Modelling of Extreme Data in a Wireless Sensor Network Through the Application of Random Field Theory

Glenn Patterson

A Thesis

in

The Department

of

Electrical & Computer Engineering

Presented in Partial Fulfillment of the Requirements for the Degree of Master of Applied Science (Electrical & Computer Engineering) at
Concordia, University
Montreal, Quebec, Canada

November 2007

Your file Votre référence
ISBN: 978-0-494-40894-0
Our file Notre référence
ISBN: 978-0-494-40894-0

ABSTRACT

Modelling of Extreme Data in a Wireless Sensor Network Through the Application of Random Field Theory

Glenn Patterson

Wireless Sensor Networks (WSNs) consist of a large number of small, simple sensor nodes which support sensing, processing, and wireless transmission capabilities in order to monitor some physical environment. The data that they collect will be transmitted to an information sink where it can be accessed by the user. Due to the vast number of nodes expected in many WSN applications, there is the possibility that at certain times the network may have potentially huge amounts of data to transmit to the sink. The fact that the nodes have limited energy resources means that when a huge amount of data is generated, the nodes' batteries will be depleted at aggressive rates as nodes try to forward this data to the sink. This problem will be particularly severe in regions of the network close to the sink, as nodes in these regions will be responsible for routing the data from large areas of the network to the sink. This phenomenon is often described as a "data-implosion" around the sink. We develop a model of the node data in a wireless sensor network based on a stochastic model of the underlying phenomenon being observed by the network. The model is based on a stationary Gaussian random field and we use this model to study the size and spatial distribution of the sets of nodes that observe statistically high data. This knowledge is exploited in order to ameliorate the data-implosion problem. Effectively, we implement a data suppression scheme that only lets nodes which sense statistically high data attempt to transmit their data to the

sink. Further, we also use our model to study network data that belongs to a given contour level and show that we can achieve further data suppression by only transmitting node data if it belongs to some predefined contour level. Finally, we show how the knowledge of the size and spatial distribution of statistically high node data in a WSN can be used to study the traffic in both schedule and contention based MAC protocols.

Acknowledgements:

I would like to thank my supervisor, Dr. M.K. Mehmet Ali for guiding me through the past two years. His patience, wisdom, and dedication were invaluable to the completion of this work. I would also like to thank my parents and brother for helping me out along the way whenever it was needed. Finally, I would like to thank all my friends I have made in Montreal and all the people I have played music with during the past two years.

TABLE OF CONTENTS

List of Figures

List of Tables

List of Notation Used

$X_i(\vec{s},t)$	Value of the random field at a node located at position $\vec{s}$ at time t.
$m_i(t)$	Mean value of the $i'th$ node at time t.
$C_S(\tau)$	The spatial covariance between the data at two nodes separated by distance, τ.
$C_T(\Delta t)$	The temporal covariance between the data samples of a node at two time instants, Δt seconds apart
α_1, α_2	Spatial correlation parameters
α_3	Temporal correlation parameter
$\sigma_i^2(t)$	Variance of the data at the $i'th$ node at time t
N	The number of nodes in the network
$\Re^2$	Two-Dimensional Euclidean Space
S	The region over which the network nodes are deployed
$A(S)$	The area of the network deployment region
λ	The (constant) density of network nodes, $\lambda = N / A(S)$
b	The critical threshold level (used to define excursion regions)
$A_{i,b}$	The area of the $i'th$ isolated region on S where data exceeds the level b
β_b	Global area on S where data exceeds the level b
K_b	The number of network nodes that exceed some threshold b. We call such nodes "excursion nodes"
ρ_j	Hop utilization factor, defined as the ratio of arrivals to the service rate at a hop.

Chapter 1 Introduction

1.1 Introduction to Wireless Sensor Networks

An emerging technology which is expected to play a great role in the future of wireless communications revolves around a simple device that unites ideas from wireless communications and embedded information processing in a unique way. This device is the wireless sensor node (or mote, in some literature). In its most fundamental form, a sensor node is capable of sensing its environment, information processing, and radio communication [1]. As sensor components become less expensive and smaller in size, a new vision in embedded computing is evolving: the vision of "Ambient Intelligence". This is where large numbers of sensors nodes will interact to gather and process data about the state of large systems and relay this data to a human user. The emergence of ambient intelligence is expected to significantly develop our experience and knowledge of person-to-environment and environment-to-environment interactions. What separates ambient intelligence from conventional embedded computer applications found in cars, washing machines, or HVAC systems, is communication: these sensor nodes are expected to be able to communicate with each other in order to relay their data to a human user [1]. Among the scientific community, a network of small sensor nodes deployed in large numbers to cooperatively observe an environment is referred to as a "Wireless Sensor Network (WSN)".

The trend towards low-cost sensor nodes means that there exists a potential to deploy these sensor nodes to monitor environments in high densities previously infeasible. By contrast, traditional environmental monitoring usually involves several expensive high-powered sensor stations, strategically deployed at large distances [2]. Each of these stations contains a data logger and the data recorded must be fetched either in person from the deployment site or through telemetry. Out of the monitoring accomplished with these traditional sensing systems, has evolved numerous physical models for various environmental phenomena. Since data gathering from large environments at high resolutions has so far been a costly undertaking, the models which exist are often rather coarse and a lot of interpolation and extrapolation techniques must be used to "guess" at the behaviour of the phenomenon at sites that are not covered by these traditional sensor stations [3], [4]. With the emergence of WSNs, for the first time scientists will be able to study complicated phenomena at previously unrealizable scales (both large and small) and to develop increasingly precise models for these phenomena. Before moving further, it is instructive to offer a more precise definition what we mean by a wireless sensor network.

1.2 Characteristics of Wireless Sensor Networks

A wireless sensor network (WSN) is a network of many sensor nodes deployed to monitor a region and an information sink(s) where the node data is sent to for processing by the application. To make it possible to deploy sensor networks at large scales and that monitor phenomenon that may only occur very infrequently, it is necessary that the nodes

have low power requirements. For example, it is expected that WSNs may monitor rare events like earthquakes, floods, and volcanoes whose recurrence time may be on the order of years or more [5],[6]. Thus, to justify its deployment in the first place, one must ensure that the network can live long enough to report on the next "big" event and does not find out after-the-fact that the network was not up to the task of monitoring the phenomenon. It is expected that the sensor nodes will often be powered by small, simple, low-power batteries. Thus, the processor, sensor, and radio unit will need to be fairly basic with correspondingly small power requirements. This means that the processing capabilities will be limited, the data storage will be small, the sensing device will be simple, and the radio will have a small transmission range. The nodes will often be deployed in an ad-hoc manner leading to unpredictable and complicated new topologies previously unexplored in communication networks. In many applications, it is envisioned that the nodes will just be "thrown" on to the field randomly for example, from an airplane [7]. The ability to deploy these nodes in such an ad hoc manner means that there is the potential to monitor regions inhospitable to direct human deployment like volcanoes, forest fires, and war zones, among others. The nodes also have the property that they are prone to failure due to battery depletion, or damage from the deployment phase or physical environment. The failure of single nodes leads to a dynamic network topology. However, this should not affect the ability of the sensor network to perform its task of monitoring a region. The replacement of individual dead nodes cannot be assumed possible in general, and these networks should be deployed with sufficient redundancy (i.e. nodes' sensing ranges overlap) to accommodate this. WSNs belong to the broader class of distributed networks. These are networks that function without

centralization through a hierarchical architecture. That is, it is expected that all the nodes will have the same capabilities in general and will have to cooperate with one another to transmit the data to the sink(s). Since the nodes have small power and therefore, small transmission ranges, it will often be necessary to use some form of multi-hop communication across the network for the nodes to deliver the data to the sink(s). The potential for different application settings for WSNs is enormous. We outline some of the more commonly proposed applications found in the literature and divide them into five broad categories [8], [1].

i) Environmental

- Disaster tracking and detection such as earthquakes, forest fires, landslides, tropical storms, and floods etc.
- Tracking the movement of endangered species.
- Monitoring phenomena that influence the output of an agricultural crop.

ii) Health

- Tracking/monitoring of patients, doctors in a hospital. By attaching a sensor to a patient, vital signs, heart rate, and blood pressure of a patient could be monitored
- WSNs could be used as an interface to the disabled and guide them in their daily actions. For example, sensors deployed about a metropolis could alert them to establishments that have the appropriate facilities to accommodate their needs (i.e. wheelchair ramps, elevators, etc).

iii) Home

- In the spirit of home automation, embedding sensors in common household devices can be used to give users remote control over these devices say via the Internet.
- Sensors can be used to track young children.

iv) Commercial/Industrial

- Control over temperature, humidity, and ventilation in industrial/office buildings via an interface between a WSN and the existing HVAC system.
- Monitoring the structural integrity of buildings or machines.
- Detecting and tracking car thefts.
- Attaching sensors to parcels of goods to monitor them as they travel to stores. Once in store, data from the sensors can be used to monitor the inventory of goods and even track sales.

v) Military

As with many new technologies, military applications were a strong driving impetus behind the initial conception of WSNs. The first application of WSNs was in the DARPA project for the U.S. Department of Defence.

- Monitoring the movement of troops (both friend and foe) by deploying sensors in the battlefield
- Attaching sensors to troops, tanks, and ammunition so that they can be distinguished from enemy troops.
- Nuclear, biological, and chemical attack detection.

1.3 WSNs and Monitoring Extreme Environmental Events

The possibility for high density deployments of sensor nodes on large regions offers enormous potential for WSNs in applications that involve monitoring extreme environmental events. These events include natural disasters (hurricanes, volcanic eruptions, landslides, avalanches, wild fires, etc.) and environmental catastrophes (oil spills, chemical leaks, radioactive contamination, etc). With their better coverage properties compared to traditional sensor networks, there is less risk that a phenomenon is undetected by the sensor nodes or is not detected fast enough (say when it has already spread beyond control). Also, the fact that these nodes are cheap means that we can deploy them redundantly enough so that small numbers of node deaths (say due to a hostile environment or battery death) will not greatly affect the ability of a WSN to monitor a region. This is in contrast to the traditional systems composed of a few nodes where the failure of a single sensing station will often cripple the network [5]. With the notion of global climate change an almost universally accepted fact in the scientific community, and the corresponding potential for more frequent occurrences of major

environmental catastrophes, the need for increasingly precise monitoring over larger scales and at cheaper costs seems more pressing than ever.

From the perspective of the application, the user will often be interested in such things as the intensity of the phenomenon at various points in space and how rapidly the phenomenon is spreading. This knowledge can be used to either prompt a response from a rescue crew, or issue a warning to individuals nearby the phenomenon of the potential danger. One of the chief problems with networks deployed in these sorts of scenarios is the bursty nature of the traffic. Due to the low frequency of these events, there are long periods of time that there is little for the network to report about. However, once the phenomenon strikes there may suddenly be large numbers of nodes that experience the phenomenon and wish to report their data to the sink. WSNs that behave in this manner are referred to as "event-driven" [1]. In fact, this great volume of data is not unique to event-driven applications of WSNs where the event is some major catastrophe. In large-scale periodic monitoring applications (say in agriculture), where the sensor nodes simply report their value at some regular time interval, there is the potential that large numbers of sensor nodes will attempt to transmit their data to the sink. A fundamental question to be addressed in this thesis is how to deal with this potentially large quantity of data. We outline several options below:

The first option is to just let every node in the network transmit its data to the sink and let the sink figure out which areas of the network area are the most critical regions. In general, this is not a practical solution in a WSN. Firstly, transmitting and receiving

are very costly operations with respect to energy in a WSN, and this approach will drain the node batteries as they not only transmit their own data packets, but also relay packets from many other nodes in the network. If a sufficient number of nodes die, then we may even experience network collapse where the network can no longer deliver any data to the sink. Even if significant numbers of nodes do not die due to the sheer volume of transmissions, this approach will likely create extreme contention for the wireless channel and consequently increase the latency of the data delivery. This latency is in addition to the latency already associated with having to route packets around dead nodes. In many applications, we can imagine that the data is time-critical and thus it is not desirable for the sink to be receiving vast quantities of data packets that are effectively out of date. Thus, high latency reflects poor network performance in delivering the data to the sink according to the needs of the application.

The second option in dealing with this large amount of data is to perform what is known as in-network processing. This is where the raw data is manipulated by the network nodes in some attempt to compress it. In WSNs, this compression can be achieved by exploiting spatial and/or temporal redundancy. However appealing this may seem, one must keep in mind that these nodes will have only basic functionality. Thus, performing in-network processing may be impossible with the limited processors found in sensor nodes, or simply too costly (say in terms of energy, time). The fact that this in-network processing must be done in a distributed manner on a large amount of data makes this task even more daunting.

The third option in dealing with this large amount of data is to distribute application specific code to the nodes so that when the phenomenon occurs, only nodes which observe values satisfying certain criteria will attempt transmission. For example, this can be done through the dissemination of queries [9] or by simply programming a set of transmission criteria into the nodes prior to deployment. This approach effectively limits the set of transmitting nodes so that only nodes who have sufficiently critical values attempt to transmit their data to the sink. This approach is appealing because of its simplicity and flexibility. It requires no complicated in-network processing and can be tailored to the user's needs by changing the transmission criteria in different application settings or at different stages of the evolution of an environmental event. This will be the approach that will be taken in this thesis. First we look at the current state of research in WSNs in environmental monitoring.

1.4 Current Applications and Visions of WSNs

We now take a look at completed or ongoing projects that seek to use WSNs to monitor potentially large scale environmental phenomenon with a large number of nodes.

In [10], [11] the researchers have developed the FireBug sensor node as part of the ITR (Information Technology Research) group through the National Science Foundation. Their research is intended to assist in the monitoring of wildfires. It is estimated that in the state of California alone, there are 5 600 wildfires every year that burn over 172 000 acres of land (in fact, as this section is being typed, [Oct 21, 2007],

there are currently reports on American news channels of a severe wildfire that is ravaging a wealthy area around Malibu, California and has begun destroying buildings on the campus of Pepperdine University). The research team envisions that these nodes will be deployed in the thousands, dropped from a helicopter or plane in front of an approaching wildfire. Traditionally, wildfires have been hard to track on account of the hostile environment they create for humans due to intense heat and smoke. The deployment of these nodes will allow precise monitoring of the fire as it evolves and enable the development of more detailed fire models to predict the speed and direction of fire spread.

In [12], [13], a team from the Department of Primary Industries (DPI) has claimed to have deployed the largest (in terms of density) functioning WSN to date. The project was launched in July, 2006 and is expected to conclude in June, 2008. The project addresses a variety of environmental and agricultural issues. Their goal was to build a long-term, spatially dense sensing environment in order to monitor the performance of a production system (in their case, a commercial orchard). The WSN consists of 273 sensor nodes strategically deployed on a nectarine orchard. The network consists of different types of sensor nodes that could collect information on soil moisture at various depths, soil and air temperature, canopy temperature, humidity, leaf wetness, trunk diameter, wind speed and direction, solar radiation, and GPS coordinates. The effect of variability in soil moisture over a growing season was used to study the corresponding effect on the variability in the canopy cover and fruit yield.

In [14], [15], Australian, John Langford, director of the Melbourne Water Research Centre and a team of researchers from other organizations have begun a project to use a WSN to increase efficiency in current irrigation systems. This project is in response to the likely reality in the near future of periods of severe water shortages. Currently, most consumers rely on food coming from so-called food-producing regions where it can be produced in large quantities and then imported to their homes, potentially thousands of miles away. The consequences of wasting what little water remains during a drought in these food-producing regions could be very dear to citizens and the economies of both regions (the food-producing regions and where the food gets imported to). Their project seeks to use a WSN to send water to a crop on demand, instead of by some scheduled roster as is currently the practice. In place, their system uses water more efficiently than conventional irrigation. As a result, better crop yields can be achieved with less water use. They predict the possibility of 10-20% water savings using their system. They have developed the NICTOR sensor node platform that uses soil moisture sensors, air temperature sensors, and leaf sensors. These nodes use gateways to interface to an irrigation drip system through the internet in order to activate the delivery of the appropriate amount of water to a site when needed.

In [16], researchers have developed a WSN to be deployed along a flood protection zone in Wilkes-Barre, Pennsylvania in order to monitor the condition of the levy system in place there. Their goal is to use the WSN as an early warning system for floods. The nodes of the WSN will monitor the water level, vibrations, and pressure that the river exerts on the levy. The researchers point out the problem with the existing

monitoring network based on traditional technologies: that is, the existing sensing network that monitors the levy is hard-wired and only covers a limited area along the protected zone of the flood plain. Also, the existing system only looks at pressure and water level. However, a more detailed understanding of hydrodynamics is required in predicting an oncoming flood. Also, because of the limited number of pressure sensors nodes, the nodes only offer partial spatial coverage along the levy. If a levy breech occurs upstream, the pressure sensed by the nodes will indicate this however, only minutes prior to the event hitting the protected zone. Thus, the current system cannot warn the user early enough that a potentially catastrophic flood is about to arrive. Their experimental research has focused on the throughput and longevity of the sensors at different data sampling rates, loads and node configurations. They are also drafting reports on the battery life as a function of the transmission rates and on the reliability of transmission. They are placing the nodes at various distances and the signal strength is monitored by looking at the data throughput. Finally, they are looking at the experimental latency of the data delivery. This experimental work is being conducted to allow WSN designers to quantify the network's ability to perform and use this to help determine the overall structure of the final network that is designed.

1.5 Research Objectives and Contributions of the Thesis

Much of the existing literature on wireless sensor networks is devoted to developing specific protocols related to MAC and routing. The usual goal of many of these protocols is to achieve a design which minimizes energy consumption in the

network and therefore prolongs network lifetime. However, despite the flurry of academic papers that come out each year on topics in sensor networks, there is little exploration of the issue of how a sensor network will nominally react to an observed phenomenon. That is, how does the underlying phenomenon being observed by our network influence such things as the nature of the data being reported, the load this data places on the network, and the consequent performance of the network when monitoring this phenomenon? The reaction of a WSN to the observed phenomenon is a fundamental issue in addressing whether or not a given network with a set of protocols will succeed in its task of monitoring a phenomenon. The sorts of phenomenon we will often be interested in are ones that act as some disruption of the steady-state of the environment being monitored. Examples of these could include volcanic eruptions, earthquakes, tropical storms, floods, a levee collapse, contamination of a water source, etc. However, our work is equally applicable to sensor networks that observe some "ambient" quantity like temperature, humidity, pollutant concentration that do not necessarily occur as some disruption to the environment. Throughout this thesis, we will use the term "phenomenon" rather loosely in order to refer to both ambient observed quantities in the environment as well as those events which occur as a disruption to the environment although our focus is often on the latter.

As noted earlier, WSNs can be anticipated to experience great traffic loads in response to observing extreme environmental events. From the perspective of the user of a sensor network, knowing whether or not the network can handle this load is of fundamental importance. If this load is too great and congestion occurs, packets

containing data and routing messages may be lost and valuable information concerning the phenomenon may arrive too late or never actually reach the sink. The consequences of this could be dire if for example a sensor network is being used to monitor a fragile environment or in order to warn of the spread of a disaster to areas inhabited by humans.

This thesis will seek to answer the following questions:

- What is the nature of the data values observed by the sensor network when observing some phenomenon?
- How many nodes will want to transmit their data to the sink in order to meet the requirements of the monitoring application?
- What is the spatial distribution of nodes who will want to transmit their data to the sink?
- What is the traffic load placed upon the network in response to the underlying phenomenon being observed in the context of the requirements of the monitoring application?

To answer these questions we will employ techniques from a branch of modern probability theory known as random field theory. This theory will provide us with a stochastic model for the data observed by the nodes as the phenomenon evolves in time and space and will offer insight into numerous quantities of interest to the designer of a WSN. Formally, the contributions of this thesis are as follows:

- The development of a model for sensor node data in a WSN observing an environmental phenomenon based on the underlying models of the physical phenomenon and the node deployment.
- An analytical expression for the distribution of number of nodes in a WSN that report a value above some application-specified threshold. As well, the spatial-distribution of these nodes on the network deployment area is analyzed.
- An analytical expression for the average number of nodes in a WSN that belong to contour lines of application-specified values.
- The application of our data modeling to study the traffic characteristics of a WSN in terms of traffic load, packet drops, single-node delay, and source-to-sink delay.

1.6 Related Work

One central theme of the current work is to find a natural way to model the data in a WSN and use this model to find the size of the set of nodes that transmit their data to the sink. As such, our work falls under the broader categories of modelling and data suppression in WSNs. Several authors have undertaken work in this same spirit and we outline these here.

In [17], the authors deal with the topic of exploiting inherent correlation (spatial and temporal) of the data in order to develop efficient communications protocols,

particularly at the MAC and transport layers. They develop a theoretical framework to model both spatial and temporal correlation of the network data. In the spatial domain, they then study the distortion incurred by representing a point-source signal with the observations taken by sensors from a subset of the total node population, located at various distances to the point-source. Their main conclusion is that the number of nodes that report data observed from a point-source can be significantly lowered from the total number of nodes in the network without incurring significant distortion at the sink. They then discuss how the locations of the representative nodes should be chosen over space so as to reduce this distortion even further. In a related paper by the same authors [18], they propose a MAC protocol (the so-called CC-MAC) which exploits the inherent spatial redundancy of the node data. As their work is concerned with limiting the set of transmitting nodes that report in the network their work is in the same spirit as ours. However, their work is designed for tracking a point-source phenomenon and deals with the accurate description of this source with a subset of the network nodes. On the other hand, our work is designed to study phenomenon that cover a large spatial region and is more concerned with finding a natural way to determine the size of the set of nodes which have record critical (i.e. "extreme") data.

In [19], the authors use the idea of representing a spatial phenomenon with contour lines to reduce number of nodes which transmit data. In their scheme, a node listens to transmissions in its listening neighbourhood for an interval of time inversely proportional to the value it has sensed. If it overhears a value within some user-defined threshold, then the node will suppress its own transmission. This has the overall

tendency that higher values get transmitted and lower values within the threshold of the overheard value are suppressed. The method they outline for forming the contour lines is completely distributed. In the network itself, nodes have no concept that they belong to a contour line and which other nodes belong to the same contour line. Instead, the set of contour lines are formed at the sink. While their work is concerned with protocol-design and ours with data modelling, the work in this thesis compliments their work. What we provide is a natural, mathematical framework for understanding how the contour lines for a given level exist in space (i.e. their size, and the number of contour lines for a level that exist on the network deployment area). This understanding will allow us to predict the number of nodes that belong to each contour line for a given level. Due to the distributed manner in which the contour lines are formed in their protocol, there is no way to determine what the values of the contours will be at the sink. Our work is more suitable to applications where the contours are instead formed according to the needs of the user say via the dissemination of application thresholds to the network nodes or through queries from the sink.

In [20], [21], the authors propose a method for constructing synthetic traces of network data at the sensor nodes. This work is primarily in response to the fact that there is little available real data sets from sensor networks since there have not been many deployments of WSNs to date. In this sense, their work assists other researchers working in WSN protocol design and allows them to be able to test their work on data that may closely resemble actual sensor network data. Many existing MAC and data aggregation protocols (see for example [18], [9]) exploit the spatial correlation of sensor network

data. The authors (in [20], [21]) then examine the performance of these protocols using synthetic correlated data that they create using their algorithm. Their method allows the data to have an arbitrary underlying first order probability distribution function, subject to a user-specified correlation structure. [20] shows how to perform this synthetic generation on a grid-topology of nodes. In [21], the same authors show how to extend this to a network with an irregular topology. While their work is more concerned with the generation of synthetic data given the underlying distribution and correlation structure, our work is more concerned with the initial modeling of the underlying phenomenon and network data and the network's response to this data.

1.7 Organization of the Thesis

Next, we present the organization of the thesis.

Chapter 1: Introduction

In chapter 1, the basic concept of a WSN and its characteristics are presented. Particular attention is paid to the motivation of WSNs from the perspective of monitoring in the physical sciences. Current projects related to environmental monitoring that use WSNs are discussed as well as related work with a similar flavour as the current thesis.

Chapter 2: Sensor Node Architecture and MAC Protocols

The individual node architecture is discussed along with considerations for the MAC protocol design.

Chapter 3: Data Modelling in a WSN

In this chapter, we introduce the use of random field theory to model the environmental phenomena that WSNs observe. Several different forms of the model for the underlying phenomenon are considered to show that this modeling can be flexibly adapted to a range of different application scenarios. We then use the model for the underlying phenomenon to construct a model for the data that the sensor nodes experience in a WSN deployment.

Chapter 4: Applications of Excursion Sets of Gaussian Random Fields to WSNs

In this chapter, we use our model for the node data in a WSN and random field theory to find the distribution of the number of nodes who will want to transmit their data to the sink under two different transmission criteria. The first criterion assumes that only nodes who observe values greater than some application threshold will transmit. The second criterion uses a contour line approach so that only nodes who report values separated by some threshold try to report their data to the sink. Relevant simulation results are presented. Then we examine how the number of nodes that observe a value greater than a threshold changes in time in response to a time-varying phenomenon.

Chapter 5: Extension to Non-Stationary Phenomenon

We show how our modelling and analysis of the number of excursion nodes can be extended to cases where the underlying phenomenon is of the point source type and the nodes experience a distance-scaled version of the phenomenon that occurs at the point source.

Chapter 6: Applications of Excursion Sets to WSN Traffic

Here we show how one can use the knowledge from our modelling and analysis in previous chapters to study the nature of traffic in a WSN that observes a stationary Gaussian random phenomenon. Applications are shown for WSNs that use either contention-based or schedule-based medium access. As well, system performance is analyzed.

Chapter 7: Conclusions and Future Work

In this final chapter, we summarize the main contributions of the thesis and offer possible insight into future research areas that can use the modelling and analysis discussed in our work.

Chapter 2 Sensor Node Architecture and MAC Protocols

The central component at the heart of a wireless sensor network is the sensor node. Thus, it is worthwhile to gain a more precise understanding of what these nodes are. As well, we will be interested in the fundamental characteristics of MAC protocol design in WSNs.

