On Maximizing the Efficiency of Multipurpose WSNs Through Avoidance of Over- or Under-Provisioning of Information

Vom Fachbereich Informatik der Technischen Universität Darmstadt
genehmigte

Dissertation

zur Erlangung des akademischen Grades eines Doktor-Ingenieur (Dr.-Ing.)
vorgelegt von

Vinay Mysore Sachidananda, M.Sc.

aus Mandya (Karnataka), India

Referenten: Prof. Dr. Kristof Van Laerhoven
Prof. Dr.-Ing. Matthias Hollick
Prof. Dr.-Ing. Ralf Steinmetz

Datum der Einreichung: 25.08.2014
Datum der mündlichen Prüfung: 13.10.2014

Darmstadt 2014, Dissertationen
D17

Abstract

A wireless sensor network (WSN) is a distributed collection of sensor nodes, which are resource constrained and capable of operating with minimal user attendance. The core function of a WSN is to sample physical phenomena and their environment and transport the information of interest, such as current status or events, as required by the application. Furthermore, the operating conditions and/or user requirements of WSNs are often desired to be evolvable, either driven by changes of the monitored phenomena or by the properties of the WSN itself. Consequently, a key objective for setting up/configuring WSNs is to provide the desired information subject to user defined quality requirements (accuracy, reliability, timeliness etc.), while considering their evolvability at the same time.

The current state of the art only addresses the functional blocks of sampling and information transport in isolation. The approaches indeed assume the respective other block to be perfect in maintaining the highest possible information contribution. In addition, some of the approaches just concentrate on a few information attributes such as accuracy and ignore other attributes (e.g., reliability, timeliness, etc.). The existing research targeting these blocks usually tries to enhance the information quality requirements (accuracy, reliability, timeliness etc.), regardless of user requirements and use more resources, leading to faster energy depletion. However, we argue that it is not always necessary to provide the highest possible information quality. *In fact, it is essential to avoid under or over provision of information in order to save valuable resources such as energy while just satisfying user evolvable requirements. More precisely, we show the interdependence of the different user requirements and how to co-design them in order to tune the level of provisioning.*

To discern the fundamental issues dictating the tunable co-design in WSNs, this thesis models and co-designs the sampling accuracy, information transport reliability and timeliness, and compares existing techniques. We highlight the key problems of existing techniques and provide solutions to achieve desired application requirements without under or over provisioning of information.

Our first research direction is to provide tunable information transport. We show that it is possible to drastically improve efficiency, while satisfying the user evolvable requirements on reliability and timeliness. In this regard, we provide a novel timeliness model and show the tradeoff between the reliability and timeliness. In addition, we show that the reliability and timeliness can work in composition for maximizing efficiency in information transport. Second, we consider the sampling and information transport co-design by just considering the attributes spatial accuracy and transport reliability. We provide a mathematical model in this regard and then show the optimization of sampling and information transport co-design. The approach is based on optimally choosing the number of samples in order to minimize the number of retransmission in the information transport while maintaining the required reliability. Third, we consider representing the physical phenomena accurately and optimize the network performance. Therefore, we jointly model accuracy, reliability and timeliness, and then derive the optimal combination of sampling and information transport. We provide an optimized model to choose the right representative sensor nodes to describe the phenomena and highlight the tunable co-design of sampling and information transport by avoiding over or under provision of information.

Our simulation and experimental results show that the proposed tunable co-design supports evolving user requirements, copes with dynamic network properties and outperforms the state of the art solutions.

Kurzfassung

Ein Sensornetz (Wireless Sensor Network, WSN) ist ein Netzwerk zufällig oder gewollt verteilter und miteinander verbundener Sensorknoten, welche nur über eingeschränkte Resourcen verfügen und mit minimaler menschlicher Überwachung agieren. Die Hauptaufgabe eines Sensornetzes ist das Sammeln von Messdaten über physikalische Ereignisse und die Übertragung der daraus folgenden und von der Anwendung erforderten Information über dieses Ereignis. Für den praktischen Einsatz ist es von Vorteil, wenn sowohl diese Anforderungen, als auch die Betriebsbedingungen der Sensoren im laufenden Betrieb verändert und an neue Bedingungen, sei es durch Veränderungen des beobachteten Ereignisses oder des WSNs selbst, angepasst werden können. Entsprechend ist die Bereitstellung der gewünschten Informationen gemäß der Nutzeranforderungen (z.B. bzgl. Genauigkeit, Zuverlässigkeit, Pünktlichkeit, etc.) bei gleichzeitiger Berücksichtigung ihrer Veränderbarkeit eines der Hauptziele bei der Einrichtung von WSNs.

Aktuelle Techniken betrachten die funktionellen Blöcke der Datenerhebung und der Übermittlung dieser Daten lediglich getrennt voneinander. Tatsächlich wird angenommen, dass der jeweils andere Block perfekt und verlustfrei operiert. Darüber hinaus betrachten einige Ansätze nur einzelne Informationsaspekte (z.B. Genauigkeit), während andere Aspekte schlicht ignoriert werden. Bestehende Forschungsarbeiten versuchen diese Blöcke für gewöhnlich dadurch umzusetzen, indem die Qualität der entsprechenden Informationsaspekte (Genauigkeit, Zuverlässigkeit, Pünktlichkeit) durch die Nutzung zusätzlicher Ressourcen verbessert wird, ungeachtet der tatsächlichen Nutzeranforderungen, was eine noch schnellere Erschöpfung der verfügbaren Energie zur Folge hat. Wir behaupten, dass es nicht notwendig ist immer die bestmögliche Informationsqualität bereitzustellen. Vielmehr ist es wichtiger die Bereitstellung sowohl zu geringer als auch zu hoher Qualität zu vermeiden, um wertvolle Ressourcen (wie Energie) dadurch einzusparen, dass Nutzeranforderungen exakt erfüllt werden. Um das zu erreichen, erarbeiten wir die Wechselbeziehungen der verschiedenen Nutzeranforderungen und zeigen wie ihre Umsetzung in Abhängigkeit voneinander gestaltet werden kann, um den Grad der Bereitstellung präzise steuern zu können.

Um die grundlegenden Probleme zu erkennen, die eine gemeinsame Gestaltung der verschiedenen funktionalen Blöcke (Co-Design) bestimmen und erforderlich machen, erstellt diese Arbeit Modelle für die Genauigkeit, die Übertragung und die Pünktlichkeit von Informationen und vergleicht bestehende Techniken. Wir zeigen die Hauptprobleme dieser Techniken auf und erarbeiten Lösungen, um die gewünschten Anwendungsanforderungen zu erreichen ohne Unter- oder eine Überversorgung von Informationen zu erzeugen.

Unser erster Forschungsbeitrag behandelt die steuerbare Übertragung von Informationen. Wir zeigen, dass es möglich ist die Energieeffizienz zu erhöhen und dabei dennoch sich verändernde Anforderungen bezüglich Zuverlässigkeit und Pünktlichkeit zu erfüllen. Dafür entwickeln wir ein neues Pünktlichkeits-Modell an

dem wir sowohl den Konflikt zwischen Zuverlässigkeit und Pünktlichkeit zeigen, als auch wie Zuverlässigkeit und Pünktlichkeit kombiniert werden können, um die Effizienz zu maximieren. Als nächstes behandeln wir die gemeinsame Gestaltung von Datenerhebung und Datenübertragung, wobei wir nur die Aspekte der räumlichen Exaktheit und der Übertragungszuverlässigkeit betrachten. Hier entwickeln wir ein mathematisches Modell zur Optimierung des Co-Designs von Genauigkeit der Datenerhebung und Zuverlässigkeit der Datenübertragung. In diesem Ansatz wird die Anzahl der Stichproben optimal so gewählt, dass die maximale Anzahl aller Übertragungen minimiert wird, während die gewünschte Zuverlässigkeit dennoch gewährleistet werden kann. Schließlich betrachten wir die genaue Repräsentation des physikalischen Phänomens und optimieren die Gesamtleistung des Netzwerks. Dafür erstellen wir ein Modell, welches sowohl Genauigkeit, Zuverlässigkeit als auch Pünktlichkeit berücksichtigt und leiten daraus die optimale Kombination von Stichproben und Übertragungsversuchen ab. Genauer gesagt ermöglicht dieses Modell genau jene Sensorknoten auszuwählen, mit denen das zu messende Ereignis am effizientesten repräsentiert werden kann, sodass die Nutzeranforderungen eingehalten werden können, ohne dass zu viele Informationen bereitgestellt werden. Damit unterstreichen wir die Bedeutung und Praktikabilität eines steuerbaren Co-Designs.

Unsere Simulation sowie die Ergebnisse unserer Experimente/Berechnungen zeigen, dass das vorgeschlagene Co-Design die sich ändernden Nutzeranforderungen unterstützt, der dynamischen Netzwerkeigenschaften gerecht wird und den Stand der Technik übertrifft.

Acknowledgements

Almost six years have passed since I started the work that eventually crystallized into this PhD thesis. Looking back from today's perspective, I must admit that there was a lot of hard work but fun was always present. It was hard as I had to learn how to do research, how to present my work and how to be critical with other people's work when it came to writing conference reviews. It has been fun when socializing with the members of GKmM, our group or celebrating birthdays over home-made, international-flavor dishes or cakes.

There are many people that helped me to whom I am very thankful. First I'd like to express my big thanks to my mentor, from whom I learned that there is always a higher gear, but also that research can be (sometimes) fun. Above all an advice not to give up and be strong. Thanks a lot, *Oskar*!. I would like to specially thank my supervisor, from whom I got the timely advise and support which gave me higher confidence and ability to make right decisions. Thanks a lot, *Kristof*!. I also like to thank to my Post-doc *Majid*, who has always supported and helped me with my research. I would be really grateful to *Prof. Dr. Reiner Hähnle*, *Prof. Dr.-Ing. Ralf Steinmetz*, *Prof. Dr.-Ing. Matthias Hollick* and *Prof. Alejandro Buchmann.* Today all these things are possible from their greatest support.

I'd like to thank *Neeraj*! for teaching me everything I know about doing research. Then, I am very thankful to all past and present DEEDS group's members for helping me, listening to my boring presentations and giving me feedback on early drafts of my papers. Hoping not to forget anyone of the past and present guys, many thanks to *Robert, Ripon, Brahim, Peter, Dan, Marco, Faisal, Matthias, Azad, Piotr, Hamza, Stefan, Daniel, Arda, Reza, Kubilay, Thorsten* and *Jesus.* Also, special thanks go to *Sabine* and *Ute* for helping me with various paperwork, for correcting and improving my German and all other circumstances related to living in Germany.

I am very thankful to all ESS group's members for helping me and giving me feedback on my presentations. Moreover, thank you for great lunch time. Hoping not to forget anyone, many thanks to *Philipp, Eugen, Marko, Agha, Nagi and Manuel.* Also, special thanks go to *Mrs. Gabriele Ploch*, *Mrs. Heike Schmitt-Spall* and *Mrs. Ursula Paeckel* for helping me with various paperwork. I would also like to thank my wonderful friends, *Philipp, Robert, Rodrigo, David, Anh, Stefan and Memo* for their unlimited support.

I would specially thank my brother, sister and sister in-law for their continual support and love which spanned the geographical distance separating us in all these years. I'd also like to thank my lovely sweethearts *Spoorthy, Aaadhya, Megha and Galina* for their support and great love. Thank you, *Vicky* and *Shubha*!

I saved the best for the last: I'd like to *dedicate* my thesis to my wonderful and lovely parents, I really thank them for giving me everything and being there for me in every moment. I love you both a lot and you are the world to me. Thank you, *Mom* and *Dad*!

Contents

4 Problem Statement 49

5 Tunable Information Transport Concerning Reliability and Timeliness 59

6 Optimizing the Sampling Accuracy and Transport Reliability 71

List of Figures

List of Tables

Chapter 1

Introduction

The field of wireless communication and sensing technologies has led to a revolutionized emerging field of Wireless Sensor Networks (WSNs). A WSN is a distributed collection of sensor nodes, which are usually battery-powered but capable of operating with minimal user attendance. Wireless sensor nodes operate in a cooperative and distributed manner. WSNs offer significant advances over traditional wired networks. WSNs can be applied in many scenarios because of their flexibility, cost-effectiveness and ease of deployment [Akyildiz et al., 2002b].

Generally, a WSN consists of static sensor nodes, which can be deployed in orderly or random fashion. A WSN has low processing capabilities with limited power. The communication is often done via radio links. The usual on-board sensors comprise of temperature, humidity, pressure, light, among others. The ad-hoc deployment of sensor nodes constitutes to cooperate to form the wireless network. The classical operation of a WSN is to sample physical attributes, process the sampled data in-the-network to extract useful information such as events, regional phenomena distributions etc., and transport the user information via hop-by-hop communication to a powerful base called sink [Sachidananda et al., 2010]. The proclaimed information is further utilized by different end users for making decisions. Here, data refer to basic monitored facts/chunks (e.g., sensor readings) and information is the collated and interpreted data systematized by purposeful acumen and processing required for an application (e.g., event occurrence).

Usually, a deployed WSN interacts with the physical environment to report the status/event to the end user. As the end user requires desired information from the WSN, it is inefficient to collect irrelevant data for any decision making. For example, ecologists might need the information of temperature and humidity of an entire national park. Providing subset of information about the national park leads to under provision of information. In addition, as energy plays an important role in WSN, it is highly inefficient to transport all the information, as communication is the dominant energy consumption operation [Akyildiz et al., 2002b]. For example, for representing a physical phenomenon such as fire, it is efficient to sample just the phenomenon area and avoid over provision of information by sampling the non-phenomenon area of the WSN. Hence, the challenge is to avoid the over provision of information to save resources such as energy and avoid under provision of information to have satisfied users.

WSNs have an endless array of potential applications in both military and civilian applications. WSNs have been proposed for a variety of applications such as localization and tracking for military purpose, motion detection for understanding earthquake patterns, intrusion detection to prevent theft, monitoring the drug administered to the patients by health applications and habitat monitoring for studying physical environments [Xu, 2002]. Other applications include sensors on buildings, sensors on vehicles, smart home, inventory management etc. One common feature shared by all of these critical applications is the vitality of the information.

The applications are interested in the desired information from the network. Sampling the desired information accurately and reliably and timely transporting the desired physical attribute is one of the key requirements in WSNs. Usually, the basic sampling methods provide samples of the data with limited consideration of the user requirements [Munir et al., 2007]. The sampling strategies alone are not sufficient for understanding the requirements of the user and causes over or under sampling. For example, if the user is interested in knowing the event map for the evolving phenomenon such as fire, we do not need the sampling of the whole deployment area, but only the border nodes of the phenomenon area need to sample. Thus, for desired sampling, various mechanisms need to be developed to satisfy the user requirements concerning the sampled information. Consequently, transporting the desired sampled information needs a design for reliably and timely transport of the information. Basic network routing provides the paths between sensor nodes and the sink for the information delivery. Nevertheless, the basic routing strategies does not guarantee to satisfy the user requirements on information transport [Sankarasubramaniam et al., 2003]. Hence, various mechanisms need to be designed on top of readily available basic solutions.

Typically, in WSNs, the state of the art considers the sampling and information transport isolated [Sachidananda et al., 2010]. Sampling protocols assume that the information transport to be perfect in delivering the sampled data to the end user [Szczytowski et al., 2010]. On the other hand, information transport considers the sampling to be perfect in providing the required information for delivery to the end user [Shaikh et al., 2010]. This isolation of sampling and information transport leads to negotiations with user requirements and hinders the deployed WSN by delivering either redundant information or under provision of information to unsatisfied user. In addition, a critical event detection application may require high sampling accuracy and transport reliability for creating event maps [Khelil et al., 2009]. Instead, a non-critical monitoring application can tolerate some loss of information and relax timeliness requirements. The evolving application requirement needs an optimal balance and co-design of the sampling and information transport. Therefore, in that respect is a necessity and the challenge in creating the co-design of sampling and information transport to optimize the information delivery by saving resources such as energy.

The commercial use of WSNs has increased with the emerge of Internet of Things (IoT) [Atzori et al., 2010]. WSNs are becoming integral part of the ubiquitous and pervasive systems, grid systems and web services [Delicato et al., 2003]. The future trend of WSNs is driving towards a multimodal environment with the integration of different systems such as robotics for rescue scenarios [http://www.gkmm.tu-darmstadt.de/] [Khelil et al., 2011]. The evolution of WSNs has led many applications to run concurrently. In addition, the same WSN application may change its requirements over time [Kuorilehto et al., 2005]. Moreover, different applications have different requirements for information. For example, biologists may tolerate timely delivery of data during monitoring the environment, firefighters need more accurate data for understanding the evolving phenomenon such as fire and engineers need more reliable data for monitoring the instruments in industrial applications. Considering the evolving requirements, the sampling and information transport should be tunable according to the application requirements.

WSNs, due to wireless communication and harsh deployment environments, are subject to a wide range of operational perturbations affecting sampling and information transport. The perturbations caused by collisions, contention and congestion lead to a deviation between the attained requirements and user requirements. If the attained information requirements are higher than required, then the valuable resources are wasted in the network. Conversely, if the attained information requirement is lower than desired, the information usefulness for the application is compromised. From Fig.1.1, there could be various designs for WSNs, some pertaining to over provide the information (Design 2) and some under providing the information (Design 1). Even so, in our work we want to achieve a design such that we can provide the desired user requirements without wasting resources (Design 3).

In our work the challenge is in developing a design for the WSN protocol suite to avoid under or over provision of information. In addition, co-design of sampling and information transport, provide tunability and adapt according

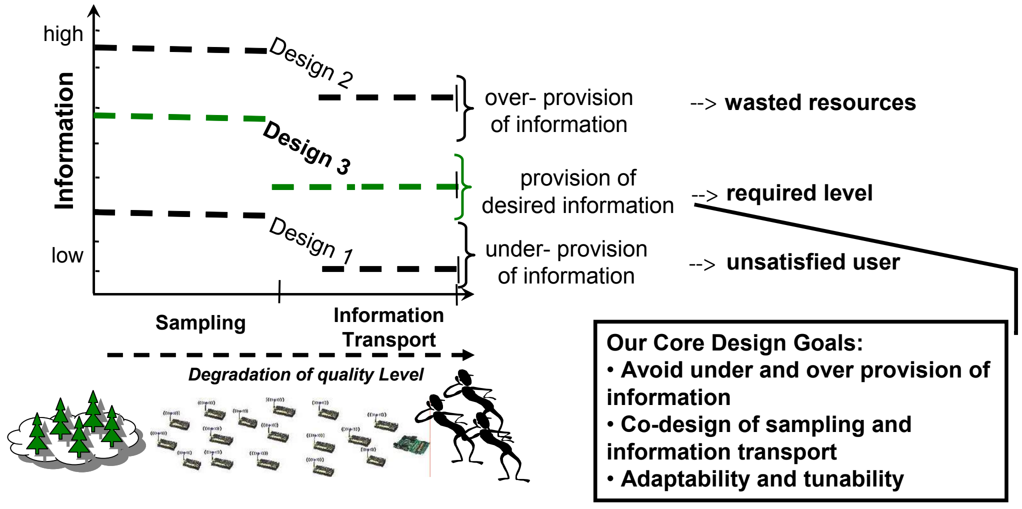


Figure 1.1: On exact provisioning of information in WSN

to various user requirements.

The remainder of this chapter is organized as follows. First, we present the functional blocks as the basis for the next sections and emphasize on the sampling and information transport functional blocks. Second, we present the motivation for further aspects of the thesis. Next, we present the main ideas driving the research in this thesis via the challenges and then present the design objectives driving the thesis. Following, the thesis is confronted in a nutshell. Later, we summarize the thesis targets refined as a set of research questions followed by the answers in the form of research contributions. Finally, the structure of the remaining thesis is outlined.

1.1 Functional Blocks

Currently, there exists justifiable work about varying aspects of WSNs activities such as data collection/sampling [Bisdikian, 2007], aggregation to further operation etc. Currently, information (quality) is addressed under different roofs, i.e., regarding fusion [Lin et al., 2008], data impact, decision making [Prasanth et al., 2004], degradation, miss association infused information and data level acquisition [Kessel, 2006] [Wälchli and Braun, 2009] [Chong and Kumar, 2003]. We refer to quality as the degree or grade of excellence, and to information quality as the quality experienced/perceived by the user concerning the received information, which (may) fully accomplish the user evolvable requirements while saving valuable resources such as energy and bandwidth. When WSNs is viewed in these aspects there are different dimensions and we classify these aspects in four different views of sampling, in-network processing, and information transport and sink operations. This classification shows that information should be considered as one important aspect throughout all operations.

A cross-layer and cross operation design considers blocks of sampling, in-network processing (compression, aggregation etc.), and information transport and sink operations. To exemplify the benefits of such design paradigms without loss of generality, in our work we consider sampling and information transport as the core functional blocks for the tunable co-design and to avoid under and over provision of information. We consider these building blocks because we expect that information should satisfy user requirements from the time of sampling till operations at the sink are conducted.

Sampling : is the process of collecting raw data samples at sensor nodes (sampling in time and space domain) [Meliou et al., 2006] [Szczytowski et al., 2010], this operation is necessary in WSN. Temporal sampling quality depends on the sensor quality and the sampling frequency. Spatial sampling has a vital role in understanding the spatial distribution of the phenomenon. As we are interested in accurately representing the physical phenomenon, in our work we consider spatial sampling.

In-Network Processing (Information Extraction) : in WSNs data collection is done by collecting raw data samples, and this raw data is processed in order to extract useful information. Aggregation is one of the most popular in-network processing techniques [Krishnamachari et al., 2002] [Zhao et al., 2003].

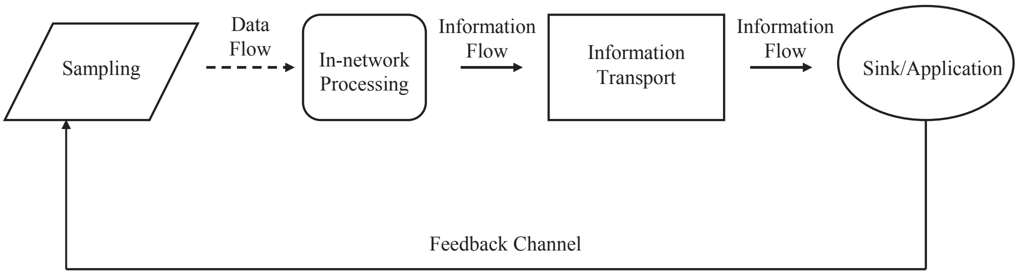


Figure 1.2: WSN functional blocks and communication channels

Information Transport : is the end-to-end transport/routing of the information from the source (where it is generated/extracted) to the sink [Al-Karaki and Kamal, 2004] [Shaikh et al., 2007]. Most of the current existing routing/transport protocols and their techniques assume that the information coming from the source is trustworthy and reliable [Gnawali et al., 2009].

Sink Operations : information is managed at the sink for further operations such as decision making. Information Risk Management (IRM) [Chang et al., 2005] is one of the approaches to minimize the risks such as information misunderstanding and inefficiencies of metrics which may affect learning quality.

In this thesis, we consider sampling and information transport as the core functional blocks. We emphasize on these blocks to move towards the optimal co-design in WSNs. Moreover, we consider the sink operations as an application or user who will use the information available at the sink for the decision making. Nevertheless, as it is very complex to reach the co-design of sampling and information conveyance, we leave in-network processing for the future workplace.

1.2 Motivation

A key functionality of WSNs consists in obtaining and transporting the information of interest (e.g., status/event) required by the applications. Despite a wide range of perturbations, the applications running on WSN also specify desired quality requirement levels (accuracy, reliability and timeliness) on the desired information. Consequently, application requirements, possibly changing over time and of tunable levels over an application, are stipulated on the sampling and information transport.

Some applications may be interested in acquiring the information periodically, while others may be interested in getting the information when some phenomenon of interest has occurred in the network. Accordingly, the generated information may also have a spatial correlation corresponding to the phenomenon of interest (e.g., the perimeter of the phenomenon area [Ghosh and Das, 2008] on the spatial distribution of the phenomena [Szczytowski et al., 2010]). Moreover, the perceived information should satisfy the application requirements (e.g., accurate form and location of the event perimeter). In integration, future WSN deployments should allow for varied concurrent applications. Usually, these applications need varied information and have evolvable requirements.

The existing solutions [Sachidananda et al., 2010] are not designed to explicitly consider evolvable application requirements, as their main design driver is to efficiently maximize the attained quality levels (accuracy, reliability and timeliness). The existing approaches over-utilize the WSN resources (e.g., energy) even when the application does not require that level of quality.

In addition, the deployment of one of the existing solutions on sensor nodes, limits the operational conditions under which sensor nodes can sample and transport information. For example, an existing solution which provides high accuracy or reliability will consume resources unnecessarily even when applications do not have high requirements. Similarly, if we deploy a solution which considers just the isolation of functional blocks, this leads to inefficient use of resources such as energy and over or under provision of information. The lack of co-design limits the communication between the functional blocks and considers other blocks to be perfect. Co-design violates these principles and uses information from different functional blocks to improve the network performance and/or lifetime.

Considering the key operations of sampling and information transport along their quality attributes (accuracy, reliability and timeliness), and delivering the gathered information with the application required quality is the key focus of our work. It is trivial to understand that, it is not necessary to provide the highest quality levels without adapting to the user requirements. Thus, for maximizing the efficiency of WSN and avoid over or under provision of information, the existing solutions are not sufficient. Therefore, necessity for a tunable co-design which takes care of all possible situations is trivial. Such an integrated approach should maximize the supported operating conditions and provide efficient mechanisms to maintain the application requirements.

1.3 Research Challenges

One of the core functions of a WSN is to observe and account status/events which can be assimilated meaningfully and responded to only if the accurate and qualitative data about the status/events is recognized. In interpreting the application context, information from sensor nodes plays a critical part. Usually, the WSN applications are data centric, i.e., they are deployed to interact with the physical environment and report the phenomenon of interest to the user via a sink.

More often than not, in WSNs the wireless communication is known for its unpredictable nature due to environmental obstacles which causes lower signal strength and consequently the various factors like reflection, scattering and dispersions reduce the radio range. Furthermore, the natural calamities cause sensor nodes to break and also sometimes the sensor nodes can be stolen causing the network coverage problem. On the other hand, due to the irregular nature of communication, the link conditions are dynamic causing the communication disruption. As communication causes the major energy consumption in WSNs, this hinders the nodes to die very fast hindering the WSN deployment. Moreover, as higher energy expenditure is due to transmission causing the short operating lifetime of WSNs, low duty cycling operation plays an important character for longer WSN life.

The radio range of nodes in WSNs pertains for deploying more number of sensor nodes. The WSNs consist of hundreds of nodes, but the sensor nodes are not cheap and are more expensive to be deployed. Deployment of WSN is a labor intensive and cumbersome activity as we do not have influence over the quality of wireless communication and also the real world puts strains on sensor nodes by interfering during communications. Sensor node calibration is also very important with larger deployment of sensor nodes. The topology of a WSN is dynamic due to perturbations in the network. Some of the important perturbations are communication failures. Failures relevant to the information transport include message loss which directly impacts the reliability and timeliness of the WSN. On the other hand, the node failures are the other perturbations causing unavailability of sensor nodes and congestion causing message loss due to high network load. Some data centric applications encounter data collection interruptions as the sensor nodes are volatile as the sensor nodes may become depleted and links between various sensor nodes may go down at any point of time. Hence, the WSN is surrounded by many issues regarding dynamic conditions.

With the multiple WSNs deployment and varying applications, there are multiple users who use different WSNs deployment and applications. The users have varying application requirements and the requirements of the users also vary based on different WSNs deployment. Fig.1.3 shows that there are multiple users accessing information from the same WSN deployment and also from different WSN deployment for different applications. Furthermore, as the IoT has been into light, the sensor nodes are connected to the Internet cloud and different users access different information from the deployed WSN. Therefore, there are multiple examples of such users using the Internet cloud for applications such as smart home, smart cities and buildings.

In WSNs, the various applications and users drive the specific information needs. The user requirements regarding information are evolvable having specified information with a certain quality. Also the achievable information quality is evolvable according to the operating conditions such as network and environmental conditions. Accordingly, the WSN functional operations should be designed while considering the fluctuating operating conditions and the user's evolvable requirements on information quality.

Common to all these observations is that in WSNs the operating conditions and/or user requirements are often desired to be evolvable, whether driven by changes of the monitored parameters or WSN properties of the

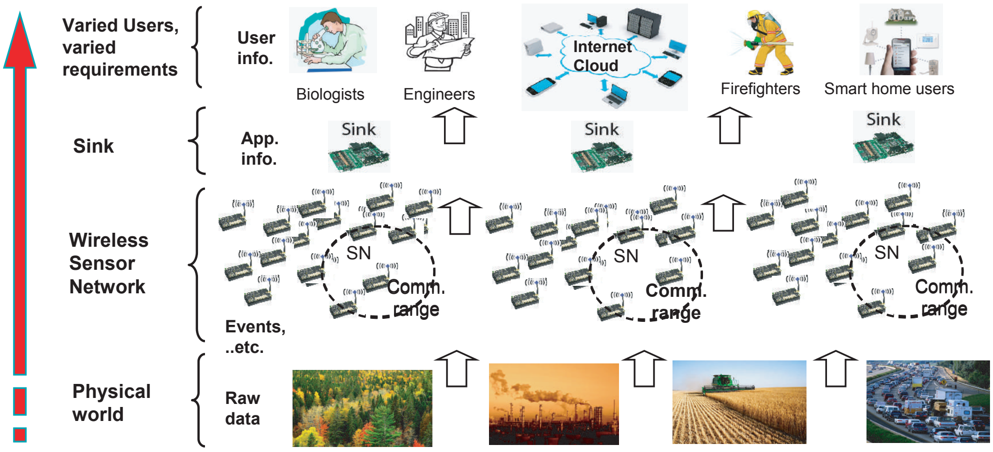


Figure 1.3: Expected integration of WSNs into IoT/sensor webs

configuration, structure, communication capacities, node density, and energy among many others. Moreover, there are multiple applications and varying WSNs deployment (Fig.1.3). While considering the evolvability of user requirement, multiple applications and varying WSN deployment, delivering the required information with the specified quality (accuracy, timeliness, reliability etc.) defined by the user and avoiding under or over provisioning of information constitutes a key objective of WSNs.

As it is vital to support the evolvable user requirements, there is a requirement for tunable adaptive co-design. Most of existing approaches address isolated functional blocks. Unfortunately, the approaches neglect the importance of the one or the other block as most of them neglect to consider co-design of these blocks to satisfy evolvable user requirements. They indeed assume the other blocks to be perfect in maintaining the highest possible information quality contribution. In addition, some of the approaches just concentrate on a few information attributes such as accuracy and ignore other attributes such as reliability and timeliness.

The existing research based on these functional blocks always tries to enhance the information leading to negotiate with quality attributes like accuracy and use more resources leading to energy depletion or activities affecting the deployed network. Hence, it is not efficient to use the naive solutions (processing techniques, protocols etc.) from the source to the sink to deliver the information with the user evolvable requirement. It is not always necessary to increase the quality, but sometimes to decrease it to save valuable resources such as energy and bandwidth, and increase the timeliness of information delivery without under-performing the required quality indicators/metrics such as accuracy. To design and deploy a WSN, one should consider the co-design to achieve the required user level while maximizing efficiency. Consequently, one requires reviewing and improving the adaptive and tuning capabilities of individual blocks and also avoiding under-over provision of information.

1.4 Design Objectives

In the following we discuss the design considerations for sampling and information transport co-design in WSNs. First, we outline the design objectives that should be followed by the sampling and information transport co-design to cope with the distinct properties of WSNs.

As stated before, the highly dynamic nature of WSNs and varying application requirements has a great impact on the design of sampling and information transport co-design. We believe that tunability, adaptation, co-design, scalability, perturbations tolerance, resource-awareness and decentralization are the key design issues for WSN applications in general and for sampling and information transport in particular.

Tunability: Due to varying, evolving and statistical nature of accuracy, reliability and timeliness requirements of WSN applications, the co-design should be able to ensure tunable spatial accuracy and tunable reliability and timeliness of information transport. The different mechanisms should adapt and tune in order to fulfil the desired user requirements.

Adaptation: Due to the diversity of WSN applications and the continuously evolving network conditions, a generalized optimal solution that is applicable for most (and ideally for all) network and application scenarios is needed. Thus, online adaptation to the key WSN characteristics should be considered towards the development of the co-design.

Co-design: Conventionally, in WSNs there exists multiple users, multiple applications, multiple dynamic conditions. However, it's very trivial that WSNs are resource constrained and it is required to use the same WSN deployment for varying aspects. As WSN is seen quite isolated with the functional blocks operations, it is important to combine the most important blocks such as sampling and information transport for a co-design operation such that we can enhance the network performance.

Scalability: Generally, WSNs are envisioned for large scale deployments. Accordingly, the co-design should scale in terms of number of nodes efficiently without excessive overhead and should provide simple mechanisms for resource constrained sensor nodes to sample and reliably and timely transport the information.

Fault Tolerance: The failures are the norm rather than the exception in WSNs. Thus, the co-design should deal with disruptions and unpredictable network conditions and other perturbations.

Resource Efficiency: The co-design and its mechanisms are supposed to run on sensor nodes with limited energy, computational power and memory. Consequently, the mechanisms should be frugal by design and resource efficient.

Decentralization: Conventionally, a sink is utilized to centrally manage the different operations of the WSN at the cost of the huge overhead of communication. With the evolving network conditions the central role of sink becomes more inefficient. Therefore, efficient decentralized or localized mechanisms should be developed for the co-design.

1.5 Thesis in a Nutshell

Our core goal of the thesis is in providing the optimized tunable co-design of sampling and information transport by avoiding under or over provision of information. Achieving the tunable co-design is complex, hence, we proceed step wise to master this complexity. Consequently, we lay a solid foundation by first identifying the existing approaches in WSNs and show that these approaches lack the tunable co-design. Furthermore, we focus on identifying the existing information attributes and define new attributes such as tunability.

From Fig.1.4, there could be different users with varying requirements (user1, user 2 and user 3). Design 1 satisfies the user requirements for sampling, however, the efficiency with information transport is ignored. Design 2 pertains to satisfy the user requirements of information transport by neglecting the functional block sampling. The Designs 1 and 2, lacks the tunability with evolving user requirements. In addition, the different attained levels with different efficiency show the lack of co-design with Design 1 and 2. In this thesis, we work on the Design 3, where the evolvable user requirements are satisfied with tunability and co-design. The user required level and attained level of Design 3 also gains a better efficiency leading to save resources such as energy, the gains are explained in the Chapters 5, 6 and 7.

