THE UNIVERSITY OF MELBOURNE

Performance Evaluation and Enhancement of Mobile and Sensor Networks

by

Malka Nishanthi Halgamuge

BSc(Eng), MSc(Eng)

Submitted in total fulfilment of

the requirements for the degree of

Doctor of Philosophy

Department of Electrical and Electronic Engineering

The University of Melbourne

Australia

2006

Produced on archival quality paper

The University of Melbourne
Australia

Abstract

Performance Evaluation and Enhancement of Mobile and Sensor Networks

by Malka Nishanthi Halgamuge
BSc(Eng), MSc(Eng)

This thesis addresses the performance evaluation and enhancement of wireless networks. Part I investigates the problem of resource allocation in cellular networks, focusing on handoff, and Part II investigates resource allocation in sensor networks focusing on power management.

In Part I, a new framework is proposed for performance evaluation and comparison between existing handoff algorithms. This framework is then extended using call dropping, handoff cost and both non-retrial and retrial options. An off-line cluster-based computationally simple heuristic algorithm is proposed to find a near optimal handoff sequence to be used as a benchmark. The benchmark reveals that existing handoff methods have room for further improvement. When the users travel on predefined paths such as roads, the proposed benchmark can be used (by pattern recognition methods such as template matching techniques) to update the possible future handoff sequence prior to each handoff.

Existing handoff algorithms are compared to identify the trade-off between signal quality and number of handoffs. This approach and the proposed cluster-based near optimal benchmark provide a realistic framework for evaluating of handoff methods in cellular networks that has not been achieved before. A method of estimating handoff cost and optimal values for retrial and non-retrial models is also investigated.

Part II investigates power management, a key area of resource allocation in

wireless sensor networks. In sensor networks using cluster based communication protocols, power management is integrated with the problem of transmitting information from one sensor to another. With no specific network infrastructure as in cellular networks, power management is crucial for survival of the sensor network. A comprehensive energy model for a wireless sensor network is proposed in Part II by considering seven key energy consumption sources. The current energy consumption models ignore many of these important sources of energy drainage. The benefit of the proposed comprehensive model that allows realistic estimation of a sensor node's lifetime is shown by comparison with other existing energy models. This work provides guidelines for efficient and reliable sensor network design that can be used to optimize energy efficiency subject to required specifications. Further, this model is applied to Low-Energy Adaptive Clustering Hierarchy (LEACH) type protocol to obtain an accurate evaluation of energy consumption and node lifetime.

Extension of the lifetime for a sensor network is important for most if not for all applications. The work therefore investigates efficient battery management for sensor networks. This study develops a method for extending the sensor network lifetime by using high capacity batteries for cluster heads. It is shown that according to different types of application, the ratio of initial battery capacities for sensors and cluster heads may vary. It is also shown that the proposed method can be used in conjunction with LEACH to increase the overall efficiency. Moreover, it is observed that parameters such as battery cost and sensor deployment cost will make a huge impact on the total network cost.

Declaration

This is to certify that

(i) the thesis comprises only my original work,

(ii) due acknowledgment has been made in the text to all other material used,

(iii) the thesis is less than 100,000 words in length, exclusive of table, maps, bibliographies, appendices and footnotes.

Signature_____

Date_____

This dissertation is dedicated to my mother
who shows me the way every day,
even today.

Acknowledgments

I would like to express my sincere gratitude to my supervisors, Prof. Moshe Zukerman, Prof. Rao Kotagiri and Dr. Hai Vu, who made this work possible by giving invaluable guidance, excellent feedback and moral support.

I would like to thank the Australian Government, for supporting my study with an Australian Postgraduate Award (APA). I also would like to thank the ARC Special Research Center for Ultra Broadband Information Networks (CUBIN), for providing a supportive academic environment to do my research and an opportunity to meet wonderful and talented people from all over the world who helped me to expand my knowledge.

I would especially like to thank all my friends in CUBIN, especially Thaya for being an "unassigned mentor". I truly appreciate their friendship to me throughout my PhD candidature, the past three and half years.

I am extremely grateful to my husband Saman, for all his patience when I was frustrated and stressed at times, and his understanding of all the difficulties that I encountered. My sincere gratitude goes to my lovely little son Nelanka and daughter Chenuli for being my stress relievers by bringing a smile to my face and love to my heart. Without the patience of my family, to whom I owe everything, it would have been impossible to accomplish this.

Finally I am forever indebted to my parents for their effort to give me the best possible education. They taught me the value of knowledge, the joy of love and the importance of family. They have stood behind me in everything I have done, providing constant support, continuous encouragement and true love. I would especially like to thank my mother whose continuous blessing and sacrifices for everything in my life, have brought me what I have.

Contents

List of Figures

List of Tables

Acronyms

ACK	Acknowledgment
ADC	Analog-to-Digital Controller
ARSS	Average Received Signal Strength
BALI	Bay Area Location Information
BHS	Best Handoff Sequence
BS	Base Station
BSC	Base Station Controller
CBN	Chemical Biological and Nuclear
CDMA	Code Division Multiple Access
CIR	Carrier to Interference Ratio
CQSL	Call Quality Signal Level
CSMA/CA	Carrier Sense Multiple Access/Collision Avoidance
CTS	Clear-to-send
ECQSL	Extended Call Quality Signal Level
EICQSL	Extended Improved Call Quality Signal Level
FDMA	Frequency Division Multiple Access
FHA	Fuzzy Handoff Algorithm
GA	Genetic Algorithm
GPS	Global Positioning System
GSM	Global System for Mobile-communication
HLR	Home Location Register
HMM	Hidden Markov Models
HPCH	High Powered Cluster Heads
ICQSL	Improved Call Quality Signal Level

IEEE	Institute of Electrical and Electronics Engineers
LEACH	Low-Energy Adaptive Clustering Hierarchy
MAC	Medium Access Control
MAHO	Mobile Assisted Handoff
MCHO	Mobile Controlled Handoff
MCU	Micro-controller Unit
MS	Mobile Station
MSC	Mobile Switching Center
NASP	Number of Acceptable Sample Points
NCHO	Network Controlled Handoff
PCS	Personal Communications Services
POMDP	Partially Observable Markov Decision Process
QoS	Quality of Service
RF	Radio Frequency
RSSI	Received Signal Strength Indication
RTS	Request-to-send
TDMA	Time Division Multiple Access
VLR	Visitor Location Register

CHAPTER ONE

Introduction

1.1 Looking Forward

Resource management and handoff are strongly interconnected issues, currently receiving considerable attention in wireless network research. Handoff, also known as handover or radio link transfer, is a process to switch an ongoing call from one base station to neighboring base station as a mobile user moves through the coverage area of a cellular system. Increased demand for mobile services led to the reduction in cell radii in congested areas. To provide the required quality of service (QoS) at an affordable cost in areas covered by micro cells, it is necessary to have efficient handoff algorithms. Having the correct criterion for evaluating handoff methods allows telecommunication providers choose the right handoff algorithm in a cost effective way. Such a selection should avoid wasting expensive radio transmission resources, as well as meeting the user QoS requirements.

Handoff in current wireless cellular systems is commonly achieved through hysteresis and threshold based methods. All such methods are centralized and managed by the base station controller.

An emerging area of wireless networks is adhoc networks, where network infrastructure is not available. Wireless sensor networks are a fast-growing area of adhoc networks that has attracted considerable attention recently. Recent successes in the miniaturization of sensors and sensor devices equipped with wireless interfaces contributed to the rapid advancement of this area. A sensor net-

work provides the much-needed interface between a base station (usually a computer) and the real world captured by itself.

The concept of handoff is only implicitly applicable to sensor networks, as sensors may communicate with the base station via other sensors with power management dependent communication protocol. Due to battery capacity limitations, power management has become a crucial issue in the design of sensor networks. The sensor network lifetime depends on power management.

1.2 Focus of the Thesis

This thesis examines the handoff problem of cellular networks in Part I and the power management problem of sensor networks in Part II. Both problems are key issues in resource allocation and in ensuring quality of service in the respective applications of wireless networks.

Handoff from one serving base station to another occurs when certain conditions on signal quality at serving base station are not met. Power management in sensor networks ensures that sensors last longer, and therefore the required quality of information collected from sensors is achieved.

Part I will show the inadequacy of existing criteria for comparing handoff methods considering many factors that influence quality, and develop a realistic and comprehensive criterion to do so.

Part II focuses on developing a comprehensive energy model for sensors and their use to enable new strategies for power management in sensor networks. The thesis provides guidelines for efficient and reliable sensor network design that can be used to optimize energy efficiency subject to required specifications. It proposes High Powered Cluster Heads to increase the sensor network lifetime. It also shows that the proposed method can be used in conjunc tion with LEACH to increase overall efficiency.

1.3 Organization of the Thesis

The thesis chapters are organized as follows:

Chapter 2: Introduction to Part I This chapter describes the key issues relating to existing handoff strategies, and provides the motivation for Part I.

Chapter 3: Optimal Handoff Sequence Firstly, this chapter proposes a new framework, based on signal quality, for performance evaluation and comparison of existing handoff algorithms. An off-line, cluster-based, computationally simple heuristic algorithm is proposed to find a near optimal handoff sequence as a benchmark. The existing handoff algorithms are then compared to identify the trade-off between signal quality and number of handoffs. The proposed method of finding a near optimal handoff sequence can be used for pattern recognition based handoffs if the user paths are known.

Chapter 4: Handoff Performance Evaluation Performance evaluation and comparison between existing handoff algorithms is conducted using a more realistic framework based on signal level, call dropping and handoff cost proposed in this work. In our evaluation, both the retrial option, where repeated call attempts are made, and the non-retrial option are considered. The results suggest two different handoff methods for urban areas with high dropping probability and suburban areas with low dropping probability. The proposed evaluation models indicate that existing handoff methods can improved. Furthermore, methods of estimating handoff cost and optimal values for retrial and non-retrial models are also investigated.

Chapter 5: Conclusions for Part I This chapter summarizes the key contributions of Part I.

Chapter 6: Introduction to Part II Part II investigates power management, an essential component of resource allocation, in sensor networks.

Chapter 7: Energy Consumption in Sensor Networks A comprehensive energy model for wireless sensor networks is proposed by considering seven sources of key energy consumption. The current energy consumption models ignore many of these important sources of energy drainage. Using the proposed model the lifetime of a sensor node is estimated. This chapter provides guidelines for efficient and reliable sensor network design that can be used to optimize energy efficiency subject to required specifications.

Chapter 8: Efficient Battery Management for Sensor Lifetime Batteries, as the widely used power provider of the sensors in the network, are considered the key factor for achieving a prolonged life. The work therefore investigates efficient battery management for sensor networks. This chapter proposes High Powered Cluster Heads to increase the sensor network lifetime. It is shown that the ratio of initial battery capacities for sensors and cluster heads changes, according to different types of application. It is also shown that the proposed method can be used in conjunction with LEACH to increase overall efficiency. Moreover, it is observed that parameters such as battery cost and deployment cost make a huge impact on total network cost.

Chapter 9: Conclusions for Part II This chapter summarizes Part II, emphasizing its key contributions.

Chapter 10: Future Research This chapter provides possible future research direction for both parts of the thesis. It includes the proposal of a method further highlighted in Appendix B. It exploits the analogy between the sensor scheduling problem to the base station assignment and the handoff problem in cellular networks. The mobile user is modeled using a Hidden Markov Model (HMM) and the handoff problem is formulated as an optimization problem of base station scheduling that minimizes cost functions containing the HMM state estimation error and base station measurement costs.

1.4 Contribution of the Thesis

Major contributions of the thesis are listed in order of appearance with the appropriate section number and citation. Some of these contributions led to publications or submitted manuscripts in journals.

1. Development of an off-line cluster-based computationally-simple heuristic handoff algorithm to find a near optimal handoff sequence as a benchmark for comparison of handoffs. (Section 3.4, 3.5 [68]).

2. Investigation of the impact of handoff users in a real system, by analyzing real data extracted from about 50,000 users (Section 3.8.4 [67]).

3. Investigation of the applicability of the developed offline handoff algorithm to pattern recognition based handoffs (Section 3.8.5 [67]).

4. Development of comprehensive handoff performance evaluation models by considering both retrial (where repeated call attempts are made after a call is lost) and non-retrial call options. (Section 4.3 [69]).

5. Investigation of the suitability of different handoff algorithms for different terrain configurations (Section 4.5 [69]).

6. Estimation of handoff cost and optimal handoff sequences for retrial and non-retrial models (Section 4.5 [69]).

7. Development of a comprehensive energy model for wireless sensor networks, by considering seven key energy consumption sources (Section 7.2 [70]).

8. Development of guidelines for efficient and reliable sensor network design, to optimize energy efficiency subject to predefined specifications (Section 7.6 [70]).

9. Estimation of the optimal number of clusters which is very sensitive to the energy model used (Section 7.4 [70]).

10. Development of the high powered cluster head concept to extend sensor network lifetime (Section 8.5.1 [66]).

11. Analysis of efficient scheduling and budgeting methods of battery power (Section 8.3, 8.4).

12. Exploitation of the analogy of the sensor scheduling problem to the base station assignment or the handoff problem in cellular networks and the use of Hidden Markov Model to describe the handoff problem. (Section B.7, [51]).

13. Formulation of the optimisation problem of base station scheduling that minimises cost functions containing the HMM state estimation error and base station measurement costs (Section B.8, [51]).

1.5 Publications by the Author Related to this Thesis

1.5.1 Journals

1. Malka N. Halgamuge, Hai L. Vu, Rao Kotagiri and Moshe Zukerman, "Signal-Based Evaluation of Handoff Algorithms", IEEE Communications Letters, pp 790-792, Volume 9, Issue 9, Sept. 2005 [68].

2. Malka N. Halgamuge, Moshe Zukerman, Rao Kotagiri, Hai L. Vu, "Sensor Energy Consumption", ACM Trans. on Sensor Networks (submitted) [70].

3. Malka N. Halgamuge, Hai L. Vu, Rao Kotagiri and Moshe Zukerman "A Call Quality Performance Measure for Handoff Algorithms in Intelligent Transportation Systems" IEEE Trans. Mobile Computing (submitted) [69].

4. Malka N. Halgamuge et al. "Base Station Scheduling for Cellular Networks with Hidden Markov Model", (to be submitted) [51].

1.5.2 International Conferences

1. Malka N. Halgamuge, "Efficient Battery Management for Sensor Lifetime", Proceedings of the 21st IEEE International Conference on Advanced Information Networking and Applications, AINA'07, Niagara Falls, Canada, 21-23 May 2007 (accepted) [62].

2. Malka N. Halgamuge, Rao Kotagiri and Moshe Zukerman, "High Powered Cluster Heads for Extending Sensor Network Lifetime", Proceedings of the 6th IEEE International Symposium on Signal Processing and Information Technology, ISSPIT06, Vancouver, Canada, 27-30 Aug. 2006 [66].

3. Malka N. Halgamuge, Rao Kotagiri, Hai L. Vu and Moshe Zukerman, "Evaluation of Handoff Algorithms Using a Call Quality Measure with Signal Based Penalties", Proceedings of IEEE Wireless Communications and Networking Conference, WCNC'06, Las Vegas, Nevada, USA, 3-6 Apr. 2006 [65].

4. Malka N. Halgamuge, Hai L. Vu, Rao Kotagiri and Moshe Zukerman, "An Ideal Base Station Sequence for Pattern Recognition Based Handoff in Cellular Networks", Proceedings of the International Conference on Information and Automation, ICIA05, pp 19-24, BMICH, Colombo, 15-18 Dec. 2005 [67].

1.6 Other Publications by the Author

1. Malka N. Halgamuge, S. M. Guru and Andrew Jennings, "Centralised Strategies for Cluster Formation in Sensor Networks", in Classification and Clustering for Knowledge Discovery, ISBN: 3-540-26073-0, Chapter 20, Springer-Verlag, pp. 315-334, Aug. 2005 [64].

2. Malka N. Halgamuge, S. M. Guru and Andrew Jennings, "Energy Efficient Cluster Formation in Wireless Sensor Networks", Proceedings of IEEE International Conference on Telecommunication, ICT'03, Volume 2, Papeete, Tahity, French Polynesia, 23 Feb.-1 March 2003 [63].

3. K. Chan, G. Lam, S.M. Guru, Malka N. Halgamuge and S. Fernando, "Development of Palm-SmartBolt Sensor System Interface", Proceedings of International Conference on Fuzzy Systems and Knowledge Discovery, FSKD'02, Singapore, Nov. 2002 [31].

4. C. Brewster, P. Farmer, J. Manners and Malka N. Halgamuge, "An Incremental Genatic Algorithm for Model Based Clustering and Segmentation", Proceedings of International Conference on Fuzzy Systems and Knowledge Discovery, FSKD'02, Singapore, Nov. 2002 [24].

Part I

Handoff Performance in Cellular Networks

CHAPTER TWO

Introduction to Part I

2.1 Problem Statement

Unlike wired communication systems, wireless communication systems require the transmission of radio waves in free space. Since radio waves attenuation increases with distance, frequencies in wireless communication systems can be reused. Such reuse is the key feature in the cellular concept. The cellular concept is supported by the cellular infrastructure that includes base stations responsible for maintaining communication links to and from cellular users. Wireless communication is also possible when no fixed infrastructure is available inside a given geographical area. This self-organizing type of wireless communication is known as adhoc communication, and is often used in sensor networks.

Part I of the thesis discusses the handoff, an important component of resource allocation of cellular networks. As there are several known handoff methods, a comprehensive framework for evaluation of these handoff methods is presented in Chapter 3. This framework is extended in Chapter 4 by considering various practical scenarios.

The rest of the present chapter provides an introduction to Part I. Section 2.2 describes the cellular concept and Section 2.3 addresses the characteristics of radio propagation. Section 2.4 describes the signal propagation model which is used in Chapters 3 and 4. Section 2.5 describe and the cellular hierarchy and Section 2.6 explains the GSM standard. Section 2.8 discusses the requirements of handoff and Section 2.9 explains how power control effect to handoff. Section 2.10

outlines different classifications of handoff methods and explains their importance. Section 2.7 presents the existing handoff methods. Section 2.12 describes a problem in handoff and how it occurs. Section 2.13 describes the relationship to resource allocation and Section 2.14 analyzes the patterns of movement of mobile users. Section 2.15 describes motivation for handoff evaluation.

To conclude this chapter, Section 2.16 summarizes the key issues involved and provides an outline for Part I.

2.2 Cellular Concept

The cellular concept in wireless communications is based on cells or smaller coverage areas, each served by a base station or a radio transmitter and a range of frequencies. The cells are intelligently assigned with radio frequency channels to allow the reuse of the spectrum without much interference. The set of adjacent cells that uses the entire allocated spectrum is called a cluster and the number of such cells is called the cluster size or the frequency reuse factor [144].

There are two types of major interferences in cellular systems. The co-channel interference is due to the use of the frequency reuse in cells of different clusters while the adjacent channel interference due to the different frequency channels within the same cluster.

2.3 Characteristics of Radio Propagation

This section provides a brief review of the major characteristics of radio propagation. This has been addressed in several studies [26, 38, 46, 111, 151, 194].

2.3.1 Why Radio Propagation?

The propagation of radio waves is strongly dependent on terrain and varies with factors that include the radio frequency, velocity of a mobile terminal and interference sources. It is important to model and accurately predict signal coverage and

interference levels, and evaluate performance parameters in comparing different signalling schemes and in finding optimum locations for base stations.

Cells can be categorized in terms of their sizes into femto-, pico-, micro-, macro and mega-cells. Different factors affect radio propagation for each cell category. Radio propagation differs from open areas to closed areas. In open areas with small distances, signal strength reduces proportionally to the square of the distance. If the distance is larger, signal strength reduces in proportionally to a larger rate of distance. When there is no line of sight between the transmitter and the receiver, signals may travel through different paths due to reflections from obstacles, and therefore, may have varied levels of strength reductions. In addition such signals may with various delays lead to the effect known as multipath delay spread [144].

2.3.2 Radio Propagation Mechanism

There is no simple model sufficient to capture all possible effects, as propagation effects are so diverse [130]. Transmission paths between the transmitter and the receiver may vary from a simple line of sight to a fully covered building. Fluctuation of the received signal will be influenced by the mobility of a receiver and/or a transmitter. Such fluctuations in received signal strength can occur in three ways [144]: reflection, diffraction and scattering.

1. **Reflection**

   When radio waves hit obstacles with large dimensions relative to the wave length, reflection can occur, leading to attenuation of rays. This attenuation is determined by factors such as radio frequency, the angle of hit and the material properties and thickness of the surface. Reflection may result from the earth's surface, buildings and walls [144].

2. **Diffraction**

   When the radio path between the transmitter and receiver has obstacles with edges, diffraction can occur. Due to bending of radio waves that can

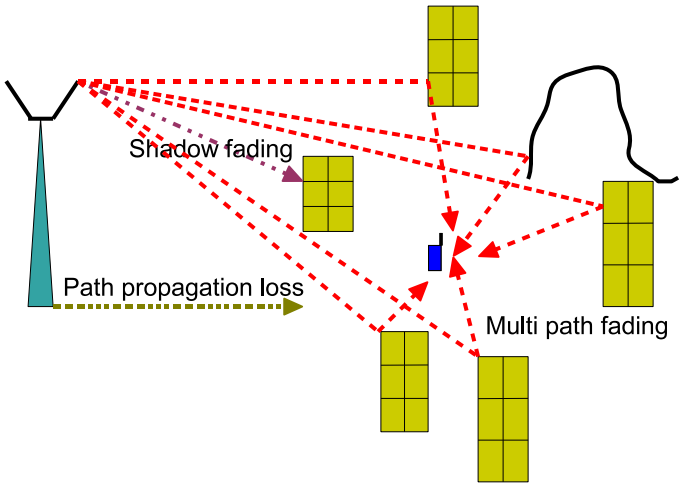


Figure 2.1: Radio propagation: path loss propagation, shadow fading, multi path fading

occur around the obstacle, the resulting secondary waves can be present every where including behind the obstacle where the line-of-sight does not exist. Diffracted rays act as a secondary source generated by edges of buildings or large objects in the propagation path.

The diffracted fields propagate away from the edge as cylindrical waves. Diffraction at high frequencies depends on the geometry of the object, amplitude and phase, among other factors.

3. **Scattering**

Scattering of rays occurs when they hit objects with dimensions less than the wave length, for example vehicles, street signs, lamp posts and indoor furniture, where the number of obstacles per unit volume is large. They scatter in the form of spherical waves in all directions.

A mobile user's received signal strength from the base station includes three components: path loss, shadow fading and multipath fading as shown in Fig. 2.1.

2.3.3 Path Loss and Attenuation

In free space, the field strength of radio signal reduces in proportion to the distance squared. The power or signal strength received at a distance d, is given by the Friss free space equation [144],

$$P_r(d) = \frac{P_t G_t G_r \lambda^2}{(4\pi)^2 r^2 L},$$

where $P_r(d)$ is the received power, P_t is the transmitted power, G_t is the transmitter antenna gain, G_r is the receiver antenna gain, r is the distance between user and the base station, L is the system loss factor not related to propagation and λ is the wavelength. Therefore, free space radio frequency propagation can be used to model point to point communications. However, this is less useful in cellular environments where point to point communication without obstacles is not always the case.

There are many different models (large scale propagation models or small scale fading models) which can be used to model signal fluctuations of received signal strength in the literature [144, 169].

2.3.4 Shadowing or Slow Fading

Shadowing or slow fading is caused by obstacles in the propagation path or line-of-sight between transmitter and receiver both outdoors and indoors [183]. They cause attenuation of waves passing through them. If the obstacle is very large (for example a building), and has structure and material that causes strong attenuation, shadowing can be extensive and depends on the radio frequency used. In such environments reception occur by non line-of-sight communication.

Diffractions around the edges of an obstacle are another cause of shadowing. Signals at radio frequencies bend around the edges of obstacles. A received signal in such an environment can have three components: line-of-sight transmitted, reflected and diffracted. When shadowing exists due to these losses, a small change of distance between transmitter and receiver may not result in variation

of the shadowing. Therefore, it is also known as slow fading.

2.3.5 Fast Fading

Fast fading is caused by multipath propagation [59]. Fast fading channels have their impulse responses changing quickly. The radio frequency signal from the transmitter may be reflected from objects such as buildings, walls and mountains. This creates multiple transmission paths between a transmitter and a receiver.

2.4 Signal Propagation Model

Radius of typical cell can various from few meters to few kilometers. As the number of users per cell is limited (due to limited bandwidth), the reduction in cell radius is a method of increasing the number of users served. A common application of this can be a busy city with a large number of mobile users per square meter. However, smaller cells will lead to more handoffs. An area with many tall buildings may need several micro cells to cover it. Considerable losses in radio energy is expected around corners or at intersections. Propagation losses in such environments were studied in [26, 38, 46, 49, 57, 111, 151, 182, 189, 195].

We consider the following log normal propagation model [60] to generate signal strengths for our work in Chapters 3 and 4.

2.4.1 Signal Strength Measurements

In most mobile communication systems, signal strength measurements are performed at regular intervals. The received signal strength with Gaussian distribution, $S(r, \rho)$, is given by [60]

$$S(r, \rho) = K_1 - K_2 log(r) + \rho, \tag{2.1}$$

where K_1 depends on transmitted power in the base station, K_2 corresponds to direct line-of-sight propagation (value ranges from 20 to 60) [60], r is the distance

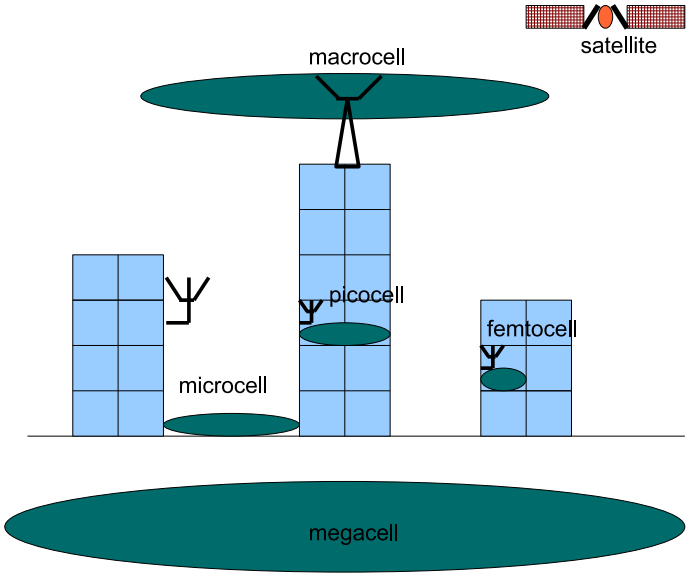


Figure 2.2: Femto-, pico-, micro-, macro- and mega-cells in the cellular hierarchy

in meters from the user to base station and ρ is the random variable with a mean zero for log-normal fading with σ^2 variance.

We use a simple decreasing correlation function to model co-relation properties as described in [60] . Correlation function of $S(r)$ is given by,

$$\begin{aligned} R_S(k) &= E[S(r), S(r+k)] \\ &= \sigma^2 a^{|k|}, \end{aligned} \tag{2.2}$$

where a is the correlation coefficient and σ^2 is the variance. Generally, the variance varies between 3-10 dB. The correlation coefficient is given by

$$a = \epsilon_D \, ^{\frac{vT}{D}},$$

where ϵ_D is the correlation between two points separated by distance D, v is the mobile user's velocity and T is the sampling interval.

2.5 Cellular Hierachy

A hierarchical cellular infrastructure can be used to support cells of different sizes. It can extend the coverage to areas not covered by large cells, provide additional support for areas with a higher density of users, and support application specific small cell coverage such as connecting laptops and cellular phones. Generally, cells in a hierarchical cellular infrastructure are categorized as femtocells, picocells, microcells, macrocells and megacells. Obviously, the smallest, femtocells are used for connecting personal equipment such as laptops and the largest Megacells cover hundreds of kilometers usually through satellites as shown in Fig. 2.2. Picocells generally only cover a single floor or part of a floor inside a building. Microcells are for urban areas and macro-cells for suburban areas. We consider microcells for our work in Chapters 3 and 4 and Appendix B.

2.6 Global System of Mobile Communications (GSM)

The Global System of Mobile Communications (GSM) is the standard for the digital second generation (2G) pan-European cellular system. In addition to the air interface, GSM also includes the definition of various interfaces between hardware and software. It is an integrated voice-data service offering three types of services: tele-services, bearer services and supplementary services.

2.7 Handoff in Cellular Networks

Handoff is the transfer of an ongoing call from one cell to another as a user moves through the coverage area of a cellular system (Fig. 2.3). In wireless cellular systems the handoff process is expected to be successful and imperceptible to users. It is also expected that the need for handoff be infrequent. In congested inner city type environments with small cell sizes, it has become a challenging task to meet these requirements.

Generally handoff or handover in GSM networks is differentiated between in-

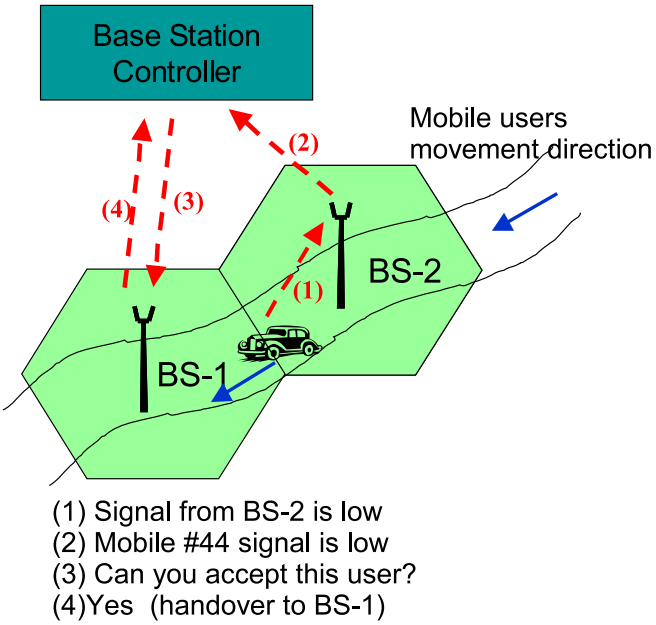


Figure 2.3: What is handoff? Transferring a radio link or switch an ongoing call from one base station to neighboring base station as a mobile user moves through the coverage area of a cellular system

ternal and external. Internal handoffs occur between base stations (BS) belonging to the same base station controller (BSC) while external handoff refers to handoff between BSCs belonging to the same mobile switching center (MSC).

Handoffs can happen between cells or within the cell (between channels). In Part I we are mainly concentrating on handoffs happening with the same network between different cells (or base stations). We are also not concerned with soft handoff [102, 134, 178, 193], where the old base station is released after a link with the new base station is established, as such handoff is mainly used with CDMA type systems.

2.8 Why Handoff?

Various network resources are needed for the handoff process, including air signaling, network signaling, database lookup and network configuration [191]. Air signaling occurs between the user and the base station while network signaling

is between the base station and other network entities like the mobile switching centers. Handoff signaling uses radio bandwidth whether it is using control channels or traffic channels. Database accesses for registration and authentication contributes to the handoff cost. The network reconfiguration is needed to provide new access users to the new base stations and terminate the user's access with the old base stations. Although in the literature handoff costs are modeled as a constant cost per handoff because of the difficulty in quantifying the cost, all the above mentioned factors are dependent on the system design and configuration, and therefore influence handoff cost.

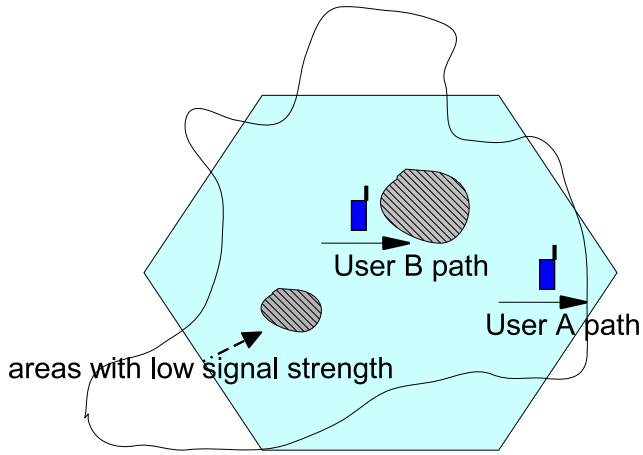


Figure 2.4: Signal strength holes that have low signal strength within a cell (adapted from [101])

A handoff is generally initiated by the signal strength deterioration, but can also be initiated by traffic balancing, where calls are moved between cells to ease traffic congestion.

Handoff is implemented on the voice channel. The importance of handoff varies with the size of the cell. Within a large cell, it is more likely that calls initiated in the cell also terminate within the cell and less likely that they are dropped as a result of reaching the cell boundary. However, a call may be terminated due to approaching a fringe area or a hole (gap) within the cell as shown in Fig. 2.4.

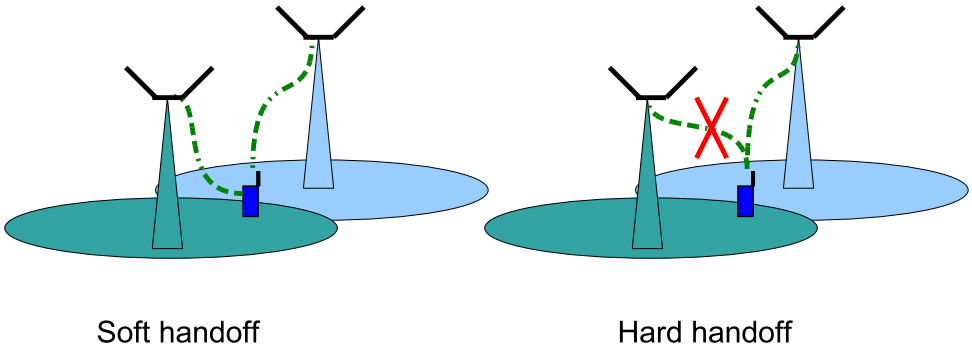


Figure 2.5: Soft and hard handoff in cellular networks

2.9 Power Control and Handoff

Power control plays an important role in reducing the number of handoffs. The base station requests the power of a mobile station to be increased or decreased depending on the strength of the received signal from the mobile station. If the user is far away from the base station or is affected by shadowing, the base station can request the mobile station to increase the emission power. Consequently, the need for handoff due to shadowing is avoided.

2.10 Classifying of Handoffs

There are various methods of classifying handoffs. Four such commonly used methods are summarized in the following subsections.

2.10.1 Classification 1

Mode of call transfer between base stations is used to classify handoffs:

1. **Soft Handoff**: Mobile users are connected to two or more base station simultaneously. They keep the connection with the old base station until a connection to the new base station is made (Fig. 2.5).

2. **Hard handoff**: A user will disconnect from the previous base station before connect to the new base station (Fig. 2.5).

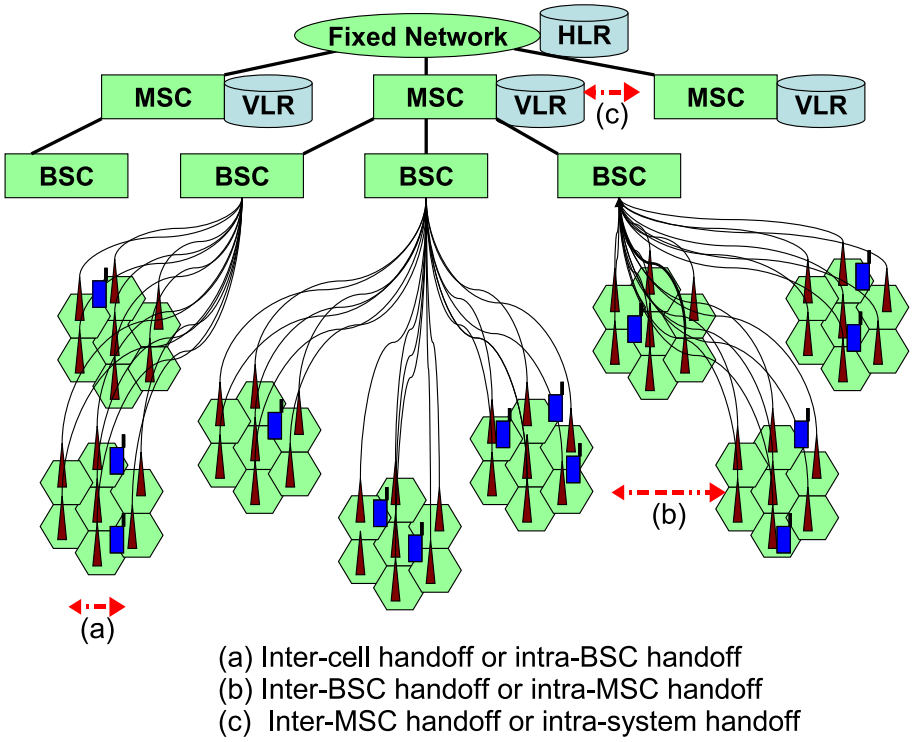


Figure 2.6: Architecture of the GSM mobile radio network

2.10.2 Classification 2

Handoff is transferring an ongoing call or radio link from an old base station to the new one. This transfer can occur between channels in the same base station or between base stations. The handoff is a expensive process that depends on network elements (BS, BSC and MSC) involved. The following steps involved in a handoff are described in Fig. 2.6.

Intra cell handoff can occur even in non cellular systems such as cordless telephones with available frequency channels [109]. Handoffs can also occur between sectors within the same cell. These handoffs are often categorized as intracell handoffs or intercell handoffs (considering sectors as cells):

1. **Intra-cell handoff**

   Handoff between two time slots or channels in the same base station.

2. **Inter-cell handoff or Intra-BSC handoff**

Handoff between two base stations connected to the same BSC.

3. **Inter-BSC handoff or Intra-MSC handoff**

   Handoff between two base stations connected to different BSCs belonging to the same MSC.

4. **Inter-MSC handoff or Intra-system handoff**

   Handoff between two base stations connected to different BSCs belonging to different MSCs.

5. **Inter-system handoff**

   Handoff between two base stations connected to different MSCs from two different PCS networks.

2.10.3 Classification 3

This classification is based on which network entities initiate/participate in the handoff [109].

1. **Network Controlled Handoff (NCHO)**

   The network decides and carries out the handoff. The base station measures the signal strength and quality from the mobile station (MS) and compares it with a predefined threshold. The network also requests the surrounding base stations to do the same and makes its decision based on the deterioration of signal strength and quality from MS. Due to the dependency on the network control, handoff time (the total time for handoff decision and execution) may be many seconds. This is also known as base station controlled handoff. The handoff time for NCHO can be up to ten seconds even more than that [109].

2. **Mobile Assisted Handoff (MAHO)**

   The main difference in this case is that the network requests the MS to monitor the signal strengths and quality from surrounding base stations and

report to the BS. This method is somewhat decentralized as both MS and BS are involved in the monitoring process and provide the network with information required to make the decision. The handoff time is less than that of NCHO and may be as low as one second.

3. **Mobile Controlled Handoff (MCHO)**

   The mobile station (MS) completely makes the decision for handoff process. It continuously measures the signal strengths and signal qualities from BSs and compares them. A handoff occurs when a predefined criterion is met. The handoff time is less than that of NCHO and in the order of 100 milliseconds [109].

2.10.4 Classification 4

As in [101], handoffs are categorized according to the following handoff criteria:

1. **Handoff based on signal strength**

   A criterion for handoff is the signal strength falling below a predefined threshold value. Implementation of this type of handoff is relatively simple as the received signal strength will be compared with the predefined threshold. However, interference is included in the received signal strength. Therefore, it is possible that the handoff does not occur where it should, due to high levels of interference contributing to the received signal strength. On the other hand, handoff may occur unnecessarily when interference is low.

2. **Handoff based on carrier-to-interference ratio (CIR)**

   An alternative method of using signal strength that contains the carrier and interference is by considering the ratio of received signal strength to interference. Often, this ratio is an approximation of the carrier to interference ratio assuming that interference is small in comparison to the carrier. If the carrier to interference ratio is lower than a certain threshold, handoff can occur, as it could be the result of low carrier or high interference. However, handoff may not occur when both carrier and interference are high or low.

This method may be difficult to implement in comparison to the handoffs based merely on received signal strengths.

2.11 Handoff Methods

Several handoff strategies have been proposed in [122,133,178,179,203] based on signal strengths, as described below.

2.11.1 Threshold Method

The *Threshold* method [122] initiates handoff when the average signal strength of the current base station falls below a given threshold value and the signal strength of a neighboring base station is greater than that of the current base station. Proper selection of threshold value is necessary here as it reduces the quality of communication link leading to call dropping. This method is recommended by GSM Technical Specification GSM 08.08 [52]

2.11.2 Hysteresis Method

The *Hysteresis* method [179] initiates a handoff only if the signal strength of one of the neighboring base stations is higher than a certain given hysteresis margin of the current base station. The advantage of this method is that it prevents the ping-pong effect (defined in Section 2.12), but this method still initiates unnecessary handoffs even when the current serving base station signal strength is strong enough.

2.11.3 Threshold with Hysteresis

The *Threshold with Hysteresis* method [203] initiates a handoff when the signal strength of the current base station drops below a given threshold and the signal strength of a neighboring base station is higher by a given hysteresis margin to that of the current base station. This method is often used in practice with $+3$dB hysteresis.

2.11.4 Fuzzy Handoff Algorithm (FHA)

The *Fuzzy Handoff Algorithm* (FHA) [119] is a complex scheme using a set of prototypes assigned to each cell to calculate the serving base station. This uses a similarity measure to calculate the closeness of the membership function of a user to that of a base station to determine the need for a handoff.

2.11.5 Pattern Recognition Based Handoff

Pattern recognition based handoff [192] is an exhaustive method of finding the best possible handoff sequence and is practical for a canonical (Manhattan) topology but involves huge computation when applied to a general network.

The pattern recognition based handoffs exploits statistical pattern recognition techniques to analyze the patterns of radio wave propagation. This involves previous knowledge about the propagation characteristics of users' known paths. These techniques assume that the user moves in previously known paths and that each measured signal strength at a given point of the path has the same deterministic and random components (with deterministic means) [192]. The deterministic components are governed by path loss and shadow fading, and the random component is governed by Rayleigh fading.

Patterns of received signal strengths can be classified into the class associated with the serving base station or that associated with a neighboring base station, which will result in a handoff. Alternatively, pattern classes can be associated with stretches of user paths. The pattern recognition methods classify a sequence of received signal strengths into a class associated with a particular section of a user's path. Among the pattern recognition methods previously used are probabilistic neural networks [126] and Hidden Markov Models [92]. Pattern recognition methods that can be used include Support Vector Machines [157, 162, 180] and template matching.

2.12 Ping-Pong Handoff

It is possible that strong shadowing caused by large obstacles found in the line of sight with the serving base station, or a highly mobile user in a boundary region between two base stations causes handoff from the serving base station to the neigbouring base station for a short period (generally for less than 10 s) until it gets back to the older serving base station [183]. This effect, called the ping pong effect, can add too many unnecessary handoffs.

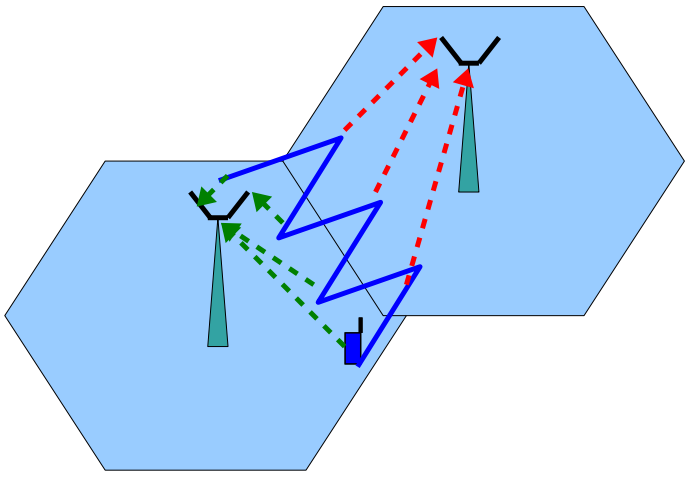


Figure 2.7: Ping-Pong handoff: serving base station changes quickly as the mobile station moves between base stations back and forth

Two commonly suggested methods to reduce this effect are:

- increase the hysteresis value,
- introduce a high averaging length for the signal strength measure to reduce small-scale fading [116].

However, neither of the above methods are practicable. A high hysteresis value may delay a necessary handover at a boundary between two cells and a high averaging time may also slow the dynamics of handoff processes to the extent that calls could be lost. Therefore, finding an appropriate solution to this problem without causing delays for necessary handoff is a question that currently has the attention of many researchers [58,100,104,107,109,131,145,146,198]. If the

hysteresis can be variable, and the ping-pong case can be distinguished uniquely from the genuine boundary crossing case, this problem can be solved.

2.13 Resource Allocation

A handoff call can be blocked if there is no free channel available in the new cell. As blocking a handoff call is more unacceptable than blocking a new call, prioritization schemes have been used. Various kinds of channel assignment strategies with handoff prioritization have been proposed to utilize the radio spectrum efficiently and reduce the probability of forced termination [77, 144, 173, 174]. Examples of such strategies are:

- **Channel Reservation:**

  In early proposals of handoff and resource allocation, a fixed portion of the radio capacity is permanently reserved for handoffs. The reserved channels cannot be used by new calls. Once the reserved channels are full, handoff and new calls compete for the remaining channels [37, 89]. Therefore, channel reservation leads to the increase the blocking probability of new calls [84].

- **Queuing Scheme:**

  Handoff requests are queued if there is no free channel available in the new cell. They maybe later admitted in case channels are freed [101,137,147,159]. The probability of successful handoff [54, 82] can be improved by queuing handoff requests at the cost of increasing the probability of blocking new calls, since new calls are not assigned a channel until all handoff requests in the queue are served [177]. Queuing can happen in the overlap region between two adjacent cells where a mobile station can communicate with more than one base station. It is very useful in macro-cells, as the mobile station can afford to wait for handoff before signal quality deteriorate to an unacceptable level.

The channel carrying approach proposed in [89] is an interesting practical approach for resource allocation combining the above two methods. A user requesting a handoff always uses a channel in the current cell. If the channel allocation scheme can support the carrying of the channel to the new cell, the channel is allocated to the handoff process.

As an alternative to prioritization schemes, mobility prediction schemes are used to support the channel allocation [15, 18, 84, 201]. In dynamic resource allocation [14, 22, 196], the amount of resources reserved for handoff at each base station should vary, supporting the mobility and the handling of variable traffic loads. Mobility prediction schemes received attention to support dynamic resource allocation, particularly for macro cell based networks. A pattern matching technique and a self-adaptive extended Kalman filter were used in [110] to predict the next cell, based on cell sequence observations, signal strength measurements and cell geometry assumptions.

As it is observed that cells are getting smaller, handoffs will occur more frequently and the movements of the user will become almost impossible to track. For the case of an ad hoc sensor network, it is reasonable to assume that users (sensors) as well as base stations (cluster heads) are randomly moving. Predicting user movement will be challenging. In [10], it was proposed that each user has a specific area called the "shadow cluster", which moves along with the user. This is a subset of the available base stations, and each base station in it expects the arrival of the user. This is a distributed method involving some communication and processing overheads at base stations which (except in the sensor network case) are generally connected to the wire line network. As the mobile terminal moves the shadow (base stations around the terminal and the direction of the movement) will be updated in real time, and the resources are dynamically allocated to the base stations in the shadow. A statistical strategy for resource allocation through the estimation of the handoff probabilities is proposed in [201]. Similarly to the strategy adopted in [10], it is not attempted to predict the long term movement of a user, and instead the handoff probabilities are used to estimate the number of resources to be reserved.

2.14 Analysis of Moving Patterns

The problem of sequential pattern analysis has been receiving considerable attention since the pioneering work [9, 172] due to its application in various areas including the analysis of moving histories of mobile users. These studies can reveal patterns in mobile user behavior that are useful for resource allocation. It is also possible to analyze on how handoff actually occurred in practice. Some of the studies use the BALI2 database [1] of moving histories for testing developed algorithms. When the movement of the mobile can be modeled as a Markov chain, there are other ways of "tracking" the mobile user described in Appendix B. Two different methods are commonly used for the analysis of moving patterns of mobile users. *A priori* based methods for sequence data mining are computationally expensive [9, 16, 123, 143] whereas projection based methods [85] may require large memory resources. These studies are aimed at finding the largest repeating subsequence with a frequency greater than a given threshold.

2.15 Motivation for Handoff Evaluation

Cellular networks use handoff methods extensively in the management of resources. The main questions Part I is trying to answer are: What is the best existing handoff method? What factors will influence the selection of a handoff? Chapters 3 and 4 revisit the existing method of handoff evaluation and show that it is not adequate, which also gives the main motivation to propose a comprehensive framework to replace the existing methods. Clearly, the proposed framework will have a performance measure, and an ideal handoff sequence for a given set of measurements based on that performance measure. It is important for the performance measure to take into consideration the factors such as cost of a terminated call, and the possibility of attempts to reconnect. Using ideal handoff sequence as a benchmark any handoff method can be evaluated.

2.16 Summary and Outline of Part I

This section provides a brief review of the relevant literature. Firstly, this chapter has explained the motivations for handoff evaluation.

Part I revises the proposed solutions to the handoff problem in cellular networks. This chapter provides a brief review of the relevant literature to the handoff problem in cellular networks, and the problem statement of Part I. Firstly, this chapter explains the cellular concept and provides a brief review of the characteristics of radio propagation. The chapter then discusses the signal propagation model used in later chapters. It also describes cellular hierachy, GSM, handoffs in detail and presents the motivation for handoff evaluation in cellular networks.

Chapter 3 proposes a new framework, based on signal quality, for performance evaluation and comparison of existing handoff algorithms and an off-line cluster-based computationally-simple heuristic algorithm to find a near optimal handoff sequence. This sequence replaces the need to use an exhaustive method to find a benchmark. When the users travel on predefined paths such as roads and footpaths, the benchmark can be used by pattern recognition methods to predict the possible future handoff sequence.

Chapter 4 extends the framework introduced in the previous chapter adopting call dropping and considering two options after a call is dropped: either repeated attempts are made to establish the call (retrial) or no such attempt is made (non-retrial). It also presents performance evaluation and comparison between existing handoff algorithms using the extended framework. Methods of estimating handoff cost and optimal values for retrial and non-retrial models are also investigated.

CHAPTER THREE

Optimal Handoff Sequence

3.1 Introduction

Increased demand for mobile services has led to a reduction in cell radius and more handoffs. The significance of having a correct criterion for evaluating handoff methods is that telecommunication providers can choose the right handoff algorithm that enables them to meet customers quality of service (QoS) requirements at competitive cost.

3.1.1 Motivation

Generally, a handoff sequence is evaluated against the cost using the number of handoffs justifiable and a measure of quality or QoS it provides. However, the use of simple average signal strength as a measure of quality considered in the thesis [191] at Stanford and in some papers [83, 119, 192] may not be the right choice. If the received signal strength falls below the acceptable level, S_{min}, the call is likely to discontinue, causing degradation of QoS. If there is a handoff sequence with a small number of assigned base stations having very high signal strengths and a large number of assigned base stations with signal strengths just below S_{min}, the average signal strength can still indicate a high value of QoS. If it is considered as part of the cost function, it can lead to a sub-optimal sequence as the best sequence. Unfortunately, no work is reported in the literature to provide a comprehensive framework to evaluate handoff methods. The best handoff sequence, if found, can be used as a benchmark. Unlike the small canonical prob-

lems considered in most research studies, where the user moves from one base station to another, the exhaustive approach to find the best handoff sequence is not practicable in realistic scenarios with multiple paths to multiple base stations.

Therefore, the first questions this work attempts to answer are: What is a realistic framework for evaluation and comparison of handoffs? How can we estimate the best handoff sequence or the benchmark for comparison for a given set of signal strength measurements without using exhaustive methods that are not practicable in most cases?

Many authors have researched handoff in cellular networks. But no literature investigates the impact of handoff by using real data values. It is essential to observe the impact of handoff in a real scenario. Would it be large enough to consider the handoff problem as important? We observe BALI-2 data in trying to answer this question in Section 3.8.4.

Once the ideal or best sequence is known, various pattern recognition methods (for example template matching techniques) can be used to find the pattern prior to each handoff. If the pattern recognition based method used is computationally simple, it can also have a low delay. Any pattern recognition method developed should be compared with other existing handoff algorithms for the handoff delay. Hence, the next obvious questions are: What are the advantages of using pattern recognition based methods? Can the best handoff sequence be used for pattern recognition based handoff?

In this chapter, research results are presented to address the above questions. The remainder of the chapter is organized as follows. Section 3.1.2 summarizes a number of recent attempts to improve existing handoff methods, Section 3.1.3 sets out the key assumptions and Section 3.1.4 presents and then describes the definitions useful for this study. The required formulae for call dropping probability are derived in Section 3.3. The initial proposal for Call Quality Signal Level measure for handoff is presented in Section 3.2. In Section 3.4, the near optimal way of finding the best handoff sequence is proposed. Section 3.5 emphasizes its application in a numerical example. The previously proposed Call Quality Signal Level measure is extended in Section 3.6. Section 3.8 presents simulation results

comparing existing handoff methods with the Best Handoff Sequence and also considering exhaustive and dynamic programming methods. It also analyzes real data samples to observe the need for handoff, and discusses how the proposed methods can be extended to pattern recognition based handoffs. Section 3.9 provides a summary of the chapter and concluding remarks.

In this chapter, we promote the user signal level as a key criterion for evaluation of handoff algorithms in addition to other handoff evaluation approaches such as delay [192, 206] and call dropping probability [45, 76, 86, 93]. As blocking probabilities are considered as a performance index for cellular networks, we attempt to keep it to a minimum. Provision of buffers, for example, can reduce the blocking probabilities [33, 61, 77, 78, 101]. We also introduce an off-line heuristic algorithm which obtains a near optimal "best" handoff sequence (BHS) that can be used as a benchmark.

3.1.2 Improvements to Existing Handoff Methods

The *Threshold* method can be easily improved by restricting the handoffs to occur only where the new base station can provide signal strength stronger than the minimum acceptable level. Otherwise, the handoff should not occur as it cannot lead to improvement in QoS. Consequently the *Threshold with Hysteresis* method can also be extended to allow such a restriction. Several recent extensions have been proposed to improve the hysteresis based strategies [108, 127]. In [127] three values were proposed for the *Threshold with Hysteresis* and *hysteresis* method. When the current base station is busy (more handoff and new call requests than available resources) the hysteresis is lowered to 2 dB to encourage quicker handoff. When the new base station is busy, then the hysteresis is increased to 10 dB to discourage the handoff. If both base stations have similar levels of activity, then a hysteresis at 6 dB is used. This approach has shown right direction, but needs to be generalized to include a more adaptive threshold and hysteresis method.

It is possible that strong shadowing caused by large obstacles found in the

line of sight with the serving base station or a highly mobile user in a boundary region between two base stations causes handoff from the serving base station to the neighboring base station only for a short period (generally for less than 10 s) until it gets back to the older serving base station. This effect, called the ping-pong effect, in which handoffs occur back and forth between base stations, can add many unnecessary handoffs. Two commonly suggested methods [127] to reduce this effect are:

- increase the hysteresis value
- introduce a high averaging length for the signal strength measure

However, neither of these methods are practical. A high hysteresis value may delay a necessary handoff at a boundary between two cells and a high averaging time may also slow the dynamics of handoff processes to the extent that calls could be lost. Therefore, finding an appropriate solution to this problem without causing delays for necessary handoff is a question that has the attention of many researchers [108, 145, 146, 177]. If the hysteresis can be varied, and the ping-pong case can be uniquely distinguished from the genuine boundary crossing case, a solution to this problem can be found. Although the specification GSM 08.08 [52] considers only the *Threshold* method, the commercial providers seems to use a +3 dB hysteresis value in addition (i.e., the *Threshold with Hysteresis* method) to minimize the ping-pong effect.