2.1 The Sensor Node

A sensor node has five primary components. These are the controller, data storage, communication and sensing devices, and a power supply. Next, we describe each of these components.

i) The Controller

The task of the controller is to process sensor network data and execute program code. A wide variety of controller types exist including the microprocessors found in common desktop computers, microcontrollers, and FPGAs. However, due to their low energy consumption and their suitability to embedded applications, microcontrollers are generally regarded as the most appropriate choice of controller-type. Microcontrollers are flexible in interfacing with a wide variety of other devices. In the context of sensor networks, we will want to be able to interface with a range of possible sensing devices (i.e. temperature, humidity, pressure, salinity, acoustic, etc.) as well as to a radio

transceiver. Microcontrollers are especially suitable to WSNs because they can enter low power modes where the parts of controller are put to sleep (are made inactive) and therefore offer a source of energy savings [1].

ii) Data Storage

The memory component to store data in a sensor node will be fairly basic. There is need for Random Access Memory (RAM) to store sensor readings at the node and data packets received from other nodes. The principal advantage is that RAM can read and write data fast. However, if the power supply is lost even temporarily, the contents of RAM are lost. The instruction set and program code of each node can be stored in Read-Only Memory (ROM), but more likely in Electrically Erasable Programmable ROM (EEPROM) [1].

iii) Communication

The first issue to address is: In which wireless medium will sensor nodes transmit and receive their data? It is expected that Radio-Frequency (RF) communication will dominate the share of transmission mediums chosen in actual applications of WSNs. The reason for using RF communication in the context of WSNs is that relatively long range communication is possible between nodes, high data rates can be achieved if necessary, and the error rates of the channel are reasonable with low energy expenditure.

The device that is responsible for communication in a sensor network is the node transceiver. A transceiver is a combined device consisting of an RF transmitter and receiver. In the transmitting direction, the transceiver's job is to convert digital data, stored in the nodes' memory, into radio waves that can be sent to other nodes. In the receiving direction, the transceiver must accept incoming radio waves and convert them to a digital stream of bits to be stored in memory. Half-Duplex operation is expected and therefore, nodes cannot transmit and receive at the same time [1].

We now discuss some properties and tasks required of transceivers for WSNs outlined in [1]. Transceivers must be able to switch between different states in order to achieve energy savings. Common states include Active, Idle, and Sleep. The active state encompasses when the transceiver is either transmitting data or has its receiver tuned in and is accepting incoming data. The idle state is when the receiver is tuned in, but not currently receiving any data. During this state, we can turn off the parts of the receiver circuitry associated with data acquisition and turn these on only when the receiver detects an incoming packet. In the sleep state, significant parts of the transceiver are switched off. Different transceivers often have different levels of sleep states. The deeper the sleep state, the more energy and delay are incurred to awaken from the sleep state. However, these costs may be justified if more energy can be saved in the long run. In general, there is always an energy expenditure associated with a transition between any two states. Thus, the frequency of switches should be reasonable. Also, the times taken to switch states incur extra delay in transmitting data to the sink so these must be taken into account as well.

Data rates in WSNs are expected to be quite low compared to typical forms of broadband wireless communication. This is due to the fact that the sensor nodes are not powerful devices. Expected data rates are on the order of tens of kilobits per second. For example, [1] shows different sensor node products with data rates between 10 kbps and 150 kbps. The choice of carrier frequency is not fixed although the ISM band has been proposed. The choice will depend in large part upon government regulations. However, it is expected that communication will take place somewhere in a band between 433 MHz and 2.4GHz. Low power constraints also imply that the transmission ranges of the nodes will be comparatively small, on the order of tens of metres [1].

iv) Sensing Device

It is hard to specify the exact form of a sensor in a WSN setting because this will depend solely on the application setting. Three broad categories of sensors are listed in [1]. Passive, omni directional sensors measure a physical quantity at the location of the node without manipulating the environment by probing. These sensors can often be powered from the physical environment alone, energy only being needed to amplify their signal for data transmission. Typical examples include thermometers, light sensors, vibration sensors, humidity sensors, chemical detectors, smoke detectors, etc. These sensors cannot discern the directional component of data. Passive, narrow-beam sensors are also passive (hence, the name) but can assign a sense of direction to measurements. An example is a camera which can take measurements (i.e. pictures) in different directions by rotating the camera. Active sensors manipulate the environment usually in some non-

intrusive way. For example, some seismic sensors will generate small shock waves through controlled explosions. Once the category of sensor is determined depending on application, one must make tradeoffs among accuracy of measurements, cost of sensors, the dependability in the application environment, delay of data delivery, and node size among other considerations. Based on the extant literature, passive, omni directional sensors seem to be the most appropriate sensor type in the widest array of applications. Generally, sensors are assumed to have a certain coverage range within which they can predict the phenomenon with reasonable accuracy. This assumption is based on the fact that many phenomena expected to be monitored will exhibit spatial correlation. This can be exploited to reduce either the number of nodes deployed in a sensor network or perhaps more realistically (if we want to design networks robust to node failures), to reduce the set of nodes which actually report about a phenomenon.

v) Power Supply

Batteries are the source of power for the sensor nodes. There are two classes of batteries; non-rechargeable and rechargeable. In the latter case, some form of energy scavenging device is needed on the node to convert energy from the physical environment (for example via solar energy, vibrations, air flow, or temperature gradients). The reality of severe energy constraints in WSNs means that the batteries must meet very demanding requirements. They must possess high capacity (of stored energy) at a small weight, size, and low price. The capacity is represented in energy per unit volume. Table 2-1 [22] shows the energy density for some common non-

rechargeable and rechargeable batteries. Also, the batteries must be able to provide sufficient instantaneous loads to accommodate higher intensity traffic and more active node functionality when required.

Table 2-1: Energy Densities for Common Batteries

Chemistry	Zinc-Air	Lithium	Alkaline
Energy Density (J/ccm)	3780	2880	1200
Rechargeable Battieries			
Chemistry	Lithium	NiMHd	NiCd
Energy Density (J/ccm)	1080	860	650

2.2 MAC Protocol Design for Wireless Sensor Networks

In our research, we will be indirectly interested in the MAC layer in order to provide design guidelines to WSN designers. Specifically, this knowledge will help us later when we consider traffic characteristics in WSNs. In communication networks, MAC protocols often fall into two broad categories and WSNs are no exception. We describe these below:

i) Random Access Protocols

Here, nodes access the channel in an uncoordinated manner. The behaviour of random access protocols is often dependent on the behaviour of packet arrivals to the nodes and node timers that dictate when to try and transmit a data packet. In WSNs, most existing contention-based protocols are of the CSMA-type. Common to all CSMA-

type protocols is that they attempt to sense whether the channel is busy (via carrier-sensing) before transmitting their packets over the channel. There are several variants of CSMA that differ in how they back-off when they detect that the medium is busy. As well, there is often some form of collision avoidance mechanism that seeks to reduce the hidden-terminal problem that is a characteristic of wireless channels. The hidden terminal problem occurs when two nodes out of range from each other try to transmit simultaneously to nodes that are in the range of both transmitting nodes. Even though the transmitting nodes both detect an idle channel, their data packets collide at the receiving nodes. This problem can be alleviated by the use of a RTS(Request-to-Send)/CTS(Clear-to-Send) handshake protocol [23].

ii) Schedule-Based Protocols

These protocols divide time into frames and further subdivide each frame into time slots of fixed duration. During each time slot, a node has exclusive access to the channel. These protocols require the dissemination of schedules to the nodes. These schedules contain the information of when each node can transmit and should listen to receive data. Time Division Multiple Access (TDMA) is a common MAC protocol based on these principles.

2.2.2 Energy Concerns in MAC Protocols for WSNs

The unique needs of WSNs make the use of traditional schemes for MAC design inappropriate. Chief among these needs is the need for nodes to conserve energy in order

to prolong the network lifetime and succeed in monitoring the underlying environment. Traditional performance metrics such as throughput efficiency, stability, user fairness, low access delay, and low end-to-end transmission delay may be traded off in order to achieve greater energy savings [1]. Indeed, in the literature of proposed MAC protocols this is almost always the case.

There are four main sources of energy waste in wireless sensor networks which we now summarize [24]. Collided data packets waste energy both through their initial failed transmission, and the need to perform (possibly unsuccessful) attempts at retransmission. In contention-based protocols, collisions are particularly prevalent when the traffic load on the network is high. If the load is low enough, collisions may not be a significant source of energy waste. The second source of energy waste is packet overhearing. Since WSNs use a broadcast medium, nodes can overhear packets that are not destined to them and then must drop these packets, thus wasting energy at the node's transceiver. Overhearing avoidance is particularly important at high network densities where a node may overhear a significant amount of irrelevant traffic. The third source of energy waste is MAC protocol overhead associated for example with disseminating transmit/listen schedules (TDMA-types) or handshaking (CSMA-types). The last source of energy waste is the idle listening. This occurs when a node is waiting to receive a packet but is not actually receiving anything wasting energy at the transceiver. When the traffic load is low, idle listening periods are particularly wasteful. A common solution is to periodically switch the transceiver between sleep and on states.

There are both advantages and disadvantages to using either a contention-based or schedule-based MAC protocol in WSNs. The principle advantage of schedule-based protocols is that theoretically, there are no collisions and the energy waste associated with them. As well, nodes can enter low-power states when it is not their slot to transmit or receive a packet. The principle disadvantages are related to complexity. All nodes' clocks must be tightly synchronized so that transmissions do not interfere. This is a rather difficult problem, especially in a WSN that may have a large number of nodes. There is also an overhead-cost associated with having to disseminate the transmit/listen schedules to all nodes. Finally, the topologies of WSNs are dynamic with nodes going to sleep, waking up, dying, and possibly being added. This means that the scheduling must respond to these changes with dead or sleeping nodes having to relinquish their time slots to new or awoken nodes. The principle advantage of contention-based MAC protocols, is that they are highly scalable and can respond to the sorts of changes in the network topology described above. The principle disadvantage of contention-based MAC protocols is the energy wasted due to collisions and overhearing when the traffic load is high. Even under low traffic conditions, there can be significant energy waste due to idle listening in contention-based MAC protocols.

An example of a contention-based MAC protocol that seeks to address these energy issues in a WSN is the S-MAC protocol [24]. The LEACH protocol is a schedule-based MAC protocol that addresses these energy issues [31].

Chapter 3 Data Modelling in a WSN

Wireless sensor networks (WSNs) are envisioned to be data-centric networks [1]. In this respect, it is not the identities of nodes who wish to transmit data to the sink that will be most important to the application, but instead the values of the data sensed at these nodes and combined with the nodes' geographic locations. In a sense, if one wishes to understand how well a sensor network will perform in its task of monitoring an environment, then it is first necessary to understand the nature of the data values that are observed in that environment and how these values are distributed in space. This understanding can aid in predicting important network performance metrics for deployed WSNs. The research contained in this chapter develops a model for sensor network data under certain fairly common application scenarios.

3.1 Application Scenario

As is often stressed in the literature of WSNs, the design and performance of a sensor network is intimately tied to the intended application of the network. Here we outline the overall intended application scenario which our model will later accommodate. We assume that we have randomly deployed a large number of wireless sensor nodes in a geographical area to monitor an environmental phenomenon that exhibits some degree of spatial and possibly temporal correlation. For the most part we will be concerned with aspects of spatial correlation since this will offer insight in terms of how the important data in the network is distributed on the deployment area. We will

develop our model with the intention of analyzing application scenarios that involve sampling the nodes at fixed intervals. At each sampling time, the sink will be particularly interested in receiving data from groups of geographically "close" nodes that report values that are extreme in either the statistical sense, (in other words, values that deviate far from the mean) or just in the physical sense (not necessarily statistically extreme, but above some fixed critical threshold). In many applications related to environmental monitoring, we envision that the sink may only be interested in receiving data from nodes that record extreme values. The reasons for this are two-fold. Firstly, extreme values at sensor nodes will correspond to points in the network deployment where the monitored phenomenon is most severe. Secondly, since the nodes are severely energy-constrained there may be scenarios where it is impossible for all nodes to transmit their data. In this latter case, it makes sense to only transmit values which are more critical. We comment that in the sciences dealing with spatial variation, it is not necessarily just a high value at a single point in space which is of interest. Rather, the user is often more interested in whether there exists a region in space of significant size that exhibits data of extreme nature [3], [25]. In the literature, the occurrence of a distinct region in space that exhibits high values is referred to as a "high-level excursion set". Thus, we are not just interested in isolated extreme values but rather, in extreme values in the network that exist over regions of some appreciable size.

Through an analysis studying how large these extreme regions are and how many nodes are deployed on them, we will essentially be addressing the nominal number of nodes in the network which will attempt to send their observed data to the sink in our

application scenarios. Of course, if the data is sufficiently correlated there may be the possibility to further restrict the set of transmitting nodes. For example, if there are already a sufficient number of nodes observing values in an extreme region to provide the desired monitoring resolution to the user, then it may be possible suppress some of the data from these extreme regions as well.

3.2 Random Fields

In this section we provide a non-mathematical description of how random fields arise in the physical sciences. For basic mathematical concepts in random field theory, we refer the reader to Appendix A for a more detailed rendering.

As discussed in the introduction, it is anticipated that one of the most ubiquitous applications of wireless sensor networks will be in environmental monitoring. Many phenomena occurring in meteorology, geosciences, and hydrology exhibit complex random variation in space and/or time [25]. Much of the literature that exists on the theory of random behaviour is centred on random processes which occur in a single dimension (usually time). The generalization of the existing theory in one dimension to higher dimensions is usually non-trivial and is a relatively new advent of probability theory beginning in the 1950's with Longuett-Higgins work studying the random properties of wave heights [26]. A rigorous probabilistic setting of theoretical results is even more recent, taking place in the 1970's and 1980's through the works of Nosko, Adler, and VanMarcke among others [25], [26]. The branch of probability theory called

random field theory can be used to extend and adapt common ideas in random process to higher dimensions. Random fields are used to model systems that exhibit random variation in a multidimensional setting.

Whereas the deterministic models which pervade much of the literature on modelling environmental phenomena tend to fail when predicting the behaviour of a phenomenon on either very small or very large scales (say extremely small/large times or spatial distances), random field models accurately capture the behaviour of a phenomenon at *all* scales. For example, the units of time could be the times between collisions of molecules or the time between major geological catastrophes in temporal phenomenon [25]. Similarly, the units of space could be as small as brain cells or as large as cities. As such, random fields have seen applications in such diverse areas as forestry, geomorphology, geology, turbulence, and seismology [26].

Consider the following phenomena that are amenable to modelling with random fields [25]:

- Depth of snowfall across a surface during a snow storm
- Pollutant concentration in a lake
- Shear stress along a fault line in the earth
- Height of the ocean surface
- Amount of recoverable solar energy
- Arial density of the population of a species
- Agricultural crop yield

- Distribution of rainfall on a crop
- Inflow of water into a reservoir
- Density, porosity, and permeability of soil
- Intensity of an earthquake

All of these phenomena can be characterized by possessing disorder in both space and/or time. In the literature, a phenomenon possessing this property is known as a distributed disorder system. The reality of random variation in space and time leads to the fact that eventually there will be occurrences of values with significant deviation from the so-called expected value of the phenomenon. In space-time phenomena, these occurrences can correspond to regions in space, time, or both simultaneously. For example, there may be a region of space which has values that significantly deviate from the mean value in a given time window. We can easily imagine that the size and magnitude of these extreme regions will be of paramount interest to the user who has deployed a sensor network to monitor environmental hazards. Fortunately, a central topic in random field theory is the analysis of the size of sets of the underlying parameter space that exceed some (preferably high) level (see Appendix A8).

3.3 Modelling the Observed Phenomenon

Recall that it is one of the goals of this thesis to study how the underlying physical phenomenon that the network observes will affect the ability of the network to perform its task of delivering data to the sink. For this reason, the purpose of this

chapter is to establish a model for the underlying nature of the phenomenon being observed by the sensor network. We will assume the reader has a basic familiarity with concepts from random field theory such as finite-dimensional distributions, covariance functions, Gaussian random fields, stationarity, isotropy, separability, and excursion sets. If necessary, this material can be found in Appendix A.

Throughout most of this work, the sorts of phenomena we will be primarily interested in are those that vary in time and act over a very large global region (larger than the network radius) and can be assumed to be fairly homogeneous in space. By homogeneous, we mean that the mean and variance of the phenomenon is more or less constant over the network deployment area over a fixed time interval. We will also assume that the phenomenon decays over time. We note that many meteorological phenomena have the aforementioned properties. It will be assumed that the phenomenon being observed is initially in some steady state and then a disturbance happens that perturbs the environment; for example, this could be the occurrence of a rain storm or hurricane. To accommodate these considerations, we suppose that the underlying phenomenon can be modelled as a three dimensional random field. The first two coordinates will parameterize the two dimensional space covered by the network deployment area, which we denote throughout this thesis by $S \subset \Re^2$. The third coordinate will parameterize time, $T \subseteq [0,\infty)$. We denote the value of the random field at position $\vec{s} \in S$ and time $t \in T$ by $X(\vec{s},t)$.

In addition to having a simple specification in terms of mean and covariance function, it is known that Gaussian random fields are reasonable models for a wide array of physical phenomena [27]. Thus, we will assume further that the underlying random field is Gaussian. We note here that we will use the same Gaussian random field model for the data at points in S during the time before and after the disturbance to the system occurs (for example, before the rain storm begins). Since Gaussian data is completely specified by its mean and covariance, the disturbance striking the network deployment area will correspond to some altering of these statistics from their initial state. To provide some momentum to creating our model, we will now specify some general possibilities on how this perturbation manifests itself in terms of its effect on the first order statistics and provide more details in the following two sections.

It will be assumed without loss of any generality that initially, the mean value of the field before the phenomenon strikes is fixed at zero. We then suppose that the disturbance strikes at time $t = 0$ and denote the time just before the disturbance as $t = 0^-$. Further, we assume that the mean value of the random field is a function of time only. This corresponds with our intuition of the phenomenon occurring on a region much larger than the network deployment and being fairly homogenous on the network deployment area. We will look at precisely how the mean value of the field will evolve in time in a later section of this chapter.

We now turn to the topic of determining some basic properties that should be encapsulated in our model for the covariance function of the random field. We first make

some preliminary definitions. For a fixed time t, the mean value of the random field at a location $\vec{s} \in S$ is given by:

$$m(\vec{s},t) = E[X(\vec{s},t)] = \int_{-\infty}^{\infty} x \cdot f_{X(\vec{s},t)}(x)dx \tag{3.1}$$

We define the space-time separation vector as:

$$\vec{d}((\vec{s}_1,t_1),(\vec{s}_2,t_2)) = \left(d_{\Re^2}(\vec{s}_1,\vec{s}_2), d_{\Re^1}(t_1,t_2)\right) \tag{3.2}$$

where $d_{\Re^2}(\cdot,\cdot) = \sqrt{(s_1^{(1)} - s_2^{(1)})^2 + (s_1^{(2)} - s_2^{(2)})^2}$ and $d_{\Re^1}(\cdot,\cdot) = |t_2 - t_1|$ are just the Euclidean distance between $\vec{s}_1$ and $\vec{s}_2$ in space, and the time difference between t_1 and t_2, respectively. To make notation more compact, we will represent the distance in space by τ, and the time difference by Δt. That is we let

$$\tau = d_{\Re^2}(\vec{s}_1,\vec{s}_2) \tag{3.3}$$

$$\Delta t = |t_2 - t_1| \tag{3.4}$$

The covariance function of a random field is used to capture the joint behaviour between any two locations $\vec{s}_1,\vec{s}_2 \in S$ at two different instants in time, $t_1,t_2 \in T$. The covariance function between two points $(\vec{s}_1,t_1),(\vec{s}_2,t_2) \in S \times T$ is defined as:

$$C((\vec{s}_1,t_1),(\vec{s}_2,t_2)) = E[X(\vec{s}_1,t_1)X(\vec{s}_2,t_2)] - m(\vec{s}_1,t_1)m(\vec{s}_2,t_2)$$

Often in modelling physical phenomenon in space and time [25], [27], it is assumed that the covariance function is separable and we assume this to be the case in our model. Separability allows us to represent the covariance function as the product of

two independent covariance functions; a spatial covariance function and a temporal covariance function. In this case, the covariance between points $(\bar{s}_1,t_1),(\bar{s}_2,t_2) \in S \times T$ in the random field is given as:

$$C\big((\bar{s}_1,t_1),(\bar{s}_2,t_2)\big) = C_S(\bar{s}_1,\bar{s}_2) \cdot C_T(t_1,t_2)$$

where $C_S(\bar{s}_1,\bar{s}_2)$ and $C_T(t_1,t_2)$ are (as yet, unspecified) spatial and temporal covariance functions. The only further requirement on these two functions is that they are positive definite, which is necessary if they are to be valid covariance functions in the first place. Recall that the product of two positive definite functions is also positive definite.

In addition to being separable, we will also assume that the covariance function is isotropic in space. This means that we can represent the spatial covariance function as function of the distance between points $\bar{s}_1,\bar{s}_2 \in S$ which we defined earlier as $\tau = d_{\Re^2}(\bar{s}_1,\bar{s}_2)$. We can therefore write:

$$C\big((\bar{s}_1,t_1),(\bar{s}_2,t_2)\big) = C_S(\tau) \cdot C_T(t_1,t_2) \tag{3.5}$$

We will specify some possibilities for $C_S(\tau)$ and $C_T(t_1,t_2)$ in the following two subsections.

3.3.1 Spatial Variability

In this section we will study how the observed phenomenon varies in space for a fixed time. In our model, all aspects of spatial variability in the data are contained in the

isotropic spatial covariance function. There are a wide variety of possible spatial covariance models available and it is even possible to construct new models from existing models. Each covariance model seeks to capture how fast the covariance dies with increasing distance between two points, any inherent symmetry in the data, and the continuity and differentiability properties of the field. The choice of model is largely dependent on the physical nature of the phenomenon being observed. Thus, there is not necessarily a single correct model and our choice is in a sense, arbitrary. However, we will want to choose a model which has certain desirable properties that are both convenient in terms of analysis and correspond to our physical intuition of an environmental random phenomenon that evolves in time on a subset of two-dimensional space as previously described. In [17], [18], [28], four popular spatial covariance models are shown which we reproduce below. These models use two correlation parameters: The range parameter, α_1, controls how fast the spatial correlation dies with distance and the roughness parameter, α_2, controls the smoothness of the sample paths of the random field.

Spherical Covariance

$$C_{Sph}(\tau) = \begin{cases} 1 - \frac{3}{2}\frac{\tau}{\alpha_1} + \frac{1}{2}\left(\frac{\tau}{\alpha_1}\right)^3 & \text{if } 0 \le \tau \le \alpha_1 \\ 0 & \text{if } \tau > \alpha_1 \end{cases} \tag{3.6}$$

In this model, two observations taken at a distance greater than α_1 are uncorrelated. This model only uses the range parameter.

Power-Exponential Covariance

$$C_{PExp}(\tau) = e^{(-(\tau/\alpha_1)^{\alpha_2})} \text{ for } \alpha_1 > 0,\ 0 < \alpha_2 \leq 2 \tag{3.7}$$

Rational Quadratic Covariance

$$C_{RQ}(\tau) = \left[1 + \left(\frac{\tau}{\alpha_1}\right)^2\right]^{-\alpha_2} ; \alpha_1 > 0, \alpha_2 > 0 \tag{3.8}$$

Matérn Covariance

$$C_M(\tau) = \frac{1}{2^{\alpha_2 - 1}\Gamma(\alpha_2)}\left(\frac{\tau}{\alpha_1}\right)^{\alpha_2} K_{\alpha_2}\left(\frac{\tau}{\alpha_1}\right) ;\ \alpha_1 > 0, \alpha_2 > 0 \tag{3.9}$$

Where $K_{\alpha_2}(\cdot)$ is the modified Bessel function of second kind and order α_2.

Note that these models feature two parameters: α_1 which is a range parameter controlling how fast the covariance dies with distance and α_2 which is a roughness parameter that controls the smoothness of the sample paths of realizations of the random field.

Earlier, we assumed that the spatial covariance function is isotropic in space for a fixed time in our model. Recall from Appendix A that isotropic covariance functions are a subset of stationary covariance functions and that strict-sense stationarity and wide-sense stationarity are equivalent for Gaussian random fields [27]. Further, as was seen in Appendix A.8, we can study high level excursion regions if the random field is mean square differentiable. Thus, we will want our model to possess this desirable property.

In Appendix B, we examine the mean-square differentiability of Gaussian random fields when they are correlated according to each of these four spatial covariance functions. For the sake of brevity in our work, we will use the rational quadratic form in our analysis because it produces a mean-square differentiable random field for all values of $\alpha_1 > 0$, $\alpha_2 > 0$.

As we will see, mean-squared differentiability of the underlying random field is a fundamental property in our modelling. Thus, we feel it is worthwhile to justify this assumption for real, physically occurring phenomena. For actual physical phenomenon, mean-squared differentiability is an extremely rare property. We may naturally ask: if most phenomena are not mean-square differentiable, then how can we justify this assumption in our model? In [25], it is shown that even a small amount of local averaging of the phenomenon in the parameter space will make the field mean-square differentiable. Local averaging refers to averaging out the signal over the local region in the parameter space. For processes which only have spatial variation, local averaging would average the signal over some spatial region around the point where each observation is taken. For processes which have only temporal variation, local averaging takes place over some averaging window centred at the time when the observation was recorded. We can even consider spatio-temporal phenomenon that vary simultaneously in space and time. Here the local averaging takes place over a cube in the three-dimensional space-time parameter set. We note that spatial local averaging is already "built-in" to many sensors. That is, the sensors report a value that represents a spatial aggregate of a small region centred about the sensor's location. This region is associated

with the so-called sensing radius of the nodes [1]. Thus, we feel that the sensor network data can be reasonably assumed to derive from an underlying mean-squared differentiable phenomenon. As well, any of the four spatial covariance functions mentioned earlier may be used in practice regardless of whether or not they create a mean-square differentiable field before local averaging takes place (see Appendix B for more details).

3.3.2 Temporal Variability

Recall that in the beginning of this chapter we specified that in our applications scenarios, the mean value of the field will only vary in time and be constant at all points in space for a fixed time. Also recall our previous assumption that the initial mean value of the field before the phenomenon strikes is zero without loss of generality. In our application scenarios, we envision that in the long run the phenomenon will decay to some steady state. To model this, we suggest considering candidate functions for the mean value that have some finite limiting value, say m_∞. We formalize our model for the mean value of the field:

$$m(\bar{s},t) = E[X(\bar{s},t)] = \begin{cases} 0 & \text{if } t = 0^-, \text{for } \forall \bar{s} \in S \\ m(t) & \text{if } t \geq 0, \text{for } \forall \bar{s} \in S \end{cases} \tag{3.10}$$

where $m(t)$ is such that $m(t) \to m_\infty$ as $t \to \infty$. The limiting value m_∞ can be chosen according to the application. For example, if we are dealing with a permanent presence of a toxic chemical in some region, then m_∞ may be some value greater than zero representing the long-term contamination level.

