Dynamic and reliable Information Accessing and Management in Heterogeneous Wireless Networks

Tobgay, Sonam

Publication date:
2013

Document Version
Accepted author manuscript, peer reviewed version

Dynamic and Reliable Information Accessing and Management in Heterogeneous Wireless Networks

a dissertation
submitted to the department
of electronic systems
of
Aalborg University
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

Sonam Tobgay
Center for TeleInFrastrutur

30^{th} September, 2013

Supervisor:
Professor Ramjee Prasad, Aalborg University, Denmark

Co-Supervisor:
Associate Professor Rasmus L. Olsen, Aalborg University, Denmark

Examination Committee:
Associate Professor Jens Myrup Pedersen (Chairman), Aalborg University, Denmark
Professor Luis Muñoz Gutierrez, Universidad De Cantabria, Spain
Professor Zary Segall, University of Maryland, Baltimore County, USA

Moderator:
Associate Professor Albena D. Mihovska, Aalborg University, Denmark

Date of Defence:
30^{th} September, 2013

Abstract

Due to rapid growth and advancement in the field of wireless technology, many portable devices which are capable of connecting each other, forming ad-hoc wireless networks; have been evolved in the recent past. When these different devices are forming different ad-hoc networks, a heterogeneous wireless network is created. This shows that, there are different sources of information that one can select from, provided one has access rights to all these available sources. One such ad-hoc network is wireless sensor network, which is applied to almost all the fields in order to collect information for various purposes. Wireless sensor network, by itself is a heterogeneous, in the sense that it can be composed of sensor nodes which have different properties and characteristics. There are also different types of sensor nodes which can measure same physical phenomenon, but exhibits different characteristics in sensing events. Due to the information dynamics at the source (which changes over time), which wireless sensor nodes are made to collect, the degree of information reliability will be determined by the type of source being chosen. In order to select the most reliable information source, it is necessary to know the network parameters like, at what rate information is being generated at the source, at what rate the sensor node sample information and how much time the information will take to reach at the destination, where information is being processed or used. In addition, it is also important that accessing information from a particular source node does not cause huge power consumption (in case if there is no access to power source or no any energy recovery mechanism). Since information access from a particular source can be done in various ways; 1) Reactive Access 2) Periodic Access and 3) Hybrid Access, it is also important to study the merits and demerits of each access strategy in relation to the information reliability and power consumption.

Various protocols for information dissemination and collection have been proposed for wireless sensor networks. Each protocol is typically designed to optimize a certain set of parameters in making routing decisions, and likewise tested for chosen network traffic as per the designed parameters.

Which means a deployment of any application needs to build its entire deployment logic based on the type of traffic for which the protocol was designed. To make information accessing from wireless networks more viable, a generic framework is required to support any application to be built on top of any arbitrary communication protocols being used. Therefore, this thesis presents information accessing and management framework and studies the behaviour of dynamic information elements in relation to information reliability and power consumption.

First, in order to study the dynamic behaviour of network parameters and to estimate the information reliability in accessing particular information source, it is crucial to know how information is being generated and detected. So, this dissertation starts by presenting a brief study on Event Process and Event Detection process based on wireless sensor node.

Secondly, a dynamic information accessing and management framework is developed to monitor the network dynamics. The framework is mainly composed of two parts, one the information source which is located at remote site and the other which monitors the information sources. Based on the captured event rate, the estimated delay rate, and frequency of sampling, the probability of mismatch of information that will occur from accessing particular information source is estimated. The framework is also capable of switching dynamically from one access strategy to another based on network parameters in order to minimize the overall power consumption of the system.

The third part of this thesis provides analytical models for information mismatch probability and power consumption under each access strategies mentioned earlier. A detailed comparative study on these three access strategies are provided with performance evaluation and analysis based on the dynamic network parameters like event rate, rate of delay and sampling rate. The validations of analytical results are done by using the developed framework.

Considering, these two requirement parameters, information reliability and power consumption, the fourth part of this thesis presents an adaptive information access mechanism thereby suggesting a trade-off point between two requirement parameters, so that energy is not extensively utilized for a minimal gain in information reliability or compromising information reliability by trying to save minimal energy. Such a mechanism will be useful as different applications have different requirements as per their purpose.

Lastly, the thesis presents a dynamic algorithm which selects the most reliable information source based on information mismatch probability and power consumption; both are relevant metrics for sensor networks. With the help of the algorithm, an analysis on how dynamic information elements and stochastic network delay influence information mismatch probability is

provided.

In summary, the thesis provides in-depth study on the impact of accessing dynamic information on information reliability and power consumption; both are relevant metrics for wireless networks. The study found out that the information mismatch probability not only depends on the information dynamics like at what rate the information is being generated at source and the delay in accessing the information, but it is also impacted by the rate at which the information is being sampled and which type of access strategy being used.

Dansk Resume

På grundlag af den hastigt voksende og udviklende markedet indenfor trådløs teknologier, er mange mobile enheder i stand til at forbinde hinanden og danne ad-hoc trådløse netværk. Når disse forskellige typer enheder danner disse ad-hoc netværks, tales der om heterogene netværk. I den forbindelse vil der være forskellige informationskilder at vælge fra når applikationer skal have fat i informationer i dets omkringliggende miljø. Dette er dog givet at der er givet tilladelse til at anvende disse informationskilder. En speciel type af sådanne ad-hoc netværk er trådløse sensor netværk, som anvendes til næsten alle områder indenfor informations- og dataopsamling med forskellige formål. Trådløse sensor netværk alene er heterogene i den forstand at de kan bestå af sensor noder med forskellige egenskaber og karakteristika. Der findes forskellige typer af sensorer som kan måle samme fysiske parametre, men udviser forskellige egenskaber i at måle hndelser i omgivelserne. På grund af informationsdynamikken hos informationskilden (som ændres over tid) vil pålideligheden af den information der bliver tilgået blive influeret af informationskildens type. For at vælge den mest pålidelige informationskilde, er det nødvendigt at kende til netvrksparametre såsom med hvilken hastighed informationer ændres ved kilden, hvor hurtig informationer samples, hvor lang tid det tager for informationer at blive transporteret over et netværk fra kilde til modtager og hvor lang tid der bliver brugt på at processere data. Yderligere er det også vigtigt at det tilgåede information fra en bestemt kilde ikke anvender for meget strøm (i det tilfælde en node f.eks. er batteridrevet og/eller ikke har adgang til at få fornyet sit energilager). Siden der findes forskellige måder at tilgå information på fra en bestemt kilde; 1) reaktiv tilgang, 2) periodisk opdateringer og 3) hybride metoder, er det også vigtigt at studere fordele og ulemper ved de enkelte strategier i forhold til informationspålideligheden og deres energiforbrug.

Forskellige protokoller til trådløse netværk har været foreslået for at kunne sprede og samle informationer til rette. Individuelle protokoller er typisk designet til at optimiere et bestem set af parametre til at bestemme datapakkeruter gennem netværket, og er ligeledes testet for udvalgte typer af

netværkstraffik jvnf. valgte optimeringsparametre. Dette betyder at udspredelsen af forskellige applikationer er nød til at basere sin indre logik baseret på typen af traffik de anvendte protokoller oprindelig var designet til. For at gøre adgang til information mere dynamisk, er der behov for et framework der kan understøtte forskellige applikationer ovenpå en række protokoller anvendt i sensor netværks.Derfor studeres der i denne afhandling eksisterende metoder og management frameworks til adgang for distribuerede informationer, og der studeres også pålidelighed for adgang til dynamiske, distribuerede informationer i relation til energiforbrug for at opnå denne tilgang. Dette gøres i flere trin:

Først, for at studere den dynamiske opførsel af netværksparametre og estimere informations pålideligheden ved tilgang af informationer fra bestemte kilder, er det vigtigt at vide noget om hvorledes informationer skabes og hndelser detekteres. Så afhandlingen starter med et overblik over hvorledes hndelser processeres og detekteres. Denæst bliver der udviklet et dynamisk informations tilgangs og management framework til at monitorere netværks og informationsdynamikken. Frameworket består hovedsagligt af to dele; informationskilden og en monitoreringsenhed der interagere med informationskilden. Baseret på estimat af hændelseshastigheder, estimerede delays i netværket og samplingsfrekvens, bliver sandsynligheden for mismatches mellem den virkelige værdi og den værdi monitoreringsenheden beregnet. Frameworket vil efterfølgende bruges til dynamisk at vælge mellem forskellige adgangsstrategier og samtidig minimere det overordnede energiforbrug for at få adgang til informationen. Som trejde del, skabes de analytiske modeller nødvendig for at beregne sandsynligheden for mismatches og effektforbrug for de forskellige typer strategier for informations adgang. En detaljeret sammenligning af de forskellige strategier bliver udført og der bliver udført en ydelsesanalyse baseret på de respektive parametre, hændelseshastigheder, delay og samplingsfrekvens. Til valideringen af modellerne anvendes det udviklede framework og emulering af data. Betragtes de to parameter; informations pålidelighed og energiforbrug, vil den fjerde del af afhandlingen præsentere den dynamisk tilpassende informationsmekanisme, som vil given anledning til at lokalisere fornuftige kompromiser mellem pålidelighed og energiforbrug set i forhold til applikationskrav omkring de to parametre, der i høj grad er forskellige fra applikation til applikation. Til sidst, præsenterer afhandlingen en dynamisk algoritme der udvælger den mest pålidelige informationskilde baseret på sandsynligheden for mismatches og effektforbrug betragtenede at forskellige typer sensorer kan vidergive samme information men under forskellige netværksforhold. Algoritmen anvendes ogstå til at analyserere hvorledes dynamiske informationselementer og stochastiske netværksdelay influere mismatch sandynlighed. Samlet set betragter afhandlingen en, i dybden, studie *for indflydelsen af adgang til dynamiske informationers pålidelighed og energiforbrug, der begge er relevante parameter for trådløse netværk. Studiet leder op til at informations mismatches ikke blot er et spørgsmål omkring informationsdynamikken, såsom med hvilken hastighed informationer ændres og delayet mellem kilde og modtager, men også i høj grad med hvilken samplingsfrekvens informationen samples og hvilken type strategi der anvendes til at sende data.*

Acknowledgement

I would like to express the deepest appreciation and heartfelt gratitude to my Supervisor Professor Ramjee Prasad. He continually and convincingly motivated me to take up this PhD study. Without his guidance and unfailing advice this dissertation would not have been possible.

I would like to thank my Co-Supervisor Dr. Rasmus L. Olsen, who taught me not only how to do research but how to be strong and think positive always. Irrespective of what situation he is, the smile on his face was enough to convince me that he has lots of concern towards me. His hard work and dedication towards work highly inspired me. He is not only a good supervisor but is a kind hearted and great person. I have no words to express my gratitude to him.

I also would like to thank Dr. Nuno Pratas for helping me in tackling programming error. I learnt a lot from him while sharing office with him for a year.

I also would like to extend my sincere gratitude to Ms. Dorthe Spare, who helped me in many ways during my PhD study. My sincere gratitude also goes to Ms. Susanne and Ms. Inga for their support.

Lastly, I extend my appreciations to all my colleagues in CTIF section for making me comfortable during my entire study period in Aalborg University.

Thank you one and all.

Dedication

A special feeling of gratitude to my loving wife Kencho Wangmo, daughter Kuenga Lhaki Pelzom and son Garab Chogyal Tobgay for their understanding and encouragement. I also dedicate this work to my father Kuenzang Wangchuk and late mother Chudu Lhamo.

Table of Contents

List of Figures

List of Tables

List of Algorithms

Chapter 1

Introduction and Problem Definition

This chapter provides an overview of the thesis and contains following sections:
Background and Motivation: This section briefly describes the topic which the study is going to explore. It provides the background on why and what motivates to study this particular topic "Dynamic and Reliable Information Accessing and Management in Heterogeneous Wireless Network."
Problem Definitions : This section brings out what challenges an information user faces while accessing to dynamic information elements. It also briefly highlights few issues that have impact on information reliability and power consumption.
Objectives of the Study: Knowing the challenges in accessing dynamic information elements, this section presents the goals which study aims to achieve at the end of the study.
Contributions of the Study: This section presents a summary of what significant study is carried out in this thesis.
Structure of the Thesis: Lastly, this section presents a brief summary on each chapter, describing how one chapter is linked to another. A block diagram, showing interactions among chapters is also provided.

1.1 Background and Motivation

Although, the benefits of wired networking technology cannot be neglected but wireless networking technology provides greater flexibility at home, work place and in the community to connect to the information source without being tied to a specific location. The flexibility in accessing information from anywhere, irrespective of location where you stay, business see it as having great impact on their operational efficiency [1]. The popularity of wireless networking over the recent years cannot be denied. With the advancement in the field of wireless technology, different types of wireless networks have evolved over the years. One of the most prominent and widely applicable wireless technologies is the Wireless Sensor Network (WSN). WSNs facilitate monitoring and collecting information from hostile physical environment both from local and remote location. As technology advances further and hardware prices drop, wireless sensor network will find its application in areas where traditional networks are inadequate [2]. Although, with increasing computing and communication capabilities, will change the role of the sensors from being mere information dissemination to more demanding task like sensor fusion, classifications, collaborative target tracking, but its role of information dissemination and collection are found to be widely used as it can be deployed even in a hostile environment where human access is impossible [3]. WSN has wide range of application domains starting from large scale global deployment like environmental monitoring to the smart home deployment. The benefits of wireless sensor network are that it has capability of large scale deployment, low maintenance, scalability and adaptability for different scenarios for different applications to mention a few. Moreover, with the evolution of WSNs, multiple applications are supported to run concurrently [4]. This type of system will reduce the deployment and administrative costs, thereby increasing the usability and efficiency of the network.

Generally, a WSN is a collection of few nodes to several hundreds of nodes deployed randomly or deterministically [5] over a region of interest, depending up on the situation and applicability. Most of the WSN applications are data centric, therefore WSNs are deployed to interact with the physical environment and report the phenomenon of interest to the user. The data generated in response to the sensed events in the environment can be accessed by user or applications in several ways, depending upon application types and their requirements. Due to the dynamic (which changes over time) nature of information, which sensor are made to collect, the information being processed and used for making decision at user end does not match with the information at the (physical) source. In addition, the

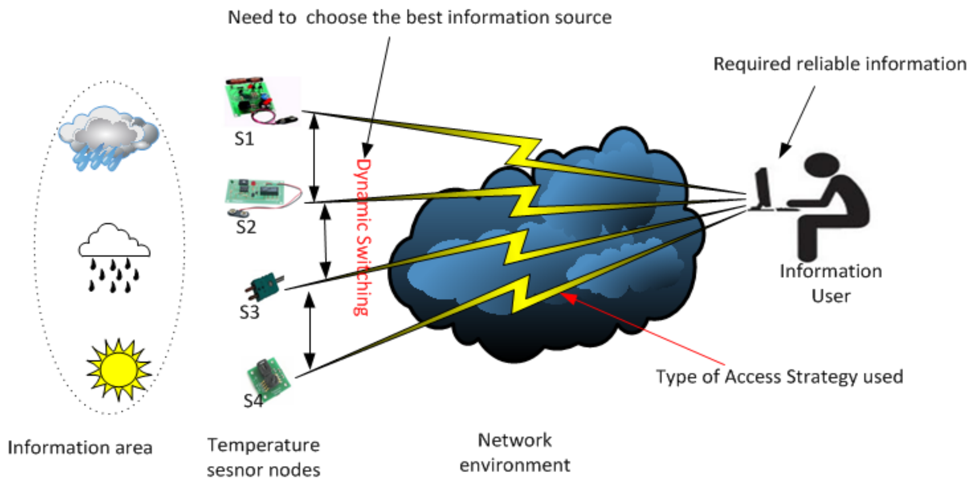


Figure 1.1: A typical scenario of information accessing and management framework.

sensor nodes which act as an information source itself is dynamic, meaning that it becomes unreachable or becomes unavailable due to several reasons like change in location of the user, power failure and change in location of the source (in case the nodes support mobility). Further, there are various challenges in accessing information from information source which are briefly described in the following section.

1.2 Information Access and its Challenges

When accessing information from WSN, there are several issues that need to be taken care. This section highlights some of the challenges and issues in WSN in relation to information accessing, power consumption and information reliability. Figure 1.1 shows a typical information accessing scenario.

Information Reliability: Sensor readings must be sent to the sink as per the transportation methods being used. Reliability depends on several factors, like how fast the information is being generated, how often the information is being sampled, the transmission delay between source and the destination and what types of access methods are being used. As an example, the information reliability would be increased by increasing the frequency of sampling information; however, this would drain out more energy from the nodes. Therefore, there should be an information access mechanism which considers trade-off between information reliability and energy consumption.

Delay: Generally, information collected by WSNs are dynamic in nature that changes over time. This means that information needs to reach at users place where information is being processed and used for making decision at certain minimum lapse of time in order to avoid use of wrong information. For applications like real time control applications, the outdated information would be of not much use. Hence, accessing information mechanism should be flexible to tune as per the applications so that it adapts in line with the application requirement.

Energy Consumption: Although, the energy consumption may not be of much concern if there are energy recovery mechanisms put in place during deployment of WSN, but the lack of battery replacement, which usually are, need energy efficient information access mechanisms. Since improving delay and information reliability will have impact on energy consumption, the access of information should be such that there is no excessive energy utilization to have little or minimal gain in information reliability and minimal delay reduction.

Adaptation: As WSNs can be applied to various application areas, the requirements of different application are different. In addition, a set of requirements of a particular applications also change depending on time and situation. As an example, take a fire monitoring WSN. When there is no fire, the information access can be slower than the one during the occurrence of the fire. Therefore, it is important to have an information access mechanism which adapts as per the time and situation.

Source of Information: Some of the above mentioned parameters are affected by the choice of information source. The transmission delay might depend on how many number of intermediate hops are there between source and the destination, and what delay rates between each intermediate hops. Information reliability will have an impact on which rate a particular information source generates information and how frequent information is being sampled from particular source. Further, if a particular source is selected too frequently, there will be over utilization of its resources, especially the energy, thereby reducing the life time of the node. So, the design of information access mechanism should take these into account so that some nodes are not over utilized whereas others are just lying idle. In addition, if there are several information sources which provide same type of information, it is desirable to switch from one information source to another whenever the user wants to change or when the current source has become unavailable, without interrupting the connection.

Type of Access Strategy: When to access and how often access should be made to an information source also depends on the applications. Some applications need frequent access whereas others do not. The access frequency has impact both on information reliability and power consumption. Different access strategies in accessing information from the WSN are possible. The access strategy should be designed as per the application requirement and should be flexible in order to choose the best access strategy at the given time and the situation.

Therefore, observing the above challenges in relation to accessing the dynamic information from a dynamic information sources, this thesis aims to address the following research objectives mentioned in Section 1.3.

1.3 Thesis Objectives

Observing the challenges mentioned in Section 1.2, with regard to accessing dynamic information from dynamic information sources, this thesis presents a dynamic and reliable information accessing and management framework, which enable application user to dynamically switch from one information source to another as per the situation (may be the information source becomes unavailable) and their requirement (users requirements might change). The framework is used to study the behaviour of the information dynamic in order to design dynamic information source selection algorithm and dynamic access strategy management. In particular, the thesis makes a comprehensive study on two important challenges mentioned above; namely power consumption and information **mismatch probability** [6]: the probability that at the time instant of using certain information for processing in the application, this information does not match the value at the (physical) source. The main objectives of the thesis are summarized as follows:

- To develop a dynamic and reliable information accessing and management framework in order to study the behaviour of information dynamics in relation to sensor networks, with capability of flexible and easy reconfigurable nodes.
- To provide a comprehensive study on the impact of information access strategies on power consumption and information mismatch probability, thereby providing analytical models for information mismatch probability and power consumption.
- To provide an adaptive information access mechanism as per the applications requirement, suggesting trade-off point between information reliability and power consumption.
- To develop a dynamic information source selection algorithm based on dynamic network parameters in order to select the source which gives better performance in terms of information reliability and power consumption.

1.4 Contributions of the Thesis

The contributions of the thesis are summarized as below:

- A dynamic and reliable information accessing and management framework is developed in order to show the concept of dynamic information elements and to study the behaviour of information dynamics.
- A comprehensive study on power consumption and information reliability (in terms of mismatch probability) in relation to different information access strategies is done. Based on this study, extended models for mismatch probability to the one defined in [6] are developed for different access strategy; reactive, periodic and hybrid. The new model captures both information sampling part and information accessing part in computing mismatch probability. The thesis also incorporates power consumption models which were never considered previously with different access strategy.
- In this study, a better way of accessing to dynamic information elements is proposed, where the power consumption is minimized in comparison to periodic and reactive information access strategy.
- Based on analytical models developed, an adaptive information access mechanism is proposed, which is useful for efficient resource management, especially energy consumption of an information source.
- An algorithm, which can dynamically choose information source based on dynamic information elements at source, is developed. This algorithm is developed by using information reliability and power consumption as the selection parameters.

1.5 Structure of the Thesis

The rest of the thesis is structured as follows:

Chapter 2 first gives brief introduction of Wireless Sensor Networks followed by classifications of wireless sensor network applications exploring their requirements and specific characteristic features. Next, it explores the design issues and characteristics features of wireless sensor network. The chapter also briefly introduces on information dissemination and routing protocols and their classifications. Lastly, the chapter brings out some of the

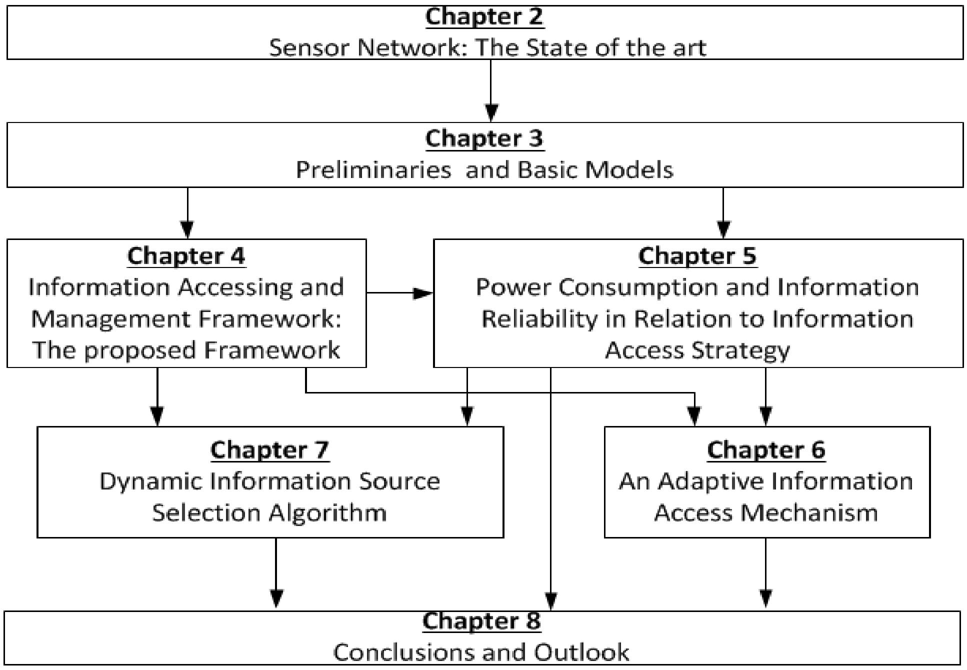


Figure 1.2: Block diagram showing inter-relations among different chapters.

parameters which the wireless sensor network applications users may like to consider during information transport over the network.

Chapter 3 provides an introduction to the basic models which are useful to get fundamentals for the models aim to develop in the later chapters. It gives a basic information access model in general, followed by the process of how events are generated and detected by sensor nodes, which act as an information source in the models considered in this thesis. The chapter also introduces the information reliability models and power consumption models for different types of information access strategies which are extensively used in this study.

Chapter 4 describes information accessing and management framework. The framework is aim at showing the concept of dynamic information source to study the behaviour of information dynamics. The framework is developed using Java based framework called Open Service Gateway initiatives (OSGi) [7].

Chapter 5 gives a comprehensive study on how the applications users requirement parameters are affected by the way how the information is being accessed over the network environment. In particular, the chapter examines three types of information access strategies; reactive, periodic and hybrid access. It explores how the power consumption and information reliability depend on the access strategies being used. This chapter provides analytical models for information reliability (in terms of mismatch probability) and power consumption for each access strategy and makes a comparative study among these three access strategies.

Chapter 6 continues from the previous chapter and explores on adaptive information access mechanism. An adaptive information access mechanism is aimed at efficient utilization of resources like energy.

Chapter 7 proposes an algorithm based on Chapter 5, which chooses information source that gives the best information reliability and minimum power consumption. The algorithm is implemented with the help of the information accessing and management framework developed in Chapter 4.

Chapter 8 concludes the thesis by revisiting the objectives of the study and summarizing the main points. The chapter also highlights the scope of future extension in relation to this study.

Chapter 2

State of the Art: Information Accessing in Wireless Sensor Network

This chapter provides the state of the art on wireless sensor network focusing on information collection and dissemination methods with respect to various applications it supports. In particular, firstly the chapter classifies the wireless sensor network applications into four broad categories; Environmental data collection, Safety monitoring, Mobility support and Hybrid and finds out their characteristics features and requirements. Next, it explores some of the design issues with respect to deployment followed by a short review of some information collections and dissemination methods that are available in the literature. Then the chapter brings out some of the existing frameworks which are developed with similar objectives to study dynamic information elements. Lastly, the chapter lists some important requirement parameters in relation to information accessing from dynamic information source and argues that a reliable and dynamic information accessing and management framework is required to study the behaviour of dynamic information elements with different network conditions.

2.1 Introduction

A wireless sensor network (WSN) is composed of few or many autonomous and integrated/compact device called sensor nodes. These sensors are generally equipped with sensing, communication and data processing or computation capabilities [8]. Recent advancement in the field of wireless communication technology and the availability of integrated low-power sensing devices, embedded processors, wireless communication kits, and power equipment, gives an enabling opportunity for the design of sensor nodes [9]. WSN is a special kind of ad hoc network where many sensor nodes are made to form a network whose sole purpose is to collect information from the region of interest. However, it differs from traditional wireless ad hoc network in many ways [10]. In terms of number of nodes, in WSN network, the number of sensor nodes can be of several magnitudes higher than the nodes in traditional wireless ad hoc network, the sensor nodes are densely deployed and prone to failures due to several reasons; WSN topology is dynamic and changes frequently and the sensor nodes have limited power, computation capabilities and memory. Besides these many shortcomings as compared to traditional wireless network, WSN has found wider application areas in various fields. The increasing usage of WSN in various application [11, 12, 13, 14] starting from wireless body sensor network [15] to highly unpredictable earthquake monitoring and alert system [16], show that few years down the line, its applicability will further grow in various fields. Its usefulness compared to traditional wireless network becomes more visible since it can be used in a hostile environment where human accessible is impossible [3].

These sensor nodes communicate over a short range via a wireless medium and collaborate to accomplish a common task, like environmental monitoring, military surveillance, and industrial process control. Wireless sensor networks have open up for new opportunities to observe and interact with the physical environment around us. They enable us now to collect and gather data that was difficult or impossible before [3]. The wireless sensor network is used for different applications with specific purpose. A different classification of wireless sensor network is described in the following section.

2.2 Sensor Network Applications

Inspired by [17], this section classifies wireless sensor network applications into four broad categories and brings out the different characteristics features in each category. Effort is made to include all the applications supported by wireless sensor network into these four broad classifications.

2.2.1 Environmental Data Collection Application

An environmental data collection application is one where wireless sensors are used to monitor a region of interest for the purpose of study like change in temperature pattern in order to monitor global warming [18]. This type of application of wireless sensor network is mostly used by researchers and scientists who want to collect several sensor readings from a set of points in a region of interest over a period of time and to study the pattern of change and variations over different seasons of the year or time in the day. They would collect data over period of time from multiple points in the region and analyze the data off-line. To be able to collect meaningful data, it has to collect at regular interval and in general the nodes are kept at the same known location. The data collection can be done spreading over a month or some even a year. Some of the characteristics features of this class of application are:

- Long lifetime of the network .
- Precision of measurement.
- Precise synchronization of time with respect to other nodes.
- Relatively static topology and low data rates.
- Data transmission can be delayed to increase or improve efficiency without much loss of performance.

2.2.2 Safety Monitoring Application

In this class of application, sensor nodes can be placed at fixed locations throughout the environment to continually monitor one or more sensor to detect an anomaly. The sensors are not collecting any data instead they send an alarm or message when there is certain safety violations in the surroundings. It is very essential that the sensor nodes are functioning. If the information being sent is over multi-hops then it is very important to check that the intermediate nodes are active and alive. For this reason, the sensor network which is deployed for this type of application must be configured such that the nodes are responsible for confirming the status of each other. An example of this type of application is forest fire detection wireless sensor network [14]. Some of the features of this class of application are:

- Immediate response low latency is critical.
- Reliable communication of alarm or message.

- Bursty nature of messages when safety violations are detected, there will be huge amount of data which will be sent by the sensors.

2.2.3 Mobility Support Application

This group of wireless sensor network is composed of tagged objects which are monitored by wireless sensor networks. Wildlife habitat monitoring wireless sensor networks described in [13] is an example of this class of application. Instead of sensing the environment data, the sensor nodes deployed for this type of applications need to sense the RF messages of the nodes attached to various objects like on the body of animal [17]. Some of the characteristics features of this type of applications are:

- Dynamic topology nodes move through the network.
- Mobile nodes connectivity will be continually changing.
- Flexibility (Scalable and dynamically adapt to changes).
- Tracking of mobile nodes - Reconfiguration of network after and joining of mobile node.

2.2.4 Hybrid Application Networks

In some scenarios, it is possible to have all the above three applications running in a single wireless sensor network architecture, provided the single architecture supports multiple applications. Take a scenario where wireless sensor network is deployed for forest fire detection as a safety monitoring purposes. The alarm message will be generated only when the nodes sense temperature exceeding a threshold value. But this same deployment could be used as an environment data collection to monitor the temperature in that region. So, when the temperature is below the threshold, the WSN can act as an environment data collection application by sending temperature readings periodically or when it is demanded. This type of sensor network architecture which supports multiple applications [19] to run on single architecture has advantages in terms of usefulness of the system.

2.3 Design Issues and Characteristics

In general, the design goals of the wireless sensor network deployment are set as per its purpose and they are application specific. As observed in

Section 2.3, the features and requirements for each application are different. In this section we bring out some of the design issues of sensor network, inspired from [8], and argue that with different characteristics features and requirement of each application, what would be the scenario when a single architecture needs to support hybrid applications.

2.3.1 Dynamicity

The dynamicity in wireless sensor network can be classified into three different levels as *topology dynamics, information dynamics and network conditions*. A sensor network is composed of sensor nodes, sink(s) and events. The dynamics in topology is caused either by mobility of certain nodes as in mobility support application or by certain nodes being removed (battery drained out, and becoming dead). Although, most of the sensor networks are considered to be static but some applications demands dynamic topology where the nodes are required to reconfigure after the topology changes. The information dynamics is due to change in information/events which the sensor network is made to monitor. One basic example of information dynamic is target detection or tracking application where the phenomenon or events change with time. The information dynamic is even relevant to the environment monitoring application. Take an example of a fire. The temperature may rise rapidly, even by a thousand degree Celsius within a minute for indoor housing [12] is not uncommon or even faster for dry, outdoor lands. This change of information value over a period of time couple by network delay in reaching the destination would have impact on the information reliability. The dynamicity is also caused by network conditions. The link conditions may change over time resulting different end-to-end delay.

2.3.2 Node Deployment

Depending upon the nature of application and the situation, sensor nodes can be deployed in two different ways: random and deterministic deployment. In random deployment, the nodes are scattered in the region of interest randomly without any specific location or in some cases with some known distributions like Poison point process [20] and [21]. These sensor nodes are generally self-organizing creating an ad hoc infrastructure. This type of deployment comes handy in getting access to information sources where human access is impossible. In some cases the sensor nodes are air dropped [3]. In this type of deployment, the location of the sink will play crucial role in terms of energy efficiency and performance [8]. There is no pre-defined path which data can be routed in this type of deployment. So, the routing protocol used in this type of infrastructure also plays greater role for performance of the system. In a network with deterministic deployment, each node has a deterministic location and data is routed through pre-determined paths.

2.3.3 Power Consumption

Generally, sensor nodes are equipped with limited power supply. In some cases, such as in [3], where sensors are air-dropped to monitor volcano eruption, replenishment of power supply would be impossible. The lifetime of the node, therefore is depended on the battery lifetime. So, if the information is collected over a multi-hop, then the network lifetime is also determined by the battery lifetime of a single node at times. Hence, effective usage and management of power consumption in sensor network take on additional importance. For this reason, we find most of the researchers are giving much attention in designing energy efficient data collection algorithms and routing protocols. It is also important to know how power is consumed in a sensor network. The power consumption can be divided into three domains: sensing, communication and pre-processing. The power consumption on sensing can vary from application to application and it depends on how sensor nodes are made to sample the data, sporadic or regular interval. The maximum power consumption takes place in data communication. This includes both data transmission and reception. The power consumption in data pre-processing is much lesser than the data communication [22].

2.3.4 Information Access Approach

Depending on the application type and their requirements, information from the information source can be accessed in different ways. Applications like safety monitoring, in general want a system of sensor or sensor to autonomously report the sensed values and raise alerts whereas applications like environment data collection may want the data to be accessed at every after certain interval. But, information can be retrieved in response to a request message from the information user whenever the information is required.

In general, there are three types of communications or access strategies: reactive access, periodic access and event-driven [6]. In reactive access the source node only sends information as and when the user demands. It is also called on-demand access or query-driven access [23]. Periodic access takes place where the source node collects the information and distributes to the user at a constant periodic interval. It is also called clock-driven. This type of access can be destined for certain subscribers or receivers. Event-driven is similar to the periodic access, in a sense that both are proactive, but in this case, the information is not distributed at periodic time interval rather the source sends when some events occur. The data access models being used will influence what type of routing protocols one need to use, especially to minimize the energy consumption and route stability [8].

2.3.5 Node Capabilities

A sensor network can be composed of sensor nodes with different capabilities, although some deployments are found to be using same nodes in terms of functionalities: computation, communication and power. But it is always preferred to deploy with different types of nodes in terms of functionality as all the nodes in a network do not have to have same capability. Such deployment makes the network more robust by supporting different requirements. Depending on the application, node can be dedicated to a particular task. As an example, the node which acts as a cluster head needs to be equipped with more power than the other nodes as it has more functions than the one which deployed as non-cluster head nodes. It depends on what type of protocol it uses; some hierarchical protocols treat cluster heads different from normal sensors, whereas some pick up cluster head among the normal sensors with random or pre-defined selection algorithm.

As seen in the literature, most of the wireless sensor network deployment supports one single application throughout the network. This type of application specific or dedicated network is simple and do not pose much technical issues related to data routing but has been seen not commercially viable where a network is shared several department with individual requirements [24]. In this situation, minimizing investment and administrative cost is vital aspect for the organization. This leads to an introduction of heterogeneous sensor networks which compose of different kinds of sensor nodes for each type of application. This raises multiple technical issues in relation to data routing and information dissemination. This is because different application has different set of requirement which leads to different data rates, which needs to handle with different data access models. Such heterogeneous environment makes data routing more challenging and would require heavier routing algorithm. Although, wireless sensor networks have given new ways to provide information from variety of applications, irrespective of the nature of physical environment, but gathering and collecting information from sensor nodes is seen as a challenge. Data dissemination and gathering are two terms used in sensor networks to describe two categories of data handling methods. Data dissemination is a process by which data and queries for data are routed in the sensor networks whereas data gathering is to transmit data that has been collected by the sensor nodes to the base stations. Data gathering protocols aim to minimize the energy consumption and delay of data gathering process [25].

2.4 Information Collection and Dissemination

Although, wireless sensor networks have given new ways to provide information from variety of applications, irrespective of the nature of physical environment, but gathering and collecting information from sensor nodes is seen as a challenge. Data dissemination and gathering are two terms used in sensor networks to describe two categories of data handling methods. Data dissemination is a process by which data and queries for data are routed in the sensor networks whereas data gathering is to transmit data that has been collected by the sensor nodes to the base stations. Data gathering protocols aim to minimize the energy consumption and delay of data gathering process [25]. Although there are differences between these two but almost all the literature described them as *routing protocols*. Unlike traditional wireless communications networks such as mobile ad hoc and cellular systems, wireless sensor networks have the following unique characteristics and constraints [26]; high density sensor node deployment, battery power sensor nodes, low memory and processor capacity, self-configurable, unreliable sensor nodes, data redundancy, application specific and dynamic topology. Due to above characteristics and constraints of wireless sensor networks, the extraction of data from the network is always a challenge. The main design challenges of routing protocols for wireless sensor network are: Energy, Processing power and Memory. Some of the design challenges as reflected in [26] and [27] are highlighted below but not limited to:

Large number of sensor nodes: Since most of the wireless sensor networks composed of large sensor nodes, it is very difficult to have an efficient addressing scheme like other wireless networks. The traditional IP scheme is not feasible to apply for wireless sensor networks. Moreover, the sensor nodes are deployed at random in hostile environment.

Limited energy capacity: Wireless sensor node, being a microelectronic device, cannot be equipped with unlimited power supply. In some deployment scenarios, replenishment of power might be impossible. So, the life time of a node has strong dependence on the battery lifetime [10]. This is the main challenge in designing wireless sensor networks. In practice, sensor network deployment makes sense only if they can run unattended for months and years without running short of energy [28].

Flow of Data: Almost all the applications of sensor network require the sensory data from multiple sources to flow towards a single destination node called sink in contrast to the traditional networks.