With the considered functional blocks sampling and information transport, we first focus on information transport and its attributes reliability and

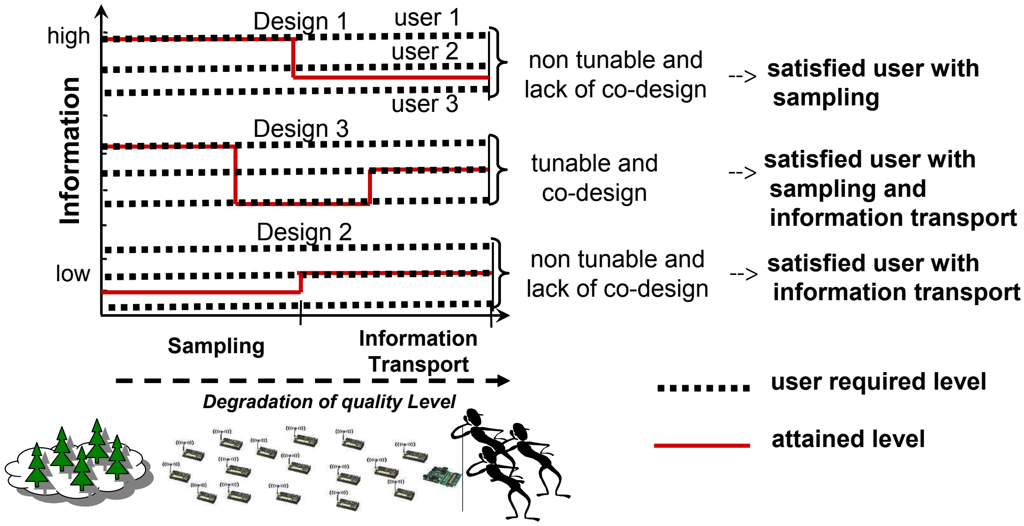


Figure 1.4: On exact provisioning of information in WSN

timeliness. Equally, we likewise want to satisfy evolvable user requirements on reliability and timeliness, we demonstrate the tunability of the information transport by trading off approach. Furthermore, we provide the tunable information transport with composite reliability and timeliness and maximizing efficiency. We strain to avoid under or over provision of information and resources by adapting the optimal number of retransmissions on a per hop basis with delay compensation and path split techniques. For instance, it is meaningless to retransmit the information if the timeliness requirement cannot be ensured; in that instance it is more beneficial to transmit on multipath that still ensures timeliness.

As the principal challenge in our thesis is in providing the tunable co-design, we induce the first attempt on considering the sampling and information transport cross fertilization. Furthermore, we go forward with optimizing the sampling accuracy and transport reliability. With the achieved co-design, we reduce the total number of retransmission by selecting the optimal number of samples to be sent to the sink. The information transport tunes to provide the desired reliability, established on the required minimum number of samples,

Finally, we consider representing the physical phenomena and optimizing the network performance. The approach is first based on modeling each single attribute (accuracy, reliability and timeliness) and then towards the optimal combination of sampling and information transport. Furthermore, the generic co-design algorithm is presented for the optimized tunable co-design to avoid under or over provision of information. We validate all our approaches viability through extensive simulations for a wide range of requirements and network conditions.

1.6 Thesis Research Questions and Contributions

In this section we briefly revisit the research targets in the form of research questions and summarize the thesis contributions. The research questions driving the research presented in this thesis belong to different aspects of achieving and maintaining the optimal tunable co-design.

Research Question 1 (RQ1): *How to provide the online tuning of reliability and timeliness in information transport?*
Chapter 5 discusses the aspects of information transport and focuses on different characteristics it should offer. Consequently, a tunable information transport providing user evolvable requirements is developed. Chapter 5 discusses the tunability aspects in information transport and the tradeoff provided with reliability and timeliness. Accordingly, we provide an online tuning design for the information transport. We evaluate the proposed tunable information transport and show its validity.

Contribution 1 (C1) – Achieving Tunable Timeliness and Composite Tradeoffs with Tunable Information Transport: We design a distinctive solution to provide tunable timeliness by exploiting the exponential and uniform model by allocating deadlines along the path from the source to the sink. On the other hand, we show how the tradeoffs can be achieved with reliability and timeliness. Stepwise we detail the composite tradeoffs by allocating reliability and timeliness by exploiting the retransmission and multipath techniques. To provide the tunabilty of reliability and timeliness in the information transport we provide an approach to satisfy the user evolvable requirements and maximizing efficiency. The approach is evaluated using widely accepted simulator TOSSIM, validating its applicability and usefulness. (RQ1)

Research Question 2 (RQ2): *How to provide optimized sampling accuracy and transport reliability in WSNs?*

The first attempt to co-designing the sampling and information transport blocks are emphasized. The attributes, accuracy and reliability are successfully bound to contribute to the optimization problem. Chapter 6 address this question by introducing and then discussing the conceptual foundation for optimized sampling accuracy and transport reliability with the achieved gains.

Contribution 2 (C2) – Optimizing Sampling Accuracy and Information Transport Reliability: In order to provide the sampling and information transport co-design, we show how the spatial correlation can be exploited in order to bind the two functional blocks. The optimal co-design of the sampling accuracy and transport reliability is achieved while maximizing the energy efficiency. The usefulness of the approach and the validation of the work are achieved through analytical and experimental evaluation. (RQ2)

Research Question 3 (RQ3): *How to provide fine tuning of accuracy and representing the physical phenomena accurately with optimized network performance?*

Chapter 7 discusses this issue and provides the basis for exploiting the spatial correlation in accord to information transport. The optimal combination of sampling accuracy and transport reliability is highlighted and shown by providing energy efficient network performance. Furthermore, in Chapter 7, we present the efficient modeling, representing the accurate physical phenomena by considering the realistic characteristics of WSNs. Accordingly, the attribute accuracy, reliability and timeliness are combined and an optimal cross-operation model is shown. The effects of the multi-attribute correlation and its effects are highlighted.

Contribution 3 (C3) – Exploiting Spatial Correlation and Optimizing Network Performance: In order to provide optimized network performance and to achieve the composite tradeoffs between accuracy, reliability and timeliness, we provide the theoretical approach on the effects of the multi-attribute correlation. We show that considering the co-design of the sampling and information transport is significant and necessary. In addition, we also emphasize that the optimal co-design with the user requirements leads to the desired WSN deployment. (RQ3)

1.7 Publications Resulting from the Thesis

The work reported in this thesis is supported by several international publications:

- **Vinay Sachidananda**, David Noack, Abdelmajid Khelil and Neeraj Suri, *Optimized Co-design of Spatial Sampling and Information Transport in Wireless Sensor Networks*, Proceedings of the Special issue in Telecommunication Systems Journal (TSMJ), (under review), 2013.
- **Vinay Sachidananda**, Abdelmajid Khelil, David Noack and Neeraj Suri, *Information Quality Aware Co-design of Sampling and Transport in Wireless Sensor Networks*, Proceedings of the sixth IEEE International Conference on Wireless and Mobile Networking Conference (WMNC), April 2013.
- **Vinay Sachidananda**, Abdelmajid Khelil, David Noack and Neeraj Suri, *Sampling and Transport Co-design in Wireless Sensor Networks*, Proceedings of the tenth IEEE International Conference on Wireless On demand Network Systems and Services (WONS), March 2013.
- **Vinay Sachidananda**, Abdelmajid Khelil, Dhananjay Umap, Matthias Majuntke, Neeraj Suri, *Trading Transport Timeliness and Reliability for Efficiency in Wireless Sensor Networks*, Proceedings of the tenth IEEE International Conference on Networking Sensing and Control (ICNSC), April 2013.
- **Vinay Sachidananda**, David Noack, Abdelmajid Khelil and Neeraj Suri, *On Co-modeling the Sampling and Transport in Wireless Sensor Networks*, Proceedings of the eleventh GI/ITG KuVS Fachgespräch "Sensornetze" (FGSN), September 2012.
- **Vinay Sachidananda**, Abdelmajid Khelil and Neeraj Suri, *Information Quality Aware Transport for Wireless Sensor Networks*, Proceedings of the European Conference on Wireless Sensor Networks (EWSN), February 2012.
- **Vinay Sachidananda**, Abdelmajid Khelil and Neeraj Suri, *Quality of Information in Wireless Sensor Networks: A Survey*, Proceedings of the fifteenth International Conference on Information Quality (ICIQ), November 2010.

Additionally, the author has been involved in the following publications that are not directly covered by the thesis:

- Philipp M. Scholl, Stefan Kohlbrecher, **Vinay Sachidananda** and Kristof Van Laerhoven, *Fast Indoor Radio-Map Building for RSSI-based Localization Systems*, Proceedings of the ninth International Conference on Networked Sensing Systems (INSS), June 2012.

- **Vinay Sachidananda**, Diego Costantini, Christian Reinl, Dominik Haumann, Karen Petersen, Parag S. Mogre, and Abdelmajid Khelil, *Simulation and Evaluation of Mixed-Mode Environments: Towards Higher Quality of Simulations*, Proceedings of the second International Conference on Simulation, Modeling and Programming for Autonomous Robots (SIMPAR), November 2010.

- Piotr Szczytowski, Faisal Karim Shaikh, **Vinay Sachidananda** Abdelmajid Khelil and Neeraj Suri, *Mobility Assisted Adaptive Sampling in Wireless Sensor Networks*, Proceedings of the ninth International Conference on Networked Sensing Systems (INSS), June 2010.

- Andrey Somov, **Vinay Sachidananda** and Roberto Passerone, *A Self-powered Module with Localization and Tracking System for Paintball*, Proceedings of the third International Workshop on Self-Organizing Systems (IWSOS), December 2008.

1.8 Thesis Structure

The rest of the thesis follows the structure of the research questions described earlier:

Chapter 2 classifies and surveys the state of the art and practice to show the lack of co-design in WSNs. Correspondingly, Chapter 2 presents the information assessment. Then, we compare the existing approaches based on functional blocks, attributes and metrics and show the approaches are lacking a tunable co-design. In addition, we also present the sampling and information transport schemes.

Chapter 3 defines and discusses the system and perturbation model used throughout this thesis. Furthermore, hotspot and sampling and information transport models are presented. Next, we abstract the key performance indicators of our thesis and the driving force for considering the co-design. Finally, the design requirements of our thesis are provided.

Chapter 4 provides a precise problem statement to show the important goals to be achieved in the thesis. The core problem is then further divided into three sub problems for further analysis.

Chapter 5 introduces our tunable information transport. First, we investigate the major considerations for the design of generalized solution through discussing the illustrative scenarios for information transport in WSNs. Next, we define and elaborate the tunable information transport in accordance with reliability and timeliness. We show the viability of our work with extensive simulations.

Chapter 6 depicts the advantages of exploiting the spatial correlation and investigates a methodology for sampling and information transport co-design. We model the co-design and compare the analytical solution to the presented sampling and transport co-design algorithm. We show the viability of our work with extensive simulations for varying network conditions.

Chapter 7 presents the representation of physical phenomena with optimal modeling of the accuracy, reliability and timeliness. We express the sampling and information transport co-design as an optimization problem with multi-attribute interaction. In addition, we present the generic co-design algorithm.

Chapter 8 provides the performance evaluation and experimentation of our thesis. We show three case studies with varying user requirements, dynamic system requirements and evolving phenomenon.

Chapter 9 concludes the thesis by revisiting the value of the contributions presented in this thesis. Finally, future research directions opened by this thesis are presented.

Chapter 2

State of the Art: Classification and Comparison

As an important basis in the context of the research presented in this thesis, this chapter starts by discussing the different approaches lacking tunable co-design in WSNs. Accordingly, the chapter provides the information assessment and classifies the attributes and metrics. Next, based on the classification we survey the state of the art. At last, we compare the existing solutions to get insights and highlight the drawbacks which hinders in the co-design of sampling and information transport. The information assessment and the comparison of state of the art presented in this chapter constitute one of the preliminary efforts to understand the approach towards the co-design of this thesis.

We believe that the tunable co-design is significant in WSNs and considered as the center of attraction for users, designers, decision makers, application planners etc. There are no previous efforts detailing the attributes/metrics/techniques related to information. We take the opportunity to review the snapshot of the state-the-art, and to discuss the pros and cons of the different existing approaches that lack tunable co-design.

This chapter forms the background and the context for the research questions posed and puts the contributions presented into perspective. The chapter concludes with a discussion on design guidelines for the efficient co-design of sampling and information transport in WSNs.

2.1 Information Provisioning

Overall, this chapter targets the ongoing research activities in a manner which provides the foundation for the design, deployment and operation of WSNs. To this end, we classify the WSN operations/functional building blocks into different classes and then map the existing approaches to them to show the lack of tunable co-design. Accordingly, we briefly summarize the existing approaches mentioning the building blocks they are concentrating and what the effects of neglecting other blocks are. In addition, we determine the way in which functional properties depend on and can be affected by various other features like deployment. Hereby, we provide an account, analysis of the design features, solutions, pros and cons that have been adopted by current frameworks and methods [Zahedi et al., 2008] [Gelenbe and Hey, 2008] [Zahedi and Bisdikian, 2007] [Bisdikian et al., 2009b].

Currently, the approaches to satisfy user evolvable requirements are addressed isolated by focusing on well-separate data processing operations/functional blocks comprising the raw data collection/sampling, in-network processing (compression, aggregation etc.), information transport and sink operation for decision making. These blocks are present from the source (raw data creation) to the sink (information delivery to the user). We argue to satisfy the user evolvable requirements by avoiding under and over provision of information when all or combination blocks are considered. Considering the different blocks as whole, the challenge lies in delivering the information just not by having the best techniques in the different functional blocks to deliver high quality, but sometimes requires tuning the techniques to deliver only required quality. We mainly (a) propose and argue for a tunable co-design, and (b) we propose to quantify the information, as the user evolvable requirements may be not satisfied while processing the data/information from the source to the sink.

2.1.1 Information Assessment

Usually, the quality of delivered/achieved information should be assessed according to the required/expected quality. For a quantitative assessment, metrics play a major role. In the following, we briefly discuss the user requirements as well as the information metrics.

Information complies with a set of attributes. These attributes are measured to give the level of detail of information. Hence, we view that user requirements are information based on some set of attributes. The user requirements can be further regarded as measured information based on a specific set of attributes. The user is not inevitably a human and can be application planner, end user, decision maker, consumer, intelligent system and so on The use of feedback channel is important here for user requirements dissemination.

Metrics are valuable at both design and deployment time as the user requirements are evolvable and the user would benefit from knowing the level of quality of received information entities for safer decision making. Measuring the information is either completed in-network or/and at the sink. A metric is a standard of measurement stated in quantitative term which captures the performance in relative to standard on the occurrence of an event. The quality of a system, such as its energy-efficiency, information attributes such as accuracy, reliability, timeliness etc. and the evaluation criterion of these qualities is judged by the term metric. The measure can be classified as happening and valuing. For example, fire detection in the forest, there is fire is the true state of event happening, there is fire with 95% accuracy is the valuing of the event. If the metric is well defined, it has to lead to actionable performance to satisfy the deployed system and also needs a capable system model to measure it. This doesn't mean to have a high rate of sampling or reliable protocol, or having non-effective metrics satisfy the user evolvable requirements. Hence, we can determine that a metric is acceptable with certain performance measure only if it has some opening limit, implying it is a boundary which is likely near/above to threshold value or real world value. In the next section, we define the necessary attributes which will be used to measure the information in our thesis.

2.1.2 Information Attributes

In order to assess the information, we first need to understand the information attributes. This section presents existing and proposes some new attributes of information. There exist many attributes for information, but we choose only those which are relevant and useful in WSN. There also exists an information model in defining information attributes [Wang et al., 1996], which benefits to define the existing attributes. To plan an application and use it in an operational perspective, one needs to give more importance on various attributes concerning the information. We define some of the existing attributes below based on information quality and functional building blocks.

Accuracy : is the degree of correctness which provides the level of detail in the deployed network. It is the value which is the closest imitation of the real world value.

Reliability : is the characteristic of information, in which information is free from change or no variation of information from all the blocks of the source to the sink. **Transport Reliability :** is the average success probability of the information to reach the sink from the source.

Timeliness : are an indicator for the time needed when the first data sample generated in the network till the information reaches the sink for decision making.

According to our knowledge from [He and Zafer, 2008] [Srinivasan, 2007] [Arnborg et al., 2000] [Tan and Gillies, 2009] [Ballou et al., 1998] there are still some missing attributes in WSNs for information, these attributes play a vital role and are useful in WSNs. The following attributes are similarly interwoven to the existing ones in the literature and also used in other fields like database management, machine learning and management studies. The following defined attributes are applicable to WSNs and also required, because of their sensible aspect in information processing.

Tunability : is the characteristic of information, where the information can be modified and undergo in-network processing based on user's evolvable requirements. Information is tunable, if the user requirements are changing to collect raw data or information at sink needs to be tailored. The user can take the advantage of the feedback channel in order to tune the information.

Affordability : is the characteristic of information to know the cost of measuring, collecting and transporting the data/information. It is the expensiveness of information. Affordability can be of raw data, and how cost effective it is to measure raw data. Affordability can be characterized to all the functional blocks based on the user's requirements.

Reusability : is the characteristic of information, where the information is reusable during its lifetime or as long as it is relevant (in the time domain) for future use in the context of WSN. In [Ballou et al., 1998], the timeliness attribute gives the similar meaning to reusability. However, the term timeliness in [Ballou et al., 1998] is mainly with information manufacturing systems. In WSN due to resource constraints and user requirements evolvability, timeliness and reusability give separate meaning.

2.2 System-level Approaches Addressing Quality of Information

In the literature, information quality is being addressed by frameworks [Zahedi et al., 2008] [Hunkeler and Scotton, 2008], methods [Kessel, 2006], models [Moody et al., 2003] and decision making techniques [Ehikioya, 1999]. In this section, we first classify the existing approaches lacking co-design. Next, we briefly describe them. Then, we qualitatively compare them. In Table 2.1, we define different classes of building blocks. Table 2.2 compares the approaches w.r.t. the covered building blocks, information attributes and metrics.

2.2.1 Classification

In this section, we classify the existing approaches that lack tunable co-design in WSN. Our classification criteria are the considered user, system and application models. We follow a step-by-step process to classify the existing approaches. First, we identify the source of information is from sensor nodes and also from the user/application. As we have already identified the functional building blocks as sampling, in-network processing, information transport and the sink operations, we take this step for granted.

Now, furthermore, we compare the existing approaches based on the building blocks, the information attributes and metrics. In Table 2.1, we define different classes of building blocks. Usually, the approaches focus on a few selected attributes. We have gathered most of the state-the-art related to lack of tunable co-design approaches in the following subsections. Usually existing approaches just focus on a single building block.

Building Blocks Classifications	**Types of Building Blocks**
Class 1: Single Building Blocks [C1S]	Sampling [D] In-network processing [IN] Information transport [T] Sink operations[S]
Class 2: Combination Building Blocks [C2C]	Sampling and In-network processing [D, IN] Sampling and Sink operations [D, S] In-network processing and Information transport [IN, T] In-network processing and Sink operations [IN, S]
Class 3: All Building Blocks [C3A]	Sampling, In-network processing, Information transport and Sink operations [D, IN, T, S]

Table 2.1: Functional building blocks classifications

2.2.2 Approaches

The approaches classified in this subsection are based on functional building blocks. Most of the existing approaches just concentrate on a selected building block like sampling or in-network processing or information transport or sink operations etc. However, they still lack to identify the effects of neglecting other blocks .i.e., for the co-design operation in WSNs.

[C1S] [D] : There are many approaches concentrating just for sampling/data collection block, we briefly summarize them here to highlight the issue on neglecting the other functional blocks. The HYBRID [Hakkarinen and Han, 2008] approach is more prominent as the variance of application requests or data change across sensor nodes increases, the model is based on push and pull method, the model dynamically switches between push and pull techniques based on system condition. However, the HYBRID model is setup for aspect of sampling, it neglects other blocks leading to negotiation with information attributes. The challenge of [Tolstikov et al., 2006] lies in considering phenomena state distribution while making application admission decision. The framework acts as an admission control scheme to decide if the WSN is able to provide the required service. Though the vital fact in this approach is sampling, the user cannot be sure of the acquired data till it reaches the sink.

The approaches [Zahedi et al., 2008] [Zahedi and Bisdikian, 2007] [Charbiwala et al., 2009] [Tolstikov et al., 2007] tend to overlook the effects of information transport. The current state of the art on a layered framework for decomposing the deployment evaluation is done in three steps of input processing, core analysis and result post-processing [Zahedi et al., 2008]. The given framework facilitates the decoupling of the three steps, the mix and match analysis and modeling approaches. It serves as a computational aid for a sensor system designer to evaluate the performance of users design based on deployment and information constraints provided by the application planner.

The results in [Charbiwala et al., 2009] demonstrate the benefit of using prior information about the event location on the probability of error. In this case the sampling phase should be accurate and also it should be relevant to the evolvable user requirements for decision making at the sink. The approach is very similar to the content centric networking, which endows the networking stack with knowledge of the intent of the communication transaction. On the same basis in [Charbiwala et al., 2009], new greedy rate control algorithm selects rates based on each nodes contribution to the information, but the drawback is that the proposed greedy rate control algorithm is unable to handle errors in wireless links. One important factor in WSNs is the process of sensor deployment and sensor selection. The later process of the sensor selection using Bayesian model is not appropriate for sensor selection as there is no notion of time in the Bayesian network and many sensor readings has to be taken to provide desired information. However, in [Tolstikov et al., 2007] this is achieved by using a dynamic Bayesian network model that provides the information quality to WSNs. The dynamic Bayesian network models optimize with one application and use very little resources in order to not to address the aspects of losses of data in the network. The setback of these kinds of models is assuming that every model is actually a complex, able of doing online data processing, which is not always true. However, what happens to data after sensor selection or to achieve information quality when it reaches the sink is not discussed.

In WSNs data models can help to combine readings from different sensor nodes to assess the information or to minimize energy consumption and thus maximize the lifetime of the WSNs while still respecting information attributes. In [Hunkeler and Scotton, 2008] the model is based on sampling. Here the framework allows several data models to run in parallel. The framework runs in offline mode, but for on-line the authors still propose future work and still lack to explain the information attribute factors within the framework. In [Hunkeler and Scotton, 2008], authors neglect the aspect of information transport and lack to provide required information. However, this leads to non-confident information reaching the sink.

[C1S] [IN] : In-network processing is the next block after sampling. The information fusion [Kessel, 2006] approaches are based on blocks of in-network processing. Identifying good candidates for information fusion is presented in [Kessel, 2006]. The analytic framework in [Kessel, 2006] is to study information fusion competition between the negative effect of disassociation and the positive effect of synthesis, to demonstrate and analyze their interplay quantity. The generic model used here is to demonstrate the varying degrees in fusion, namely increased quality versus decreased quality. In [Kessel, 2006] the approach concentrates on in-network processing and is not sure about what are the data collected and how the information is transported. Though one can assume a good underlying routing protocol, but the facts of violating attributes with accurately collected data and saving resources makes the approach still primitive.

[C1S] [T] : In the block of information transport the attributes reliability and timeliness are highly regarded. Though most of the routing protocols always assume the data/information coming from sensor nodes are accurate enough, but can't place a certain level of confidence in this data/information. Hence, in this regard, the attributes such as accuracy, precision are ignored. This issue is identified in the approach [Nichols, 2009] which is focusing on information transport. Disregarding the fusion process and not focusing on the sensor fusion aspects, assuming that those processes have been completed, the framework as in [Nichols, 2009] handles the quality assigned message in the network. However, in [Nichols, 2009] the mere aspects of in-network processing and sink operations and attributes related to this blocks are violated.

[C1S] [S] : Information plays different roles and has different values for decision makers at different levels. For characterizing the information quality spectrum the techniques like fuzzy values [Ehikioya, 1999] are used. Here, the approach is just concentrating on sink operations. In [Ehikioya, 1999], authors discuss the issues of uncertain data, imprecision. The main aim is in determining acceptance regions, similarity functions to determine the similarities between components and the confidence measures to rate attributes. Therefore, exploiting the tolerance for imprecision and uncertainty when precise information carries a cost or unavailable on the decision making process.

2.2.3 Classification Based on Information Attributes and Metrics

This sub-section is the classification of approaches based on attributes and metrics, which are used to characterize and quantify information. We always argue that to have achievable information pertaining to user evolvable requirements, user needs to respect the characteristics of information. Moreover, we present some of the approaches concentrating only on some attributes and measuring them.

The principle based framework [Zahedi et al., 2008] is a strategy of principles and steps to achieve ideology of deployment planning, decision making, and quality enhancement. The current state of art on a layered framework for decomposing the deployment evaluation is done in three steps of input processing, core analysis and result post-processing. The framework uses the probability of error to measure detection probability and false alarm rate. The main aspect of detection in WSNs is any event, in [He and Zafer, 2008] detection performance is measured by average sampling rate with characteristics such as accuracy and timeliness. The information aware route control in [Charbiwala et al., 2009] uses the probability of error as metric to measure accuracy. It explicitly optimizes application relevant information metrics during network resource allocation decision. The approach presented in [Gelenbe and Hey, 2008] focuses on accuracy and measure it with Peak Signal Noise Ratio (PSNR). However, though the information here is measured, other attributes such as timeliness for the timely arrival of information for decision making have been not discussed.

Exploiting the tolerance for characterization of information quality using fuzzy logic [Ehikioya, 1999] some attributes such as accuracy, completeness, relevance, timeliness and usability are explored. However, though the work considers some attributes relevant for the information, never quantifies it. With the characterization of information, Information Risk Management (IRM) [Chang et al., 2005] is also proposed in the literature to minimize the risks such as information misunderstanding and insufficiency of metrics which may affect learning quality. Dimension extension (DIME) is a framework to accommodate local and prior knowledge into learning course by measuring the accuracy through the dot product as metric.

To achieve better results, data processing is used in current trends in information quality. Usually, in resource constrained framework a real good data processing is a key precondition for analysis decision and data integration. One of the frameworks [Yan et al., 2008] addressing this is based on rule base, scheduling and log management. The attributes such as consistency, accuracy, extensibility and interactivity are used for data cleaning and measured by metrics such as recall and false-positive rate. The overall design fully shows the features of extensibility and interactivity, meaning the framework allows users to add rules, and at the same time allows users to form strategies in the needs of different data cleaning. The concept of operational context to ease the dynamic binding of sensor resources to applications represents information needs of an application and the capabilities of the sensor resources by the 5WH (why, where, when, what, who, how) principle [Bisdikian et al., 2009a]. In the interpretation of the 5WH primitives provided, spatial and temporal relevance are used as a metric to measure data completeness.

The evolution of the context may be used to adjust dynamically the weights of the sensor nodes that ease selecting the right set of sensor nodes given the dynamic context change as the one in [Anwar Hossain et al., 2008]. Some attributes such as certainty, accuracy/confidence and timeliness are used for context aware information computation. Still here the information is not measured. Relative to this the selection of sensors can be made by using metrics such as information gain and using other attributes missing in [Anwar Hossain et al., 2008]. By targeting all the building blocks and attributes related to each block, we now brief a strategy [Kannan et al., 2003] that develops a game-theoretic metric called path weakness to measure the qualitative performance of different routing mechanisms. The approach uses qualitative performance as an information characteristic and uses a sensor - centric concept.

Considering the information transport, prioritizing traffic has been studied for a long time. Disregarding the fusion process and not focusing on the sensor fusion aspects, if those processes have been completed, the framework as in [Nichols, 2009] handles the quality assigned message in the network. Based on this the key metric transient information level is defined, which is the product of information and projected physical distance of that information from destination node. This approach is very relevant to information transport block as attribute related to information transport such as timeliness of information are used. The information level is also measured, but the approach neglects the effects of other building blocks and some attributes.

2.3 Sampling and Information Transport Schemes

Providing the optimized co-design of sampling and transport is not straightforward due to the dynamic requirements and operational conditions. Traditional network design investigates the sampling and communication co-design from the simplistic view that the application data rate usually exceeds the capacity of the network and therefore the rate should be adapted accordingly. The additive increase/multiplicative decrease of the Transmission Control Protocol (TCP) is a renowned example of these efforts. Further efforts focus on varied user requirements and provide a QoS based design of network transport that allocates variable data rates to varied users. Also in WSN, QoS provisioning [Chen and Varshney, 2004] [Martínez et al., 2007] [Abbas and Kure, 2010] [Karenos et al., 2005] [Tan et al., 2006] focus on network capacity and consider a simplistic model of sampling. In networked process control community, a co-design of sampling and transport has been addressed. This co-design has been driven by the limited capacity of the network. In WSN, in addition to the network capacity constraint, the co-design should take into consideration the energy constraint, which is of higher priority.

The state-of-the-art in WSN focus either on the sampling accuracy (e.g., [Li et al., 2006] [Szczytowski et al., 2010] [Vuran and Akyildiz, 2006] [Sankarasubramaniam et al., 2003] [Kim and Yang, 2010]) or transport reliability (e.g., [Zhang et al., 2007] [Felemban et al., 2006] [Huang and Fang, 2008] [Shaikh et al., 2010] [Barbancho et al., 2007] [Morita et al., 2008] [Ortiz et al., 2011]) or transport timeliness (e.g., [Jiang et al., 2009] [Sahoo and Baronia, 2007] [Karenos and Kalogeraki, 2006] [Lu et al., 2002] [Hey and Gelenbe, 2011] [Kwon et al., 2010]). However, there is no prior work addressing a co-design of sampling and transport in composition along online adaptation to satisfy user evolvable requirements while maximizing energy efficiency and minimize loss of information.

In [Li et al., 2006], the authors address, the node selection for optimizing accuracy in WSN. However, the information transport is assumed to be reliable. In [Szczytowski et al., 2010], the authors propose an adaptive sampling approach to achieve user required accuracy and to avoid over-/under-sampling. While this poses an efficient and adaptive approach to model, sampling accuracy, reliable transport is not considered in this work and reliability is implicitly assumed to be perfect. In [Vuran and Akyildiz, 2006], the authors address the spatial correlation based on MAC protocol called Correlation based Collaborative Medium Access Control (CC-MAC). However, though the authors address the optimized solution for accuracy, the transport reliability and timeliness are neglected. In [Kim and Yang, 2010], authors discuss about sampling and how efficiently the data can be stored. Storing the deviating data than the normal data is highlighted. Hence, in this regard in [Kim and Yang, 2010] the bit-vector based information storage method has been proposed and authors discuss about the importance of accuracy. However, the authors neglect the information transport and also the attributes reliability and timeliness.

In [Sankarasubramaniam et al., 2003], the sampling of convergecast applications is addressed. However, adapting the sampling rate is independent of the application requirements. In [Zhang et al., 2007], the authors focus on bursty convergecast where the key challenges are reliable and real-time error control and the resulting contention control. However, [Zhang et al., 2007] does not offer mechanisms to adapt to changing user requirements and neglect the aspect of sampling accuracy. In [Felemban et al., 2006], probabilistic techniques are applied for service differentiation. However, the solution aims at providing strict conditions for messages. In [Huang and Fang, 2008], the authors propose multi-path forwarding to ensure end-to-end delays. Also [Huang and Fang, 2008] is not adaptable to fluctuating network conditions to make routing decisions. However, optimizing accuracy and reliability for maximizing efficiency are missing in [Felemban et al., 2006] [Huang and Fang, 2008]. In [Wang et al., 2011], the authors propose metrics to measure the quality of a path. However, they do not address tunability. GIT [Shaikh et al., 2010], aims at satisfying the end-to-end reliability by dividing the reliability per hop. The proposed transport protocol is tunable regarding the achievable reliability. Providing a solid basis for reliability, [Shaikh et al., 2010] yet has to be extended to consider sampling accuracy. In [Barbancho et al., 2007], introduce a new algorithm named SIR which is based on the concept of using artificial intelligence in WSNs. However, while considering the reliability, latency, and maximizing energy efficiency, the authors neglect the accuracy to be considered as one of the key aspect.

In [Morita et al., 2008], authors discuss about data transmission in sensor and actuator network and introduce an efficient data transmission protocol called RT. In addition, authors also discuss how to reduce the total amount of energy consumptions of sensor nodes. The loss ratio of sensor values can be reduced without retransmission of lost messages in the RT protocol. However, in [Morita et al., 2008], authors neglect the accuracy and overlook the timeliness attributes. In [Ortiz et al., 2011], authors discuss about a new role-based routing protocol named NORIA that makes use of fuzzy logic to make decisions. The reliability of the packets reaching the sink and energy efficiency is highlighted in the work. However, the authors in [Ortiz et al., 2011] neglect to consider the accuracy and overlook the co-design of sampling and information transport.

In CFLOOD [Jiang et al., 2009], the authors address the problem of flooding and improve it with a new concept of controlled flooding. Due to the controlled flooding timely delivery of the packets is possible. However, the authors miss the important aspect of reliability and target maximum reliability as they focus on the detection of critical events. By means of a Time Division Multiple Access (TDMA) scheme at the expense of limiting the length of routing paths delay guarantees are provided in [Sahoo and Baronia, 2007]. Traffic regulation mechanisms are explored as a means to provide end to end guarantees with a combination of queuing models and message schedulers in [Karenos and Kalogeraki, 2006]. In [Lu et al., 2002], the packets are scheduled with high/low priority in the velocity-monotonic order without any guarantee in the end to end sense. However, the above protocols overlook to provide the user defined timeliness and consequently to tune reliability and timeliness in composition. In [Hey and Gelenbe, 2011], the packet prioritization is discussed with detecting the important events and a new protocol named RRR has been introduced. The authors consider the mean latency achieved for different priority packets and events. However, reliability and accuracy attributes are neglected and overlook considering both sampling and information transport in combination. In [Kwon et al., 2010], data gathering in WSNs is highlighted and how constructing a routing tree can be cost effective is discussed. Authors in [Kwon et al., 2010] propose a distributed and localized framework for tree construction called Local Parent Designation (LPD) and then extend to so called LPD-Local Fix (LPD-LF) to reduce further the construction cost with less local information exchange. The latency and energy efficiency are considered, however, the authors neglect sampling accuracy and transport reliability.

Considering accuracy, reliability and timeliness mutual dependencies are not as straightforward. Hence, for a co-design it is not sufficient to just superpose a tunable sampling scheme with another tunable transport scheme.

The challenge is still to provide efficient composite tunability of both data operations. In [Han and Venkatasubramanian, 2007], the authors propose accuracy-aware context data collection and queries for heterogeneous mobile ubiquitous computing environments. However, the approach overlooks the transport reliability. The authors in [Park et al., 2011] present a transport protocol with tunable timeliness and reliability. However, the work is optimized for a specific domain, i.e., real-time control and ignores the sampling quality. In [Sachidananda et al., 2013], we consider the tuning of transport reliability and timeliness in composition, but without addressing the sampling accuracy. In [Tan et al., 2010], the authors present a co-design of data aggregation and data transport in WSN, ignoring the sampling operation. Summarizing, to the best of our knowledge, there is no prior work on sampling and transport co-design for providing application required quality with optimized tradeoffs spanning accuracy, reliability, timeliness and energy efficiency in WSNs. In this work, we build first steps to fill this research gap.

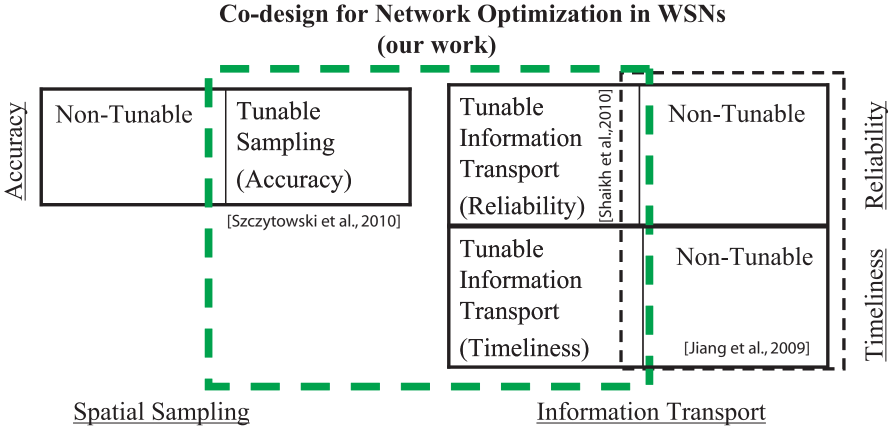


Figure 2.1: The design space for tunable co-design and state of art

2.4 Chapter Summary

Users are mainly interested in the information from the WSN. This highlights the importance of understanding the quality of sampling and information transport. Consequently, we provided a comprehensive survey/review of most of the existing work related to lack of tunable co-design. In addition, we show the approaches that neglect one or the other functional blocks and just work in isolation. Our study was performed on the functional blocks, users, information attributes and its application perspective. Information across all WSN functional operations from the raw data generation to the information extraction and delivery to the user was the base of our study. Furthermore, we have also classified building blocks into different classes.

