Preparatory Definitions for Understanding Target Localization in Wireless Sensor Networks

Everette Adams

Abstract—In order to fully understand the concepts given in research journals that explain the vast amounts of processes for locating targets using Wireless Sensor Networks (WSNs), a list of definitions with detailed explanations of the various terms are provided in this paper to enhance the process of explanatory comprehension within this subject. The terms and definitions are the most basic level of understanding for novice calculations and procedures within this field of research.

Index Terms—Method of Moments (MME), Stochastic Process, Monomolecular Growth Curve, Nonlinear Regression.

I. INTRODUCTION

Wireless sensor networks are a popular research topic [1]-[56]. The definitions within this document are written in paragraph form with italics highlighting each definition for better realization.

II. PROCEDURE

A Sensor Network is a group of specialized transducers with a communications infrastructure intended to monitor and record conditions at diverse locations. Commonly monitored parameters are temperature, humidity, pressure, wind direction and speed, illumination intensity, vibration intensity, sound intensity, power-line voltage, chemical concentrations, pollutant levels and vital body functions. A sensor network consists of multiple detection stations called sensor nodes, each of which is small, lightweight and portable. Every sensor node is equipped with a transducer, microcomputer, transceiver and power source. The transducer generates electrical signals based on sensed physical effects and phenomena. The microcomputer processes and stores the sensor output. The transceiver, which can be hard-wired or wireless, receives commands from a central computer and transmits data to that computer. The power for each sensor node is derived from the electric utility or from a battery [1]. Target Localization modules provide detection of a specified target to calculate the position of the camera/robot based on the orientation and size of the target. In order for the target to be accurately detected, these assumptions are made:

1. The target is composed of horizontal and vertical lines. Some curved edges are acceptable but the majority of the target is composed of straight lines.
2. The target’s edges are sharp and of high contrast with respect to the rest of the image.
3. The target is perpendicular to the plane of movement. (i.e. the camera/robot moves on the ground/horizontal plane in front of the target.) While some tilt and roll is permitted, significant tilt or roll will cause detection failure. It is assumed that your camera/robot stays right side up.
4. Most of the target is in view. Some obstructions are permitted but enough corners must be visible for detection to work.
5. The target is planar, (i.e. it must be on a flat surface). If the target is bent or curved in any way, detection will fail. Given these assumptions the module will return X and Y coordinates of the camera relative to the target such that you can know an approximate location. The module has no specific units specified (i.e. meters, feet, etc.). The results will be in whatever units you choose to use for the size of the target (i.e. meters, feet, etc.) [2][3].

Wireless Sensor Networks (WSNs) consist of spatially distributed autonomous sensors to monitor physical or environmental conditions, such as temperature, sound pressure, etc., and to cooperatively pass their data through the network to a main location. The more modern networks are bi-directional, also enabling control of sensor activity. The development of wireless sensor networks was motivated by military applications such as battlefield surveillance; today such networks are used in many industrial and consumer applications, such as industrial process monitoring and control, machine health monitoring, and so on. A WSN is built of “nodes,” from a few to several hundreds or even thousands, where each node is connected to one, or sometimes several sensors. Each sensor network node has typically several parts: a radio transceiver with an internal antenna or connection to an external antenna, a microcontroller, an electronic circuit for interfacing with the sensors, and an energy source such as a battery or an embedded form of energy harvesting. A sensor node might vary in size from that of a shoebox down to the size of a grain of dust, although functioning “motes” of genuine microscopic dimensions have yet to be created. The cost of sensor nodes is similarly variable, ranging from a few to hundreds of dollars, depending on the complexity of the individual sensor nodes. Size and cost constraints on sensor nodes result in corresponding constraints on resources such as energy, memory, computational speed and communications bandwidth. The topology of the WSNs can vary from a simple
Star Network to an advanced Multi-Hop Wireless Mesh Network [4][2].

Fig. 1. Example of a Wireless Sensor Network.