The last consideration in choosing a model function for the evolution of the mean value of the field in time is the rate of the mean function (in other words the derivative). Intuitively, this approximately corresponds to how fast the phenomenon decays. Again, this will also depend on what phenomena the network is observing. We now offer some fairly general examples that the form of the mean function, $m(t)$, could take.

- $m(t) = m_\infty \pm ce^{-kt}$ where $c > 0,\ k > 0$
- $m(t) = m_\infty \pm \dfrac{c}{k_1 t^{k_2}}$ where $c > 0,\ k_1 > 0,\ k_2 > 0$
- $m(t) = m_\infty + t^{k_1} ce^{-k_2 t}$ where $c > 0,\ k_1 > 0,\ k_2 > 0$

We comment that the first example function is common in modeling the decay of radioactive phenomena.

Note that the general form of the possibilities we listed for $m(t)$, are expressed in the form of the sum of the limiting value, m_∞, and a time-varying function, $g(t)$ (i.e. $m(t) = m_\infty + g(t)$). In order for $m(t) \to m_\infty$ as $t \to \infty$, all we require is that $g(t)$ has a limiting value of zero.

We now turn to the topic of temporal covariance between two samples from the same site $\bar{s} \in S$ at two different times. Recall that we represented this function earlier by $C_T(t_1, t_2)$ in our separable covariance model (3.5). In many applications, the temporal data at a site may exhibit the Markov property. That is, the future value of the process is

determined only by the present value and is independent of the past. The following positive-definite function captures this property [29].

$$C_T(t_1,t_2) = e^{-\alpha_3|t_2-t_1|} - e^{-\alpha_3(t_1+t_2)} \tag{3.11}$$

We compute the variance of the phenomenon at a site $\bar{s} \in S$ as follows.

$$\begin{aligned}\sigma^2(t) &= Var(X(\bar{s},t)) \\ &= C_S(\tau=0)\cdot C_T(t,t) \\ &= C_S(\tau=0)\cdot(1-e^{-2\alpha_3 t})\end{aligned}$$

We wee that the variance is a function of time for a fixed location.

In [17], the temporal covariance used is just of simple power-exponential form:

$$C_T(t_1,t_2) = e^{-\alpha_3|t_2-t_1|} \tag{3.12}$$

This temporal covariance function implies that the each point $\bar{s} \in S$ has an identical variance for all time:

$$\begin{aligned}\sigma^2(t) &= Var(X(\bar{s},t)) \\ &= C_S(\tau=0)\cdot C_T(t,t) \\ &= C_S(\tau=0)\cdot e^{-\alpha_3|t-t|} \\ &= C_S(\tau=0) \quad , \text{ for all } \bar{s} \in S\end{aligned}$$

In our model, we will consider the covariance model given by (3.12)

3.4 Modelling Data in a WSN

In this section, we provide the basic model for the node data in a wireless sensor network. Our data model consists of two stages. The first stage is a network deployment model which describes how the nodes are spatially distributed on the network deployment region S. The second stage uses the network deployment model and the random field model for the underlying phenomenon to create the network data model.

3.4.1 Network Deployment Model

We will assume that N sensor nodes are randomly deployed on a region $S \subset \Re^2$, according to the uniform distribution. We will call S the deployment area and represent its area by $A(S)$. We denote each node by a vector $\bar{s}_i$ in S corresponding to its physical location. The state of the network will be given by the vector, $\mathcal{N}=\{\bar{s}_1, \bar{s}_2, \ldots, \bar{s}_N\}$. The fact that the nodes are uniformly deployed means that the distribution of the number of nodes, K, in a given area, $A_0 \leq A(S)$, follows the spatial Poisson distribution. Thus, we have:

$$\Pr(K = k; A_0) = \frac{(\lambda A_0)^k e^{-\lambda A_0}}{k!}, k = 0,1,2,\ldots,N \qquad (3.13)$$

where $\lambda = \frac{N}{A(S)}$ is the fixed node density per unit area in the network.

For simplicity, it will be assumed that there are no dead or malfunctioning nodes in the network. That is, all nodes have sufficient battery power and can accurately sense,

process, and transmit data. It will also be assumed that the observation noise at each node is negligible and that each node observes the exact signal of the underlying phenomenon at that point in space.

3.4.2 Network Data Model

Recall that S is the network deployment area and that $T = [0, \infty)$ is the time interval after the phenomenon starts. We may expect that in many applications, the nodes take samples at discrete time intervals of $t_k = k\Delta t, k = 1, 2, \ldots$ where $t_k \subset T$. Let $\bar{T}_S$ denote the set of sampling times, $\bar{T}_S = \{t_1, t_2, \ldots, t_k, \ldots\}$. Thus, the network data occurs on the domain $S \times \bar{T}_S$ and corresponds to samples of the three-dimensional joint temporal-spatial Gaussian random field we described earlier.

Suppose we fix the time to $t \in T$. From the theory of Gaussian random fields, we know that the finite-dimensional joint distribution of the values of the field at any number of points on the parameter space $S \times T$ is jointly Gaussian. This assures us that the data at the network nodes is jointly Gaussian at each sampling time.

At a fixed sampling time $t \in \bar{T}_s$, we construct the vector of Gaussian random variables corresponding to the value of the random field at each node $i = 1, \ldots, N$: $\bar{X}(t) = [X_1(t), X_2(t), \ldots, X_N(t)]$. We know that any joint Gaussian distribution is specified completely by its mean and covariance. Recall from our random field model of

the observed phenomenon, that the data at all nodes at fixed time t have some identical mean value, $m(t) = m_\infty + g(t)$, where $g(t)$ can be any function with zero limit as $t \to \infty$. Further, we have assumed that the covariance function is separable (3.5) and composed of the product of an isotropic rational quadratic spatial covariance function, $C_S(\tau)$, (3.8) and an exponential temporal covariance, $C_T(t_1, t_2)$, (3.12). Thus, we have the following covariance between the data at two nodes i and j at fixed sampling time $t \in \bar{T}_s$:

$$\begin{aligned} C(X_i(t), X_j(t)) &= C_s(\tau_{i,j}) \cdot C_T(t,t) \\ &= C_s(\tau_{i,j}) \cdot \sigma^2(t) \\ &= \left(1 + \left(\frac{\tau_{i,j}}{\alpha_1}\right)^2\right)^{-\alpha_2} \end{aligned} \tag{3.14}$$

where $\tau_{i,j}$ denotes the distance between nodes i and j. We now construct the $N \times 1$ mean vector for the joint distribution as:

$$\bar{m}_{\bar{X}(t)} = \begin{bmatrix} m(t) \\ m(t) \\ \vdots \\ m(t) \end{bmatrix} \tag{3.15}$$

The $N \times N$ covariance matrix is given by

$$K_{\bar{X}(t)} = \sigma^2(t) \cdot \begin{bmatrix} C_S(0) & C_S(\tau_{1,2}) & \cdots & C_S(\tau_{1,N}) \\ C_S(\tau_{2,1}) & C_S(0) & \cdots & C_S(\tau_{2,N}) \\ \vdots & & \ddots & \vdots \\ C_S(\tau_{N,1}) & \cdots & & C_S(0) \end{bmatrix} \tag{3.16}$$

$K_{\bar{X}(t)}$ is both symmetric (since the distance operator is symmetric i.e. $\tau_{i,j} = \tau_{j,i}$) and positive-definite by virtue of the component entries being generated from the product of a

positive definite functions $C_S(\tau_{i,j})$ and a positive constant $\sigma^2(t)$. Thus, $K_{\vec{X}(t)}$ is an admissible covariance matrix for the joint Gaussian distribution of the data at the sensor nodes. We can then write the joint distribution of $\vec{X}(t)$ as:

$$f_{\vec{X}(t)}(x_1,\ldots,x_N) = \frac{1}{(2\pi)^{N/2}\left|K_{\vec{X}(t)}\right|^{1/2}} \exp\left\{-\frac{1}{2}(\vec{x}-\vec{m}_{\vec{X}(t)})^{Tr}(K_{\vec{X}(t)})^{-1}(\vec{x}-\vec{m}_{\vec{X}(t)})\right\} \quad (3.17)$$

This completes the representation of the data in a WSN at a fixed sampling time.

Chapter 4 Applications of Excursion Sets of Gaussian Random Fields to WSNs

In our application scenarios, we envision that the sink will only be interested in data from nodes that sense a sufficiently high ("extreme") value. The purpose of this chapter is to study the size and spatial distribution of the sets of nodes which observe extreme data. We will use techniques from excursion set theory of random fields to find the distribution of the number of nodes in a WSN that experience data values above levels b which can be considered extreme. From this we then derive the average number of nodes that experience extreme data in the network. We will also determine the average number of nodes that fall on contour lines of non-zero width that we construct using the underlying random field model of the phenomenon. Some background in random field theory, in particular the theory of high-level excursion sets is necessary to understand this chapter. Appendix A provides a quick overview of important concepts in random field theory. However for the benefit of the reader, we repeat relevant results contained in Appendix A when necessary so that the material in this chapter is fairly self-contained.

4.1 Normalization of Data

The analysis will assume that both the network data and the level b has been standardized to have zero mean and unit variance. There is no loss of generality in

performing the analysis on standardized data. The standardization of Gaussian data is achieved by the following linear transformation:

$$x_{std}(t) = \frac{x(t) - m(t)}{\sigma(t)} \tag{4.1}$$

Assume that we have some absolute level b^* and we want to examine the average number of nodes in the network report data above this level during the phenomenon. The analysis requires standardizing the threshold to:

$$b = \frac{b^* - m(t)}{\sigma(t)}$$

Clearly this basic, invertible transformation simply subtracts the mean value and divides by the standard deviation.

4.2 The Distribution of the Number of Excursion Nodes Above Different Thresholds

In this section, we will set down the basic probabilistic framework to describe the number of nodes that exceed some high level b at a given time. It will assumed that $m(t) = m$ and $\sigma^2(t) = \sigma^2$ at all locations in the network deployment area.

Recall that $S \subset \Re^2$ denotes the network deployment area. In this section, we will be looking at properties of the set of nodes on S that are above the level b. We call a node that has value greater than b an *excursion node*. There is no exact criterion for how high the level b should be in order to be considered extreme, but often $b \geq 1$ is taken as a starting point. Often, we will not be too strict in our choice of b because we will be

interested in seeing if our derived results for the average number of nodes above b will hold even for moderate values (say $0 \le (b-m)/\sigma \le 1$). Assume that the data at all nodes has been standardized so that $m = 0$ and $\sigma = 1$.

In random field theory, an *excursion region* is defined as an isolated subset of S where the value of the random field $X(\bar{s})$ exceeds the level b. For a level b, we denote the $i'th$ excursion region in S for this level by $A_{i,b}$. We will occasionally abuse notation slightly by using $A_{i,b}$ to also denote the area of the $i'th$ excursion region. Since the underlying field is random, the area of each $A_{i,b}$ is also a random quantity.

In Appendix A8, an expression is given for mean-square differentiable Gaussian random fields that characterizes the average area of an isolated excursion area, denoted by $E[A_{i,b}]$, for $b \to \infty$. This expression is:

$$E[A_{i,b}] = \left(\frac{1-\Phi(b)}{\phi(b)}\right)^2 \frac{2\pi}{\sigma_{\dot{X}_1}\sigma_{\dot{X}_2}} \frac{1}{|\mathbf{V}|^{1/2}} \tag{4.2}$$

The parameters $\sigma_{\dot{X}_1}$, $\sigma_{\dot{X}_2}$ correspond to the mean-square value of the directional derivatives of the field in the direction of each axes in 2-dimensional space. The matrix $\mathbf{V}$ contains information regarding the correlation between the gradient fields in the direction of each axis in 2-dimensional space. Anyhow, we will not be too concerned with the particulars of the values in this matrix because our earlier assumption on the isotropy of the underlying random field means $\mathbf{V}$ is given by the identity matrix. Thus,

its determinant is $|\mathbf{V}| = 1$, [25]. Isotropy of the underlying random field also guarantees that the values of $\sigma_{\dot{X}_1}$, $\sigma_{\dot{X}_2}$ are equal to one another. These values can be computed as:

$$\sigma_{\dot{X}_i}^{\ 2} = -\left.\frac{d^2 C_S(\tau)}{d\tau^2}\right|_{\tau=0}, \quad i = 1, 2$$

When the random field has a rational quadratic correlation function as we assumed earlier for our data model, we have:

$$\sigma_{\dot{X}_i} = \sqrt{\frac{2\alpha_2}{\alpha_1^{\ 2}}}, \quad i = 1, 2$$

Thus, the average area of a single isolated excursion region for a high level b when the random field is described by our model from Chapter 3 can be computed as:

$$\begin{aligned} E[A_{i,b}] &= \left(\frac{1-\Phi(b)}{\phi(b)}\right)^2 \frac{2\pi}{(2\alpha_2/\alpha_1^{\ 2})} \\ &= \left(\frac{1-\Phi(b)}{\phi(b)}\right)^2 \frac{\pi \cdot \alpha_1^{\ 2}}{\alpha_2} \end{aligned} \tag{4.3}$$

where $b \to \infty$, $\Phi(b)$ and $\phi(b)$ are the CDF and PDF respectively for a standard normal random variable, and α_1, α_2 are the spatial correlation parameters introduced in Chapter 3. We note that the general distribution of the area of each $A_{i,b}$ is unknown in current literature. Nonetheless, we can still approximate the average area of each $A_{i,b}$ for high b using the asymptotic formula shown above [25]. In Figure 4.1, we plot the average area of an isolated excursion area as a function of the correlation range parameter α_1 for different levels $b = 2, 2.5, 3$, and a fixed correlation roughness parameter of $\alpha_2 = 1$. As we would expect, the average size of isolated excursions regions is an increases with the

correlation range parameter α_1. As well, we see that the isolated excursion regions become smaller on average as b is increased. In Figure 4.2 we reproduce a diagram from [35] showing the excursion regions of a realization of a random field for the levels $b = \{0.8, 1.2, 1.6\}$. This figure clearly demonstrates that the average size of an excursion area becomes smaller as the level b is increased. Also observe that the actual shapes of these excursion regions become more "regular" with increasing b.

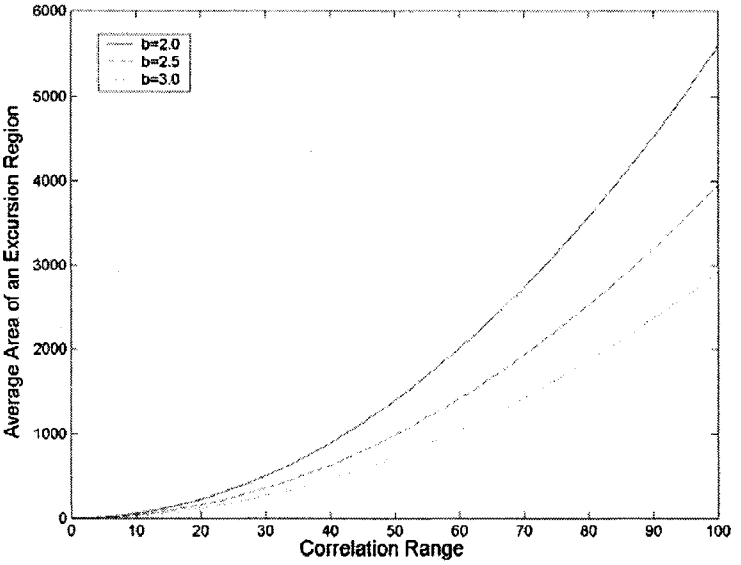

Figure 4.1: The average area of an isolated excursion region as a function of the correlation range parameter, α_1, for levels $b = \{2, 2.5, 3\}$, and fixed roughness parameter $\alpha_2 = 1$

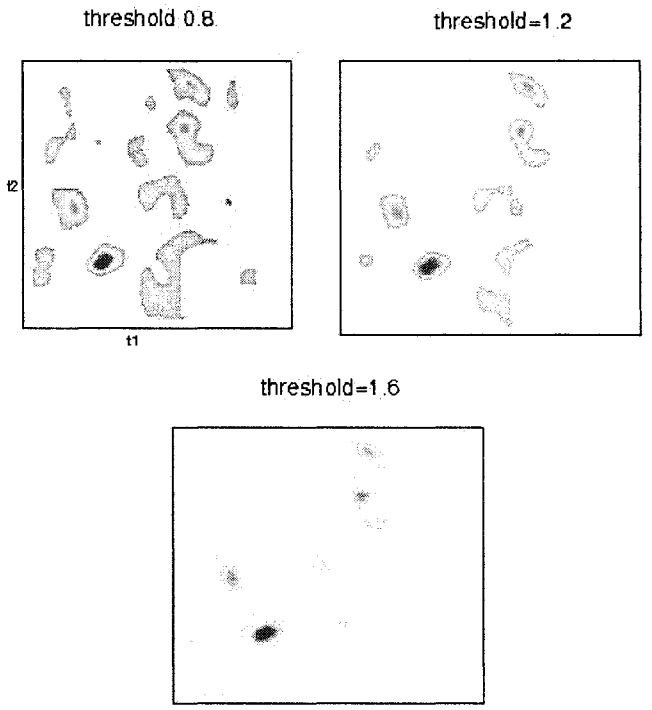

Figure 4.2 : Excursion regions of a realization of a random field for different levels $b = \{0.8, 1.2, 1.6\}$ (reproduced from [35]).

We are now prepared to analyze the number of nodes in the network that observe a value greater than b at a given time. Recall that an excursion node is a node in the network that observes a value greater than the level b. Because the nodes sample the underlying random field, excursion nodes are characterized by the property that their locations in S are necessarily within an isolated excursion region, $A_{i,b}$. Equivalently, any node that falls in a region, $A_{i,b}$, is necessarily an excursion node.

The first quantity we will be interested in is the number of nodes that fall on each region, $A_{i,b}$. Because, the nodes have been deployed uniformly on the network

deployment area with density λ, the number of nodes on each $A_{i,b}$ is governed by the spatial Poisson distribution with parameter λ. Denote by $K_{i,b}$, the number of excursion nodes that fall on the $i'th$ excursion region, $A_{i,b}$. We note that the general stochastic distribution of the size of $A_{i,b}$ is unknown and the best we can do is approximate it with its area $E[A_{i,b}]$. We have the following approximation for the distribution of $K_{i,b}$:

$$\Pr(K_{i,b} = k) \approx \frac{(\lambda E[A_{i,b}])^k \cdot e^{-\lambda E[A_{i,b}]}}{k!}, \quad k = 0, 1, 2, \ldots \tag{4.4}$$

where $b \to \infty$, and $E[A_{i,b}]$ is given by (4.3). The average number of excursion nodes for each region $A_{i,b}$ can be computed as:

$$\begin{aligned} \overline{K}_{i,b} &= E[\lambda \cdot A_{i,b}] \\ &= \lambda \cdot E[A_{i,b}] \end{aligned} \tag{4.5}$$

The second quantity we will be interested in is the total number of excursion nodes in the entire network. In [25] it is shown that as the level b is increased towards the asymptote, the number of local maxima of the random field approaches a spatial Poisson process in the plane (see Appendix A8). The parameter of this Poisson process is given by:

$$\mu_b = \frac{1 - \Phi(b)}{E[A_{i,b}]} \tag{4.6}$$

where $b \to \infty$. It is then shown that as b increases towards the asymptote, each isolated excursion area contains exactly one local maximum. Thus, for high b there is a one-to-one correspondence between the number of local maxima and the number of isolated

excursion regions on the network deployment area. We then conclude that for high b, the number of isolated excursion regions on the network deployment area approaches the spatial Poisson process with parameter μ_b given in (4.6).

Let β_b represent the total global area of the random field on S that has data above the level b. β_b corresponds to the sum of the areas of each individual excursion region $A_{i,b}$. We will refer to β_b as the *global excursion area.* Let $\overline{\beta}_b$ be the associated average global excursion area. Using (4.2) and (4.6), we compute $\overline{\beta}_b$ as:

$$\begin{aligned}\overline{\beta}_b &= (A(S)\cdot\mu_b)(E[A_{i,b}]) \\ &= \left(A(S)\cdot\frac{(1-\Phi(b))}{E[A_{i,b}]}\right)(E[A_{i,b}]) \\ &= A(S)\cdot(1-\Phi(b))\end{aligned} \tag{4.7}$$

where $b\to\infty$ and $\Phi(b)$ is the CDF of a standard normal random variable. Looking at (4.7) we see that $\overline{\beta}_b$ is only a function of the level b and is independent of the degree of spatial correlation of the random field. The spatial correlation only affects the spatial distribution of the isolated excursion regions. Note that by expression (4.6), the rate of isolated excursion regions in the plane, μ_b, is inversely proportional to the average area of an isolated excursion region. Physically, as the correlation level is increased two things happen in the random field. Firstly, the average area of each isolated excursion region increases. Simultaneously however, the frequency of these regions in the plane is decreased in equal proportion to the increase in the average area of isolated excursion regions. This creates a sort of "global balancing" effect on the quantity $\overline{\beta}_b$.

Let K_b be the total number of excursion nodes in the entire network for a level b. Using (4.7), the mean value of K_b is given by:

$$\begin{aligned}\overline{K}_b &= \lambda \cdot \overline{\beta}_b \\ &= \left(\frac{N}{A(S)}\right) \cdot [A(S) \cdot (1 - \Phi(b))] \\ &= N \cdot (1 - \Phi(b))\end{aligned} \tag{4.8}$$

where $b \to \infty$. Looking at (4.8), we observe that for asymptotically high b, this expression is identical to the probability that N independent nodes sense data above the level b. What this shows is that the expected number of excursion nodes on the entire network area is asymptotically independent of the correlation level. This means that at least as far as determining the number of excursion nodes in a WSN, we can treat the data as if it was independent. The invariance of $\overline{K}_b$ to the correlation level follows simply from the invariance of $\overline{\beta}_b$ to the correlation level and the fact that $\overline{K}_b$ is determined by multiplying $\overline{\beta}_b$ by the node density, λ .

Physically, what these previous comments on the effect of correlation mean is that if the underlying phenomenon exhibits strong spatial correlation, then the excursion nodes will be occur a small number of relatively large isolated excursion areas. On the other hand, if the underlying spatial correlation of the field is low, then the excursion nodes will occur on a large number of relatively small isolated excursion areas.

Recall that in our application scenarios, we envision that the sink will only be interested in data that is extreme. Thus, if the phenomenon is considered extreme when it exceeds some high level b, then expression (4.8) is a way of understanding the total amount of data packets that the network will have to handle after a sampling instant occurs.

We comment here that for high values of b, the factor $1-\Phi(b)$ seen in expressions (4.7) and (4.8) can be thought of as representing the proportion of the network that has extreme data. In (4.7), this factor represents the proportion of the total *area* of the network deployment that experiences extreme data while in (4.8), this factor represents the proportion of the total *number of nodes* that observe extreme data.

4.2.1 Numerical Results

In this section we compare the asymptotic expression (4.8) we derived for the average number of excursion nodes for a level b with results from simulations of standardized correlated Gaussian data at N points in space. The purpose here is to study how well our expression (4.8) predicts simulation values for different levels of b. The rational quadratic covariance function given by (3.7) was used in these simulations. In all simulations, the number of nodes was $N = 500$ and the nodes were deployed on a circle of radius $R_{Net} = 50$ metres. The results for the mean number of excursion nodes in the network were obtained by averaging over 1000 simulation runs. This experiment was repeated for three different correlation levels (low, medium, and high) created by varying

the scale parameter α_1 and keeping the roughness parameter α_2 fixed ($\alpha_2 = 1$) in the covariance function. A useful metric for understanding the amount of data that has been suppressed in the network is the percentage of the average number of excursion nodes from simulation to the total number of nodes deployed. We call this simulation metric the *percentage excursion nodes*. Tables 4-1, 4-2, and 4-3 summarize our simulation results.

Table 4-1: Theoretical and Simulated Average Number of Excursion Nodes (Low Correlation)

1. N=500, Rnet=50, alpha1=1, alpha2=1 (low correlation) Average Number of Excursion Nodes				
Level (b)	**Theory**	**Simulation**	**% Difference**	**% Excursion Nodes**
0.00	250.0000	249.5810	0.1676	49.9162
0.20	210.3701	209.3950	0.4635	41.8790
0.30	191.0443	190.8440	0.1048	38.1688
0.50	154.2688	155.0650	0.5161	31.0130
0.70	120.9818	120.3900	0.4892	24.0780
0.90	92.0301	91.9030	0.1381	18.3806
1.00	79.3276	78.6490	0.8554	15.7298
1.20	57.5348	57.8970	0.6295	11.5794
1.50	33.4036	33.3680	0.1066	6.6736
1.70	22.2827	22.1940	0.3981	4.4388
1.80	17.9652	18.1210	0.8672	3.6242
1.90	14.3583	14.4650	0.7431	2.8930
2.00	11.3751	11.2320	1.2580	2.2464
2.10	8.9322	9.0910	1.7778	1.8182
2.20	6.9517	6.9500	0.0245	1.3900
2.30	5.3621	5.4850	2.2920	1.0970
2.40	4.0988	4.0790	0.4831	0.8158
2.50	3.1048	3.1750	2.2610	0.6350
3.00	0.6749	0.6540	3.0968	0.1308

Table 4-2: Theoretical and Simulated Average Number of Excursion Nodes for (Medium Correlation)

2. N=500, Rnet=50, alpha1=5, alpha2=1 (medium correlation) Average Number of Excursion Nodes				
Level (b)	**Theory**	**Simulation**	**% Difference**	**% Excursion Nodes**
0.00	250.0000	251.9180	0.7672	50.3836
0.20	210.3701	209.0010	0.6508	41.8002
0.30	191.0443	191.8300	0.4113	38.3660
0.50	154.2688	154.3790	0.0714	30.8758
0.70	120.9818	120.8060	0.1453	24.1612
0.90	92.0301	92.6090	0.6290	18.5218
1.00	79.3276	80.0590	0.9220	16.0118
1.20	57.5348	57.1820	0.6132	11.4364
1.50	32.7830	33.4036	1.8931	6.6807
1.70	22.1410	22.2827	0.6400	4.4565
1.80	17.6100	17.9652	2.0170	3.5930
1.90	13.7600	14.3583	4.3481	2.8717
2.00	11.3420	11.3751	0.2918	2.2750
2.10	9.1440	8.9322	2.3163	1.7864
2.20	7.0220	6.9517	1.0011	1.3903
2.30	5.2860	5.3261	0.7586	1.0652
2.40	4.1610	4.0988	1.4948	0.8198
2.50	3.0950	3.1048	0.3166	0.6210
3.00	0.6940	0.6749	2.7522	0.1350

Table 4-3: Theoretical and Simulated Average Number of Excursion Nodes for (High Correlation)

3. N=500, Rnet=50, alpha1=25, alpha2=1 (high correlation) Average Number of Excursion Nodes				
Level (b)	**Theory**	**Simulation**	**% Difference**	**% Excursion Nodes**
0.00	250.0000	235.6180	5.7528	47.1236
0.20	210.3701	213.1390	1.3162	42.6278
0.30	191.0443	189.1380	0.9978	37.8276
0.50	154.2688	151.6270	1.7125	30.3254
0.70	120.9818	117.9940	2.4696	23.5988
0.90	92.0301	90.5360	1.6235	18.1072
1.00	79.3276	80.4900	1.4653	16.0980
1.20	57.5348	53.6530	6.7469	10.7306
1.50	33.4036	36.0010	7.7758	7.2002
1.70	22.2827	21.2870	4.4685	4.2574
1.80	17.9652	16.6550	7.2930	3.3310
1.90	14.3583	14.4650	0.7431	2.8930
2.00	11.3751	10.0140	11.9656	2.0028
2.10	8.9322	8.9560	0.2665	1.7912
2.20	6.9517	6.6550	4.2680	1.3310
2.30	5.3621	5.0470	5.8764	1.0094
2.40	4.0988	4.6070	12.3988	0.9214
2.50	3.1048	3.9040	25.7408	0.7808
3.00	0.6749	0.6350	5.9120	0.1270

Some comments on the results from these tables are now in order. Firstly, for all three levels of correlation the asymptotic formula for $\overline{K}_b$ predicts the number of network nodes above level b from simulation results very well for a broad range of levels b. As is seen, for low and medium levels of correlation the percentage difference between the

theoretical and simulated values are most of the time less than 1 percent. On the other hand, for a high level of correlation the percentage difference is somewhat larger. We also note that for very high levels of b that the percentage difference is usually quite large compared to this same difference at lower levels of b. We attribute this to the fact that as the level b is increased, the isolated excursion regions for b are so small that no nodes fall in these regions when the network is deployed. Thus, there is no data collected from these regions and this explains the higher percentage difference often seen at higher levels b. In applications with hundreds or even thousands of nodes deployed, we suggest that using (4.8) to estimate the average number of excursion nodes will likely be "good enough". Finally, the percentage of excursion nodes in the simulation shows how we can drastically reduce the set of excursion nodes by increasing the level b.