Sensor node locations: Most of the proposed routing protocols assume that the sensors nodes are either equipped receivers with global positioning system (GPS) [26] or use some localization techniques to estimate location [29]. But equipping every sensor node with GPS however, makes sensor node expensive and the objective of making sensor affordable would be defeated [30].

Data redundancy: Data collected by various sensor nodes are typically based on common phenomenon; hence the probability of data redundancy is very high. The routing protocol needs to incorporate data aggregation techniques, where data from different source nodes are combined according to a certain aggregation function, e.g. average, max, or min, to decrease the number of transmission, thereby achieving energy efficiency.

Application Specific: The sensor networks are application specific. The requirement of routing protocol changes as per the specific application. It is very challenging to design routing protocols which can meet the requirements of all applications due to diverse applications.

Scalability: The size of the network grows, so the routing protocols need to be scalable to support the addition of sensor nodes. All sensor nodes may not necessarily have same capabilities of energy, processing, sensing and communications. Therefore, the protocol needs to be scalable to support sensor nodes with different capability features.

2.4.1 Classifications of Routing Protocols in WSN

Routing protocols play an important role in sensor networks because of its unique characteristics. It is different from the traditional network routing since it does not have fixed infrastructure, communicate via wireless link which is unreliable, and energy constrained nodes deployed in hostile environment etc. Many routing algorithms for wireless sensor network have been developed since the inception of wireless sensor networks, all aiming towards challenges posed by its unique features as oppose to the traditional networks. Different authors have classified into different group depending on their usage, functionality and applications [8, 26, 31]. We have classified into four board categories based on the network structure inspired by [17, 32].

Data Centric Routing

The protocols are differentiated into two categories called data-centric and address-centric. The address-centric routing protocols find the shortest path between source and the destination with addressing scheme like IP whereas in data-centric routing protocols focus is made to search routes from multiple source nodes to a single destination node. In the sensor networks, data-centric routing is preferred where data consolidation and aggregation is done by the intermediate nodes on the data coming from multiple sources before sending to the sink node. This way, number of data transmission can be optimized, thereby achieving energy efficiency.

Hierarchical Routing

The hierarchical-based routing tries to build layers so that each layer does some specific job. It mainly aims in clustering the nodes where the cluster head does data aggregation and filtering of redundant data in order to conserve energy. This forms two-layer routing, where one layer is used to select cluster head while other is used for routing [16]. All nodes forward a message for a node that is in higher hierarchy than the sender. The set of nodes which send message to the same aggregator is called cluster, while the aggregator is also referred as cluster head [32]. The node which acts as a cluster head is considered to have more resources than other cluster members as in [33] but however, some protocols like LEACH [34] change their cluster head with some scheduling policies within the cluster members in order to evenly distribute the energy load among the sensors in the network.

QoS- based Routing

The Quality of Service (QoS) required by applications will influence the selection of routing protocols. In terms of requirements, the quality of service can be lifetime of a node, reliable information, energy efficiency, mobility support and so on. These requirements are different for different applications type. For instance, in some critical applications, the information should be delivered within certain period after it is being sensed. Further, applications like safety monitoring would be demanding information to be delivered as soon as it is being sensed. On the other hand, applications like usual environment data collection over the period to study the melting of ice in the Himalayan region [18] can effort to delay but its one of the main quality of service could be accuracy with accurate time stamps so that a useful information is drawn.

Location-based Routing

Since sensor nodes have limited energy capacity, most of the routing protocols aim to reduce the consumption of energy in routing processes. In most of the protocols location of the sensor nodes are used to find the distance between two communicating pairs in order to find the best possible path with low energy usage. If location of a particular sensor node is known, query can be sent to that particular location only without sending to other regions which will reduce the number of transmission significantly [9]. Location-based protocol makes use of the position information to relay data to certain region of the network rather than the whole network.

Table 2.1: Shows trade-off metrics of each routing protocol category

Category of Protocol	**Trade-off Metric**
Data-Centric Routing Protocol	Power dissipation is adapted with the traffic pattern. A trade-off between power consumption and reliability. E.g. a family of data centric protocol like SPIN [35], if the nodes interested in data (destination) are far away from source node, then the intermediate nodes will not relay such data.
Hierarchical Routing Protocol	Efficient energy management, but scalability is restricted as nodes directly communicate to cluster head. An example of hierarchical routing protocol is LEACH [34]
Location-based Routing Protocol	Number of transmission is reduced as the information is sent to particular location only. But usage of GPS makes expensive and the contrition of cheap sensors nodes are not suitable [30]
Quality of Service based Routing Protocol	A trade-off between energy consumption and data quality. But some protocols in this category like SAR [36] suffer from overhead of maintaining tables and states of each node, where there are huge numbers of nodes.

2.5 Related Information Management Frameworks

In the recent past, an increasing number of research works towards context aware information management framework have been found in the research community. This is due to availability of number of devices which are capable of supplying information. Many of these devices are able to communicate and interact with global networks like Internet or any other network environment. In this section, few such information management frameworks are reviewed. Based on this review, this section describes how the study in this thesis is different and which requirement parameters are addressed in the study.

A high level description of a context management framework aiming to make components and applications in Personal Networks (PNs) context aware is provided in [37]. It has also provided with a list of requirements to context management framework. Although, a high level description and number requirements to context management framework are provided but in order really implement, it is necessary to investigate how these requirements can be achieved. However, the article has provided high level description which really motivates in designing context management framework. In particular, this work serves as a basic fundamental in finding out requirement parameters with regards to information accessing from remote dynamic information sources.

Another high level description of a generic context management framework is provided in [38], providing how interactions among context information sources and context aware services, components and applications can be supported. Although, this work has provided an insight to initial design and specifications of a generic context management framework, but it lacks the analytical investigations. Investigations on how information dynamics and network dynamics affect the context values would help in designing good context management framework.

In [39] a mobile device oriented framework for context information management to solve the problems on context storage and communication inefficiency is provided. The framework composed of four main components, data collector, data processor, context manager and local context consumer. The framework uses ontology based context model where each component is defined by using Resource Description Framework (RDF) [40] language. The authors have claimed that the framework deals with real time applications with quick data access, power efficiency with privacy protection. However, the framework is restricted mobile phones and it is not applicable to other devices like PCs. Moreover, detail studies on how information reliability and end to end delay in accessing context information are not explored.

Architecture for distributed context management framework is proposed in [41]. Its aim is to establish an open framework to support development of ubiquitous services which can seamlessly interoperate with different devices. It is based on the standard protocol like Universal Plug and Play [42] so that all kinds of devices could connect to the ubiquitous network and all services can be dynamically deployed by any vendors into the network. However, the communication model which affects the information reliability and power consumption are not investigated in the proposed framework.

Although number of information management framework are seen in the literature all aiming to providing context aware management system but most of the framework proposed needs to be supported by investigating how different parameters like the way how information is being accessed, the nature of information element at the source and the delay in accessing dynamic information do impact the system performance.

2.6 Requirements in Relation to Accessing Dynamic Information Source

Knowing the classifications of wireless sensor network applications and its characteristic features from the design issues to application requirements point of view, this section identifies the requirements of application with respect to information access to dynamic information sources. It is also observed that various routing protocols are developed for dissemination and collection of information from wireless sensor networks. However, most of these routing protocols are typically designed based on a certain set of chosen parameters, depending on the application, in making routing decisions. Due to different features and requirement of each application, a single routing protocol which can be used with any application in general is hard to design. Therefore, in order to make sensor networks more feasible, there should be a system whereby any applications can be built on top of any arbitrary communication protocol stack [43]. To design such an information management framework, it is important to know the information attributes in relation to information access from dynamic information sources. From the short literature review done on existing information management frameworks in Section 2.5, it is observed that there are number of requirement parameters that need to study while designing information access management framework with regard to dynamic information elements. Motivated from the requirement parameters listed in [37], this section brings out the following requirement parameters with respect to accessing dynamic information sources. How these parameters are affected by the dynamic information elements, changing network conditions (different delays) and the way how information is being accessed are investigated in later chapters.

2.6.1 Information Reliability

All applications the best possible information reliability, although the level of information reliability requirement would be different for different application. In the literature [44], reliability is defined as the characteristic of information, in which information has not change its value when it is transported from the source to the destination. This definition of information reliability does not take into account the freshness of information. The delay in getting information coupled by the information dynamic at source, there is a probability that information being used by the information user does not match the current true value at the remote source. Therefore, this thesis investigates information reliability in terms of mismatch of information in order to quantitatively support in designing and developing information accessing and management framework in relation to dynamic information at source. Although, the particular information reliability metric, mismatch probability is relatively new in the research community but off late its usage and study has gain some attention as seen from [26, 45, 46, 47]. Out of these existing works, probabilistic models for access strategies to dynamic information elements are developed in [6]. Three different access strategies namely; reactive approach, periodic and triggered by changes of the information elements are investigated. The study has investigated three performance parameters; network overhead, access delay and mismatch probability. In doing so, a mathematical model for mismatch probability under each access strategy is provided. Based on this work, in this thesis, an extension of mismatch probability models developed in [6] is provided. This thesis also investigates and studies the power consumption model under each information access strategy which is not explored in the existing works.

2.6.2 Timeliness

Timeliness with respect to information accessing from a source can be defined as the time required reach the information from the source to destination. In terms of information source like sensor node it can be defined as the time needed to reach the information at the information user after being generated in response to sensed event from information area. This

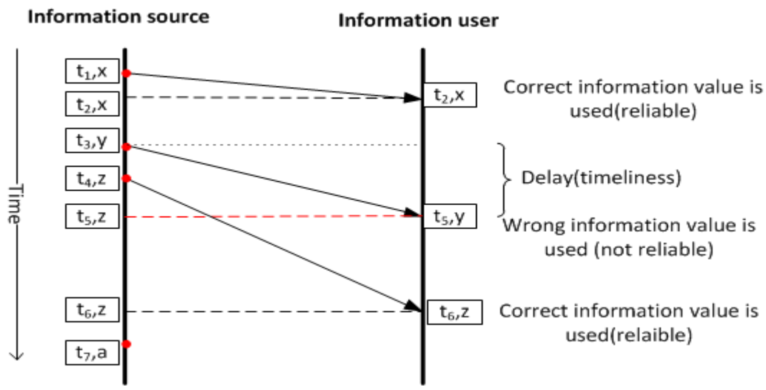


Figure 2.1: Pictorial definition of information reliability and timeliness with respect to dynamic information elements.

basically is same as the end-to-end delay and relates to the freshness of information. In order to reduce the mismatch of information, less end-to-end delay is desirable by all applications or information users. Owing to different network conditions due to various reasons like different routes and different delay rates at different instant of time, the end-to-end delay does not remain constant. Depending on applications, the end-to-end delay will have greater impact for some applications while some can bear the longer delay. Therefore, depending on applications accessing of information can be tuned so that the application satisfies with their requirement. Such adaptive information access mechanism based on network conditions and dynamic information elements are investigated in this thesis.

Figure 2.1, show how these two parameters can be explained in the form of diagram in relation to dynamic information elements. t_i denotes time and the letters denote information value. So, t_1, x means, at time t_1, the value of information is x. The figure shows that the probability of using wrong information not only depends on the end-to-end delay but also depends on the rate at which the event is generated. This is depicted in the figure where although, the delay between t_3 and t_5 is less than the delay between t_4 and t_6 but mismatch of information occurred in the former case. This is because the rate of change of event is faster in case of the former.

2.6.3 Power Consumption

As highlighted in Section 2.3.3, study on power consumption is essential in wireless sensor network as sensor nodes are (usually) powered by battery and are deployed in a hostile terrain for some applications. For some application scenarios, recharging of battery is impossible; therefore, the lifetime of sensor node has strong dependence on battery lifetime. The power consumption in WSN can be divided into three domains; sensing, communication and data processing [9]. Among these three, communication consumes maximum power, which accounts for information transmission and reception. It is therefore, there is strong correlation between power consumption and the way how information is being accessed from the source. As different applications have different level of requirement, depending up on the application, the system can either increase or decrease the power consumption. In particular, in this thesis investigation of how information access can be adapted as per the application requirement is provided in order to support in designing the information accessing and management framework.

2.7 Conclusions

This chapter explored the state of the art in wireless sensor networks with regards to various applications and the information collection and dissemination protocols. Special focus was made on the impact of information dynamics on information reliability and power consumption and to design dynamic and reliable information accessing and management framework to study the behaviour of information dynamics. The chapter also explored some of the existing information management frameworks, from which many useful points in terms of requirement parameters in relation to accessing dynamic information elements were acquired. As a conclusion of this chapter, the observations are revisited here:

- There are various types of wireless sensor network applications with different characteristics and requirements.

- Several number of routing protocols are available. However, each protocol is designed for a specific application.

- No single routing protocol can satisfy all the applications because of different characteristics and requirements.

- There are some information management frameworks which aim to provide interoperability within different applications irrespective of what routing protocols or requirement they have.

- The dynamic information element, which changes over time at source, is seen as a challenge in designing information management framework.

- Lastly, the chapter lists three main requirement parameters in relation to accessing to dynamic information elements; information reliability, timeliness (delay) and power consumption.

Therefore, in order to support in designing information management framework that manages dynamic information elements, it is necessary to study how the dynamic information elements coupled by changing network conditions do impact the above mentioned requirement parameter. So, this thesis aims at designing information accessing and management framework and makes an extensive study on the impact of information dynamics and network conditions on information reliability and power consumption, both are relevant performance metrics in wireless networks.

Chapter 3

Preliminaries and Basic Models

This chapter starts by describing generalized information access model being considered in the study. It also briefly introduces the process of event generation and detection in wireless sensor nodes. This helps in knowing the behaviour of information dynamics with regards to how it is generated, at what rate it is likely to change and so on, which are basic requirements to modeling information reliability. The chapter also briefly describes the power consumption model and different types of information access strategies that are considered in the study. The main objective of this chapter is to introduce the basic models, based on which further models can be developed depending on application and its purpose.

3.1 Information Access Process

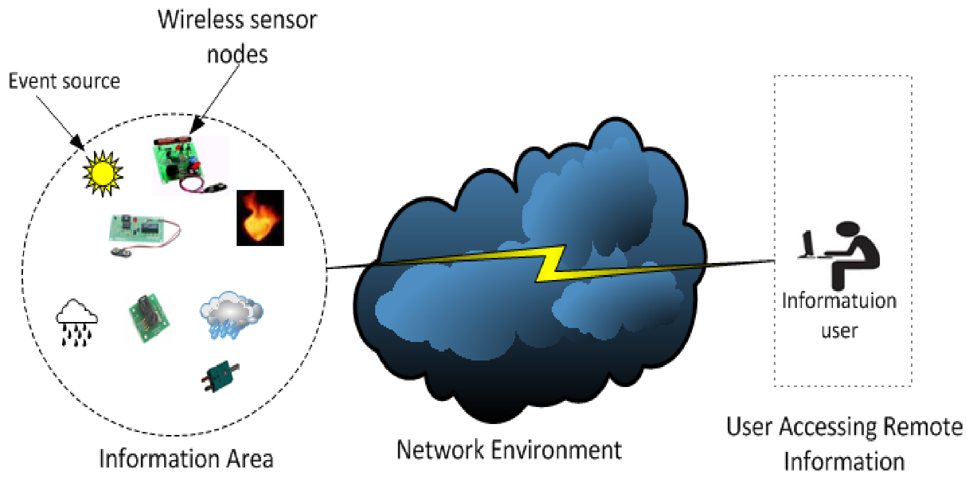


Figure 3.1: Abstract view of information access model.

Figure 3.1 shows an abstract view of how information accessing from wireless sensor network is modeled in order to study the information dynamic in this thesis. The model consists of wireless sensor nodes deployed in a region of interest called information area. The sensor nodes can be of same type or different types depending up on the purpose and applications. Throughout the thesis, these sensor nodes are referred as information sources, which generate information in response to sensed events in an information area. As information dynamics is more applicable to wireless sensor network, the model here describes information sources as wireless sensor nodes, but it should be noted that the model can be applied to any device which provides information. The model assumes that the generated information is accessed by the information user via network environment. This network environment could be any; internet or wireless local area network or wireless wide area network, etc. The only requirement is that the information user should have connection to access remote information sources.

Generally, a typical multi-hop wireless sensor network generates two types of information [5] 1)*locally generated packets* and 2) *relay packets*. The locally generated packet consists of information which is locally sampled by sensors, whereas relay packets are the ones which are received from neighbouring nodes. The local packets mainly depend upon the physical phenomenon of the region of interest and application type, while the relay packets depend on the network parameters. Further, the information generated in response to sensed events can be classified into two categories [48] as *atomic* (information entity is realized through a single message) and *composite* (which considers aggregation or composition of information from different nodes) information. The study in this thesis considers atomic information and mainly focuses on locally generated packets.

Information accessing in wireless sensor networks is the process of sending individual reports from sensors to the sink node. One way of sending individual report is with data aggregation, where each sensor on the way combines (fuses) reports from other sensors with its own reports and forwards a single report, whereas in some cases, the individual reports are just relayed by the intermediate nodes without any data aggregation. However, there is also a middle way, where collecting correlated data with partial aggregation. In this thesis, the model considers the later one; where an individual sensor sends its report and the intermediate nodes only relay the reports without any data aggregation. This means that the nodes or devices on the route in between the information source and information user are not considered for the analysis. The study assumes that the intermediate nodes relay information without altering the value; however the delay due to relaying information is the key performance factor the study will focus on at later chapters. This means, that any routing or retransmissions will be assumed being a part of the resulting end-to-end delay distribution that study focuses on. Further description on how the information sources are modeled can be found in Chapter 4, however, in the next section, a brief study on the event process and event detection is given. This is because the event process plays a crucial role, as it can be seen later, on information reliability.

3.2 Event Process and Event Detection

Definition of an Event: *An event is said to occur when a change in information value exceeds some preset threshold values*

In order to analyze information reliability in the later chapters, it is necessary to investigate how events are detected and processed by the sensor node. However, it is important to realize that this subsection is not meant to be rigorous mathematical analysis, but rather as a motivation for the parameters of the event process that will be used later in Chapter 4. The basic way of obtaining data from sensors is that some electronic component changes the physical attribute of interest into an electrical signal which is lead through some analog filters to the A/D converter of the sensors microprocessor part as shown in Figure 3.2. In this process several sources of noise exist, as well

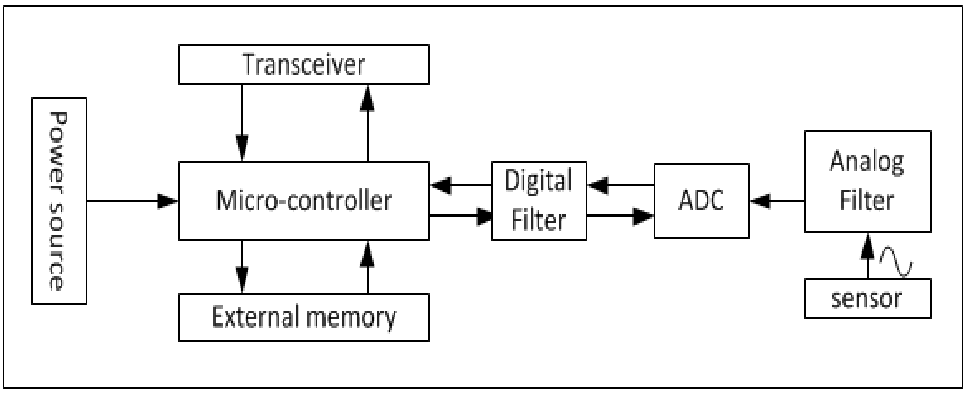


Figure 3.2: Architecture of typical sensor node and how information is generated in response to sensed events.

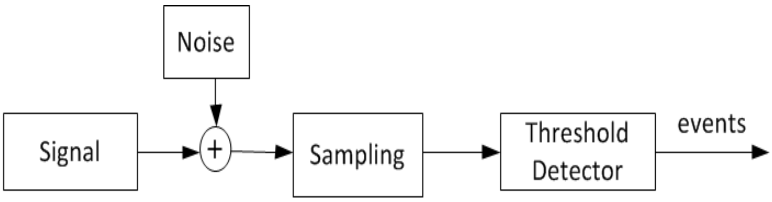


Figure 3.3: Process of an event detection.

as in the A/D conversion procedure. Therefore, often signals are filtered digitally afterwards; to cut-off undesired higher frequencies of the signal caused by the added noise (which usually can be considered white noise). The nature of the process being sampled is very different from among different physical attributes, e.g. temperature or lighting behaves very differently. However, since the focus of the study is in the rate of change of a discretized signal, consider a sinusoidal function with some white noise added to it, in order to evaluate when an exponential distributed inter event rate a good assumption, that is used in later chapters of this study. Consider a sum of number of sinusoidal signals as

$$f(t) = \sum_{i=1}^{N} A_i sin(\omega_i t + \phi_i) + N(0, \sigma^2) \tag{3.1}$$

with A_i as the amplitude, ω_i as the frequency, ϕ_i, a random phase indicating a random moment of the process which the sampling procedure starts, σ^2, the variance of the white noise being added due to the various noise sources and N is the number of sinusoidal signals used. A graphical representation of the process of an event detection is shown in Figure 3.3.

In order to get different signal scenarios, using different frequency and amplitude components, a different signal is created from the base sinusoidal wave and the variance of the white noise of σ being added. For example, consider an area under temperature surveillance by which the temperature can be expected to deviate by 10°C over time from some average temperature. Weather change may be one reason for such change, and may happen over tens of minutes. Such temperature deviation scenarios are created in this section where the temperature goes between +10°C and -10°C over a certain time interval in in a different shapes, and sampled different time intervals. Some noise is added in the sampling process and then digitally filtered with a 5^{th} order Butterworth filter with a cut-off frequency of 2.5 Hz. These different scenarios give a better picture to realize real signal rather than assuming a simple sinusoidal signal, thereby closely representing the real application scenarios where different applications generate events with different distributions. Therefore, the study is interested in the following: In any given time of the signal, take a sample, and then investigate how long it takes for this value to change significantly. By a significant change, it means that an event has occurred. Then basically, it is to solve for each t for dt in the following equation:

$$|f(t) - f(t + dt)| > threshold \tag{3.2}$$

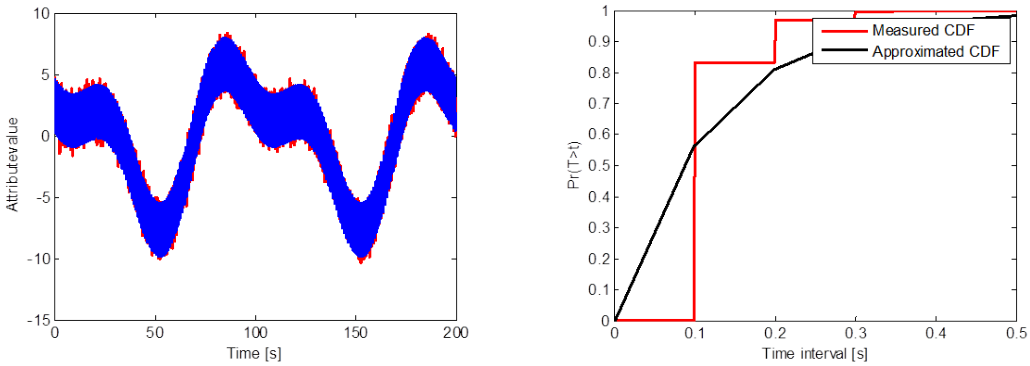
If f(t) is the noisy signal, then the resulting cumulative density function of the *dt's* can be obtained by using particular threshold values with a certain noise level σ. The resulting cumulative density function of the *dt's* can be influenced by the different parameters like threshold values being set, sampling and cut-off frequencies being chosen and the level of noise in the signal.

From (3.1) and (3.2) several parameters can be identified that influence the output CDF which are needed at later stage of the study. Those are: the amplitude (A), the frequency (ω), the noise level (σ^2) and the threshold level. This section provides some simple analysis of modeling process of event detection and find out at what parameters influences event inter arrival time.

In particular, the section explores the following three parameters : 1) Sampling and cut-off frequencies with respect to signal frequency 2) Noise level and 3) Threshold values.

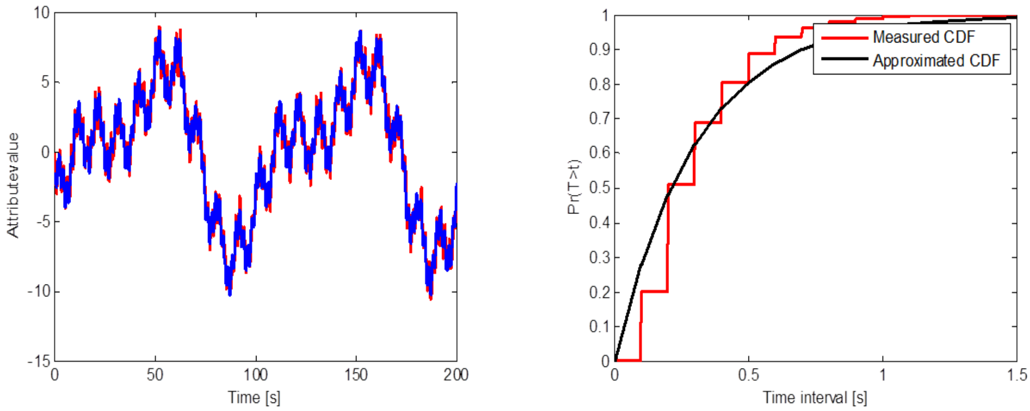
1) *Sampling and cut-off frequency*

In order to show how different frequencies value chosen do affect the distribution of inter event arrival rate, a different frequency and amplitude components are used, thereby getting a signal which is a sum of number of sinusoidal signals. The approximated CDF with fitted exponential distribution is compared with the measured (constructed) CDF, thereby showing the



(a) (b)

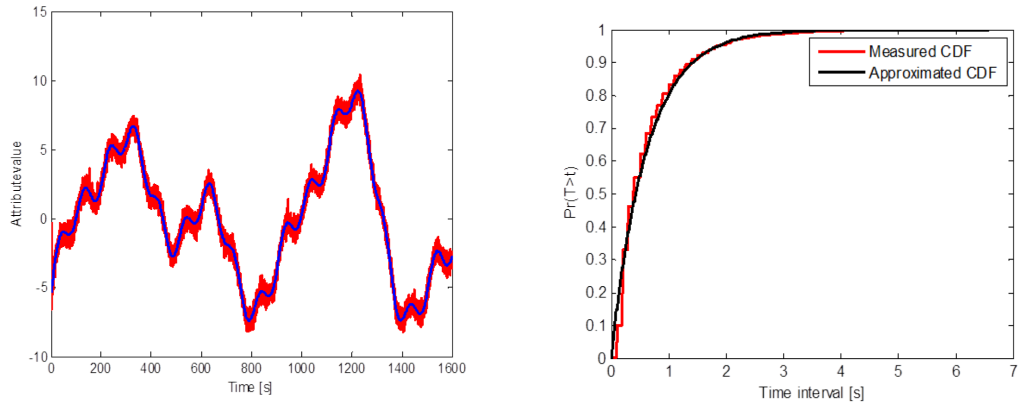
Figure 3.4: (a) True signal and filtered signal with noise ($\sigma = 0.5°C$), with frequencies of [1/100 1/50 1 3] Hz and corresponding amplitude components of [5 3 2 1], using a threshold of 1 and sampling frequency = 10 Hz (filter cut-off frequency = 2.5 Hz).(b) Corresponding CDF of events based on the filtered signal and the approximated CDF using an exponential distribution.



(a) (b)

Figure 3.5: (a) True signal and filtered signal with noise ($\sigma = 0.5°C$),(b) CDF of events based on the filtered signal and the approximated CDF using an exponential distribution. A frequencies of [1/100 1/50 1/10 1/2] Hz and corresponding amplitude components of [5 3 2 1], using a threshold of 1 and sampling frequency = 10 Hz (filter cut-off frequency = 2.5 Hz).

deviation from exponential distribution.



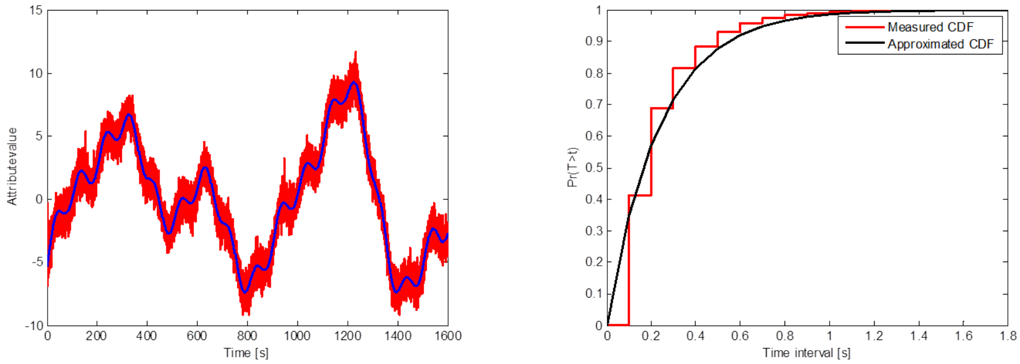
(a) (b)

Figure 3.6: (a) True signal and filtered signal with noise ($\sigma = 0.5°C$), with frequencies of [1/800 1/500 1/300 1/100] Hz and corresponding amplitude components of [5 3 2 1], using a threshold of 1 and sampling frequency = 10 Hz (filter cut-off frequency = 2.5 Hz). (b) Corresponding CDF of events based on the filtered signal and the approximated CDF using an exponential distribution.

In Figure 3.4(a), the cutt-off frequency which is greater greater than the highest frequency component of the signal is set. The comparison between the measured CDF and approximated exponential CDF is shown in Figure 3.4(b). In Figure 3.5(a), the frequency of the signal component is reduced so that the cut-off frequency is greater than the highest frequency component of the signal and the corresponding CDFs comparison is shown in Figure 3.5(b). The frequency of the signal component is further reduced so as to see how the the CDF of the inter event arrival time is affected as shown in Figures 3.6(a) and 3.6(b). Observing the CDF plots in Figures 3.4(b), 3.5(b) and 3.6(b), it is clear that the sampling frequency and the cut-off frequency have to carefully be selected for an inter event arrival time to behave as an exponential distribution. Further, assume memoryless property, although it may not hold completely due to representativeness of the signal. It is therefore, important to investigate on the limits for the approximation for an exponential distribution, at which parameter setting, does the inter event arrival time can be assumed as exponentially distributed.

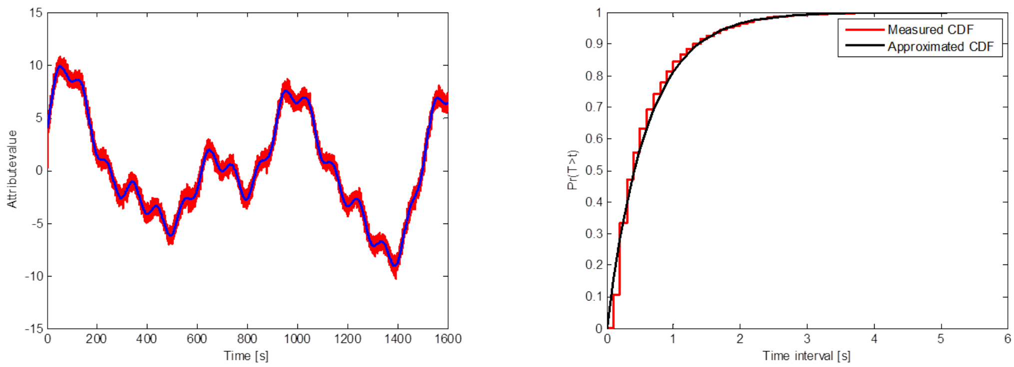
2) *Noise level of the signal*

Assuming that the signal is sampled with desired frequency values,what is the impact of noise level in the signal? In order to study this, various val-



(a) (b)

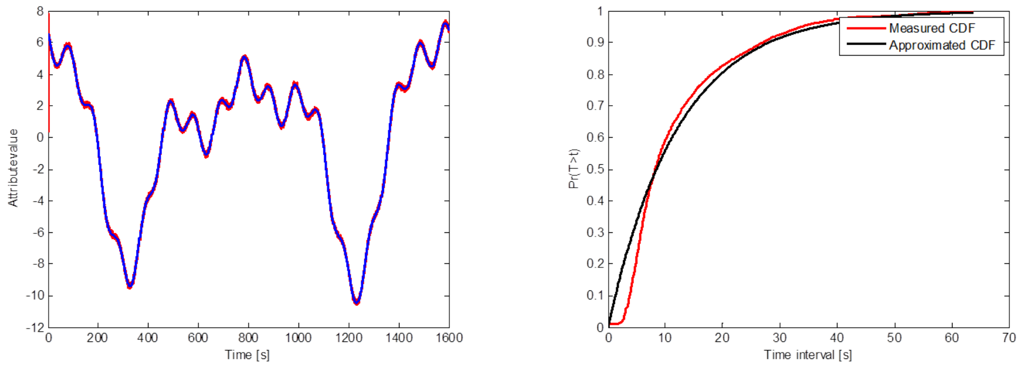
Figure 3.7: (a) True signal and filtered signal with noise (σ = 1.0°C), with frequencies of [1/800 1/500 1/300 1/100] Hz and corresponding amplitude components of [5 3 2 1], using a threshold of 1 and sampling frequency = 10 Hz (filter cut-off frequency = 2.5 Hz).(b) Corresponding CDF of events based on the filtered signal and the approximated CDF using an exponential distribution.



(a) (b)

Figure 3.8: (a) True signal and filtered signal with noise (σ = 0.5°C), with frequencies of [1/800 1/500 1/300 1/100] Hz and corresponding amplitude components of [5 3 2 1], using a threshold of 1 and sampling frequency = 10 Hz (filter cut-off frequency = 2.5 Hz).(b) Corresponding CDF of events based on the filtered signal and the approximated CDF using an exponential distribution.

ues of σ are used and analyzed how the measured CDF of inter event arrival time deviates from the approximated CDF to exponential distribution. It is observed that the exponential assumption is appropriate with low frequency signal with low noise level. The observation is shown in Figures 3.7(b),3.8(b) and 3.9(b).



(a) (b)

Figure 3.9: (a) True signal and filtered signal with noise (σ = 0.1°C), with frequencies of [1/800 1/500 1/300 1/100] Hz and corresponding amplitude components of [5 3 2 1], using a threshold of 1 and sampling frequency = 10 Hz (filter cut-off frequency = 2.5 Hz).(b) Corresponding CDF of events based on the filtered signal and the approximated CDF using an exponential distribution.

3) *Threshold*

Assuming that the signal has permissible noise level and sampled with desired sampling frequency (with correct cut-off frequency), analysis is done here to show how the threshold value affects the distribution of the inter event arrival time. The results are plotted in Figures 3.10(a), 3.10(b), 3.10(c) and 3.10(d). It is observed that a threshold value of 1 and above but below 3 results measured CDF as comparable with the approximated CDF of exponential distribution, assuming the rest of the parameters are chosen appropriately.

It can also be shown using probability plot that the distribution of inter event arrival rate more or less follows exponential distribution, provided the parameters mentioned above (frequency, threshold and noise level) are chosen appropriately. Figure 3.11 shows probability plot showing that the measured CDF of event is not far from the approximated CDF considering exponential distribution.

To summarize, in order to assume inter event arrival time as expo-

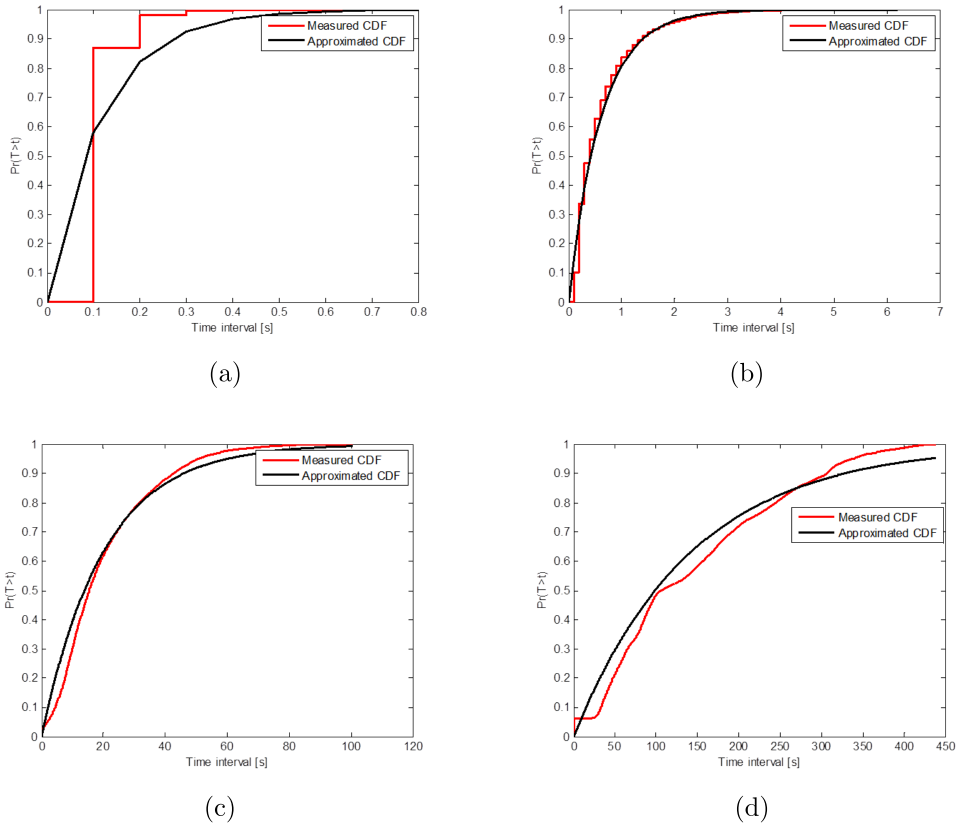


Figure 3.10: CDF of events based on the filtered signal and the approximated CDF using an exponential distribution. Frequencies of [1/800 1/500 1/300 1/100] Hz and corresponding amplitude components of [5 3 2 1], using a signal with noise ($\sigma = 0.5°C$) and sampling frequency = 10 Hz (filter cut-off frequency = 2.5 Hz). a) Threshold = 0.1, b) Threshold = 1.0 and c) Threshold = 3.0 and d) Threshold = 10.0

nential distribution, there is a need to appropriately choose the parameters which influences the cumulative density function of inter event arrival rate. The following points are observed from the analysis carried out above in order to assume exponential distributed inter event arrival rate:

- Low frequency and/or low amplitudes types of signal, with appropriate noise (say, $\sigma = 0.5$) component
- An appropriate threshold value (Threshold value of above 1 and below 3, observed to be appropriate as shown in above analysis)
- Proper digital filtering, i.e. cut-off frequencies do not impact the event rate (or at least is fixed and same for all subsequent evaluations).

