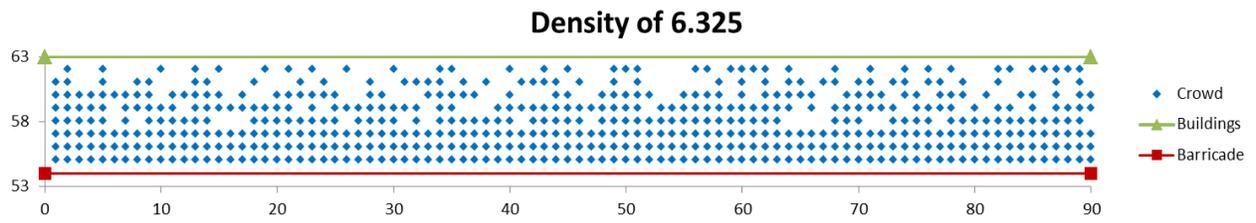


Chart 10: Crowd Density of 6.325

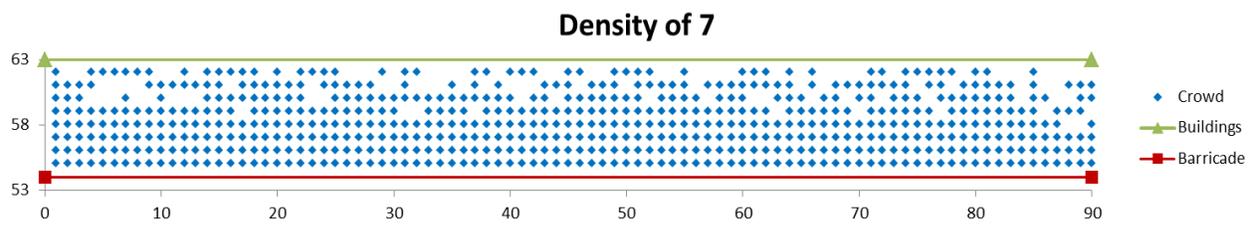


Chart 11: Crowd Density of 7

Appendix 3: VBA Model

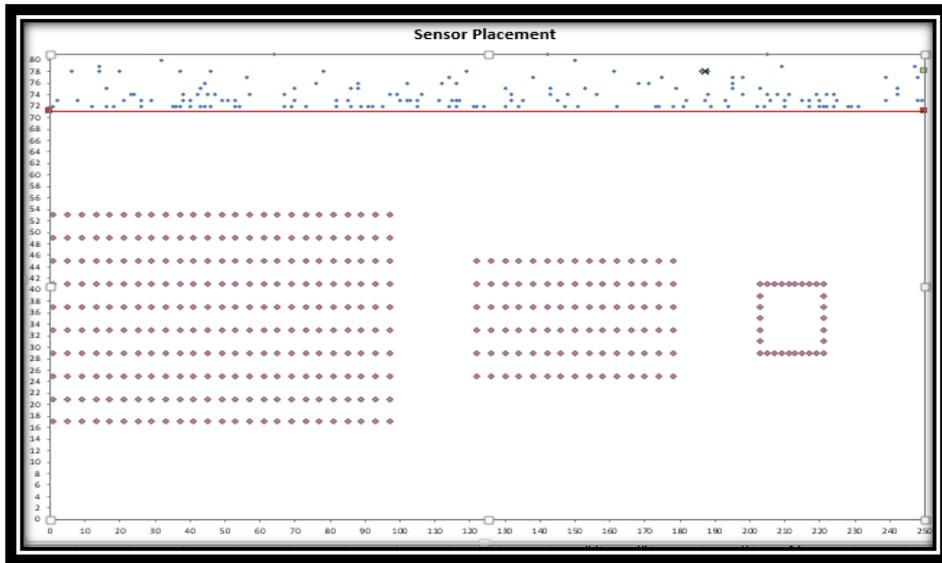


Figure 1: Excel Chart resulting from a run with a low density distribution of spectators and placement of participants

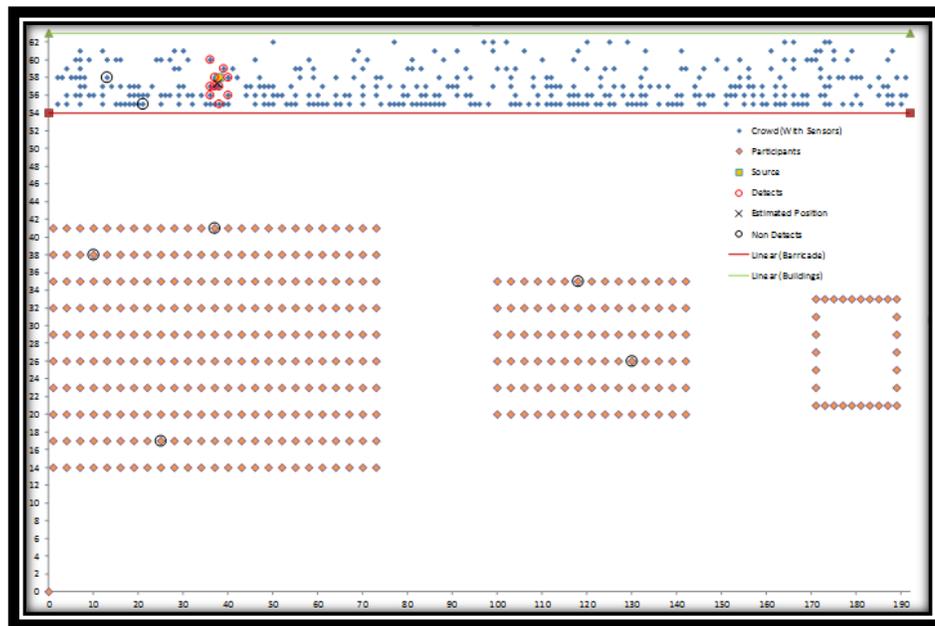


Figure 2: Excel Chart shows high spectator density distribution in blue, participant distribution in pink, source with a black x, and sensors that report source detection with a red highlight.

Appendix 4: QGIS Representation

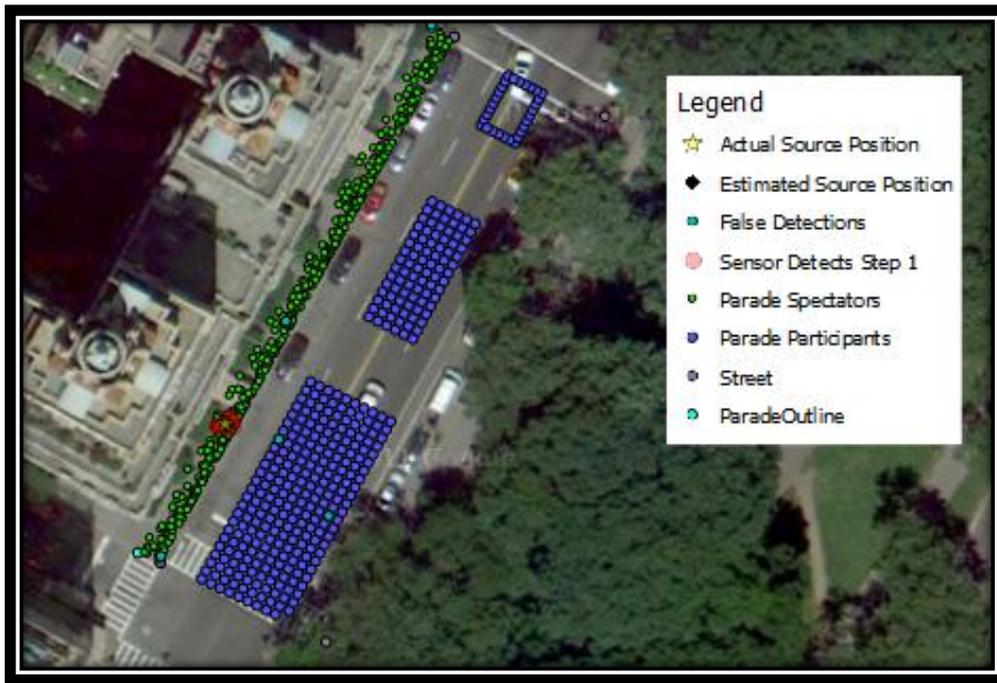


Figure 3: Quantum GIS model of Parade Viewing and Participant Area from 75th Avenue to 74th Avenue