A Wireless Mesh Network (WMN) is a communications network made up of radio nodes organized in a mesh topology. Wireless mesh networks often consists of mesh clients, mesh routers and gateways. The mesh clients are often laptops, cell phones and other wireless devices while the mesh routers forward traffic to and from the gateways, which may, but need not, connect to the Internet. The coverage area of the radio nodes working as a single network is sometimes called a mesh cloud. Access to this mesh cloud is dependent on the radio nodes working in harmony with each other to create a radio network. A mesh network is reliable and offers redundancy. When one node can no longer operate, the rest of the nodes can still communicate with each other, directly or through one or more intermediate nodes. The animation below illustrates how wireless mesh networks can self-form and self heal. Wireless mesh networks can be implemented with various wireless technology including 802.11, 802.15, 802.16, cellular technologies or combinations of more than one type. A wireless mesh network can be seen as a special type of wireless ad-hoc network. A wireless mesh network often has a more planned configuration, and may be deployed to provide dynamic and cost effective connectivity over a certain geographic area. An ad-hoc network, on the other hand, is formed ad hoc when wireless devices come within communication range of each other. The mesh routers may be mobile, and be moved according to specific demands arising in the network. Often the mesh routers are not limited in terms of resources compared to other nodes in the network and thus can be exploited to perform more resource intensive functions [5].

Fig. 2. Example of a Wireless Mesh Network.

Star Networks are one of the most common computer network topologies. In its simplest form, a star network consists of one central switch, hub or computer, which acts as a conduit to transmit messages. This consists of a central node, to which all other nodes are connected; this central node provides a common connection point for all nodes through a hub. In Star topology every node (computer workstation or any other peripheral) is connected to a central node called a hub or a switch. The switch is the server and the peripherals are the clients. Thus, the hub and leaf nodes, and the transmission lines between them, form a graph with the topology of a star. If the central node is passive, the originating node must be able to tolerate the reception of an echo of its own transmission, delayed by the two-way transmission time (i.e. to and from the central node) plus any delay generated in the central node. An active star network has an active central node that usually has the means to prevent echo-related problems. The star topology reduces the chance of network failure by connecting all of the systems to a central node. When applied to a bus-based network, this central hub rebroadcasts all transmissions received from any peripheral node to all peripheral nodes on the network, sometimes including the originating node. All peripheral nodes may thus communicate with all others by transmitting to, and receiving from, the central node only. The failure of a transmission line linking any peripheral node to the central node will result in the isolation of that peripheral node from all others, but the rest of the systems will be unaffected. It is also designed with each node (file servers, workstations, and peripherals) connected directly to a central network hub, switch, or concentrator. Data on a star network passes through the hub, switch, or concentrator before continuing to its destination. The hub, switch, or concentrator manages and controls all functions of the network. It also acts as a repeater for the data flow. This configuration is common with twisted pair cable. However, it can also be used with coaxial cable or optical fiber cable [6].

Fig. 3. Depiction of a Star Network.

A Bayesian Network is a probabilistic graphical model (a type of statistical model) that represents a set of random variables and their conditional dependencies via a Directed Acyclic Graph (DAG). For example, a Bayesian network could represent the probabilistic relationships between diseases and symptoms. Given symptoms, the network can be used to compute the probabilities of the presence of various diseases. Formally, Bayesian networks are DAGs whose nodes represent random variables in the Bayesian sense: they may be observable quantities, latent variables, unknown parameters or hypotheses. Edges represent conditional dependencies; nodes, which are not connected, represent variables that are conditionally independent of each other. Each node is associated with a probability function that takes as input a particular set of values for the node’s parent variables and gives the probability of the variable represented by the node [7]. In mathematics and computer science, a
Directed Acyclic Graph (DAG), is a directed graph with no directed cycles. That is, it is formed by a collection of vertices and directed edges, each edge connecting one vertex to another, such that there is no way to start at some vertex \( v \) and follow a sequence of edges that eventually loops back to \( v \) again [8].

Fig. 4. Example of Directed Acyclic Graph.

In statistics, Method of Moments (MME) is a method used for estimating a population’s parameters. The parameters can be its mean, variance, median, or any other mathematical calculation involving a specified population. A brief moment of time is sampled within the experimental population, and that time is equated in sample moments of unobservable population moments. The overall solved equations that are generated from the samples are totaled as quantities of estimation. Estimates by the method of moments may be used only as a starting point for likelihood approximations [9]. In probability theory, a Stochastic Process, or random process is a collection of random variables; this is often used to represent the evolution of some random value, or system, over time. This process does not describe a process that evolves in an instant. In this process, there is some indeterminacy: even if the initial condition is known, there are several directions in which the process may evolve over time [10]. A Monomolecular Growth Curve involves one molecule. The growth curve of that molecule is an empirical model of the evolution of a quantity over time. Values for measured data can be plotted on a graph as a function of time. In mathematical statistics, growth curves are often modeled as being continuous Stochastic Processes [11].

