UNIVERSITY OF TECHNOLOGY SYDNEY

Generic Benchmarking for Application Specific Wireless Sensor Networks Multi Criteria Performance

A thesis submitted for the degree of Master by Research

By

Dong YU

2012

CERTIFICATE OF AUTHORSHIP/ORIGINALITY

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of requirements for a degree except as fully acknowledged within the text.

I also certify that the thesis has been written by me. Any help that I have received in my research work and the preparation of the thesis itself has been acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

Signature of Student

Production Note:
Signature removed prior to publication.

Acknowledgements

I would like to express my sincere appreciation to Dr. Priyadarsi Nanda and Professor Xiangjian He, my supervisors, for their encouragement and guidance. I would like to express my thanks to my colleagues and friends in the Centre for Innovation in IT Services and Applications (iNEXT), for the inspiration I got from their innovative ideas presented in workshops and presentations.

ABSTRACT

Due to stringent energy constraint and demand for performance requirement, a generic architecture like TCP/IP or Internet is not feasible with sensors used across various applications. Instead, application specific design methodology is the de facto consensus accepted among Wireless Sensor Network (WSN) community. While it wins WSN performance gains for individual applications, the methodology sacrifices all plausible attributes a generic architecture can contribute. Without a unified reference model as comparing foundation, the profound problem in true protocols contribution evaluation and comparison remains challenging. Moreover, the stochastic and statistical nature of WSNs makes realistic performance analysis fairly complex. In multi criteria QoS context, this problem is further magnified by big design space with not yet fully understood parameters and the competing relationship between multi objective performance metrics. This work introduces a generic wireless-benchmarking methodology not only qualitatively evaluation from high level abstraction, concerning only profound pros and cons from a general viewpoint of tradeoffs between generality, performance and cost, but also a set of practical workflows that are designed to support quantitative evaluation and analysis of WSN protocols for application-specific objectives. This methodology and the accompanying new benchmark concepts, such as performance efficiency, development efficiency and performance stability, are designed to gain new insight of the dynamic behavior of WSN protocols in a systematical way

compared to the current ad-hoc evaluation approaches applied by most of the community.

LIST OF FIGURES

1. Transitional Region with Probability Link. 20
2. Uncertainty Modeling in WSN 21
3. A Typical PDF Bell Curve Graph 23
4. A Typical Cumulative Distribution Function (CDF) Graph 24
5. Competing Performance Metrics under Wide Design Space. 27
6. Single Performance Index (SPI) Composition. 29
7. Triangular Constraints. 34
8. Conflict between Three Main Metrics. 37
9. Architecture Efficiency in Two Aspects. 38
10. SDLC and Architecture Efficiency. 40
11. Workflow of Proposed Benchmarking Solution. 46

Contents

Acknowledgements iv

Abstract v

List of figures vii

1 Introduction 1

2 Literature Review. 7

3 Fundamentals Issues in WSNs Evaluation and Design 11

3.1 Application specific design as a challenge in system evaluation . ..11

3.2 Uncertainty attribute of WSN performance 15

3.3 WSNs performance dynamic modeling at system level. 19

3.4 Characterizing WSNs performance with statistical concepts. 22

3.5 Multi QoS metrics and energy constraint conflicts in WSNs. 26

4 A Generic WSNs Evaluation Framework 32

4.1 System architecture evaluation fundamental concepts 33

4.2 Triangle constraints in general system architecture evaluation . .. .33

4.3 Triangular constraints reduce to bilateral constraints tradeoff . . . 39

5 A Practical Benchmarking Solution for QoS Performance Index 45

5.1 A practical workflow of benchmarking solution 45

5.2 Metrics development and interdependency analysis 48

6 Parameter Reduction and Interdependency Analysis. 52

6.1 Interaction effect of multiple variables 52

6.2 The limitation of existing work. 54

6.3 General problem formulation 55

6.4 P-value with linear regression model 55

6.5 The Choquet model: a nonlinear model 60

6.6 Comparison of the two models 67

7 Conclusion and Future Prospect. 69

References

CHAPTER 1

INTRODUCTION

Wireless Sensor Networks (WSNs) comprise of set of micro sensor nodes which have capabilities of sensing specific phenomenon, storing and processing raw data, delivering data from source nodes to sink nodes using wireless channels. Comparing with traditional wired network, WSNs have three distinguished characteristics which define its uniqueness in networked world, namely, the miniature size of the node, stochastic natured wireless channel and different application requirements.

Firstly a node is a reduce-formed computer system with extra sensing task but limited energy. The reduced volume of sensor node is vital in its design consideration: computing resources including processing and storage are stringently restricted to what current practical MEMS (Micro Electro-Mechanical Systems) technology can offer. Small size of battery in the sensor node is the only source of energy supply and it is impractical to replace like other mobile electrical devices due to economic reason and rough environment. From technology evolving perspective, the challenge here is that Moore’s Law does not apply to battery capacity, though the density of transistors on a chip has consistently doubled every 18 months, the energy density of batteries only seems to double every 10 years [1]. The need to conserve energy leads WSNs research community to focus on what is possible to get things done in an efficient way. Application specific design methodology is the natural result of such endeavor in which performance is achieved only within the scope of specific application by vertical integration of best matching protocols. The best performance for the specific application scenario is achievable due to the fact that design space is small constrained only by the assumptions of the targeted application.

The property of wireless radio channel is notably different to wired link in traditional network. The distinction from wired link (twisted-pair, coaxial cable, optic fiber) is that the radio channel is unpredictable (random) and could vary over very short time and space. The randomness is inherent in the form of fading - temporal and spatial random variation in signal propagation between networked nodes. This leads to new challenges in signal transmission and credible performance prediction and evaluation. How to characterize and model the WSNs stochastic link performance is critical in WSNs evaluation. New metrics should be introduced to reflect the individual inherent uncertainty and collective nonrandom regularity [39].

The third unique aspect about WSNs is that sensor networks have a wide variety of applications and systems with vastly varying requirements and characteristics. The sensor networks can be used in military environment, disaster management, habitat monitoring, medical and health care, industrial fields, home automation networks, detecting chemical, biological, radiological, nuclear, and explosive material etc. As WSNs are evolving into a more mature research field, with more and more protocols being developed and publicly released, the use of WSNs are equally fast expanding into new domains with stringent performance requirement. Optimization efforts have been focused on trying to find out the optimal configuration for specific application domain without possibility of generalization applied to wider application area. Generality is ideal for commercialization but not suitable for vastly different application scenarios given the aforementioned resource constraints.

The uniqueness of WSNs in its resources limitation, transient channel state and drastic application requirement, brings in application specific system design methodology. From WSNs research kickoff at early stage, research efforts are mostly focused on the isolated programming issue of single layer protocols with little regard for other layer functionality. This leads to protocols that exist in a vacuum and perform well on a theoretical basis, but have problems when deployed under real-life circumstances [40, 41]. Many of these protocols are further validated using ad-hoc created experimental tests, specifically aimed at the strengths of a specific protocol, leaving little room for objective comparison with other protocols. Up to now, there exists no fixed set of accepted testing methods, scenarios, parameters or metrics to be applied on a protocol under test. This lack of standardization significantly increases the difficulty for a developer to assess the relative performance of their protocols compared to the current state of the art.

System designers usually adopt application specific design methodology to compromise between multiple competing Quality of Service (QoS) requirements and energy constrains. Application specific design methodology has advantage over strict layered architecture like TCP/IP or Internet under energy constraint, as they can focus on one specific application performance while ignoring generality cost of encapsulation and abstraction. But loose or vague layered design methodology also has serious challenge in system design as well as system evaluation. As we all know that layered architecture provide standard interface between neighboring layers, designers can focus on the problem on one specific targeted layer, the architecture itself will take care of all the compatibility and combination issues. In another words, there is no compatibility and protocol combination issue once the work is done in a strict layered context. Application specific design methodology unfortunately cannot afford such luxury. In terms of system evaluation, the problem is even exaggerated due to the fact that each designer has their specific vision of functionality distribution among components of referenced architecture, such as Physical Layer, MAC layer, Network Layer. When researchers design new routing protocols, they normally describe their ideal protocol contexts in the form of certain assumptions, that makes comparison of alternative solutions difficult without a common context like a strict layered structure. Moreover, the dynamic wireless medium involving spatial elements, temporal elements and load pattern uncertainty, makes wireless sensor network performance stochastic and statistical in nature. To characterize the performance, new metrics need to be identified and defined to capture the uncertain nature. Another critical challenge is to determine the unknown design space. How to decide the most significant parameters which contribute to final system performance? Such question is vital in efficiently generating performance curve for the evaluated system. WSNs research have come to a critical point where a generic benchmarking methodology is required to compensate the side effect of application specific design, to provide authentic protocol evaluation and comparison between competing solutions, to further stimulate widespread use of WSNs application.

If we consider all of the specific scenarios and different metrics in the architecture evaluation, it will be difficult to compare with each other. We cannot see the big picture as we immerse ourselves totally into complicated details. Nevertheless, a system can be considered at various abstraction levels, allowing to evaluate only relevant properties and behaviors at different abstraction levels, whose relevancy will depend on the context.

In this work, we will firstly introduce an evaluation method from high level abstraction, concerning only profound pros and cons from viewpoints of tradeoffs between generality, performance and cost. We will also introduce a benchmarking work flow to make it possible to quantitatively comparison of alternative solutions.

This thesis is organized in six chapters. In Chapter 2, existing work done in the area of interest is presented. Chapter 3 establishes a background understanding of key issues that serves as the foundation for building our proposed solutions. The chapter presents characteristics of application specific design methodology, uncertainty attribute of WSN performance, performance dynamic modeling at signal level and system level, performance evaluation with statistical concepts, multi QoS metrics and energy constraint conflict in WSNs. We emphasize that while application specific design is necessary to build efficient application based on limited energy budget, but for a fair comparison of alternative design solutions, a generic evaluation framework is needed to deal with all of the components aforementioned. Our generic evaluation framework with practical benchmarking solution is presented in Chapter 4. In this chapter, we define the purpose of our evaluation framework is to provide high level abstract information regarding architecture inherent performance pros and cons based on three main competing metrics, namely QoS performance, cost and generality potential. We show how the evaluation model guides the performance evaluation in a balanced way of combining development efficiency and performance efficiency in the context of System Development Life Cycle (SDLC). Chapter 5 provides an example practical benchmarking solution for QoS performance evaluation. Workflow and algorithm are given to get a single QoS performance index. We have pay special attention to metrics development and interdependency analysis due to big design space and unknown parameter effects on performance metrics. Chapter 6 further discusses of parameters reduction methodologies for practical purpose. Finally, we summarize our work in Chapter 7 and outline how our evaluation methodology can evolve to reflect the dynamics of future technology advance.

CHAPTER 2

LITERATURE REVIEW

During the last decade, wireless experimentation has received significant attention. Various test beds have been setup, toolkits and frameworks have been developed and data repositories have been deployed to efficiently deal with the complexities of wireless experimentation and facilitate analysis. Emulab [35], ORBIT (Open-Access Research test bed for Next-Generation Wireless Networks) [33], MiNT (a miniaturized mobile multi-hop wireless network test bed [37], Ad-hoc Protocol Evaluation (APE) test bed [36], EXC toolkit [34], etc are available to study WSN based research projects.. Each of these platforms is tailored to meet the requirements of a specific area of focus while few of them initially aim at generic benchmarking for wider wireless research community. Even though currently there are multiple European research projects targeting at benchmarking of wireless networks, such as CREW [26], BonFIRE [27], or OneLab(2) [28], few resulting benchmarking frameworks for wireless network are found in literature. In this Chapter, we focus primarily on experimentation methodologies that have been designed to make evaluation of contribution of protocols and alternative solutions comparison trust worthy.

Among the work that can be found, TinyBench [29] focuses solely on the internal metrics of a single sensor node. The work of Kim [30] benchmark execution is observed only at single layer. The works cited above both ignored layer incompatibilities problems between different sets of protocols. Thus they cannot give overall performance of a network based on the protocols under test. Authors in [38] convinced that individual network protocols may perform very differently in combination with different protocols in different layers. Their experiments have included two very different MAC protocols: SCP-MAC [52], a synchronized MAC protocol designed for extremely low duty cycles, and X-MAC [53], an optimized random access protocol, and three routing protocols: CTP [54], a collection tree routing protocol, and two point-to-point routing protocols, TYMO [55] and TinyLUNAR [56].