Appendix 5: Data Charts

Appendix 5.1: Base Case Analysis Charts

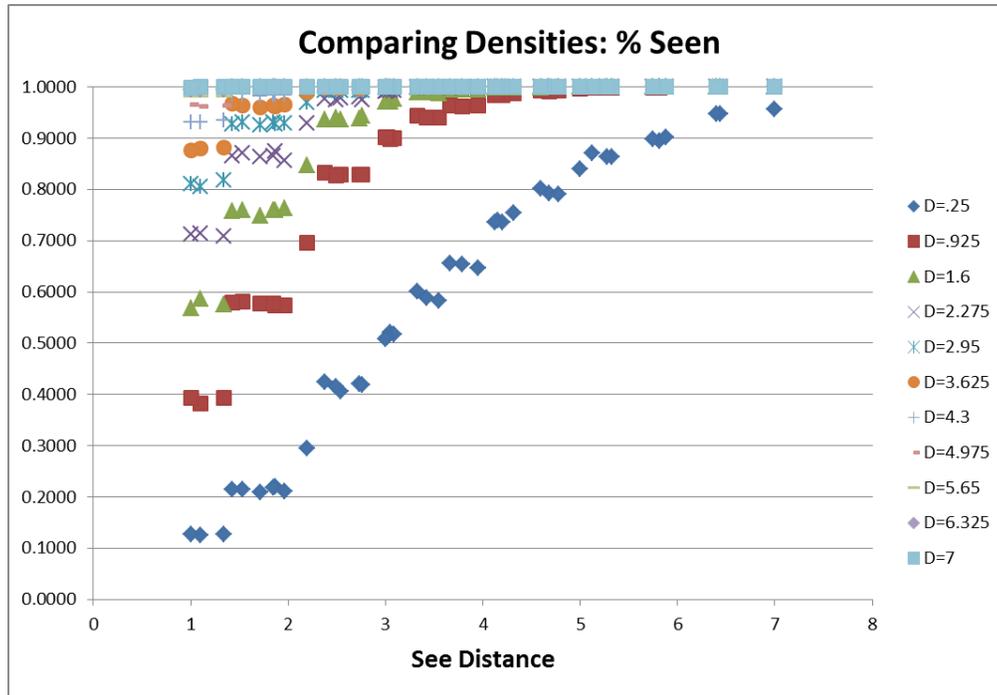


Chart 12: Percent of Time the Source is Seen for Base Case

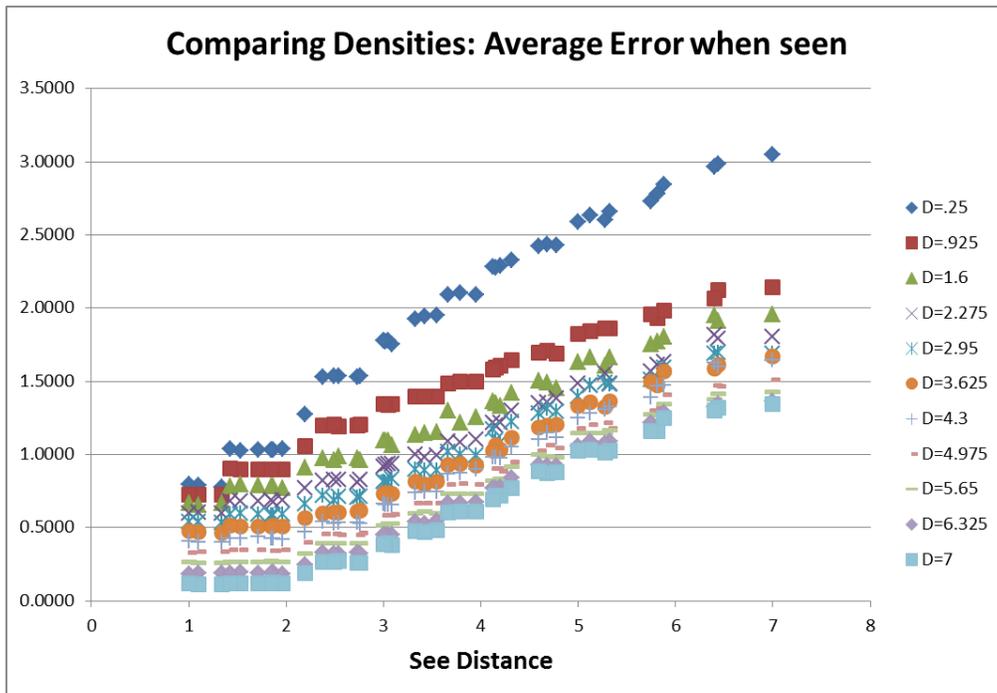


Chart 13: Average Error When the Source is Seen for Base Case

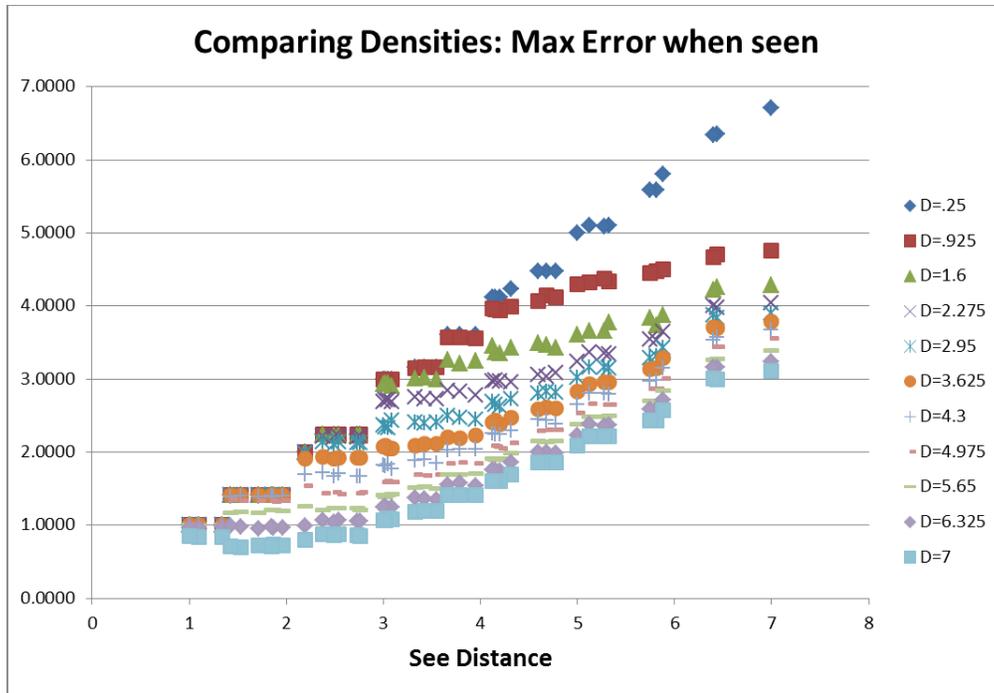


Chart 14: Maximum Error When the Source is Seen for Base Case

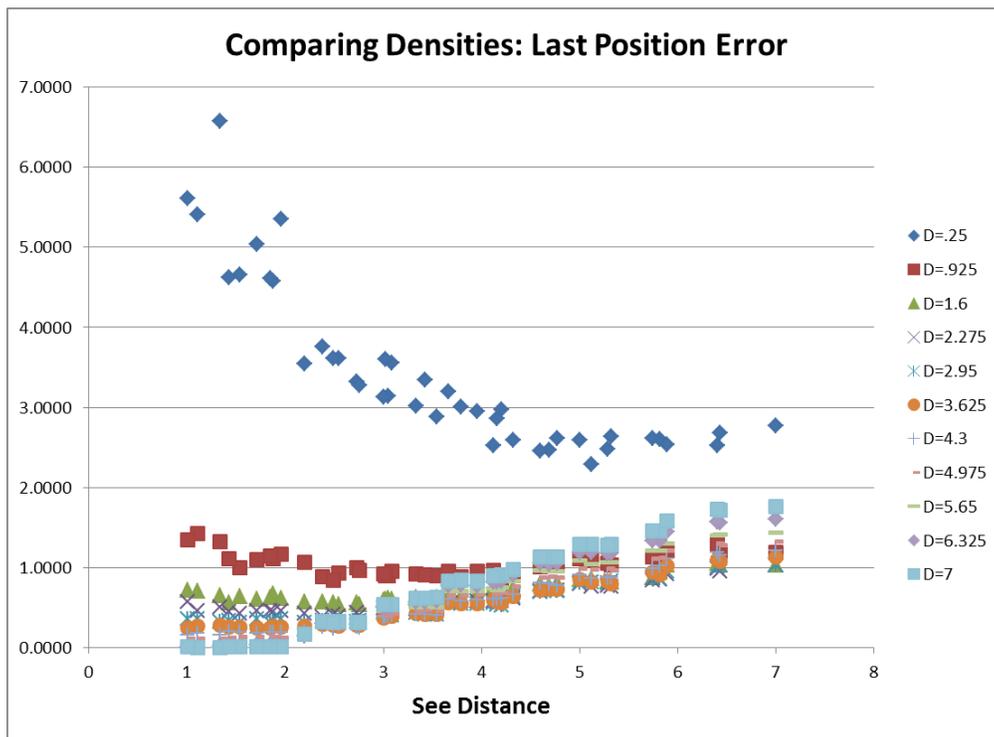


Chart 15: Last Position Error for Base Case (All Densities)

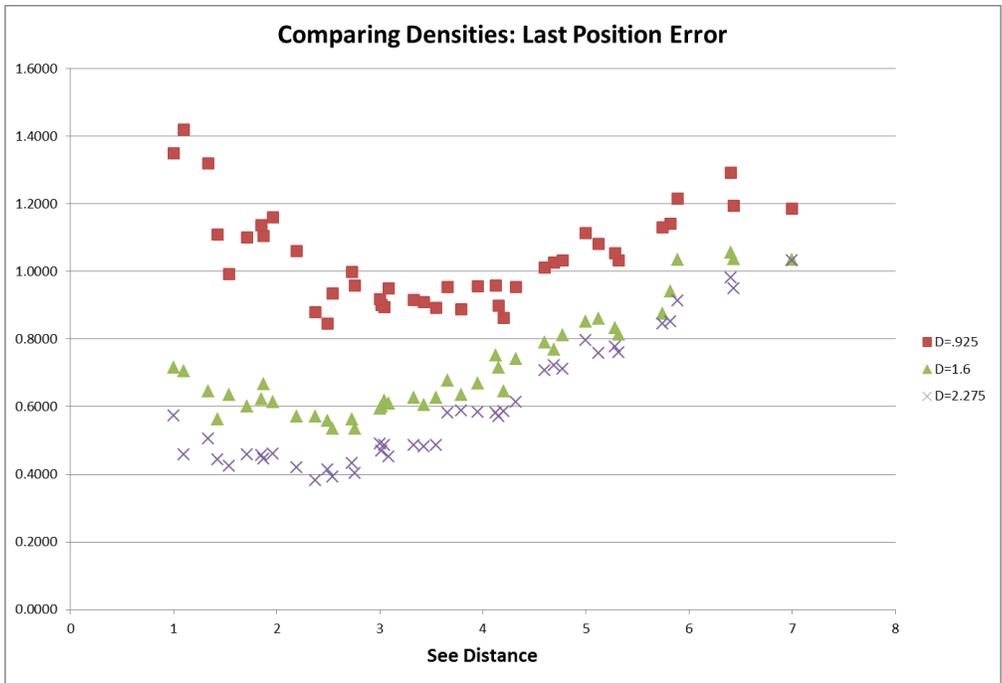


Chart 16: Last Position Error for Base Case (Low Densities)

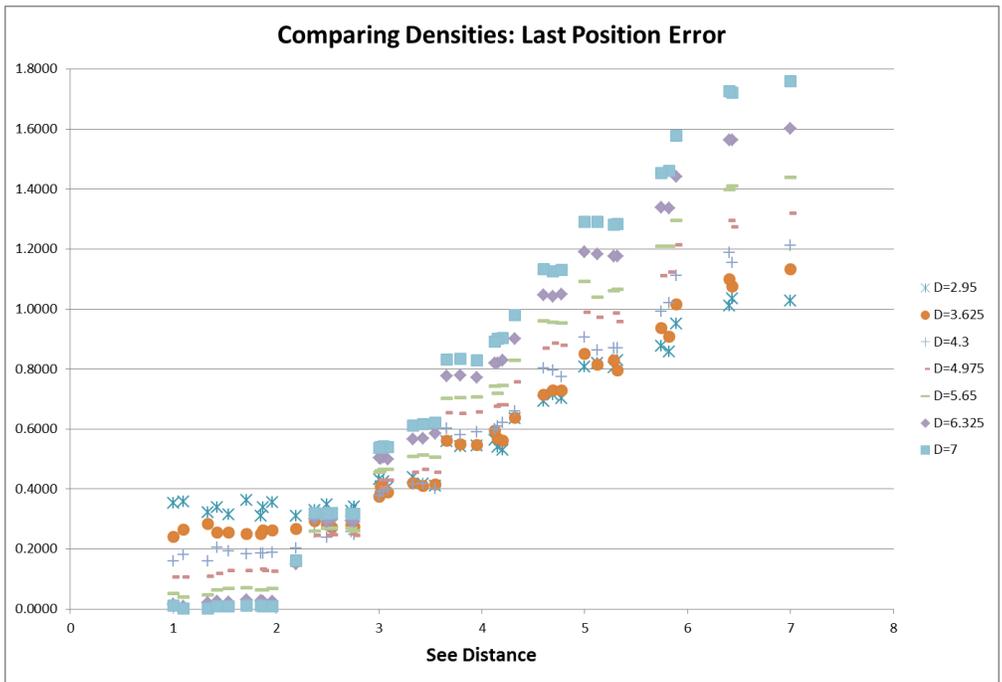


Chart 17: Last Position Error for Base Case (High Densities)

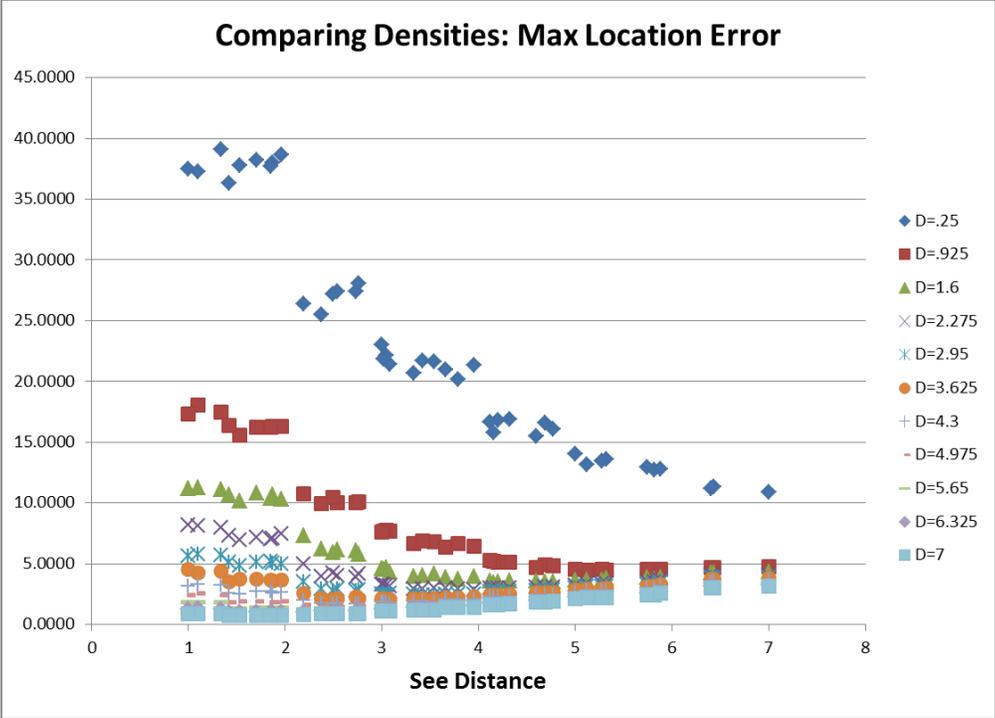


Chart 18: Maximum Location Error for Base Case

Appendix 5.2: Negative Detection Analysis Charts

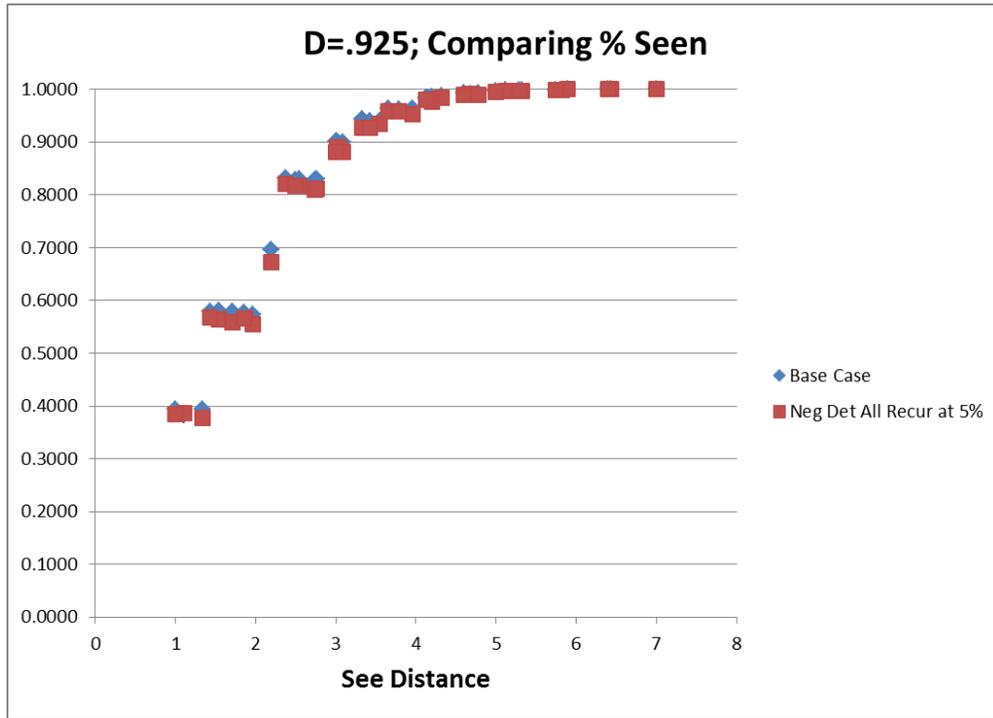


Chart 19: Percent of Time the Source is Seen for Negative Detections (5%) – Density of .925

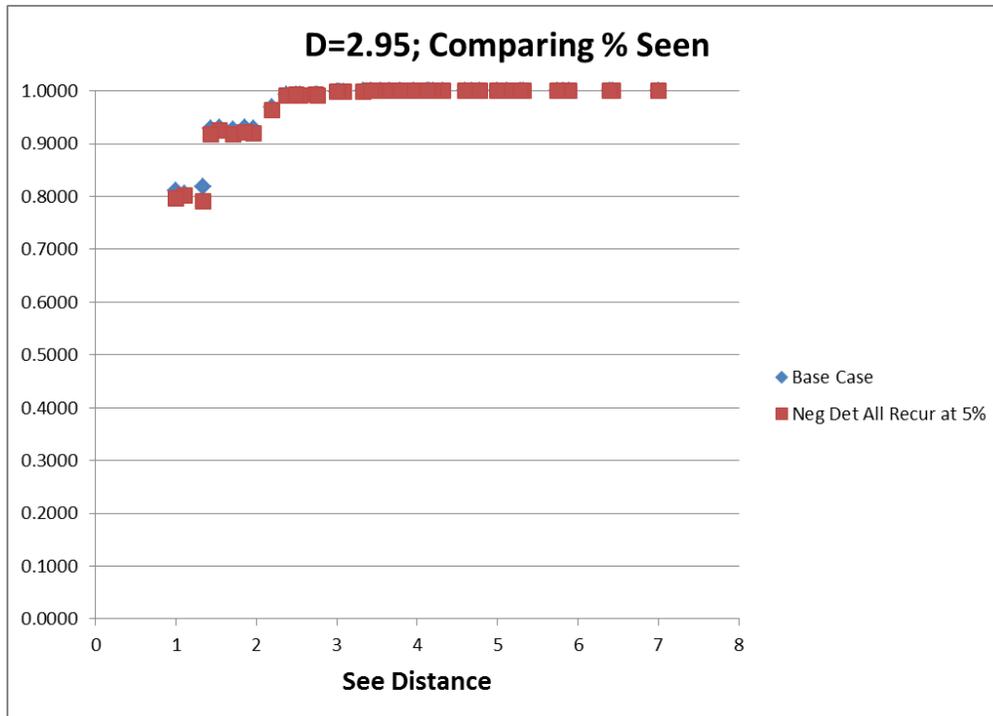


Chart 20: Percent of Time the Source is Seen for Negative Detections (5%) – Density of 2.95