Fig. 5. Plot of a True Monomolecular Growth Curve.

In statistics, Nonlinear Regression is a form of regression analysis in which observational data are modeled by a function, which is a nonlinear combination of the model parameters and depends on one or more independent variables. The data are fitted by a method of successive approximations. The data consist of error-free independent variables (explanatory variables), \( x \) and their associated observed dependent variables (response variables), \( y \). Each \( y \) is modeled as a random variable with a mean given by a nonlinear function \( f(x,B) \). Systematic error may be present but its treatment is outside the scope of regression analysis. If the independent variables are not error-free, this is an errors-in-variables model, also outside of this scope [12].

Fig. 6. Plot of a graphical nonlinear regression.

The term Occupancy Estimation is used to describe a system that generates occupancy estimates based on Kinetic-Motion (KM). The model predicts the movements of occupants through a region divided into a plurality of segments. The system includes a controller for executing an algorithm representing the KM-based model. The KM-based model includes state equations that define each of the plurality of segments as containing congested portions and uncongested portions. The state equations define the movement of occupants based, in part, on the distinctions made between congested and uncongested portions of each segment [13]. Parameter Estimation refers to the process of using sample data to estimate the value of a population parameter (mean, variance, or \( t \) score) or a model parameter (for example, a weight in a regression equation). Surveys are often used for the purposes of parameter estimation [14]. In mathematical statistics and information theory, Fisher Information (sometimes called information) can be defined as the variance of the \( t \) score, or as the expected value of the observed information [15]. Cramer-Rao Lower Bound in its simplest form states that the variance of any unbiased estimator is at least as high as the inverse of the Fisher information. An unbiased estimator that achieves this lower bound is said to be (fully) efficient. Such a solution achieves the lowest possible mean squared error among all unbiased methods, and is therefore the Minimum Variance Unbiased (MVU) estimator. However, in some cases, no unbiased technique exists which achieves the bound. This may occur even when an MVU estimator exists. The Cramer-Rao bound can also be used to bound the variance of biased estimators of a given bias. In some cases, a biased approach can result in both a variance and a mean squared error that are below the unbiased Cramer-Rao lower bound [16][2]. The Kálmán Filter, also known as Linear Quadratic Estimation (LQE), is an algorithm that uses a series of measurements observed over time, containing noise (random variations) and other inaccuracies, and produces estimates of unknown variables that tend to be more precise than those based on a single measurement alone. More formally, the Kalman filter operates recursively on streams of noisy input data to produce a statistically optimal estimate of the underlying system state. The Kálmán filter has numerous applications in technology. A common application is for guidance, navigation and control of vehicles, particularly aircraft and spacecraft. Furthermore,
the Kalman filter is a widely applied concept in time series analysis used in fields such as signal processing and econometrics. The algorithm works in a two-step process. In the prediction step, the Kalman filter produces estimates of the current state variables, along with their uncertainties. Once the outcome of the next measurement (necessarily corrupted with some amount of error, including random noise) is observed, these estimates are updated using a weighted average, with more weight being given to estimates with higher certainty. Because of the algorithm's recursive nature, it can run in real time using only the present input measurements and the previously calculated state; no additional past information is required [17].