In cellular or micro cellular environments in cities, users often move on predetermined paths. As the built environment is not changed often, this regularity can be exploited in a pattern recognition based handoff method. The most suitable sequence of assigned base stations or the Best Handoff Sequence (BHS) can provide the basis for pattern recognition based handoff methods. This chapter describes in detail a computationally simple method to estimate BHS which can also be used as a benchmark for comparing different handoff algorithms. Further, it uses the proposed Call Quality Signal Level (CQSL) evaluation to compare various well known handoff methods.

3.1.3 Assumptions

- We assume that channel capacity is unlimited, and therefore we do not consider handoff queuing.
- We assume a homogeneous network where all cells are identical in size, user mobility and cell coverage as in [77, 148].
- Each cell is assumed to have an equal number of neighbors as in [142].
- A log normal propagation model is assumed and no power control is assumed to exist [160].
- Mobile users are uniformly distributed in the region [125].
- Users' directions are randomly chosen between $[0, 2\pi]$ uniformly [125].
- Here we assume that we know the users' locations and base stations, and give information about signal strengths in each point.
- A call is dropped after observing the drop level signal level for a number of consecutive sample points [158].

3.1.4 Definitions

- S_{min}: minimum signal strength below which the signal quality is unacceptable to the user ($S_{max} > S_{min}$).
- S_{max}: signal strength beyond which the marginal benefit is considered negligible.
- S_{drop}: dropping level signal below which the call is dropped, if that level is maintained for a certain period. We also refer to this signal range as "drop level signal level".

Consider a cellular mobile network with M base stations designated $B_1, B_2, ..., B_M$. Define $\mathcal{B} = \{B_1, B_2, ..., B_M\}$. Let a sample path l be an arbitrary path in which a mobile user is traveling. Consider a set of paths denoted Θ for the purpose of

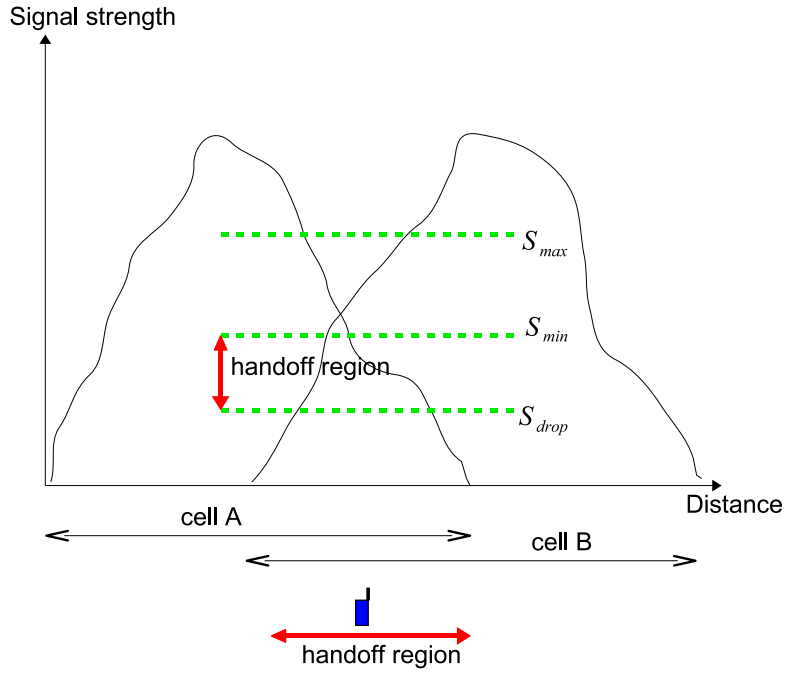


Figure 3.1: Difference signal strength values: S_{min}, S_{max} and S_{drop}

evaluating handoff algorithms. Sample points are points on the sample path for which the signal strength received from base stations are measured. Let S_{ij} be the signal strength at sample point i received from base station B_j. Define a handoff sequence x or $x(l)$ of sample path l, thus represents a sequence of base stations assigned to the sample points in l, assigning $b_i \in \mathcal{B}$ to the i^{th} sample point, i.e., $x =< b_1, b_2, ..., b_N >$ where N is the number of sample points. (Note that b_i and b_j, $\forall i, j$ may designate the same base station.)

For every sample path, the set of all possible handoff sequences is denoted by

$$X = \{x_i | 0 < i \leq M^N\}.$$

The number of handoffs $\gamma(x)$ in a handoff sequence x equals the number of changes in the base station sequence. For example, the handoff sequence $x = \{B_1, B_1, B_2, B_3, B_3, B_3\}$ has $\gamma(x) = 2$.

For a given handoff sequence $x \in X$, $S_i(x) = S_{ij}$ is defined such that $B_j = b_i$, $b_i \in x$. Let S_{min} be the minimum signal strength below which the signal quality

is unacceptable to the user.

3.2 Call Quality Signal Level Measure (CQSL)

In this section we propose a new framework, based on signal quality, for performance evaluation and comparison between existing handoff algorithms. It includes new call quality measures and an off-line cluster-based computationally-simple heuristic algorithm to find a near optimal handoff sequence used as a benchmark. We then compare existing handoff algorithms and identify the trade-off between signal quality and number of handoffs.

For a given sample path and its associated handoff sequence, we define the following signal quality measures.

1. Average Received Signal Strength **(ARSS**(x)**)** is defined by

$$\frac{1}{N}\sum_{i=1}^{N} S_i(x).$$

2. Number of Acceptable Sample Points **(NASP**(x)**)** represents the number of sample points of the handoff sequence with signal strength above S_{min}. Let

$$\Omega_x = \{i | S_i(x) \geq S_{min}\},$$

then $\mathrm{NASP}(x) = |\Omega_x|$, where $|\mathrm{Y}|$ denotes the number of elements (cardinality) in the set Y.

3. The concept of Call Quality Signal Level (CQSL(x)) proposed in this chapter combines the above two measures and is defined by

$$\mathrm{CQSL}(x) = \frac{\sum_{i\in\Omega_x} A_i(x)}{|\Omega_x|} - CN(x), \tag{3.1}$$

where $A_i(x) = S_i(x)$ if $S_i(x) < S_{max}$, otherwise $A_i(x) = S_{max}$, and $N(x) = (N - |\Omega_x|)$ is the number of samples with signal strength lower than S_{min},

and C is the cost (or the penalty) for an unacceptable sample point. We assign $\sum_{i\in\Omega_x} A_i(x)/|\Omega_x|$ to zero when $|\Omega_x| = 0$.

Let p be the maximum allowed proportion of sample points with signal quality below S_{min}, i.e., $N(x)/N \leq p$. The p value may be agreed between the service provider and the user. Assuming $|\Omega_x \neq 0|$, the minimum value that CQSL(x) can take is when

$$\frac{N(x)}{N} = p, \tag{3.2}$$

$$\sum_{i\in\Omega_x} \frac{A_i(x)}{|\Omega_x|} = S_{min}, \tag{3.3}$$

in (3.1). We choose C such that the above minimum value is greater than or equal to zero. The parameter C in (3.1) can be bounded as follows:

$$C \leq \frac{\sum_{i\in\Omega_x} A_i(x)/|\Omega_x|}{N(x)} = \frac{S_{min}}{pN}. \tag{3.4}$$

Here we choose the cost to be linear. However, we could also set it up dynamically to reflect the fact that consecutive unacceptable sample points are worse than a single unacceptable sample point. Using (3.1) and (3.4), we obtain

$$\text{CQSL}(x) \geq \frac{\sum_{i\in\Omega_x} A_i(x)}{|\Omega_x|} - \frac{S_{min}(N - |\Omega_x|)}{pN}. \tag{3.5}$$

The measures ARSS(x), NASP(x) and CQSL(x) are defined for any $x \in X$ on an arbitrary sample path $l \in \Theta$. For a given handoff algorithm there is at least one optimal handoff sequence for a given l according to the algorithms criteria. Assuming that all sample paths are independent, and equally important, different handoff algorithms will be evaluated by averaging the values of these measures over all the sample paths. For example we use the average:

$$\overline{\text{CQSL}} = \sum_{l} \frac{[\text{CQSL}(x(l))]}{\eta}, \tag{3.6}$$

where $\eta = |\Theta|$.

In (3.5) the parameter p is related to call dropping probability. In practice, a call is dropped if either the high co-channel interference, or the signal level below a certain threshold $S_{drop} < S_{min}$ (call dropping condition) is maintained for d consecutive samples in the handoff sequence. (We use this simple criteria to model duration of bad connections.)

Furthermore we introduce the signal quality per handoff:

$$\lambda = \frac{\overline{\text{CQSL}}}{\overline{\gamma}}, \tag{3.7}$$

where

$$\overline{\gamma} = \sum_{l} \frac{[\gamma(x(l))]}{\eta}.$$

We will use (3.7) to compare different handoff methods in section 3.8.

3.3 Call Dropping Probability

Call dropping occurs after the call is established but before the call is properly terminated. It is also possible that a call is dropped due to unavailability of voice channels which is considered as a blocked call and not a dropped call.

The call dropping rate and the specified minimum voice quality level required are considered as inversely proportional to each other. In (3.5) the parameter p is related to a call dropping probability. In practice, a call is dropped if either there is a high co-channel interference, or the signal level is maintained for d consecutive samples in the handoff sequence below a certain threshold $S_{drop} < S_{min}$.

Let P_{drop} be the probability that a call is dropped. For a sample path l, let δ_l be the probability of receiving a signal strength below S_{drop}. Let μ_l be the probability of receiving co-channel interference above a specified value. Therefore, the call dropping probability in the sample path l with $N_l > d$ consecutive sample points

is given by the following recursive formulae:

$$P_{drop_l}(N_l) = P_{drop_l}(N_l - 1) + P_l^d(1 - P_l)(1 - P_{drop_l}(N_l - d - 1)), \tag{3.8}$$

where the probability of having a single sample point satisfying the call dropping condition is

$$P_l = 1 - (1 - \delta_l)(1 - \mu_l),$$

$$P_{drop_l}(N_l < d) = 0,$$

and

$$P_{drop_l}(N_l = d) = P_l^d.$$

If the co-channel interference is neglected then $P_l = \delta_l$. The average call dropping probability over η sample paths is:

$$P_{drop}(N_l, d) = \frac{1}{\eta} \sum_{l=1}^{\eta} P_{drop_l}(N_l). \tag{3.9}$$

Given a number of sample points N_l, the value of d can be determined such that $P_{drop}(N_l, d)$ in (3.9) is below a certain threshold according to a QoS requirement. The value of p in (3.5) is chosen to be $\geq d/N_l$. Knowing p, for a given handoff sequence of a sample path with $N_l = N$ sample points, we can calculate the minimum value of CQSL(x) using (3.5).

3.4 The Best Handoff Sequence (BHS)

Our aim in this section is to obtain the best handoff sequence (BHS) which minimize the number of handoffs, and maintain the signal strengths $\geq S_{min}$ at all times [68]. Herein, we provide an off-line algorithm to find this sequence which

will represent a benchmark value. A brute force method (exhaustive search) is impractical because of a large number of possible sample paths M^N involved. Due to this reason we use a heuristic method based on a cluster approach specified in [68] to find the optimal BHS by maximizing CQSL and minimizing the number of handoffs (γ).

Therefore, our aim here is to obtain the BHS defined by the following multiple objective unconstrained optimization problem:

$$\max_{x \in X} \ (1-a)\psi(x) - a\gamma(x), \tag{3.10}$$

where

$$\psi(x) = \frac{\sum_{i \in \Omega_x} A_i(x)}{|\Omega_x|} - \frac{S_{min}(N - |\Omega_x|)}{pN},$$

and $a \in [0,1]$, the weight factor indicates the relative importance of the two objectives $\psi(x)$ and $\gamma(x)$. Since all the paths are independent, maximization of CQSL$(x(l))$ for any path l also maximizes $\overline{\text{CQSL}}$. Note that in [192] only two base stations are considered and therefore an exhaustive search for the optimal handoff sequence is used. In our case exhaustive search is impractical because of a large number of possible sample paths (M^N) involved. Therefore, we propose a heuristic method based on the so-called cluster approach.

Let G_{ij}, referred to as a cluster, be a set of signal strengths $\geq S_{min}$ from base station j associated with a group of consecutive sample points $\{i, i+1, ..., i+L_{ij}-1\}$, where $1 \leq L_{ij} = |G_{ij}| \leq N - i + 1$. Let

$$W_{ij} = \sum_{r=i}^{i+L_{ij}-1} \frac{S_{rj}}{L_{ij}}. \tag{3.11}$$

In order to solve (3.10) our heuristic algorithm maximizes $\psi(x)$ by finding maximum average signal level W_{ij} and at the same time minimizes the number of handoffs $\gamma(x)$ by choosing longest clusters (max L_{ij}).

Let H_{ij} be the parameter associated with a cluster G_{ij} which is defined as the

weighted value

$$H_{ij} = \alpha L_{ij} + (1 - \alpha) W_{ij}, \tag{3.12}$$

where $\alpha \in [0, 1]$. According to (3.12) when $\alpha = 0$, $H_{ij} = W_{ij}$, and hence maximizes signal quality, and when $\alpha = 1$, $H_{ij} = L_{ij}$, and hence maximizes the cluster length.

Let Φ be a signal strength matrix of $N \times M$, received from M base stations with N sample points for a particular sample path. Here, we aim to find the BHS as a set of G_{ij} starting from the first row ($i = 1$) until the last row ($i = N$) in Φ, which maximizes the optimization function (3.10). Our heuristic algorithm is described as follows.

Step 0: Set $i = 1$.

Step 1: At the i^{th} sample point in Φ the algorithm finds all clusters which start from this i^{th} sample point, it then selects the cluster G_{ij^*} with maximum value of H_{ij}, i.e.,

$$G_{ij^*} = \arg \max_{G_{ij}} H_{ij}. \tag{3.13}$$

If there is only a single cluster starts from the i^{th} sample point, then automatically it will be selected. If there is no cluster starting from i then go to **Step 2**. The algorithm assigns the base station j^* associated with the selected cluster as the serving base station for the $\{i, i+1, .., i+L_{ij^*}-1\}$ sample points. If $i + L_{ij^*} < N$ then return to **Step 1** with

$$i = i + L_{ij^*}$$

until $i = N$. **Step 2:** Starting from row i the algorithm skips all the rows in Φ until it finds a new row u with a cluster G_{uv} containing at least one sample point with signal strength $> S_{min}$ (i.e. $L_{uv} \geq 1$). Return to **Step 1** with this u^{th} sample point to find the G_{uv^*} cluster with maximum value of H_{uv}. In addition, for the above skipped sample points between $\{i, ..., u-1\}$, the algorithm assigns the previous serving base station j^* (no handoff) or the new serving base station

v^* (handoff) such that the average signal strength over all the skipped sample points is maximized. If no such u^{th} sample point is found, then we assign the previous serving base station j^* as the new serving base station. If $u + L_{uv^*} < N$ then return to **Step 1** with $i = u + L_{uv^*}$ until $i = N$.

3.5 BHS with a Numerical Example

Here we consider a scenario involving a user moving through a road and consider three base stations. The task is to determine if the user is getting good connection all the time with less handoff. The choice of deciding between various base stations arises because signal strength from the better base station tend to make the good quality call. At the same time constantly changing of base stations is also costly. Finding optimal solution is the challenge.

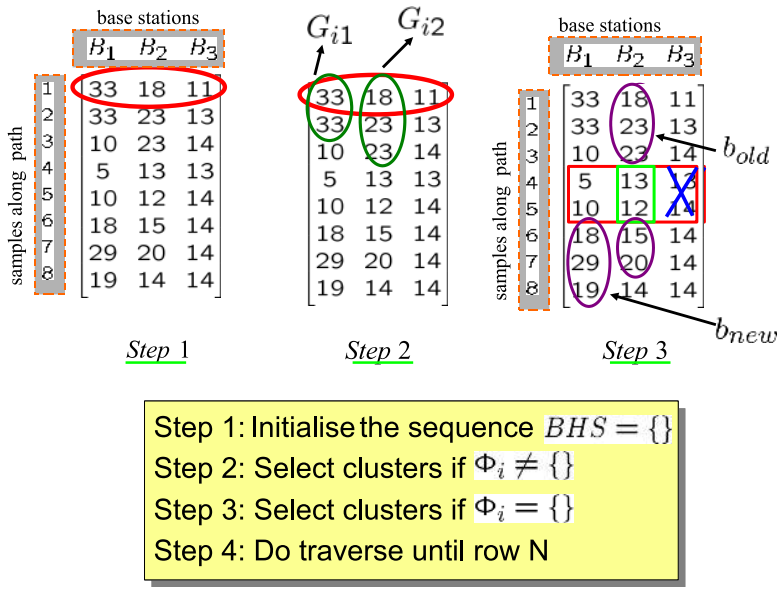


Figure 3.2: Best Handoff Sequence (BHS) algorithm, when $S_{min} = 15$ dB, base stations $M = 3$, and sample points $N = 8$, for a particular sample path

Consider a $(N \times M)$ signal strength matrix, Φ, received from $M = 3$ base stations, with $N = 8$ sample points for a particular sample path as given in Fig.

3.2. Let G_{ij}, referred to as a cluster, be a set of signal strengths $\geq S_{min} = 15$ from base station j associated with a group of consecutive sample points $\{i, i+1, ..., i+L_{ij}-1\}$, where $1 \leq L_{ij} = |G_{ij}| \leq N - i + 1$.

Let W_{ij} be the average signal strength of the cluster. Let H_{ij} be the parameter associated with a cluster G_{ij} which is defined as the weighted value $H_{ij} = \alpha L_{ij} + (1-\alpha)W_{ij}$, where $\alpha \in [0,1]$. The heuristic algorithm maximizes signal quality by finding maximum average signal level W_{ij} and at the same time minimizes the number of handoffs $\gamma(x)$ by choosing longest clusters. In the following we demonstrate the algorithm based on the example given in Fig. 3.2.

Set BHS $= \{\}$ and $i = 1$.

Step 1: Find the subset Φ_i as a set of the base stations from which its signal strength in the i^{th} row of the matrix Φ is $\geq S_{min} = 15$. So according to Fig. 3.2, $\Phi_i = \{1, 2\}$. If Φ_i is empty we need to proceed to Step 3.

Step 2: Find all the G_{ij} clusters which starts from each of the element of the set Φ_1. In our example, we can identify two clusters G_{i1}, G_{i2}. The algorithm assigns the base station 2 associated with larger H_{ij} value as the serving base station for the first three sample points in the path. Therefore, we obtain BHS $= \{B_2, B_2, B_2\}$. We proceed to Step 1 with $i = 4$.

Step 3: With $i = 4$, Step 1 produces an empty set which need to be addressed in this Step. Because Φ_4 is empty, we skip two rows (4^{th} and 5^{th}) until finding a row in a matrix Φ which contains a signal strength > 15 (in our example this is the 6^{th} row). We then proceed again from Step 1 by setting $i = 6$ to find the optimal base station for the 6^{th} sample point and onwards. After repeating Step 1 and 2 we obtain $\{B_1, B_1, B_1\}$ as an optimal handoff sequence for the $6^{th}, 7^{th}$ and 8^{th} sample points in our example. For the skipped sample points 4^{th} and 5^{th} the algorithm then assigns the previous serving base station, $b_{old} = B_2$, (no handoff) or the new serving base station, $b_{new} = B_1$ (handoff) such that the average signal strength over all the skipped sample points is maximized. In our example $b_{old} = B_2$ is chosen because the average signal strength from B_2 is greater than the average signal strength from B_1 over the skipped sample points 4^{th} and 5^{th}. If after the skipping rows in matrix Φ we cannot continue with Step 1, i.e., there is no new

serving base station, we then simply continue with the old base station over all the skipped sample points. Repeat Steps 1, 2 and 3 until the last row of matrix Φ. In our example shown in Fig. 3.2 we obtain BHS $= \{B_2, B_2, B_2, B_2, B_2, B_1, B_1, B_1\}$.

Recalling our aim of the BHS is to minimize the number of handoffs and maintain the signal strengths $\geq S_{min}$ at all times, we observe that in the above numerical example, the number of handoffs $\gamma(x) = 1$.

It should be noted that the Best Handoff Sequence BHS contains indirectly the following features not considered by any other handoff method:

- when the strength of a serving base station is well above S_{min}, further reduction in the number of handoffs can be achieved by adapting a *flexible* or *variable* hysteresis. Clustering all the signal strengths above the minimum acceptable level and then selecting the base station corresponding to the best cluster as the serving base station will help to achieve a nonlinear variable hysteresis.

- unnecessary handoffs are avoided in sample points where all base stations provide signal strengths below S_{min}.

3.6 New CQSL

However, the above CQSL measure does not effectively distinguish between two sequences with the same average signal strength of good sample points, where one has a large number of good sample points with a relatively small signal strength, and another has only few good sample points but with a large signal strength. Because of it, we slightly modify the CQSL measure in this section by deducting the penalty before getting the average.

New CQSL is described as follows:

$$\text{CQSL}(x) = \frac{1}{N} \left\{ \sum_{i \in N_g(x)} A_i(x) - CN_b(x) \right\}, \tag{3.14}$$

where $\forall x \in N_g(x)$,

$$A_i(x) = \begin{cases} S_i(x) & \text{if } S_i(x) \leq S_{max} \\ S_{max} & \text{otherwise} \end{cases},$$

N is the number of sample points,

$$N_g(x) = \{i | S_i(x) \geq S_{min}\},$$

C is the cost (or the penalty) for an unacceptable sample point,

$$N_b(x) = \left(N - |N_g(x)|\right),$$

is the number of samples with signal strength lower than S_{min}.

Similar to [68], we can obtain the lower bound as

$$\text{CQSL}(x) \geq \frac{\sum_{i \in N_g(x)} A_i(x)}{N} - \frac{S_{min}|N_b(x)|\,|N_g(x)|}{pN^2}, \tag{3.15}$$

where p is the maximum allowed proportion of sample points with signal quality below S_{min}, i.e., $N_b(x)/N \leq p$.

We consider here the lower bound for CQSL for comparison of different handoffs. In current practice, service providers do not associate 'p' the maximum bound of the proportion of "bad" sample points with the QoS requirement, however, the framework proposed herein provides such a parameter to support differentiated services. As every handoff incurs cost, for comparison purposes we also use λ, the quality per handoff. To compute λ, firstly we apply equation (3.15) to (3.6) and then to equation (3.7).

If a cost of single handoff is estimated as H_{cost} \$, the signal quality per dollar is

$$\frac{\lambda}{H_{cost}}.$$

We will therefore use (3.7) to compare different handoff methods in section 3.8.

A pattern recognition method can detect a pattern of serving base stations that may appear before an imminent handoff. The first step in this process is to determine the best handoff sequence for a given sample path. For example, in a path of 100 sample points and 3 serving base stations, there are 3^{100} possible sample paths and one of them is the Best Handoff Sequence (BHS). Signal level quality and the number of handoffs for such sequences should be evaluated before selecting the BHS. As this requires high computational complexity, a cluster based approach is proposed in [68] to find the ideal or the best handoff sequence. In the next section we describes in detail a computationally simple method to estimate BHS which can also be used as a benchmark for comparing different handoff algorithms.

3.7 Simulation Set-up

We ran wireless network simulations using MatLab. For this experiment we consider $M = 3$ adjacent cells with 100 m radius. The base stations are located in a plane with the following coordinates: $(100, 150)$, $(250, 75)$ and $(250, 250)$ [meters]. We randomly generate $\eta = 1000$ sample paths, each with $N = 100$ where each pair of consecutive points are one meter apart.

A log-normal propagation model [60] was assumed to generate signal strengths in each sample point along all the sample paths, i.e., $S_{ij} = K_1 - K_2 log(r) + \rho$, where $K_1 = 85$; $K_2 = 35$ are constants, r is the distance in meters from the user to the base station, and ρ is Gaussian distributed ($N(0, \sigma^2)$) representing the shadowing effect. We set $\sigma = 3,5$ dB as in [192], shadowing correlation distance equals 20 m, $S_{min} = 15$ dB as in [119] and $S_{max} = 1.5 S_{min}$.

All the sample paths are straight lines that start from points in the square area $\{(100, 100), (200, 100), (200, 200), (100, 200)\}$. Their directions are randomly chosen between $[0, 2\pi]$ uniformly. Note that P_{drop} is computed by (3.9) as a decreasing function of d, where $d \geq 3$ gives the call dropping $\leq 1\%$. The p value in (3.5) is selected such that $p \geq d/N = 0.03$. Here we use $p = 0.1$.

Table 3.1: Parameter values used in handoff simulation

Symbol	Description	Value
M	Number of base stations	3 [119]
N	Number of sample points	20, 100
r	Cell radius	100 m [192]
T	Threshold	variable
η	Number of sample path	1000 [192]
K_1	Path-loss constant	85 dB
K_2	Path-loss exponent	35 dB
σ	Standard deviation of shadow fading	3 dB [192] and 5 dB [135,178]
S_{min}	Minimum acceptable signal strength	15 dB [119]
S_{max}	Maximum acceptable signal strength	$1.5S_{min}$ dB
S_{drop}	Dropping signal strength	14.5 dB
	Correlation distance of shadow fading	20 m [114]
	Spatial sampling interval	1 m [114]

3.8 Simulation Results

Here we compare the different handoff methods introduced in Chapter 2, using the quality measures of (3.5) and (3.7) for Figures 3.10, 3.3 and (3.15) for Figures 3.4, 3.6.

The values of $\overline{\text{CQSL}}$ and $\overline{\gamma}$ in all figures are obtained by varying the T_{HO} threshold in the Threshold method and Threshold + Hysteresis method (with +3 dB hysteresis margin), and the H hysteresis threshold in the Hysteresis method, respectively, from 1 to 30 dB. A similar range was used when varying the so-called similarity threshold in the FHA method.

3.8.1 Comparison of Handoff Algorithms with BHS

As a benchmark value, we show in Fig. 3.3 the $\overline{\text{CQSL}}$ and $\overline{\gamma}$ of the BHS for different values of $\alpha \in [0, 1]$ in (3.12). Observe that these values are almost unchanged (insensitive) for different α values which indicates that we can use either L_{ij} or W_{ij} for the cluster selection.

The complexity of the proposed heuristic algorithm is $O(MN)$ in comparison

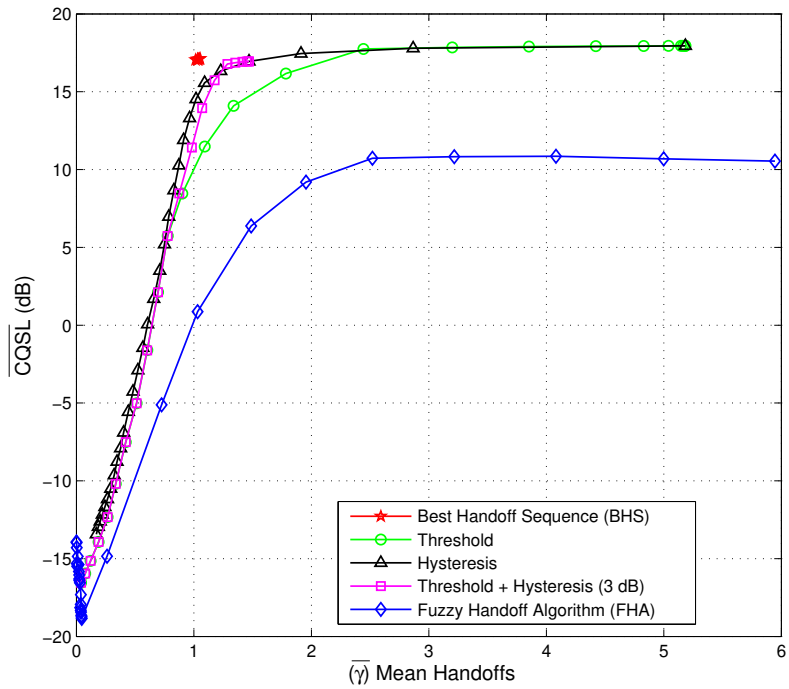


Figure 3.3: Comparison of the various handoff algorithms: $\overline{\text{CQSL}}$ versus $\overline{\gamma}$, when standard deviation of shadow fading $\sigma = 3$ dB, with 1000 users

to the complexity $O(M^N)$ using an exhaustive search. Using the same network but with $N = 20$ sample points, the difference between a solution resulting from an exhaustive search and ours was never more than 0.24% as in Table 3.3.

Figure 3.3 shows that the Hysteresis method and Threshold + Hysteresis method (with +3 dB hysteresis margin) provide the best values that are closest to the BHS. When high numbers of handoffs can be tolerated the Threshold method will be as efficient as the above two. Our simulations indicate FHA is less desirable than other methods.

Based on Fig. 3.3, optimum parameter settings for each handoff method can be obtained from the "knee" point of the corresponding curves (similar to [192]). For example the "knee" point for the Threshold handoff method is a point with $\overline{\text{CQSL}} = 14.09$ dB and $\overline{\gamma} = 1.33$ values according to Fig. 3.3. This is when the threshold T_{HO} is set to 14 dB which produces highest $\overline{\text{CQSL}}$ with lowest average number of handoffs.

In Table 3.2 we compare all handoff methods at their "knee" points using

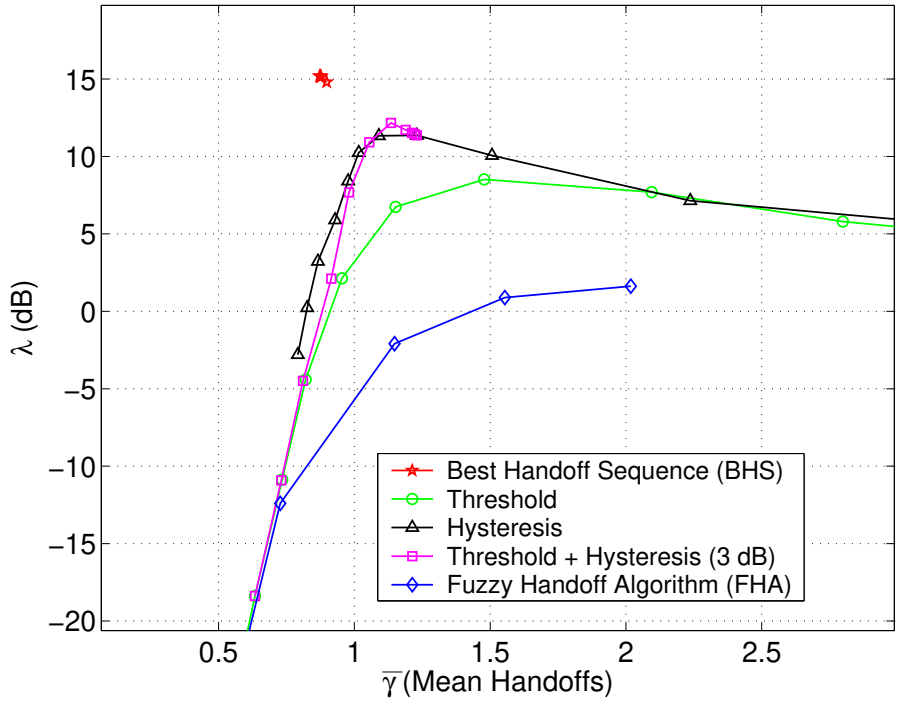


Figure 3.4: Comparison of the various handoff algorithms: λ versus mean handoffs ($\overline{\gamma}$), when $\sigma = 5$ dB, with 1000 users by varying the handoff threshold from 0-30 dB

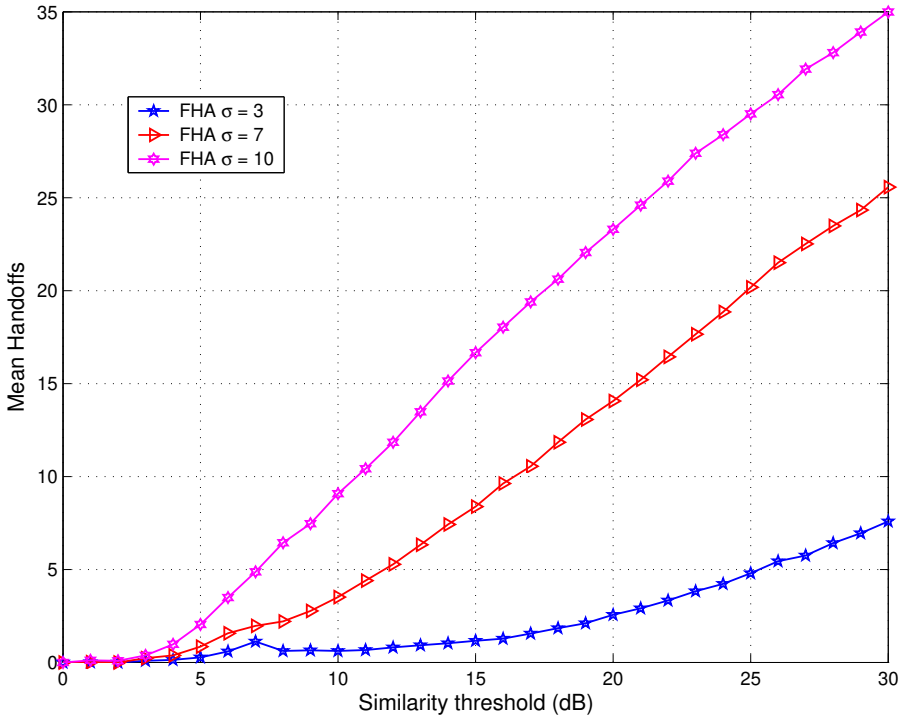


Figure 3.5: Mean handoffs for 1000 users in the FHA algorithm for different σ, standard deviation of shadow fading values

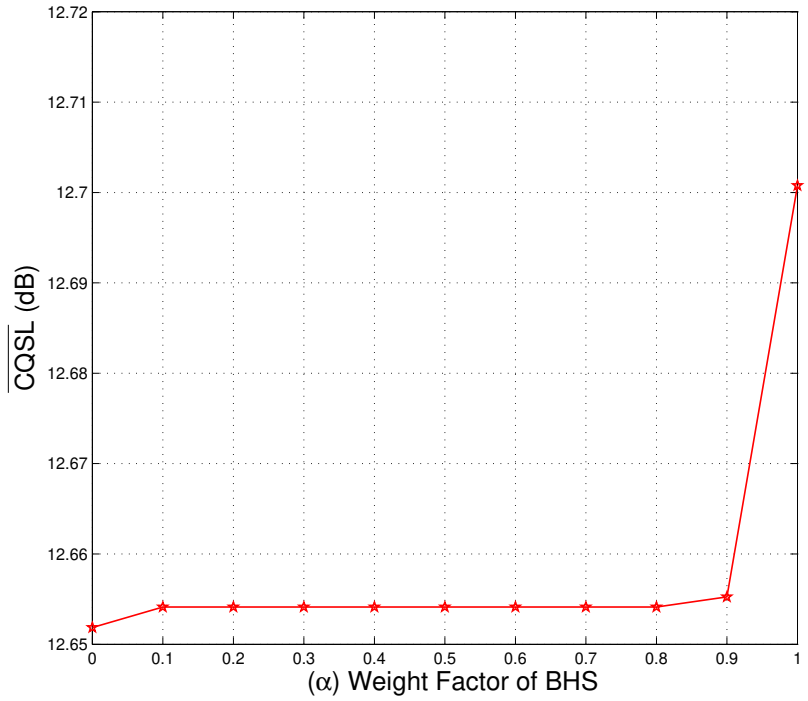


Figure 3.6: The effect of parameter α, for CQSL in the BHS: CQSL versus weight factor of BHS

the following quantities: $\overline{\text{CQSL}}$ in (3.5), λ in (3.7) and $\overline{\text{NASP}}/\overline{\gamma}$, where $\overline{\text{NASP}} = \sum_l [\text{NASP}(x(l))]/\eta$. The optimal threshold for each method at the "knee" point is presented in column identified with T_k in dB. The benchmark BHS has the highest λ and $\overline{\text{NASP}}/\overline{\gamma}$ values. We repeated our experiment for various N values ($N = 50, 100, 200$) and found the results to be consistent.

Table 3.2: "Knee" parameter values for all handoff algorithms

Method	T_k (dB)	$\overline{\text{CQSL}}$ (dB)	$\langle\, \lambda\,,\, \overline{\gamma}\,\rangle$	$\overline{\text{NASP}}/\overline{\gamma}$
BHS	15	17.05	$\langle 16.88\,,\, 1.01\rangle$	93.90
Thres+Hys	15	15.73	$\langle 13.44\,,\, 1.17\rangle$	78.11
Hysteresis	6	15.56	$\langle 14.27\,,\, 1.09\rangle$	83.73
Threshold	14	14.09	$\langle 10.59\,,\, 1.33\rangle$	68.25
FHA	25	9.18	$\langle 4.70\,,\, 1.95\rangle$	42.15

Figure 3.5 confirms that the mean number of handoffs for FHA increases with the increase of σ as reported in [119].

Figure 3.4 compares the number of handoffs of four handoff algorithms: Threshold, Hysteresis, Threshold with Hysteresis at 3dB and FHA, with BHS.

Figure 3.6 shows that the performance of BHS only slightly varies (12.65 to 12.7 dB) as α is varied. The best performance is observed when $\alpha = 1$, which justifies the use of cluster length rather than the weighted value H_{ij} in Step 2. It can be observed that the variation is not significant. The minimum number of handoffs needed to guarantee $p \leq 0.1$ can be found in Fig 3.4 by setting CQSL $= 0$ for each handoff method. It is clear that FHA in the present form fails to reach this quality level with the low number of handoffs. Similar to the approach taken by [153] and [192], optimum parameter settings can be obtained from the "knee" of the curves. Clearly the benchmark BHS with $S_{min} = 15$ dB and $\alpha = 1$ provides the most efficient parameter setting or the highest λ value. When high numbers of handoffs can be afforded the Threshold method with $T_{HO} = 14$ dB will be as efficient as the above two traditional handoff methods. Our simulations indicate FHA (similarity threshold at 21 dB) is less desirable in comparison to other methods.

A realistic framework for evaluation and comparison of handoffs was presented recently in [68]. It uses a computationally simple benchmark for comparison of various handoff strategies using a new realistic signal quality measure. This section illustrates in detail the algorithm to obtain the benchmark sequence and presents a comparison on commonly used handoff techniques and a fuzzy rule based handoff method using a modified form of the framework presented in [68].

Our results show that there is substantial room for improvement in existing handoff algorithms with respect to the signal level measures as well as the number of required handoffs. Previous studies have concentrated mostly on reducing the number of handoffs, neglecting quality of signal. It is therefore important to develop new handoff methods that would take into account both signal strength quality and number of handoffs as proposed in this work. The performance of any new proposed algorithm can be compared with our benchmark solution.

The effect of the selection of p on CQSL is also obvious from (3.5), as the CQSL

measure increases with the increase of p. We can call p the probability of call failure due to unavailability of a suitable base station if N is sufficiently large.

3.8.2 Optimal Value for BHS via Exhaustive and Dynamic Programming Methods

1. **Exhaustive Search:**

   The optimal handoff sequence can be defined as one which maximizes all evaluation methods. The following equation can be used recursively to find the optimal handoff sequence:

   $$\max_{i\in 1,..,N} \max_{(n,m)\in C} \left\{ \sum_{k=1}^{i} S_{kn} + \sum_{k=i+1}^{N} S_{km} \right\}, \tag{3.16}$$

   where $C = \{(n,m)|n \in \{1,..,M\}, m \in \{1,..,M\}, n \neq m\}$.

2. **Dynamic Programming:**

   Let S_{ij} be the signal strength at sample point i received from base station j. Consider a $(N \times M)$ Φ signal strength matrix received from M base stations with N sample points for a particular sample path, where

   $$1 \leq i, j \leq N.$$

   We obtain

   $$D_{max}[S(i,j), M] = \max_{k\in\theta} D_{max}[S(i,k), 0] + D_{max}[S(k+1,j), M-1], \tag{3.17}$$

   where $S(i,j)$ is a subset of S and M is number of handoffs and $\theta = i, i+1, ..., j-M+1$.

We claimed that BHS provides a near optimal solution. For example under the same conditions (same signal strength matrix) with same γ, the BHS algorithm provide 99.52% of signal strength of the optimal solution achieved by exhaustive search. Under the same conditions (same signal strength matrix and γ)

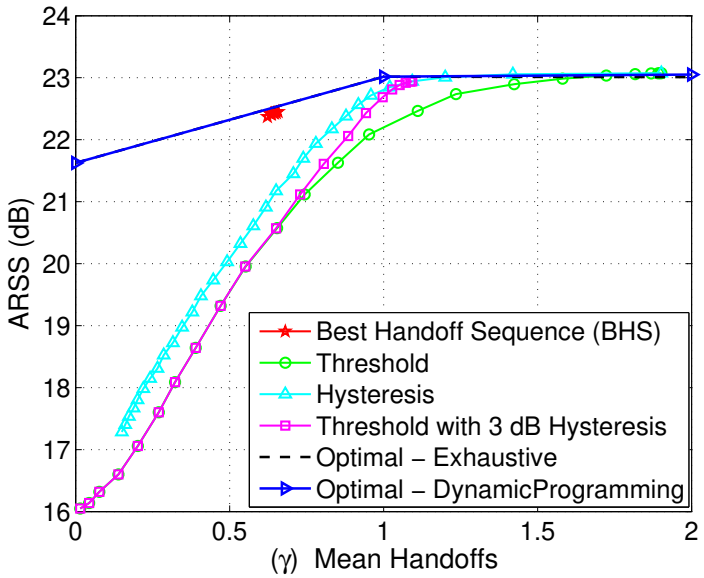


Figure 3.7: Comparison of the BHS, exhaustive method and optimal dynamic programming solution

Table 3.3: Comparison of exhaustive method and optimal dynamic programming solution

Evaluation Method	No. of Handoff (γ)		
	γ **= 0**	γ **= 1**	γ **= 2**
Exhaustive Method	21.7268	23.0351	23.0320
Dynamic Programming	21.7268	23.0351	23.0702

the difference in average signal quality between our heuristics and the optimal solution achieved by exhaustive search was never more than 0.48% for the 1000 cases we considered. Fig. 3.7 and Table 3.3 show those comparisons. We note that the results received from the exhaustive and Dynamic Programming methods are almost identical as expected.

3.8.3 Analysis of Call Dropping Probability

It is of interest to observe how CQSL varies with p. Figure 3.10 shows how it varies for the handoff method *Hysteresis*. As p is related, it is also valuable to observe how the dropping probability varies.

According to Fig. 3.8 and Fig. 3.9 dropping probability increases when the number of consecutive samples (d) decreases or the mean number of handoffs increases. Approximately after 1.4 mean handoffs, the dropping probability increases rapidly.

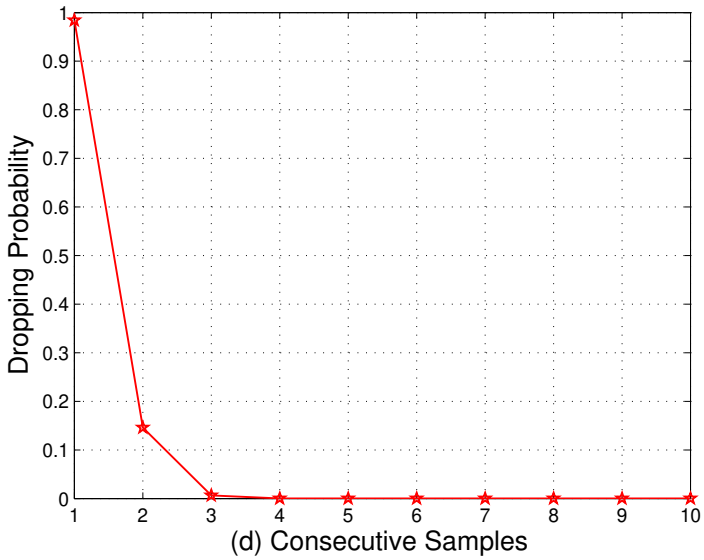


Figure 3.8: Dropping probability versus number of consecutive samples (d)

3.8.4 Analysis of Real Data found in BALI-2

Figure 3.13 shows a realistic scenario in which both new and handoff calls appear over 24 hours (1440 minutes) using the data base BALI-2 (Stanford University Mobile Activity Traces) [1]. BALI-2: Bay Area Location Information (real-time) dataset records the mobile users' moving and calling activities in a day. We extract 50577 users from BALI-2 to analyze impact of handoff, in a real system. It is interesting to observe that both handoff calls and new calls are higher around the lunch time, and less around early morning hours. It also shows that most users in that database need around 4 handoffs per day. This database was created using the real traffic data in San Fransisco Bay area shown in Fig. 3.11. A more simple data set is created for our simulations to illustrate the use of BHS as a benchmark.

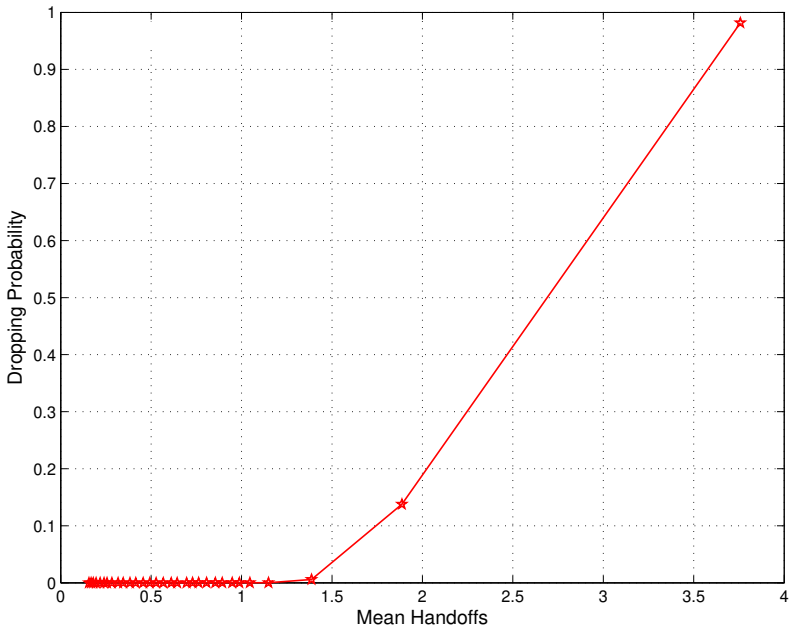


Figure 3.9: Dropping probability versus mean number of handoffs

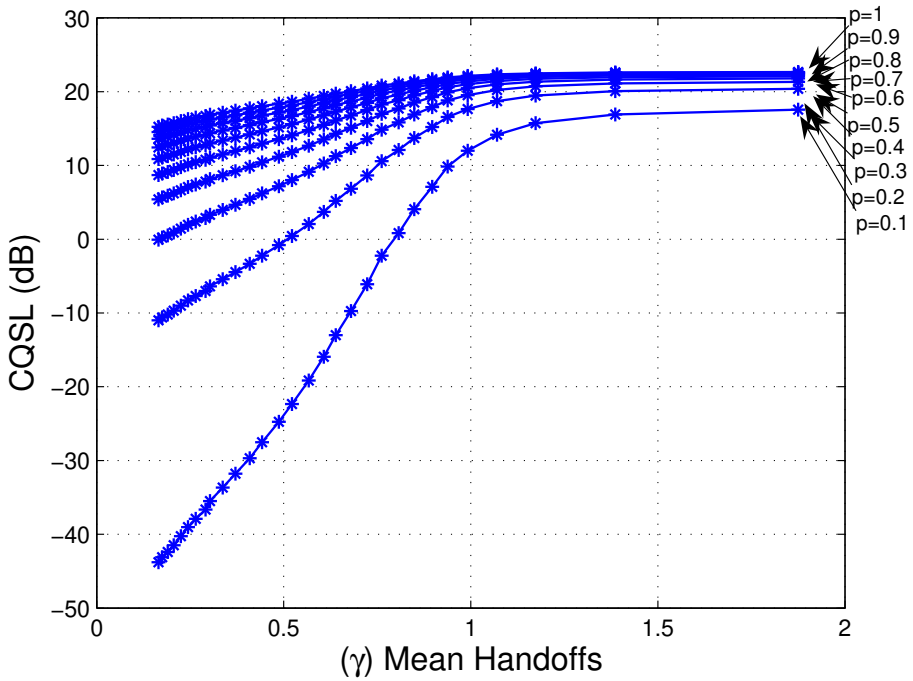


Figure 3.10: CQSL for different p for hysteresis handoff method, where p is the maximum allowed proportion of sample points with signal quality below S_{min}

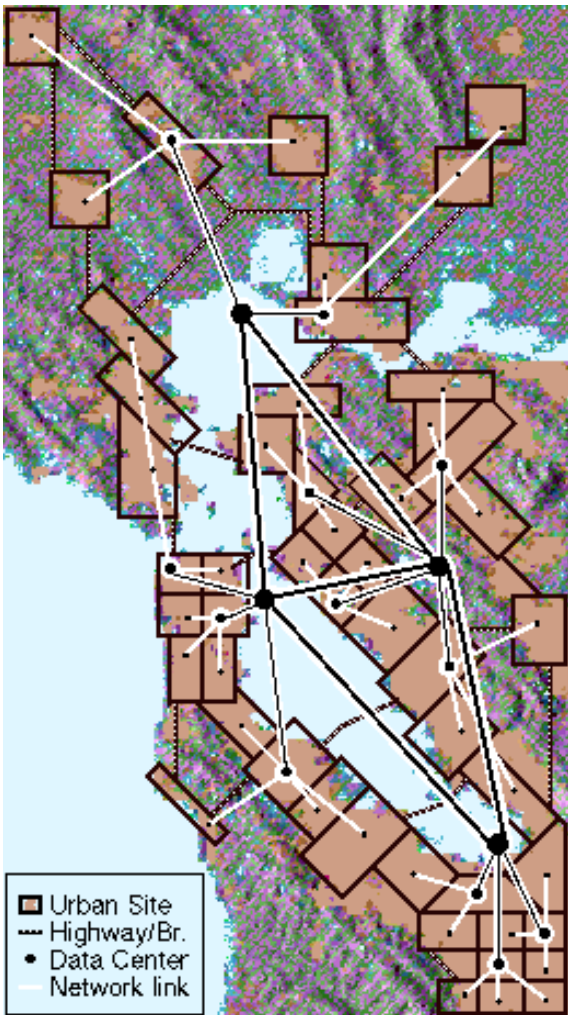


Figure 3.11: Map of San Francisco bay area

3.8.5 How BHS Can Be Used for Pattern Recognition Based Handoff

In cellular or micro cellular environments in cities, users move on predetermined paths such as roads and sidewalks. As the buildings and trees remain static, received signal strengths at a point on such a path will not fluctuate much. Considering sample points located on a straight line perpendicular to the road, it is estimated that received signal strengths belong to the same distribution [192]. This regularity is not exploited in current handoff methods. In order to use this regularity, the signal strengths need to be measured along sample points in all predetermined paths such as roads. The most suitable base station assignment at

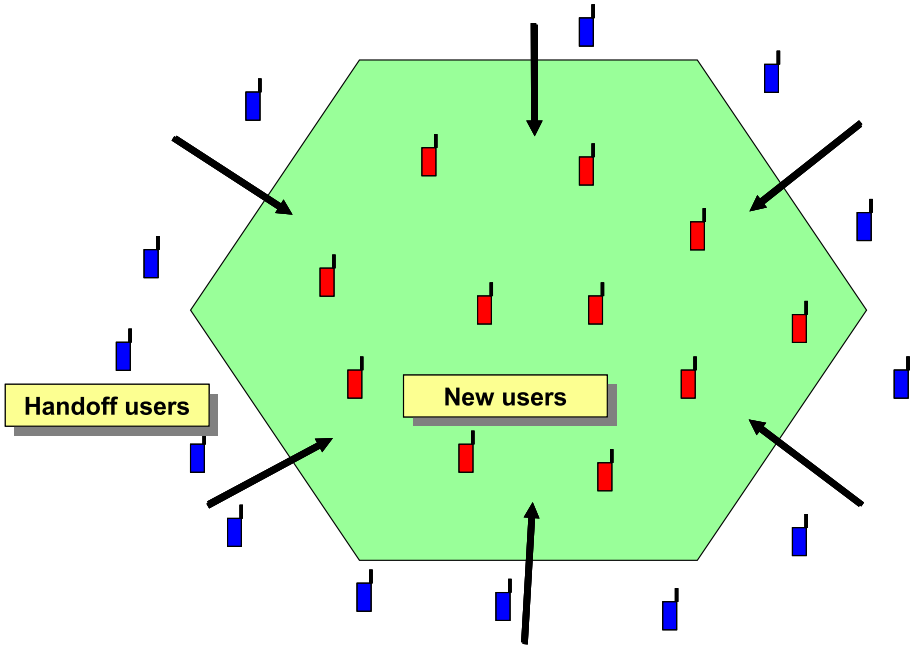


Figure 3.12: New users and handoff users, at a single cell

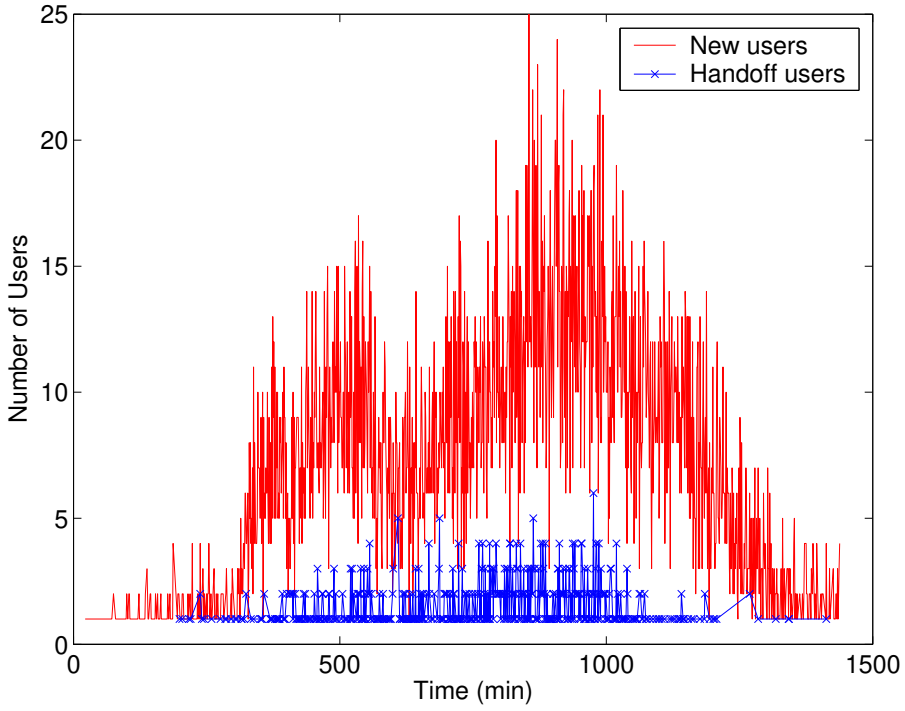


Figure 3.13: Ratio of the number of new users to handoff users at a single cell. (Analysis of real data found in BALI-2: Bay Area Location Information dataset)

each sample point should be determined considering handoff costs and QoS parameters. The most suitable sequence of assigned base stations or the best handoff sequence can provide the basis for pattern recognition based handoff methods.