Based on these classes we have mapped the existing information frameworks, methods and models, approaches and highlighted the pros and cons. Moreover, our classification technique is more general to be adapted to most of the applications. On the other hand, we also mapped the existing work related to information attributes and metrics.

Nevertheless, this chapter presented an overview on information based on building blocks, attributes, and metrics. However, we have identified the tunable co-design related problems, i.e., to focus on all or combination building blocks and satisfy user requirements. We have also mentioned that violating of least required attributes may affect information. We built our co-design view on these factors and proposed that information should satisfy user evolvable requirements by saving resources.

Approaches	**Building Blocks**	**Attributes**	**Metrics**
Analysis [Zahedi et al., 2008]	[C1S], [D]	Detection probability and false positive rate	Probability of error
Information in DTN [Nichols, 2009]	[C1S], [T]	Timeliness, integrity and consistency	Transient information level
Context Aware Computation [Anwar Hossain et al., 2008]	[C1S], [D]	Certainty, accuracy, integrity	x
Characterization of Information [Ehikioya, 1999]	[C1S], [S]	Accuracy, completeness, relevance, usability	x
Rate Control [Charbiwala et al., 2009]	[C1S], [D]	Accuracy	Probability of error
Information Fusion [Kessel, 2006]	[C1S], [IN]	Uncertainty	x
Data Model Framework [Hunkeler and Scotton, 2008]	[C1S], [D]	Accuracy, reliability	x
Resource Management [Tolstikov et al., 2006]	[C1S], [D]	Completeness, uncertainty, accuracy	x
Data Driven Sensor Reporting [Hakkarinen and Han, 2008]	[C1S], [D]	x	x
Information Management [Tolstikov et al., 2007]	[C1S], [D]	Certainty, completeness, accuracy	Entropy
Analysis for WSN Deployment [Zahedi and Bisdikian, 2007]	[C1S], [D]	Accuracy	Probability of detection and false positive rate
Empirical Approach [Gelenbe and Hey, 2008]	[C1S], [D]	Accuracy	Peak Signal to Noise Ratio
Information Risk Minimization [Chang et al., 2005]	[C1S], [I]	Accuracy	Dot product
Sensor Sampling [Bisdikian, 2007]	[C1S], [D]	Accuracy and confidence	x
Detection Performance [He and Zafer, 2008]	[C1S], [D]	Accuracy and robustness	Average sampling rate
Data Cleaning [Yan et al., 2008]	[C1S], [IN]	Consistency, accuracy, extensibility, interactivity	Recall and False-Positive Error
Quality of Routing [Kannan et al., 2003]	[C1S], [T]	Qualitative performance	Path weakness
Information Awareness [Arnborg et al., 2000]	[C1S], [IN]	Precision, quality and usability	x
Dynamic Target Tracking [Tan and Gillies, 2009]	[C1S], [S]	Accuracy	Entropy, information gain, residual likelihood
Letter Soup for Information [Bisdikian et al., 2009a]	[C1S], [D]	Accuracy, timeliness	Spatial and temporal relevancy

Table 2.2: Classification of approaches based on functional building blocks, attributes and metrics

Chapter 3

Preliminaries

This chapter first presents the system model. Subsequently, perturbation model is presented which covers the various failures encountered in the stated system. Then, the hotspot model is developed keeping the generalization intact. Furthermore, we propose the sampling and information transport model by considering the co-design as the key aspect. Next, this chapter presents with the key performance indicators, which in turn accords the terminologies for the optimal co-design of sampling and information transport in WSNs. Later, the fundamentals driving our co-design are presented. Finally, the design requirements are discussed.

3.1 System Model: Generalized WSNs

We consider the conventional model of a WSN having N sensor nodes [1, .. $N-1$] and a single sink (S). Typically, each node is equipped with one or more sensing device, short range transceivers with limited processing, memory and energy capabilities. We consider the sink to be adequate in power (ideally up to entire expected life of the network), memory and higher processing capabilities as compared to the sensor nodes. We assume that all nodes are static in nature (including the sink) and are placed in a finite area. However, the topology of WSN is dynamic due to perturbations in the network. Each sensor node maintains a limited buffer of size Q. The sensor nodes communicate with each other via bi-directional multi-hop wireless links employing a Carrier sense multiple access (CSMA)-based Medium Access Control (MAC) protocol. For any two nodes X and Y we define their link quality $LQ = p_{(X,Y)} . p_{(Y,X)}$, where $p_{(X,Y)}$ and $p_{(Y,X)}$ indicate the probability that a message sent by Node X is received correctly by Node Y and vice versa. Let the sequence of hops (X, H_1), $(H_1, H_2) \cdots (H_f\ , S)$ create a path $Path_i$ from Node X to the sink. All sensor nodes know their hop distance $h_{(X)}$ from the sink and their one hop neighbors. Based on hop distances, the neighbors of a node can be classified as upstream neighbors, downstream neighbors and equal neighbors.

The information is generated in the network from sensor nodes (e.g., the sensor nodes detecting the event) and forwarded towards the sink. Without loss of generality, we assume the vicinity of the event detecting node will be congested due to activities such as event detection and aggregation. We consider multi-purpose WSNs, i.e., different applications are running simultaneously in the network. In addition, we allow that applications may change their requirements during the operation of the WSN. The requirements are disseminated to all nodes, e.g., through an efficient flooding protocol such as [Khelil and Suri, 2007]. We assume that the most strict user requirements do not exceed the maximal capacity of the WSN [Gupta and Kumar, 2000].

Furthermore, the assumption of a single sink is not the requirement in our work. The WSN may also consist of few designated sinks. In this work we assumed, for the sake of simplicity, the existence of a single sink. In the case of multiple sinks, the sensor nodes can be associated with any of the sink based on different criteria, e.g., shortest distance towards the sink.

3.2 Perturbation Model

WSN is obviously subject to a wide range of computing and communication level faults. Low-cost hardware, limited resources and severe environmental conditions lead to frequent perturbations in WSNs [Arora et al., 2004]. Our perturbation model is based on the ability of sampling and information transport co-design to tolerate the effects of these perturbations [Walter and Suri, 2003].

We mainly emphasize the temporal evolvability of the perturbations in WSNs, which hinders in maintaining the required level of application requirements (reliability, timeliness and accuracy). We classify the tolerable perturbations as communication and node failures.

Communication Failures : Communication failures constitute the most frequent failures in the WSN. Failures relevant to the information transport include message loss which directly impacts the reliability and timeliness of the WSN. Collisions and contention constitute the major causes of message loss in WSNs. Collisions occur when two or more sensor nodes transmit messages simultaneously assuming the channel is clear and available for transmission. Once, the collision happen the message is lost. On the other hand, contention refers to the situation when the offered load on the link reaches a value close to the capacity of the link. In such a situation, the sensor nodes sense the channel to transmit the messages and find it busy. The sensor node keeps waiting and trying until the channel is clear for transmission. During contention either the sensor nodes after certain attempts discard the messages or they may receive more messages causing buffer overflows leading to message loss.

Node Failures : At a node level message loss is caused by congestion and unavailability of sensor nodes. Usually, the congestion is due to increasing network load. When the buffer capacity at a sensor node is exceeded, congestion occurs resulting in message losses. On the other hand unavailability of sensor nodes can be due to many reasons (1) Sensor nodes usually operates on batteries, which limits the operational lifetime of the sensor nodes. Typically, the drained batteries cannot be recharged or replaced, thus they can not be part of the network. (2) Also, the sensor nodes are often deployed in harsh environments and may suffer physical damage. (3) Furthermore, energy saving schemes that are based on duty cycles [Strasser et al., 2007] may be utilized, resulting again in temporary or prolonged sensor unavailability. For the sensor node unavailability, this thesis relies on the underlying routing protocols to provide the alternate good neighbor to route the information.

Intolerable Perturbations : Intolerable perturbations are those whose effects cannot be handled by the sampling and information transport co-design. WSNs may be deployed in harsh environments such as for fire detection, tracking of people in catastrophic areas. These environments can permanently destroy the sensor nodes on a large scale or the entire WSN, which obviously can not be handled. Other intolerable faults include crash failure of the sink and network partitioning. The sink plays an important role and acts as a bridge between the user and the WSN. Therefore, if the sink crashes, the network will not be able to communicate with the user resulting in an intolerable perturbation. Network partitioning is considered as an intolerable perturbation too, since source nodes and the sink may belong to different network partitions. These intolerable faults can be transformed into tolerable ones, if the maintenance and reconfiguration of the WSN is possible.

3.3 Application Model

Hotspot : We consider a physical phenomenon of interest that spans a specific small sub area of the WSN field. In general, the application is interested in information about this spatial phenomenon, e.g., the perimeter of its area. We consider that the sink is interested in estimating the hotspot S with certain accuracy value. We consider that the sensor node N is at a certain distance with variance σ_s due to the hotspot and also with certain noise σ_N i.i.d Gaussian random variables of zero mean and variance. We also consider that any signals Gi and Gj between any sensor nodes N_i and N_j can be with certain correlation coefficient $\rho(i,j)$ and also certain correlation $\rho(S,i)$ between hotspot epicenter and sensor node N. Here, the correlation is the statistical relationship between the signals. We acclaim that the information reaching the sink is with certain contortion (reflecting how close is the achieved spatial phenomenon in the real world) which should satisfy the user required contortion, i.e. the accuracy threshold in our case. The contortion function is the composition of the signal magnitude along with the variance between the sensor nodes. The various differences between the variance and the signal magnitude results in higher or lower contortion.

Sampling and Information Transport : A minimum set of spatial samples is required to reconstruct the information on the sink. To this end, sensor nodes sample this spatial phenomenon and transmit them towards the sink. We assume that the sampling sensor nodes have different sampling qualities. Apart from detecting the hotspot with certain accuracy, we also consider that the information reaching the sink is with certain reliability. We assume that the link quality differs and even the number of hops (h) to the sink from the source are different for each path. On the other hand, we also acclaim to consider user defined end-to-end timeliness to be maintained from the source nodes to the sink.

We assume an underlying routing protocol, which provides a path to the sink for all sensor nodes. Without loss of generality, we assume that all paths end at the sink. We consider an underlying routing protocol, which provides a sensor node with a next hop along the $Path_i$ towards the sink. The sensor nodes generate message(s) corresponding to the information to form convergent traffic in the upstream direction, i.e., from sensor nodes to the sink. A sensor node can be either a source node and/or relay node. The single path routing protocols can be utilized as a multipath with the knowledge of neighbor nodes [Kim et al., 2004].

In order to acquire hop count $h(X)$ the sensor nodes may utilize the underlying routing protocol as well. When the network initializes and the sink creates spanning tree, it includes a hop counter to the beacon messages, which allows sensor nodes to acquire knowledge of $h(X)$, i.e., how far they are from the sink. The sink periodically sends beacon messages to maintain the routing tree. Thus, the sensor nodes have an almost consistent view of $h(X)$. It should be noted that the hop distance is dependant upon the underlying routing protocol and the beacon messages follow the path from the sink towards the sensor nodes. Therefore, $h(X)$ always reflect the optimal distance between the sensor nodes and the sink.

In order to ensure the desired requirements the co-design must overcome the above mentioned perturbations using both temporal and spatial redundancy techniques for information transport. Temporal redundancy addresses certain actions to be performed over time, e.g., retransmissions to overcome communication failures. Using spatial redundancy we assume that redundant source nodes or paths are available for information transport towards the sink.

3.3.1 Key Performance Indicators and Terminology

Given the typical WSN applications, we now investigate application requirements for sampling and information transport. We identify the following key requirements of applications on sampling and information transport and these performance indicators are:

Reliability: The conventional requirement of applications is to receive all information generated in WSN via the sink. The reliability of information transport quantifies the ability of the network to deliver the information using appropriate mechanisms. Typically, application requirements are not absolute, i.e., some information loss can be tolerated and are statistical in nature.

Timeliness: WSN applications further demand the availability of information in-time. Moreover, some applications such as tracking require strict timeliness behavior. This can be understood as a requirement on the information transport mechanism to deliver the information in time.

Accuracy: Considering the application and its trivial for a WSN to represent a phenomenon, accuracy quantifies to represent this phenomenon based on the deployed sensor nodes and number of samples. Furthermore, considering the co-design the accuracy requirement is directly related to the reliability requirement of information transport. Moreover, the spatial accuracy is controlled by the deployed sensor nodes and over or under sampling can deviate the reliability and timeliness requirements of the information transport.

We consider the energy as the information cost:

Energy Efficiency: Application requirement for energy efficiency is directly related to the lifetime of the WSN. As the sensor nodes posses limited energy sources, the information creation and delivery solution must be energy efficient. Since, the transmissions are the major factor in energy depletion in WSN, the information transport mechanisms must utilize a minimal number of transmissions to deliver the information to the sink. On the other hand, activating the right number of sensor nodes for the spatial sampling would maximize the efficiency.

We assume in this thesis, that the raw sensor data typically has a temporal and spatial correlation. The information transport strategy should be aware of redundant samples and accordingly, take measures for assuring the reliability of information transport. We assume that whenever the information is generated it is relevant for the application and should be transported to the sink. We consider the spatial correlation of samples and appropriately manage it in order to efficiently transport the information. We also assume that the application's timeliness requirement can be strict or flexible depending on the WSN application.

Terminology for Tunable Information Transport

Consider Sensor Node S that is h hops from the sink and has an information to send to the sink with a user-specified reliability and timeliness. H_1, H_2,......H_h denote the h hops from S to the sink where H_i is the i^{th} hop from the sink.

1. **Transport Reliability** (R) is the average success probability of the information to reach the sink.
2. **User Desired Reliability** (R_d) is the average reliability as required by the user.
3. **Link Reliability** (R_{L_i}) is the achieved success probability of one single message transmission on Hop H_i.
4. **Hop Reliability** (R_{H_i}) is the achieved success probability after r transmissions of the same message on H_i.
5. **Desired Hop Reliability** ($R_{d_{hop}}$) is the hop reliability to be maintained in order to achieve the overall user required reliability R_d.
6. **Transport Latency** (L) is the time needed for the information to reach the sink.
7. **User Tolerated Latency** (L_{tol}) is the maximum delay allowed for the information to reach the sink.
8. **Hop Latency** (L_{H_i}) is the delay experienced on H_i.
9. **Tolerated Link Latency** ($L_{tol_{H_i}}$) is the maximum delay allowed on H_i.

Terminology for Sampling and Information Transport Co-design

In this section, we provide the terminology for sampling and information transport co-design as the preliminary requirement for next sections.

1. **Transport Reliability** (R_{path}): We define the end-to-end transport reliability as the success rate of one sample from one specific sampling node to reach the sink. Moreover, considering R_{link} on the lowest level, varying number of retransmissions affects R_{path} directly.

2. **Sensing Accuracy**: The sensing accuracy is the accuracy of sampling as perceived by the application/user/sink. Accordingly, sensing accuracy is the ratio of the number of samples received at the sink S_{rx} to the minimum required number of samples S_{min}. The sensing accuracy depends on the optimized combination of transport reliability R_{path} and activating the right number of sensor nodes S_{tx} for sampling accuracy.

Terminology	**Definitions**
R_{link}	The achieved success probability of one message transmission on one link
R_{hop}	The achieved success probability of message transmissions on one Hop after specific number of retransmissions
R_{path}	Reliability of one path
R_{inf}	The achieved success probability of the information (S_{min} samples) to reach the sink
Fi_{acc}	The sensing accuracy fidelity, i.e., is the expectation that the perceived sensing accuracy is equal to the desired sensing accuracy
S_{min}	The application desired number of samples from the phenomenon area
S_{tx}	The number of samples transmitted from the phenomenon area
S_{rx}	The number of samples received at the sink
h	Number of hops from sampling nodes to the sink
$\#ret_h$	Total number of retransmissions on one hop
$\#ret_{total}$	Total number of retransmissions induced by the transport of S_{tx} samples

Table 3.1: Important notations and their meanings

Terminology for Representing the Physical Phenomenon

In the following we define important terms for representing the physical phenomenon.

1. **Contortion Accuracy** (CA): The contortion accuracy is the spatial accuracy which depends on the phenomenon and the required information. Usually, the application expects a certain spatial sample distribution. The contortion accuracy reflects how close the achieved distribution to the required one is. We choose the notion of sampling accuracy as the contortion accuracy.

2. **Contortion Experienced** (CE): The contortion experienced is the perceived contortion accuracy of sampling by the application/user/sink.

3. **Contortion Required** (CR): The contortion required is the desired contortion accuracy of sampling by the application/user/sink.

4. **Fidelity** (Fi_{acc}): The fidelity is the lower bound for the expectation that the contortion experienced (CE) is equal or less than the contortion required (CR).

5. **Transport Reliability** (R): The end-to-end transport reliability is defined as the success rate of the samples to reach the sink.

6. **Transport Timeliness (L)**: is the time needed for the samples to reach the sink.

3.4 Driving Force for a Co-design: Fundamentals

Considering our design objectives and requirements, the co-design is the most complex and core driving problem of our thesis. Hence, in this section we provide the fundamentals of the tunable co-design as the driving force of our core problem statement. Here we refer to the functional blocks of sampling and information transport.

The common theme in all the wireless networks is the use of the wireless channel for communication. The wireless channel has several unique characteristics, unlike the wire-line networks. The most important one is the transmitted signal that proliferates through the wireless medium is affected by attenuation, dynamic behavior of the wireless links and degrades more rapidly with distance as compared to the wire-line channels. One typical assumption is that each isolated functional block (sampling and information transport) can be optimized independently and performance gains within each functional block will be sufficient for the wireless networks as in the equivalent wire-line networks. This might often lead to a sub-optimal solution and inefficient use of network resources.

The unique problems and opportunistic exploitation of wireless links, and the new modalities offered by WSNs paradigm make a strong case for co-design and optimization. An example of a unique problem is that the TCP in layered architectures implicitly assumes that a packet loss is caused due to collision which is not true for WSNs where a packet loss may occur because of other phenomenon like fading or varying link quality. Potentially harsh environmental conditions, unattended operation, and operating in free frequency band make WSNs even more prone to errors by interference or fading.

An important aspect of WSNs is their dynamic behaviour. The conventional WSNs approach of addressing isolated functional blocks is inflexible as they communicate in a strict manner. In such a case the functional blocks are designed to operate under the worst conditions as opposed to adapting to changing conditions. This leads to inefficient use of resources such as energy. Adaptation represents the ability of the functional blocks to observe and respond to the dynamic conditions of WSNs.

Co-design may be best understood by explaining the isolated functional blocks. The latter limits the communication between the functional blocks and considers other blocks to be perfect. Co-design violates these principles and uses information from different functional blocks to improve the network performance and/or lifetime. A co-design due to the inherent resource constraints of WSNs and application/user specific requirements is conjectured. However, it is a further requirement to have the tuning capability of the functional blocks and its attributes for an optimal network performance. Tunable co-design exploits the best features of each functional block, with the goal of achieving flexible and efficient design solutions. Co-design is promising for exploring to optimize the performance in WSNs. Co-design aims to achieve gains in overall system performance, such as increase in network capacity, energy efficiencies and support to a variety of applications. The primary objective to investigate the tunable co-design is to increase the fidelity of the functional blocks while minimizing the cost such as energy. The tunable co-design optimization is a technique to improve the performance in WSNs.

Before moving on, it is important to note that co-design is not only motivated by the characteristics of the WSNs. Other factors such as multiple applications, multiple users, limited energy, memory, bandwidth, and the need to satisfy user requirements on accuracy, reliability and timeliness also play important roles. In fact, it is the combination of all the variations and constraints that gives rise to tunable co-design. In summary, the central idea of the tunable co-design is to optimize the network performance by exploiting the characteristics of sampling and information transport. All the above characteristics, coupled with the need to conserve energy for sensor nodes, make it important to allow more inter-dependencies, more information sharing, and more flexibility in the design of energy-constrained WSNs. This motivates the concept of tunable co-design.

3.5 Design Requirements

In the following we discuss the design requirements for sampling and information transport co-design in WSNs. First, we outline the design requirements for tunable information transport. Next, we highlight the requirements for the sampling and transport co-design. Finally, we outline the requirements for representing the physical phenomenon.

3.5.1 Design Requirements for Tunable Information Transport

The transport reliability is the ability of the transport protocol to meet the desired reliability, i.e., $R = R_d$. The transport timeliness is the ability of the transport protocol to meet the tolerated end to end deadline, i.e., $L = L_{tol}$. The transport tunability is the ability of the transport protocol to just meet the required reliability without violating the tolerated end to end requirements, i.e., to ensure that $R = R_d \textit{ AND } L = L_{tol}$. Being close to the requirements allows maximizing efficiency, which represents the key reasoning behind our approach.

For information quality aware transport we derive the basic design requirements based on the application requirements, WSN characteristics and the presented design objectives.

- The information transport has to deal with generic characteristics of WSNs and diversified applications. Therefore, it is required that the information transport should be realized keeping in view the limited sensor node capabilities and should be as general as possible.

- Maximizing efficiency is a major requirement of WSNs due to limited energy resources, therefore, the information transport have to reduce message overhead as much as possible. The message overhead is also a good indicator of energy consumption, bandwidth utilization and the storage overhead.

- Node level localized data transport without global topology information, and overcoming un-reliable and unnecessary re-transmission, saving energy.

3.5.2 Design Requirements for Tunable Sampling and Information Transport Co-design

A minimum number of spatial samples S_{min} is required to reconstruct the information on the sink. To this end, S_{tx} sensor nodes sample this spatial phenomenon and transmit the samples towards the sink. We assume that the S_{tx} sampling sensor nodes have the same number of hops h to the sink. The hops are considered as the average hop count from all the active sources to the sink.

- Variations of one or two hops do not affect the model and the end result, as we are interested in the small sub area of the phenomenon. This is the case if the phenomenon area is small compared to the WSN field which is often the case for event-driven applications.

- The application requirements should be distributed from the sink to the sensor nodes.

- The number of sampling sensor nodes S_{tx} can be controlled, e.g., through an existing duty cycling algorithm that interacts with the sampling scheme, e.g., [Szczytowski et al., 2010] to decide on which nodes to keep active.

3.5.3 Design Requirements for Representing the Physical Phenomenon

We consider a physical phenomenon of interest that crosses a specific small sub area of the WSN field. In general, the application is intrigued with information about this spatial phenomenon, e.g., the perimeter of its area. We consider that the sink is interested in estimating the hotspot with certain accuracy value. We consider that the sensor node in the phenomenon area is at a certain distance with variance due to the hotspot and additionally with certain noise. We also consider that any signals between any sensor nodes can be with certain correlation coefficient and additionally certain correlation between hotspot epicenter and sensor node. Here, the correlation is the statistical relationship between the signals.

The sampling sensor nodes have different sampling qualities. The link quality differs amongst the sensor nodes and even the number of hops (h) to the sink from the source are different for each path. The most strict user requirements do not exceed the maximal capacity of the WSN [Gupta and Kumar, 2000].

On the other hand, we also acclaim to consider user defined end-to-end timeliness to be maintained from the source nodes to the sink. We assume that the most strict user requirements do not exceed the maximal capacity of the WSN [Gupta and Kumar, 2000].

- The information reaching the sink is with certain contortion which should gratify the user required contortion, i.e. the accuracy threshold in our case.
- A minimum set of spatial samples is required to reconstruct the information on the sink. To this end, sensor nodes sample this spatial phenomenon and transmit their samples towards the sink.
- Apart from detecting the hotspot with certain accuracy, the information reaching the sink should satisfy certain reliability.
- The user defined end-to-end timeliness should be maintained from the source nodes to the sink.

Chapter 4

Problem Statement

As an important base in the context of the research represented in this thesis, this chapter begins by discussing the fundamental requirements for a co-design. The problem statement is given out with a precise analysis and discussion. This chapter forms the basic step stone and the context for the research questions posed and puts the contributions presented into perspective.

4.1 Problem Statement

Considering the spatial phenomena of interest, the user/application view requires a certain *sensing task* (e.g., perimeter of the phenomenon area [Ghosh and Das, 2008] or the spatial distribution of the phenomena [Szczytowski et al., 2010]). Moreover, the perceived contortion accuracy should satisfy the application requirements (e.g., accurate form and location of the event perimeter). In addition, future WSN deployments should allow for varied concurrent applications. Usually, these applications need varied information and have evolvable requirements.

Moreover, achieving the best possible sampling accuracy and transport reliability and timeliness is concerned to a large resource overhead, particularly, because sensor nodes rely on batteries. A higher quality level is frequently associated with higher deployment costs and higher resource overhead. A higher accuracy of spatial sampling of a spatial physical phenomenon of interest is usually achieved through a higher number of active sampling sensor nodes in the area of the phenomenon resulting in a higher energy/bandwidth overhead. On the other hand, higher transport reliability usually is achieved through a higher number of retransmissions, and lower transport timeliness may require multipath transmissions to ensure using the fastest possible path. Hence, besides attaining the required quality levels, it is indispensable to maximize energy/bandwidth efficiency and minimize loss of information. Considering the design view, the sampling accuracy can be tuned by injecting some redundancy (e.g., activating more sensor nodes on the perimeter for higher accuracy) and using sampling protocols that allow for over-sampling such as [Szczytowski et al., 2010]. Generally, transport reliability is tunable through the number of transmissions.

As discussed in Chapter 2, the state-of-the-art on information quality [Sachidananda et al., 2010] and Quality of Service (QoS) [Chen and Varshney, 2004] [Martínez et al., 2007] in WSN lacks the online composite adaptation of sampling accuracy and transport reliability and timeliness to the network conditions and application requirements. The performed sampling accuracy satisfies the application requirements only if the information transport is perfect, which is not true in WSNs. On the other hand, the transport reliability and timeliness assume the sampling block to be perfect while addressing the application requirements. The optimized tunable co-design of sampling and information transport that maximizes the energy efficiency and minimize loss of information while satisfying the user requirements is lacking in the literature. In particular, there are no efforts in WSN addressing the composite tunability of sampling accuracy and transport reliability and timeliness.

Usually, the sensing application (users, services, feedback controller, etc.) has a specific requirement for the contortion accuracy. The contortion accuracy experienced at the sink fundamentally depends on the transport reliability. The key challenge has consequently been to tune sampling accuracy and transport reliability and timeliness in composition so that the requirement is met. The naive approach of massive over-sampling and allowing an arbitrary number of retransmissions might indeed result in high contortion accuracy. However, such a solution would be highly ineffective as it is not required to provide higher quality than the user requirements. On the other hand, this naive solution results in unacceptable energy overhead, which significantly limits the availability of the entire WSN. Our work emphasizes that sampling accuracy cannot be considered without transport reliability and timeliness for an optimized efficiency.

Common to all these observations is that the application requirements have to be exactly satisfied by considering the co-design of the functional blocks sampling and information transport. In addition, the right tradeoff between sampling accuracy and transport reliability, timeliness should be considered in all real-world applications to ensure satisfied applications. The challenge is finding methods for combining these attributes and localized algorithms for implementing the tunable co-design efficiently.

Achieving sampling accuracy, transport reliability as well as timeliness while maximizing efficiency requires a sophisticated technique, which is the core problem of this thesis. For our optimal tunable co-design problem, we aim to find the best balance between sampling accuracy and transport reliability and timeliness. The same user experience could be achieved by different combinations of all three attributes. For example, providing higher sampling accuracy would allow for lower transport reliability. As it is complex to provide the optimized solution, we progress stepwise to master the complexity. Using probabilistic analytical expressions for relating sampling accuracy, transport reliability, timeliness and efficiency, the desired outcome is a composition of the number of retransmissions per hop and the number of nodes to sample the phenomenon. The key challenging problem relies on minimizing the overall number of retransmissions and loss of information given the network topology (i.e., the number of paths and the number of hops per path), the user-required contortion accuracy, the link quality and transport timeliness.

In this thesis, we investigate some of the key solutions for supporting optimal tunable co-design in WSNs. However, as the core problem is explained, we break down the core problem into three sub problems to master the complexity of optimal tunable co-design in WSNs and are as follows:

4.1.1 Tunable Information Transport with Reliability and Timeliness

To satisfy the evolvable user requirements, the transport schemes have to be carefully designed in order to reliably and timely deliver the information to the sink. To proceed towards the optimal tunable co-design primarily we focus on the key operations of information transport and their quality attributes, i.e., transport reliability and timeliness.

Achieving the best possible reliability and timeliness is related to large overhead regarding resources, particularly because sensor nodes rely on batteries. A higher reliability usually is achieved through a higher number of retransmissions resulting in a higher energy/bandwidth overhead. Timeliness may require path splitting instead of simple retransmissions on the same path, thus, causing higher traffic related to higher energy/bandwidth overhead. Hence, besides attaining the required quality levels, it is indispensable to maximize energy/bandwidth efficiency.

As discussed before, there are applications that may be satisfied with lower reliability and timeliness. For example, biologists may tolerate the delayed and lossy delivery of forest temperature data. Lower application requirements represent an opportunity to increase the WSN efficiency. Varied applications may be satisfied with varied reliability and timeliness levels. To reduce deployment costs, WSNs are more and more required to serve multi-users for multi-purposes. For instance, the purpose of a WSN deployment may suddenly need to be changed. For instance, upon a catastrophic event, the WSN should support rescue operation and stop unnecessary monitoring activities. In future smart cities and rural areas, public WSNs should deliver different information entities to varied authorities or users.

Typically, users have different requirements on transport reliability and timeliness. Timeliness requirements may range from strict and real time to soft deadline that can vary from seconds to minutes to hours. Varied WSN users usually require the best effort reliability with different levels of efficiency. Best effort reliability requirements can be expressed in message delivery success rate or a ratio of event detection false positives or false negatives.

As WSN is more and more used for multi-purpose deployments, the WSN protocol suite such as information transport should provide for tunability in order to support these applications with varied/evolvable reliability and timeliness requirements while maximizing efficiency. Available approaches usually optimize for best effort reliability or timeliness. As it is not always required to provide best effort reliability or timeliness, it is challenging to just provide the user required performance. Unfortunately, there are no efforts addressing the tunability of both reliability and timeliness in composition. In this thesis, we address this tradeoff by providing the user required evolvable reliability and timeliness levels while maximizing efficiency.

Achieving both transport reliability and timeliness while maximizing efficiency requires a sophisticated tradeoff technique, which is one key contribution of this thesis.

4.1.2 Optimizing the Sampling Accuracy and Transport Reliability

A higher accuracy of spatial sampling is usually achieved through a higher number of sampling sensor nodes in the area of the physical phenomenon resulting in a higher energy/bandwidth overhead. However, the sampling accuracy satisfies user requirements assuming the information transport to be perfect, which is not true in WSNs. On the other hand, transport reliability usually achieved through a higher number of retransmissions assume the sampling block to be perfect while satisfying user requirements.

The user/application view considering the spatial phenomena of interest requires a certain sensing accuracy (e.g., perimeter of a hole in the phenomena [Ghosh and Das, 2008], spatial distribution of the phenomena [Szczytowski et al., 2010]). Considering the design view, the sampling accuracy can be tuned by injecting some redundancy (e.g., activating more sensor nodes on the border of the coverage hole) and protocols such as [Szczytowski et al., 2010] allow for over-sampling. On the other hand, transporting the samples from the spatial phenomena of interest with a certain designed reliability requires a co-design of sampling and information transport with certain sampling accuracy, transport reliability, best effort timeliness and maximizing efficiency. In this order, the sensing accuracy is the co-design of spatial sampling accuracy and transport reliability. Hence, for a given sensing accuracy, we provide the optimal solution with online adaptation by maximizing energy efficiency.

Fulfilling the evolvable user requirement while achieving both sampling accuracy and transport reliability and while maximizing efficiency requires a sophisticated tradeoff technique, this is the second key contribution of this thesis. In our approach, we aim to find an optimal tradeoff between sampling accuracy and transport reliability.

Mathematical Formulation

Providing a specific requirement of S_{min} samples, the application actually expects exactly $S_{rx} = S_{min}$ samples to be delivered. However, this guarantee is hard to be satisfied in WSNs. Therefore, we assume the application requires to meet the requirements with certain fidelity $Fi_{acc} \in [0, 1]$. Furthermore, generating only S_{min} samples and delivering all of them to the sink would require a large number of retransmissions.

Preliminary investigations have shown that by slightly increasing the number of generated samples S_{tx} we can significantly reduce the total number of transmissions needed to deliver S_{min} samples to the application. However, sending too many additional samples will finally result in an unnecessary high number of retransmissions. Hence, we aim to find the optimal number of additional samples and the optimal path reliability that result in a minimal number of total retransmissions. Such an optimization allows co-designing sampling and transport for a maximized message efficiency, which transforms into maximized energy efficiency, as usually radio is the most energy consuming module on a sensor node.

Summarizing, we formulate the problem as follows:

Minimize $\#ret_{total}$

subject to

$P(S_{rx} \geq S_{min}) \geq Fi_{acc}$

More precisely, $\#ret_{total}$ can be expressed depending on the network characteristics and the application requirements as we will elaborate in the next section. The expected result is to determine the optimal (S_{tx}, R_{path}) tuples for given network conditions (link reliability R_{link}, hop distance h) and application requirements (S_{min}).

4.1.3 Representing the Physical Phenomena with Optimized Network Performance

For a wide class of WSNs applications, it is crucial to satisfy the user requirements. Co-design of the functional blocks of sampling and information transport for having maximized efficiency and the protocol suite for real time adaptation to the varying network conditions is vital. Therefore, it is necessary to bind the sampling accuracy and information transport reliability and timeliness.

In the co-design of sampling and information transport it is trivial to consider the requirements of the application. We consider spatial accuracy, which can be understood as spatial resolution of event (hotspot) detection. Here, we not only determine how many number samples, but which samples are required to represent the physical phenomenon. Thus, accuracy can be tuned in terms of adding or removing active sensor nodes and in that way spatial resolution is increased or decreased, respectively. Increasing or decreasing the number of nodes to send a sample will increase, respectively, decrease the accuracy. As accuracy and reliability are closely related, reliability can be seen as a hidden requirement, derived from the accuracy requirement. Moreover, timeliness is closely related to both accuracy and reliability, increasing or decreasing one of the accuracy or reliability will directly affect the timeliness. Packet transmission is the most expensive atomic operations in WSNs [Shnayder et al., 2004]. We abstract energy efficiency in terms of the number of retransmissions, as energy consumption impacts the sensor nodes the most.

Furthermore, we consider the contortion (accuracy, reflecting how close the achieved spatial phenomenon in the real world is) between the sensor nodes is given as the difference between the signal magnitude. To have the desired contortion the signal differences are taken between every sensor node in the hotspot. Considering the evolvable user requirements, the application actually expects exactly certain contortion at the sink. However, this guarantee is hard to be satisfied in WSNs due to the reliability of information transport. On the other hand, considering the timeliness makes the optimal co-design more complex.