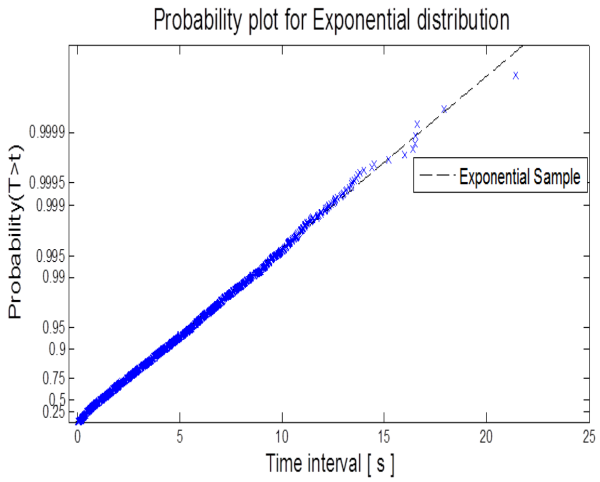


Figure 3.11: Probability plot for exponential distribution, with frequencies of [1/800 1/500 1/300 1/100] Hz and corresponding amplitude components of [5 3 2 1], using a signal with noise ($\sigma = 0.5°C$) and threshold of 1.5 and sampling frequency = 10 Hz (filter cut-off frequency = 2.5 Hz).

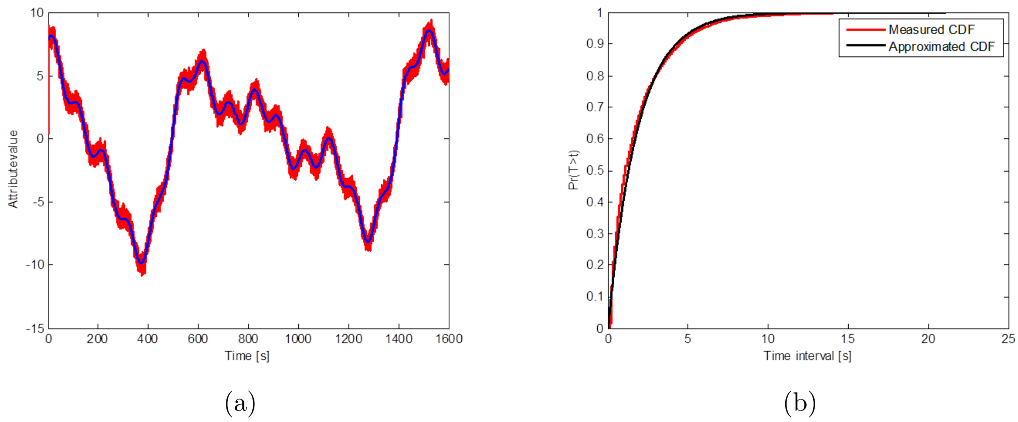
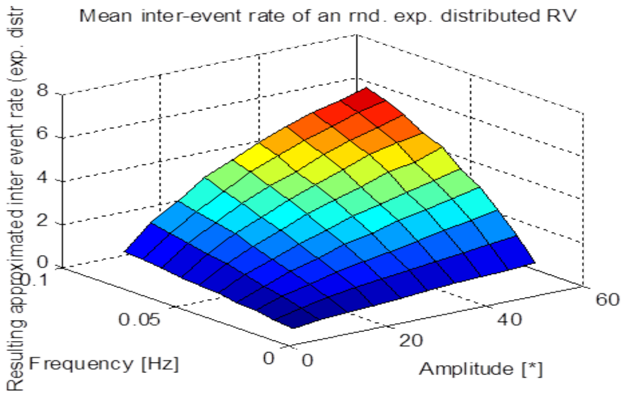
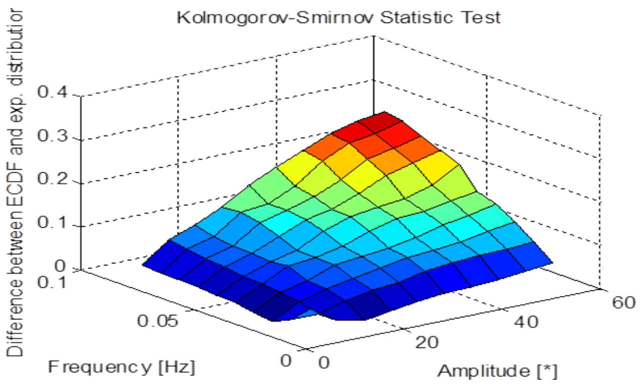


Figure 3.12: (a) True signal and filtered signal with noise ($\sigma = 0.5°C$), with frequencies of [1/800 1/500 1/300 1/100] and corresponding amplitude components of [5 3 2 1], with threshold of 1.5.(b) Corresponding CDF of events based on the filtered signal and the approximated CDF using an exponential distribution.

After choosing appropriate parameters with the help of analysis done above, in the following, a further analysis is carried out by using Two- sample Kolmogorov-Smirnov Statistic to find out the difference between measured CDF and approximated CDF (exp.) in order to show the deviation from the assumed distribution. The KolmogorovSmirnov to test whether two under-



(a)



(b)

Figure 3.13: Mean values of approximated exponential distribution of inter event times for frequencies [1:2:20] Hz of signal frequency, and amplitudes [1:2:20] of signal amplitude, using a threshold of 1.5 and sampling frequency of 10 Hz (and filter cut-off frequency at 2.5Hz).

lying one-dimensional probability distributions difference is given by

$$D_{n,n'} = \sup_x |F_{1,n}(x) - F_{2,n'}(x)|,$$

where $F_{1,n}$ and $F_{2,n'}$ are the empirical distribution functions of the first and the second sample respectively.

If f(t) is the noisy signal in Figure 3.12(a), the Figure 3.12(b) contains the resulting cumulative density function of the *dt's* found for particular example with parameters reflected in the caption. One can approximate the distributions of *dt's* with the smooth curve which in this case has a mean inter event time of 2.4 sec (or the rate of 0.4 events/sec). The noise level is fixed to 0.5 and the threshold is set at 1.5. The signal is sampled at 10

Hz and the data is run through a digital 5_{th} order Butterworth filter with a cut-off frequency of 2.5Hz.

Figure 3.13(a), shows the approximated rate value for an exponential distribution for different frequencies and amplitudes (i.e. demonstrating different types of information dynamics) for the signal shown in Figure 3.12(a). Figure 3.13(b) shows the Kolmogorov-Smirnov Statistic Test plot, showing the difference between measured (constructed) CDF and approximated CDF with exponential distribution. The plot reveals that the difference is less significant at low frequencies and low amplitude area, which are also the scenarios with low event rate distributions.

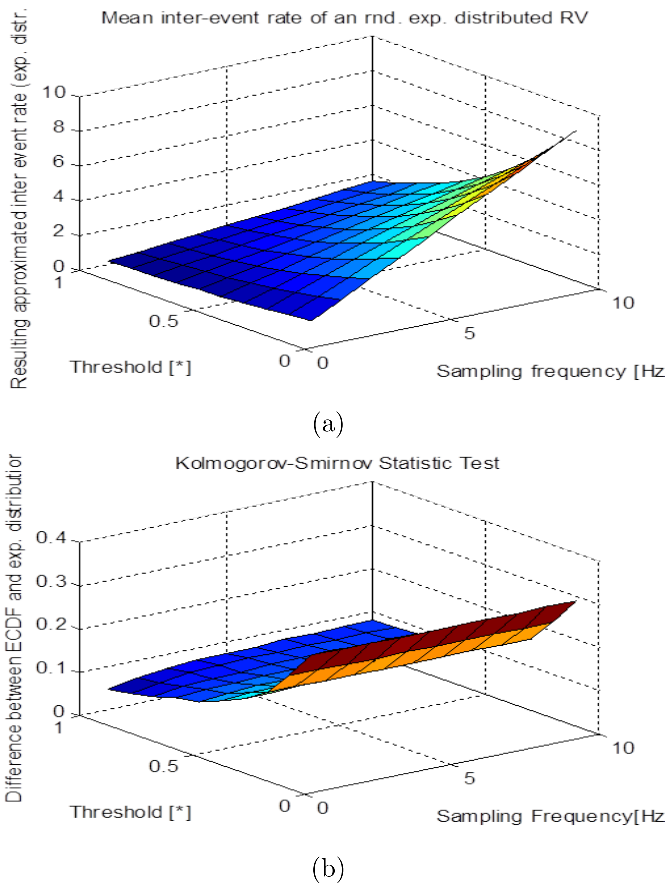


Figure 3.14: Mean values of approximated exponential distribution of inter event times for threshold level [0.1:0.1:1] degree, and sampling frequency [1:1:10] Hz of a signal of amplitude [5 3 2 1] and frequency [1/800 1/500 1/300 1/100] Hz (16 minute) and $\sigma = 0.5°C$.

Figure 3.14(a) and Figure 3.14(b) show the relation of the signal shown in Figure 3.12(a) with varying threshold level and sampling frequency (with a cut-off frequency of the digital filter of 1/4th of the sampling frequency).

These plots reveal that although the sampling frequency does have an impact on the event rate, but then if the threshold level is chosen appropriately (from the analysis done before), e.g. at 1.5, the impact is relatively low. Therefore, choosing an appropriate threshold value is emphasized so that the approximations error is kept low.

As a conclusion on this analysis, in later chapters, the study focuses on event processes which have the following properties:

- Low frequency and/or low amplitudes types of signal, with relatively low noise component
- A threshold value chosen appropriately using the described threshold detection mechanisms for indicating events
- Proper digital filtering, i.e. cut-off frequencies do not impact the event rate (or at least is fixed and same for all subsequent evaluations).
- No (or insignificant) influence of sampling frequency to the approximated event rate
- The resulting discrete event process is assumed Poison process.

3.3 Information Reliability Model

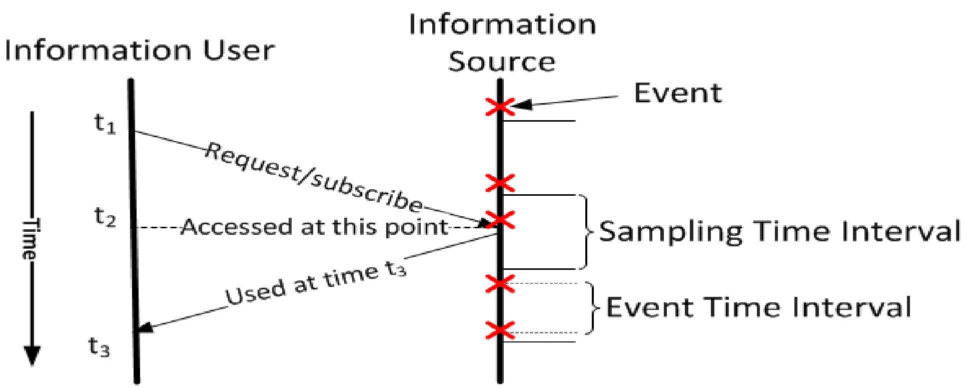


Figure 3.15: Message chart diagram of information access process.

The information accuracy or reliability is measured in terms of different parameters in different ways. In [49] and [50], the information accuracy is measured in terms of mean square error which calculates the error between the expected value at the receiver and the real value at the source. The information which is generated in response to the sensed events at source has to be transported to the destination, where it is processed or used. Most of the existing routing protocols assume that the information coming from the source is reliable and trustworthy [44]. But, the routing protocols are not able to estimate what is the probability that the received information value is the same as the value at the source at the time when the information is being used. Therefore, in this thesis, a parameter called mismatch probability defined from [6], which is an event based reliability measurement metric is used. In the context of this study, the *mismatch probability* is defined as the probability that information value used and processed by the client at time t_3 is not same as the one which is available at (physical) source at time t_3, (see Figure 3.15). This reliability parameter is based on the freshness of information rather than the difference between the information value sent from the source and the information value received at the destination. There can be several parameters that needs to be estimated and investigated in relation to information reliability but only an event rate (λ) is considered here. Although, estimating more parameters may lead to more accurate in some sense but requires potentially much more parameters to be estimated, thereby making the model very complex to implement. Therefore, dealing with single parameter leads to the low model complexity, however requires some appropriate signal tuning prior to the usage. At the later stages in the study, it shows that things are more simplified considering exponential distribution for event rate. This also simplifies computations that may need to be executed in near real time. Therefore, whenever nothing is specified, the study assumes all exponential distributions and focuses on the following parameters in extending reliability models described in [6] in later chapters.

- λ, event rate assumed as exponentially distributed
- T, a sampling time interval
- τ, a sampling rate, assuming exponential distributed time intervals
- ν, stochastic network delay rate

3.4 Information Access Strategies

Communications between information source and the information user can occur in different ways depending on the quality of service requirements of the application or the purpose of the network as a whole. In general, there are three types of communications or access strategies: reactive access, periodic access and event-driven [6].

1. *Reactive Access Strategy*: In reactive access the source device only sends information as and when the user demands. It is also called *on-demand* access or query-driven access [23].

2. *Periodic Access Strategy*: Periodic access takes place where the information source collects the information and sends or distributes to the user at a constant periodic interval. It is also called *clock-driven*. This type of access can be destined for certain subscribers or receivers.

3. *Event Driven Access*: Event-driven is similar to the periodic access strategy, in a sense that both are proactive, but in this, the information is not distributed at periodic time interval rather the source sends when some events occur. However, this study explores an access method which is not purely an event driven. The information is retrieved only if the information value has changed from the previously retrieved value. This is called as *Hybrid Access Strategy* in this study. For checking whether the information has change in value, a small constant delay is introduced in the analysis at later stage in the study. This is a new model of access strategy discussed in this study which was not considered in [6].

Referring to Figure 3.15, information user can use any of these access strategies to access information source. An investigation of how the way information is being accessed does impact the information mismatch probability is done at later chapters in this study.

3.5 End-to-End Delay Distribution

The information reliability in terms of mismatch probability can be affected by the time the information takes to reach the destination. Therefore, in order to analyze an impact of information access strategy on information reliability, it is necessary to study how the end-to-end delay affects information reliability. For simplicity, the end-to-end delay distribution is assumed to be exponentially distributed in our analysis later. This is reasonable as the objective of this study is not to develop the delay models and do the deeper analysis, rather to show the proof-of-concept on how dynamic information elements do impact information reliability and power consumption. Nevertheless, compares the assumed exponential delay distributions with more complex delay distribution model in order to show how much deviation does exponential distribution has. Consider the network as shown in Figure 3.16. The source node(S) and the destination node(D) has several intermediate nodes. If the route chosen is S-2-6-D, then for such scenario, inspired from [5], the delay can be modeled by discrete time Markov chain as shown in Figure 3.17. Assuming, the delay rate from each node is exponentially distributed [51, 52] the overall mean delay is computed using matrix exponential distribution as described. Considering same delay rate($a_1 = a_2 = a_3 = \nu$) from each node and retransmission delay rate of ν' as shown in Figure 3.17. The progress rate matrix as per Figure 3.17 given by β, starting vector by P

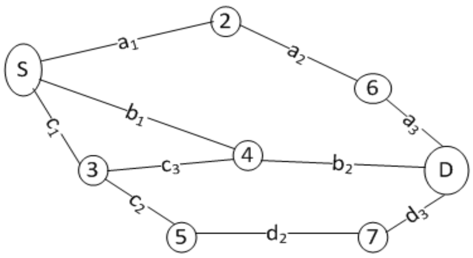


Figure 3.16: Considered Network scenario.

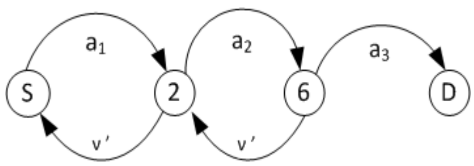


Figure 3.17: Markov Model, Transition diagram.

and vector e' to denote transpose.

$$\beta = \begin{pmatrix} a_1 & -a_1 & 0 \\ \nu' & a_2 & -a_2 \\ 0 & \nu' & a_3 \end{pmatrix}, \qquad P = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}, \qquad e' = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

The cumulative distribution function of delay rates can be computed as

$$F(t) = 1 - Pexp(-\beta t)e' \tag{3.3}$$

The delay distribution in case of without retransmission can be modeled by setting $\nu' = 0$. As per Figure 3.17 the retransmission delay rate ν', with probability of retransmission p can be computed as

$$\nu' = (\nu - \nu(1-p))/(1-p)$$

There are possibilities of developing complex end-to-end delay models using complex Markov Chain Model as provided in [5]. In [5], end-to-end delay distribution is modeled by considering entire network as a queueing network and the communication system at each node modeled by a discrete time Markov chain. However, in this study, such complex delay distributions are not considered, so as to keep the system more simple and to stick to the objective of the study.

3.6 Power Consumption Model

Inspired from [53], in wireless sensor network applications, the power used by a node consists of the power consumed by receiving, transmitting, listening for events on the radio channel, sampling data and sleeping. Therefore, the total power consumption can be modeled as

$$P_t = P_{samp} + P_{tx} + P_{listen} + P_{sleep} + P_{recv} \tag{3.4}$$

where, P_{samp}, P_{tx} are the amount of power dissipated for sampling and transmitting respectively, P_{listen} for listening to radio channel, P_{sleep} for sleep mode and P_{recv} for receiving. This power consumption model is used in later chapters for investigating the impact of information access strategies on power consumption and information reliability. P_{samp} and P_{tx} are two parameters which this study explores in relation to information accessing from a remote dynamic information source. In relation to Figure 3.15, P_{samp} is power expend when an information element is sampled. This happens at a rate given by the sampling rate τ, which is stochastic and assumed to be exponentially distributed. The information is sample at an average rate of τ and stored in cache. P_{tx} is the power expend on sending/accessing information from cache to information user. This is affected by the end-to-end delay between information source and the information user. This delay will be given by $t_3 - t_2$ as per the Figure 3.15.

At later stages in this study, the impacts of these access strategies on power consumption and information reliability are explored both analytically and with the help of simulation.

3.7 Conclusions

The preliminaries models which are fundamental and basic requirement for the further investigations of information dynamics are described here. In summary,

- the chapter described the main information access process being considered
- a brief analysis of the event process and event detection were also provided in order to understand the behaviour of dynamic information elements.
- different strategies in accessing information from information source were introduced.

- an introduction to end-to-end delay distribution is provided. This introduction provides an insight on how delay distributions can be modeled as per the level of complexity and scope of the study.
- a general power consumption model, which is used for further development of power consumption models under different information access strategy was described.

As mentioned in Section 3.2, the chapter is not meant for rigorous mathematical analysis on event process and detection, but rather a motivation for the following chapters. Therefore, there are possibilities to explore further analysis with rigorous mathematical analysis. With the simple analysis carried out on the event process and detection, in later chapters the study assumes exponential distributed event inter arrival rate with low frequency and low amplitude signals with appropriate noise component. The chapter also provided some comparative study by considering different distributions for inter arrival time and investigating how far is the exponential distribution considered in this study from other distributions.

Chapter 4

Proposed Framework

This chapter describes the proposed framework architecture for dynamic and reliable information accessing and management in heterogeneous wireless network. The chapter consists of several components which correspond to particular function and purpose. The key components are: information source model - which deals with event generation and sampling of events, event monitoring model - which monitors information sources and network parameters, and access management unit - which manages how information is being accessed from the source. To evaluate the performance of the framework, a computation delay with respect to number of components running within the framework is provided.

4.1 Motivation

Generally, the information which wireless sensors are made to collect is dynamic and it changes over time. This dynamic behaviour of information leads to using wrong value of information at the user end. In order to study this information dynamics, it is necessary to have a framework which monitors these dynamic parameters like event rate, rate of change of delay etc. Although the existing simulation packages like Network Simulator-2, Matlab etc. are widely used and largely accepted, however, it is found that it is not trivial to model which can monitor these information dynamics in order to investigate their impacts; particularly, on information mismatch probability. But designing robust and dynamic simulation package is not an objective of this study; rather its aim is to facilitate in investigating the information dynamics with different network conditions. In particular, the framework is being developed in order to accomplishing following tasks, which are described as real life scenarios.

Scenario A

As technology in wireless networks advances, there is growing number of service providers. For example, Personal Area Network (PAN), Wireless Hotspots, Wireless Local Area Network (WLAN), etc. Look at below scenarios:

- The security department wants to set up wireless sensor network based intrusion detection system using surveillance camera in the corporate premises.
- The maintenance department wants to set up wireless sensor network to monitor the temperature in the premises in order to prevent fire.
- The top level management wants to monitor the air quality (pollution level) in the premises.

For the above three applications, the organization requires three different wireless sensor networks to provide three different services. The total investment cost will be reduced if all the information sources can be accessed using the same system.

Scenario B

Contrary to *Scenario A*, sometimes, we face a situation where in the middle of accessing particular information from a particular source, the source of information either gets disconnected or not being reachable. This may be caused by several reasons like network signal might have problem, or the user is moved to a region where source cannot be reached, or it can be other way round, that the source might have moved, if the information source is mobile. In such situation, if there is another information source which can provide the same information, it is desirable to dynamically switch to the later without disrupting the continuity or getting disconnected.

Therefore, the reason behind developing this framework is to show the concept of dynamic switching from one information source to another based on the information dynamics and delay. To test the working of the framework, the information mismatch probability in accessing remote dynamic information is estimated using this framework. Further to validate the results, these estimated mismatch probability values are verified using the analytical models in Chapter 5. In addition, the framework is also used to implement reliable information source selection and management algorithm based on information mismatch probability and power consumption in Chapter 7. The detailed description of the framework architecture is given in the following section.

4.2 Proposed Framework Architecture

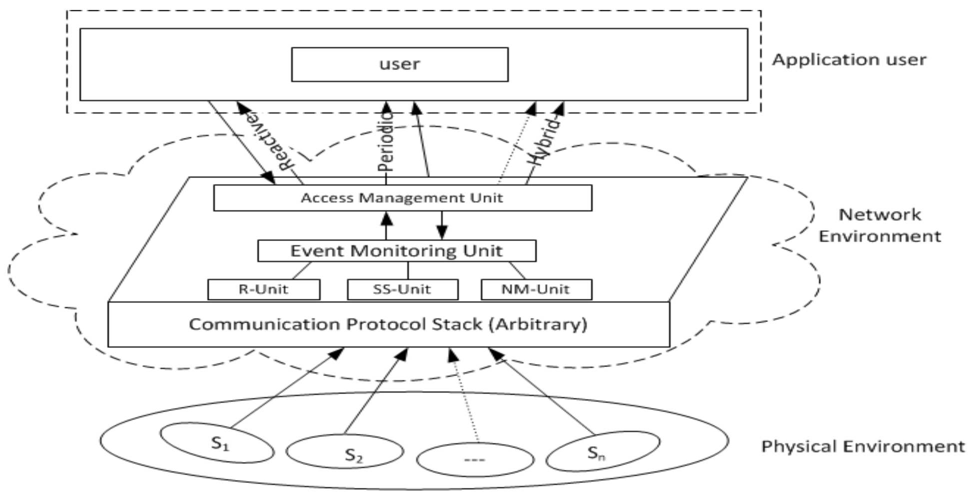


Figure 4.1: Proposed Framework Architecture ($S_1, S_2...S_n$ are information sources).

Before describing the framework, it is necessary to study what features does the framework should have. The main objective in developing this framework is to study the dynamic behaviour of information elements by way of managing the information sources and monitoring the network conditions. Based on this objective, the framework need to support the following features:

- *Modularity*: In general, modularity is a desirable system property. An individual component of the framework should be examined, modified and maintained independently. The framework should able to control and monitor information source individually.

- *Dynamic Updates*: As the framework is aim to study the dynamic behaviour of information elements, the framework should able to manage information sources dynamically at run time. It should able to remotely start, stop or install any updates dynamically at run time.

- *Generic*: As the proposed framework is aim to develop on top of an arbitrary communication protocol as shown in Figure 4.1, it should be generic; meaning that it should support any application irrespective of what communication protocol it uses.

- *Facilitate Interactions*: The framework should be able to support interactions between information sources and information user. This is an important requirement of the proposed framework. Multiple information sources providing same information are possible. In such situation, interactions between information sources and information user. This property of the proposed framework is used for selecting the best from multiple information sources in Chapter 7.

Figure 4.1 illustrates the proposed conceptual framework architecture for information accessing and management framework. However, it is to be noted that it is a proof-of-concept and emulation tool for algorithm later proposed is in its place. Closely observing the features that the proposed framework should support, a Java based framework called Open Services Gateway initiatives (OSGi) [7]. OSGi works on the concept of bundles, meaning that each and every software module describing different components in the framework is developed as OSGi bundle. Each bundle in OSGi can be controlled individually without affecting the rest of the system, thereby supporting modular application development. These bundles communicate through the services provided by the OSGi framework at run time. The installation, starting, stooping, updating and reinstalling of OSGi bundles can be done dynamically at run time. OSGi is an open service platform for the management of multiple applications and services of all types like networked devices in buildings, at home, in vehicles, in industry, mobile telephony and other environments. However, the standard OSGi framework does not support communication among distributed bundles if the bundles are located in different machines [54]. R-OSGi (Remote-OSGi) is a middleware platform which implements the OSGi standard in distributed environment. Therefore, the R-OSGi is used facilitate the communication among OSGi bundles and services which are located in different remote locations. Each of the component shown in the framework architecture like information sources, or the components shown in network environment in Figure 4.1 is developed as an OSGi service bundle. The proposed framework takes advantage of the *ServiceListener* method of OSGi in order to dynamically react to the change in state of services. For, example, if a service becomes unavailable; the system dynamically finds other service provider bundles and fetches the information without the knowledge of the service consumer. This method facilitates in listening to the events which are published by the service registry. Whenever, a service is registered or unregistered, the OSGi framework publishes a *ServiceEvent* to all registered *ServiceListener*. So, the OSGI framework facilitates in monitoring the services dynamically. Each component does a specific function and is briefly described below:

4.2.1 Event Monitoring Unit (EMU)

EMU is responsible for monitoring information dynamics and is the main block of the proposed framework. It consists of three different subunits; Registration Unit (R-Unit). Source Selection Unit (SS-Unit) and Network Monitoring Unit (NM-Unit).

Registration Unit(R-Unit)

To be able to access information by the users, all the information sources must be registered with the OSGi service registry. So, when the information sources are created, the information sources are made to register with the OSGi registry system. This unit keeps record of every registered information source that can be consumed by the information user. The R-Unit looks for published services (information sources) continuously and registers with the service registry whenever it finds one. It is also responsible for removing the service references from the service registry whenever they become unavailable due to various reasons like, network unreachable, or node is switched-off.

Source Selection Unit (SS-Unit)

The framework monitors the registered information sources by monitoring the status of the event called *ServiceEvent*. There are three types of ServiceEvent: *registered*, *unregistered* and *modified*. This unit invokes source selection algorithms in two cases: 1) when more than one information sources are registered, 2) when the current information source being used is unregistered and there are more than one information sources available. The un-registration happens when the particular information source becomes unavailable due to several reasons. The selection and switching of information source takes place based on requirement parameter. One such requirement parameter is information mismatch probability. Further details on this selection algorithm are provided in Chapter 7. The framework is also does dynamic re-registration and un-registration of information source as and when the information sources reappear and disappear.

Network Monitoring Unit (NM-Unit)

NM-Unit is responsible for monitoring the network parameters like event rate, rate of change of network delay and information sampling rate. Based on the network parameters collected by this unit, the computation is done as per the requirement criteria for specific purposes like information source selection, estimating mismatch probability or access strategy selection. Besides the dynamic network parameters, this unit also makes use of some static parameters like maximum power consumption limit in order to make selection decision.

4.2.2 Access Management Unit (AMU)

The function of AMU is to select a suitable information access strategy based on applications requirement. As per the applications requirement, one of the access strategy is selected among the three access strategies namely; reactive, periodic and hybrid access strategies. It is assumed that the information are sampled and stored in the cache memory from where the information user gets access as per the selected access strategy. In later chapters, these three access strategies are explored in relation to information mismatch probability and power consumption. In particular, an analytical model for information mismatch probability and power consumption under each access strategy is developed in Chapter 5. The analysis of how different access strategy do affect information mismatch probability and power consumption is carried out both analytically (using models developed) and simulation (using proposed framework).

4.2.3 Information Source Unit

The information sources which are supposed to be sensor nodes are depicted as S_1, S_2 and so on. The function of this source is to generate information in response to the sensed data from environment. Each information source is modeled as an OSGi service bundle and made to register with the

EMU with the help of RR-Unit. In particular, both the information inter-arrival rate and the service delay rates are pick up from random numbers. This can be modeled as any distribution depending on the application. Each information source model has four sub modules which are used to accomplish specific task to mimic dynamic behaviour of the source.

Inter-arrival Time

As mentioned in Chapter 3 in Section 3.1, a node in a typical wireless sensor network generates two types of traffic; locally generated and relay traffic. The inter-arrival time of locally generated traffic depends on applications. In [55], it is mentioned that the WSN applications can be categorized as event-driven and periodic data generation. For applications like environmental data collection, where sensors nodes repeatedly poll their sensed data, the data generation is periodic while for applications like safety monitoring, the data generation is not periodic, but rather sporadic. For such applications the sensor nodes send data whenever certain things occur only. So modeling information source depends on what type of applications one is interested.

End-to-End Delay

As mentioned in Section 3.1, this study focuses on locally generated atomic information. Although, the study considers a single delay distribution from source to destination and the intermediate nodes relay without altering information value, the resulting delay due to relaying, routing and retransmissions are assumed to be the part of end-to-end delay distribution that the study focuses on.

Sampling time interval

The sampling time interval is one of the main parameter that is focused on this study. This is because, how fast or how slow the information is being sampled from the information source has greater impact on the information reliability and power consumption.

4.3 Framework Implementation

Figure 4.2 shows the building blocks of the framework. For simplicity, a detailed composition of an information source is shown only with a single information source labelled as Information Source 2. However, all information sources are composed of all the components shown in the sample module. This section provides the implementation details regarding how each component is modeled to facilitate in investigating the information dynamics

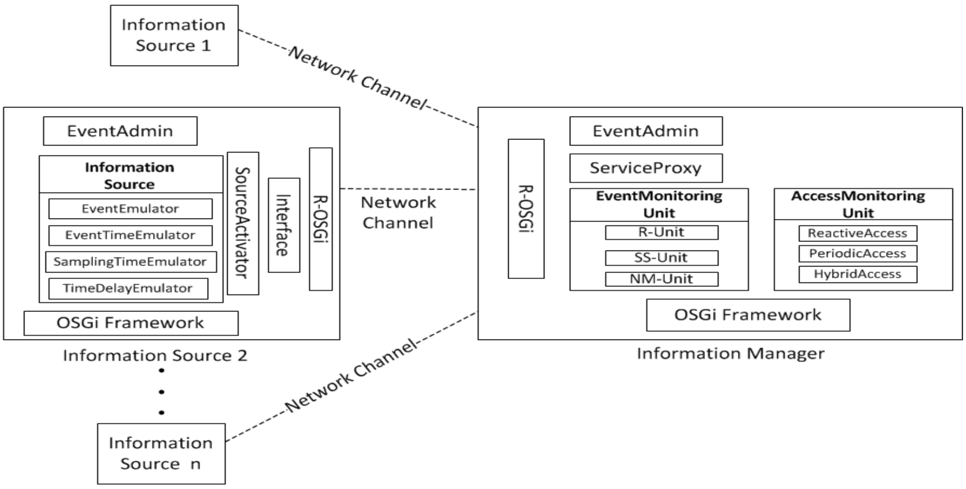


Figure 4.2: Building blocks of the proposed framework.

affecting information reliability and power consumption.

4.3.1 Information Source

Each information source model is composed of five modules which does specific task towards emulation of dynamic information source. Each module description is given below:

EventEmulator

As mentioned in Section 4.2, the proposed framework is developed using Java based OSGi framework. The *EventEmulator* is a method that acts as a physical environment where events are occurring, such as temperature monitoring in a region. Let us say that an event is generated whenever temperature value exceeds certain threshold. Considering such an event does not occur frequently, the probability that the event occurs at any time is assumed to be governed by Poison process. However, to make event process generic, the method here picks up any random value from a probability distribution function (PDF) whenever the method is invoked. This method is invoked according to the event inter-arrival time generated by the *EventTimeEmulator* method. In particular, it has a timer function that triggers as per the inter-arrival time generated by the *EventTimeEmulator*.

EventTimeEmulator

This method is responsible for generating inter-arrival time of an event. It continuously generates an independent identical exponentially distributed random numbers with parameter λ. These exponentially distributed random values are fed as an input to the timer function within the *EventEmulator* method, which then picks up a value from a PDF. Both the information value and the time at which the information is generated are stored in a separate linked list, called *sourceeventlist* and¡emphsourcetimelist respectively. The sampling of information is done from this linked list. In order to sample with fresh and up-to-date information, the last item in the list is extracted every time the information is sampled. The information extraction is accomplished with the help of a timer function which accepts the sampling time interval as an input.

Sampling Time Emulator

The sampling of information is done with the help of this method. The method generates random numbers which are independent identical exponentially distributed with parameter τ. These random numbers are fed as input to the timer function within an *EventTimeEmulator* which when the timer expires, picks up an information value (the last item) from the linked list of information (*sourceeventlist*). The sampled information is assumed to be stored in a cache memory also in a form of linked list as *eventatcache*, from where the user can access depending on the access strategy selected. Corresponding to the sampled information from *sourceeventlist*, time value is also extracted from *sourcetimelist* and stored in a *timeatcache* linked list.

A block diagram representation of interaction among above three modules is depicted in Figure 4.3.

SourceActivator

Each and every information source has to be published in order to get access by the information user. *SourceActivator* is a Java class which registers the particular information source with the OSGi service registry system. It is accomplished with the help of *registerService* method by providing its properties. An interface block shown in Figure 4.3 is used to specify the channel to the monitoring system where EMU is hosted. If the sources (emulated one) have to be replaced with real sensors, then it is possible by changing this interface.

TimeDelayEmulator

This is a method whose task is to generate a random numbers which is also assumed to be independent identical exponential distributed with parameter . Every time the information is accessed, the method picks up a value from this exponential distribution and supply to the network monitoring unit (NM-Unit). This value is used as an end-to-end network delay for computation of information mismatch probability.

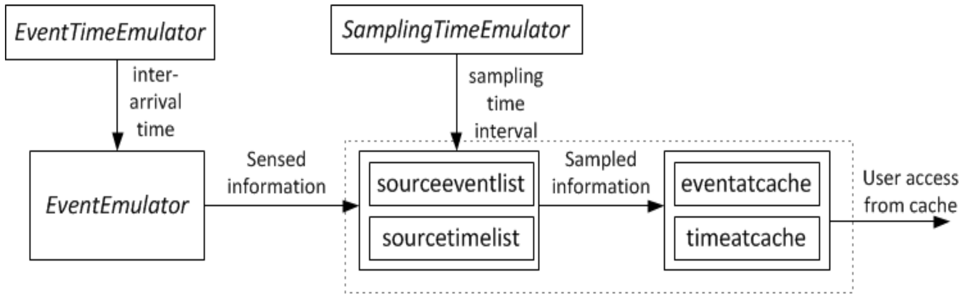


Figure 4.3: Interaction block diagram of information generation and sampling.

4.3.2 Information Manager

Information manager mainly consists of two major units; EMU and AMU. The function of EMU is to monitor the events and manage information sources depending up on the information dynamics while AMU is responsible for managing access strategy as per the user preference.

EMU is modeled with three subunits and how they are modeled is presented here:

R-Unit listens for registered information sources with the help of *addServiceListener* method. When information sources are found, this unit gets the references to all the information sources by using the interfaces though which they are published and constructs a list of sources identified by their properties. This is accomplished by using the *getRemoteServiceReferences* and *getRemoteService* methods.

SS-Unit monitors the status of registered information sources by using *serviceChanged* method. This method takes three types of arguments; *registered*, *unregistered* and *modified*. Whenever the information source is registered or unregistered the list in the *R-Unit* is updated. However, the modified status is not utilized in this framework. *SS-Unit* has an algorithm which selects the best available information source based on the set criteria. One such criterion is information mismatch probability, based on which this algorithm is implemented in Chapter 1.