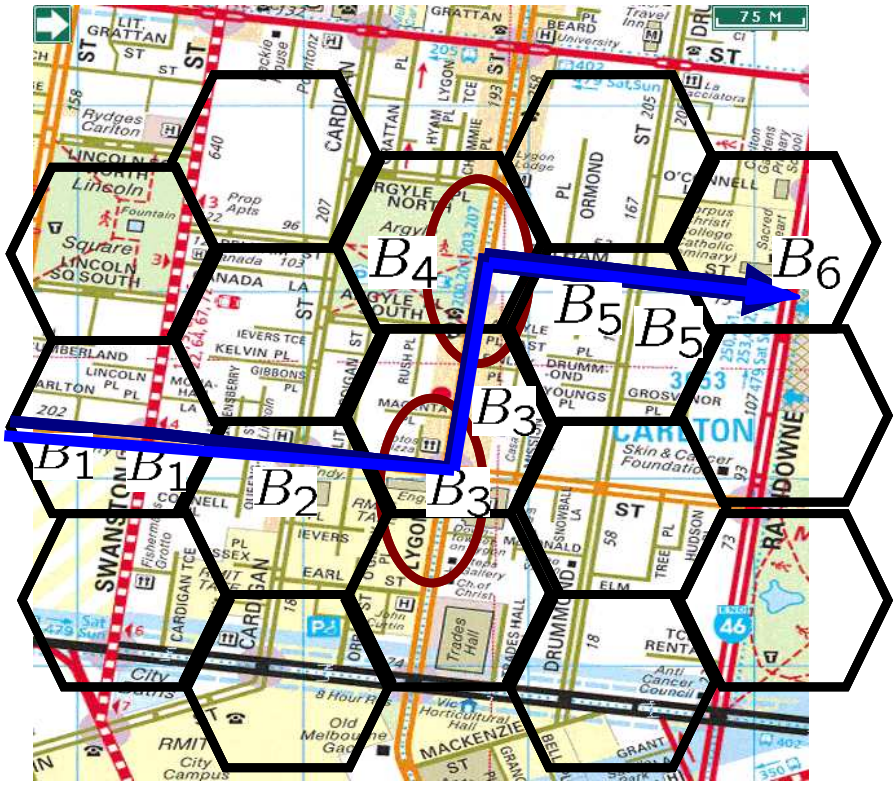


Figure 3.14: How BHS can be used for pattern recognition based handoff [Map is taken from Melway, Australia]

For example, consider the canonical case with 10 sample points involving only two base stations B_1 and B_2. When the user is moving from B_1 to B_2, the ideal or the best handoff sequence could be: $\{B_1, B_1, B_1, B_2, B_2, B_2, B_2, B_2, B_2, B_2\}$. The sequence entry at each sample indicates the serving base station at that point. When the cost function for the sequences is known, the ideal or best sequence that optimizes the cost function can be found.

Template matching is often used as a pattern recognition method. BHS provides a near optimal handoff sequence for a given path. BHS of various paths can be used as templates for determining the handoff sequence for a path. Once a matching template is found, the next base station can be predicted. Therefore, pattern recognition with BHS can be used to determine the path of the user.

3.9 Chapter Summary

A signal level based criterion exists for the evaluation of handoff algorithms. We have proposed a new call quality measure, considering existing measures and call dropping probability.

This chapter has developed an off-line cluster-based computationally-simple heuristic algorithm to find a near optimal handoff sequence that can be used as a benchmark. It has shown that BHS is near optimal by using exhaustive and dynamic programming methods. These results show that there is substantial room for improvement in existing handoff algorithms with respect to the signal level measures as well as number of required handoffs. This chapter has also shown that BHS can be used as a reference for pattern recognition based handoff.

Handoff in current wireless cellular systems is commonly achieved through hysteresis and threshold based methods. All such methods are centralized and managed by the base station controller assisted by the mobile station and the base station. Increased integration in electronic hardware makes it possible to include many complex features in mobile stations. Therefore, it is timely and useful to develop handoff algorithms that can be managed or processed by mobile terminals.

Several important conclusions arise from this chapter:

- A more efficient Call Quality Signal Measure can be derived considering existing measures and call dropping probability
- The proposed near optimal Best Handoff Sequence is a computationally simple yet accurate enough solution for obtaining the best handoff sequence (to be used as a benchmark) which would otherwise be obtained by exhaustive or dynamic programming methods.
- Simulations show that the Hysteresis method is the best handoff method.
- Analysis of real data by extracting about 50,000 users shows that there is a impact of handoff users in a real system.

- The proposed near optimal Best Handoff Sequence can be used for pattern recognition based handoff.

Following these conclusions the next chapter considers the use of this new approach for performance evaluation and comparison between existing handoff algorithms.In that evaluation both the retrial option where repeated call attempts are made and the non-retrial call option are considered.

CHAPTER FOUR

Handoff Performance Evaluation

4.1 Introduction

Evaluating the performance of handoff methods is particularly important because it can be used to compare existing and future handoff procedures, so that telecommunications providers can select the best handoff algorithm to optimize handoff management functions.

In Chapter 3, we attempt to model the Call Quality Signal Level (CQSL) measure, taking into consideration signal levels, unacceptable sample points and call dropping. However, to model call quality realistically one should consider other aspects such as the number of handoffs, and should differentiate between various levels of low signal strengths in calculating the penalties. This helps to increase the call quality and reduce call drop in inner city areas congested by many mobile users.

4.1.1 Motivation

Handoff costs are modeled in the literature as a constant cost per handoff, because of the difficulty in quantifying the cost. A comprehensive evaluation framework for handoff methods will reveal the tradeoff between handoff strategies for network designers. Ideally, such a framework would include the practical considerations of handoff cost, call dropping and associated penalties. In practice, different systems may have different drop rates, caused by either coverage problems because of irregular terrain configurations or inadequate channel availability. The

questions we try to answer are: What is the most suitable comprehensive framework to replace the existing way of comparing handoff methods? What is the best existing handoff method? What factors will influence the selection of a handoff? What options should we consider in case of a call drop? What impact does the system dropping probability have on the selection of handoff method? How do we select a suitable handoff method for a given terrain configuration? Clearly, the framework should have a performance measure and possibly an ideal solution for a given set of measurements that needs to be approximately achieved by handoff methods. It is important to evaluate the handoff methods considering the cost of a terminated call, and the possibility of attempts to reconnect. How can we estimate the optimal handoff sequence and how well do existing handoff methods perform in comparison to the estimated optimal handoff sequence?

4.1.2 Handoff Methods

A wireless/mobile environment gives rise to challenging design problems due to user movement and limited bandwidth on the wireless link. Performance evaluation and comparison of different handoff algorithms are needed to determine the appropriate handoff algorithm with respect to call quality and number of handoffs.

An analytical model is proposed in [117] to estimate the performance of handoff protocols using buffering policies at the base station. In [191, 192], the cost function for evaluating the handoff methods consists of two components: signal level based quality and the cost associated with the handoff. The proportional weight between these components was selected arbitrarily. In Chapter 3, we argued that the signal level based quality as well as the proportional weight given are not optimal. In [178], the cost function was defined as a linear combination of expected values of the number of service failures and the number of handoffs without indicating how the weight parameters between the two costs are found. The effect of user retrial in wireless communication has been studied in [103, 108, 118, 176, 181]. Some of these studies involve user behaviors [176] and were in-

spired by similar studies for wired networks [181].

Efforts have been reported to improve handoff efficiency in various ways [19, 28, 39, 55, 60, 83, 114, 115, 124] such as:

1. minimize the number of handoffs on a path;
2. maximize the call quality by maximizing the signal level or the received signal strength;
3. minimize the unnecessary handoffs in situations where the existing base station provides a signal strength above a drop level, or the selected new base station does not provide it;
4. minimize handoff delay caused by the complexity of the handoff algorithm.

Considering the fact that most mobile systems are interference limited, it is widely assumed that the received signal strength is an adequate indicator of call quality [20]. The tradeoff between the expected number of handoffs and expected number of drop level signal levels is considered in [178], suggesting a locally optimal handoff algorithm that uses the ideas of hysteresis and hysteresis with threshold handoff methods. The signal levels are considered either as drop level (below a certain signal level) or non-drop level. They indicate that the globally optimal handoff sequence that can be found by dynamic programming is too costly to implement, and depends on prior knowledge of the mobile user's path. The locally optimal solution is obtained by restricting the path of the user to two consecutive points.

It is argued in [178] that the first two criteria would lead to minimizing unnecessary handoffs. The handoff delay that may lead to up-link interference to other mobiles, and therefore, additional costs, is considered as a handoff selection criterion in [192].

Handoff probabilities are used in [203] as part of a proposed analytical model to compare threshold plus hysteresis handoff methods with various thresholds and hysteresis values. According to simulations carried out using a simple two base station scenario (canonical case), the model can be used as a design tool.

However, it is yet to be validated in more complex scenarios involving multiple base stations.

In this chapter, we compare various handoff methods using a performance measure based on a new Call Quality Measure that reflects practical call dropping considerations including signal based penalties and the number of handoffs. Simulations are carried out considering more complex scenarios than the canonical case.

4.1.3 Handoff Evaluation

Telecommunications providers should be able to choose the best handoff algorithm to optimize their handoff management functions. From the perspective of the user, it is essential to maintain low call failures (may be guaranteed or agreed) and achieve strong signal strengths for successful calls. From the perspective of the network operator, the priority is to minimize call failures, to provide good QoS and to minimize the cost.

We propose two measures to quantify the performance of handoff algorithms. The increase in call quality is quantified using the proposed Call Quality Signal Level (CQSL). Reduction in call quality leads to call drop, and therefore, it is integrated as a cost in CQSL. A nearly optimal benchmark solution, Best Handoff Sequence (BHS) is used for comparison of various handoff algorithms.

The remainder of this chapter is organized as follows. In Sections 4.1.4 and 4.1.5, we describe assumptions and definitions. In Section 4.2, we describe our handoff evaluation algorithm, Call Quality Signal Level. In Section 4.3, we extend our evaluation algorithm as ECQSL, by considering all practical considerations: signal level, call dropping and handoff cost. In Section 4.4, we apply this model for four alternatives. Then, in Section 4.5 we provide simulation results and discussion where we demonstrate the benefit of our handoff evaluation model. Section 4.6 concludes the chapter.

4.1.4 Assumptions

- We assume that channel capacity is unlimited and therefore, we do not consider handoff queuing.
- We assume a homogeneous network where all cells are identical in size, user mobility and cell coverage as in [77, 148].
- Each cell is assumed to have an equal number of neighbors as in [142].
- A log normal prorogation model is assumed and no power control is assumed to exist [160].
- Mobile users are uniformly distributed in the region [125].
- Users directions are randomly uniformly distributed over $[0, 2\pi]$ as in [125].
- We know the users' locations and base stations, and give information about signal strengths in each point.
- A call is dropped after observing the drop level signal level for a number of consecutive sample points [158].
- After a call is dropped there are two possible scenarios that could occur in a sample path;
  - in the retrial mode, another call will be placed after some time that may be required to place the call
  - in the non-retrial mode no such retrial attempt occurs.

The retrial model has a significant impact on the network performance, as it reflects the utilization of resources otherwise wasted. It also allows a fairer comparison between various handoff methods than the non-retrial model.

4.1.5 Definitions

Consider a cellular mobile network with M base stations designated $B_1, B_2, ..., B_M$. Define $\mathcal{B} = \{B_1, B_2, ..., B_M\}$. Let a sample path l be an arbitrary path in which

a mobile user is travelling. Consider a set of paths denoted Θ for the purpose of evaluating handoff algorithms. Sample points are points on the sample path for which the signal strength received from base stations are measured. Let S_{ij} be the signal strength at sample point i received from base station B_j. A handoff sequence x or $x(l)$ for sample path l, is defined as a sequence of base stations assigned to the sample points in l, assigning $b_i \in \mathcal{B}$ to the i^{th} sample point, i.e., $x =< b_1, b_2, ..., b_N >$ where N is the number of sample points. (It is possible that b_i and b_j, $\forall i, j$ may designate the same base station.)

For every sample path, the set of all possible handoff sequences is defined as $X = \{x \in \mathcal{B}^N\}$. The number of handoffs $\gamma(x)$ in a handoff sequence x equals the number of changes in the base station sequence. For example, the handoff sequence $x = \{B_1, B_1, B_2, B_3, B_3, B_3\}$ has $\gamma(x) = 2$.

For a given handoff sequence $x \in X$, the signal at i^{th} sample point for handoff sequence x, $S_i(x) = S_{ij}$ is defined such that $B_j = b_i$, base station used at sample point i. Let S_{min} be the minimum signal strength below which the signal quality is unacceptable to the user. Let $S_{max} > S_{min}$ be the signal strength beyond which the marginal benefit is considered negligible and $S_{drop} < S_{min}$ is the dropping signal level below which the call is dropped, if that level is maintained for a certain period.

Let $N_g(x) = \{i | S_i(x) \geq S_{min}\}$, and $N_b(x) = \big(N - |N_g(x)|\big)$ the number of samples with signal strength lower than S_{min}, where $|\mathrm{Y}|$ denotes the number of elements (cardinality) in the set Y.

The signal based penalty is the penalty that differs with various levels of unacceptable or low signal strength which is less than S_{min}. For example, there is a penalty if signal level is $S_{drop} < S_i(x) < S_{min}$, a higher penalty if signal level is below S_{drop}, and a much higher penalty with an increasing number of consecutive sample points.

In Section II, we describe initial approaches to the problem, and in Section III, we describe our handoff evaluation algorithm. Simulation results and discussion are presented in Section IV and, finally, a chapter summary is given in Section V.

Figure 4.1 shows the previously proposed CQSL and a new suggestion de-

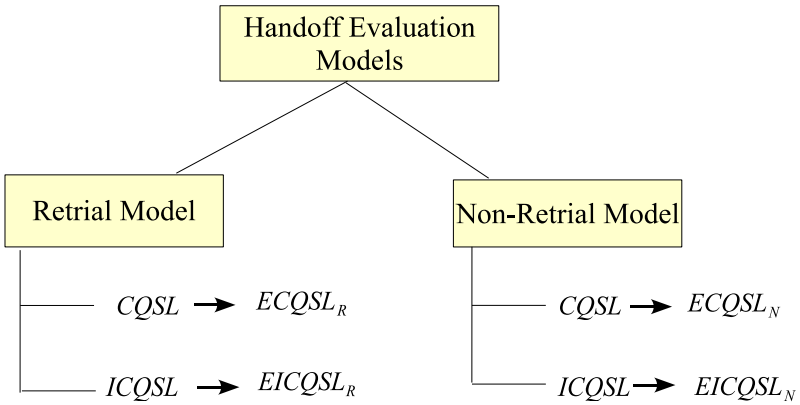


Figure 4.1: Handoff evaluation measures, CQSL, ICQSL and ECQSL, under the retrial and non-retrial models

scribed in Section 4.2.2 as ICQSL. For both retrial and non-retrial models, CQSL and ICQSL measures can be further extended as shown later.

Extensions to CQSL and ICQSL for retrial and non-retrial models are given as ECQSL_R, ECQSL_N and EICQSL_R and EICQSL_N respectively.

4.2 Call Quality Signal Level (CQSL)

4.2.1 Previously Proposed CQSL

The concept of Call Quality Signal Level (CQSL(x)) proposed in [68] uses a combination of the following signal quality measures:

- Average Received Signal Strength **(ARSS(x))** is defined by $\frac{1}{N}\sum_{i=1}^{N} S_i(x)$.

- Number of Acceptable Sample Points **(NASP(x))** represents the number of sample points of the handoff sequence with signal strength above S_{min}. Here NASP$(x) = |N_g(x)|$, where $|\text{Y}|$ denotes the number of elements (cardinality) in the set Y.

The CQSL(x) associated with a handoff sequence x of a path l is the average signal strength of acceptable sample points minus the penalty assigned to unacceptable sample points on that path, i.e.:

$$\mathrm{CQSL}(x) = \frac{\sum_{i \in N_g(x)} A_i(x)}{|N_g(x)|} - CN_b(x), \tag{4.1}$$

where $\forall_i \in N_g(x)$

$$A_i(x) = \begin{cases} S_i(x) & \text{if } S_i(x) \leq S_{max} \\ S_{max} & \text{otherwise,} \end{cases}$$

and C is the cost (or the penalty) for an unacceptable sample point. We assign $\sum_{i \in N_g(x)} A_i(x)/|N_g(x)|$ to zero when $|N_g(x)| = 0$.

Let p be the maximum allowed proportion of sample points (N) with signal quality below S_{min}, i.e., $N_b(x)/N \leq p$. The p value may be agreed between the service provider and the user. Assuming $|N_g(x)| \neq 0$, the minimum value that $\mathrm{CQSL}(x)$ can take is when (i) $N_b(x)/N = p$ and (ii) $\sum_{i \in N_g(x)} A_i(x)/|N_g(x)| = S_{min}$ in (4.1). We choose C such that the value of the proposed measure is greater or equal to zero. It is equivalent to setting a bound on C as follow:

$$C \leq \frac{\sum_{i \in N_g(x)} A_i(x)/|N_g(x)|}{N_b(x)} = \frac{S_{min}}{pN}. \tag{4.2}$$

Here we choose the cost to be linear with $N_b(x)$. Using (4.1) and (4.2), we can obtain

$$\mathrm{CQSL}(x) \geq \frac{\sum_{i \in N_g(x)} A_i(x)}{|N_g(x)|} - \frac{S_{min} N_b(x)}{pN}. \tag{4.3}$$

However, the above CQSL measure does not effectively distinguish between two sequences with the same average signal strength of good sample points, where one has a large number of good sample points with a relatively small signal strength, and another has only a few good sample points but with a large signal strength. In the next section the CQSL is improved by considering these problems.

4.2.2 Improved Call Quality Signal Level (ICQSL)

We slightly modify the CQSL measure (refer to as Improved CQSL or ICQSL) by deducting the penalty before getting the average as follows:

$$\text{ICQSL}(x) = \frac{1}{N}\left\{\sum_{i\in N_g(x)} A_i(x) - CN_b(x)\right\}. \tag{4.4}$$

As previous we choose C such that the minimum possible value is equal to zero. The parameter C in (4.4) can be bounded as follows:

$$C \leq \frac{\sum_{i\in N_g(x)} A_i(x)}{N_b(x)} = \frac{S_{min}|N_g(x)|}{pN}, \tag{4.5}$$

where $S_{min}|N_g(x)| \leq \sum_{i\in N_g(x)} A_i(x) \leq S_{max}|N_g(x)|$ and $N_b(x)/N \leq p$.

Using (4.4) and (4.5), we can obtain the lower bound as

$$\text{ICQSL}(x) \geq \frac{\sum_{i\in N_g(x)} A_i(x)}{N} - \frac{S_{min}N_b(x)|N_g(x)|}{pN^2}. \tag{4.6}$$

Furthermore, in Section 4.2.3, we try to investigate suitability of the both measures by evaluating minimum and upper bounds of the each proposed measures.

4.2.3 Reference Values for CQSL and ICQSL Measures with Numerical Example

We define a worst case reference level, where we can obtain minimum CQSL value when $N_b(x) = pN$. This should guarantee a CQSL better than the worst case. By using this reference level, then, we can evaluate the suitability of both measures. Note that, by taking difference reference levels we can evaluate different CQSL value ranges, but, here we obtain worst case values. Figure 4.2 shows the comparison of the two measures. It is preferable to have a smaller and positive range for the reference level CQSL $= 0$.

At the reference level, from equation (4.3) and when $N_b(x) = pN$, we can

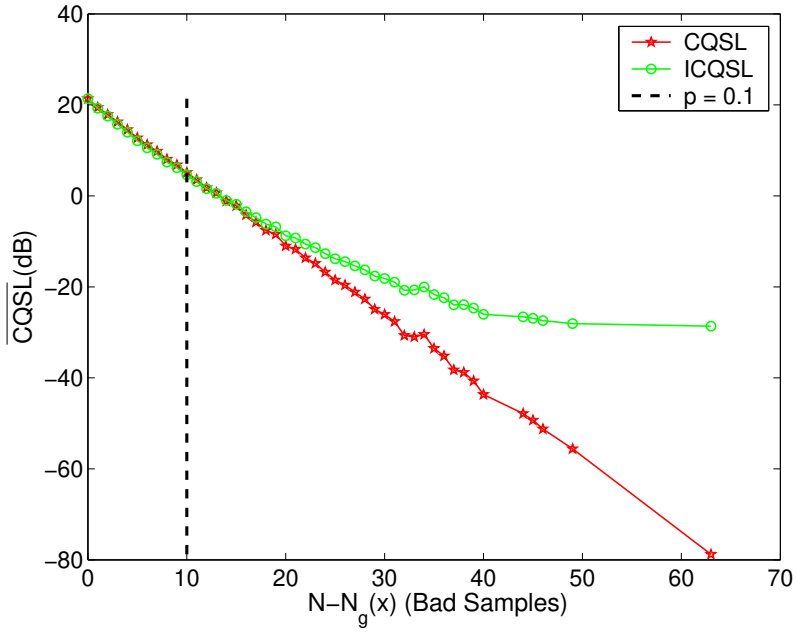


Figure 4.2: Different handoff evaluation measures, $\overline{\text{CQSL}}$, versus Bad Samples $(N - N_g)$ when $p = 0.1$

observe that

$$\text{CQSL}(x) \geq \frac{\sum_{i \in N_g(x)} A_i(x)}{|N_g(x)|} - S_{min}. \tag{4.7}$$

To obtain lower and upper bound for CQSL measures, we consider the minimum and maximum of the individual terms. The first term in (4.7) corresponds to the average of good sample points with signal strengths $S_i(x) \geq S_{min}$, therefore we have

$$S_{min} \leq \frac{\sum_{i \in N_g(x)} A_i(x)}{|N_g(x)|} \leq S_{max}. \tag{4.8}$$

Finally, we can obtain the possible range for the reference level when CQSL $= 0$.

$$0 \leq \text{CQSL}^{\text{ref point}}(x) \leq S_{max} - S_{min}. \tag{4.9}$$

When, $S_{min} = 15$, $S_{max} = 1.5S_{min}$, $p = 0.1$, $N = 20$,

$$0 \leq \text{CQSL}^{\text{ref point}}(x) \leq 7.5. \tag{4.10}$$

Similarly at the reference level, from equation (4.6) and when $N_b(x) = pN$, we can observe that

$$\text{ICQSL}(x) \leq \frac{\sum_{i \in N_g(x)} A_i(x)}{N} - \frac{S_{min}\ |N_g(x)|}{N}. \tag{4.11}$$

To obtain lower and upper bounds for ICQSL, we consider minimum and maximum of the individual terms:

$$\frac{S_{min}|N_g(x)|}{N} \quad \leq \quad \frac{\sum_{i \in N_g(x)} A_i(x)}{N} \leq \frac{S_{max}|N_g(x)|}{N}. \tag{4.12}$$

By substituting $N_g(x) = N - N_b(x)$ and $N_b(x) = pN$ which is the definition of the reference level, we can obtain

$$0 \quad \leq \quad \text{ICQSL}^{\text{ref point}}(x) \leq (S_{max} - S_{min})(1 - p).$$

Considering the same numerical values as before, we obtain

$$0 \leq \text{ICQSL}^{\text{ref point}}(x) \leq 6.75. \tag{4.13}$$

According to equations (4.10) and (4.13), we can use either $\text{CQSL}(x)$ or $\text{ICQSL}(x)$, because both of them have non negative margin range for the reference (minimum but still acceptable) signal quality. It can be observed that there is not a big difference in values between (4.10) and (4.13) when p is small. Both provide acceptable and reasonable ranges for "zero" reference level.

In both options, we choose the cost to be linear with $N_b(x)$. However, we could differentiate between drop level ($S_i \geq S_{drop}$) sample points and an unacceptable, but still non drop level ($S_{min} \geq S_i \geq S_{drop}$) sample points. We may set a cost for drop level sample point dynamically to reflect the fact that consecutive drop level sample points are worse than a single drop level sample point. Such an extension will represent the scenario more realistically as described in Section 4.3.

4.3 Extended Call Quality Signal Level (ECQSL)

In this section we propose improvements to the CQSL and ICQSL presented in Section 4.2. The reduction in call quality leads to the call drop, and therefore it is integrated as a penalty in CQSL and ICQSL. Here we assume both retrial and non-retrial models and for the retrial model we assume that a handoff sequence may contain multiple droppings, i.e., that a dropped call is immediately replaced by another call when the drop occurs.

We propose in this section three extensions to the CQSL and ICQSL which we presented in Section 4.2:

- Differentiation of the penalties based on different levels of signal quality associated with $N_b(x)$,
- Introduction of higher penalties for consecutive sample points with signal strengths below a drop level,
- Inclusion of the handoff cost.

In order to take into account the different levels of quality impairment caused by unacceptable signal strengths, we define the cost (penalty) as a function of signal strengths as follows:

$$C(S_i(x)) = \begin{cases} C_1 & \text{if } 0 \leq S_i(x) \leq S_{drop} \\ \frac{C_1}{2}(J) & \text{if } S_{drop} \leq S_i(x) \leq S_{min} \\ 0 & \text{if } S_i(x) \geq S_{min}, \end{cases} \tag{4.14}$$

where C_1 is a predefined parameter and $J = \left[1 + \cos \pi \left(\frac{S_i(x) - S_{drop}}{S_{min} - S_{drop}}\right)\right]$. The above function is illustrated in Fig. 4.3.

4.3.1 No Penalty Region

The first term in (4.4), $\sum_{i \in N_g(x)} A_i(x)$, corresponds to sample points with acceptable signal strengths, and therefore there is no cost (penalty) involved. These sample points belong to the no penalty region as shown in Fig. 4.3.

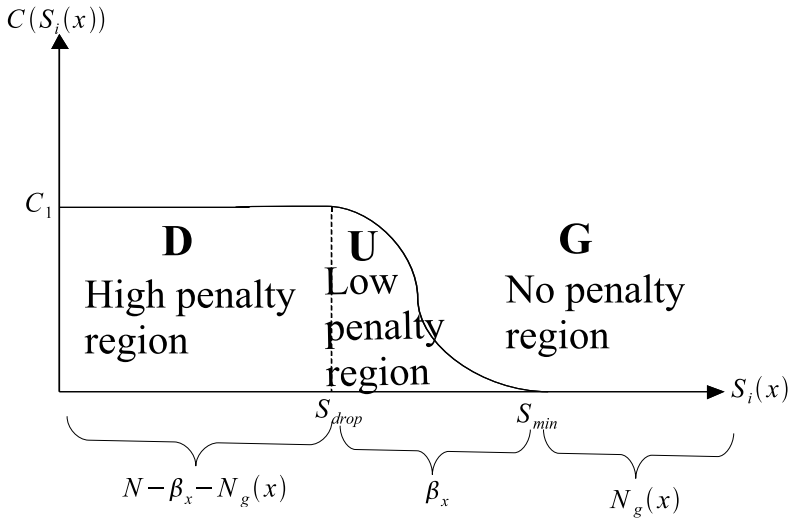


Figure 4.3: Penalty function $C(S_i(x))$ with signal strength $S_i(x)$, considering the different levels of quality caused by unacceptable signal strength

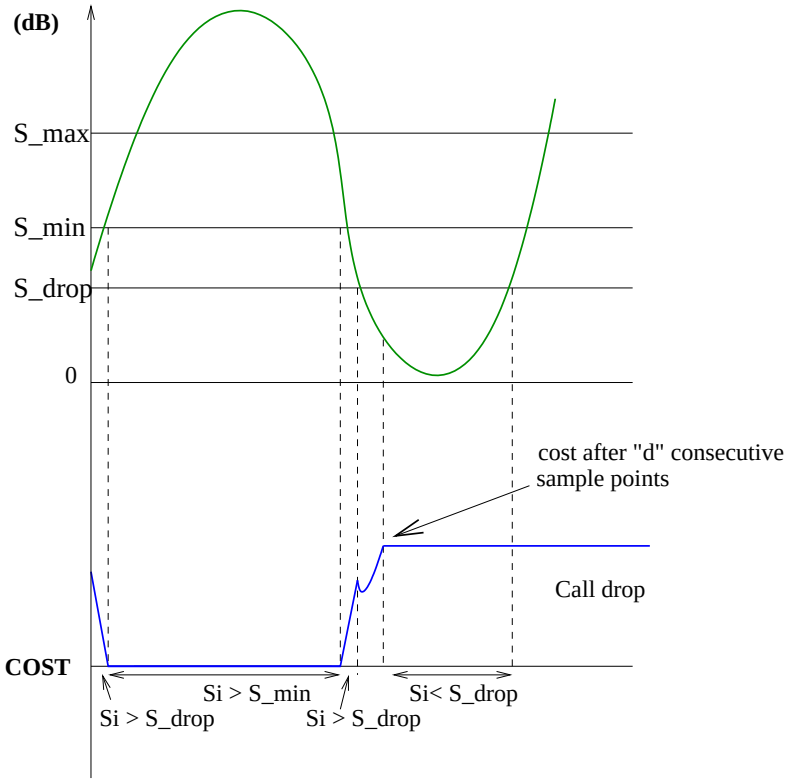


Figure 4.4: Call is dropped if the signal strength is below the call dropping level (S_{drop}), for d consecutive sample points

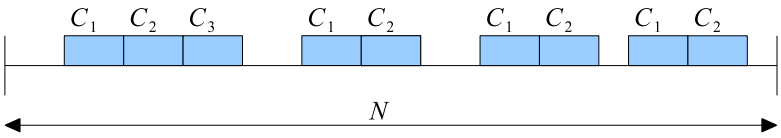


Figure 4.5: Example of the number of consecutive sample points associated with signal strength less than S_{drop}, assuming $d = 3$, $a = 1$ and (4.15)

4.3.2 Low Penalty Region

The second term in (4.4), $CN_b(x)$, corresponds to sample points with unacceptable signal strengths. We characterize these sample points into two groups. The first group consists of sample points with signal strengths between $S_{drop} < S_i(x) < S_{min}$ (low penalty region).

For any handoff sequence x the total cost associated with its sample points in the low penalty region is $\frac{C_1}{2}\left[1+\cos\pi\left(\frac{S_i-S_{drop}}{S_{min}-S_{drop}}\right)\right]$, where $\beta(x) = \{i|S_{min} > S_i(x) > S_{drop}\}$.

4.3.3 High Penalty Region

The second group is a set of sample points with signal strengths below a call dropping level ($0 \leq S_i(x) \leq S_{drop}$), which correspond to the high penalty region in Fig. 4.3. In any handoff sequence x, a call is dropped as explained in Fig. 4.4, if the signal strength is below the call dropping level (S_{drop}), for d consecutive sample points (dropping points). In practice these d consecutive sample points can be varied with user speed, therefore we assume users move with constant speed. The cost assigned for each of the sample points among these d dropping points is defined as follows

$$C_r = \begin{cases} a^{r-1}C_1, & \text{if } 2 \leq r \leq d-1, \\ a^{d-1}C_1, & \text{if } d \leq r \leq N, \end{cases} \tag{4.15}$$

where a is a scaling factor. One example, as shown in Fig. 4.5, is a sample path with three sets of two consecutive sample points and one set of three consecutive

sample points associated with signal strength less than S_{drop} assuming $d = 3$, $a = 1$ and (4.15). The cost associated for this sample path is: $4C_1 + 4C_2 + C_3$.

The parameter a is chosen such that the cost associated with the i^{th} dropping point, weighted by the probability that there are i consecutive dropping sample points, is equal to the weighted cost of the $(i+1)^{th}$ dropping point. The probability of having i consecutive dropping sample points in an arbitrary handoff sequence x is given by

$$p(i) = (N - i + 1)\delta^i(1 - \delta)^{N-i}, \tag{4.16}$$

where δ is the probability of receiving a signal strength below S_{drop}. Knowing $p(i)$, $\forall\, 1 \leq i \leq d$ the value of parameter a is given by assuming,

$$a = \frac{p(d-1)}{p(d)} + const = \left(\frac{N - d + 2}{N - d + 1}\right)\frac{(1-\delta)}{\delta} + const. \tag{4.17}$$

In practice $N \gg d$, and therefore we obtain $a = \frac{1-\delta}{\delta} + const$, which will be used as a scaling factor in our cost function. Considering that $a \geq 1$, and obvious choice of $a \Rightarrow 1$ when $\delta \Rightarrow 1$, therefore we can set $const = 1$.

Now we can write,

$$a = \frac{(1-\delta)}{\delta} + 1. \tag{4.18}$$

In practice, different systems may have different drop rates, caused by either coverage problems because of irregular terrain configurations or inadequate channel availability [101]. Therefore, here we consider the weighting factor, a, as fixed for a given terrain configuration. According to Fig. 4.6, we choose the weighting factor, a, once we know the dropping probability.

State Transition Diagram for Cost Model

We can model this cost value in a state transition diagram as shown in Fig. 4.7. We may also consider $d + 2$ possible states a sample point may belong to:

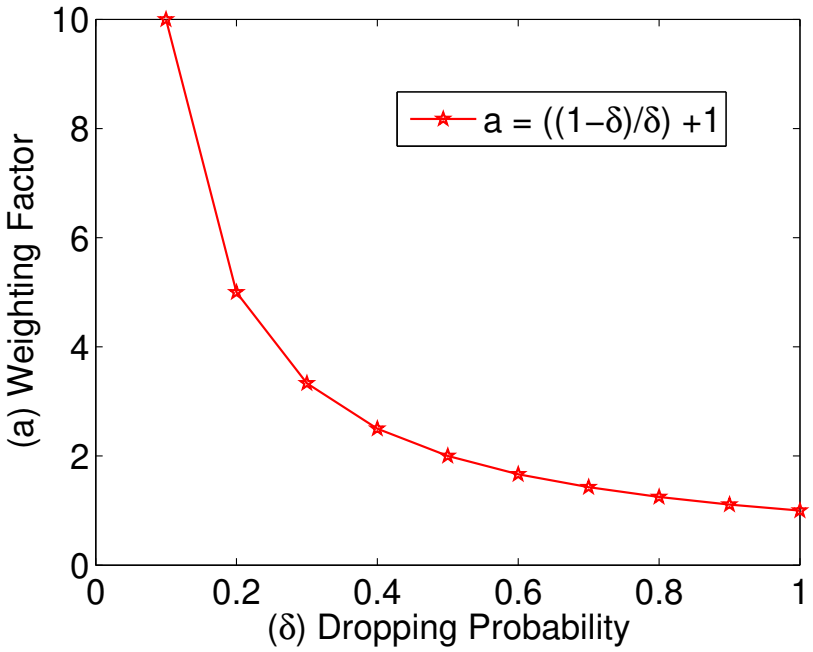


Figure 4.6: Weight factor a versus dropping probability δ. (Different systems would have different drop rates, caused by either coverage problems or inadequate channel availability. Therefore, we consider that the weighting factor, a is fixed for a particular terrain configuration)

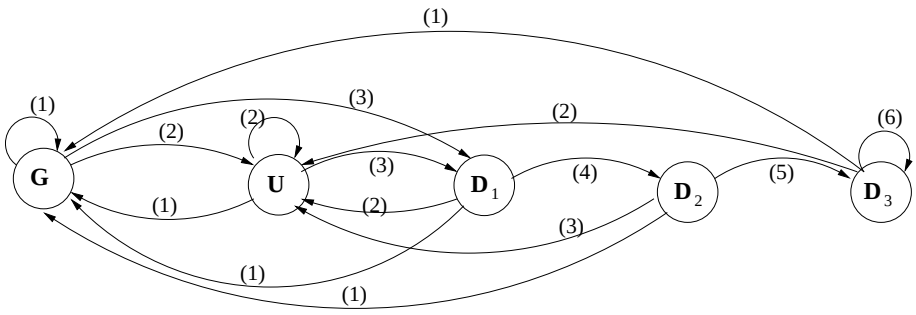


Figure 4.7: State transition diagram for cost model, for $d = 3$ consecutive sample points

- State G - a good (penalty free) with acceptable signal quality ($S_i \geq S_{min}$)
- State U - unacceptable but non drop level signal quality ($S_{drop} \leq S_i < S_{min}$)
- States $D_1..D_d$ - with drop level signal quality ($S_i < S_{drop}$) in blocks of 1 to d consecutive sample points.

We can identify $d+3$ different types of costs associated with state transitions. Figure 4.7 shows the estimate of the cost function associated with transition between different states when moving along a sample path assuming $d = 3$. Let C^{old} and C^{new} represent the costs before and after the state transition. We can define $d+3 = 6$ different costs associated with state transitions from (4.14), (4.15) and Fig. 4.7:

$(1) C^{new} = C^{old}$
$(2) C^{new} = C^{old} + \frac{C_1}{2}\left[1 + \cos\left(\frac{S_i - S_{drop}}{S_{min} - S_{drop}}\right)\right]$
$(3) C^{new} = C^{old} + C_1$
$(4) C^{new} = C^{old} - C_1 + 2aC_1$
$(5) C^{new} = C^{old} - 2aC_1 + 3a^2C_1$
$(6) C^{new} = C^{old} + a^2C_1.$

The cost component associated with each state transition is labeled in Fig. 4.7, and shown as a $(d+2) \times (d+2)$ matrix,

$$Cost = \begin{pmatrix} (1) & (2) & (3) & - & - \\ (1) & (2) & (3) & - & - \\ (1) & (2) & - & (4) & - \\ (1) & (2) & - & - & (5) \\ (1) & (2) & - & - & (6) \end{pmatrix}.$$

4.3.4 Handoff Cost

Another important issue is the cost for handoff. Therefore, finally, we include the handoff cost, C_h, as a linear function of the number of handoffs γ_x. In the following two sections we suggest the adaptation of the above three penalties to

retrial and non-retrial models. Let K denote the sample point at which the first call drop happens. Note that, for retrial and non-retrial models $N_g(x)$ can be varied as in following Table 4.1.

Table 4.1: $N_g(x)$ values used in retrial and non-retrial models

Model	$N_g(x)$
Retrial Model	$i = 1, ..., N$
Non-Retrial Model	$i = 1, ..., K$

In the Section 4.4.3 we propose how these extensions apply to CQSL and IC-QSL for retrial model and Section 4.4.5 for non-retrial model.

4.4 ECQSL for Four Cases: Retrial & Non-Retrial Models

In this section we explain derivations of ECQSL$_R$, and improved version, EICQSL$_R$, for retrial and non-retrial models.

4.4.1 Connection between C_1 and C_h

Considering that the maximum cost of having $(d-1)$ consecutive drops followed by an unacceptable but non drop level signal should be less than the cost of having d consecutive drops leading to the call drop, we derive:

$$\begin{aligned}
&\max \text{cost}(DD.....U) \leq \min \text{cost}(DD.....D) \\
&C_1 + \text{cost}(DD) + (d-1)C_h \leq \text{cost}(DD) + a^{d-1}C_1 \\
&\text{Therefore, } C_1 + (d-1)C_h \leq a^{d-1}C_1 \\
&C_h \leq \frac{(a^{d-1}-1)C_1}{(d-1)}.
\end{aligned} \tag{4.19}$$

4.4.2 Gamma Function for Retrial Model

In [118, 176, 181] various retrial models were proposed considering user retrial patterns. For example, it takes only a few seconds to redial a number that was

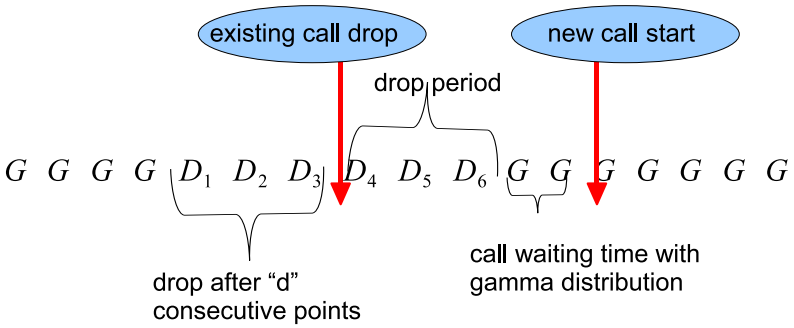


Figure 4.8: Call waiting time to connect next call just after call drop from previous existing call, with retrial model where repeated call attempts are made

not successfully connected in the previous attempt. In this model we assume that call discontinuation (drop) does not mean the end of consideration of the sample path for the proposed extended quality measure ECQSL_R including the handoff cost for handoff sequence x. A sample path may have multiple call drops. We will still consider good components between those call drops for the ECQSL_R and EICQSL_R. As the retrial behavior of users may vary, it is not straightforward to estimate the call waiting time after a call is dropped. In this work we employ gamma distribution to model the call waiting time just after call drop from existing call as described in Fig. 4.8. The advantage of using the gamma distribution function is that it uses only positive real numbers.

$$X_i \sim \text{Gamma}(\omega_i, \zeta), \tag{4.20}$$

where $i = 1, 2, ..N$, mean is $N_g\zeta$ and variance is $\omega\zeta^2$. We select $\omega = 2$ and $\zeta = 1$ based on the assumption that a user can retry and connect the call within a few seconds after call drop. The measures ECQSL_R and EICQSL_R can be used to differentiate between the various lengths of "call drop regions" in the sample path.

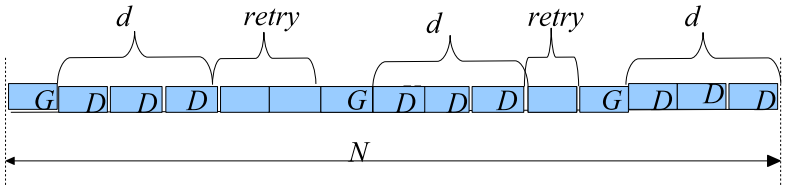


Figure 4.9: Retrial model where repeated call attempts are made. (Here G represents a good sample point where $S_i > S_{min}$ and D a dropping sample point where $S_i < S_{drop}$, with $d = 3$)

4.4.3 ECQSL$_R$ Measure for the Retrial Model

The extended expression of (4.1) for CQSL using this cost function is given by

$$\begin{aligned}\text{ECQSL}_R(x) &= \frac{\sum_{i\in N_g(x)} A_i(x)}{|N_g(x)|} \\ &-\frac{C_1}{2}\left[\sum_{i\in\beta_x}\left(1+\cos\pi\left(\frac{S_i(x)-S_{drop}}{S_{min}-S_{drop}}\right)\right)\right] \\ &-\sum_{r=1}^{d_{max}} h_r \sum_{j=1}^{r} C_r - C_h\gamma_x, \end{aligned} \tag{4.21}$$

where $\beta(x) = \{i | S_{min} > S_i(x) > S_{drop}\}$, d_{max} is the largest number of consecutive dropping points in a sample path, and h_r is the number of r consecutive dropping points in the same sample path.

The constant C_1 in (4.21) is chosen such that the lower bound of ECQSL$_R(x)$ is never less than zero. In order to find the lower bound of ECQSL$_R(x)$ we consider each individual term in (4.21). The first term in (4.21) corresponds to good sample points with signal strengths $S_i(x) > S_{min}$. Therefore we have

$$S_{min} \leq \frac{\sum_{i\in N_g(x)} A_i(x)}{|N_g(x)|}. \tag{4.22}$$

The maximum penalty for the retrial model occurs when the retrial is used for the maximum number of times or $N/(d+1)$ as in the following sequence with $d = 3$ (Fig. 4.9).

We assume that a retrial can only occur with a good level signal (G). Therefore, we can obtain: $N_g = \left\lceil \frac{N}{d+1} \right\rceil$ and $\gamma = \left\lceil \frac{N}{d+1} \right\rceil d$, and derive:

$$\text{Max penalty} = \underbrace{C_h \left\lceil \frac{N}{d+1} \right\rceil d}_{handoff} + \underbrace{\left(\frac{a^d - 1}{a - 1} \right) C_1 \left\lceil \frac{N}{d+1} \right\rceil}_{dropping}. \tag{4.23}$$

The next three terms are interdependent as bad sample points are distinguished between those sample points with $S_{drop} \leq S_i(x) < S_{min}$, and those sample points where there are consecutive points with $S_i(x) < S_{drop}$.

The lowest of these three negative terms in (4.21) is the maximum cost which corresponds to a case when all the bad sample points fall within the high penalty region, and is given by $\left\lceil \frac{N_b(x)}{d} \right\rceil (a^0 C_1 + a^1 C_1 + a^2 C_1 +, .., + a^{d-2} C_1 + a^{d-1} C_1) = \left\lceil \frac{N_b(x)}{d} \right\rceil C_1 \left(\frac{a^d - 1}{a - 1} \right)$. We assume that handoff should have a lower penalty than the dropping. Thus

$$\begin{aligned} - \quad & C_h \left\lceil \frac{N}{d+1} \right\rceil d - \left(\frac{a^d - 1}{a - 1} \right) C_1 \left\lceil \frac{N}{d+1} \right\rceil \leq \\ - \quad & \frac{C_1}{2} \left[\sum_{i \in \beta_x} \left(1 + \cos \pi \left(\frac{S_i(x) - S_{drop}}{S_{min} - S_{drop}} \right) \right) \right] \\ - \quad & \sum_{r=1}^{d_{max}} h_r \sum_{j=1}^{r} C_r - C_h \gamma_x. \end{aligned} \tag{4.24}$$

From (4.21), (4.22) and (4.24), we obtain

$$S_{min} - \left\lceil \frac{N}{d+1} \right\rceil C_1 \left(\frac{a^d - 1}{a - 1} \right) - C_h \left\lceil \frac{N}{d+1} \right\rceil d \leq ECQSL_R(x).$$

The C_1 constant is then determined by setting the lower bound value to zero. Therefore we can compute C_1 for ECQSL$_R$ as,

$$C_1 = \frac{S_{min}}{\left\lceil \frac{N}{d+1} \right\rceil \left[\left(\frac{a^d - 1}{a - 1} \right) + \frac{(a^{d-1} - 1)}{(d-1)} d \right]},$$

and, handoff cost C_h, for ECQSL$_R$ as

$$C_{h1} = C_h = \frac{S_{min}(a^{d-1}-1)}{\left\lceil \frac{N}{d+1} \right\rceil (d-1)\left[\left(\frac{a^d-1}{a-1}\right) + \frac{(a^{d-1}-1)}{(d-1)}d\right]}. \tag{4.25}$$

4.4.4 EICQSL$_R$ Measure for the Retrial Model

The extended expression for ICQSL, (4.4) cost function using the retrial model is given by

$$\begin{aligned}\text{EICQSL}_R(x) = \frac{1}{N}\Bigg\{ & \sum_{i\in N_g(x)} A_i(x) \\ & -\frac{C_1}{2}\left[\sum_{i\in\beta_x}\left(1+\cos\pi\left(\frac{S_i(x)-S_{drop}}{S_{min}-S_{drop}}\right)\right)\right] \\ & -\sum_{r=1}^{d_{max}} h_r \sum_{j=1}^{r} C_r - C_h\gamma_x\Bigg\}. \end{aligned} \tag{4.26}$$

Similarly to Section 4.4.3, the constant C_1 in (4.26) is chosen such that the lower bound of EICQSL$_R(x)$ is never less than zero. The first term in (4.40) corresponds to good sample points with signal strengths $S_i(x) > S_{min}$, therefore we have

$$|N_g(x)|S_{min} \leq \sum_{i\in N_g(x)} A_i(x). \tag{4.27}$$

From (4.24), (4.26) and (4.27) we obtain

$$\begin{aligned}\frac{1}{N}\Bigg[& |N_g(x)|S_{min} - \left\lceil \frac{N}{d+1} \right\rceil C_1\left(\frac{a^d-1}{a-1}\right) \\ & -C_h\left\lceil \frac{N}{d+1} \right\rceil d\Bigg] < \text{EICQSL}_R(x). \end{aligned}$$

As in Section 4.4.3, we consider the lower bound is equal to zero. Therefore we can compute C_1 and C_h for EICQSL$_R$ as,

$$C_1 = \frac{|N_g(x)|S_{min}}{\left\lceil \frac{N}{d+1} \right\rceil \left[\left(\frac{a^d-1}{a-1} \right) + \frac{(a^{d-1}-1)}{(d-1)} d \right]},$$

and,

$$C_{h2} = C_h = \frac{|N_g(x)|S_{min}(a^{d-1}-1)}{\left\lceil \frac{N}{d+1} \right\rceil (d-1) \left[\left(\frac{a^d-1}{a-1} \right) + \frac{(a^{d-1}-1)}{(d-1)} d \right]}. \tag{4.28}$$

In the next Section we propose how extensions proposed in this section apply to CQSL and ICQSL for non-retrial model which has only one attempt.

4.4.5 ECQSL$_N$, EICQSL$_N$ for the Non-retrial Model

Here we clearly explain both ECQSL$_N$ and EICQSL$_N$ for non-retrial model.

4.4.6 Measure for the Non-retrial Model, ECQSL$_N$

In this model, we assume that call discontinuation (drop) mean the end of consideration of the sample path for the proposed ECQSL$_N$ estimation. We are only interested in a sample path until the first call drop. We will not consider good components between call drops. The ECQSL$_N$ will not be able to differentiate between the various lengths of "call drop regions" in the sample path.

$$\begin{aligned} \text{ECQSL}_N(x) &= \frac{\sum_{i \in N_g(x)} A_i(x)}{|N_g(x)|} \\ &- \frac{C_1}{2} \left[\sum_{i \in \beta_x} (1 + \cos \pi \left(\frac{S_i(x) - S_{drop}}{S_{min} - S_{drop}} \right) \right] \\ &- \sum_{r=1}^{d_{max}} h_r \sum_{j=1}^{d-1} C_r - C_k(N-K) - C_h \gamma_x, \end{aligned} \tag{4.29}$$

where $1 < K < N$ and C_k is a constant cost associated with the first call drop at the K^{th} sample point, therefore giving an opportunity for the longest calling. Note that $K = 0$ is not valid due to the assumption $|N_g(x)| \geq 1$.

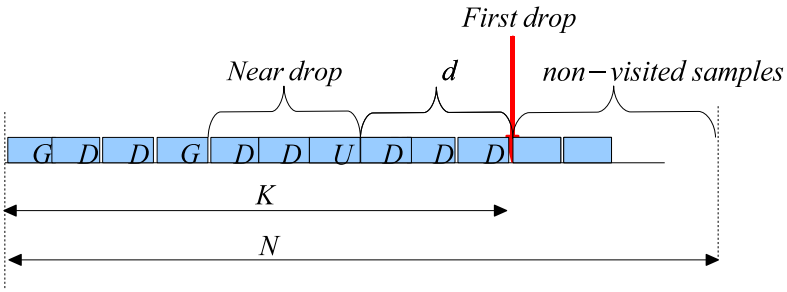


Figure 4.10: First call drop at the K^{th} sample point, under the non-retrial model

$N_g = 1 \quad \gamma = (K-1)$

The minimum of these three negative terms in (4.29) is the maximum cost which corresponds to a case when all the bad sample points fall within the high penalty region, and is given by

$$\begin{aligned}\underbrace{\text{Max penalty}}_{non-retrial} &= \underbrace{C_h(K-1)}_{handoff} \\ &+ \underbrace{\left(\frac{a^{d-1}-1}{a-1}+1\right) C_1 \left\lceil \frac{K-d-1}{d} \right\rceil}_{near-drop} \\ &+ \underbrace{\left(\frac{a^{d}-1}{a-1}\right) C_1}_{drop} + \underbrace{C_k \left\lceil \frac{N-K}{d} \right\rceil}_{after\ K^{th}\ point}. \end{aligned} \tag{4.30}$$

The penalty associated with (4.29) should be less than the maximum penalty calculated in (4.30). Therefore,

$$\begin{aligned} &-C_h(K-1) - \left(\frac{a^{d-1}-1}{a-1}+1\right) C_1 \left\lceil \frac{K-d-1}{d} \right\rceil \\ &- \left(\frac{a^{d}-1}{a-1}\right) C_1 - C_k \left\lceil \frac{N-K}{d} \right\rceil \\ &\leq -\frac{C_1}{2}\left[\sum_{i\in\beta_x} (1+\cos\pi\left(\frac{S_i(x)-S_{drop}}{S_{min}-S_{drop}}\right)\right] \\ &- \sum_{r=1}^{d_{max}} h_r \sum_{j=1}^{d-1} C_r - C_k(N-K) - C_h\gamma_x. \end{aligned} \tag{4.31}$$

By substituting (4.22) and (4.31) in (4.29) we obtain

$$S_{min} - C_h(K-1) - \left(\frac{a^{d-1}-1}{a-1}+1\right) C_1 \left\lceil \frac{K-d-1}{d} \right\rceil - \left(\frac{a^d-1}{a-1}\right) C_1 - C_k \left\lceil \frac{N-K}{d} \right\rceil \leq \text{ECQSL}_N(x).$$

The C_1 constant is then determined by setting the above minimum cost value to zero, and is given by

$$C_1 = \frac{S_{min} - C_h(K-1) - C_k \left\lceil \frac{N-K}{d} \right\rceil}{\left(\frac{a^{d-1}-1}{a-1}+1\right) \left\lceil \frac{K-d-1}{d} \right\rceil + \left(\frac{a^d-1}{a-1}\right)}. \tag{4.32}$$

So far we have used (4.19) and (4.31) for determining 3 parameters C_1, C_h and C_k. We need a third equation to determine the parameters for the non-retrial model. (We compare the cost of near drop with cost of the non-visited samples using the group of samples with the same number of samples.)
Here we assume the cost of non-visited samples is greater than the cost of near drop sample points and is given by

$$C_k \geq \left(\frac{a^{d-1}-1}{a-1}+1\right) C_1. \tag{4.33}$$

From (4.19), (4.32) and (4.33) C_1 for ECQSL_N can be obtained

$$C_1 = \frac{S_{min}}{\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4} \tag{4.34}$$

where $\alpha_1 = \left(\frac{a^{d-1}-1}{d-1}\right)(K-1)$,
$\alpha_2 = \left(\frac{a^{d-1}-1}{a-1}+1\right) \left\lceil \frac{K-d-1}{d} \right\rceil$, $\alpha_3 = \left(\frac{a^d-1}{a-1}\right)$
and $\alpha_4 = \left(\frac{a^{d-1}-1}{a-1}+1\right) \left\lceil \frac{N-K}{d} \right\rceil$.

Therefore, the handoff cost, C_h, for ECQSL_N can be obtained

$$C_{h3} = C_h = \frac{S_{min}\left(a^{d-1}-1\right)}{(d-1)\left[\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4\right]}. \tag{4.35}$$

We can derive the signal quality measure for the non-retrial case by substituting (4.33), (4.34) and (4.35) in (4.29).

