Habitat Monitoring Using Wireless Sensor Networks

A project report submitted in partial fulfillment of the requirements

for the degree of

Bachelor of Technology (Electronics and Communication Engineering)

by

ADYASHA PANDA (109EC0227)
Session 2012-2013
Electronics and Communication Engineering

Under the guidance of

Prof. S. K. Patra
Professor, Department of Electronics and Communication Engineering

National Institute of Technology, Rourkela
Rourkela-769008, ODISHA

National Institute of Technology, Rourkela

Rourkela-769008, Odisha

Declaration

I hereby declare that the Project entitled ***"Habitat Monitoring Using Wireless Sensor Networks"*** submitted in partial fulfillment of the requirements for the degree of Bachelor of Technology in Electronics and Communication Engineering at National Institute of Technology, Rourkela for the academic session 2012 – 2013 is a record of my original work done under Dr. S. K. Patra, Professor, National Institute of Technology, Rourkela. Wherever contributions of others are involved, every endeavor has been made to acknowledge the same with due reference to literature.

ADYASHA PANDA

(109EC0227)

National Institute of Technology, Rourkela

Rourkela-769008, Odisha

Certificate

This is to certify that the Project entitled ***"Habitat Monitoring Using Wireless Sensor Networks"*** submitted by **Adyasha Panda** in partial fulfillment of the requirements for the award of Bachelor **of Technology Degree** in **Electronics and Communication Engineering** Session 2009-2013 at **National Institute of Technology, Rourkela** is a credible and authentic work carried out by her under my supervision and guidance.

Place: NIT, Rourkela

Date:

(**Prof. S.K.Patra**)

Professor, Electronics and Communication Engineering

National Institute of Technology, Rourkela

Acknowledgement

I feel privileged to express my heartiest thanks and gratitude to **Prof. S. K. Patra, Professor, Department of Electronics and Communication Engineering** for his precise and inspiring guidance, valuable and generous suggestions throughout this project. He contributed to a great measure in surmounting all hardships, I faced during the project work. His tolerance in dealing the problems and encouragement were a constant source of inspiration.

I am profusely thankful to **Mr. Trilochan Panigrahi,** for guiding me and helping me in understanding the concepts of source bearing estimation of wireless sensor networks.

I am grateful to Prof S. K. Meher, Head of the Department, Electronics and Communication Engineering, for providing the necessary facilities in the department.

Last but not the least; I would like to thank to all the staff of NIT, ROURKELA those who have directly or indirectly helped me to successfully complete the project.

ADYASHA PANDA

(109EC0227)

Abstract

The deployment of wireless sensor networks in habitat monitoring is gaining importance as the manpower cost is increasing day by day. The positions of the cattle is detected and if detections at successive time intervals indicate that the position of the cattle is hardly changing, there is a chance that the cattle is sick or injured and a warning message is issued to the owner of the farm. The positions have been estimated using the Direction of Arrival estimation by maximum likelihood and MUSIC (MUltiple SIgnal Classification) algorithms. The performance of the system has been evaluated in terms of minimum root mean square error and probability of resolution. The results of direction of arrival have been improvised using the averaging process and the multimodal problem has been optimized using differential evolution.

Since Direction of Arrival estimation gives only the direction and not the precise position, the phase detection of the signals is done to differentiate different positions having the same direction of arrival. Finally analysis is done regarding the movement of cattle. If it is found that they do not move and occupy the same position for a considerably large period of time, warning message is issued to the owner of the farmland.

Table Of Contents

List of Figures

List of Abbreviations

ESPIRIT	Estimation of Signal Parameters Via Rotational Invariance Techniques
ISM	Industrial, Scientific and Medical
MUSIC	MUltiple SIgnal Classification
DOA	Direction of Arrival
CRB	Cramer Rao Bound

Chapter 1. Introduction

1.1. Introduction

The estimation of source bearing (also known as direction of arrival) is one of the challenging problems in wireless sensor networks. Accurate direction of arrival is a major problem for tracking and localization of sources. In fields of radar, sonar, radio-astronomy, underwater surveillance and seismology the direction of arrival is a well-known problem. It can also be applied to the field of habitat monitoring to monitor the health of animals based on their movements in successive intervals of time, where a warning signal can be issued to the owner of the farmland if an animal is found to be stationary for a considerable amount of time. Wireless sensor networks consist of tiny, inexpensive, low-powered sensor nodes which are connected with each other through a wireless link. The major advantage of wireless sensor networks are that it allows for gathered spatially independent observation of certain physical phenomenon and process this observed data in a distributed fashion where the objective is to utilize all sensor observations to estimate certain parameters.

In the recent pasts many methods have been proposed which include estimation of signal parameters via rotational invariance techniques (ESPIRIT). Maximum Likelihood is found out to be one of the best techniques used for source localization problems. But the methods based on maximum likelihood estimation fail to provide global performance.

Evolutionary algorithms like particle swarm optimization, adaptive particle swarm optimization, bacteria foraging optimization methods have been used for maximum likelihood source localization. But these methods failed to achieve good performance at lower signal to noise ratio. So, to obtain the exact Maximum Likelihood solutions, the direction of arrival must be estimated by estimated by optimizing the multimodal problem over a high dimensional problem space. By considering the effectiveness of Differential evolution algorithm as compared to other evolutionary optimization algorithms, a differential algorithm based solution is proposed to compute the maximum likelihood functions and explore the performances of superior performances over the traditional multiple signal classification algorithm.

Frequencies of signals used are in the ISM (Industrial, Scientific and Medical) radio bands and hence do not interfere with the frequency bands used for commercial purposes. After direction of arrival estimations are estimated for successive intervals of time, they are compared with the preceding directions. If directions of arrival are found out to be the same, phase detection of the carrier signal is done to differentiate between different positions on the same angle of arrival. If phase as well as direction of arrival of estimation is found out to be the same for considerable amount of time a warning message is issued to the owner of the farmland that the animal is not moving which might be due to some injury or illness. The architecture of a general habitat monitoring system where there are number of patches to be monitored is as follows:

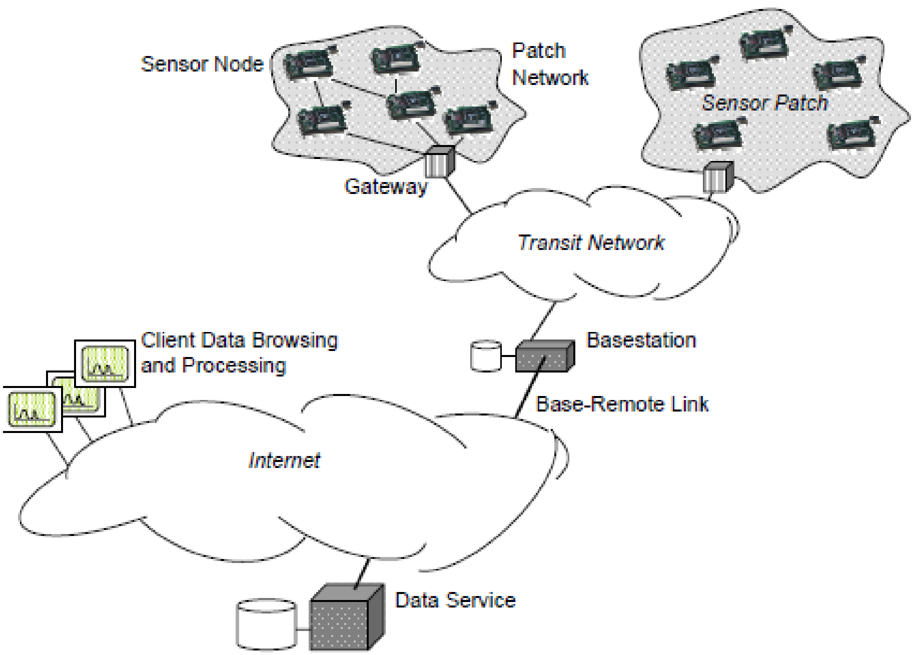


Figure 1:1 : Architecture of a Habitat Monitoring System

1.2. Motivation

As already mentioned one of the most important problems in the field of wireless sensor networks is source localization. A very important application of this is in the field of habitat monitoring which on being implemented cuts down the manpower cost needed to monitor the health of the animals in a farmland. Apart from this, this system is quite robust and reliable. Differential Evolution based Maximum Likelihood estimation provides better results than any of the existing source localization. The performance of this system when evaluated in terms of minimum mean square error and probability of resolution was quite good even at low signal to noise ratios. So when this type of intelligence is incorporated in habitat monitoring system accurate estimations of the positions of animals in a farmland can be done which eliminates the need to monitor them manually and saves a great deal of man-hours.

Chapter 2. Background Theory

2.1. Wireless Sensor Networks

Wireless sensor networks consist of large number of tiny inexpensive nodes which are generally low powered and are connected with each other through a wireless link. Processing Units are present in each sensor node locally. The advantage of such type of sensor nodes is that it allows for independent observation of spatially gathered data and uses this data to estimate certain parameters. These sensor nodes may be arranged as either a uniform linear array or a uniform circular array. They can also be arranged in a random manner. The sensor nodes can be arranged in any of the three following manner.

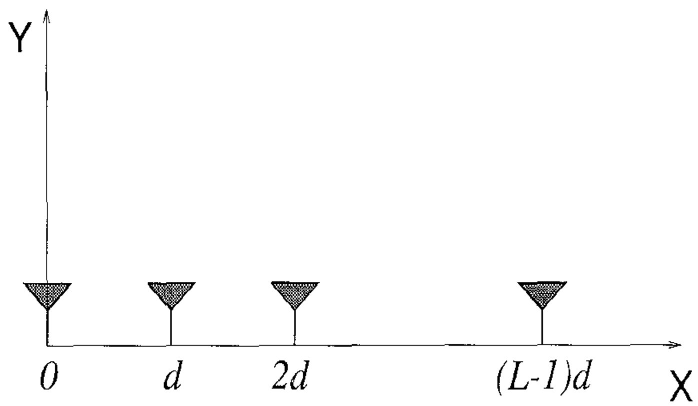
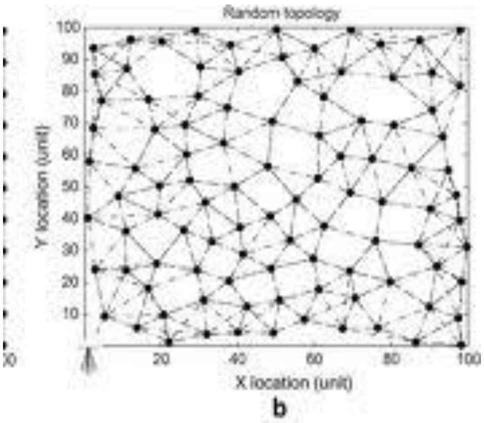


Figure 2:1 : Uniform Linear Array

The major advantages of these wireless sensor networks are that it allows for spatially gathered independent observation of certain physical phenomenon. Since each node possesses a local processor, it can calculate and independently estimate certain parameters. This reduces overhead communication problems. Also, it is a power efficient process. However, localisation of sources is an important problem in wireless sensor networks. In fields of underwater surveillance, radio astronomy tracking of sources is an important issue.