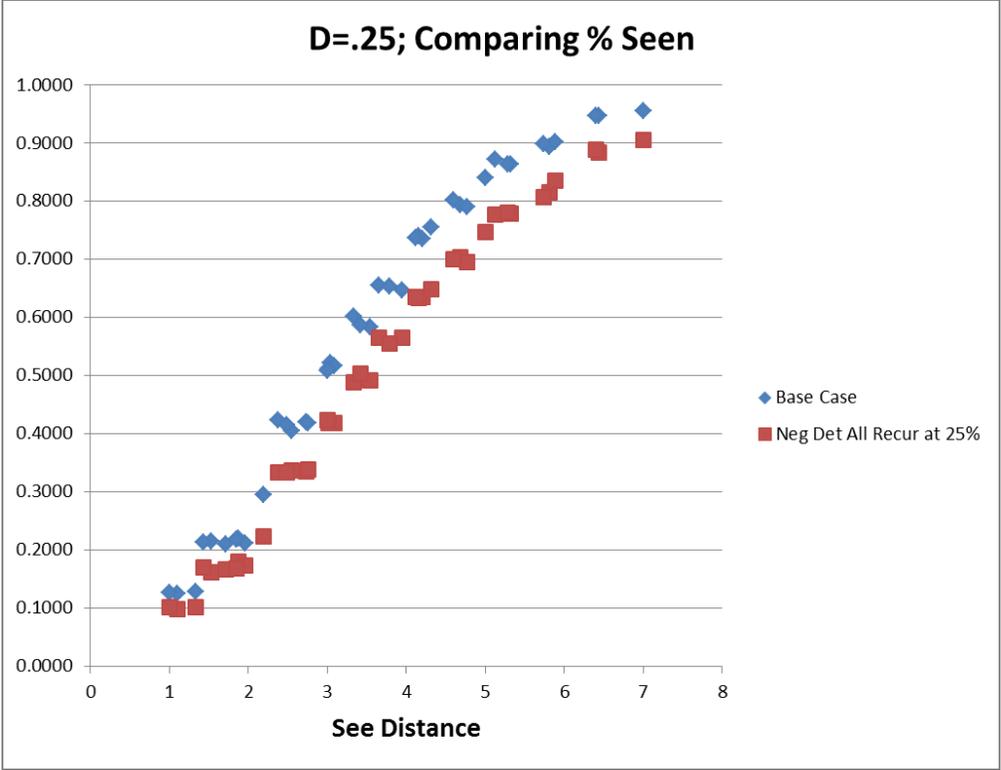


Chart 21: Percent of Time the Source is Seen for Negative Detections (25%) – Density of .25

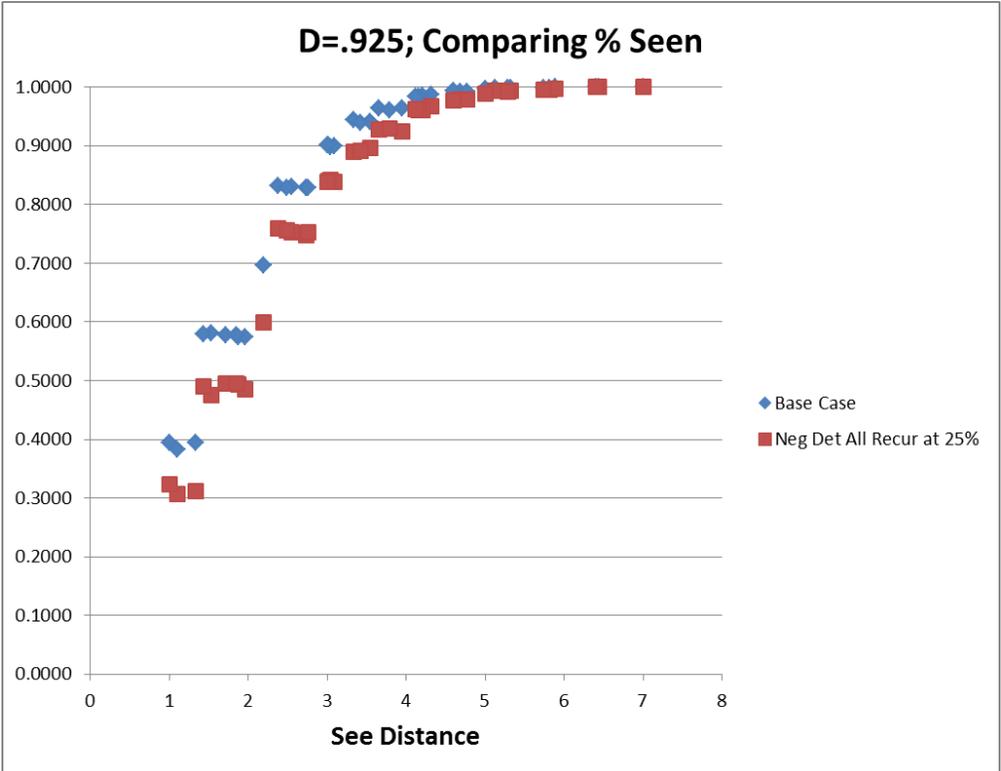


Chart 22: Percent of Time the Source is Seen for Negative Detections (25%) – Density of .925

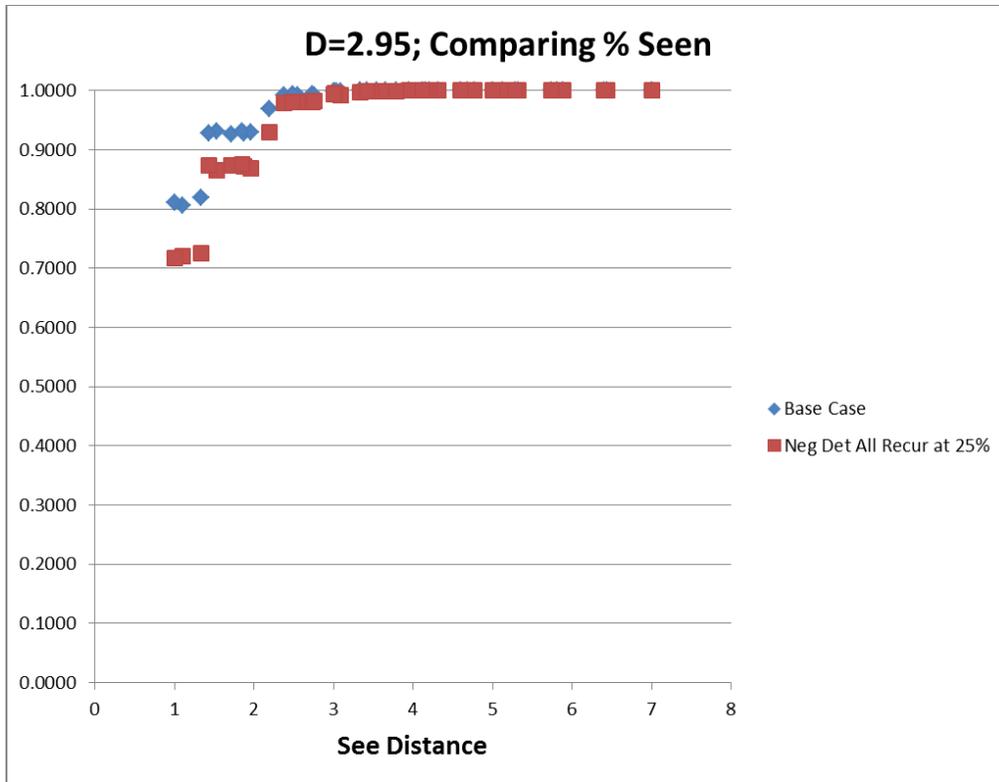


Chart 23: Percent of Time the Source is Seen for Negative Detections (25%) – Density of 2.95

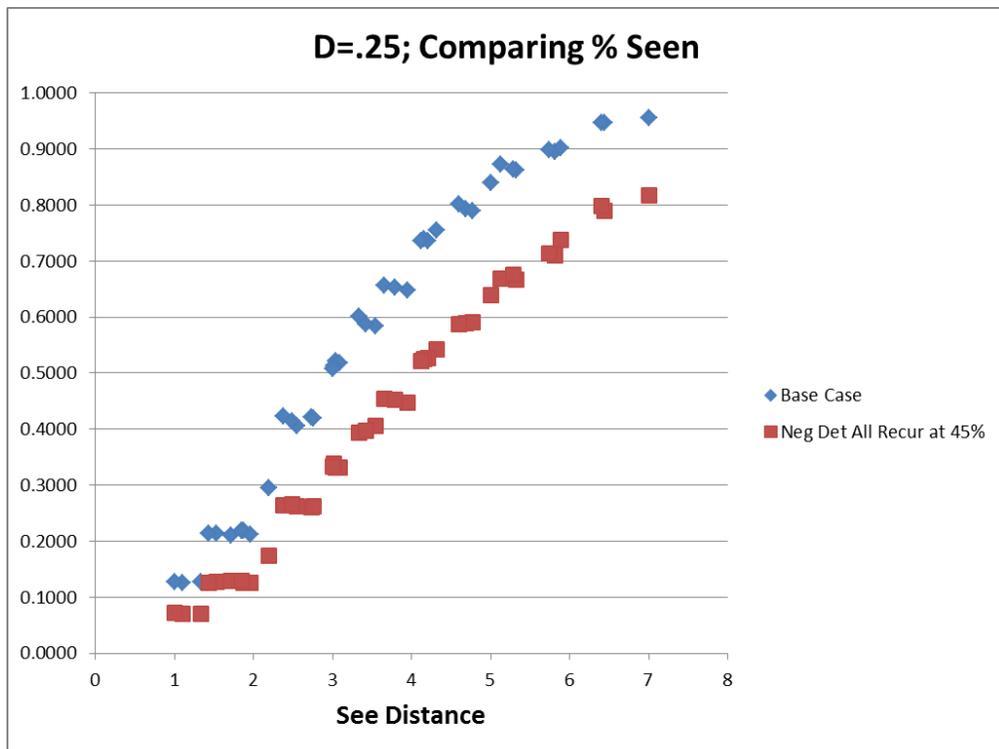


Chart 24: Percent of Time the Source is Seen for Negative Detections (45%) – Density of .25

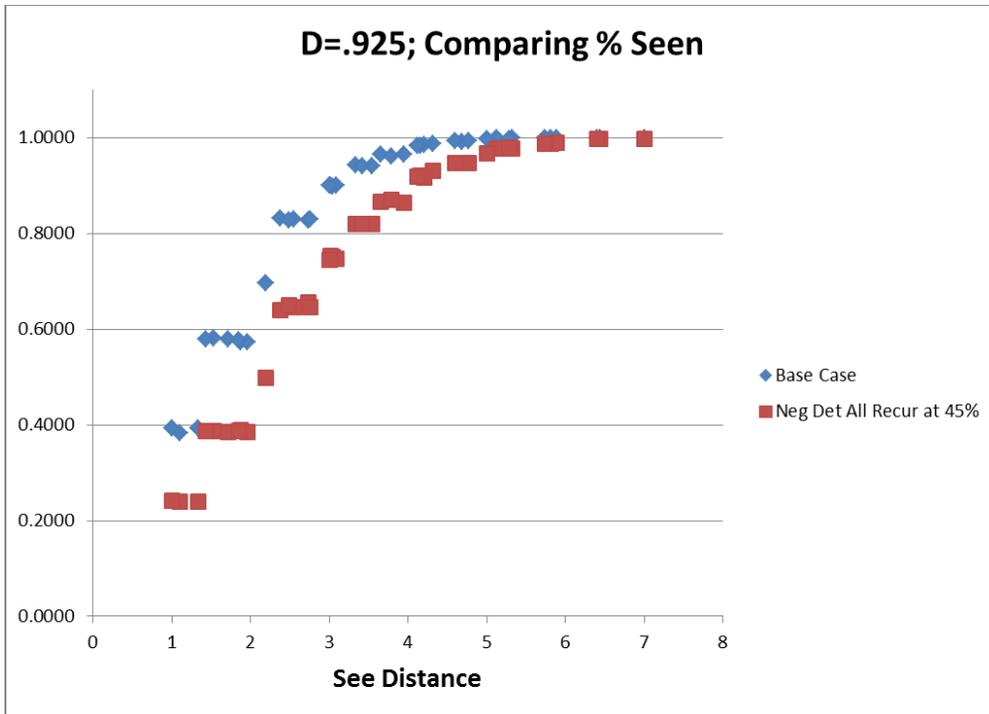


Chart 25: Percent of Time the Source is Seen for Negative Detections (45%) – Density of .925