4.4.7 Measure for the Non-retrial Model, EICQSL$_N$

The extended expression for ICQSL, (4.4), cost function using the non-retrial model, is given by

$$\begin{aligned}\text{EICQSL}_N(x) = \frac{1}{N}\Bigg\{ & \sum_{i\in N_g(x)} A_i(x) \\ & -\frac{C_1}{2}\left[\sum_{i\in\beta_x}\left(1+\cos\pi\left(\frac{S_i(x)-S_{drop}}{S_{min}-S_{drop}}\right)\right)\right] \\ & -\sum_{r=1}^{d_{max}} h_r \sum_{j=1}^{d-1} C_r - C_k(N-K) - C_h\gamma_x\Bigg\}. \end{aligned} \tag{4.36}$$

From (4.27), (4.30) and (4.36) we obtain

$$\begin{aligned}\text{EICQSL}_N(x) > \frac{1}{N}\Big[& S_{min}|N_g(x)| - C_h(K-1) \\ & -\left(\frac{a^{d-1}-1}{a-1}+1\right) C_1\left\lceil\frac{K-d-1}{d}\right\rceil \\ & -\left(\frac{a^d-1}{a-1}\right) C_1 - C_k\left\lceil\frac{N-K}{d}\right\rceil\Big]. \end{aligned}$$

Finally C_1 and C_h for EICQSL$_N$ can be obtained

$$C_1 = \frac{S_{min}|N_g(x)|}{\alpha_1+\alpha_2+\alpha_3+\alpha_4}, \tag{4.37}$$

$$C_{h4} = C_h = \frac{S_{min}|N_g(x)|(a^{d-1}-1)}{(d-1)\left[\alpha_1+\alpha_2+\alpha_3+\alpha_4\right]}. \tag{4.38}$$

For retrial and non retrial models we obtain:

$$\mathrm{ECQSL}_R(x) = \frac{\sum_{i\in N_g(x)} A_i(x)}{|N_g(x)|} - \frac{C_1}{2}\left[\sum_{i\in\beta_x}\left(1+\cos\pi\left(\frac{S_i(x)-S_{drop}}{S_{min}-S_{drop}}\right)\right)\right] - \sum_{r=1}^{d_{max}} h_r \sum_{j=1}^{r} C_r - C_{h1}\gamma_x, \quad (4.39)$$

$$\mathrm{EICQSL}_R(x) = \frac{1}{N}\left\{\sum_{i\in N_g(x)} A_i(x) - \frac{C_1}{2}\left[\sum_{i\in\beta_x}\left(1+\cos\pi\left(\frac{S_i(x)-S_{drop}}{S_{min}-S_{drop}}\right)\right)\right] - \sum_{r=1}^{d_{max}} h_r \sum_{j=1}^{r} C_r - C_{h2}\gamma_x\right\}, \quad (4.40)$$

$$\mathrm{ECQSL}_N(x) = \frac{\sum_{i\in N_g(x)} A_i(x)}{|N_g(x)|} - \frac{C_1}{2}\left[\sum_{i\in\beta_x}(1+\cos\pi\left(\frac{S_i(x)-S_{drop}}{S_{min}-S_{drop}}\right)\right] - \sum_{r=1}^{d_{max}} h_r \sum_{j=1}^{d-1} C_r - C_k(N-K) - C_{h3}\gamma_x, \quad (4.41)$$

$$\mathrm{EICQSL}_N(x) = \frac{1}{N}\left\{\sum_{i\in N_g(x)} A_i(x) - \frac{C_1}{2}\left[\sum_{i\in\beta_x}\left(1+\cos\pi\left(\frac{S_i(x)-S_{drop}}{S_{min}-S_{drop}}\right)\right)\right] - \sum_{r=1}^{d_{max}} h_r \sum_{j=1}^{d-1} C_r - C_k(N-K) - C_{h4}\gamma_x\right\}, \quad (4.42)$$

Handoff cost for ECQSL and EICQSL for retrial and non-retrial models we obtain:

For retrial model,

$$C_{h1} = \frac{S_{min}(a^{d-1}-1)}{\left\lceil\frac{N}{d+1}\right\rceil (d-1)\left[\left(\frac{a^d-1}{a-1}\right)+\frac{(a^{d-1}-1)}{(d-1)}d\right]},$$

$$C_{h2} = \frac{|N_g(x)|S_{min}(a^{d-1}-1)}{\left\lceil \frac{N}{d+1} \right\rceil (d-1)\left[\left(\frac{a^d-1}{a-1}\right) + \frac{(a^{d-1}-1)}{(d-1)}d\right]},$$

and for non-retrial model,

$$C_{h3} = \frac{S_{min}\left(a^{d-1}-1\right)}{(d-1)\left[\left(\frac{a^{d-1}-1}{d-1}\right)(K-1) + \left(\frac{a^{d-1}-1}{a-1}+1\right)\left\lceil \frac{K-d-1}{d} \right\rceil + \left(\frac{a^d-1}{a-1}\right) + \left(\frac{a^{d-1}-1}{a-1}+1\right)\lceil \frac{N-K}{d} \rceil\right]},$$

$$C_{h4} = \frac{S_{min}|N_g(x)|(a^{d-1}-1)}{(d-1)\left[\left(\frac{a^{d-1}-1}{d-1}\right)(K-1) + \left(\frac{a^{d-1}-1}{a-1}+1\right)\left\lceil \frac{K-d-1}{d} \right\rceil + \left(\frac{a^d-1}{a-1}\right) + \left(\frac{a^{d-1}-1}{a-1}+1\right)\lceil \frac{N-K}{d} \rceil\right]}.$$

4.4.8 Optimal Value for ECQSL

Exhaustive Search

The optimal handoff sequence can be defined as one which maximizes all evaluation methods. The following equation can be used to find the optimal handoff sequence:

$$\max_{i\in 1,..,N} \max_{(n,m)\in C} \left\{ \sum_{k=1}^{i} S_{kn} + \sum_{k=i+1}^{N} S_{km} \right\}, \tag{4.43}$$

where $C = \{(n,m)|n \in \{1,..,M\}, m \in \{1,..,M\}, n \neq m\}$.

In our exhaustive method, instead of all possible options we consider only the following options

$$= \sum_{d=1}^{L} M\left(Nd\right), \tag{4.44}$$

where $L =$ maximum number of handoffs and $L <<<< (N-1)$. The complexity of our exhaustive method is $O(MN^d)$ in comparison to the complexity $O(M^N)$ using a normal exhaustive search.

The following illustrated procedure is used to obtain optimal handoff sequences carried out in Section 4.5.

Algorithm 1 Exhaustive search for optimal values

for $x \in \{x_1, .., x_{1000}\}$ **do**
  best $= 0$
  next $=$ exhaustive $(X, CQSL, k)$
  while next $>$ best and $k <$ maxlimit **do**
    best $=$ next
    $k = k + 1$
    next $=$ exhaustive $(X, CQSL, k)$
  end while
end for

4.5 Simulation Results and Discussion

Table 4.2: Parameters used in simulation

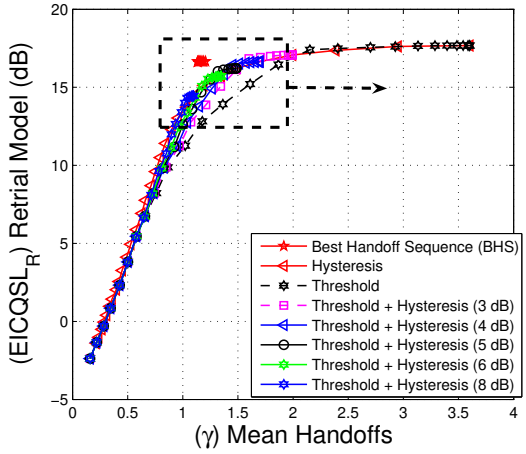
Cell radius (r)	100 m [192]
Number of base stations (M)	3 [119]
Number of sample paths (η)	1000 [119, 192]
Spatial sampling interval	1 m [114]
Standard deviation of shadow fading (σ)	5 dB [178]
Correlation distance of shadow fading	20 m [114]
Sample points (N)	100
Threshold (T)	variable
Minimum acceptable signal strength (S_{min})	15 dB [119]
Dropping signal strength (S_{drop})	14.5 dB
Maximum signal strength (S_{max})	$1.5S_{min}$ dB
Path-loss constant (K_1)	85
Path-loss exponent (K_2)	35 [119]
Consecutive sample points (d)	3

In this work we assume that channel capacity is unlimited. Users move in any random direction. Further, as in [148] we assume a homogeneous network where

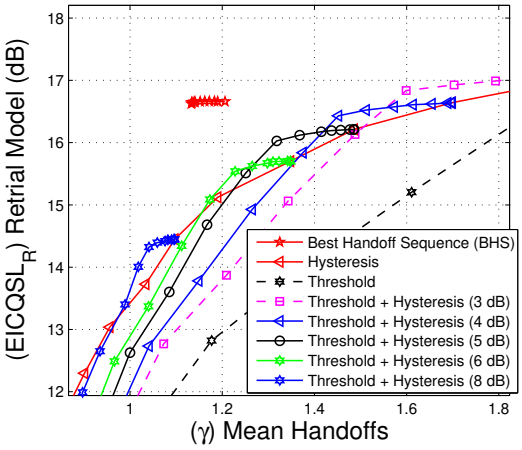
all cells are identical in size, user mobility and cell coverage. Each cell is assumed to have an equal number of neighbors as in [142]. A log normal propagation model is assumed, and no power control exists.

Here we compare the different handoff methods introduced in Section 4.1 using different quality measures. We randomly generate $\eta = 1000$ sample paths, each with a number of sample points $N = 100$ where each pair of consecutive points are one meter apart. For a more realistic view, we add shadowing to the simulation following a log-normal propagation model, as described in [60]. This was assumed to generate signal strengths in each sample point along all the sample paths, i.e., $S_{ij} = K_1 - K_2 log(r) + \rho$, where $K_1 = 85$; $K_2 = 35$ are constants, r is the distance to the base station, and ρ is Gaussian distributed ($N(0, \sigma^2)$) representing the shadowing effect. We set $\sigma = 5$ dB, shadowing correlation distance equals 20 m, $S_{min} = 15$ dB as in [119] and $S_{max} = 1.5 S_{min}$. All the sample paths are straight lines that start from points in the square area $\{(100, 100), (200, 100), (200, 200), (100, 200)\}$ and are distributed uniformly in the region [125]. Their directions are randomly uniformly distributed over $[0, 2\pi]$ as in [125]. In our simulations, we assume that a call will be dropped after $d = 3$ consecutive dropping sample points. Assuming that the user is traveling at a constant speed, it is possible to calculate the speed corresponding to $d = 3$. In practice, often a time period (for the GSM this is 6 s) with dropping signal strength level is considered as a service failure or call dropping. When the user speed increases, it is possible to increase d, leading to lower call dropping. Simulation parameters are summarized in Table 4.2. The values in figures are obtained by varying the threshold in the Threshold method, as well as the hysteresis threshold in both the Hysteresis and the Threshold with Hysteresis methods, respectively, from 1 to 30 dB, to see the most efficient threshold value.

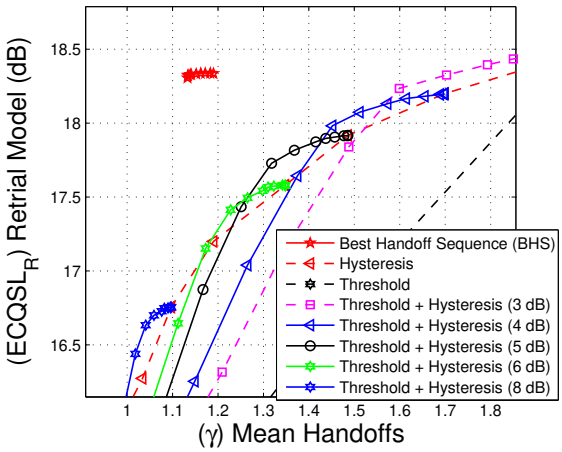
The performance of different handoff methods was evaluated as shown in Fig. 4.11 for the retrial model, where repeated call attempts are made assuming $\delta = 0.9$. It is indicate that the Threshold method with 5 dB or 6 dB performs well as we need to minimize the number of handoffs as well as maximizing the signal quality. It can be observed that with the non-retrial case, there is less difference in



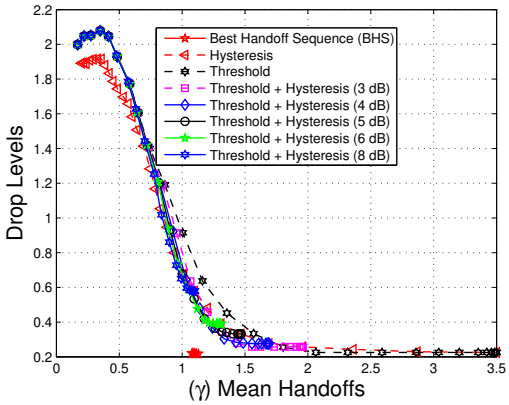
(a) EICQSL versus mean handoffs ($\overline{\gamma}$) values for retrial model, when $\delta = 0.9$ (full figure)



(b) EICQSL versus mean handoffs $\overline{\gamma}$ values for retrial model, when $\delta = 0.9$



(c) ECQSL versus mean handoffs ($\overline{\gamma}$) values for retrial model, when $\delta = 0.9$



(d) Drop levels versus mean handoffs $\overline{\gamma}$ for retrial model, when $\delta = 0.9$

Figure 4.11: Performance evaluation for different handoff methods using the retrial model, where repeated call attempts are made. Here the dropping probability $\delta = 0.9$.

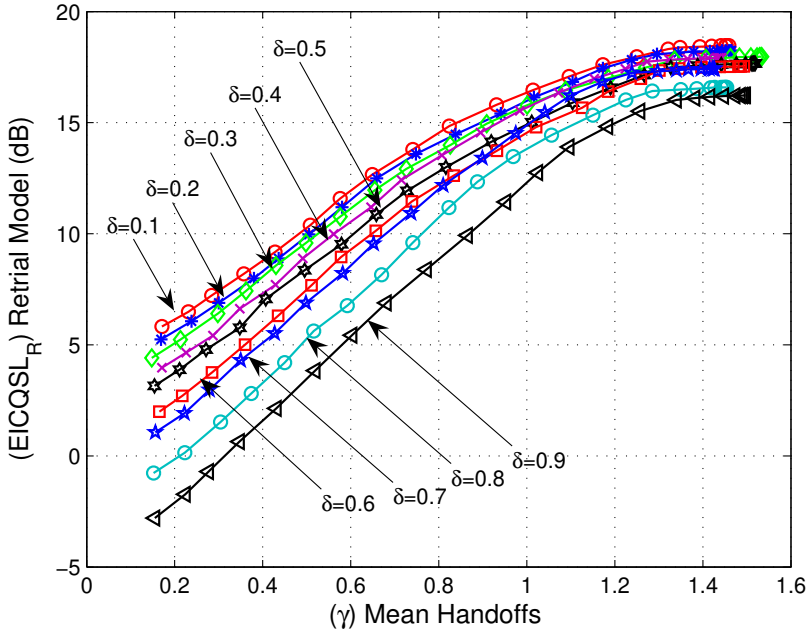


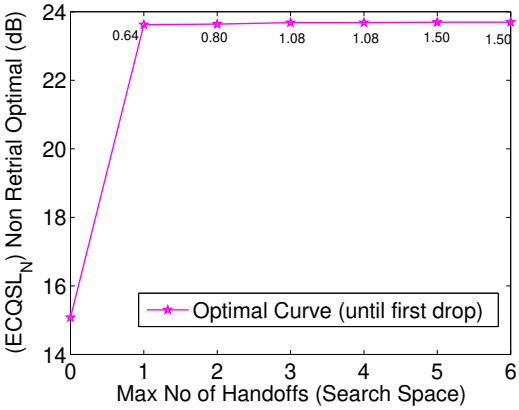
Figure 4.12: EICQSL$_R$ for retrial model when δ varies

quality between existing handoff methods.

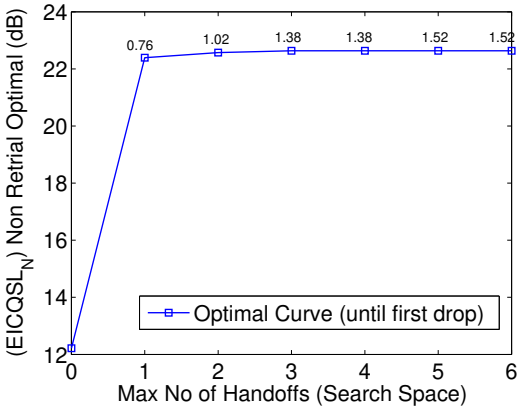
Using the parameter values from Table 4.2, and varying the system dropping probability δ from 0.1 to 0.9 with the Threshold with 6 dB Hysteresis method, we find that the total quality of the call decreases with increasing δ, as shown in Fig. 4.12. For this reason, we compare existing handoff methods by varying δ, within the interval $0 < \delta < 1$ and observe that the handoff method with the best performance varies when δ is increased as shown in Table 4.3. The simulation results indicate that the threshold method with 4 dB hysteresis performs well for urban areas with high dropping probability, while the threshold method with 6 dB hysteresis performs well for suburban areas with low dropping probability.

Results in Fig. 4.13 and Fig. 4.14 are generated using the Algorithm 1 (procedure) proposed in Section 4.4.8. The values associated with each point along the curves in the graph are the average number of handoff per user. Observe that the quality ($ECQSL_R$ and $ECQSL_N$) does not improve as the maximum allowable number of handoff per user increases.

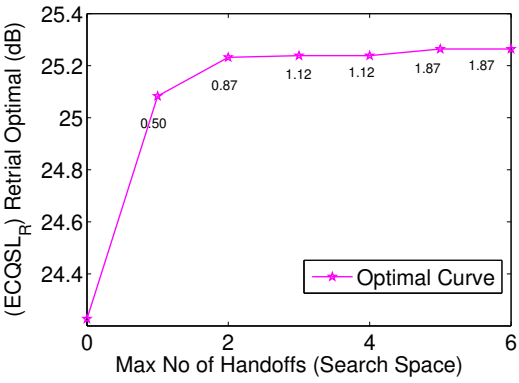
In Fig. 4.15, we compare existing handoff methods with optimal values for the retrial and the non-retrial models. We found that the existing handoff methods



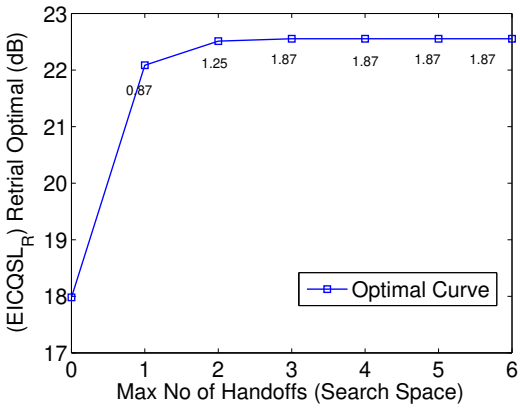
(a) ECQSL versus search space for non-retrial models



(b) EICQSL versus search space for non-retrial models



(c) EICQSL versus search space for retrial models



(d) ECQSL versus search space for retrial models

Figure 4.13: Optimal curves for the retrial and the non-retrial models when considering ECQSL and EICQSL

Table 4.3: Recommended handoff methods for different system dropping probabilities

system	Retrial Model		Non-retrial Model	
δ	ECQSL$_R$	EICQSL$_R$	ECQSL$_N$	EICQSL$_N$
0.1	Hysteresis	Hysteresis	T+H 6 dB	T+H 6 dB
0.2	Hysteresis	Hysteresis	T+H 6 dB	T+H 6 dB
0.3	Hysteresis	Hysteresis	T+H 6 dB	T+H 6 dB
0.4	Hyst or T+H 6 dB	Hyst or T+H 6 dB	T+H 5 dB	T+H 5 dB
0.5	T+H 6 dB	T+H 6 dB	T+H 4 dB	T+H 4 dB
0.6	T+H 5 or 6 dB	T+H 5 or 6 dB	T+H 5 dB	T+H 5 dB
0.7	T+H 5 or 6 dB	T+H 5 or 6 dB	T+H 4 or 5 dB	T+H 4 or 5 dB
0.8	T+H 5 or 6 dB	T+H 5 or 6 dB	T+H 4 dB	T+H 4 dB
0.9	T+H 5 or 6 dB	T+H 5 or 6 dB	T+H 4 dB	T+H 4 dB

are less efficient than the optimal handoff sequence by a margin of 29-45 % for retrial model and by 34-77 % for non-retrial model, as shown in Fig. 4.13.

We also observed that the handoff distributions of various handoff methods are different when 1000 users are considered, as shown in Fig. 4.16. Moreover, according to Fig. 4.17 when handoff cost decreases, the consecutive dropping points increases. Therefore, handoff cost decreases with increasing mobile user speed.

4.6 Chapter Summary

This chapter has proposed a new measure for performance evaluation and comparison between existing handoff algorithms, taking into consideration signal level, call dropping and handoff cost for cellular networks, for both the retrial (where repeated call attempts are made) and the non-retrial call options. The increase in quality of the calls was quantified using the proposed quality measure. Moreover, this measure can be used by network operators to set suitable values for the hysteresis margin, and the handoff threshold to obtain optimal quality while reducing the number of handoffs and call dropping.

The results indicate that, for urban areas with high dropping probability, the

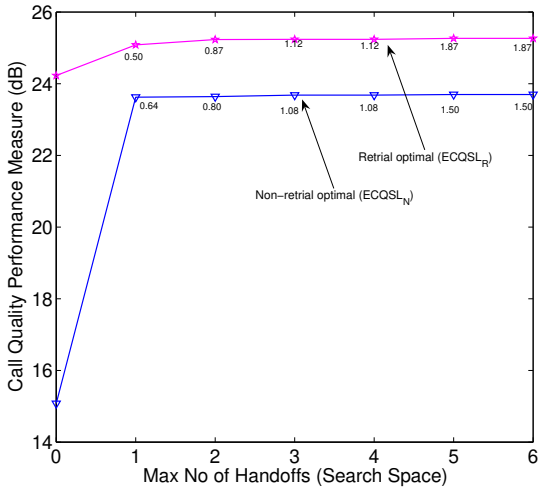
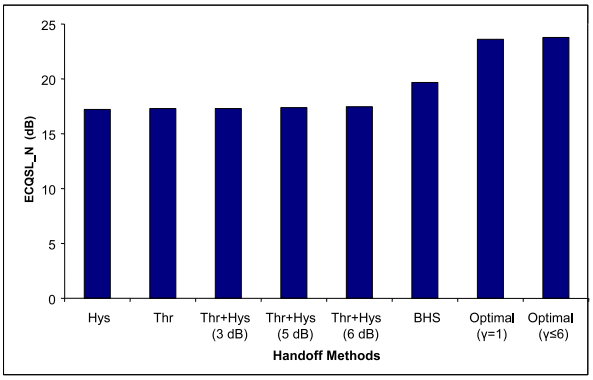


Figure 4.14: Optimal curves for retrial and non-retrial for *ECQSL*

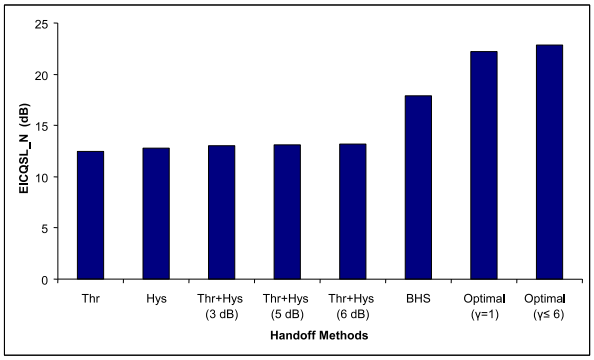
Threshold with 4 dB Hysteresis performs well, while for suburban areas with low dropping probability, the Threshold with 6 dB Hysteresis performs well. This chapter has found that the existing handoff methods are less efficient than the optimal handoff sequence for the retrial model by 29-45 % and for the non-retrial model by 34-77 %. We have also proposed a method to estimate handoff cost and optimal values for retrial and non-retrial models. This chapter has also provided recommendations for specific parameter values to improve the performance of currently used handoff methods. Designers can now optimize quality of the call based on efficient handoff algorithm and using other recommended parameter values.

Several important conclusions arise from this chapter:

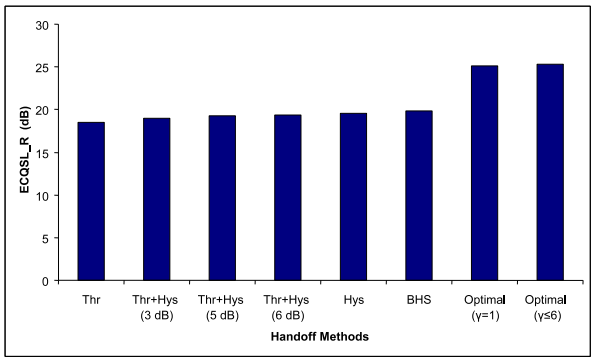
- A more realistic method to evaluate performance of a handoff method can be derived by considering signal level, call dropping and handoff cost for both retrial (where repeated call attempts are made after a call is lost) and non-retrial call options.
- The results suggest two different handoff methods for urban areas with high dropping probability and suburban areas with low dropping probability. (Suitable handoff methods for given terrain configuration as suggested.)



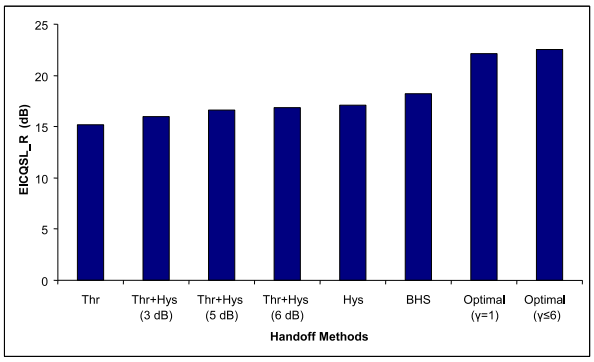
(a) Different handoff methods with optimal values ($ECQSL_N$ for non-retrial models)



(b) Different handoff methods with optimal values ($EICQSL_N$ for non-retrial models)



(c) Different handoff methods with optimal values ($ECQSL_R$ for retrial models)



(d) Different handoff methods with optimal values ($EICQSL_R$ for retrial models)

Figure 4.15: Comparing different handoff methods with optimal values when number of handoff ($\gamma = 1$ and $\gamma \leq 6$)

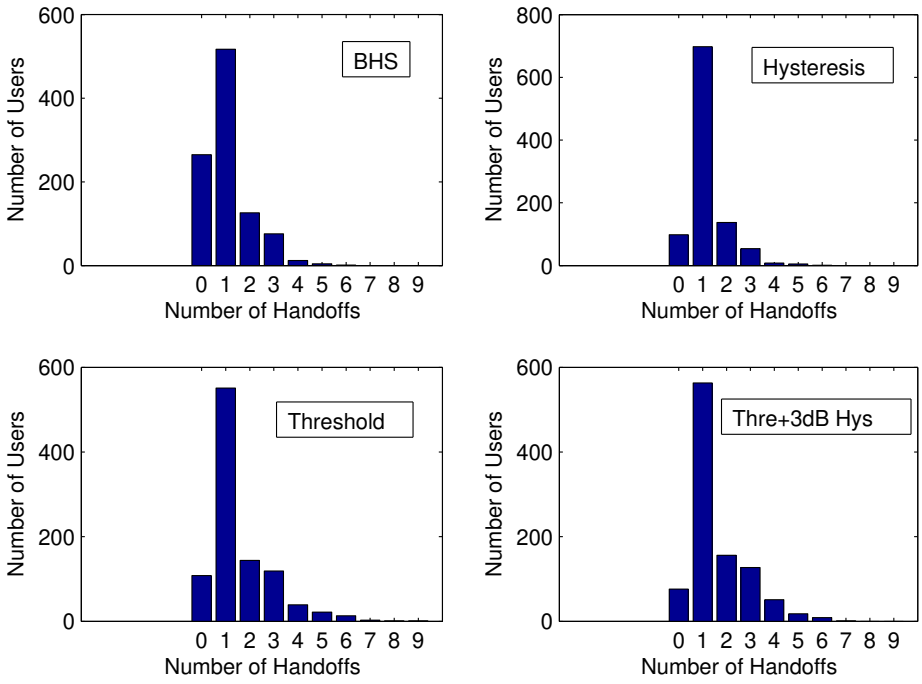
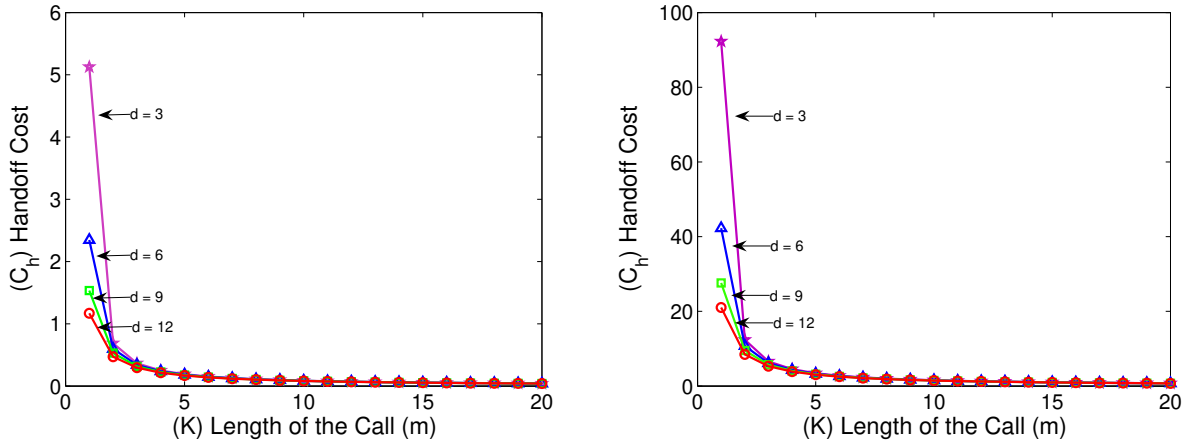


Figure 4.16: User distribution of various handoff evaluation methods considering 1000 users



(a) Handoff cost, C_h, versus length of the call, K, in meters, for ECQSL on non-retrial model with different consecutive dropping points, d

(b) Handoff cost, C_h, versus length of the call, K, in meters, for EICQSL on non-retrial model with different consecutive dropping points, d

Figure 4.17: Handoff cost, C_h, versus length of the call, K, in meters, for ECQSL and EICQSL on non-retrial model with different consecutive dropping points, d. This graph shows that handoff cost decreases when consecutive dropping points increases. Handoff cost thus decreases with mobile users' increasing speed

- Both handoff cost and optimal handoff sequence can be estimated.
- The proposed evaluation models indicate that existing handoff methods have room to improve by comparison to the optimal handoff sequence.

CHAPTER FIVE

Conclusions for Part I

Part I has investigated the evaluation of handoff algorithms in cellular networks, leading to a new framework for evaluation that is primarily based on signal quality, but also considers call dropping probability, quality of service and possible options of retrial after call failure. The limits of a critical parameter were determined using dropping probability, as presented in Appendix A. An off-line cluster-based computationally simple heuristic algorithm was proposed to find a near optimal handoff sequence. This sequence can be used as a benchmark to compare existing handoff algorithms to identify the trade-off between signal quality and number of handoffs.

In an inner city wireless communication network where microcells may be necessary and handoff is a considerable problem in resource allocation, users normally travel on predefined paths such as roads and foot-paths. In these scenarios, the best sequence approximated by the proposed benchmark can be used with pattern recognition methods (for example template matching techniques) to update the possible future handoff sequence prior to each handoff. The benchmark's accuracy was compared to those of computationally expensive dynamic programming and exhaustive methods.

This framework for performance evaluation and comparison of existing handoff algorithms was further extended to include for both retrial (where repeated call attempts are made) and non-retrial call options. A method of estimating handoff cost and optimal values for retrial and non-retrial models was also presented. The extension included a realistic and practical way of providing penal-

ties when a call is dropped. The increase in quality of the calls is quantified using the proposed quality measure.

The results suggested two different handoff methods for urban areas with high dropping probability and suburban areas with low dropping probability. The proposed evaluation model indicated that existing handoff methods have room to improve by comparison to the optimal handoff sequence.

The work also has provided recommendations for specific parameter values to improve performance of currently used handoff methods. Designers can now optimize quality of the call based on efficient handoff algorithm and using other recommended parameter values.

Part II

Power Management in Sensor Networks

CHAPTER SIX

Introduction to Part II

6.1 Problem Statement

Sensor networks consist of many sensor nodes that can be deployed in random positions. Different aspects of sensor networks such as data aggregation or fusion [23, 30], packet size optimization [154], target localization [207] and network protocols [73, 80, 132, 199, 200] are discussed in the literature with respect to crucial energy limit and maximizing network lifetime [98,99,105,156,161,168,205].

The lessons learned from work on handoff for cellular networks, discussed earlier, inspired this work on infrastructureless or ad hoc wireless sensor networks. The lack of an infrastructure requires a sensor network to be self organizing. Therefore, the handoff solutions discussed in Part I cannot be applied directly. The concept of handoff is very much integrated into sensor networks' communication protocols and design requirements. In turn, these depend strongly on power management, and therefore, it is a crucial design consideration in sensor networks that are mostly battery powered.

Power management starts with the task of finding a comprehensive energy model considering all possible sources of energy drainage. We should then investigate methods of estimating sensor network lifetime. Effective power management requires the development of more efficient communication protocols with existing battery technology. Ultimately, effective management methods should lead to sensor network design guidelines capable of prolonging a sensor network's life.

In this introduction to Part II, we elaborate the motivation for power management considering available radio resources in Section 6.2. Section 6.3 summarizes some applications of wireless sensor networks that demand efficient power management strategies. Section 6.4 describes the three well-known fixed assignment multiple access methods, emphasizing the method used in wireless sensor networks. Section 6.5 outlines the performance objectives of a sensor network and Section 6.6 sets out the key considerations of network design triggered by these objectives. Section 6.7 discusses methods of power management and Section 6.9 summarizes the chapter and provides an outline for Part II.

6.2 Motivation for Power Management

In the last decade, commercial radio technology has advanced and commercial standards such as Bluetooth, developed by the Bluetooth consortium [2], have started to appear. Ad hoc networks have been gaining popularity for military, space, biomedical and manufacturing applications in recent years because their easy deployment and lack of infrastructure requirements.

Unlike cellular wireless networks, ad hoc wireless networks do not need any fixed communication infrastructure. Three main networking protocols are known in wireless communications: direct communication, multi-hop communication and clustering. The routes can be single or multi-hop and the nodes which may be heterogeneous and communicate via packet radio.

The heterogeneity of the nodes would allow some nodes to be servers and others to be clients. The ability of an ad hoc node to act as a server or service provider will depend on its energy, memory and computational capacities. Each node should estimate its own battery life before committing to a task. Even relaying packets for others may result in deteriorating its own limited battery power, and the node may not accept the task when it is devoted to another important activity.

Direct (one-to-one) communication between the base station and a large number of sensors is extremely energy consuming because of direct transmission and

the multi-hop communication. This is considered globally inefficient as data is routed through individual sensor nodes to the base station, making clustering an appropriate method to use. Clustering reduces the data to be transmitted to the base station by processing all data locally. It is widely accepted that aggregation of sensor nodes into clusters reduces the energy required for long distance radio transmissions, especially when the radio ranges of individual sensors are expected to be short. Data aggregation techniques can be used to combine correlated data from sensor nodes into a small set of data which contains only relevant information [73]. Using cluster-based communication protocols, sensor nodes send their data to the Cluster Head (CH) which then forwards the data to the sink node or base station. All sensor nodes within the cluster may be identical, however, CH may have in some instances additional features such as more computational power, longer-range radio and location awareness using the Global Positioning System (GPS). Obviously, power management may also depend on the design requirements of the sensor network, generally dictated by the application concerned. Cluster-based sensor systems can be used when the sensor network is fully wireless but only slightly mobile, hybrid (wireless and wired), or fully wireless with a known sensor location.

6.3 Applications of Wireless Sensor Networks

Several research papers [29, 95, 136] have discussed the applications of wireless sensor networks and their challenges. Sensor network applications can be divided into two groups: querying applications and tasking applications [79]. In querying, information collected by sensors will be processed based on the query that triggered the data collection, for example, to obtain data about an event in the environment. To minimize the communication cost, the data must be aggregated before it is passed back to the origin of the query.

In tasking applications, an event to be observed or monitored by the network triggers the data collection. Sensor nodes perform some actions if an event triggers them. As in the querying method, data is aggregated to avoid many nodes

forwarding the same data. Sensors can also be coordinated to get a better idea about the event, for example, some sensors can be moved closer to the event. Some well-known applications of sensor networks are summarized below.

1. **Security and Military Sensing**

   Traditionally, sensors are used in defense technologies, and therefore, many developments in related areas, such as multi sensor data fusion, are associated on their infancy with military applications. Sensor networks can allow remote monitoring of sensitive information important for security. Examples include research and developments in Chemical, Biological and Nuclear (CBN) sensors [6] and battlefield intelligence regarding the numbers, locations and movement of troops [95].

2. **Habitat Monitoring**

   One of the earliest known civil applications of sensor networks is in ecological habitat monitoring. A team from University of California Berkeley [113, 170, 171] used a wireless sensor network to observe birds on an island, using a base station connected over the web via a satellite communication link. This kind of ”unattended” monitoring minimizes disruption to the objects of study by an observer walking around the island to collect data.

3. **Industrial Control and Monitoring**

   Networked sensors are used to monitor and control manufacturing processes and are considered as a part of a factory automation. Particularly in sensitive industries such as chemical plants, various types of sensors (chemical, temperature and other types) can provide information to control the process. Other common examples of industrial applications include monitoring, lighting, heating, ventilation and air conditioning in large commercial buildings.

4. **Health Monitoring**

In medical terms, health monitoring refers to non-life-threatening situations. Particular applications include tracking and monitoring the performance of an athlete using wearable sensors for various types of information (for example illicit performance enhancing drug use), monitoring medical implants and for investigations of the digestive system by using sensors that can be swallowed.

Health monitoring of large structures such as aircrafts, mining excavators, or road bridges is also associated with sensor networks. Networking the various sensors embedded in such structures can provide valuable information for monitoring safety and durability.

5. **Home Automation and Consumer Goods**

   Intelligent homes equipped with networked sensors monitored or controlled by the user are already a commercial reality. In some cases automated lighting, curtain control, ventilation and heating depend on sensor networks. A new application in this area is in aged care homes, where sensors allow residents to manage their own health and safety.

   Networked sensors also have applications in toys and other consumer goods such as cars and computers. Multiple wireless sensors communicate with each other in toy robots. Wireless peripherals in computers are already popular. The integration of toys with personal computers is an interesting area in wireless sensor networks.

6. **Intelligent Agriculture and Farming**

   Sensors can detect soil moisture, need for fertilizer, level of pesticides, received levels of sunshine and other information and through a network provide valuable information for intelligent use of resources such as water and fertilizers.

   Animals on a farm can be tagged with sensors that allow a base station to monitor their location and raise an alarm when the animal needs attention such as treatment of parasites or when it wanders on to a road or any other

dangerous environment.

6.4 Selection of the Access Method

The choice of access method affects the network QoS. There are three Fixed Assignment Multiple Access methods which have a fixed allocation of channel resources.

1. **Frequency Division Multiple Access (FDMA)**

   The frequency is used to separate simultaneously transmitted signals. The FDMA is based on frequency division multiplexing known from analog technology and radio/TV broadcasting. The design of FDMA systems needs to consider issues such as adjacent channel interference and the near-far problem.

2. **Time Division Multiple Access (TDMA)**

   In TDMA, many users can share the same frequency in different time slots. The time slot adjustments provide the flexibility to allow different access rates. Clearly, TDMA's inherent digital compatibility makes it more applicable for wireless communication. A user may be assigned a time slot that can be synchronized with the receiver.

3. **Code Division Multiple Access (CDMA)**

   In a CDMA environment, multiple users use the same band simultaneously. The users are identified by a code or key. The receiver uses the code to distinguish between users. CDMA is a useful method when integrating voice, data and video into wireless communication. Its strength lies in its capability to accommodate users with technical diversity (for example different bandwidth requirements and switching methods) without the need for coordination.

Our work uses TDMA as the access method, commonly used in sensor networks.

6.5 Network Performance Objectives

A sensor network design may have stricter restrictions imposed by the intended application than any other wireless network. Therefore, such design requirements involve technical challenges, as described below [29].

1. **Low Energy Dissipation**

   In wireless sensor networks the source of power has to be isolated from the main grid power. Batteries are the obvious choice of power, although in some cases energy is taken from an alternative source (for example solar power in habitat monitoring, or conversion from thermal or mechanical energy in pressure sensors). Therefore low energy dissipation is an essential objective of wireless sensor network design.

2. **Low Cost**

   The relative cost of a sensor unit and usage cost should be low for a sensor network. Communication protocols should be designed to lower the implementation cost by reducing the microelectronics (or silicon area) including the memory. The ad hoc or self organizing nature of the sensor network also lowers the network administration costs.

3. **Security**

   In the security of sensor networks, a main concern is to ensure message integrity i.e., to avoid information being modified by an intruder. The problems associated with this task include the implementation of low cost hardware, and managing of key distribution.

4. **Network Type**

   As mentioned earlier, a widely used star network with a single master (or base station) and many slaves may not be ideal, as the network's range is then limited to that of the base station. Most such networks, therefore, are multi hop networks with clusters of sensors implemented as local star networks. A cluster head acts as the master for each cluster. These cluster heads

communicate with the base station using multi hop transmission. It is challenging to design such a network considering the low power requirement and minimal microelectronics available.

5. **Data Throughput**

   In most real sensor networks, data throughput is not expected to be high (few bits per second). This will affect the protocol design, as sensor networks will be much less efficient (due to essential headers, the need for security and so on) than general wireless networks having much higher data throughput. This requirement possesses a challenge to the protocol design for wireless sensor networks.

6.6 Design Challenges Posed by Network Performance Objectives

In contrast to cellular type wireless network designs where infrastructure costs play a significant role in the design, the node cost is the main hardware cost associated with wireless sensor networks. A sensor network's main energy consumption occurs in sensing, data processing and communications. Minimizing the energy consumption is a primary goal in sensor network design, and is addressed in this thesis in detail later. This goal introduces some other design challenges.

The nodes can be designed to have sleep periods, and an event may trigger a node to wake up and start processing that information. This can, however, reduce the responsiveness and therefore effectiveness of a node due to possible latency in the waking up process. Nevertheless, if the event is reported rapidly enough, this strategy can still work in applications with a very high sensor density.

Scalability is another major challenge in the design of a sensor network with a large number of sensors. In such situations protocols must involve localized communication and distributed processing and may support hierarchical sensor network architectures.

It is likely that the design involves heterogeneous sensors in the network. A

realistic scenario (used later in this work) is to have a small number of nodes with high computational capacity and high battery power, and a large number of devices with lower computational capacity and low battery power. A key design criterion is to obtain the right proportion of each group of sensors and the proportion of the computational power between the two types of nodes.

The network design can support self configuration as they are ad hoc networks with no central management. The network should be capable of configuring its own topology, self-calibrating and coordinate it's own inter-node communication.

A sensor network deployment is primarily based on the requirements of the coverage and connectivity needed. Requirements for coverage will depend on the environment and the quality and the safety of information to be collected. Requirements for connectivity depend on the topology of information routing selected.

In deploying a sensor network, the following issues should be considered:

1. **Sensor Placement:**

   Sensors can be randomly scattered, or carefully placed in tactically important locations. Several research efforts [94, 106, 186, 187] have proposed optimal sensor placement.

2. **Number of Sensors to be Placed:**

   It may be possible to deploy more sensors than needed from the beginning coupled with sensor sleep times. This may reduce sensor deaths from draining power. Another option is to add sensors incrementally. In both cases environmental considerations play a role as the increased number of dead sensors should not interfere with the environment.

3. **Topology of Information Routing:**

   This is an important consideration as power management also depend on it. Possible options include cluster based topology with a cluster managing

local information processing. Cluster heads may use either a direct connection or multi-hop to communicate with the base station.

6.7 Power Management Methods

As stated in the previous section, topology of information routing is an important issue in sensor deployment. It is also important in power management, as energy usage depends on topology.

1. **Rotating Cluster Heads**

   One well-known strategy is to rotate cluster heads to achieve a balanced energy dissipation among the nodes. A PhD project [71] conducted at MIT proposed an application specific communication protocol "Low-Energy Adaptive Clustering Hierarchy" or LEACH based on clustering of mobile sensor nodes. The core idea behind LEACH protocol is that sensor nodes located closer to each other will have high correlation in their measured data, and therefore, it is not necessary for the nodes to communicate with the central base station. Instead, neighbouring nodes are grouped into clusters. Each cluster has a cluster head (CH) that collects data from other members of the cluster, aggregates them, and sends to the central base station, which may be located far outside the sensor field. In order to prevent a premature battery failure of the sensors selected as CHs, all the clusters are reconfigured and CHs are reassigned after a certain period of time. Before the assignment of each new CH, there is a set up time where each sensor independently decides whether to be a CH or not according to a probability based criterion. Unfortunately, such a system may create too many unevenly distributed cluster heads and this in turn result in strongly varying cluster sizes and clearly leading to rapid dissipation of the energy in the sensor network. Therefore, there is clearly substantial room for improvement of this protocol. A simple solution would be to add a cluster merging step to the set up phase before CHs announce their election as CHs to the rest of the network.

The handoff issue in such a dynamic network is strongly integrated with the selection of the cluster and the CH. Each time the clusters are formed, a sensor has to judge whether a handoff should take place or not. Unlike the cellular wireless networks, there will be no central decisions regarding handoff, nor it is feasible to have any resource allocation in most cases.

A hybrid system in which mobile systems themselves select whether to communicate directly or via base stations, the "cellular ad-hoc united communication system", has been proposed in [7, 8].

2. **Hierarchical Clustering Architectures**

   It is known that energy used in communication is far higher than the energy used for sensing and computation. For example, the energy needed for communicating 1 bit over wireless is about 1000-10000 higher than that needed to process the information [29]. Dividing the entire network into clusters and using CHs to process local information before communicating it to the base station (generally using multiple hops) reduces energy consumption. Cluster hierarchies include various levels of clusters, for example, level $k-1$ cluster heads forms clusters and selects cluster heads for level k, and the remaining level $k-1$ cluster heads become member nodes of the level k clusters [29].

3. **Traffic Distribution and System Partitioning**

   The usual strategy of sending information via the shortest path is not suitable for wireless sensor nodes with limited battery power. Therefore, traffic distribution should take the availability of nodes and the expected network lifetime into account. For example, system partitioning [32, 121, 184] may allow the sharing of intensive communications by remotely located sensor nodes that are not used often.

4. **Information Processing and Data Aggregation**

   Information processing and data aggregation [30, 81, 129] methods have to support many tasks, including low power communication; dense spatial

sampling of important events; asynchronous and distributive computation; data fusion; and querying and routing. Data fusion and aggregation should lead to minimizing of traffic loads by reducing redundant information. It is likely that multiple sensors report the same event to intermediate nodes, which then fuse the data before forwarding it. An example of such a data-centric protocol is sensor protocols for information via negotiation (SPIN) [74,88].

5. **Cross Layer Design**

   Network protocols in wireless networks are normally designed as a layered stack, which enables the simplification of the network, and the development of robust and scalable protocols [79]. A main disadvantage of this approach is that the design and operation of each layer are isolated from the rest. Therefore the interface between the layers remains static and independent from individual network constraints [79]. In a light of this disadvantage, various approaches have been proposed using cross layer design [56,71,72, 152,164]. Principles and strategies of such approaches are discussed in [56]. Major issues in this active research area include the information exchange between layers and negotiating specific application requirements subject to global restrictions.

6. **Energy Scavenging**

   In addition to the two commonly known approaches to the problem of keeping the sensor nodes alive (improving batteries' energy density of batteries and improving energy usage by using new protocols), an interesting new approach has been reported. Self-generation of power by nodes is known as energy scavenging. This is an alternative way to solve the problem of power management. As nodes attempt to use the environment to generate power, there cannot be a universally applicable method of energy scavenging. A survey on power or energy scavenging methods is reported in [150]. Both solar power and vibration are promising methods of power scavenging as they offer relatively high power density [150]. As opposed to bat-

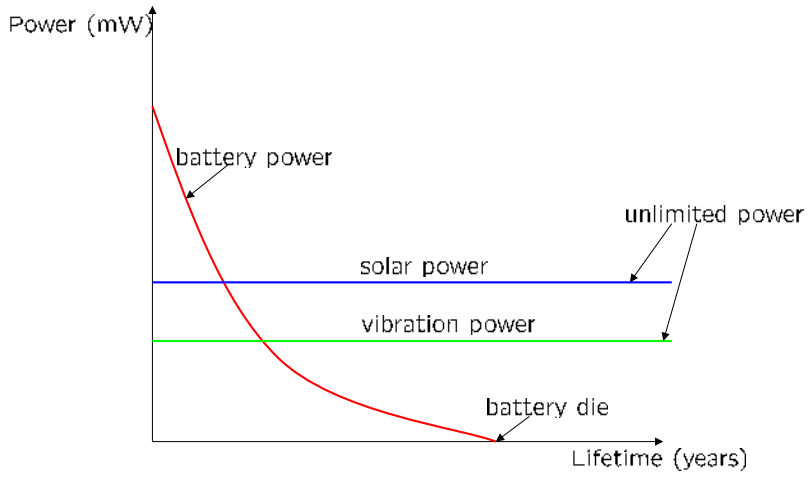


Figure 6.1: Different power sources verses lifetime

teries that have a limited lifetime, both solar and vibrations based power have unlimited lifetime as shown in Fig. 6.1. The use of environmental energy to power wireless sensor networks has been proposed in several studies [12,17,87,90,91,128,141,149,204]:

(a) **Solar Power** [11,87,139]

Solar power is a commercially available mature technology with efficiency between 12% to 25% and may be suitable for most outdoor applications. The available power depends on the amount of sunlight available.

(b) **Vibration** [13]

In a built environment, vibrations can be the source of power. However, use of this source may be limited to selected applications. A particular use of this technology is in indoor environments, where vibrations of many surfaces inside a building can be used.

(c) **Passive Human Power**

This can be a source of energy for wearable sensors. Some experiments suggest that the foot during heel strike and the bending of the ball of the foot may be easily used to scavenge power. The energy generated in the foot should be transferred to other parts of the body where a wearable node is likely to be located.

(d) **Active Human Power**

In this case, humans generate the power required. Some examples are radios powered by human action and flashlights powered by a bicycle or by squeezing a lever. This method may be limited to certain less critical applications of wearable sensor networks.

(e) **Acoustic Noise**

It is often debated whether acoustic noise can be used as a source of power. For most sensor network applications, the power level is too low.

6.8 Network Lifetime

The definition of network lifetime of a sensor network is application-specific. Network lifetime is defined as the maximum time limit before certain conditions are satisfied. Commonly used conditions for this purpose are [48]:

- the time for the first node to die [34, 35, 155, 165] and
- the time for a certain percentage of network nodes to die [50, 188, 190, 197].

Alternatively, there are other conditions such as:

- mean expiration time [175].
- in terms of the packet delivery rate [36],
- in terms of the number of alive flows [25],
- the time to the first loss of coverage [21],

therefore consider quality of communication the alive nodes achieve.

6.9 Summary and Outline of Part II

This chapter provides a brief review of the relevant literature and introduces Part II. Firstly, this chapter presents the motivation for power management in sensor

networks and summarizes major sensor network applications. It then discusses the associated network performance objectives and design challenges. The chapter then discusses power management methods, a major design challenge.

The work in Part II investigates power management in sensor networks. In Chapter 7, a comprehensive energy model for a wireless sensor networks is proposed by considering seven key energy consumption sources. The current energy consumption models ignore many of these important sources of energy drainage. Using the proposed model the lifetime of a sensor node is estimated. The benefit of using the proposed comprehensive model is shown by comparison with other existing energy models. Further, this model is applied to LEACH protocol to obtain its accurate evaluation of energy consumption and node lifetime. Chapter 7 also investigates the impact of the optimal number of clusters, for example, free space fading energy, different energy models, physical area, duty cycles and number of sensors. This chapter provides guidelines for efficient and reliable sensor network design that can be used to optimize energy efficiency subject to required specifications.

Extending the sensor network's lifetime is important for most of applications. Batteries are the most widely used power source for sensors in the network. By efficiently managing capacity of the batteries the lifetime of the network can be prolonged. We therefore investigate efficient battery management in sensor networks. In Chapter 8, the simulation results show that the average energy consumption ratio of the normal sensor node to the cluster head is very low (0.0429). We therefore propose high energy packs for CHs. This study develops a method based on High Powered Cluster Heads that can alleviate this problem while extending network lifetime. Furthermore, by adding multiple high powered batteries to a single cluster, the sensor network's lifetime can be significantly increased. It is shown that the ratio of initial battery capacities of sensors and cluster heads changes according to different types of application. It is also shown that the proposed method can be used in conjunction with LEACH to increase overall efficiency. Moreover, it is observed that factors such as battery cost and sensor deployment cost make a huge impact on the total network cost.

CHAPTER SEVEN

Energy Consumption in Sensor Networks

7.1 Introduction

Sensor networks play a major role in many aspects of society including home automation, consumer electronics, military application, agriculture, environmental monitoring and health monitoring [29]. Usually sensor devices are small and inexpensive, so that they can be produced and deployed in large numbers. Their resources of energy, memory, computational speed and bandwidth are severely constrained [4]. Therefore, it is important to design sensor networks to maximize their life expectancy.

Many simple energy models have proposed in literature by considering mainly computational and communication energy dissipation at the sensor in wireless sensor networks. These energy models ignore other important energy dissipated sources and, therefore, may not actually produce the real network lifetime.

7.1.1 Motivation

Current energy consumption models ignore many important sources of energy drainage. What is the comprehensive energy model for wireless sensor networks that considers all important energy consumption sources? It is essential to consider the impact of such an energy model on the sensor network lifetime. What is the impact when the comprehensive energy model is considered in place of the

commonly used energy model? We compare existing energy models and trying to answer this question in a later section. LEACH is a protocol which rotates the CH among all the sensors. By rotating CHs, LEACH distribute the energy load among all the nodes, therefore network lifetime can be increased. However, rotation of CHs in communication rounds dissipate battery power unnecessarily in LEACH type protocol. Can such a comprehensive energy model be applied to a rotating CH based protocol like LEACH?

Independent of LEACH, it should be possible to calculate the optimum number of clusters. How can we estimate the optimal number of clusters? What is the effect on sensor network lifetime, if we use the optimal number of clusters? Would the optimal number of clusters be changed substantially if we use different energy models? There are applications with densely populated sensors in relatively small sensor field areas. Would clustering still lead to higher efficiency when the deployment area is small, for example, when a million sensors are placed over a 10^5 m^2 area? How do the distances, duty cycles and number of sensors affect the sensor network lifetime?

Sensor numbers also have an effect on the network lifetime. How much network lifetime is increased when sensors are distributed uniformly and number of sensors is doubled? How does the optimal number of clusters change with free space fading? How does the size of the network change with different free space fading?

Designers can optimize energy efficiency subject to required specifications. What are the most important guidelines for efficient and reliable sensor network design?

In this chapter my research results will address the above questions. The chapter is organized as follows. Section 7.1.3 sets out the key assumptions of this chapter. In Section 7.2, we describe our sensor energy model and relevant energy consumption sources. In Section 7.3, extend our model to a network. We apply our results to a LEACH [73] type protocol to obtain sensor energy consumption and network lifetime. Section 7.5 outlines the simulation set up and Section 7.6 states the numerical results where we demonstrate the benefits of our comprehen-

Table 7.1: Different energy models

Energy Sources	Heinzelman et al. [73]	Millie et al. [120]	Zhu et al. [205]	Our Model
Processing	√	√	√	√
Communication	√	√	√	√
Sensing	–	–	√	√
Transient	–	√	–	√
Logging	–	–	–	√
Actuation	–	–	–	√
Initial	–	–	–	√

sive energy model and its application to a general sensor network, with particular focus on a LEACH-type protocol. Section 7.7 provides a summary of the chapter and concluding remarks.

7.1.2 Energy Models

Accurate network lifetimes can be predicted using a comprehensive energy consumption model. There have been various attempts at modeling sensor node energy consumptions.

1. Heinzelman et al. [73] use a simple energy model by considering only microcontroller processing and radio transmission and reception. This model does not consider other important sources of energy consumption, such as transient energy, sensor sensing, sensor logging and actuation.