We comment that these simulations also show the invariance of the mean number of excursion nodes in the network to the correlation level as previously described. If one compares the three Tables 4-1, 4-2, and 4-3 for a fixed value of the threshold b, then it is evident that the mean number of excursion nodes is essentially the same in all three simulations.

4.3 High Level Contour Lines

Here we consider an application of excursion theory for random fields that is essentially an extension of the preceding section concerning the number of excursion nodes in the network. In many scientific monitoring applications, the sink may only be

interested in the properties of the contour lines of a phenomenon being observed by the WSN. A contour line of a random field is simply all points in the plane that have the same value. We will show in this section that we can accurately determine the number of nodes belonging to a set of contour lines in the plane for a high level b and that this is an effective means of further reducing the set of nodes that attempt to transmit their data to the sink.

We note that that for a random field on $S \subset \Re^2$, the boundaries of excursion sets for a given level b are equivalent to a set of contour lines for the level b that occur on S. Suppose then that the sink is only interested in nodes that belong to some contour line for a predefined level b. The question we will seek to address is how many nodes fall on this contour line. Since the nodes have been deployed uniformly on the network deployment area, the probability that they fall exactly on a contour line defined by the underlying random field is obviously zero (since a line has zero area). So it is clear that we will have to redefine what it means for a node to "belong" to a particular contour line for a given value b.

Definition: *We will say that a node at location $\bar{s} \subset S$ belongs to a contour line c_b associated with the level b, if its value $X(\bar{s})$ satisfies:*

$$b - \delta \leq X(\bar{s}) \leq b + \delta$$

for some $0 < \delta << b$.

In effect, we have defined a "quantization band" of half-width δ, centred at the value b. Although this definition may seem ad hoc and opportunistic it is not without some practical justification. Recall that we are sampling a continuous random field at discrete points in space. The value realised by the random field at a given point in space is some number from the real line. In general, the sensing devices will not be able represent the exact value of the random field due to physical memory constraints of the node and limitations of the sensing device itself. Instead, the value of the random field at the node can only be stored with limited precision and therefore there is a built-in rounding error. This "rounding" performed on the true data of the underlying phenomenon essentially represents a quantization of continuous values on the real line to a set of discrete values that can be stored in the nodes memory. This quantization may also be necessary from the perspective that a sensor network node transmits small information packets to conserve energy. In this sense, the energy/precision trade-off may be desirable. Formally, our quantization rule operates as follows:

Quantization Rule:

Suppose that $0 < \delta << b$. *Any node that measures a value* $y \in [b-\delta, b+\delta)$ *is quantized to the value* $\hat{y} = b$

We assume that there is no further measurement error (say from observation noise). Then the problem of determining how many nodes in the network belong to a given contour line then reduces to finding the set of nodes who observe values that are quantized to some level b.

At this point, we make some observations concerning the underlying "smoothness" of the random field. Here we define "smoothness" of the random field to mean that the random field has continuous sample paths. Recall from the conjecture in Appendix A7 that we can essentially be certain that a stationary Gaussian random field has continuous sample paths. Intuitively, this means that the underlying field is "nice" in the sense that it does not have any unexpected jumps or points where the spatial rate of change of the phenomenon is infinite.

Suppose we consider the $i'th$ excursion regions $A_{i,b-\delta}$ and $A_{i,b+\delta}$ associated the levels $b-\delta$ and $b+\delta$ respectively. For small δ, the continuity of the sample paths assures us that in general, each isolated excursion region for high level $b+\delta$ is contained within a larger isolated excursion region for level $b-\delta$. All nodes that are located between the boundaries of these two regions of the random field measure values $y \in [b-\delta, b+\delta)$ and assuming our quantization rule applies, they store the data value b in memory. So from the sinks perspective, the number of nodes "on" the contour line corresponds to the number of nodes between these two boundaries of the random field. Denote by $\overline{N}_{c_b,i}$, the average number of nodes that form the $i'th$ contour line for level b. Since the nodes are deployed uniformly on the deployment area with intensity $\lambda = N / A(S)$, we have that:

$$\overline{N}_{c_b,i} = \lambda \cdot \left(E[A_{i,b-\delta}] - E[A_{i,b-\delta}]\right) \tag{4.9}$$

The term in parenthesis is just the average area enclosed between the two boundaries of the contour line. It should be noted that this last result is hard to verify in practice because unless we know the value of the random field at every point in the deployment

area (thus, requiring an infinite number of nodes to gather samples), there is always some guesswork involved in forming the contour lines. [4] shows examples where very different pictures of the contour lines are constructed by the same data set. In this sense, constructing the contour lines from discretely spaced data locations is a subjective science.

Now we analyze the number of contour nodes for a level b over the entire network, which we denote by N_{c_b}. Based on our previous comments, we can express the average number of contour nodes on the network as the average number of nodes that contain data between the levels $b-\delta$ and $b+\delta$ on the entire network. Using (4.8) we get that:

$$\begin{aligned}\overline{N}_{c_b} &= N \cdot \left(\overline{K}_{b-\delta} - \overline{K}_{b+\delta}\right) \\ &= N \cdot \left[1 - \Phi(b-\delta)\right] - N \cdot \left[1 - \Phi(b+\delta)\right] \\ &= N \cdot \left[\Phi(b+\delta) - \Phi(b-\delta)\right]\end{aligned} \tag{4.10}$$

Although the previous derivation used an intuitive argument based on the apparent smoothness of the random field, simulation will later show that (4.10) is indeed an accurate expression of the average number of nodes in the network that belong to a set of contour lines for level b.

Since the contour lines for a level b are formed by the excursion regions for the levels $b-\delta$ and $b+\delta$, it follows that the distribution of the isolated contour lines in space follows the Poisson distribution. Using our definition for a contour line, the

occurrence of an isolated contour line in the plane is brought about by any occurrence of an isolated excursion region for the level $b-\delta$. Thus, the parameter in the Poisson distribution for the number of isolated contours for the level b is simply $\mu_{b-\delta}$. This implies that on average, there are $\mu_{b-\delta} \cdot A(S)$ isolated contour lines in the plane for the level b.

4.3.2 Numerical Results

In this section, we will show that (4.10) can be accurately used to predict the average number of nodes that will be quantized to the contour line of level b over the entire network. We perform simulation with levels from the range $b \geq 0$. We perform 1000 simulations for each contour level b and use $N = 500$ nodes deployed on a circular area of radius $R_{Net} = 50$ metres. Again, the simulations are repeated for three different levels of spatial correlation (low, medium, and high) by varying the scale parameter α_1 and keeping the roughness parameter α_2 fixed ($\alpha_2 = 1$). For all simulations, the quantization half-interval was chosen as $\delta = 0.05$. So for example, any data measurement $y \in [1.75, 1.85)$ gets "conceptually" rounded to $\hat{y} = 1.8$ and therefore belongs to the contour line for $\hat{y} = 1.8$. To understand the amount of data suppression achievable by only letting contour nodes transmit their data, we define the simulated *contour node percentage*. This is simply the ratio of the number of contour nodes in simulation to the total number of nodes deployed. We show the results of our simulation in the Tables 4-4, 4-5, and 4-6:

Table 4-4: Theoretical and Simulated Average Number of Contour Nodes (Low Correlation)

1. N=500, Rnet=50, alpha1=1, alpha2=1 (low correlation),delta=0.05 Average Number of Contour Nodes				
Level (b)	**Theory**	**Simulation**	**% Difference**	**% Contour Nodes**
0.00	19.9388	20.0640	0.6240	3.875
0.20	19.5443	19.3750	0.8738	3.8402
0.30	19.0622	19.2010	0.7229	3.485
0.50	17.5978	17.4250	0.9917	3.1056
0.70	15.6094	15.5280	0.5242	2.6684
0.90	13.3032	13.3420	0.2908	2.429
1.00	12.0985	12.1450	0.3829	1.9578
1.20	9.7111	9.7890	0.7958	1.2938
1.50	6.4793	6.4690	0.1592	0.946
1.70	4.7062	4.7300	0.5032	0.786
1.80	3.9512	3.9300	0.5394	0.6448
1.90	3.2844	3.2240	1.8734	0.5264
2.00	2.7029	2.6320	2.6938	0

Table 4-5: Theoretical and Simulated Average Number of Contour Nodes (Medium Correlation)

2. N=500, Rnet=50, alpha1=5, alpha2=1 (medium correlation), delta=0.05 Average Number of Contour Nodes				
Level (b)	**Theory**	**Simulation**	**% Difference**	**% Contour Nodes**
0.00	19.9388	19.7035	1.1942	3.9407
0.20	19.5443	19.7745	1.1641	3.9549
0.30	19.0622	19.3140	1.3037	3.8628
0.50	17.5978	17.8570	1.4515	3.5714
0.70	15.6094	15.7490	0.8864	3.1498
0.90	13.3032	13.3040	0.0060	2.6608
1.00	12.0985	12.0590	0.3276	2.4118
1.20	9.7111	9.7030	0.0835	1.9406
1.50	6.4793	6.2970	2.8950	1.2594
1.70	4.7062	4.7410	0.7340	0.9482
1.80	3.9512	3.7830	4.4462	0.7566
1.90	3.2844	3.2990	0.4426	0.6598
2.00	2.6760	2.7029	0.9952	0.54058

Table 4-6: Theoretical and Simulated Average Number of Contour Nodes (High Correlation)

3. N=500, Rnet=50, alpha1=25, alpha2=1 (high correlation), delta=0.05 Average Number of Contour Nodes				
Level (b)	**Theory**	**Simulation**	**% Difference**	**% Contour Nodes**
0.00	19.9388	19.5900	1.7805	3.9180
0.20	19.5443	19.6820	0.6996	3.9364
0.30	19.0622	18.4620	3.2510	3.6924
0.50	17.5978	17.9620	2.0276	3.5924
0.70	15.6094	15.6110	0.0102	3.1222
0.90	13.3032	13.1300	1.3191	2.6260
1.00	12.0985	11.8040	2.4949	2.3608
1.20	9.7111	9.4390	2.8827	1.8878
1.50	6.4793	6.6460	2.5083	1.3292
1.70	4.7062	5.0770	7.3035	1.0154
1.80	3.9512	4.0080	1.4172	0.8016
1.90	3.2844	2.9760	10.3629	0.5952
2.00	2.6760	2.8290	5.4083	0.5658

As may be seen, the theoretical values for the average number of contour nodes in the network obtained using (4.10) closely match the simulation results. In fact, even for moderate contour level $0 \le b \le 1$, the results more or less coincide with the theory and the percentage difference is generally very small. The accuracy of the results for the average number of contour nodes using (4.10) is not surprising since we already saw that the theoretical prediction of the average number of excursion nodes is quite good.

Simple comparison of the tables for the average number of nodes on a contour line for level b with the tables for the average number of excursion nodes above b show that if only the contour nodes transmit, then the set of nodes which will transmit their data to the sink can be drastically reduced. For instance, under medium correlation, the percentage of excursion nodes in the network for level $b = 1.5$ is 6.6 % of the total node population. Meanwhile, if we limit transmissions in the network to nodes which fall on a

contour lines for $b = 1.5$ using the same degree of correlation, then only about 1% of the total node population will transmit to the sink.

4.4 The Average Number of Excursion Nodes During a Time-Varying Phenomenon

In this section, we let time vary in order to study how the average number of excursion nodes in the network changes in response to the underlying phenomenon. Here, we assume that there is a single application-specific threshold b used to define the excursion regions. This value of b can be interpreted as some threshold such that when a group of nodes report a value above b, this represents the phenomenon in some critical or dangerous state. No new theory is required in order to generalize the average number of excursion nodes when the phenomenon varies in time. Thus we will proceed directly with simulation results. We will use the notation $E[K_b(t)] = \overline{K}_b(t)$ to denote the mean number of excursion nodes in the network at time t.

We now produce simulation data showing how the average number of nodes above some fixed threshold b varies in time. The simulation parameters here are essentially the same as our previous simulations. The spatial covariance function used was of rational quadratic form and the temporal covariance was the exponential type according to our network data model. In all simulations, $N = 500$ nodes were deployed over a radius of $R_{Net} = 50$ metres. 1000 simulation runs were performed in determining the average number of nodes above the fixed threshold level of $b = 5$. This experiment was only performed for a single correlation level (high) created by setting $\alpha_1 = 25$,

$\alpha_2 = 1$ as before. We assumed that the mean value of the field determined by $m(t) = 15 \cdot e^{-0.05t}$ (see Figure 4.3). This can correspond to the mean decay of some pollutant or chemical in a medium whose initial mean is $c = 15$ and has a rate of decay $k = 0.05$. We simulated the network data at $\Delta t = 5$ second intervals. Since we are mainly interested in the implications of spatial correlation in terms of local suppression of data, we here assumed that the value of the temporal correlation parameter is $\alpha_3 = 1$. This choice makes the samples of a single node taken at $\Delta t = 5$ seconds intervals essentially independent of each other. This allows us to generate the network data at each sampling instant independently of the data from previous sampling instants. With $\alpha_3 = 1$, the covariance between the node data at a sampling time t was calculated using $C(X_i(t), X_j(t)) = \left(1 + (\tau_{i,j} / 5)^2\right) \cdot \left(1 - e^{-2t}\right)$. In Table 4-7, we show the numerical results of our simulation.

Table 4-7: Simulated and theoretical average number of excursion nodes during a "fast" phenomenon. Nodes have high spatial correlation $\alpha_1 = 25$.

N=500, Rnet=50, alpha1=25, alpha2=1 (high correlation) alpha3=1, b=5, m(t)=15exp(-0.05*t)				
	Average Number of Excursion Nodes			
Time	Theory	Simulation	% Difference	% Excursion Nodes
5	500.0000	500.0000	0.0000	100.0000
10	499.9896	499.9650	0.0049	99.9930
15	490.7440	490.4150	0.0671	98.0830
20	348.9189	348.6820	0.0679	69.7364
25	120.6031	124.4020	3.0537	24.8804
30	24.5803	21.8100	12.7020	4.3620
35	4.1734	3.8020	9.7685	0.7604
40	0.7446	0.7370	1.0312	0.1474

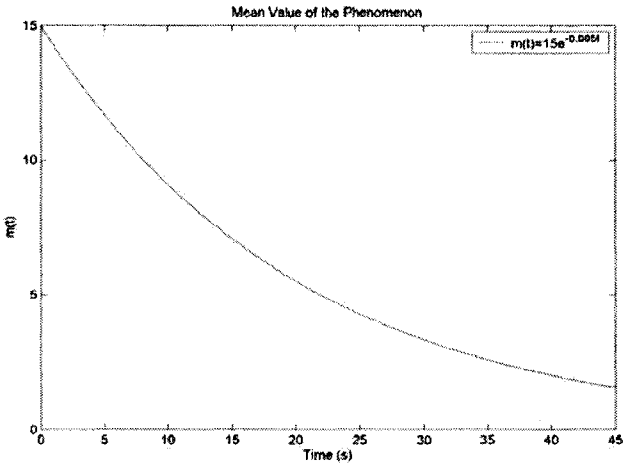

Figure 4.3: The mean value of the simulated "fast" phenomenon, $m(t) = 15e^{-0.05t}$

We show the temporal evolution graphically, with the simulated and theoretical average number of nodes who sense a data value greater than b plotted against time in Figure 4.4.

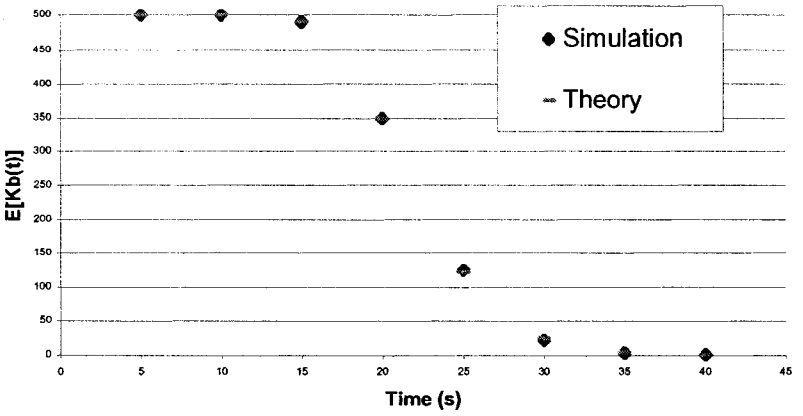

Figure 4.4: The simulated and theoretical values for the average number of excursion nodes for critical threshold value of b=5 against time for a "fast" phenomenon. The mean value of the phenomenon is $m(t) = 15e^{-0.05t}$.

From our simulation results in Figure 4.3, we see a drastic transition at between 20 seconds and 25 seconds. There are an average of 350 excursion nodes at $t = 20$ seconds while only 119 excursion nodes at $t = 25$ seconds. This is clearly a sharp drop and corresponds to a 66% decrease in the average number of critical excursion nodes. To put this in terms of a network application, if all nodes in the network with data above b were to transmit their values at each sampling time, then only 23% of all nodes will transmit at 25 seconds while 70% of all nodes transmit in the previous sampling time at 20 seconds. Clearly, the nominal load on the network has changed considerably between these two sampling times.

The above scenario could correspond to a phenomenon that evolves very rapidly and reaches its steady-state value very fast. Examples of these phenomena could include earthquakes, explosions, bridge collapses, etc. However, equally valid are scenarios where the phenomenon evolves over much greater time scales. We now repeat the previous simulation, but change the mean function to $m(t) = te^{-0.06t}$ and the units of time to minutes rather than seconds (see Figure 4.5). In this case, the phenomenon does not begin with a large impulse to the environment but instead gradually climbs out of the pre-phenomenon state, reaches its maximum, then gradually decays again to the steady-state. Phenomena characterized by this type of evolution could include rain storms, forest fires, volcanic eruptions, contamination of a body of water, etc. In our simulations, we will sample the network every 5 minutes. We show our results in Table 4-8:

Table 4-8: Simulation and theoretical values for the average number of excursion nodes during a "slow" phenomenon. The phenomenon has high spatial correlation

N=500, Rnet=50, alpha1=25, alpha2=1 (high correlation) alpha3=0.05, b=5, m(t)=t*exp(-0.06t)				
Average Number of Excursion Nodes				
Time (minutes)	**Theory**	**Simulation**	**% Difference**	**% Excursion Nodes**
5	48.7492	47.0540	3.6027	9.4108
10	343.6331	345.2050	0.4554	69.0410
15	432.0083	431.7390	0.0624	86.3478
20	423.5275	423.9630	0.1027	84.7926
25	359.2268	358.1700	0.2951	71.6340
30	241.8173	245.4720	1.4888	49.0944
35	118.8029	122.2450	2.8157	24.4490
40	42.5718	41.4100	2.8056	8.2820
45	12.0457	11.9700	0.6324	2.3940
50	3.0128	2.6060	15.6101	0.5212

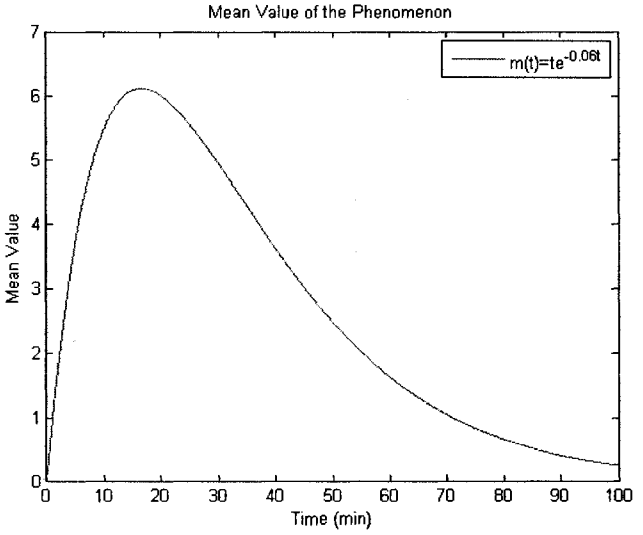

Figure 4.5: The mean value of the simulated "slow" phenomenon, $m(t) = t \cdot e^{-0.06t}$.

The corresponding plot of our theoretical and simulation results from Table 4-8 is shown in Figure 4.6:

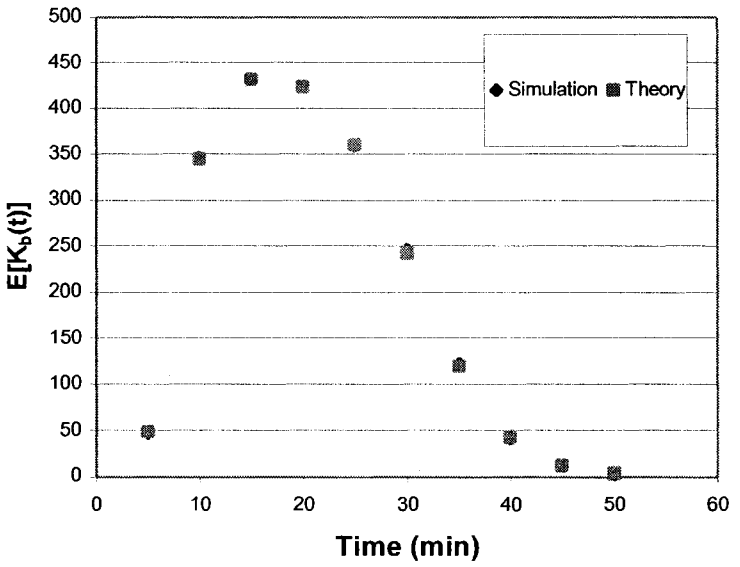

Figure 4.6: The simulated and theoretical values for the average number of excursion nodes for critical threshold b=5 during a "slow" phenomenon. The mean value of the phenomenon is $m(t) = te^{-0.006t}$.

For the two preceding simulations scenarios we comment as usual that the analytical formulas predict the average number of excursion nodes accurately.

Assume as usual that the transmission criterion for a node to attempt to transmit its sensor reading is that it records a value greater than the critical threshold *b*. Both of these simulations demonstrate that in WSNs deployed to monitor when a phenomenon exceeds the threshold *b*, there may exist sampling instants where a huge percentage of the network nodes will want to transmit their sensor readings. Despite the fact that we would like to transmit all data that is above the level *b*, the question to be addressed is whether the network can handle this nominal load. If the network cannot handle these huge amounts of data, then we offer two possible solutions. In the first solution, we could

define a series of n different contour levels $b_k = b + k \cdot \Delta$ where $k = 0, 1, 2, \ldots, n$ and Δ is some application-defined threshold difference between the contour lines. Then we could define the transmission criterion in the network such that only nodes that belong to one of these contour lines transmit their data. A second possible solution is to adaptively define the threshold b as time changes. That is, we can use a time-dependent threshold $b(t)$ at the nodes. As the nodes detect the phenomenon increasing in between sampling instants, the value of the threshold can be raised. Similarly, as they detect the phenomenon decreasing the value can be lowered. If $b(t)$ is chosen appropriately, then we can ensure that only the nodes with the "statistically" high sensor readings will attempt to transmit their data to the sink during the phenomenon thereby reducing the data load on the network to a level where it can be handled.

Chapter 5 Extension to Non-Stationary Phenomena

Many physically occurring phenomenon exhibit spatial non-stationarity. In this chapter, we illustrate how the analysis from the previous chapter may be extended to WSNs in this type of environment.

5.1 Introduction to Spatially Non-Stationary Random Fields

In the context of random fields, spatial non-stationarity can arise if at least one of the following is true:

- The mean of the field is not constant at all points across the parameter space
- The covariance is not a function of only the separation vector between two points in the parameter space.

Many phenomena can be characterized as occurring at an "epicentre" that affect the surrounding points in the environment in proportion to the point's distance to the epicentre. Examples of these phenomena could include volcanic eruptions, geysers, forest fires, earthquakes, among others. We will say that these sorts of phenomena are of the "point-source" type. In these cases, the point-source phenomenon will induce a random field in its local environment. Using random field theory, we can model these sorts of phenomenon by introducing a dependency between the mean value of the random field at some point and its physical distance to the epicentre. In this chapter, we will be

interested in the sorts of scenarios where the field exhibits spatial non-stationarity *only* with respect to the mean. Thus, we will assume that the covariance is still only a function of the physical distance between two points (in other words, the field is isotropic).

The question arises, can we somehow extend our existing results concerning the average number of excursion nodes in the network to scenarios where the underlying phenomenon is of the point-source type and therefore does not exhibit a constant mean value across the network deployment area for a fixed time.