NM-Unit uses the list of information sources constructed in the *R-Unit* to get access to each and every information sources available. It extracts the parameters like inter-arrival time, sampling rate, the delay in accessing information from each source. Basically, this unit gathers the required parameters and compute information mismatch probability in accessing each information source based on which the *SS-Unit* can choose the information source. A block diagram showing the interactions among these units is shown in Figure 4.4 and the steps of how interactions are accomplished are summarized below:

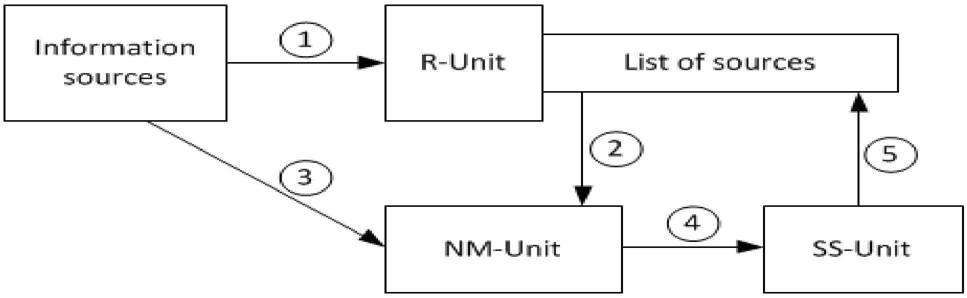


Figure 4.4: Interaction among the unit within information manager.

1. *R-Unit* prepares a list of information sources, identified by their properties.

2. *NM-Unit* gets the references for all the information sources in the list prepared by R-Unit.

3. *NM-Unit* accesses network parameters and compute required metric, e.g. information mismatch probability

4. *SS-Unit* uses the above computed metric, based on which an information source is selected in step 5.

Similarly, AMU has also three thread of control, which is responsible for implementing the particular access strategy. As per the access strategy selected by the application user, corresponding thread of control is triggered.

4.4 Performance Evaluation

The information manager unit within EMU needs to execute lines of code to accomplish the task event monitoring like registration of information source, monitoring the network conditions, switching the information sources, etc. In doing this it takes some time (computation delay) which is assumed to be negligible when the estimation of mismatch probability was carried out. This computation delay will increase when the number of information sources increases. In this study, the number of information sources is equivalent to the number of OSGi components included in the framework. In Figure 4.5,

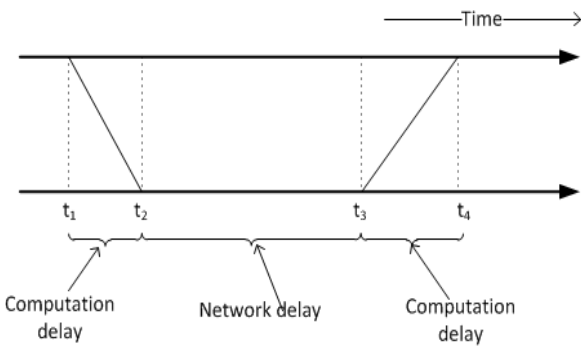


Figure 4.5: Shows computation delay with network delay.

the computation delay (δ)= (t_2-t_1)+ (t_4-t_3) and the network delay (D_ν) is shown as t_3-t_2. So total delay is given by

$$D_{Total} = \delta + D_\nu$$

Assuming $D_\nu >> \delta$, but δ depends on number of nodes (OSGi components) added to the framework.

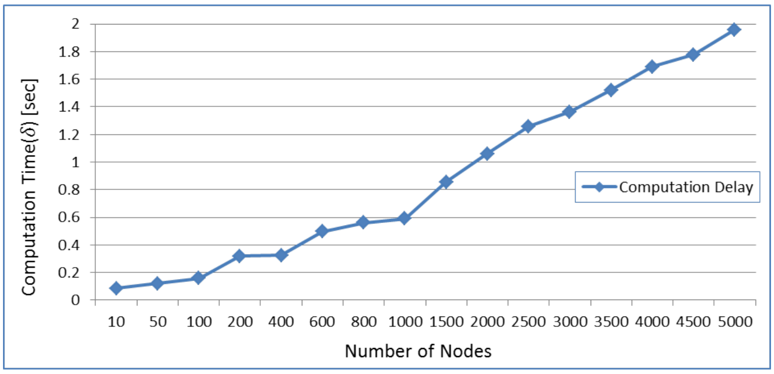


Figure 4.6: Computation delay verses number of nodes.

Figure 4.6 shows how this computation delay (δ) increases with increase in number of nodes. This is to investigate what is the best number of nodes the framework can support without affecting the performance of the system. Otherwise, if the number of nodes is such that it results a significant computation delay, then it should be added to the network delay to get the correct total delay.

As an example, if the system has to support 800 number of nodes, it takes a computation delay of 0.6 sec (refer Figure 4.6). In case if the network delay is also about 0.6 sec, then the computation delay has equal significant to mismatch probability. Therefore, in order to estimate information mismatch probability correctly, it is necessary to consider this computation delay.

4.5 Limitations of the Framework

In this section, some limitations of the proposed framework are highlighted. The main limitation of this framework is that the real deployment of physical sensors is implemented. But using OSGi, which provides modularity and catches the dynamic updates in different modules, the study showed the concept of dynamic information source in order to study the behaviour of dynamic information elements. Although, the performance evaluation on computation time is tested but there is also requirement to see how much memory is being used, CPU time in running the framework. Therefore, it deserves further study on these issues, which will be explored as a future work. Although, making interactive, a graphical user interface (GUI) is desirable, however, at present stage the proposed framework does not have GUI.

4.6 Conclusions

To conclude, the main points are summarized below:

- The system architecture of the proposed framework and its implementation details were presented.
- The information source modeling details were also presented, which is not a trivial to emulate as to show the dynamic behaviour of information.
- As computation delay was assumed to be negligible compared to network delay, computation delay was estimated to show how many number of nodes can be supported before the computation delay reaches a point where it cannot be neglected.

This chapter showed the concept of dynamic information sources with the help of dynamic information accessing and management framework. The developed framework is further used for studying the behaviour of dynamic information elements throughout the later chapters.

Chapter 5

Power Consumption and Mismatch Probability in Relation to Information Access Strategy

The framework for dynamic information accessing and management was proposed in Chapter 4. In order to back up the proposed framework and to further analyze how information access strategy do impact information reliability and power consumption, this chapter presents an extended analytical model for mismatch probability and power consumption under three information access strategies; reactive, periodic and hybrid from the one defined in [6]. In particular, an extended model presented here captures a combined probability of information mismatching that will occur during sampling of information and information accessing, in contrary to the one defined in [6], where it was considered separately. Furthermore, this chapter also introduces power consumption models for each access strategies. A comparison between the mismatch probability values estimated with the help of the proposed framework in Chapter 4 and the values obtained using analytical models are made in order to validate the results. Lastly, this chapter also investigates the impact of shifted delay on information reliability. In particular, a trade-off between deterministic preprocessing delay and stochastic network delay is investigated.

5.1 Motivation

Accessing information from a remote dynamic information source is always a challenge as mention in Section 2.6. There are several reasons why it is always a challenge starting from design issues to the routing techniques used to gather information as described in Chapter 2. Although, there are several issues that are of important with respect to accessing remote dynamic information elements, this chapter focuses on three requirement parameters in relation to information access strategy. These three requirement parameters are briefly revisited here, but more description can be found in Section 2.6.

Information Reliability: As most of the wireless sensor networks are made to collect the information, which is dynamic in nature, it is important that the information being received and used by the application user is fresh enough so that the decision made based on the information received is reliable.

Timeliness: Receiving information on time is the desire of all the applications irrespective of types of network. But the severity and impact of using outdated information varies depending on the types of application. Therefore, it is important to investigate how different applications requirement can be fulfilled by using different types of information access strategy.

Power Consumption: Generally, wireless sensor nodes are powered by battery. When such nodes are deployed without any energy recovery mechanism, the lifetime of the node is determined by how many hours the battery can last before it becomes dead. Therefore, efficient management of power usage while accessing information from the sensor node will help to extend the lifetime of the node.

In Chapter4, the information reliability in terms of mismatch probability is found to be affected by the way how information is being sampled and remotely accessed. Inspired by this observation, in this chapter, a comprehensive study on power consumption and information reliability in relation to information access strategy is provided. In summary, this chapter aims at providing the following in relation to power consumption and information mismatch probability:

- To develop an analytical models for mismatch probability and power consumption under each information access strategy.

- To investigate the impact of information access strategy on power consumption and mismatch probability under varying network conditions and dynamic information elements.

- Lastly, to investigate the impact of time shifted delay distribution on information mismatch probability to find a trade-off between deterministic preprocessing delay and stochastic network delay.

Based on the above studies, the chapter makes some recommendations when and where, which access strategy is suitable depending up on the types of application.

5.2 Related Works: Power Consumption and Reliability

Much of the works in the literature is either concentrated on energy consumption or on reliability and not much work on how information access strategies have an impact on these two parameters are found. However, both these parameters are somehow influenced by how information is being accessed [53]. A theoretical and simulation based power consumption on WSN has been explored in [56]. They have mainly focused on how data aggregation can reduce the energy consumption. However, it is mainly concentrated on the computation of power consumption and the reliability issues of information are neglected. Similar work is done in [53], where they also use the same approach as in [56] in calculating the energy consumption in WSN. In both the works, no study is carried out on how the way one access the information does impact the reliability and power consumption. In [23], attempt is made on analysis of lifetime and energy consumption based on access strategies but it failed to discuss about all the three access strategies and concentrated only on periodic access strategy. Data collection in wireless sensor network with mobile sink with respect to energy consumption is studied in [57] and [58]. In [57], a two tier distributed hash table-based scheme data collection for event-driven wireless network is proposed. The data is collected via a single hop routing between a node and the mobile sink. It is said that when an event is occurred in the region, the nodes which are in the region will announce the event which will trigger the mobile sink to move towards the region where event has occurred. In [58], the movement of the mobile sink is made arbitrary with assumption that sink has no resource limitation. This assumption is not really practical when wireless sensor networks are deployed in harsh terrain and hostile environment conditions. The study have not considered any particular access strategy and investigated only energy consumption during transmission. An event-driven wireless sensor networks using energy-saving data collection is proposed in [59]. They have considered wireless sensor networks with high density of nodes. They divide data collection method into four phases: extracting edge-nodes, setting freeze nodes, collecting data and making branch tables. The data is collected to the stationary sink from the edges nodes via the relays nodes.

The concept of mismatch probability and access strategies is extensively explored in [6, 45, 46, 47]. However, none of the works considered the effect of access strategies on power consumption. Therefore, this chapter investigates more on mismatch probability considering both sampling of event part and the information transmission coupled with power consumption for each access strategy.

5.3 System Model Description

5.3.1 Abstract Scenario

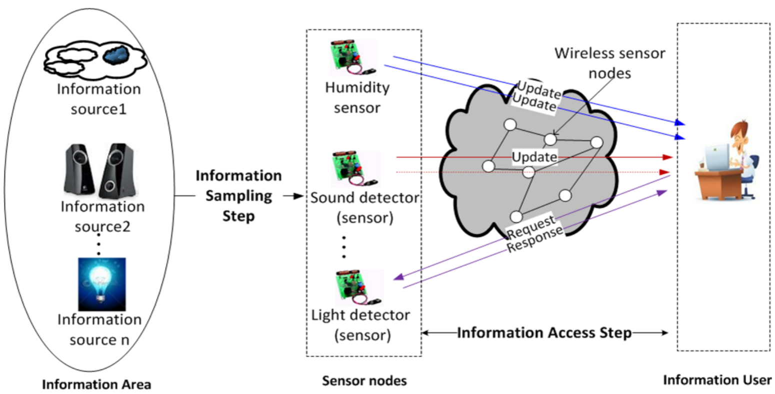


Figure 5.1: An abstract view of remote dynamic information access with different access strategies.

Figure 5.1 shows an abstract view of remote dynamic information access model considered in this study. This model is based on the basic information access model described by Figure 3.1, but with incorporation of different access strategies. Further, all assumptions regarding the modeling of information source in Chapter 4 are valid and will be applying for developing analytical models in this chapter.

Consider some dynamic events occur within an *information area* which is within the sensing range of sensor nodes (information source). These sensor nodes generate information in response to the sensed events from the information area, which act as information sources for the *information user*. In all the access strategies, we model information access as a two-step process: 1) *information sampling step*, where the information is collected by the sensor nodes from the information area and transferred to a cache at every sampling period, and 2) *information access step*, where the information collected by the nodes is accessed by the remote information user using different access strategies over a network environment. The model is based on the locally generated information packet as mentioned in Section 3.1. Therefore, model does not take into account the packets which generated within intermediate nodes between information source and the information user. However, the delay due to relaying packets within intermediate nodes and retransmission are assumed to be a part of resulted end-to-end delay.

As described in Chapter 4, the model assumes a sampling process of some dynamic information which changes only at discrete points in time, considered as a Poison process for simplicity. The inter-arrival time of information and the end-end delay are assumed to be exponentially distributed

5.4 Analytical Model

Out of various factors which impact the performance and life time of wireless sensor node, the model focuses on information reliability and power consumption. In this section, analytical models for mismatch probability and power consumption for the two steps; information sampling and access step are presented.

5.4.1 Information Sampling Step

Since the sampling of information is done periodically and is common for all the access strategies, the mismatch probability and power consumption model would be same for all the three access strategies during information sampling step.

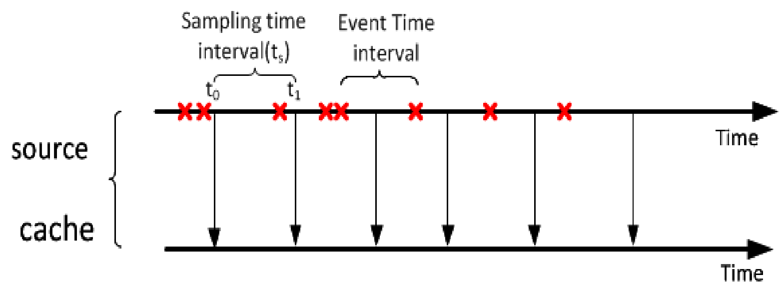


Figure 5.2: Sampling process model; Sampling time intervals ($t_s = t_1 - t_0$) are exponentially distributed with rate τ

Mismatch Probability

The information source samples information with exponential distributed time intervals, but periodically and assumed to be stored in its cache to be accessed and used by information user locally and later it is from this cache, remote access fetches data. But it can be argued that due to clock drift, scheduling effects etc., the sampling process is not perfect, but shows some stochastic behaviour. Simulation studies in [45] show that if the period is deterministic, then the mismatch probability will be lower but follow same behaviour as the exponentially distributed interval. Thus, the stochastic sampling assumption can be seen as a worst case scenario. The remote information user accesses the information as per access strategy being used. This assumption allows the mismatch probability to expressed as given by mismatch probability of periodic access derived in [6] as

$$mmPr_{Per} = \int_0^\infty exp\left(-\int_0^t \tau F_D(s)ds\right) A_E(dt) \tag{5.1a}$$

with A_E as the distribution of backward recurrence times of the event process [6, sec.2, p.45]. For deterministic delays, and in particular delays of $t = 0$, and exponentially distributed event time process, (5.1a) simplifies to

$$mmPr_{Samp} = \frac{\lambda}{\lambda + \tau} \tag{5.1b}$$

where λ is the rate of information update and τ is the sampling rate.

Power Consumption

As described in Chapter 3 in Section 3.6, the power consumption model for an information source can be expressed as

$$P_t = P_{samp} + P_{tx} + C \tag{5.2a}$$

where, P_{samp}, P_{tx} are the amount of power dissipated for sampling and transmitting respectively. The term C (constant) is introduced for other power terms like receiving power, idle power, etc. Since, we are interested in comparing power consumption with different access strategies, we only consider P_{samp} and P_{tx} since the rest of the components remain same for all of our information access strategies. The term C is neglected and is never used in our following analysis. The power consumption for the sampling process will be same for all the three access strategies but the power consumption in transmission would depend on how information is being accessed. If t_{samp} is the time required for each sampling with rate τ , the average power consumption for sampling over time, T can be expressed as

$$\bar{P}_{samp} = \frac{1}{T}\int_0^T \tau P_{samp}(t)dt \tag{5.2b}$$

With an assumption of sampling time interval as exponentially distributed, then (5.2b) simplifies to

$$\bar{P}_{samp} = P_{samp} t_{samp} \tau \tag{5.2c}$$

To summarize, information mismatch probability during information sampling step is decided by at what rate the information is changing at the source and the how fast the information is being sampled, here with assumption that the event rate follows exponential distribution. The average power consumption during the sampling step is governed by sampling rate alone and not by the rate at which the information changes. The key point here is therefore, both the power consumption and information mismatch probability can be controlled by varying sampling rate as per the requirement, i.e. how much information reliability an application needs to maintain or what amount of power an information source can manage to spend.

5.4.2 Information Access Step

As mentioned in Section 5.3.1, although, the models described in this chapter are applicable to both types of packets; our analysis generally assumes locally generated packets. Access to information that requires multi hop or cases where data needs to be retransmitted due to e.g. bad link conditions can be modeled via the statistical description (cumulative density functions) of the access delays. Once the sensor node gets the information, it either waits for the remote information user to request for information (reactive access) or it proactively distributes the sampled information to its subscribers (proactive access). The following sections analyze how these different access strategies do affect the mismatch probability and power consumption.

Reactive Access Strategy

Reactive access strategy is one, where the information is accessed on-demand. This access strategy is depicted in Figure 5.3

Mismatch Probability

The mismatch probability for reactive access assumes that information in the cache is correct at time t_2 and measures the probability that the information changes in the time interval t_2 to t_0 (see also Figure 5.3). In [6], this definition of mismatch probability for reactive access strategy is derived as given below.

$$mmPr_{Rea} = \frac{1}{\mathbb{E}(E)} \int_0^\infty \bar{F}_D(t) \bar{F}_E(t) dt \tag{5.3a}$$

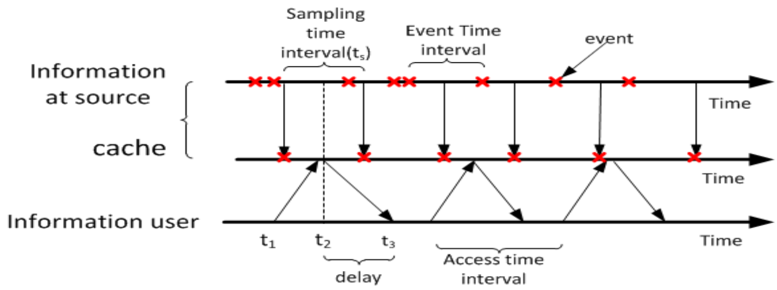


Figure 5.3: Reactive Access Model with Periodic Sampling

With F_D as the distribution of the time intervals (delays i.e. $t_3 - t_2$) and F_E the event time interval distribution, and the over line indicating the reliability function, $1 - F_x$, assuming request arrives equally likely in time intervals.
With an assumption of delay and event time interval are all independent and identical exponentially distributed (with rates λ for event and ν for delay), then (5.3a) can be simplified as

$$mmPr_{Rea} = \frac{\lambda}{\lambda + \nu} \tag{5.3b}$$

The combined mismatch probability with the sampling, cannot assume correct value at time t_2 due to the sampling scheme. It can be calculated as a weighted average of two scenarios; 1) the information is already mismatching when being accessed at time t_2 (see Figure 5.3), which happens with probability $mmPr_{Samp}$. 2) information was correct at time t_2 but information at source changed during time t_2 to t_3 , which happens with probability $mmPr_{Rea}$. By reshaping the existing mismatch probability model of [6] to a sampling process by reducing otherwise included network delay into a deterministic zero delay, the overall mismatch probability for reactive access can be expressed as

$$mmPr_{ComRea} = mmPr_{Rea} + mmPr_{Samp}(1 - mmPr_{Rea}) \tag{5.4a}$$

Substituting (5.1b) and (5.3b) in (5.4a), we get the expression for combined mismatch probability, under the given exponential assumption as

$$mmPr_{ComRea} = \frac{\lambda(\lambda + \tau + \nu)}{(\lambda + \tau)(\lambda + \nu)} \tag{5.4b}$$

When the delay rate approaches to infinity (meaning that the mean network delay (D) goes to zero) the overall mismatch probability $mmPr_{ComRea}$ approaches to the mismatch probability of the periodic access. This shows that the lower bound of mismatch probability is, as expected, given by the sampling process as shown in below equation:

$$mmPr_{Samp,D=0} = \lim_{\nu \to \infty} \frac{\lambda(\lambda+\tau+\nu)}{(\lambda+\tau)(\lambda+\nu)} = \frac{\lambda}{\lambda+\tau} \tag{5.4c}$$

Similarly, it can be shown that as sampling rate (τ) approaches infinity (meaning that information is sampled at faster rate, or in other words sampling time approaches zero), the combined mismatch probability approaches to mismatch probability given by pure reactive (where information is correct at time t_2 access.

$$mmPr_{Rea,T=0} = \lim_{\tau \to \infty} \frac{\lambda(\lambda+\tau+\nu)}{(\lambda+\tau)(\lambda+\nu)} = \frac{\lambda}{\lambda+\nu} \tag{5.4d}$$

It is natural from (5.4b), when $\lambda \to \infty$, $mmPr_{ComRea} = 1$.

Power Consumption

Since the transmission of information takes place only when user makes a request, the reactive transmission power consumption depends on the information access rate and the transmission delay. If μ is the average information access rate, and P_{tx} is the power consumption for each transmission, then the average power consumption over time, T can be expressed as

$$\bar{P}_{tx} = \frac{1}{T}\int_0^T \mu P_{tx}(t)dt \tag{5.5a}$$

With same concept and exponential assumption as in (5.2c), (5.5a) reduces to

$$\bar{P}_{tx} = P_{tx} t_{tx} \mu \tag{5.5b}$$

Using (5.2c) and (5.5a) in (5.2a), the total average power consumption for both information sampling and access steps can be written as

$$\bar{P}_{Rea} = P_{samp} t_{samp} \tau + P_{tx} t_{tx} \mu \tag{5.5c}$$

To summarize on reactive access strategy, the main points are listed here:

- The mismatch probability under pure reactive i.e. $mmPr_{Rea}$ is influenced by the rate at which information is changed and the rate of delay between information source and the information user. So, it is not affected by the rate at which the information is sampled.

- The combined information mismatch probability i.e. $mmPr_{ComRea}$ is influenced by all the three parameters, event rate, rate of delay and sampling rate.
- The combined mismatch probability, $mmPr_{ComRea}$ approaches mismatch probability at the information sampling step as the delay rate goes to infinity, i.e. when delay (D) goes to zero.
- The combined mismatch probability, $mmPr_{ComRea}$ approaches mismatch probability given by pure reactive as the sampling rate goes to infinity, i.e. when sampling time (T) goes to zero.
- The average power consumption under reactive access strategy depends both on sampling rate and the information access rate.
- Therefore, the main take away point from this section is how to vary sampling rate in order to keep mismatch probability same, given event rate (λ) or delay rate (ν) changes.

Periodic Access Strategy

In this access strategy, instead of information user initiating with a request message, the information source proactively distributes any updates to the potential information users in periodic time interval. Under this proactive periodic access, the time interval at which the information source periodically distributes information to the potential information users is assumed to be equal to the sampling rate. This means that every time the information is sampled, the information user gets updated information. In other words, the information sampling rate is equal to the information access rate as shown in Figure 5.4. So, the alternative access rates do not make sense as the information is updated and stored in a cache at a rate given by sampling rate.

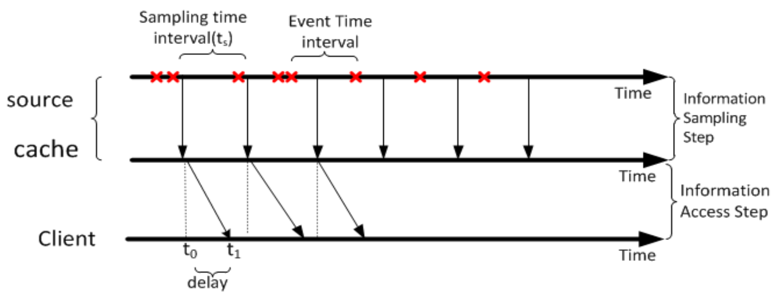


Figure 5.4: Periodic Access with Periodic Sampling

Mismatch Probability

The mismatch probability for proactive periodic access is derived in [6] and given by (5.1a). With assumption of event being Poison process with intensity λ and downstream delays are i.i.d exponentially distributed with mean $1/\nu$, (5.1a) can be re-written as [6]

$$mmPr_{Per} = \lambda \int_0^\infty exp\left(\frac{\tau}{\nu}(1 - e^{-\nu t})\right) exp(-(\tau + \lambda)t)dt \qquad (5.6)$$

The above expression (5.6) is alternatively expressed in [6] as (by change of variable t and $e^{-\nu t}$, and using notation $\phi = \frac{\lambda}{\nu}; \psi = \frac{\tau}{\nu}$).

$$mmPr_{ComPer} = \phi e^{\psi} \frac{\Gamma(\psi + \phi)}{\psi^{\psi+\phi}} F_{\Gamma(\psi+\phi,\psi)}(1)$$

where $F_{\Gamma(a,b)}$ is the cdf of a gamma distribution with parameters a and b.

Power Consumption

In this access strategy, the power consumption for information access step will not depend on information access rate; instead it will depend on sampling rate,τ . This is because the average information access rate is equal to the sampling rate. Using same concept as in reactive access strategy, the average power consumption for both the sampling and information access steps can be can be expressed as

$$\bar{P}_{Per} = (P_{samp}t_{samp} + P_{tx}t_{tx})\tau$$

To summarize on periodic access strategy,

- The combined information mismatch probability under periodic access strategy is influenced by all three parameters, event rate, sampling rate and the delay rate.

- In contrary to reactive access strategy, the power consumption under periodic access only depends on the sampling rate. This is due to the fact that, every time the information is sampled, the information source distributes it to the information user.

Hybrid Access Strategy

As described briefly in Section 3.4, proactive access can be categorized into two sub access strategies; 1) proactive periodic and 2) proactive event-driven in general. In proactive periodic access (see Figure 5.4), the information is sent or accessed at each time the information is sampled irrespective of whether there is any update or change of information value from previous access. Usually, in proactive event-driven the information is sent only when some event has taken place (e.g. like temperature exceeds the set threshold value). But in this model, a slight modified form of proactive event-driven access is being considered, where the information is sent to the client only when the information value to be sent is not the same as the previous value which is already sent (see Figure 5.5). This access method is called as proactive hybrid access. Therefore, this access will avoid the redundant information being sent consecutively, thereby reducing the power consumption of the information source.

Mismatch Probability

The expression for mismatch probability for information sampling part would be given by the same expression as in (5.1a) with mean sampling rate τ, as information is sampled at every τ time interval. The reason for this is that the model that is being used here is from [6], which considers only useful updates (updates which have new information) through a thinned Poison update process. Since this model sends updates only when they contain useful updates, the mismatch probability is exactly the same, however, traffic load is different. But, the delay distribution for the information access step will be affected by a small additional delay introduced for checking whether there is any update for every sampling period (in turn affecting the mismatch probability). If d_0, which is equal to $t' - t_1$ (see Figure 5.5) is an average

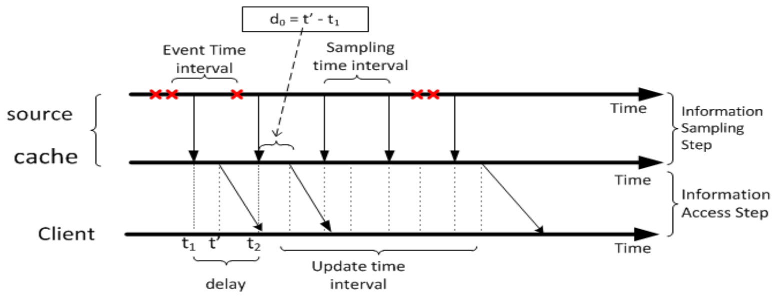


Figure 5.5: Periodic Access with Periodic Sampling

additional delay introduced by checking whether there is any update after every sampling period plus additional delay for preprocessing and D is the downstream network delay, then the total delay introduced is

$$D = d_0 + D_\nu$$

With $F_D(s - d_0)$ as the distribution of the shifted delay, the mismatch probability of hybrid access strategy will be given by (5.1a). Since it does not easily simplify, numerical integration of (5.1a) with a time shifted delay distribution is performed in order to compare mismatch probability with other access strategies (see Figure 5.17).

Power Consumption

The power consumption for proactive hybrid access would be affected by the rate at which events that exceed a threshold occurs. It is also important to consider the power required to check for an update and processing or filtering as described in event process and detection in Section 3.2. If P_{pro} is the power consumed for each checking, then average power consumption in checking an update over a time, T can be expressed as

$$\bar{P}_{pro} = \frac{1}{T}\int_0^T \tau P_{pro}(t)dt$$

The total average power consumption in hybrid access strategy can be expressed as

$$\bar{P}_{Hybrid} = (P_{samp}t_{samp} + P_{pro}d_0)\tau + P_{tx}t_{tx}\tau' \tag{5.7a}$$

The update rate (τ') can be computed as

$$\tau' = 1 - P(no\text{−}update)\tau$$

The probability that no any update has occurred within a sampling period can be derived if the process is stationary and memory less. Consider f_e the event arrival distribution and f_s the sampling time interval distribution (which previously assumed exponentially distributed), then under these conditions, this probability can be expressed as

$$P(no\text{−}update) = \int_0^\infty \left(1 - \int_0^t f_e(s)ds\right) f_s(t)dt \tag{5.7b}$$

With assumption of all processes being exponentially distributed (with rates τ for sampling and τ' for update) are all independent and identical exponentially distributed, then (5.7b) simplifies to

$$\tau' = \frac{\lambda\tau}{\lambda + \tau} \tag{5.7c}$$

When λ approaches to infinity (meaning that information arrival rate increases), τ' approaches to τ . This means that the upper bound of update rate is given by the sampling process as it is expected.

$$\tau' = \lim_{\lambda\to\infty} \frac{\lambda\tau}{\lambda + \tau} = \tau$$

Using (5.7c) in (5.7a), the final expression for average power consumption for hybrid access can be written as

$$\bar{P}_{Hybrid} = (P_{samp}t_{samp} + P_{pro}d_0)\tau + P_{tx}t_{tx}\left(\frac{\lambda\tau}{\lambda+\tau}\right) \tag{5.7d}$$

To summarize, the main points are listed here:

- The combined mismatch probability under hybrid access is impacted by event rate, sampling rate, and delay rate.
- In addition, it is also impacted by the time spent for checking for an update (d_0). Due to this additional delay, the information mismatch probability under hybrid access strategy will be always greater than the mismatch probability that would result under periodic access strategy.
- But, there is a positive impact on power consumption. Since, information is retrieved only when there is change in value, this update rate (τ') is always less than or equal to the sampling rate (τ).
- When update rate becomes equal to the sampling rate, the power consumption under hybrid access strategy would be same as under periodic access strategy. However, the information mismatch probability will still be affected by the constant delay (d_0) even under this condition ($\tau = \tau'$).

5.5 Validation of the Models

In order to validate the working of the proposed framework described in Chapter 4, an algorithm is designed to estimate information mismatch probability for which the steps are provided in 5.1. A comparison between the mismatch probability values estimated with the help of the proposed framework (implementation) and the one which is obtained using the analytical models is provided here. The same parameters are used for both analytical and implementation for comparison purposes. For testing purpose, the event process is modeled as Poison process; however, it can be modeled with any distribution depending on the application type.

The steps for estimating information mismatch probability is provided here in Algorithm 5.1.

In all the three access strategies under study, the two values of mismatch probability are found to be more or less same, thereby validating the

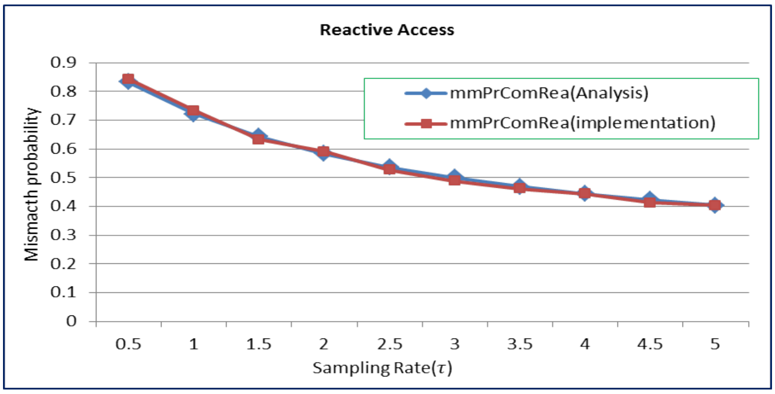
Algorithm 5.1 Steps for estimating information mismatch probability

1: The information value and time at which the information is generated are continuously recoded as a two separate linked lists, *sourceeventlist* and *sourcetimelist*
2: At every sampling time interval, the last item in the *sourceeventlist* is extracted and put into another linked list, called *eventatcache* and corresponding time value is copied to another linked list, *timeatcache*
3: At every access (as per the selected access strategy), the last item from the *eventatcache* linked list is extracted and put into another linked list called *receivereventlist*. The *receivertimelist* is formed by adding delay (with specified distribution) with the last item in the *timeatcache*.
4: For every *receivereventlist* item (with the corresponding item in the *receivertimelist*), we track back to the *sourceeventlist*. If *receivereventlist* item equals the *sourceeventlist* item, then it is a match, otherwise it is a mismatch.
5: Steps 3 and 4 are repeated for 1000 times and average value of number of mismatch is computed to get the mismatch probability.

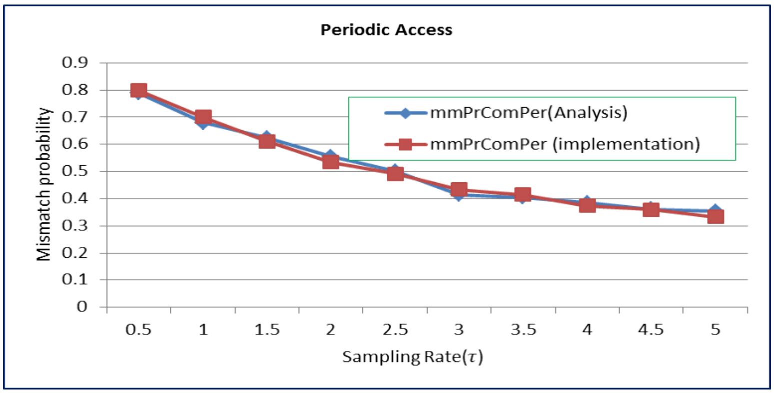
working of the proposed framework as well as showing that the analytical models are also valid. These comparisons are shown in Figure 5.6(a) to Figure 5.6(c).

Figure 5.7 is provided to show comparison among all the three access strategies. In hybrid access strategy, a constant value of $d_0 = 0.1 sec$ is added to the end-to-end delay as mentioned in our hybrid access strategy definition in Section 3.4. It is observed that the mismatch probability for hybrid is more than the other two access strategies. This is because of the extra delay incurred in checking for an update.

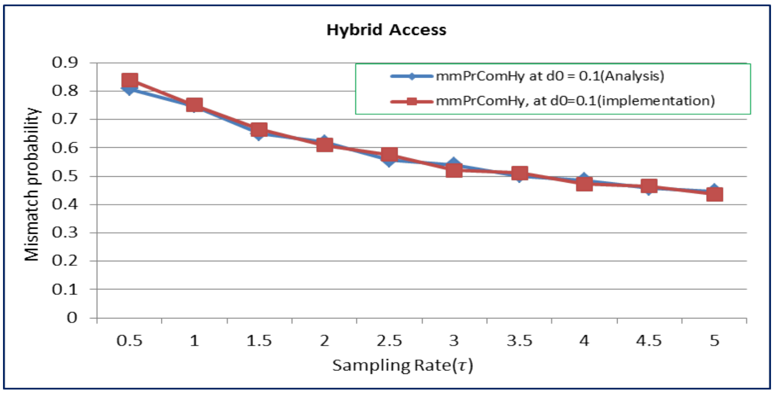
The result which is shown in Figure 5.7, the mismatch probability values were computed based on assumption that the information sampling rate is exponentially distributed. However, the sampling rate can also be kept constant instead of assuming exponentially distributed. Therefore, in order to see how mismatch probability behaves with constant sampling rate, Figure 5.8(a) and Figure 5.8(b) show comparative results between exponential sampling rate and constant sampling rate for reactive and periodic access respectively. The figures also show the difference between them. Looking at the plot (Exp.− Const.), it is observed that the difference in mismatch probability increases upto certain values and then starts decreasing. This shows that when sampling rate decreases the mismatch probability approaches 1 whereas when sampling rate increases the mismatch probability values approaches 0 for both the rates (exponential and constant).



(a)



(b)



(c)

Figure 5.6: Comparisons of mismatch probability values obtained by analytical models and using framework at $\lambda = 2.0$ and $\nu = 10$ (a) Reactive, (b) Periodic and (c) Hybrid Access.