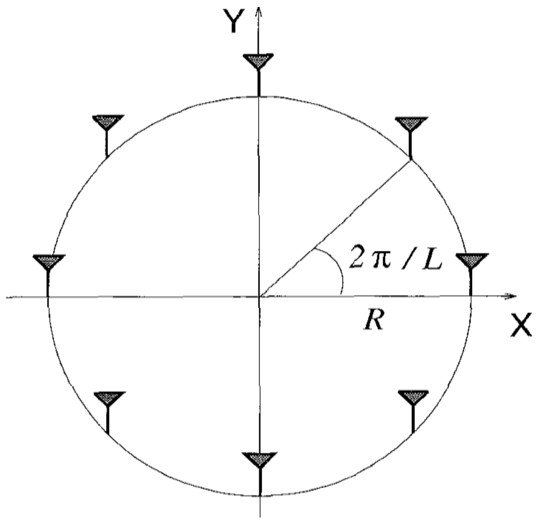


Figure 2:2 : Random Array

Figure 2:3 : Uniform Circular Array

In centralized approach all sensor nodes send their observation to any central processor for parameter estimation. But this centralized approach possesses excess overhead communication problems. A decentralized method has been proposed here. The sensor nodes are assumed to receive information being transmitted in the ISM (Industrial, Medical and Scientific) radio band. These portions of the radio band are reserved internationally for purposes other than communication. Also, these observed data can be processed in a distributive fashion where the objective is to utilize all sensor observations to estimate certain parameters.

2.2. Source Localization

Finding the location of sources emitting the signals which are being received by the sensors is quite challenging because the signals are of unknown amplitudes, unknown phase and are also corrupted by noise. Several techniques have been used for Direction of Arrival estimation-Capons beam forming model, correlation, maximum-likelihood, MUSIC (MUltiple SIgnal Classification), ESPIRIT, and matrix pencil. In the fields of radar, sonar, radio-astronomy, underwater surveillance and seismology the DOA (Direction of Arrival) estimation is a well-known problem.

Capon's beam forming method being one of the most primitive one, has a number of disadvantages which have been overcome by the new techniques.

Of all these techniques listed above maximum likelihood is one of the best techniques used in source localization problems. However, due to very high computational load of the multivariate nonlinear optimization problem required in maximum likelihood estimation several high resolution suboptimal techniques have been suggested which include the MUSIC (MUltiple SIgnal Classification). Basics like how to model the signal being emitted by the source, how to account for the propagation delay for the successive sensors, how to model the noise being added are understood in the next chapter where the details of the modelling are presented.

DOA(Direction of Arrival) being one of the estimated parameters, the Cramer Rao Bound (CRB) gives the minimum variance of this parameter that is it sets the lower bound for this parameter.

2.2.1. Maximum Likelihood Estimation

The main focus of this technique is to maximize the probability that the signal came from a particular angle. Maximum likelihood methods can be broadly divided into two types:

1. Stochastic Maximum Likelihood Technique
2. Deterministic Maximum Likelihood Technique.

The difference between the two methods lies in the way the signals are modeled. In stochastic maximum likelihood technique the signals have been modeled as Gaussian random processes while in deterministic maximum likelihood technique these have been modeled as deterministic but unknown and are estimated together with the direction of arrival. Maximum Likelihood Estimators are in most cases very consistent and quite often, a variable change can provide an unbiased estimator which the most desirable results.

2.2.1.1. Stochastic Maximum Likelihood Technique

One of the reasonable models used for maximum likelihood, this models signal waveforms as Gaussian random processes. However, it is also applicable even if the data is not modeled as Gaussian. For large samples the precision to which the signals have been estimated is found to depend on the second order statistics like powers or the correlations of the signal waveforms. The fact that the modeling of signal waveforms is done as Gaussian only to obtain a tractable Maximum Likelihood. Let the signal waveforms have zero mean and have second order statistics.

$$E\big(s(t)s^H(s)\big) = P(s,t); \qquad (2.1)$$

$$E\big(s(t)s^T(s)\big) = 0; \qquad (2.2)$$

The above equations lead to an observation vector $x(t)$ to be a white, zero-mean, circularly symmetric vector Gaussian random vector with the covariance matrix

$$A(\Phi)PA^H(\Phi) + \sigma^2 I = UAU^H \qquad (2.3)$$

The set of parameters modeled are different from the deterministic model used. The function is now shown to depend on Φ, P and σ^2.The negative log likelihood function is found to be proportional to :

$$(1/N)\sum_{t=1}^{N} \parallel \Pi_A x(t) \parallel = Tr\{\Pi_A R\}; \qquad (2.4)$$

Although this is a non-linear function this criterion allows for explicit separation of some of the parameters. For a fixed value of Φ, the minimum value of the function can be found to be:

$$\sigma_{SML}^2(\Phi) = \frac{1}{(L-M)Tr\{\Pi_A R\}} \qquad (2.5)$$

$$P_{SML}(\Phi) = A(R - \sigma_{SML}^2(\Phi)I)A^H \qquad (2.6)$$

When these are substituted to the above equations, a compact form is obtained.

$$\Phi_{SML} = \arg\left(min_{\Phi}\ log \parallel AP_{SML}(\Phi)A^{H} + \ \sigma^{2}_{SML}(\Phi)I)A^{H} \ \parallel\right); \qquad (2.7)$$

The determinant is termed as the generalized variance in the statistical literature. It measures the volume of a confidence interval for the data vector. The observation with the lowest cost and in harmony with the Maximum Likelihood principle is considered. The above function is a highly non-linear function of its argument Φ.

Stochastic maximum likelihood techniques work better in case of larger sample than the deterministic likelihood procedures. This property holds good regardless of the property of the signals being used. The signals need not be Gaussian in particular. For Gaussian signals the stochastic maximum likelihood techniques attain the Cramer Rao lower bound on the estimated error variance. This follows from the general theory of maximum likelihood estimation. All unknowns in the stochastic models are finally estimated consistently. This is entirely in contrast to the deterministic model for which the number of signal waveform parameters s(t) grows without bound as the sample size increases which implies that they cannot be consistently estimated.

2.2.1.2. Deterministic Maximum Likelihood Technique

Noise can be assumed to be emanating from a large number of independent uncorrelated noise sources. But the same cannot be considered for signals emitted from the sources. Noise can be modeled as a stationary Gaussian white random process with the signal waveforms being deterministic and unknown. The carrier frequencies are known. Assuming spatially white and circularly symmetric noise, the second order moments of noise take the form:

$$E\{n(t)n^{H}(s)\} = \sigma^{2}I;$$
$$E\{n(t)n^{T}(s)\} = 0;$$

As a result of the statistical observation vector x(t) is also circularly symmetric. In addition it is also temporally white Gaussian random process, with mean $A(\Phi)s(t)$

and covariance matrix $\sigma^2 I$.The maximum likelihood function is basically a probability density function of all the observation given the unknown parameters. The probability distribution function is the complex L-variate Gaussian:

$$\frac{1}{\pi\sigma^{2^L}} e^{-\|x(t)-As(t)\|^2/\sigma^2} \tag{2.8}$$

where ‖.‖ denotes the Euclidean norm and the argument of $A(\Phi)$ being dropped for convenience. The likelihood function is obtained as:

$$L_{DML}\,(\Phi, s(t), \sigma^2) = \prod_{t=1}^{N} \frac{1}{\pi\sigma^{2^L}} e^{-\|x(t)-As(t)\|^2/\sigma^2} \tag{2.9}$$

The unknown parameters indicated above in the likelihood function are the signal parameters Φ, the signal waveforms $s(t)$ and the noise variance σ^2. The deterministic maximum likelihood estimates of the unknowns are found out as the maximizing arguments of $L(\Phi, s(t), \sigma^2)$.

The logic behind this is that the probability of observations for these values be as large as possible. For reducing the complexity the negative logarithm of the function is minimized. So normalizing by N and ignoring the parameter independent L log π-term we get

$$L_{DML}\,(\Phi, s(t), \sigma^2) = L\ log\sigma^2 + (\frac{1}{\sigma^2 N}) \sum_{n=1}^{N} \|\ x(t) - As(t)\ \|^2 \tag{2.10}$$

whose minimizing arguments give the deterministic maximum likelihood estimates. Explicit minima with respect to the variance σ^2 and $s(t)$ are given by:

$$\sigma^2 = \left(\frac{1}{L}\right) Tr\{\Pi\, R\}; \tag{2.11}$$

where R is the sample covariance matrix

and Π is the orthogonal projector onto null space of A^H .

$$\Phi_{DML} = \arg\{min_{\Phi} Tr\{\Pi \text{ R}\}\} \qquad (2.12)$$

The interpretation of the above equations is that the measurements $x(t)$ are projected into a model subspace which is orthogonal to all anticipated signal components and a power measurement is evaluated.

The energy should be the smallest when the projector has removed all the true signal components. Since only a finite number of noisy samples is available, the energy is not perfectly measured and then Φ will deviate from Φ_{DML} .However, the error will converge to zero as the number of samples considered is increased ton infinity.

This deterministic likelihood technique is valid even for correlated or even coherent signals. To calculate the deterministic likelihood which is non-linear, numerical methods are used to solve it. A Gaussian Newton technique is used which rapidly to the minimum. The only disadvantage is that to obtain a sufficient accurate estimate large time would have been required.

Deterministic Maximum likelihood technique gives satisfactory results even in presence of noise.It is definitely a more robust algorithm as compared to stochastic maximum likelihood approach.Thus ,for the case of habitat monitoring this variant of maximum likelihood technique was chosen.

When this was simulated using MATLAB the following results were obtained for sources at 160 and 40 degrees respectively.The points of minima indicate the location of sources.The plot is basically as shown below with the source location being given by the minima of the 2-D surface plot.It is basically a probability distribution plot.