Recent works on wireless benchmarking [23-25, 32, 38] have shown good accounts of their respective workflow. A workflow consists of a sequence of connected steps. Emphasis is on the flow paradigm, description of the tasks, procedural steps, organizations involved, required input and output information. Rehman et al in [23] further recognize the challenges in benchmarking of wireless network, especially "deciding what metrics to calculate and what parameters have direct or indirect impact on the low-level or high-level metrics is challenging". Here, Rehman has defined adjustable internal element affecting final performance as parameters and aspects of performance measurement as metrics. In [24], authors define a set of scenarios, metrics and parameters; also provide a more critical insight into the selection of performance metrics for benchmarking. They emphasize that selection of critical metrics and effective design space reduction is the necessary first step for producing useful benchmarks. However, these contributions are not completed by methodologies to resolve the problem of how to achieve the targeted goal. Moreover ad hoc workflow and metrics selection cannot be hold for general purpose fair comparison. We have to establish a set of systematic metrics as common ground to reflect the true value of proposed solution. Meanwhile, the context in which metrics are applied to cannot be ignored. System evaluation based on evaluation framework is desirable to systematically establish key metrics. Work in [60, 61] try to do so in a qualitative way. Nevertheless, their efforts only concern the metrics at performance level, while the higher level metrics from architecture perspective are missing. For system designers, higher level abstraction metrics, such as development efficiency, generality, resource cost, are crucial in decision making. In addition, although qualitative evaluation is necessary part of system evaluation, it is not enough to support management decisions.

Another class of work is quantitative evaluation, such as in [62], introduces a performance score model based on criteria weight and measurement. In general, this kind of work is strong in how to synthesis the final performance score, but fall short in systematic metric development like qualitative evaluation does.

In contrast, our proposed solution includes not only an evaluation framework to develop key performance metrics at both architecture and QoS performance levels, but also a practical benchmarking workflow of achieving synthetic single performance index. Moreover we have explored the feasible methodologies for effective parameter design space reduction.

Our work has four folds of contribution: We adopt the concept of protocol performance stability as one of the performance metrics to reflect the statistical nature of wireless network performance; secondly we propose an effective solution to reduce the design and evaluation space, decide most significant set of parameters for system evaluation by interdependency analysis; thirdly we propose a methodology for application specific quantitative evaluation purpose, in which the users specify and weigh the performance metrics for their very specific multi criteria evaluation; finally we have develop a simple triangular constraints qualitative evaluation framework, concerning only three genetic important elements: performance, cost and generality. It provides a universal effective tool for WSNs designers to juggle between performance efficiency and development efficiency from System Development Life Cycle (SDLC) perspective. To the best of our knowledge, this is the first work in wireless benchmarking which has aforementioned realistic concepts along with algorithms implementable in real life practice.

CHAPTER 3

FUNDAMENTALS ISSUES IN WSNS EVALUATION AND DESIGN

In this chapter, we will introduce fundamental issues we envisioned in realistic performance evaluation and fair comparison of different design solutions in WSNs context. Application specific design methodology and ad hoc evaluation practice widely adopted in WSNs research community make competing solution difficult to compare with each other. Transient wireless channel among other factors contributes to a wide span of dynamic performance measurement. Competing multi objectives constraint between QoS metrics and energy limitation further complicates the system performance evaluation. Accordingly, we suggest some ways to mitigate the difficulties of characterizing WSNs performance. Deviation is introduced as a performance metric which its relative importance is specified by application specific requirements according to application scenarios. We introduce a single performance index algorithm to incorporate all the measurement mean, deviation and user specified weight, as a common comparable score among alternative solutions.

3.1 Application specific design as a challenge in system evaluation

Due to scarce of resources and energy stringent constraint, generic architecture similar to TCP/IP is unaffordable in WSNs. WSNs based research have adopted application-specific ad hoc design methodology in practice. Application specific design focuses on performance of each individual application while sacrificing generality. Still referring to layered structure, but the distribution of the functionality in each layer depends on the preference of individual designers on specific application circumstances. Standard interfaces between layers like TCP/IP are no longer available. While this methodology has proved its capability to provide light weight performance gain for individual application, various problems have emerged as critical challenge for industry and academia in WSNs field.

1.) Generality lost: Generality are realized through abstractions to cover a variety of different protocols under the same layer, abstractions will cost computing resources and hinder the performance. We have to choose performance when trading between performance and generality in WSNs context.
2.) Research community so far provided a wealth of innovative solutions, each solution has its own assumptions about the other layers, not referring to a standard benchmarking architecture. So integration and reusability are challenging.
3.) Moreover the protocols are evaluated in ad hoc manner by respective authors, their evaluation processes are not reproducible in most cases, true contribution and alternative solution comparison are partially based on metrics selection favorable to their own solution. As a final product, factors influencing the performance are huge in design space.

On the other hand, innovative algorithm design in academia normally targets at least a wider application domain, and is not limited to one specific application scenario. Researchers have to focus on one specific problem at hand while referring to a de facto architecture as consensus to cooperate with each other. The fact is that there exists no such standard architecture with clearly defined criteria and assumptions for its structured interaction, similar to standard interface defined in TCP/IP. Without such standard definition of architecture for which the algorithm can fit in, algorithm designers individually have to define such criteria and assumptions as context before algorithm design. Without satisfaction of these assumptions, the protocol's performance is unknown. Within these assumptions, the performance is tested by ad hoc simulation or small scale experimentations using test bed. Its final credibility can only be tested by real life implementation in the hands of application engineers.

In [9] Wang et al. realized how assumptions and objectives were important in WSN design for energy conservation technique due to lack of a strict standard architecture. The surveyed energy conserving mechanisms have vast span of assumptions, and they make different assumptions regarding the sensors and the network. They also have different objectives that are determined by their applications. Hence a fair comparison among the surveyed mechanisms should take into consideration all these factors. We summarize the assumptions and objectives as follows:

1. Assumptions:

1) network structure, 2) sensor deployment strategy, 3) detection model, 4) sensing area, 5) transmission range, 6) time synchronization, 7) failure model, 8) location information, and 9) distance information.

2. Objectives:

1) maximum lifetime, 2) sensor coverage, 3) network connectivity, 4) data delivery ratio, 5) quality of surveillance, 6) stealthiness, 7) balanced energy consumption, 8) scalability, 9) reliability, and 10) timeliness

Researcher must understand the effect of assumptions on the performance objective metric as well as interaction effect on other coexisting metrics. Yu et al [10] point out that omitting the important factors involved will cause the overall System Under Test (SUT) behaves differently in real implementation as in simulation experiment. Thus all metrics that have influence on performance and energy consumption in wireless sensor networks should be evaluated.

We understand that application specific design methodology is indispensable in WSNs design and integration before an era of battery and chip technology have a big leap breakthrough. But for system evaluation, a generic benchmarking methodology is urgently required to provide true contribution and fair comparison between alternative solutions. Such generic evaluation solution maybe complicated but, necessary for industry to setup as a common platform and investigate the feasibility of new proposals integration into product. Well crafted evaluation results are provided to suggest best and worst performance scenarios for implementation engineer to consider. What are the pros and cons of the proposed solution? Under dynamic and wide range of metrics and parameters, what is the best performance scenario of the proposed solution and what is the worst case scenario for implementation engineer to avoid? What are the other layer assumptions that the solution requires to fulfill its functionality? Without a benchmarking architecture to contain all the diverse application specific assumptions, and to test diverse performance under wide range of metrics and parameters, we cannot answer any of the questions asked by implementation engineer.

A generic benchmarking methodology is the first step to diminish the gap between industry and academia, and promote widespread adaptation of novel ideas emerged in research circle.

3.2 Uncertainty attribute of WSN performance

Several factors contribute to the fact that wireless sensor networks often do not work as expected when deployed in a real-world setting. Firstly, there is possibility of wrong expectation from system designer side: analytical model does not fit into the problem in hands. That is often a problem for inexperienced designers. For all simulation or other experiments methods, first step is to eliminate the possibility of this kind of profound design problem in preliminary stage. Secondly, there is possible wrong expectation from simulation results: Simulation modeling can not faithfully reflect the System Under Test (SUT), simulation credibility in WSNs has been studied by many researchers, Yu et al in [10] has recognized the source of simulation credibility problem involves levels of abstraction (approximation) in WSNs:

1). Higher level of abstraction is more concept proof natured; more effort has to put on development of abstraction for relevancy. It is hard to reduce the design space, to decide the effective parameters and interdependency between them. An extreme high level of abstraction is analytical mathematical formula with only parameters of concerned . All the other details are ignored as default or explicit assumptions. Even though some of these assumptions have a spectrum of variations and have effect on the final output performance, there are no explicit functions describing the relationship. Like topology, load pattern, node density, network size, environmental setup, noise level etc, just enumerate a few. No research has been done to prove that the effects of these elements are secondary or negligible to the main functions in SUT.

2). Lower level abstraction needs more efforts on detailed modeling; the truth of the individual components. The lowest level of abstract: emulation is a clone of the real system. It needs advanced skills on application specific programming. It is prone to programming error or specific component fidelity problem. Even though design space considerably diminishes comparing to higher level abstraction, the unspecified function between environmental elements and final output problem still remain as an issue for future research.

Except the designer's preliminary problem of analytical model mismatching design target, we can further identify fault point of performance disagreement between expectation and real world implementation into components of WSNs hierarchy.

1). Environmental influences which may lead to non-deterministic behavior of radio transmission.

2). Node level problem: Malfunction of the sensors, or even the complete failure of a sensor node due to cheap manufacturing cost. Scarce resources and missing protection mechanisms on the sensor nodes may lead to program errors: operating system reliability and fault-tolerance.

3). Network level problem: Network protocols especially (MAC and Routing) are not robust to link failure, contention, topology dynamics.

4).Unforeseen traffic load pattern: A common cause for network problems is an unforeseen high traffic load. Such a traffic burst may occur for example, when a sensor network observes a physical phenomenon, and all nodes in the vicinity will try to report at once, causing the occurrence of packet collisions combined with a high packet loss.

All these factors contribute to the uncertainty of the sensor network behavior and function. These elements increase the probability of network functionality deviation from its normal operation and affects its' collected data accuracy.

In order to effectively develop parameters, we will congregate hierarchical possible points of deviation into four groups.

1). Spatial elements uncertainty: include site specific characteristic: fading, signal attenuation, interferences, and network scale: topology, network size and density.

2) Temporal elements uncertainty: even on one particular spot, link state flips with time. 3) Data communication pattern uncertainty: include load burst pattern uncertainty (the volume, frequency of the data burst), communication interval difference (how often is the data communication happening, how long is the interval between two adjacent communications) , and different communication modes (inquiry triggered, regular report, or event triggered communication).

4). Algorithm internal programming uncertainty: include malfunctioned models, assumption realization problem and other normal cooperation problem in programming,.

There is also a great deal of uncertainty involved with evaluation of aforementioned random performance. The uncertainty comes from three mappings in the evaluation synthesis process from architecture level.

1). The uncertainty of understanding how architectural decisions map into quality attributes response. That is to say, there is uncertainty in knowing exactly, how well it will perform or adapt to changes. Especially in WSNs, there is vast spectrum of the spatial and temporal characteristics of environment which the architecture will operate in and the algorithm will take as input space. Sensitivity analysis is impractical in big design space to evaluate the mapping between the input decisions and the resulting performance which is statistical and nondeterministic in nature.

2). The uncertainty of understanding how architectural decisions map to cost. Cost modeling is not precise, at best only gives a range of value. A small sample setup of experiments can not estimate the true cost in real implementation.

3). The uncertainty of understanding how quality response map into final goals (performance gains). Even one who knows perfectly how architecture responses to its stimuli and distributes the stimuli inside the architecture, it is still quantitatively unknown that how much performance gains contributing to the final goals of the system.