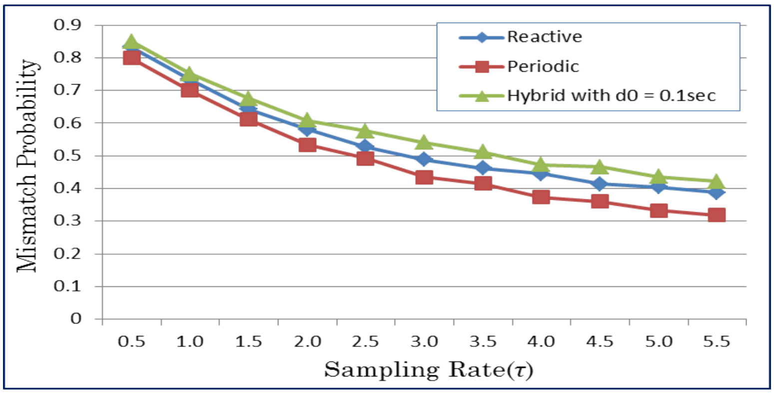
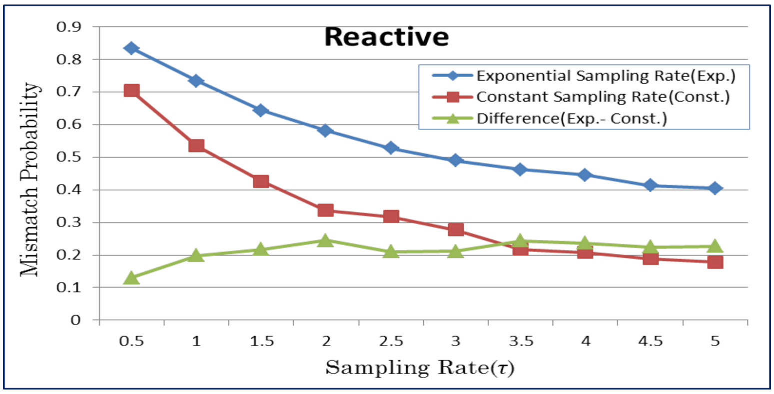
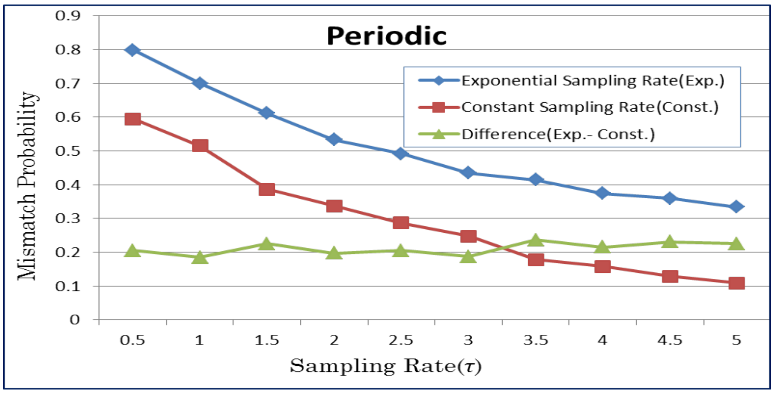


Figure 5.7: Information mismatch probability for different access strategies at $\lambda = 2.0$ and $\nu = 10$.



(a)



(b)

Figure 5.8: Information mismatch probability comparison between exponential sampling rate and constant sampling rate at $\lambda = 2.0$ and $\nu = 10$. a) Reactive Access b) Periodic Access

The result in Figure 5.7 is under assumption that the delay from the source to destination is exponentially distributed, irrespective of number of intermediate nodes en route to destination. To be more realistic, a further analysis with Erlang delay distribution is carried out, which was introduced in Section 3.5 in Chapter 3. Considering two intermediate nodes between source

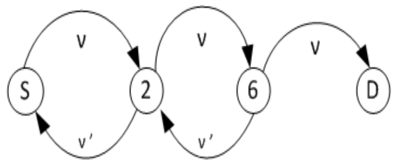


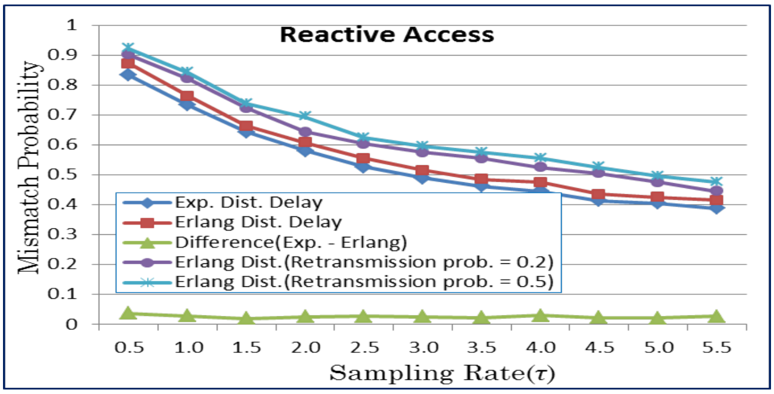
Figure 5.9: Markov Model, Transition diagram.

and destination node, with transmission delay rate of ν and retransmission delay rate of ν' as shown in Figure 5.9. Inspired from [5], the delay can be modeled by discrete time Markov chain as shown in Figure 5.9. Assuming, the delay rate from each node is exponentially distributed [51, 52] the overall mean delay is computed using matrix exponential distribution as in Section 3.5 in Chapter 3.

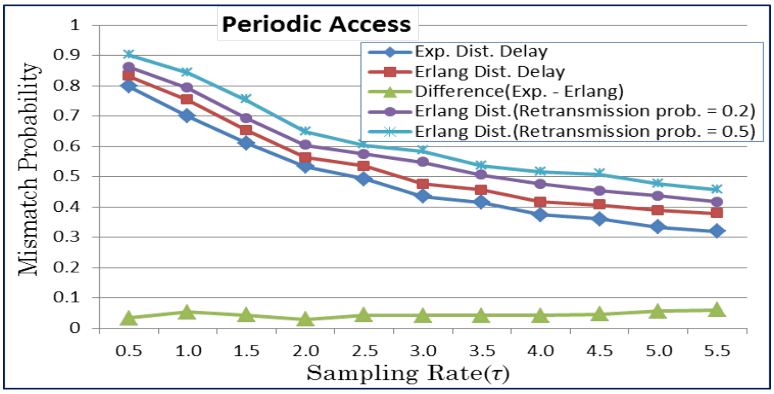
The cumulative distribution function of delay rates are computed using (3.3) both for with and without retransmission. In addition, a constant link delay offset of 0.05 sec and node processing delay of 0.01 sec are also added at each link and node to get the overall delay.

The delay obtained by considering Erlang distribution (both with and without retransmission) is fed as an input to the algorithm 5.1. The mismatch probability values obtained by assuming exponential delay distribution and Erlang distribution is compared. The results are shown in Figures 5.10(a), 5.10(b) and 5.10(c). The mismatch probability value in case of retransmission is greater than the one without retransmission.

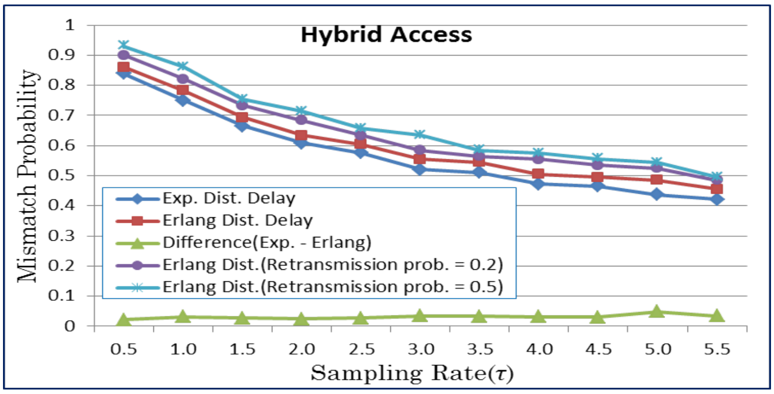
Observing the difference between the Erlang and exponential delay distribution plots, there is not much significant variation. For periodic access strategy, the difference in mismatch probability value is about 4%, in reactive access and hybrid access strategies the difference is about 2%. The difference is expected as Erlang distribution takes into account of delays caused by intermediate nodes as well as offsets (link offsets and node offsets). The difference between the mismatch probability values obtained assuming exponential delay distribution and assuming Erlang delay distribution delay is not much significant. Therefore, the analytical mismatch probability model developed (extending the one developed in [6]) in Section 5.4 considering exponential delay distribution can be considered valid to show the concept of dynamic information behaviour.



(a)



(b)



(c)

Figure 5.10: Mismatch probability comparison between Erlang distributed delay and exponential distributed delay at $\lambda = 2.0$ and $\nu = 10$. a) Reactive Access, b) Periodic Access, c) Hybrid Access at $d_0 = 0.1$sec

5.6 Performance Evaluation and Result Analysis

In this section, the performance of each access strategy with respect to mismatch probability and power consumption are analyzed. The section also evaluates the models and investigates how the mismatch probability parameter behaves with other network performance parameters; delay, sampling rate and event inter arrival time.

5.6.1 Mismatch Probability

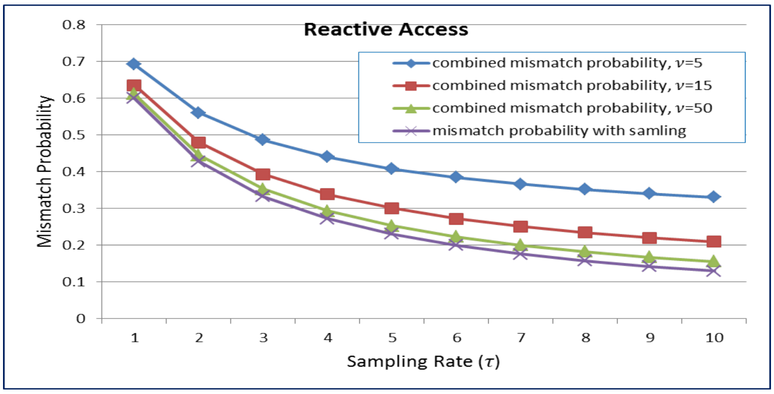
The analysis of how mismatch probability behaves with varying event rate and delay rate for all the three access strategies are provided here.

Figure 5.7 shows how mismatch probability varies with sampling rate at various mean delays (different network conditions) for three access strategies (a) Reactive, (b) Periodic and (c) Hybrid.

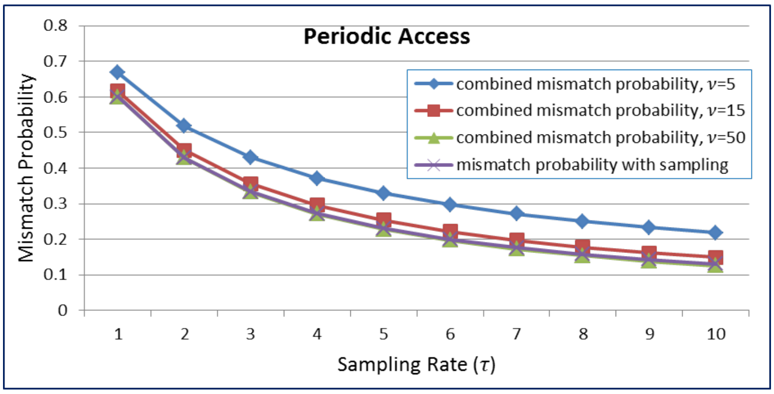
In all the three access strategies, as ν approaches infinity (meaning that the mean delay approaches to zero), the mismatch probability approaches to mismatch probability given by information sampling step ($mmPr_{Samp}$). But an interesting observation seen from Figures 5.11(a),5.11(b), and 5.11(c) is that the mismatch probabilities do not decrease beyond $mmPr_{Samp}$. This is due to the fact that the mismatch probability in information sampling step is not influenced by network delay. Therefore, the combined mismatch probability cannot be reduced beyond $mmPr_{Samp}$ irrespective of what the network conditions are. This is also shown by (5.4c) for reactive access, however such equations could not be shown in case of periodic and hybrid as the (5.1a) is not easily solvable.

Figures 5.12(a),5.12(b), and 5.12(c) show the variation of mismatch probability with sampling rate with different values of event rate (λ). As expected, the value of mismatch probability increases as λ increases for all the three access strategies. In all the three access strategies, it is observed that as event rate (λ) approaches infinity, the combined mismatch probability approaches 1. This is natural to observe as higher values of event rate (λ) means the events are changing very fast at the source.

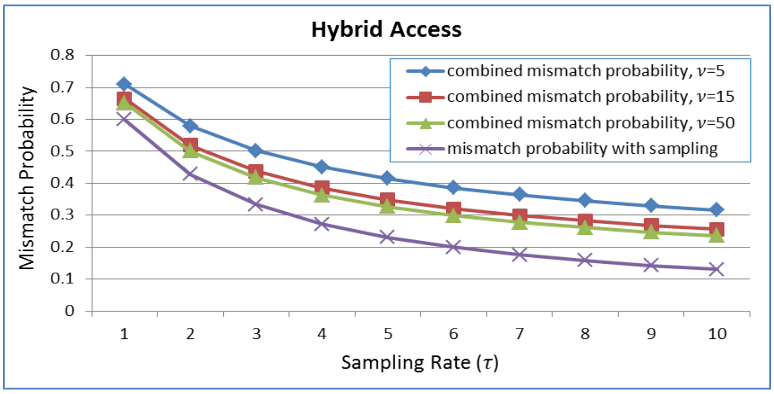
Figures 5.13(a),5.13(b), and 5.13(c) show how mismatch probability varies with event rate for different values of sampling rate (τ). In reactive access strategy, as sampling rate approaches infinity (meaning the sampling time interval approaches 0), the combined mismatch probability decreases and approaches to mismatch probability given by the pure reactive i.e. $mmPr_{Rea}$. Referring back to Figure 5.3, $mmPr_{Rea}$ is computed by as-



(a)

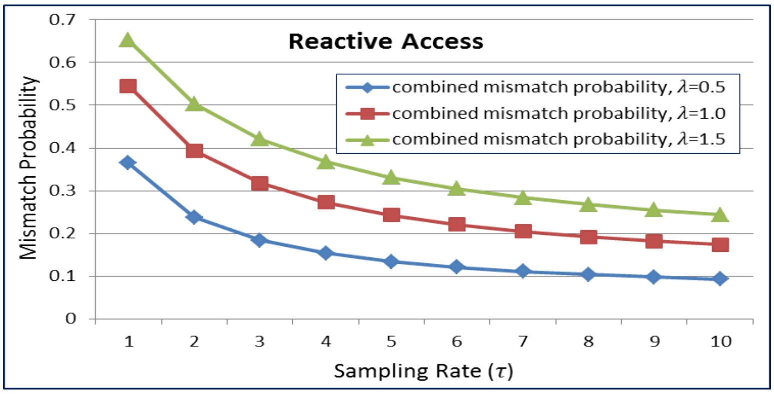


(b)

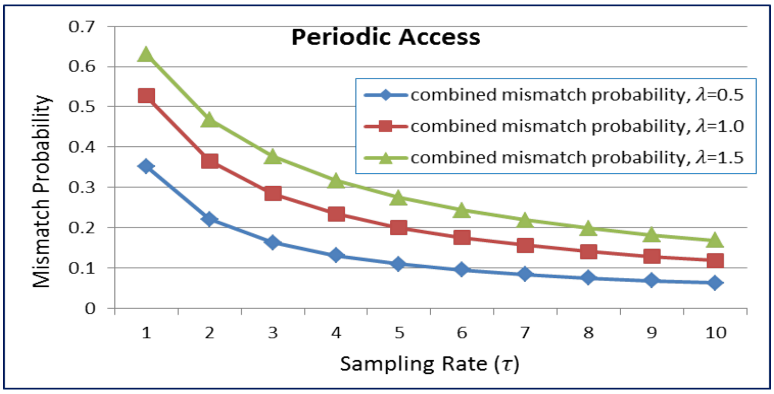


(c)

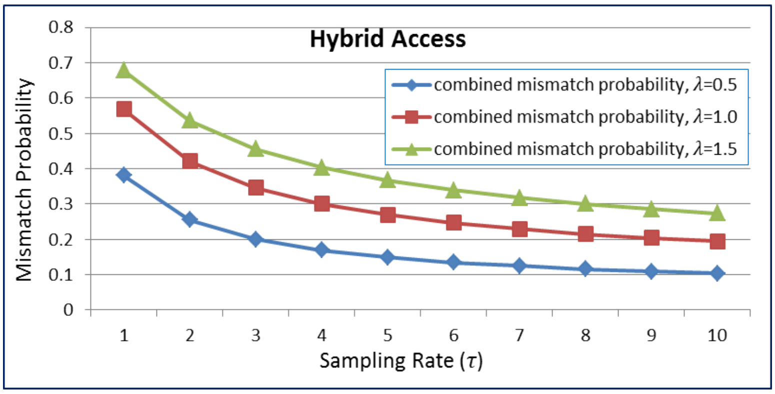
Figure 5.11: Mismatch probability verses sampling rate with varying rate of delay (ν) at $\lambda = 1.5$ (a) Reactive Access, (b) Periodic Access and (c) Hybrid Access at $d_0 = 0.1sec$.



(a)



(b)



(c)

Figure 5.12: Mismatch probability verses sampling rate with varying event rate (λ) at $\nu = 10$ (a) Reactive Access, (b) Periodic Access and (c) Hybrid Access at $d_0 = 0.1sec$.

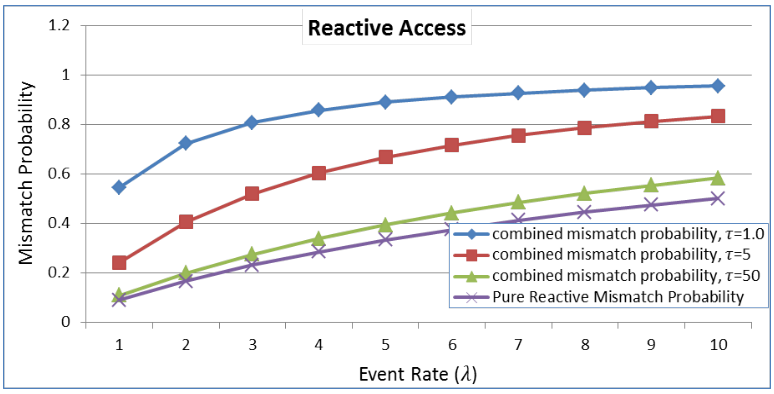
suming the information was correct at time t_2, but the information at source has changed during time interval t_2 to t_3. So, the pure reactive mismatch probability ($mmPr_{Rea}$) is not affected by sampling rate. It is therefore, the combined mismatch probability in case of reactive access does not drop beyond $mmPr_{Rea}$. This is seen from Figure 5.13(a). The same is shown mathematically in 5.4d by letting τ goes to infinity. In contrary, such behaviour where the lower limit of the combined mismatch probability is decided by $mmPr_{Rea}$ is not observed in case of periodic and hybrid access strategy (in hybrid case, Figure 5.13(c) does not show this due to the constant delay d_0 introduced, but it can be shown by reducing this delay or by further increasing the value of sampling rate). From Figure 5.13(b), it is clearly seen that the combined mismatch probability drops beyond $mmPr_{Rea}$. This is because, the under periodic and hybrid, the information update is assumed to be sent as per the sampling rate. Therefore, faster the information is sampled, the less information mismatch probability will occur, provided the mean delay remains constant.

Figure 5.14 to Figure 5.16 show the derivative of mismatch probability with respect to different parameters. These types of analysis show the level of sensitivity on information mismatch probability due to changes in information elements or network conditions.

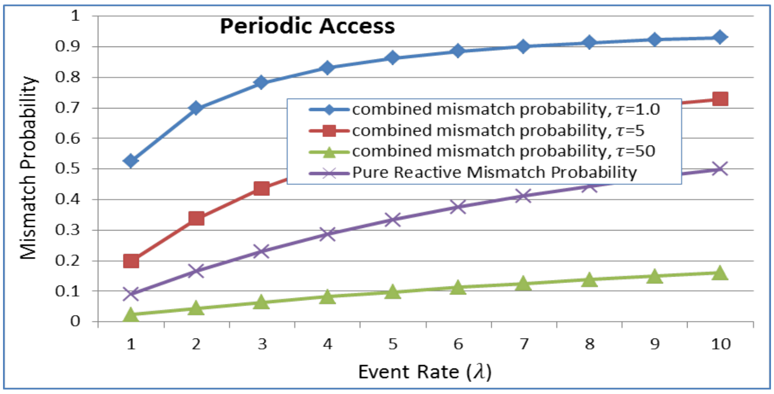
Figure 5.14 shows derivatives of mismatch probability with respect to event rate (λ). It shows that variation of mismatch probability is more at lower values of event rate (λ). In other words, information mismatch probability is more sensitive to changes in event rate when the change rate is slower compare to faster change. This observation applies to all the three access strategies. From the figure, it is observed that the Hybrid access is more sensitive to change in event rate than the other two. This also shows that if the scenario where event rate is like to change, then the hybrid will perform worst compare to other two access strategies in terms of information mismatch probability.

Figure 5.15 shows derivatives of mismatch probability with respect to mean delay rate (ν). Here also, it shows that variation of mismatch probability is more at lower values of mean delay rate (ν). Therefore, irrespective of what information access strategy being used, the information mismatch probability is more sensitive to changing network delays at lower values of delay rates than at higher values of delay rates. Contrary to Figure 5.14, here the reactive access strategy is more sensitive to the change on network delays. This means that if the scenarios is such that the network delay is likely to change, hybrid gives better performance in terms of mismatch probability.

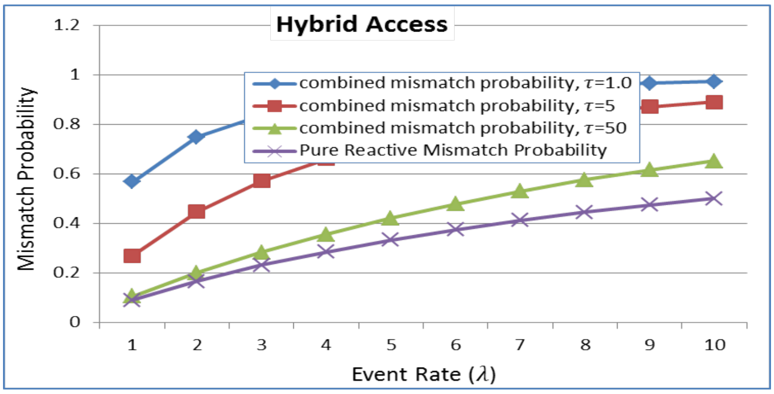
Similarly, a derivative of mismatch probability with respect to information sampling rate is shown in Figure 5.16. Here too, the variation of



(a)



(b)



(c)

Figure 5.13: Mismatch probability verses event rate with varying sampling rate (τ) at $\nu = 10$ (a) Reactive Access, (b) Periodic Access and (c) Hybrid Access at $d_0 = 0.1sec$.

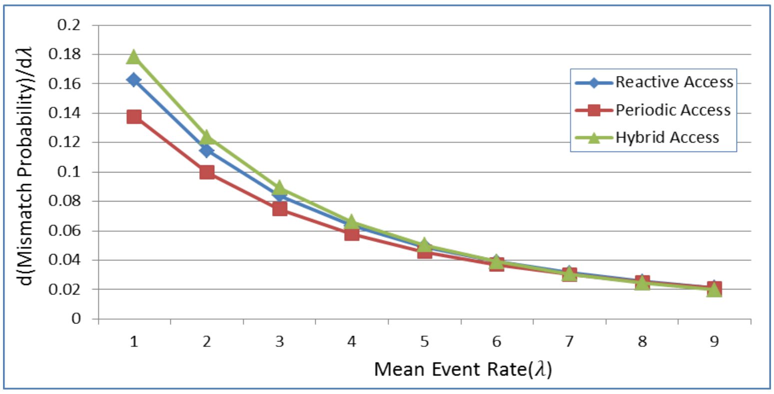


Figure 5.14: Derivatives of mismatch probability with respect to mean event rate (λ) at $\nu = 10$ and $\tau = 5$

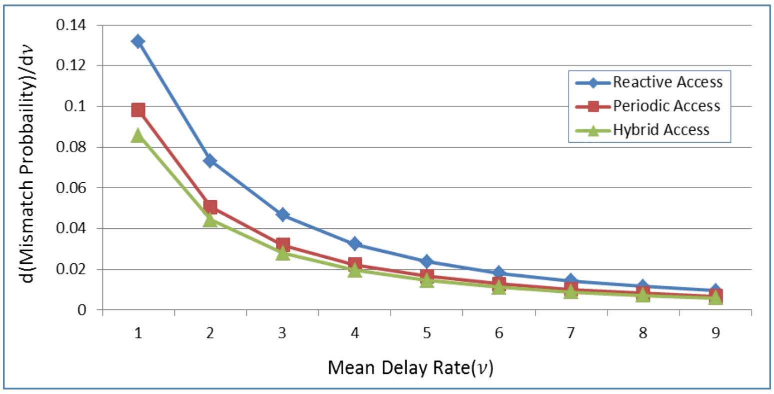


Figure 5.15: Derivatives of mismatch probability with respect to mean delay rate (ν) at $\lambda = 1.5$ and $\tau = 5$

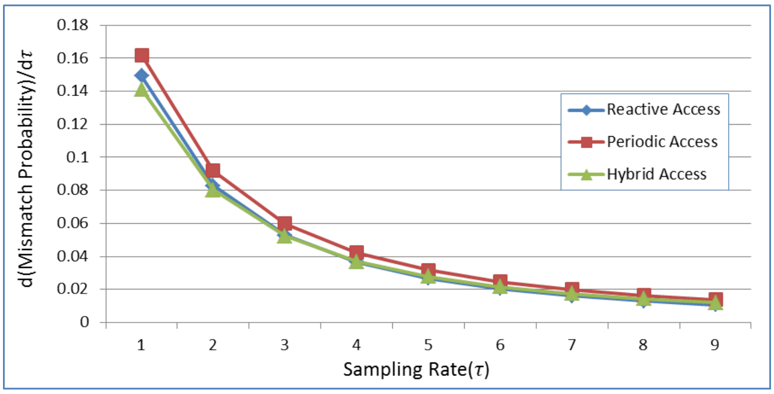


Figure 5.16: Derivatives of mismatch probability with respect to sampling rate(λ) at $\nu = 10$ and $\lambda = 1.5$

information mismatch probability is observed to be more when the sampling rate changes take place at lower values of sampling rate for all information access strategies. From this figure, it shows that the change in sampling rate has more impact on periodic access than other two. This is due to the fact that in periodic access the information is distributed periodically (at every sampling time) irrespective of whether there is demand or not.

Figure 5.17 shows the comparisons of mismatch probabilities. The comparison also gives an idea how the delay, d_0 (see Figure 5.5) introduced for checking whether there is an information update and other processing delay like filtering and detect events, does affect the mismatch probability. It is logical to see that smaller value of d_0 gives more reliable information in terms of mismatch probability. In particular, as d_0 approaches zero, the mismatch probability of hybrid access approaches to $mmPr_{ComPer}$. Not much difference is observed between $mmPr_{ComPer}$ and $mmPr_{ComRea}$. A more detailed analysis on time shifted delay is further provided in Section 5.7, where a trade-off between whether it is better to process information before sending (thereby not sending any information that are sampled) or just send information immediately without processing and checking is investigated. Figure

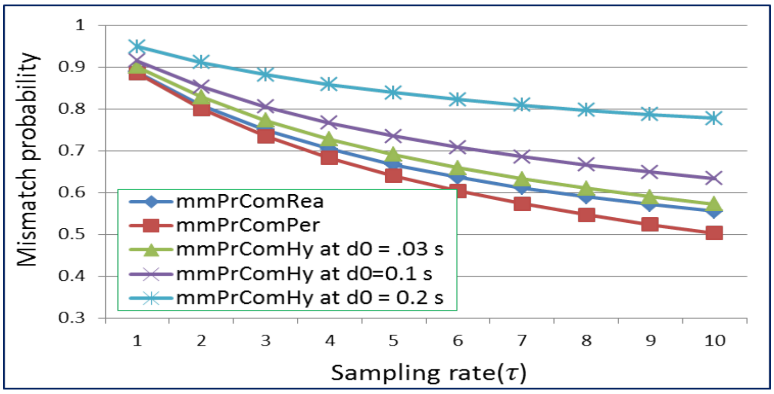


Figure 5.17: Comparisons of mismatch probability at $\nu = 10$ and $\lambda = 5.0$

5.18 shows derivative of mismatch probability with respect to the constant deterministic delay introduced for checking and processing information after it is sampled with various values of information sampling rates. As seen from Figure 5.18, mismatch probability is more sensitive to value of d_0 at high sampling rate compare to lower values of sampling rate. This can be observed by comparing with τ=1 and τ = 10 or 5. This is obvious because d_0 is added at every time the information sampled. Therefore, when the information is sampled at faster rate that many number of times the checking and processing take place.

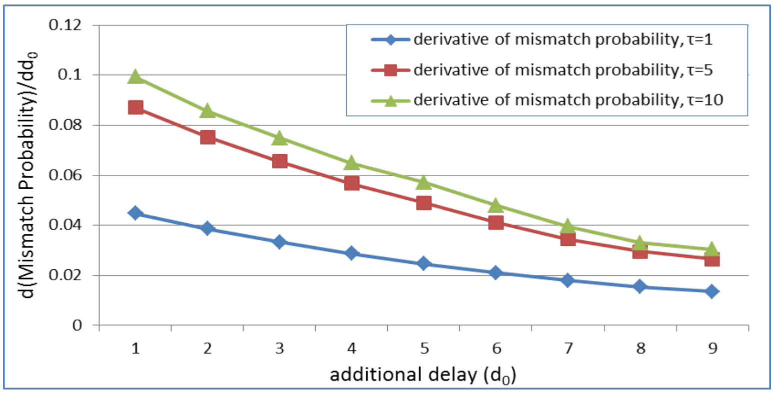


Figure 5.18: Derivatives of mismatch probability (Hybrid) with respect to constant additional delay (d_0) at $\nu = 10$ and $\lambda = 1.5$

5.6.2 Power Consumption

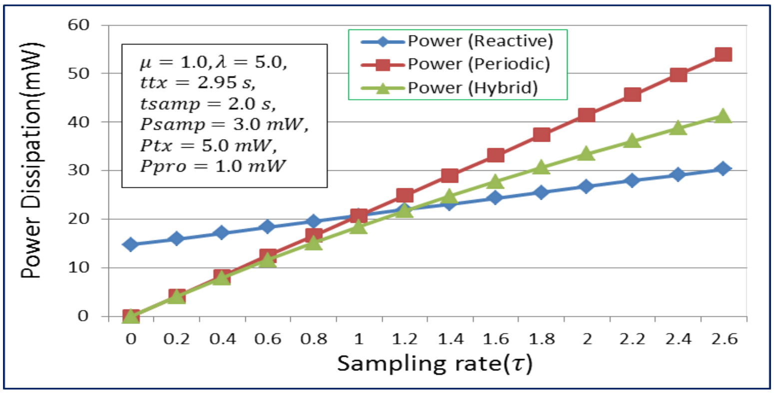


Figure 5.19: Comparisons of Power Dissipation in different access strategies.

Figure 5.19 shows the comparison of power dissipation for different access strategies. It is observed that the power dissipation is smaller in case of hybrid access compare to periodic access, as expected. The power dissipation characteristic of reactive and periodic strategy depends greatly on two parameters: namely μ, the information access rate andτ, the sampling rate. Assuming, information access rate is constant ($\mu = 1$), as shown in the Figure 5.19, it is seen that power dissipation in periodic is smaller than reactive till the value of τ, the sampling rate reaches at value of μ.

In particular, it can be seen that the power dissipation depends on the ratio of μ/τ as summarized below:

$$\begin{cases} \bar{P}_{Per} = \bar{P}_{Rea} & \text{for } \mu/\tau = 1 \\ \bar{P}_{Per} > \bar{P}_{Rea} & \text{for } \mu/\tau < 1 \\ \bar{P}_{Per} < \bar{P}_{Rea} & \text{for } \mu/\tau > 1 \end{cases} \tag{5.8a}$$

As expected, similar observation is made between power dissipation in reactive and hybrid access as shown in Figure 5.19 with the ratio between μ and τ'. As sampling rate(τ) increases, the value of update rate (τ') (see (5.7c)) also increases. When update rate(τ') becomes faster than the access rate(μ) (as per Figure 5.15, it happens at $\mu = 1.3$), the hybrid power dissipation overtakes the reactive power dissipation.

$$\begin{cases} \bar{P}_{Hybrid} = \bar{P}_{Rea} & \text{for } \mu/\tau' = 1 \\ \bar{P}_{Hybrid} > \bar{P}_{Rea} & \text{for } \mu/\tau' < 1 \\ \bar{P}_{Hybrid} < \bar{P}_{Rea} & \text{for } \mu/\tau' > 1 \end{cases} \tag{5.8b}$$

However, it should be noted that the value of τ' depends on the sampling rate(τ) and the information arrival time(λ) as seen from (5.7c). Since the update rate (τ') will never cross the value of sampling rate(τ), therefore, the power consumption in case of hybrid access strategy ($\bar{P}_{Hybrid}$) will be always less than or equal to($\bar{P}_{Per}$).

$$\begin{cases} \bar{P}_{Hybrid} = \bar{P}_{per} & \text{for } \tau'/\tau = 1 \\ \bar{P}_{Hybrid} < \bar{P}_{Per} & \text{for } \tau'/\tau < 1 \end{cases} \tag{5.8c}$$

5.7 Time Shifted Delay Distribution

In some scenarios, it is necessary to consider a time shifted delays. This is because it is inefficient for source nodes to transmit all the raw data to the sink, especially when sensed data has high redundancy or correlation. To reduce the size of packet, thereby reducing the transmission delay, several approaches are proposed. The authors in [60] and [61] argue that communication of data is one of the main energy consuming task that sensor node undertakes. They suggest conventional compression techniques to compress data without sacrificing the salient information. In [62], a Secure Grid-Sensor Integration Architecture (SGSIA) is proposed aiming to reduce the overall data transmitted to the grid system. To do so, SGSIA pre-processes data at the source nodes to filter raw data streams, so that the total amount of data generated per unit time at the source is filtered. Similar approach is also proposed in [63], where preprocessing for data to check for error or missing data in done before transmitting. In [64], compressive sensing for data gathering is investigated. This method tries to reduce sensor data traffic over the network through collecting far fewer measurements than the number of original sensor data. The compressive sensing technique enables to reconstruct a sparse signal from a small number of measurements. However, all the above mentioned preprocessing approaches introduce some delays and also need to expend energy for preprocessing. So, in order to reduce the total power consumption, the sum of power consumed for transmission and computation for preprocessing has to be less than the power consumed for transmission without preprocessing [61]. Furthermore, there should be significant reduction in transmission delay by preprocessing the data. In other words, in order to see the advantage of preprocessing data, the end-to-end delay reduction should be greater than the delay incurred in preprocessing the data. This section investigates a trade-off between preprocessing and sending information immediately as soon as it is sampled.

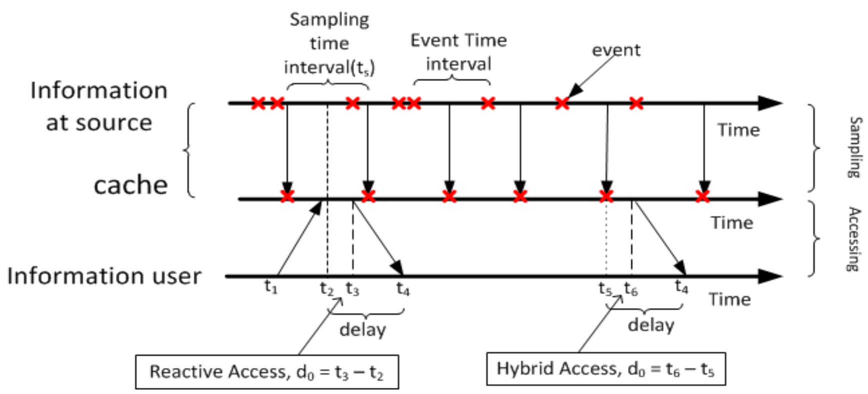


Figure 5.20: Information access with time shifted delay (d_0).