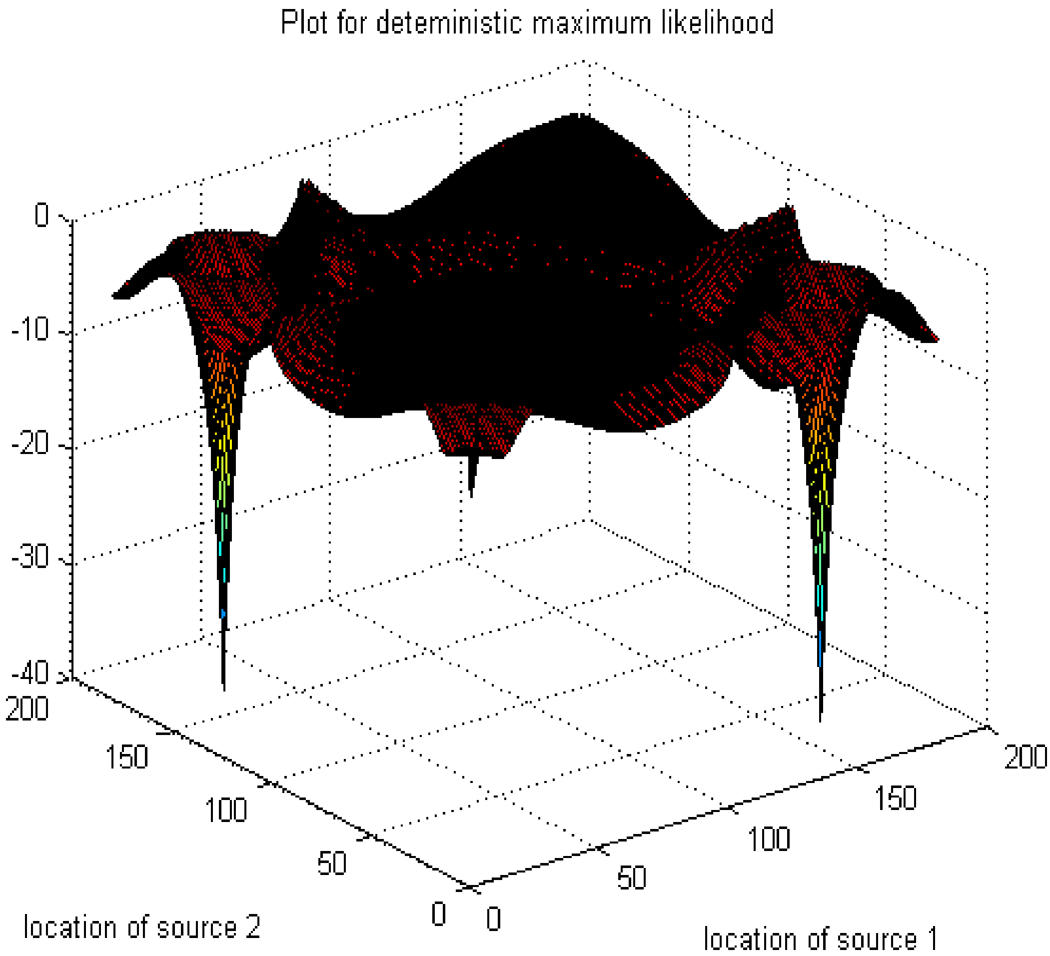


Figure 2:4 : 2D Surface Plot for Deterministic Maximum Likelihood Technique

2.2.2. Music Algorithm

As compared to Capon's beam forming method, this MUSIC algorithm provides greater resolution. The peaks at the source locations in the graph are pronounced. MUltiple SIgnal Classification technique is a sub-space based method. The received signal is calculated from the emitted signals and the steering matrix with white noise added to it. The covariance matrix of the received signal is calculated .Let it be denoted by R.The spectral decomposition of the covariance matrix can be expressed as :

$R = APA^H + \sigma^2 I = U_S \Lambda_s U_S^H + \sigma^2 U_n U_n^H;$ (2.13)

where it is assumed that APA^H is of full rank and

Λ_S is the diagonal matrix having the largest M Eigen values.

(with M being the number of sources)

Since the eigenvectors in U_n(the noise eigenvectors) are known to be orthogonal to A, we have

$$U_n^H a(\Phi) = 0, \quad \Phi \in \{\Phi_1, \Phi_{2,\ldots\ldots\ldots\ldots}, \Phi_M\}$$

The array is assumed to be unambiguous to allow the direction of arrival to be unique. The steering vectors which correspond to distinct direction of arrival should form a linearly independent set that is if

$$\{a(\eta_1), \ldots\ldots\ldots\ldots, a(\eta_L)\}$$

(where L denotes the number of sensor nodes) ,then P has full rank and APA^H also has full rank.

So from the above equation $\Phi_1, \ldots.., \Phi_M$ are the solutions for the direction of arrival. Thus after obtaining the covariance matrix it is separated into the signal eigenvectors and the noise eigenvectors. The orthogonal projector into the noise subspace is estimated as

$$\Pi = U_n U_n^H;$$

Finally the MUSIC spectrum is defined as:

$P_M(\Phi) = \frac{a^H(\Phi) a(\Phi)}{a^H(\Phi) \Pi a(\Phi)}$ (2.14)

The peaks of the MUSIC spectrum give the location of the sources.

The improvement in performance of MUSIC was so significant that it replaced all the existing methods and was the most favored choice for direction of arrival estimation. In addition, MUSIC provides statistically consistent estimates. The peaks at the source locations are significantly pronounced than in the Capon's beam forming method.

Though MUSIC spectrum does not in reality represent the spectral components, its limitation is that it cannot resolve between sources placed at small separations from each other. Moreover its performance in presence of noise is not at all satisfactory. This is an important issue and cannot be overlooked. In situations requiring high accuracy in the direction of arrival estimation, the mitigation of these issues are very important. MUSIC spectrum plots for two cases have been shown below.

For the case where the sources are separated by 10 degrees, the peaks are clearly visible .But for the case where the sources are separated by 5 degrees; the peaks are not clearly visible. Only one peak is formed. This indicates that the MUSIC algorithm cannot locate sources even if they are 5 degrees apart from each other.

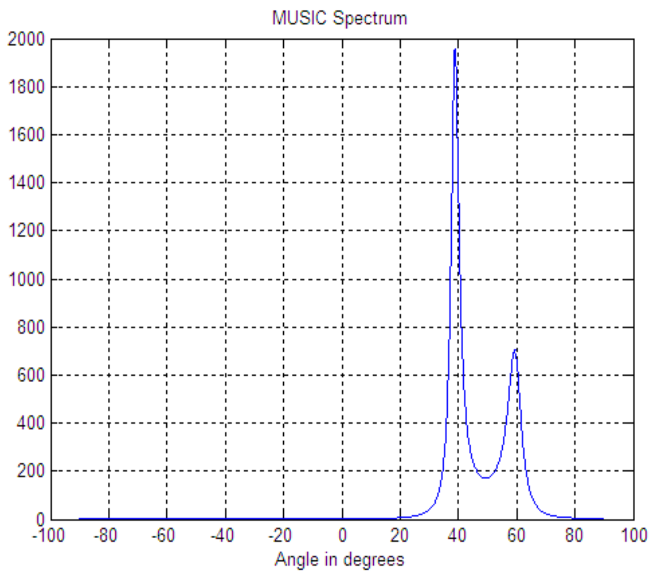


Figure 2:5 : MUSIC Spectrum for Sources at 40 and 60 degrees

MUSIC spectrum for sources separated by 5 degrees or less is shown below. It is clearly visible that the only one peak is formed.

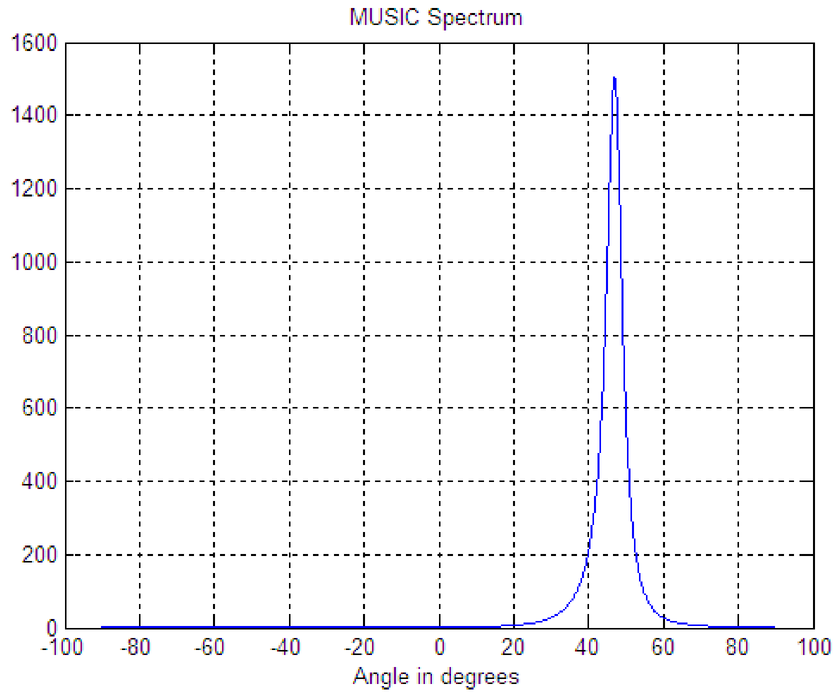


Figure 2:6 : MUSIC Spectrum for Sources at 45 and 50 degrees.

2.3. Differential Evolution

Differential Evolution was developed by Price and Storn in the year 1995.Being a population based stochastic search algorithm this is faster than the other variants of genetic algorithm. It is widely used for optimization in the fields of Mechanical engineering and communication Systems.

Differential Evolution has been considered as a novel evolutionary computation technique. In the recent past Differential evolution has gained importance among other evolutionary computations due to a number of factors which may be like easy implementation, simple concept and quick convergence.

The difference between Differential Evolution and other evolutionary algorithm is that it uses a less stochastic and greedy algorithm in its approach [2]. In an N-dimensional search space that is in a population of potential solutions to an optimization problem, a fixed number of vectors are initialized randomly and then after that evolution of new population takes place over the time to search the space and locate the minima of the function.

The unique thing about differential evolution is that it combines simple arithmetical operators along with the classical operators like recombination, selection and mutation to evolve from a starting population which was randomly generated to the final solution.

Differential evolution is a scheme with the fundamental idea of generating the trial parameter vectors. Then differential evolution mutates the vectors by adding to them the random vector differentials. Then fitness of the trial and the target vector is evaluated. The one with the highest fitness is considered for the next generation.

The basic steps followed are given briefly:

1. **Parameter Set-up**:

   The parameters like population size, maximum and minimum values of the optimization variables or boundary constraints of the optimization variables, the mutation factor or the constant factor, the crossover rate and the number of stopping generations are chosen and set.

2. **Initialization of Population**

   Generation is set to zero. The population is initialized with random values which are generated according to a uniform probability distribution. These are chosen within given bounds. These bounds may be of any type depending on the constraints imposed on the parameter chosen .

3. Evaluation of Population

Evaluation of fitness value of each individual is done and if fitness satisfies the predefined condition the result is saved otherwise the next step is followed.

4. **Mutation Operation**

In mutation operation a vector differential is added to a population vector of the individuals. For a target vector $x_{i,g}$ a mutant vector is produced according to the relation:

$$v_{i,g} = x_{r1,g} + F\left(x_{r2,g} - x_{r3,g}\right); \qquad (2.15)$$

where F is the mutation factor or the constant factor.

5. Recombination Operation

Recombination operation is employed to generate a trial vector by replacing certain parameters of the target vector with the corresponding parameters of the mutant vector that is the randomly generated vector. There are many methods of recombination like binomial recombination and exponential recombination.

$$t_{j,i,g} = v_{j,i,g} \; if \; \left(rand_j \le C\right) or \; j = j_{rand} \qquad (2.16)$$

$$= x_{j,i,g} \qquad otherwise$$

6. Selection Operation

The process of producing better offspring is called selection. If the trial vector $t_{i,g}$ has better fitness value then it replaces the target vector in the next generation, otherwise the target vector retains it place in the population. Once the new population is installed the processes of mutation, recombination and selection is replaced until optimum is found out. Selection operation is the link between the present and the future generations.