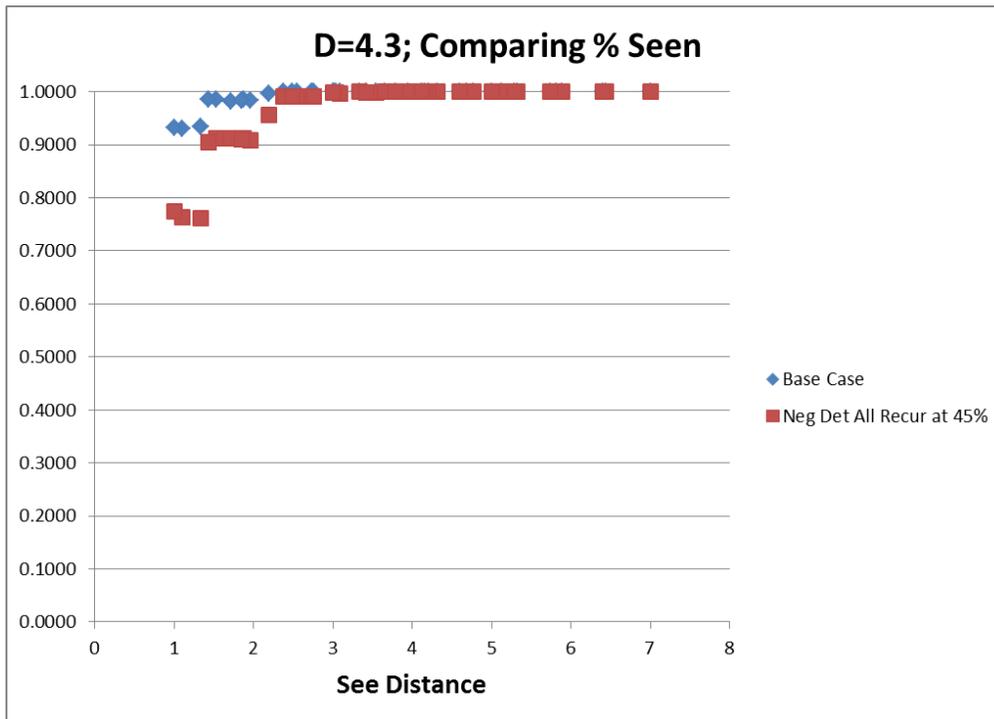


Chart 26: Percent of Time the Source is Seen for Negative Detections (45%) – Density of 4.3

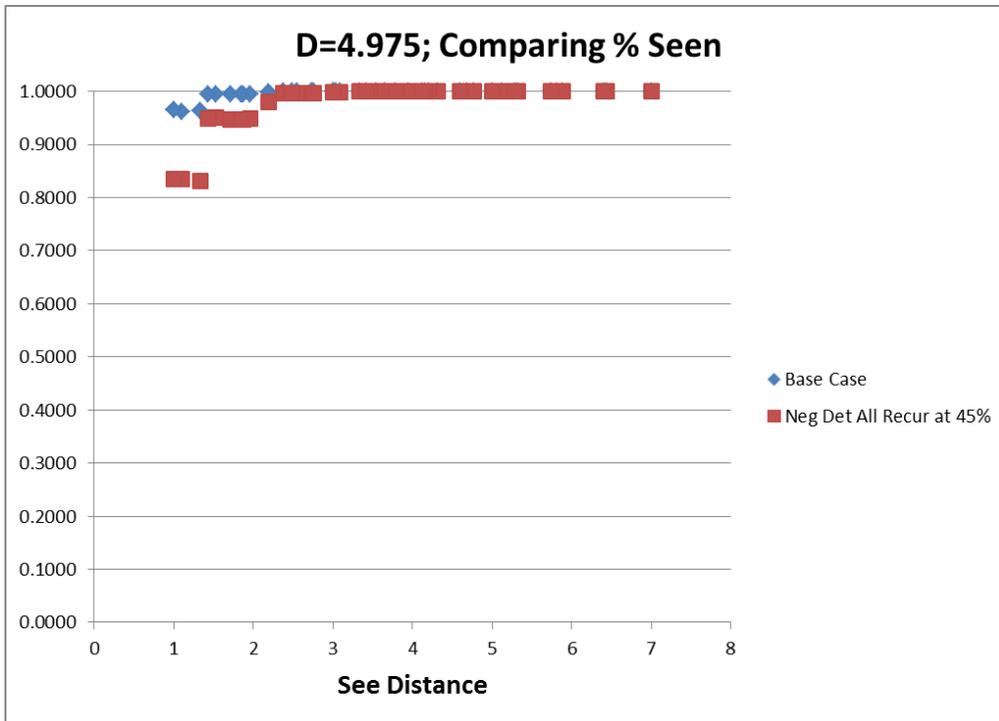


Chart 27: Percent of Time the Source is Seen for Negative Detections (45%) – Density of 4.975

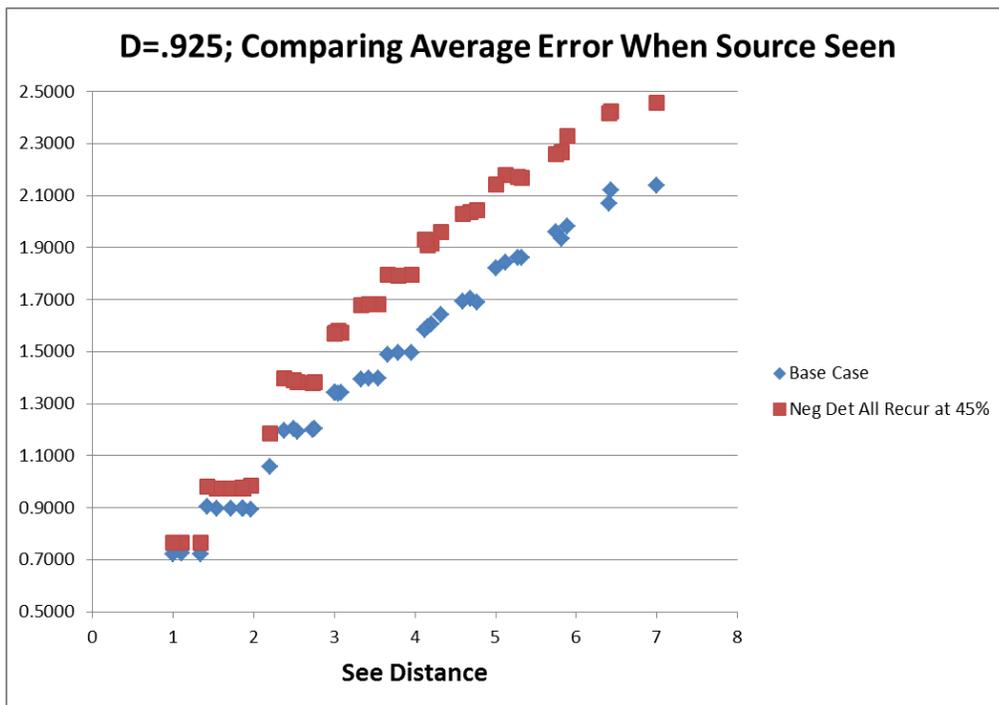


Chart 28: Average Error When the Source is Seen for Negative Detections (45%) – Density of .925

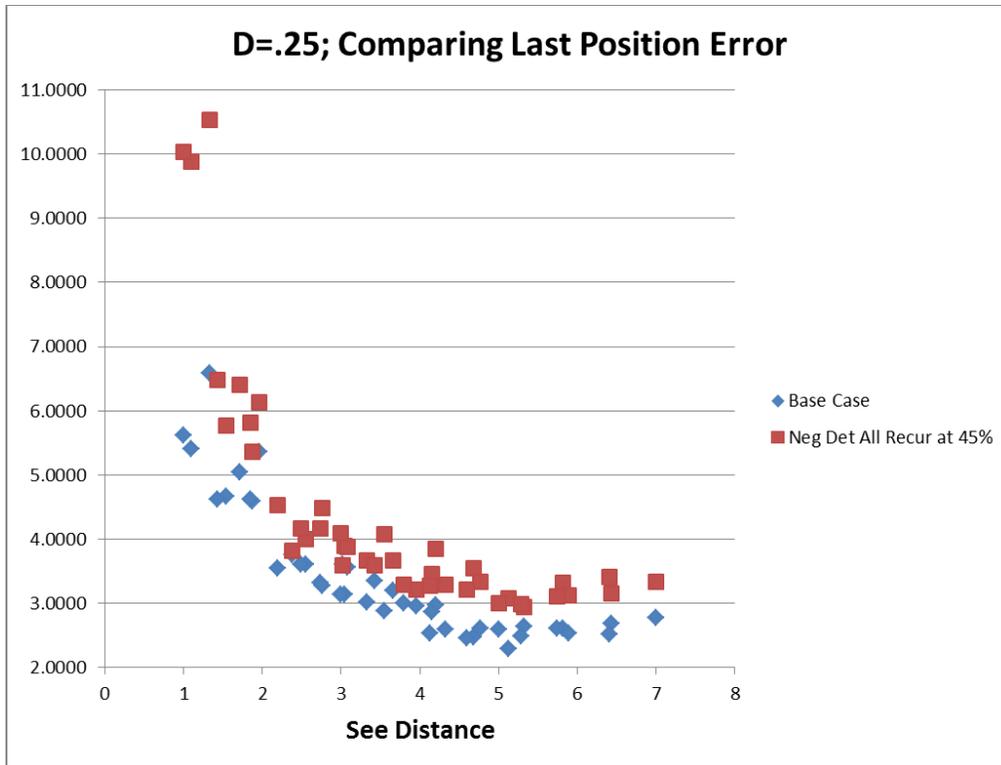


Chart 29: Last Position Error for Negative Detections (45%) – Density of .25

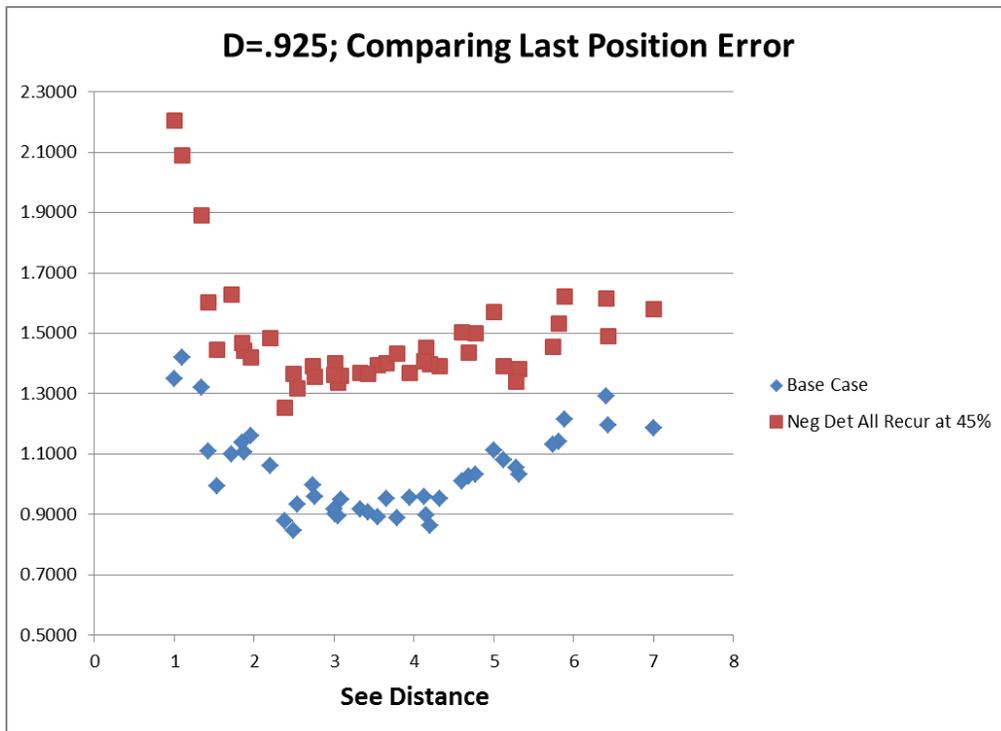


Chart 30: Last Position Error for Negative Detections (45%) – Density of .925

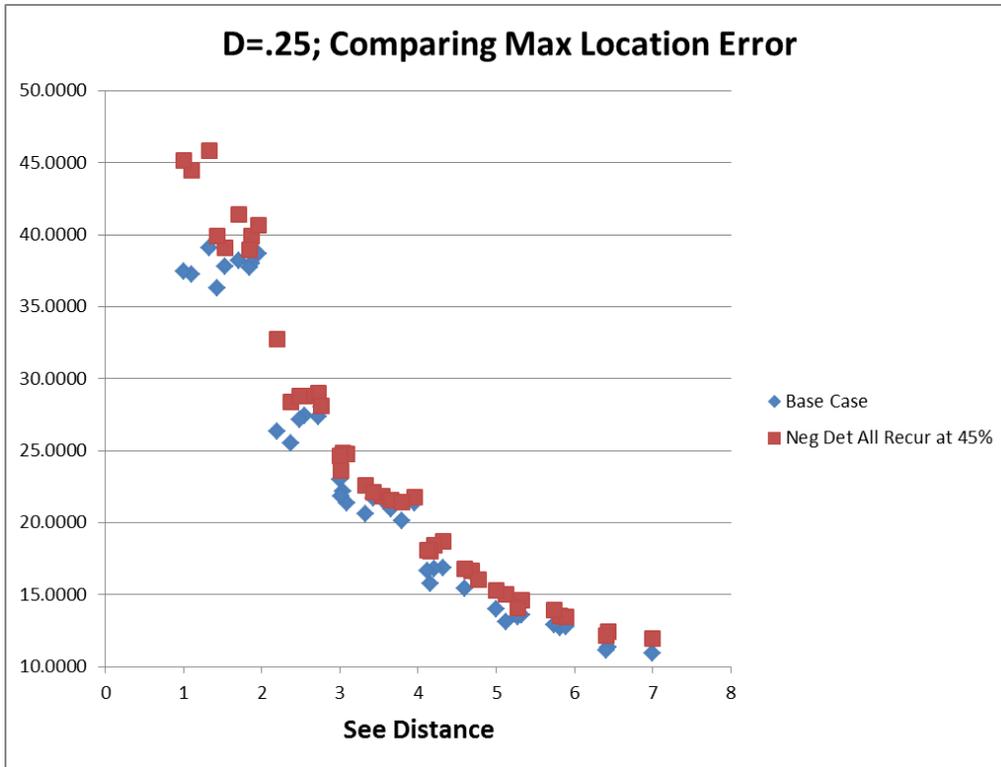


Chart 31: Maximum Location Error for Negative Detections (45%) – Density of .25

Appendix 5.3: Random Source Movement Analysis Charts

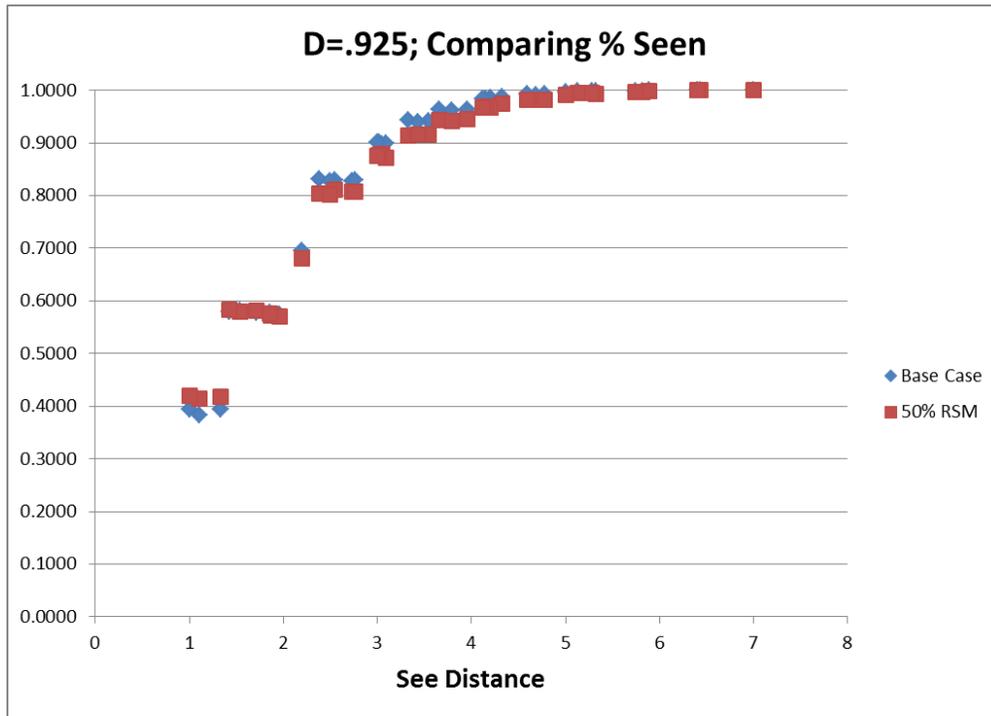


Chart 32: Percent of Time the Source is Seen for Random Source Movement – Density of .925