3.3 Performance dynamic modeling at system level

Modeling communications in wireless networks is a challenging task since it asks for a simple mathematical object on which efficient algorithms can be designed, but that must also reflect complex physical constraints inherent to wireless networks. As an idealized mathematical object, Unit Disk Graph (UDG) is a popular model that enabled the development of efficient algorithms for crucial networking problems. The UDG model assumes perfect reception within circled range and no interference from outside the range. A lot of study has based their algorithms on this model. However, recent experimental studies [2-5] have shown that Packet Reception Rate (PRR), as an index of environmental influence on radio signal, demonstrates temporal and spatial characteristic with statistical variation. Illustrated Figure 1 in [5] demonstrates transitional region with unreliable links with probability characteristic, and therefore the idealized UDG model can be very misleading.

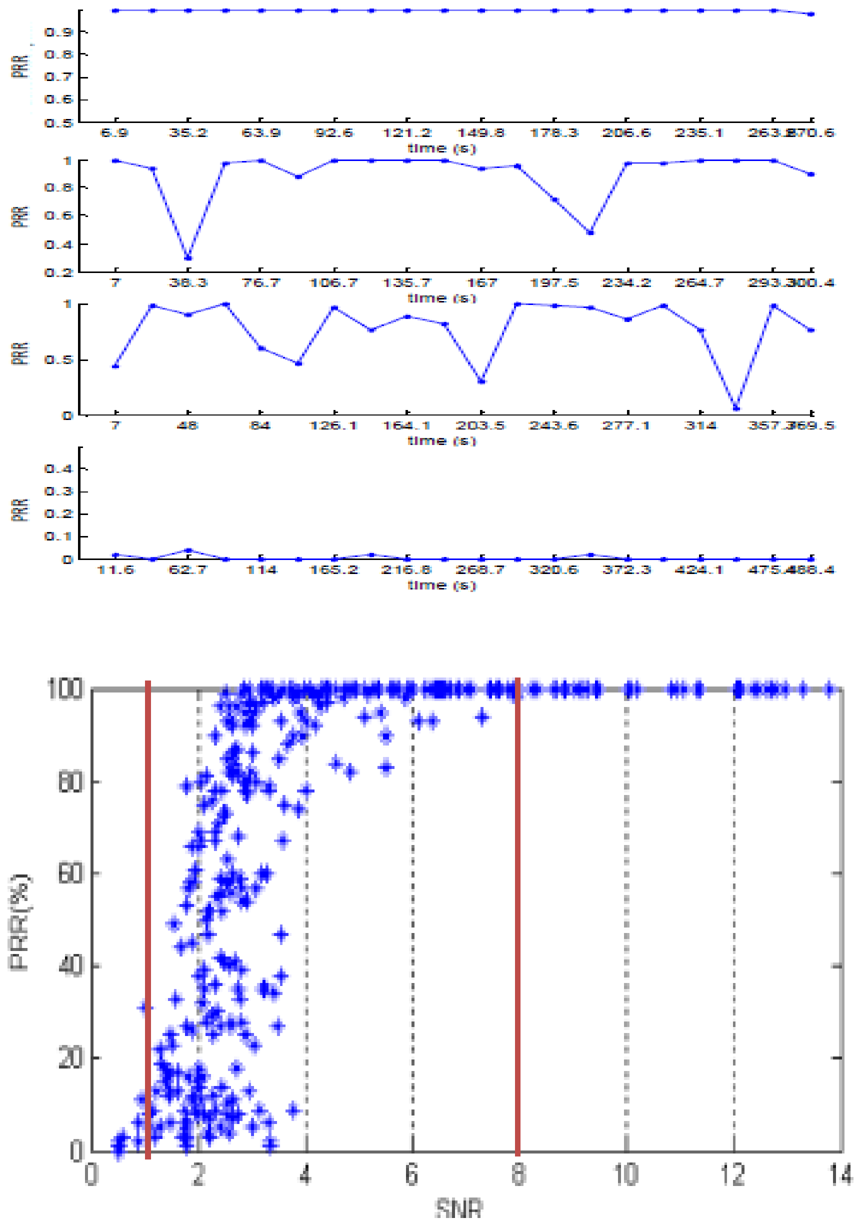

Figure 1, Transitional region with probability link [5]

We summarize above observation into two equations:

$$PRR = F\ (\text{Packet Size, SIR, Distance, Time}) \tag{1}$$

$$P = F\text{ (spatial, temporal, traffic load pattern, SUT)} \qquad (2)$$

Equation (1) represents the signal level performance dynamics with only environmental influence on antenna receiving end. In equation (2), "P" represents system level performance dynamics involving all four factors as shown in Figure 2. As the input elements display statistical behaviors, output performance definitely will have a statistical distribution pattern with certain norm and deviations for specific scenarios. Since wireless performance is inherently statistical in nature, accurate performance testing must account for these random components [6]. More over, comparing performance curves produced by a number of metrics makes it difficult to evaluate how well a given protocol suits for the purpose of an application. It may also be difficult to estimate, which of the protocols at hand would perform the best with respect to that application [7].

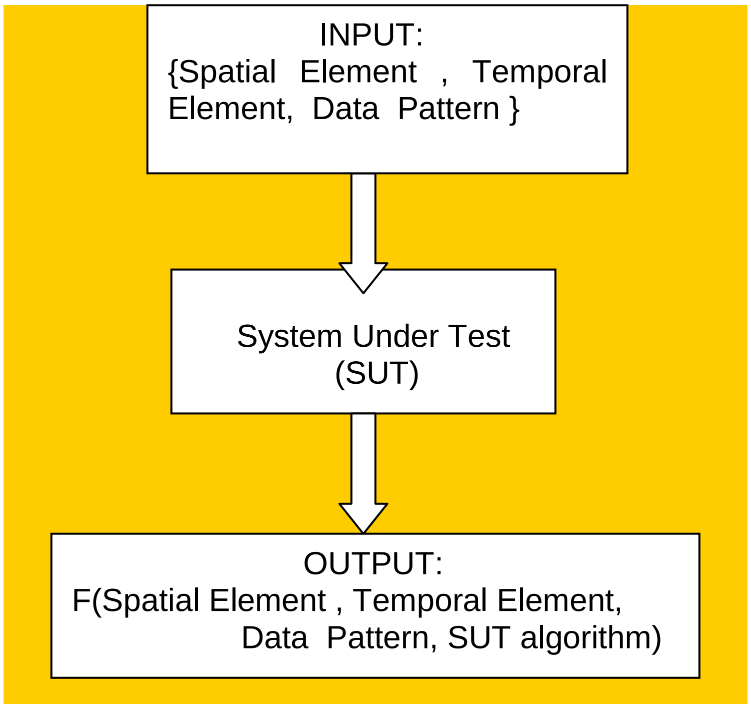

Figure 2. Uncertainty modeling in WSN

3.4 Characterizing WSNs Performance with Statistical Concepts.

WSNs sense the targeted phenomenon, collect data and make decision to store, aggregate or send the data according to distributed local algorithm. The modulated electromagnetic waves propagate in free-space; interact with the environment through physical phenomenon such as reflection, refraction, fast fading, slow fading, attenuation and human activities. Even with the best wireless protocols, the best chipsets, the best RF design, the best software, wireless performance is going to vary. Wireless performance is inherently statistical in nature, and accurate performance evaluation must reflect this nature.

We observed that currently most ad hoc evaluations in wireless network field, especially in WSNs research, no matter in the form of test bed experiment or simulation, only focus on mean value of performance metrics, and do not pay much attention on performance deviation. For applications like inhabitant monitoring, average performance is sufficient for data gathering and collective data analysis. However, average ‘throughput’, ‘lifetime’, ’reliability’ or ‘delay response’ are not sufficient enough to predict performance on certain application scenarios. Any dip in performance, no matter how short, can result in dropped packets that cause visual artifacts or pixilation of the image of wireless video monitoring application. In extreme cases like in health monitoring application, any dropped packet may cause life or death difference.

Consequently, the viewer/user experience is completely dependent on the wireless system's 'worst-case' performance like the 99th percentile.

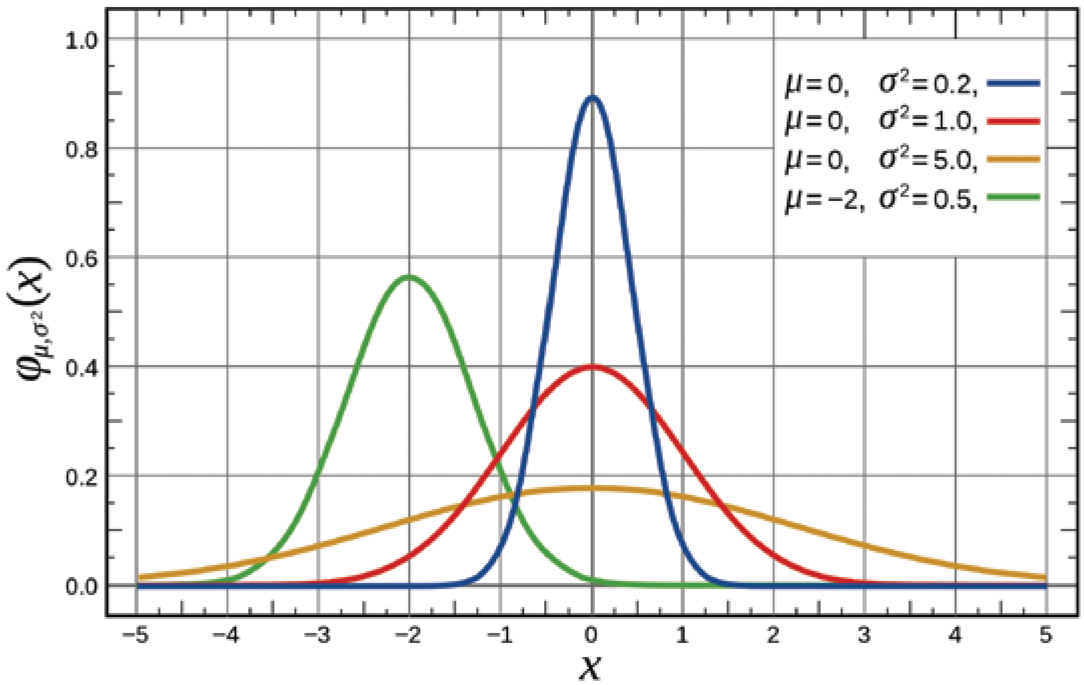

Figure 3. An example PDF Bell curve graph depicting different average and standard deviation parameters. From [6]

Figure 3. represents Probability Density Function' (PDF) of a sample performance metric with four different Systems Under Test (SUT), each with different average and standard deviation (variability) parameters. The graph illustrates 'normal' probability distribution revealing statistical characteristics of metric X, representing at least approximately, any variable that tends to cluster around the mean as norm , shown as, but not necessary, the familiar 'bell curve'. It shows the relative probabilities of getting different values. It answers the question,

- .What is the chance I will see a certain result?
- What is the mean value or norm of the respective SUT at this specific performance metric?
- How cohesive and stable is the performance for each SUT?

Examining the random process represented by the red curve in figure 3, we would expect outcomes with a value around ‘0’ to be twice as prevalent as outcomes of around 1.25 (40% versus 20%). However, in some cases, we are more interested in a threshold performance value as benchmark value than individual probability point, what is the probability of having performance being less or greater than a threshold value? A transformed PDF ,Cumulative Distribution Function (CDF) (Figure 4.) helps answer this question.