Nonlinear Filtering in signal processing is a filter that has an output that is not linear to the function of its input. That is, if the filter outputs signals R and S for two input signals r and s separately, but does not always output aR + bS when the input is a linear combination ar + bs. Both continuous-domain and discrete-domain filters may be nonlinear [18]. In statistics, a Particle Filter, also known as a Sequential Monte Carlo Method (SMC), is a sophisticated model estimation technique based on simulation. Particle filters are usually used to estimate Bayesian models in which the latent variables are connected in a Markov chain — similar to a Hidden Markov model (HMM), but typically where the state space of the latent variables is continuous rather than discrete, and not sufficiently restricted to make exact inference tractable (as, for example, in a linear dynamical system, where the state space of the latent variables is restricted to Gaussian distributions and hence exact inference can be done efficiently using a Kalman filter). In the context of HMMs and related models, "filtering" refers to determining the distribution of a latent variable at a specific time, given all observations up to that time; particle filters are so named because they allow for approximate "filtering" (in the sense just given) using a set of "particles" (differently weighted samples of the distribution) [19]. Information Fusion, or data fusion is a formal framework in which are expressed the means and tools for the alliance of data originating from different sources [20]. Field Estimation is a fundamental requirement in realistic computational geophysical fluid dynamics, and is the optimal estimation of gridded fields and of spatial-temporal scales directly from the spatially irregular and multivariate data sets that are collected by varied instruments and sampling schemes [21]. Supervised Learning is the machine-learning task of inferring a function from labeled training data. The training data consist of a set of training examples. In supervised learning, each example is a pair consisting of an input object (typically a vector) and a desired output value (also called the supervisory signal). A supervised learning algorithm analyzes the training data and produces an inferred function, which is called a classifier (if the output is discrete) or a regression function (if the output is continuous). The inferred function should predict the correct output value for any valid input object. This requires the learning algorithm to generalize from the training data to unseen situations in a "reasonable" way [22]. In statistics, Bootstrapping is a method for assigning measures of accuracy to sample estimates. This technique allows estimation of the sampling distribution of almost any statistic using only very simple methods. Generally, it falls in the broader class of re-sampling methods. Bootstrapping is the practice of estimating properties of an estimator (such as its variance) by measuring those properties when sampling from an approximating distribution. One standard choice for an approximating distribution is the empirical distribution of the observed data. In the case where a set of observations can be assumed to be from an independent and identically distributed population, this can be implemented by constructing a number of re-samples of the observed dataset (and of equal size to the observed dataset), each of which is obtained by random sampling with replacement from the original dataset. It may also be used for constructing hypothesis tests. It is often used as an alternative to inference based on parametric assumptions when those assumptions are in doubt, or where parametric inference is impossible or requires very complicated formulas for the calculation of standard errors [23].

Dependent and Independent Variables are variables used in an experiment or modeling that can be divided into three types: "dependent variable", "independent variable", or other. The "dependent variable" represents the output or effect, or is tested to see if it is the effect. The "independent variables" represent the inputs or causes, or are tested to see if they are the cause. Other variables may also be observed for various reasons [24].

Copula Theory, in probability theory and statistics, is a kind of distribution function. Copulas are used to describe the dependence between random variables. They are named for their resemblance to grammatical copulas in linguistics. The cumulative distribution function of a random vector can be written in terms of marginal distribution functions and a copula. The marginal distribution functions describe the marginal distribution of each component of the random vector and the copula describes the dependence structure between the components. Copulas are popular in statistical applications as they allow one to easily model and estimate the distribution of random vectors by estimating marginal and copula separately. There are many parametric copula families available, which usually have parameters that control the strength of dependence [25].

In statistics, Maximum-Likelihood Estimation (MLE) is a method of estimating the parameters of a statistical model. When applied to a data set and given a statistical model,
Maximum-likelihood estimation provides estimates for the model’s parameters [26][27].

**Model Selection** is the task of selecting a statistical model from a set of candidate models, given data. In the simplest cases, a pre-existing set of data is considered. However, the task can also involve the design of experiments such that the data collected is well suited to the problem of model selection. Given candidate models of similar predictive or explanatory power, the simplest model is most likely to be correct [28].

**Bayesian Estimation** in estimation theory and decision theory is an estimator or decision rule that minimizes the posterior expected value of a loss function (i.e., the posterior expected loss). Equivalently, it maximizes the posterior expectation of a utility function. An alternative way of formulating an estimator within Bayesian statistics is maximum a posteriori estimation [29].

In Bayesian statistics, Maximum a Posteriori (MAP) estimation is a mode of the posterior distribution. The MAP can be used to obtain a point estimate of an unobserved quantity on the basis of empirical data. It is closely related to Fisher’s method of maximum likelihood (ML), but employs an augmented optimization objective, which incorporates a prior distribution over the quantity one wants to estimate. MAP estimation can therefore be seen as a regularization of ML estimation [30].