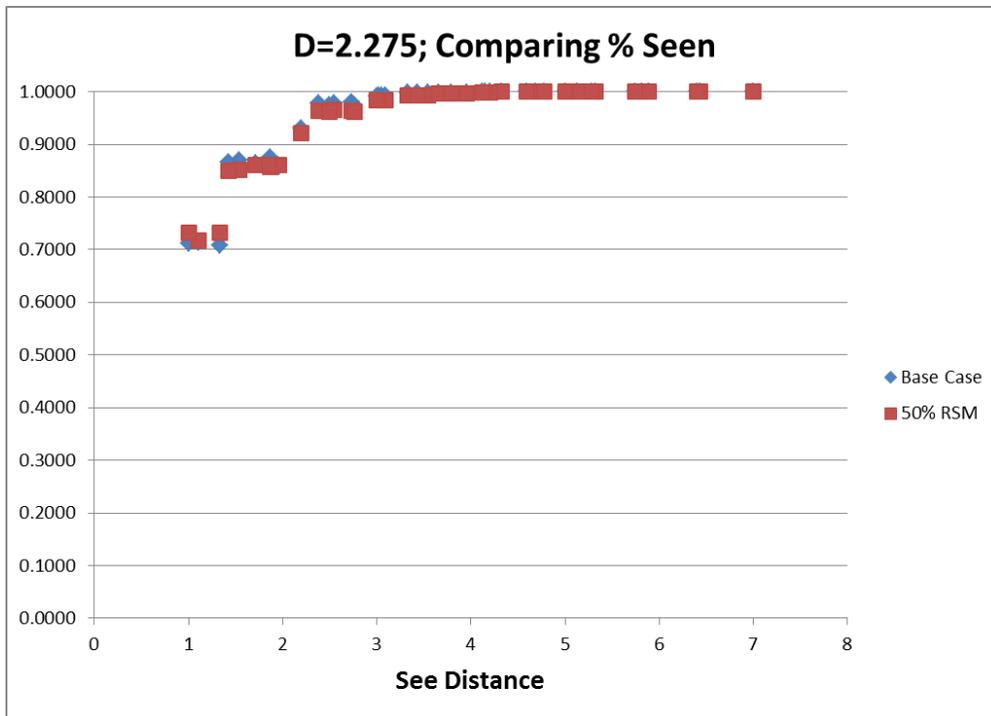


Chart 33: Percent of Time the Source is Seen for Random Source Movement – Density of 2.275

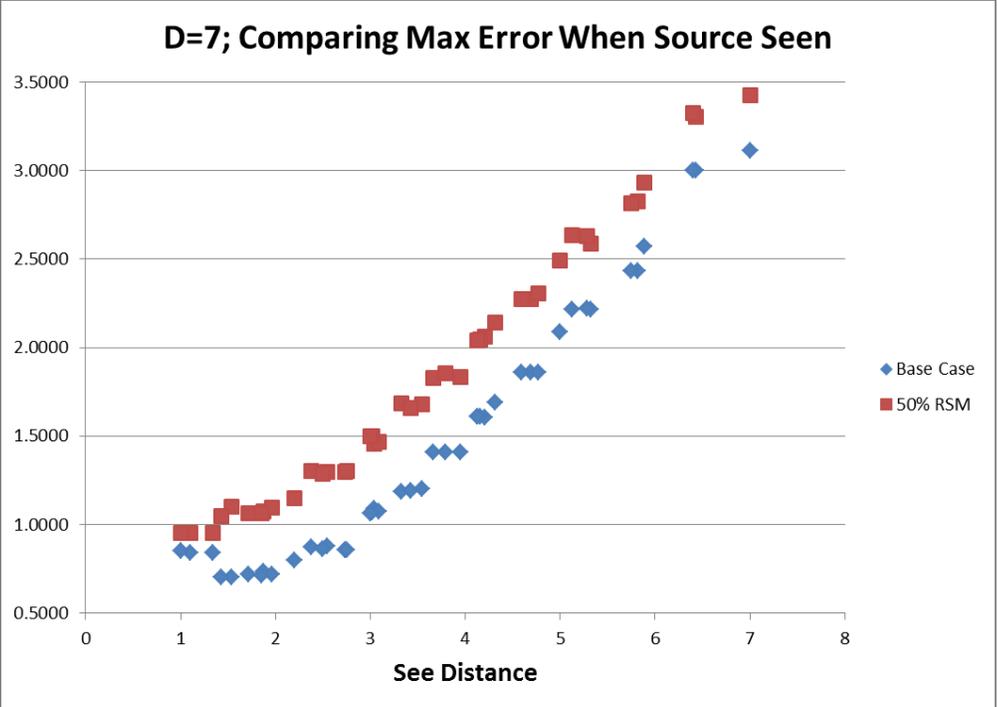


Chart 34: Maximum Error When the Source is Seen Random Source Movement – Density of .925

Appendix 5.4: Crowd Movement Analysis Charts

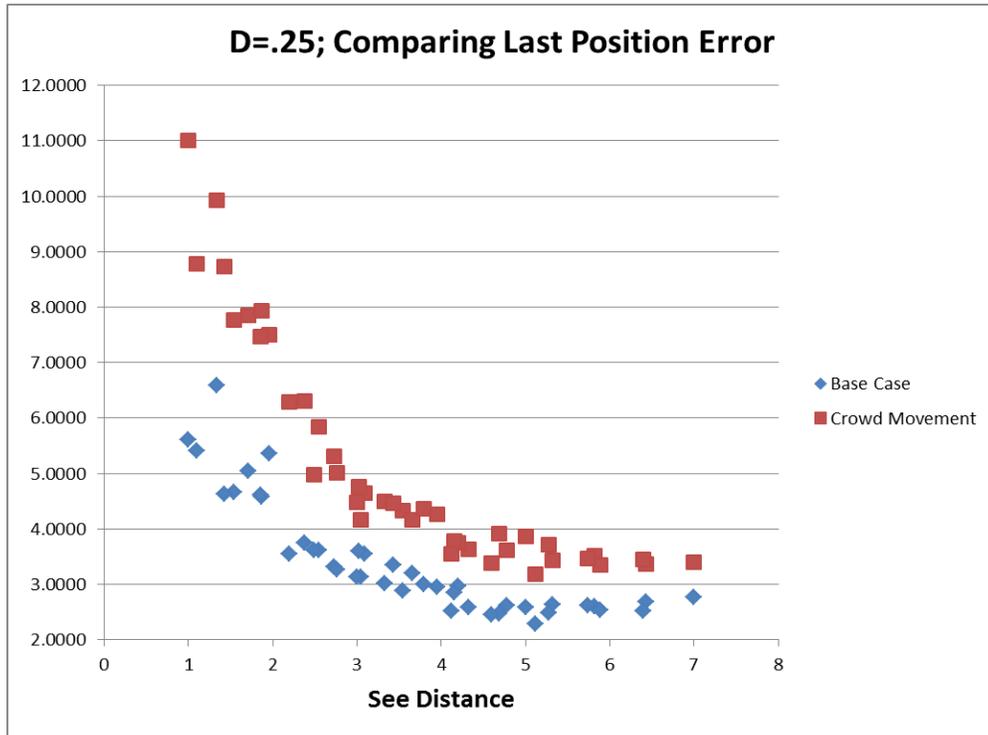


Chart 35: Last Position Error for Crowd Movement – Density of .25

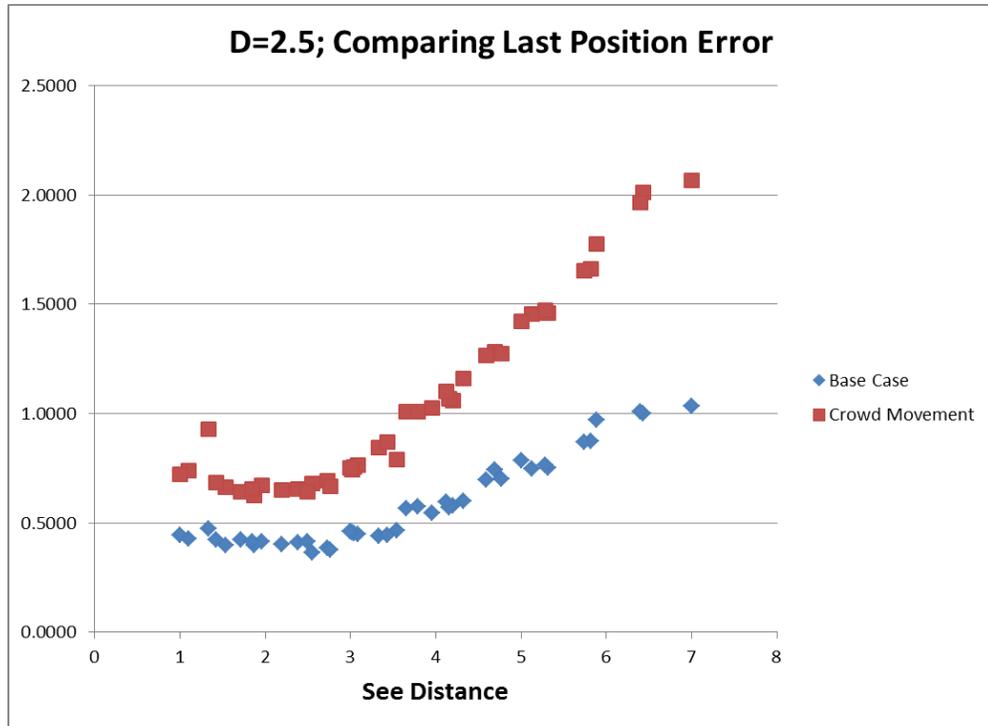


Chart 36: Last Position Error for Crowd Movement – Density of 2.5

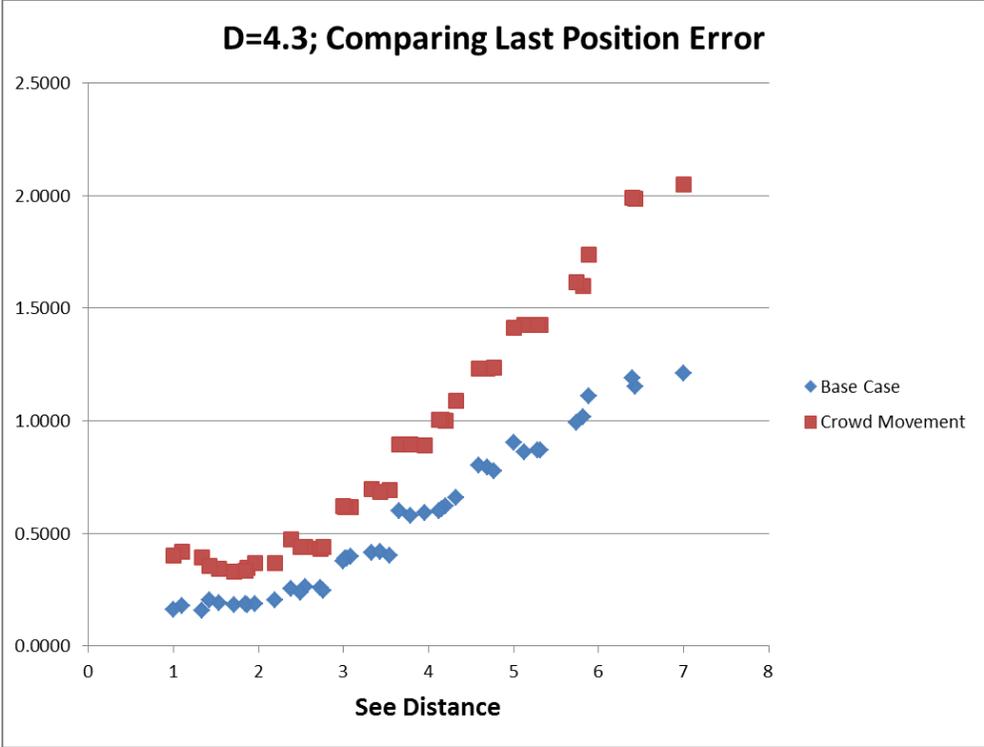


Chart 37: Last Position Error for Crowd Movement – Density of 4.3

Appendix 5.5: Distributions of Efficiencies Charts

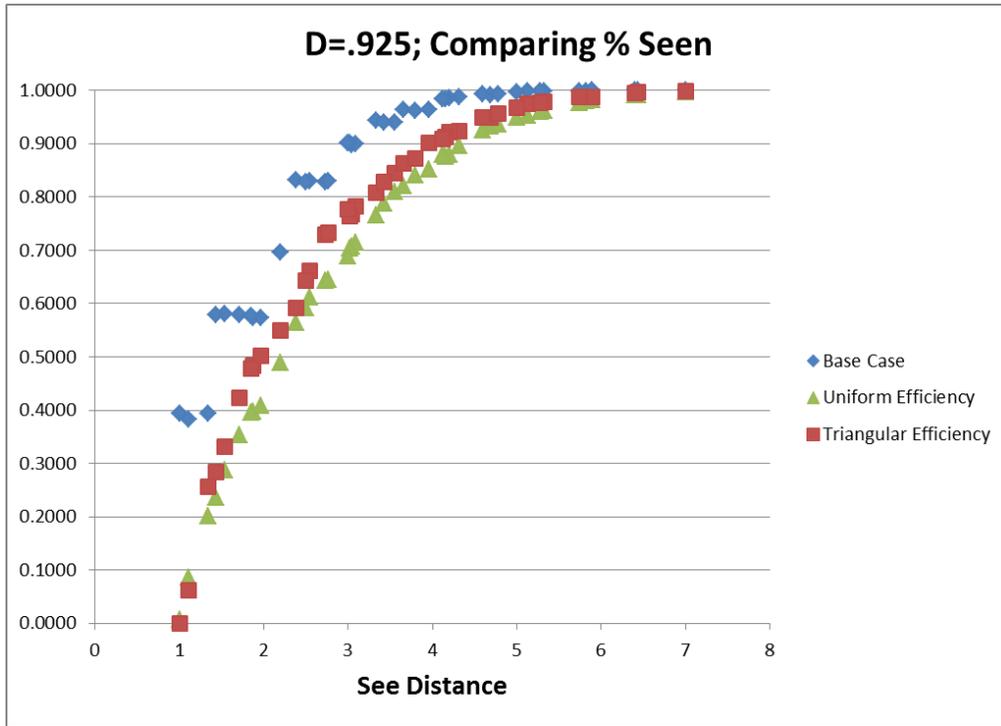


Chart 38: Percent of Time the Source is Seen for Distribution of Efficiencies – Density of .925