For random fields where the mean is constant across all points in space at fixed times, recall our critical observation regarding (4.8); that for an asymptotically high level b, the average number of nodes that experience data above the b in a sensor network observing correlated data is the same as if the network were observing independent and identically distributed data. Now our simulations showed us that at least in the context of sensor networks with hundreds of nodes, this simple formula works well over a range of levels b and not just for asymptotic levels. So at least from the perspective of the average number of excursion nodes, this quantity is well-approximated using results from an analogous independent and identically distributed system.

From the forgoing comments, we conjecture that if the mean value is not identical at all points in S for a fixed time, then the average total number of excursion nodes for a spatially correlated, mean-square differentiable Gaussian random field can be computed

as if the data were independent (i.e. uncorrelated) but *not* identically distributed. Let X_i denote the data value of the $i'th$ node at a fixed time. Then we expect:

$$\begin{aligned}\overline{K}_b &= \sum_{i=1}^{N} \Pr(X_i > b) \\ &= \Pr(X_1 > b) + \Pr(X_2 > b) + \ldots + \Pr(X_N > b)\end{aligned} \tag{5.1}$$

In order to compute (5.1) in practice, all node data must be standardized by subtracting the mean value at each node and dividing by the standard deviation. Note that since we assumed the covariance structure is unchanged from our earlier modelling, we have that the variance of all nodes is the same for any fixed time. For a fixed time *t*, suppose that the $i'th$ node has mean $m_i(t) = m_i$ and variance $\sigma^2{}_i(t) = \sigma^2$ and that the data between nodes is *correlated* as before. Then we can compute the average number of excursion nodes at a fixed time when the data is correlated and has a location-dependent mean as follows:

$$\begin{aligned}\overline{K}_b &= \sum_{i=1}^{N} \Pr(X_i > b) \\ &= \Pr(X_1 > b) + \Pr(X_2 > b) + \ldots + \Pr(X_N > b) \\ &= \Pr\left(Z_1 > \frac{b - m_1}{\sigma}\right) + \Pr\left(Z_2 > \frac{b - m_2}{\sigma}\right) + \ldots + \Pr\left(Z_N > \frac{b - m_N}{\sigma}\right)\end{aligned} \tag{5.2}$$

where Z_i, $i = 1, \ldots, N$, are *independent* standardized normal random variables. As usual, this is an asymptotic expression. In this case, we formally require each $(b - m_i)/\sigma \to \infty$, $i = 1, \ldots, N$.

We comment here that although we can still compute the average number of excursion nodes in the network when the phenomenon has a location-dependent mean value, our result concerning the Poisson distribution for the number of isolated excursion regions for high b will not apply to the non-stationary field. This is because the Poisson result implies that the excursion regions occur randomly on S. However, a location-dependent mean in the field means that nodes can belong to a high level excursion region simply by virtue of their physical locations (for example, by being close to the epicentre of the phenomenon) and thus the locations of these regions is non-random.

5.2 Numerical Results (Average Number of Excursion Nodes)

We now produce some simulation results which show that (5.2) does indeed accurately approximate the average number of excursion nodes when the data is correlated and has a spatially varying mean. In our simulations, we suppose that as usual we have deployed $N = 500$ nodes on a circular area of radius $R_{Net} = 50$ metres. For the sake of simulation, we then create a phenomenon epicentre at a location chosen randomly on a circle of radius $R^* = 2 \cdot R_{Net}$ that encompasses the network deployment area. Further, we suppose that this epicentre is a "virtual" node that also produces Gaussian data. We call the original network of nodes augmented with this virtual epicentre node the extended network, $\mathcal{N}^+ = \mathcal{N} \cup N_E$ where N_E represents the epicentre node and $\mathcal{N}$ represents the original network of sensor nodes. We assume that the data in $\mathcal{N}^+$ is correlated according to the isotropic rational quadratic covariance function with correlation parameter $\alpha_1 = 50$ (strong correlation). For simulation purposes, we assume

that at any time t the mean value of the epicentre is given by $m_E(t) = 15e^{-0.05t}$ and that the corresponding mean value at each node is a distance-scaled version of the mean at the epicentre and is given by $m_i(t) = m_E(t)/\sqrt{\tau_{i,E}}$, $i = 1,\ldots,N$, where $\tau_{i,E}$ is the distance between the $i'th$ node and the epicentre of the phenomenon. This choice of mean value is purely illustrative. We show the average number of excursion nodes as time evolves under these conditions for the fixed level $b = 2$. The results of our simulation are shown in Table 5-1.

Table 5-1: Comparison of simulated and theoretical average number of excursion nodes in WSN when the data exhibits a spatially varying mean

N=500, Rnet=50, alpha1=50, alpha2=1 (medium correlation) alpha3=1, b=2, m(t)=15exp(-0.05*t) at epicentre					
	Average Number of Excursion Nodes				
Time (minutes)	Theory	Simulation	% Difference	Min(Node Mean)	Max(Node Mean)
5	236.0378	234.3130	0.7307	1.2200	10.7188
10	148.7199	149.5840	0.5810	0.9421	5.7853
15	52.0771	54.2810	4.2320	0.5219	1.5053
20	36.9228	33.2570	9.9283	0.3992	1.0511
25	29.3223	28.1380	4.0389	0.3122	0.8296
30	24.2640	23.7700	2.0359	0.2443	0.6202

Table 5-1 shows that the values produced by our analytical expression (5.2) closely match our simulation results. Suppose that at a fixed sampling time, we look at the mean value of each node in the network, $m_1(t), m_2(t),\ldots, m_N(t)$. In Table 5-1, we have also included the maximum and minimum of these mean values at each sampling time, that is $\text{Max}_{\forall i}\,[m_i(t)]$ and $\text{Min}_{\forall i}\,[m_i(t)]$ where $i = 1,\ldots,N$. Comparing the difference between the maximum and minimum mean node values over the set of all nodes in the network is a way of understanding the "degree" of spatial non-stationarity in the phenomenon at each sampling time. Effectively, there is greater spatial non-

stationarity if this difference is large. As is seen in the results, the phenomenon exhibits stronger spatial non-stationarity at the earlier sampling times. As time evolves and the phenomenon at the epicentre dies, the degree of spatial non-stationarity of the random field also dies. Despite the stronger non-stationarity at early sampling times, the analytical results still closely match the simulation results. This confirms that having a location dependent mean at each node really does not affect the ability to determine the average number of excursion nodes in the network.

We comment here that the above analytical and numerical results assume that the location of the epicentre of a point-source phenomenon is known so that we can compute the mean value of each node at any time instant. This may be a reasonable assumption in a wide variety of applications where the WSN is monitoring a static point-source. For example, a volcano vent or a fault line in the earth can both be regarded as static point sources. It is reasonable that this knowledge will be available to sensor network designers. On the other hand, some point-source phenomena may occur randomly in space. Examples of these may be the location of an explosion in a terrain or the storm centres found in hurricanes and tornados. If a stochastic description of the location of the epicentre is known for the phenomenon, then designers can uncondition on the location of the phenomenon in order to predict the number of excursion nodes in the sensor network. However, we will not pursue these scenarios in our work.

Chapter 6 Applications of Excursion Sets to WSN Traffic

The purpose of this chapter is to show how WSN designers might use the results of our modelling in order to understand the traffic properties in WSNs that only transmit statistically extreme data. Further, we show how this understanding can guide the WSN design process. The principle concept we will exploit from our modelling is that we can accurately predict the size and spatial distribution of the regions of the sensor deployment area which experience statistically extreme values when the network observes a *stationary* Gaussian random field. We will consider two types of MAC protocols, schedule-based and contention-based. Our analysis will lend itself naturally to deriving various performance measures. When a schedule-based MAC protocol is used, we will determine the packet loss probability distribution. When a contention-based MAC protocol is used we will determine the mean delay of a packet at each hop and the mean delay of a packet from source to sink.

6.1 Traffic Characteristics in Schedule-Based MAC

We assume that every Δt seconds, all network nodes sample their data. As usual, suppose that the criterion for a node to transmit its data after being sampled is that it has a value greater than some statistically high threshold b. In our scenario, the set of regions in the plane containing nodes which transmit their data at some fixed sample time corresponds to the set of isolated excursion regions, $A_{i,b}$, on the deployment area S. Recall the fundamental property that for high b, the frequency of excursion regions in the

plane approaches a Poisson distribution with parameter μ_b. Thus, we can view the production of data traffic in our sensor network as originating at a series of traffic-generating regions corresponding to the isolated excursion regions. The Poisson result for the number of isolated traffic-producing regions implies that their spatial distribution is uniform on S. In this sense, these traffic regions are well separated in space and do not experience clustering. We assume that excursion nodes access the channel in scheduled slots in a TDMA frame of duration Δt seconds. Next, we will elaborate on the packet arrival and service processes in the network.

6.1.1 Packet Loss Model for WSN Traffic using TDMA

We now propose a simple queuing model that describes excursion node traffic in a WSN when nodes access the channel through a TDMA frame. This model will allow us to study the number of packets that are lost in each frame.

i) Arrival Process

Our basic idea is to treat each isolated excursion region like a "virtual" queue. We define a customer as a data packet from an excursion node. This data packet contains the sensor reading of the node. We assume that all data packets produced by excursion nodes are of fixed length. Thus, at each sample time $j\Delta t$, each isolated excursion region queue experiences the simultaneous batch arrival of $K_{i,b}(j\Delta t)$ customers. Since we assumed earlier that the nodes are uniformly deployed on the deployment region, the

number of customers that constitute this batch arrival is distributed according to the Poisson distribution with parameter $\lambda = N / A(S)$ on each isolated excursion region $A_{i,b}$. The average the number of customers in each excursion region is therefore:

$$\overline{K}_{i,b}(j\Delta t) = \lambda E[A_{i,b}]$$

To make notation more compact we will drop the dependence on time and just write $K_{i,b}(j\Delta t) = K_{i,b}$ for the number of nodes on each excursion region from now on.

Our previous comments tell us that the arrivals to each queue occur as a batch at deterministic time intervals every Δt seconds (note however, that the size of the batch arrivals to each queue are random according to the Poisson distribution as mentioned above). Thus, the arrival rate to the $i'th$ excursion region queue is $K_{i,b} / \Delta t$.

ii) Service

We will assume some sort of TDMA frame is used to service data packets from the nodes in each isolated excursion region. We take advantage of the fact that the location of each excursion region queue is uniformly distributed on the deployment area. So in this sense, the queues are well spread out and do not cluster close to one another. This fact allows us to posit that we can expect spatial reuse of the channel will be possible so that users in one excursion region can use the same time slots as the users in another excursion region [30]. Thus, we will assume that the transmissions that occur in isolated excursion regions do not interfere with one another. As a first-cut assumption, we will also assume that transmissions from each queue to the sink occur in a single hop.

This could happen through the use of an aggregator node such as a higher-powered device (see 802.15.4 full-function devices [1]) or a regular node that acts as a gateway to the sink and has its neighbouring excursion nodes transmit their data to it before it routes the aggregated packet to the sink (this is in the spirit of the LEACH protocol [30]). Thus, we model each queue as having a single server. Assume that the TDMA frame divides the Δt seconds between sample epochs into slots of duration δ such that a single data packet from each excursion node can fit into a single slot. This slot duration corresponds to the service of a single customer in our queuing model and is determined by the propagation delay of a single packet transmission. Thus, the service rate is deterministically $1/\delta$.

Next we summarize the server, arrival process, and service processes of our model :

- Each isolated excursion region, $A_{i,b}$, has a virtual TDMA server queue.
- The arrivals to each queue occur as a batch of random size at deterministic times (given by the sampling epochs).
- Service is deterministic and happens at a rate of one packet every δ seconds (the slot time). Transmissions to the sink occur in a single hop or through a gateway node in a single hop.

The following sections show some aspects of the performance of a WSN using our model.

6.1.2 Packet Loss at Each Queue

We can view our current model for each virtual TDMA server as a D/D/1 queue. At a fixed sample time, each queue can only support a limited number of customers given by the number of slots in the frame. We denote the number of slots in each time frame as $s = \left\lfloor \frac{\Delta t}{\delta} \right\rfloor$ (we use the $\lfloor \cdot \rfloor$ operator to denote the value of the argument rounded down to the nearest integer). Since there are only s slots in the frame before the next sampling instant, we can only support s customers at each queue. We ask what happens if the number of excursion node data packets at a D/D/1 queue exceeds s? Since by the next sample instant Δt seconds later, any leftover packets will be "out-of-date", we will assume that these packets are dropped.

If it happens that a queue drops packets, we will be interested in knowing the distribution of the number of packets lost at each queue and the average number dropped in the whole network. Let $L_{i,b}$ correspond to the number of packets that the $i'th$ D/D/1 queue drops. $K_{i,b}$ represents the number of packets that need service at the $i'th$ D/D/1 queue during each frame. We can express the probability that j packets are dropped at this queue as:

$$\Pr(L_{i,b} = j) = \begin{cases} \Pr(K_{i,b} \le s) \text{ if } j = 0 \\ \Pr(K_{i,b} = s + j) \text{ if } j > 0 \end{cases} \tag{6.1}$$

The expected number of dropped packets is derived as follows:

$$\begin{aligned}\bar{L}_{i,b} &= E[L_{i,b}] \\ &= \sum_{j=1}^{\infty} j \cdot \Pr(L_{i,b} = j) \\ &= \sum_{j=1}^{\infty} j \cdot \Pr(K_{i,b} = s + j) \\ &= \sum_{j=1}^{\infty} j \cdot \frac{(\lambda E[A_{i,b}])^{s+j} e^{-(\lambda E[A_{i,b}])}}{(s+j)!}\end{aligned}$$

We perform a change of variables and let $n = s + j$. We then get

$$= \sum_{n=s+1}^{\infty} (n-s) \cdot \frac{(\lambda E[A_{i,b}])^{n} e^{-(\lambda E[A_{i,b}])}}{n!}$$

$$= \sum_{n=s+1}^{\infty} n \cdot \frac{(\lambda E[A_{i,b}])^{n} e^{-(\lambda E[A_{i,b}])}}{n!} - s\sum_{n=s+1}^{\infty} \frac{(\lambda E[A_{i,b}])^{n} e^{-(\lambda E[A_{i,b}])}}{n!}$$

$$= \left[\sum_{n=0}^{\infty} n \cdot \frac{(\lambda E[A_{i,b}])^{n} e^{-(\lambda E[A_{i,b}])}}{n!} - \sum_{n=0}^{s} n \cdot \frac{(\lambda E[A_{i,b}])^{n} e^{-(\lambda E[A_{i,b}])}}{n!}\right] - s\sum_{n=s+1}^{\infty} \frac{(\lambda E[A_{i,b}])^{n} e^{-(\lambda E[A_{i,b}])}}{n!}$$

$$= \lambda E[A_{i,b}] - \sum_{n=0}^{s} n \cdot \frac{(\lambda E[A_{i,b}])^{n} e^{-(\lambda E[A_{i,b}])}}{n!} - s\Pr(K_{i,b} > s)$$

Therefore, our final expression for the mean packet drop is:

$$\bar{L}_{i,b} = \lambda E[A_{i,b}] - \sum_{n=0}^{s} n \cdot \frac{(\lambda E[A_{i,b}])^{n} e^{-(\lambda E[A_{i,b}])}}{n!} - s\Pr(K_{i,b} > s) \qquad (6.2)$$

This shows how the number of nodes dropped depends on the density of the network, the size of each excursion region (a random quantity), and the number of slots, s, in the time frame.

We also define P_{loss} as the probability that a data packet from an excursion node is dropped from the TDMA frame. We have:

$$P_{loss} = \frac{\overline{L}_{i,b}}{\overline{K}_{i,b}} \tag{6.3}$$

Since the number of D/D/1 queues on the network deployment area is Poisson with intensity μ_b, the average number of packets that are dropped in the entire network from all D/D/1 queues at each sampling time is then given by:

$$\overline{L}_b = (\mu_b A(S)) \cdot \overline{L}_{i,b} \tag{6.4}$$

where $A(S)$ is the network deployment area.

6.1.3 Numerical Results for TDMA MAC

In this section, we wish to provide designers with some guidelines for designing WSNs to monitor spatially varying phenomenon when they use a TDMA-style MAC protocol like the one described by our D/D/1 queuing model. The fundamental design parameter is the sampling interval, Δt.

We want to study the effect of the sampling interval Δt on the packet loss at each virtual queue as described in the last section. Thus, we will assume that the node density is constant for now. Assume that the data rate of the channel is 30 kbps, and data packets have a fixed length of 10 bytes (80 bits). These are all fairly typical values in expected WSNs [1], [18]. Using these values for packet length and data rate, we will assume that the time is slotted into slots of duration 80 bits/30 kbps = 2.67 msec so that a single data

packet from an excursion node at a D/D/1 queue can be transmitted in each time slot. From the preceding section, we know that we experience packet loss at a queue as soon as the number of customers at the queue exceeds the number of slots. It makes sense to choose the number of slots as some value greater than the average number of nodes on each isolated excursion region. However, as a worst-case analysis assume that we first choose the number of slots as equal to the average number of excursion nodes per excursion region. That is, assume that $s = \lfloor \lambda E[A_{i,b}] \rfloor$. Suppose that we have deployed $N = 5000$ nodes on a semi-circular deployment area of radius $R_{Net} = 300$ m with a transmission threshold of $b = 1$. Further, assume a correlated standardized Gaussian underlying phenomenon with $\alpha_1 = 15$ and $\alpha_2 = 1$. These assumptions produce the following values for the average isolated excursion regions, and the approximate average number of nodes on each of these areas (see Table 6-1):

Table 6-1: Excursion Region Properties for D/D/1 Model

N=5000, R_{Net}=300 m, b=1 $\alpha_1 = 15$, $\alpha_2 = 1$	
$E[A_{i,b}]$ (m^2)	303.8895
λ	0.0354
$\overline{K}_{i,b}$ (nodes)	10.7577

Using the values from Table 6-1, we will first assume that there are then $s = 10$ slots in each TDMA frame which gives the duration of a single frame as $\Delta t = 10 \times 2.67$ msec = 26.7 msec. In Figure 6.1, we analytically produce equation (6.1), the probability mass function of the number of packet losses in a TDMA frame assuming $s = 10$ slots.

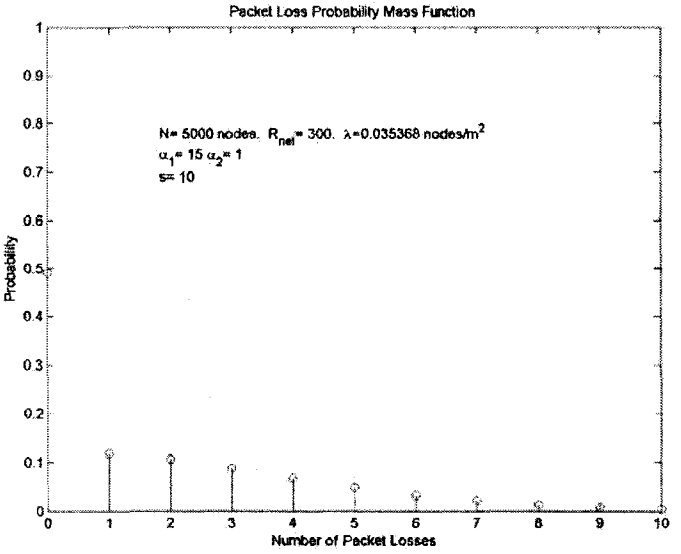

Figure 6.1: Probability Mass Function of the Number of Packets Lost in a TDMA frame when s= $\overline{K}_{i,b} \approx 10$ slots.

This plot shows that roughly half of the time when sampled, the D/D/1 queues do not experience dropped packets ($\Pr(L_{i,b} = 0) = 0.4903$). This corresponds to our intuition that approximately half of the time when sampling, the number of excursion nodes will be greater than its average quantity. The probability that there are any packets dropped in a TDMA frame is $\Pr(L_{i,b} \geq 1) = 1 - 0.4903 = 0.5097$. This is quite a high probability of at least a single packet being dropped and may be unacceptable to designers of WSNs in many applications.

Analytical computations using (6.2) show that the average number of packets dropped from a single D/D/1 queue during a sampling interval of $\Delta t = s \cdot \delta = 26.7$ msec in the current scenario is $\overline{L}_{i,b} = 1.6905$ packets. We can then compute the loss probability for a packet using (6.3) as:

$$P_{loss} = \frac{\overline{L}_{i,b}}{\overline{K}_{i,b}}$$
$$= \frac{1.6905}{10.7577}$$
$$= 15.71\%$$

In order to ameliorate this high loss probability of a data packet, we suggest amending our worst-case scenario assumption that set the number of slots equal $\lfloor \overline{K}_{i,b} \rfloor$. Let us then examine the effect of increasing the sampling interval so that the number of slots is chosen as $s = \lfloor \overline{K}_{i,b} + 2 \cdot \overline{L}_{i,b}(s = 10) \rfloor = 14$ slots, where $\overline{L}_{i,b}(s = 10)$ is the packet loss probability when the number of slots in a TDMA frame was $s = 10$ (our former worst-case scenario). When we use $s = 14$ slots, the duration of a single TDMA frame is now $\Delta t = 14 \times 2.67$ msec = 37.38 msec. We show the probability mass function for packet drop when this is the case in Figure 6.2.

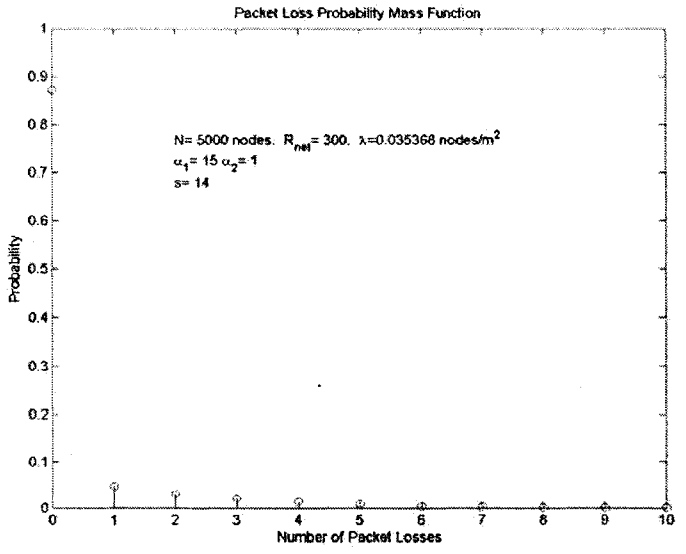

Figure 6.2: Probability Mass Function of the Number of Packets Lost in a TDMA frame when $s = 14$ slots.

In Figure 6.2 we see a drastic performance improvement in our TDMA system. Here the probability that no packets are lost in a TDMA frame is now about 87%, which is significantly better than our previous scenario. As well the corresponding probabilities for $L_{i,b} = 1, 2, 3, \ldots$ packet losses are significantly lower than their values in the former scenario when s was chosen as 10 slots.

The average number of packets dropped in a TDMA frame when $s = 14$ slots is analytically computed using (6.2) as $\overline{L}_{i,b} = 0.3102$ packets. The packet loss probability computed using (6.3) is then:

$$\begin{aligned} P_{loss} &= \frac{\overline{L}_{i,b}}{\overline{K}_{i,b}} \\ &= \frac{0.3102}{10.76} \\ &= 2.88\% \end{aligned}$$

Clearly by increasing the number of slots in a TDMA frame to $s = 14$, we realize a significantly better packet loss probability. In Table 6-2, we present a performance comparison between using $s = 10$ and $s = 14$ slots in a TDMA frame.

Table 6-2 : Performance comparison of TDMA system using two different frame lengths, Δt .

N=5000, R_{Net}=300 m, b=1 $\alpha_1 = 15$, $\alpha_2 = 1$	
Δt	$P_{loss} = \overline{L}_{i,b} / \overline{K}_{i,b}$
26.70 msec	15.71%
37.38 msec	2.88%

The question to be addressed is whether the increased sampling interval meets the needs of the application. We suggest that in many applications this trade-off will be acceptable and network designers should choose the sampling frequency Δt such that the number of slots in this interval is at least $\lfloor \overline{K}_{i,b} + 2 \cdot \overline{L}_{i,b} \rfloor$. Of course, the needs of the application will determine how much greater than $\lfloor \overline{K}_{i,b} + 2 \cdot \overline{L}_{i,b} \rfloor$ the number of slots in a TDMA frame must be.

6.2 *Traffic Analysis in Contention-Based MAC*

In this section, we show how we can apply our theory of excursion sets of stationary Gaussian random fields in modelling the performance of a WSN using multihop communications with a contention-based MAC protocol. In what follows, we are simply looking for a rough first-cut analysis that can give WSN designers some guidelines when designing a network. The usefulness of our analysis will lie in its simplicity and lack of dependency on the exact specifics of the routing and MAC protocols. We state the basic assumptions for our scenario as follows:

- The nodes are sampled at discrete instants in time every Δt seconds. Time is slotted into slots of duration δ seconds, determined by the propagation delay of a transmission. After each sampling time, only excursion nodes attempt to transmit their data to the sink.

- The rate of temporal change of the underlying phenomenon is relatively slow compared to the sampling interval Δt. Thus, we assume that once a node becomes an excursion node, it tends to stays an excursion node for a long period of time relative to Δt.

- Each excursion node will encapsulate its sensor reading in a single packet of fixed length. Non-excursion nodes only participate in relaying the data packets from the excursion nodes to the sink. Downstream excursion nodes may also assist in routing packets from upstream excursion nodes.

- At each hop in the routing path, a data packet moves towards to the sink by a physical distance of r_{hop}.

- Sensor nodes access the channel through a CSMA-type MAC protocol

As part of our first cut model, we need a way to roughly describe how data flows in the network. Let us assume without loss of generality that the network is densely deployed on a semi-circle of radius, R_{Net}, with the sink located at the origin as in Figure 6.3. Suppose that we conceptually divide up the network deployment region into J rings of constant thickness corresponding to the physical 1-hop distance of a data packet (earlier assumed to be constant, r_{hop}). These rings will correspond to a series of regions in the plane between concentric circles of radii $r_1, r_2, \ldots, r_J$. Denote ring i by S_i and its area by $A(S_i)$, $i = 1 \ldots J$. We wish to emphasize that these rings are purely conceptual. Thus, we can choose them in whatever manner convenient to our analysis. For us, that

convenient choice is $r_i - r_{i-1} = r_{hop}$. Suppose that R_{net} is divisible by r_{hop}; we have $r_1 = r_{hop}$, $r_2 = 2 \cdot r_{hop}$, ... , $r_J = R_{Net}$. Using this construction, we assume that when a data packet is routed from source to sink, that it passes through a single ring in each hop. We then suppose that all data travels roughly in a straight line from its source to the sink at the origin creating the data implosion at rings closer to the sink as described earlier.