Generating only required samples for satisfying contortion and delivering all of them to the sink would require a large number of retransmissions. Sending too many additional samples will finally result in an unnecessary high number of retransmissions which hinders timeliness. Moreover, the complexity will grow as the spatial samples have different qualities and as we also consider the different link quality and different hops from the source towards the sink. As the key problem, we aim to find the optimal number of sensor nodes and the optimal number of retransmission that result in a minimal number of total retransmissions and satisfy all requirements. Furthermore, the timeliness requirement is met accordingly to the optimal combination of accuracy and reliability. Providing a multi-attribute correlation and representing the accurate physical phenomena with optimized network performance marks as one of the main contribution of this thesis.

Mathematical Formulation

In the co-design of sampling and information transport it is trivial to consider the requirements of the application. Fig. 4.1 shows that the hotspot at center is severe and then degrades moving away from the center. The contortion between the sensor nodes (D, E, and F) is given as the difference between the signal magnitudes. To have the desired contortion the signal differences are taken between every sensor node in the hotspot.

Given a specific requirement CR that is satisfied with certain contortion accuracy CA, the application actually expects exactly $CE = CA$ at the sink. More precisely, due to this information loss Δ_{acc} the experienced contortion CE will always be worse than the contortion at the source, i.e. $CE = CA + \Delta_{acc}$. On the other hand, instead of strict requirements, we assume the application requires to meet the requirements with certain fidelity $Fi_{acc} \in [0, 1]$ gearing it more towards the probabilistic nature of WSNs. Hence, the requirement at the source is to satisfy $P(CE \leq CR) \geq Fi_{acc}$. Based on the sampling value closer to the hotspot, Fig. 4.1 shows one such example, as sensor node (A)>sensor node (B)>sensor node(C).

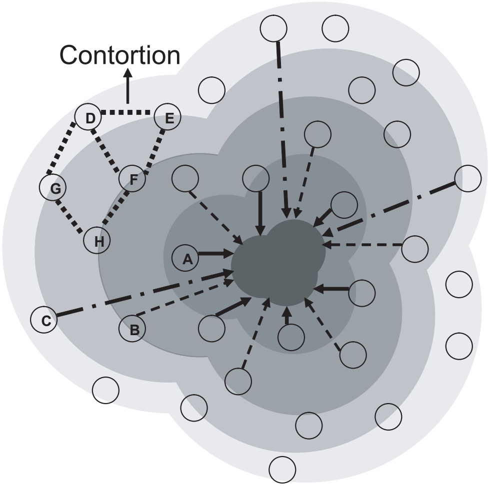


Figure 4.1: Hotspot growth and measuring of the contortion from the signals of sensor node (D), sensor node (E) and sensor node (F)

Generating only required samples for satisfying CR and delivering all of them to the sink would require a large number of retransmissions. Preliminary investigations have shown that by slightly increasing the number of generated samples we can significantly reduce the total number of transmissions needed. However, sending too many additional samples will finally result in an unnecessary high number of retransmissions which hinders timeliness. Moreover, the complexity will grow as the spatial samples have different qualities and as we also consider the different link quality and different hops from the source towards the sink.

As the key goal, we aim to find the optimal number of additional samples and the optimal $\#ret_h$ that result in a minimal number of total retransmissions and satisfy all requirements. We assume in this work that the sampling scheme is able to achieve the desired over sampling. The application requirement is to receive any set of samples that satisfies the required contortion.

Summarizing, we can formulate the co-design as:

$$Min\{C(\#ret_{total}) : CA, R, L\}$$

More precisely, the cost function C can be expressed depending on the individual network characteristics and the application requirements such that the total number of retransmissions ($\#ret_{total}$) are minimized.

4.2 Chapter Summary

In this chapter we have explained our core problem statement on optimized tunable co-design in WSNs. As the key goal of our thesis, we show that the optimal tunable co-design and multi-attribute optimization is an important problem for optimized network performance. Through this thesis, we show the mutual dependency of spatial accuracy, transport reliability and timeliness constraints and show the binding of all the three attributes through formulating it as the optimization problem.

Chapter 5

Tunable Information Transport Concerning Reliability and Timeliness

The continuous variation of application requirements and dynamic operational perturbations complicates the design of information transport in WSNs. This chapter targets a comprehensive solution for information transport in WSNs. Many applications require the delivery to be reliable and timely. However, increasing reliability/timeliness comes at the cost of higher energy consumption as in both cases, additional messages have to be sent: Retransmissions to increase reliability and information delivered via a second, faster path to ensure timeliness. Existing transport protocols over- or under-provide reliability and/or timeliness and lack optimized efficiency. This work aims in tuning reliability and timeliness in composition for a maximized efficiency. Our approach's takes the reliability/timeliness requirements as input and features a message efficiency that optimally meets user requirements. Information transport proceeds in two steps in a fully distributed way: (i) Finding the optimal number of retransmissions on a per hop basis with delay compensation, and (ii) path split and/or replication if reliability or timeliness requirements are violated.

In particular, this chapter makes the following contributions.

- The *rT* algorithm that provides tunable timeliness with best effort reliability. This algorithm finds the optimal number of retransmissions and implements delay compensation on a per hop basis. If delay compensation is not effective, a path replication is conducted.

- The *RT* algorithm that provides tunable reliability and timeliness in composition. *RT* extends *rT* by path replication if either retransmissions, or delay compensation at the same path are not effective, i.e., replicating the path if either timeliness or reliability requirements are violated.

- We show the performance of our algorithms against previous efforts through extensive simulations. To the best of our knowledge, we are not aware of any other algorithm that achieves the tuning of both reliability and timeliness satisfying the varying user evolvable requirements.

5.1 Overview

Considering the generalized WSN models discussed in Chapter 3, we now present a generic solution that dynamically and autonomously adapts to maintain the desired information transport reliability and timeliness. First, we define the reliability and timeliness which are specific to the tuning of information transport and as required by the user or achieved by a transport protocol. Furthermore, we provide illustrative scenarios on how our solution progress towards tuning both reliability and timeliness for information transport.

5.1.1 Illustrative Scenarios for the Proposed Information Transport

In Fig. 5.1, we illustrate three typical scenarios for information transport. These scenarios are the drivers to develop our algorithms.

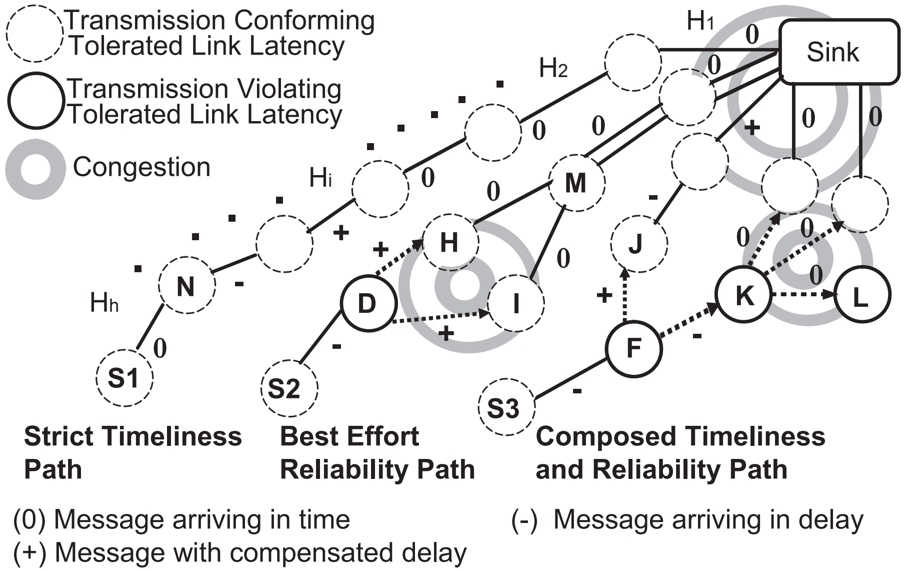


Figure 5.1: Three illustrative scenarios for the proposed information transport

In order to allow for a fully distributed solution, we propose to make per hop decisions. For instance, it has been proven that the per hop reliability in WSN outperforms the end-to-end (e2e) acknowledgment and retransmissions [Shaikh et al., 2010]. Accordingly, hop-by-hop retransmissions towards the sink are the standard approach. To this end, the overall path reliability is equally divided among all hops on the path. Similarly, we design a timeliness strategy on a per hop basis. Our approach provides the desired application reliability despite evolving application requirements and dynamic network conditions by adopting the adaptive retransmission techniques for tunable reliability from [Shaikh et al., 2010]. We modify the tunable reliability scheme in [Shaikh et al., 2010] to couple the selection of appropriate retransmissions per hop to the allowed tolerated link latency.

In case all required retransmissions can be performed within the tolerated link latency on all hops along the path no modifications to [Shaikh et al., 2010] are required. If on a hop H_i the number of required transmissions are not possible without violating the $L_{tol_{H_i}}$, then appropriate countermeasures are needed. In the following, we briefly discuss these developed countermeasures, which represent our main contributions. As mentioned before and in order to master the complexity, we proceed progressively by considering the three basic scenarios illustrated in Fig. 5.1, i.e., the information entities sent by S_1, S_2 and S_3.

Delay Compensation: Consider $S1$ that generates information and sends it to the sink. We assume that Node N requires a number of retransmissions which would violate the tolerated link latency. If the caused delay does not exceed a portion (say δ) of the tolerated link latency of the next hop, we propose a scheme for delay compensation. This strategy ensures strict timeliness notion while providing the best effort reliability.

Delay Compensation with Path Split: Consider $S2$ has made delay compensation, however, Node D can not conduct delay compensation anymore as the link latency would exceed the δ of next hops tolerated latency. Accordingly, we propose a mechanism to split the path to ensure $R_{d_{hop}}$ within the required $L_{tol_{H_i}}$. We refer to the path split by sending the same message to two neighboring nodes.

Delay Compensation with Path Replication: Consider the scenario of $S3$. Node F requires delay compensation and path split into two neighboring sensor nodes J and K. However, delay compensation and path split are not sufficient at Node K. Hence, Node K has to conduct a path replication to three neighbors (the number three is based on the number of remaining retransmissions). We refer to path replication by the fact of sending the same message to three or more neighboring nodes.

In all the scenarios above, we briefly explained how our approach efficiently finds the tradeoff between provisioning the reliability and timeliness on one side and minimizing the number of retransmissions on the other side, through delay compensation, then path split, then path replication if required. We note that path split and path replication are local decisions and the paths may converge to the same path after a certain number of hops (this means a node may forward the same message more than once, e.g., Node M).

5.1.2 Mapping User Requirements

Our aim is to satisfy user required reliability and timeliness. As we follow a hop-by-hop reliability and timeliness assurance, we should carefully map the e2e user requirements for the single hops. Obviously, the hop-by-hop selection of requirements should satisfy:

$$1 - \prod_{i=1}^{h}(1 - R_{d_{hop}}) \geq \mathrm{R}_d \text{ and } \sum_{i=1}^{h} L_{tol_{H_i}} \leq \mathrm{L}_{tol}.$$

For satisfying the user required reliability R_d we adopt the per-hop decisions which are equally distributed to every hop according to Eq. (5.3). Recall that h is the total number of hops from the information source to the sink.

To satisfy the required timeliness, we need a mechanism to perform per-hop decisions. Usually, the per-hop deadline computation can follow a constant, increasing or decreasing function. A constant function allocates the e2e deadline evenly to all the hops from the source to the sink, implicitly assuming that a packet would suffer the same delay at each hop.

Intuitively, in a convergecast network, the closer a node to the sink, the greater will be the traffic that the node has to forward towards the sink. Thus, longer will be the delay that a packet will suffer at nodes closer to the sink. Accordingly, a longer hop deadline should be assigned for the hops closer to the sink. Thus, the partitioning/mapping function should be increasing. This assumes that congestion occurs only in the surrounding of the sink (e.g., path from $S2$ in Fig. 5.1). The growth of deadlines can be then linear, polynomial or exponential. Inspired by exponential back-off algorithms that double the retry time upon an unsuccessful medium access, we propose to use an exponential growth for deadlines.

Similarly, the information source area usually undergoes high communication activities (event detection, aggregation, etc.). In some scenarios, more than one node from the event area will report information to the sink. This increases the contention level in that area. Accordingly, an information source should select higher hop deadlines. Usually, the contention at the source node is lower than at the sink that would receive data from different information source areas simultaneously. Between the information source area and the sink shorter hop deadlines can be allocated as messages may select different disjoint less loaded paths. In the following, we introduce a novel deadline partition model.

Considering both contention effects above, the hop deadline allocation can be calculated as an exponential decrease with the distance from the source ($\epsilon * e^{-\alpha*(h-h_i)}$) and an exponential increase towards the sink ($e^{\alpha*h_i}$). Accordingly, we propose to compute the tolerable latency on hop H_i using Eq. (5.1)

$$L_{tol_{H_i}} = \frac{\epsilon * e^{\alpha*(-h_i+(h/2))} + e^{\alpha*(h_i-(h/2))}}{\tau} + \beta \tag{5.1}$$

$\epsilon \in [0.5,1]$ is a constant to address the fact that deadlines at the sink should be higher than at the source; α is a constant to control the gradient of increase/decrease; β is the minimum deadline that should be allocated to a hop; τ is the time scale factor to be able to select deadlines so that $\sum_{i=1}^{h} L_{tol_{H_i}} = L_{tol}$.

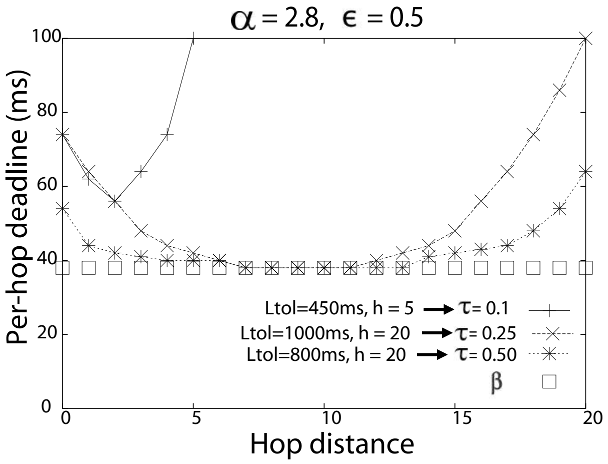


Figure 5.2: Hop deadline distribution along a path

Fig. 8.8 exemplifies the deadline assignment for 3 paths. Given the constants ϵ, α and β, a source node that is h hops from the sink can compute an appropriate τ, so that $\sum_{i=1}^{h} L_{tol_{H_i}} = L_{tol}$ is valid. The source node forwards τ and h values along the information so that every node on the path to the sink can calculate its own deadline using Eq. (5.1) without violating the e2e timeliness requirement.

Algorithm 1 Tunable Timeliness with Best Effort Reliability (*rT* Alg. at Hop H_k)

```
Const: $\epsilon$, $\beta$, $\alpha$, $TO$
Var: $L_{tol_{H_k}}$, $L_{C_k}$
start timers T1, T2;
if $k{=}{=}h$ then
    /*Source node*/
    $L_{C_k}$=0; $\delta$=0; $h$= No. of hops to the sink;
else
    Upon receiving a data $msg$
    extract $\tau$, $h$, $R_{d_{hop}}$, $L_{C_{k+1}}$
    $L_{C_k} = L_{C_{k+1}}$;
    send ACK to $H_{k+1}$;
    $\delta = L_{C_k} - \sum_{i=h}^{k} L_{tol_{H_k}}$;
end if
/*If message is not delayed*/
if $\delta < 0$ then
    $rT\text{-}Transport(msg, H_{k-1})$; exit();
else
    /*message is delayed*/
    if ($\delta \leq 0.3 * L_{tol-H_k}$) then
        /*Delay $\leq$ threshold $\rightarrow$ Delay Compensation Scheme*/
        $L_{tol_{H_k}} += 0.3 * L_{tol_{H_{k-1}}}$;
        $rT\text{-}Transport(msg, H_{k-1})$; exit();
    else
        if ($0.3 * L_{tol-H_k} < \delta \leq L_{tol-H_k}$) then
            Select a second next-hop $H'_{k-1}$
            $rT\text{-}Transport(msg, H_{k-1})$;
            $rT\text{-}Transport(msg, H'_{k-1})$; exit();
        else
            exit(); /* e2e deadline violated*/
        end if
    end if
end if
/*Upon receiving an ACK */
Stop timers T1, T2;

/*Function $rT\text{-}Transport()$*/

transport($msg$, $H_{k-1}$); {
/*Do $\#ret$ that are allowed in tolerated hop latency*/
while (T2.value() < $L_{tol_{H_k}}$) do
    for (i=0, i<$r$, i++) do
        $L_{C_k} += T.value()$;
        T1.reset();
        msg.append($\tau$, $h$, $R_{d_{hop}}$, $L_{C_k}$);
        send msg to $H_{k-1}$;
        wait for ACK or $TO$ expiration;
    end for
end while
}
```

5.2 Tunability of Reliability

In this work, we adopt the tunable reliability concepts proposed in GIT [Shaikh et al., 2010]. Therefore, we briefly summarize the GIT approach on providing tunable reliability. To ensure the desired reliability on one hop, more than one transmission may be required to overcome node and communication level perturbations. Given r the number of transmissions required, then the information transport reliability across Hop H_i is:

$$R_{H_i} = 1 - (1 - R_{L_i})^r \tag{5.2}$$

Since r is the total number of transmissions, $\#ret_{max} = r - 1$. For an R_d imposed by the application and known number of hops h from the sink, a source sensor node can calculate the desired reliability requirement across one hop as:

$$R_{d_{hop}} = R_d^{1/h} \tag{5.3}$$

$R_{d_{hop}}$ is forwarded by the source node along the path to the relay nodes. Once the decision of sending the information is taken from the sensor node, it calculates the maximum number of transmissions required to maintain the $R_{d_{hop}}$ using Eq. (5.2) [Shaikh et al., 2010] as follows:

$$r = \lceil \frac{\log(1 - (R_{d_{hop}}))}{\log(1 - R_{H_i})} \rceil \tag{5.4}$$

With Eq. (5.4) we can conclude that the achieved e2e reliability is a function of the path length and $\#ret$. The desired number of retransmissions required to satisfy the reliability R_d is without any time bound L_{tol}. Moreover, in the case of prioritizing reliability to timeliness either the timeliness is under- or over-provided. Prioritizing reliability to timeliness is appropriate for applications that are not sensitive to timeliness. Hence, in order to satisfy both reliability and timeliness, we need a supplemental mechanism.

5.3 Tunability of Timeliness

Now, we investigate prioritizing timeliness to reliability. The result is our first contribution, the *rT* algorithm, which is the first step towards a composite tunability, i.e., the *RT* algorithm. We discuss on how a possible delay can be compensated with and without path split. We first calculate the hop deadline distribution for satisfying the user specified requirements. Knowing, the value of L_{tol}, the source node divides the L_{tol} into tolerated hop latencies. Though, we divide L_{tol} into $L_{tol_{H_i}}$, we take into consideration the retransmission probability on a relay node, where reliability $R_{d_{hop}}$ will be time bound by L_{tol}. We append the user messages with $R_{d_{hop}}$, L_{C_i}, τ and h, where L_{C_i} is the cumulative latency from the source to Hop H_i.

We compensate the delay when the tolerated hop latency for intermediate an hop H_i is not met, as explained below. Unfortunately, in the case of prioritizing timeliness to reliability the reliability may be either under- or over-provided, i.e., $R_{H_i}{<}R_{d_{hop}}$ or $R_{H_i}{>}R_{d_{hop}}$. Thus, prioritizing timeliness to reliability, leads to best effort reliability.

We now present how delay compensation and path split can be achieved to better satisfy the e2e deadline. For this, we propose the rT algorithm. In any intermediate hop H_k, if $L_{C_k} > \sum_{i=h}^{k}\mathrm{L}_{tol_{H_i}}$, the process of compensating the delay and path split is conducted. On the other hand, the mechanism of path split into two neighboring nodes is also done. We split the path to two neighboring nodes because L_{tol} is without any bound R_d. However, it is not always true that we can compensate the delay and meet the e2e requirements of reliability and timeliness.

During tolerated link latency calculation, a node calculates the number of retransmissions that can be achieved for that current hop and tolerated link latency. Usually, we have sensor nodes meeting the tolerated link latency (i.e., $L_{C_k} < \sum_{i=h}^{k}\mathrm{L}_{tol_{H_i}}$) (Alg. 1, L. 15-16). However, if the link delay for a hop H_k is larger than the tolerated link latency, it is unfortunate to receive the ACK from hop H_{k-1} at H_k. Hence, when $L_{C_k} > \sum_{i=h}^{k}\mathrm{L}_{tol_{H_i}}$, our algorithm compensates the delay by borrowing the time from the next hop (Alg. 1, L. 20-22). The compensation condition is that $\delta = L_{C_k} - \sum_{i=h}^{k}\mathrm{L}_{tol_{H_i}}$ varies from $0{<}\delta{<}0.3 * L_{tol_{H_{k-1}}}$.

If the δ condition is violated, compensating the delay at Hop H_k, may not allow to send the required number of retransmissions. Hence, the path split approach decides to forward the information to two neighboring nodes H_{k-1} and H'_{k-1} (Alg. 1, L. 24-27). Though, tolerated link latency $\sum_{i=h}^{k}\mathrm{L}_{tol_{H_i}}$ for the intermediate hop H_k is not met, the reliability R_d across the hops is increased. After the message is forwarded to both next hops, the receivers will still check the condition $L_{C_k} > \sum_{i=h}^{k}\mathrm{L}_{tol_{H_i}}$, for compensating the delay. However, if the delay is higher than L_{tol}, the message is just dropped (Alg. 1, L. 29).

Though, we compensate the delay and split the path, we may still fail to ensure both the desired reliability and timeliness. Hence, in this case, we need a further mechanism.

5.4 Composite Tunability of Reliability and Timeliness

In this section we propose a solution which would provide the composite tunability of reliability and timeliness i.e., the *RT* algorithm.

5.4.1 Composite Reliability and Timeliness

A sensor node at Hop H_k includes $R_{d_{hop}}$, L_{C_k}, τ and h to the message when it forwards it to next hops. However, to achieve tunability and to reach a suitable tradeoff between reliability and timeliness, we need a holistic investigation of r, $\sum_{i=h}^{k} \mathrm{L}_{tol_{H_i}}$, R_d, $R_{d_{hop}}$, L_{tol} and L_{C_k}. To achieve the tradeoff between reliability and timeliness, the decision is based on nodes local network conditions and application requirements.

The path replication approach ensures for compensating the loss of reliability in any intermediate hop H_k when $L_{C_k} > \sum_{i=h}^{k} \mathrm{L}_{tol_{H_i}}$ and $R_{d_{hop}}$ is not satisfied. In order to maintain the required information transport reliability and timeliness, each node along the path dynamically adapts r according to its local timeliness and reliability requirements.

5.4.2 Trading Reliability and Timeliness

Now, we provide the composite tunability of the optimal timeliness along with improving the reliability of the information reaching the sink. Fig. 5.1 illustrate the algorithm execution.

The sink spreads the user defined e2e R_d and L_{tol} to all nodes. The source nodes (e.g.,S_1, S_2 and S_3 in Fig. 5.1) accordingly calculate the tolerated link latency and per-hop reliability. Node S_1 retransmits the message until it receives an implicit ACK by listening to a forward of the same message.

Now, consider Node S_2 during its first hop to D. The transmission of S_2 meet the tolerated link latency and also the per-hop reliability. For Node D, $L_{C_k} > \sum_{i=h}^{k} \mathrm{L}_{tol_{H_i}}$ and the per-hop reliability is lower than that required by the user, therefore, Node D first compensates the delay and also starts path split (Alg. 2, L. 22-32). Nodes H and I send implicit ACKs to Node D after the information is correctly received and forwarded. Nodes H and I forward the information to the next hop meeting the tolerated link latency $\sum_{i=h}^{k} \mathrm{L}_{tol_{H_i}}$. The next hop nodes forward the received information to the sink, until the tolerated link latency expires or an ACK is received.

Considering Node S_3, the information is forwarded to Node F. However, $L_{C_k} > \sum_{i=h}^{k} \mathrm{L}_{tol_{H_i}}$ at Node F. Apart from delay compensation, Node F decides on path split (Alg. 2, L. 22-32). The information forwarded to Node J meets the next tolerated link latencies and information reaches the sink. However, Node K suffers from supplemental delay. Hence, Node K conducts path replication to three nodes to have the required tradeoff between reliability and timeliness. The number of neighboring nodes for path replication are decided based on the required number of retransmissions and L_{C_k}, and if $r < \sigma$, then the path replication is carried on with adapting the reliability and timeliness (Alg. 2, L. 35-40). Except Node L, the other two nodes, which received the information from Node K send an implicit ACK to Node K. As Node L suffers from an additional delay, Node L sends a negative ACK to Node K or tolerated link latency expires before receiving ACK. If Node K receives positive ACK from other two neighboring nodes the retransmission to Node L is canceled while fulfilling the user timeliness and reliability requirements. Node K forwards the information to the next hops meeting tolerated link latency and delivering the information to the sink.

5.5 Chapter Summary

By introducing the tunable information transport this chapter established the necessary basis for the composite tuning of reliability and timeliness as per the application requirements. We have introduced the tunable timeliness, which efficiently assign the tolerable hop latencies on the path, compensates delays, and splits the path when needed. The optimized solution combines the re-transmission approach meeting the tolerable hop timeliness and the path replication approach when the tolerable hop timeliness is violated. This is the first instance of tuning when the combination of both the reliability and timeliness is implemented.

Algorithm 2 : Composite Tunability of Reliability and Timeliness (*RT* Alg. at Hop H_k)

```
Const: $\epsilon$, $\beta$, $\alpha$, $TO$
Var: $R_d$, $R_{d_{hop}}$, $L_{tol_{Hk}}$, $L_{C_k}$
start timers T1, T2;
if ($k$==$h$) then
    /*Source node*/
    $R_{d_{hop}} = R_d^{1/h}$;
    calculate $r$ using Eq. (4);
    $L_{C_k}$=0; $T2T = L_{tol_{Hk}}$; $h$= No. of hops to the sink;
else
    /*Upon receiving a data message $msg$ */
    extract $\tau$, $h$, $R_{d_{hop}}$, and $L_{Ck+1}$
    $L_{C_k} = L_{C_{k+1}}$;
    send ACK to $H_{k+1}$;
    $T2T = L_{tol_{Hk}} - (L_{C_k} - \sum_{i=h}^{k} L_{tol_{Hk}})$;
    $R_{d_{hop}} = R_d^{1/h}$;
    calculate $r$ using Eq. (4);
end if
/*If desired hop reliability can be satisfied and msg is not delayed*/
if (S of possible trans in $T2T \geq r$) then
    $RT\text{-}Transport(msg, H_{k-1}, r)$; exit();
else
    $\delta = T4r\text{-}T2T$;
    if ($\delta \leq 0.3 * L_{tol-H_k}$) then
        $L_{tol_{H_k}} += 0.3 * L_{tol_{H_{k-1}}}$;
        $RT\text{-}Transport(msg, H_{k-1}, r)$; exit();
        if ($0.3 * L_{tol-H_k} < \delta < L_{tol-H_k}$) then
            if ($r > \sigma$) then
                /*Path Split*/
                select a second next-hop $H'_{k+1}$;
                compute $r1$ for $H_{k-1}$ and $r2$ $H'_{k+1}$; /*$r = r1 + r2$*/
                $RT\text{-}Transport(msg, H_{k-1}, r1)$;
                $RT\text{-}Transport(msg, H'_{k+1}, r2)$; exit();
            end if
        else
            if ($r < \sigma$) then
                /*Path Replication*/
                Compute remaining $r_k$ for $H_{k-1}$;
                select $H_n$ neighbors;
                compute $r_n$ for $H_{n-1}$;
                $RT\text{-}Transport(msg, H_{n-1}, r_n)$; exit();
                if ($R_{d_{hop}} \geq R_d$) then
                    send Implicit ACK to $H_{k-1}$;
                    $RT\text{-}Transport(msg, H_{k-1}, r)$; exit();
                end if
            else
                Exit();
            end if
        end if
    end if
end if
/*Upon receiving an ACK */
Stop timers T1, T2;
/*Function $RT\text{-}Transport()$*/

$RT\text{-}Transport(msg, H_{k-1}, r)$ {
while ($T2.value() < L_{tol_{Hk}}$) do
    for (i=0, i<$r$, i++) do
        $L_{C_k} += T.value()$;
        T1.reset();
        msg.append($\tau$, $h$, $R_{d_{hop}}$, $L_{C_k}$);
        send msg to $H_{k-1}$;
        wait for ACK or $TO$ expiration;
    end for
end while
}
```

Chapter 6

Optimizing the Sampling Accuracy and Transport Reliability

In this chapter, we provide the sampling and information transport co-design for enhanced information transport by exploiting inherent spatial correlation of information in WSN. To overcome dynamic network conditions and evolving application requirements an adaptive retransmission mechanism based on spatial correlation is proposed. The presented solution relies on local adaptation mechanisms which ensure scalability. Analytical results show that the proposed solutions provide application specific spatial accuracy and reliability and save expensive retransmissions leading to energy efficient solution.

A key task in Wireless Sensor Networks (WSNs) is to deliver specific information about a spatial phenomenon of interest. To this end, a few sensor nodes sample the phenomenon and transmit the acquired samples, typically multihop, to the application through a gateway called a sink. Many applications require the spatial sampling to be accurate and the delivery to be reliable. However, providing a higher accuracy/reliability comes at the cost of higher energy overhead as additional messages are required: increasing the number of samples to increase the accuracy of sampling and increasing the number of retransmissions to increase the transport reliability. Existing design approaches overlook optimized spatial sampling accuracy and transport reliability in combination for minimizing energy consumption. This work aims at providing the optimized solution for sampling accuracy and transport reliability in composition for a maximized efficiency. Our approach features a message efficiency that optimally meets application requirements with the online adaptation and appropriate tradeoff between accuracy and reliability. The sampling and transport co-design proceeds by finding the optimal number of sensor nodes for the accuracy of the spatial sampling with the effect of reducing the number of retransmissions and still satisfying the application requirements. We validate the approach viability through analytical modeling and extensive simulations for a wide range of requirements.

This chapter in particular makes the following contributions.

- We provide a mathematical model for composite investigation of accuracy, reliability and efficiency.
- We formulate and solve a constrained optimization problem to determine the optimal combination of sampling accuracy and transport reliability that maximizes efficiency. Our solution relies on the proposed analytical model and considers varied levels of fidelity w.r.t. to exactly meet the application requirements for achieving a certain sensing accuracy.
- Through extensive simulations, we confirm the tunability and optimized performance of our sampling and transport co-design approach.

The rest of the chapter is organized as follows. First, we present an overview. Next, we give an overview of the sampling and information transport co-design, followed by the terminology and the problem statement. Next, we detail our approach on sampling accuracy and reliable information transport co-design, i.e., interlinking sampling accuracy and transport reliability for developing the optimal solution. Furthermore, we present the integrated sampling and transport algorithm. We provide the performance evaluation results in the end of the chapter followed by giving the summary of the chapter.

6.1 Overview

In WSNs delivering the gathered information with the application required quality is the main concern. To satisfy the required quality, it is crucial to carefully design the core functional blocks, such as (a) the sampling scheme in order to accurately represent the physical phenomena, and (b) the transport scheme in order to reliably deliver the information to the sink. In our work, we focus on the key operations of spatial sampling and transport along their quality attributes, i.e., accuracy and reliability respectively.

The user/application view considering the spatial phenomena of interest requires a certain *sensing task* (e.g., the perimeter of the phenomenon area [Ghosh and Das, 2008] on the spatial distribution of the phenomena [Szczytowski et al., 2010]). Moreover, the perceived sensing accuracy should satisfy the application requirements (e.g., accurate form and location of the event perimeter). In addition, future WSN deployments should allow for varied concurrent applications. Usually, these applications need varied information and have evolvable requirements.

Moreover, achieving the best possible sampling accuracy and transport reliability is related to a large resource overhead, particularly, because sensor nodes rely on batteries. A higher quality level is often related to higher deployment costs and higher resource overhead. A higher accuracy of spatial sampling of a spatial physical phenomenon of interest is usually achieved through a higher number of active sampling sensor nodes in the area of the physical phenomenon resulting in a higher energy/bandwidth overhead. On the other hand, the transport reliability usually is achieved through a higher number of retransmissions. Hence, besides attaining the required quality levels, it is indispensable to maximize energy/bandwidth efficiency. Considering the design view, the sampling accuracy can be tuned by injecting some redundancy (e.g., activating more sensor nodes on the perimeter for higher accuracy) and using sampling protocols that allow for over-sampling such as [Szczytowski et al., 2010]. Generally, transport reliability is tunable through the number of transmissions.

Usually, the sensing application (users, services, feedback controller, etc) has a specific requirement on the sensing accuracy. The sensing accuracy experienced at the sink fundamentally depends on the transport reliability. The key challenge has consequently been to tune transport reliability and sampling accuracy in composition so that the requirement is met. The naive approach of massive over-sampling and allowing an arbitrary number of retransmissions might indeed result in high sensing accuracy. However, such a solution would be highly ineffective as it is not required to provide higher quality than the application requirements. On the other hand, this naive solution results in unacceptable energy overhead. Our work emphasizes that sampling accuracy cannot be considered without transport reliability for an optimized efficiency.

Achieving both sampling accuracy and transport reliability while maximizing efficiency requires a sophisticated tradeoff technique, this is one of the main contributions of this thesis. In our solution, we aim to find an optimal tradeoff between sampling accuracy and transport reliability. The same user experience could be achieved by different combination of both attributes. For example, providing higher sampling accuracy would allow for lower transport reliability. As it is complex to provide the optimized solution, we progress stepwise to master the complexity. Our solution considers energy in terms of retransmissions and sampling accuracy in terms of samples needed at the sink. Using probabilistic analytical expressions for relating sampling accuracy, transport reliability and efficiency, the desired outcome is a composition of the number of retransmissions per hop and the number of nodes to sample the phenomenon. The key challenge relies on minimizing the overall number of retransmissions given the number of hops, samples required, the user-required sampling accuracy and the link quality.

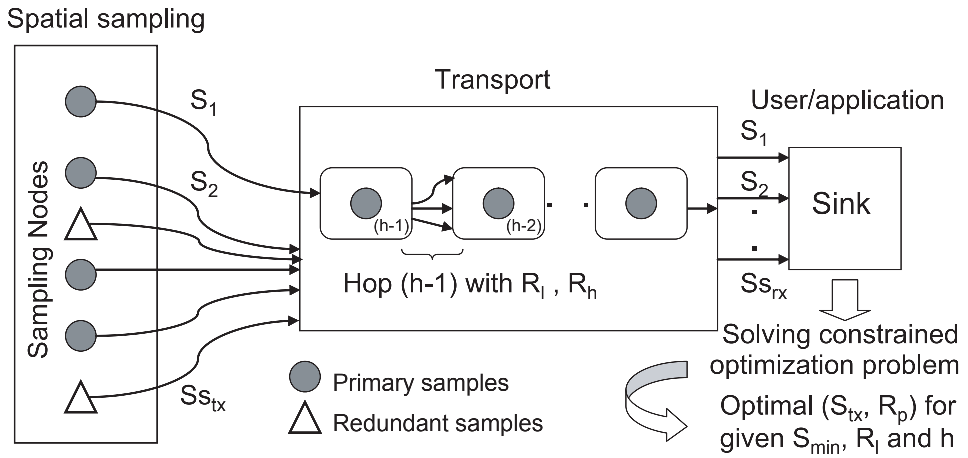


Figure 6.1: A holistic view of sampling, transport and application interactions

6.2 Sampling and Transport Co-design

We need to progress towards the optimal solution in a stepwise manner, Fig. 6.1 illustrates the two operations, spatial sampling and information transport. For readability, we emphasize one sample (S_1) and one path towards the sink. The main reasoning behind the targeted sampling and transport co-design is to online tune both operations using optimized S_{tx} and R_{path} values. To this end, we first solve the formulated optimization problem. This requires to analytically express the total number of retransmissions $\#ret_{total}$ as a function of the sampling accuracy S_{min} and transport reliability R_p and to select those pairs that globally minimize the $\#ret_{total}$.