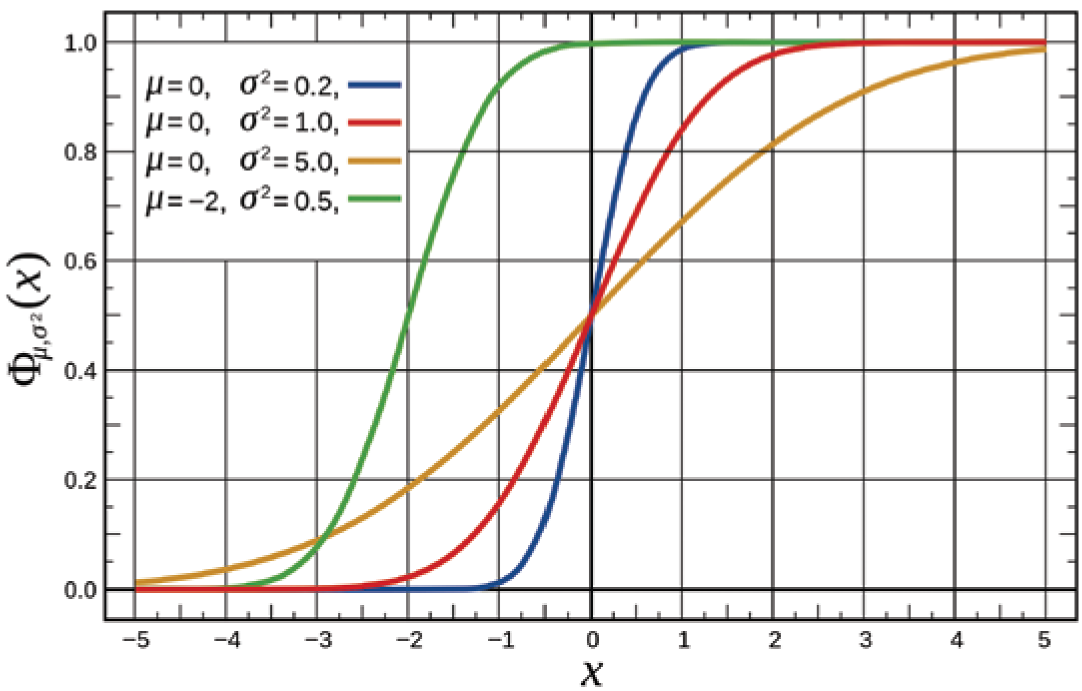

Figure 4 A typical cumulative distribution function (CDF) graph. From [6]

Examining the random performance metric represented by the red curve in figure 4,

- The probability of producing a result less than -0.75 is about 20%.
- The probability of producing a result less than 0 is 50% and
- The probability of producing a result less than 2 is about 95%.

To characterize wireless performance, CDF graphs can provide immensely useful information for system designers to compare performance of alternative solutions. Ideally, we prefer better mean value and smaller deviation, but if the ideal choice is unavailable, we have different optional solution to choose. All the choices should be put in application specific context to choose the right protocol for right application. The principle is:

1. Drastic different mean value, if mean value represent positive metrics, like throughput or lifetime, bigger is better, ideally we prefer SUT with bigger mean value and less deviation.
2. Same mean value, different deviation: long tail means performance not stable, we prefer smaller deviation.
3. Slightly different mean value, but one with long tail, we prefer stable over slightly improved peak performance.
3. But if the optimal solution is not available, we have choice over performance stability and higher norm performance according to different application scenario.

   (opt1.) Higher performance potential but less predictable performance

   (opt2.) Less performance potential but higher stable performance

Reference [6] presents practical guidelines on how to actually acquire the statistical performance PDF and CDF curve of a SUT, nevertheless sampling is the key to recover statistical performance and drawing the PDF and CDF curve of a wireless system. Furthermore, to predict real-life performance accurately, researchers ideally should conduct sampling tests across all relevant dimensions and variable if possible. However, in most cases, the design space is too big to exhaustively investigate all factors influencing the final performance. But planners must at least consider three rough dimensions, as we have mentioned above, to characterize wireless performance accurately: time, space, and data pattern. Under each category, there are vast known or unknown parameters that can affect the performance. Hence it is worthwhile to investigate the effect of parameter change to specific performance metric (sensitivity analysis). The effective way to deal with the vast design space is parameter reduction and inter-dependency analysis. To find the smallest set of the parameters that most significantly contribute to performance objectives**,** we will further explore effective mathematical models dealing with this aspect in later chapter.

3.5 Multi QoS performance metrics and energy constraint conflicts in WSNs

WSNs energy-oriented research originates from conflict between application performance requirements and limited energy supply by battery. In foreseeable future, it will remain as a bottleneck for its widespread development unless a breakthrough at relevant material science field occurs. More recently, WSNs have been used for more advanced applications such as wireless building automation, industrial process automation, security monitoring, disaster intervention and medical interventions. We can not overemphasize energy conservation while ignoring application's targeted functionality. To what extent we should emphasize the importance of energy aspect comparing other QoS objectives depends on application scenarios. Tradeoffs have to be made on per application basis.

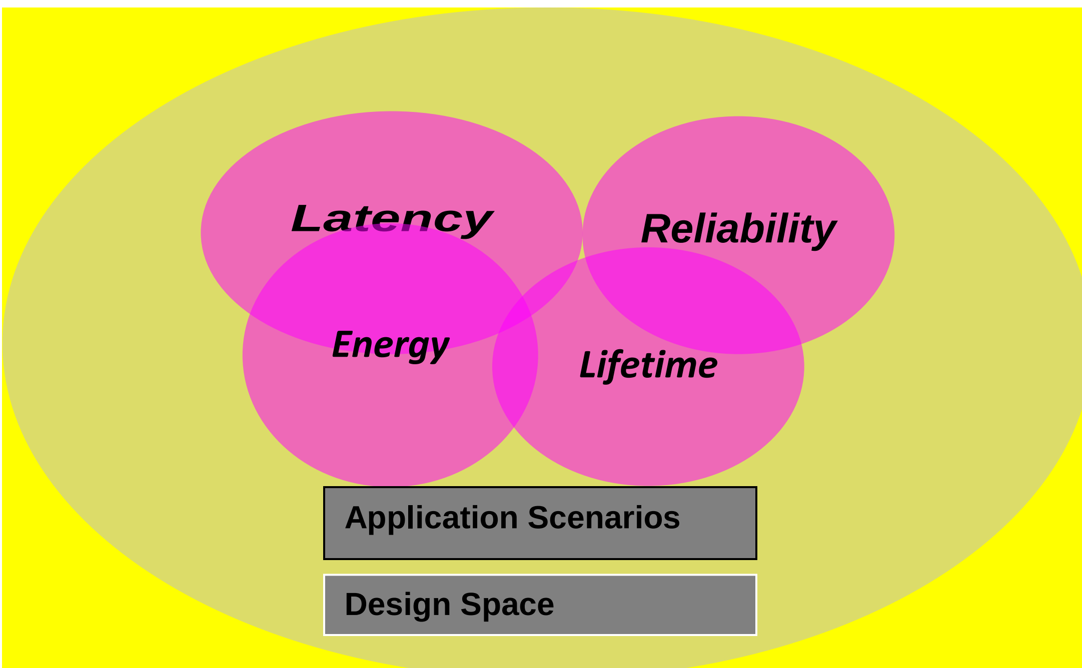

Figure 5. Multi competing performance metrics under wide design space for different applications

The goal of wireless performance evaluation and testing is either; 1). Test a single system to determine if it meets some minimal application specific performance criteria or; 2). Compare two systems to determine 'which is better' under certain application specific scenario. In both cases, the intent is to predict the real-life systems performance under application context. There is a tendency in WSN application specific design and evaluation, that proposed protocols normally focus on optimizing one or two targeted performance criteria, while ignoring the impact on other performance criteria. Typically WSN lifetime (energy efficiency), response time (delay), reliability, and throughput are among the main concerns in the design and evaluation process. Under the constraint of wireless sensor node size, the energy budget and computing resources are unfeasible to afford any luxury algorithms. Under such constraint, there does not exist a perfect optimal solution satisfying all performance metrics in the problem (NP hard problem), rather, the question we sought is how to consider multiple criteria explicitly and structure complex problems which would rather lead to more informed and better decisions (Figure 5).The methodology should be general enough to contain different application scenario according to decision maker's preferences.

There have been few works on application-driven protocol stack evaluation for WSNs. Our evaluation methodology, similar to analytic hierarchy process (AHP) [22,7], using a Single Performance Index (SPI) for each alternative solution or System Under Test (SUT), as the final quantified goodness measurement for alternative solutions comparison (Figure 6).

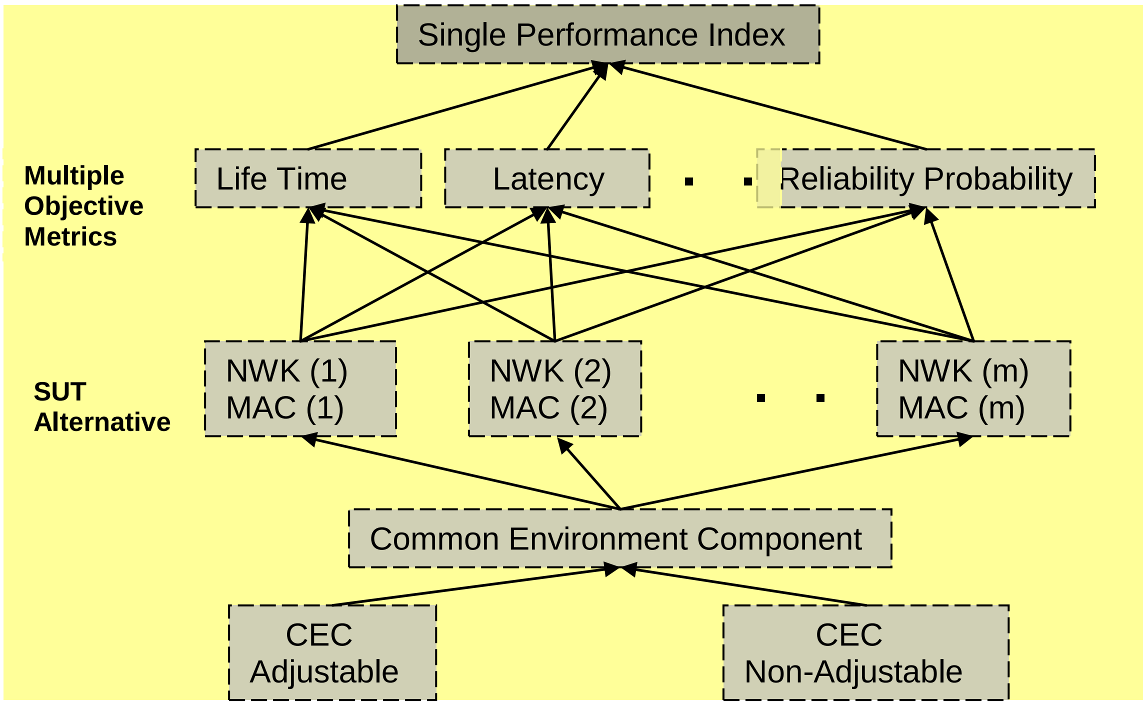

Figure 6. Single Performance Index (SPI) composition

The end-user configures the relative importance of the individual design metrics in a function that maps the individual metric values to a single SPI value. Firstly we define the default overall objective function as a weighted sum of the individual design metric normalized values as other AHP methodologies normally do: separation of defining design metric, and weighing those functions' importance in an overall objective function.

$$SPI_{norm} = a * m(L) + b * m(R) + c * m(T) ;$$

Here m (L), m (R) and m (T) are three examples of performance measurement, representing performance of Life Time, Reliability and timeliness respectively. A key feature of our approach is that, we introduce the statistical analysis of the resulting experiment data, not only using measurement mean value instead of one supposed normalized value (which is not realistic representation of the dynamic truth of the wireless network nature), we introduce deviation of performance measurement PDF as a critical secondary performance metric to emphasis the importance of stability of SUT protocol as aforementioned. Even a higher mean performance metrics, given the performance spreads over a wide spectrum of measurement, not cohesive to so called norm performance (mean) value, it will be problematic for certain application scenarios which require consistent performance , such as health monitoring application and multimedia application.