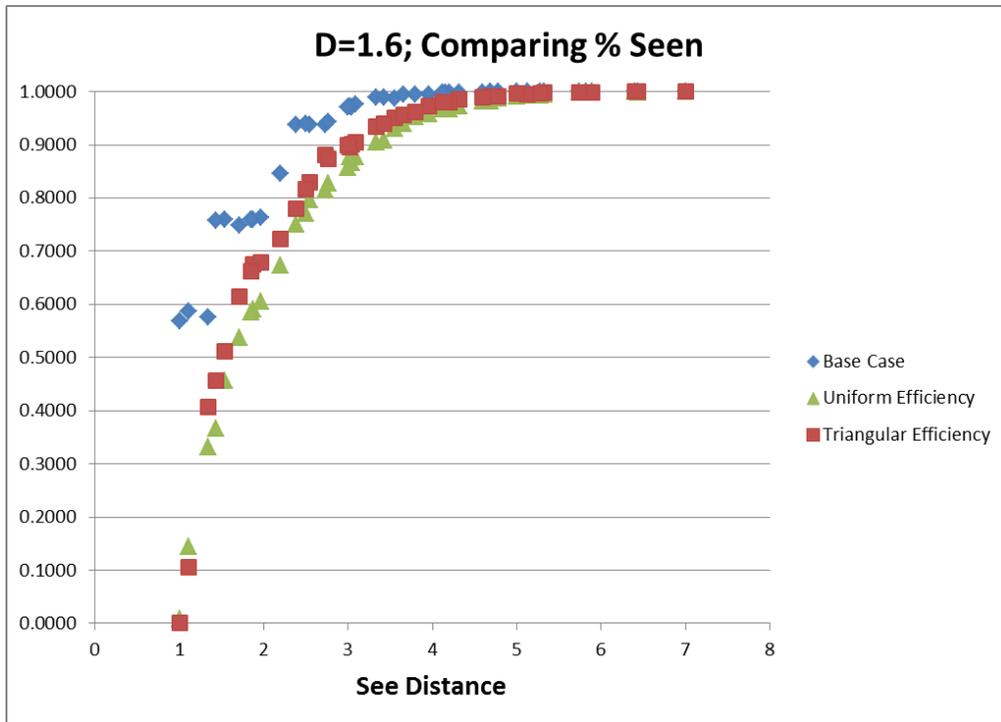


Chart 39: Percent of Time the Source is Seen for Distribution of Efficiencies – Density of 1.6

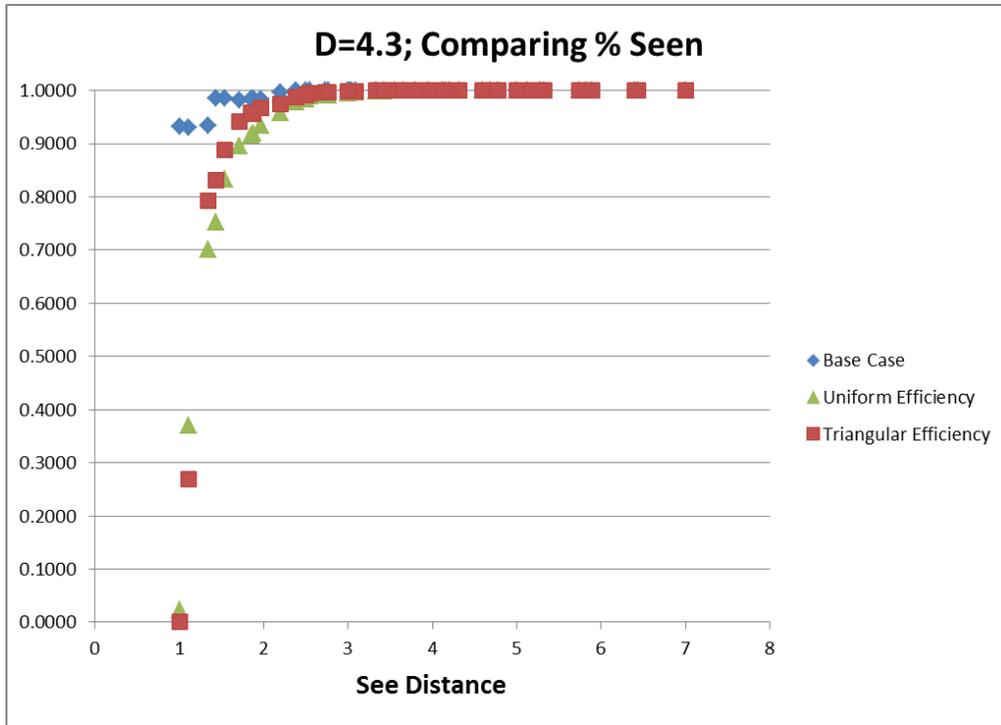


Chart 40: Percent of Time the Source is Seen for Distribution of Efficiencies – Density of 4.3

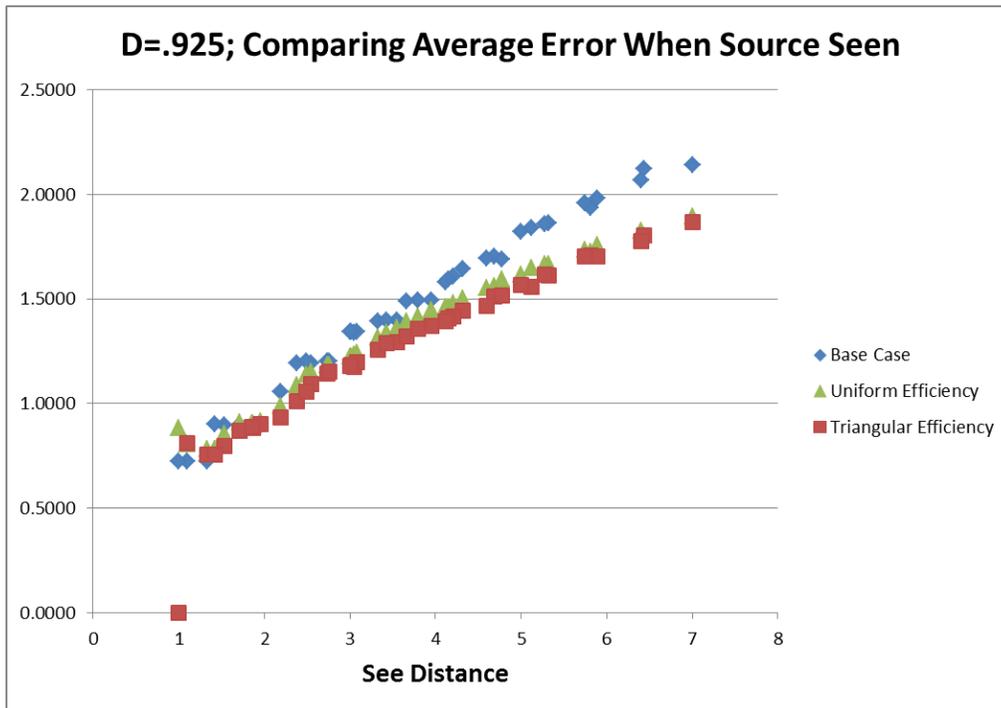


Chart 41: Average Error When Source is Seen for Distribution of Efficiencies – Density of .925

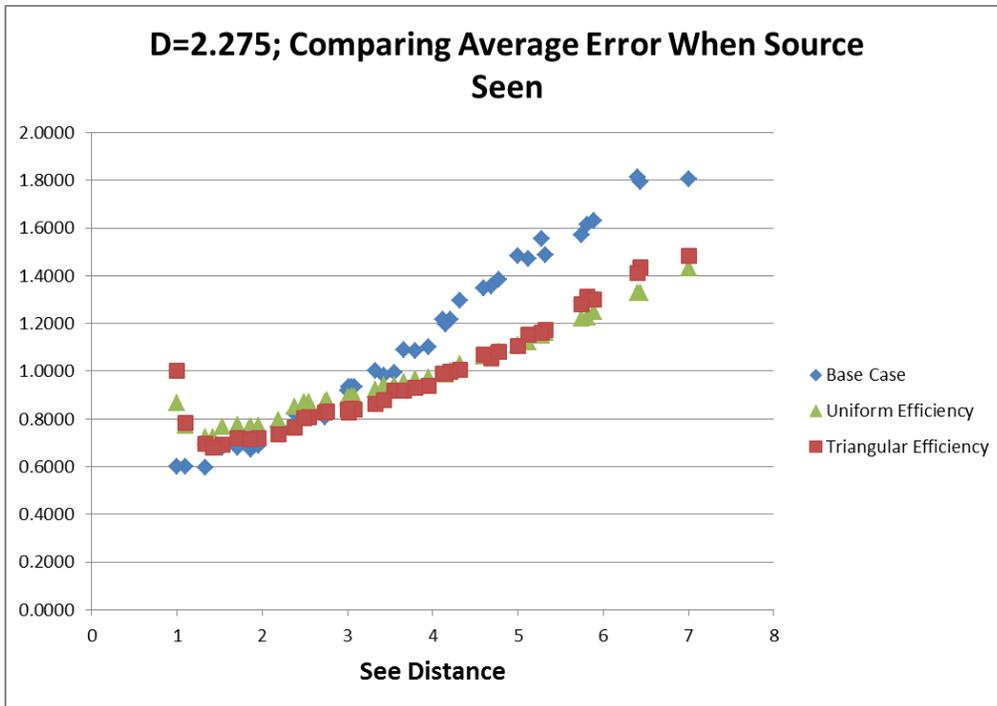


Chart 42: Average Error When Source is Seen for Distribution of Efficiencies – Density of 2.275

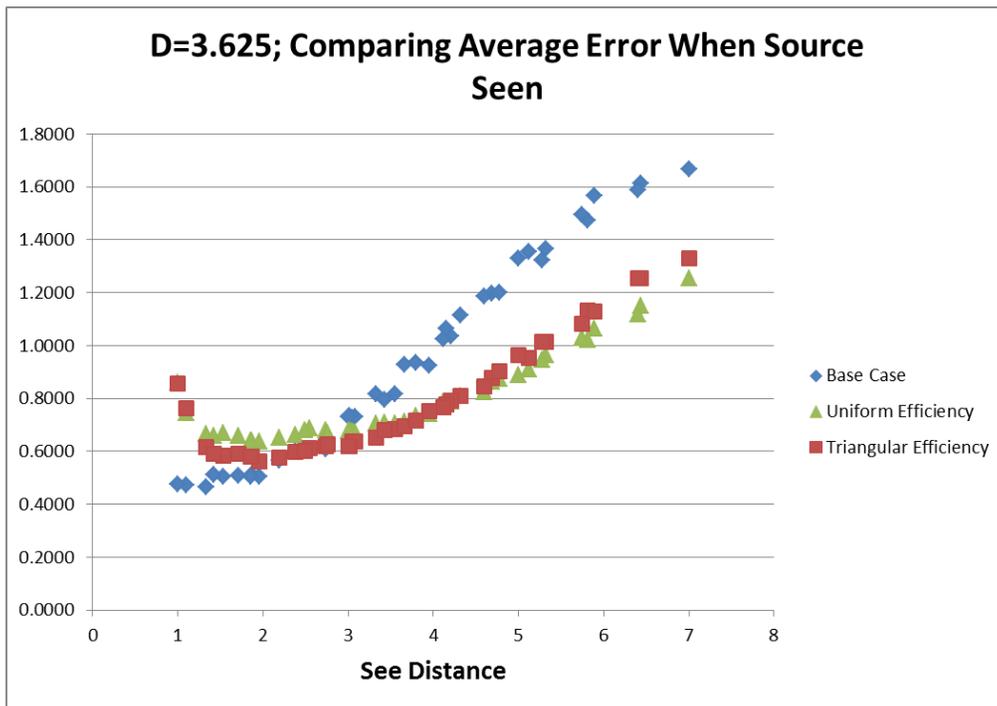


Chart 43: Average Error When Source is Seen for Distribution of Efficiencies – Density of 3.625