The source nodes, which are in spatial phenomena of interest is considered for the spatial sampling. First, we need to calculate the required number of $\#ret_h$ per hop. Moving on we need then to find the total number of retransmissions per path and finally the overall retransmissions. On the other hand, we need to make sure that with the given optimal selection the right S_{tx} samples are transmitted from the spatial phenomena are. For example the source node S_1 now transmits its sample towards the next relay node, where on hop $h-1$ we need to satisfy the R_{link} and $\#ret_h$. The forwarding nodes now forward this information towards the sink. Now, this process continues with all the S_{tx} being transmitted which include the required and exact redundant samples. As now the S_{tx} samples are being transmitted from the spatial phenomena of interest. The transport channel has to make sure that given the application requirement of S_{min} and R_{link} from the sink reaches the sink. Finally, the success probability will be to receive the required S_{rx} show in Fig. 6.1 at the sink. This step wise complexity towards to optimal solution follows next.

6.2.1 Function of Tunable Sampling and Information Transport

The total number of retransmissions occurring for a certain number of samples transmitted for the phenomenon area is the sum of all retransmissions on all traversed hops. Then, the expected maximum $\#ret_{total}$ can be computed as:

$$\#ret_{total} = S_{tx} * h * \#ret_h \tag{6.1}$$

The number of retransmissions per hop is determined by the achieved hop reliability R_{hop} and the underlying link quality R_{link} as shown in

Fig. 6.1. The $\#ret_h$ is then computed as:

$$\begin{aligned} R_{hop} &= 1 - (1 - R_{link})^{\#ret_h} \\ \#ret_h &= \frac{log(1 - R_{hop})}{log(1 - R_{link})} \end{aligned} \tag{6.2}$$

Deriving Eq. (2) was the basic first step towards calculating the number of retransmissions per path. The achieved path reliability R_{path} depends on the achieved reliability at its hops as follows:

$$\begin{aligned} R_{path} &= R_{hop}^{h} \\ R_{hop} &= R_{path}^{\frac{1}{h}} \end{aligned} \tag{6.3}$$

In its turn, the achieved reliability R_{inf} depends on the achieved reliabilities of the paths the information (S_{tx} samples) traverses:

$$\begin{aligned} R_{inf} &= 1 - (1 - R_{path})^{S_{tx}} \\ R_{path} &= 1 - (1 - R_{inf})^{\frac{1}{S_{tx}}} \end{aligned} \tag{6.4}$$

The reliability of path R_{path} indeed derived from the average effective reliability reaching the sink, in turn, provides the success probability of at least S_{tx} samples will be delivered to the sink. In the primary case of just sending the required samples, regardless of which sensor nodes in the spatial phenomena of interest, the sensor nodes always send the samples to its next hop along the path R_p. Along each path, the S_{tx} out of N samples is trivial as the success probability of R_{inf} with p trials does not hold the negation of the probability at least S_{min} samples are received. However, we drive the given success rate based on Bernoulli process while providing the optimal solution.

Now we can express $\#ret_h$ as a function of h, S_{tx}, R_{inf} and R_{link} by substituting (6.4) in (6.3) and the resulting equation in (6.2):

$$\begin{aligned}
1-(1-R_{link})^{\#ret_h} &= R_{path}^{\frac{1}{h}} \\
1-(1-R_{link})^{\#ret_h} &= (1-(1-R_{inf})^{\frac{1}{S_{tx}}})^{\frac{1}{h}} \\
(1-R_{link})^{\#ret_h} &= 1-(1-(1-R_{inf})^{\frac{1}{S_{tx}}})^{\frac{1}{h}} \\
\#ret_h * log(1-R_{link}) &= log(1-(1-(1-R_{inf})^{\frac{1}{S_{tx}}})^{\frac{1}{h}}) \\
\#ret_h &= \frac{log(1-(1-(1-R_{inf})^{\frac{1}{S_{tx}}})^{\frac{1}{h}})}{log(1-R_{link})} \qquad (6.5)
\end{aligned}$$

However, we do not know the required reliability R_{inf} since it depends on the accuracy requirement. Next we will elaborate on two possibilities how to correlate accuracy and reliability.

In Eq. (6.5) only R_{inf} is still not determined. As we have pointed out before, the combined accuracy and reliability application requirement consists of the number of samples S_{min} that have to be delivered to the sink. Our approach is to allow for a controlled degree of over-sampling and transport reliability that minimizes the total number of retransmissions while delivering the required S_{min} samples. In this work, we assume that any S_{min} samples from the generated S_{tx} samples fulfill the application requirement. The relation between reliability and the number of samples received at the sink accordingly can be defined as:

$$R_{inf} = \frac{S_{rx}}{S_{tx}}$$

which can be modeled as the expectation value of a Bernoulli process with S_{tx} trials and a success probability of R_{inf}. Hence, Eq. (6.5) becomes

$$\#ret_h = \frac{log(1-(1-(1-\frac{S_{min}}{S_{tx}})^{\frac{1}{S_{tx}}})^{\frac{1}{h}})}{log(1-R_{link})} \qquad (6.6)$$

Substituting (6) in (1), we obtain the total number of retransmissions as a function of sampling accuracy and transport reliability. This represents a fundamental basis for solving a crucial optimization problem.

Lemma 1. *Assume that S_{min} samples is being transported on a path p, let h denote the number of hops from a node to the sink. If the success* *probability of R_{inf} with p trials does not hold the negation of the probability at least S_{min} samples are received, then the probability of receiving S_{min} samples are always around 50% and cannot be manipulated by the user.*

6.2.2 Optimal Sampling Accuracy and Transport Reliability

Using the example of a Bernoulli process, the equation for R_{inf} can also be written as:

$$1 - \left(\binom{S_{tx}}{0} R^0_{path} (1 - R_p)^{S_{tx} - 0} \right) = R_{inf}$$

and can be described as $P(at\ least\ 1\ out\ of\ S_{tx}\ samples\ is\ received)$ or according to the original notion as $1 - P(all\ S_{tx}\ samples\ are\ lost)$. It is obvious that no information about the expected number of samples or the probability of receiving them can be given.

In order to be more flexible and to meet the application requirements we need to express the probability that at least S_{min} samples are received, which should be greater than or equal to the fidelity requirement Fi_{acc}. We describe it as the complementary probability of the event, where up to $S_{min} - 1$ samples are lost:

$$1 - \left(\sum_{i=0}^{S_{min}-1} \binom{S_{tx}}{i} R^i_{path} (1 - R_p)^{S_{tx} - i} \right) \geq Fi_{acc} \tag{6.7}$$

Using this equation, the user can specify the reliability requirements more intuitively and precisely by providing Fi_{acc}. In order to use the above equation to derive our reliability requirements from the accuracy requirements, we need to solve it for R_{path}.

In the following we solve this by using the incomplete Beta function [Dutka, 1981]. Obviously, the equation above describes the cumulative distribution function:

$$1 - F(S_{min} - 1) \geq Fi_{acc}$$

With the following relation of the distribution function to the Beta distribution:

$$\sum_{i=0}^{k} \binom{n}{i} \cdot S_{tx}^{i} \cdot (1 - S_{tx})^{n-i} = I_{1-S_{tx}}(n-k, k+1)$$

where $I_x(a,b)$ is the regularized incomplete Beta function

$$I_{S_{tx}}(a,b) = 1 - I_{1-S_{tx}}(b,a)$$

we get the following derivation for R_{path}:

$$\begin{aligned}
1 - \left(\sum_{i=0}^{S_{min}-1} \binom{S_{tx}}{i} R_{path}^{i} (1 - R_p)^{S_{tx}-i} \right) &\geq Fi_{acc} \\
1 - F(S_{min} - 1) &\geq Fi_{acc} \\
1 - I_{1-R_{path}}(1 + S_{tx} - \lfloor S_{min} \rfloor, \lfloor S_{min} \rfloor) &\geq Fi_{acc} \\
I_{R_{path}}(\lfloor S_{min} \rfloor, 1 + S_{tx} - \lfloor S_{min} \rfloor) &\geq Fi_{acc} \\
I_{Fi_{acc}}^{-1}(S_{min}, 1 + S_{tx} - S_{min}) &= R_{path}
\end{aligned}$$

Therefore, the new expression for the number of retransmissions per hop, depending on the accuracy requirements is:

$$\begin{aligned}
1 - (1 - R_{link})^{\#ret_h} &= R_{path}^{\frac{1}{h}} \\
1 - (1 - R_{link})^{\#ret_h} &= I_{Fi_{acc}}^{-1}(S_{min}, 1 + S_{tx} - S_{min})^{\frac{1}{h}}
\end{aligned}$$

$$\#ret_h = \frac{log(1 - (I_{Fi_{acc}}^{-1}(S_{min}, 1 + S_{tx} - S_{min}))^{\frac{1}{h}})}{log(1 - R_{link})} \tag{6.8}$$

The optimal number of samples S_{tx} for a certain parameter setting can be found at the local minimum of $f(S_{tx}) = \#ret_{total}$, hence, the optimal number of active sensor nodes is:

$$\begin{aligned}\#ret_{total} &= \min\{f(S_{tx}) : S_{tx} \in \mathbb{N}\} \\ &= \min\{h \cdot S_{tx} \cdot \\ &\left\lceil \frac{log(1-(I^{-1}_{Fi_{acc}}(S_{min}, 1+S_{tx}-S_{min}))^{\frac{1}{h}})}{log(1-R_{link})} \right\rceil \\ &: S_{tx} \in \mathbb{N}\}\end{aligned} \tag{6.9}$$

Note that $\#ret_{total}$ will always be an integer value due to the ceiling function applied to the number of retransmissions per hop (Eq. (6.8)), since non-integer values obviously cannot be applied in practice.

6.3 Analytical Evaluation of the Sampling Accuracy and Transport Reliability

Based on the design goal, the objective function is to satisfy the application requirement given by the minimum number of samples S_{min} and a fidelity value Fi_{acc}, as indicated in the problem formulation and Eq. (6.9). Optimization and visualization of analytical results were conducted using Wolfram Mathematica [Wolfram, 1999]. Eq. (6.9) is plotted for selected settings in Fig. 8.1. Each graph consists of several linear segments resulting from the corresponding $\#ret_h$ value, which is highest for $S_{tx} = S_{min}$ and lowest as soon as so many samples have been added that retransmissions per hop are reduced to one (see Table 6.1 for examples of $\#ret_h$). Jumps from one segment to the next occur as soon as the reliability has been increased by redundant samples that much that Fi_{acc} is still satisfied when decreasing $\#ret_h$ by one. Note that there is always a small range where providing reliability using additional samples is more effective than using more retransmissions.

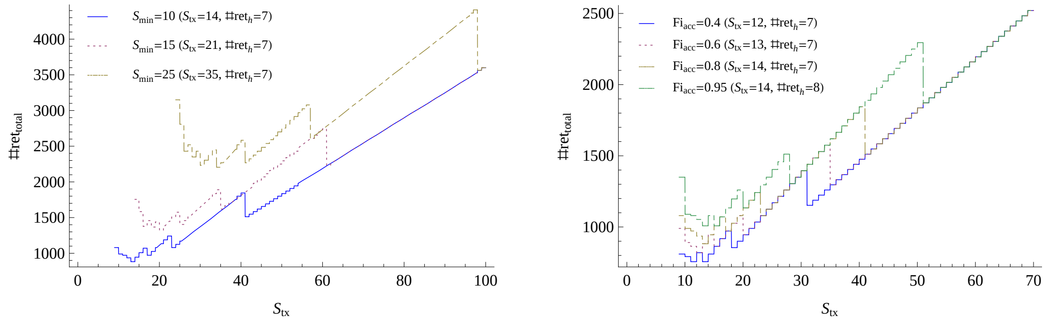
The main impact of the minimal number of samples required S_{min} is that at least S_{min} sensor nodes need to sample the phenomenon. Furthermore, by increasing S_{min} the steps become larger until the $\#ret_h$ can be reduced. Higher requirements on Fi_{acc} obviously need high sampling/transmission redundancy. Hence, the threshold to reduce the $\#ret_h$ is higher for stronger requirements.

Furthermore, a higher requirement on Fi_{acc} generally decreases the potential gain in efficiency by activating more nodes. Besides the linear impact on $\#ret_h$, determination of the slope of the graph and of the initial number of retransmissions, the number of hops per path h also impacts how fast the next lower $\#ret_h$ can be achieved. Finally, the link quality has a significant

S_{min} / Fi_{acc}		10		25	
	R_L / h	0.4	0.6	0.4	0.6
0.8	10	(S_{tx} | $\#ret_h$ |)	(S_{tx} | $\#ret_h$ |)	(S_{tx} | $\#ret_h$ |)	(S_{tx} | $\#ret_h$ |)
		14 | 7	14 | 4	35 | 7	34 | 4
	15	14 | 8	14 | 5	31 | 9	31 | 5
0.95	10	14 | 8	16 | 4	33 | 8	37 | 4
	15	14 | 9	14 | 5	33 | 9	33 | 5

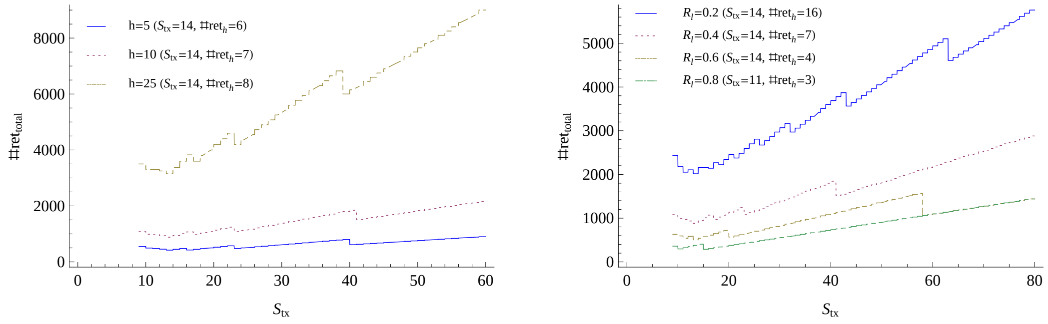
Table 6.1: Optimal tuples of the number of samples (S_{tx}) and the number of retransmissions per hop ($\#ret_h$) for a selection of parameter settings

impact, especially on the $\#ret_h$, since the low reliability has to be compensated by either more retransmissions or more samples. For a limited selection of WSN settings, we show the optimization results in Table 6.1.



(a) Constant values are: $Fi_{acc} = 0.8$, $h = 10$, $R_{link} = 0.4$

(b) Constant values are: $S_{min} = 10$, $h = 10$, $R_{link} = 0.4$



(c) Constant values are: $Fi_{acc} = 0.8$, $S_{min} = 10$, $R_{link} = 0.4$

(d) Constant values are: $Fi_{acc} = 0.8$, $S_{min} = 10$, $h = 10$

Figure 6.2: Graphs for different values of (a) S_{min}, (b) Fi_{acc}, (c) h and (d) R_{link}

6.4 Integrated Sampling and Transport

So far, we determined the optimal accuracy and reliability settings using global view. In the following, we present a localized integrated sampling and transport algorithm in generalized WSNs and its practicality.

After the phenomenon detection and notification from the source to the sink, the sink immediately knows about the important properties such as link reliability and hop count. Fidelity and accuracy requirements are provided by the user or the application and always accessible to the sink. Having this information the sink can then solve Eq. (6.9) for the phenomenon area. The attained optimal values (S_{tx} and $\#ret_h$) are reliably transmitted to the sources in the phenomenon area. The overhead induced by the reliable communication is negligible since only a single message has to be transported reliably. After the source has received the values for the sampling and transport co-design, it can use the existing duty-cycling algorithm to a) activate the right number of sensor nodes and b) to notify them about the number of retransmissions for the information transport. As soon as a sensor node is activated the user required sample is transported towards the sink with optimal number of retransmissions. As for the information transport each sensor node forwards the optimal number of retransmissions to upstream nodes by appending the number to the actual sample.

Assuming the sink knows R_{link} for a certain WSN deployment. In addition, we consider the sink knows the application requirements concerning the sampling (S_{min} can be either passed by the application directly or can be obtained from the sampling scheme). As the basic step, the sink solves the optimization problem constrained by the constraint Fi_{acc} and determines f_1 and f_2 for easy computation of optimized (accuracy, reliability) tuples.

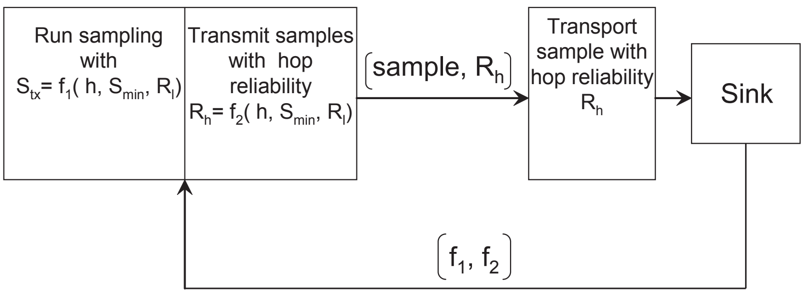


Figure 6.3: Basic message flow in integrated sampling and transport

The sink, then disseminates optimal f_1 and f_2 sensor nodes as show in

Fig. 6.3. The sampling sensor nodes need to cooperate to increase the number of samples from S_{min} to S_{tx} (according to $f_1(h,\ S_{min},\ R_{link}) = S_{tx}$) (Line 7 in Alg. 1). The sampling scheme controls the level of over-sampling in localized efficient manner (Line 8 in Alg. 1). Once the sampling nodes are determined, these nodes need to transmit the samples to the sink with a reliability according to $f_2(h,\ S_{min},\ R_{link}) = R_l$ (Line 11, 12 in Alg. 1). The computed R_{hop} should be disseminated along the message so that the tunable transport protocol can self-tune to perform the desired reliability.

Algorithm 3 The Design of Integrated Sampling and Transport

```
 1: Const: $h$, $R_l$, $S_{min}$
 2: Var: $f_1$, $f_2$, $S_{tx}$, $R_p$
 3: ————————————————
 4: /*Sampling SN → Tunable Sampling with $S_{tx}$:*/
 5: if $SN_i$ in phenomenon area then
 6:     $S_{min}$ = Sample();
 7:     $S_{tx} = f_1(h,\ S_{min},\ R_l)$;
 8:     execute Sample($S_{tx}$);
 9:     if $SN_i$ is a sampling node then
10:         take a sample $S_i$;
11:         $R_h = f_2(h,\ S_{min},\ R_l)$;
12:         transport($S_i$, $R_h$);
13:         Exit();
14:     end if
15: end if
16: ————————————————
17: /*Transport SN → Tunable Transport with $R_h$:*/
18: UPON reception of msg (sample, $R_h$):
19: transport(sample, $R_h$);
20: {
21:     $\#ret_h = \frac{log(1-R_h)}{log(1-R_l)}$
22:     Transmit msg (sample, $R_h$) $\#ret_h$+1 provided an implicit ACK is not
    received
23:     Exit();
24: }
```

6.5 Chapter Summary

Through this chapter we have achieved important steps towards the co-design of sampling and transport as per the application requirements. We have developed an analytical model for the case that no differences between sensor readings have to be regarded. This simplifies the problem of finding a specific subset of nodes to the problem of merely finding the optimal number of nodes that have to send a sample. Our analytical model gives the optimal number of sensor nodes, so that the specific application requirements are satisfied. The optimized solution provided depending on the application requirements, reduces the total number of retransmissions by adding redundancy and sending more samples than required. This is the first instance of real time adaptation when an integrated sampling and transport solution is implemented. The present work is just focusing on the accuracy and reliability attributes and is further being extended for additionally considering the timeliness attribute.

Chapter 7

Efficient Capture of Physical Phenomena with Optimized Network Performance

Wireless Sensor Networks (WSNs) constitute a rapidly growing research area, spanning both a wide variety of devices and applications. Typically, a WSN comprises a number of sensor nodes possessing limited processing and power capabilities, often communicating over unreliable and low bandwidth radio links [Akyildiz et al., 2002a]. Empirically, the core operation of a WSN is to satisfy user required information and transport from the network to the application via a gateway node termed as a sink. A primary design objective of WSN is to provide quality of information with *spatially accurate and responsiveness*, i.e., accuracy, reliability and timeliness.

7.1 Overview

A key task in wireless sensor networks is to deliver information from the sensor nodes to the sink. Many applications require the delivery to be accurate, reliable and timely. However, improving any of the accuracy, reliability and timeliness come at the cost of higher energy consumption as in all the cases, additional messages has to be collected and transported: (a) activating more sensor nodes for improving the data collection for accuracy, (b) retransmissions to increase reliability and information delivery via second, faster path to ensure timeliness. On the other hand, factors such as noise, signal variance and the correlation between sensor nodes also affect the accuracy of the information. Moreover, the information attributes are the functions of active sensor nodes. Elevating or improving any one of the information attributes i.e. accuracy, reliability and/or timeliness can have a direct impact on the other attributes and hence on the number of sensor nodes. Existing design approaches overlook multi-attribute correlation of spatial sampling accuracy and transport reliability and timeliness in combination for maximizing efficiency. This work proposes a co-design that proceeds by providing the mathematical notion for accuracy, reliability and timeliness and a cross-operation interaction of the attributes and their effects on each other. Furthermore, we provide the generalized holistic co-design algorithm and validate the approach viability through analytical modeling and simulations for a range of requirements.

WSNs are the key to gathering the information needed by smart environments. A WSN is required that is fast, accurate, reliable and easy to install and maintain. The challenges in the hierarchy of detecting the relevant quantities, monitoring and collecting the data, assessing and evaluating the information, formulating meaningful user displays, and performing decision-making and alarm functions are enormous. The information needed by smart environments is provided by distributed WSNs, which are responsible for sensing as well as for the first stages of the processing hierarchy.

In WSNs delivering the information with the user required information to the user/sink is of main concern. WSNs are mainly characterized by dense deployment of sensor nodes, which collectively transmit information about sensed events to the sink. Sensor nodes report the sensed information to the sink, which is usually stationary. Then from the sink user can access the performance of Network. Due to severely constrained resources, sensor nodes are subject to frequent failures. Therefore, WSNs are typically designed with a large number of redundancies to achieve fault tolerance and to maintain the desired network lifetime and coverage.

Typical WSN applications require spatially dense sensor deployment in order to achieve satisfactory coverage. As a result, several sensor nodes record information about a single event in a sensor field. But too much redundancy is also not useful as it hampers the parameters like reliability and timeliness by increasing the number of nodes. Thus, spatial correlation has to be exploited up to such extent that accuracy is not changed. Also itâs necessary to check the timeliness parameter because only those bits that are transferred prior to their deadlines contribute towards useful information. Deadlines could arise for various reasons, for example, the necessity to react to external events in a timely manner, and the need to deliver dynamically changing data prior to the expiration of their respective validity intervals. For time-sensitive applications, it is useful to understand delay in the network. Reliability is also an important factor that cannot be overlooked. The user is always interested in the probability that the information is delivered from source to sink. Hence the system should be reliable. Reliability can be maintained by many methods like, detecting the information loss and then recovering it, by using ACK reply message, but the method we are using is by controlling the number of retransmissions which controls the reliability. The recovery of lost information can be done by increasing the number of retransmissions, in such a way that it does not add in a delay in our end to end deadline for timeliness constraint and also does not affect accuracy of the information.

Depending on different applications, the traffic in the WSNs may be mixed with time sensitive packets, accuracy and reliability demanding packets. For periodic temperature record packets, as long as it arrives at the sink accurately, path delay is not critically significant. On the other hand, for multimedia packets, i.e. video packets, if most of them are received at a critical time, some loss is acceptable. Accuracy also plays an important role in both of the above scenarios. Another kind of traffic poses strict requirements for accuracy, delay and reliability. For example, for a danger warning packet, it should be delivered to the destination as soon as possible without loss and accurately.

To satisfy the user required information, we should carefully design the information attributes such as the accuracy of the samples to represent the real phenomena, the accuracy of the event detection from these samples, and the timeliness and reliability of the data/information transport from the sources towards the sink. From the above explanation, we note that these attributes may be orthogonal to each other. The intrinsic properties of WSN such as their energy constraints, and limited availability of resources, constitute an unfavourable environment for end-to end timeliness guarantees. Many existing solutions are based on a timeliness notion borrowed from real time systems, which can only express strict end-to-end deadlines. However, it is practically infeasible to impose these timeliness requirements in WSN without overestimating the network capacity. On the other hand, it is just infeasible to attain better timeliness without considering the accuracy of the samples and reliability of the data reaching the sink from the sources. So it only makes sense when we consider all these three constraints together in the same network and observe their effect on one another. As to satisfy user requirements, varying one constraint directly influences other constraints, which may not be desirable, it is always important and vital to consider these three constraints together and provide an optimized solution to satisfy user requirements. In this work, we address this multidimensional optimization problem by providing exactly the user required levels while efficiently performing trade offs between accuracy, reliability and timeliness. To the best of our knowledge, this is the first effort which considers all these three fundamental constraints together and proceeds further to optimize them by performing a perfect trade off to guarantee the user required information.

Achieving sampling accuracy, transport reliability as well as timeliness while maximizing efficiency requires a sophisticated technique, which is the main contribution of this article. In our optimal solution, we aim to find the best balance between sampling accuracy and transport reliability and timeliness. The same user experience could be achieved by different combinations of all three attributes. For example, providing higher sampling accuracy would allow for lower transport reliability. As it is complex to provide the optimized solution, we progress stepwise to master the complexity. Using probabilistic analytical expressions for relating sampling accuracy, transport reliability, timeliness and efficiency, the desired outcome is a composition of the number of retransmissions per hop and the number of nodes to sample the phenomenon. The key challenge relies on minimizing the overall number of retransmissions and loss of information given the network topology (i.e., the number of paths and the number of hops per path), the user-required contortion accuracy, the link quality and transport timeliness.

In summary, our contributions are as follows:

- We provide a mathematical model for composite investigation of accuracy, reliability, timeliness and efficiency.
- We formulate a constrained optimization problem to determine the optimal combination of sampling accuracy and transport reliability and timeliness that maximizes efficiency.
- We define a generic holistic co-design algorithm to provide for maximized efficiency while delivering the required sampling accuracy and transport reliability and timeliness.

- Through extensive simulations, we confirm the performance of our generic co-design approach.

The structure of the chapter is as follows. First, we describe some of the preliminaries with terminology and the problem statement. Next, we detail our approach on sampling accuracy and reliable and timely information transport co-design, i.e., interlinking sampling accuracy and transport reliability as well as timeliness for developing the optimal solution. Furthermore, we show the multi-attribute correlation and provide the analytical evaluation. Finally, we present the generalized holistic co-design algorithm and provide the performance evaluation results.

7.2 Theoretical Analysis Towards Responsive Co-design

In this section, we analytically express the accuracy, reliability and timeliness. Furthermore, we are going to bind the accuracy and reliability to proceed towards the multi-attribute co-design.

7.2.1 Sampling Accuracy

The notion of sampling accuracy is chosen as the contortion function to accurately represent the spatial phenomenon. Sampling accuracy in this work is defined by the $CA(M)$ function as expressed in Eq. (7.1) [Vuran and Akyildiz, 2006].

$$\begin{aligned} CA(M) = \sigma_s^2 - \frac{\sigma_s^4}{M(\sigma_s^2+\sigma_N^2)}(2\sum_{i=1}^{M}\rho(s,i)-1) + \\ \frac{\sigma_s^6}{(M)^2(\sigma_s^2+\sigma_N^2)^2}\sum_{i=1}^{M}\sum_{j\neq 1}^{M}\rho(i,j) \end{aligned} \tag{7.1}$$

As according to the study [Vuran and Akyildiz, 2006], more the contortion function value, the less is the accuracy. The sampling accuracy function mainly depends on M, σ and ρ. As according to our problem formulation, the contortion accuracy $CA(M)$ has to always satisfy the contortion required CR ($CA(M) < CR$), which is the threshold value/application requirement. In order to investigate the contortion achieved when a smaller number of nodes sending information, we assume that only M out of N packets are received by the sink, where N is the total number of sensor nodes in the event area. However, considering the information transport, it is not trivial that we always receive the M packets at the sink. In addition, due to the packet loss we always make the information out of scope for the application. On the other hand, if the contortion accuracy is not satisfied and the reliability of information transport of the packets is never determined, the considered spatial phenomenon at the source cannot be represented according to the ground truth at the sink.

7.2.2 Transport Reliability

Reliability builds on the link quality of each hop. Therefore, let RH_h be the Bernoulli variable indicating the successful delivery of a message at hop h with probability

$$\mathbf{P}\{RH_h = 1\} = 1 - (1 - l_h)^{\#ret_h} \tag{7.2}$$

where l_h is the link quality and $\#ret_h$ the maximum number of transmission attempts at hop h. Then, RP_p is 1 if the sample traversing path p arrives at the sink and 0 otherwise.

Thus,

$$\begin{aligned}\mathbf{P}\{RP_p = 1\} &= \prod_{h \in H_p} \mathbf{P}\{RH_h = 1\} \\ \mathbf{P}\{RP_p = 1\} &= \prod_{h \in H_p} (1 - (1 - l_h)^{\#ret_h}).\end{aligned} \tag{7.3}$$

Here H_p specifies the hops taken on path p. Let the information reliability RI be the Bernoulli variable that is 1 if and only if all the sent samples arrive at the sink.

$$\begin{aligned}\mathbf{P}\{RI = 1\} &= \prod_{p \in P} \mathbf{P}\{RP_i = 1\} \\ \mathbf{P}\{RI = 1\} &= \prod_{p \in P} \prod_{h \in H_p} (1 - (1 - l_h)^{\#ret_h})\end{aligned} \tag{7.4}$$

Note that reliability depends on the link quality and the number of retransmissions at each hop, however, only the latter can be tuned in deployed networks.

7.2.3 Transport Timeliness

To satisfy the required timeliness, we need a mechanism to perform per-hop decisions. Usually, the per-hop deadline computation can follow a constant, increasing or decreasing function. A constant function allocates the end-to-end deadline evenly to all the hops from the source to the sink, implicitly assuming that a packet would suffer the same delay at each hop. Intuitively, in a convergecast network, the closer a node to the sink, the greater will be the traffic that the node has to forward towards the sink. Thus, the longer will be the delay that a packet will suffer at nodes closer to the sink. Accordingly, a longer hop deadline should be assigned for the hops closer to the sink. The growth of deadlines can be then linear, polynomial or exponential. Inspired by exponential back-off algorithms that double the retrial time upon an unsuccessful medium access, we propose to use an exponential growth for deadlines. However, using various other forms of timeliness will fit into our optimization function. Accordingly, we compute the tolerable timeliness on hop h using Eq. (7.5)

$$LH_h = \frac{L}{2^h} \tag{7.5}$$

where h is the number of hops from the sink node and L represents the end-to-end deadline. However, considering this global model we can compute an upper bound for the timeliness for a particular setting. Assuming that the transmission behaviour is uniform throughout the whole network (especially over greater time) let d be a constant for the average delay a package experiences from the time it starts existing on one node until one transmission attempt is done. Then, the largest delay a package might experience on hop h is

$$LH_h = d \cdot \#ret_h$$

and

$$LP_p = \sum_{h \in P_p} L_h$$

is the largest delay on path p respectively. Thus, a timeliness requirement L is guaranteed to be satisfied, if $max\{LP_p | p \in paths\} \leq L$.

7.3 Binding the Sampling Accuracy and Transport Reliability

Considering just the contortion accuracy from Eq. (7.1) [Vuran and Akyildiz, 2006] and the drawback of loss of packets in information transport, we now wire the reliability with the contortion accuracy.

$$\begin{aligned} CE(M) = \sigma_s^2 - &\frac{\sigma_s^4}{E[X](\sigma_s^2+\sigma_N^2)}(2\sum_{i\in M}(\rho(s,i)-1)\cdot RP_i) \\ &+\frac{\sigma_s^6}{E[X]^2(\sigma_s^2+\sigma_N^2)^2}\sum_{i\in M}\sum_{j\in M\setminus\{i\}}\rho(i,j)\cdot RP_i\cdot RP_j \end{aligned}$$

However, as according to the problem formulation we need to satisfy $P(CE \leq CR) \geq Fi_{acc}$.

$$\begin{aligned} &\mathbf{P}(CE \leq CR) \\ &= \mathbf{P}\left\{\sigma_s^2 - \frac{\sigma_s^4}{E[X](\sigma_s^2+\sigma_N^2)}(2\sum_{i\in M}(\rho(s,i)-1)\cdot RP_i)\right. \\ &\left.+\frac{\sigma_s^6}{E[X]^2(\sigma_s^2+\sigma_N^2)^2}\sum_{i\in M}\sum_{j\in M\setminus\{i\}}\rho(i,j)\cdot RP_i\cdot RP_j \leq CR\right\} \\ &= \ldots \\ &= \sum_{\left(x_1,\ldots,x_{|M|}\right)\in\{0,1\}^{|M|}} \\ &\mathbf{P}\left\{\sigma_s^2 - \frac{\sigma_s^4}{E[X](\sigma_s^2+\sigma_N^2)}(2\sum_{i\in M}(\rho(s,i)-1)\cdot x_i)\right. \\ &\left.+\frac{\sigma_s^6}{E[X]^2(\sigma_s^2+\sigma_N^2)^2}\sum_{i\in M}\sum_{j\in M\setminus\{i\}}\rho(i,j)\cdot x_i\cdot x_j \leq CR\right\} \\ &\cdot\prod_{i=1}^{n}\mathbf{P}\left\{R_i = x_i\right\} \end{aligned}$$

Here X is a random variable for the number of samples received at the sink. The particular values which X can take are denoted by the corresponding combination of the x_n's which in turn describe one particular combination

of active nodes (i.e., node n is activated iff $x_n = 1$).
Furthermore,

$$E[X] = \sum_{n=1}^{|M|} n \cdot \mathbf{P}(X = n)$$

and

$$\mathbf{P}(X = n) = \sum_{(i,j) \in C(|M|,n)} \prod_i \mathbf{P}(RP_i = 1) \prod_j \mathbf{P}(RP_j = 0)$$

where $C(|M|, n)$ is a set of tuples of sets, the first of each denoting a particular combination of indices for n successful paths and the second denoting the corresponding combination of indices of $|M| - n$ unsuccessful paths.

7.4 Multi-attribute Sampling and Transport Co-design

Now, let M be the set of nodes available in the area of interest.Let $x_i, i \in M$ be 1 if node i will send its sample via path P_i and 0 otherwise, where P_i is the set of all nodes on the path from x_i to the sink. The function $p(i) = \sum_{h \in P_i} \#ret_h$ computes the number of retransmissions on the path the sample sent by node i will require. Let $z_k, k = 0, ..., 2^{|M|} - 1$ be 1 if the particular combination $(x_i x_{i-1} ... x_1 x_0)$ of the ordered x_i values corresponding to the binary representation of k satisfies the accuracy requirement and 0 otherwise.