The various steps are pictorially represented in the following block diagram:

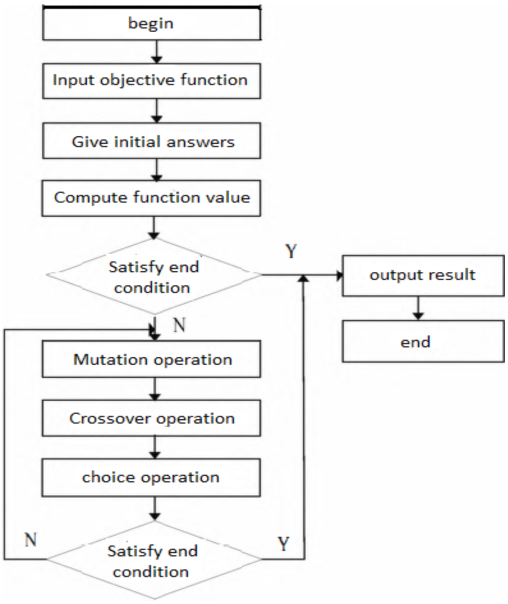


Figure 2:7 : Block Diagram for Differential Evolution

The more the number of generations, the more the computation time and the greater is the precision. Greater precision is desirable whereas higher computation time is undesirable. A tradeoff between the two has to be achieved.

2.4. Phase Detection

Phase detection is done additionally to ensure that the animals in the farmland are actually moving or not.

Direction of arrival estimation is done every five hours to ensure the movement of the animal. Same direction of arrival for consecutive intervals indicates that the signals emitted by the sources attached to the body of the animals emit signals which arrive in the same angle. Such a scenario has been illustrated below:

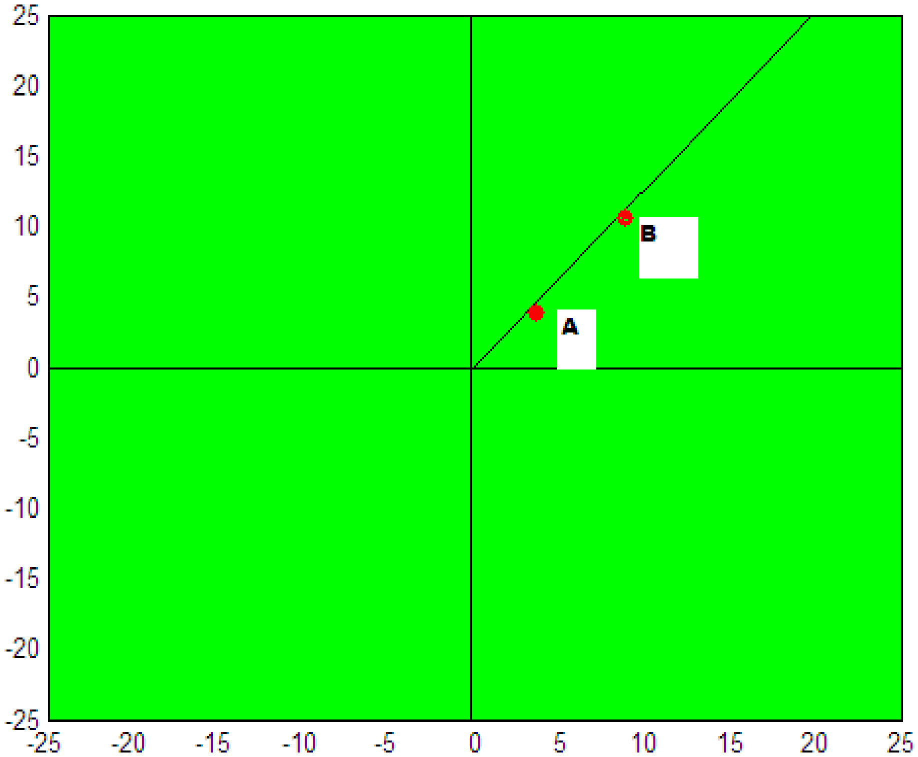


Figure 2:8 : Figure depicting Animals on Same of Arrival but diff. Positions

So the same direction of arrival estimations for two or more consecutive locations may point to two facts:

1. The animal is not moving
2. The animal is moving along the same direction of arrival

But the fact that there could be many positions on a given direction of arrival necessitates the need for some other parameter estimation that could be different for different positions on the same direction of arrival. One such parameter is the phase of the carrier signal.

The received signals are assumed to have frequency in the ISM bands so that they do not interfere or disrupt vital communication processes taking place. The frequency range of the ISM band selected for the signals is 122-123 MHz. Thus the carrier frequency is expected to be double of the signal frequency. So carrier frequency is given by:

$$f_c = 2 \times (122 - 123) \quad kHz$$
$$= \ 244 - 246 \, kHz$$

Carrier phase estimation is an important issue in the field of communication systems as it is essential for correct synchronization at the receiver end.

There are basically two criteria which can are widely used for signal parameter estimation:

1. Maximum Likelihood criterion
2. Maximum a posteriori probability criterion

The received signal received may be modeled as a delayed version of the input with some noise being added to it.

$$r(t) = s(t - \tau) + \ n(t); \qquad (2.17)$$

where τ is the propagation delay

n (t) is the noise introduced

$$s(t) = Re\left[s_l(t)e^{j2\pi f_c t}\right]; \qquad (2.18)$$

where $s_l(t)$ is the equivalent low-pass signal.

The received signal may be expressed as

$$r(t) = Re\{[s_l(t - \tau)e^{j\Phi} + \ z(t)] \quad e^{j2\pi f_c t} \}; \qquad (2.19)$$

where the carrier phase Φ due to propagation delay τ is given by $\Phi = -2\pi f_c \tau$

The carrier phase is not only dependent on the time delay. The precision to which the phase can be estimated is given by the symbol period T. Generally the error is usually a small fraction of the symbol period T. But a small error in τ can cause a large phase error because f_c is usually large.

In Maximum a posteriori probability the phase to be estimated that is Φ is modeled as random and characterized by probability density function while in maximum likelihood it is assumed to be random but deterministic.

There are two approaches for synchronization of carrier at the receiver:

1. One way is to multiplex in frequency a pilot signal which allows the receiver to extract and synchronize its local oscillator's frequency and phase to that of the received signal during transmission of an un modulated carrier component with the information bearing signal generally a phase locked loop is employed to track and acquire the carrier component with the phase locked loop so designed that it has a narrow band width which is not affected by presence of spectral components from the information bearing signal.

2. This approach is more prevalent in practice where the carrier estimate is derived directly from the modulated signal the advantage of this process is that the total transmitter power is allocated for transmission of information bearing signal .In cases of carrier recovery this process is generally followed.

Chapter 3. Modeling & Design

3.1. Basics of Modeling

For simplicity smaller number of sources is assumed at a time. Since the objective is to estimate the direction at which signals impinge on the sensor arrays, which may be linearly, circularly or randomly arranged the sources are assumed to be at certain angles from the sensor array. After simulation applying specific techniques that is calculating the received signal from the emitted signals and performing certain mathematical operations on it, the angles of arrival are found it. These are compared with the angles we had assumed the sources to be. If the values coincide it implies that the direction of arrival estimation is accurate.

3.1.1. Calculation of Emitted Signals

The sources are attached to the body of the animals in the farmland. They are assumed to emit binary phase shift keying (bpsk) signals in the ISM (Industrial, Scientific and Medical) radio bands. Let the number of sources considered at a time be R and the number of sensors used be N. Let the number of snapshots taken be L. Since bpsk signals are antipodal signals they are generated in MATLAB by rounding off a random number to its nearest integer, multiplying it by 2 and then subtracting 1 from it.

$$sig = round\big(rand(R, L)\big) \times 2 - 1; \qquad (3.1)$$

The frequency of the signals is in the MHz range .They have been chosen such that the wavelengths corresponding to these frequencies are a very small fraction of a centimeter. Generally frequency is in the range 122-123 GHz. Generally the length of the sources is comparable to the wavelength of the signal being emitted. The wavelength corresponding to 122MHz is given by:

$$\lambda = \frac{c}{f} = (3 \times 10^8)/(122 \times 10^9)\ in\ metres$$

$$= 0.00246\ metres \quad = 0.246\ centimetres$$

The wavelength of the signal is calculated to be 0.246 centimeters. The length of the source which will emit signals of such wavelength is comparable to the wavelength.

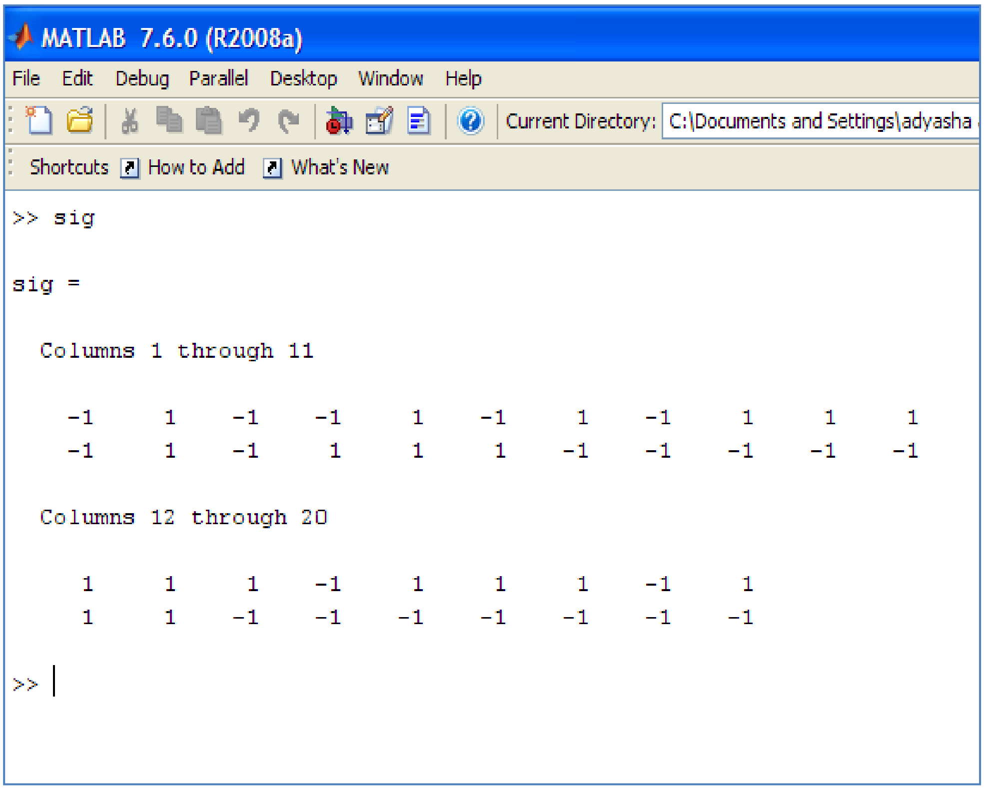


Figure 3:1 : Snapshot of emitted signals generated in MATLAB

3.1.2. Steering Matrix

The steering matrix is incorporated to manifest the incremental delay introduced when the emitted signal is received by a sensor array. As we progress towards the other sensor nodes from the reference sensor node the incremental delay increases. The steering matrix is obviously different for different type of arrangement of sensor nodes. It has the dimensions of number of sensors nodes and the number of sources. A simple signal which is at the direction of arrival Φ results in a multiple of the emitted signal.