If d_0 is the constant delay (time required for every information packet compression, in case of compression scenario) then we can express total delay D as $D = d_0 + D_\nu$, where D_ν is the stochastic downstream delay; the probability density function of D can be expressed as [6]

$$f_D(t) = \begin{cases} 0 & \text{for } 0 \leq t < d_0 \\ f_{D_\nu}(t - d_0) & \text{for } t \geq d_0 \end{cases}$$

Mismatch probability for reactive access is given by [6]

$$mmPr_{Rea,sd} = 1 - \frac{1}{\mathbb{E}(E)} \int_0^\infty \left(\int_{t+d_0}^\infty \bar{F}_E(s) ds \right) f_{D_\nu}(t) dt \qquad (5.9a)$$

With an assumption of delay and event interval are all independent and identical exponentially distributed with event arrival rate λ and rate of delay ν, then (5.9a) simplifies to

$$mmPr_{Rea,sd} = 1 - \frac{\nu e^{-\lambda d_0}}{\lambda + \nu} \tag{5.9b}$$

Substituting (5.1b) and (5.9b) in (5.4a), we get the expression for combined mismatch probability with compressed information, under the given assumption as

$$mmPr_{ComRea,sd} = \frac{\lambda(\lambda + \nu + \tau) + \tau\nu[1 - e^{-\lambda d_0}]}{(\lambda + \nu)(\lambda + \tau)} \tag{5.9c}$$

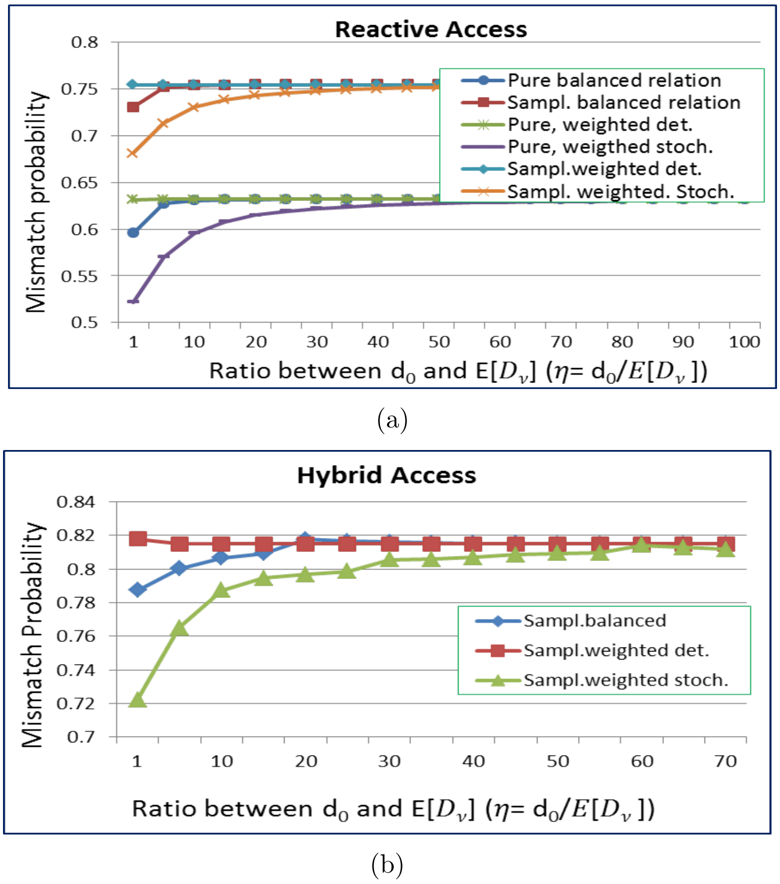
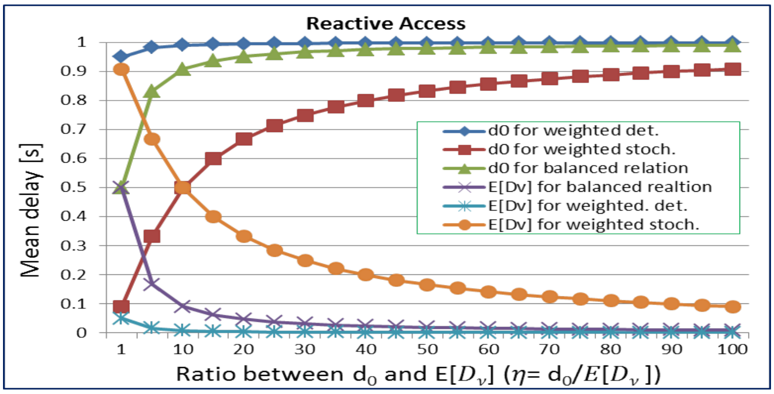


Figure 5.21: Impact of η on mismatch probability at $\lambda = \tau = 1$ and mean delay (E[D]) of 1 second (a) Reactive access and (b) Hybrid Access.

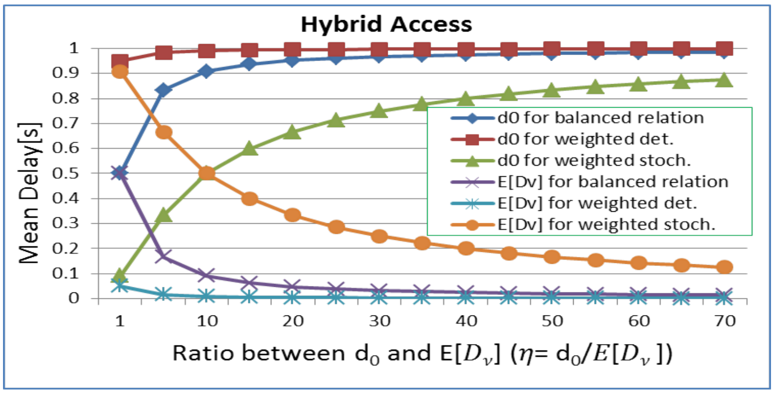
In the following the trade-off between a deterministic preprocessing delay and a stochastic network delay (caused by buffering, retransmissions etc.) that is influenced by the packet size of the data are considered. However, the total mean delay is kept constant as the study wishes to investigate whether it is better to have a deterministic behaviour or a stochastic behaviour. Hence, the following constraints are to the evaluation.

$$E[D] = d_0 + E[D_\nu] = constant$$

Finally, the ratio between d_0 and $E[D_\nu]$ as $\eta = d_0/E[D_\nu]$ are used and



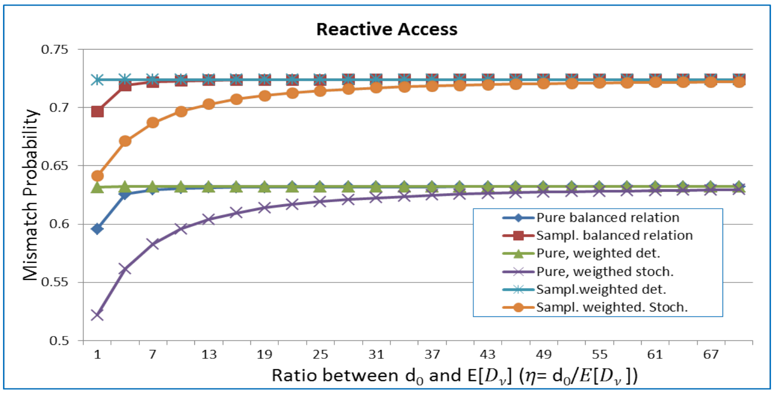
(a)



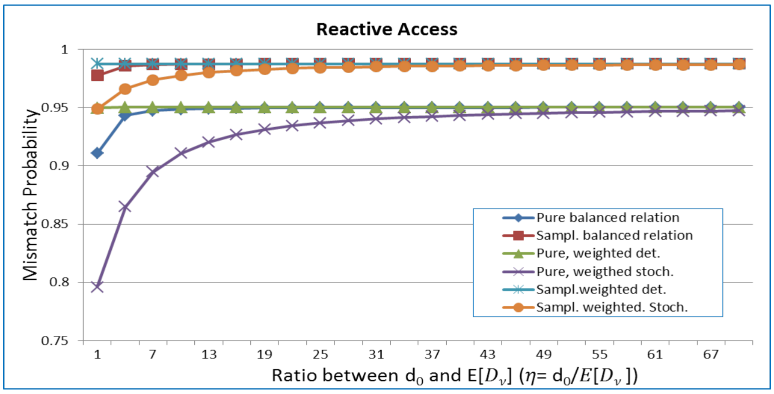
(b)

Figure 5.22: Impact of η on mean delay at $\lambda = \tau = 1$ and mean delay (E[D]) of 1 second (a) Reactive access and (b) Hybrid Access.

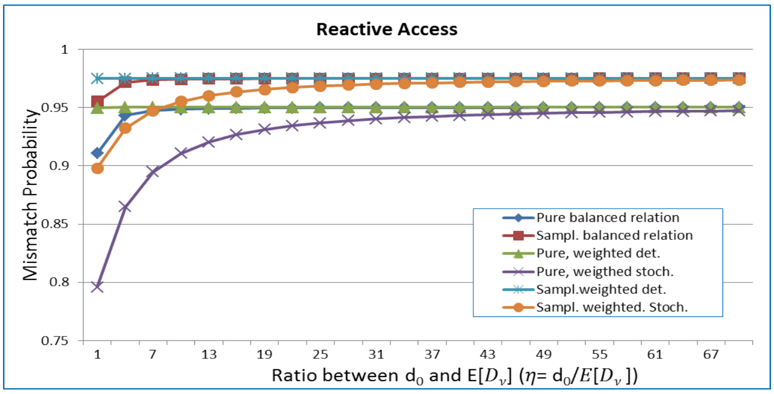
analyze its impact on mismatch probability. The weight (k) between deterministic delay and stochastic delay (the rate) is set to 10 for illustration purposes only. In both reactive and hybrid access strategies, it is observed that, increasing the ratio (η), that is by setting d_0 (the deterministic delay) increases the mismatch probability to some point basically showing it is better to have the stochastic network delay impacting the update stream. These observations are reflected in Figure 5.21(a) for reactive case and, Figure 5.21(b) for hybrid case. The corresponding delay plots are shown in



(a)



(b)



(c)

Figure 5.23: Impact of η on mismatch probability at mean delay (E[D]) of 1 second for Reactive Access (a) at $\lambda = 1, \tau = 3$ (b) at $\lambda = 3, \tau = 1$ (c) at $\lambda = \tau = 3$.

Figure 5.22(a) and Figure 5.22(b) for respective access strategies. Obviously, that also supports that it is then better to have algorithms running which only reduces traffic lightly and uses less time on preprocessing rather than having a heavy preprocessing algorithm that drastically reduces a stochastic network delay. In any case, the conclusions are also limited by the assumption that a preprocessing algorithm is producing the constant delay d_0 and the network is producing the stochastic exponential delay with rate ν.

For reactive access strategy further analysis have been done to see how different values chosen for event rate (λ) and sampling rate (τ) changes the impact on , thereby on mismatch probability. Comparing Figure 5.21(a) and Figure 5.23(a) when sampling rate (τ) increases the mismatch probability in the case of combined reactive ($mmPr_{ComRea,sd}$) decreases while there is no effect in case of pure reactive ($mmPr_{Rea,sd}$), irrespective of whether it is balanced relation or weighted. This is because the sampling rate has no influence on mismatch probability under pure reactive scenario. But, when the event rate (λ) is increased, its impact on mismatch probability (increase in mismatch probability) is observed in both pure reactive and sampled reactive cases. This can be observed by comparing Figure 5.23(b) and Figure 5.21(a).

Table 5.1: Shows how d_0 values are computed for different cases, at k = 10 and $\eta = d_0/E[D_\nu]$.

Cases	d_0	**Mismatch probability equation**
Pure balanced relation	$\frac{E[D]}{1+\frac{1}{\eta}}$	Equation 5.9b
Sample balanced relation	$\frac{E[D]}{1+\frac{1}{\eta}}$	Equation 5.9c
Pure weighted deterministic	$E[D] - \frac{1}{\nu}$	Equation 5.9b
Pure weighted stochastic	$\frac{E[D]}{1+\frac{k}{\eta}}$	Equation 5.9b
Sample weighted deterministic	$E[D] - \frac{1}{\nu}$	Equation 5.9c
Sample weighted stochastic	$\frac{E[D]}{1+\frac{k}{\eta}}$	Equation 5.9c

Figure 5.23(c) is plotted to compare which parameter, event rate (λ) or sampling rate (τ) has greater influence on mismatch probability by changing both the values by the same quantity. Although, λ and τ have opposite

impact, meaning event rate (λ) has negative impact while sampling rate(τ) has positive impact on mismatch probability, the impact of event rate (λ) seems to have more than the impact of sampling rate (τ). It is therefore, an increase in mismatch probability is observed. This can be easily noticed by comparing the mismatch probability for sampled reactive case from Figure 5.21(a) and Figure 5.23(c).

5.8 Conclusions

The chapter started by exploring information reliability model and power consumption model in wireless sensor network with respect to accessing information. Basically, three types of information access strategies were explored. For each access strategies, mismatch probability models and power consumption models were presented. With these analytical models, the investigation on how different information access strategies affect the information reliability in terms of mismatch probability and power consumption of information source were conducted. In particular, the following topics were explored in this chapter:

The chapter presented a detailed study on the effect of information access strategies namely: reactive, periodic and hybrid on power consumption and reliability (mismatch probability) on wireless sensor network.

- As expected, the power consumption in hybrid access was observed to be always less as compared to the periodic access. The power consumption between reactive and periodic access depends on the ratio between the access rate and the sampling rate. Similarly, the power consumption between reactive and hybrid access depends on the ratio between the access rate and the update rate.

- Information mismatch probability is more sensitive to all these three parameters; sampling rate (τ), event rate (λ) and delay rate (ν) at lower values. Therefore, at lower rates, small change in these parameters will have greater impact on mismatch probability.

- An event rate(λ) was observed to have more influence on mismatch probability than the sampling rate(τ)

- The mismatch probability with hybrid access was observed always higher than the other two access strategies because of an additional delay (d_0) introduced for checking update and preprocessing. Information mismatch probability is more sensitive to d_0 at higher sampling rates.

Both power consumption and information mismatch probability are influenced by rate at which information is sampled. Sampling rate has positive impact on mismatch probability whereas it has negative impact on power consumption. Therefore, it is possible to find a trade-off between these two parameters to adapt information access as per the requirement.

The information mismatch probability under different access strategy is estimated using the framework proposed in Chapter 4 and the values are compared with the mismatch probability values obtained by using analytical models.

- Information mismatch probability under three information access strategies; reactive, periodic and hybrid were estimated with the help of the proposed framework in Chapter 4. As expected it was observed that the information mismatch probability under hybrid access strategy were greater than the other two.
- The information mismatch probability comparison between reactive and periodic access strategies shows that the reactive access suffers more in terms of information reliability.
- The information mismatch probability by keeping sampling rate at constant is also estimated and comparison with exponential sampling rate is made.
- A comparison of information mismatch probability estimation between exponential and Erlang delay distribution is provided. The difference in mismatch probability value between these two distributions is found to be about 4% in periodic access strategy case and about 2% in case of reactive and hybrid access strategies.
- A comparison of information mismatch probability estimation between exponential, Erlang delay distribution without and with retransmission consideration is provided for all the three access strategies.

The impact of adding deterministic processing delay on information reliability and trade-off between compression time and reduction in stochastic network delay were presented.

- It was observed that, in order to have better information reliability, it is better to have compression algorithm which only reduces traffic lightly and uses less time on compression rather than having heavy compression algorithm which drastically reduces stochastic network delay

To summarize, the information mismatch probability under all the three access strategies under consideration is impacted by the rate at which the information is generated at source, the rate at which information is sampled and the rate of delay between information source and the information user. Therefore, it is understood that the information reliability in terms of mismatch probability can be maintained at required level by varying sampling rate. In addition, it was observed that the power consumption of the information source is impacted by the information sampling rate. If we compare power consumption and mismatch probability and see how they are affected by information sampling rate, it will be observed that sampling rate as positive impact on mismatch probability whereas it has negative impact on power consumption. This means that if we increase the sampling rate, the information reliability will be improved whereas with the increase in sampling rate, the information source will have to spend more power.

Therefore, an information access mechanism which can adapt as per the application requirement can be possible, thereby finding a trade-off between information reliability and power consumption. This adaptive information access mechanism by varying information sampling rate is presented in Chapter 6.

Chapter 6

An Adaptive Information Access Mechanism

Wireless sensor network is being used for different types of applications today. The requirements are not same for all the applications. As an example, information reliability may be essential for one application but for another, power consumption may be of greater concern than information reliability. In Chapter 5, it was observed that these two requirement parameters; information reliability and power consumption can be varied by varying information sampling rate. If a single wireless sensor network is to support more than one application, there should be an accessing management system that adapt as per the requirement of the application. The information access mechanism should be such that the resources are not either over utilized or underutilized. Therefore, this chapter proposes adaptive information access mechanism based on power consumption and information reliability. The framework proposed in Chapter 4 can be used to adapt information access as per these two parameters. Further, the chapter also points out a trade-off point between power consumption and information reliability beyond which adaptation gains very minimal in information reliability but has to spend huge amount of power.

6.1 Motivation

Be it wireless network or any other traditional network, satisfying applications requirement with minimum utilization of resources is utmost important to harvest better benefits both in terms of investment and performance. In particular, when wireless sensor network have the ability to supporting various applications, it would be cost effective and yields better overall performance if it could satisfy multiple applications requirements with single network deployment. Further, the requirements for applications are also influence by the availability of resources. As an example, if the sensor nodes have direct power supply instead of battery powered, then there is no need to worry about the lifetime of the nodes. Contrary to this, if the information area is such that, replacing or recharging of battery is not possible, it is important that the energy is not wasted by making unnecessary access to the information source. This section brings out a scenario and argues why adaptive information access mechanism is beneficial in such a scenario.

Consider that a wireless sensor network is set up in the Himalayan region to monitor and study the temperature variation over the years. For this type of application, it is more crucial to have longer lifetime than the more reliable and up-to-date information. So, information access frequency to such type of application should be adapted such that the energy consumption is always maintained at certain required level. But this will have an impact on the information reliability, meaning as frequency of accessing information is controlled to maintain required power consumption level, the information may not be as fresh as in the case where there is no control over frequency of accessing the information.

On the other hand, consider a scenario where wireless sensor network is installed inside a building for safety monitoring purposes. This type of application does not need to worry about the power consumption as the nodes can be directly connected to the power supply or the battery can be replaced or recharged easily. But, for such an application, timeliness and reliable information is very much crucial. So, access to information can be adapted so that required level of information reliability is maintained at all the time.

Looking at the above two situations, it is therefore, interested in investigating how this type of adaptive access can be achieved. Furthermore, an investigation on trade-off between power consumption and information reliability will provide support in designing adaptive information access mechanism.

6.2 Related Works: Adapting Sampling

Since information quality and energy consumption are important performance parameters for sensor network, a number of research articles can be found in the literature. We briefly review few of these articles which are mainly focused on improving information quality and low energy consumption with respect to sampling rate.

In [65], a reward and punishing based cooperative adaptive sampling is proposed. This method considers cooperation among nodes to find a balance between energy consumption and the information quality. It uses clustered based sensor network system, where a node which detects frequent phenomenal change increases the sampling rate while the others decrease their sampling rates. In this way, the approach is able to keep the overall network data quality high and minimize the energy consumption by reducing the sampling rate of other nodes.

A two-step approach of adapting sampling called backcasting is proposed in [66] aiming at maintaining high information quality with reduced energy consumption. In the first step, called the preview step, an initial estimate of the environment is formed by using subset of nodes. This subset of sensor nodes communicate with the fusion center with this estimated information which indicates that some sensor may not be required in order to achieve a required level of accuracy. So the fusion center backcasts information based on the estimate received and selectively activates additional sensor nodes in the next step, called refinement step. Here, the term adaptive sampling is used how many sensor nodes are activated simultaneously rather than how fast or slow the information is transmitted. In this approach, adaptive sampling helps in reducing energy consumption by activating fractions of the available nodes.

An adaptive sampling scheme that responds to the characteristics of the streaming data is developed in [67]. Their approach is based on predictions of future values of a data stream based on Kalman filter. It is based on centralized approach where the activities of the nodes are monitored. The bandwidth allocation is done automatically as per the activity of the nodes. The nodes which show an increase in activity is allocated bandwidth which gives shorter time interval (shorter sampling interval) between successive measurements whereas the nodes which do not show an increase in activity are made to have longer time interval (higher sampling interval).

However, most of the previous works were done based on application specific sensor networks. When single sensor network has to support more than one applications, the challenges arise as each individual application has its own set of requirements that needs to be given due importance. As we have described in Chapter 2, the preference for requirements differ from one application to another. In this chapter, an adaptive information access mechanism which can satisfy different applications without much compromising the overall performance of the system is proposed and presented.

6.3 Adaptive Access Mechanisms

As highlighted in Section 6.1, an adaptive information access will be useful when there are multiple applications with different requirements. This chapter mainly investigates two requirement parameters, namely; information mismatch probability and power consumption. The proposed adaptive information access mechanism based on the analytical models provided in Chapter 5. The adaptation is basically done by varying information sampling rate.

6.3.1 Adapting Information Sampling Rate

Referring back to Figure 5.17 in Chapter 5, it is observed that the information reliability in terms of mismatch probability can be varied by varying the rate of information sampling. Therefore, it is possible to maintain mismatch probability at certain required level by adapting sampling rate. Consider a situation, where an application (say safety monitoring) needs to maintain an information reliability of 0.80 (mismatch probability = 0.20) at all time. Although, it is possible to make similar analysis with all the three access strategies, we have mainly focused on reactive access as it is easier and simple to show the concept. However, we have made a comparative analysis of all the three access strategies, which equally show how adaptive information access is possible with all the access strategies.

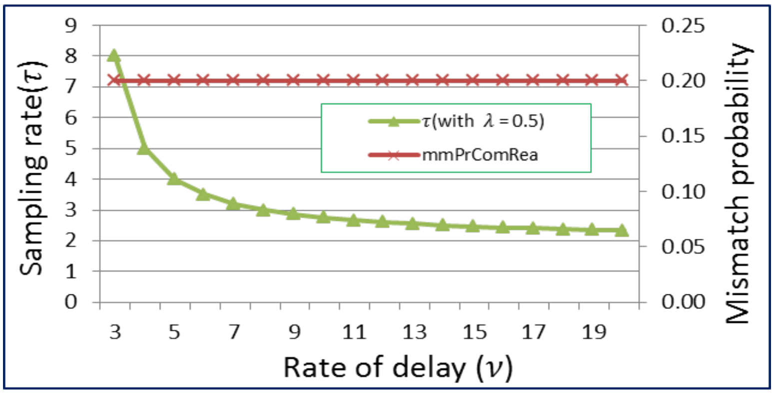
Reactive Access Strategy

Revisiting (5.4b) and solving it, sampling rate (τ) can be expressed as

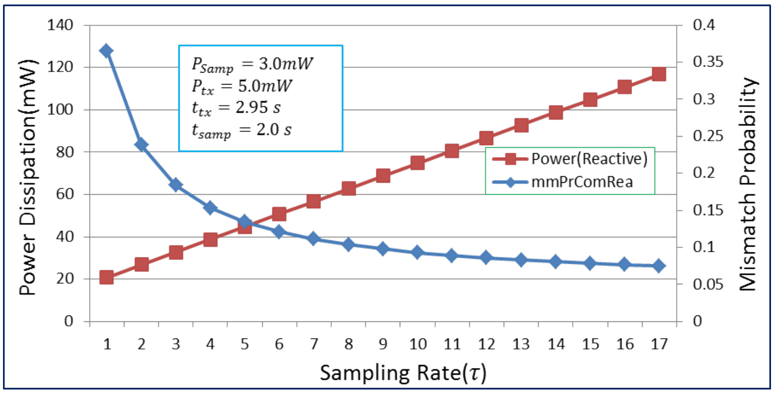
$$\tau = \frac{\lambda(\lambda+\nu)(1-mmPr_{ComRea})}{mmPr_{ComRea}(\lambda+\nu)-\lambda} \tag{6.1}$$

Also solving (5.5c) for sampling rate (τ), it can be expressed in terms of power consumption as

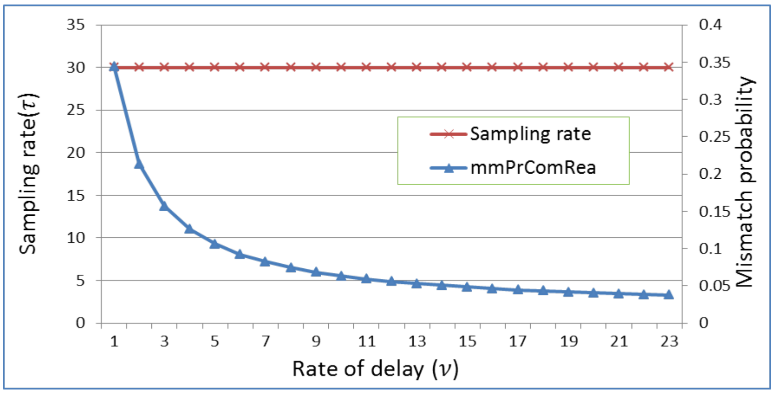
$$\tau = \frac{\bar{P}_{Rea} - P_{tx}t_{tx}\mu}{P_{samp}t_{samp}} \tag{6.2}$$



(a)



(b)



(c)

Figure 6.1: Reactive Access (a). Adaptation of sampling rate to maintain mismatch probability of 0.20, (b).Mismatch probability and power dissipation variation with sampling rate at $\lambda = 0.5$ and $\nu = 10$ and $\mu = 1$ (c). Mismatch probability variation keeping τ constant and $\lambda = 0.5$.

This requirement of 0.80 (mismatch probability of 0.20) information reliability can be satisfied by adapting the sampling rate assuming the event rate is constant (although it is not always true) as shown in Figure 6.1(a). However, this comes at a cost of increase in power consumption, because when information is sampled with faster rate, it draws more power. This can be seen from Figure 6.1(b). Therefore, we need to find out a trade-off point between power consumption and mismatch probability. From Equation 6.1, one can see that as the denominator term approaches zero, (at $mmPr_{ComRea} = \frac{\lambda}{(\lambda+\nu)} = mmPr_{Rea}$) one would need to increase sampling frequency to infinite and gain only little in terms of information reliability (confirmed by the asymptote in Figure 6.1(b). This is thereby the trade-off point where instead of trying to achieve required information reliability, it is economical to reduce power consumption.

Figure 6.1(c) shows how the mismatch probability behaves with network delay if the sampling rate is kept constant (at some more or less arbitrarily chosen value). This illustrates that the reliability will vary highly depending on link conditions or end-to-end delay (network conditions) if sampling frequency is not adapted.

Based on Figure 6.1(b), the derivatives of the power consumption and mismatch probability with respect to sampling rate (τ) can be computed in order to find how sensitive is power consumption and information mismatch probability(mmPr) with change in sampling rate.

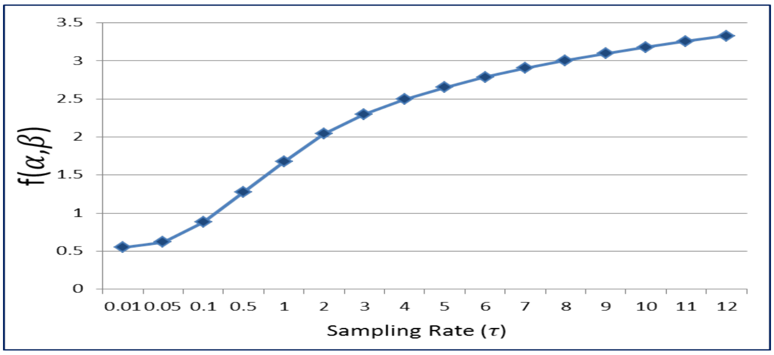


Figure 6.2: Sensitivity plot: Showing how sensitive is power consumption and mismatch probability to the change in sampling rate(τ) with Y-axis plotted with log scale.

$$\alpha = \frac{dmmPr}{d\tau} \qquad \text{and} \qquad \beta = \frac{dPower}{d\tau}$$

Then, $f(\alpha, \beta) = \beta/\alpha$ will show how sensitive is the power consumption and mismatch probability to the change in sampling rate (τ). Figure 6.2 shows the $f(\alpha, \beta)$ plotted against sampling rate (τ). From Figure 6.2, it is observed that the power consumption is more sensitive to the change in sampling rate at higher values of sampling rate, while mismatch probability is more sensitive to change in sampling rate at lower values of sampling rate.Therefore at higher sampling rates, one should worry about power, and at lower sampling rates, one should worry about mismatch probability (as a rule of thumb, i.e. there are as seen no sharp line between the two parameters. At later section, an adaptive information access algorithms are provided where, switching from one algorithm to another is suggested based on the result shown in Figure 6.2.

Comparison between different Access Strategies

As adaptive information access by varying sampling rate is applicable to all the three access strategies, a comparative study is made here in order to find out which access strategy is better among the three. Figure 6.3 below shows the comparative study of three access strategies with respect to adaptive information access by varying sampling rate. The target mismatch probability to be maintained is set at 0.20. If we compare between periodic

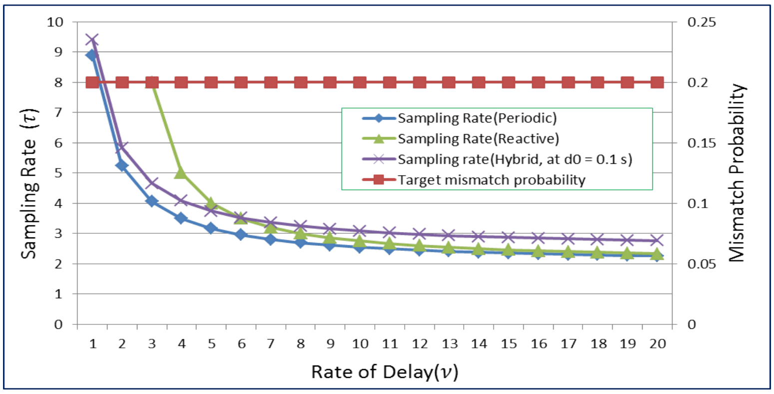


Figure 6.3: Comparison between the three access strategies in adaptive information access by varying sampling rate at $\lambda = 0.5$.

and hybrid access strategies in Figure 6.3, we see that the sampling rate for hybrid should be always greater in order to maintain same level of information reliability in terms of mismatch probability. This is due to the fact that, in hybrid access, there is an additional delay of d_0. So, if the application needs to maintain high information reliability and does not care much about the power consumption, it is better to use periodic access while applying this type of adaptive information access. Among the three, it is observed that the periodic access out performs than the other two.

The analysis in case of periodic access and hybrid access, are done numerically as the mismatch probability Equation (5.1a) and Equation (5.6), which give expressions for mismatch probability, are not easily solvable mathematically, to get expressions for sampling rate (τ).

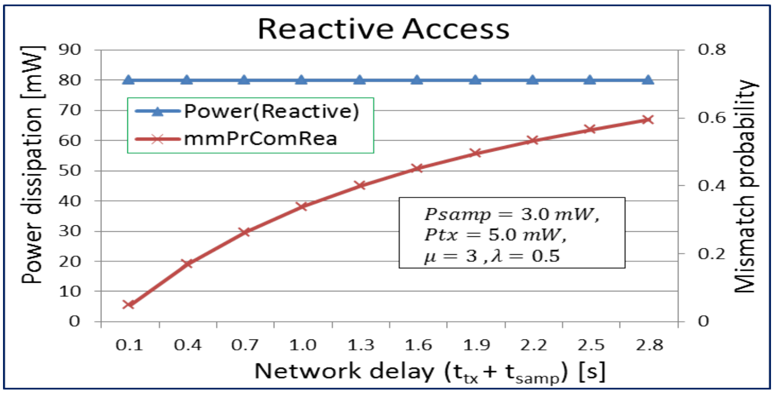
In Figure 6.3, in the case of Sampling Rate (Reactive), it is observed that the plot ends at $\nu = 3.0$ and does not go beyond unlike other two plots. The reason for this observation is, at this setting ($\lambda = 0.5$), the mismatch probability does not depend on ν when ν is less than 3. This also shows that at this point the mismatch probability is influenced by sampling process only.

6.3.2 Trade-off between Reliability and Power Consumption

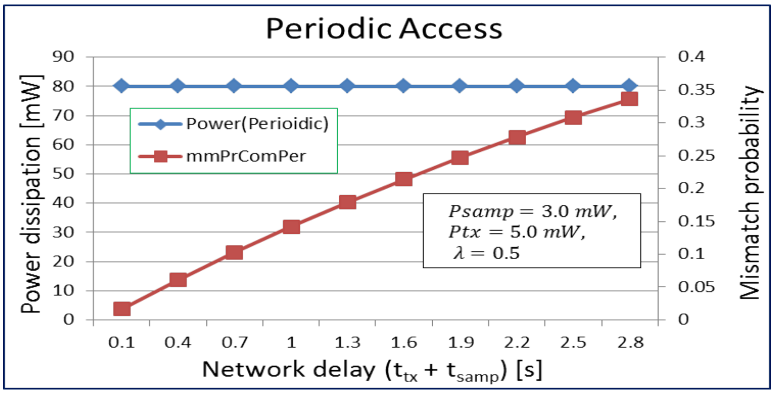
As we know from our state-of-the-art study in Chapter 2 the requirements of wireless sensor network would be different as per the purpose and application type. As an example, sensor network application which monitors environment data collection over a time will not be concerned much about reliability rather it is important to extend lifetime of the sensor nodes, whereas sensor network which is used for safety monitoring purposes (e.g. fire detection in building) needs to have high reliability. Therefore, depending on the applications and their purposes, we can have a trade-off between power consumption and information reliability. So, in this section, we analyze the trade-off mechanism with the three access strategies and study the differences on the amount of trade-off needs to be made (power consumption and mismatch probability). We try to keep our parameters same for all access strategies.

For reactive access strategy, Equations (6.1) and Equation (6.2) can be used to analyze how trade-off between power consumption and information reliability can be made in order to fulfil applications requirement. In case of periodic and hybrid access strategies, numerical integration is performed to get the values of sampling rate (τ), as done in Section 6.3.1 in order to analyze the trade-off mechanisms.

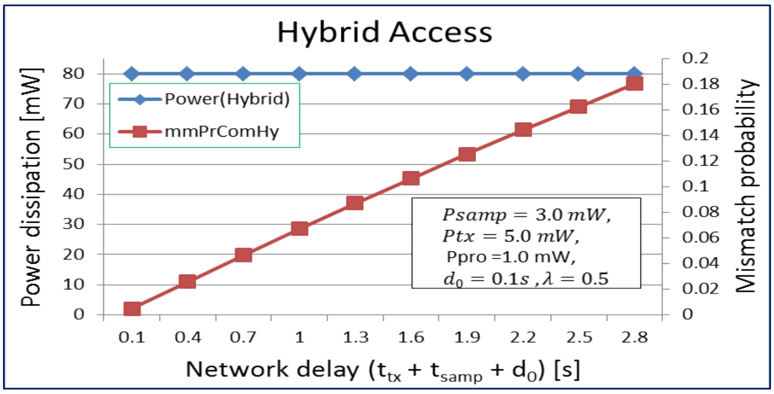
Suppose, if an application has to maintain a power dissipation of 80mW for all time, then we need to compromise on the information reliability (mismatch probability). This means that we need to increase the mismatch probability by lowering the sampling rate. Figure 6.4(a), 6.4(b) and 6.4(c) show how mismatch probability is varied in order to maintain power dissipation of 80mW, considering network delay ($\frac{1}{\nu}$) is the sum of sampling time (t_{samp}) and transmission time (t_{tx}) (for hybrid constant d_0 is also added). This



(a)



(b)



(c)

Figure 6.4: Maintaining constant power consumption by varying mismatch probability (a) Reactive Access, (b) Periodic Access and (c) Hybrid Access.

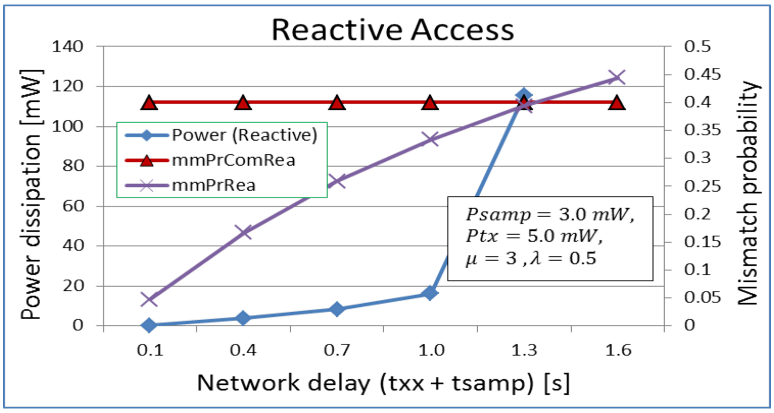
type of adaptation could be used for WSN application like environment data collection where reliability is not much critical.

Figure 6.5(a), 6.5(b) and 6.5(c) show how mismatch probability of 0.40 is maintained at the cost of power consumption. The power dissipation is varied by varying sampling rate in order to maintain the required mismatch probability. This type of adaptation would be useful for an application (e.g. safety monitoring), where reliability cannot be compromised but the power consumption less an issue.

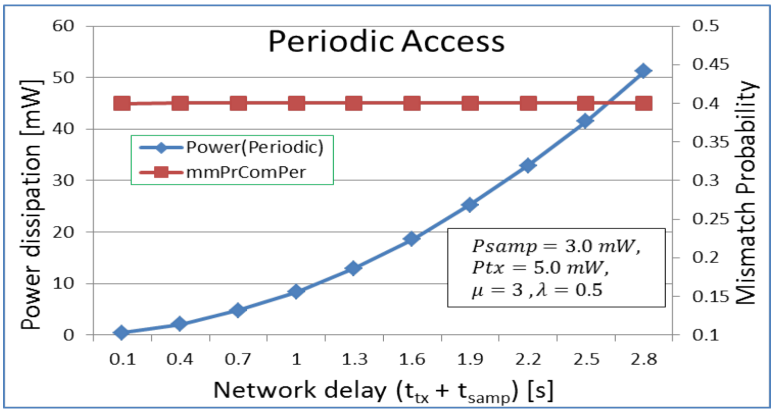
If we compare Figure 6.4(a), 6.4(b) and 6.4(c), it is observed that among the three access strategies, hybrid access shows that less trade-off in information reliability need to be made, followed by periodic access and then reactive access in order to maintain same required power consumption level. Similar behaviours are observed in terms of trade-off on power consumption.

Similarly, how the required information reliability in terms of mismatch probability is maintained by adapting power consumption are shown in Figure 6.5(a), 6.5(b) and 6.5(c) for different access strategies. It shows that more energy needs to be spent with reactive access, followed by periodic and then hybrid access. This analysis shows that hybrid access strategy outperforms in making trade-off between power consumption and information reliability compare to periodic access strategy, as expected. But the power consumption of reactive access will depend on how fast the information is being accessed as explained in Chapter 5. When access rate (μ) increases the power consumption under reactive case will also increase.