Thus,

$$\begin{aligned} z_k = 1 \Leftrightarrow \sigma_s^2 - \frac{\sigma_s^4}{E[X](\sigma_s^2 + \sigma_N^2)} (2 \sum_{i \in M} (\rho(s,i) - 1) \cdot r_i) \\ + \frac{\sigma_s^6}{E[X]^2 (\sigma_s^2 + \sigma_N^2)^2} \cdot \sum_{i \in M} \sum_{j \in M \setminus \{i\}} \rho(i,j) \cdot r_i \cdot r_j \leq CR \end{aligned}$$

Using these notations and the constraints derived in the prior sections, we find the following optimization problem:

$$
\begin{aligned}
&\text{Minimize} \sum_{i\in M} x_i \cdot p(i) \\
&\text{subject to} \\
&\sum_{k=0}^{2^{|M|}-1} \left(z_k \cdot \prod_{i=1}^{|M|} \mathbf{P}\left\{R_i = x_i\right\} \right) && \geq Fi_{acc} \\
&max\{\sum_{h\in P_p} d \cdot \#ret_h\} && \leq L
\end{aligned}
$$

7.4.1 Analytical Evaluation of the Multi-attribute Correlation

In this section, we show the basic characteristics of the single attributes and proceed with the results of the optimal co-design with attributes and their effects on each other. This section gives us the basis for the approach of the holistic co-design and shows that there is always an impact on attributes when varying just one of them.

Fig. 7.1 shows how the path reliability is impacted by a different number of retransmissions per hop. As it is intuitive, a higher number of retransmissions can maintain a higher reliability. Accordingly, in Fig. 7.1 it can be seen that the probability that all sent samples arrive at the sink decreases with increasing number of samples. However, this notion of information reliability is no suitable metric for representing the user's requirements.

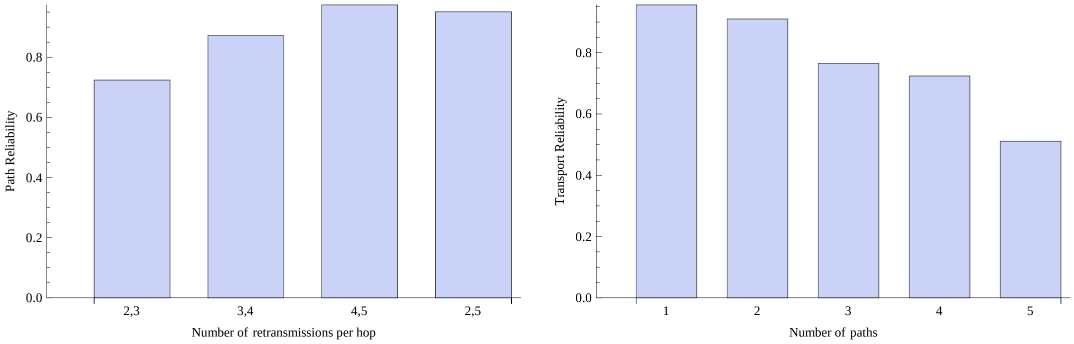


Figure 7.1: Top left figure: Impact of number of retransmissions per hop on path reliability. $\#ret_h$ have been chosen randomly out of intervals {2,3}, {3.4}, {4,5} and {2,5} respectively. Top right figure: Effect of number of parallel paths/active nodes on overall transport reliability

Similarly, Fig. 7.2 show the impact of network topology and a varying number of retransmission on the fidelity achieved at the sink. While in the practical scheme fidelity is given as a fixed requirement, we chose this representation to emphasize the impact of different parameters on the experienced sampling quality. Here each bar denotes analysis of a random network setup. As it can be seen, differences in network setup have more than just statistical impact.

Finally, in 7.3 show the impact of network topology and the impact of network size on attaining contortion accuracy. The sampling quality here tends to follow a statistical impact, while the fidelity requirement is fixed.

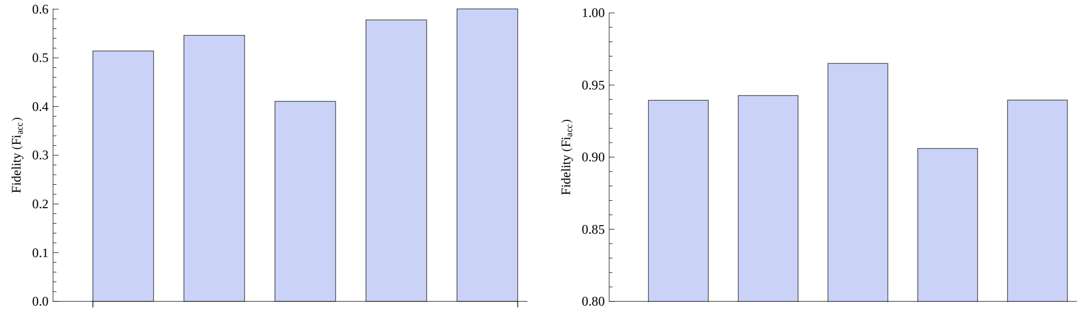


Figure 7.2: Impact of different network setups on the achieved fidelity. Top left figure: (n = 10, CR = 30, hops $\in_R$ {4,5,6}, $\#ret_h \in_R$ {2,3}). Top right figure: (n = 12, CR = 50, hops $\in_R$ {3,5,7}, $\#ret_h \in_R$ {3,4})

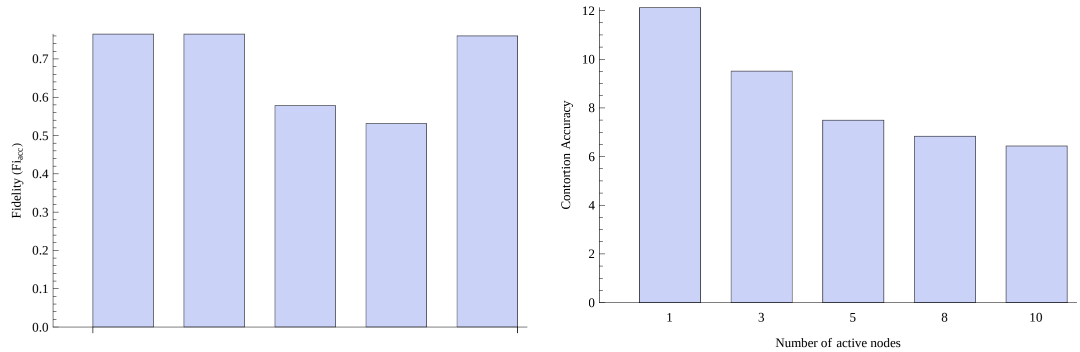


Figure 7.3: Top left figure: Impact of different network setups on the achieved fidelity. (n = 10, CR = 30, hops $\in_R$ {4,5,6}, $\#ret_h \in_R$ {2,4}). Top right figure: Impact of network size on attained contortion accuracy

We refer to [Vuran and Akyildiz, 2006] for characteristics of the accuracy definition. In a nutshell, efficiency and accuracy are opposed properties, as efficiency decreases when accuracy is increased.

7.5 Generic Holistic Co-design Algorithm

So far, we determined correlation of contortion accuracy and transport reliability and timeliness using a global view and showed how to optimally consider it analytically. In the following, we present generic holistic co-design algorithm in generalized WSNs.

After the phenomenon detection and notification from the source to the sink, the sink immediately knows about the important properties such as link reliability and hop count (together implying L) and M, e.g. by a n-hop neighbor restricted hello-protocol. σ and ρ are specified by assumption or knowledge about the phenomenon. Fidelity and contortion accuracy requirements are provided by the user or the application and always accessible to the sink (Line 5-8 in Alg. 1). Since concrete values for contortion accuracy vary from phenomenon to phenomenon and have no inherent meaning other than a comparative one, in practice a further abstraction level would be applied for better usability.

To calculate the optimal values to provide the required contortion accuracy, transport reliability and timeliness, we need to solve the constrained optimization problem. The optimized values known by the sensor nodes pertain to satisfy the user requirements obtained from the sink. We consider the sink to know the application requirements concerning the sampling and information transport. Moreover, the requirement ($P(CE \leq CR) \geq Fi_{acc}$) is the key aspect to be solved by the sink. As the basic step, the sink solves the optimization problem constrained by Fi_{acc} and determines the optimal parameters (set of active nodes N; $ret_h, \forall h \in P_i, \forall i \in N$) for the given network state (Line 11-13 in Alg. 1).

The attained optimal values are reliably transmitted to the sources in the phenomenon area (Line 14 in Alg. 1). The overhead induced by the reliable communication is negligible since only a single message has to be transported reliably. After the source has received the values it can implement the optimal sampling and transport co-design, e.g. by using an existing duty-cycling algorithm to a) activate the right number of sensor nodes and b) to notify them about the number of retransmissions with given timeliness for the information transport. As for the information transport, upon receiving the data message, each sensor node forwards the controlling information including the sample value if selected for sampling (Line 32-30 in Alg. 1). Note that no information regarding timeliness is needed to be distributed since it is implicitly satisfied by applying the optimal number of retransmissions on the selected paths. The information transport is handled with *GHC* transport function (Line 34-42 in Alg. 1).

The *GHC* algorithm enables dynamic application of the sampling and transport co-design under varying conditions while keeping the message overhead low. In Chapter 8, we provide the results and show the validity of *GHC* Algorithm.

Algorithm 4 : Generic Holistic co-design Algorithm (*GHC* Alg.)

```
Const: $CR$, $M$, $\sigma$, $\rho$, $Fi_{acc}$
Var: $CE$, $\#ret_h$, $h$, $R$, $L$, $x_i$, $N$
————————————————————
/*Sampling SN → Sampling with $CE$:*/
/*After the phenomenon detection:*/
if $SN_i$ in phenomenon area then
        Transmit msg (sample, $h$, $R$, $M$) to the sink
end if

/*Sink solves the optimization problem*/
    Min($x_i \cdot p(i)$) s.t.
    ($P(CE \leq CR) \geq Fi_{acc}$
     $max(LP_i) \geq L$)
    Transmit msg ($ret_h$, $N$) to the source

/*Sampling SN → Upon receiving msg:*/
while active neighbors $\leq |N|$ do
   activate neighbor nodes
end while

for $SN_i$ in network do
   /*Upon receiving a data message msg */
   if msg is activation message then
      sample ← doSampling();
   else
      sample ← NULL;
   end if
   reply with ACK
   GHC.transport(msg, $SN_i$, sample);
   Exit();
end for

/*Function $GHC.transport()$*/
$GHC.transport(msg, SN, sample)$ {
  $\#ret_h$ ← msg.getRetrans(SN.getId());
if sample != NULL then
   msg.setSampleData(sample);
end if
for (r = 0; r $\leq$ $\#ret_h$; r++) do
      Forward msg to next hop;
      wait for ACK or expiration;
end for
}
```

7.6 Chapter Summary

Through this chapter, we have provided the tunable co-design to efficiently capture the physical phenomena with optimized network performance. We have shown the mutual dependency of spatial accuracy, transport reliability and timeliness constraints and have shown the binding of all the three attributes through formulating it as the optimal co-design. We have provided the mathematical model for representing the physical phenomenon as a user requirement. We have also provided the notions for all the three attributes and how to move towards the optimal combination. Through analytical results we have observed that all the three attributes are interrelated and have an impact on each other. On the other hand, with our GHC Algorithm, we represent as a first algorithm which considers all the three attributes and also satisfies the user requirements. For future work, we aim to consider the in-network processing techniques for co-design and to check if we can minimize the loss of information. On the other hand, we also plan for tuning and adapting the existing MAC-protocols for the generic co-design.

Chapter 8

Performance Evaluation and Experimentation

In this thesis, we developed tunable co-design for assuring accuracy, reliability and timeliness and avoiding under and over provisioning of information in Wireless Sensor Networks (WSNs). This chapter provides the performance evaluation and experimentation of the work presented in this thesis. As the basis first we present the simulation environment we have been using. Next, we present the simulation results and then finally we conclude this chapter with the experimentation on the WSN testbed.

8.1 Simulation Environment

The main goal of the thesis was to develop optimized sampling and information transport co-design to support multi-attribute interaction of accuracy, reliability and timeliness in WSNs. Our research effort was driven by the current need of generic solution for sampling and information transport co-design to fulfil evolving application requirements and dynamic network conditions.

We simulate between 20 to 200 sensor nodes in an area of 75×75 unit^2 which is partitioned in a grid topology. The distance between two neighbouring nodes is 5 units. The sink is located in one corner. The information is generated from one corner and from the middle of the network and transported towards the sink. Furthermore, for the case of representing an accurate physical phenomenon, the information is generated from the phenomenon area and transported towards the sink.

The performance of our protocol is measured in terms of contortion accuracy and transport reliability, timeliness and average number of transmissions (which includes all the transmissions and retransmissions from the source node to the sink. The accuracy is measured in terms of contortion, on how close the evaluation information is to the real world value. The reliability is measured as the successful delivery of messages in behold of the end user requirements. The timeliness are measured in *ms.*

8.1.1 Simulation Studies

For the evaluation of our work we perform three case studies. The first case study deals with the varying user requirements. As its complex solution of the tunability of accuracy, reliability and timeliness and co-design of sampling and information transport, we break the first case study in three different parts. We proceed in a stepwise manner towards the final solution. First, we perform simulations for information transport with reliability and timeliness. Second, we perform simulations for the co-design of the sampling and information transport with accuracy and reliability. Third, we present the evaluation for the tunabality of accuracy, reliability and timeliness with co-design of sampling and information transport.

In the first case study, first we perform simulations for information transport with varying R_d and L_{tol}. To evaluate the tunability of our algorithms, we vary the desired reliability and keep the desired timeliness constant. Next, we keep the desired reliability constant and vary the desired timeliness. We select representative protocols from the existing literature as discussed in Chapter 2 and compare them with our work. The competitor protocols we have chosen are GIT [Shaikh et al., 2010], $CFLOOD$ [Jiang et al., 2009] and *MMSPEED* [Felemban et al., 2006].

Second, we perform simulations for the co-design of sampling and information transport for the performance of our protocol measured in terms of accuracy (samples received at the sink) and reliability (success probability of messages from source to sink). Any of these combinations, computable by using the optimal solution for the $\#ret_h$ are valid, i.e., they guarantee the delivery of S_{min} samples with a probability greater than or equal to Fi_{acc}. To show the online adaptation of our work, we first vary the link reliability R_{link}, keeping Fi_{acc} and S_{min} as constant. Finally, we vary S_{min} while keeping Fi_{acc} and R_{link} as constant. In all figures the optimal tuple is highlighted with a square box and have considered the median.

Finally, we perform simulations for different combinations of accuracy (CA), reliability (R) and timeliness (L) on varying user requirements. For representing the accurate physical phenomenon first show the simulation results for varying the desired accuracy and keep the desired reliability and timeliness constant. Next, we keep the desired accuracy and timeliness constant and vary the desired reliability. Finally, we vary timeliness and keep accuracy and reliability constant. Furthermore, we also show the results considering the efficiency with varying accuracy, reliability and timeliness. We measure the efficiency by the average number of transmissions (which includes all the transmissions and retransmissions from the source node to the sink). We select representative protocols from the existing literature as discussed in Chapter 2 and compare them with our work. The competitor protocols we have chosen are *ASample* [Szczytowski et al., 2010], *GIT* [Shaikh et al., 2010], *CFLOOD* [Jiang et al., 2009] and *MMSPEED* [Felemban et al., 2006].

In case study 2 we show the results for dynamic system requirements. First, we vary the network size for desired accuracy, reliability and timeliness. Second, we also vary the traffic with varying the information rate, i.e., the number of messages sent from the source per second. Finally, in order to address various perturbation levels, we varied the Bit Error Rate (BER).

Furthermore, we proceed towards the case study 3 for evolving phenomenon. First, we show the impact of fidelity on attaining accuracy, reliability and timeliness. Next, we show the simulation results from the varying area of the WSN. Finally, in the case study 3, we show the impact of phenomenon intensity on accuracy, reliability and timeliness.

8.2 Case Study 1: Simulation Results For Varying End User Requirements

In this section we provide the simulation results for the varying end user requirements. First, we provide the simulation results for the tunability of information transport.

8.2.1 Tunable and Adaptable Information Transport with Tradeoffs Between Reliability and Timeliness

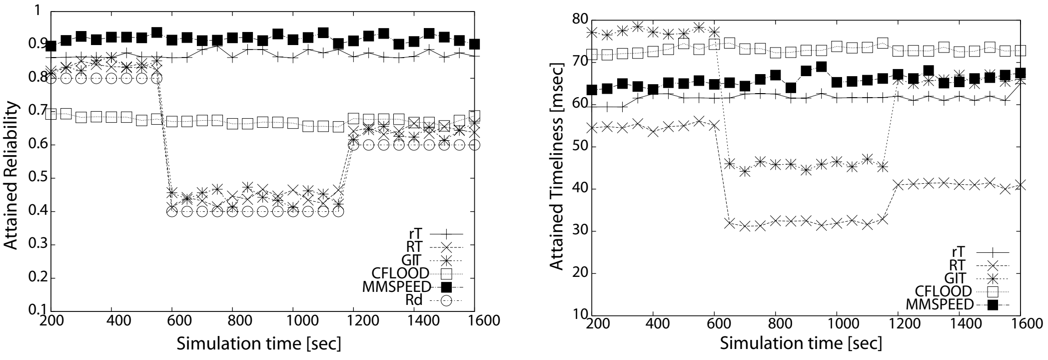


Figure 8.1: Tunability for varying desired reliability

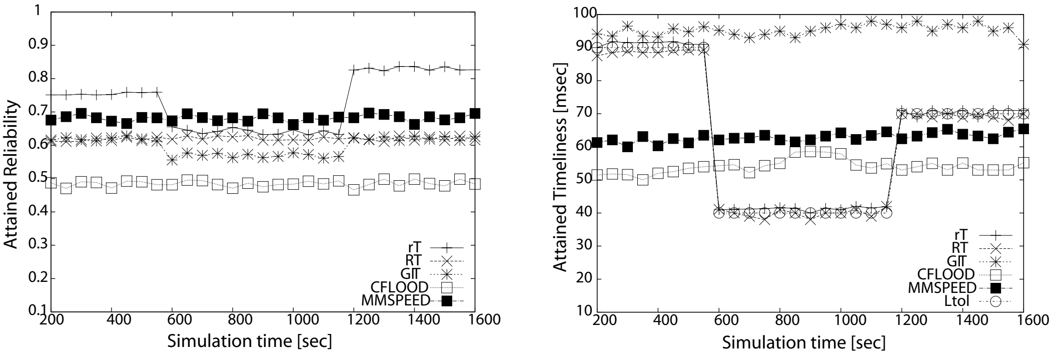


Figure 8.2: Tunability for varying desired timeliness

Fig. 8.1 shows the adaptation to user requirements with varying reliability, i.e., $R_d = 0.8$, 0.4 and 0.6 and constant desired timeliness $L_{tol} = 60$ms (τ

= 1.75). RT and GIT adapt according to the application requirements and provide tunable reliability. $CFLOOD$ and $MMSPEED$ do not adapt to the varying reliability requirements. RT outperforms other algorithms w.r.t timeliness, thanks to the tunability of RT.

Fig. 8.2 shows the adaptation to application requirements with varying timeliness, i.e., constant desired reliability R_d = 0.5 and L_{tol} = 90ms, 40ms and 70ms and corresponding values of τ = 1.75, 2.65 and 1.65. RT adapts according to the application requirements with varying timeliness and provides the desired reliability (Fig. 8.2). The timeliness of RT are always satisfying and outperforms GIT, $CFLOOD$ and $MMSPEED$, because of its adaptability and path split mechanism.

8.2.2 Cross Fertilization of Sampling Accuracy and Information Transport Reliability

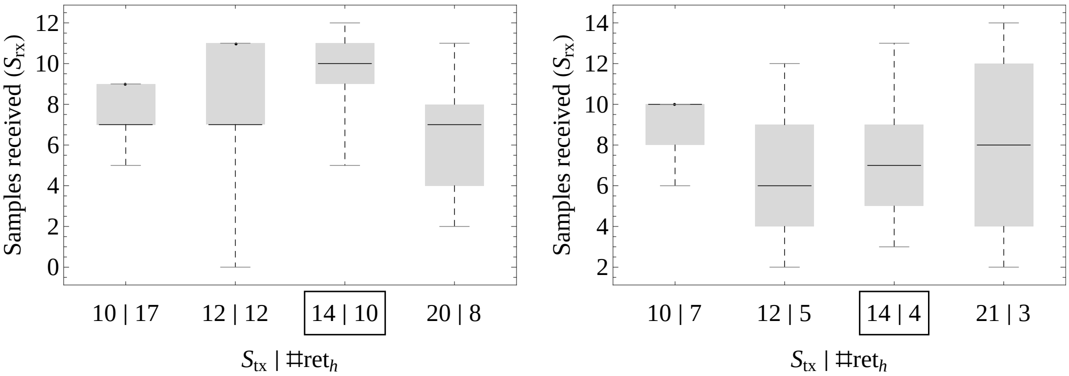


Figure 8.3: Impact of varying link reliability (R_l) on sensing accuracy, [R_l = 0.3, Fi_{acc} = 0.8] and [R_l = 0.6, Fi_{acc} = 0.8]

In order to verify the analytical optimal solutions (from Chapter 6)complying to optimal solutions in the simulations, we selected combinations around the optimal combination, i.e., combinations with less samples and more $\#ret_h$ as well as combinations with more samples and less $\#ret_h$.

Hence, regarding the total number of transmissions, if simulations for all combinations yield worse results than for the optimal ones, it is a strong indicator that this combination has indeed been optimized. As for the samples received, we expect very similar results throughout all variations as an indicator for the tunability of our model. In some cases, when an optimal solution existed for more than one combination, the first one, i.e. using less active sensor nodes, was chosen. Furthermore, in practice, the number of

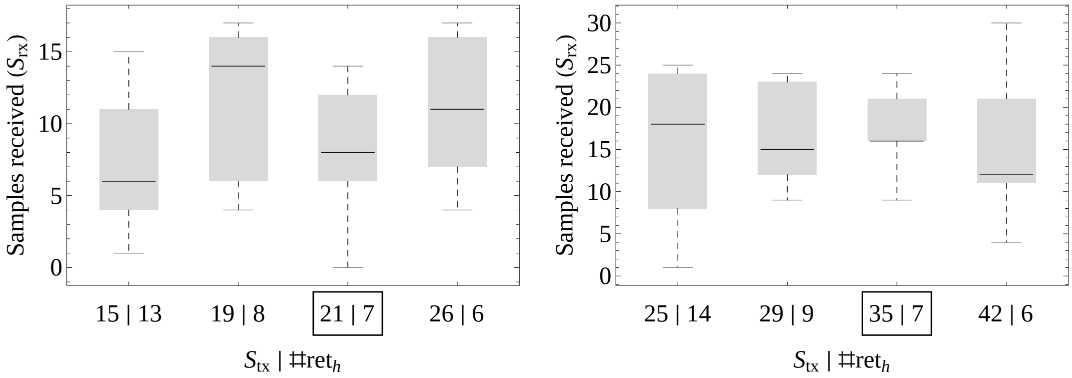


Figure 8.4: Impact of varying samples application requirements (S_{min}) on sensing accuracy, [$S_{min} = 15$] and [$S_{min} = 25$]

transmissions is very likely to be less than the computed values, as it is dependent on whether delivery was successful or not. In all figures the optimal tuple is highlighted with a square box.

Representative results of the simulations for varying the constraint R_{link} are shown in Fig. 8.3. As expected, the choice of R_{link} has influence on the delivery of S_{min} samples, as all combinations denote valid options to satisfy the requirements. Furthermore, the delivery rate has been slightly higher for the simulations with higher requirements for Fi_{acc}, completely conforming to our expectations. In all the cases of varying R_{link} we achieve the desired number of S_{min} samples to reach the sink.

Representative results of the simulations for varying the S_{min} samples are shown in Fig. 8.4. The results are very close to our analytical solution. The impact of varying S_{min} samples does not affect the application requirements and completely satisfies the probability of satisfying the Fi_{acc} and S_{min}.

Finally, it remains to observe that the variations in hop length and link reliability introduced in simulations in contrast to the analytical model result in a noticeable noise throughout simulations, blurring the differences between the different variations.

8.2.3 Tunable and Adaptable Co-design for Enhancing Network Performance with Accuracy, Reliability and Timeliness

In this section, we present the simulation results for representing the accurate physical phenomenon. We show the tunability of our work with varying accuracy, reliability and timeliness. The application requirements for varying accuracy are $CA = 80$, 60, 40, $R = 0.8$ and $L = 60$ms. The requirements on varying reliability are $CA = 80$ $R = 0.8$, 0.6, 0.4 and $L = 80$ms. Finally, the requirements for varying timeliness are $CA = 70$ $R = 0.6$ and $L = 80$ms, 50ms, 70ms.

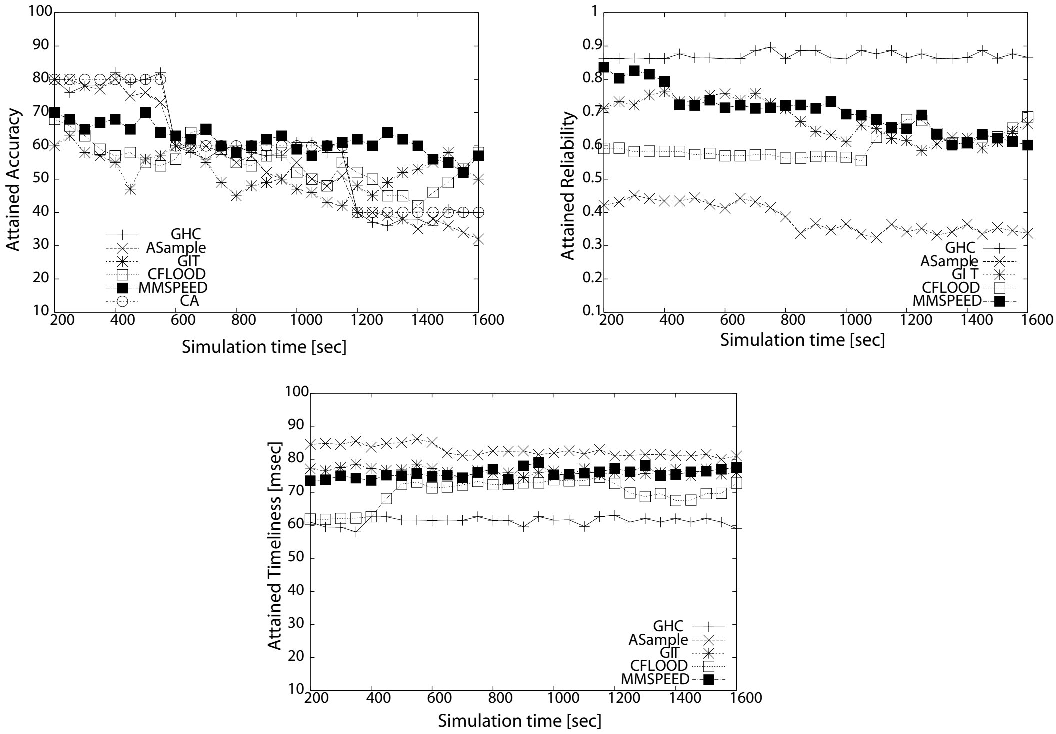


Figure 8.5: Tunability for varying desired accuracy

Fig. 8.5 shows the tunability of accuracy with varying user requirements on accuracy $CA = 80$, 60, 40. For the fair evaluation, we have considered the definition of the contortion accuracy to measure the contortion achieved by different protocols. We observe that GHC attain the desired contortion accuracy with a slight difference than the user requirement. However, other compared protocols shows decline in their performance as compared to the user requirement. And *ASample* shows a better performance than the compared GIT, $CFLOOD$ and $MMSPEED$, but not comparable to GHC. The contortion accuracy attained by GIT, $CFLOOD$ and $MMSPEED$ are independent of the desired contortion accuracy. On the other hand, GIT, $CFLOOD$ and $MMSPEED$ also fail to represent the real phenomenon at the sink. Considering reliability with requirement of $R = 0.8$, GHC performs according to the requirement and overcomes the perturbations and satisfies the user requirements. Though *ASample* performs slightly better

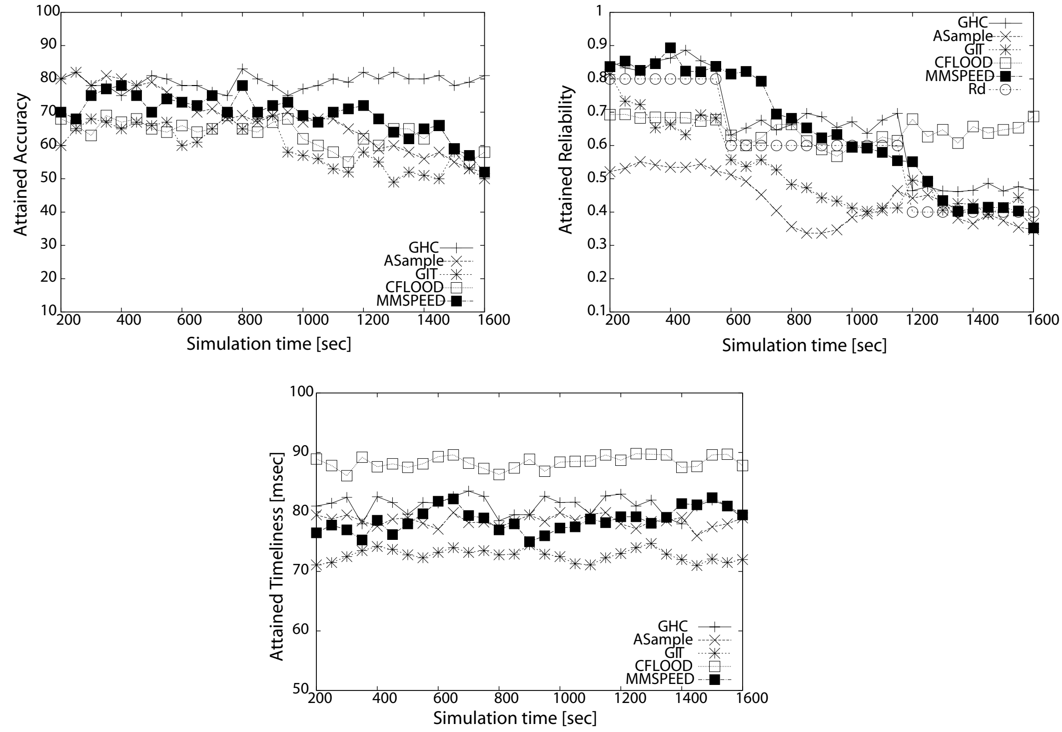


Figure 8.6: Tunability for varying desired reliability

with accuracy, *ASample* fails and performs worst considering the reliability requirements and fails to come close to the desired reliability. However, the *GIT* and *MMSPEED* performs better than the other protocols, but still fails to satisfy the user required reliability. On the other hand, *CFLOOD* declines with its performance and lacks to compete with *GHC*. Considering timeliness with requirement of $L = 60$ms, *GIT* and *MMSPEED* takes much time for successful message delivery and inclines much than the desired timeliness. *GHC* and *CFLOOD* are comparable and performs better than other compared protocols. However, *GHC* is fairly close to desired timeliness and performs better than other protocols.

Fig. 8.6 shows the tunability of reliability with varying user requirements on reliability $R = 0.8$, 0.6, 0.4. We observe that *GHC* performs fairly well and satisfies the desired reliability requirement. *GHC* also adapts and tunes accordingly to the varying reliability requirement. However, the attained reliability by *GHC* is due to the achieved desired accuracy as shown in Fig. 8.5. As we bind the accuracy and reliability, there are always a correlation and effects on each other. The results attained by *GIT* and *MMSPEED* are not satisfying the desired reliability. However, *GIT* and *MMSPEED* adapts to

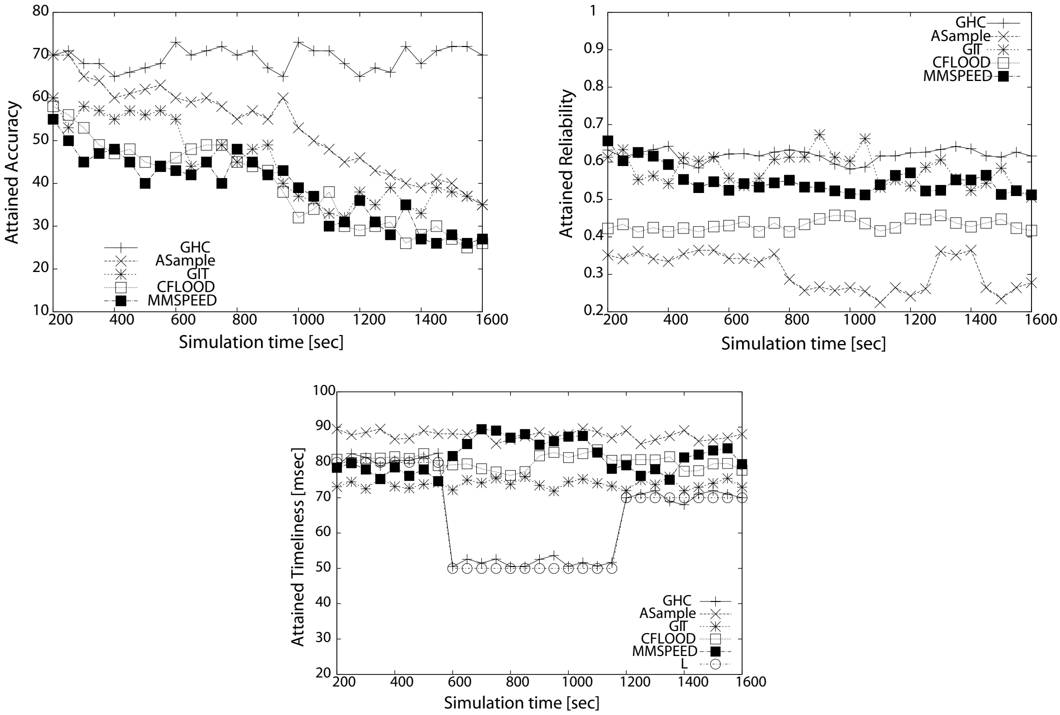


Figure 8.7: Tunability for varying desired timeliness

come closer to the desired reliability and fails to tune to the user required reliability. The result from $CFLOOD$ deviates from the user required reliability. *ASample* performs poorly in consideration with the other protocols and falls low from the desired reliability. Considering accuracy with requirement of $CA = 80$, *ASample*, GIT, $CFLOOD$ and $MMSPEED$ declines and lack to cope up with the co-design of sampling and information transport. On the other hand, GHC maintains the desired reliability due to the co-design and the combination effect of the optimal result of accuracy and reliability. Considering timeliness with requirement of $L = 80$ms, GHC and $MMSPEED$ performs fairly close with the timeliness requirement and $CFLOOD$ takes much time while delivering the required reliable information about the phenomenon. On the other hand, GIT and *ASample* shows poor performance and not comparable with desired timeliness.