We introduce stability performance index, as:

$$SPI_{stability} = a^{'} * (1/\delta^{2}\ (L)) + b^{'} * (1/\ \delta^{2}\ (R)) + c^{'} * (1/\delta^{2}\ (T));$$

So overall we have:

$$SPI = SPI_{norm} + SPI_{stability} =$$

$\{ a * m\ (L) + a^{'} * (1/\delta^{2}\ (L))\ \} + \{ b * m\ (R) + b^{'} * (1/\delta^{2}\ (R)) \} + \{ c * m\ (T) + c^{'} * (1/\delta^{2}\ (T)) \}$

={[weighted Lifetime performance score] + [weighted Reliability performance score]} + …

$$SPI = \sum_{i=1}^{n} (W_i * Metric_i(mean) + W_i^{'} * (1/\delta_i^{2}));$$

Here W_i = (*a, b, c…)* represents respectively the user specified relative importance of the performance metrics norm values. And $W_i^{'}$ = ($a^{'}$, $b^{'}$, $c^{'}$ …) indicates the relative importance of the metrics cohesive characteristic represented by deviation ($1/\delta^{2}$). The relative importance of each design metric as weight is assigned by considered application specific scenarios, how important in your application is certain metric (network lifetime, reliability, throughput, delay, etc) respectively? How important is cohesive characteristic of performance to your application? Which metric is utmost important for you?

In general, in this chapter, some key issues involved in WSNs performance evaluation are identified. We have explored uncertainty elements contributing to statistical natured performance. Trying to characterize the stochastic performance, we introduced deviation as an additional performance metric to indicate the cohesive characteristic of the measurement. We further established a single performance index to make different solutions based on application specific assumptions comparable. The quantified index is constructed by sum of product of each metric measurement value and its relative importance specified by user preference according to application requirements. In this way, with fully consideration of feature of the application, we setup a generic index to distinguish the quality of different System Under Test (SUT).

Up to now, we only concerns the QoS performance index, however, from WSNs designers' perspective, performance gain is only one pillar of the WSNs evaluation. In next chapter, we will establish an evaluation framework from higher level abstraction, to reveal other important pillars of WSNs evaluation.

CHAPTER 4

A GENERIC WSNS EVALUATION FRAMEWORK

While evaluating proposed architectures in WSNs literature, we realize that over the last few years of WSNs evolvement, a large number of architecture solutions have been introduced according to de facto application specific methodology targeting a wide range of applications such as environmental tracking applications, medical and industrial applications, home automation applications, surveillance systems, etc. This diversity requires a specific design methodology and settings for the targeted application. If we consider all of the specific scenarios and different metrics in the architecture evaluation, it will be difficult to compare with each other. We cannot see the big picture as we immerse ourselves totally into complicated details. Nevertheless, a system can be considered at various abstraction levels, allowing considering only relevant properties and behaviors at different abstraction levels, whose relevancy will depend on the context. In this chapter, we will introduce an evaluation methodology from high level abstraction, concerning only profound pros and cons from viewpoints of tradeoffs between generality, performance and cost in architecture design. A practical benchmarking workflow is followed to illustrate how to obtain a single performance index value.

4.1 System Architecture Evaluation Fundamental Concepts

System architecture is the conceptual model that defines the system structure, externally visible properties of components, and internal interaction among them. Architecture design is a process which defines and refines components based on reasoning about the goal and structure of the system. An architecture evaluation framework provides a comparison platform for different architectures. A system can be evaluated through different aspects of concerns about the system. Moreover evaluation framework should be a dynamic evolution process as new science and technology advancement would make certain constraints relax and push new frontier challenges.

Since any evaluation framework can only provide one view of the possible perspectives which focus on certain metrics, it is important that first one should focus on 1) clarifying the purpose of the evaluation; 2) developing key evaluation questions or metrics accordingly. The purpose of our evaluation framework is to: 1) provide high level abstract information regarding architecture inherent performance pros and cons based on three main competing metrics; 2) demonstrate the contribution of the evaluated architecture towards broader goals.

4.2 Triangle Constraints in General System Architecture Evaluation

Traditionally the Project Constraint Model recognized three key constraints; "Cost", "Time" and "Scope", these are also referred to as the "Project Management Triangle" or "Iron Triangle", where each side represents a constraint. One side of the triangle cannot be changed without affecting others. It is useful to help with intentionally choosing system biases according to the goals of your project.

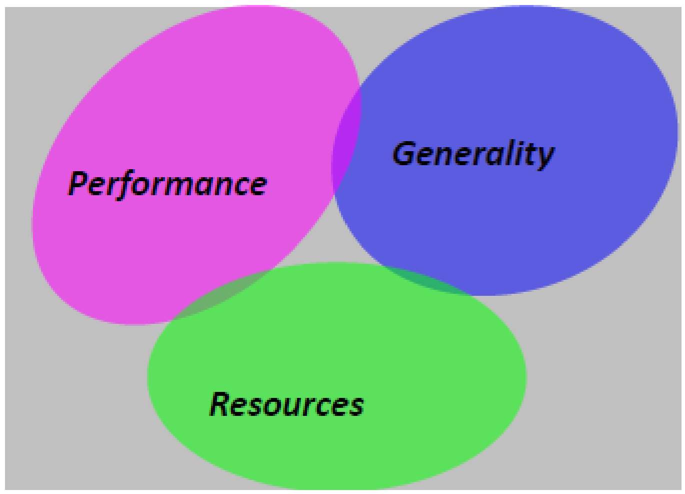

Figure 7 Triangular Constraints

Inspired by *Project Constraint* Model, when we investigate WSNs architecture we observe corresponding “Iron Triangle” in the architecture evaluation modeling: three key *constraints,* namely performance gain, generality and resource cost. These constraint metrics cannot be optimized perfectly simultaneously. In another words, optimum optimization of one metric from the triple constraints cannot be achieved without penalties from others. The shift focus of the three constraints can capture the main characteristic of the architecture behavior.

1) Resources Cost: Energy and computing resources (CPU and memory) limitation is the primary concerns when WSNs just started due to the cubic

centimeter level form of the nodes. Basically battery is irreplaceable in the fields if we consider the harsh environment, amount of nodes and the human resource cost of the replacing work. Energy consumption mainly involves signal transmitting/receiving and CPU computing, it is estimated that the energy consumption when transmitting 1 bit of data on the wireless channel is similar to energy required to execute thousands of cycles of CPU instructions [48]. The implication for this phenomenon is that node processing is far more energy efficient than only signal relay. Nevertheless, CPU and memory limited capacity indicates that WSNs cannot bear the luxury of any complex and advanced algorithm. Without battery (supply), antenna technology (consumption) and IC technology (computation) big breakthrough, resource cost will remain critical in architecture design and evaluation.

2) Performance: Performance is the primary goal of any application. As high level abstraction, here performance is a collective word covering diverse Quality of Service (QoS) metrics. As applications evolve, more advanced requirements to underlying architecture are emerging. Under current resource cost restriction, for the purpose of optimal performance, WSNs system designers adopt a practical methodology: i.e. application specific design. To realize the required performance of WSN, many research groups have utilized every means to make their work done. Roughly they tried to follow the TCP/IP reference model, but applications were in general vertical integrated from layered perspective by their own set of components which are specifically designed to work together in this very application. Performance is gained by the cost of generality and component reuse.

3) Generality: Generality requirement for WSNs architecture designers is an extra metrics in addition to resource cost and performance gain from historical view of the WSNs development. It is secondary for designers to consider in individual designing practice. Though, for widespread application and commercialization purpose, generality is critical as much as the other two metrics. We can draw lessons from the success of Internet. TCP/IP is successful in its own modularity instead of previous monolithic solution, where, applications need not be built up form bottom up: from hardware, hardware drivers to application functionality, rather, different innovations in layering module can fit in the layers which they reside in with assumption of the lower layer support and higher layers compatibility, which facilitate the new innovation being easily invented by individual efforts and adopted by existing architecture. Separation of concerns is crucial to the design of the layering architecture in which great efforts have been made to separate concerns into well-defined layers. This allows protocol designers to focus on the concerns in one layer, and ignore other layers. The cost of generality is inefficiency: from application raw data (the meaningful payload) each layer' header is attached to original raw data from top layer all the way down to the physical layer signal bits ready for the transmitting. Notice that it is the part of the application raw data which is the fundamental purpose of the transmission, whereas the information sent with it, such as headers or metadata, solely as overhead to facilitate delivery. In WSNs context, this overhead has significant impact on overall energy consumption and further network lifetime. However, Header encapsulation is essential for

designing modular communication protocols in which logically separate functions in the network are abstracted from their underlying structures.

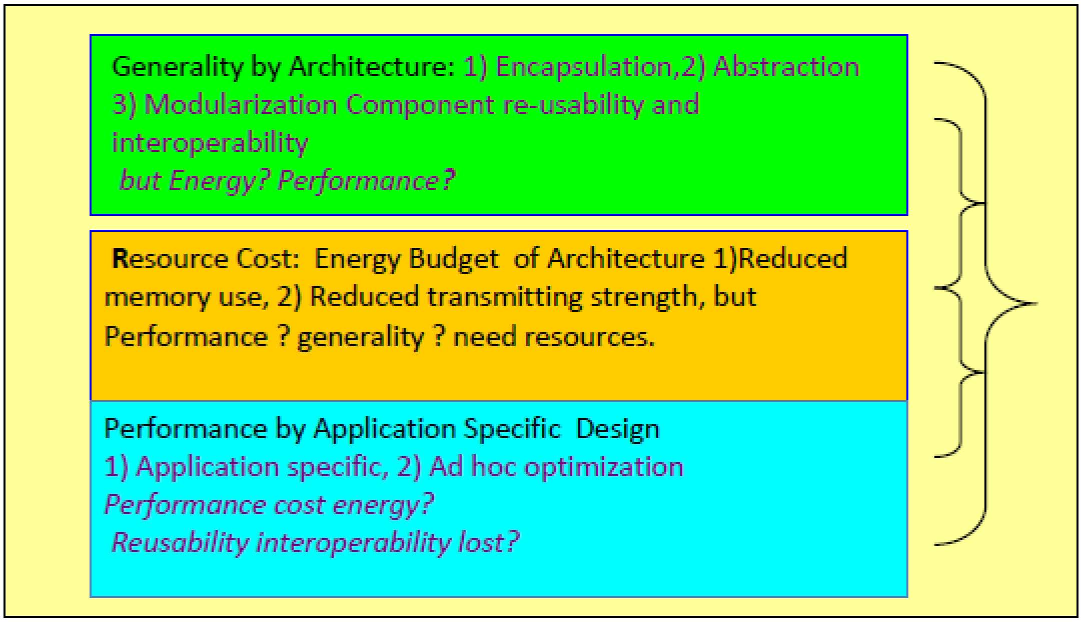

Figure 8 Conflicts between Three Main Metrics

We summarize that aforementioned three metrics deeply interweave in such a way as Figure 8. The cost of the layering and generality is more layers of 1) Encapsulation,2) Abstraction 3) Modularization to contain differences of coexisting protocols in the same layer, the performance efficiency of the overall network stack will drop considerably due to more layers of wrapping and unwrapping around original raw data. Resource cost on the other hand increases when dealing with transmitting and receiving more overhead signal bits from antenna. When we try to reduce resource cost as a single purpose, there is various means to reduce transmitting power and memory footprint at same time maintaining basic operation. In this case, performance and generality are too luxury to expect. Lower transmitting power means less signal coverage and more interferences noise; lower memory footprint means complex algorithms like nonlinear programming, evolutionary multi objective optimization are infeasible. Even routine algorithms need more skillful programmers to develop. In such a simple basic settings, generality is out of the question.

The third spectrum of the constraint triangle is performance gain. Industry experiences show that performance and generality are natural enemies under energy resource constraint. In the case of TCP/IP, generality prevails at the cost of performance. Performance drops when layers of encapsulations try to gain more generality and hide the difference of coexisting protocols. That is why application specific methodology prevails over other methodologies in WSNs design field. Whichever means are used in an ad hoc way to optimize performance for one specific application.