The steering matrix is also known as the action vector, array propagation vector or signal replica vector. The figure shows a uniform linearly arranged array and the incremental delay as we move from the first sensor node.

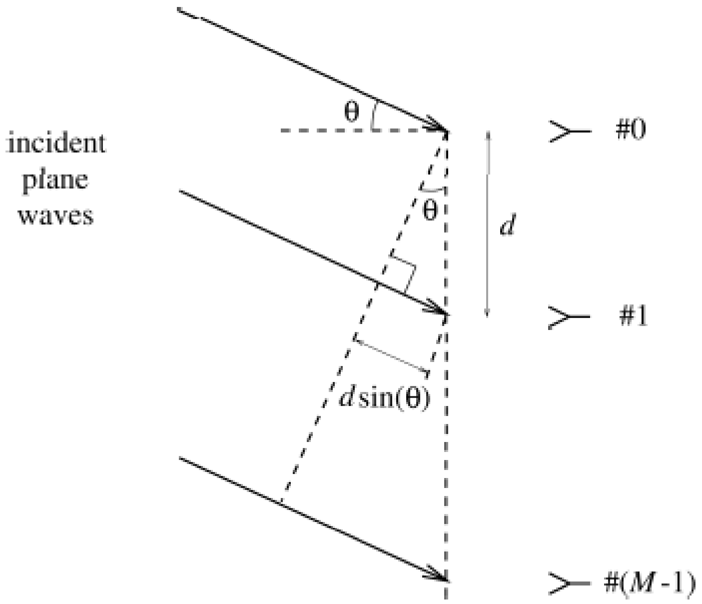


Figure 3:2 : Path Difference in an Uniform Linear Array

For a uniform linear array,

At sensor #0, the path difference will be given by

$e^{jwt} = \cos(wt) + jsin(wt);$

At sensor #1, the path difference will be given by

$e^{jw(t-(d\div c)cos\psi))}$ = $\cos\,(wt - (d \times \cos\,\psi\,) \div c)$ +$jsin(wt - (w \times d \times \cos\,\psi\,) \div c)$

Since $w = 2 \times \Pi \times f = 2 \times \Pi \times c \div lambda$; the path difference will be modified to

$e^{j(wt-\Phi)}$

where Φ denotes the electrical angle of incidence and $\Phi = \frac{(2\Pi \times d)}{lambda} \times\ \cos\boldsymbol{\psi}$;

Similarly at sensor #2 ,the path difference will be given by $e^{j(wt-2\Phi)}$ and so on.Thus the steering matrix will be given by

$$S = e^{jwt}\left[1\ \ e^{-j\Phi}\ \ e^{-j2\Phi}\ e^{-j3\Phi}\ \ \ldots\ldots\ldots\ldots..\right];$$

Hence,the steering matrix can be effectively calculated in this manner.

3.1.3. Modelling of Noise

Since the signals are being emitted in air at ISM radio band frequencies in free air medium they are not free from noise.Noise is modelled as additive white gaussian noise which has a zero mean and the signal to noise ratio is specified according to requirement.The performance of the algorithm is evaluated in terms of probability of resolution and mean square error at different values of signal to noise ratio.

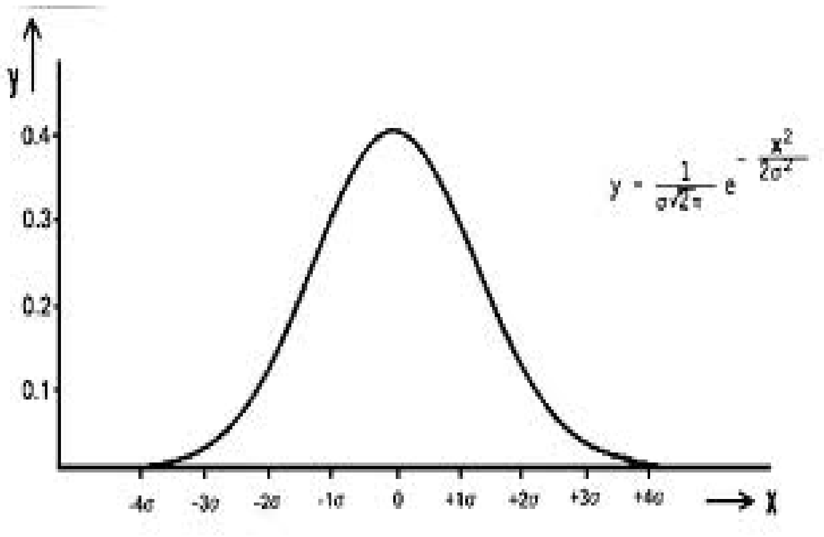


Figure 3:3 : Plot of Probability Distribution Function

The gaussian function is given by:

$$y = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}; \qquad (3.2)$$

where μ denotes the mean of the random variable

and σ^2 denotes the variance .

Since the noise is white , it has all spectral components .The spectral components are present for all frequencies.When viewed in a spectrum analyser the signal can be viewed as shown in the following figure:

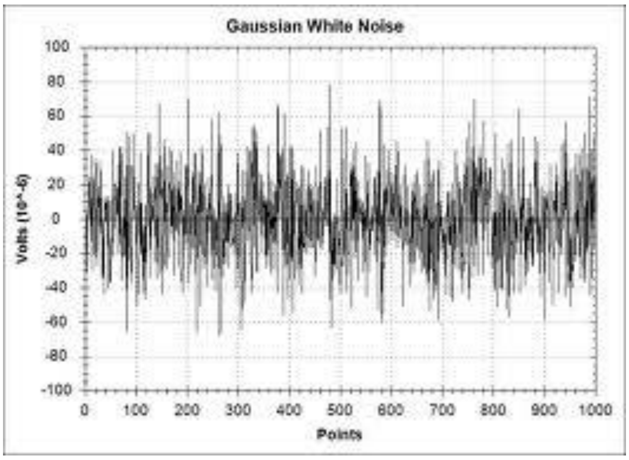
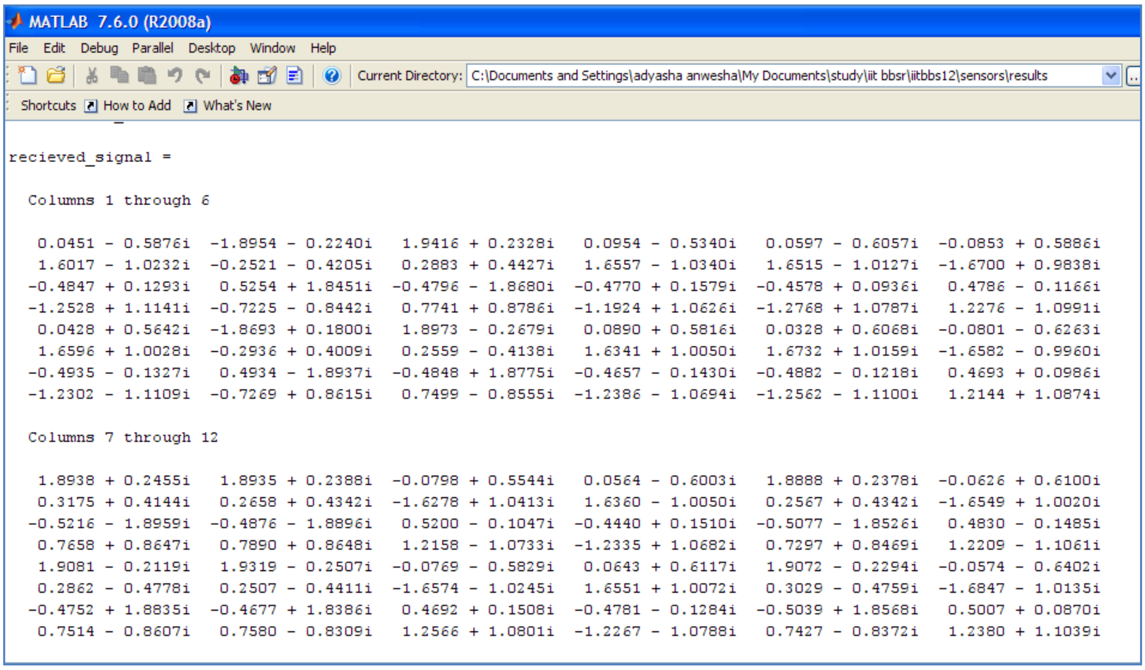


Figure 3:4 : View of a Gaussian Noise in a Spectrum Analyser

3.1.4. Modelling the Received Signal

In a real life scenario of habitat monitoring the received signals are retrieved from the sensor nodes.But in simulation we need to calculate the received signals so that they can be further mathematically operated on to give the accurate direction of arrival.

$$received\ signal = emitted\ signal\ \times steering\ matrix + noise;$$



```
MATLAB 7.6.0 (R2008a)
File  Edit  Debug  Parallel  Desktop  Window  Help
Current Directory: C:\Documents and Settings\adyasha anwesha\My Documents\study\iit bbsr\iitbbs12\sensors\results
Shortcuts  How to Add  What's New

recieved_signal =

  Columns 1 through 6

   0.0451 - 0.5876i  -1.8954 - 0.2240i   1.9416 + 0.2328i   0.0954 - 0.5340i   0.0597 - 0.6057i  -0.0853 + 0.5886i
   1.6017 - 1.0232i  -0.2521 - 0.4205i   0.2883 + 0.4427i   1.6557 - 1.0340i   1.6515 - 1.0127i  -1.6700 + 0.9838i
  -0.4847 + 0.1293i   0.5254 + 1.8451i  -0.4796 - 1.8680i  -0.4770 + 0.1579i  -0.4578 + 0.0936i   0.4786 - 0.1166i
  -1.2528 + 1.1141i  -0.7225 - 0.8442i   0.7741 + 0.8786i  -1.1924 + 1.0626i  -1.2768 + 1.0787i   1.2276 - 1.0991i
   0.0428 + 0.5642i  -1.8693 + 0.1800i   1.8973 - 0.2679i   0.0890 + 0.5816i   0.0328 + 0.6068i  -0.0801 - 0.6263i
   1.6596 + 1.0028i  -0.2936 + 0.4009i   0.2559 - 0.4138i   1.6341 + 1.0050i   1.6732 + 1.0159i  -1.6582 - 0.9960i
  -0.4935 - 0.1327i   0.4934 - 1.8937i  -0.4848 + 1.8775i  -0.4657 - 0.1430i  -0.4882 - 0.1218i   0.4693 + 0.0986i
  -1.2302 - 1.1109i  -0.7269 + 0.8615i   0.7499 - 0.8555i  -1.2386 - 1.0694i  -1.2562 - 1.1100i   1.2144 + 1.0874i

  Columns 7 through 12

   1.8938 + 0.2455i   1.8935 + 0.2388i  -0.0798 + 0.5544i   0.0564 - 0.6003i   1.8888 + 0.2378i  -0.0626 + 0.6100i
   0.3175 + 0.4144i   0.2658 + 0.4342i  -1.6278 + 1.0413i   1.6360 - 1.0050i   0.2567 + 0.4342i  -1.6549 + 1.0020i
  -0.5216 - 1.8959i  -0.4876 - 1.8896i   0.5200 - 0.1047i  -0.4440 + 0.1510i  -0.5077 - 1.8526i   0.4830 - 0.1485i
   0.7658 + 0.8647i   0.7890 + 0.8648i   1.2158 - 1.0733i  -1.2335 + 1.0682i   0.7297 + 0.8469i   1.2209 - 1.1061i
   1.9081 - 0.2119i   1.9319 - 0.2507i  -0.0769 - 0.5829i   0.0643 + 0.6117i   1.9072 - 0.2294i  -0.0574 - 0.6402i
   0.2862 - 0.4778i   0.2507 - 0.4411i  -1.6574 - 1.0245i   1.6551 + 1.0072i   0.3029 - 0.4759i  -1.6847 - 1.0135i
  -0.4752 + 1.8835i  -0.4677 + 1.8386i   0.4692 + 0.1508i  -0.4781 - 0.1284i  -0.5039 + 1.8568i   0.5007 + 0.0870i
   0.7514 - 0.8607i   0.7580 - 0.8309i   1.2566 + 1.0801i  -1.2267 - 1.0788i   0.7427 - 0.8372i   1.2380 + 1.1039i
```