**Stochastic Resonance** (SR) is a phenomenon that occurs in a threshold measurement system (e.g., a man-made instrument or device; a natural cell, organ or organism) when an appropriate measure of information transfer (signal-to-noise ratio, mutual information, coherence, etc.) is maximized in the presence of a specific non-zero level of stochastic input noise thereby lowering the response threshold; the system resonates at a particular noise level. Stochastic resonance is observed when noise added to a system changes the system’s behavior in some fashion. More technically, SR occurs if the signal-to-noise ratio of a nonlinear system or device increases for moderate values of noise intensity. It often occurs in bi-stable systems or in systems with a sensory threshold and when the input signal to the system is “sub-threshold”. For lower noise intensities, the signal does not cause the device to cross threshold, so little signal is passed through it. For large noise intensities, the output is dominated by the noise, also leading to a low signal-to-noise ratio. For moderate intensities, the noise allows the signal to reach threshold, but the noise intensity is not so large as to swamp it. Thus, a plot of signal-to-noise ratio as a function of noise intensity shows a γ shape [31].

**Data Fusion** is the process of integration of multiple data and knowledge representing the same real-world object into a consistent, accurate, and useful representation. Fusion of the data from 2 sources (dimension #1 & #2) can yield a classifier superior to any classifiers based on dimension #1 or dimension #2 alone [32].

**Dependent Observations** are observations that are somehow linked or clustered; they are observations that have a systematic relationship to each other. Clustering typically results from research and sampling design strategies, and the clustering can occur within space, across time, or both [33].

**Non-Subtractive Dithering** is a detailed mathematical investigation of multi-bit quantizing systems using non-subtractive dither. It is shown that by the use of dither having a suitably-chosen probability density function, moments of the total error can be made independent of the system input signal, but that statistical independence of the error and the input signals is not achievable. Similarly, it is demonstrated that values of the total error signal cannot generally be rendered statistically independent of one another, but that their joint moments can be controlled and that, in particular, the error sequence can be rendered spectrally white [34].

**Quantization**, in mathematics and digital signal processing, is the process of mapping a large set of input values to a smaller set—such as rounding values to some unit of precision. A device or algorithmic function that performs quantization is called a quantizer. The error introduced by quantization is referred to as quantization error or round-off error. Quantization is involved to some degree in nearly all-digital signal processing, as the process of representing a signal in digital form ordinarily involves rounding. Quantization also forms the core of essentially all lossy compression algorithms [35].

**Distributive Estimation**, or **Estimation of distribution algorithms** (EDAs), sometimes called **probabilistic model-building genetic algorithms** (PMBGAs), are stochastic optimization methods that guide the search for the optimum by building and sampling explicit probabilistic models of promising candidate solutions. Optimization is viewed as a series of incremental updates of a probabilistic model, starting with the model encoding the uniform distribution over admissible solutions and ending with the model that generates only the global optima.

EDAs belong to the class of evolutionary algorithms. The main difference between EDAs and most conventional evolutionary algorithms is that evolutionary algorithms generate new candidate solutions using an **implicit** distribution defined by one or more variation operators, whereas EDAs use an **explicit** probability distribution encoded by a Bayesian network, a multivariate normal, or another model class [36].

**Dithering** is defined as trembling, or vibration [37]. **Probabilistic Quantization** algorithms are used for distributed computation of averages of the node data over networks with bandwidth/power constraints or large volumes of data. Distributed averaging algorithms fail to achieve consensus when deterministic uniform quantization is adopted. It is a dynamical system that generates sequences achieving a consensus, which is one of the quantization values. In addition, it shows that the expected value of the consensus is equal to the average of the original sensor data. Reports of the results from simulations are conducted to evaluate the behavior and the effectiveness of the proposed algorithm in various scenarios [38].

**Location Estimation** in wireless sensor networks is the problem of estimating the location of an object from a set of noisy measurements, when the measurements are acquired in a distributed manner by a set of sensors. Many civilian and
military applications require monitoring that can identify objects in a specific area, such as monitoring the front entrance of a private house by a single camera. Monitored areas that are large relative to objects of interest often require multiple sensors (e.g., infra-red detectors) at multiple locations. A centralized observer or computer application monitors the sensors. The communication to power and bandwidth requirements call for efficient design of the sensor, transmission, and processing [39].

REFERENCES


AUTHOR’S PROFILE

Everette B. Adams received his Associates degree from Colorado Technical University, Colorado Springs, CO, USA, in 2009, where he studied Business. He also attended the University of Alabama at Birmingham, Birmingham, Alabama, USA, and is currently enrolled, and majoring in Health Related Professions. His personal research areas are in Astro-Biology, Wireless Sensor Networks, and Target Localization. He's also a card holding member of the American Physical Society.