2. The model proposed by Millie and Vaidya [120] does not consider energy consumption of sensor sensing, sensor logging and actuation.

3. Zhu and Papavassiliou's model [205] does not consider energy consumption of transient energy, sensor logging and actuation.

In Table 7.1, we compare these energy models with the new method proposed in this chapter.

7.1.3 Assumptions

This section list the assumptions that are used in this chapter and Chapter 8.

1. All sensor nodes communicate directly with the CH in one hop, as in [73], and a CH communicates with the neighboring CH or base station depending on the distance.

2. The base station is far away from the sensor field and at a fixed location.

3. All sensor nodes (except CHs) are homogeneous. Therefore energy consumption for all activities, excluding for communication energy due to different transmit distances to their CHs, are the same for each sensor node.

4. Transmission and reception energy used by a CH is higher than that of a normal sensor node because of the additional data processing and aggregation tasks associated with the cluster head.

5. As in [72], we assume for a given signal to noise ratio, a symmetric radio channel, making the energy needed to transmit from one point to another and in the reverse direction identical.

6. All sensors acquire information at a fixed rate, making data available to be sent to the sink at any time.

7.2 Sensor Energy Model

We consider a wireless sensor network with a cluster topology, as shown in Fig. 7.1, in which sensors are grouped into clusters, and individual sensors sense data and transmit to CH using single hops as in [73]. Here we assume that all sensor nodes within a cluster use time-division multiple-access (TDMA) to access their CH. Data is generated in individual sensor nodes. The CH aggregates this data and forwards it to the base station or sink via other CHs with multi-hop communication.

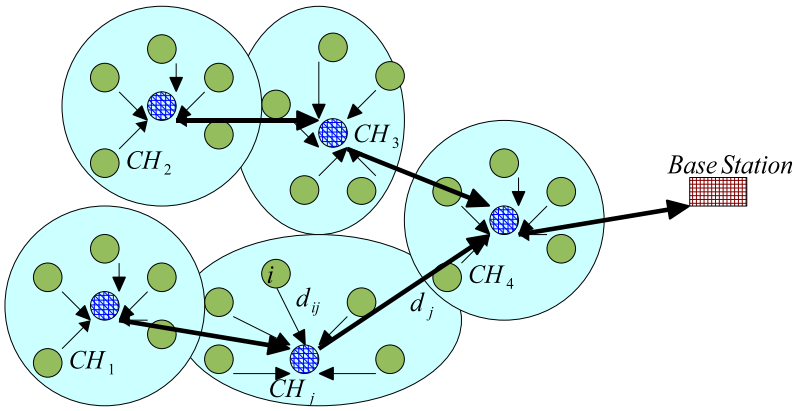


Figure 7.1: Cluster topology of a Sensor Network. (Sensors are grouped into clusters, and individual sensors sense data and transmit to cluster heads (CH). Cluster heads aggregate this data and then forward it through a unique root, depending on the tree structure, to the base station or sink node)

As in [168], a sensor lifetime can be divided into rounds. In each round a sensor node performs steps 1-5 as shown in Fig. 7.2(a), and any CH performs steps 1-7 as shown in Fig. 7.2(b).

We assume that every sensor node generates a fixed-sized packet and forwards this to its CH. All generated packets are forwarded to the base station by the CH in each round. The base station schedules transmission time based on TDMA to avoid collisions.

Here, we define *distance* as the physical length between a sender and its receiver. Sensors may not be placed at equal distances as they are generally placed around CHs randomly. However, by calculating the total energy consumption of each sensor we can suggest average energy consumptions for all the nodes while accounting for load balancing. In the literature [48], *network lifetime* is defined as either the time until the first (or last) node dies or the time until a given percentage (P_{node}) of the nodes dies. We adopt the latter definition.

A node/CH can die either because it runs out of battery, or because the death of other CHs isolates it from its base station. We define C_{opt}, *optimal number of clusters*, as the number of clusters that minimizes energy dissipation.

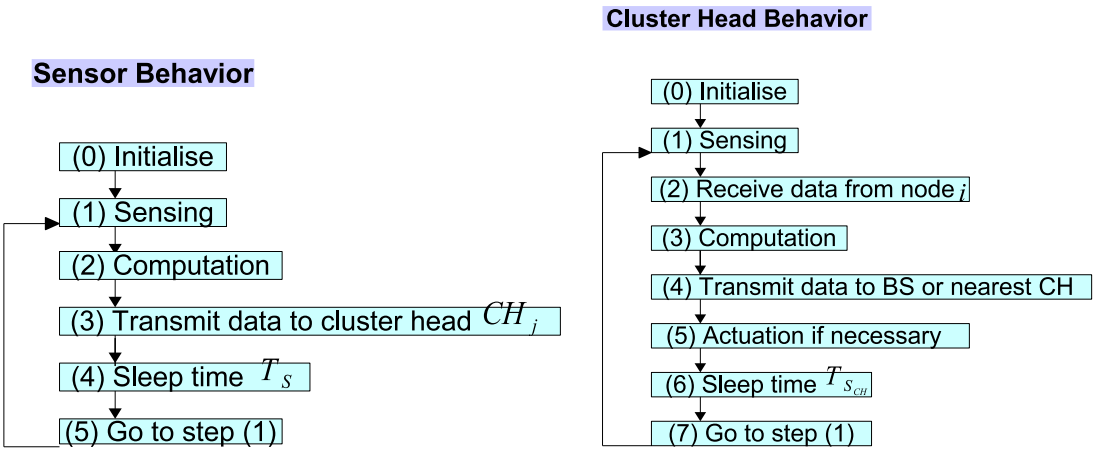
As in [72], we assume a symmetric radio channel making the energy needed to

transmit from one point to another in both directions identical. We also assume that all sensors acquire information at a fixed rate, making data available to be sent to the sink every round.

The energy consumed by a sensor node can be attributed to seven main basic energy consumption sources: micro controller processing, radio transmission and reception, transient energy, sensor sensing, sensor logging and actuation.

We assume that all sensor nodes (except CHs) are homogeneous, therefore energy consumption for all activities (excluding for communication energy due to different transmit distances to their CHs,) are the same for each sensor node. It is also assumed that the transmission and reception energy used by a CH is higher than that of a normal sensor node because of the additional data processing and aggregation tasks associated with it.

Let $h_i > 1$ be a weighting factor that applies to a CH to indicate by how much it consumes more energy than a regular sensor node for energy source i, with $i = 1, 2, 3, 4$ for processing, transmission and reception, sensing and sensor logging respectively. Sections 7.2.1 to 7.2.8 describe the energy needed in each step in



(a) Sensor node operation

(b) Cluster head operation

Figure 7.2: Sensor node and cluster head operations

detail (Fig. 7.2(a) and 7.2(b)).

7.2.1 Initial Energy

The energy dissipation for initial set up or to turn on the sensor, E_{ini}, is considerable. The initial or start up energy becomes comparable to transition energy in short range communication [202]. Here we assume that the initial energy of regular sensor node is equal to the initial energy of the CH.

7.2.2 Micro-controller Processing

The energy for processing and aggregation of the data mainly consumed by the micro-controller is attributed to two components: energy loss from switching, E_{switch}, and energy loss due to leakage current, E_{leak}. This energy, E_{leak} dissipated by leakage current, occurs when a sub-threshold leakage current flows between the power source and the ground.

The total energy dissipation by the sensor node used for data processing and aggregation for b bit packet, E_{pro_N}, is given by [185]:

$$E_{pro_N}(b, N_{cyc}) = \underbrace{bN_{cyc}C_{avg}V_{sup}^2}_{switching} + \underbrace{bV_{sup}\left(I_0 e^{\frac{V_{sup}}{n_p V_t}}\right)\left(\frac{N_{cyc}}{f}\right)}_{leakage},$$

and total energy dissipation by the cluster head (CH), $E_{pro_{CH}}$, is given by

$$E_{pro_{CH}}(h_1, b, N_{cyc}) = \underbrace{h_1 bN_{cyc}C_{avg}V_{sup}^2}_{switching} + \underbrace{bV_{sup}\left(I_0 e^{\frac{V_{sup}}{n_p V_t}}\right)\left(\frac{N_{cyc}}{f}\right)}_{leakage},$$

where weighting factor, $h_1 > 1$, N_{cyc} is the number of clock cycles per task, C_{avg} is the average capacitance switched per cycle, V_{sup} is the supply voltage, I_0 is the leakage current, n_p is the constant, V_t is the thermal voltage and f is sensor frequency.

Assuming that sensor nodes only sense data and transmit to their CH once during each round, we ignore energy dissipation from data processing and aggregation in sensor nodes.

7.2.3 Radio Transmission and Reception

Communication between neighboring sensor nodes is enabled by a sensor transceiver. Energy dissipation by a sensor node can be attributed to transmitting and receiving data. According to [185] the energy dissipation due to transmit b bit packet, in a distance d from sensor node to the CH is given by

$$E_{tx_N}(b,d) = \underbrace{bE_{elec}}_{electronics} + \underbrace{bd_{ij}^{n}E_{amp}}_{amplifier}, \tag{7.1}$$

where E_{elec} is the energy dissipated to transmit or receive electronics, E_{amp} is the energy dissipated by the power amplifier and n is the distance based path loss exponent (we use $n = 2$ for free space fading[1], and $n = 4$ for multi path fading [73]). Energy dissipation due to receiving b bit packet from the sensor node is given by $E_{rx_N}(b) = bE_{elec}$. Therefore energy dissipation due to transmission of a b bit packet over a distance d from the CH to the neighboring CH can be estimated by

$$E_{tx_{CH}}(h_2,b,d) = h_2 \underbrace{bE_{elec}}_{electronics} + \underbrace{bd_{j}^{n}E_{amp}}_{amplifier},$$

where $h_2 > 1$ and the energy dissipation due to receiving a b bit packet from the CH estimated by $E_{rx_{CH}}(b) = h_2 bE_{elec}$.

7.2.4 Control Packet Overheads

These include energy dissipation from transmitting and receiving RTS, CTS, ACK packets and retransmissions. This is only relevant to contention based protocols like CSMA/CA and not to TDMA.

[1]Free space fading refers to the attenuation of received signal strength when transmitter and receiver have a clear unobstructed line-of-sight path between them [144].

7.2.5 Transient Energy

In each node, radio and micro-controller units (MCU) support different operating modes including active, idle and sleep. Transitions between operating modes involve significant energy dissipation [120]. Changes in radio operating mode can caused significant amount of power dissipation. These are often ignored in the literature. Let T_{tranON} and $T_{tranOFF}$ be the times required for sleep-to-idle and idle-to-sleep transitions, respectively. A sensor node will listen to a busy tone

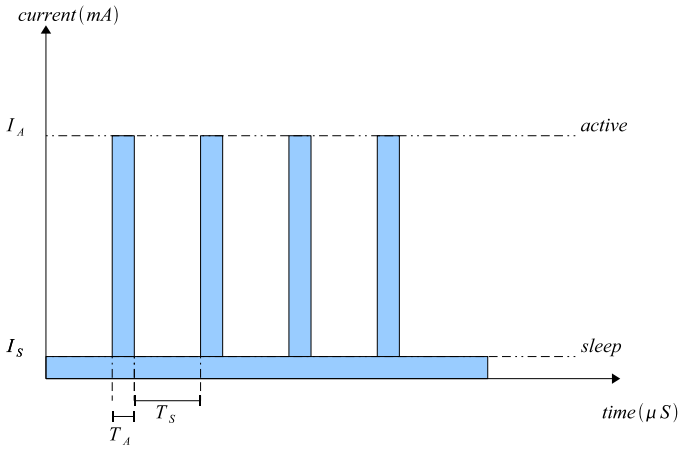
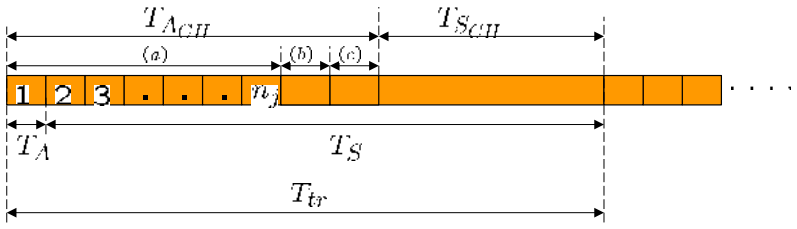


Figure 7.3: Duty cycle for a sensor node



(a) time to receive packets from its own n_j sensors

(b) time to receive packets from neighbouring CH

(c) time to transmit packets to neighbouring CH

Figure 7.4: Wake-up and sleeping times of the sensor nodes and the CHs

of the channel, wake up for a duration of T_A and then sleep for T_S, assuming $T_S \gg T_A$. Similarly, CH wakes up for duration $T_{A_{CH}}$, which will be discussed in

Section 7.3.3, and then sleeps for $T_{S_{CH}}$. Let T_{tr} be the time between consecutive packet transmissions. The CH will transmit all the packets it receives in one batch every T_{tr} seconds (Fig. 7.4). This is given by

$$T_{tr} = T_{A_{CH}} + T_{S_{CH}} = T_A + T_S. \tag{7.2}$$

The duty cycle for the sensor node, d_N, can be defined as in [120]:

$$d_N = \frac{T_{tranON} + T_A + T_{tranOFF}}{T_{tranON} + T_A + T_{tranOFF} + T_S}.$$

Similarly, the duty cycle for the CH, d_{CH} is defined by:

$$d_{CH} = \frac{T_{tranON} + T_{A_{CH}} + T_{tranOFF}}{T_{tranON} + T_{A_{CH}} + T_{tranOFF} + T_{S_{CH}}}. \tag{7.3}$$

The average current for a sensor node is given by $I_N = d_N I_A + (1 - d_N) I_S$. The total energy dissipation by a sensor node in operating mode is evaluated by

$$E_{tran_N} = T_A V_{sup} \left[d_N I_A + (1 - d_N) I_S \right],$$

where I_A and I_S are current for active and sleeping mode. Similarly, the average current for a CH is given by $I_{CH} = d_{CH} I_A + (1 - d_{CH}) I_S$ and the energy dissipation due to operating mode at the CH is evaluated by

$$E_{tran_{CH}} = T_{A_{CH}} V_{sup} \left[d_{CH} I_A + (1 - d_{CH}) I_S \right].$$

7.2.6 Sensor Sensing

The sensing system links the sensor node to the physical world. Sources of sensor power consumption are: signal sampling and conversion of physical signals to electrical signals, signal conditioning and analogue to digital conversion (ADC). Let I_{sens} be the total current required for sensing activity and T_{sense} be the time duration for sensor node sensing. We evaluate the total energy dissipation for

sensing activity for b bit packet, E_{sens_N} at the sensor node by

$$E_{sens_N}(b) = bV_{sup}I_{sens}T_{sens},$$

and the total energy dissipation for sensing activity at the CH by

$$E_{sens_{CH}}(h_3, b) = h_3 E_{sens_N}(b),$$

where weighting factor, $h_3 > 1$.

7.2.7 Sensor Logging

Sensor logging consumes energy used for reading b bit packet data and writing it into memory [166]. Sensor logging energy consumption for a sensor node is evaluated by

$$E_{logg_N}(b) = E_{write} + E_{read} = \frac{bV_{sup}}{8}\left(I_{write}T_{write} + I_{read}T_{read}\right),$$

where E_{write} is energy consumption for writing data, E_{read} is energy consumption for reading b bit packet data, I_{write} and I_{read} are current for writing and reading 1 byte data. Energy consumption for logging sensor readings at the CH, can be evaluated by

$$E_{logg_{CH}}(h_4, b) = h_4 E_{logg_N}(b),$$

for weighting factor, $h_4 > 1$.

7.2.8 Actuation

Energy dissipation for actuation, E_{actu}, is hard to estimate in general because this is highly dependent on application. Total energy dissipation for actuation is $E_{actu}N_{act}$ where N_{act} is the number of actuations per CH. For example if we use temperature sensors to drive a fan that needs two motors, there can be a com-

mand to switch on the two motors when the temperature is beyond some value. In this case N_{act} would be 2. In practice, however, actuation may not apply to all sensors. Actuation is performed by dedicated cluster heads, except for LEACH type protocols where actuation can be performed by any sensor node.

7.2.9 Rotation of Cluster Heads

These include energy dissipation due to rotating CH, E_{rotate}. This is only relevant to rotating CH based protocols like LEACH, discussed in section 7.3.2.

7.3 Network Energy consumption

7.3.1 Applying the Proposed Energy Model to a Fixed Cluster Head

In this section we apply the previously defined energy model to a sensor network, assuming that N_s sensor nodes are randomly and uniformly distributed in a $M \times M$ region.

Consider a sensor network with k clusters where the clusters are laid out in a directed tree topology whose root is a base station (sink node). Cluster j comprises one CH denoted CH_j and n_j sensor nodes, $j = 1, 2, ..., k$. Thus, on average the number of sensor nodes (including the CH) in a cluster is (N_s/k), as some clusters will have $\lceil N_s/k \rceil$ and some $\lfloor N_s/k \rfloor$ sensor nodes.

The total number of sensors is $N_s = \sum_{j=1}^{k}(n_j + 1)$. Sensors transmit information to their respective CH. The CH will then forward the packet through a unique route of CHs to the sink node (the uniqueness results from the tree structure). All transmissions are from the leaves through the intermediate nodes towards the root which is the sink node. Let d_j be the distance between CH_j and the next CH (or the sink node) that it transmits to, and d_{ij} be the distance between node i in cluster j and its cluster head. Here, the total energy consumed by sensor node i

in cluster j is

$$E_N(ij) = \left[\underbrace{E_{ini}}_{initial} + \underbrace{bE_{elec} + bd_{ij}^n E_{amp}}_{transmit} + \underbrace{T_A V_{sup} \left[d_N I_A + (1 - d_N) I_S \right]}_{transient} \right.$$
$$\left. + + \underbrace{bV_{sup} I_{sens} T_{sens}}_{sensing} \underbrace{bV_{sup} \left(I_{write} T_{write} + I_{read} T_{read} \right)}_{data-logging} \right]. \quad (7.4)$$

Similarly, the total energy consumed by cluster CH_j is

$$E_{CH}(j) = \left[\underbrace{E_{ini}}_{initial} + \underbrace{h_1 b N_{cyc} C_{avg} V_{sup}^2}_{switching} \left(n_j \right) + \underbrace{bV_{sup} \left(I_0 e^{\frac{V_{sup}}{n_p V_t}} \right) \left(\frac{N_{cyc}}{f} \right)}_{leakage} \left(n_j \right) \right.$$
$$+ \underbrace{h_2 bE_{elec}}_{receive} \left(n_j - 1 \right) + \underbrace{h_2 bE_{elec} + bd_j^n E_{amp}}_{transmit} + \underbrace{h_3 bV_{sup} I_{sens} T_{sens}}_{sensing} + \underbrace{E_{actu} N_{act}}_{actuation}$$
$$\left. + \underbrace{T_{CH} V_{sup} \left[d_{CH} I_A + (1 - d_{CH}) I_S \right]}_{transient} + \underbrace{h_4 bV_{sup} \left(I_{write} T_{write} + I_{read} T_{read} \right)}_{data-logging} \right], \quad (7.5)$$

where $n_j = (N/k), \forall j$, for the case of equi-sized clusters (all clusters have the same number of sensor nodes). For LEACH, energy consumption due to CH rotation, E_{rotate} is added to (7.5).

To compute the energy consumption of all $E_N(ij)$ and $E_{CH}(j)$ values, we apply equation (7.4) and (7.5) first to the leaf clusters and progress recursively down the tree until we reach the root.

Therefore the total energy consumed by the entire network is given by

$$E_{tot} = \sum_{j=1}^{k} \left(E_{CH}(j) + \sum_{i=1}^{n_j} E_N(ij) \right). \quad (7.6)$$

7.3.2 Applying the Proposed Energy Model to Rotating Cluster Heads (LEACH)

Here we apply our energy model to a rotating CH based protocol like LEACH [73].

LEACH is a protocol which rotates the CH among all the sensors in its cluster. It is an example that directly extends the cellular TDMA model to sensor networks [199]. By rotating cluster heads, LEACH distribute the energy load among all the nodes, so that the network's lifetime is increased. Unfortunately, rotating cluster heads in every communication round dissipates battery energy unnecessarily.

Let S be the set of network scenarios. As there are $n_j + 1$ sensors in a cluster j, the total number of network scenarios (the cardinality of S) is given by $|S| = \prod_{j=1}^{k}(n_j + 1)$. Consider a network scenario in which a particular choice of sensor nodes are CHs – one in each cluster. Let $d_j(s)$ be the distance from the CH_j to the next CH (or sink) in a network scenario s ($s \in S$). For each $s \in S$ and each cluster j, let $E_{CH}(s_j)$ be the energy consumed by CH_j and $E_N(s_{ij})$ be the energy consumed by sensor i in cluster j, both in scenario s. Replacing d_j with $d_j(s)$ in (7.4) and (7.5) enables us to obtain $E_N(ij)$ and $E_{CH}(j)$ values respectively.

Let T_s be the proportion of time the system spends in network scenario s, $s \in S$. The average energy consumption of any sensor in cluster j in such a LEACH-type protocol is estimated by

$$E_L(j) = \frac{1}{(n_j + 1)} \sum_{s=1}^{|S|} T_s \left[E_{CH}(s_j) + \sum_{i=1}^{n_j} E_N(s_{ij}) \right], \tag{7.7}$$

and the total energy consumed by the entire network is given by

$$E_{tot_L} = \sum_{j=1}^{k} E_L(j).$$

This gives rise to many interesting questions of how to optimize the T_s values and the sleep and active times to maximize network lifetime.

7.3.3 Wake up Time for Cluster Head $(T_{A_{CH}})$

The above energy models require the evaluation of the sensor's duty cycle d_{CH} defined in (7.3). In this subsection, we derive an expression for CH wake up time, $T_{A_{CH}}$, which is then used to determine d_{CH}. Fig. 7.5 shows an example of

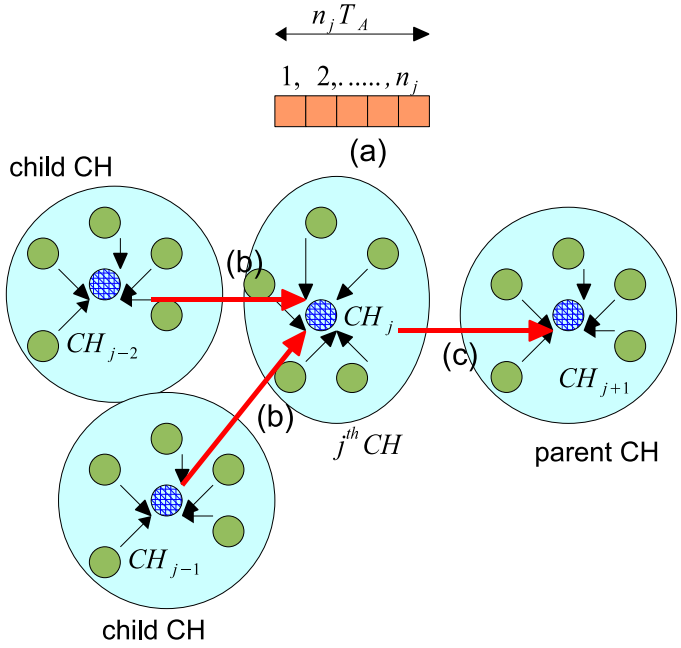


Figure 7.5: Wake up time for cluster head. (This has three components: time taken to receive data from its own sensors, receive data from its children CH, and transmit data to its parent CH. Data can be received only from child CHs and can be transmitted to their parent CH)

a sensor network which we will use to illustrate the data transmission between clusters. With each CH_j, we can associate a parent $p(j)$ and a set $c(j)$ of children. A child $c \in c(j)$ is a CH that forwards data to CH_j. A parent $p(j)$ is a CH that will transmit data from CH_j towards to the base station or sink. For example, in Fig. 7.5, the parent of CH_j is CH_{j+1} and the children of CH_j are CH_{j-1}, CH_{j-2}, therefore, $c(j) = \{CH_{j-1}, CH_{j-2}\}$ and $p(j) = \{CH_{j+1}\}$. The total wake up time for j^{th} CH in one round is the total time taken to:

(a) receive data packets from the sensor nodes (total of n_j sensors) in its own cluster, with each sensor having T_A as the wake up time,

(b) receive incoming data packets from child CHs and

(c) transmit data packets to its parent CH, $p(j)$.

For a LEACH-type protocol, wake up time for CH_j, $T^j_{A_{CH}}$ is given by

$$T^j_{A_{CH}} = \underbrace{\max n_j T_A}_{(a)} + \underbrace{\sum_{c \in c(j)} T^{CH}_{c,j}}_{(b)} + \underbrace{T^{CH}_{j,p(j)}}_{(c)}, \tag{7.8}$$

where $T^{CH}_{c,j}$ is the time taken for transmission from its child CH c to CH j and $T^{CH}_{j,p(j)}$ is the time taken for transmission from CH j to its parent CH $p(j)$. Knowing $T_{A_{CH}}$, the duty cycle of the CH in (7.3) can be determined.

We find in Sections 7.3.1 and 7.3.2 that the sensor node's energy consumption depends on the number of clusters in a network. In the following we seek to find the optimal number of clusters to maximize the network lifetime.

7.4 Finding the Optimal Number of Clusters

In this section we apply the principles discussed in the previous sections to develop a technique for increasing network lifetime by choosing the optimal number of clusters. Generally speaking, if we have more clusters while maintaining the same load per CH, the transmission distance from a sensor to its parent CH is reduced. Therefore, the overall energy consumption is also reduced. On the other hand, increasing the number of clusters means that the transmission path between a sensor and the BS will include more CH-to-CH hops which means higher overall energy consumption. The aim is therefore to find the optimal number of clusters so that the overall energy consumption is minimized. Note that this optimal clustering depends highly on the energy model used [167]. Therefore, it is important to use the right energy model, as this chapter aims to do.

We will now demonstrate the use of our energy model for optimal clustering and compare the results with other approaches. Assume that each sensor node transmits data to its CH only once during each round. Therefore, from (7.4), total

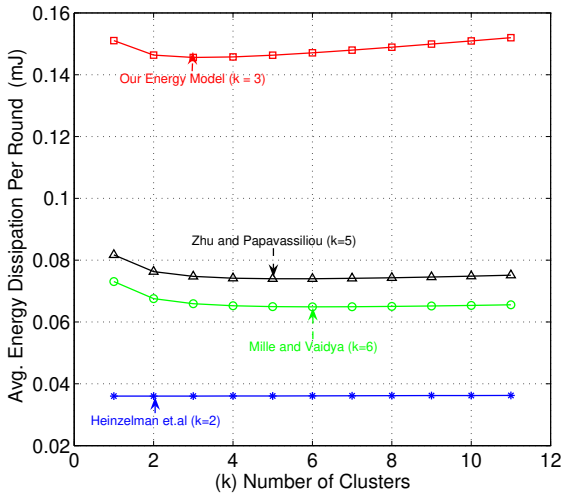


Figure 7.6: Average energy dissipation versus number of clusters when $E_{fs} = 7$ nJ/bit/m^2 and $N_s = 100$ nodes, $M = 100$ by using (7.13). Optimal number of clusters, based on their energy models, are indicated with arrows. Here the average distance from CH to base station or sink node is 22 m. This shows the difference between the energy models does have significant effect to the sensor energy dissipation.

energy consumed by a sensor node is

$$E_{node} = \left[E_{ini} + bE_{elec} + bd^2_{toCH}E_{fs} + E_{tran_N} + bE_{sens_N} + bE_{logg_N}\right], \tag{7.9}$$

where b is the number of bits in every packet, d_{toCH} is the distance between node and CH, E_{fs} is the free space fading energy.

Similarly, from (7.5), the total energy consumed by a CH during each round is

$$\begin{aligned} E_{head} = & \left[E_{ini} + bE_{pro_{CH}}\left(\frac{N}{k}\right) + h_2 bE_{elec}\left(\frac{N}{k} - 1\right)\right. \\ & \left. + h_2 bE_{elec} + bd^4_{toBS}E_{mp} + E_{tran_{CH}} + bE_{sens_{CH}} + bE_{logg_{CH}}\right], \end{aligned} \tag{7.10}$$

where and E_{mp} is the multi path fading energy. Note that we consider a multi path model with d^4 power loss, and assume that actuation is not performed.

Consider a square of area $M \times M$ with k clusters, that is, the area covered by each cluster is approximately M^2/k. As in [73], we assume that the CH is at

the center of mass of its cluster, and we acknowledge that the cluster area can be arbitrary shaped, but for simplicity we assume that it is a square. For $k = 1$, assuming sensors are randomly uniformly distributed over the square area, the mean square distance from a sensor to its CH is given by

$$\begin{aligned} E[d_{toCH}] &= \int_0^M \int_0^M d(x,y)\rho(x,y)dxdy \\ &= \frac{1}{M^2}\int_0^M \int_0^M \left(x - \frac{M}{2}\right)^2 + \left(y - \frac{M}{2}\right)^2 dxdy, \end{aligned} \tag{7.11}$$

where $\rho(x,y)$, $0 \leq x, y \leq M$, is the joint probability density function. (If sensors are placed uniformly then we have $\rho(x,y) = \frac{1}{M^2}$.)

For $k > 1$, the mean square distance is given by

$$\begin{aligned} E[{d_{toCH}}^2] &= \frac{k}{M^2}\int_0^{\frac{M}{\sqrt{k}}} \int_0^{\frac{M}{\sqrt{k}}} \left(x - \frac{M}{2\sqrt{k}}\right)^2 + \left(y - \frac{M}{2\sqrt{k}}\right)^2 dxdy, \\ &= \frac{M^2}{6k}. \end{aligned} \tag{7.12}$$

The above calculation (7.11) and (7.12) is for a square area, therefore $k = i^2$ where i is an integer. As an approximation, we evaluate the mean square distance using (7.11) and (7.12) with an arbitrary value of k. Knowing the mean square distance, we can now derive the optimal number of clusters.

From (7.9) and (7.10) the energy dissipation in a single cluster during each round is given by

$$E_{cluster} = E_{head} + \left(\frac{N}{k} - 1\right) E_{node} \approx E_{head} + \left(\frac{N}{k}\right) E_{node}. \tag{7.13}$$

The total energy for k clusters during each round based on our energy model is obtained using (7.9), (7.10), (7.12) and (7.13) as

$$\begin{aligned} E_{our} &= kE_{cluster} \\ &= b\left(E_{elec}N_s + E_{pro_{CH}}N_s + d_{toBS}^4 E_{mp}k + E_{sens_{CH}}k + E_{tran_{CH}}k + E_{logg_{CH}}k\right. \\ &\left. + E_{elec}N_s + E_{fs}\frac{M^2}{6k}N_s + E_{tran_N}N_s + E_{sens_N}N_s + E_{logg_N}N_s\right). \end{aligned} \tag{7.14}$$

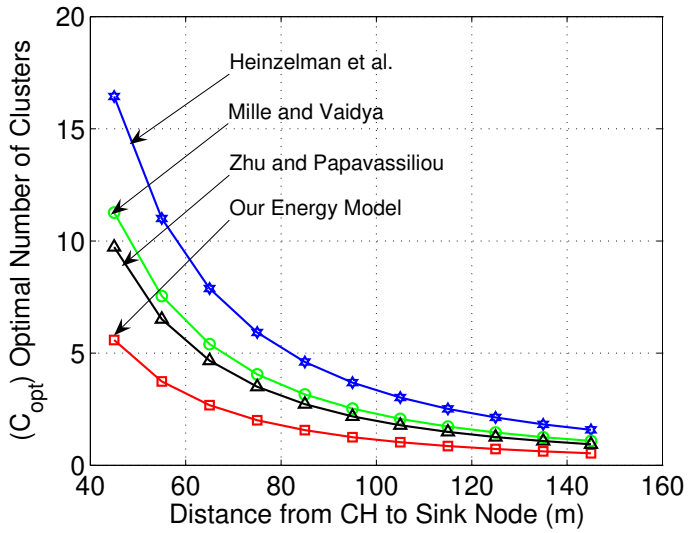


Figure 7.7: Optimal number of clusters with distance from CHs to sink node or base station. Analytical results for $N_s = 100$ nodes, $M = 100$, $E_{fs} = 7$ nJ/bit/m^2 and $E_{mp} = 0.0013$ pJ/bit/m^4 when $45 <$ distance < 145 is maintained.

We adopt the assumption of [73] that the BS is far from sensor nodes and therefore the distance between CH to the BS for all CHs can be considered equal. By differentiating (7.14) with respect to k and equating to zero,

$$\frac{\partial(E_{our})}{\partial(k)} = 0,$$

the resulting optimal number of clusters, C_{opt}, is

$$C_{opt} = \left\lceil \frac{\sqrt{N}}{\sqrt{6}} \frac{M}{d_{toBS}^2} \sqrt{\frac{E_{fs}}{D_\alpha}} \right\rceil, \tag{7.15}$$

where $D_\alpha = (E_{mp} + E_{sens_{CH}} + E_{tran_{CH}} + E_{logg_{CH}})$. Knowing C_{opt}, for a given network, we can evaluate the average radius of a circular cluster as $M/\sqrt{\pi C_{opt}}$, the average length of square cluster as $M/\sqrt{C_{opt}}$, and the circum-radius of the hexagonal cluster is $\sqrt[4]{\frac{4}{27}} M/\sqrt{C_{opt}}$. Providing these cluster shape alternatives allows designers to choose the one appropriate for their work.

According to Heinzelman et al. [73], the total energy during each round, E_{Hein},

Table 7.2: Optimal number of clusters with different energy models, for $N_s = 100$ sensor nodes and $M = 100$, when the distance between the CH and the sink node is $45 - 145$ m

Energy Model	$E_{fs} = 7$ [nJ/bit/m^2]		$E_{fs} = 10$ [pJ/bit/m^2]	
	C_{opt} $45 < d_{toBS} < 145$	C_{opt} Simulation	C_{opt} Range $45 < d_{toBS} < 145$	C_{opt} Simulation
Our Energy Model	1-6	3	0-2	1
Zhu et al.	1-12	5	0-4	2
Mille et al.	2-13	6	0-5	2
Heinzelman et al.	2-19	11	1-7	3

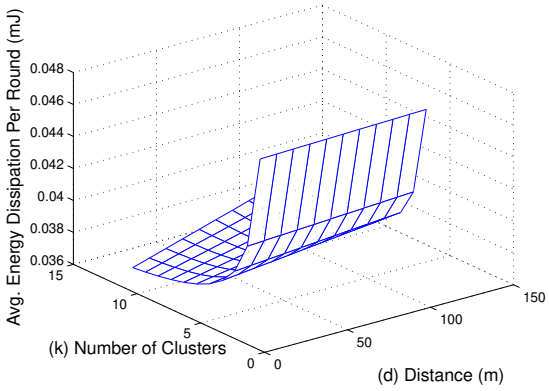
is given by

$$E_{Hein} = b\left(E_{elec}N_s + E_{pro_{CH}}N_s + d_{toBS}^4 E_{mp}k + E_{elec}N_s + E_{fs}\frac{1}{2\pi}\frac{M^2}{k}\right). \quad (7.16)$$

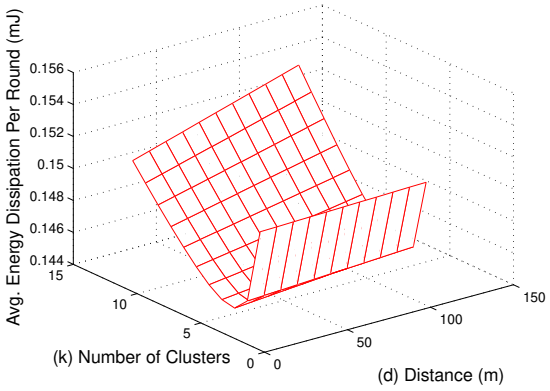
In the following we compare the average energy dissipation of our energy model and that of other energy models in [73, 120, 205].

In Fig. 7.7 and Table 7.2, we show that in our proposed energy model, the range of the optimal number of clusters is $1 < C_{opt} < 6$, while according to the energy models in [73, 120, 205], the range is within $2 < C_{opt} < 16$, when the distance between the CH and the sink node is between $45 - 145$ m. Our simulation results agree with this analysis. For example, for a distance of 55 m, the overestimation of the optimal number of clusters is 194.15% [73], 101.62% [120], and 74.06% [205]. This shows the difference between the models. Nevertheless, the difference in optimal number of clusters between these energy models gets closer as the distance between the sink and the network increases, as shown in Fig. 7.7. This is because, as this distance increases the energy dissipation for communication becomes more and more dominant in the cost function.

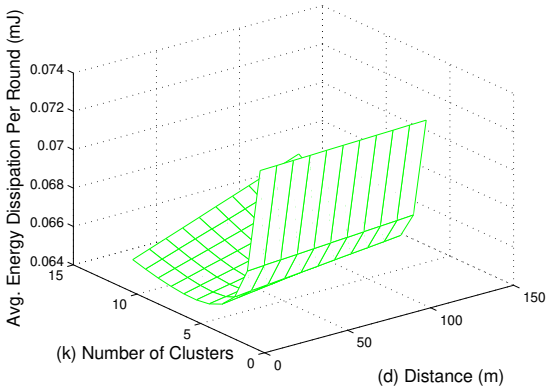
An optimal number of clusters for a given number of sensors is found by varying the distance and comparing the energy dissipation per round. From the simulation results, it is confirmed that the optimal number of clusters is three for [73], two for [120] and [205], and one for our energy model when $E_{fs} = 10$ pJ/bit/m^2 with 100 node network. Therefore, we can conclude that clustering



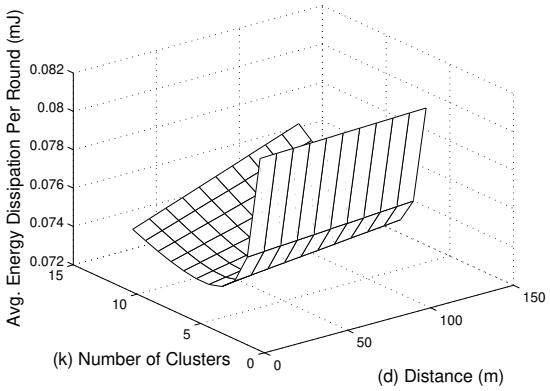
(a) Average energy dissipation versus no. of clusters and distance for Heinzelman et al. energy model.



(b) Average energy dissipation versus no. of clusters and distance for our energy model.



(c) Average energy dissipation versus no. of clusters and distance for Mille and Vaidya energy model.



(d) Average energy dissipation versus no. of clusters and distance for Zhu and Papavassiliou energy model.

Figure 7.8: Average energy dissipation versus number of clusters and distance with $N_s = 100$ nodes, $M = 100$, $E_{fs} = 7$ nJ/bit/m^2 and $E_{mp} = 0.0013$ pJ/bit/m^4 for different energy models

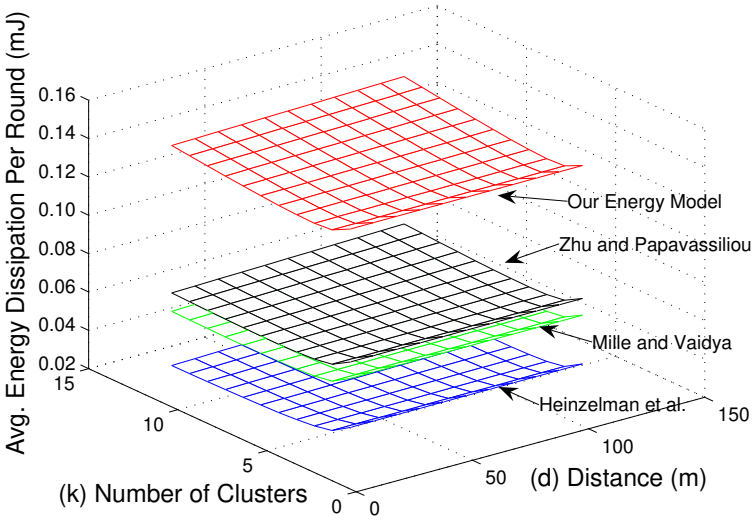


Figure 7.9: All energy models with average energy dissipation versus number of clusters and distance with $N_s = 100$ sensor nodes, $M = 100$, $E_{fs} = 7$ nJ/bit/m^2 and $E_{mp} = 0.0013$ pJ/bit/m^4

will not increase efficiency when the deployment area is small.

Figure 7.6 shows the average energy dissipation versus number of clusters when $E_{fs} = 7$ nJ/bit/m^2 and $N_s = 100$ nodes, $M = 100$ using (7.13). The optimal number of clusters, based on their energy models, are indicated with arrows. According to Fig. 7.6, optimizing the number of clusters does have significant effect on sensor network lifetime.

We also observe different results for the optimal number of clusters by comparison to [73]. In Fig. 7.6, for example, we observe that the optimal number of clusters of the our energy model is 3 while the energy models of [73], [120], and [205] will lead to optimal number of clusters to be 11, 6, and 5, respectively, for the same free space fading energy. The main reasons for this variation lie in our use of a more comprehensive energy model with a more realistic estimation of processing energy which turned out to be higher than the value considered in [73, 120, 205].

From Figures 7.8(a)− 7.8(d), we can observe that:

1. the number of clusters varies with the energy model used, as well as the distance from the CH to the base station,

2. energy dissipation varies with the number of clusters, and therefore, we can find the number of clusters that minimizes energy dissipation.

As shown in Fig. 7.9, the energy dissipation varies with the energy model used.

7.5 Simulation Set-up

In our simulations, we consider a sensor network with $N_s = 100$ sensor nodes. Consider our deployment area as the square $\{(0,0),(0,100),(100,0),(100,100)\}$ as in [27, 73, 164]. The base station or sink node is located in the coordinate $(50, 175)$ which is outside the deployment area and connected to an external power supply. Initially, CHs are randomly placed within an 50 m $\times$ 50 m square placed in the middle of the 100 m $\times$ 100 m deployment area. All other sensor nodes in each cluster are randomly uniformly distributed in a circle of 25 m radius around their respective CH.

We generated 1000 random setups, each with the above simulation setup. Therefore, each simulation's data point is obtained by averaging over 1000 random setups. We assume that the total number of sensors in the entire network N_s is 100, and each node reports data once every 300 ms ($T_{tr} = 0.3$ s). The channel bandwidth was set to 1 Mb/s as in [73], and each single packet size is $b = 2$ kb, as in [72], which maintains a low average data rate requirement per node (< 12 bps). We assume the energy dissipation for actuation, E_{actu}, is 0.02 mJ as in [141] and energy for starting up the radio, E_{ini} is 1 μJ as in [140].

Note that we do not account for energy dissipation in re-transmitting because of the packets collided in the simulations. As in [120], for our simulation, we used Mica2 Motes hardware values [5] and time values are based on radio's data sheet [44]. We assume that, consistent with a LEACH application, a CH and a sensor node have the same radio. Sensor node's sleeping time, $T_S = 299$ ms, and

Table 7.3: Parameter values used in energy model

Symbol	Description	Value
N_{cyc}	Number of clock cycles per task	0.97×10^6 [75]
C_{avg}	Avg. capacitance switch per cycle	22 pF [44]
V_{sup}	Supply voltage to sensor	2.7 V [44]
f	Sensor frequency	191.42 MHz [185]
n_p	Constant: depend on the processor	21.26 [75]
n	Path loss exponent	2 or 4 [75]
I_0	leakage current	1.196 mA [75]
V_t	Thermal voltage	0.2 V [185]
b	Transmit packet size	2 Kb [72]
E_{elec}	Energy dissipation: electronics	50 nJ/bit [75]
E_{amp}	Energy dissipation: power amplifier	100 pJ/bit/m^2 [75]
T_{tranON}	Time duration: sleep $\rightarrow$ idle	2450 μs [120]
$T_{tranOFF}$	Time duration: idle $\rightarrow$ sleep	250 μs [120]
I_A	Current: wakeup mode	8 mA [5]
I_S	Current: sleeping mode	1 μA [5]
T_A	Active time	1 ms [120]
T_S	Sleeping time	299 ms [120]
T_{tr}	Time between consecutive packet	300 ms
T_{sens}	Time duration: sensor node sensing	0.5 ms
I_{sens}	Current: sensing activity	25 mA
I_{write}	Current: flash writing 1 byte data	18.4 mA [163]
I_{read}	Current: flash reading 1 byte data	6.2 mA [163]
T_{write}	Time duration: flash writing	12.9 ms [163]
T_{read}	Time duration: flash reading	565 μs [163]
E_{actu}	Energy dissipation: actuation	0.02 mJ [141]
E_{ini}	Energy dissipation: initial set up	1 μJ [140, 202]

Table 7.4: Weighting factor values h_i

h_i	Sensor Node	CH
h_1 (processing)	1	1.2
h_2 (communication)	1	1.2
h_3 (sensing)	1	1.1
h_4 (logging)	1	1.1

wakeup time, $T_A = 1$ ms, are considered as in [120]. Self-discharge of batteries is considered as 3% per year as in [166]. We conducted Matlab simulations with different parameter settings, described later.

Furthermore, we assume selection of the weighting factor, h_i, as in Table 7.4. Here we consider processing and communication energy of the CH is 20% more, and sensing and logging energy is 10% more than that of regular sensor nodes.

7.6 Simulation Results

7.6.1 Energy Comparison

In Fig. 7.10 we present a pie chart describing the energy consumption for communication, processing, transient, sensor loggings and sensing energy as 51%, 12%, 10%, 14% and 6% of the total energy respectively. All of these sources of energy consumption are considerable.

The same parameters, namely $N_s = 100$ sensor nodes, $M = 100$, $k = 10$ clusters, $E_{fs} = 10$ pJ/bit/m^2 and $E_{mp} = 0.0013$ pJ/bit/m^4, are used to generate Fig. 7.11, where we compare the effect of the difference energy models on the sensor network lifetime. Each simulations data point is obtained by averaging over 1000 random setups, but observe that the results does not change with single simulation.

We consider the actuation performed only for this pie chart (Fig. 7.10) and exclude it from all other simulations, for the purpose of fair comparison with the other energy models that also exclude it.

For all our experimental computation, we used an AAA size alkaline battery

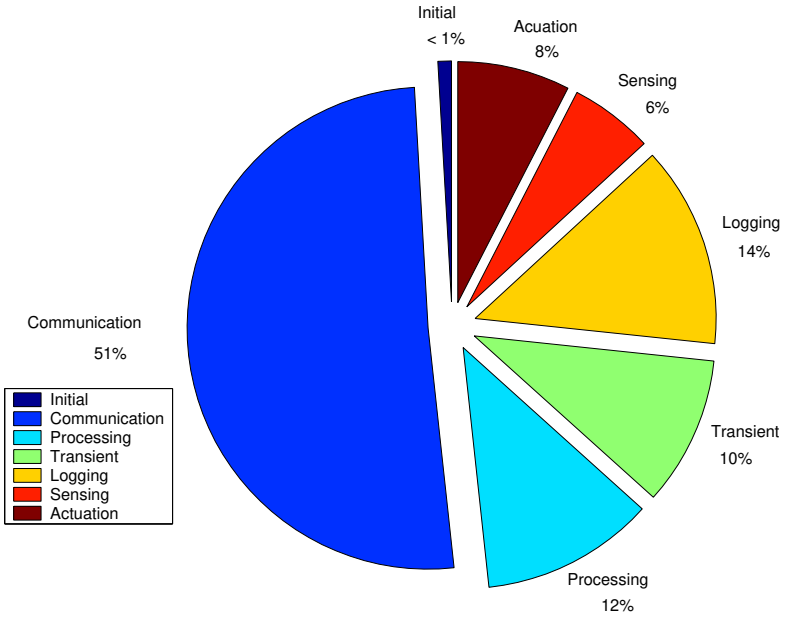


Figure 7.10: Energy consumption pie chart for any sensor in cluster j, when actuation is considered. (Here $N_s = 100$ sensor nodes, $M = 100$, $k = 10$ clusters, $E_{fs} = 10$ pJ/bit/m^2 and $E_{mp} = 0.0013$ pJ/bit/m^4)

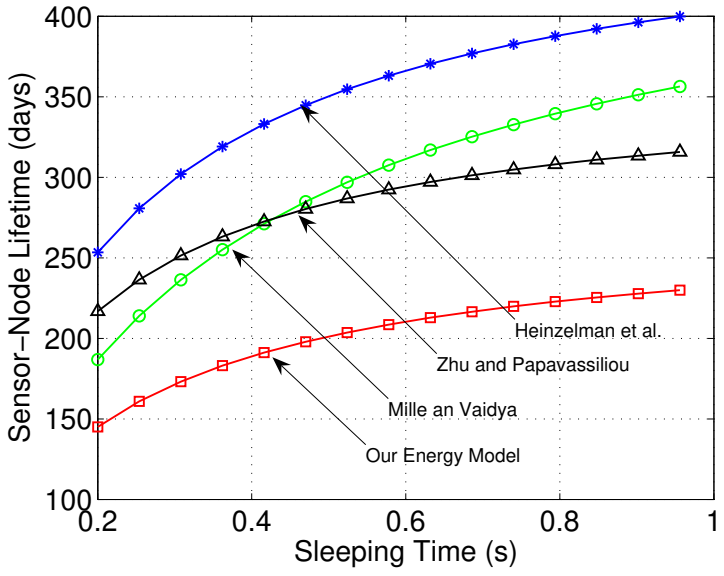


Figure 7.11: Sensor node lifetime verses sleeping time of the sensor node, with different energy models with AA alkaline batteries by using (7.4) and (7.5). (Here we consider for $N_s = 100$ sensor nodes, $M = 100$, $k = 10$ clusters, $E_{fs} = 10$ pJ/bit/m^2 [73] and $E_{mp} = 0.0013$ pJ/bit/m^4 [73])

with 700 mAh. However, we also repeated our simulation for an AA size alkaline battery with 1500 mAh, a C-cell battery with 5000 mAh and D-cell battery with 9000 mAh for 1.5 V and found that all results were consistent. Node lifetime can be computed by

$$\text{Node lifetime} = \frac{\text{initial battery capacity}}{\text{avg. current} \times 365 \times 24} \quad [\text{years}],$$

where the units of the initial battery capacity is mAh and the average current is mA.

In Fig. 7.11, we show that existing energy models overestimate life expectancy of a sensor node by 30-58%.

7.6.2 Effect of Free Space Fading Energy (E_{fs})

The optimal number of clusters derived in [73] is only applicable if the free space fading energy is assumed to be constant, which may not be the case in practice.

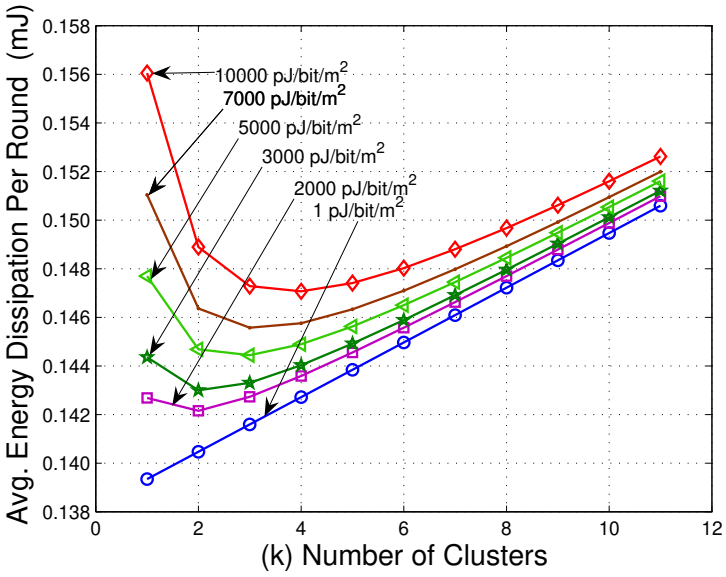
For this reason, we repeated the above simulation by varying free space fading energy, E_{fs} within the interval $(1 - 10^4)$ [pJ/bit/m^2] and observed, in Fig. 7.12(a) that when E_{fs} increases, the optimal number of clusters also increases.

We found that energy dissipation per round increases from 7.11% to 12.81% as the optimal number of clusters changes from 1 to 3. Moreover, we observed that when free space fading energy, $E_{fs} < 1670$ pJ/bit/m^2 the optimal number of clusters needed is one and thus clustering is not necessary.

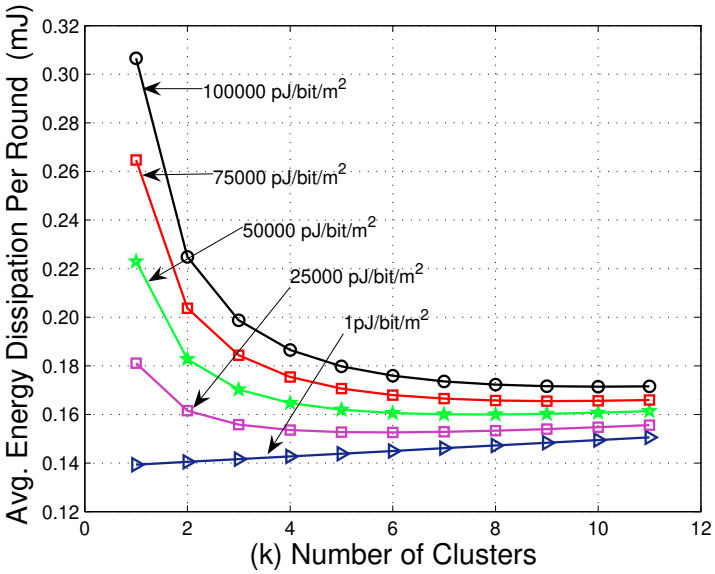
We repeated the simulation by increasing E_{fs} further, up to 10^5 pJ/bit/m^2, and confirm that the optimal number of clusters becomes more important with higher free space fading energy dissipation as shown in Fig. 7.12(b).

According to Figs 7.13(a)$-$ 7.13(d), the number of optimal clusters increases with the increase of free space fading energy, E_{fs}, for all the above mentioned four energy models.

Now we consider the same network but we vary the number of clusters k, free space fading energy, E_{fs}, and the distance from the CHs to the base station or sink node. Here we investigate the analytical results which derived in Section

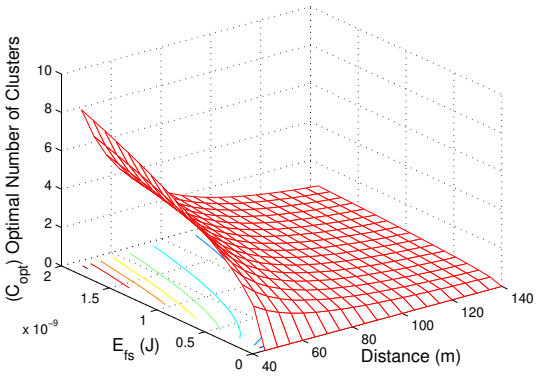


(a) When E_{fs} varies between $(1 - 10^4)$ pJ/bit/m^2.

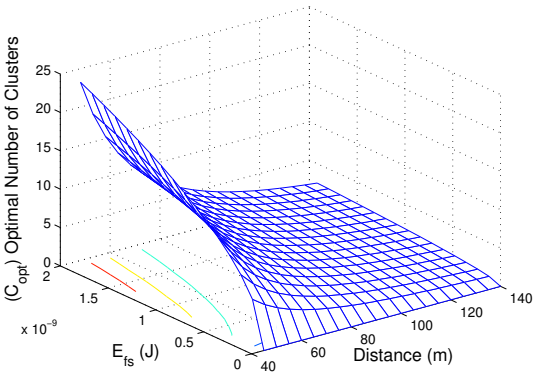


(b) When E_{fs} vary between $(1 - 10^5)$ pJ/bit/m^2.