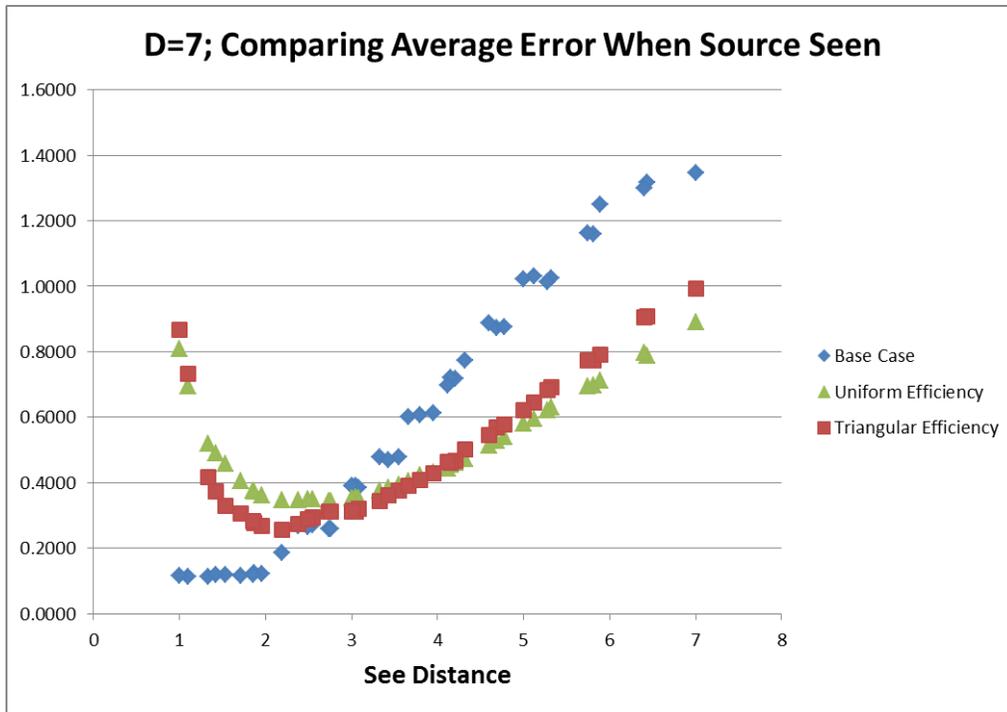


Chart 44: Average Error When Source is Seen for Distribution of Efficiencies – Density of 7

Appendix 5.6: False Detection Analysis Charts

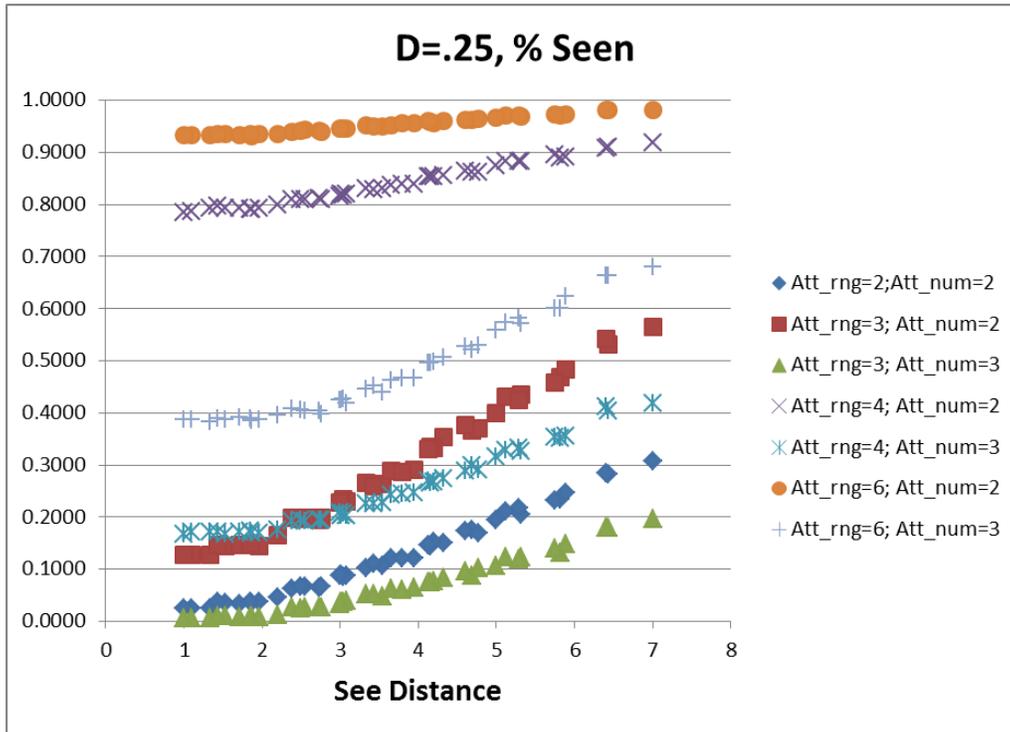


Chart 45: Percent of Time Source is Seen for False Detections – Density of .25

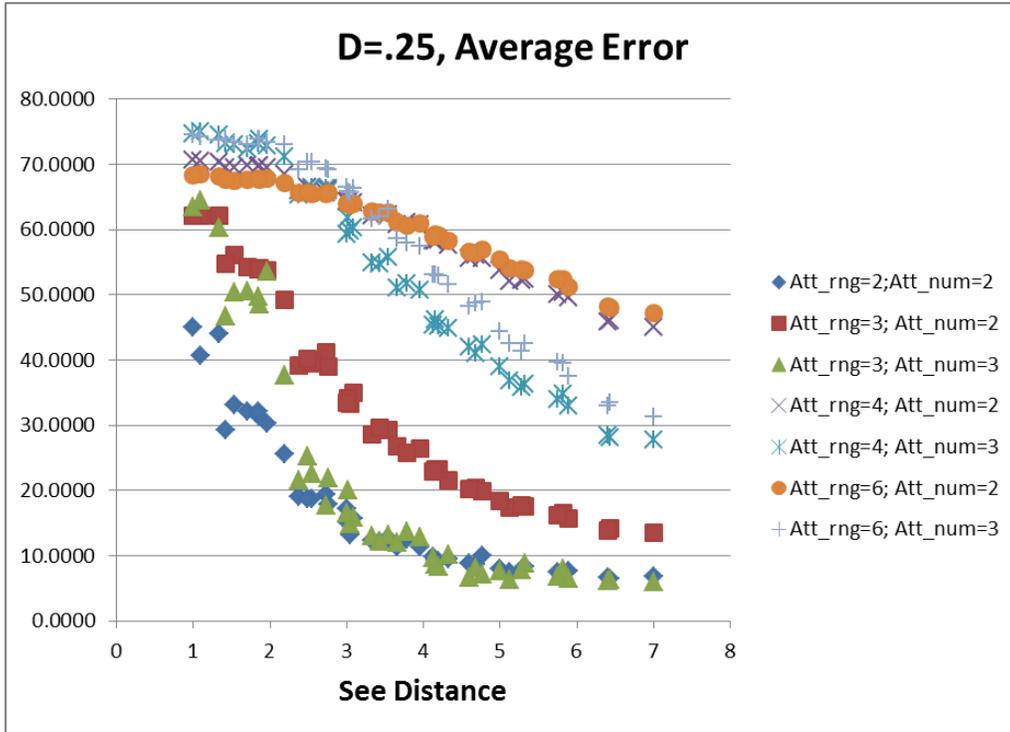


Chart 46: Average Error When Source is Seen for False Detections – Density of .25

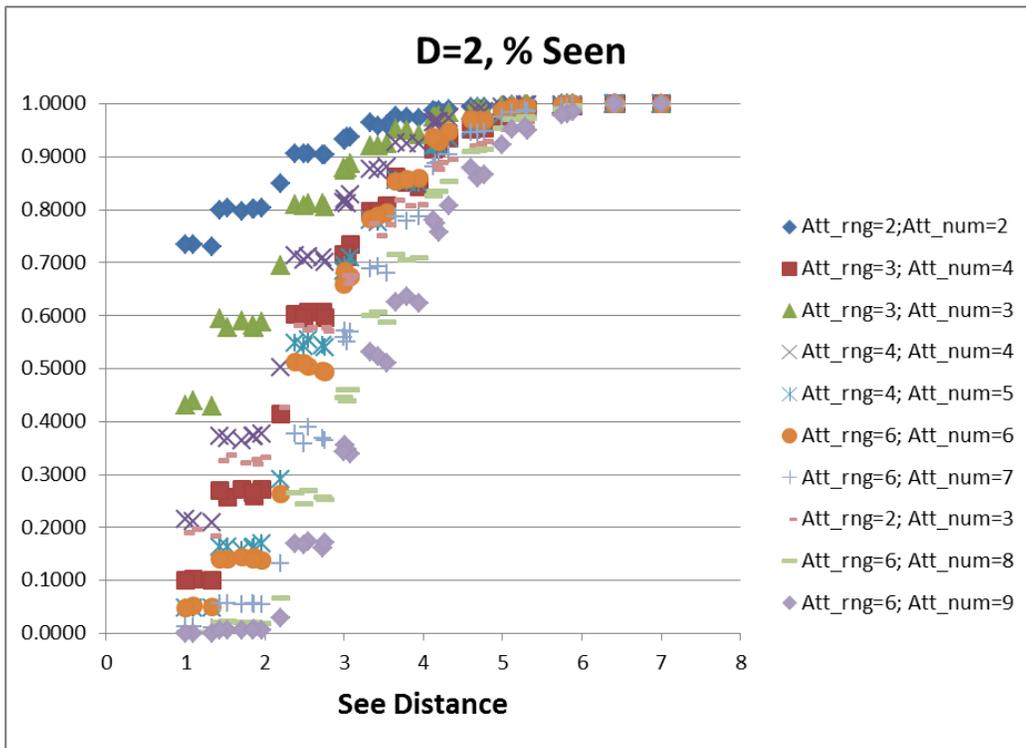


Chart 47: Percent of Time Source is Seen for False Detections – Density of 2

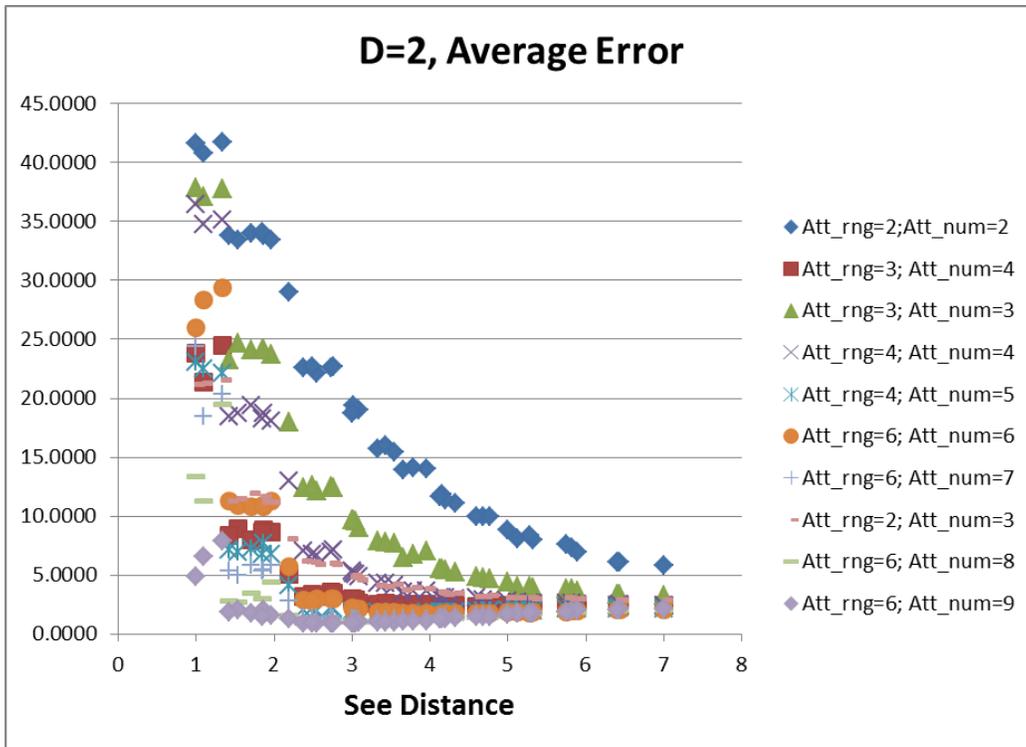


Chart 48: Average Error When Source is Seen for False Detections – Density of .25