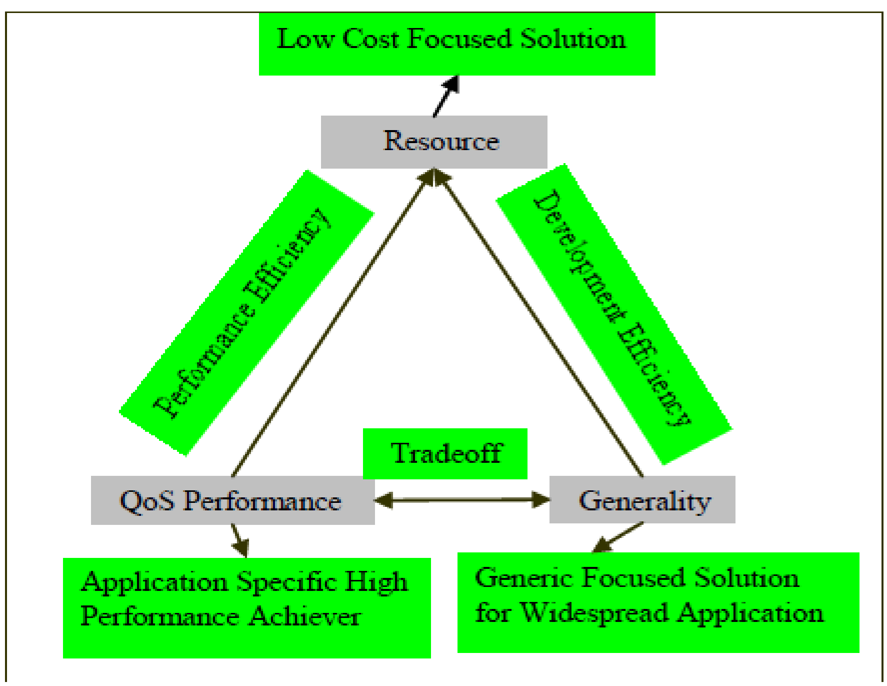

Figure 9. Architecture Efficiency in Two Aspects

4.3 Triangular Constraints Reduce to Bilateral Constraints Tradeoff

We further elaborate the triangular constraints tradeoffs as presented in Figure 9. Due to the inherent constraints of sensor nodes, resource efficiency is one of the key requirements and critical focus of WSN research. When combining generality and performance to resource cost consideration, two important types of efficiency are identified: Network QoS Performance Efficiency (PE) and Deployment Efficiency (DE). Network QoS performance efficiency is measured by the ratio of performance gain to resource cost; while development efficiency is measured by the resource cost of certain level of generality.

$$PE = QoS\ Performance\ Gain\ /\ Resource\ Cost \qquad (3)$$

$$DE = Level\ of\ Generality\ /\ Resource\ Cost\ for\ Generality \qquad (4)$$

A good architecture can improve the resource efficiency of the network and then enhance network performance efficiency without the cost of architecture efficiency, e.g., the ability to accommodate heterogeneity and adapt to a wide range of underlying communication mechanisms at diverse scenarios. An architecture built on static technologies is destined for obsolescence [42]. Just like ZigBee as an architecture cannot provide a viable solution due to “ZigBee proposes a classic layered architecture, but each layer assumes a specific instance of the surrounding layers: e.g., the routing layer assumes the IEEE 802.15.4 link and physical layers.”[42]. Nevertheless, the two efficiencies are corresponding to the interest of focus for application specific performance-pursuit solution and generic widespread-pursuit solution.

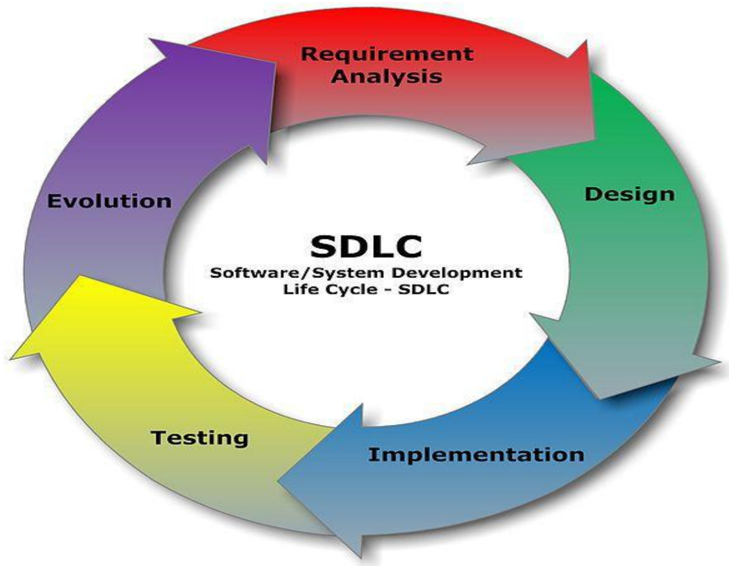

Figure 10. SDLC and Architecture Efficiency

Let us first consider performance and cost as two factors together and we define a new combined metric called Performance Efficiency (PE) as presented in equation (3). Here QoS performance is a unified quantity representing performance gain all the way through architecture evaluation and comparison. Resource costs consist of memory /CPU utilization and energy expenditure. There needs a mechanism to weigh different parameters of interest to reveal the relationship between the coexisting elements. Most importantly, the performance and cost indicators should have threshold values assigned to each metric as essential part of evaluation algorithm. Once resource bottleneck metric has come to the threshold value, or performance metrics come to the lower bound threshold, no further performance efficiency evaluation is needed. Since QoS performance is not an absolute value which is relative to evaluator's preference of metrics of interest and weighing indices of metrics, so the measure of goodness is subjective relevant to evaluators' intention. Further more, QoS performance is not a static value which dynamic changing at spatial and temporal space, thus PE will demonstrate statistical characteristics as workload, underlying environment and system utilization change as shown in [44].

Secondly let us consider generality and cost as two factors together. We define a new combined metric called Development Efficiency (DE) as shown in equation (4). Here Level of Generality can be an ordinary number or a relative percentage to represent the extent to which generality is preserved in the design. Resource Cost for Generality is the extra resource cost of additional coding for generality purpose. We put Development Efficiency (DE) in System Development Life Cycle (SDLC) context as shown in Figure 10. Architecture DE concerns beyond the QoS performance metrics such as throughput, delay and reliability of current implementation. DE in SDLC context concerns metrics that can facilitate the openness and widespread application of the architecture, which is related to architecture quality assurance attributes such as updatability, compatibility, maintainability, testability, flexibility, portability, reusability, interoperability. According to McCall model [47], these attributes are grouped into three categories in SDLC process: operation {updatability, reliability, compatibility}, transition {portability, interoperability, reusability} and revision {maintenance, testability, flexibility}. Among these metrics, we explain some of the metrics specifically relevant to architecture development efficiency.

- Architecture portability: the level of efforts needed to transfer a software program from one platform to another platform.
- Architecture flexibility: the level of efforts required to modify an operational program to accommodate changing needs or a changing environment.
- Architecture interoperability: the effort required to couple one software system to another.
- Architecture compatibility: protocols capable of co-existing in harmony under one system for the potential response to wide application scenarios

Thirdly let's consider the tradeoff of the two identified efficiency concepts Performance Efficiency (PE) and Deployment Efficiency (DE) in Figure 9. Through aforementioned concepts extraction, a triangle constraint three dimensional problem is reduced to a two dimensional tradeoff problem. No absolute merits are possible in tradeoff evaluation, only *relative merits* can be obtained on preference bases as introduced in Analytic Hierarchy Process [45, 46]. The challenge here is how to quantify the cost and benefits of architecture decisions, specifically the tradeoff of two identified efficiency concepts.

Typically we categorize communication architecture according to layering status, namely layered architecture, cross-layered architecture and layer-less architecture. The classification method is from structure point of view, but not power enough to fully explore the essence of WSNs. Never before was there such a wide range of network architectures and deployment options available to network designers as in WSNs. How to classify and evaluate qualitatively and quantitatively to decide its suitability for specific application is tediously hard only from structure organization. Now our triangular constraint model provides a new way to classify and evaluate WSNs architectures according to three critical metrics, cost, performance and generality. As shown in Figure 9, there are three extreme forms of architecture which emphasizes one metric only: 1) Generality, 2) High Performance and 3) Low Cost. From this we can identify inherent merits of structured architecture [50, 51] as such as:

1) Layered: generality best, performance and cost poor;

2) Cross-layer: generality modest, better performance and cost;

3) Layer-less generality lost, possible better performance and energy cost, but with unseen problem in development process.

Of course, this roughly classification lost granularity. We can further quantitatively weigh each of the three metrics to reflect its importance in architecture evaluation to suit specific application domain scenarios. Combining experiment measurement results, such that the final single score SPI value in Eq. (5) will reveal the application requirement and granularity. Here W for weighing, M for Measurement from experiments:

$$SPI = W_{generality} * M_{generality} + W_{\cos t} M_{\cos t} + W_{QoSperformance} * M_{QoSperformance} \qquad (5)$$

Or we can quantitatively weigh the two identified efficiency instead as (6) to reflect the designer's preference over development efficiency and performance efficiency for perceived application specific design . Here E is for efficiency measurement.

$$SPI = W_{development} * E_{developemt} + W_{QoSperformance} * E_{QoSperformance}. \qquad (6)$$

In summary, in this chapter we present a qualitative evaluation framework at system designer's perspective. This high level abstraction only concerns common fundamental issues in system design and evaluation. In this framework, evaluation metrics are not treated at the same level. A hierarchy of metrics is introduced to distinguish its effective scope. QoS performance, generality and cost are at the top of the hierarchy as most important factors in the evaluation process. Under QoS performance, there is second level of metrics: response time, reliability, throughput etc. which is usually main concerns of other QoS related work. In next chapter, we will show example of how to quantify QoS performance index based on measurement.

CHAPTER 5

A PRACTICAL BENCHMARKING SOLUTION FOR QOS PERFORMANCE INDEX

Last chapter we qualitatively develop a high level abstraction for effective evaluation of competing solutions. To show how to quantify the evaluation framework, we will introduce a practical workflow to acquire single QoS performance index for each solution. Beware that as an example, this chapter aims only to gain QoS performance index. However, as a guideline, the methodology can be modified to apply to wider scope as indicated in chapter 4 Eq. 5 and Eq.6.

5.1 A practical workflow of benchmarking solution

System evaluation process starts with the end users as application experts who know very well what kind of performance they needs; they specify the most concerned QoS performance metrics, and the weights of each metric.

The WSNs designers decide the initial parameters according to the literature studies and previous experiment experience. Then for each performance metric start the iterative experiment process as such:

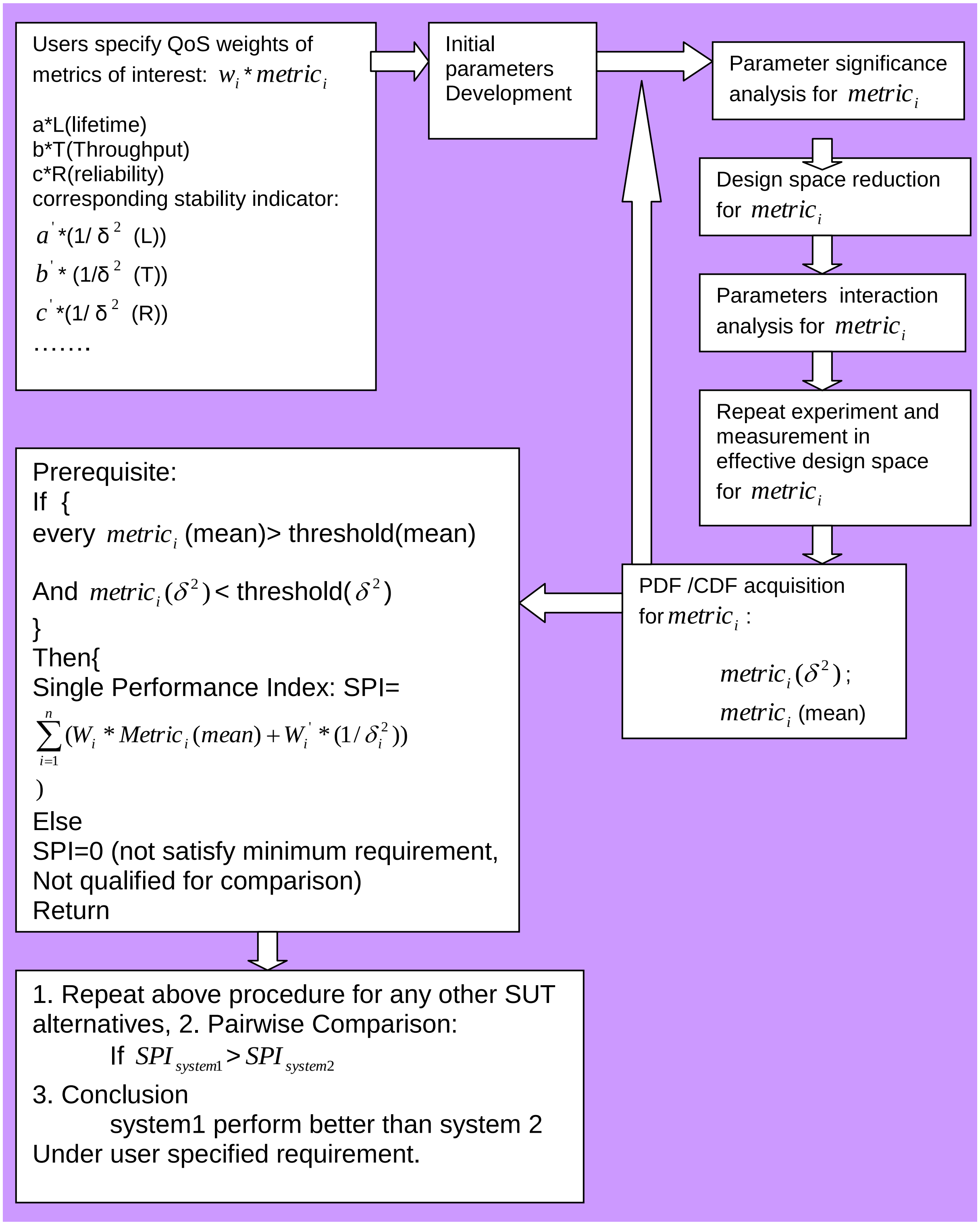

Figure 11. Workflow of proposed benchmarking solution

(1). Parameters significance analysis for $metric_i$.