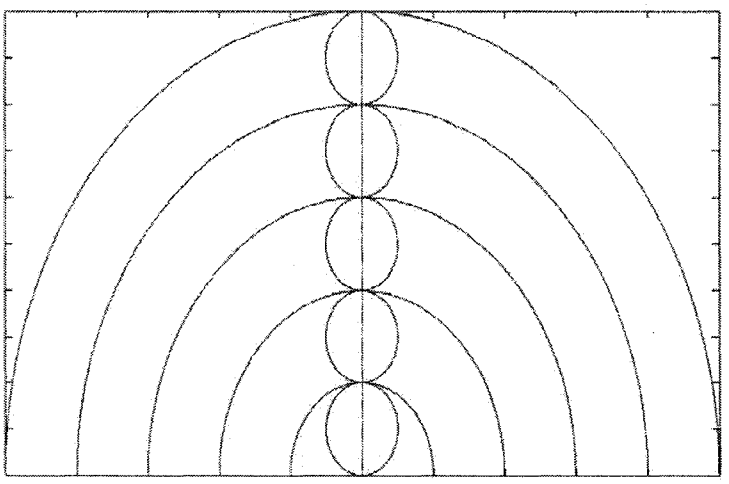

Figure 6.3: Illustration of conceptual ring construction. The thickness of each ring is the physical 1-hop distance and the smaller circles correspond to hop-neighbourhoods in a typical routing path.

6.2.1 Contention Model

Here we describe our basic modelling of how nodes in a dense sensor field contend for the channel. Recall that we assume that nodes in any hop neighbourhood access the channel using a CSMA-type protocol where time is slotted into slots of duration δ seconds. Suppose that the probability that any node (relay or excursion node) with a packet to transmit tries to capture the channel in a slot is p. We will call this the "attempt" probability. Note that because there is the possibility that multiple nodes will pick the same time slot, a node must wait some interval of time before it successfully

captures the channel. We express the probability that a single node successfully captures the channel in a time slot as P_s. The number of slots until a node successfully captures the channel in a slot can be modelled with the geometric distribution with mean value $1/P_s$ slots. The attempt probability p can be chosen to maximize the success probability, P_s. In [34], it is shown that P_s is maximized for $p = 1/n$, where n is the number of nodes contending for the channel in the hop neighbourhood. If p is chosen as this maximizing value, then as the number of nodes trying to access the channel increases, then the probability of successful transmission quickly stabilizes to a constant value of $1/e$ (stabilization is reached after about $n = 4$ contending nodes). We will exploit this fact later.

6.2.2 Traffic Modelling

External Arrivals

Since only excursion nodes will attempt to transmit their data to the sink, we will treat each isolated excursion region, $A_{i,b}$, in the plane as a source of new arrivals into the network. Recall from Appendix A.8, that the local maxima of a stationary Gaussian random field occur in the plane according to the spatial Poisson process with intensity μ_b and there is a one-to-one correspondence between isolated excursion regions and local maxima for high b. For the sake of simplicity then, we will assume that packets in an isolated excursion region are routed from the location of the excursion region's local

maximum. This assumption allows us to be able to think of the $A_{i,b}$ regions as occurring "in" the rings $i = 1,2,\ldots,J$ according to which ring its local maxima lies in.

We comment that in many applications these excursion nodes may actually choose to suppress their data transmission. The reason for this is that even within excursion regions where the most "critical" data is, there is spatial redundancy among excursion nodes [1], [9], [19]. In order to achieve this further data suppression, each excursion node must be able to listen to the transmitted values from other excursion nodes in its local area. If it overhears a packet transmission from an excursion node in its local neighbourhood then we will assume that it will suppress its transmission during the current sampling interval. To model this behaviour, we will assume that within each Δt second sampling interval, an excursion node will suppress the transmission of its packet from the last sampling instant with some probability, q. Define M as the number of Δt second sampling intervals that an excursion node suppresses its packet transmissions before finally attempting to transmit a data packet. M is geometrically distributed with parameter $1-q$. Thus, the mean number of sample intervals before an excursion node attempts to transmit a data packet is given by $\overline{M} = 1/(1-q)$. We will say that an arrival has occurred at an excursion node at the instant when the excursion node finally attempts to transmit a data packet. Denote by u, the mean time interval between successive arrivals at an excursion node. We have

$$u = \frac{\Delta t}{1-q} \tag{6.5}$$

We will approximate the time interval by the exponential distribution with mean value $u = \Delta t/(1-q)$. That is:

$$f(t) = (1/u)\cdot e^{-(1/u)\cdot t}$$

where u is given in (6.5). Now assume that there are $K_{i,b} = k$ nodes on an isolated excursion region $A_{i,b}$ and that all of these excursion nodes can interfere with one another. We wish to treat each region $A_{i,b}$ like a "virtual" queue. Thus, we are interested in the inter-arrival time between instants where any of the k excursion nodes on $A_{i,b}$ attempt to transmit a data packet. Define y as this inter-arrival time. We have the following conditional cumulative probability distribution function:

$$\Pr(y > t \mid K_{i,b} = k) = e^{-k\cdot(t/u)} \tag{6.6}$$

In order to determine the cumulative probability distribution function of y, we will uncondition on $K_{i,b} = k$ nodes. Note that the probability that $K_{i,b} = k$ is approximately given by the Poisson distribution:

$$\Pr(K_{i,b} = k) \approx \frac{e^{-\lambda E[A_{i,b}]}\cdot(\lambda\cdot E[A_{i,b}])^k}{k!}, k = 0,1,\ldots \tag{6.7}$$

The corresponding probability generating function for $K_{i,b}$ is given by:

$$\begin{aligned} G_{K_{i,b}}(z) &\approx \sum_{k=0}^{\infty} z^k \Pr(K_{i,b} = k) \\ &= e^{-\lambda E[A_{i,b}](1-z)} \end{aligned} \tag{6.8}$$

Using (6.6), we may express the cumulative probability density function of y as:

$$
\begin{aligned}
F_y(t) &= \Pr(y > t) \\
&= \sum_{k=0}^{\infty} \Pr(y > t \mid K_{i,b} = k) \cdot \Pr(K_{i,b} = k) \\
&= \sum_{k=0}^{\infty} e^{-k \cdot (t/u)} \cdot \Pr(K_{i,b} = k)
\end{aligned}
$$

Substituting $z = e^{-t/u}$, we get:

$$
\begin{aligned}
F_y(t) &= \sum_{k=0}^{\infty} z^k \cdot \Pr(K_{i,b} = k) \\
&\approx e^{-\lambda E[A_{i,b}](1-z)} \Big|_{z=e^{-t/u}} \\
&= e^{-\lambda E[A_{i,b}](1-\exp(-t/u))} \\
&\approx e^{-\lambda E[A_{i,b}](1-(1-t/u))} \\
&= e^{-\lambda E[A_{i,b}](t/u)}
\end{aligned}
\tag{6.9}
$$

The first line above is just the probability generating function of $K_{i,b}$, given in (6.8). The fourth line above uses a Taylor expansion to approximate $e^{-t/u} \approx 1 - t/u$. Thus, we have the following approximate expression for the cumulative distribution function of y:

$$
F_y(t) \approx 1 - e^{-\lambda E[A_{i,b}](t/u)} \tag{6.10}
$$

where u is given by (6.5). (6.10) shows that the inter-arrival time between attempts to transmit excursion node packets on an isolated excursion region $A_{i,b}$ is approximately exponentially distributed with parameter

$$
\begin{aligned}
\nu &= \lambda E[A_{i,b}]/u \\
&= \overline{K}_{i,b}/u
\end{aligned}
\tag{6.11}
$$

This allows us to state that the arrivals to each virtual excursion region queue are approximately Poisson with rate $\nu = \lambda E[A_{i,b}]/u$.

Service

Once a data packet arrives to a virtual excursion region queue server, the excursion node that is holding the packet must contend for the channel with other excursion nodes who also want to transmit their packets. Once a packet's contention period is over and it has successfully captured the channel, the packet will then need to be transmitted to the next hop in the packet's routing path. We will view each hop-neighbourhood that a packet passes through in its journey to the sink as a virtual queue as before. Denote by X_j the total service time of a packet in a hop-neighbourhood in ring j. We can express the service time of a packet at each hop as being composed of the sum of a contention period, T_{con}, and a packet transmission period, T_{tx}. That is:

$$X_j = T_{con} + T_{tx} \tag{6.12}$$

We first analyze the contention period, T_{con}, at each hop. Assume that each sampling interval of Δt seconds is divided into slots of duration δ seconds. Recalling our previously described contention model, we will assume that the number, *n,* of contending nodes in each hop-neighbourhood is known to all nodes with data packets and they set their attempt probability to $p = 1/n$ in each slot. Thus, we can assume that $P_s = 1/e$ is the probability that a single node successfully captures the channel in each slot. Since the number of slots between consecutive successful channel captures is geometrically distributed with parameter $P_s = 1/e$, the mean number of slots in a packet's contention period is just $1/P_s = e$ slots. Thus, the mean time that a node with a packet to transmit spends contending for channel is $\delta \cdot e$ seconds. Like before, we will approximate the

geometric distribution for the number of slots in a contention period with the exponential distribution. That is, we will assume that the contention period of a packet in each hop is exponentially distributed with mean value $\delta \cdot e$ seconds. We then have that:

$$\overline{T}_{con} = \delta \cdot e \tag{6.13}$$

We now analyze the transmission period, T_{tx}, of a packet at each hop in its routing path. Recall that due to the data implosion that occurs in WSNs, that multiple streams of data may converge at hop-neighbourhoods as we move up the routing path towards the sink. In our case, each stream of data corresponds to the data packets coming from each isolated excursion region $A_{i,b}$ in the network. We expect that these streams will need to be multiplexed. As discussed earlier, this merging of streams will increase contention in the hop-neighbourhood and therefore increase the total service time of a packet there. We model the increased service time of a packet in a hop due to merging of different streams by assuming that at each hop-neighbourhood, the fixed length packets from each stream are combined into a super-packet. The length of this super-packet will be proportional to the number of streams that are multiplexed in the hop-neighbourhood. Denote by $\overline{M}_j$ the average number of external streams that merge at a hop-neighbourhood. We will assume that the average transmission time of a single (non-combined) packet is $1/\mu_{packet}$ seconds. Thus, the average transmission time of a packet in the hop-neighbourhoods in rings $j = 1,\ldots,J$ is given as :

$$\overline{T}_{tx}(S_j) = \frac{\overline{M}_j}{\mu_{packet}} \tag{6.14}$$

Using (6.13) and (6.14), we can express the mean service time of a packet in a hop-neighbourhood in ring S_j as:

$$\begin{aligned} \overline{X}_j &= \overline{T}_{con} + \overline{T}_{tx}(S_j) \\ &= \delta \cdot e + \frac{\overline{M}_j}{\mu_{packet}} \end{aligned} \tag{6.15}$$

We will assume that the service time of a packet at a hop in ring S_j in its routing path is exponentially distributed with mean $\overline{X}_j$ seconds as given in (6.15).

We will now offer a heuristic approach to estimating the number of merged streams at a hop-neighbourhood, $\overline{M}_j$. Without loss of generality, we imagine that the network is deployed on a semi-circular region of radius R_{Net}. Intuitively, we expect that the number of streams that pass through a given downstream hop-neighbourhood increases in some exponential manner as we get closer to the sink. Finally, at the last hop in ring S_1, we expect that all streams will merge. We model this phenomenon by creating an imaginary wedge W_j at each hop-neighbourhood in rings $S_1, S_2, \ldots, S_J$. We assume that the angle associated with this wedge is given by the following:

$$\theta_j = \frac{\pi}{2^{j-1}}, \mathrm{j} = 1, \ldots, J$$

This corresponds to halving the total angle spanned by the semi-circle at each hop-neighbourhood as we move *away* from the sink (we note that this is merely our solution to modelling the data implosion phenomenon and more elegant choices of the wedge angles may be possible by detailed consideration of the employed routing protocol).

Further, suppose that each wedge will have radius given by $r_{W_j} = R_{Net} - r_{j-1}$ which the distance from the lower boundary of the ring to the outer edge of the network deployment area R_{Net}. This wedge construction is illustrated in Figure 6.4. As noted earlier, we will assume that all streams travel roughly in a straight-line towards the sink. In fact, we will amend this notion a little and assume here that all streams which originate at excursion regions in a wedge will pass through the tip of the wedge that exists at each ring. This allows us to compute the number of streams which have been merged at a hop neighbourhood in ring S_j as simply the number of isolated excursion regions which occur in the wedge W_j at this hop. By our construction of the wedge angles and radii, the area of a wedge at ring j is given by

$$A(W_j) = \frac{\pi\left(R_{Net} - r_{j-1}\right)^2}{2^j} \tag{6.16}$$

Since the isolated excursion regions occur according to a spatial Poisson process in the plane with intensity μ_b, we can then compute the number of streams that have been merged at a hop neighbourhood in S_j as:

$$\overline{M}_j = \mu_b \cdot A(W_j) \tag{6.17}$$

We can then express the mean service time of a packet at a hop-neighbourhood in S_j as:

$$\begin{aligned} \overline{X}_j &= \delta \cdot e + \frac{\overline{M}_j}{\mu_{packet}} \\ &= \delta \cdot e + \frac{\mu_b \cdot A(W_j)}{\mu_{packet}} \end{aligned} \tag{6.18}$$

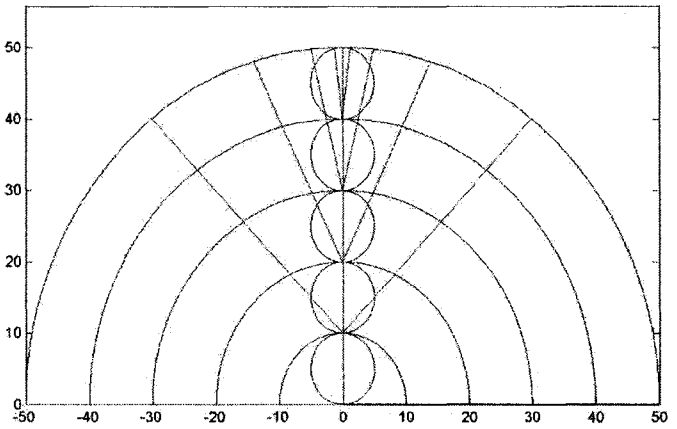

Figure 6.4: Wedge construction for determining the average number of merged streams in a hop-neighbourhood. The small circles correspond to hop-neighbourhoods in each ring.

Traffic Modelling Summary

We now provide a concise summary of the previous traffic modelling. Since the arrival processes of data packets from the isolated excursion regions $A_{i,b}$ to the network are Poisson with rate $\nu = \lambda E[A_{i,b}]/u$ and the corresponding service time in each subsequent hop-neighbourhood is exponential with mean given by (6.18), each hop-neighbourhood on the routing path from an excursion node to the sink can be modelled as a "virtual" M/M/1 queue. Thus, the hop-neighbourhoods that a packet passes through on its routing path to the sink corresponds to a series of M/M/1 queues in tandem. We show the topology of a routing path in Figure 6.5.

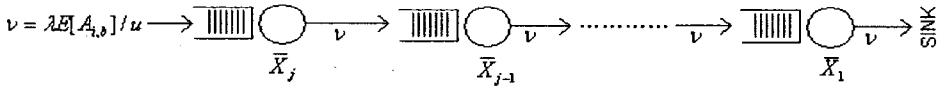

Figure 6.5: M/M/1 queue model for the hop-neighbourhoods along a packet's routing path.

Our first-cut queuing model permits analysis of several performance measures which we explore in the following sections.

6.2.3 Server Utilization and Stability

In order that the virtual queues in our network our stable, it is critical that the utilization of the server in a hop-neighbourhood in ring j, defined as $\rho_j = \nu \cdot \overline{X}_j$, is less than one. Physically, a utilization factor less than one ensures that the hop-neighbourhoods do not overload due to packet arrivals happening faster than they can be serviced there. Since the last hop ring, $j = 1$, is a bottleneck where all streams are merged (and hence, service is slowest), it is sufficient to just analyze the server utilization in this hop. This is because intuitively, stability in the last hop implies stability in all earlier hops further from the sink. Using formulas (6.5) and (6.11), the last-hop server utilization is:

$$\begin{aligned} \rho_1 &= \nu \cdot \overline{X}_1 \\ &= \frac{\lambda E[A_{i,b}](1-q)}{\Delta t} \cdot \overline{X}_1 \end{aligned} \tag{6.19}$$

In order for the utilization to be less than 1 in (6.19), the sampling interval must satisfy the following lower bound for a given probability of suppression q at an excursion node:

$$\Delta t > \lambda E[A_{i,b}](1-q) \cdot \overline{X}_1$$

Denote by $\Delta t_{\min}$ the lowest possible value of the sampling interval Δt such that $\rho_1 < 1$. We have:

$$\Delta t_{\min}(q) = \lambda E[A_{i,b}] \cdot (1-q) \cdot \overline{X}_1 \tag{6.20}$$

Clearly for a given value of q, we have last-hop stability (i.e. $\rho_1 < 1$) if and only if $\Delta t > \Delta t_{\min}(q)$.

6.2.4 Average Packet Delay

In this section, we will use our M/M/1 queuing model from the previous section to derive the delays associated with a packet generated at an excursion node.

Average Delay of a Packet in a Single Hop

Assume that all hop-neighbourhood queues are stable by choosing both Δt so that the utilization factor of the hop-neighbourhoods in every ring are less than 1. To compute how much time a packet spends in each hop on average, we use the well known result from M/M/1 queues. Suppose that a packet is currently in a hop-neighbourhood in the *j'th* ring. We have:

$$\bar{d}_j = \frac{\overline{X}_j}{1-\rho_j} \tag{6.21}$$

where $\rho_j = \nu \cdot \overline{X}_j$. This value is clearly dependent on which ring the packet is currently in. We can find the average delay at any hop-neighbourhood by unconditioning on the

probability that the packet is in a hop neighbourhood in ring S_j. Recall, we assumed that the network deployment area is a semi-circle of radius R_{Net} Denote

$$\begin{aligned} p_j &= \Pr(\text{Hop is in } S_j) \\ &= \frac{A(S_j)}{A(S)} \\ &= \frac{r_j^{\,2} - r_{j-1}^{\,2}}{R_{Net}^{\,2}} \end{aligned} \tag{6.22}$$

Using (6.22), the unconditional average delay of a packet at a hop-neighbourhood is then:

$$\begin{aligned} \bar{d} &= \sum_{j=1}^{J} p_j \cdot \bar{d}_j \\ &= \frac{1}{A(S)} \sum_{i=1}^{J} A(S_j) \cdot \bar{d}_i \end{aligned} \tag{6.23}$$

Average Source-to-Sink Delay of a Packet

We reiterate our assumption that data packets traverse a single ring at a time in the direction of the sink. For a packet originating at an excursion region in ring S_j, the total average delay along the routing path is then:

$$\bar{D}_j = \sum_{i=1}^{j} \bar{d}_i, \text{ for } j = 1, \ldots, J \tag{6.24}$$

Since the isolated excursion regions are uniformly distributed on the plane with intensity μ_b, the probability that the original excursion region was in ring S_j is given by p_j as in expression (6.22). As before, we can derive the unconditional value of the total average delay in the following manner:

$$\overline{D} = \sum_{j=1}^{J} p_j \cdot \overline{D}_j \tag{6.25}$$

6.2.5 Numerical Results

The purpose of this section is to plot some curves for the previous derivations. First we need to make some assumptions about the channel, packet sizes, and slot time. We will assume data packets are $L = 10$ bytes (or 80 bits) long and that the data rate is 30 kbps. Using these two values, the mean transmission time of a packet is $1/\mu_{packet} = 80$ bits/ 30 kbps ≈ 2.67 msec. Let the time slots have of duration $\delta = 10\,\mu$sec. These are all fairly typical values in expected WSNs [1], [18].

We assume that the network consists of $N = 5000$ nodes deployed on a semi-circular region with a radius 300 metres and that the hop distance in each transmission is $r_{hop} = 10$ metres. We suppose the level associated with the excursion regions is $b = 1.0$ and all data has been standardized to have zero mean and unit variance with correlation level $\alpha_1 = 15$ ($\alpha_2 = 1$ as usual). In Table 6-3 , we get the following values for excursion set quantities under the conditions described above:

Table 6-3: Excursion Region Properties in M/M/1 Model .

N=5000, R_{Net}=300 m, b=1 $\alpha_1 = 15$, $\alpha_2 = 1$	
$E[A_{i,b}]$ (m^2)	303.8895
λ	0.0354
$\overline{K}_{i,b}$ (nodes)	10.7577

The first order of business is to ensure last-hop stability by choosing the sampling interval sufficiently high. In Figure 6.6, we plot the minimum achievable sampling interval from (6.20) as a function of the suppression probability q for the current scenario.

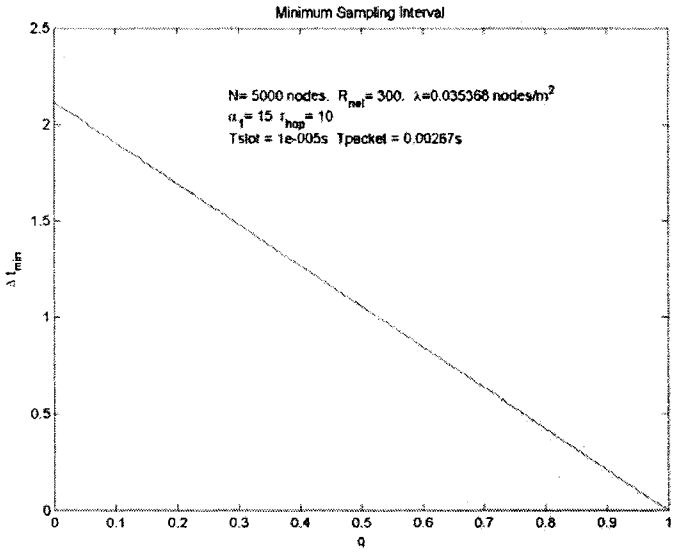

Figure 6.6: Minimum sampling interval $\Delta t_{\min}$ using (6.20).

Based on Figure 6.6, we comment that as the excursion node suppression probability q is increased, we can sample the network nodes more often by using a smaller sampling interval, Δt. This shows that in terms of the maximum achievable sampling rate $1/\Delta t$, WSNs that implement some form of data suppression can outperform WSNs that do not suppress data.

Suppose that we choose $\Delta t = 2.4$ seconds so that the bound in (6.20) is satisfied and we have last-hop stability for $\forall q \in (0,1]$. In Figure 6.7, we show the mean delay at a hop in each ring for three different values of the suppression probability of an excursion node; $q = \{0.2, 0.5, 0.8\}$. Then, in Figure 6.8 we show the total source-to-sink delay of a data packet originating at an excursion region in rings $j = 1,\ldots,J$ for these same suppression probabilities.

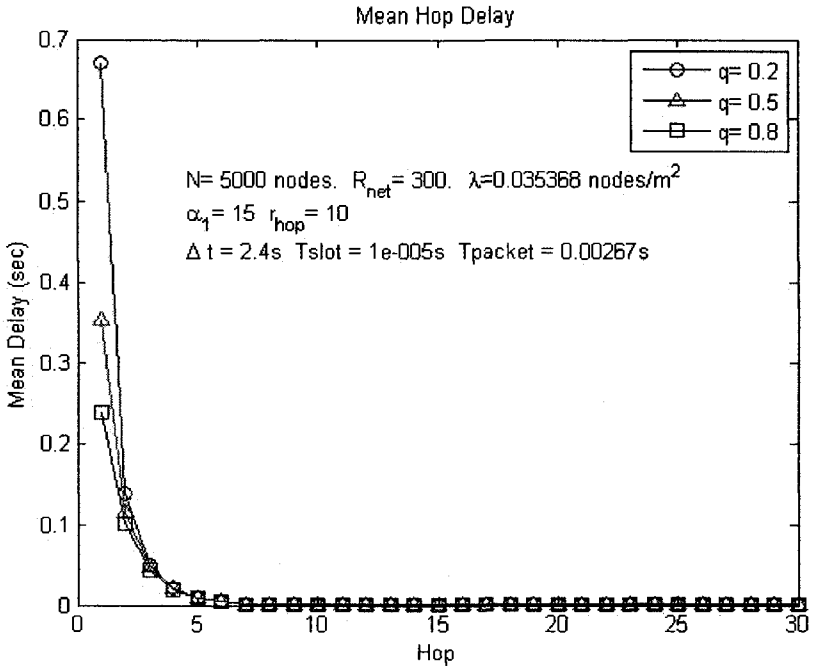

Figure 6.7: The mean delay of a packet at each hop in the network for different values of the excursion node suppression probability, q.

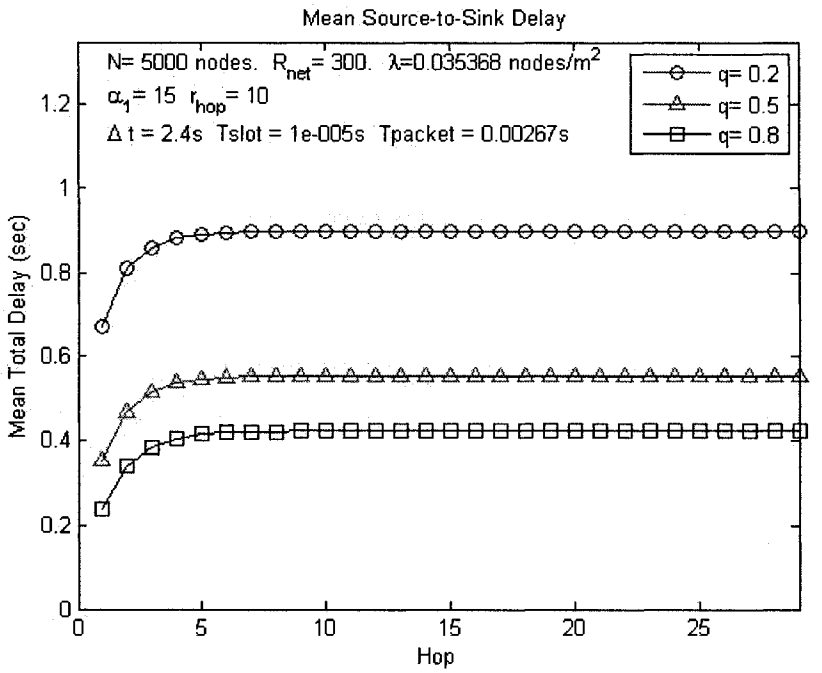

Figure 6.8: The mean source-to-sink delay of a packet originating at an excursion region in ring $j = 1,\ldots 30$ for different values of the excursion node suppression probability, q.