Fig. 8.7 shows the tunability of timeliness with varying user requirements on timeliness $L = 80$ms, 50ms, 70ms. We observe that GHC tunes and adapts accordingly with varying desired timeliness, thanks to the co-design of sampling and information transport. On the other hand, the optimize function is combined with accuracy, reliability and timeliness, hence, any

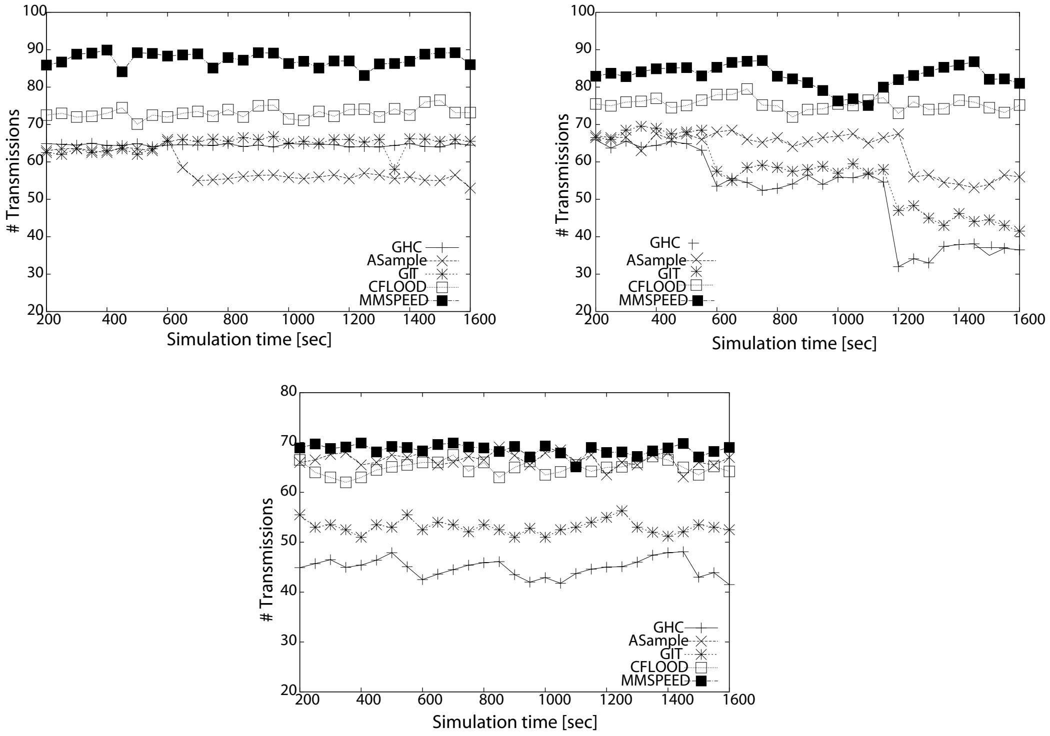


Figure 8.8: Efficiency for varying accuracy, reliability and timeliness

variation with the accuracy and reliability directly also affects the timeliness. All the compared protocols perform worse and fail to tune to the desired timeliness. However, $CFLOOD$ and $MMSPEED$ performs fairly equal and lacks to tune to the desired timeliness. GIT and *ASample* performs in very different proximity from the user required timeliness, both the protocols deviate and lacks to adapt and tune towards to user requirements. Considering accuracy with requirement of $CA = 70$, $CFLOOD$, $MMSPEED$, GIT and *ASample* decline with the accuracy and deviate from the required accuracy requirement. GHC performs fairly well and copes up with the desired accuracy. Though *ASample* has a good start with the accuracy behaviour, but lack to cope up with user requirements due to the missing co-design. Considering reliability with requirement of $R = 0.6$, GHC results are desirable to the end user requirements of the reliability. GIT and $MMSPEED$ performs fairly close enough with the desired user requirement. $CFLOOD$ lack to adapt to the desired reliability and *ASample* simply performs poorly.

Fig. 8.8 shows the performance with number of transmissions considering the varying accuracy, reliability and timeliness. We observe that GHC and GIT perform a similar number of transmissions with varying accuracy and reliability. As an observation from Fig. 8.6, we achieve a slightly better reliability than *GIT* and still perform equally well with the transmissions compared to *GIT*. Moreover, as our optimization function combines with accuracy and reliability, we emphasize to minimize the number of retransmissions with our co-design model. The transmissions attained by *CFLOOD* and *MMSPEED* are higher than *GHC* and *GIT*. On the other hand, *ASample* performs with more transmissions with reliability, however, performs fairly well with accuracy. The transmissions with the varying timeliness has a better effect on *GHC* and our protocols perform better than competing protocols. *GIT* is close enough to *GHC*, but still takes much transmissions due to the lack of co-design. *ASample*, *CFLOOD* and *MMSPEED* performs worse and the efficiency is poor compared to *GHC*.

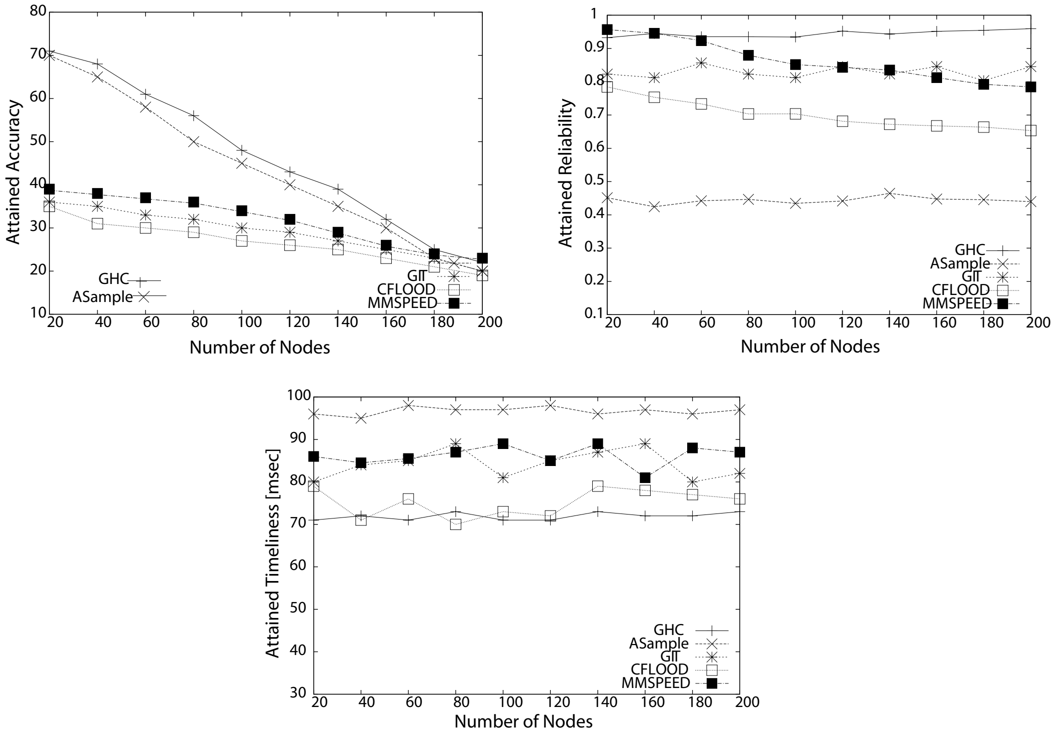


Figure 8.9: Impact of network size on attained contortion accuracy, transport reliability and transport timeliness

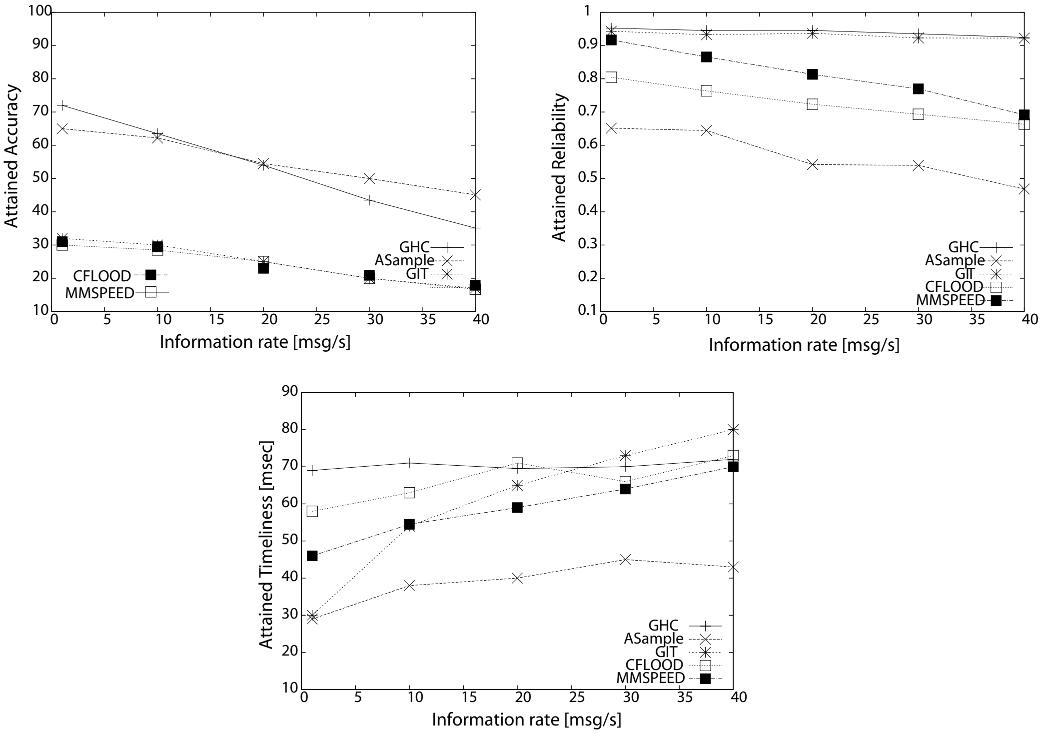


Figure 8.10: Impact of network load on attained contortion accuracy, transport reliability and transport timeliness

8.3 Case Study 2: Simulation Results For Dynamic System Requirements

In this section we provide the simulation results for the dynamic system requirements with varying network size, network load and BER. We fix the user requirements for sampling and information transport, i.e., $CA = 70$, $R = 0.9$ and $L = 70$ms.

Fig. 8.9 shows the performance for different number of nodes. We observe that as the number of nodes are increased the contortion function decreases with GHC. GHC results shows that it performs accordingly to the optimal function and contortion accuracy value decreases from 70 to 30 providing better performance. On the other hand, *ASample* also performs accordingly with GHC and show a better contortion performance. GIT, $CFLOOD$ and $MMSPEED$ does not pertain to the desired user requirement and starts poorly at the value below 40 and fails to represent any accurate physical phenomenon. From Fig. 8.9 we can also observe that GHC attain the de-

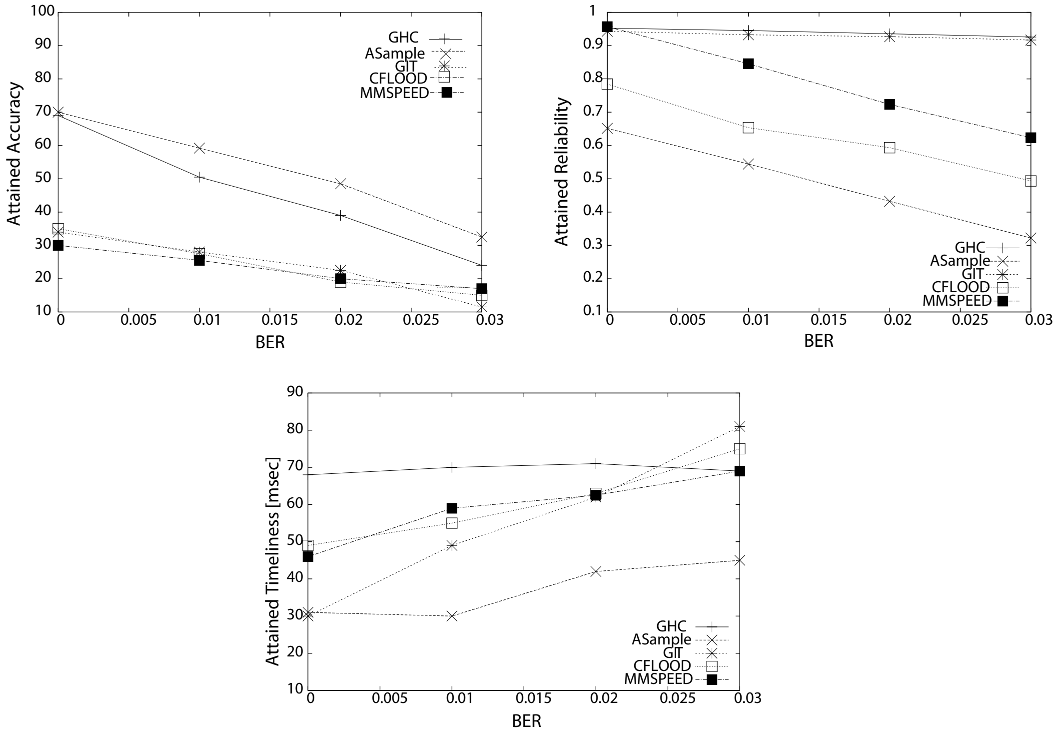


Figure 8.11: Impact of BER on attained contortion accuracy, transport reliability and transport timeliness

sired reliability. *GIT* and *MMSPEED* fails to satisfy the desired reliability with a slight difference. The reliability attained by *CFLOOD* is lower than the desired reliability. Reliability of *ASample* is independent of the desired reliability. Regarding timeliness, *GHC* meet the tolerated timeliness independent from node density and outperforms competing protocols. The competitor protocols just deviate considering the desired timeliness and fail to satisfy the user requirements.

Fig. 8.10 shows the performance for different information rates, i.e., the impact of network load. We observe that *GHC* and *ASample* attains the desired accuracy and performs according to the contortion function. *GIT*, *CFLOOD* and *MMSPEED* results are poorly performed and deviates from the desired accuracy requirement. Furthermore, considering the desired reliability *GHC* attains desired reliability with varying information rate. *GIT* performs accordingly with the desired reliability requirement. The other competitor protocols just fail to compete with *GHC* and lacks to satisfy user requirements. Regarding timeliness, the *GHC* outperforms competing protocols. Moreover, *ASample* performs very poor considering the timeliness requirement.

Fig. 8.11 shows the performance for varied BER, i.e., the impact of perturbation levels. The performance of GHC considering accuracy is satisfying with desired accuracy and the protocol behaviour is tuned according to the contortion function. GIT, $CFLOOD$ and $MMSPEED$ does not pertain to the desired user requirement and fails to represent any accurate physical phenomenon. From Fig. 8.11 we can also observe that GHC and GIT attains the desired reliability. On the other hand, the other competitor protocols perform quite poor considering the desired reliability. Furthermore, GHC outperforms competitor protocols by meeting desired timeliness. Fig. 8.11 also shows that other competitor protocols are independent from the desired timeliness and fail to cope with the co-design. We can also claim the total number of transmissions required to attain the desired transport reliability are quite efficient considering GHC. The overall performance while considering accuracy, reliability and timeliness are completely satisfied by GHC, while the competitor protocols fail with one or the other attributes and lack to co-design.

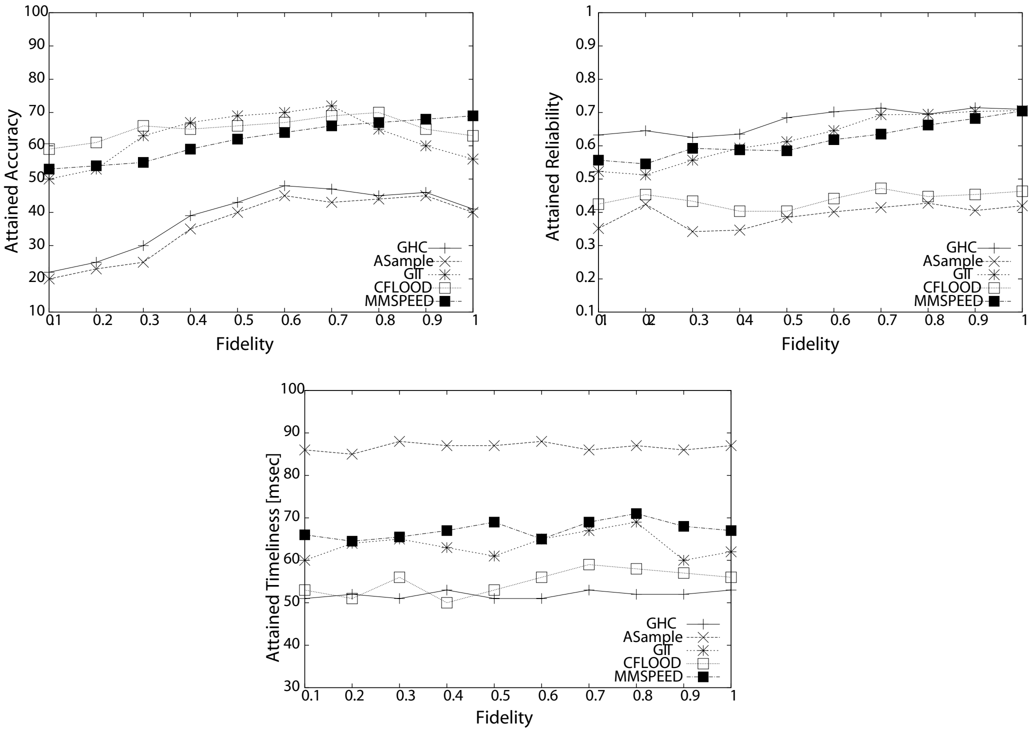


Figure 8.12: Impact of fidelity on attained contortion accuracy, transport reliability and transport timeliness

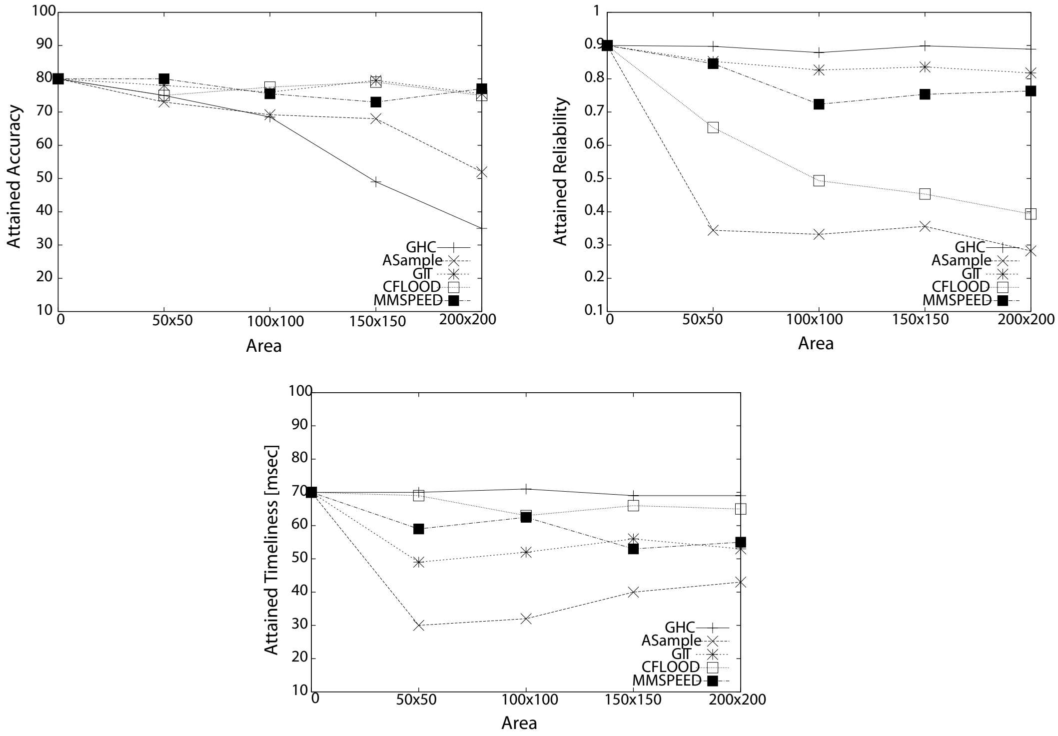


Figure 8.13: Impact of area on attained contortion accuracy, transport reliability and transport timeliness

8.4 Case Study 3: Simulation Results For Evolving Phenomenon

In this section we provide the simulation results for the evolving phenomenon with varying fidelity, area of the WSN and intensity of the phenomenon. We fix the user requirements on varying fidelity, i.e., CA = 40, R = 0.7 and L = 50ms. The requirements on varying the area of the WSN are CA = 80, R = 0.9 and L = 70ms. For the intensity of the phenomenon we observe the impact of the varying phenomenon on accuracy, reliability and timeliness.

Fig. 8.12 shows the performance for varying fidelity with attained contortion accuracy. We observe that GHC and $ASample$ fairly reach the user requirement on contortion accuracy. Moreover, as the fidelity requirement is low the protocols tend to show low contortion accuracy and then improve towards the desired contortion accuracy. Fig. 8.12 shows the performance for varying fidelity with attained reliability. We observe that GHC, GIT and $MMSPEED$ attain the desired reliability with a slight difference. More-

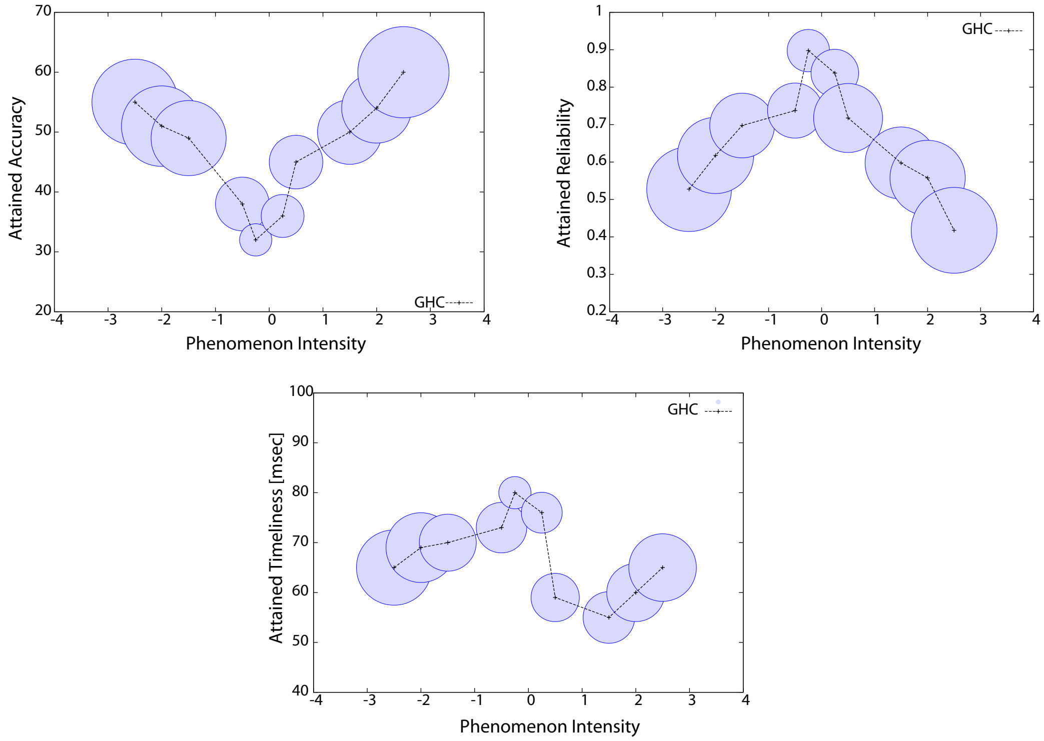


Figure 8.14: Impact of varying phenomenon on attained contortion accuracy, transport reliability and transport timeliness

over, *ASample* performs poorly in regard to other protocols for satisfying the desired reliability. Fig. 8.12 shows the performance for varying fidelity with attained timeliness. The timeliness attained by *GIT* and *MMSPEED* are higher than the desired timeliness. On the other hand, the performance of *ASample* regarding timeliness is poor and fails to satisfy the desired timeliness.

Fig. 8.13 shows the impact of the different WSN area on desired accuracy, reliability and timeliness. The performance of *GHC* considering the accuracy is meeting the desired requirements and also in bound to the contortion function. As the area being increased with number of sensor nodes, the contortion function has its direct effect on the accuracy and pertains to decrease. While other competitor protocols *GIT*, *CFLOOD* and *MMSPEED* performs fairly poorly and is out of scope information for representing the accurate physical phenomenon. While *ASample* has a fair performance with desired accuracy, but does not perform close to *GHC*. From Fig. 8.13 we can also observe that *GHC* attains desired reliability with varying area. The rest of the competitor protocols just fails and lacks to satisfy the user required reliability. Furthermore, considering timeliness, the *ASample* completely fails to cope with the desired timeliness and performs poorly. On the other hand, *GIT*, *CFLOOD* and *MMSPEED* deviates and are independent from the desired timeliness. However, *GHC* takes over the competitor protocols and performs fairly well.

Fig. 8.14 shows the performance for a varied phenomenon (hotspot). The phenomenon is more intense in the center and then spreads to other regions. We consider evaluating the varying phenomenon just with our work to see how our protocol behaves in order with the behavior of representing the physical phenomenon. As observed in Fig. 8.14, the accuracy has a desired effect with the center of the phenomenon and the contortion function increases while the phenomenon spreads. Furthermore, the reliability of our work tends to be high and effective when the phenomenon is started and then is relaxed when the phenomenon is spreading. Finally, the timeliness behavior is more liable with the intense phenomenon and is fairly decreased while moving away from the center of the phenomenon.

8.5 Experimentation

In this section we emphasize on experimentation. We show the results of our work by experimenting on real world deployment on TUDμNet [Guerrero et al., 2012] testbed.

8.5.1 TUDμNet: Experimentation Environment

We experiment in two different deployments of TUDμNet testbed at the Technische Universität Darmstadt: (1) In computer science department building (Piloty) on floors 1 and 2 with 51 TelosB sensor nodes. Sensor Nodes are placed on windows or over fluorescent tubes inside the department's offices. The physical area of the deployment is 30 x 20 x 8 with an average distance of 13.9 (m). The nodes are sparsely deployed with density size of 0.01 (n/m^3) and partitioned in a grid topology. (2) In GKmM lab (TIZ building) with 60 TelosB sensor nodes, the deployment here is abundant with density size 0.09 (n/m^3), covers an area of approximately $220m^2$ and also partitioned in a grid topology. The physical area of the deployment is 31 x 7 x 3 with an average distance of 10.4 (m). The gateway is located in one corner. The information is generated from one corner and from the middle of the network and transported towards the gateway. The Piloty testbed has a more normal distribution (higher number of intermediate length links), whereas the TIZ building testbed shows a heavy tailed distribution (higher

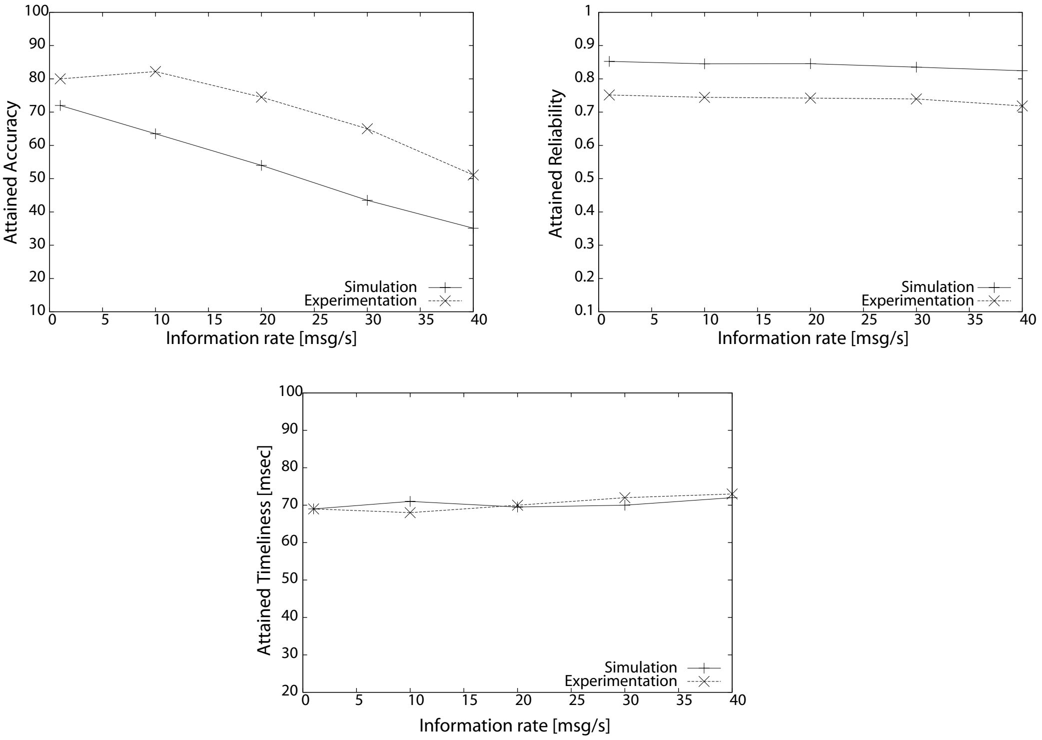


Figure 8.15: Experimentation of network load on attained contortion accuracy, transport reliability and transport timeliness

number of short links, fewer long links).[1]The further information about the testbed can be found in [Guerrero et al., 2013].

The performance of our protocol is measured in terms of accuracy and reliability, timeliness. The accuracy is measured in terms of contortion, on how close the evaluation information is to the real world value. The reliability is measured as the successful delivery of messages in behold of the end user requirements. The timeliness are measured in *ms*.

8.5.2 Experimentation Results

From the experimentation of our work we show the results of our experiments on sparse network deployment (Piloty) and abundant network deployment (TIZ building). Considering the Piloty testbed, we show the results for network load and BER. On the other hand, we show the results for user require-

---

[1] The explanation of the testbed was the current status during the time of the experimentation.

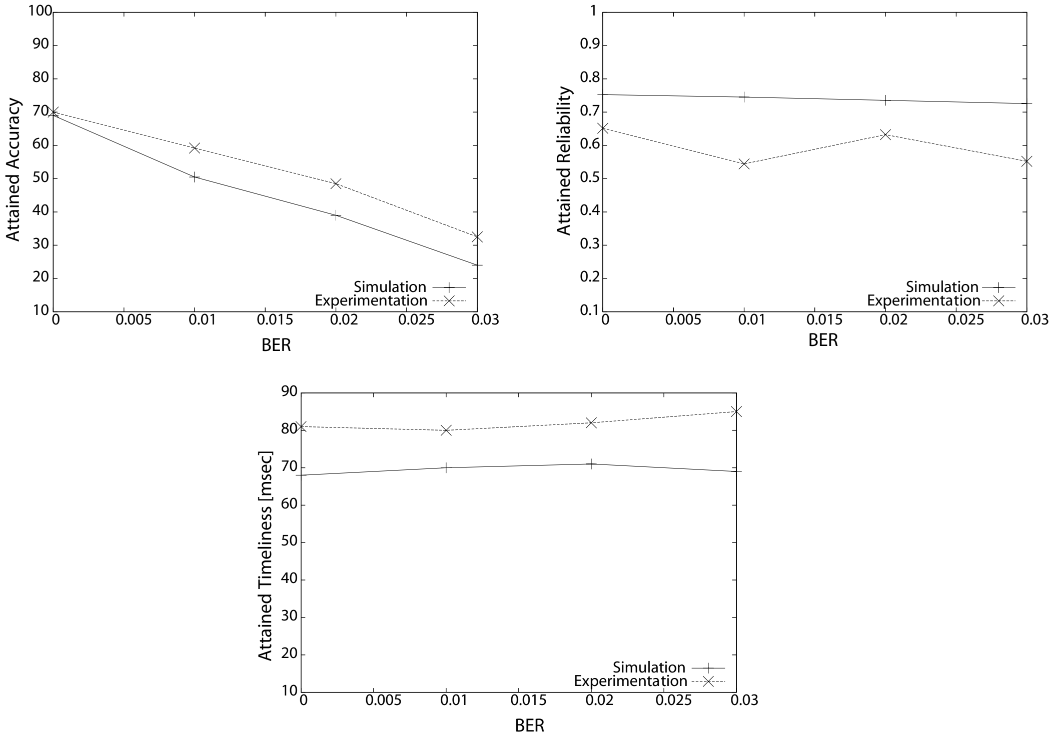


Figure 8.16: Experimentation of BER on attained contortion accuracy, transport reliability and transport timeliness

ments and number of sensor nodes on TIZ building testbed. Furthermore, we compare our experimental results with simulation results and show the gains.

Fig. 8.15 shows the performance for different information rates, i.e., the impact of network load. The impact of the network load is experimented on the sparse network of Piloty testbed. The requirements are $CA = 70$, $R = 0.8$ and $L = 70$ms. We observe that simulation results attain the desired accuracy and perform according to the contortion function. On the other hand, the experimental results perform very close to the desired requirement of the accuracy. Here, due to the sparse deployment and realistic network conditions, the experimental results show a small deviation from the user requirement. In a sparse network the accuracy reaches to 51, because our contortion function is based on the variance and the distance between the signal magnitude of the sensor nodes. Furthermore, considering the desired reliability the simulation and experimental results performs closely to the desired reliability with varying information rate. However, the reliability from the experimental results are little low and because of the sparse deployment

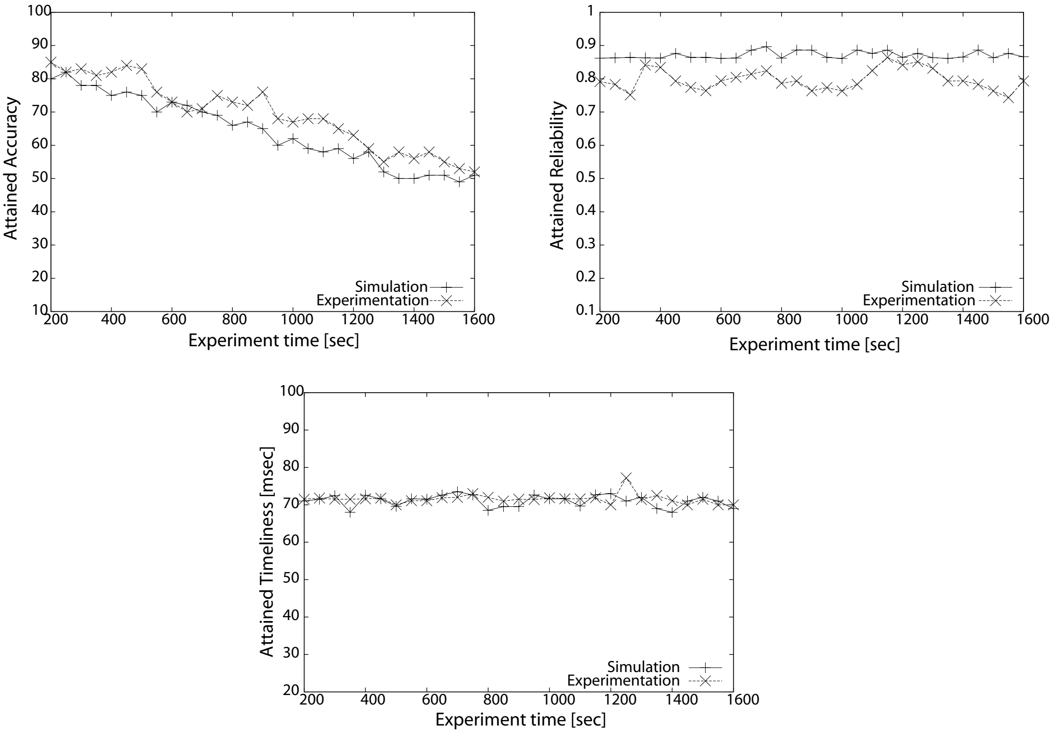


Figure 8.17: Experimentation of user requirements on attained contortion accuracy, transport reliability and transport timeliness

and as well of the link quality in realistic conditions. Regarding timeliness, the experimental results take performs quite close to the user requirements and to the simulation results.