Figure 3:5: Snapshot of the Command Window Depicting Emitted Signal

3.2. System Design

The farmland is considered to be a square of dimensions 50m×50m.There are animals in the farmland.Some of them are moving while others are stationary while others are moving.The Direction of arrival estimation of animals is done by using maximum likelihood optimised by differnential evolution at regular time intervals say every 5 hours.If the animal is found to have the same direction of arrival for two consequtive estimations,then the phase of the carrier signal is checked.If the phase of the carrier also turns out to be the same the the animal is indeed stationary and a warning signal is issued to the owner of the farmland.

3.2.1. Design Parameters of the Farmland

The farmland has been designed in the shape of a square of 50m×50m with the centre of the field being at origin.For the case of simplicity we have assumed four animals in the farmland and estimated their directions of arrival at six consequtive time intervals.The directions of arrival have been estimated by maximum likelihood approach optimised by differential evolution and improvised by averaging process. The farmland designed using MATLAB looks like

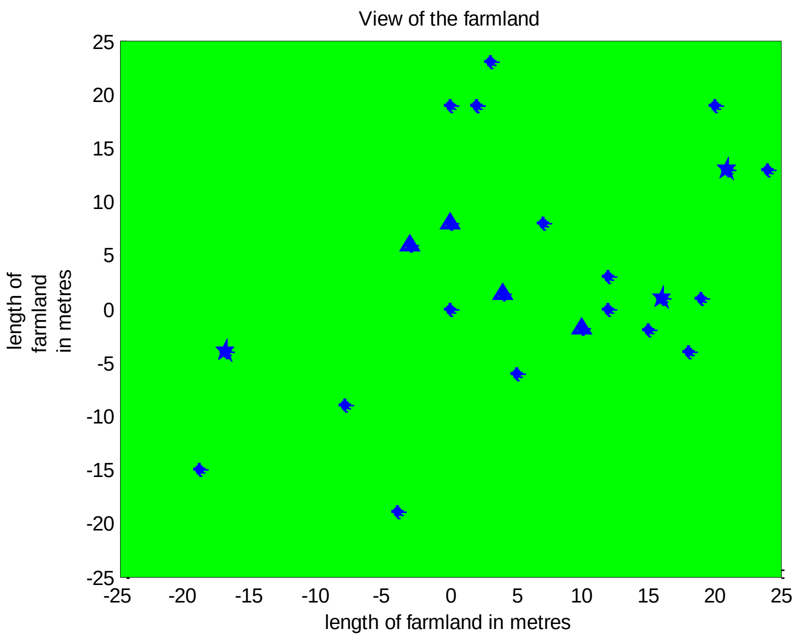


Figure 3:6 : View of Farmland Depicting Random Position of Animals

The triangles denote the initial position of the animals and stars the next five consecutive locations. The final positions have been denoted by pentagons.

3.2.2. Design of Maximum Likelihood Algorithm

Maximum likelihood approach was used for direction of arrival estimation of the maximum likelihood algorithm.It is basically a log-likelihood function defined in terms of noise and signal parameters.The advantage is that it requires only a few iterations to converge. It is free of any further structural constraints or parametric model restrictions. These restrictions are usually imposed on noise covariance matrix and received signals in most prevalent ML-based approaches to DOA estimation in spatially correlated noise. It is based on unification of eigenvectors. The proposed algorithm concentrates the ML estimation problem with respect to all nuisance parameters.

An array of N sensor nodes is placed at a certain distance from a group of R sources. Since observations are not made continuously, they are made a number of times in a given time interval. Suppose observations are taken for M snapshots. The sensor nodes may be placed in any arrangement –linear, circular or random. The only thing that varies with the arrangement is the steering matrix. The steering matrix is of dimensions $N \times R$ and that of received signal will be $\times M$. The received signal is given by the equation described above. Let the received signal, after noise being added to it be X. The dimensions of X will be $N \times M$

$\max\ \_likelihood = \quad abs(\log_e \Big(\det \Big(s \times S_{cap} \times s1' + sigma_{sq} \times$
$eyeN,N;$(3.3)

Where S = steering matrix;

$$A = inv(s' \times s) \times s'; \tag{3.4}$$

$$Pa = s \times A; \tag{3.5}$$

$$Projection = eye(N, N) - Pa; \tag{3.6}$$

$$sigma_{sq} = \left(1 \div (N - R)\right) \times trace(projection \times R); \qquad (3.7)$$

$$S_{cap} = A \times (R - sigma_{sq} \times eye(N, N)) \times A'); \qquad (3.8)$$

The value of the above function is plotted against the angles that is the value of the function is calculated for all possible angles pairs and then the surface plot is obtained.

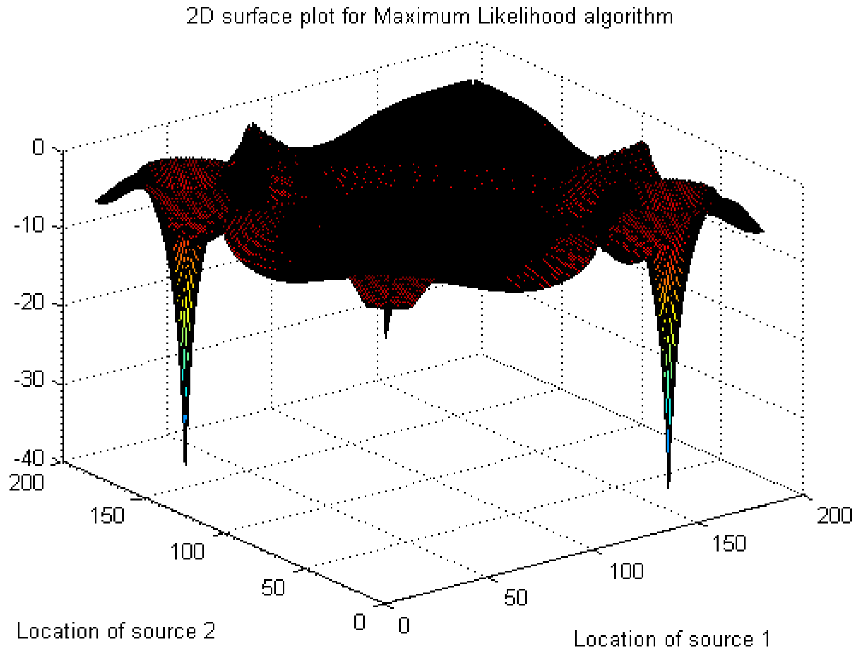


Figure 3:7 : MATLAB Simulation Plot for Maximum Likelihood

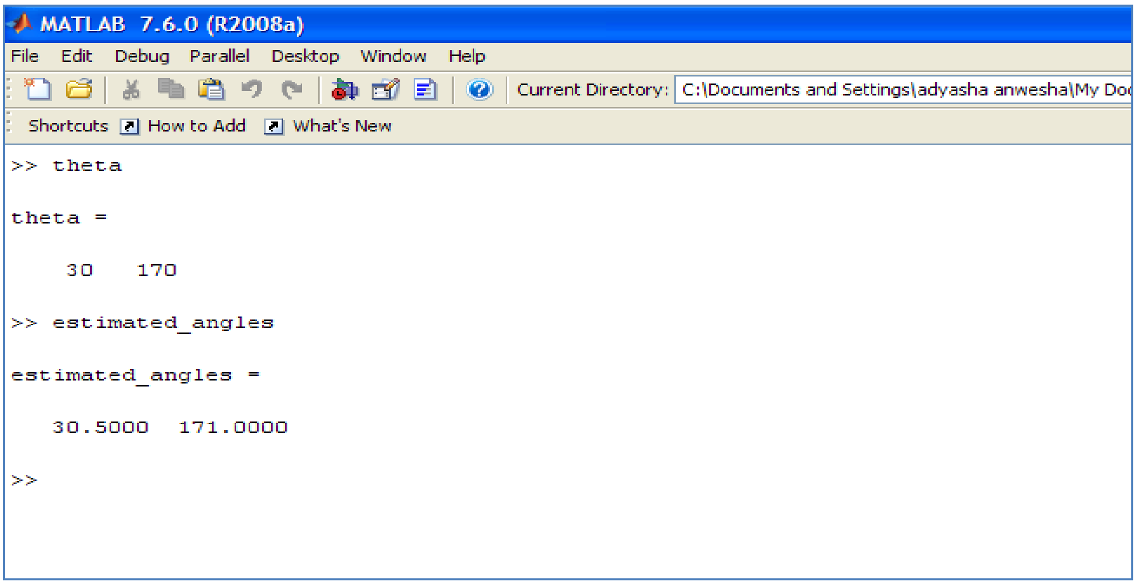


Figure 3:8 : Snapshot for Command Window for Phase Estimation

3.2.3. Design of Differential Evolution Algorithm

Being a stochastic, population based optimisation algorithm, differential evolution is used in this case for optimisation of maximum likelihood function[2].Maximum likelihood function being a multi-variate non linear function is optimised best by differential evolution.the details of the various parametes used are as follows:

Table 3-1 : Parameters and corresponding constants for differential evolution

Parameters	Constants
Population size	50
Number of generations	40
Constant factor	1.2
Cross over ratio	0.6

The parameters mentioned above have been set after a lot of experimentation.Precision upto the second decimal place have been achieved with these parameters.Though greater precision could have been achieved by incearsing the population size as well as the number of generations but computation time has also been taken into consideration and thus these parameters were set[2].