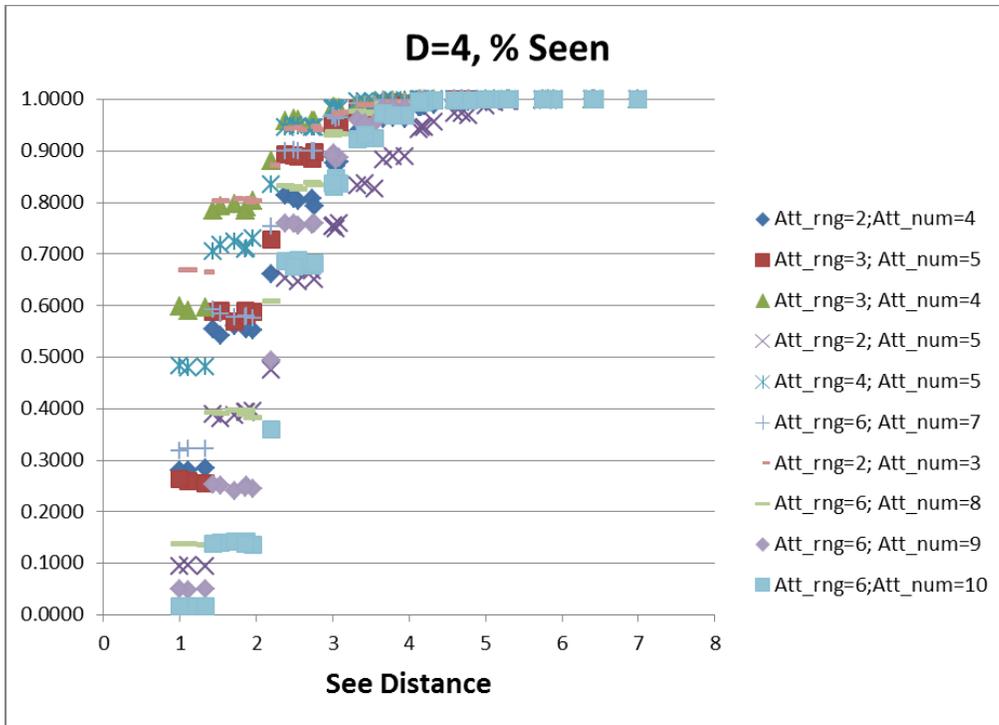


Chart 49: Percent of Time Source is Seen for False Detections – Density of 4

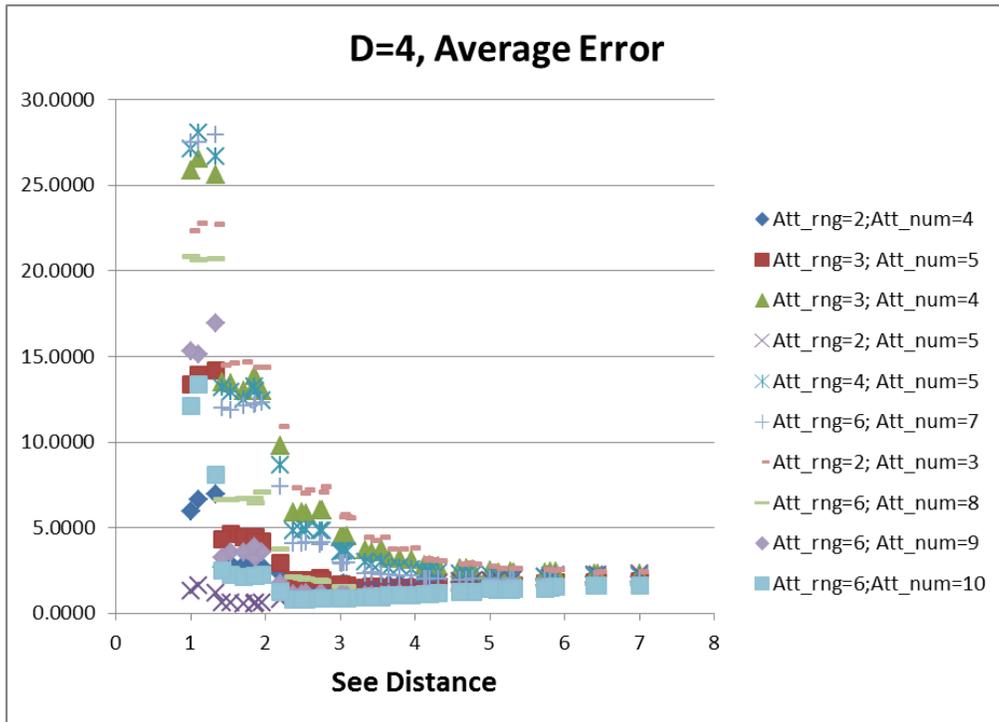


Chart 50: Average Error When Source is Seen for False Detections – Density of 4

Appendix 6: Earned Value Management Chart

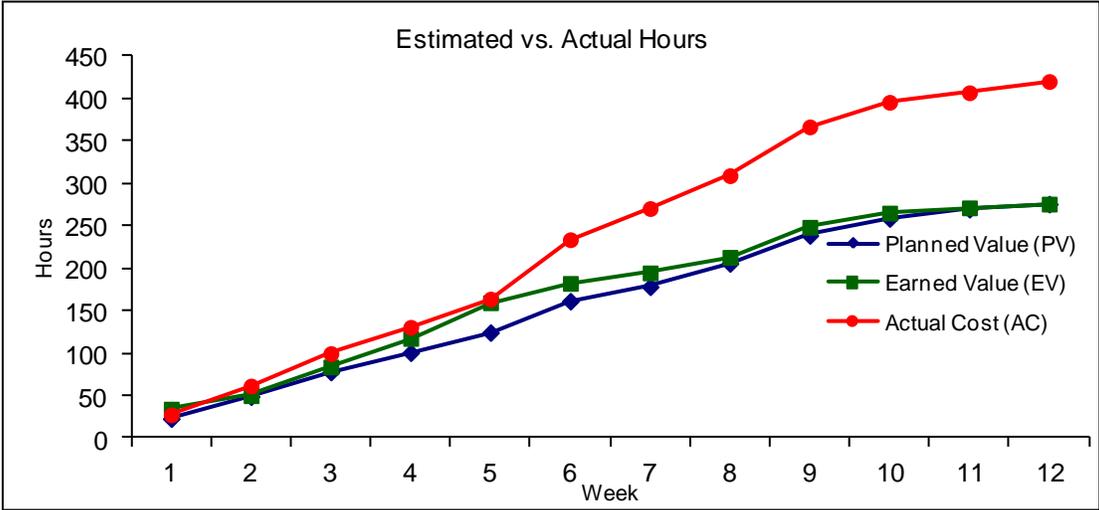


Chart 51: Earned Value Management

Appendix 7: “See Distance” Intervals for Data

Interval	Number of data points in interval for each density
[1, 1.4142)	3
[1.4142, 2)	6
[2, 2.2361)	1
[2.2361, 2.8284)	5
[2.8284, 3)	0
[3, 3.1623)	4
[3.1623, 3.6056)	3
[3.6056, 4)	3
[4, 4.1231)	0
[4.1231, 4.2426)	3
[4.2426, 4.4721)	1
[4.4721, 5)	3
[5, 5.099)	1
[5.099, 5.3851)	3
[5.3851, 5.6569)	0
[5.6569, 5.831)	2
[5.831, 6)	1
[6, 6.0828)	0
[6.0828, 6.3246)	0
[6.3246, 6.4031)	0
[6.4031, 6.7082)	2
[6.7082, 7)	0
[7,)	1

Chart 52: “See Distance” Intervals for Data

Appendix 8:Gantt Chart

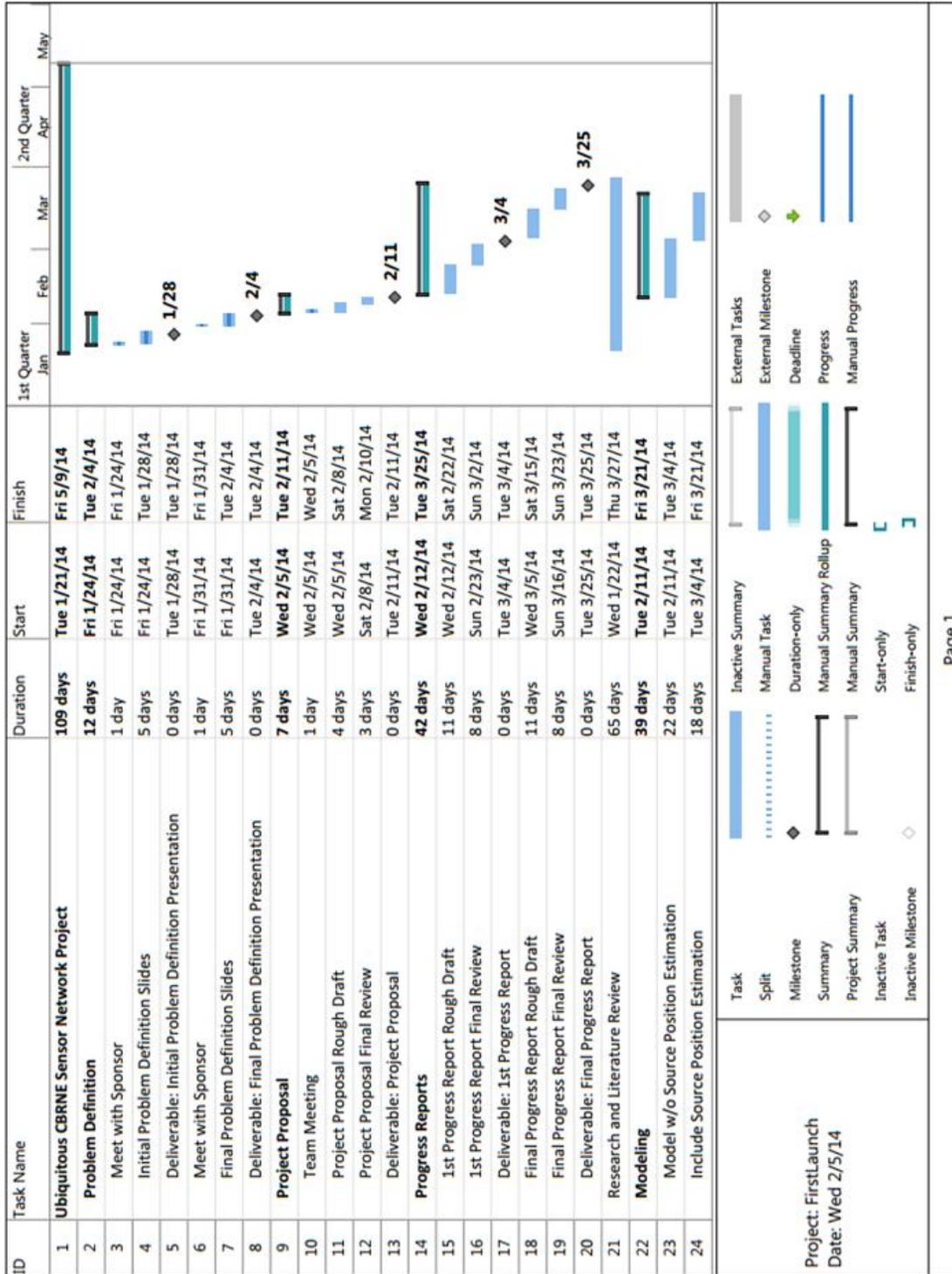


Chart 53: Gantt Chart (1)

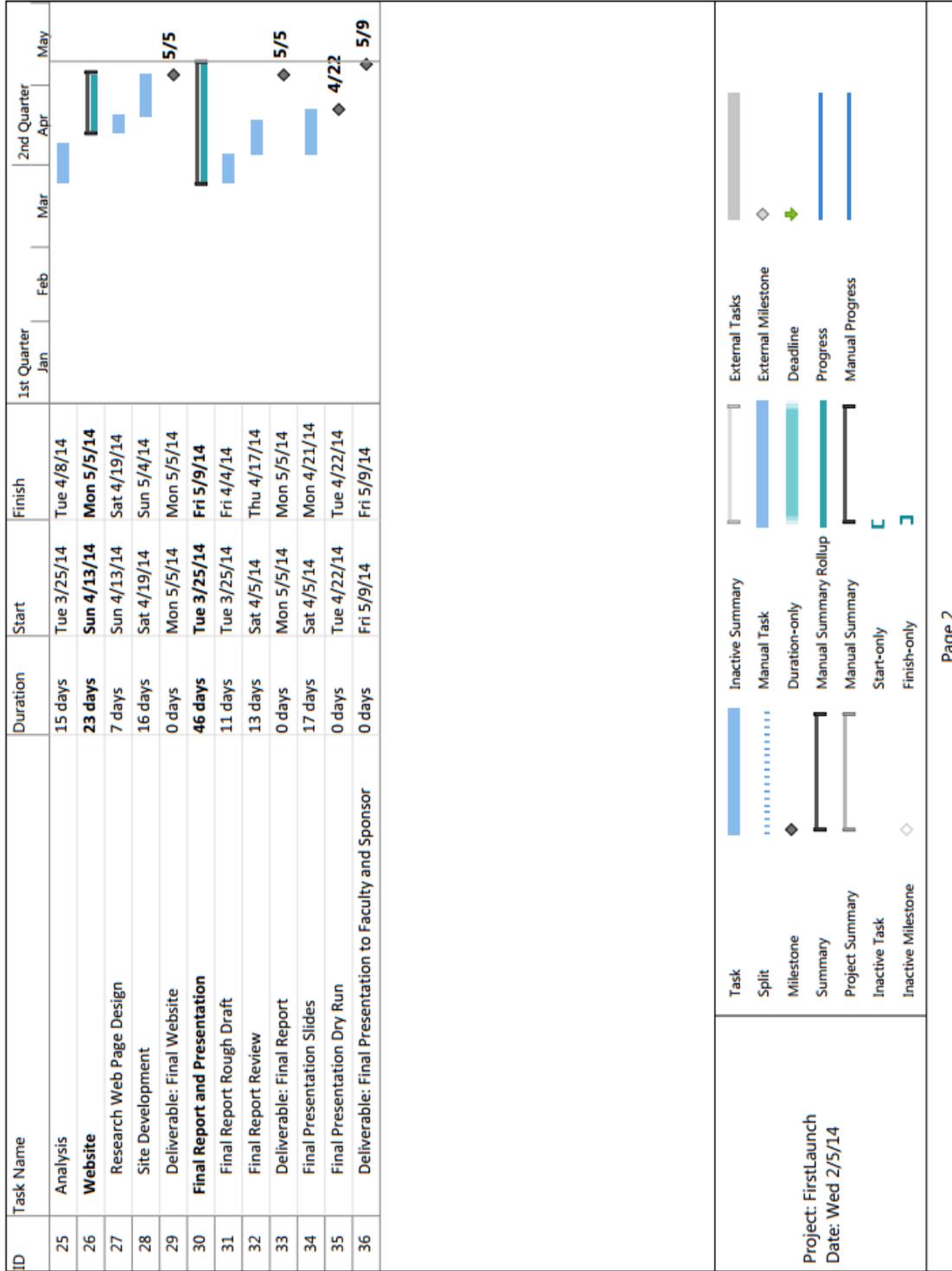


Chart 54: Gantt Chart (2)

Appendix 9: Reference of Key Terms

Attenuation Number - The number of sensors that must be detecting within a the attenuation range for the sensor to be counted towards estimated position.

Attenuation Range - The distance a sensor will check to see if other sensors are also detecting .

Average Error - The average of the distance from the estimated source position to the actual position of the source over a single run.

CBRNE - Chemical, Biological, Radiological, Nuclear, Explosive.

Efficiency - The original design strength of a sensor is considered to be 100% efficiency. Over time, sensors degrade. A sensor with a 50% efficiency has .5 the strength of its original design strength.

False detection - A sensor showing a detection when there is no source present within its “see distance”.

Last Position Error - The difference of the last estimated source position and the source drop off location.

Maximum Error - The maximum of all the distances from the estimated source position to the actual position of the source over a single run.

Maximum Location Error - The maximum distance the source is able to move away from the point that the network last detected it.

Negative Detection - A sensor’s failure to see the source when it should see it.

Percent Seen - The fraction of time a source is seen by the sensor network.

Recur Rate - The individual chance a sensor has a negative or false detection.

See Distance - The largest distance a sensor can be from a source in order for a detection to occur.

Ubiquitous - Anytime, anywhere, anything.

WSN- wireless sensor network; a collection of sensor nodes organized into a cooperative network.