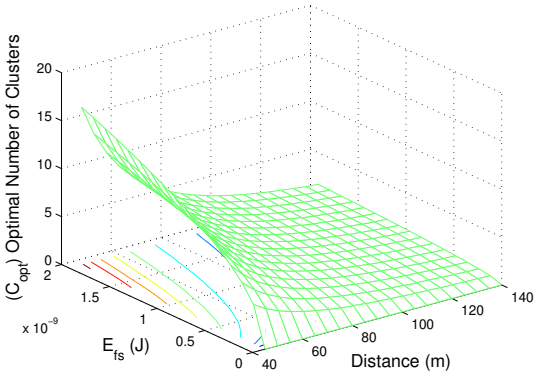
Figure 7.12: Average energy dissipation versus number of clusters when E_{fs} varies for our energy model. Both graphs show that when E_{fs} increases, the optimal number of clusters also increases



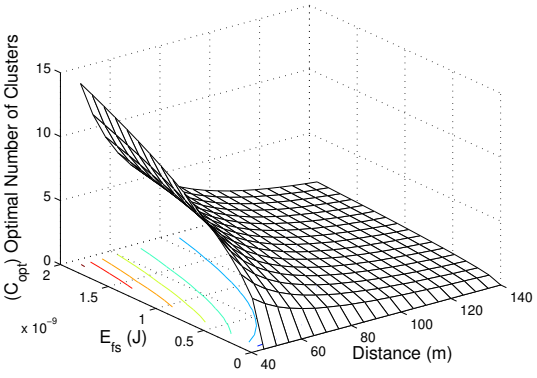
(a) Relationship between E_{fs}, C_{opt} and the distance for our energy model.



(b) Relationship between E_{fs}, C_{opt} and the distance for Heinzelman et al. energy model.

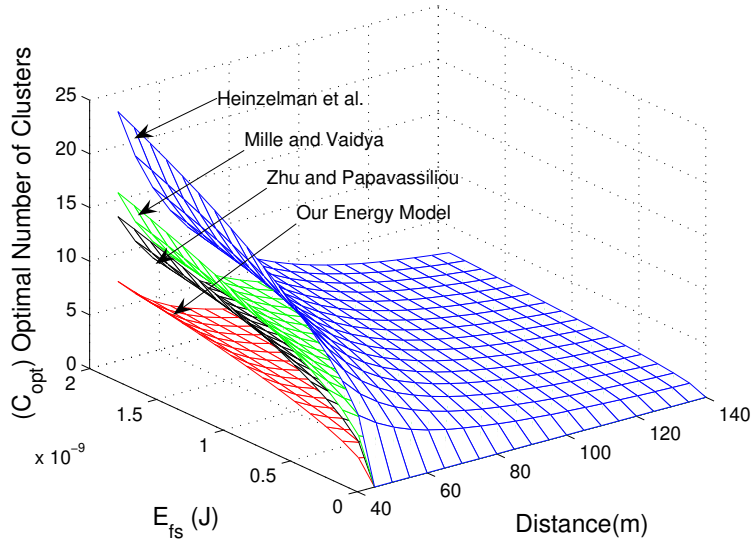


(c) Relationship between E_{fs}, C_{opt} and the distance for Mille and Vaidya energy model.

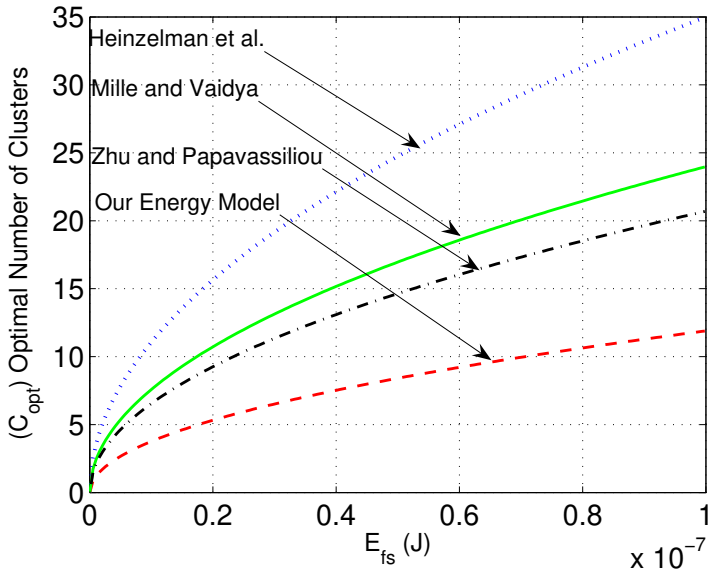


(d) Relationship between E_{fs}, C_{opt} and the distance for Zhu and Papavassiliou energy model.

Figure 7.13: Relationship between free space fading energy, E_{fs}, optimal number of clusters, C_{opt} and the distance for different energy models



(a) Relationship between free space fading energy, E_{fs}, optimal number of clusters, C_{opt} and the distance.



(b) Relationship between optimal number of clusters, C_{opt}, and free space fading energy, E_{fs}.

Figure 7.14: Variation between optimal number of clusters, C_{opt}, for different energy models

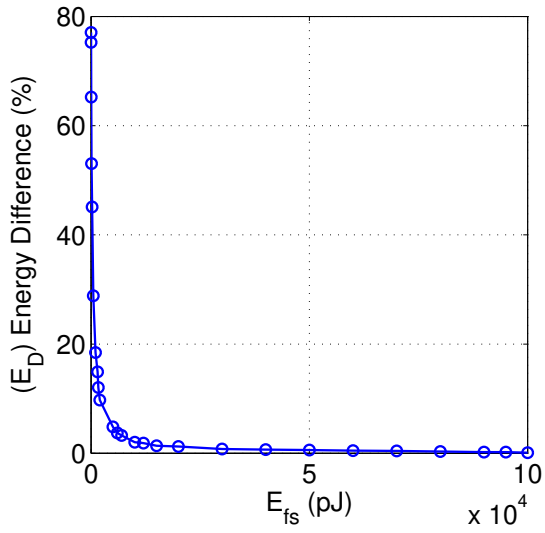


Figure 7.15: Different free space fading energy values, E_{fs}, versus percentage of total energy difference in the sensor node, E_D considering our and Heinzelman energy models with 100 sensors and $M = 100$ square root of the physical area. (This shows that there is a significant energy difference between energy models, when free space fading energy is low. When $E_{fs} = 10$ pJ/bit/m^2 [73], the energy difference $E_D = 75.27$ %)

7.3, 7.3.3 and 7.4 (see Fig. 7.13 − 7.15)

We observed that the optimal number of clusters decreases dramatically with the increasing distance. Energy differences between the models are shown in Fig. 7.14(a). For the particular case when distance = 100 m, the variation of the optimal number of clusters with free space fading energy is shown in Fig. 7.14(b).

7.6.3 Energy Difference E_D

Let E_D represent the percentage of energy difference between our energy model and the energy model in [73]. We found that the energy difference dramatically decreases when E_{fs} is contained within the interval $(1 - 5 \times 10^3)$ [pJ/bit/m^2] according to Fig. 7.15. We kept the optimal number of clusters as three and free space fading energy E_{fs} as 7×10^3 pJ/bit/m^2, and repeated the simulation. As shown in Fig. 7.16, we observed that the sensor lifetime is increased by 12.74% when the sleeping time is 0.2 s, and by 13.92% when the sleeping time is 1 s,

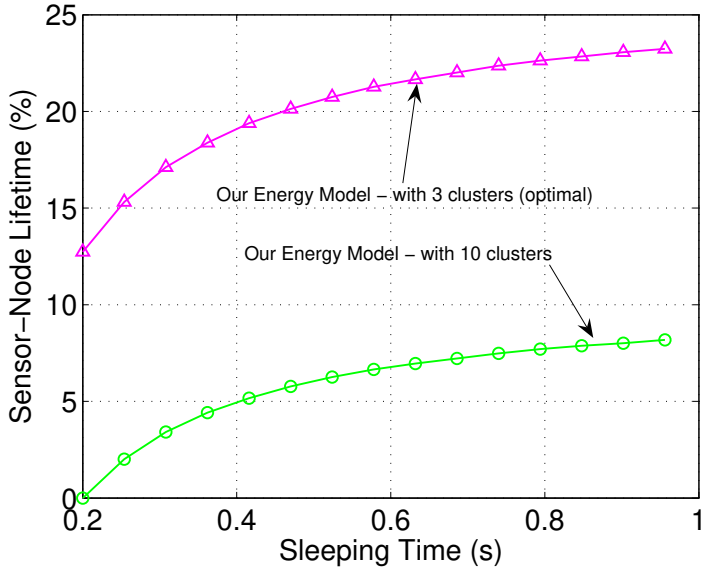


Figure 7.16: Sensor node lifetime, comparing our energy model using ten clusters and three clusters (optimal number of clusters $C_{opt} = 3$) when free space fading energy $E_{fs} = 7 \times 10^3$ pJ/bit/m^2 with an AAA alkaline battery.

when the number of clusters used is reduced from 10 (non optimal clusters) to 3 (optimal clusters).

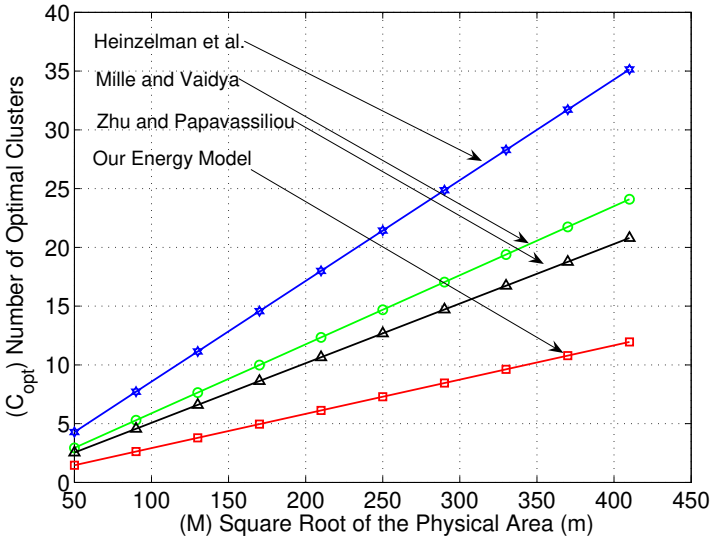
7.6.4 Effect of Physical Area of Sensor Network

Next, we vary the number of sensor nodes N_s and the physical area M to their effect on the distance, free space fading energy and the number of optimal clusters. In addition, we consider how sensor network lifetime can be maximized and the number of sensors required to design a network for a given lifetime.

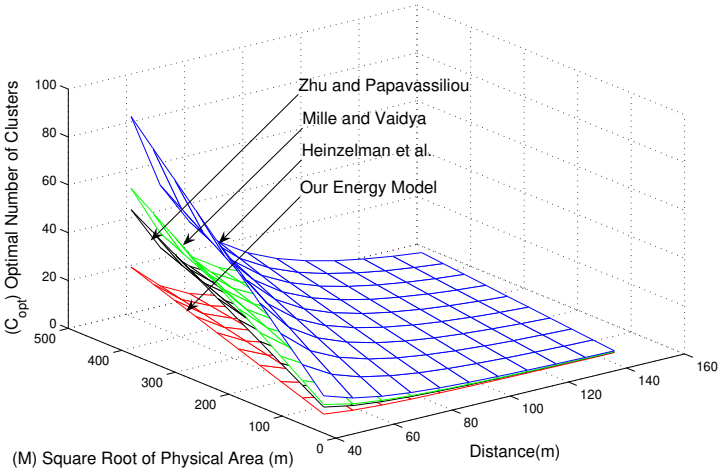
As demonstrated in Fig. 7.17(a), the optimal number of clusters varies linearly with M (the square root of the physical area). Importantly, the effect of free space fading energy, E_{fs}, becomes less by increasing the distance as in Fig. 7.17(b).

7.6.5 Effect of the Duty Cycle

Finally, we vary the number of duty cycles to investigate the affect on energy consumption (see Fig. 7.18). As shown in Fig. 7.18, when the number of duty



(a) C_{opt} versus square root of the physical area in all energy models.



(b) Variation in C_{opt} with the square root of the physical area in all models when the distance from CH to base station varies.

Figure 7.17: Variation between optimal number of clusters, C_{opt}, for different energy models, when $E_{fs} = 10$ pJ/bit/m^2

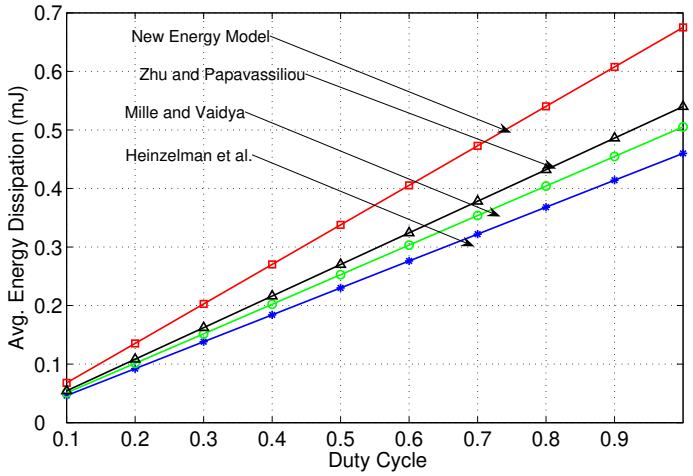


Figure 7.18: Average energy dissipation per sensor versus duty cycle in energy models

cycles increases, the average energy consumption of all models diverge. When the number of duty cycles is 1, the maximum overestimation of our energy model relative to the Heinzelman energy model is 46.77%.

7.6.6 Percentage of Alive Nodes

The lifetime of a sensor network depends on the application where the sensors are deployed. Therefore, we investigate how the number of live sensor nodes varies with the number of rounds or time. An overestimation is shown in Fig. 7.19, when the number of sensor nodes live is plotted against the number of rounds. The maximum overestimation of the death of the last node of our energy model relative to the Heinzelman energy model is 30.1%. We repeated the all above simulations for AA, C-cell and D-cell batteries and found the results to be consistent

7.6.7 Effects of Number of Sensors and Distance of Cluster Heads from Base Station

We investigate how the number of sensors affects the optimal number of clusters and the sensor node lifetime in a given physical area. The optimal number

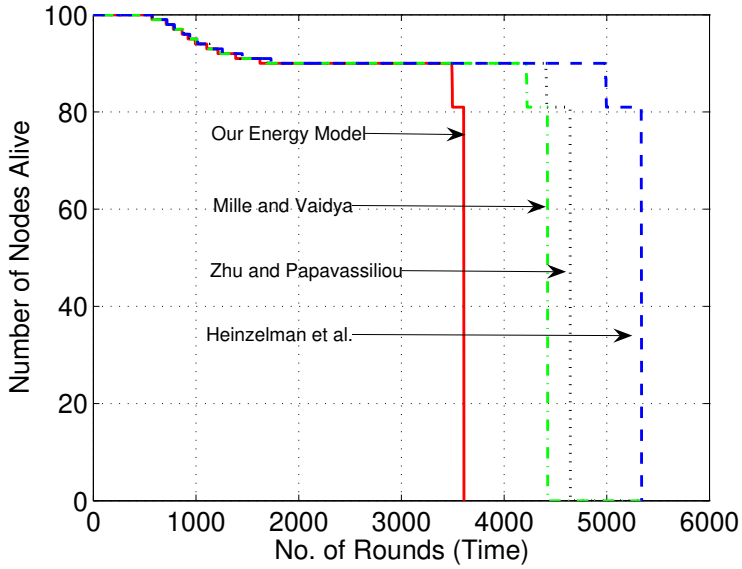


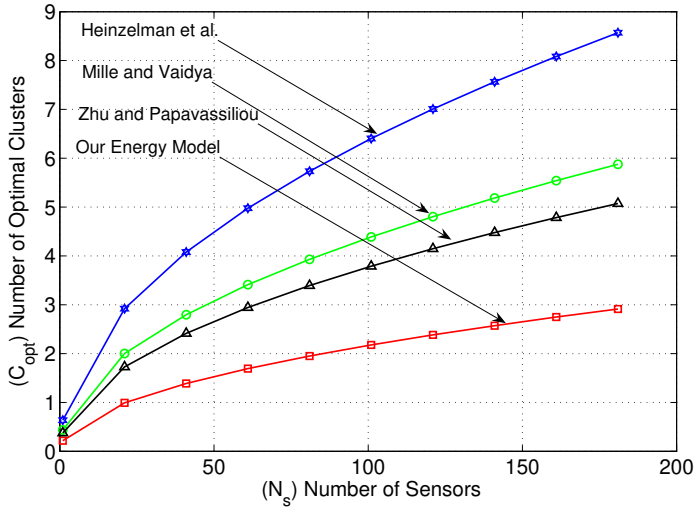
Figure 7.19: Variation in number of live nodes depending on number of rounds (time), comparing all energy models.

of clusters increases with the number of sensors used for all energy models as shown in Fig. 7.20(a) and decreases with the distance as in Fig. 7.20(b). It is clear that this change is far less for the proposed energy model in comparison to the Heinzelman method. According to our energy model, a moderate increase in the number of sensors used may not change the design parameters concerning the optimal number of clusters.

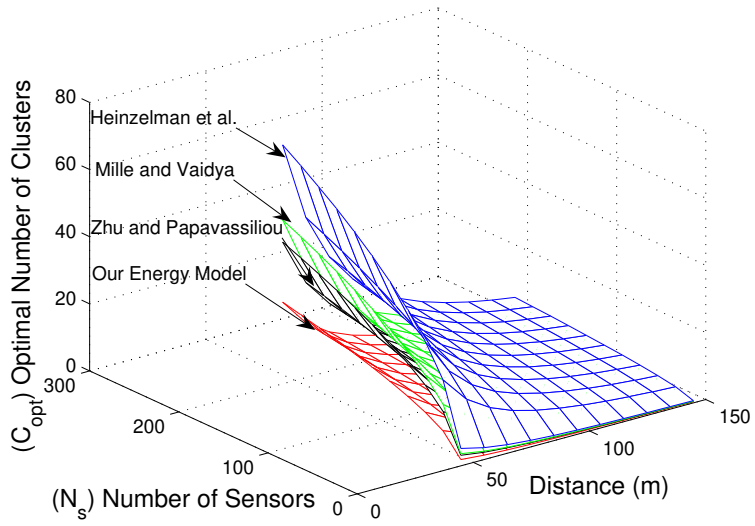
By (7.1), $E \propto d^2$, one may expect that when the number of sensors is doubled the sensor lifetime is multiplied by four. However, as we show this is not the case. Consider sensor deployment with uniform distribution.

According to Fig. 7.21, all energy dissipations converge when the number of sensors is increased. Therefore, the change in energy dissipation with respect to the change in the number of sensors becomes less. It should be noted that sensor node lifetime is inversely proportional to the total energy consumption of the sensor node:

$$\text{Sensor lifetime} \propto \frac{1}{\text{Total Energy Consumption}},$$



(a) C_{opt} versus number of sensors (analytical results).



(b) C_{opt} versus the number of sensors and the distance from CH to base station.

Figure 7.20: Variation in optimal number of clusters, C_{opt}, for different energy models, for a square root of the physical area $M = 100$ and $E_{fs} = 10$ pJ/bit/m^2

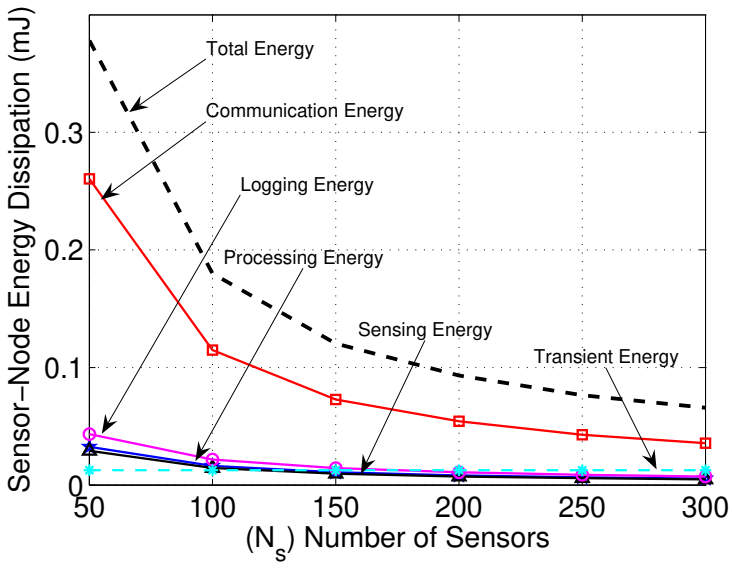


Figure 7.21: All energy components with sensor node lifetime versus number of sensors for $E_{fs} = 10$ pJ/bit/m^2 and a square root of the physical area $M = 100$, for uniform deployment, with increasing number of sensors

Table 7.5: How different energy sources change with increasing number of sensors with uniform sensor deployment, when $E_{fs} = 10$ pJ/bit/m^2

N_s	Communication Energy [μJ]	Processing Energy [μJ]	Sensing Energy [μJ]	Logging Energy [μJ]	Transient Energy [μJ]
50	0.2605	0.0326	0.2916	0.0434	0.0126
100	0.1478	0.0164	0.1458	0.0217	0.0126
150	0.0727	0.0108	0.0097	0.0144	0.0126
200	0.0543	0.0081	0.0073	0.0108	0.0126
250	0.0428	0.0065	0.0058	0.0087	0.0126
300	0.0357	0.0054	0.0049	0.0072	0.0126

Table 7.6: Sensor node lifetime when increasing number of sensors, with uniform sensor deployment, when $E_{fs} = 10$ pJ/bit/m^2

N_s	C_{opt}	Radius $\frac{M}{\sqrt{\pi C_{opt}}}$ [m] (Circular Cluster)	Length $\frac{M}{\sqrt{C_{opt}}}$ [m] (Square Cluster)	Length $\sqrt[4]{\frac{4}{27}}\frac{M}{\sqrt{C_{opt}}}$ [m] (Hexagonal Cluster)	Avg. Distance [m] Simul. d_{toCH}	Sensor Node Lifetime without E_{tran_N} (days)	Sensor Node Lifetime with E_{tran_N} (days)	Relative Increase in Lifetime
50	1	56.4190	100	62.0403	36.9804	118.86	109.70	1
100	2	39.8942	70.71	43.8691	25.5901	256.35	217.25	1.9802
150	3	32.5735	57.73	35.8190	20.1241	395.88	309.76	2.8236
200	3	32.5735	57.73	35.8190	19.1248	529.37	385.91	3.5176
250	4	28.2095	50	31.0202	17.5014	667.43	454.44	4.1423
300	4	28.2095	50	31.0202	17.5125	800.86	512.59	4.6723

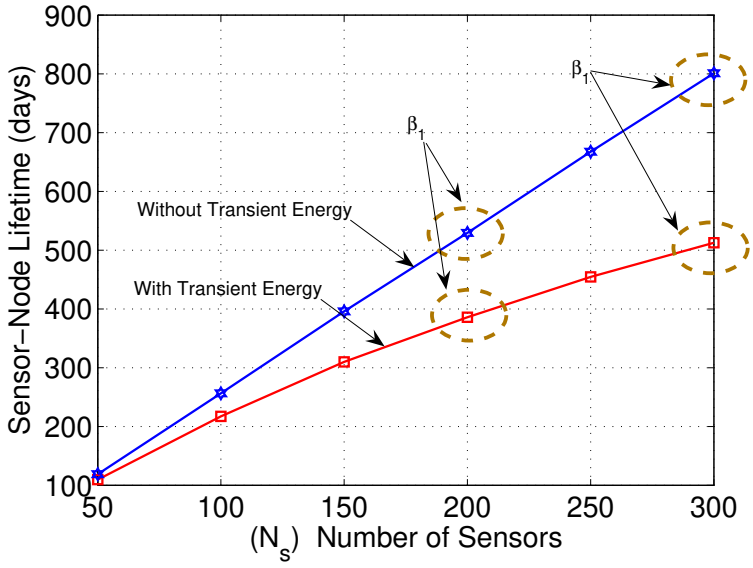


Figure 7.22: Senor node lifetime versus number of sensors with and without transient energy. (Here $E_{fs} = 10$ pJ/bit/m^2 and a square root of the physical area $M = 100$, for uniform sensor deployment)

according to which, the plot of sensor node lifetime should approximately resemble the inverse of the plot of total energy consumption. As can be seen from Fig. 7.21, all energy consumption types, except transient energy, decrease with the increasing number of sensors.

According to Fig. 7.22, two plots describing the sensor node lifetime versus the number of sensors can be observed for the two cases, with and without transient energy. Furthermore, according to Fig. 7.22, it can be observed that points such as β_1 and β_2 characterize the deviation from a linear relationship between sensor node lifetime and the number of sensor nodes. This can be verified when observing the same number of optimal clusters assigned with increasing number of sensors in Table 7.6. For example, the same number of optimal clusters ($C_{opt} = 4$) is assigned when the number of sensors are 150 and 200, making the cluster radius stay the same. Interestingly, the simulation also shows that the average cluster radius increases from 17.5014 to 17.5125 m, also increasing the sensor lifetime. This is due to the decrease in the number of bits to be transmitted.

Note that when the number of sensors increases, the average number of bits sent by each sensor decreases so that the total amount of information in a network is kept constant at 10^5 bits.

Knowing C_{opt} and M for a given network, we evaluate and present the average radius of a circular cluster in the 4^{th} column of Table 7.6, the average length of a square cluster in the 5^{th} column, and the average circum radius of a hexagon in 6^{th} column. Network designers can then use these values to optimize network lifetime.

We then investigate how the optimal number of clusters, C_{opt}, varies as a function of the number of sensors and area when one million sensors are deployed, as shown in Table 7.7 and Fig. 7.23. We observe that the optimal number of clusters increases with the increased number of sensors and area. For example, it is observed that 58 clusters are needed for a million sensors and no clustering for 100 sensors, with $M = 100$ m, when the distance from the CH to the sink node or base station is 145 m when $E_{fs} = 10$ pJ/bit/m^2 is maintained. We also observe

that network lifetime can be increased up to 7.3337 years for a million sensors by assuming that the minimum required number of bits to be transmitted is 200 per round. We cannot decrease the number of bits of the data to be transmitted, as the number of sensors grow, or sensors will fail to transmit the information. Depending on the application we may increase the sensor sleep time instead of reducing the number of bits.

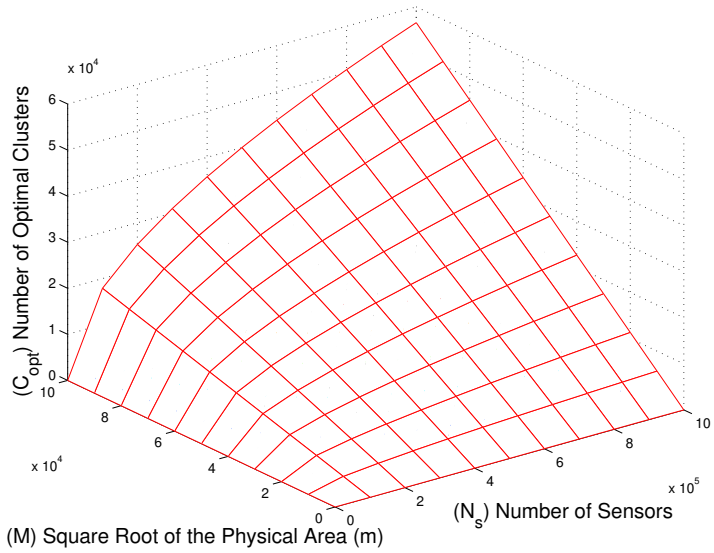


Figure 7.23: Optimal number of clusters (C_{opt}) versus square root of the physical area (M), and number of sensors (N_s), when distance from CHs to sink node or base station is 145 m and, $E_{fs} = 10$ pJ/bit/m^2 is maintained

Table 7.7: Optimal clusters (C_{opt}) with physical area (M) and number of sensors (N_s), when $E_{fs} = 10$ pJ/bit/m^2

(M) Square Root of Physical Area [m]	(N) Number of Sensors					
	100	400	800	10×10^3	10×10^4	10×10^5
100	1	2	2	6	19	58
200	1	3	4	12	37	116
300	1	4	5	18	55	174
400	2	5	7	24	74	232
500	2	6	9	29	92	290

7.7 Chapter Summary

We developed a new, realistic and comprehensive energy model for wireless sensor networks. The energy consumption between different sources in the considered set up of a sensor node were analyzed. The results indicate that simple energy models overestimate the real sensor node lifetime. This chapter has applied our model to a LEACH-type protocol to obtain an accurate evaluation of the energy consumption and node lifetime. This work leads to new, interesting and useful ways to maximize sensor network lifetime. It shows that the energy consumption estimated by Heinzelman et al. [73] overestimated the average life of a sensor node by 51-58%, Zhu and Papavassiliou [205] by 32-41%, and Mille and Vaidya [120] by 30-35%.

This chapter concludes that the number of clusters does not play a significant role for moderately sized sensor networks if the free space fading energy is low. For large networks, on the other hand, cluster optimization is still important even if the free space fading energy is low. Moreover, it has shown that that the optimal number of clusters is very sensitive to the energy model used. This chapter has observed that overestimation of the last node death is 30.1% when the number of alive sensor nodes is plotted against the number of rounds (time). It also provides an estimation of the number of sensor nodes needed to design a network for a given lifetime, with all the important factors that influence the life expectancy of a sensor network. Based on the work presented in this chapter, designers can optimize energy efficiency subject to their required specifications.

Several important conclusions arise from this chapter:

- Less comprehensive energy models used in many papers overestimate the real sensor node lifetime.
- For moderately sized sensor networks the number of clusters do not play a significant role if the free space fading energy is low.
- For large networks, on the other hand, cluster optimization is still important

even if free space fading energy is low.

- The optimal number of clusters is very sensitive to the energy model used.

Following these conclusions the next chapter considers the use of the comprehensive energy model developed in this chapter.

Table 7.8: Optimal cluster with increasing E_{fs} with 100 sensors and $M = 100$ physical area

E_{fs} [pJ/bit/m^2]	C_{opt}	Radius = $\frac{M}{\sqrt{\pi C_{opt}}}$ [m] (Circular Cluster)	Length = $\frac{M}{\sqrt{C_{opt}}}$ [m] (Square Cluster)	Circum Radius = $\sqrt[4]{\frac{4}{27}}\frac{M}{\sqrt{C_{opt}}}$ [m] (Hexagonal Cluster)
1	1	56.4190	100	62.0403
10 [73]	1	56.4190	100	62.0403
1000	1	56.4190	100	62.0403
1500	1	56.4190	100	62.0403
1670	1	56.4190	100	62.0403
2000	2	39.8942	70.7107	43.8691
5000	2	39.8942	70.7107	43.8691
5010	2	39.8942	70.7107	43.8691
7000	3	32.5735	57.7350	35.8190
10000	3	32.5735	57.7350	35.8190
12000	4	28.2095	50	31.0202
15000	4	28.2095	50	31.0202
20000	5	25.2313	44.7214	27.7453
30000	6	23.0329	40.8248	25.3279
40000	7	21.3244	37.7964	23.4490
50000	8	19.9471	35.3553	21.9346
60000	8	19.9471	35.3553	21.9346
70000	9	18.8063	33.3333	20.6801
80000	10	17.8412	31.6228	19.6189
90000	10	17.8412	31.6228	19.6189
95000	11	17.0110	30.1511	18.7059
100000	11	17.0110	30.1511	18.7059

CHAPTER EIGHT

Efficient Battery Management for Sensor Lifetime

8.1 Introduction

Extending the lifetime of a sensor network is important for most if not for all applications. The sensor network life expectancy depends on various factors including the initial battery capacity of individual sensor nodes, the amount of processing that can occur and the amount of information that can be collected, the layout of the sensor network, the number of sensors involved and the location of the sink node or base station. The battery, as the widely used power provider of the sensors in the network, is considered the key factor for achieving a prolonged life. Carefully scheduling and budgeting battery power in sensor networks has become a critical issue in network design.

Assuming that all other parameters are equal, we consider two ways of increasing the sensor network life expectancy:

1. Use of higher energy batteries for selected nodes

2. Efficient scheduling and budgeting battery power

Several research efforts [40–43] have proposed energy-efficient battery management techniques that can improve the energy efficiency of radio communication devices. We investigate the effect of two levels of batteries, one level for cluster head (CH) and the other level for nodes, on the sensor network's lifetime.

8.1.1 Motivation

It is a challenging task to accurately monitor a remote environment by combining data from thousands of micro sensors. It is also challenging to maintain a long network lifetime if the sensors are battery powered.

The LEACH is an application specific communication protocol based on clustering of sensor nodes developed in a recent PhD thesis at MIT [71]. Clustering with rotating cluster heads, as proposed in LEACH, is a well-known way to meet these challenges. It is clear in this case that the "handoff problem" in sensor networks cannot be isolated from the "cell selection" (in this case cluster placement) problem, if the cells (clusters) are updated dynamically, which is the case in an ad hoc wireless sensor network. It is useful to investigate whether commercially available batteries can be used for such a battery management strategy; for example taking one level of batteries for CH and the other level for nodes. The questions this work attempts to answer are: How does the network lifetime varies with the ratio of the two levels of batteries? Is it a good option to use multiple batteries per cluster head? How does the number of sensors and the position of the sink node influence the network lifetime? How can LEACH (or a similar communication protocol) be supplemented to obtain a longer network lifetime using an effective battery management strategy? What are the limitations? We may supplement LEACH by adopting the two levels of batteries. How does the ratio of the two levels of batteries influence the use of LEACH for different applications? For example, sensors may be dropped from an aircraft, increasing their high deployment cost. The last question this chapter attempts to answer is: What is the total network cost if we vary parameters such as battery cost and deployment cost?

In Section 8.2, the system model is described followed by Section 8.5 describes how the sensor network lifetime can be extended using commercially available high energy batteries in High Powered Cluster Heads (HPCH). The network cost is observed by varying different parameters such as battery cost and deployment cost. In Section 8.6 the simulation setup is described. Section 8.7 provides the

simulation results and the chapter is concluded in Section 8.8.

8.1.2 Assumptions

The key assumptions used in this chapter are the same as the ones used in Chapter 7. Refer to Section 7.1.3 for details.

8.2 System Model

Consider a k cluster sensor network where the clusters are laid out in a directed tree topology so that its root is a base station (sink node) as shown in Fig. 7.1. Cluster j comprises one CH, denoted CH_j, and n_j sensor nodes, $j = 1, 2, ..., k$. Hence, the total number of sensors is $N_s = \sum_{j=1}^{k}(n_j + 1)$.

Here we assume that N_s sensor nodes are randomly and uniformly distributed in a $M \times M$ region. Thus, on average number of sensor nodes (including CH) in a cluster is (N_s/k), meaning that some clusters may have $\lceil N_s/k \rceil$ or $\lfloor N_s/k \rfloor$ sensor nodes. Let d_j be the distance between CH_j and the next CH (or the sink node) that CH_j transmits to, and let d_{ij} be the distance between node i in cluster j and CH_j.

Using (7.5) and $n_j = (N_s/k), \forall j$, we can rewrite the total energy consumed by cluster CH_j, $E_{CH}(j)$ as:

$$
\begin{aligned}
E_{CH}(j) = \Bigg[& \underbrace{E_{ini}}_{initial} + \underbrace{h_1 b N_{cyc} C_{avg} V_{sup}^2}_{switching} \left(\frac{N_s}{k}\right) \\
& + \underbrace{b V_{sup} \left(I_0 e^{\frac{V_{sup}}{n_p V_t}}\right) \left(\frac{N_{cyc}}{f}\right)}_{leakage} \left(\frac{N_s}{k}\right) + \underbrace{h_2 b E_{elec}}_{receive} \left(\frac{N_s}{k} - 1\right) \\
& + \underbrace{h_2 b E_{elec} + b d_j^n E_{amp}}_{transmit} + \underbrace{h_3 b V_{sup} I_{sens} T_{sens}}_{sensing} \\
& + \underbrace{T_{CH} V_{sup} \left[d_{CH} I_A + (1 - d_{CH}) I_S\right]}_{transient} + \underbrace{E_{actu} N_{act}}_{actuation} \\
& + \underbrace{h_4 b V_{sup} \left(I_{write} T_{write} + I_{read} T_{read}\right)}_{data-logging} \Bigg], \qquad (8.1)
\end{aligned}
$$

where $d_{CH} = (T_{tranON} + T_{A_{CH}} + T_{tranOFF})/(T_{tranON} + T_{A_{CH}} + T_{tranOFF} + T_{S_{CH}})$, $T_{A_{CH}} =$ CH's wakeup time, $T_{S_{CH}} =$ sleeping time, N_{act} is the number of actuations per CH and $h_1, h_2, h_3, h_4 > 1$ as in Table 7.4. It should also be noted that for LEACH, energy consumption for cluster head rotation E_{rotate} is added to (8.1).

LEACH [72] also supports single hop transmission with periodically rotating cluster heads balancing the energy consumption. However, it assumes that all nodes are capable of data processing and long distance communication. Furthermore, the nodes should be able to communicate between clusters and support different MAC protocols. As in [168], time is divided into rounds. During each round, we assume that every sensor node generates a fixed-sized packet and that all generated packets are forwarded to the base station by the cluster head. The base station ensures collision avoidance by scheduling transmission time based on time division multiple access (TDMA). Sensors are generally placed around cluster heads randomly and hence are not equidistant from cluster heads. However, by calculating the total energy consumption of each sensor, we can estimate the average energy consumptions for all the nodes considering load balancing.

In the next section the network lifetime and the factors influencing it are observed when High Powered Cluster Heads (HPCHs) are used, without using LEACH. We describe the selection of a suitable number of sensors and the energy ratio between a sensor and a CH, that gives maximum sensor lifetime as well as how the distance from CH to base station affects sensor node lifetime.

8.3 Effect of Energy Ratio on Sensor Lifetime

Here we consider that each sensor node transmits data to its CH during each round. Therefore, from (7.4), the total energy consumed by a sensor node during each round, as in [70], is

$$E_{node} = \left[E_{ini} + bE_{elec} + bd_{toCH}^2 E_{fs} + E_{tran_N} + bE_{sens_N} + bE_{logg_N}\right], \qquad (8.2)$$

where b is the number of bits in every packet, d_{toCH} is the distance between node and CH, E_{fs} is the free space fading energy. Similarly, from (8.1), the total energy consumed by a CH during each round is

$$\begin{aligned} E_{head} &= \left[E_{ini} + bE_{pro_{CH}}\left(\frac{N_s}{k}\right) + h_2 bE_{elec}\left(\frac{N_s}{k} - 1\right)\right. \\ &\quad + h_2 bE_{elec} + bd_{toBS}^4 E_{mp} + E_{tran_{CH}} \\ &\quad \left. + bE_{sens_{CH}} + bE_{logg_{CH}}\right], \end{aligned} \tag{8.3}$$

where and E_{mp} is the multipath fading energy. Note that we consider a multipath model with d^4 power loss, and assume that actuation is not performed. From (8.2) and (8.3) the energy dissipation in a single cluster during each round is

$$E_{cluster} = E_{head} + \left(\frac{N_s}{k} - 1\right) E_{node}. \tag{8.4}$$

Here R_E represents the average energy consumption ratio of the normal sensor node to the cluster head as

$$R_E = \frac{E_{node}}{E_{head}}, \tag{8.5}$$

allowing us to investigate how energy ratio affects the sensor network lifetime. As proposed in [66], it is appropriate to use high energy batteries for cluster heads. In this chapter we attempt to find the optimum energy ratio between a CH and a sensor for a given lifetime.

From (8.4) and (8.5), we obtain

$$E_{cluster} = E_{head}\left[1 + \left(\frac{N_s}{k} - 1\right) \times R_E\right]. \tag{8.6}$$

The total energy during each round based on the above energy model is obtained using (8.2) and (8.6), the total energy is given by

$$\begin{aligned} E_{network} &= kE_{cluster}, \\ &= [k + (N-k)R_E]\left(E_{ini} + bE_{pro_{CH}}\left(\frac{N_s}{k}\right)\right. \end{aligned}$$

$$+ \quad h_2 b E_{elec} \left(\frac{N_s}{k} \right) + b d_{toBS}^4 E_{mp} k$$
$$+ \quad E_{tran_{CH}} + b E_{sens_{CH}} + b E_{logg_{CH}} \Big) . \tag{8.7}$$

We adapt the assumption of [73] that the distance between CH to the base station or sink node for all CHs can be considered identical as the base station is far from sensor nodes. By using (8.7), we can evaluate network lifetime by varying the number of sensors and the energy ratio.

It should be noted that the number of sensors is inversely proportional to the number of transmit bits per sensor:

$$\text{no of sensors} \propto \frac{1}{\text{transmit bits per sensor}}.$$

In next Section 8.4, we investigate how energy ratio and battery ratio, the ratio of initial battery capacities for sensors and cluster heads, affects sensor node lifetime in a given physical area using the same network layout.

8.4 Effect of Battery Ratio and Energy Ratio on Sensor Lifetime

Battery life greatly affects the overall network communication performance [112]. In this section we analyze the effect of the ratio of initial battery capacities for sensors and cluster heads, R_B, on sensor network lifetime. As in [99], the corresponding ratio R_E of energy used between a sensor node and a CH is determined by the application. In contrast to R_E, R_B does not depend on the application. We define R_B of the sensor node, b_{node}, and the CH, b_{head},

$$R_B = \frac{b_{node}}{b_{head}}. \tag{8.8}$$

Lifetime of the sensor node, L_{node}, can be calculated as

$$L_{node} = \frac{b_{node}}{\left(\frac{E_{node}}{V_{node} T_{node}}\right)}, \tag{8.9}$$

where V_{node} is the supply voltage of sensor node with battery capacity b_{node} mAh and T_{node} is the time taken to transmit one packet from sensor to CH. Similarly, the lifetime of the cluster head, L_{head} is given by

$$L_{head} = \frac{b_{head}}{\left(\frac{E_{head}}{V_{head} T_{head}}\right)}, \tag{8.10}$$

where V_{head} is the supply voltage of CH with battery capacity b_{head} mAh and T_{head} is the time taken to transmit one packet from CH to neighboring CH or sink node.

In an ideal network, the lifetime of sensor nodes and CHs should be equal. Therefore,

$$L_{node} = L_{head}. \tag{8.11}$$

Using (8.5), (8.8), (8.9), (8.10) and (8.11), we obtain

$$\frac{R_B}{R_E} = \frac{V_{head} T_{head}}{V_{node} T_{node}}.$$

The supply voltage of the sensor and the CH should be same within one system. If the base station is far away from the sensor field and at a fixed location, we obtain $T_{head} > T_{node}$,

$$R_B = k_1 R_E,$$

where $k_1 = T_{head}/T_{node} \geq 1$ is a constant. Therefore, we can conclude that $R_B \geq R_E$.

Conversely, when $R_B > k_1 R_E$, a possible conclusion is that the CH may completely drain off its battery and die before the nodes, and nodes cannot transmit data to the sink node or base station due to isolation.

This also may occur, when $R_B < k_1 R_E$ leads to the death of nodes before the cluster heads, and there is no information to transmit to the sink node.

We also analyze how the distance from CH to sink node affect the sensor network lifetime. As in [73],

$$E_{tx_{CH}}(h_2, b, d) = \underbrace{h_2 b E_{elec}}_{electronics} + \underbrace{b d_j^n E_{amp}}_{amplifier}, \tag{8.12}$$

where $h_2 > 1$ is a weighting factor, E_{elec} is the energy dissipated to transmit or receive electronics, E_{amp} is the energy dissipated from the power amplifier, $n = 2$ for free space fading and $n = 4$ for multipath fading. We can show that the energy consumption increases exponentially with increase in distance, especially for distances greater than 20 m.

We can further analyze how the number of sensor nodes and distance from CH to sink node, d, affect the sensor network lifetime, for a given physical area. We define the percentage difference of sensor network lifetime, R_L, using 100 to 200 sensors, as

$$R_L = \left(\frac{L_{200} - L_{100}}{L_{100}} \right) \times 100\%,$$

where L_{100} and L_{200} represent the network lifetime using 100 and 200 sensor nodes respectively.

8.5 Supplementing LEACH with Power Management

A major drawback of the LEACH algorithm is the overheads due to cluster head rotation, which we attempt to reduce by the new method proposed in this chapter. If the cluster heads at the beginning of the network operation can be selected by the user, they can be equipped with stronger batteries than the sensor nodes. The rotation of cluster heads may then be unnecessary, as long as the battery capacity of the cluster heads is higher than the sensor nodes.

We propose a way to extend network lifetime by introducing several special

sensors with higher battery power than normal sensors. We use these special sensors as cluster heads until their battery capacity is reduced to that of a normal sensor node before adopting a LEACH type method.

8.5.1 High Powered Cluster Heads (HPCH)

Consider a sensor network with a number of n_S sensors and a number of n_{CH} cluster heads. Let l denote the number of rounds and let L_n be the lifetime of the sensor network or time (in number of rounds) taken until the remaining battery power of the sensor node reaches zero. We consider alkaline AAA size batteries with b_{AAA} power and AA size batteries with b_{AA} power [66].

In this work, we consider HPCH with high battery power, b_{AA}, and normal sensor nodes with low power, b_{AAA}, as opposed to homogeneous, medium-power sensor nodes in a sensor network. We use HPCH as non-rotating cluster heads until the average mAh of HPCH reaches the level of average mAh of normal sensors or the threshold. After the threshold is reached LEACH is used. Let us assume that the threshold is reached at the round when $l = l^{th}$. At the threshold point or $l = l^{th}$ round, $HPCH(l^{th})$ is

$$HPCH(l^{th}) = b_{AAA}(l^{th}),$$

where $HPCH(l^{th})$ is the battery capacity of HPCH in mAh at the l^{th} round. Here we can calculate the initial battery capacity of the LEACH, b_{LEACH} as follows:

$$b_{LEACH} = \frac{n_{CH}b_{AA} + n_S b_{AAA}}{n_{CH} + n_S}. \tag{8.13}$$

From (7.4) the average energy consumed by the normal sensor nodes is defined by

$$E_1 = \frac{\sum_{j=1}^{k}\sum_{i=1}^{n_j} E_N(ij)}{n_S}, \tag{8.14}$$

and from (8.1) the average energy consumed by the cluster heads is defined by

$$E_2 = \frac{\sum_{j=1}^{n_{CH}} E_{CH}(j)}{n_{CH}}. \tag{8.15}$$

Using (8.14), the remaining battery capacity of the normal sensor node, P_n at round l is defined by

$$P_n^{(l)} = b_{AAA}^{(l-1)} - \frac{E_1^{(l)}}{V_{sup}}, \tag{8.16}$$

where $l = 1, 2, ..., L_n$. Similarly, using (8.15), the remaining battery capacity at round l up to round l^{th} at the threshold point is defined by

$$HPCH^{(l)} = b_{AA}^{(l-1)} - \frac{E_2^{(l)}}{V_{sup}}, \tag{8.17}$$

where $l < l^{th}$. The remaining battery capacity at round l after threshold point ($l \geq l^{th}$), with LEACH is defined by

$$HPCH^{(l)} = b_{AAA}^{(l-1)} - \frac{E_2^{(l)}}{V_{sup}}. \tag{8.18}$$

For a HPCH scheme with j CHs ($j = 1, 2, 3, 4$), we define $R_D(j)$ as

$$R_D(j) = \left(\frac{T_{HPCH}(j) - T_{LEACH}}{T_{LEACH}} \right) \times 100\ \%, \tag{8.19}$$

where $T_{HPCH}(j)$ is the number of rounds or the lifetime of HPCH for j^{th} cluster head and T_{LEACH} is the number of rounds or the lifetime of LEACH.

8.5.2 Network Cost Minimization with Different Battery Capacities

Our aim in this section is to analyze minimum network cost by varying different parameters, such as battery cost and deployment cost.

The total cost of the above network is given by $\sum_i N_i C_i$, where N_i is the number

of sensors of battery type i and C_i is the purchase cost of a node with battery type i. However, if sensors are dropped remotely (for example from an aircraft), then the deployment cost needs to be increased. Therefore total cost is defined by

$$C_{tot} = \sum_i N_i C_i + \sum_i D_i(N_i), \tag{8.20}$$

where $D_i(N_i)$ is a function representing the deployment cost of N_i sensors of battery type i. Moreover, here we define α_{cost} as the ratio of the cost of a AA battery C_{AA} and the cost of a AAA battery C_{AAA}, i.e.,

$$\alpha_{cost} = \frac{C_{AA}}{C_{AAA}}. \tag{8.21}$$

The total dollar cost minimization is subjected to the following conditions:

1) the minimum number of sensors required should be greater than or equal to the number of sensors needed to cover the target area;

2) the required network lifetime should be less than the actual network lifetime. This is equivalent to the requirement that the total number of rounds required should be less or equal to the maximum number of rounds possible until the battery capacity of the network reaches zero level as in Fig. 8.6(a).

8.6 Simulation Set-up

We use a comprehensive energy model, as in [70], for our experiments and conducted Matlab simulations. The supply voltage V_{sup} is 1.5 V. It is assumed that each node reports data once every $T_{tr} = 300$ ms. The channel bandwidth was set to 1 Mb/s as in [73], and each single packet size is $b = 2$ Kb, which maintains the average data rate requirement per node (< 12 bps). The energy for starting up the radio, $E_{ini} = 1$ μJ as in [140].

For our simulation we use the above proposed energy model. We do not account for energy dissipation in re-transmitting due to the packets colliding in the simulations. As in [120], we used Mica2 Motes hardware values [5] and time

Table 8.1: Parameter values used in all experiments

Symbol	Description	Value
N_{cyc}	Number of clock cycles per task	0.97×10^6 [75]
C_{avg}	Avg. capacitance switch per cycle	22 pF [44]
V_{sup}	Supply voltage to sensor	2.7 V [44]
f	Sensor frequency	191.42 MHz
n_p	Constant: depend on the processor	21.26 [75]
n	Path loss exponent	2 or 4
I_0	leakage current	1.196 mA [75]
V_t	Thermal voltage	0.2 V
b	Transmit packet size	2 Kb [72]
E_{elec}	Energy dissipation: electronics	50 nJ/bit [75]
E_{amp}	Energy dissipation: power amplifier	100 pJ/bit/m^2 [75]
T_{tranON}	Time duration: sleep $\rightarrow$ idle	2450 μs [120]
$T_{tranOFF}$	Time duration: idle $\rightarrow$ sleep	250 μs [120]
I_A	Current: wakeup mode	8 mA [5]
I_S	Current: sleeping mode	1 μA [5]
T_A	Active time	1 ms [120]
T_S	Sleeping time	299 ms [120]
T_{tr}	Time between consecutive packet	300 ms
T_{sens}	Time duration: sensor node sensing	0.5 ms
I_{sens}	Current: sensing activity	25 mA
I_{write}	Current: flash writing 1 byte data	18.4 mA [163]
I_{read}	Current: flash reading 1 byte data	6.2 mA [163]
T_{write}	Time duration: flash writing	12.9 ms [163]
T_{read}	Time duration: flash reading	565 μS [163]
E_{ini}	Energy dissipation: initial set up	1 μJ [140]
E_{actu}	Energy dissipation: actuation	0.05 mJ [141]
E_{rotate}	Energy dissipation: CH rotation	0.02 mJ
n_S	number of sensors	90
n_{CH}	number of cluster head	10
k	number of clusters	10
b_{AAA}	Initial battery capacity alkaline AAA	700 mAh
b_{AA}	Initial battery capacity alkaline AA	1500 mAh

values based on the radio data sheet [44].

A sleeping time of $T_S = 299$ ms, and a wakeup time of $T_A = 1$ ms, are considered for each sensor node. The self-discharge of a battery is assumed as 3% per year as in [166]. We generate 1000 random setups, each with the above experiment setup. Therefore each simulations data point is obtained by averaging over 1000 random set-ups.

Network lifetime can be calculated by dividing the mAh of the battery by the total average current, multiplied by the number of hours in a single year.

8.6.1 Energy Ratio and Battery Ratio for HPCH

For our experiments, we consider a 100 sensor network ($N_s = 100$) with 10 clusters ($k = 10$) each of which has 10 sensors (including the CH). Consider a deployment area to be the square $\{(0,0),(0,100),(100,0),(100,100)\}$ [metres] as in [73] and [164]. The base station or sink node is located in the coordinate $(50,175)$ which is outside the deployment area and connected to an external power supply. Initially, CHs are randomly placed within an 50 m $\times$ 50 m square placed in the middle of the 100 m $\times$ 100 m deployment area. All other sensor nodes in each cluster are randomly uniformly distributed in the circle of 25 m radius of their respective CH. We did not consider actuation for our simulation.

8.6.2 HPCH with LEACH

We considered AAA size alkaline batteries with 700 mAh. In our simulations, we consider a sensor network with $N = 100$ number of sensors and a deployment area to be the square $\{(0,0),(0,160),(160,0),(160,160)\}$. The base station or sink node is located in the coordinate $(80,170)$, which is outside the deployment area and connected to an external power supply. Initially, the 10 CHs are randomly placed within an 80 $\times$ 80 square placed in the middle of the 160 $\times$ 160 deployment area. All other 9 sensor nodes in each cluster are randomly and uniformly distributed in the circle of 40 m radius of their respective CH. Clusters are formed randomly, but fixed numbers of clusters are formed with equal sizes.

Table 8.2: Optimum values for different distances

Distance (m)	(N_s) Sensors	No. of Rounds (lifetime)	(R_L) Ratio (%)
5	200	7.244×10^5	3.14
10	200	7.243×10^5	3.18
50	300	6.973×10^5	27.18
100	500	6.830×10^5	257.55
150	700	6.623×10^5	522.19
200	900	6.412×10^5	631.91
300	1400	6.114×10^5	685.41
400	1700	6.099×10^5	695.31

We assume that a CH and a sensor node have the same radio. This is consistent with a LEACH application. We considered AAA size alkaline batteries with 700 mAh and AA size alkaline batteries with 1500 mAh for 1.5 V. We admit that the choice of these batteries are only for the purposes of simulation, and may not be the ideal choice for many sensor networks.

We conducted Matlab simulations with two different parameter settings: Experiment 1 and Experiment 2. For HPCH, in both experiments, the initial battery capacity of a cluster head is assumed to be 1500 mAh. Once the battery power of the cluster head is reduced to that of a normal sensor node, the LEACH method will be used.

8.7 Simulation Results

8.7.1 Analysis of Energy Ratio and Battery Ratio for HPCH

The energy consumption increases exponentially after about 20 m as shown in Fig. 8.1. Therefore, if the distance is less than 20 m, a relatively long network lifetime can be expected. However, the variation of the energy consumption decreases with the increasing distance according to (8.12). This can be due to very low constant values involved in (8.12).