Figures 6.7 and 6.8 show that the delay of a packet is decreased as the suppression probability, q, increases at the excursion nodes. We note however that the suppression probability is not a design parameter that can be tuned by network designers in order to achieve performance improvements. Instead, the achievable value of q is determined by the spatial correlation exhibited by the underlying phenomenon that the network is observing and the density of the node deployment. In general, we can expect to be able to achieve higher values of q when the network observes a highly correlated phenomenon. This is because when the phenomenon has higher correlation, neighbouring nodes will tend to overhear more similar values.

For applications where the correlation of the phenomenon is moderate or even low, the designer may effectively trade-off resolution of the data set at the sink for lower packet delay. To do this, suppose that the criterion for a node suppressing its data is that it overhears a neighbouring node transmit a value within some threshold, ε, of its sensor reading. If it is known beforehand that the phenomenon has relatively low spatial correlation, than the threshold parameter, ε, can be increased. Of course this will tend to tend to suppress data transmissions that are more dissimilar than if we used a lower value of ε. Since the suppressed transmissions contain less spatial redundancy than if a lower value of ε was used, the resolution of the data set at the sink is lower. However, for WSNs deployed to report on some emergency phenomenon the resolution of data set at the sink may be less important than achieving a lower latency of data delivery.

Another trade-off that can be made in WSNs using a contention-based model is increasing the sampling interval, Δt, in order to achieve lower delay for a data packet. We show this by plotting the source-to-sink delay for a fixed value of $q = 0.5$ for increasing values of Δt in Figures 6.9, 6.10, and 6.11.

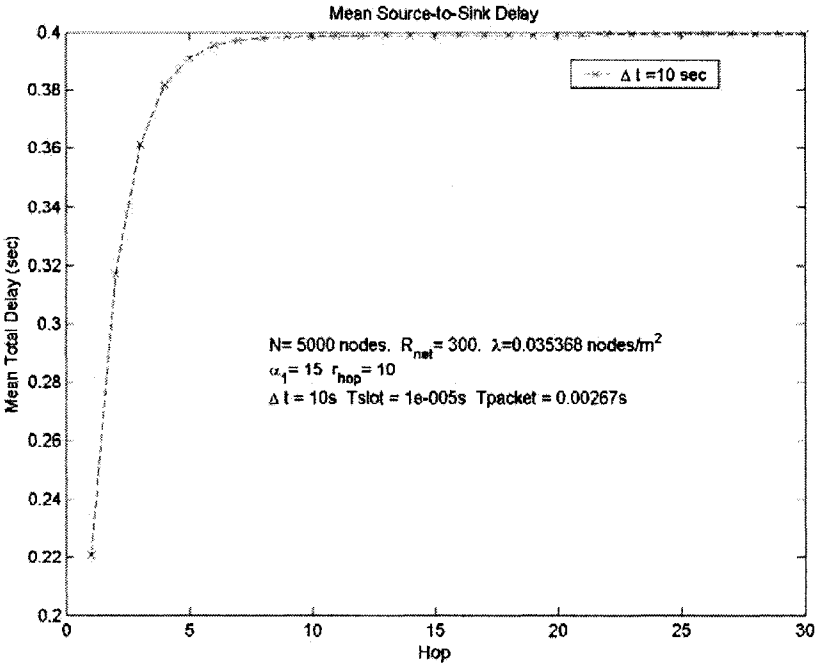

Figure 6.9: Source-to-sink delay for packets originating at an excursion node in ring $j = 1,\ldots 30$ for $q = 0.5$ and $\Delta t = 10$ seconds.

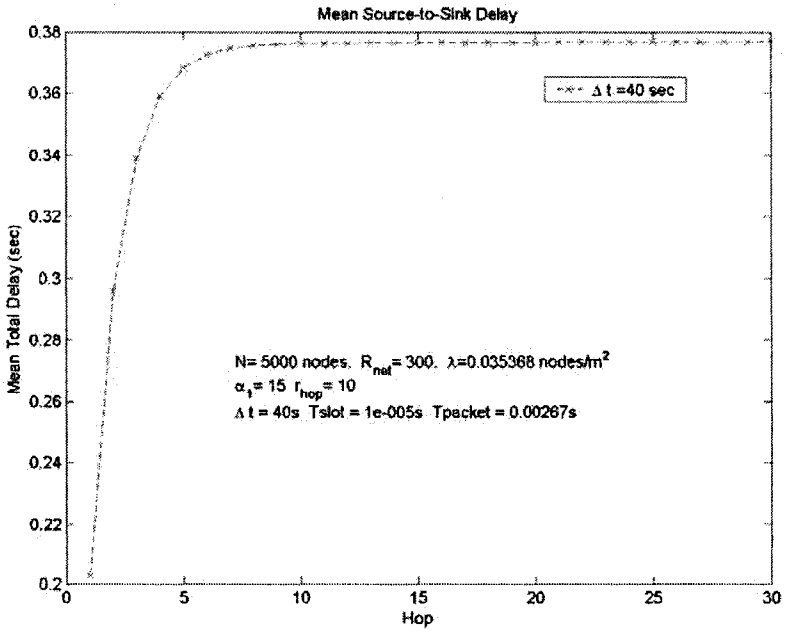

Figure 6.10: Source-to-sink delay for packets originating at an excursion node in ring $j = 1,\ldots 30$ for $q = 0.5$ and $\Delta t = 40$ seconds.

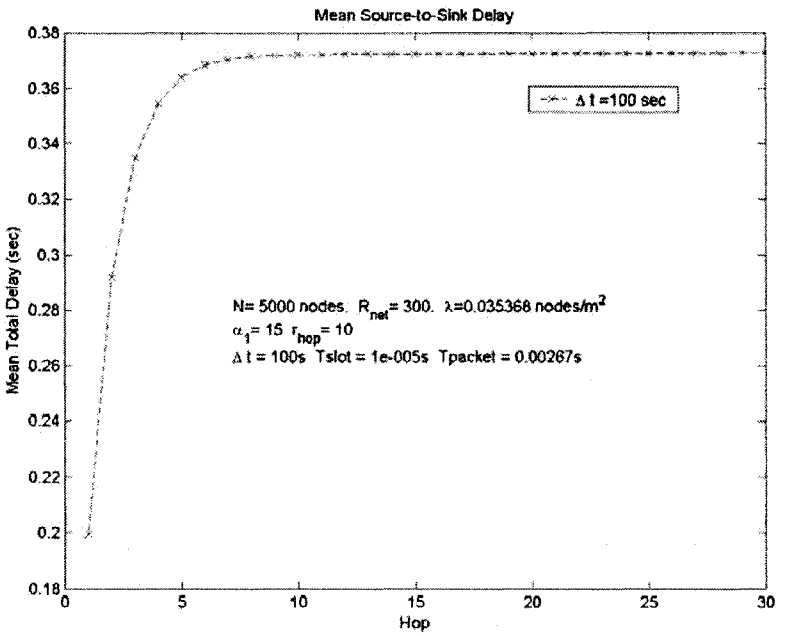

Figure 6.11: Source-to-sink delay for packets originating at an excursion node in ring $j = 1,\ldots 30$ for $q = 0.5$ and $\Delta t = 100$ seconds.

Comparison of Figures 6.9, 6.10, and 6.11 shows that increasing the sample interval cannot arbitrarily decrease the source-to-sink delay. Instead, this delay stabilizes after about $\Delta t = 40$ seconds. The minimum achievable end-to-end delay under our network assumptions is about 0.37 seconds for packets from excursion nodes located in ring $j = 30$ (the furthest ring from sink) and 0.20 seconds for packets from excursion nodes in ring $j = 1$ (the ring closest to sink). As usual in WSN design, the application scenario will determine whether an increased sampling interval of $\Delta t = 40$ seconds is justified in order to decrease the source-to-sink delay for our contention-based system.

6.2.6 TDMA and Contention-Based Performance Comparison

In this section, we will compare the performance of a WSN using a contention-based MAC protocol with the same network using a TDMA-based protocol. We will use packet loss and mean end-to-end delay as metrics for performance comparison.

Before we can compare packet loss of the contention-based system to our TDMA system, we need to first define a metric for both systems which will permit objective comparison. We define the packet loss rate, r_L, as the number of packets that are dropped (in the TDMA system) or suppressed (in the contention-based system) from an isolated excursion region, $A_{i,b}$, per unit time. Recall that in both system models, we will sample the network every Δt seconds. In our contention-based system, this packet loss rate is given by:

$$r_L^{cont} = \overline{K}_{i,b} \cdot q \cdot \frac{1}{\Delta t} \tag{6.26}$$

In our TDMA system, the packet loss rate is:

$$r_L^{TDMA} = \frac{\overline{L}_{i,b}}{\Delta t} \tag{6.27}$$

where $\overline{L}_{i,b}$ is given by (6.2)

From (6.26), we see that in a contention-based system the rate of packet loss is a function of the average number of nodes in an isolated excursion region, the suppression

probability, q, and the sampling interval, Δt. Recall, that q is determined by the spatial correlation of the underlying phenomenon. Thus, we will arbitrarily assume $q = 0.5$. Although, increasing the sampling interval for our contention-based system will arbitrarily reduce the packet loss rate, this is not really an option in settings that require periodic monitoring or up-to-date information about the phenomenon. Thus, we will assume as before that in our contention-based system we have arbitrarily set $\Delta t = 2.4$ seconds. As before, we assume that $N = 5000$ nodes have been deployed on a semi-circle with a radius of 300 metres, b=1 is the excursion threshold, and the underlying phenomenon has correlation parameters $\alpha_1 = 15$, $\alpha_2 = 1$. This gives $r_L^{cont} = 2.24$ packets/second in each excursion region.

Now consider the TDMA system in the same deployment described above with $s = 14$ slots each with duration of 2.67 msec. These values imply a TDMA frame duration of $\Delta t = 37.38$ msec and $\bar{L}_{i,b} = 0.3102$ packets. This creates a packet loss rate of $r_L^{TDMA} = 8.30$ packets/second in each excursion region. Clearly, this is quite bad compared to our contention-based system. However, if we increase the number of slots to $s = 17$, the packet loss rate becomes $r_L^{TDMA} = 1.2107$ which now outperforms the contention-based system. The cost of achieving this performance improvement is an increase in the sampling interval to $\Delta t = 17 \times 2.67$ msec = 45.39 msec. Recall, that for a similar packet loss rate in the contention-based system we must sample the network nodes at least every 2.4 seconds. In this scenario, we can therefore sample the network using the TDMA based system at least 52 times more frequently than in the contention-

based system and simultaneously achieve a lower packet loss rate. Thus, we conclude that the TDMA system can *always* outperform the contention-based system in terms of rate of packet loss and simultaneously sample the network nodes significantly more frequently.

We now compare the two systems in terms of source-to-sink delay. For a TDMA system the mean delay of a packet is just $\overline{D}_{TDMA} = \frac{\Delta t}{2}$ because a packet can be scheduled in any of the s slots with equal probability. Under our assumptions above using $s = 17$ as a worst-case value for the delay, we have $\overline{D}_{TDMA} = 45.39$ msec / 2 = 22.69 msec. For our contention-based system, we saw earlier in Figures 6.9 , 6.10, and 6.11 that the minimum achievable source-to-sink delay for $q = 0.5$ (recall q is fixed by the underlying phenomenon) stabilizes after about $\Delta t = 40$ seconds to approximately $\overline{D}_{con} = 0.20$ seconds for a packet from an excursion node close to the sink and $\overline{D}_{con} = 0.37$ seconds for a packet from an excursion node furthest from the sink. Despite the delay improvement of our contention-based system when the sampling interval is increased to $\Delta t = 20$ seconds, the delay is still greater than the delay of our TDMA frame using $s = 17$ slots (which we saw above was 22.69 msec). In general, we can expect the delay of the TDMA system to outperform our contention-based system. The main reason for lower source-to-sink delay in the TDMA system is that communication between excursion nodes and the sink occurs in a single hop where as our contention-based system uses multihop communication to reach the sink and experiences high hop delays close to the sink due to the data implosion problem.

Although the TDMA system can simultaneously outperform the contention-based system in terms of packet loss rate, minimum sampling interval, and source-to-sink delay, there is a cost that is hidden from our performance analysis. In our model of the TDMA system, we either require high-powered devices to act as gateways for the excursion nodes so that data can be transmitted to the sink in a single hop, or that regular-powered excursion nodes make a high-powered transmission to the sink in a single hop. In our models, single transmissions are therefore more energy costly in the TDMA system. Whether or not high-powered devices are available and practical in a WSN or excursion nodes can even transmit to the sink in a single hop will depend primarily on the available technology at the time of deployment.

Chapter 7 Conclusions and Future Work

7.1 Conclusions

In this thesis we have assumed that in many application scenarios, where there are potentially thousands of nodes that sense an environment, it will be necessary to limit the set of transmitting nodes. Our solution to this problem is to impose the following transmission criterion in the network: only allow a node to transmit its sensor reading if it records an "extreme" (i.e. statistically high) value when sampled. In the natural sciences, many phenomena are modelled using stationary Gaussian random fields. Through personal communications with Dr. Erik Vanmarcke of Princeton University (an expert in random field theory and author of [25]), it was noted that Gaussian random fields have been previously used to model a diverse array of phenomena in the natural sciences including the ground motion of an earthquake at small distances (say less than 1km), shear wave velocities in rock masses, wind speeds during severe storms, the heights of ocean waves, and salinity in oceans. Thus, we have assumed throughout this thesis that the WSN has been deployed to monitor a large-scale environmental phenomenon that can be modelled as a stationary Gaussian random field. A central topic in the theory of random fields is the study of the size and spatial distribution of the isolated regions of the field that experience extreme values. In this thesis, we have used the theory of high level excursion regions for stationary Gaussian random fields in order to gain insight into the size and spatial distribution of the set of nodes which will want to transmit their data to the sink.

A common set of results in this thesis concern properties of the underlying phenomenon observed by the network. These results follow directly from the theory of random fields. They are:

- A model is provided for the underlying phenomenon observed by the network using a stationary Gaussian random field.
- An expression is given for the average area of an isolated region that experiences extreme values on the network deployment area. We called these extreme regions "excursion regions".
- The number of isolated excursion regions on the deployment area follows the Poisson distribution and the parameter of this distribution has been determined.
- The average area of an isolated excursion region is inversely proportional to the frequency of the occurrence of these regions on the network deployment area.
- The total average area of the deployment region that experiences values above some high level b is only a function of b and is independent of the correlation level.

Our work then used these results to understand various properties of nodes which observe extreme data in the network when the underlying phenomenon is modelled as a stationary Gaussian random field. Formally, our contributions in this area are:

- Derivation of the joint probability distribution of the data observed at the nodes in the network.

- The distribution of the number of nodes on an isolated excursion region approximately follows a Poisson distribution. The parameter of this distribution has been determined. We called nodes in an isolated excursion region "excursion nodes".
- An expression for the average number of excursion nodes on the entire network deployment area.
- Derivation of the average number of nodes who observe data belonging to some (preferably) high level contour line of non-zero width. We called these "contour nodes". We showed that we can further limit the set of transmitting nodes by only allowing contour nodes to transmit their data.
- An extension of our result for the average number of excursion nodes on the entire network deployment area to scenarios where the underlying phenomenon exhibits a location-dependent mean value and is thus, non-stationary has been provided.

Then we used our results regarding excursion regions and excursion nodes to study the nature of traffic in a WSN using either a TDMA-based or a CSMA-based MAC protocol. In both cases, it is assumed that only extreme data will be transmitted in the network. In the TDMA-based system, transmissions of data packets from excursion nodes to the sink occur in a single hop. In the CSMA-based system, the transmission of data to the sink is done in a multi-hop manner and non-excursion nodes assist with routing the data packets from excursion nodes. Our contributions in this area are:

- A description of the size and spatial distribution of the regions in a wireless sensor network that produce traffic characterized by containing extreme data values.
- Derivation of the distribution of the packet loss in a TDMA-based system.
- Derivation of the minimum achievable sampling rate, stability, and delay in WSNs that use a CSMA-based MAC protocol.

As a final conclusion, the work in this thesis shows that limiting the set of transmitting nodes to those which sense extreme values will reduce the network traffic load significantly. This will in turn reduce the latency in the delivery of data to the end user and will increase the lifetime of the network.

7.2 *Future Work*

Our work really only serves as an introduction to the use of random field theory in understanding the performance of a WSN deployed to monitor some phenomenon. Much of our modelling is purely illustrative. This is in order to show WSN designers the potential of using insight in random field theory in understanding the response of the network to the observed phenomenon. In practice, many environmental phenomena have been modelled using random field theory. As well there exists literature on excursion sets for non-statistically high levels b and for non-Gaussian phenomenon. Thus, we expect that the results from the aforementioned literature can be applied to application scenarios that meet these criteria (i.e. non-statistically high levels b and/or non-Gaussian phenomenon). Deployed WSNs are not one-size-fits-all type networks. Instead, they

must be designed according to the specific application scenario. The basic idea we wish to impart to sensor network designers is that in order to really know how a WSN with a given protocol suite will perform when it is deployed in a real environment, one needs to be able to describe the underlying phenomenon with some degree of accuracy. Random fields are an easy way to describe many naturally occurring phenomena. Future work may look at existing random field models for specific physical phenomena and study the performance of WSN in these applications. As well, modelling and analysis using random fields can be used to study the spatial and temporal distribution of energy consumption in WSNs in a similar vein to our work in chapter 6.

Appendix A Random Field Background Theory

A1 Preliminaries

When it comes to random processes in a single dimension, there is a wealth of literature for the analysis of such important topics as continuity and differentiability of sample paths, number of upcrossings of a high threshold, maximum values, first passage times, etc. However, the generalization of the existing theory in one dimension to higher dimensions is non-trivial and will be the topic of the following primer on some essentials of random field theory. We will often be interested in random fields defined on *n*-dimensional Euclidean space $\Re^n$ or a subset of *n*-dimensional space, $S \subseteq \Re^n$. The contents of this appendix have been distilled from [25] and [27].

Definition: *A random field* $X(\bar{s})$ *on* $S \subseteq \Re^n$ *is a function whose values are random variables for any* $\bar{s} = (s_1, s_2, \ldots, s_n) \in S$

A2 Finite-Dimensional Distributions

A random field is described by its *finite-dimensional cumulative distribution* at an arbitrary number (say *k*) points in S:

$$F_{\bar{s}_1,\ldots,\bar{s}_k}(x_1,\ldots,x_k) = \Pr(X_{\bar{s}_1} \le x_1,\ldots,X_{\bar{s}_k} \le x_k)$$

The finite-dimensional joint density function is as usual, found by differentiating the cumulative distribution function with respect to each $x_i, i = 1 \ldots k$:

$$f_{\bar{s}_1,\ldots,\bar{s}_k}(x_1,\ldots,x_k) = \frac{\partial^k F_{\bar{s}_1,\ldots,\bar{s}_k}(x_1,\ldots,x_k)}{\partial x_1 \cdots \partial x_k}$$

Note that whereas for random processes the parameter space is indexed by a parameter in a single dimension $\Re$ (usually time), the parameter space in random fields can be indexed by a general vector $\bar{s} \in S \subseteq \Re^n$. For example, this vector could represent 2-dimensional space, 2-dimensional space coupled with time ($S \times T \subseteq \Re^2 \times \Re^1$). In general, any subset of n-dimensional space is possible. Denote by $X(\bar{s}) = x(\bar{s}), \forall \bar{s} \in S$ a realization of a random field in some subset of *n*-dimensional space. Note that both the parameter-space and the possible outcomes of a random field may be either discrete or continuous.

We can define the expectation at a point $\bar{s} \in S$ as follows. Assume that the outcomes of the random variable $X(\bar{s})$ are continuous and denote by $f_{X(\bar{s})}(x)$, the probability density function of the random field at $\bar{s}$. Then we have:

$$m(\bar{s}) = E[X(\bar{s})] = \int_{-\infty}^{\infty} x \cdot f_{X(\bar{s})}(x) dx$$

The covariance function between two points $\bar{s}, \bar{r} \in S$ is defined as:

$$C(\bar{s}, \bar{r}) = E[X(\bar{s})X(\bar{r})] - m(\bar{s})m(\bar{r})$$

From the covariance function, we determine the variance of the random field at a point $\bar{s}$ as:

$$\sigma^2(\bar{s}) = C(\bar{s}, \bar{s})$$

With the covariance and variance functions, the correlation function between points $\bar{s}, \bar{r}$ is:

$$\rho(\bar{s},\bar{t}) = \frac{C(\bar{s},\bar{r})}{\sigma(\bar{s})\sigma(\bar{r})}$$

A3 Positive Definiteness

A fundamental concept in random field theory is positive definiteness. This arises from the fact that the set of positive definite functions coincides with the set of covariance functions [27]. So, a necessary condition for a function $C(\bar{s},\bar{r})$ to serve as an admissible covariance function is that it is positive definite.

Definition: *Let k be a positive integer, and let* $\bar{s}_i \in S$ *and* $c_i \in \Re^1$ *be an arbitrary constant for* $i = 1,\ldots,k$. *Then the function C on* $S \otimes S$ *is said to be positive definite on* $\Re^n$ *if:*

$$\sum_{i=1}^{k}\sum_{j=1}^{k} c_i c_j \cdot C(\bar{s}_i,\bar{s}_j) \geq 0$$

For any choice of k, $\{\bar{s}_i,\ldots,\bar{s}_j\}$, *and* $\{c_i,\ldots,c_j\}$.

We now list several properties of positive definite functions which follow from the definition and are worth knowing:

- The product of two positive definite functions is also positive definite
- The sum of two positive definite functions is positive definite
- Positive definiteness is preserved under multiplication of a positive scalar

A4 Gaussian Random Fields

Gaussian random fields occupy a central role in random field theory. This is because they are natural models for many physical phenomena, and the corresponding finite-dimensional distribution is completely specified by the mean and covariance functions [25], [27].

Suppose we sample a random field at k points, $\bar{s}_1,\ldots,\bar{s}_k$, on $S \subseteq \Re^n$. Then the random field is Gaussian if the finite-dimensional distribution $F_{\bar{s}_1,\ldots,\bar{s}_k}(x_1,\ldots,x_k)$ is multivariate normal for any choice of the k points, $\bar{s}_1,\ldots,\bar{s}_k$.

Let $\bar{x} = [x_1,\ldots,x_k]$ be the vector of values of the random field at points $\bar{s}_1,\ldots,\bar{s}_k$ and let $\bar{m} = [m(\bar{s}_1),\ldots,m(\bar{s}_k)]$ be the respective mean vector. Further, let $\mathbf{K}$ be the matrix of covariances between the values at the points $\bar{s}_1,\ldots,\bar{s}_k$. Then the joint probability density function (pdf) associated with the k sample points of the random field is

$$f_{\bar{s}_1,\ldots,\bar{s}_k}(x_1,\ldots,x_k) = \frac{1}{2\pi|\mathbf{K}|^{1/2}} \exp\left\{-\frac{1}{2}(\bar{x}-\bar{m})^T \mathbf{K}^{-1}(\bar{x}-\bar{m})\right\}$$

In general the mean vector $\bar{m}$ can be chosen arbitrarily. However, to be a valid covariance matrix, $\mathbf{K}$ must be positive definite. Thus, the construction of an admissible Gaussian random field model is really about specifying a valid positive definite covariance structure.

A5 Symmetries

The class of admissible covariance functions is often restricted to reflect physical symmetry of the random field and simultaneously simplify the verification of positive definiteness. The covariance function is said to have symmetry if it is invariant under some transformation of the parameter space $S \subseteq \Re^n$.

Stationarity

Assume that S is a vector space in $\Re^n$ (in other words, a subspace of $\Re^n$ closed under vector addition and scalar multiplication).

Definition: (Strict Sense Stationarity)

A random field is stationary in the strict sense if all its finite-dimensional distributions are invariant under arbitrary translations by a vector $\vec{r}$. That is:

$$F_{\vec{s}_1+\vec{r},\ldots,\vec{s}_k+\vec{r}}(x_1,\ldots,x_k) = F_{\vec{s}_1,\ldots,\vec{s}_k}(x_1,\ldots,x_k) \text{ for all } \vec{r} \in S$$

Definition: (Wide Sense Stationarity)

A random field is stationary in the wide sense if $m(\vec{s}) = m$ *and* $C(\vec{s}_1,\vec{s}_2) = C(\vec{\tau})$

where $\vec{\tau} = \vec{s}_1 - \vec{s}_2$ *is the separation vector.*

Note that strict-sense stationarity implies wide-sense stationarity although the converse is not necessarily true. For Gaussian random fields however, the two are equivalent.

Isotropy

An even more restricted set of covariance functions are the so-called isotropic covariance functions. These are a subclass of stationary covariance functions and possess two additional symmetries besides translation invariance: rotation and reflection invariance. First, assume that the vector space $S \subseteq \Re^n$ is equipped with a well-defined distance measure (or norm) denoted by $d(\vec{s},\vec{r})$. The most common example of this is $\Re^n$ equipped with the Euclidean *n*-norm,

$$\tau = d(\vec{s},\vec{r}) = \| \vec{\tau} \| = \sqrt{(s^{(1)} - r^{(1)})^2 + \ldots + (s^{(k)} - r^{(k)})^2} \qquad \text{(A.1)}$$

Definition: *(Isotropic Random Field)*

A stationary random field on $S \subseteq \Re^n$ is isotropic in the wide sense if the covariance function depends only on the distance between the two points. That is:

$$C(\vec{s},\vec{r}) = C(\tau)$$

A6 Separability

In modelling spatial uncertainty of phenomenon in space and time in physical sciences like meteorology and geosciences, it is common to assume that the covariance structure of naturally occurring phenomenon is separable in space and time [25]. This means we can write the covariance function as the product of a purely spatial covariance function and a purely temporal covariance function.

Definition: *(Separable Space-Time Covariance Structure)*

A covariance function $C((\vec{s},t_1),(\vec{r},t_2))$ *of a spatio-temporal random field on* $S \times T$ *(where* $T = [0,\infty]$ *) is separable in space and time if it can be expressed as:*

$$C((\vec{s},t_1),(\vec{r},t_2)) = C_S(\vec{s},\vec{r}) \cdot C_T(t_1,t_2)$$

A7 Analytical Properties

Continuity

Although there are different ways to define the continuity of a random field, the form we will be interested in, especially for Gaussian random fields, is mean-squared continuity.