The performance of the system was also checked for conditions when the signal to noise ratio was as low as zero decibles.It was evaluated in terms of minimum mean square error and probability of resolution.The variable for probability of resolution was incremented by one if the the difference of the estimated values of direction of arrival from the actual values was less than half the difference between the actual values. The plot of the signal to noise ratio versus mimimum mean square error and signal to noise ratio versus probility of error is as folows:

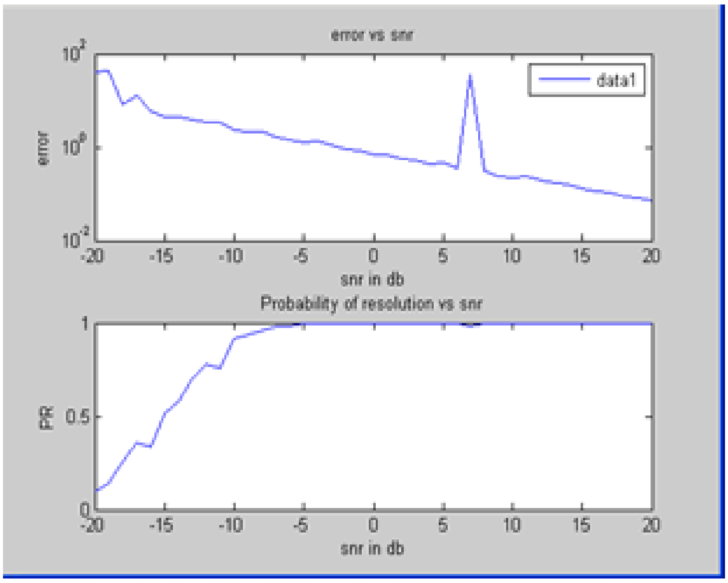


Figure 3:9 : Performance of Differential Evolution in Presence of Noise

It is observed from the graph that probability of resolution converges to one even in very low signal to noise ratio conditions.

After getting satisfactory results for even high noise conditions,the parameters were set and used in the final algorithm.The objective function is the maximum likelihood function or the log likelihood function.

3.3. Averaging of the Maximum Likelihood Performance

One of the methods to minimise the errors is also by averaging.There are of various types: forward averaging, backard averaging and forward-backwad averaging. This involves dividing the entire signal space into number of subspaces and then estimating their direction of arrival taking into account.All possible combination of signal subspaces are considered and finally direction of arrival estimation is done for each signal subspace.Since direction of arrival for each subspace has been done multiple times,the average of these multiple observations is considered for the final estimated value.

As the number of terms for averaging increases the error reduces.This has been illustrated in the following table:

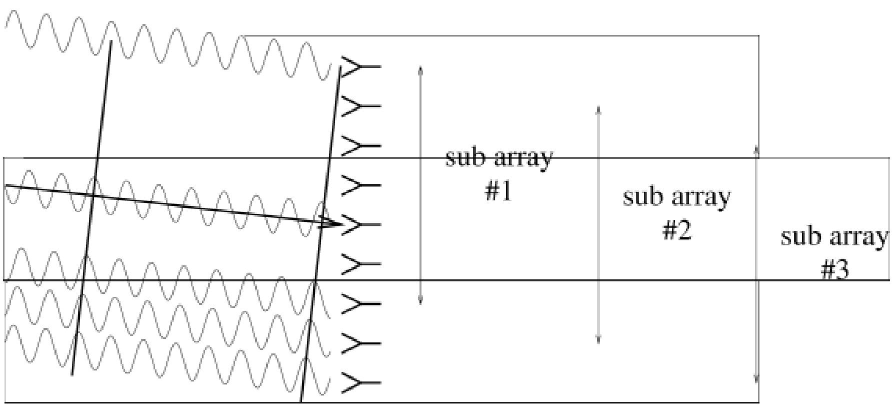
Table 3-2 : Table depicting number of averaging terms and corresponding minimum mean square error

Number of times DOA of a given subspace is estimated	Minimum mean square error
2	0 .78
3	0.64
4	0.59

Thus it is evident from the above table that as the number of terms in the averaging process increases the,the error is minimised.Further,it is observed increasing the averaging terms also increases the computation time.So, a trade off between the two has to be made.

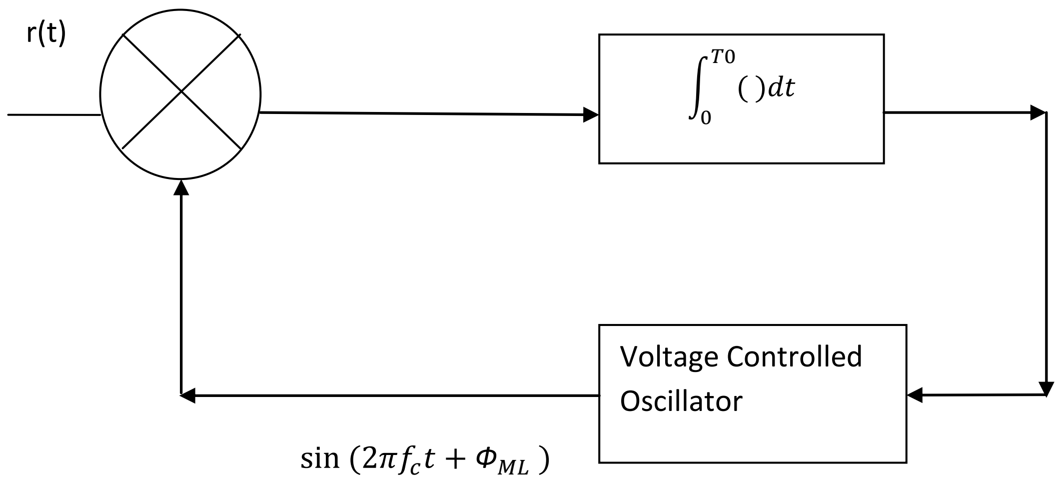
In this case for simplicity let four sources be assumed.Let the sources be named as A,B,C,D.Maximum likelihood function for two sources is considered for the time being.Instead of estimating the direction of arrival for sources AB and CD only,the direction of arrival of AB, BC, CD, DA, AC, BD are considered.As clearly evident direction of arrival for each source has been estimated thrice.So all these three observations for a given source were averaged and the final estimation was obtained.

Table 3-3 : Foward Backward Averaging of Sensor Observations in Direction of Arrival Estimation



3.4. Maximum Likelihood Phase Detection

Phase detection of a received signal is very important for synchronisation purposes.The received signal is a delayed version of the transmitted signal with some phase added to it.In addition ,some noise may also be added to the signal.The receiver observes the received signal over a time interval greater than the time period of the symbol and detects the phase of the signal.In this case there has to be a carrier to detect the phase of the signal.Maximum likelihood phase detection is also known as log likelihood phase detection. The reasons will become evident once the equations are known. In maximum likelihood phase detection the signal parameter Φ is treated as deterministic but unkown. The reciever extracts the estimate[1].Observations made in a singe time interval are also called as one shot estimates. Generally tracking loops are used that update the estimates.



Block Diagram for Log Likelihood Estimates

Here since the signals are transmitted in air,the noise being added to it is modelled as additive white gaussian noise. Our objective is to maximise the probability that the signal being transmitted has a certain given phase.Let $r(t)$ be the reieved signal and

$s(t)$be the transmitted signal. A continuous time equivalent of the maximization of p(r/Φ) is developed in this section.

Since the noise is modelled as additive white gaussian noise whose mean is zero the joint probability distribution function is expressed as:-

$$p(r/\Phi) =) = (\frac{1}{\sigma\sqrt{2\Pi}})^N \times e^{\{-\sum_{n=1}^{N}(r-s)^2/2\sigma^2\}} \qquad (3.9)$$

Where $$r(t) = \int_0^{T0} r \times \Phi(t)\, dt \qquad (3.10)$$

$$s(t) = \int_0^{T0} s(t;\psi) \times \Phi(t)\, dt \qquad (3.11)$$

Where $T0$ represents the integration interval.Maximising the above equation implies maximising of $p(r/\Phi)$ with respect to Φ which is same as maximisation of maximisation of log likelihood function.

The maximum likelihood function can be represented as:-

$$\Lambda(\Phi) = e^{-\left(\frac{1}{N0}\right)\int_0^{T0}[\{r(t)-s(t;\Phi)\}^2]} dt; \qquad (3.12)$$

Finally to maximise $\Lambda(\Phi)$ its is differentiated with respect to theta and equated to zero.

In this case,a carrier of twice the signal frequency has been considered .The signal frequency lies in the 122-123MHz range of the ISM band.Thus, the signal frequency is assumed to be in the 244-246 MHz range.For simplicity a cosine wave with an unknown phase is taken as the carrier wave and the received signal is this cosine signal added with some noise.

$r(t) = A \times cos(2\Pi \times f_c \times t) + n(t);$

Where f_c denotes the carrier frequency lying in the 244-246MHz range.

Simplifying the maximum likelihood function we get;

$$\Lambda(\Phi) = e^{-\left(\frac{1}{N0}\right)\int_0^{T0}[\{r(t)-s(t;\Phi)\}^2]} dt; \qquad (3.13)$$

$=e^{\{-1/N0\int_0^{T0} r(t)\,^2 dt + 2/N0\int_0^{T0} r(t)s(t;\Phi)dt - 1/N0\int_0^{T0} S(t;\Phi)^2 dt\}}$

From the above equationthe integral of $r(t)^2$,the first term of the exponential factor does not involve any signal parameter. The third tem of the exponential factor $s(t;\Phi)^2$ is a constant which is equal to the energy of the signal over a time interval equal to the observation period $T0.$ Only the second term which includes the cross correlaion of the received signal $r(t)$ with the transmitted signal $s(t;\Phi)$, has to be maximised because it depends on the choice of Φ. Thus the function reduces to:

$$\Lambda(\Phi) = Ce^{(\frac{2}{N0})\int_0^{T0} r(t)s(t;\Phi)dt} \qquad (3.14)$$

Where C=constant independent of Φ.

During maximisation the constant term may also be ignored.Also,exponential function and logarithmic functions are exponential functions,maximising an exponential function and maximising its logarithm is the same.So we need to maximise:

$$\Lambda(\Phi) = \left(\frac{2}{N0}\right)\int_0^{T0} r(t)s(t;\Phi)dt \qquad (3.15)$$

Where the constant term C has been ignored.

Putting $r(t) = A \times cos(2\Pi \times f_c \times t) + n(t)$ in the above equation ,the maximum likelihood function becomes:

$$\Lambda(\Phi) = \left(\frac{2}{N0}\right)\int_0^{T0} r(t)\cos(2\Pi \times f_c \times t + \Phi)\,dt \qquad (3.16)$$

A necessary condition for this function to be maximum at Φ is given by

$$\frac{d\Lambda(\Phi)}{d\Phi}=0$$

This condition gives: $\int_0^{T0} r(t)\sin(2\Pi \times f_c \times t + \Phi)dt = 0;$ (3.17) which gives: $$\varPhi = -tan^{-1}\frac{\int_0^{T0} r(t)\sin(2\Pi f_c t\,)dt}{\int_0^{T0} r(t)\cos\,(2\Pi f_c t)dt} \tag{3.18}$$

Thus the hase of the signal was detected in the following manner.