Fig. 8.16 shows the performance for varied BER, i.e., the impact of perturbation levels. The impact of BER is experimented on the sparse network of Piloty testbed. The requirements are $CA = 70$, $R = 0.7$ and $L = 70$ms. The performance of our work on the testbed considering accuracy is satisfying with desired accuracy and the protocol behavior is tuned according to the contortion function. The simulation and experimental results are closely comparable considering the accuracy. From Fig. 8.16 we can also observe that our simulated work attains the desired reliability. On the other hand, the experimental results provide reliability close to 0.6 and still performing very close to the desired reliability. Fig. 8.16 also shows the timeliness behavior of work, the experimental results show that on realistic network condition, the timeliness perform close to the user requirements. We can also claim the total number of transmissions required to attain the desired transport reliability are quite efficient considering our work. The overall performance

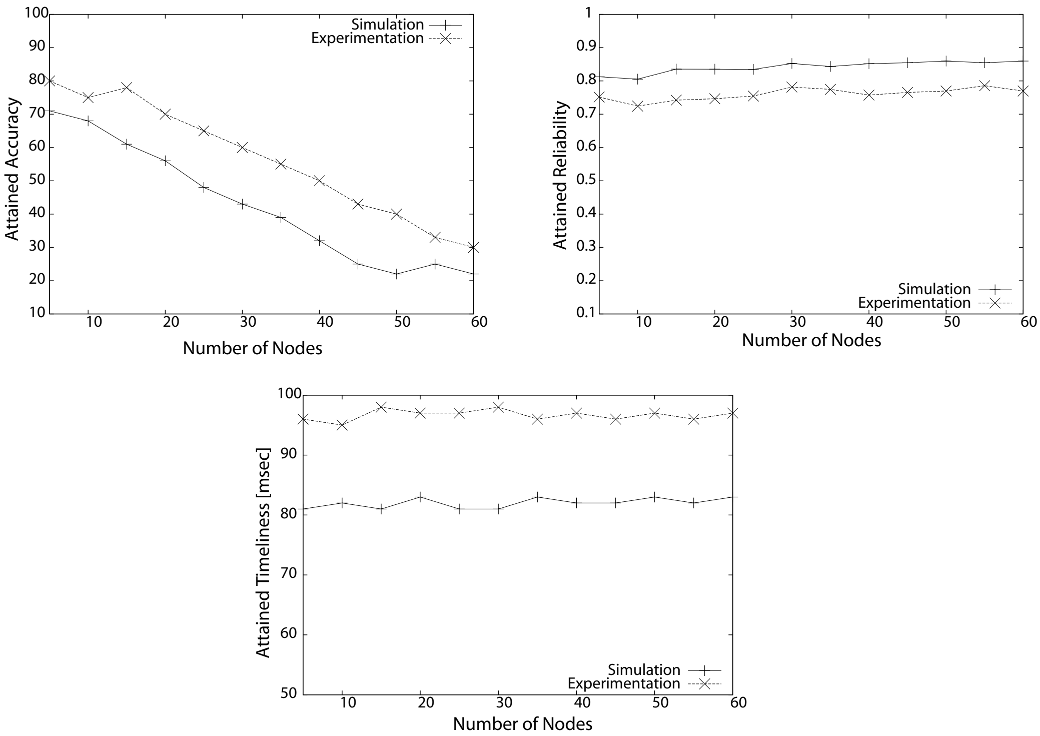


Figure 8.18: Experimentation of network size on attained contortion accuracy, transport reliability and transport timeliness

considering accuracy, reliability and timeliness are completely satisfied with experimental results and as well compared to the simulation results.

Fig. 8.17 shows the experimental results of our work for the user requirements on accuracy, reliability and timeliness. The impact of user requirements is experimented on the abundant network of TIZ building testbed. The requirements are $CA = 80$, $R = 0.8$ and $L = 70$ms. The experimental results for accuracy show that the user requirements on the accuracy are satisfied. Here, the experimental results of accuracy are close to simulation results because of the abundant testbed. Hence, the variance and the difference between the signal magnitudes results in lower contortion accuracy. Considering reliability with requirement of $R = 0.9$, simulation results performs according to the user requirement and overcomes the perturbations. Here, the reliability of the experimental results also performs closer to user requirements because of the good link quality and as sensor nodes are quite close to each other. Considering timeliness with requirement of $L = 70$ms, both simulation and experimental results are satisfying the user requirements.

Fig. 8.18 shows the experimental results for different number of nodes.

The impact of the number of nodes is experimented on the abundant network of TIZ building testbed with 60 sensor nodes. The requirements are $CA = 70$, $R = 0.8$ and $L = 80$ms. We observe that as the number of nodes are increased the contortion function decreases, i.e., with lower contortion accuracy, we achieve a better representation of the phenomena. Both the experimental and simulation results yield similar results with a slight difference. Considering the reliability, in an abundant network, the reliability is well maintained due to better link quality and the experimental results yielding the desired results as according to user requirements. Regarding timeliness, both experimental and simulation results show that our work meets the tolerated timeliness independent from node density.

Experimentation Gains

After performing the experimentation and comparing with our simulation results, some observation are explained. Some of the results of experimentation show a deviation from the simulation due to long radio links and their fashion of deployment. When the experimentation was conducted on the sparse network of Piloty building, the radio range and the link quality plays a major role and the difference in the results are due to these aspects and other operational conditions. However, due to the different techniques adapted in our solutions such as optimal retransmissions, path split mechanisms, we still gain a very close result with the simulation values.

For the experimentation on the abundant network of TIZ building, we see that the reliability is well maintained and also the message loss is lower due to the short radio links and availability of nodes and paths for the information transport. We can claim that when our solutions are running on sensor nodes in real environment, we still over come dynamic operational conditions and perform very closely to the user requirements. In addition, thanks to the tunable co-design of sampling and information transport, and the optimal interaction of accuracy, reliability and timeliness for maximizing efficiency and optimal network performance.

8.6 Chapter Summary

As the observation of the simulation and experimental results, the validity of our work performs fairly well while considering all the three attributes (accuracy, reliability and timeliness). Moreover, our work out performs other protocols with a tunable co-design, combining sampling and information transport. On the other hand, as our optimization function at its core was designed to also minimize the number of transmissions, our work performs fairly better than the competitor protocols regarding efficiency. Our work avoids over or under provision of information and adapts and tunes accordingly to user requirements. In addition, our work solely represents as a work considering accuracy, reliability and timeliness and satisfying the user requirements.

Chapter 9

Conclusions and Future Research

In this thesis, we developed tunable co-design to maximize efficiency by avoiding under or over provision of information in Wireless Sensor Networks (WSNs). This chapter concludes the thesis by summarizing our main contributions and discussing their extendability.

We believe that the work presented in this thesis opens up new interesting research directions. Therefore, this chapter discusses the key issues, and presents ideas for further enhancing the information quality aware co-design in WSNs introduced in this thesis.

9.1 Overall Thesis Contributions

The main goal of the thesis was to develop optimized sampling and information transport co-design to support multi-attribute interaction of accuracy, reliability and timeliness in WSNs. Our research effort was driven by the current need of generic solution for sampling and information transport co-design to fulfill evolving application requirements and dynamic network conditions. Accordingly, this section discusses the key contributions driven by the research questions listed in Section 1.6.

9.1.1 Tunable Information Transport

Through tunable information transport, we have achieved the composite tuning of reliability and timeliness as per the application requirements. We have introduced the tunable timeliness, which efficiently assigns the tolerable hop latencies on the path, compensates delays, and splits the path when needed. The optimized solution combines the re-transmission approach meeting the tolerable hop timeliness and the path replication approach when the tolerable hop timeliness is violated. This is the first instance of tuning when the combination of both the reliability and timeliness are implemented.

We have provided three instances of the tunable information transport with the strict timeliness notion, best effort reliability and composite reliability and timeliness. We have also presented two algorithms (i) the tunable timeliness algorithm provides with best effort reliability. This algorithm finds the optimal number of retransmissions and implements delay compensation on per hop basis. If delay compensation is not effective, a path split is conducted, i.e., the path is replicated to two parts. (ii) the reliability and timeliness algorithm that provides tuning in composition and satisfies the user evolvable requirements.

Overall, the tunable information transport with reliability and timeliness in composition, satisfying the user evolvable requirements represent the contribution **C1** of our work, as defined in the introductory chapter of this thesis (Section 1.6).

Resultant publications

- **Vinay Sachidananda**, Abdelmajid Khelil and Neeraj Suri, *Quality of Information in Wireless Sensor Networks: A Survey*, Proceedings of the fifteenth International Conference on Information Quality (ICIQ), November 2010

- **Vinay Sachidananda**, Abdelmajid Khelil and Neeraj Suri, *Information Quality Aware Transport for Wireless Sensor Networks*, Proceedings of the European Conference on Wireless Sensor Networks (EWSN), February 2012.

- **Vinay Sachidananda**, Abdelmajid Khelil, Dhananjay Umap, Matthias Majuntke, Neeraj Suri, *Trading Transport Timeliness and Reliability for Efficiency in Wireless Sensor Networks*, Proceedings of the tenth IEEE International Conference on Networking Sensing and Control (ICNSC), April 2013.

9.1.2 Optimizing Sampling and Information Transport

As proceeding towards the information quality aware co-design we have achieved important steps towards the co-design of sampling and information transport as per the application requirements. We have developed an analytical model for the case that no differences between sensor readings have to be regarded. This simplifies the problem of finding a specific subset of nodes to the problem of merely finding the optimal number of nodes that have to send samples. Our analytical model gives the optimal number of sensor nodes, so that the specific application requirements are satisfied.

The optimized solution provided depending on the application requirements, reduces the total number of retransmissions by adding redundancy and sending more samples than required. This is the first instance of real time adaptation when an integrated sampling and transport solution is implemented.

The development of the sampling and transport co-design and highlighting the *STC* algorithm considering the spatial accuracy and transport reliability represent the contribution **C2** of our work, as defined in the introductory chapter of this thesis (Section 1.6).

Resultant publications

- **Vinay Sachidananda**, Abdelmajid Khelil, David Noack and Neeraj Suri, *Sampling and Transport Co-design in Wireless Sensor Networks*, Proceedings of the tenth IEEE International Conference on Wireless On demand Network Systems and Services (WONS), March 2013.

- **Vinay Sachidananda**, Abdelmajid Khelil, David Noack and Neeraj Suri, *Information Quality Aware Co-design of Sampling and Transport* *in Wireless Sensor Networks*, Proceedings of the sixth IEEE International Conference on Wireless and Mobile Networking Conference (WMNC), April 2013.

9.1.3 Representing the Physical Phenomena with Optimized Network Performance

As the core thesis goal, we have achieved the generic information quality co-design and multi-attribute optimization. We have shown the mutual dependency of spatial accuracy, transport reliability and timeliness constraints and have shown the binding of all the three attributes through formulating it as the optimal co-design. We have provided the mathematical model for representing the physical phenomenon as a user requirement. We have also provided the notions for all the three attributes and how to move towards the optimal combination. Through analytical results we have observed that all the three attributes are interrelated and have an impact on each other. On the other hand, with our *GHC* Algorithm, we represent as a first algorithm which considers all the three attributes and also satisfies the user requirements.

The presented optimal co-design and the multi-attribute co-relation of the attributes accuracy, reliability and timeliness represent the contribution **C3** of our work, as defined in the introductory chapter of this thesis (Section 1.6).

Resultant publications

- **Vinay Sachidananda**, David Noack, Abdelmajid Khelil and Neeraj Suri, *On Co-modeling the Sampling and Transport in Wireless Sensor Networks*, Proceedings of the eleventh GI/ITG KuVS Fachgespräch "Sensornetze" (FGSN), September 2012.
- **Vinay Sachidananda**, David Noack, Abdelmajid Khelil and Neeraj Suri, *Optimized Co-design of Spatial Sampling and Information Transport in Wireless Sensor Networks*, Proceedings of the Special issue in Telecommunication Systems Journal (TSMJ), (under review), 2013.

9.2 Lessons Learned

The work presented in this thesis has developed a basis for information quality aware co-design in WSNs. The basic fact from our information assessment to consider more than one functional block for the co-design has been vital. Considering the sampling and information transport blocks and its attributes accuracy, reliability and timeliness has proved to be important in the co-design. Moreover, it proves to be a necessity for any information quality aware design to consider the user requirements and also the respective attributes and metrics. On the other hand, it is trivial to make sure to have the user view and design view for the optimal deployment of the WSNs and to achieve the optimal results.

Second, as it has been quite complex to achieve the goal of information quality aware co-design, it is necessary to first tackle the information transport block. It is also trivial to know that the tunability of the information transport makes the co-design more user oriented. The reliability and timeliness attributes are always interdependent and whenever one or the other attribute is modified, there is always an effect on the other. Moreover, the tunable information transport lays out the first path to consider the sampling block and also it proves that reliability is a hidden requirement of accuracy.

Third, to ensure the co-design, exploiting the spatial correlation is important. The complex problem of sampling and information transport co-design provides the first basic step towards the optimal combination of the both functional blocks. On the other hand, considering accuracy as a user requirement and combining it with reliability proves the first cross-operation of both the functional blocks.

Finally, to represent the physical phenomenon accurately by selecting the representative nodes is necessary for optimal network performance. The optimal co-design of sampling and information transport needs a very complex technique and indeed needs a realistic modeling of each of the attribute (accuracy, reliability and timeliness). For the optimal deployment of WSNs, the key goal of our thesis, which was achieved through the multi-attribute co-relation and the optimized function of sampling and information transport, is trivial. On the other hand, the optimal co-design also maximizes the efficiency by saving resources such as energy. Finally, the core part of the thesis has been achieved through these lessons which we have learned.

9.3 Open Ends - Basis for Future Work

While the work presented in this thesis addressed the research questions driving it towards making the discussed contributions, it also opened new and interesting research perspectives along its way. In the following, we briefly present some of the most promising ones.

In-network Processing Co-design

As we have shown in our information assessment, it would be interesting to consider the in-network processing functional block for the co-design. The challenge lies in how could the information be processed within the network and also will it be possible to maximize the efficiency by saving resources. On the other hand, mapping the relevant attributes to the in-network processing block and combining it with the provided co-design along with the accuracy, reliability and timeliness would be a bigger challenge. Moreover, working on the optimal co-design of sampling, information transport and in-network processing to minimize the loss of information, to satisfy evolvable user requirements and maximize the efficiency is very interesting and challenging.

Information Quality Attributes and Metrics

Though information attributes are relatively well discussed, information metrics definition and their efficient computation are still in their infancy. Accordingly, the future research directions may progress on the aspects of defining attributes and metrics and the techniques to efficiently compute them on the fly in all information extraction stages. However, as one need to narrow research into fewer attributes, one will take some must considered attributes during the flow of information from the source to the sink. One can define and defend how it is relevant and necessary to use these attributes and violation of this lead to information which does not satisfy the user requirements. Metrics and their run-time quantification represent a powerful tool to assess the dependability of WSN, which allows for efficient and tunable information quality provisioning.

Tuning and Adapting MAC-protocols

Tuning and adapting the existing MAC-protocols for the generic co-design. As the MAC-protocols are very important in the co-design, it would be a challenge in itself to tune and adapt the MAC-protocols according to the evolvable user requirements. On the other hand, it would be also interesting to try various different MAC-protocols in combination with our provided optimal co-design.

Minimize Information Loss

As we have always tried to maximize the efficiency, it would be also interesting to minimize the loss of information. It would be a challenge to consider the information loss as the cost function in the optimal co-design. Trying to adapt with different evovlable user requirements of accuracy, reliability and timeliness and satisfying them while minimizing the loss of information would play a vital and challenging role. As an example, Akaike's information criterion (Akaike's information criterion, is a measure of the goodness of fit of an estimated statistical model grounded on the concept of entropy) can be used to measure information quality when certain information is lost from the source to the sink.

Heterogeneous WSNs and Applications

As a preliminary effort we have started to explore the possibilities of using our co-design for mobility assisted WSNs. Alongside mobility, heterogeneous sensing as well heterogeneous mobile nodes can be a part of WSNs. It is worth investigating to include different modules in our co-design which keeps the heterogeneity of the environment and devices intact while providing application specific accuracy, reliability and timeliness.

Bibliography

A. M. Abbas and O. Kure. Quality of service in mobile ad hoc networks: a survey. *International Journal of Ad Hoc and Ubiquitous Computing*, 6(2): 75–98, 2010.

I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci. Wireless sensor networks: a survey. *Computer Networks*, 38(4):393–422, 2002a.

I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci. Wireless sensor networks: a survey. *Computer networks*, 38(4):393–422, 2002b.

J. N. Al-Karaki and A. E. Kamal. Routing techniques in wireless sensor networks: a survey. *Wireless Communications, IEEE*, 11(6):6–28, 2004.

M. Anwar Hossain, P. K. Atrey, and A. El Saddik. Context-aware qoi computation in multi-sensor systems. In *Mobile Ad Hoc and Sensor Systems, 2008. MASS 2008. 5th IEEE International Conference on*, pages 736–741. IEEE, 2008.

S. Arnborg, J. Brynielsson, H. Artman, and K. Wallenius. Information awareness in command and control: Precision, quality, utility. In *Information Fusion, 2000. FUSION 2000. Proceedings of the Third International Conference on*, volume 2, pages THB1–25. IEEE, 2000.

A. Arora, P. Dutta, S. Bapat, and V. K. et al. A line in the sand: a wireless sensor network for target detection, classification, and tracking. *Computer Networks*, 46(5):605–634, 2004.

L. Atzori, A. Iera, and G. Morabito. The internet of things: A survey. *Computer networks*, 54(15):2787–2805, 2010.

D. Ballou, R. Wang, H. Pazer, and G. K. Tayi. Modeling information manufacturing systems to determine information product quality. *Management Science*, 44(4):462–484, 1998.

J. Barbancho, C. León, F. J. Molina, and A. Barbancho. A new qos routing algorithm based on self-organizing maps for wireless sensor networks. *Telecommunication Systems*, 36(1-3):73–83, 2007.

C. Bisdikian. On sensor sampling and quality of information: A starting point. In *Pervasive Computing and Communications Workshops, 2007. PerCom Workshops' 07. Fifth Annual IEEE International Conference on*, pages 279–284. IEEE, 2007.

C. Bisdikian, J. Branch, K. K. Leung, and R. I. Young. A letter soup for the quality of information in sensor networks. In *Pervasive Computing and Communications, 2009. PerCom 2009. IEEE International Conference on*, pages 1–6. IEEE, 2009a.

C. Bisdikian, L. M. Kaplan, M. B. Srivastava, D. J. Thornley, D. Verma, and R. I. Young. Building principles for a quality of information specification for sensor information. In *Information Fusion, 2009. FUSION'09. 12th International Conference on*, pages 1370–1377. IEEE, 2009b.

Q. Chang, X. Wang, and X.-L. Wang. Local information embedding by dimension extension and information risk minimization. In *Machine Learning and Cybernetics, 2005. Proceedings of 2005 International Conference on*, volume 7, pages 4335–4339. IEEE, 2005.

Z. Charbiwala, S. Zahedi, Y. Kim, Y. Cho, and M. Srivastava. Toward quality of information aware rate control for sensor networks. In *Fourth International Workshop on Feedback Control Implemenation and Design in Computing Systems and Networks*, 2009.

D. Chen and P. K. Varshney. Qos support in wireless sensor networks: A survey. In *International Conference on Wireless Networks*, volume 233, 2004.

C.-Y. Chong and S. P. Kumar. Sensor networks: evolution, opportunities, and challenges. *Proceedings of the IEEE*, 91(8):1247–1256, 2003.

F. C. Delicato, P. F. Pires, L. Pinnez, L. Fernando, and L. da Costa. A flexible web service based architecture for wireless sensor networks. In *Distributed Computing Systems Workshops, 2003. Proceedings. 23rd International Conference on*, pages 730–735. IEEE, 2003.

J. Dutka. The incomplete beta functionâa historical profile. *Archive for history of exact sciences*, 24(1):11–29, 1981.

S. A. Ehikioya. A characterization of information quality using fuzzy logic. In *Fuzzy Information Processing Society, 1999. NAFIPS. 18th International Conference of the North American*, pages 635–639. IEEE, 1999.

E. Felemban, C.-G. Lee, and E. Ekici. Mmspeed: multipath multi-speed protocol for qos guarantee of reliability and. timeliness in wireless sensor networks. *Mobile Computing, IEEE Transactions on*, 5(6):738–754, 2006.

E. Gelenbe and L. Hey. Quality of information: An empirical approach. In *Mobile Ad Hoc and Sensor Systems, 2008. MASS 2008. 5th IEEE International Conference on*, pages 730–735. IEEE, 2008.

A. Ghosh and S. K. Das. Coverage and connectivity issues in wireless sensor networks: A survey. *Pervasive and Mobile Computing*, 4(3):303–334, 2008.

O. Gnawali, R. Fonseca, K. Jamieson, D. Moss, and P. Levis. Collection tree protocol. In *Proceedings of the 7th ACM Conference on Embedded Networked Sensor Systems*, pages 1–14. ACM, 2009.

P. E. Guerrero, A. P. Buchmann, A. Khelil, and K. Van Laerhoven. Tudμnet, a metropolitan-scale federation of wireless sensor network testbeds. *sensors*, 20:Z1s, 2012.

P. E. Guerrero, I. Gurov, S. Santini, and A. Buchmann. On the selection of testbeds for the evaluation of sensor network protocols and applications. In *Signal Processing Advances in Wireless Communications (SPAWC), 2013 IEEE 14th Workshop on*, pages 495–499. IEEE, 2013.

P. Gupta and P. R. Kumar. The capacity of wireless networks. *Information Theory, IEEE Transactions on*, 46(2):388–404, 2000.

D. Hakkarinen and Q. Han. Data quality driven sensor reporting. In *Mobile Ad Hoc and Sensor Systems, 2008. MASS 2008. 5th IEEE International Conference on*, pages 772–777. IEEE, 2008.

Q. Han and N. Venkatasubramanian. Timeliness-accuracy balanced collection of dynamic context data. *Parallel and Distributed Systems, IEEE Transactions on*, 18(2):158–171, 2007.

T. He and M. Zafer. Adaptive sampling for transient signal detection in the presence of missing samples. In *Mobile Ad Hoc and Sensor Systems, 2008. MASS 2008. 5th IEEE International Conference on*, pages 760–765. IEEE, 2008.

L. Hey and E. Gelenbe. Adaptive packet prioritisation for large wireless sensor networks. *Telecommunication Systems*, 48(1-2):125–150, 2011.

X. Huang and Y. Fang. Multiconstrained qos multipath routing in wireless sensor networks. *Wireless Networks*, 14(4):465–478, 2008.

U. Hunkeler and P. Scotton. A quality-of-information-aware framework for data models in wireless sensor networks. In *Mobile Ad Hoc and Sensor Systems, 2008. MASS 2008. 5th IEEE International Conference on*, pages 742–747. IEEE, 2008.

B. Jiang, B. Ravindran, and H. Cho. Cflood: A constrained flooding protocol for real-time data delivery in wireless sensor networks. In *Stabilization, Safety, and Security of Distributed Systems*, pages 413–427. Springer, 2009.

R. Kannan, S. Sarangi, S. S. Iyengar, and L. Ray. Sensor-centric quality of routing in sensor networks. In *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies*, volume 1, pages 692–701. IEEE, 2003.

K. Karenos and V. Kalogeraki. Real-time traffic management in sensor networks. In *Real-Time Systems Symposium, 2006. RTSS'06. 27th IEEE International*, pages 422–434. IEEE, 2006.

K. Karenos, V. Kalogeraki, and S. V. Krishnamurthy. A rate control framework for supporting multiple classes of traffic in sensor networks. In *Real-Time Systems Symposium, 2005. RTSS 2005. 26th IEEE International*, pages 11–pp. IEEE, 2005.

R. T. Kessel. The dynamics of information fusion: Synthesis versus misassociation. In *Information Fusion, 2006 9th International Conference on*, pages 1–6. IEEE, 2006.

A. Khelil and N. Suri. Gossiping: Adaptive and reliable broadcasting in manets. In *Dependable Computing*, pages 123–141. Springer, 2007.

A. Khelil, F. K. Shaikh, A. Ali, and N. Suri. gmap: efficient construction of global maps for mobility-assisted wireless sensor networks. In *Wireless On-Demand Network Systems and Services, 2009. WONS 2009. Sixth International Conference on*, pages 189–196. IEEE, 2009.

A. Khelil, C. Reinl, B. Ayari, F. K. Shaikh, P. Szczytowski, A. Ali, and N. Suri. Wireless sensor cooperation for a sustainable quality of information1. *Pervasive Computing and Networking*, pages 71–100, 2011.

S. Kim and S. O. Yang. Wireless sensor gathering data during long time. In *Ubiquitous Information Technologies and Applications (CUTE), 2010 Proceedings of the 5th International Conference on*, pages 1–5. IEEE, 2010.

S. Kim, R. Fonseca, and D. Culler. Reliable transfer on wireless sensor networks. In *IEEE Sensor and Ad Hoc Communications and Networks (SECON)*, pages 449– 459, 2004.

L. Krishnamachari, D. Estrin, and S. Wicker. The impact of data aggregation in wireless sensor networks. In *Distributed Computing Systems Workshops, 2002. Proceedings. 22nd International Conference on*, pages 575–578. IEEE, 2002.

M. Kuorilehto, M. Hä, T. D. Hä, et al. A survey of application distribution in wireless sensor networks. *EURASIP Journal on Wireless Communications and Networking*, 1900(5):774–788, 2005.

S. Kwon, J. Shin, J. Ko, and C. Kim. Distributed and localized construction of routing structure for sensor data gathering. *Telecommunication Systems*, 44(1-2):135–147, 2010.

H. Li, S. Jiang, and G. Wei. Information-accuracy-aware jointly sensing nodes selection in wireless sensor networks. In *Mobile Ad-hoc and Sensor Networks*, pages 736–747. Springer, 2006.

H. Lin, M. Lu, N. Milosavljevic, J. Gao, and L. J. Guibas. Composable information gradients in wireless sensor networks. In *Proceedings of the 7th international conference on Information processing in sensor networks*, pages 121–132. IEEE Computer Society, 2008.

C. Lu, B. M. Blum, T. F. Abdelzaher, J. A. Stankovic, and T. He. Rap: A real-time communication architecture for large-scale wireless sensor networks. In *Real-Time and Embedded Technology and Applications Symposium, 2002. Proceedings. Eighth IEEE*, pages 55–66. IEEE, 2002.

J.-F. Martínez, A.-B. Garcí, I. Corredor, L. López, V. Hernández, and A. Dasilva. Qos in wireless sensor networks: survey and approach. In *Proceedings of the 2007 Euro American conference on Telematics and information systems*, page 20. ACM, 2007.

A. Meliou, D. Chu, J. Hellerstein, C. Guestrin, and W. Hong. Data gathering tours in sensor networks. In *Proceedings of the 5th international conference on Information processing in sensor networks*, pages 43–50. ACM, 2006.

D. L. Moody, G. Sindre, T. Brasethvik, and A. Sølvberg. Evaluating the quality of information models: empirical testing of a conceptual model quality framework. In *Proceedings of the 25th International Conference on Software Engineering*, pages 295–305. IEEE Computer Society, 2003.

K. Morita, A. Aikebaier, T. Enokido, and M. Takizawa. A data transmission protocol for reliable and energy-efficient data transmission in a wireless sensor-actuator network. *Telecommunication Systems*, 38(3-4):71–82, 2008.

M. F. Munir, A. A. Kherani, and F. Filali. On stability and sampling schemes for wireless sensor networks. In *Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks and Workshops, 2007. WiOpt 2007. 5th International Symposium on*, pages 1–10. IEEE, 2007.

R. A. Nichols. Qoi in dtn-based directional networks. In *Sarnoff Symposium, 2009. SARNOFF'09. IEEE*, pages 1–5. IEEE, 2009.

A. M. Ortiz, F. Royo, T. Olivares, J. C. Castillo, L. Orozco-Barbosa, and P. J. Marron. Fuzzy-logic based routing for dense wireless sensor networks. *Telecommunication Systems*, pages 1–11, 2011.

P. Park, C. Fischione, A. Bonivento, K. H. Johansson, and A. Sangiovanni-Vincent. Breath: an adaptive protocol for industrial control applications using wireless sensor networks. *Mobile Computing, IEEE Transactions on*, 10(6):821–838, 2011.

R. Prasanth, J. Cabrera, J. Amin, R. Mehra, R. Purtell, and R. Smith. Quality of information measures for autonomous decision-making. In *American Control Conference, 2004. Proceedings of the 2004*, volume 2, pages 1002–1007. IEEE, 2004.

V. Sachidananda, A. Khelil, and N. Suri. Quality of information in wireless sensor networks: A survey. *ICIQ*, 2010.

V. Sachidananda, A. Khelil, D. Umap, M. Majuntke, and N. Suri. Trading transport timeliness and reliability for efficiency in wireless sensor networks. In *Networking, Sensing and Control (ICNSC), 2013 10th IEEE International Conference on*, pages 720–727. IEEE, 2013.

A. Sahoo and P. Baronia. An energy efficient mac in wireless sensor networks to provide delay guarantee. In *Local & Metropolitan Area Networks, 2007. LANMAN 2007. 15th IEEE Workshop on*, pages 25–30. IEEE, 2007.

Y. Sankarasubramaniam, Ö. B. Akan, and I. F. Akyildiz. Esrt: event-to-sink reliable transport in wireless sensor networks. In *Proceedings of the 4th ACM international symposium on Mobile ad hoc networking & computing*, pages 177–188. ACM, 2003.

F. K. Shaikh, A. Khelil, and N. Suri. On modeling the reliability of data transport in wireless sensor networks. In *Parallel, Distributed and Network-Based Processing, 2007. PDP'07. 15th EUROMICRO International Conference on*, pages 395–402. IEEE, 2007.

F. K. Shaikh, A. Khelil, B. Ayari, P. Szczytowski, and N. Suri. Generic information transport for wireless sensor networks. In *Sensor Networks, Ubiquitous, and Trustworthy Computing (SUTC), 2010 IEEE International Conference on*, pages 27–34. IEEE, 2010.

V. Shnayder, M. Hempstead, B.-r. Chen, G. W. Allen, and M. Welsh. Simulating the power consumption of large-scale sensor network applications. In *Proceedings of the 2nd international conference on Embedded networked sensor systems*, pages 188–200. ACM, 2004.

J. Srinivasan. The role of trustworthiness in information service usage: The case of parry information kiosks, tamil nadu, india. In *Information and Communication Technologies and Development, 2007. ICTD 2007. International Conference on*, pages 1–8. IEEE, 2007.

M. Strasser, A. F. Meier, K. Langendoen, and P. Blum. Dwarf: Delay-aware robust forwarding for energy-constrained wireless sensor networks. In *Proceedings of the 3rd IEEE International Conference on Distributed Computing in Sensor Systems (DCOSS 2007)*, pages 64–81, June 2007.

P. Szczytowski, A. Khelil, and N. Suri. Asample: Adaptive spatial sampling in wireless sensor networks. In *Sensor Networks, Ubiquitous, and Trustworthy Computing (SUTC), 2010 IEEE International Conference on*, pages 35–42. IEEE, 2010.

C. H. J. Tan and D. F. Gillies. Generating reliable quality of information (qoi) metrics for target tracking. In *SPIE Defense, Security, and Sensing*, pages 733607–733607. International Society for Optics and Photonics, 2009.

H.-X. Tan, M.-C. Chan, W. Xiao, P.-Y. Kong, and C.-K. Tham. Information quality aware routing in event-driven sensor networks. In *INFOCOM, 2010 Proceedings IEEE*, pages 1–9. IEEE, 2010.

W. L. Tan, O. Yue, and W. C. Lau. Performance evaluation of differentiated services mechanisms over wireless sensor networks. In *Vehicular Technology Conference, 2006. VTC-2006 Fall. 2006 IEEE 64th*, pages 1–5. IEEE, 2006.

A. Tolstikov, C.-K. Tham, and J. Biswas. Quality of information assurance using phenomena-aware resource management in sensor networks. In *Networks, 2006. ICON'06. 14th IEEE International Conference on*, volume 1, pages 1–7. IEEE, 2006.

A. Tolstikov, W. Xiao, J. Biswas, S. Zhang, and C.-K. Tham. Information quality management in sensor networks based on the dynamic bayesian network model. In *Intelligent Sensors, Sensor Networks and Information, 2007. ISSNIP 2007. 3rd International Conference on*, pages 751–756. IEEE, 2007.

M. C. Vuran and I. F. Akyildiz. Spatial correlation-based collaborative medium access control in wireless sensor networks. *Networking, IEEE/ACM Transactions on*, 14(2):316–329, 2006.

M. Wälchli and T. Braun. Efficient signal processing and anomaly detection in wireless sensor networks. In *Applications of Evolutionary Computing*, pages 81–86. Springer, 2009.

C. J. Walter and N. Suri. The customizable fault/error model for dependable distributed systems. *Journal of Theoretical Computer Science*, 290(2): 1223–1251, 2003.

J. Wang, Y. Liu, M. Li, W. Dong, and Y. He. Qof: Towards comprehensive path quality measurement in wireless sensor networks. In *INFOCOM, 2011 Proceedings IEEE*, pages 775–783. IEEE, 2011.

R. Y. Wang, D. M. Strong, and L. M. Guarascio. Beyond accuracy: What data quality means to data consumers. *J. of Management Information Systems*, 12(4):5–33, 1996.

S. Wolfram. *The MATHEMATICA® Book, Version 4*. Cambridge university press, 1999.

N. Xu. A survey of sensor network applications. *IEEE Communications Magazine*, 40(8):102–114, 2002.

H. Yan, X.-c. Diao, and K.-q. Li. Research on information quality driven data cleaning framework. In *Future Information Technology and Management Engineering, 2008. FITME'08. International Seminar on*, pages 537–539. IEEE, 2008.

S. Zahedi and C. Bisdikian. A framework for qoi-inspired analysis for sensor network deployment planning. In *Proceedings of the 3rd international conference on Wireless internet*, page 28. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2007.

S. Zahedi, M. B. Srivastava, and C. Bisdikian. A computational framework for quality of information analysis for detection-oriented sensor networks. In *Military Communications Conference, 2008. MILCOM 2008. IEEE*, pages 1–7. IEEE, 2008.

H. Zhang, A. Arora, Y.-r. Choi, and M. G. Gouda. Reliable bursty convergecast in wireless sensor networks. *Computer Communications*, 30(13): 2560–2576, 2007.

J. Zhao, R. Govindan, and D. Estrin. Computing aggregates for monitoring wireless sensor networks. In *Sensor Network Protocols and Applications, 2003. Proceedings of the First IEEE. 2003 IEEE International Workshop on*, pages 139–148. IEEE, 2003.

Index

Curriculum Vitae

Personal Data

Name: Vinay Mysore Sachidananda

Date of birth: May 31st, 1984

Place of birth: Karnataka, India

School Education

1997-2000 Matriculation from St. Joseph's& Secondary Education, Mysore, India. (secured 95%)

2000-2002 Pre-university from Board of Intermediate, Mysore, India. (secured 89%)

University Education

2002-2006 *Bachelor of Engineering (Computer Systems)* – Visweswaria & Technology University, Karnataka, India. (secured Higer distiction)

2006-2008 *Maste of Science (Networking and Mobility)* – University of Trento, Trento, Italy. (secured Higer distiction)

2008-2014 *Ph.D. in Computer Science* – Technische Universität Darmstadt, Darmstadt, Germany.