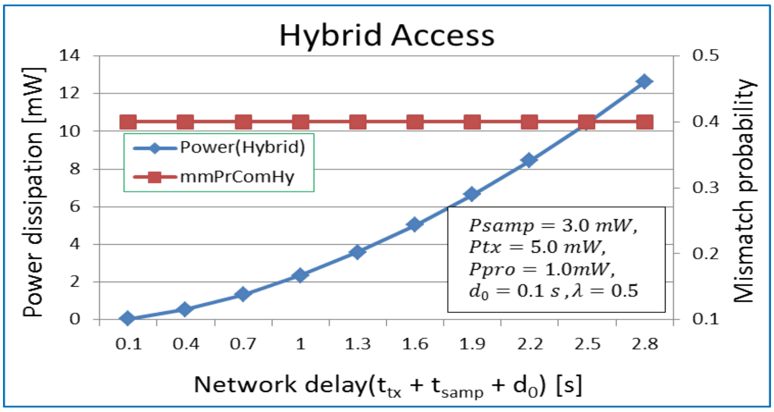
In Figure 6.5(a), it is observed that when network delay reaches at 1.3 seconds, the combined mismatch probability ($mmPr_{ComRea}$) becomes equal to mismatch probability of pure reactive ($mmPr_{Rea}$). Beyond this point the mismatch probability of pure reactive access, given by Equation 5.4c, becomes higher than the target required mismatch probability. This is the point where mismatch probability no more depends on sampling rate and purely decided by the stochastic network delay and the rate at which information changes at source. This shows that mismatch probability cannot be adapted below mismatch probability ($mmPr_{Rea}$) given by pure reactive access. This is due to the fact that ($mmPr_{Rea}$) does not depend on sampling rate (τ).



(a)



(b)



(c)

Figure 6.5: Maintaining the required reliability by varying power consumption (a) Reactive Access, (b) Periodic Access and (c) Hybrid Access.

6.4 Adaptive Information Access Mechanism Algorithms

In order to implement adaptive information access mechanism described in Section 6.3, an algorithm can be designed and implemented using the proposed framework described in Chapter 4. The pseudo codes of the algorithms are shown in Algorithm 6.1 (maintaining power consumption) and Algorithm 6.2 (maintaining information reliability). Referring back to Figure 4.1, the algorithm can be executed in the Event Monitoring Unit (EMU) and particularly within NM-Unit. As per the requirement level, e.g. level of information reliability required maintaining or maximum power consumption limit permissible, the sampling rate can be adapted.

The algorithms for adaptive information access are shown for reactive information access only. With similar approach, the adaptive information access algorithms can be implemented for periodic and hybrid access strategies. Choosing of an appropriate information source mentioned in step 3 of Algorithm 6.1 and Algorithm 6.2 can be done as per the Algorithm 7.1 described in Chapter 7. These two algorithms are based on trade-off between power consumption and information reliability. To maintain power consumption, the algorithm compromises on information reliability. Similarly, to maintain information reliability, the algorithm compromises on power consumption. The trade-off between these two parameters is achieved by varying sampling rate.

Algorithm 6.1 Adaptive Information Access Mechanism to Maintain Required Power Consumption Limit

1: Initiate the request for information
2: Check for the registered information sources
3: Choose an appropriate information source
4: Get the maximum power consumption limit ($\bar{P}_{Rea}$)
5: Get mean delay rate (ν)
6: Compute sampling rate as per the power consumption level as

$$\tau = \frac{\bar{P}_{Rea} - P_{tx} t_{tx} \mu}{P_{samp} t_{samp}}$$

7: Adapt sampling rate (τ) as computed above

Algorithm 6.2 Adaptive Information Access Mechanism to Maintain Information Reliability

1: Initiate the request for information
2: Check for the registered information sources
3: Choose an appropriate information source
4: Get the level of information reliability required ($mmPr_{ComRea}$)
5: Get the information arrival rate(λ)
6: Get mean delay rate (ν)
7: Compute sampling rate as per required $mmPr_{ComRea}$ level

$$\tau = \frac{\lambda(\lambda + \nu)(1 - mmPr_{ComRea})}{mmPr_{ComRea}(\lambda + \nu) - \lambda}$$

8: Adapt sampling rate (τ) as computed above

6.5 Conclusions

This chapter mainly focused on an adaptive information access mechanisms and the main points are summarized below:

- An adaptive information access mechanism by varying sampling rate.
  - It is found that an application demanding a certain level of information reliability could be satisfied by adjusting the rate at which the information is sampled.
  - It was learnt that sampling rate cannot keep on increasing, expecting improvement in information reliability. This is because after certain value, the gain in information reliability is so minimal as compare to the increase in power consumption due to increase in sampling rate, thereby giving a trade-off point where instead of achieving information reliability, it is economical to reduce power consumption.
- A trade-off mechanism between power consumption and information reliability.
  - It is learned that an application which requires certain level of information reliability could be satisfied by increasing sampling rate, which increases the power consumption.
  - Similarly, certain applications like environmental data collection over a period of time, which may not require very high reliable

information, the lifetime of a sensor node could be extended by reducing the sampling rate, thereby reducing the power consumption.

  – As expected, for reactive access strategy, it is observed that the combined mismatch probability requirement level cannot be lower than the mismatch probability given by the pure reactive access. This is the point where mismatch probability no more depends on sampling rate and purely decided by the stochastic network delay and the rate at which information is changed.

- Algorithms for implementing adaptive information access mechanisms by using the proposed framework are described.

Chapter 7

Dynamic Information Source Selection Algorithm

Given a situation, where a particular piece of information can be obtained from multiple sources, provides information user a choice to use the information source that best suits his/her requirement like what level of information reliability or how much information source node is ready to spend power. However, when these information sources have different characteristics features; like different event rates, different technologies, having possibility of configuring with different sampling rate, located in different locations- which results in different access delays because of different routes, there is a need for an algorithm which dynamically selects and configures access to information source based on information dynamics. Therefore, this chapter presents a dynamic information source selection algorithm based on information mismatch probability metric. The algorithm is design to facilitate in selecting the information source which gives the best information reliability in terms of mismatch probability. The performance analysis of the algorithm is evaluated by creating different scenarios which represent different network conditions.

7.1 Motivation

Due to high demand and high competition in electronics manufacturing companies, various types of sensors, which can measure same phenomenon, are manufactured by different companies. As an example, various types of temperature sensors are manufactured by Microchip Technology Inc. [68], OMEGA Engineering Inc. [69] and Applied Sensor Technologies [70], to name a few. As temperature sensor is commonly used for various purposes, this group of sensors is considered as an example for the sack of easy understanding in this chapter. There are different types of sensors which can measure same phenomenon, i.e. temperature. A few popular temperature sensors commonly used are: Thermocouple, Platinum Resistive Temperature Device (RTD), Thermistor, and other integrated Silicon Based Sensors [71]. Each of these sensors is designed for specific temperature ranges and environment conditions with different costs, level of sensitivity, and response time among others (refer Table 7.1). Sensitivity of a sensor node is defined as a rate of change in output as the input varies [71, 72]. As observed from Table 7.1 the sensitivity of different types of temperature sensors is different. Since sensitivity is the rate of change of output signal, depending on the input signal, it is directly related to the event arrival rate.

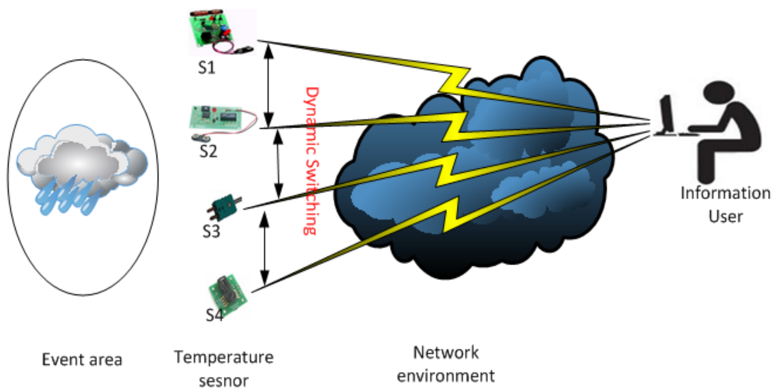


Figure 7.1: An abstract view of information accessing with source selection possibility.

Consider a scenario, where an organization wants to monitor the temperature of a certain locality and due to some reasons, like costs, robustness towards failure of one technology (not to rely on specific technology), existing sensor nodes (already available) and different threshold values, the organization has deployed four types of temperature sensors, which are mentioned in the Table 7.1 [68, 72]. These temperature sensors have different characteristics which lead to different behaviour in sensing temperature information from the surroundings. This means, although, they all measure same temperature information but the rate at which they sense temperature from surrounding is different, resulting a scenario where different sources supplying the same information. Or all information sources give equally useful information but the reliability in terms of information mismatch probability differs as they are placed in different locations. In this type of scenario, when the same information (temperature, in this case) can be obtained from different information sources, the users are provided with choice to select the best suitable information source as per their preferences and requirement. Moreover, having multiple information sources would provide information user with continuous supply of information even if the current information source becomes unavailable. However, in order to take advantage of having multiple information sources, there should be a system put in place where information source switching is dynamically done as and when it is required. Therefore, in this chapter, the framework which was proposed and described in Chapter 4 is used to implement a dynamic information source switching and selection. An abstract view of information accessing from multiple information sources is shown in Figure 7.1.

From previous chapters, particularly from Chapter 5, it was learned that event arrival rate has an impact on the information mismatch probability and it has direct relationship, meaning that higher the event arrival rate, the more the information mismatch occurs. But, it is not necessarily true that the information source which has lowest event arrival rate always gives lowest mismatch probability. This is because, mismatch probability does not only depend on event rate, but it also depends on other parameters like network delay, access configuration and sampling rate. Therefore, based on the algorithm, this chapter also explores and investigates how these information dynamics which have impact on information mismatch probability, do influence the information source selection decision. In particular, a dynamic and reliable information source selection algorithm based on information mismatch probability and power consumption is presented in this chapter.

7.2 Proposed Algorithm

The proposed algorithm is based on the comprehensive study done in Chapter 5 on the topic information mismatch probability and power consumption. This algorithm for dynamic information source selection chooses information source based on the information mismatch probability. The mismatch probability depends on dynamic information elements; event arrival

Table 7.1: Comparisons of common temperature sensors specifications.

Characteristics	**Platinum RTD**	**Thermistor**	**Thermocouple**	**Integrated Silicon**
Temperature Range	-200°C to 500°C	-40°C to 260°C	-270°C to 1750°C	-55°C to 150°C
Changing Parameter	Resistance	Resistance	Voltage	Voltage/ Current
Accuracy	±0.5°C	±0.1°C	± 0.5°C	± 1°C
Sensitivity (ρ)	2mV/°C	40mV/°C	0.05mV/°C	-2mV/°C
Response Time	2-5 sec	1-2 sec	2-5 sec	4-60 sec
Cost	$ 60-215	$ 10-350	$ 20-235	$ 1-10

rate, sampling rate and end-to-end delay between information source and the information user. The algorithm computes mismatch probability based on the dynamic information elements whenever the information user access information. The sampling rate is tuned as per the maximum power consumption limit of the node. So, the algorithm takes into account both information reliability in terms of information mismatch probability and power consumption of the node, both are relevant performance parameters for wireless sensor networks.

Although, this algorithm is applicable to all the three information access strategies discussed in this thesis, but it is implemented for the reactive access strategy only. The only reason for choosing reactive as a sample analysis is that the mismatch probability equation in this case is easily solvable thereby keeping minimum time for selection of information source. Whereas for subscription based access (hybrid and periodic), they incur more time for selection of information source (due to numerical integration) thereby adding to access delay (unless it is a matter of reconfiguration, in which selection can be done by reconfiguring according to the network delay). The mismatch probability and sampling rate under reactive access strategy can is given by (5.4b) and (6.2) respectively. And they are reproduced in the pseudocode of the algorithm provided in Algorithm 7.1 for easy understanding of the algorithm.

The following values are used for testing and evaluation of the algorithm in Section 7.3, $\bar{P}_{Rea} = 80mW$, $P_{tx} = 5mW$, $P_{samp} = 3mW$, $t_{samp} = 2sec$ and $\mu = 3sec^{-1}$. Since,($t_{tx} + t_{samp}$) is the total delay, it is taken as $\frac{1}{\nu}$ in this

Algorithm 7.1 Pseudocode for Information Source Selection based on Mismatch Probability and Power Consumption

Initiate the request for information
Check for the registered information sources
procedure ArraySource($NoOfSource$)
  for each source i in $NoOfSources$ **do**
    Get reference ID of each source ($RefID$)
    Get information arrival rate (λ)
    Get the average rate of delay (ν)
    Compute sampling rate as per the required power
    consumption level as

$$\tau = \frac{\bar{P}_{Rea} - P_{tx} t_{tx} \mu}{P_{samp} t_{samp}}$$

    Compute mismatch probability($mmPr_{Rea}$) as

$$mmPr_{ComRea} = \frac{\lambda(\lambda + \tau + \nu)}{(\lambda + \tau)(\lambda + \nu)}$$

  end for
  Sort the $RefID$ by $mmPr_{Rea}$
  Choose the first element in the list
end procedure

algorithm.

7.3 Algorithm Implementation

In order to implement the proposed algorithm, the framework which was proposed and described in Chapter 4 is used. In Figure 7.2, information sources S1, S2 and so on were assumed to be any sensor nodes which can provide different information or same information. In order to suite this particular algorithm implementation, the information sources are modeled as same type of sensor node (temperature sensors) but with different characteristics which behave as different information sources for the same type of information (temperature information). The modified scenario is shown in Figure 7.1, with four temperature sensors shown as S1, S2, S3 and S4 acting as four different information sources which provide temperature information. It can be seen at later stage in this section, how these sources are modeled to

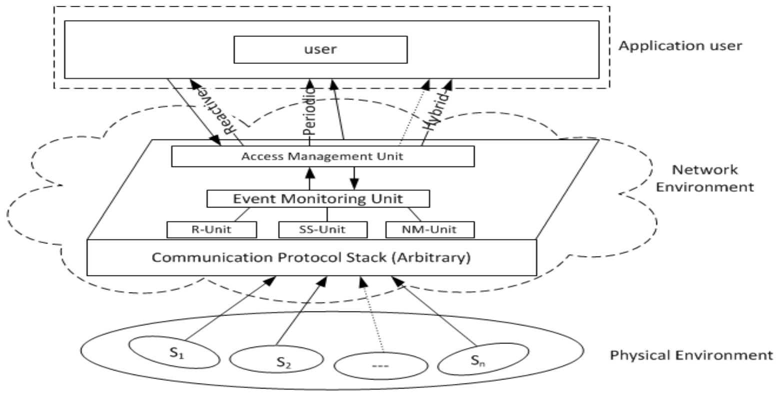


Figure 7.2: Proposed Framework Architecture (reproduced from Chapter 4).

represent difficult cases for selection of information source by modeling with different delay rates between each information source and information user. By assigning different delay rates, they represent different network traffics; different routes, congestion, retransmission, that a packet would encounter to reach the final destination.

As these four temperature sensors have different characteristic features as shown in Table 7.1, each source is modeled such that they generate information with different rates, assumed to be exponentially distributed without diverting from the proposed framework and assumptions described in Chapter 3. By doing so, all the information sources represent temperature sensors but generating information at different rates. This modeling is based on the notion that due to difference in sensitivity values for different sensors, the event rate will differ, although further exploration and study is encouraged to find out how the different properties of the devices really affects event rate. Nevertheless, in this chapter, the event rate (λ) is modeled based on the sensitivity values of each sensor and is computed as below (for effectively filtered signal):

As defined in Chapter 3, an occurrence of an event can be expressed as

$$|E(t) - E(t + dt)| > threshold$$

Considering a temperature change of ϕ°C at every t sec, and a sensor node which has a sensitivity of ρ mV per degree Celcius, assumming effectively filtered signal and constant change rate, the rate(θ) of change for temperature T can be computed as

$$\theta = \frac{d\rho}{dt} = \left(\frac{d\rho}{d\phi}\right)\left(\frac{dT}{dt}\right) \tag{7.1}$$

Using (7.1), we can get the time interval and information arrival rate (λ) as

$$\bar{t} = \frac{threshold}{\theta} \qquad \text{and} \qquad \lambda = 1/\bar{t}$$

If V is the generated voltage, then the sensitivity (ρ) can be expressed as

$$\rho = \frac{dV}{dT}$$

As an example, event rate (λ) for Platinum RTD temperature sensor as per the values shown in Table 7.1 is computed here. The Platinum RTD sensor has $\rho = 2mV/°$C. Say that a temperature change of $\phi = 1°$C/sec, then using Equation (7.1), rate of change (θ) will be equal to $2mV/sec$. This would give an average time interval, $\bar{t} = 1sec$, i.e $\lambda = 1\text{sec}^{-1}$, considering a threshold value of $2mV$. The threshold value selected has influence on event rate as seen from Chapter 3. It was learnt that for inter event arrival time to follow exponential distribution, the threshold value must be between 1 and 3. This is from the analysis carried out in Chapter 3. Similarly, event rate for other temperature sensors are obtained and the values are tabulated in Table 7.2. In order to have better comparison, four values of event rates (λ_1 to λ_4) are tabulated in the Table 7.2, with different values of thresholds and temperature change rate (ϕ). Two different values of temperature change rates are used 1)$\phi = 1°$C/sec (faster change), 2) $\phi = 0.03°$C/sec (slower change). A daily normal (fine weather) temperature may result slower change whereas abnormal weather with rain/clouds and wind may result faster change in weather condition.

Table 7.2: Event Rates for different values of ϕ and thresholds.

Threshold Value	**Temp. Change Rate(ϕ)**	**Event Rate**	**Platinum RTD**	**Thermistor**	**Thermocouple**	**Integrated Silicon**
$2mV$	1.0°C/sec	λ_1	1.0	20.0	0.025	1.0
	0.03°C/sec	λ_2	0.03	0.6	0.00075	0.03
$1mV$	1.0°C/sec	λ_3	2.0	40.0	0.05	2.0
	0.03°C/sec	λ_4	0.06	1.2	0.0015	0.06

Figure 7.3 shows how event rate(λ) varies against threshold values with different sensitivity values (ρ). It is clearly observed from the figure that as the threshold values increases, the event rate decreases and the variation is more at lower values of threshold. This means that if the chosen threshold value is high, most of the events will not be detected which will affect the

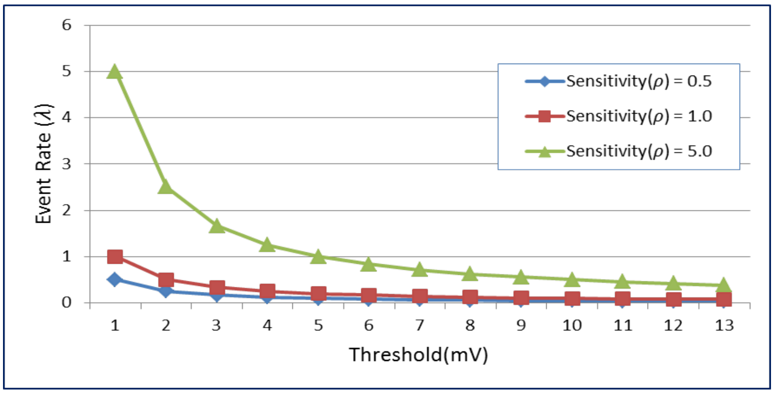


Figure 7.3: Event rate (λ) variation against threshold values with different values of sensitivity (ρ).

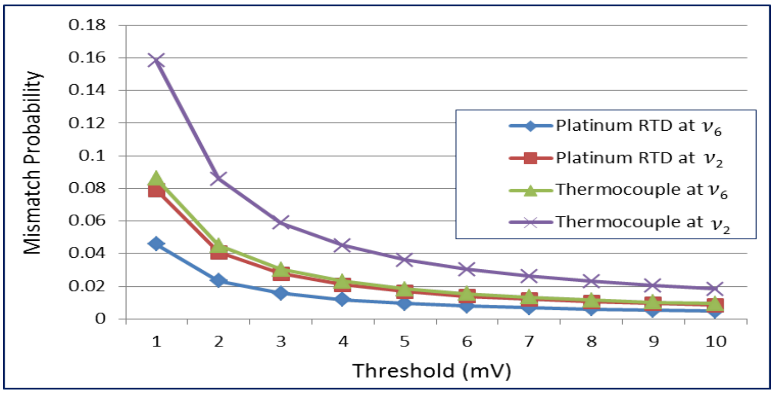


Figure 7.4: Mismatch probability variation with different values of mean delay rates (refer Table 7.3) with constant sampling rate ($\tau = 3.0$) and temperature change rate ($\phi = 0.03°C/sec$).

accuracy of information. The figure also shows that higher the value of sensitivity, more events can be detected.

As event rate (λ) is computed based on the sensitivity (ρ) of the sensor, the threshold value chosen will have impact on information mismatch probability. An analysis on the impact of different threshold values on mismatch probability is provided as shown in Figure 7.4. The variation of mismatch probability is observed to be more at lower threshold values as compared to higher threshold values. An interesting point that can be learnt from Figure 7.4 is the trade-off between information accuracy and information reliability. As shown from Figure 7.3, the accuracy (in terms of detecting events) is high at lower threshold values. So, in order to have greater accuracy, lower values of thresholds are desired but this will result high information mismatch probability (high chances of using wrong information). This is observed from

Figure 7.4. However, it is to be noted that this analysis is based on the assumption that the event rate is exponentially distributed. However, the short analysis carried out in Chapter 3 shows that inter event arrival time can be assumed as exponential distribution, provided appropriate parameters like thresholds and sampling frequencies are chosen. Nevertheless, this deserves further study to find out what types of distributions are more relevant and how they affect the information mismatch probability.

Table 7.3 shows different mean delay rates (assumed exponentially distributed) assigned in order to examine how different delay rates affect the selection of information sources. As mentioned earlier in this section, each delay rate represents a different network condition for each type of sensor, which means different access delays due to several reasons like different routes, network congestion and so on.

Table 7.3: Mean Delay Rates values used for evaluation purposes

Mean Delay Rates	**Platinum RTD**	**Thermistor**	**Thermocouple**	**Integrated Silicon**
ν_1	0.9	45	0.009	1.2
ν_2	0.8	40	0.008	0.8
ν_3	0.7	35	0.007	0.6
ν_4	0.6	30	0.006	0.4
ν_5	0.5	25	0.005	0.2
ν_6	0.4	20	0.004	0.1

7.4 Evaluation and Analysis

The impact of information dynamics on information source selection is evaluated and analyzed under four cases 1) with uniform average delay rates, 2) with different average delay rates, 3) with combinations of different average delay rates and 4) with different threshold values.

As expected, when the delay between information sources and information user is uniform, the information source (thermocouple,$\rho = 0.05mV/°\mathrm{C}$) which has lowest sensitivity is selected the most. This is because it has the least λ associated to it. This is seen from Figure 7.5. But in reality, when information sources are deployed in different locations and separated by different number/kinds of intermediate networking devices, their delays will not be uniform since they take different routes. Therefore, the scenario shown in Figure 7.5 represents a very specific case. So in order to get a clear picture on

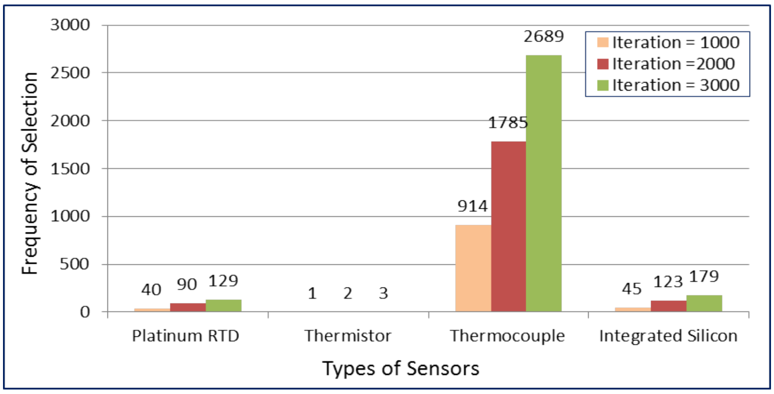


Figure 7.5: Source selection scenario at λ_1 from Table 7.2 with uniform delay rate of $\nu = 0.5$.

how the selection of information sources would be in real scenario, different delay rates are used for each information source and the resulted information source selection observation is shown in Figure 7.6

Contrary to Figure 7.5, where the selection of information sources are mainly decided by the event rate, Figure 7.6 shows that the selection of information sources are greatly influenced by the network conditions between each source and the information user. Consider a case where the information mismatch probability is not known. Which information source can be chosen? To choose among many information sources would be very difficult and selection is not necessarily the one which has the lowest delay (Thermistor sensor- which actually gives worse information reliability, generally due to its higher value of sensitivity). Therefore, when the decision is affected by combinations of different network parameters, mismatch probability plays an important role in deciding the selection of the most reliable information source and the requirement of an algorithm such as this is felt in such situations.

The delay rates in a particular access route may not remain same all the time. So, investigations on how different delay rates do impact on the selection of information sources are presented here. Figure 7.7 shows how information sources are selected with combinations of different average delay rates between each information source. The corresponding plot for average mismatch probability variations with combinations of different average delay rates is shown in Figure 7.8

From Figure 7.7, it is observed that at ν_1, thermistor sensor is selected about 420 out of 5000 samples. And integrated silicon sensor is selected about 2000 out of 5000 samples. Similarly, other two sensors are also se-

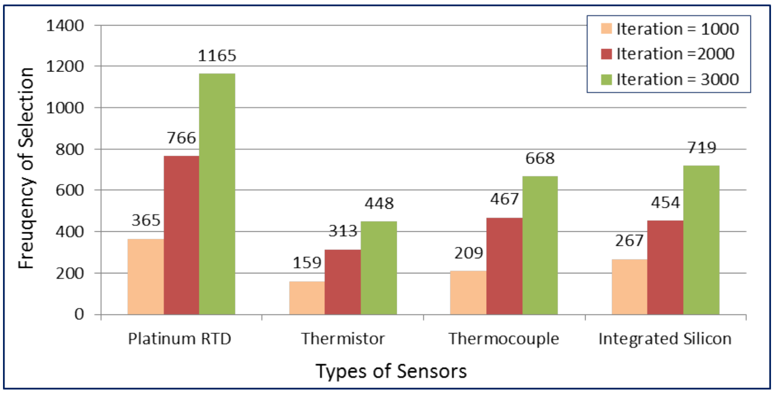


Figure 7.6: Source selection with different average delay rates (ν_4) from Table 7.3 and with λ_1 from Table 7.2.

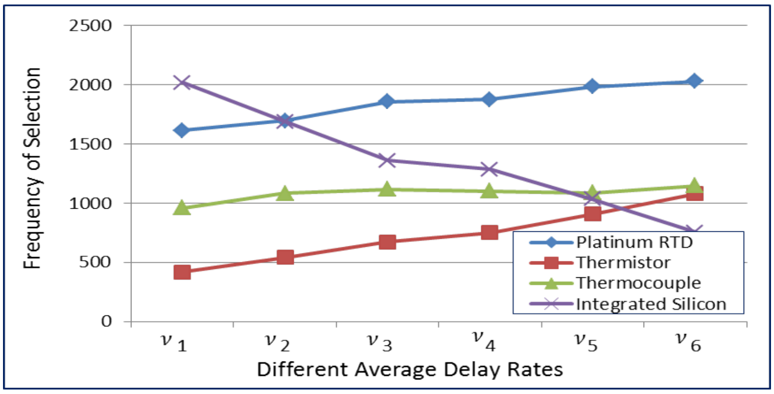


Figure 7.7: Information source selection scenario with combination of different delay rates from Table 7.3 with λ_1 as shown in Table 7.2.

lected for some runs. Now look at the corresponding mismatch probability plot in Figure 7.8, it shows that at ν_1, integrated silicon sensor has the lowest mismatch probability. Without this information dynamic source selection algorithm, one would never pick up alternative sensors for the best information reliability. This would land up picking one particular sensor for all samples as shown in Figure 7.9 where at ν_1 integrated silicon is chosen for all 5000 samples, and from ν_2 onwards platinum RTD is chosen for all 5000 samples.

The scenario such as shown in Figure 7.9 is based on static setting, where selection is based on average mismatch probability. Such static setting algorithm, where event rates and delay rates are kept static (keeping same for all 5000 samples) would land up using the resources of one source extensively while others are never being used at all. Therefore, by using the proposed dynamic information source selection algorithm, which is based

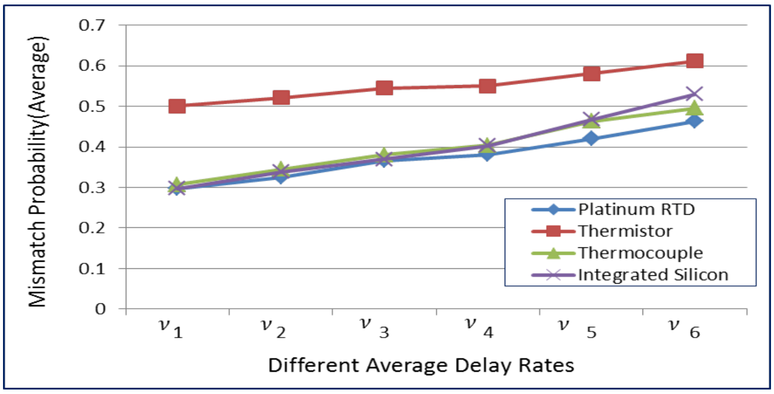


Figure 7.8: Mismatch probability variation with combination of different delay rates from Table 7.3 with λ_1 as shown in Table 7.2.

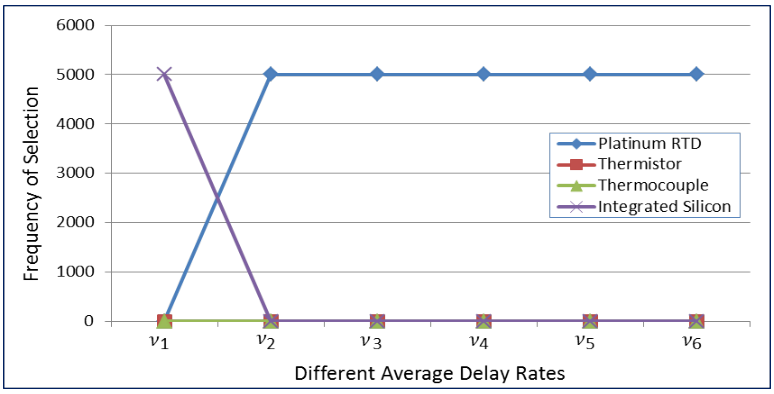


Figure 7.9: Information source selection variation based on average mismatch probability (w.r.t. Figure 7.8) with different average delays rates shown in Table 7.3.

on the information dynamics and access delays (which change over time), it is possible to select information source which gives the best information reliability.

Although, the mismatch probability plot shown in Figure 7.8 is based on average, which means at some runs, the mismatch probability becomes lower than the average while at some runs it becomes higher than the average. However, if one compares frequency of source selection plot in Figure 7.7 and mismatch probability plot in Figure 7.8, one will observe that the sensor which has lowest average mismatch probability value is selected the most. In particular, at ν_1 integrated silicon sensor is selected the most as it has the lowest mismatch probability value, and so on and so forth with the most frequently selected ones. But one should remember at this point that those different values of ν represent different network conditions. Therefore, slightly after ν_2 in Figure 7.7, it is observed that the frequency of selection of integrated silicon sensor drops unlike other sensors. This is because the delay for integrated silicon sensor degrades relatively faster than the others (see also Table 7.3). At ν_6, a different behaviour is observed, where the thermistor sensor is selected more frequently than the integrated silicon, although the mean mismatch probability of thermistor is greater than the integrated silicon. This clearly shows that the decision of selection of particular information source is not decided solely by the average mismatch probability values or any single network parameter for that matter, but by the combinations of different parameters. Therefore, it depends on individual scenario, which is often decided by the information dynamics and network delays.

In order to investigate whether the different values of temperature change

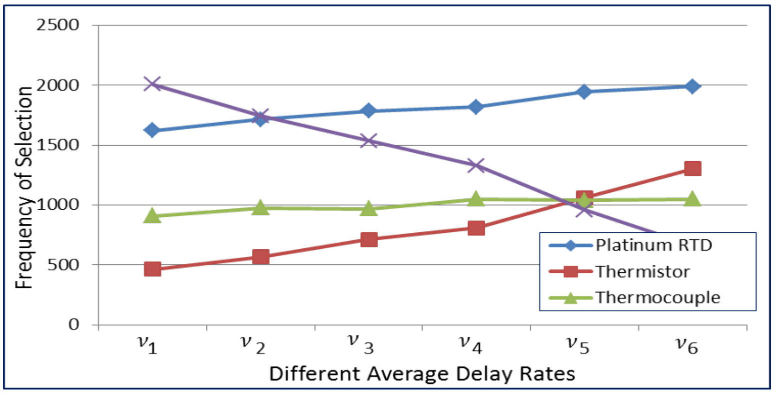


Figure 7.10: Information source selection scenario with combination of different delay rates from Table 7.3 with λ_2 as shown in Table 7.2.

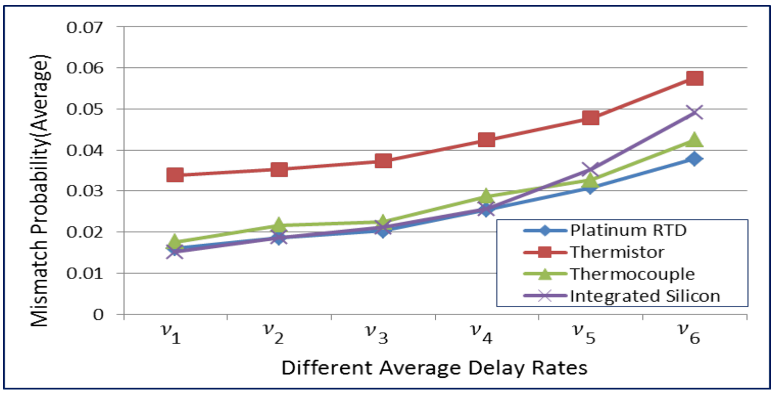


Figure 7.11: Mismatch probability variation with combination of different delay rates from Table 7.3 with λ_2 as shown in Table 7.2.

rate (ϕ) has any impact on the frequency of selection, a similar study is done by taking different values of $\phi = 0.03$°C/sec. The corresponding values of event rate for each sensor are shown as λ_2 in Table 7.2.

As expected, information source selection decision is not affected by using different value of ϕ. This can be observed by comparing Figure 7.10 with Figure 7.7. However, average information mismatch probabilities have improved. This is due to the decrease in event rate (λ) as a result of decrease in ϕ value. This observation can be seen from Figure 7.11.

As shown in Figure 7.3, different *threshold* values impact information mis-

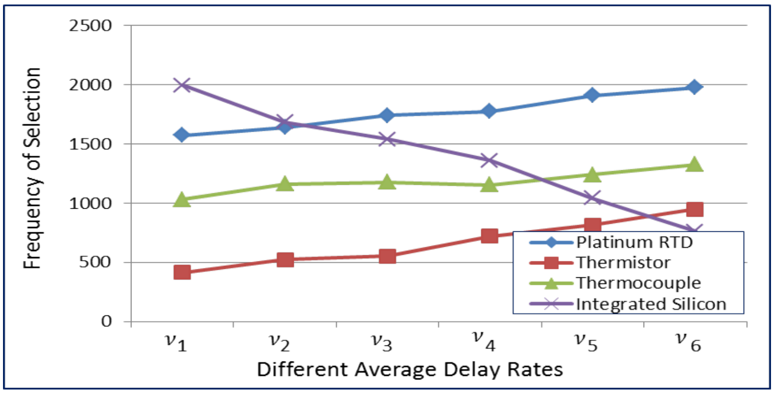


Figure 7.12: Information source selection scenario with combination of different delay rates from Table 7.3 with λ_3 as shown in Table 7.2.

match values as event rate depends on the value of threshold chosen. In order to investigate if the threshold value chosen has any impact on the information source selection decision, a similar set of study is presented here as above but using different threshold value ($threshold = 1mV$) as reflected in Table 7.2.

Figure 7.12 shows the information source selection variation based on network condition shown in Table 7.3. Figure 7.13 shows the corresponding information mismatch probability variation. Comparing Figure 7.7 and Figure 7.12, it is observed that these two results are more or less same; thereby showing that the information source selection decision is not affected by using different threshold values. However, as expected the information mismatch probability values are affected. This can be seen by comparing Figure 7.8 ($threshold = 2mV$) and Figure 7.13 ($threshold = 1mV$). Information reliability in terms of mismatch probability decreases with lower values of thresholds and this is also observed in Figure 7.4.

Similarly, investigation is done by using λ_4 as event rate. The results are shown in Figure 7.14 and Figure 7.15. Comparing Figure 7.14 and Figure 7.10, here also it is observed that the threshold value chosen does not have

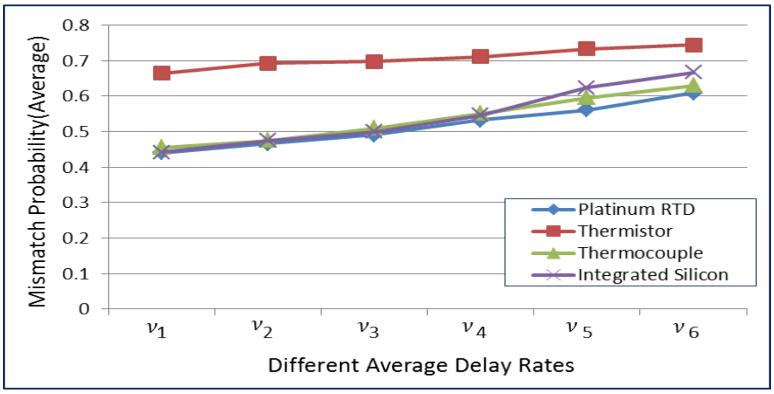


Figure 7.13: Mismatch probability variation with combination of different delay rates from Table 7.3 with λ_3 as shown in Table 7.2.