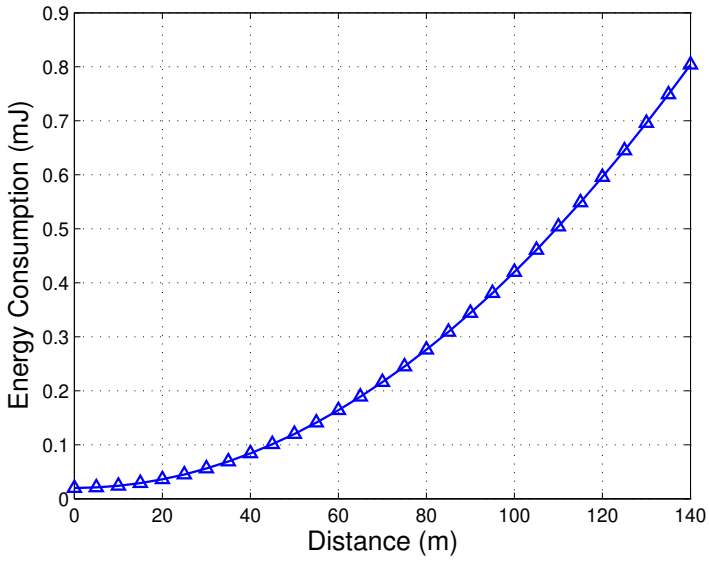


Figure 8.1: Distance versus energy consumption for communication

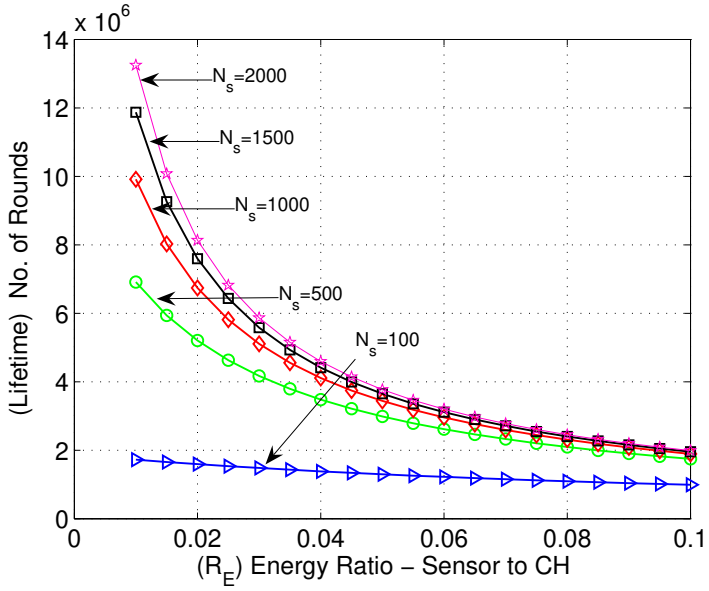


Figure 8.2: Number of rounds (sensor node lifetime) versus the energy ratio, R_E, between sensor to CH, for different number of sensors.

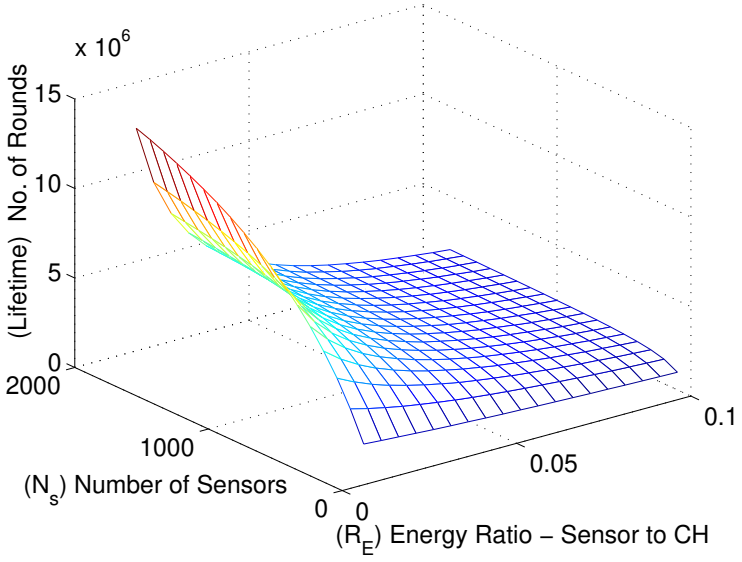


Figure 8.3: Number of rounds (sensor node lifetime) versus energy ratio, R_E, between sensor node and the CH, where the number of sensors in a given physical area $M = 100$. Here the average distance from CH to base station or sink node is 22 m.

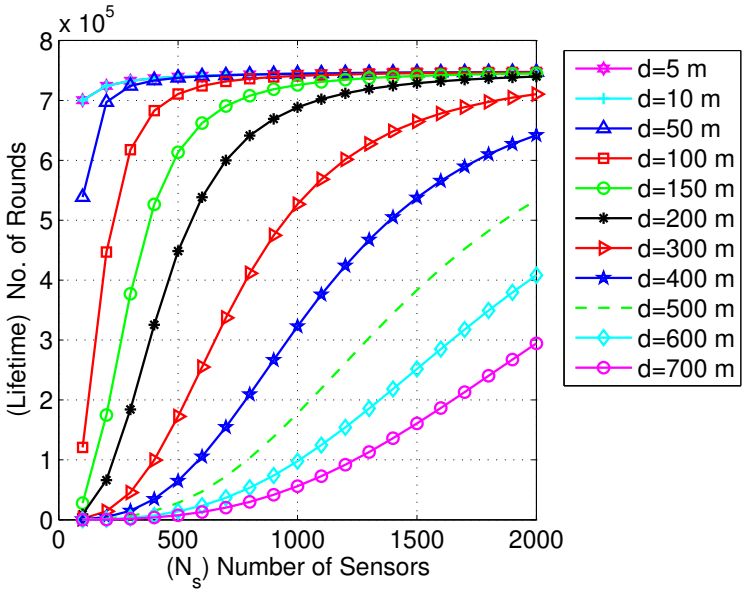


Figure 8.4: Number of rounds (sensor node lifetime) versus the number of sensors in a given physical area $M = 100$ by varying the distance from CH to base station or sink node

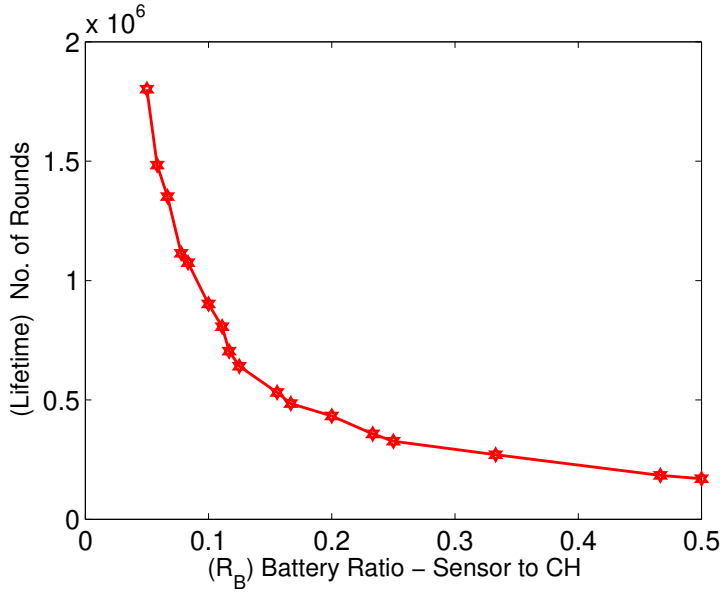


Figure 8.5: Battery ratio, the ratio of initial battery capacities for sensors and CH versus number of rounds (sensor node lifetime)

In Fig. 8.2, sensor node lifetime increases with the increasing number of sensors and decreases with increasing energy ratio, R_E. Therefore, node lifetime can be calculated for the different energy packs used. According to this result, suitable batteries can be selected for sensors and CHs, for example, the AAA battery for sensors and the AA for CHs.

Node lifetime is illustrated in Fig. 8.3. Node lifetime always increases with the increasing number of sensors but decreases with increasing energy ratio, R_E. In our simulations, the average distance from the CH to the base station or sink node is 22 m. As shown in Fig. 8.1 energy dissipation dramatically increases after about 20 m. Therefore, after about 20 m we can expect the sensor node lifetime not to increase heavily.

Moreover, optimum parameter setting for each distance can be obtained from the "knee" point of the corresponding curves (similar to [68, 192]). For example the "knee" point for the distance 150 m, is 700 sensors and 6.623×10^5 number of rounds (lifetime) as shown in Fig. 8.4. In Table 8.2 we compare the different distances from CH to base station or sink node at their "knee" point values.

Table 8.3: Sensor network lifetime with different battery ratio by considering 700 mAh node battery capacity and same application

CH Battery Capacity (mAh)	Battery Ratio $R_B = \frac{b_{node}}{b_{head}}$	No. of Rounds (lifetime)
1400	0.5000	1.7009×10^5
1500	0.4667	1.8412×10^5
2100	0.3333	2.7066×10^5
2800	0.2500	3.2732×10^5
3000	0.2333	3.5758×10^5
3500	0.2000	4.5250×10^5
4200	0.1667	4.8480×10^5
4500	0.1556	5.3103×10^5
5600	0.1250	6.4111×10^5
6000	0.1167	7.0449×10^5
6300	0.1111	8.0691×10^5
7000	0.1000	9.0209×10^5
8400	0.0833	10.7500×10^5
9000	0.0778	11.1390×10^5
10500	0.0667	13.5170×10^5
12000	0.0583	14.8320×10^5
14000	0.0500	18.0130×10^5

According to the distance from CH to sink node we can choose the number of sensors needed to obtain the maximum network lifetime. Similarly, if we know the time period we need to obtain information, we can select the maximum number of sensors needed. This may save the total cost of the network.

Furthermore, we investigate how battery capacity influences the network's lifetime. We keep the application the same to maintain the same energy ratio and the same sensor battery at 700 mAh, while changing the battery of the CH from 1400 mAh to 14000 mAh. According to Fig. 8.5 and Table 8.3, sensor network lifetime dramatically decreases with increasing battery ratio, R_B. Based on this work we can choose the required number of sensors and the suitable batteries for sensors and CHs, if we use high powered CHs, within a required lifetime.

A substantial amount of energy can be wasted if we do not carefully select batteries for battery powered sensors. More importantly, this investigation allows the network designer to specify the required CH selection which optimizes energy usage, and therefore can save the total network cost.

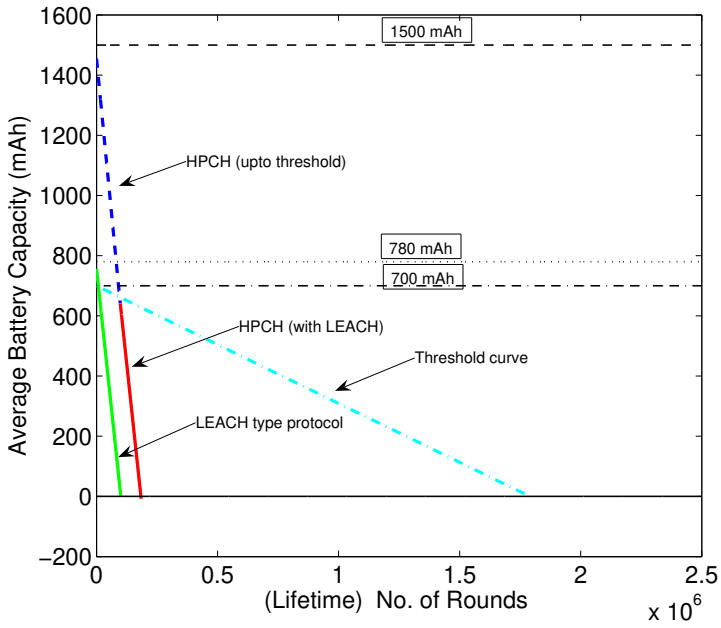
8.7.2 Analysis of High Powered Cluster Heads with LEACH

Figure 8.6 illustrates the network lifetime for various schemes. Equation (8.16) was used to obtain the curve for normal sensor node, (8.17) for HPCH up to the l^{th} threshold point, and (8.18) for HPCH after l^{th} threshold point in Fig. 8.6.

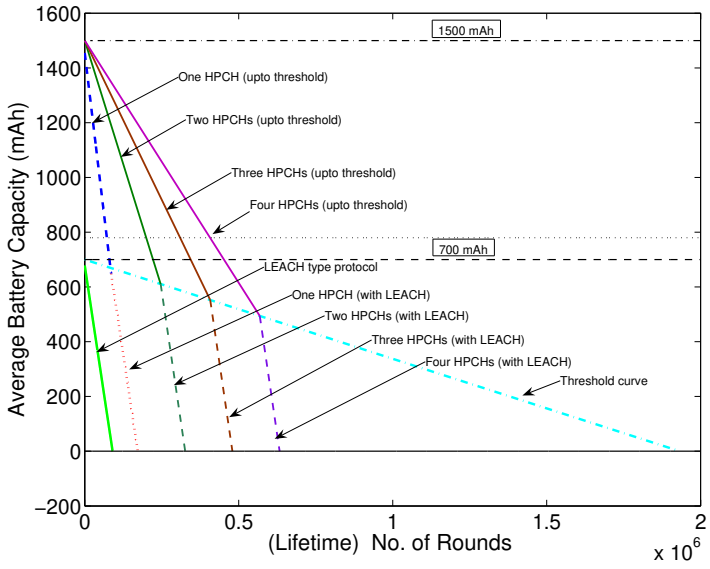
Therefore, energy is saved by keeping the same topology until LEACH is applied. The initial battery capacity of a normal sensor is assumed to be 700 mAh for HPCH in Experiment 1. As the 10 high powered nodes are equipped with 1500 mAh and the other 90 sensors with 700 mAh, the battery capacity of a sensor in LEACH is computed based on (8.13), as $(10 \times 1500 + 90 \times 700)/100 = 780$ mAh in Experiment 1.

The use of 780 mAh may not be a perfect choice as there is no such battery available in the market. In order to conduct a fair comparison, it is assumed that batteries are available with any number of mAh. Figure 8.6(a) shows the average battery capacity versus the number of rounds for various options in Experiment 1. It is shown that the lifetime of a normal sensor node is much higher than that of a cluster head. Based on Fig. 8.6(a), it can be concluded that the network lifetime with a the HPCHs is much higher than that of the LEACH-type protocol. Since the price difference between a AAA size alkaline battery with 700 mAh and a AA size alkaline battery with 1500 mAh is about 30 cents, the network lifetime increase per dollar can be as large as $4.2/0.30 = 14\%$. Note that prices of batteries may vary as they depend on market demand and various other factors.

We acknowledge that there is no standard size alkaline battery with 780 mAh. We assume the availability of such a battery only for the purposes of comparison, and to demonstrate the benefit of our approach. Since there is no 780 mAh battery currently available, we replaced 780 mAh batteries with 700 mAh for LEACH in



(a) Experiment 1



(b) Experiment 2

Figure 8.6: Network lifetime comparison with experiments 1 and 2

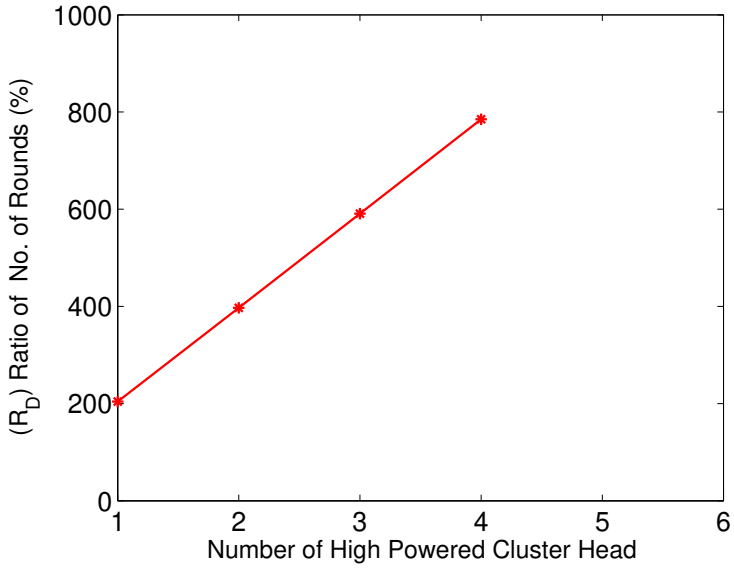


Figure 8.7: Network lifetime comparison; number of HPCH with j cluster heads versus $R_D(j)$

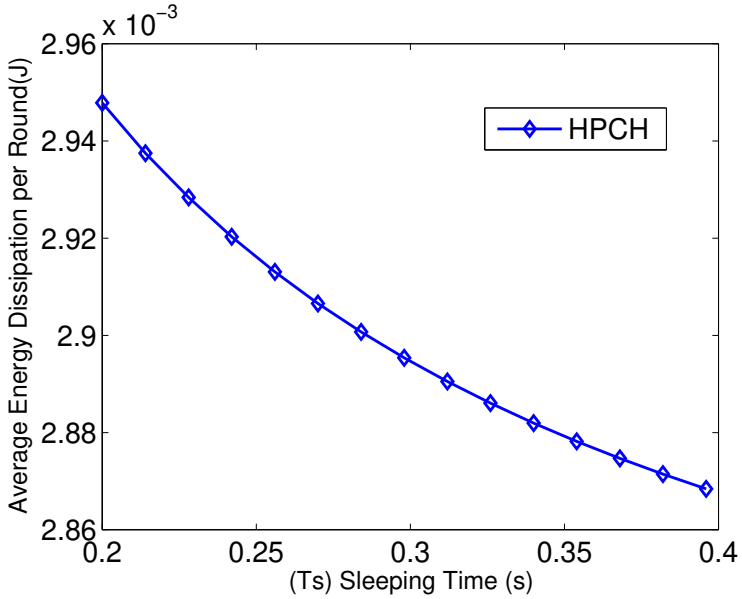


Figure 8.8: Average energy dissipation per round versus sleeping time

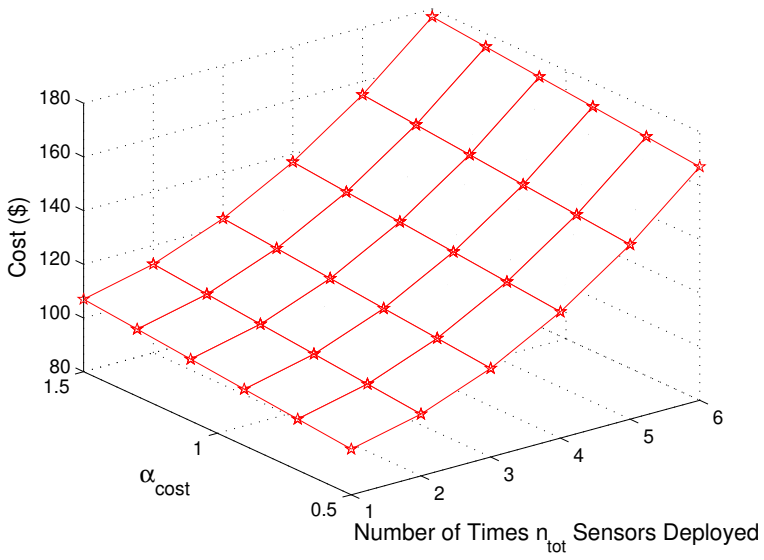


Figure 8.9: Total cost with cost ratio α_{cost} (between AA and AAA batteries), versus number of times N_s sensors are deployed

obtaining Fig. 8.6(b) in Experiment 2. Up to 4 high powered cluster heads are included in a single cluster to increase the network lifetime, as shown in Fig. 8.6(b). This shows that network lifetime can be increased by considering more than one HPCH per cluster. Average energy consumption ratio of the normal sensor node to the cluster head (E_1/E_2) is 0.0429. As this is very low, we can conclude that it is appropriate to use high energy packs for cluster heads.

As shown in Fig. 8.7, by having one, two, three and four HPCHs, we achieve 204%, 397%, 591% and 785% times of the network lifetime, respectively by using (8.19). Moreover, the number of high-powered CHs is proportional to R_D.

In Fig. 8.8, the relationship of the average energy dissipated per round to sensor sleeping time is plotted. As expected, the former reduces nonlinearly with the latter. We did these experiments using several energy models, ours [70], Wang and Chandrakasan [185], Zhu and Papavassiliou [205] and Mille and Vaidya [120] and found the results to be consistent.

We consider the following values. The total cost for a network of $N = 100$ sensors, is computed based on equation (8.20). The cost of a AA battery is $C_{AA} =$

1 $ and the deployment cost per sensor is $D_i = 0.02$ $. The computed total cost is illustrated in Fig. 8.9, versus the cost ratio α_{cost} (as in (8.21)), and number of times N_s sensors are deployed. The cost increases nonlinearly with the number of times N_s sensors are deployed. Moreover, the total cost increases marginally with α_{cost}.

8.8 Chapter Summary

This chapter has investigated the problem of energy wastage due to the selection of unsuitable batteries. We have shown how the energy ratio between a CH and a sensor can affect the sensor network lifetime. Based on this work we can choose the required number of sensors and the suitable batteries for sensors and CHs, if we use high powered CHs. These results save on the total cost of the network, an important factor in wireless sensor networks.

Unnecessary overheads, due to cluster head rotation in every communication round, are a significant problem in LEACH. Our simulation results show that the average energy consumption ratio of the normal sensor node to the cluster head (E_1 / E_2) is 0.0429. As this is very low, this chapter has concluded that it is acceptable to use high energy packs for cluster heads. This chapter proposed a method, called the High Powered Cluster Head (HPCH) method, that can alleviate this problem while extending network lifetime. Comparing this approach with an equivalent LEACH system, where the initial total battery capacity is equal to that of our system, the results indicate an increase in lifetime of 104.23% It is shown that the decision on the batteries should be linked to their actual cost as well as the performance. If the price difference between AAA size and AA size alkaline batteries is small, for example, a high network lifetime increase per dollar can be achieved by replacing the former with the latter.

Important conclusions drawn from this chapter are:

- From the simulation results, the average energy consumption ratio of the normal sensor node to the cluster head (E_1 / E_2) is 0.0429. As this is very low, this chapter has concluded that it is appropriate to get high energy

packs for cluster heads.

- Network lifetime can be dramatically increased by using multiple high powered batteries or by adding a high powered battery to each CH.
- According to the different types of application, the optimal ratio of the batteries for sensors and CHs may change.
- The distance from CH to sink node influences total network lifetime.
- Parameters such as battery cost and deployment cost make a huge impact on the total network cost.
- Inclusion of the HPCH and the use of the proposed method is significantly more effective than using LEACH only [67].

CHAPTER NINE

Conclusions for Part II

Part II has investigated the power management in sensor networks. A comprehensive energy model for wireless sensor networks was proposed by considering seven key energy consumption sources. Many of these have been ignored in current energy consumption models. Using our proposed model the lifetime of a sensor node was estimated and compared with other existing energy models, demonstrating the benefit of using such a comprehensive model.

We have also applied our model to a LEACH type protocol to obtain an accurate evaluation of energy consumption and node lifetime. We have shown that existing energy models overestimate life expectancy of a sensor node by 30-58% and also result in an "optimized" number of clusters that is too large.

This work also investigated the impact on the optimal number of clusters made by factors such as free space fading, different energy models, physical area, duty cycles and number of sensors. We thus, have obtained guidelines for efficient and reliable sensor network design. Based on this work, designers can optimize energy efficiency subject to their required specifications.

We have observed that: 1) the optimal number of clusters increases with the increase of free space fading energy, 2) for sensor networks with 100 sensors over an area of $10^4 - 10^5$ m^2, finding the optimal number of clusters becomes less important when free space fading energy is very low (less than 1670 $pJ/bit/m^2$), while for larger networks, on the other hand, cluster optimization is still important even if free space fading is low.

The design of sensor network determines its lifetime, which is important for

most applications. Efficient energy management for sensor networks depends on the selection of batteries for individual sensors. Our simulation results have shown that the average energy consumption ratio of the normal sensor node to the cluster head is very low. Therefore, an obvious choice is to use high energy packs for cluster heads and regular energy packs for other sensor nodes.

The unnecessary overheads, due to cluster head rotation in every communication round, is a significant problem in LEACH. This study has developed a method based on high powered cluster heads, that can alleviate this problem while extending network lifetime. We have shown that depending on the application, an optimal ratio of initial battery capacities for sensors and cluster heads can be found.

Part II also proposed a hybrid method, using high powered cluster heads until their energy packs are drained to the level of other sensor nodes, and then using LEACH to increase overall efficiency. This hybrid method also accounted for battery cost and deployment cost, which make a huge impact on cost of the total network.

CHAPTER TEN

Future Research

The handoff problem of base station scheduling in cellular networks resembles the sensor scheduling problem often considered with Hidden Markov Model sensors, where a single Markov chain is observed by a set of noisy sensors. Each time one or more sensors that can provide the next measurement are selected. Each measurement is associated with a cost, which is calculated using a cost function. The best sensor selection should minimize both this cost and the cost of any state estimation error. Some recent work [47, 53, 96, 97] inspired this partially completed study to take a similar approach towards the handoff problem reported in Appendix B. In the handoff problem we see the similarity between base stations and noisy sensors, and the need for selecting of a serving base station. In the Markov process we need to specify the resulting next state for each starting state and action. The Markov assumption that the next state is solely determined by the current state (and current action) is true for handoff decision making, if we ignore the ping-pong effect. We show that the most probable base station sequence can be estimated if the user movement can be modeled using a Hidden Markov Model (HMM). This chapter has investigated the application of sensor scheduling to the scheduling of observation modes in the handoff problem, leading to the estimation of the assigned base station sequence. It has evaluated the methods of finding the optimal schedule to minimize a set of cost functions.

In contrast to previously presented methods of base station sequence estimation [68, 192], the proposed method considers observation mode options and creates an optimal schedule of observation modes, which can then be used to es-

timate the hidden states of HMM identical to the base station sequence. This chapter has also discussed two methods for solving the optimization problem. Both dynamic programming based recursion with some limitations in defining the cost functions, and the GA based optimization with some limitations with computational cost, are widely applied to various problems in sensor scheduling.

Several important conclusions and future directions arise from this work:

- This work exploits the analogy of the sensor scheduling problem to the base station assignment or the handoff problem in cellular networks.
- The handoff problem is formulated as an optimization problem of base station scheduling that minimizes cost functions containing the HMM state estimation error and base station measurement costs.
- This optimization problem can be solved using dynamic programming techniques.

Considering the recent developments in Global Positioning System (GPS) technology and the regulations of the FCC in the United States that wireless providers must pinpoint the wireless emergency call's location within 125m, it can be assumed that GPS will become a standard in future mobile terminals. Road topology based mobility prediction of users leads to handoff that can be accurately identified using pattern recognition methods. In such methods base stations are required to maintain a database of the roads area's. The database entries can be extracted from previously created digital maps suitable for GPS based navigational devices. The roads can be stored as road segments identified with "junction pairs". The transition between road segments can be modeled as a second-order Markov process. Based on previous experience, certain road segments can be identified as "handoff probable". Such mobility prediction schemes are successful in macrocells or in cases where the users travel on known paths.

This work is also motivated by opportunities to use and possibly extend existing work in sensor networks and optimization. An analogy between the sensor

scheduling and handoff (as described earlier) has never been considered before. The possible application of research results to be developed in this work to the new area of ad hoc sensor networks is also an exciting and new approach.

Bibliography

[1] BALI-2: Stanford university mobile activity traces, BALI - Bay Area Location Information (real-time). Stanford University. [online] http://www-db.stanford.edu/pleiades/BALI2.html.

[2] Bluetooth. [online] http://www.bluetooth.com.

[3] POMDP tutorial. [online] http://www.cs.brown.edu/research/ai/pomdp /tutorial /pomdp-solving.html.

[4] Wikipedia. [online] http://www.wikipedia.org.

[5] MICA2 Mote Datasheet, 2004. [online] http:// www.xbow.com /Products /Product-pdf-files/Wireless-pdf / MICA2-Datasheet.pdf.

[6] Science and security. [online] www.dest.gov.au/meetings/documents/, 2006.

[7] T. Adachi and M. Nakagawa. Cellular ad-hoc united communication system. In *Workshop on mobile communication IWMC*, pages 278–282, 1996.

[8] T. Adachi and M. Nakagawa. Battery consumption and handoff examination of a cellular ad-hoc united communication system for operational mobile robots. In *Proc. IEEE Personal, Indoor and Mobile Radio Communications*, volume 3, pages 1193–1197, September 1998.

[9] R. Agrawal and R. Srikant. Mining sequential patterns. In *Proc. 11th Int. Conf. on Data Engineering*, March 1995.

[10] D. A. Levine I. F. Akyildiz and M. Naghshineh. A resource estimation and call admission algorithm for wireless multimedia networks using the shadow cluster concept. *IEEE/ACM Trans. Networking*, 5(1):1–12, February 1997.

[11] C. Alippi and C. Galperti. An adaptive maximum power point tracker for maximising solar cell efficiency in wireless sensor nodes. In *Proc. IEEE int. Symp. Circuits and Systems ISCAS'06*, pages 21–24, May 2006.

[12] R. Amirtharajah, J. Collier, J. Siebert, B. Zhou, and A. Chandrakasan. DSPs for energy harvesting sensors: Applications and architectures. *IEEE Pervasive Computing*, 4(3):72–79, Sep. 2005.

[13] R. Amirtharajah, J. Wenck, J. Collier, J. Siebert, and B. Zhou. Circuits for energy harvesting sensor signal processing. In *Proc. ACM/IEEE Design Automation Conf.*, pages 639–644, July 2006.

[14] Y. Argyropoulos, S. Jordan, and S. P. R. Kumar. Dynamic channel allocation in interference-limited cellular systems with uneven traffic distribution. *IEEE Trans. Veh. Technol.*, 48(1):224–232, January 1999.

[15] M. D. Austin and G. L. Stuber. Velocity adaptive handoff algorithms for microcellular systems. In *Proc. IEEE Veh. Technol. Conf.*, volume 43 of *3*, pages 549–561, August 1994.

[16] J. Ayres, J. Gehrke, T. Yu, and J. Flannick. Sequential pattern mining using a bitmap reprensentation. In *SIGKDD*, Canada, 2002.

[17] H. Balakrishnan, R. Baliga, D. Curtis, M. Goraczko, A. Miu, B. Priyantha, A. Smith, K. Steele, S. Teller, and K. Wang. Lessons from developing and devolping the cricket indoor location system. Technical report, MIT, USA, Nov. 2003.

[18] D. Bansal, A. Chandra, R. Shorey, A. Kulshreshtha, and M. Gupta. Mobility models for cellular systems: cell topography and handoff probability. In *Proc. Veh. Technol. Conf.*, volume 3, pages 1794–1798, May 1999.

[19] F. Barcelo. Performance analysis of handoff resource allocation strategies through the state-dependent rejection scheme. *IEEE Trans. Wireless Commun.*, 3(3):900–909, May 2004.

[20] R. Beck and F. W. Ho. Evaluation and performance of field strength related handover strategies for micro-cellular systems. In *3rd Nordic Sem. Digital Land Mobile Radio Commun.*, Copenhagen, Denmark, 1988.

[21] M. Bhardwaj and A. P. Chandrakasan. Bounding the lifetime of sensor networks via optimal role assignments. In *Proc. IEEE INFOCOM Conf.*, volume 3, pages 1587–1596, June 2002.

[22] G. Boggia and P. Camarda. Modeling dynamic channel allocation in multicellular communication networks. *IEEE J. Select. Areas Commun.*, 19(11):2233–2242, November 2001.

[23] A. Boulis, S. Ganeriwal, and M. B. Srivastava. Aggregation in sensor networks: an energy-accuracy trade-off. In *Proc. Int. Sensor Network Protocols and Applications*, pages 128–138, 2003.

[24] C. Brewster, P. Farmer, J. Manners, and M. N. Halgamuge. An incremental approach to model based clustering and segmentation. In *Proc. 1st Int. Fuzzy Systems and Knowledge Discovery FSKD'02*, volume 2, pages 586–590, 2002.

[25] T. Brown, H. Gabow, and Q. Zhang. Maximum flow-life curve for a wireless ad hoc network. In *Proc. of MobiHoc Conf.*, pages 128–136, California, USA., Oct. 2001.

[26] R. Bultitude and G. K. Bedal. Propagation characteristics on microcellular urban mobile radio channels at 910 MHz. *IEEE J. Select. Areas Commun.*, 7(1):31–39, Jan. 1989.

[27] M. Busse, T. Haenselmann, and W. Effelsberg. TECA: a topology and energy control algorithm for wireless sensor networks. In *Proc. ACM int.*

Symp. on Modeling analysis and simulation of wireless and mobile systems, MSWiM '06, pages 317–321, New York, NY, USA, 2006.

[28] K. J. Bye. Handover criteria and control in cellular and microcellular systems. In *Int. Conf. Mobile Radio Personal Commun.*, pages 94–98, December 1989.

[29] E. H. Callaway. *Wireless Sensor Networks, Architectures and Protocols*. Auerbach Publications, 2004.

[30] E. Cayirci. Data aggregation and dilution by modulus addressing in wireless sensor networks. *IEEE Commun. Lett.*, 7(8):355–357, Aug. 2003.

[31] K. Chan, G. Lam, S. M. Guru, M. N. Halgamuge, and S. Fernando. Development of a smartbolt prototype with energy model for clustered sensor systems. In *Proc. 1st Int. Fuzzy Systems and Knowledge Discovery FSKD'02*, volume 1, pages 280–284, 2002.

[32] A. Chandrakasan. Design considerations for distributed microsensor systems. In *Proc. IEEE Custom Integrated Circuits Conf.*, pages 279–286, San Diego, May 1999.

[33] C. Chang, C. J. Chang, and K. R. Lo. Analysis of a hierarchical cellular system with reneging and dropping for waiting new and handoff calls. *IEEE Trans. Veh. Technol.*, 48(4):1080–1091, July 1999.

[34] J.-H. Chang and L. Tassiulas. Routing for maximizing system lifetime in wireless ad-hoc networks. In *Proc. 37th Annual Allerton Conference on Communication, Control, and Computing*, Illinois, Sep. 1999.

[35] J. H. Chang and L. Tassiulas. Energy conserving routing in wireless ad-hoc networks. In *Proc. IEEE INFOCOM Conf.*, pages 22–31, Israel, Mar. 2000.

[36] B. Chen, K. Jamieson, H. Balakrishnan, and R. Morris. Span: An energy-efficient coordination algorithm for topology maintenance in ad hoc wireless networks. *Wireless Networks*, 8:481–494, 2002.

[37] J. S. Chen and C. R. Dow. Fixed channel allocation scheme performance enhancement for cellular mobile systems. *Communications, IEE Proceedings*, 149(4):232–236, August 2002.

[38] S. T. S. Chia, R. Steele, E. green, and A. Baran. Propagation and bit error ratio measurements for a microcellular system. *Institute of Electronic and Radio Engineers*, 57(6):255–266, Nov./Dec. 1987.

[39] S. T. S. Chia and R. J. Warburton. Handover criteria for city microcellular radio systems. In *Proc. IEEE 40th Veh. Technol. Conf.*, pages 276–281, 1990.

[40] C. F. Chiasserini and R. R. Rao. Energy efficient battery management. In *Proc. IEEE INFOCOM Conf.*, volume 2, pages 396–403, March 2000.

[41] C. F. Chiasserini and R. R. Rao. Stochastic battery discharge in portable communication devices. In *Fifteenth Annual Battery Conference on Applications and Advances*, pages 27–32, January 2000.

[42] C. F. Chiasserini and R. R. Rao. Stochastic battery discharge in portable communication devices. *IEEE Aerospace and Electronic Systems Mag.*, 15(8):41–45, August 2000.

[43] C. F. Chiasserini and R. R. Rao. Energy efficient battery management. *IEEE J. Select. Areas Commun.*, 19(7):1235–1245, July 2001.

[44] Chipcon. CC1000 datasheet, Aug. 2005. [online]http://www.chipcon.com/filies/CC1000-Data-Sheet-2-1.pdf.

[45] F. A. Cruz-Perez and L. Ortigoza-Guerrero. Regulated call dropping scheme for handoff prioritisation in wireless communication systems with heterogeneous platform types. In *Proc. IEEE PIMRC Conf.*, volume 2, pages 1451–1455, September 2003.

[46] D. Devasirvatham. Radio propagation studies in a small city for universal portable communications. In *Proc. IEEE Veh. Technol. Conf.*, pages 100–104, June 1988.

[47] S. Dey. Reduced-complexity filtering for partially observed nearly completely decomposable markov chains. *IEEE Trans. Signal Processing] [see also IEEE Trans. Acoustics, Speech, and Signal Processing]*, 48(12):3334–3344, December 2000.

[48] Q. Dong. Maximizing system lifetime in wireless sensor networks. In *Proc. IEEE Int. Symp. Info. Processing in Sensor Networks Conf.*, pages 13–19, Apr. 2005.

[49] P. Eggers, P. Mogensen, C. Jensen, and C. Andersen. Urban area radio propagation measurements at 955 and 1845 MHz for small and microcells. In *Proc. IEEE Global Telecommunications Conf.*, pages 1297–1301, Dec. 1991.

[50] A. Ephremides. Energy concerns in wireless networks. *IEEE Trans. Wireless Commun.*, 9(4):48–59, Aug. 2002.

[51] M. N. Halgamuge et al. Base station scheduling for cellular networks with hidden markov model. *IEEE Com. Lett.*, (to be submitted).

[52] ETSI. *GSM Technical Specification*. GSM 08.08, France, version 5.12.0 edition, June 2000.

[53] J. Evans and V. Krishnamurthy. Optimal sensor scheduling for hidden markov models. In *Proc. IEEE Acoustics, Speech and Signal Processing Conf.*, volume 4, pages 2161–2164, 1998.

[54] Y. Fang and I. Chlamtac. Handoff probability in PCS networks. In *Universal Personal Communications*, volume 1, pages 765–768, ICUPC '98, October 1998.

[55] E. A. Frech. Cellular models and hand-off criteria. In *Proc. IEEE 39th Veh. Technol. Conf.*, pages 128–135, 1989.

[56] A. J. Goldsmith and S. B. Wicker. Design challenges for energy-constrained ad hoc wireless networks. *IEEE Wireless Commun.*, 9(4):8–27, Aug. 2002.

[57] E. Green. Path loss and signal variability analysis for microcells. In *Proc. IEE int. Mobile Radio and Personal Communications Conf.*, pages 38–42, Dec. 1989.

[58] F. Guangbin, I. Stojmenovic, and Z. Jingyuan. A triple layer location management strategy for wireless cellular networks. In *11th Int. Conf. on Computer Communications and Networks*, pages 489–492, October 2002.

[59] M. Gudmundson. Analysis of handover algorithms. In *Proc. IEEE Veh. Technol. Conf.*, pages 537–542, May 1991.

[60] M. Gudmundson. Correlation model for shadow fading in mobile radio systems. *Electronics Lett.*, 27(23):2145–2146, November 1991.

[61] L. O. Guerrero and A. H. Aghvami. A prioritized handoff dynamic channel allocation strategy for PCS. *IEEE Trans. Veh. Technol.*, 48(4):1203–1215, July 1999.

[62] M. N. Halgamuge. Efficient battery management for sensor lifetime. In *Proc. IEEE 21st Int. Advanced Information Networking and Application Conf. AINA'07*, May 2007 (accepted).

[63] M. N. Halgamuge, S. M. Guru, and A. Jennings. Energy efficient cluster formation in wireless sensor networks. In *Proc. IEEE Int. Telecommunications Conf. ICT'03.*, volume 2, pages 1571–1576, Feb.-Mar. 2003.

[64] M. N. Halgamuge, S. M. Guru, and A. Jennings. *Centralised strategies for cluster formation in sensor networks*, pages 315–334. Classification and Clustering for Knowledge Discovery. Springer-Verlag, Aug. 2005. ISBN: 3-540-26073-0.

[65] M. N. Halgamuge, R. Kotagiri, H. L. Vu, and M. Zukerman. Evaluation of handoff algorithms using a call quality measure with signal based penalties. In *Proc. IEEE WCNC Conf.*, volume 2, LasVegas, USA, April 2006.

[66] M. N. Halgamuge, R. Kotagiri, and M. Zukerman. High powered cluster heads for extending sensor network lifetime. In *Proc. IEEE ISSPIT Conf.*, pages 64–69, Vancouver, Canada, Aug 2006.

[67] M. N. Halgamuge, H. L. Vu, R. Kotagiri, and M. Zukerman. An ideal base station sequence for pattern recognition based handoff in cellular networks. In *Proc. Int. Conference on Information and Automation, ICIA05*, pages 19–24, BMICH, Colombo, Dec. 2005.

[68] M. N. Halgamuge, H. L. Vu, R. Kotagiri, and M. Zukerman. Signal based evaluation of handoff algorithms. *IEEE Commun. Lett.*, 9(9):790–792, Sep. 2005.

[69] M. N. Halgamuge, H. L. Vu, R. Kotagiri, and M. Zukerman. A call quality performance measure for handoff algorithms in intelligent transportation systems. *IEEE Trans. Mobile Computing*, (submitted).

[70] M. N. Halgamuge, M. Zukerman, R. Kotagiri, and H. L. Vu. Sensor energy consumption. *ACM Trans. on Sensor Networks*, (submitted).

[71] W. B. Heinzelman. *Application-Specific Protocol Architectures for Wireless Networks*. PhD Thesis, Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science, June 2000.

[72] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan. Energy-efficient communication protocol for wireless microsensor networks. In *Proc. 33rd Annu. Int. Hawaii System Sciences Conf.*, volume 2, pages 1–10, Jan. 2000.

[73] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan. An application-specific protocol architecture for wireless microsensor networks. *IEEE Tran. on Wireless Comm.*, 1(4):660–670, Oct. 2002.

[74] W. R. Heinzelman, J. Kulik, and H. Balakrishnan. Adaptive protocols for information dissemination in wireless sensor networks. In *Proc. Mobicom' 99 Conf.*, pages 174–185, 1999.

[75] W. R. Heinzelman, A. Sinha, A. Wang, and A. P. Chandrakasan. Energy-scalable algorithms and protocols for wireless microsensor networks. In *Proc. IEEE Acoustics, Speech and Signal Processing Conf.*, volume 6, pages 3722–3725, June 2000.

[76] P. L. Hiew and M. Zukerman. Efficiency comparison of several channel allocation schemes for digital mobile communication networks. *IEEE Trans. Veh. Technol.*, 49(3):724–733, May 2000.

[77] D. Hong and S. S. Rapport. Traffic model and performance analysis for cellular mobile radio telephone systems with prioritized and non prioritized handoff procedures. *IEEE Trans. Veh. Technol.*, 35, August 1986.

[78] D. Hong and S. S. Rapport. Priority oriented channel access for cellular systems serving vehicular and portable radio telephones. *IEE Proc. Part I*, 136(5):339–346, 1989.

[79] M. Ilyas and I. Mahgoub. Handbook of sensor networks: Compact wireless and wired sensing system. CRC Press, 2005.

[80] C. Intanagonwiwat, R. Govindan, and D. Estrin. Directed diffusion: A scalable and robust communication paradigm for sensor networks. Tech. rep. 00-732, 2000, University of Southern California, Los Angeles.

[81] C. Intanagonwiwat, R. Govindan, D. Estrin, J. Heidemann, and F. Silva. Directed diffusion for wireless sensor networking. *IEEE/ACM Trans. Networking*, 11(1), Feb. 2003.

[82] K. Ioannou, S. Louvros, I. Panoutsopoulos, S. Kotsopoulos, and G. K. Karagiannidis. Optimizing the handover call blocking probability in cellular networks with high speed moving terminals. *IEEE Commun. Lett.*, 6(10):422–424, October 2002.

[83] K. I. Itoh, S. Watanabe J. S. Shih, and T. Sato. Performance of handoff algorithm based on distance and RSSI measurements. *IEEE Trans. Veh. Technol.*, 51(6):1460–1468, November 2002.

[84] A. Jayasuriya. Handover channel allocation based on mobility. In *IST Mobile and Wireless Communications*, Summit 2003, June 2003.

[85] P. Jian, H. Jiawei, B. Mortazavi-Asl, H. Pinto, C. Qiming, U. Dayal, and H. Mei-Chun. Prefixspan,: mining sequential patterns efficiently by prefix-projected pattern growth. In *Proc. 17th Int. Conf. on Data Engineering*, pages 215–224, April 2001.

[86] J. Jiang and T. H. Lai. Call admission control vs. bandwidth reservation: reducing handoff call dropping rate and providing bandwidth efficiency in mobile networks. In *International Conference Parallel Processing*, pages 581–588, August 2000.

[87] X. Jiang, J. Polastre, and D. Culler. Perpetual environmentally powered sensor networks. In *Proc. Int. Symp. Information Processing in Sensor Networks*, pages 463–468, Apr. 2005.

[88] J. Julik, W. Heinzelman, and H. Balakrishnan. Negotiation-based protocols for disseminating information in wireless sensor networks. *Wireless Networks*, 8:169–185, 2002.

[89] L. Junyi, N. B. Shroff, and E. K. P. Chong. Channel carrying: a novel handoff scheme for mobile cellular networks. *IEEE/ACM Trans. Networking*, 7(1):38–50, February 1999.

[90] A. Kansal, D. Potter, and M. B. Srivastava. Performance aware tasking for environmentally powered sensor networks. In *Proc. int. Measurement and modeling of computer systems Conf.*, pages 223–234, New York, NY, USA, 2004. ACM Press.

[91] A. Kansal and M. B. Srivastava. An environmental energy harvesting framework for sensor networks. In *Proc. Int. Symp. Low Power Electronics and Design Conf.*, pages 481–486. ACM Press, 2003.

[92] O. Kennemann. Pattern recognition by hidden markov models for supporting handover decisions in the GSM system. In *Proc. Nordic Seminal Digital Mobile Radio Commun.*, pages 195–202, June 1994.

[93] M. I. Kim and S. J. Kim. A 2-level call admission control scheme using priority queue for decreasing new call blocking and handoff call dropping. In *Proc. IEEE Veh. Technol. Conf.*, volume 1, pages 457–461, April 2003.

[94] A. Krause, C. Guestrin, A. Gupta, and J. Kleinberg. Near-optimal sensor placements: Maximizing information while minimizing communication cost. In *Proc. IEEE Int. Information Processing in Sensor Networks*, 2006.

[95] B. Krishnamachari. *Networking Wireless Sensors*. Cambridge University Press, 2005.

[96] V. Krishnamurthy. Algorithms for optimal scheduling and management of hidden markov model sensors. *IEEE Trans. Signal Processing*, 50(6):1301–1312, June 2002.

[97] V. Krishnamurthy and B. Wahlberg. Algorithms for scheduling of hidden markov model sensors. In *Proc, 40th IEEE Conf. on Decision and Control*, volume 5, pages 4818–4819, 2001.

[98] H. Kwon, T. H. Kim, S. Choi, and B. G. Lee. Lifetime maximization under reliability constraint via cross-layer strategy in wireless sensor networks. In *Proc. IEEE WCNC Conf.*, volume 3, pages 1891–1896, March 2005.

[99] J. J. Lee, B. Krishnamachari, and C. C. J. Kuo. Impact of heterogeneous deployment on lifetime sensing coverage in sensor networks. In *Proc. IEEE SECON Conf.*, pages 367–376, Oct. 2004.

[100] W. C. Y. Lee. Link capacities between GSM and CDMA. In *Proc. IEEE ICC Conf.*, volume 3, pages 1335–1339, June 1995.

[101] W. C. Y. Lee. *Mobile Cellular Telecommunication*. McGraw-Hill, 2 nd edition, 1995.

[102] A. E. Leu and B. L. Mark. Dicrete-Time Level-Crossing analysis of soft handoff performance in cellular networks. *IEEE Trans. Information Theory*, 52(7), July 2006.

[103] K. K. Leung, W. A. Massey, and W. Whitt. Traffic models for wireless communication networks. *IEEE J. Select. Areas Commun.*, 12:13531364, October 1994.

[104] S. Li, F. Cheng, Y. Yuan, and T. Hu. Adaptive frame switching for UMTS UL-EDCH - ping-pong avoidance. In *Proc. IEEE Veh. Technol. Conf.*, volume 5, pages 2469–2473, 2006.

[105] Y. Lim, H. Seo, and B. G. Lee. Lifetime maximization under data convergence latency constraint in wireless sensor networks. In *Proc. IEEE GLOBECOM Conf.*, volume 5, pages 3187–3192, Nov.-Dec. 2004.

[106] F. Y. S. Lin and P. L. Chiu. A near-optimal sensor placement algorithm to achieve complete coverage-discrimination in sensor networks. *IEEE Commun. Lett.*, 9(1):43 – 45, Jan. 2005.

[107] H. Lin, R. Juang, and D. Lin. Validation of an improved location-based handover algorithm using GSM measurement data. *IEEE Trans. Mobile Computing*, 4(5):530–536, Sept. 2005.

[108] Y. Lin. Modeling techniques for large-scale PCS networks. *IEEE Commun. Mag.*, 35(2):102–107, Feb. 1997.

[109] Yi-Bing Lin and I. Chlamtac. *Wireless and Mobile Network Architectures*. Wiley, USA, 2001.

[110] T. Liu, P. Bahl, and I. Chlamtac. Mobility modeling, location tracking, and trajectory prediction in wireless ATM networks. *IEEE J. Select. Areas Commun.*, 16(6):922–936, August 1998.

[111] F. Lotse and A. Wejke. Propagation measurements for microcells in central stockholm. In *Proc. IEEE Veh. Technol. Conf.*, pages 539–541, May 1990.

[112] C. Ma and Y. Yang. Battery-aware routing for streaming data transmissions in wireless sensor networks. *Mobile Networks and Applications*, 11(5):757–767, 2006.

[113] A. Mainwaring, J. Polastre, R. Szewczyk, D. Culler, and J. Anderson. Wireless sensor networks for habitat monitoring. In *Proc. ACM Int. Wireless Sensor Networks and Applications Conf.*, USA, Sep. 2002.

[114] P. Marichamy and S. Chakrabarti. On threshold setting and hysteresis issues of handoff algorithms. In *Proc. IEEE Personal Wireless Communication Conf.*, pages 436–440, Feb 1999.

[115] A. Markopoulos, P. Pissaris, S. Kyriazakos, and E. D. Sykas. Performance analysis of cellular networks by simulating location aided handover algorithms. 2003.

[116] A. Markopoulos, P. Pissaris, S. Kyriazakos, and E. D. Sykas. *Efficient Location-Based Hard Handoff Algorithms for Cellular Systems*, volume 3042 of *Series Lecture Notes in Computer Science*, pages 476–489. Springer Berlin / Heidelberg, 2004.

[117] M. A. Marsan, C. F. Chiasserini, and A. Fumagalli. Performance models of handover protocols and buffering policies in mobile wireless ATM networks. *IEEE Trans. Veh. Technol.*, 50(4):925–941, July 2001.

[118] M. A. Marsan, G. de Carolis, E. Leonardi, and R. Lo Cigno amd M. Meo. Efficient estimation of call blocking probabilities in cellular mobile telephony networks with customer retrials. *IEEE J. Select. Areas Commun.*, 19(2):332–346, Feb 2001.

[119] H. Maturino-Lozoya, D. Munoz-Rodriguez, F. Jaimes-Romero, and H. Tawfik. Handoff algorithms based on fuzzy classifiers. *IEEE Trans. Veh. Technol.*, 49(6):2286–2294, Nov 2000.

[120] M. J. Mille and N. H. Vaidya. A mac protocol to reduce sensor network energy consumption using a wakeup radio. *IEEE Trans. Mobile Computing*, 4(3):228–242, May 2005.

[121] R. Min, M. Bhardwaj, C. Seong-Hwan, E. Shih, A. Sinha, A. Wang, and A. Chandrakasan. Low-power wireless sensor networks. In *Proc. 14th Int. VLSI Design Conf.*, pages 205–210, Jan. 2001.

[122] Shirvani Moghaddam, Vakili Tabataba, and A. Falahati. New handoff initiation algorithm (optimum combination of hysteresis and threshold based methods). In *Proc. IEEE Veh. Technol. Conf.*, volume 4, pages 1567–1574, Sept 2000.

[123] J. Mohammad and J. Zaki. *An efficient algorithm for mining frequent sequences, Machine Learning*. Kluwer Achademic Publishers, Boston, 2000.

[124] A. Murase and I. C. Symington. Handover criterion for macro and microcellular systems. In *Proc. IEEE Veh. Technol. Conf.*, pages 276–281, 1990.

[125] S. Nanda. Teletraffic models for urban and suburban microcells: cell sizes and handoff rates. *IEEE Trans. Veh. Technol.*, 42(4):673–682, Nov. 1993.

[126] R. Narasimhan and D. Cox. A handoff algorithm for wireless systems using pattern recognition. In *Proc. IEEE int. Symp. Personal, Indoor, Mobile Radio Commun.*, Boston, Sept. 1998.

[127] S. Ohzahata, S. Kimura, and Y. Ebihara. Adaptive handoff algorithm for efficient network resource utilization. In *17th Int. Conf. Advanced Information Networking and Applications, AINA 2003.*, pages 743–748, March 2003.

[128] J. A. Paradiso and T. Starner. Energy scavenging for mobile and wireless electronics. *IEEE Pervasive Computing*, 4(1):18–27, Jan. 2005.

[129] S. Pattem, B. Krishnamachari, and R. Govindan. The impact of spatial correlation on routing with compression in wireless sensor networks. In *Proc.*

Int. Symp. Information Processing in Sensor Networks, pages 28–35, Berkeley, California, USA, 2004. ACM Press.

[130] M. C. H. Peh. *Multiple Access Communications in Wireless Networks*. PhD Thesis, University of Melbourne, Melbourne, Australia, August 2004.

[131] X. Peng, X. Zhang, L. Ma, and S. Zhang. An improved handoff algorithm for using mobile ipv6 in wireless environment with overlapping areas. In *Proc. IEEE Circuits and Systems Symp. on Emerging Technologies Conf.*, volume 1, pages 1–4, 2004.

[132] J. Polastre, J. Hill, and D. Culler. Versatile low power media access for wireless sensor networks. In *Proc. int. Embedded networked sensor systems Conf., SenSys '04*, pages 3–5, Baltimore, Maryland, USA, Nov. 2004.

[133] G. P. Pollini. Trends in handover design. *IEEE Communication Magazine*, 34(3):82–96, March 1996.

[134] R. Prakash and V. V. Veeravalli. Locally optimal soft handoff algorithms. *IEEE Trans. Veh. Technol.*, 2(3):347–356, March 2003.

[135] R. Prakash and V. V. Venugopal. Adaptive hard handoff algorithm. *IEEE J. Select. Areas Commun.*, 18(11):2456–2464, Nov. 2000.

[136] D. Puccinelli and M. Haenggi. Wireless sensor networks: applications and challenges of ubiquitous sensing. *IEEE Circuits and Systems Mag.*, 5(3):19–31, 2005.

[137] C. Purzynski and S. Rappaport. Multiple call handoff problem with queued handoffs and mixed platform types. *Proceedings IEE Communications*, 142(1):31–39, February 1995.

[138] L. R. Rabiner. A tutorial on hidden markov models and selected applications in speech recognition. *IEEE Proceeding on*, 77:257285, February 1989.

[139] V. Raghunathan, A. Kansal, J. Hsu, J. Friedman, and M. Srivastava. Design considerations for solar energy harvesting wireless embedded systems. In

Proc. Int. Symp. Information Processing in Sensor Networks, pages 457–462, Apr. 2005.

[140] V. Raghunathan, C. Schurgers, S. Park, and M. B. Srivastava. Energy-aware wireless microsensor networks. *IEEE Signal Processing Mag.*, 19(2):40–50, March 2002.

[141] M. Rahimi, H. Shah, G. S. Sukhatme, J. Heideman, and D. Estrin. Studying the feasibility of energy harvesting in a mobile sensor network. In *Proc. IEEE int. Robotics and Automation Conf. ICRA '04*, volume 1, pages 19–24, Sep. 2003.

[142] M. Rajaratnam and F. Takawira. Nonclassical traffic modeling and performance analysis of cellular mobile networks with and without channel reservation. *IEEE Trans. Veh. Technol.*, 49(3):817 – 834, May 2000.