Repeat *l* experiment measurements, recode each experiment the state of each parameter x_i as f_{ji} ($1<i<n, 1< j<l$) and corresponding performance measurement ψ_j ($1<j<l$). Then use linear aggregation and P-value to decide significant parameters to the $metric_i$ as detailed in next chapter.

x_1	x_2	$\cdots$	x_n	Ψ
f_{11}	f_{12}	$\cdots$	f_{1n}	Ψ_1
f_{21}	f_{22}	$\cdots$	f_{2n}	Ψ_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
f_{l1}	f_{l2}	$\cdots$	f_{ln}	Ψ_l

(2). Design space reduces from *n* parameters to *m* parameters for $metric_i$ according (1).

(3). *m* parameters interaction analysis for $metric_i$.

Tune the parameters based on the reduced parameters set, Repeat *l* experiment measurements, recode the state of each parameter x_i as f_{ji} ($1<i<m$, $1<j<l$) and corresponding performance measurement ψ_j ($1<j<l$). *then use the Choquet nonlinear aggregation model as described in later chapter to decide the most effective parameters set including interaction effect of individual parameter.*

(4). Now we have finally approach the effective parameters set. Tune the effective parameters set, repeat measurement and get the performance curve, get the $metric_i(\delta^2)$ *and* $metric_i$ (mean) for $metric_i$.

(5). Change another metric of interest, start over again from (1).

(6). When all metrics of interest finish evaluation, calculate the Single Performance index as aforementioned formula.

$$SPI=\sum_{i=1}^{n}(W_i * Metric_i(mean) + W_i^{'} * (1/\delta_i^2));$$

(7). For competing solution for pair wise comparison, repeat the above process and get SPI value and compare:

If $SPI_{system1} > SPI_{system2}$ then system1 perform better than system 2

Notice that we can setup threshold value for minimum requirement, any time if $Metric_i$ mean or deviation is less than the threshold value, the candidate solution is not qualified for further comparison due to unsatisfactory for minimum user specification

5.2 Metrics Development and interdependency Analysis

Environment component is an indispensable part of WSN system evaluation. Environment elements play important role on system /protocol behavior. Common Environment Component (CEC) Non-adjustable in Figure 4 is the physical environment component, including fading, energy model which is normally nonadjustable in the targeted application scenario, is same for all the alternatives of interest at one test, but changeable according to typical application scenario. CEC Adjustable is used as sensitivity test parameters to get a range of performance in the reasonable parameter design space. Typical CEC Adjustable should include but not limited to:

network topology _ network density _ network size _ traffic pattern _ transmission power _ duty cycle _ modulation type _ SNR _ data rate _ BER _ packet length _ contention window size _ etc.

They are roughly under MAC layer in a standard protocol standard. These parameters are key parameters in protocol evaluation, but not exhaustive list of all key elements in WSN evaluation as new parameters are introduced by ongoing new proposals. We put these parameters in the context of equation (2), either preplanned structured topology or ad hoc topology, various density value, different number of nodes, show spatial influences and scalability performance on objective functions; various traffic patterns, different predefined duty cycle and different interference characteristics should show temporal influences on system function output.

CEC Adjustable ideally should be orthogonal parameters, independent to other parameters, reduce the complexity of the evaluation and sensitivity analysis. Under this orthogonal assumption, sensitivity analysis then can adopt a common approach which is that of changing parameter One-factor-At-a-Time (OAT), to see what effect this parameter may produce on the output.

- Moving one factor at a time and,
- Going back to the central/baseline point after each movement.

This appears a logical approach as any change observed in the output will unambiguously be due to the single factor changed. Furthermore by changing one factor at a time one can keep all other factors fixed to their central or baseline value. This increases the comparability of the results**.** However, despite its simplicity, this approach cannot detect the presence of interactions between input factors [8]. Previous researches have shown the complex relationship of interdependency exists between the parameters. How these parameters are collectively related to and individually contribute to one specific performance objective metric function is critical in performance evaluation and sensitivity analysis. The design space is enormous if we further consider the combining effect of some parameters. To reduce the design space to manageable size it depends on how we derive the interdependency among system parameters and identify the most critical and significant effect of subset of parameters on the performance metrics for system objective function**.** It is important to distinguish here the notion of criticality of parameter with that of importance or weight of performance metrics***.*** By *critical,* we mean that a parameter with small change (as a percentage) may cause a significant change of the final performance. It is objective observation of parameter different effect on SUT, on the other hand, the importance or weight of performance metric is user specified subjective preference to emphasize the application requirement.

CHAPTER 6

PARAMETERS REDUCTION METHODOLOGY

There is a big design space in wireless network evaluation. In the literature [57,58], some research works have focused on the sensitivity analysis of individual parameter to the performance metrics. Metrics ideally should be orthogonal to each other which make it possible to evaluate the effect of each metric independently. But in real world systems, interdependency and combination among the parameters with combined effect on objective function are not rare and hence, complicate the issue. After presenting a practical workflow in previous chapter, in this chapter we present parameters reduction methodologies to reduce the evaluation space and further make the proposed solution implementable.

6.1 Interaction effect of multiple variables

For interdependency measure and significance analysis, we would like to know which subset of parameters has the most significant effect, individually or collectively, on the performance metric of interest. Thus we can get rid of the system parameters with insignificant effect on performance metric and reduce the design space accordingly.

We are interested in orthogonal parameters individually and combined parameters collectively. We define the property of orthogonal in parameter selection as: If two parameters, a and b, are orthogonal between each other, they independently contribute to the performance metric, their effect on performance metric is additive such that:

i). Effect of x_1 is the same at all values of x_2 (independent principle); ii). The effect of all the X variables are simply added together (f (a + b) =f (a) +f (b)).

However, many parameters are interdependent in nature, their combined effect either positively or negatively affect the objective function. Hence, the interaction among predicative attribute cannot be ignored and would furthermore add new dimension to the design space. They are called non addictive measure. We define non-additive equations as:

i).The effect of x_1 varies according to the values of x_2 (dependent principle 1);

ii).The effect of each variable in the interaction depends on the values of other variables (dependent principle 2);

iii). Combination effect dose not add up for individual effects:

if f (a + b) ≠ f (a) + f (b); either: f (a + b) > f (a) + f(b) (in which case combination has positive effect more than two separate individuals) or f(a + b) < (f(a) + f(b) (in this case, the combination has conflicting effect, weakening each other).

6.2 The limitation of existing work

System design and evaluation normally include many system parameters, making the multidimensional optimization and evaluation problem very difficult to solve**.** P-value and correlation analysis are the most common ways in statistics to analyze the significance of effect between two random variables. However only pure statistical answer to this problem is unsatisfactory. One attempt is to measure the importance of a variable purely by its observed significance level (P value). The distinction between statistical significance and practical importance applies here**.** Some variable in practice is difficult or meaningless to change randomly once the network application scenario is specified**,** like topology, network size, etc. Another attempt is to measure the importance of a variable only by the magnitude of its regression coefficient. This approach fails in fidelity because the regression coefficients depend on the underlying scale of measurements. The question involved can be answered only in the context of a specific research question by using subject matter knowledge.

We will introduce some existing analytical and experiment works which show promising potentials in interdependency measure and significance analysis under uncertainty. To make evaluation operational, we try to reduce parameters to an effective subset using quantitative significance analysis and linear regression methodology. Moreover, to realistically reflect the combined effect of interdependency among some system parameters on final performance metric, nonlinear regression and non-additive measure theory are used.

6.3 General Problem Formulation

First let's start with a general problem formulation. Without loss of generality, let *X(n)* be the original design vector with n system parameters $\{x_1,\ x_2\ ,\ldots, x_n\}$; *Ψ (m)* be the m system performance metrics of interest; for each metric ψ_i in *Ψ (m)* we will establish a linear (in chapter 6.4) and a nonlinear (in chapter 6.5) relationship to identify its significant subset individual parameters and combination parameters.

6.4 P -Value with Linear Regression Model

We introduce this approach similar to in [11, 12, 13, 14], P-value and linear correlation analysis are combined to establish the relationship between some of the parameters and the output. highest values of correlation and the sign of the correlation are captured as index of parameter significant to performance metric. P-value is specifically used to obtain the most statistically significant variables influencing the output variable, and linear correlation is used for both influential variables and the variables direction effect. A null hypothesis H_0 is tested by gathering data and then measuring how probable is the occurrence of the data. Assuming the null hypothesis is true, which means there is no significant effect of tested parameter on specified performance metric. If the resulting p value <

specified threshold (here conventionally as 0.05), then the null hypothesis H_0 is rejected and conclusion made accordingly. With marginal probability of type I error, the variable under test has statistically significant effect on the performance metric.