Definition: (Mean-Squared Continuity)
A random field X is mean-square continuous in $S \subseteq \Re^n$ *if for every sequence* $\{\vec{s}_k\}$ *such that* $\|\vec{s}_k - \vec{s}\| \to 0$ as $n \to \infty$, *then*

$$E[|X_i(\vec{s}_k) - X_i(\vec{s})|^2] \to 0 \text{ as } n \to \infty \text{ for } \forall \vec{s} \in S$$

Note that in general, mean-squared continuity of a random field does not imply that the sample paths of a random field will be continuous. However, for Gaussian random fields, mean-squared continuity is necessary and almost sufficient condition for the continuity of the sample paths.

Theorem: *Assume that* $E[X(\bar{s})]$ *is continuous. Then a random field* $X(\bar{s})$ *is mean-square continuous at* $\bar{s} \in \Re^n$ *if and only if its covariance function* $C(\bar{s},\bar{r})$ *is continuous at* $\bar{s} = \bar{r}$.

A consequence of this theorem is given by the following corollary:

Corollary: *A stationary random field* $X(\bar{s})$ *is mean-square continuous at any* $\bar{s} \in \Re^n$ *if and only if its correlation function* $\rho(\bar{\tau})$ *is continuous at* $\bar{0}$.

For a Gaussian random field we have the following theorem which states when mean-square continuity will guarantee the continuity of the sample paths:

Theorem: *Let* $X(\bar{s})$ *be a stationary Gaussian random field with a continuous correlation function. Then if for some finite* $c > 0$, *and some* $\varepsilon > 0$,

$$1 - \rho(\bar{\tau}) \leq \frac{c}{|\log(\tau)|^{1+\varepsilon}}$$

For all $\bar{\tau}$ *with* $\tau = \|\bar{\tau}\| < 1$, *then the random field* $X(\bar{s})$ *will have continuous sample paths with probability one.*

It is worth noting that this bound is almost trivially satisfied for stationary Gaussian random fields. The bound is an ever increasing function of τ and approaches infinity as $\tau \to 1$. [27] states that it is almost impossible to find a continuous correlation function violating this bound. The following conjecture is then posited:

Conjecture: *Stationary Gaussian random fields with continuous expectations and covariance functions posses continuous sample paths with probability one.*

Differentiability

Let $X(\bar{s})$ be a Gaussian random field on $\Re^n$. Assume that X has differentiable sample paths. Then we are interested in the associated gradient field $\dot{X}(\bar{s})$ which is a space vector in $\Re^n$ with components defined with respect to the standard Cartesian coordinate system:

$$\dot{X}_i(\bar{s},w) = \frac{\partial X(\bar{s})}{\partial s_i} = \lim_{\Delta \to 0} \frac{X(\bar{s} - \Delta \cdot \bar{e}_i, w) - X(\bar{s},w)}{\Delta}$$

Where w is kept fixed and $\bar{e}_i$ is a unit vector in the same direction as the i'th coordinated axis. Note that since differentiation is a linear operator, the gradient field is also Gaussian (recall that a linear transformation of jointly Gaussian random variables is also Gaussian).

Similar to continuity of a random field, we will often be interested in whether a random field (not necessarily Gaussian) is mean-square differentiable. As usual, we let $S \subseteq \Re^n$ and proceed with the following definition:

Definition: (Mean-Square Differentiability)

A random field X (not necessarily Gaussian) is mean-square differentiable in $S \subseteq \Re^n$ *if for every sequence* $\{\bar{s}_k\}$ *such that* $\|\bar{s}_k - \bar{s}\| \to 0$ as $n \to \infty$, *then*

$$E\left[\left|\dot{X}_i(\vec{s}_k)-\dot{X}_i(\vec{s})\right|^2\right]\to 0 \text{ as } n\to\infty \ \forall \mathrm{i}=1,\ldots,\mathrm{n},\ \forall \vec{s}\in S$$

Similar to mean-square continuity, mean-square differentiability is a necessary condition for differentiable sample paths. There is a simple relationship between mean-square differentiability and the covariance function as detailed in the following theorem.

Theorem: *Consider a random field* $X(\vec{s})$ *on* $\Re^n$ *with covariance function C. If the derivative* $\partial^2 C(\vec{s}_1,\vec{s}_2)/\partial s_1^{(i)}\partial s_2^{(i)}$ *exists and is finite for all* $i=1,\ldots,n$ *at the point* $(\vec{s},\vec{s})$, *then* $X(\vec{s})$ *is mean-squared differentiable at all* $\vec{s}\in\Re^n$. *The covariance function of the gradient field in the direction of the i'th coordinate axis is then given simply by* $\partial^2 C(\vec{s}_1,\vec{s}_2)/\partial s_1^{(i)}\partial s_2^{(i)}$.

For stationary random fields, the following corollary simplifies the theorem further.

Corollary: *Consider a stationary random field* $X(\vec{s})$ *on* $\Re^n$ *with covariance function C. If the derivative* $\partial^2 C(\vec{\tau})/\partial(\tau^{(i)})^2$ *exists and is finite for all* $i=1,\ldots,n$ *at the point* $\vec{0}$, *then* $X(\vec{s})$ *is mean-squared differentiable at all* $\vec{s}\in\Re^n$. *The covariance function of the gradient field in the direction of the i'th coordinate axis is then given simply by* $-\partial^2 C(\vec{\tau})/\partial(\tau^{(i)})^2$.

Note that the negative sign comes from taking the partial derivative of $C(\vec{\tau})=C(\vec{s}-\vec{t})$ with respect to $\vec{t}$. Since the covariance function of a stationary random

field must have a maximum value at $\bar{0}$, it must be that $\partial C(\bar{\tau})/\partial \tau_i = 0$ for all $i = 1,\ldots,n$ in order for the second order derivatives to exist. For an isotropic random field the partial derivatives in each direction are all equal and so it is enough to consider the existence and finiteness of $d^2C(\tau)/d\tau^2\big|_{\tau=0}$.

A8 Excursion Set Theory

Excursion sets of a random field are the *n*-dimensional generalization of level crossings of a 1-dimensional random process. The study of their properties is one of the central topics of random field theory and is very much still a "work in progress". For a random process in time, one is often interested in the amount of time that the random process spends above some fixed level *b* for a given up-crossing/down-crossing pair. This so-called excursion time is clearly a closed interval on the real line. For a random field in *n*-dimensions, this generalizes to an *n*-dimensional hyper-sphere. For our purposes we will be mainly interested in the 2-dimensional case, and so the excursion sets above a level *b* form a series of contour lines in the plane. The properties of these excursion sets have been studied by many authors but we will draw primarily from the work on [25].

It will be seen that properties of the partial derivatives of the covariance function at $\bar{0}$ will allow us to predict the average size and frequency of the excursion sets in the plane. Henceforth in this section, we shall be considering a random field in 2-dimensional space. That is, $X(\bar{s}) = X(s_1, s_2)$ with $(s_1, s_2) \in S \subseteq \Re^2$.

Before proceeding we will need to declare some quantities related to derivatives of the covariance function. We call the value of the second partial derivative of the gradient field in the direction of the i'th coordinate, the mean-square value of the gradient field and denote it as:

$$\sigma_{\dot{X}_i}^2 = \lambda_2^{(1)} = -\partial^2 C(\bar{\tau})/\partial \bar{\tau}_i^2 \Big|_{\bar{\tau}=\bar{0}} \tag{A.2}$$

The $\lambda_2^{(1)}$ notation comes from the alternate spectral representation of random fields. Also, we define:

$$\lambda_{11} = -\frac{\partial^2 C(\bar{\tau})}{\partial \tau_1 \partial \tau_2}\Big|_{\bar{\tau}=(0,0)} \tag{A.3}$$

Again, λ_{11} is a parameter from the spectral theory of random fields and is referred to as the joint spectral moment of second order.

We now form a matrix of spectral moments as:

$$\mathrm{V} = \begin{bmatrix} \lambda_2^{(1)} & \lambda_{11} \\ \lambda_{11} & \lambda_2^{(2)} \end{bmatrix} \tag{A.4}$$

In [25], the expected area of an isolated excursion region is approached by considering the marginal random processes $X(s_i)$, $i = 1, 2$ which are obtained from $X(s_1, s_2)$ by keeping one coordinate fixed and varying the other. Let $\phi(\cdot)$ and $\Phi(\cdot)$ be the standard Normal PDF and CDF functions respectively. It is then shown that:

1. The mean rate of crossings of a level b by the marginal process $X(s_i)$, $i = 1, 2$ is:

$$v_b^{(i)} = \phi(b) \cdot \frac{\sqrt{2}}{\pi} \cdot \sigma_{\dot{X}_i} \tag{A.5}$$

2. The mean length, in the direction of the s_i axis, of an isolated excursion region above b by $X(s_i)$ is:

$$E[\mathrm{T}_b^{(i)}] = \sqrt{2\pi} \cdot \frac{1-\Phi(b)}{\phi(b)} \cdot \frac{1}{\sigma_{\dot{X}_i}} \tag{A.6}$$

The quantities $E[T_b^{(i)}]$ may be thought of as the average dimensions of an isolated excursion region above b in the plane. In [25], it is then conjectured that this interpretation is valid and exact in the limit as $b \to \infty$. The expected size of an isolated excursion region in the plane is then given by:

$$\begin{aligned} E[A_{i,b}] &= \frac{E[T_b^{(1)}]E[T_b^{(2)}]}{|\mathbf{V}|^{1/2}} \\ &= \left(\frac{1-\Phi(b)}{\phi(b)}\right)^2 \frac{2\pi}{\sigma_{\dot{X}_1}\sigma_{\dot{X}_2}} \frac{1}{|\mathbf{V}|^{1/2}} \end{aligned} \tag{A.7}$$

where $|\mathbf{V}|$ is the determinant of the matrix $\mathbf{V}$ defined previously (A.4).

Some comments on the preceding theory are now in order. For "extreme" levels b (roughly, $b/\sigma \geq 2$), we are almost guaranteed a one-to-one correspondence between a b-upcrossing in the plane (i.e. an isolated excursion region) and a local maxima of the field [25]. The first statement above shows that occurrences of isolated excursion regions in the plane become increasingly rare since $\phi(b)$ is a decreasing function in b. As b increases, the excursion regions become more isolated and the local maxima of the field tend towards a Poisson process in the plane.

Now that we know there is roughly a one-to-one correspondence between local maxima and isolated excursion regions and that the local maxima tend to a Poisson process on the plane, we are interested in the frequency of the local maxima (equivalently, the excursion regions) in the plane. Denote by μ_b the mean number of local maxima per unit area in the plane, which is also the asymptotic spatial mean rate of the Poisson process of excursion regions. In [25], it is shown that:

$$\mu_b = \frac{(1-\Phi(b))}{E[A_{i,b}]} = \frac{1}{2\pi}\frac{[\phi(b)]^2}{(1-\Phi(b))}\sigma_{\dot{X}_1}\sigma_{\dot{X}_2}|\mathbf{V}|^{1/2} \qquad \text{(A.8)}$$

We can use the preceding result to develop the probability distribution of the maxima of the random field on a region with area a_0. Let $N_b(a_0)$ be the number of local maxima above the level b in an area of size a_0. As b increases, $N_b(a_0)$ tends towards a Poisson probability mass function with rate $\mu_b \cdot a_0$:

$$\Pr[N_b(a_0) = k] = \frac{(\mu_b a_0)^k}{k!}e^{-\mu_b a_0},\ \mathrm{k} = 0,1,2,\ldots,\mathrm{b} \to \infty \qquad \text{(A.9)}$$

Note the equivalence between the following two events:

The maximum of the random field is less than or equal to b on an area size a_0

$\leftrightarrow$*There are 0 local maxima above b on an area size a_0.*

Therefore we have the following relationship:

$$\Pr[\underset{a_0}{Max}\{X(s_1,s_2)\} \le b] = \Pr[N_b(a_0) = 0] = e^{-\mu_b a_0},\ b \to \infty$$

Appendix B Spatial Covariance Functions and Mean-Square Differentiability

Here we look at the mean-square differentiability of random fields using four spatial covariance functions commonly found in the literature.

Recall from the corollary in Appendix A (section 2), that for an isotropic random field to be mean-square differentiable we need $dC(\tau)/d\tau^2\big|_{\tau=0}$ to exist and be finite. Also, recall the comment to the corollary that stated that for an isotropic random field, the covariance function must have a maximum at $\tau = 0$ (thus, the first order derivatives must be zero at $\tau = 0$) in order for the second order derivatives to even exist.

Spherical Covariance

$$C_{Sph}(\tau) = \begin{cases} 1 - \frac{3}{2}\frac{\tau}{\alpha_1} + \frac{1}{2}\left(\frac{\tau}{\alpha_1}\right)^3 & \text{if } 0 \le \tau \le \alpha_1 \\ 0 & \text{if } \tau > \alpha_1 \end{cases}$$

The first derivative at $\tau = 0$ is:

$$C'_{Sph}(0) = \begin{cases} 1 - \frac{3}{2\alpha_1} + \frac{3}{2}\left(\frac{\tau}{\alpha_1}\right)^2\Big|_{\tau=0} & \text{if } 0 \le \tau \le \alpha_1 \\ 0 & \text{if } \tau > \alpha_1 \end{cases}$$

$$= \begin{cases} 1 - \frac{3}{2\alpha_1} & \text{if } 0 \le \tau \le \alpha_1 \\ 0 & \text{if } \tau > \alpha_1 \end{cases}$$

Thus, we see that only when $\tau > \alpha_1$ (i.e. when the fields are uncorrelated) the second derivative exists. In our work, this represents a somewhat pathological case

because we want to study the network data under spatially correlated data. Thus, the spherical covariance function will not be suitable in our model.

Power-Exponential Covariance

$$C_{PExp}(\tau) = e^{(-(\tau/\alpha_1)^{\alpha_2})} \text{ for } \alpha_1 > 0,\ 0 < \alpha_2 \le 2$$

We have that :

$$C'_{PExp}(0) = -\frac{\alpha_2}{\alpha_1}\tau^{\alpha_2 - 1}e^{(-(\tau/\alpha_1)^{\alpha_2})}\Big|_{\tau=0}$$
$$= \begin{cases} -\infty & \text{for } 0 < \alpha_2 < 1 \\ -1 & \text{for } \alpha_2 = 1 \\ 0 & \text{for } 1 < \alpha_2 \le 2 \end{cases}$$

Because isotropic random fields must have a maximum (i.e. first derivative equal to zero) at $\tau = 0$, the only potential for mean-square differentiability is for $1 < \alpha_2 \le 2$. We now examine the second derivatives here:

$$C_{PExp}''(\tau) = \frac{\alpha_2}{\alpha_1}\tau^{\alpha_2 - 1}(1 - \alpha_2 + \frac{\alpha_2}{\alpha_1}\tau^{\alpha_2})e^{(-(\tau/\alpha_1)^{\alpha_2})}$$

We have to be a little careful evaluating this expression at $\tau = 0$. We use the limit as $\tau \to 0$. Note that $\lim_{\tau \to 0}(1 - \alpha_2 + \frac{\alpha_2}{\alpha_1}\tau^{\alpha_2}) = (1 - \alpha_2)$. Then we see:

$$C_{PExp}''(0) = \begin{cases} -\infty & \text{for } 1 < \alpha_2 < 2 \\ -2 & \text{for } \alpha_2 = 2 \end{cases}$$

Thus, we conclude that an isotropic random field is mean-square differentiable with the power-exponential covariance function for $\alpha_1 > 0$ and $\alpha_2 = 2$.

Rational Quadratic Covariance

$$C_{RQ}(\tau)=\left[1+\left(\frac{\tau}{\alpha_1}\right)^2\right]^{-\alpha_2};\alpha_1>0,\alpha_2>0$$

For the first derivative we have:

$$C'_{RQ}(\tau)=-2\alpha_2\frac{\tau}{\alpha_1^{\,2}}\left[1+\left(\frac{\tau}{\alpha_1}\right)^2\right]^{-\alpha_2-1}$$

So that

$$C'_{RQ}(0)=0 \text{ for all } \alpha_1>0,\alpha_2>0$$

Thus, the second order derivative exists for all $\alpha_1>0,\alpha_2>0$. Simple, but tedious calculus gives that:

$$C''_{RQ}(0)=-2\frac{\alpha_2}{\alpha_1^{\,2}}$$

So, the second derivative exists and is finite at zero separation distance. Thus, an isotropic random field is mean-square differentiable if it has a rational quadratic covariance function.

Matérn Covariance

$$C_M(\tau)=\frac{1}{2^{\alpha_2-1}\Gamma(\alpha_2)}\left(\frac{\tau}{\alpha_1}\right)^{\alpha_2}K_{\alpha_2}\left(\frac{\tau}{\alpha_1}\right);\ \alpha_1>0,\alpha_2>0$$

Where $K_{\alpha_2}(\cdot)$ is the modified Bessel function of second kind and order α_2.

For the first order derivative we have:

$$C'_M(0) = \left[\frac{(\tau/\alpha_1)^{\alpha_2}\alpha_2 K_{\alpha_2}(\tau/\alpha_1)}{2^{\alpha_2-1}\Gamma(\alpha_2)\tau} + \frac{(\tau/\alpha_1)^{\alpha_2}\left(\frac{\alpha_2\alpha_1 K_{\alpha_2}(\tau/\alpha_1)}{\tau} - K_{\alpha_2-1}(\alpha_1/\tau)\right)}{2^{\alpha_2-1}\Gamma(\alpha_2)\alpha} \right]_{\tau=0}$$

$$= \infty$$

Since the first order derivative is not zero at zero distance, the second order derivatives do not exist. Thus, an isotropic random field with the Matérn covariance function is not mean-square differentiable and will not be suitable for our model.

So, we have three possible spatial covariance models (under the appropriate conditions on α_1 and α_2) that will produce mean-square differentiable isotropic random fields and be compatible with our model and analysis. The choice of model is of course dependent on the inherent correlation of physical phenomenon being observed. For the sake of brevity in this work, the rational quadratic covariance model has been used exclusively in our simulations. In Figure A.1, we show some sample covariance functions of the rational quadratic kind for three levels of spatial correlation. We fix the roughness parameter $\alpha_2 = 1$ and vary α_1. Higher values of α_1 correspond to higher spatial correlation.

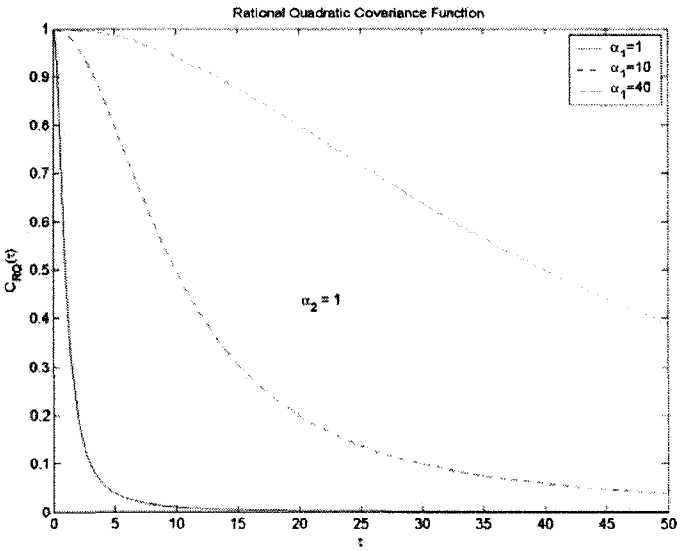

Figure A.1: Three plots of the rational quadratic covariance function versus distance for low, medium, and high levels of correlation.

REFERENCES

[1] H. Karl, A. Willig, "Protocols and Architectures for Wireless Sensor Networks", Wiley, 2005.

[2] A. Mainwaring, J. Polastre, et al, "Wireless Sensor Networks for Habitat Monitoring", from *WSNA 02* (Atlanta, Georgia), pp. 88-97, 2002.

[3] J. Chiles, P. Delfiner, "Geostatistics: Modelling Spatial Uncertainty", Wiley, 1999.

[4] H.Wackernagel, "Multivariate Geostatistics: An Introduction with Applications", Springer, 1998.

[5] G. Werner-Allen, K. Lorincz, et al , " Deploying a Wireless Sensor Network on an Active Volcanoe" in *IEEE Internet Computing*, Vol.10, issue 2, pp. 18-25, March-April 2006.

[6] M. Youssef, A. Yousif, N. El-Sheimy, "A Novel Earthquake Warning System Based on Virtual MIMO-Wireless Sensor Networks", *Canadian Conference on Electrical and Computer Engineering (CCECE 2007),* pp. 932-935, April 2007.

[7] E. Duarte-Melo, M. Liu, "Analysis of Energy Consumption and Lifetime of Heterogeneous Wireless Sensor Networks", *Global Telecommunications Conference 2002 (GLOBECOM 02)*, vol. 1, pp. 17-21, 2002

[8] I. Akyildiz, W. Su, et al , "A Survey on Sensor Networks", in *IEEE Communications Magazine*, vol. 40, issue 8, pp. 102 – 114, August 2002.

[9] S. Yoon, C. Shahabi, "The Clustered Aggregation Technique for Leveraging Spatial and Temporal Correlations in Wireless Sensor Networks", in *ACM Transactions of Sensor Networks*, vol. 3, issue 1, March 2007.

[10] D. Doolin, N. Sitar, "Wireless Sensors for Wildfire Monitoring", in *Proceedings of SPIE Symposium on Smart Structures & Materials*, vol 5765, pp. 477-484, March 2005 .

[11] News item at Crossbow Corporation's blog: http://blog.xbow.com/xblog/2007/09/researchers-cre.html

[12] I. McCauley, B. Matthews, et al . (from the Integrated Smart Sensing Group) http://dpi.projectforum.com/isss/11 .

[13] News item at Crossbow Corporation's blog: http://blog.xbow.com/xblog/2007/10/worlds-largest-.html

[14] E. Skafidas, "Precision Agriculture Information Networks: Nictor Sensor Network Platform" through NICTA (Australia) http://www.cse.unsw.edu.au/~sensar/hardware/pictures/Nictor.pdf

[15] News Item in *The Age*: "Irrigators Sense Its Time To Ignore Clock" http://www.theage.com.au/news/business/irrigators-sense-its-time-to-ignore-clock/2006/10/15/1160850812657.html

[16] News item in Emerging Technologies For Virtual Instrumentation blog: http://emertech.blogspot.com/2006/11/david-carey-uses-wireless-sensor.html

[17] M. Vuran, O. Akan, I. Akyildiz, "Spatio-Temporal Correlation: Theory and Applications for Wireless Sensor Networks", in *Computer Networks J (Elsevier),* vol. 45, no. 3, pp. 245-259, June 2004.

[18] M. Vuran, I. Akyildiz, "Spatial Correlation-Based Collaborative Medium Access Control in Wireless Sensor Networks", in *IEE?ACM Transactions of Networking* , vol. 14, no. 2, April 2006.

[19] X. Meng, T. Nandagopal, et al, "Contour Maps: Monitoring and Diagnosis in Sensor Networks", in *Computer Networks: The International Journal of Computer and Telecommunication Networking (ACM)*, vol. 50, issue 15, pp. 2820-2838, October 2006.

[20] A. Jindal, K. Psounis, "Modelling Spatially-Correlated Data of Sensor Networks", in *ACM Transactions of Sensor Networks"*, vol. 2, issue 4, pp. 466-499, November 2006.

[21] A. Jindal, K. Psounis, "Modelling Spatially-Correlated Data of Sensor Networks with Irregular Topologies", in *2nd Annual IEEE Comm. Society Conference on Sensor and Ad Hoc Communications and Networks (SECON),* pp. 305-316, September 2005.

[22] S. Roundy, D. Steingart, L. Frechette, et. al , "Power Sources for Wireless Sensor Networks", in *Proceedings of 1st European Workshop on Wireless Sensor Networks (EWSN)*, vol. 2920, pp. 1-17, Springer, Berlin, January 2004.

[23] C.S. Raghavendra, S. Singh "PAMAS: Power Aware Multi-Access Protocol with Signalling for Ad-Hoc Networks", *ACM Computer Communication Review,* vol. 27, pp. 5-26, 1998.

[24] W. Ye, J. Heidemann, D. Estrin, "Medium Access Control with Coordinated, Adaptive Sleeping for Wireless Sensor Networks", in *IEEE/ACM Transactions on Networking,* vol. 12, issue 2, pp. 493-506, June 2004.

[25] E. Vanmarcke, "Random Fields: Analysis and Synthesis", MIT Press, 1983.

[26] R. Adler, "Random Fields", *Article entry in Wiley Encyclopedia of Statistical Sciences,* 2006. (Online at: http://www.mrw.interscience.wiley.com/emrw/9780471667193/ess/article/ess2164/current/html).

[27] P. Abrahamsen, "A Review of Gaussian Random Fields and Correlation Functions", 2nd Edition, Norwegian Computing Center, April 1997. (Online at: www.math.ntnu.no/~omre/TMA4250/V2007/abrahamsen2.ps).

[28] J.O. Berger, V. de Oliviera, B. Sanso, "Objective Bayesian Analysis of Spatially Correlated Data", *Journal of American Statistical Association*, vol. 96, no. 456, pp. 1361-1374, 2001.

[29] W. Feller, "An Introduction to Probability Theory and its Applications", vol. 2, John Wiley and Sons, 1966.

[30] M.K. Mehmet Ali, H. Gu, "Performance Analysis of a Wireless Sensor Network", *Proceedings of WCNC 2006,* vol. 2, pp. 1166-1171, 2006.

[31] W. Heinzelman, J. Kulik, H. Balakrishnan, "An Application-Specific Protocol Architecture for Wireless Microsensor Networks", *IEEE Transactions on Wireless Networking*, vol. 1, no. 4, pp. 660-670, 2002.

[32] W. Heinzelman, A. Chandrakasan, H. Balakrishnan, "Adaptive Protocols for Information Dissemination in Wireless Sensor Networks", *Proceedings of the 5th Annual Conference on Mobile Computing and Networking*, pp 174 – 185, Seattle, WA, August 1999.

[33] C.F. Chiasserini, R. Gaeta, M. Garreto, et. al, "Fluid Models for Large-Scale Wireless Sensor Networks", *Performance Evaluation,* vol. 64, pp. 715-736, 2007.

[34] A. Leon-Garcia, I. Widjaja, "Communication Networks: Fundamental Concepts and Key Architectures", 2nd Edition, McGraw-Hill, 2004

[35] Cao, J., "The Excursion Set of Random Fields with Applications to Human brain Mapping", Ph.D. Thesis, Dept. Mathematics and Statistics, McGill University, 1997. Available at http://cm.bell-labs.com/stat/cao.