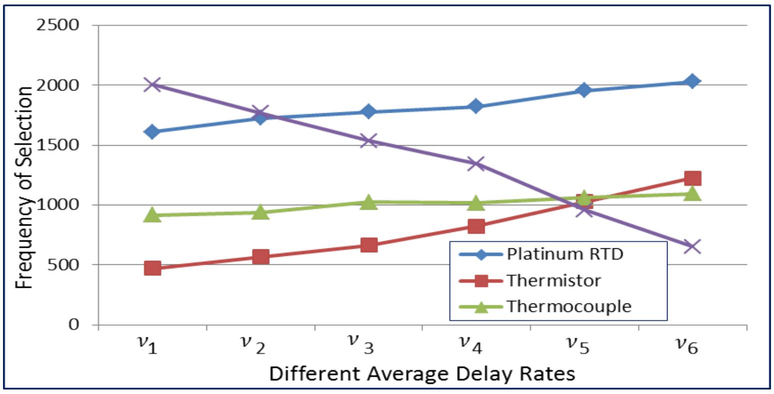


Figure 7.14: Information source selection scenario with combination of different delay rates from Table 7.3 with λ_4 as shown in Table 7.2.

any impact on information source selection decision. Same observation is made on the impact on mismatch probability values. These can be seen by comparing Figure 7.11 with Figure 7.15.

If you compare all the results, with respect to information source selection verses different average delay rates, the frequency of selection of particular information source is same for all the cases, irrespective of different event rate and different threshold values being used. This shows that the proposed dynamic information source selection algorithm selects the information source with the best mismatch probability as per the network conditions and information dynamics.

The selection of information source can be done based on lowest event rate (λ) or highest delay rate (ν) in addition to selection based on information mismatch probability (which takes into account both λ and ν) described Section 7.4. The Algorithm 7.1 can be modified to set the criteria of selection

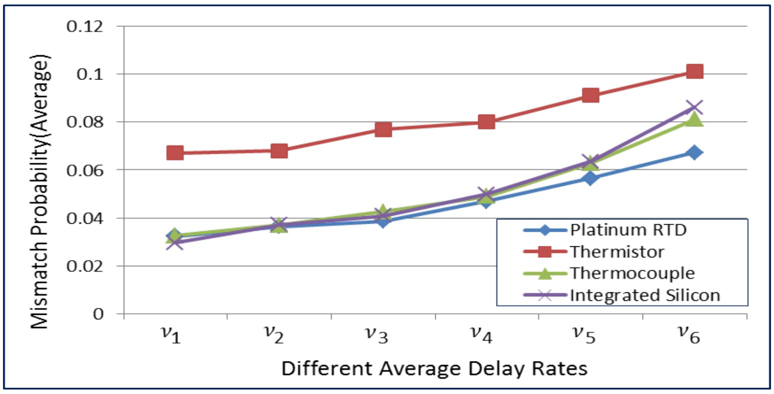


Figure 7.15: Mismatch probability variation with combination of different delay rates from Table 7.3 with λ_4 as shown in Table 7.2.

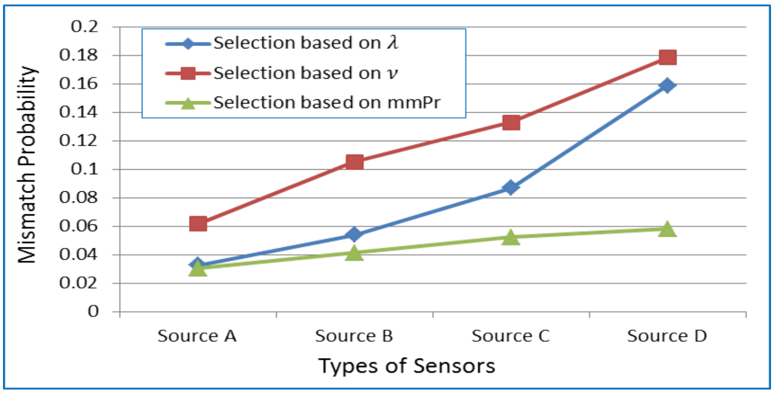
to lowest λ or highest ν. This section provides a comparative study on the following three types of selection methods:

1. Selection based on lowest mean delay (highest delay rate (ν))
2. Selection based on lowest event rate(λ)
3. Selection based on lowest information mismatch probability (mmPr)

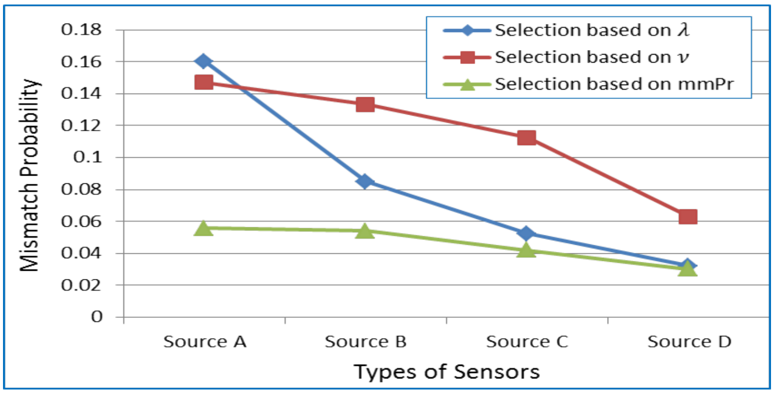
Table 7.4: Parameters used for performance comparisons of selections methods.

Cases	**Para-meter**	**Source A**	**Source B**	**Source C**	**Source D**
I	λ	0.1	0.2	0.3	0.4
	ν	4.0	3.0	2.0	1.0
II	λ	0.4	0.3	0.2	0.1
	ν	1.0	2.0	3.0	4.0
III	λ	0.1	0.2	0.3	0.4
	ν	1.0	2.0	3.0	4.0
IV	λ	0.4	0.3	0.2	0.1
	ν	4.0	3.0	2.0	1.0

In order to show the performance comparison among different information source selection methods, three cases with different values of event rate (λ) and delay rate (ν) are used as shown in Table 7.4. Figure 7.16(a) to Figure 7.17(a) show the performance (information mismatch probability) comparison of different selection methods for three different cases.



(a)

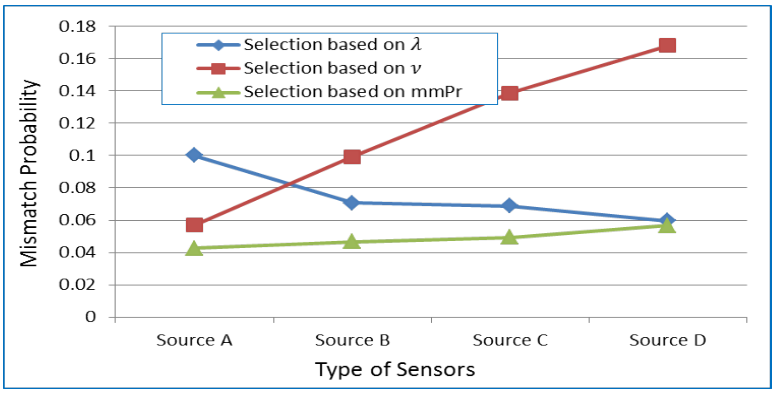


(b)

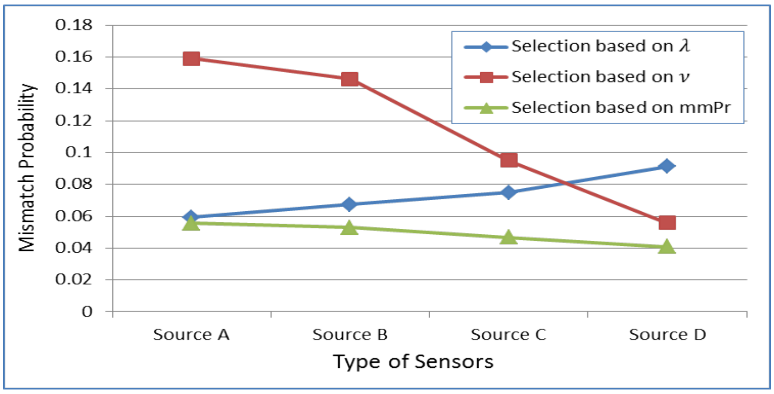
Figure 7.16: Performance comparisons for different selection methods for different cases as shown in Table 7.4, a) Case I b) Case II

In Case I, as expected, the mismatch probability values increases moving from source A to source D. This is because the values of λ increases while value of ν decreases (refer Table 7.4, which results in increasing mismatch probability for all the three selection methods.This is shown in Figure 7.16(a). When selection is made based on lowest value of λ, the mismatch probability mainly influenced by value of ν. Similarly, when the selection is based on highest value of ν, the mismatch probability is mainly influenced by λ. Since both the parameters are set such that they have negative (mismatch probability increases) impact mismatch probability, the mismatch probability resulted due to selection based on mismatch probability also show an increase in mismatch probability value as it moves from source A to source D.

In Case II, the values of λ is set so that it increases from source A to source D, whereas values of ν are set such that it increases from source A to source D (refer Table 7.4). So this results in decrease in mismatch probability as it moves from source A to source D irrespective of selection methods being used.This is shown in Figure 7.16(b). This is due to decrease in event rate(λ) and decrease in delay rate(ν) from source A to source D.



(a)



(b)

Figure 7.17: Performance comparison for different selection methods for different cases as shown in Table 7.4, a) Case III b) Case IV

In Case III, both the values of λ and ν are set such that they both increase from source A to source D. As explained in case I, the mismatch probability value (where the selection is based on highest value of ν) increases, this is due to increase in λ, which has greater impact on mismatch probability as seen from Case IV. In case of selection based on lowest of λ, the mismatch probability decreases. This is due to an increase in ν. However, this result may change as per the values chosen for λ and ν. Since these two parameters are such that, they influence the mismatch probability in different directions (meaning when λ increases, mismatch probability increases, whereas when ν increases, mismatch probability decreases), the mismatch probability value for selection based on lowest mismatch probability will either increase or decrease as per the values of λ and ν. This is shown in Figure 7.17(a).

In case IV, both the values of λ and ν are set such that they both decrease from source A to source D. As expected, mismatch probability decreases in case of selection based on highest value of ν while it increases in case of selection based on lowest value of λ. This is shown in Figure 7.17(b). Comparing Figure 7.17(b) with Figure 7.17(a), it is learnt that λ has more influence on mismatch probability than ν. That is why, the mismatch probability in case of selection based on lowest mismatch probability decreases in Case IV.

For all cases, the selection method which is based on mismatch probability gives better information reliability compared to other two methods.

7.5 Conclusions

In this chapter, a dynamic information source selection algorithm is presented, which selects the information source which gives the best mismatch probability. To conclude this chapter, the main lessons learnt are listed here:

- By using the information mismatch probability as information reliability parameter, the proposed algorithm was found to be useful in making decision which information source to use in an otherwise difficult decision scenario. Without the mismatch probability metric, it would be difficult to determine which information source to use, and for various delay and event rates (λ).

- The strength of the algorithm was tested by setting different mean network delay rates (making scenarios difficult for selection), and letting the algorithm decide which information source to use. This can be seen from Section 7.4, where algorithm chooses the information source which gives the best mismatch probability.

- This type of algorithm is required and essential, since static setting is not useful (as shown from Figure 7.7 as it tends to extensively use particular information source thereby using the resources of the particular source heavily. Moreover, the selection decision is not considering the information dynamics which has impact on information reliability.

- The consistency and reliability of the algorithm is tested by using different parameters like different values of thresholds, event rates (λ). The results with regards to selection decision were found to be consistent with different threshold and λ values.

- A comparative study of information source selection based on lowest event rate (λ), highest delay rate (ν), and based in lowest mismatch probability value was carried out. It is found out that the selection based on mismatch probability gives better information reliability in terms of mismatch probability.

- It was observed that the event rate(λ) influences mismatch probability more than the mean delay rate(ν). Therefore, the information dynamics at source has greater impact on information reliability in terms of information mismatch probability.

- With the help of the proposed dynamic information selection algorithm it was possible to select alternative information source depending on the information dynamic element at source and network delay between source and the destination.

Although, the proposed algorithm was found to be useful for making selection decision in an otherwise difficult decision scenario compare to static settings, however, it deserves further study to investigate the relationship between average mismatch probability and the selection frequency. In this chapter, the proposed algorithm is implemented only with reactive information access strategy. Therefore, implementation under periodic and hybrid information access strategies could be explored in order to investigate if they show any different result with regard to information source selection.

Chapter 8

Conclusions and Outlook

This chapter revisits some of the main points which were studied in this thesis. In particular this chapter contains following three sections:
Reflections- A brief recapture of the whole thesis is provided in this section by revisiting the main points and reflecting on how the study was carried out.
Concluding Remarks- In this section, the main findings and contributions of the study are provided. The main purpose of this section is to revisit the research objectives defined in the beginning of the thesis and to reflect back on the findings to see if the objectives are fulfilled.
Outlook and Future Directions- This section brings out some of the opinions of the author on the topic. It also gives some future directions, should there be any researcher inspired from this work and wishing to carry on further research on this topic.

8.1 Reflections

One cannot avoid remotely accessing to dynamically changing information elements and it has become a required functionality with various network services. One common problem to remotely accessing such information element is the high probability of using incorrect information in making decision at the user end. This problem becomes more challenging with some sensor network applications like safety monitoring, where using wrong information have severe impact on the action taken based on the decision. In order to support design decision, it is essential to know the characteristics features of dynamic information elements and how they affect the information reliability. Therefore, in order to study the behaviour of dynamic information element with regard to information reliability, it is essential that there is an information management framework which catches dynamic information parameters. Generally, when the sensors are used to collect information from a hostile environment, an efficient usage of their battery power is utmost important in order not to reduce the lifetime of the network. Therefore, in this thesis, comprehensive studies on information accessing to remote dynamic information sources are made. A brief revisit of what each main chapter was all about is provided here.

Chapter 3 provided basic models with respect to information access to a dynamic information sources. It was a foundation chapter, which introduced the basic fundamentals with regard to information access model, which were investigated and used in later chapters. The chapter described a general information access model bringing out three types of information access strategies; reactive, periodic and hybrid. Following the description of information access model, a brief study on an event process and detection was carried out as it was necessary to investigate how events are detected and processed by a sensor node. Lastly, the chapter presented a general power consumption model, which was one of the main focus areas in this study.

Chapter 4 presented a dynamic and reliable information accessing and management framework. The framework was developed using Java-based framework called OSGi, which has capability of providing an environment for modularization of applications and services. The developed framework facilitated in studying the behaviour of dynamic information elements, which were useful for designing adaptive information access mechanism and information source selection algorithms in later chapters. The working of the framework was evaluated by estimating the information mismatch probability for three different access strategies discussed in Chapter5.

Chapter 5 presented a comprehensive study on information mismatch probability models and power consumption models under different information access strategies. The analytical mismatch probability models defined in [6] were extended by reshaping to a sampling process by reducing otherwise included network delay into a deterministic zero delay. Analytical investigations on how dynamic information elements affect the information mismatch probability and power consumption were done with respect to all the three information access strategies.

Chapter 6 presented an information access mechanism which manages information reliability and power consumption based on the requirement of the application. In particular, an adaptive information access mechanism by way of finding trade-off point between information reliability and power consumption was provided.

Chapter 7 presented a dynamic information source selection algorithm. The algorithm was implemented and evaluated using the framework developed in Chapter 4. The algorithm was developed using mismatch probability as a metric to decide the selection decision. As this metric is a combination of different dynamic information elements, the algorithm was found to be useful in selecting the information source that gives the best information reliability compare to static setting selection method.

8.2 Concluding Remarks

An OSGi based dynamic and reliable information accessing and management framework was developed. The framework was used to study the dynamic behaviour of information, in relation to wireless sensor networks. The flexible and easy reconfigurable property that OSGi provides made the framework to be dynamic which facilitated in studying the dynamic characteristics and features of information collected by wireless sensors. By monitoring the information dynamics at source and the access delay between information source and the information user, the information mismatch probability, a metric that was used to measure information reliability was estimated. Without this type of framework, an estimation of information mismatch probability was not possible, unless one resort to compute mathematically. The usefulness of information reliability metric (mismatch probability) was shown throughout the thesis.

A comprehensive study on power consumption and information mismatch probability in relation to information access strategies; reactive, periodic and hybrid was carried out. Based on this study, an extension to the mismatch probability models defined in [6] were provided. For all the three access strategies, it was observed that the information mismatch probability was affected by the rate at which the information is generated at the source, rate at which the information is being sampled and the access delay or end-to-end delay between the information source and the information user. Furthermore, it was found out that the information mismatch probability was also impacted by the way how information is being accessed; meaning which information access strategy is used. Using the same parameters settings, the information mismatch probability under hybrid access strategy was observed to be always greater than the other two. This observation was expected as the processing delay introduced for checking for an update adds to the access delay. Comparing between reactive and periodic access strategy, reactive was found to be suffering in terms of mismatch probability. Regarding on the impact of information access strategy on power consumption, the hybrid access showed advantages over the other two. Although, the advantages over reactive access depends on the ratio between the information access rate (μ) and the information update rate (τ'), but the power consumption by using this access strategy was always less than that periodic access. As the highest update rate(τ') can go up to the sampling rate, the maximum power that hybrid access strategy has to spent will be always less than or equal to the power that has to spent with periodic access.

As a result of the comprehensive study done on power consumption and mismatch probability, an adaptive information access mechanism to dynamic information sources was presented. It was interesting to learn that information access can be adapted as per the level of requirement and the available resources. For instance, the information reliability requirements in terms of information mismatch probability could be maintained by varying sampling rate. In addition, it was also found out that by increasing the information sampling rate expecting gain in information reliability had huge impact on power consumption. This study presented a trade-off point between information reliability and power consumption at which an increase in information sampling rate did not have significant improvement on information reliability but had huge increase in power consumption. Therefore, it was learnt that it is important to find a trade-off between gains in one parameter at the cost of other.

The usefulness of mismatch probability metric had shown when it came to the design of a dynamic information source selection algorithm. As mismatch probability depends on the dynamic information elements like event rate and access delays, the metric was used for making decision on information source selection. It was demonstrated that the presented algorithm could select the information source which gives the best information mismatch probability in an otherwise difficult situation. Although, the observation was restricted to the assumption that the delays are exponentially distributed, the algorithm's ability to select which information source to use among the different information sources having different network conditions, showed its strength.

8.3 Outlook and Future Directions

Dynamic and reliable information accessing and management framework presented in this thesis was developed with an objective to study the behaviour of dynamic information elements. To develop such a framework and to evaluate it, it was found that, one needs to know how the characteristic features of dynamic information elements. Modeling information source that mimics dynamic information elements was found to be not a trivial task. However, with the help of an OSGi framework, which supports modularity with dynamic and flexible reconfigurable properties facilitated in modeling the information source which emulates dynamic information elements. While studying the dynamic information element's behaviour, both analytical and with the help of the framework, the end-to-end delay distribution considered was simple exponential distribution. This is a worst case scenario and further study considering different distributions would be one of the future directions that need to mention here.

The process of event generation and detection done in this study was not meant for vigorous mathematical analysis. However, there is a possibility of making in depth study so as to get the clear picture how dynamic information changes its behaviour in changing time and environment.

The information source selection algorithm presented in this thesis selects the source which gives the best information mismatch probability. Doing so, the framework needs to compute information mismatch probability metric for each access made by the information user. The time required to compute this metric will also add to the end-to-end delay. Although, it was not considered in this study, however, this will have impact on the resources like power usage. Therefore, a novel way to estimate mismatch probability without the need of computation would give better performance. One possible future direction in this regard is to estimate the ratio between rate of information generation and rate of delay for each source and study at which point the dynamic switching of information source is taking place. However, it would not be trivial as it also depends on the rate at which information is sampled, which in turn is decided by the power consumption limit.

Bibliography

[1] C. Maple, G. Williams, and Y. Yue, "Reliability, availability and security of wireless networks in the community," *Informatica*, vol. 31, pp. 201–208, 2007.

[2] D. Geogoulas and K. Blow, "Wireless sensor network management and functionality: An overview," *Wireless Sensor Network*, vol. 1, no. 4, pp. 257–267, 2009.

[3] WZ Song, R. Huang, M. Xu, A. Ma, B. Shirazi, and R. LaHusen, "Air-dropped sensor network for real-time high-fidelity volcano monitoring," in *In Mobisys'09*, 2009, pp. 305–318.

[4] Y. Yu, LJ Rittle, V. Bhandari, and JB LeBrun, "Supporting concurrent applications in wireless sensor networks," in *Proceedings of the 4th international conference on Embedded networked sensor systems*, 2006, pp. 139–152.

[5] Yunbo Wang, M. C. Vuran, and S. Goddard, "Cross-layer analysis of the end-to-end delay distribution in wireless sensor networks," *Networking, IEEE/ACM Transactions on*, vol. 20, no. 1, pp. 305–318, 2012.

[6] M. Bogsted, RL Rasmus, and HP Schwelfel, "Probabilistic models for access strategies to dynamic information elements," *Performance Evaluation*, vol. 67, pp. 43–60, 2010.

[7] Open Service Gateway Initiative, "About the osgi alliance," Available: http://www.osgi.org, accessed on 11-Dec-2010.

[8] Kemal Akkaya and Mohamed Younis, "A survey on routing protocols for wireless sensor networks," *Ad Hoc Networks*, vol. 3, no. 3, pp. 325–349, 5 2005.

[9] Marcos A. M. Vieira, Adriano B. da Cunha, and Diógenes C. da Silva, "Designing wireless sensor nodes," in *Proceedings of the 6th international conference on Embedded Computer Systems: architectures, Modeling, and Simulation*, Berlin, Heidelberg, 2006, SAMOS'06, pp. 99–108, Springer-Verlag.

[10] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: a survey," *Computer Networks*, vol. 38, no. 4, pp. 393–422, 3/15 2002.

[11] T. Ahonen, R. Virrankoski, and M. Elmusrati, "Greenhouse monitoring with wireless sensor network," in *Mechtronic and Embedded Systems and Applications, 2008. MESA 2008. IEEE/ASME International Conference on*, 2008, pp. 403–408.

[12] N. Artim, *An Introduction to Fire Detection, Alarm, and Automatic Fire Sprinklers*, 2007.

[13] AJ Garcia-Sanchez, F. Garcia-Sanchez, F. Losilla, P. Kulakowski, J. Garcia-Haro, A. Rodrguez, JV Lpez-Bao, and F. Palomares, "Wireless sensor network deployment for monitoring wildlife passages," *Sensors*, vol. 10, pp. 7236–7262, 2010.

[14] J. Lloret, M. Garcia, D. Bri, and S. Sendra, "A wireless sensor network deployment for rural and forest fire detection and verification," *Sensors*, vol. 9, pp. 8722–8747, 2009.

[15] Y. Hao and R. Foster, "Wireless body sensor networks for health-monitoring applications," *Physiological Measurement*, vol. 29, no. 11, pp. R27–56, 2008.

[16] Liang Cheng and S. N. Pakzad, "Agility of wireless sensor networks for earthquake monitoring of bridges," in *Networked Sensing Systems (INSS), 2009 Sixth International Conference on*, 2009, pp. 1–4, ID: 1.

[17] J. Lester Hill, "System architecture for wireless sensor networks," *University of California, Berkeley*, 2003.

[18] B. R. Shrestha and P. Manandhar, "Monitoring glacial lake in himalayas through sensor and wireless technology," Available: http://www.apan.net/meetings/kualalumpur2009/proposals/Agriculture/, 2009.

[19] A. Majeed and T. A. Zia, "Multi-set architecture for multi-applications running on wireless sensor networks," in *Advanced Information Networking and Applications Workshops (WAINA), 2010 IEEE 24th International Conference on*, 2010, pp. 299–304.

[20] S. Bandyopadhyay and E. J. Coyle, "An energy efficient hierarchical clustering algorithm for wireless sensor networks," in *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies*, 2003, vol. 3, pp. 1713–1723 vol.3.

[21] V. P. Mhatre, C. Rosenberg, D. Kofman, R. Mazumdar, and N. Shroff, "A minimum cost heterogeneous sensor network with a lifetime constraint," *Mobile Computing, IEEE Transactions on*, vol. 4, no. 1, pp. 4–15, 2005.

[22] I. F. Akyildiz, T. Melodia, and K. R. Chowdhury, "A survey on wireless multimedia sensor networks," *Computer Networks*, vol. 51, pp. 921–960, 2007.

[23] E. J. Duarte-Melo and Mingyan Liu, "Analysis of energy consumption and lifetime of heterogeneous wireless sensor networks," in *Global Telecommunications Conference, 2002. GLOBECOM '02. IEEE*, 2002, vol. 1, pp. 21–25 vol.1.

[24] Song Guo, Chunxia Fan, and T. D. C. Little, "Supporting concurrent task deployment wireless sensor networks," in *Network Computing and Applications, 2008. NCA '08. Seventh IEEE International Symposium on*, 2008, pp. 111–118.

[25] D. Lagutin, "Data dissemination and gathering in sensor networks," Available: http://www.tcs.hut.fi/Studies/T-79.7001/2007SPR, accessed on 11-Dec-2012.

[26] S. K. Singh, M. P. Singh, and D. K. Singh, "Routing protocols in wireless sensor networks- a survey," *International Journal of Science and Engineering Survey (IJCSES)*, vol. 1, no. 2, 2010.

[27] A. K. Dwivedi and O. P. Vyas, "Network layer protocols for wireless sensor networks: Existing classifications and design challenges," *International Journal of Computer Applications.*, vol. 8, no. 12, pp. 30–34, 2010.

[28] D. Chu, A. Deshpande, J. M. Hellerstein, and Wei Hong, "Approximate data collection in sensor networks using probabilistic models," in

Data Engineering, 2006. ICDE '06. Proceedings of the 22nd International Conference on, 2006, pp. 48–48.

[29] N. Bulusu, J. Heidemann, and D. Estrin, "Gps-less low-cost outdoor localization for very small devices," *Personal Communications, IEEE*, vol. 7, no. 5, pp. 28–34, 2000.

[30] J. N. Al-Karaki and A. E. Kamal, "Routing techniques in wireless sensor networks: a survey," *Wireless Communications, IEEE*, vol. 11, no. 6, pp. 6–28, 2004.

[31] T. P. Lambrou and C. G. Panayiotou, "A survey on routing techniques supporting mobility in sensor networks," in *Mobile Ad-hoc and Sensor Networks, 2009. MSN '09. 5th International Conference on*, 2009, pp. 78–85.

[32] V. Biradar Rajashree, V. C. Patil, S. R. Sawant, and R. R. Mudholkar, "Classifications and comparisons of routing protocols in wireless sensor networks," *Special Issue on Ubiquitous Computing Security Systems*, pp. 704–711.

[33] M. Younis, M. Youssef, and K. Arisha, "Energy-aware routing in cluster-based sensor networks," in *Modeling, Analysis and Simulation of Computer and Telecommunications Systems, 2002. MASCOTS 2002. Proceedings. 10th IEEE International Symposium on*, 2002, pp. 129–136.

[34] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *Proceedings of the 33rd Hawaii International Conference on System Sciences(HICSS'00)*, 2000.

[35] W. Heinzelman, J. Kulik, and H. Balakrishnan, "Adaptive protocols for information dissemination in wireless sensor networks," in *MobiCom '99 Proceedings of the 5th annual ACM/IEEE international conference on Mobile computing and networking*, 1999, pp. 174–185.

[36] K. Sohrabi, J. Gao, V. Ailawadhi, and G. J. Pottie, "Protocols for self-organization of a wireless sensor network," *Personal Communications, IEEE*, vol. 7, no. 5, pp. 16–27, 2000.

[37] M. Bauer, R. L. Olsen, M. Jacobsson, L. Sanchez, J. Lanza, M. Imine, and N. Prasad, "Context management framework for magnet and beyond," in *In Workshop on "Capturing Context and Context Aware Systems and Platforms", Proceedings of IST Mobile and Wireless Summit.*, 2006.

[38] L. Sanchez, J. Lanza, R. Olsen, M. Bauer, and M. Girod-Genet, “A generic context management framework for personal networking environments,” in *Mobile and Ubiquitous Systems: Networking and Services, 2006 Third Annual International Conference on*, 2006, pp. 1–8.

[39] Yongkai Zhan, Shuangquan Wang, Zhuang Zhao, Canfeng Chen, and Jian Ma, “A mobile device oriented framework for context information management,” in *Information, Computing and Telecommunication, 2009. YC-ICT '09. IEEE Youth Conference on*, 2009, pp. 150–153.

[40] W3C, “Resource description framework (rdf) schema specification,” Available: http://www.w3.org/TR/2000/CR-rdf-schema-20000327/, 2000, accessed on 11-Feb-2013.

[41] Jun-Hong Lu, Chiung-Ying Wang, and Ren-Hung Hwang, “An open framework for distributed context management in ubiquitous environment,” in *Ubiquitous, Autonomic and Trusted Computing, 2009. UIC-ATC '09. Symposia and Workshops on*, 2009, pp. 88–93.

[42] UPnP, “Upnp forum,” http://www.upnp.org, accessed on 11-Feb-2013.

[43] M. Venkataraman, M. Chatterji, and M. Kwiat, *Dynamic Routing Framework for Wireless Sensor Networks*, In Tech, yen kheng tan edition, 2010.

[44] V. Sachidananda, A. Khelil, and N. Suri, “Quality of information in wireless sensor networks:a survey.,” in *In 15th International Conference on Information Quality (ICIQ 2010)*, 2010, pp. 193–207.

[45] R. L. Olsen, H. P. Schwefel, and M. B. Hansen, “Qrp01-5: Quantitative analysis of access strategies to remote information in network services,” in *Global Telecommunications Conference, 2006. GLOBECOM '06. IEEE*, 2006, pp. 1–6.

[46] H. P. Schwefel, M. B. Hansen, and R. L. Olsen, “Adaptive caching strategies for context management systems,” in *Personal, Indoor and Mobile Radio Communications, 2007. PIMRC 2007. IEEE 18th International Symposium on*, 2007, pp. 1–6.

[47] A. Shawky, R. L. Olsen, J. Pedersen, and H. P. Schwefel, “Optimizing the quality of dynamic context subscriptions for scarce network resources,” in *Proceedings of the 1st European Workshop on AppRoaches to MObiquiTous Resilience, ARMOR '12*, 2012, pp. 6.1–6.6.

[48] Shaikh F. Karim, "Tunable reliability of information transport in wireless sensor networks," Available: http://tuprints.ulb.tu-darmstadt.de/2184/, 2010.

[49] Kang Cai, Gang Wei, and Huifang Li, "Information accuracy versus jointly sensing nodes in wireless sensor networks," in *Circuits and Systems, 2008. APCCAS 2008. IEEE Asia Pacific Conference on*, 2008, pp. 1050–1053.

[50] J. Karjee and H. S. Jamadagni, "Data accuracy model for distributed clustering algorithm based on spatial data correlation in wireless sensor networks," *CoRR*, 2011.

[51] L.Lipsky, Ed., *Querying Theory: A Linear Algebric Approach*, McMillan and Company, New York, 1992.

[52] M.F.Neuts, Ed., *Matrix-Geometric Solutions in Stochastich models: An Algorithmic Approach*, John Hopkins, University press, Baltimore, 1981.

[53] P. Joseph, H. Jason, and C. David, "Versatile low power media access for wireless sensor networks," in *Proceedings of the 2nd international conference on Embedded networked sensor systems*, 2004, pp. 95–107.

[54] J. S. Rellermeyer, G. Alonso, and T. Roscoe, "R-osgi: distributed applications through software modularization," in *Proceedings of the ACM/IFIP/USENIX 2007 International Conference on Middleware*, 2007.

[55] Q. Wang, "Traffic analysis and modeling in wireless sensor networks and their applications on network on network optimization and anomaly detection," *Network Protocols and Algorithms*, vol. 2, no. 1, pp. 74–92, 2010.

[56] S. Croce, F. Marcelloni, and M. Vecchio, "Reducing power consumption in wireless sensor networks using a novel approach to data aggregation," *The Computer Journal*, vol. 51, no. 2, pp. 227–239, 2007.

[57] H. Sabbineni and K. Chakrabarty, "Data collection in event driven wireless sensor networks with mobile sinks," *International Journal of Distributed Sensor Networks*, 2010.

[58] D. Puthal, B. Sahoo, and S. Sahrma, "Dynamic model for efficient data collection in wireless sensor networks with mobile sink," *International Journal of Computer Science and Technology*, vol. 3, no. 1, pp. 623–628, 2012.

[59] Sakiko Kawai and Takuya Asaka, "Event-driven wireless sensor networks using energy-saving data collection," in *Communications (APCC), 2012 18th Asia-Pacific Conference on*, 2012, pp. 300–305.

[60] M. Chen and M. Fowler, "Data compression trade-offs in sensor networks," *Proceedings of SPIE*, 2004.

[61] N. Kimura and S. Latifi, "A survey on data compression in wireless sensor networks," in *Information Technology: Coding and Computing, 2005. ITCC 2005. International Conference on*, 2005, vol. 2, pp. 8–13 Vol. 2.

[62] Z. A. Baig, M. Baqer, and A. I. Khan, "Sgsia-in-network data preprocessing for secure grid-sensor integration," in *e-Science and Grid Computing, 2006. e-Science '06. Second IEEE International Conference on*, 2006, pp. 160–160.

[63] Ming Xu, "Research and design of data preprocessing of wireless sensor networks based on multi-agents," in *Network Infrastructure and Digital Content, 2009. IC-NIDC 2009. IEEE International Conference on*, 2009, pp. 50–53.

[64] Haifeng Zheng, Shilin Xiao, Xinbing Wang, Xiaohua Tian, and M. Guizani, "Capacity and delay analysis for data gathering with compressive sensing in wireless sensor networks," *Wireless Communications, IEEE Transactions on*, vol. 12, no. 2, pp. 917–927, 2013.

[65] A. Masoum, N. Meratnia, Z. Taghikhaki, and P. J. M. Havinga, "Reward and punishment based cooperative adaptive sampling in wireless sensor networks," in *Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP), 2010 Sixth International Conference on*, 2010, pp. 145–150.

[66] R. Willett, A. Martin, and R. Nowak, "Backcasting: adaptive sampling for sensor networks," in *Information Processing in Sensor Networks, 2004. IPSN 2004. Third International Symposium on*, 2004, pp. 124–133, ID: 1.

[67] A. Jain and E. Y. Chang, "Adaptive sampling for sensor networks," in *Proceedings of the 1st international workshop on Data management for sensor networks: in conjunction with VLDB 2004*, 2004, pp. 10–16.

[68] Microchips Technology Inc., "Temperature sensor products," Available: http://www.microchip.com, accessed on 11-Feb-2013.

[69] OMEGA Engineering Inc., "Temperature sensors," Available: http://www.omega.com/cservice/omeganet.html, accessed on 11-Dec-2011.

[70] Applied Sensor Technologies, "Innovation in temperature measurement," Available: http://www.appliedsensortech.com/contact.htm, accessed on 11-Feb-2013.

[71] B. Bonnie, "Temperature sensing technologies," Available: http://www.microchip.com/stellent/idcplg, accessed on 11-Feb-2013.

[72] A. Hollinger, "Temperature sensing devices and optimization of temperature acquisition systems," Available: http://sesnorwiki.org/doku.php/sesnors/temperature, accessed on 11-Feb-2013.

List of Publications

1. Sonam Tobgay, Rasmus L. Olsen, Ramjee Prasad. "Architecture for running multiple applications on a single wireless sensor network: a proposal." *First International Conference on Advances in Computing and Communication,* Kochi, India, Springer Verlag, 2011, p. 37-45

2. Sonam Tobgay, Rasmus L. Olsen, Ramjee Prasad. "Adaptive Information Access in Multiple Applications Support Wireless Sensor Network." *8th International Symposium on Communication Systems, Networks and Digital Signal Processing (CSNDSP),*, Poznan, Poland, IEEE Xplore, 2012, p. 1-4.

3. Sonam Tobgay, Rasmus L. Olsen, Ramjee Prasad. "The Effect of Information Access Strategy on Power Consumption and Reliability in Wireless Sensor Network." *The 2nd IEEE International Conference on Parallel, Distributed and Grid Computing (PDGC'12).*, Shimla, India, IEEE Xplore, 2012, p. 325-330.

4. Sonam Tobgay, Rasmus L. Olsen, Ramjee Prasad. "Information Source Selection and Management Framework in Wireless Sensor Networks." *Wireless Personal Multimedia Communications Symposium (WPMC'13)*, Princeton, USA, IEEE Xplore, 2013 (presented).

5. Sonam Tobgay, Rasmus L. Olsen, Ramjee Prasad. "Adaptive Information Access to Remote Dynamic Information Source." *Submitted to Mediternean Journal of Computer Networks*, SoftMoter Ltd, 2013.

Mapping of Published Papers with Chapters

Paper Number	Chapter Numbers							
	1	2	3	4	5	6	7	8
Paper 1	☐	☐	☐	☑	☐	☐	☑	☐
Paper 2	☐	☐	☑	☐	☑	☑	☐	☑
Paper 3	☐	☐	☐	☐	☑	☑	☐	☑
Paper 4	☐	☐	☐	☑	☐	☐	☑	☑
Paper 5	☐	☐	☑	☑	☑	☑	☐	☑

Note: ☑ indicates that the contents of the papers are included in the corrsponding chapters.