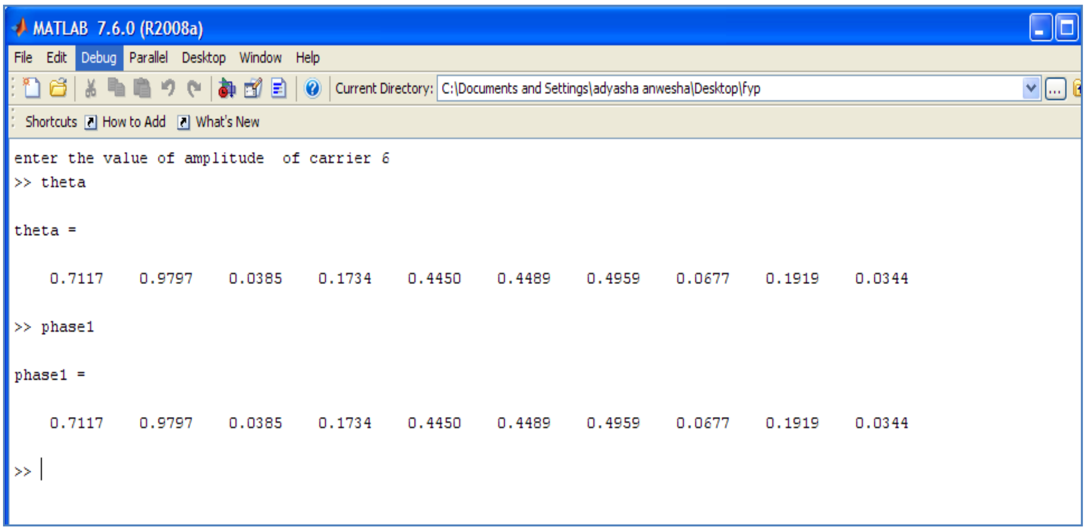


Figure 3:10 : Snapshot for MATLAB Simulation of Phase Detection

Chapter 4. Result

4.1. Final Program and Results Obtained

The final program for direction of arrival was written .For the prototype, a group of four animals were considered and the direction of arrival was estimated for six consecutive time intervals with each interval being of a duration of five hours. If the direction of arrival was found to be same for at least two consecutive time intervals a message was displayed that "The animal has same direction of arrival. We need to check its phase". On the other hand if the direction of arrival of the animals were found to be different for each of the time intervals a message was displayed that "The animal is moving".

The farmland and position of the animals for the final simulation was as follows:

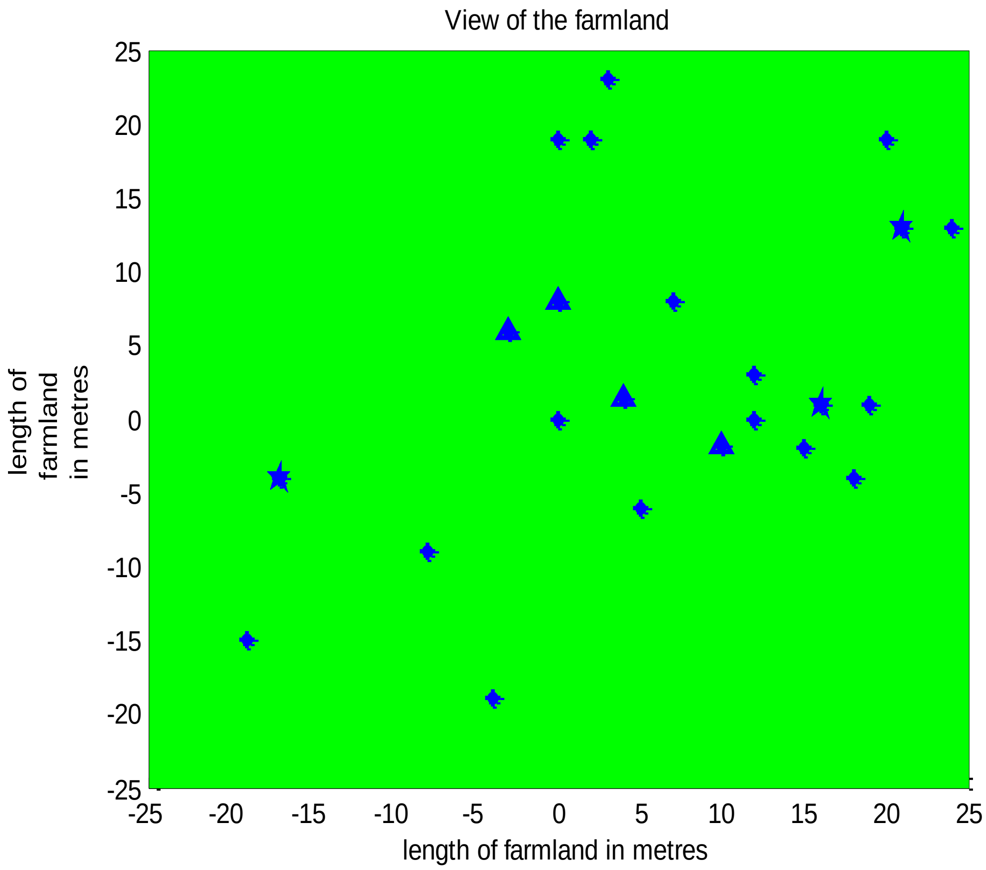


Figure 4:1 : View of Farmland Depicting Random Position of Animals

For the final simulation the position of animals were not randomly generated because for the case of random generation the position of each animal for all the five successive intervals will be different. The triangles denote the initial positions, the stars the next five consecutive locations and finally the pentagons represent the final positions. So, to check for the case of same direction of arrival the positions of the animals were given as input in a manner such that the direction of arrival for at least one animal will be the same so that the phase can be detected and the reliability of the phase detection algorithm can be checked. In this case for the second animal, the third and the fourth positions were given in the same direction of arrival. But the phases for these positions have been generated randomly. Thus when simulated the result was obtained as:

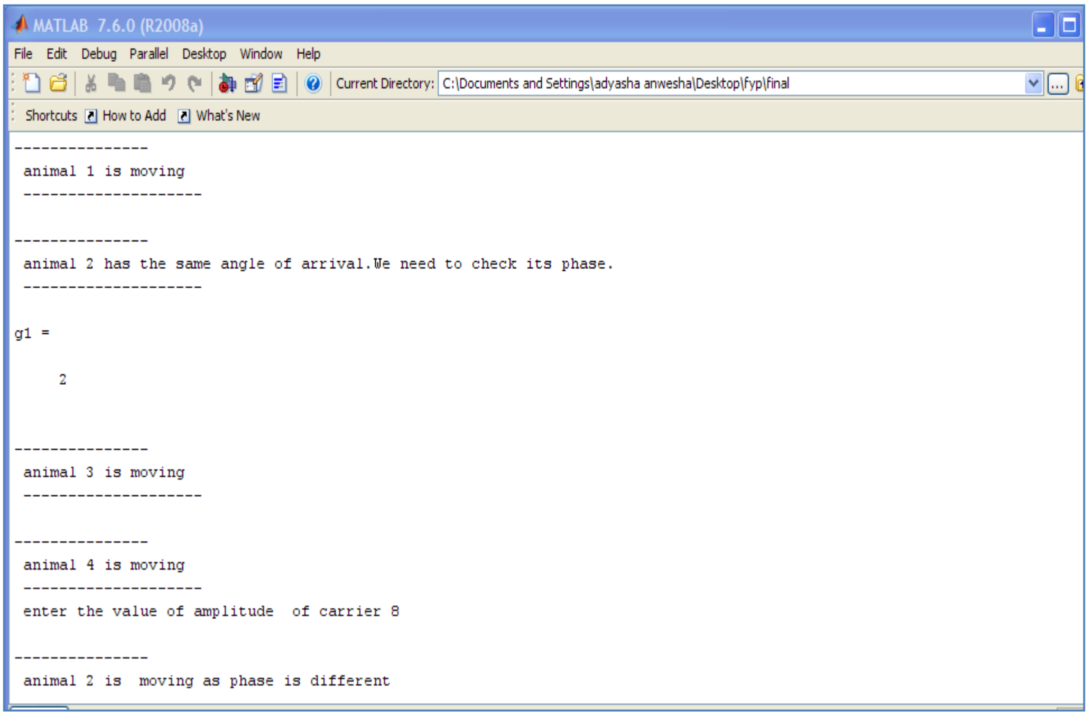


Figure 4:2 : Snapshot of the Screen Displaying Final Simulation Result

Chapter 5. Conclusion

5.1. Conclusion

The various methods of estimating the direction of arrival of animals were studied. The advantages and disadvantages of each of these methods were studied and some of the methods were simulated. MUSIC (MUltiple SIgnal Classification) and Maximum likelihood methods were simulated using MATLAB. After studying the nature of graphs generated which indicate the position of the sources it was seen that MUSIC (MUltiple SIgnal Classification) method is not suitable for closely spaced sources. Hence, deterministic maximum likelihood technique was chosen for the final simulation. After the direction of arrival estimation was done for a given animal these estimates were compared with each other and if two consecutive locations were found to have the same estimate, phase detection of the carrier signal was carried out. Same phase of the carrier indicates same position while different phase indicates different position. Thus the movement of animals was monitored saving a lot of man hours that would have involved in monitoring.

5.2. Future scope

This algorithm of habitat monitoring system can be incorporated to monitor patch of habitats. Hardware architecture to implement this can be developed and the system as a whole can be implemented. A small geographic region can be monitored by developing sensor nodes which can communicate and co-ordinate with each other An online system can be developed which will track the movement of the animals and update the database. A base station may also be installed and the messages may be transmitted using a Yagi - Uda antenna with a link of about 350metres.

Reference

[1] M.Adjrad, A.Belouchrani, "Estimation of Multi-Component Polynomial Phase Signals Impinging on a Multisensory Array Using State-Space Modeling", IEEE Transactions on Signal Processing , volume 55,no. 1, pp. 32-45, August 2007

[2] M. Li and Y. Lu, "A Refined Genetic Algorithm for Accurate and Reliable DOA Estimation with a Sensor Array," Wireless Personal Communication, volume 43, no. 2, pp. 533–547, 2007.

[3] A.Ferreol, P. Larzabal, and M .Viberg, "On the Asymptotic Performance Analysis of Subspace Based DOA Estimation in the Subspace of Modeling Errors: Case of MUSIC," IEEE transactions on Signal Processing, volume 54, no.3, pp-907-920, March 2006.

[4] I.F. Akyilidz, Weilman Su, Y Sankarasubramaniam and E.Cayiri , "A Survey on Sensor Networks," IEEE Commuication Magazine, volume 40, no. 8, pp. 102-114, August 2002.

[5] D. Estrin, L. Girod, G. Pottie, and M. Srivastava, "Instrumenting the World with Wireless Sensor Networks," in *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, volume 4, (Salt Lake City, UT), pp. 2033–2036, May 2001.

[6] B. Flanagan and K. Bell, "Improved array self-calibration with large sensor position errors for closely spaced sources," in *Proceedings of the First IEEE Sensor Array and Multichannel Signal Processing Workshop*, (Cambridge, MA), pp. 484 – 488, March 2000.

[7] John G. Proakis and Masoud Salehi, "Digital Communications", Fifth edition, Mc Graw Hill publications.