[143] S. Ramakrishnan and R. Agrawal. Mining sequential patterns: Generalizations and performance improvments. In *International Conference in Extending Database Technology, EDBT*, 1996.

[144] T. S. Rappaport. *Wireless Communications: Principles and Practice*. Prentice Hall, New Jersey, 1996.

[145] L. K. Rasmussen. On ping-pong effects in linear interference cancellation for CDMA. In *Spread Spectrum Techniques and Applications*, volume 2, September 2000.

[146] L. K. Rasmussen and I. J. Oppermann. Ping-pong effects in linear parallel interference cancellation for CDMA. *IEEE Trans. Wireless Commun.*, 2(2):357–363, March 2003.

[147] E. Del Re, R. Fantacci, and G. Giambene. Different queuing policies for handover requests in low earth orbit mobile satellite systems. *IEEE Trans. Veh. Technol.*, 48(2):448–458, March 1999.

[148] H. L. Rong and S. S. Rappaport. Personal communication systems using multiple hierarchical cellular overlays. *IEEE J. Select. Areas Commun.*, 13(2):406–415, Feb 1995.

[149] S. Roundy, B. P. Otis, Y.H. Chee, J. M. Rabaey, and P. Wright. A 1.9 GHz RF transmit beacon using environmently scavenged energy. In *Proc. IEEE int. Symp. Low Power Elec. and Devices*, Feb. 2003.

[150] S. J. Roundy. *Energy Scavenging for Wireless Sensor Nodes with a Focus on Vibration to Electricity Conversion*. PhD Thesis, The University Of California, Berkeley, 2003.

[151] A. J. Rustako, N. Amitay, G. J. Owens, and R. S. Roman. Radio propagation at microwave frequencies for line-of-sight microcellular mobile and personal communications. *IEEE Trans. Veh. Technol.*, 40(1):203–210, Feb. 1991.

[152] A. Safwat, H. Hassanein, and H. T. mouftah. Power-aware virtual base stations (PW-VBS) for wireless mobile ad hoc communications. *J. Computer Networks*, 41(3):331–346, 2003.

[153] A. Sampath and J. M. Holtzman. Estimation of maximum doppler frequency for handoff decisions. In *Proc. IEEE Veh. Technol. Conf.*, pages 859–862, May 1993.

[154] Y. Sankarasubramaniam, I. F. Akyildiz, and S. W. McLaughlin. Energy efficiency based packet size optimization in wireless sensor networks. In *Proc. IEEE Int. Sensor Network Protocols and Applications Conf.*, pages 1–8, 2003.

[155] A. Santuari. Steiner tree NP-completeness proof. Technical report, University of Trento, May 2003.

[156] C. Schurgers, V. Tsiatsis, S. Ganeriwa, and M. Srivastava. Optimizing sensor networks in the energy-latency-density design space. *IEEE Trans. Mobile Computing*, 1(1):70–80, Mar. 2002.

[157] D. J. Sebald and J. A. Bucklew. Support vector machine techniques for nonlinear equalization. *IEEE Trans. Signal Processing*, 48(11):3217–3226, Nov. 2000.

[158] G. N. Senarath and D. Everitt. Combined analysis of transmission and traffic characteristics in micro-cellular mobile communication systems. In *Proc. IEEE Veh. Technol. Conf.*, pages 577–580, May 1993.

[159] G. N. Senarath and D. Everitt. Performance of handover priority and queueing systems under different handover request strategies. In *Proc. IEEE Veh. Technol. Conf.*, volume 2, pages 897–901, July 1995.

[160] G. N. Senarath and D. Everitt. Reduction of call drop-outs during handoffs using efficient signal strength prediction algorithms for personal communication systems. In *Proc. IEEE GLOBECOM Conf.*, volume 3, pages 2308–2312, November 1995.

[161] E. Shih, S. Cho, N. Ickes, R. Min, A. Sinha, A. Wang, and A. Chandrakasan. Physical layer driven protocol and algorithm design for energy-efficient wireless sensor networks. In *Proc. MOBICOM Conf.*, Rome, Italy, July 2001.

[162] A. Shilton, M. Palaniswami, D. Ralph, and T. Chung. Incremental training of support vector machines. *IEEE Trans. Neural Networks*, 16(1):114–131, Jan. 2005.

[163] V. Shnayder, M. Hempstead, B.Chen, G. W.r Allen, and M. Welsh. Simulating the power consumption of large-scale sensor network applications. In *Proc. 2nd int. Embedded networked sensor systems Conf. SenSys '04*, pages 188–200, NY, USA, 2004.

[164] M. L. Sichitiu. Cross-layer scheduling for power efficiency in wireless sensor networks. In *Proc. IEEE INFOCOM Conf.*, volume 3, pages 1740–1750, 2004.

[165] S. Singh, M. Woo, and C. S. Raghavendra. Power-aware routing in mobile ad hoc networks. In *Proc. IEEE MOBICOM Conf.*, pages 181–190, Oct. 1998.

[166] S. Skafidas, K. Saleem, M. Halpern, W. Qiu, and H. Gan. Wireless sensor networks. Lecture notes, University of Melbourne, Australia, 2005.

[167] G. Smaragdakis, I. Matta, and A. Bestavros. SEP: A stable election protocol for clustered heterogeneous wireless sensor networks. In *Proc. 2nd Int. Workshop on Sensor and Actuator Network Protocols and Applications, SANPA '04*, Boston, Aug. 2004.

[168] S. Soro and W. B. Heinzelman. Prolonging the lifetime of wireless sensor networks via unequal clustering. In *Proc. IEEE int. Parallel and Distributed Processing Symposium Conf.*, pages 236b–236b, April 2005.

[169] G. L. Stuber. *Principl of Mobile Communication*. Kluwer Academic Publishers, Norwell, Massachusetts 02061, USA, 2 nd edition, 1958.

[170] R. Szewczyk, A. Mainwaring, J. Polastre, J. Anderson, and D. Culler. An analysis of a large scale habitat monitoring application. In *Proc. int. Embedded networked sensor systems Conf. SenSys '04*, pages 214–226, Baltimore, MD, USA, 2004. ACM Press.

[171] R. Szewczyk, E. Osterweil, J. Polastre, M Hamilton, A Mainwaring, and D. Estrin. Habitat monitoring with sensor networks. 47(4):34–40, June 2004.

[172] S. Tabbane. An alternative strategy for location tracking. *IEEE J. Select. Areas Commun.*, 13(5):880–892, June 1995.

[173] S. Tekinay and B. Jabbari. Handover and channel assignment in mobile cellular networks. *IEEE Communications Magazine*, 29(11):42–46, November 1991.

[174] S. Tekinay and B. Jabbari. A measurement-based prioritization scheme for handovers in mobile cellular networks. *IEEE J. Select. Areas Commun.*, 10(8):1343–1350, October 1992.

[175] C.-K. Toh. Maximum battery life routing to support ubiquitous mobile computing in wireless ad hoc networks. *IEEE Commun. Mag.*, pages 138–147, June 2001.

[176] P. Tran-Gia and M. Mandjes. Modeling of customer retrial phenomenon in cellular mobile networks. *IEEE J. Select. Areas Commun.*, 15(8):1406–1414, October 1997.

[177] N. D. Tripathi, J. H. Reed, and H. F. Van Landingham. Handoff in cellular systems. *IEEE Personal Communications*, 5(6):26–37, December 1998.

[178] V. V. Veeravalli and O. E. Kelly. A locally optimal handoff algorithms for cellular communication. *IEEE Trans. Veh. Technol.*, 46(3):603–609, August 1997.

[179] R. Vijayan and J. M. Holtzman. A model for analysing handoff algorithms. *IEEE Trans. Veh. Technol.*, 42(3):351–356, August 1993.

[180] M . Viswanathan and R. Kotagiri. Comparing the performance of support vector machines to regression with structural risk minimisation. In *Proc. Int. Intelligent Sensing and Information Processing Conf.*, pages 445–449, Dec. 2004.

[181] N. Vlajic, C. D. Charalambous, and D. Makrakis. Performance aspects of data broadcast in wireless networks with user retrials. *IEEE/ACM Trans. Networking*, 12(4):620–633, August 2004.

[182] J. F. Wagen. Signal strength measurements at 881 MHz for urban microcells in downtown tampa. In *Proc. IEEE Global Telecommunications Conf.*, pages 1313–1317, May 1991.

[183] B. H. Walke. *Mobile Radio Networks Networking, Protocols and Traffic Performance*. Willey, New York USA, 2002.

[184] A. Wang and A. Chandrakasan. Energy efficient system partitioning for distributed wireless sensor networks. In *Proc. IEEE Acoustics, Speech and Signal Processing Conf.*, volume 2, pages 905–908, May 2001.

[185] A. Wang and A. Chandrakasan. Energy-efficient DSPs for wireless sensor networks. *IEEE Signal Processing Mag.*, 19(4):68–78, July 2002.

[186] H. Wang, K. Yao, and D. Estrin. Information-theoretic approaches for sensor selection and placement in sensor networks for target localization and tracking. Technical Report 52, CENS, UCAL, Aug. 2005.

[187] H. Wang, K. Yao, and D. Estrin. Information-theoretic approaches for sensor selection and placement in sensor networks for target localization and tracking. *Journal of Commun. and Networks*, 7(4):438–449, Dec. 2005.

[188] R. Wattenhofer, L. Li, P. Bahl, and Y. Wang. Distributed topology control for power efficient operation in multihop wireless ad hoc networks. In *Proc. IEEE INFOCOM Conf.*, Anchorage, AK, Apr. 2001.

[189] J. Whitteker. Measurements of path loss at 910 MHz for proposed microcell urban mobile systems. *IEEE Trans. Veh. Technol.*, 37(3):125–129, Aug. 1988.

[190] J. E. Wieselthier, G. D. Nguyen, and A. Ephremides. Resource management in energy-limited, bandwidth-limited, transceiver-limited wireless networks for session-based multicasting. *Computer Networks*, 39(2):113–131, June 2002.

[191] K. D. Wong. *Handoff Algorithms using Pattern Recognition*. PhD Thesis, Stanford University, Department of Electrical and Electronic Engineering, June 1998.

[192] K. D. Wong and D. C. Cox. A pattern recognition system for handoff algorithms. *IEEE J. Select. Areas Commun.*, 18(7):1301–1312, July 2000.

[193] K. D. Wong and T. J. Lim. Soft handoffs in CDMA mobile systems. In *Proc. IEEE Personal Communications Conf.*, pages 6–17, December 1997.

[194] H. H. Xia, H. Bertoni, L. R. Maciel, A. L. Stewart, R. Rowe, and L. Grindstaff. Radio propagation measurements and modelling for line-of-sight microcellular systems. pages 349–354, May 1992.

[195] H. H. Xia, S. Kim, and H. Bertoni. Microcellular propagation measurements in Dallas city. In *Proc. IEEE Veh. Technol. Conf.*, pages 593–597, May 1993.

[196] M. Xiao, N. B. Shroff, and K. P. Chong. A utility-based power-control scheme in wireless cellular systems. *IEEE/ACM Trans. Networking*, 11(2):210–221, April 2003.

[197] Y. Xu, J. Heidemann, and D. Estrin. Geography-informed energy conservation for ad hoc routing. In *Proc. Mobicom Conf.*, 2001.

[198] Z. Yang, Y. Wang, D. Zhao, J. He, and X. Fu. Fast seamless handover scheme and cost performance optimization for ping-pong type of movement. In *Proc. IEEE PIMRC Conf.*, volume 4, pages 2131–2135, Sept. 2005.

[199] W. Ye and J. Heidemann. Medium access control in wireless sensor networks. Technical Report ISI-TR-580, USC/ISI TECHNICAL REPORT, Oct. 2003.

[200] W. Ye, J. Heidemann, and D. Estrin. An energy-efficient MAC protocol for wireless sensor networks. In *Proc. IEEE INFOCOM Conf.*, volume 3, pages 1567–1576, June 2002.

[201] W. W. H. Yu and H. Changhua. Resource reservation in wireless networks based on pattern recognition. In *International Joint Conference on Neural Networks*, volume 3, pages 2264–2269, IJCNN '01, July 2001.

[202] Y. Yu, B. Krishnamachari, and V. K. Prasanna. Energy-latency tradeoffs for data gathering in wireless sensor networks. In *Proc. IEEE INFOCOM Conf.*, volume 1, March 2004.

[203] N. Zhang and J. M. Holtzman. Analysis of handoff algorithms using both absolute and relative measurements. *IEEE Trans. Veh. Technol.*, 45(1):174–179, Feb 1996.

[204] P. Zhang, C. M. Sadler, S. A. Lyon, and M. Martonosi. Hardware design experiences in zebranet. In *Proc. SenSys*, Feb. 2003.

[205] J. Zhu and S. Papavassiliou. On the energy-efficient organization and the lifetime of multi-hop sensor networks. *IEEE Commun. Lett.*, 7(11):537–539, Nov. 2003.

[206] M. Zonoozi and P. Dassanayake. *Handover delay and hysteresis margin in microcells and macrocells*. PhD Thesis, Sep 1997.

[207] Y. Zou and K. Chakrabarty. Target localization based on energy considerations in distributed sensor networks. In *Proc. IEEE Int. Sensor Network Protocols and Applications Conf.*, pages 51–58, May 2003.

APPENDIX A

Determining the Value of Handoff Constant a

A.1 Method 1

The probability of having exactly one drop level sample point at arbitrary position X of the sequence, $p_X(d_1)$, where $1 \leq X \leq N$ is:

$$p_X(d_1) = (1-\delta)^{X-1}\delta(1-\delta)^{N-X} = \delta(1-\delta)^{N-1}, \tag{A.1}$$

where δ is the probability of receiving a signal strength below S_{drop}.

From (A.1), the probability of having one drop at any point in the sequence is

$$= \sum_{X=1}^{N} p_X(d_1) = N\delta(1-\delta)^{N-1}, \tag{A.2}$$

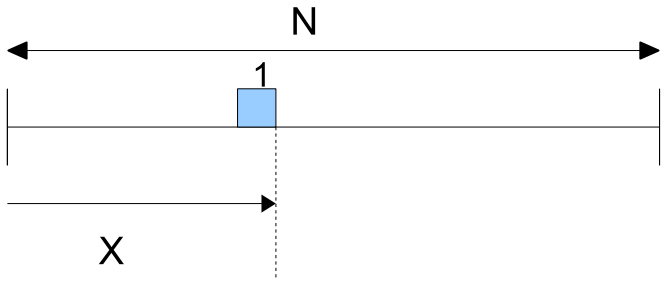


Figure A.1: Probability of having exactly one drop level sample point at arbitarary position X of the sequence

The probability of having two consecutive drop levels at X is:

$$p_X(d_2) = (1-\delta)^{X-1}\delta^2(1-\delta)^{N-X-1} = \delta^2(1-\delta)^{N-2}. \tag{A.3}$$

From (A.3), the probability of having two consecutive drops anywhere in the sequence is,

$$= \sum_{X=1}^{N-1} p_X(d_2) = (N-1)\delta^2(1-\delta)^{N-2}, \tag{A.4}$$

so, the probability of having d consecutive drops anywhere is,

$$= (N-d+1)\delta^d(1-\delta)^{N-d}.$$

We also note that

$$p_X(d_d) < p_X(d_{d-1}).$$

We make then the assumption that $a \propto \frac{p_X(d-1)}{p_X(d)}$.
The constant, *const*, can be set by considering that a tends to reach 1, when δ tends to 1, and therefore $const = 1$.

$$\begin{aligned} a &= \frac{p_X(d_{d-1})}{p_X(d_d)} + const. = \left(\frac{N-d+2}{N-d+1}\right)\frac{\delta^{d-1}}{\delta^d}\frac{(1-\delta)^{N-d+1}}{(1-\delta)^{N-d}} + const, \\ &= \left(\frac{N-d+2}{N-d+1}\right)\frac{(1-\delta)}{\delta} + const. \end{aligned}$$

Let us assume $N >>> d$. Therefore we choose

$$a = \frac{(1-\delta)}{\delta} + 1.$$

This value shows in Fig. 4.6. From (4.15), sequence of d dropping points is given as

$$C_1, aC_1, ..., a^{d-1}C_1.$$

The rare occurrence of bad sample points is to avoid dropping.

A.2 Method 2

The probability of having the first drop at an arbitrary position X of the sequence, $p_X(d)$, where $1 \leq X \leq N$ is:

$$p_X(d) = (1-\delta)^X \delta^d,$$

where δ represents the probability of receiving a signal strength below S_{drop}.

The probability of the first drop occurring at any point is given by

$$= \sum_{X=1}^{N-d} p_X(d) = \sum_{X=1}^{N-d} (1-\delta)^X \delta^d,$$

therefore, the probabilities of $d-1$ and d are given by

$$p(d-1) = (1-(1-\delta)^{N-d+1})\delta^{d-2}, \tag{A.5}$$

$$p(d) = (1-(1-\delta)^{N-d})\delta^{d-1}. \tag{A.6}$$

We make the asumption that $a \propto \frac{p(d-1)}{p(d)}$.
From (A.5) and (A.6),

$$\frac{p(d-1)}{p(d)} = \frac{(1-(1-\delta)^{N-d+1})\delta^{d-2}}{(1-(1-\delta)^{N-d})\delta^{d-1}}$$

The constant, *const*, can be set by considering that a tends to reach 1, when δ tends to 1, and therefore $const = 1$.

$$a = \frac{p(d-1)}{p(d)} + const. = \frac{(1-(1-\delta)^{N-d+1})}{(1-(1-\delta)^{N-d})\delta} + const. \tag{A.7}$$

Therefore, from (A.7)

$$a = \frac{(1-\delta)}{\delta} + 1.$$

This value shows in Fig. 4.6. From (4.15), the sequence of d dropping points is given as

$$C_1, aC_1, ..., a^{d-1}C_1.$$

In practice, different systems may have different drop rates, caused either by coverage problems because of irregular terrain configurations or inadequate channel availability [101]. Therefore, here we consider the weighting factor a as fixed for a given terrain configuration. Using Fig. 4.6, we choose the weighting factor a once we know the dropping probability.

APPENDIX B

Base Station Scheduling with HMM

B.1 Motivation

The first question I address in this Appendix is: How can we fully adapt known theoretical results for optimal sensor scheduling and management to the handoff problem? The second question I address is: How can we formulate the handoff problem as an optimization problem of base station scheduling with HMM states? If we use the signal strength data received from all base stations and the benchmark solution, we may be able to interpret the "estimation cost" as a cost associated with providing good QoS. The "measurement cost" consists of the handoff cost (if applicable), and the cost inversely proportional to the signal strength of the assigned base station plus the base station usage cost. Note that the measurement cost varies in cellular networks (for example when there is a need to use another cellular provider, such as a satellite).

B.2 Assumptions

- The system is a Markov process.
- There are two possibilities for selecting the observations.
- Using received signal strengths or base stations is more costly than using the prediction method.

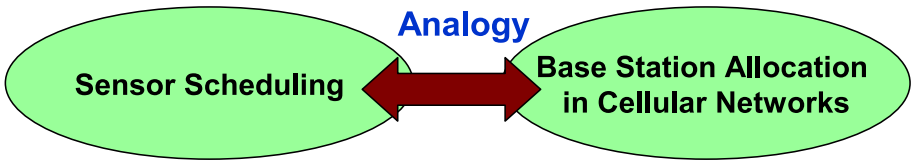


Figure B.1: Analogy established in this work

- We assume a homogeneous network where all cells are identical in size, user mobility and cell coverage as in [77, 148].
- Each cell has an equal number of neighbors as in [142].

The problem of sensor scheduling described in [47,53,96,97] is similar in many ways to the handoff problem. No literature addresses these similarities, but they are outlined in the following sections.

In present handoff schemes, it is assumed that the action taken (for example in favour of handoff) will have the expected effect (for example continuation of the call). This is not always true and hence we should model the handoff decision making to include the probabilities involved. Partially observable Markov models enable this extension. Various algorithms described in [3] can solve such partially observable Markov processes.

B.3 What is a Hidden Markov Model (HMM) ?

If the system to be modeled is assumed to be a Markov process, and the parameters in particular states are unknown, a hidden Markov model (HMM) can be used. The unknown or hidden parameters are extracted from the observable parameters, which can then be used for further analysis of the system. A general Markov model has a state transition matrix associated with directly visible states. A HMM is a Markov model in which the state transitions are not directly visible but are observed through feature vectors that have a probability distribution within each state. X_i is a finite-state Markov chain with transition probability matrix A and initial distribution π_0 taking values in a discrete set of states $\{e_1, e_2, .., e_{N_s}\}$, where N_s is the number of states. Y_i is another known stochastic

process such that the filtering prediction in a HMM can be defined as (A, B, π_0).
A = State transition probability matrix,
B = Observation symbol probability matrix,
π_0 = Initial probability vector.

In a hidden Markov model, the state is not directly visible, but output from the state transition is visible. Output symbols are described by a probability matrix of states to output symbol transitions. Therefore it is expected that the sequence of symbols generated by a HMM provide information about the sequence of hidden states. Applications of HMM vary from speech, handwriting and gesture recognition to bioinformatics.

B.4 Similarities Between Sensor and Base Station Scheduling

B.4.1 Sensor Scheduling

1. Each time one or more sensors that can provide the next measurement is selected.

2. Each measurement is associated with a cost calculated using a cost function.

3. The best sensor selection should minimize the cost function

B.4.2 Base Station Scheduling

1. Base stations and noisy sensors are similar in the Markov model.

2. We need to select the serving base station of each sample point.

3. In this case, discrete distance intervals are considered (instead of time intervals).

4. Costs comprise the direct cost of using all modes, the cost of receiving lower signal strengths of selected serving base stations and the handoff cost of all

Table B.1: Similarities between sensor scheduling and base station scheduling

Sensor Scheduling	**Base Station Scheduling**
The best sensors that minimize the cost function are selected	The best observation modes that minimize the cost function are selected (example: base stations B_1, B_2, B_3 is an observation mode).
Observations are made from the selected sensors or prediction based on the past experience.	Observations are made from observations modes or prediction (existing handoffs do not consider this option).
Hidden states may represent distance to the target (close or far away) or friend or enemy status.	Hidden states may represent any combination of • the assigned base station selected from the set of base stations • an immediate handoff occurred or not • received signal strength is greater than S_{min} or not.

modes.

5. The best base station scheduling should minimize the cost function.

B.5 Basic Steps of Base Station Scheduling

Assume that there are two possibilities for selecting the observations, $\theta_m = \in \{\theta_1, \theta_2\}$:

- **Step 1:** Select θ_1 or θ_1, as in Fig. B.2
- **Step 2:** Observe the symbols
- **Step 3:** Estimate state at $i + 1$

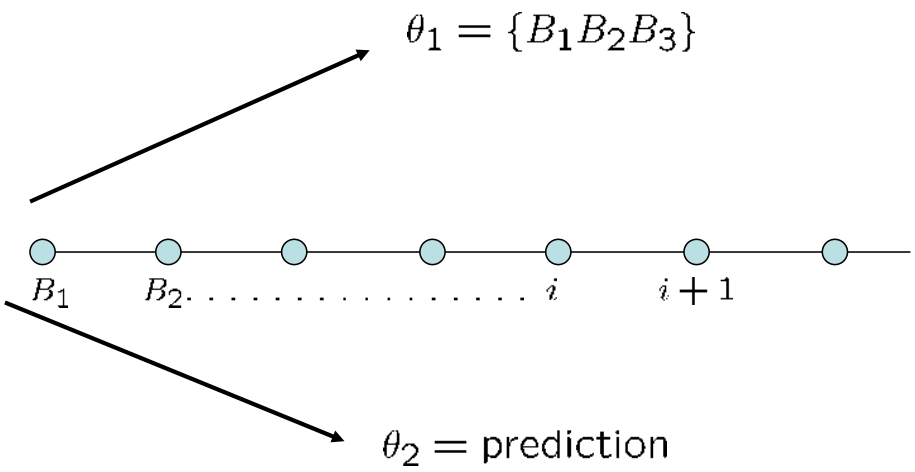


Figure B.2: Observation mode of the base station scheduling, where $\theta_m =\in \{\theta_1, \theta_2\}$

B.6 Observation Mode Selection

We consider a cellular mobile network with M (noisy) base stations designated $B_1, B_2, ..., B_M$. Let the sample path be an arbitrary path in which a mobile user is traveling. Sample points are points on the sample path for which the signal strength values received from base stations are measured. Let $i = 1, .., N$ denote discrete sample points. Let S_{iB_j} be the signal strength at sample point i received from base station B_j. Let $B^i \in \{B_1, B_2, ..., B_M\}$ denote the base station assigned to the i^{th} sample point. Let S_{min} be the minimum signal strength below which the signal quality is unacceptable for the user, meaning that a successful link must fullfill $S_{iB^i} \geq S_{min}$.

The base station assignment at each sample point can be compared with the sensor selection in the sensor scheduling problem. For example, target tracking via sensor scheduling may involve an active mode sensor (an active radar that may disclose its position to the target) or a passive mode sensor (such as sonar), or a predictive mode where the target is predicted from past observations only. For example, if the active radar is close to the target, then it may be very costly to use it, as it may disclose the location of the radar to the target. Therefore a passive sensor or the predictive mode would be a better selection [96]. In handoff, we consider three modes of observation analogous to the selection of sensors in

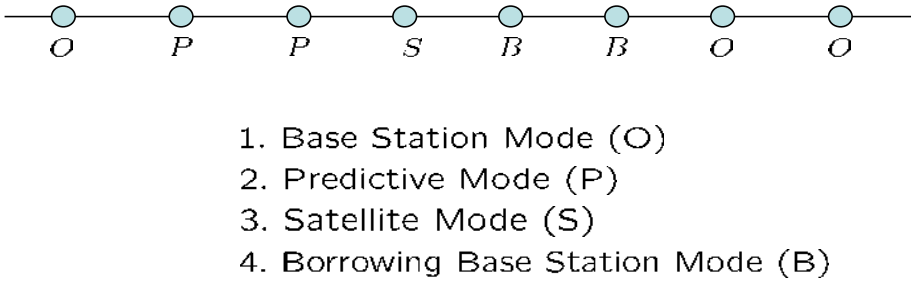


Figure B.3: Different modes of observation that can be used

sensor scheduling. In the base station mode, we observe signal strengths from various base stations at the sample point to decide the base station to be assigned. In the satellite mode, a satellite determines the next assigned base station. In the predictive mode, we can also predict the next assigned base station without observing any base stations or satellites (Fig. B.3).

All on-line solutions for the handoff problem observe signal strengths only from base stations in the vicinity, and there is no option to select the observation mode. However, the signal strengths received from base stations at the sample point may not be useful and may even lead to suboptimal decisions in some situations. For example

- when the base station assigned to the previous sample point is still suitable for the new sample point and no handoff is required, so we do not need to observe the signal strengths at the new sample point, or
- when none of the base stations in the present sample point is suitable, and it is likely to stay that way for the new sample point.

From a practical perspective, if the signal strengths observed (usually from all base stations in the vicinity of the user) provide either noisy or unacceptable or almost equally acceptable levels of signal strengths, they may be useless for selecting the next base station. Therefore, base station selection based on prediction instead of observation may be useful in these cases.

The base station or satellite observation mode is identical to the HMM (optimal filtering problem) and the predictive mode is identical to the prediction problem known in HMM. The cost function for both modes can be formulated, includ-

ing the state estimation error and mode selection cost. The stochastic backward dynamic programming recursion can be used for estimating the optimal HMM mode scheduling policy. The dynamic programming recursion can be transferred to practical solutions using the heuristic algorithms reported in [3, 96]. Some linear programming based algorithms for solving partially-observable Markov decision processes (POMDPs) are described in [3]. Most optimal scheduling solutions are based on linear cost functions. Another option is the approximation of non-linear cost functions by piecewise linear functions as in [97].

B.7 Base Station Assignment with HMM

Assume X_i is an N_s-state HMM with state space $\{e_1, e_2, .., e_{N_s}\}$, where e_k denotes the N_s-dimensional unit vector with 1 in the k^{th} position and zeros elsewhere. This choice simplifies the subsequent notation [96]. A HMM is characterised by number of states (N_s): total number of finite symbols (N_O), the discrete set of possible observation symbols $\{O_1, .., O_{N_O}\}$, and the length of the observation sequence (T). The initial probability vector π_0 is,

$$\pi_0 = [\pi_o(k)]_{N_s \times 1}, \; where \; \pi_o(k) = P(X_o = e_k) \quad and \; k \in \{1, .., N_S\}, \tag{B.1}$$

where $\pi_o(k)$ is the probability of being in state k at the beginning of the state sequence and X_0 is an initial state. The state transition probability matrix is $A = [a_{kl}]_{N_s \times N_s}$, where

$$a_{kl} = P(X_i = e_l | X_{i-1} = e_k),$$
$$k, l \in \{1, .., N_s\} \;\; and \;\; \sum_{k=1}^{N_s} a_{kl} = 1.$$

We define the observation symbol probability distribution,

$$Z(u_i = \theta_m) = [b_i(O_r(\theta_m))]_{N_S \times N_O}$$

where

$$b_i(O_r(\theta_m)) = P(y_i(u_i) = O_r(u_i)|X_i = e_k, u_i = \theta_m). \tag{B.2}$$

Here u_i is the mode of observation at sample point i selected as $\theta_m \in \{\theta_1, ..., \theta_{N_m}\}$, where N_m is the total number of available observation modes and $1 \geq m \geq N_m$. For example $u_i = B_1, B_2, .., B_M$ can be the set of M base stations selected from a group of L base stations for observation where ($M \leq L$) and $y_i(u_i)$ denote the observations from the selected mode, $r \in \{1, .., N_O\}$, and $k = 1, 2, .., N_S$. The formula $Y_i = \{u_1, u_2, .., u_i, y_1(u_1), y_2(u_2), .., y_i(u_i)\}$ represents the information available at the sample point i, and the selection of the observation mode at the sample point $i + 1$ is based on Y_i. As with the three basic steps considered in sensor scheduling, we can proceed with the following stages:

- **Scheduling:**
  Generate $u_{i+1} = f_{i+1}(Y_i)$, where f_{i+1} denotes the policy to determine the next observation mode

- **Observation:**
  Observe the symbols $y_{i+1}(u_{i+1})$.

- **Estimation:**
  Find the optimal state estimate π_{i+1} (which is a column vector of dimension N_S) of the Markov model X_{i+1} using the forward algorithm [138]:

  $$\pi_{i+1} = E\{X_{i+1}|Y_{i+1}\}. \tag{B.3}$$

  Since we consider that X_i is a unit vector,

  $$\pi_i(k) = P(X_i = e_k|Y_i),$$

  where $k \in \{1, 2, ..N_s\}$.

B.8 Cost Estimation

Let us model the user movement through the sample points as a HMM. Assume that M base stations are selected for observation, and $N_s = 2M$. When the user at a sample point i is served by base station B^i, we consider two possibilities: no handoff occurred at the i^{th} sample point (i.e. $B^i = B^{i-1}$) or a handoff occurred at the i^{th} sample point (i.e. $B^i \neq B^{i-1}$). In order to distinguish between these two cases, we denote the serving base station in the first case by $B^i = B_j$ and in the second case by $B^i = B_j^H$, where $j \in \{1, ..., M\}$. We choose 2M states as: $e_1 = B_1$,...., $e_M = B_M$, $e_{M+1} = B_1^H$,....,$e_{2M} = B_M^H$. The transitional probabilities of the Markov Chain can be provided with matrix $A = [a_{kl}]_{2M \times 2M}$ where $a_{kl} = P(X_i = e_l | X_{i-1} = e_k)$, and $k, l \in \{1, ., 2M\}$.

We consider the cost functions $c(X_i = e_k, u_{i+1} = \theta_m)$. The observation mode selection is associated with two types of costs:

- **State estimation cost:**
  As in the sensor scheduling problem, state estimation error can be computed as the weighted distance $E_i^m = \alpha_i(\theta_m)||X_i - \pi_i||_D$, where $\alpha_i(\theta_m)$ are known positive weights for all observation modes, and the distance function is assumed to be a convex function. For example D can represent the l_2 norm.

- **Mode usage costs:**
  The cost $c_i(X_i, \theta_m)$ is the cost of using the mode $u_{i+1} = \theta_m$, when the state in the Markov model is X_i. In our application, if $X_i \in \{e_{M+1}, ..., e_{2M}\}$, there will be handoff related costs associated with the mode selected for the subsequent sample point. Further, a cost that is inversely proportional to the signal strength of the assigned base station will be added to the constant cost associated with each mode.

We can express the cost function as the summation of these two components:

$$c(X_i = e_k, u_{i+1} = \theta_m) = \alpha_i(\theta_m)||X_i - \pi_i||_D + c_i(X_i, \theta_m). \tag{B.4}$$

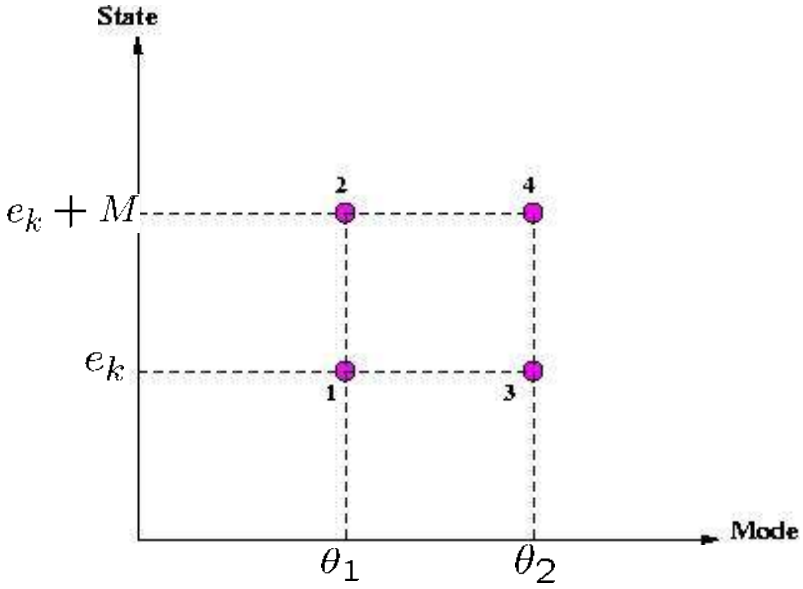


Figure B.4: Cost estimation: separate cost functions for 4 cases

Dynamic programming is often used to find the most appropriate mode schedule along the sample points that minimizes the cost, as in [96]. However, this approach may not lead to practical solutions, unless approximations are used. For linear or piecewise linear cost functions, such approximations may yield good results. But the problem is still unresolved for cost functions that cannot be approximated by piecewise linear functions. A Genetic Algorithm (GA) can be used to obtain the optimal mode schedule without such restrictions limiting the formulation of cost functions. A population of solutions can be initialized with an individual member of the population representing a possible mode schedule (or sequence). For example with $N = 5$ sample points and two modes (encoded as "1" and "0") available for scheduling: "10101" represents a possible individual. Crossover and mutation steps of the GA will improve the costs of the individuals, and finally we can approximately find the best individual (i.e. the best mode schedule) that minimizes the cost. GA based methods are known to be time consuming, but similar less computationally intensive methods such as Particle Swarm Optimization can be applied.

When the best mode schedule is known, we can estimate the states at all sample points, and obtain the optimal base station schedule.

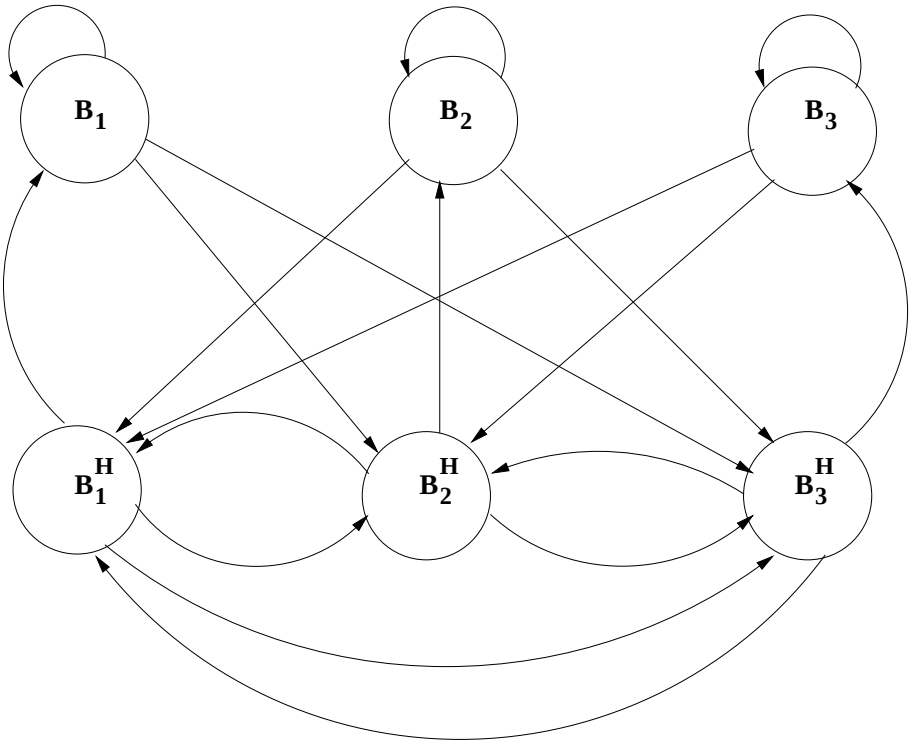


Figure B.5: HMM state transition diagram for $M = 3$

B.9 An Application Example

In order to simplify the problem, let us assume $M = 3$ (i.e., $N_s = 6$). Let the probability of handoff between two consecutive samples be represented by p_1, and assume that both handoff options (for example from B_2 to B_1^H or B_3^H) are equiprobable. By the definition of the states all transition probabilities from B_j and B_j^H to B_j^H are zero, where $j \in \{1, 2, 3\}$. The matrix A is given as follows (see also Fig. B.5):

$$A = \begin{pmatrix} 1-p_1 & 0 & 0 & 0 & \frac{p_1}{2} & \frac{p_1}{2} \\ 0 & 1-p_1 & 0 & \frac{p_1}{2} & 0 & \frac{p_1}{2} \\ 0 & 0 & 1-p_1 & \frac{p_1}{2} & \frac{p1}{2} & 0 \\ 1-p_1 & 0 & 0 & 0 & \frac{p_1}{2} & \frac{p_1}{2} \\ 0 & 1-p_1 & 0 & \frac{p_1}{2} & 0 & \frac{p_1}{2} \\ 0 & 0 & 1-p_1 & \frac{p_1}{2} & \frac{p_1}{2} & 0 \end{pmatrix}.$$

A reasonable estimate for the initial state vector is:

$$\mathfrak{B}_0 = \begin{pmatrix} 1/3 \\ 1/3 \\ 1/3 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Assume that there are two possible ($N_m = 2$) observation modes $\theta_m \in \{\theta_1, \theta_2\}$. In θ_1, we observe the signal strengths received from three noisy base stations. It is then possible to use a HMM state filter (also known as the forward algorithm) to estimate the state at the $(i+1)^{th}$ sample point [96]. In θ_2 we use a HMM state predictor to estimate the next state, without using the signal strengths received from the base stations. This is similar to using an "active" sensor or "passive-prediction" sensor in the aircraft identification problem described in [96].

When using θ_1, the observation (signal strength) S_{i+1B_j} measured from base station B_j at the $(i+1)^{th}$ sample point is discretized into $y_{i+1}(B_j) \in \{1, 0\}$, with two corresponding ranges of signal strength: 1 if $S_{i+1B_j} \geq S_{min}$, 0 if $S_{i+1B_j} < S_{min}$. There are 2^3 possible observation symbols $y_{i+1}(B_1B_2B_3)$ with $T = 3$ in the case of mode θ_1, and there is an additional observation symbol "nothing" when θ_2 is selected. All possible observation symbols ($N_O = 9$) are: $O_1 = 000$, $O_2 = 001$, $O_3 = 010$, $O_4 = 011$, $O_5 = 100$, $O_6 = 101$, $O_7 = 110$, $O_8 = 111$, $O_9 = nothing$.

Let us assume that the probability of receiving $S_{i+1B^i} \geq S_{min}$ is p_2, where S_{i+1B^i} is the signal strength received from B^i, which is the base station assigned to the i^{th} sample point, at the sample point $i+1$. The prediction (in θ_2) obviously leads to the observation symbol "nothing" with the probability of 1. We can define 6×9 matrix of symbol probabilities $Z(u_i = \theta_m) = [P\{y_i(u_i) = O_r(u_i) | X_i = e_k, u_i = \theta_m\}]_{6\times 9}$ for each observation option:

$$Z(u_i = \theta_1) = \begin{pmatrix} \frac{(1-p_2)}{4} & \frac{(1-p_2)}{4} & \frac{(1-p_2)}{4} & \frac{(1-p_2)}{4} & \frac{p_2}{4} & \frac{p_2}{4} & \frac{p_2}{4} & \frac{p_2}{4} & 0 \\ \frac{(1-p_2)}{4} & \frac{(1-p_2)}{4} & \frac{p_2}{4} & \frac{p_2}{4} & \frac{(1-p_2)}{4} & \frac{(1-p_2)}{4} & \frac{p_2}{4} & \frac{p_2}{4} & 0 \\ \frac{(1-p_2)}{4} & \frac{p_2}{4} & \frac{(1-p_2)}{4} & \frac{p_2}{4} & \frac{(1-p_2)}{4} & \frac{p_2}{4} & \frac{(1-p_2)}{4} & \frac{p_2}{4} & 0 \\ \frac{(1-p_2)}{4} & \frac{(1-p_2)}{4} & \frac{(1-p_2)}{4} & \frac{(1-p_2)}{4} & \frac{p_2}{4} & \frac{p_2}{4} & \frac{p_2}{4} & \frac{p_2}{4} & 0 \\ \frac{(1-p_2)}{4} & \frac{(1-p_2)}{4} & \frac{p_2}{4} & \frac{p_2}{4} & \frac{(1-p_2)}{4} & \frac{(1-p_2)}{4} & \frac{p_2}{4} & \frac{p_2}{4} & 0 \\ \frac{(1-p_2)}{4} & \frac{p_2}{4} & \frac{(1-p_2)}{4} & \frac{p_2}{4} & \frac{(1-p_2)}{4} & \frac{p_2}{4} & \frac{(1-p_2)}{4} & \frac{p_2}{4} & 0 \end{pmatrix},$$

and

$$Z(u_i = \theta_2) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

We define our cost function as:

$$\begin{aligned} c(X_i = e_{k'}, u_{i+1} = \theta_1) &= \rho^s + \frac{r^{signal}}{S_{iB^i}} + E_i^1 \\ c(X_i = e_{k'+M'}, u_{i+1} = \theta_1) &= \rho^s + \frac{r^{signal}}{S_{iB^i}} + r_1^h + E_i^1 \\ c(X_i = e_{k'}, u_{i+1} = \theta_2) &= \rho^p + \frac{r^{signal}}{S_{iB^i}} + E_i^2 \\ c(X_i = e_{k'+M'}, u_{i+1} = \theta_2) &= \rho^p + \frac{r^{signal}}{S_{iB^i}} + r_2^h + E_i^2, \end{aligned}$$

where $k^{'} \in \{1, 2, 3\}$, ρ^s and ρ^p denote the direct cost of using θ_1 and θ_2 respectively, r^{signal} is the cost associated with receiving lower signal strengths, and r_1^h and r_2^h are the handoff costs associated with using the modes θ_1 and θ_2. We can assume that $\rho^s > \rho^p$, meaning that using received signal strengths is more costly than using prediction. The cost is inversely proportional to the received signal strength with the constant r^{signal} in both modes, as we wish to maximize the re-

ceived signal strength. We may select the handoff costs $r_1^h > r_2^h$ to encourage the use of prediction immediately after a handoff. The state estimation error cost can be defined using l_2 cost ($D = 2$):

$$E_i^1 = E_i^2 = \alpha_i(1 - \pi_i^{'}\pi_i), \tag{B.5}$$

as derived in [96], where $\alpha_i = \alpha_i(\theta_1) = \alpha_i(\theta_2)$.

B.10 Three States

We used a HMM with $2M$ states in this work. Another option is to consider $3M$ states extending the above HMM: $e_1 = B_1^+, ..., e_M = B_M^+$, $e_{M+1} = B_1^H, ..., e_{2M} = B_M^H, e_{2M+1} = B_1^-, ..., e_{3M} = B_M^-$, where B_j^+ denote the assigned base station j with signal strength $\geq S_{min}$ and no handoff occurring, B_j^- denotes signal strength $\leq S_{min}$ and no handoff, and B_j^H denotes signal strength $\geq S_{min}$ and a handoff occurring (Fig. B.6). This three states model simplifies the cost functions.

B.11 Summary

This work exploits analogies between sensor scheduling and the base station assignment, which is also similar to the handoff problem in cellular networks. The mobile user is modeled using a Hidden Markov Model (HMM). The handoff problem is formulated as an optimization problem of base station scheduling that minimizes cost functions containing the HMM state estimation error and base station measurement costs. The formulated problem can be solved using either algorithms known in partially observed Markov decision processes or population based optimization methods such as Genetic Algorithms.

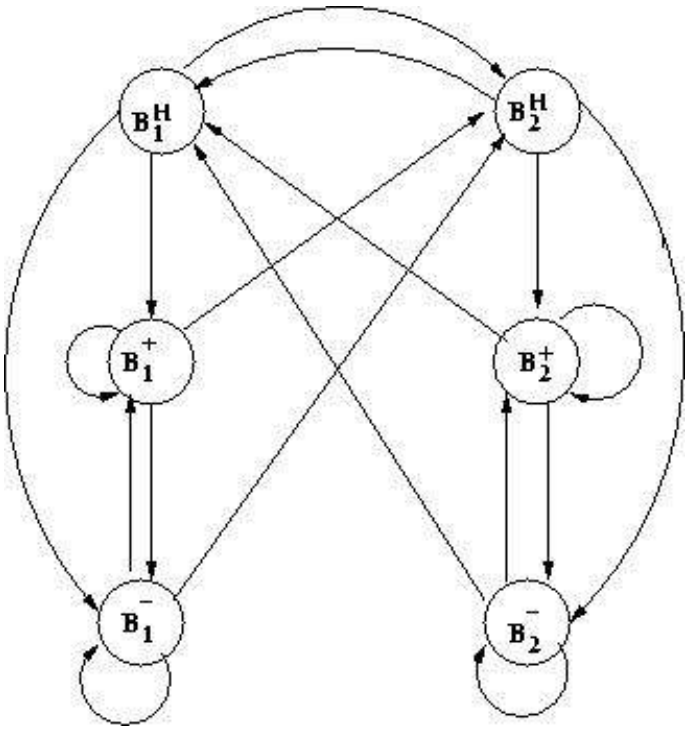


Figure B.6: HMM three state transition diagram for $M = 3$

APPENDIX C

Average Distance

C.1 Estimation of the Mean Square Distance from Sensor to CH

The average linear distance from sensor to CH,

$$d_{toCH} = \frac{M}{2(k+1)}, \tag{C.1}$$

where k is the number of clusters with an M length straight line.

When we consider area with $M \times M$, the area occupied by each cluster is

$$\approx \frac{M^2}{k},$$

where M^2 is the area of the network and k is the number of clusters.

The distance from the sensor (assumed to be at the center of the area) to any point in the square area is given by

$$d_{toCH}(x,y) = \sqrt{\left(x - \frac{M}{2}\right)^2 + \left(y - \frac{M}{2}\right)^2}.$$

When we consider a square with $M \times M$, the area covered by each cluster is nearly M^2/k.

As in [73], we assume that the CH is at the center of mass of its cluster, and we

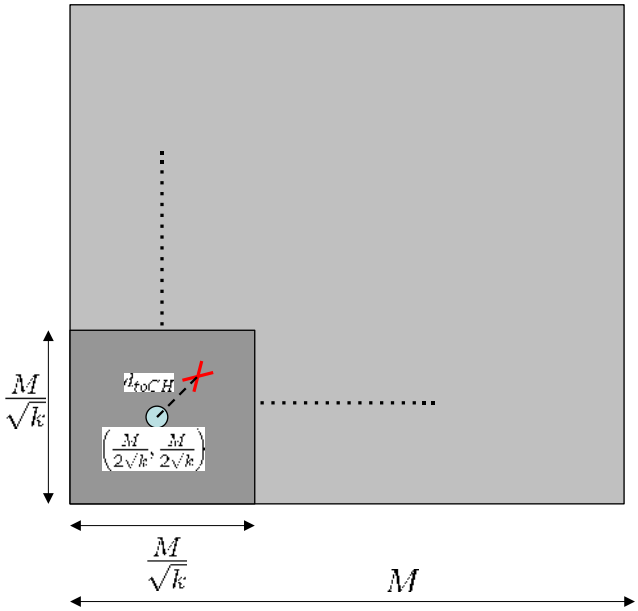


Figure C.1: Mean square distance from sensor to CH, when the area measures $M \times M$ with k number of clusters

acknowledge that the cluster area can be arbitrarily shaped, but for simplicity, we assume that it is a square. For $k = 1$, assuming sensors are randomly uniformly distributed over the square area, the mean square distance from a sensor to its CH is given by

$$\begin{aligned} E[d_{toCH}] &= \int_0^M \int_0^M d(x,y)\rho(x,y)dxdy \\ &= \frac{1}{M^2}\int_0^M \int_0^M \left(x - \frac{M}{2}\right)^2 + \left(y - \frac{M}{2}\right)^2 dxdy, \end{aligned}$$

where $\rho(x,y)$, $0 \leq x,y \leq M$, is the joint probability density function. (If sensors are placed uniformly then we have $\rho(x,y) = \frac{1}{M^2}$.)

The mean square distance for one ($k = 1^2$) cluster is given by

$$E[{d_{toCH}}^2] = \frac{1}{M^2}\int_0^M \int_0^M \left(x - \frac{M}{2}\right)^2 + \left(y - \frac{M}{2}\right)^2 dxdy$$

$$E[{d_{toCH}}^2] = \frac{1}{M^2}\int_0^M \left(\left[\frac{x}{6}(3M^2 - 3(x+2y)M + 2(x^2+3y^2)) \right]_0^M \right) dy$$

$$E[{d_{toCH}}^2] = \frac{1}{M^2}\int_0^M \left(\left[\frac{M}{6}(3M^2 - 3(M+2y)M + 2(M^2+3y^2)) \right] \right) dy$$

$$E[{d_{toCH}}^2] = \frac{1}{M^2}\left[\frac{My}{6}(2M^2 - 3My + 2y^2) \right]_0^M$$

The mean square distance for one cluster ($k = 1^2$) is given by

$$E[{d_{toCH}}^2] = \frac{1}{M^2}\frac{M^4}{6} = \frac{M^2}{6}. \tag{C.2}$$

Similarly the mean square distance for four cluster ($k = 4$) is given by

$$E[{d_{toCH}}^2] = \frac{4}{M^2}\int_0^{\frac{M}{2}}\int_0^{\frac{M}{2}} \left(x - \frac{M}{2\sqrt{4}}\right)^2 + \left(y - \frac{M}{2\sqrt{4}}\right)^2 dxdy,$$

and the mean square distance for nine cluster ($k = 9$) is given by

$$E[{d_{toCH}}^2] = \frac{9}{M^2}\int_0^{\frac{M}{3}}\int_0^{\frac{M}{3}} \left(x - \frac{M}{2\sqrt{9}}\right)^2 + \left(y - \frac{M}{2\sqrt{9}}\right)^2 dxdy$$

The mean square distance for k cluster ($k = n^2$) (as in Fig. C.1) is given by

$$E[{d_{toCH}}^2] = \frac{k}{M^2}\int_0^{\frac{M}{\sqrt{k}}}\int_0^{\frac{M}{\sqrt{k}}} \left(x - \frac{M}{2\sqrt{k}}\right)^2 + \left(y - \frac{M}{2\sqrt{k}}\right)^2 dxdy$$

$$E[{d_{toCH}}^2] = \frac{k}{M^2}\int_0^{\frac{M}{\sqrt{k}}} \left(\left[\frac{x^3}{3} - \frac{Mx^2}{2\sqrt{k}} + \frac{M^2x}{2k} + y^2x - \frac{Myx}{\sqrt{k}} \right]_0^{\frac{M}{\sqrt{k}}} \right) dy$$

$$E[{d_{toCH}}^2] = \frac{k}{M^2}\int_0^{\frac{M}{\sqrt{k}}} \left\{ \frac{y^2M}{\sqrt{k}} - \frac{M^2y}{k} + \frac{M^3}{3k^{\frac{3}{2}}} \right\} dy$$

$$E[{d_{toCH}}^2] = \frac{k}{M^2}\left[\frac{2M^3y - 3\sqrt{k}M^2y^2 + 2kMy^3}{6k^{\frac{3}{2}}} \right]_0^{\frac{M}{\sqrt{k}}}$$

Finally, the mean square distance for k cluster ($k = n^2$) is given by

$$E[{d_{toCH}}^2] = \frac{M^2}{6k}, \tag{C.3}$$

when $n = 1, 2, 3, ...$

Although (C.3) applies to $k = 1, 4, 9, 16, ...$ under our original assumption of square shaped clusters, we use here (C.3) to evaluate the mean distance between the CH and a node in its cluster for any integer value of k. Here we assume the density of nodes is uniform through out the cluster area.

C.2 Obtain the Distance from Sensor to CH when Number of Sensors are Doubled

From eq (7.15), we can obtain the optimal number of clusters, C_{opt} as

$$C_{opt} = \left\lceil \frac{\sqrt{N_s}}{\sqrt{6}} \frac{M}{d_{toBS}^2} \sqrt{\frac{E_{fs}}{D_\alpha}} \right\rceil, \tag{C.4}$$

where $D_\alpha = (E_{mp} + E_{sens_{CH}} + E_{tran_{CH}} + E_{logg_{CH}})$.

Let C_{opt_N} represent the optimal number of clusters for N sensors

$$C_{opt_{N_s}} = \sqrt{\frac{N_s}{6}} \frac{M}{d_{toBS}^2} \sqrt{\frac{E_{fs}}{D_\alpha}}, \tag{C.5}$$

Similarly, let $C_{opt_{2N}}$ represent the optimal number of clusters for N sensors

$$C_{opt_{2N_s}} = \sqrt{\frac{N_s}{3}} \frac{M}{d_{toBS}^2} \sqrt{\frac{E_{fs}}{D_\alpha}}, \tag{C.6}$$

At the optimal value $k \rightarrow C_{opt}$.

By substituting (C.5) for (C.3), the average distance for N sensors is given by

$$\begin{aligned} {d_{toCH}}^2(N_s) &= \frac{M^2}{6C_{opt_{N_s}}}, \\ &= \frac{M^2}{6\sqrt{\frac{N_s}{6}} \frac{M}{d_{toBS}^2} \sqrt{\frac{E_{fs}}{D_\alpha}}}. \end{aligned} \tag{C.7}$$

The average distance for $2N$ sensors is given by

$$\begin{aligned} {d_{toCH}}^2(2N_s) &= \frac{M^2}{6C_{opt_{2N_s}}}, \\ &= \frac{M^2}{6\sqrt{\frac{N_s}{3}} \frac{M}{d_{toBS}^2} \sqrt{\frac{E_{fs}}{D_\alpha}}}. \end{aligned} \tag{C.8}$$

From (C.7) to (C.8)

$$\frac{{d_{toCH}}^2(N_s)}{{d_{toCH}}^2(2N_s)} = \sqrt{2}. \tag{C.9}$$

Therefore, when the number of sensors are doubled with uniform distribution, the distance between sensor and CH is reduced by a factor of the square root of two.