The problem of modeling based on multivariate linear regression involves choosing the suitable coefficients of the modeling such that the model's response well approximates the real system response in terms of the performance metric of interest. A general linear model can be established as such:

$$
\begin{aligned}
\psi_i^{(1)} &= a_0 + a_1 p_1^{(1)} + a_2 p_2^{(1)} + \ldots + a_n p_n^{(1)} \\
\psi_i^{(2)} &= a_0 + a_1 p_1^{(2)} + a_2 p_2^{(2)} + \ldots + a_n p_n^{(2)} \\
&\cdots \\
\psi_i^{(l)} &= a_0 + a_1 p_1^{(l)} + a_2 p_2^{(l)} + \ldots + a_n p_n^{(l)}
\end{aligned}
\qquad (7)
$$

Where ψ_i is the i^{th} metric of Ψ(m) which is under consideration, $\psi_i^{(1)}$ to $\psi_i^{(l)}$ is the l measurement values of performance metric ψ_i in repeating experiments on simulator and ($(p_1^{(k)}, p_2^{(k)}, \ldots p_n^{(k)})$) are the values of *n* chosen effective parameters for the k^{th} experiment, respectively. a_j *(j<=n) is the coefficient for each parameter.* Equation (7) can be rewritten in matrix format as: finally

$$\underbrace{\begin{bmatrix} \psi_i^{(1)} \\ \psi_i^{(2)} \\ \cdots \\ \psi_i^{(l)} \end{bmatrix}}_{\psi} = \underbrace{\begin{bmatrix} 1P_1^{(1)}\, p_2^{(1)} \ldots p_n^{(1)} \\ 1p_1^{(2)}\, p_2^{(2)} \ldots p_n^{(2)} \\ \cdots \\ 1p_1^{(l)}\, p_2^{(l)} \ldots p_n^{(l)} \end{bmatrix}}_{P} \underbrace{\begin{bmatrix} a_0 \\ a_1 \\ \cdots \\ a_n \end{bmatrix}}_{A} \quad \text{Or} \quad \psi = P * A \qquad (8)$$

Using the above formulation, the approximation problem becomes to estimate the values of $a_0^{'}, a_1^{'}, \ldots a_n^{'}$ to optimize a cost function between the approximation and real values of performance metric of interest. . Then, an approximated performance metric of the application for the j^{th} experiment is predicted as

$$\hat{\psi}_i^{(j)} = a_0^{'} + a_1^{'} p_1^{(j)} + a_2^{'} p_2^{(j)} + \ldots + a_n^{'}\ p_n^{(j)} \qquad (9)$$

One of these methods used widely in computer science is the Least Square Regression which calculates the parameters in equation. (8) by minimizing the least square error as follows:

$$\text{LSE}=\sqrt{\sum_{m=1}^{l} (\hat{\psi}_i^{(j)} - \psi_i^{(m)})^2} \qquad (10)$$

The set of coefficients $a_0^{'}, a_1^{'}, \ldots a_n^{'}$ approximately describes the relationship between the performance metric of interest and its configuration parameters.

$$\psi_i = F(p_{1,} p_{2,} \ldots, p_n) \cong a_0^{'} + a_1^{'} p_1 + a_2^{'} p_2 + \ldots + a_n^{'} p_n \qquad (11)$$

It can be mathematically proved [21] that the least square error between real and approximated values is minimized when

$$A = (P^{\psi} P)^{-1} P^{\psi} \psi \qquad (12)$$

Due to temporal changes in system, an experiment is repeated *l* times and then the average performance metric $\hat{\psi}_i^{j}$ is calculated. After finishing all experiments of the application, matrixes *A, P* and ψ are formed from equation (8). Finally, the model coefficients of the application are calculated by equation (12).

To show how the concept of above linear regression approach is practicable in parameter reduction and variable significance analysis in WSN , We find a case in literature [11] how the authors use Residual Energy as the only performance metric of interest and use: *Transmission Interval, Num of hop, Sensing Interval, Sensing Radius, Net Density, Transmission Radius and Num of Sink* as basic design variables, to conduct the range of simulation experiments to model the relationship of residual energy with respect to the most effective parameters. Identification of which parameters more significant and effective are based on their P-value, and then linear correlation is used for both influential variables and the variables direction effect. A subset of experiment data are shown in TABLE 1.

From terms of effective parameters identification, P-value and linear analysis are consistent and supportive each other. Based on P-value of parameters, the Transmission interval, the Number of hops, the sensor interval, and the Sense Radius are the most effective parameters in the experimental context because of their lower P-values than threshold P-value 0.05. The direct relationship between residual energy and transmission interval, sensing interval can be achieved from positive sign of the linear correlation in Table 1.The reverse relationship between the residual energy and the number of hops, sensing radius also can be resulted from negative sign of the linear correlation in table 1.

The experimental results are cohesive to theoretical analysis. For example if the number of hops increases, the residual energy will decrease.

After effective subset parameters are identified, the overall residual energy of a sensor is modeled using linear regression in term of the most effective parameters as:

$$E = a_0 + a_1 Tran\_Interv + a_2 Num\_Hop + a_3 Sens\_Interv + a_4 Sens\_Radius$$

Parameters	P-value	Linear Correlation
Transmission interval	3.7979e-005	0.2842
Num hop	0.00051	-0.2411
Sensor interval	0.02397	0.1580
Sense Radius	0.04933	-0.1355
Net density	0.11896	-0.1095
Transmission Radius	0.32401	-0.0694
Sink	0.42896	-0.0557

Table 1. P-value , linear analysis of effective parameters and Residual Energy [11]

Through this experiment, we can see the approach is practical to identify effective parameters and reduce redundant parameters accordingly. But we also noticed that this approach has an assumption that parameters involved in the evaluation are orthogonal or independent, no interaction effect should be further considered. However this assumption simply cannot be guaranteed from researchers. In next nonlinear model, the Choquet model, the interaction can be further identified.

6.5 The Choquet model: a nonlinear model

We introduce a nonlinear statistical model [15, 17, 18, 19, 20] based on non-additive integrals [16] for interdependency model and significance analysis. The distinguishing feature of this model is that the interaction among system parameters toward the performance metrics of interest can be properly measured through the nonlinear integral such as Choquet integral.

The nonlinear data measurement consists of l observations of *X(n)* and *Ψ*, and have a form as

x_1	x_2	$\cdots$	x_n	Ψ
f_{11}	f_{12}	$\cdots$	f_{1n}	Ψ_1
f_{21}	f_{22}	$\cdots$	f_{2n}	Ψ_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
f_{l1}	f_{l2}	$\cdots$	f_{ln}	Ψ_l

Each row f_j (,j<= l) is observation of system parameters f_{ji} (i <=n) and their corresponding collective system performance output ψ_j.

The interdependency among system parameters toward the performance metrics of interest is described by a set function μ defined on the power set of X

$$\Psi = c + \int f d\mu + N(0, \delta^2),$$

Where *c* is a regression constant, *f* is an observation of x_1, x_2 ,…, x_n , *μ* is a non-additive measure, and $N(0, \delta^2)$ is a normally distributed random perturbation with expectation 0 and variance δ^2. Given observation data, the optimal regression coefficients can be determined by using the least square. The basic idea to solve the Choquet model is a two step procedure.

1).To reduce the non-linear multi regression model to the traditional linear multi regression model by converting each n-dimensional vector attribute datum to a 2^n-dimensional vector datum, which is defined over the power set of attributes;

2). To solve the linear model by using the standard least-square method.

(1) Data Transformation:

The non-additive measure in the Choquet model is defined over the power set P(X). The key point here is the data transformation from n-dimensional *X* to 2^n-dimensional power set over X: P(X).

Let suppose we have a system with a small design space of three parameters, then the attribute set X=*{x*1*, x*2*, x*3*}* , the power set over X be *P(X)* =({x1}, {*x2*}, {*x1,x2*}, {*x3*}, {*x1,x3*}, {*x2, x3*}, {*x1,x2,x3*}). Corresponding to binary coding expression as: *P(x)*=({001},{010},{011},{100},{101},{110},{111}) ; Where k=$k_3 k_2 k_1$ represent combination of parameters *x3, x2*, *x1*.

If k_i = 1 then the corresponding x_i in the combined effective parameter set, otherwise the corresponding x_i not in the combing parameter set. We use μ_k to denote μ(A), where A=$\bigcup_{k_i=1}\{x_i\}$.

x_1	x_2	x_3
3	10	7
-6	-9	-3
7	0	2

Table 2

$\{x_1\}$ K=001	$\{x_2\}$ K=010	$\{x_1, x_2\}$ K=011	$\{x_3\}$ K=100	$\{x_1, x_3\}$ K=101	$\{x_2, x_3\}$. K=110	$\{x_1, x_2, x_3\}$ K=111
0	3	0	0	0	4	3
0	0	0	3	3	0	-9
5	0	0	0	2	0	0

Table 3

Consider a small set of sample data with the attribute set *{x*1*, x*2*, x*3*}* as shown in *Table 2*. The first vector (3*,* 10*,* 7) is aggregated to an equivalent subset-based representation as shown in *table 3*, so that the contribution to the objective can be calculated using the non-additive measure in a linear way:

For $\{x_1\}$, k=001, $\min\limits_{k_i=1}(f_{ji}) - \max\limits_{k_i=0}(f_{ji}) = 3 - 10 < 0$, so $z_{11} = 0$;

For $\{x_2\}$, K=010, $\min\limits_{k_i=1}(f_{ji}) - \max\limits_{k_i=0}(f_{ji}) = 10 - 7 = 3$, so $z_{12} = 3$;

For $\{x_1, x_2\}$, K=011, $\min\limits_{k_i=1}(f_{ji}) - \max\limits_{k_i=0}(f_{ji}) = 3 - 7 < 0$, so $z_{13} = 0$;

For {x_3}, K=100, $\min\limits_{k_i=1}(f_{ji}) - \max\limits_{k_i=0}(f_{ji}) = 7 - 10 < 0$, so $z_{14} = 0$;

For {x_1, x_3}, K=101, $\min\limits_{k_i=1}(f_{ji}) - \max\limits_{k_i=0}(f_{ji}) = 3 - 10 < 0$, so $z_{15} = 0$;

For {x_2, x_3}, K=110, $\min\limits_{k_i=1}(f_{ji}) - \max\limits_{k_i=0}(f_{ji}) = 7 - 3 = 4$, so $z_{16} = 4$;

For {x_1, x_2, x_3}, K=111, $\min\limits_{k_i=1}(f_{ji}) - \max\limits_{k_i=0}(f_{ji}) = 3 - 0 = 3$, so $z_{17} = 3$.

(2) Apply aforementioned procedure to solve the Choquet model to determine $c, \mu_1, \mu_2, ...\mu_{2^n-1}$ as:

Construct the $l \times (2^n)$ augmented matrix $Z = [z_{jk}]$

z_{j0}=1;

$$z_{jk} = \begin{cases} \min\limits_{k_i=1}(f_{ji}) - \max\limits_{k_i=0}(f_{ji}); if > 0 \\ \\ 0; otherwise \end{cases}$$

$$\underbrace{\begin{bmatrix} \psi_i^{(1)} \\ \psi_i^{(2)} \\ ... \\ \psi_i^{(l)} \end{bmatrix}}_{\psi} = \underbrace{\begin{bmatrix} 1, z_{11}...z_{1(2^n-1)} \\ 1, z_{21}...z_{2(2^n-1)} \\ ... \\ 1, z_{l1}...z_{l(2^n-1)} \end{bmatrix}}_{Z} \underbrace{\begin{bmatrix} c \\ \mu_1 \\ ... \\ \mu_{2^n-1} \end{bmatrix}}_{U} \quad \text{or} \quad \boldsymbol{\psi = Z * U}$$

Find the least square solution of the system of linear equations having the above augmented matrix for unknown variables $c, \mu_1, \mu_2, ... \mu_{2^n-1}$, similar as aforementioned linear regression problem.

$$\delta^2 = \frac{1}{l}\sum_{j=1}^{l}(\psi_j - c - \sum_{k=1}^{2^n-1} z_{jk}\mu_k)^2$$

The non-additive model based on Choquet integral has been used in many applications since its first introduction half century ago. Due to its distinguished feature that the interaction among predictive attributes toward the objective attribute can be properly reflected through a set of non-additive measures, recently the model has been introduced into wireless network to tackle the parameters interaction problem, area such as wireless cross layer design, wireless multimedia communication, Wireless LAN performance analysis.

To show the practical uses of the Choquet model in wireless network, we adopt an experiment from [19] as a case study where the Choquet model is applied for interdependency measure and significance analysis on parameter set of MAC layer in IEEE 802.11 for throughput metric of WLANs. The parameter set includes the number of users, the minimum contention window size, MAC-frame size, retransmission times and data rate, etc.

CWmin	Frame Size (Byte)	Data rate (Mbps)	Retransmission times	Throughput (bps)
16	500	1	2	0.445E+06
16	1000	2	4	0.903E+06
16	1500	11	10	3.177E+06
…	…	…	…	…
32	500	1	2	0.438E+06
32	1000	2	4	0.918E+06
32	1500	11	10	3.086E+06
…	…	…	…	…
64	500	1	2	0.451E+06
64	1000	2	4	0.878E+06
64	1500	11	10	3.088E+06
…	…	…	…	…
256	500	1	2	0.432E+06
256	1000	2	4	0.852E+06
256	1500	11	10	2.704E+06

Table. 4 results from [19]

The experimental data sets have been collected by running simulations with different combinations of parameters. To reduce the complexity of the Choquet model, investigations into the relevant analytical and experimental conclusions of variables sensitivity in literature is critical. Four parameters have been implemented as interaction analysis target. They are 1). Minimum Contention_Window size (*cw*), 2). Frame_Size(*fm*), 3). Data_Rate(*dr*) and 4). Retransmission_Times(*r*). Other two parameters are adopted as context parameters, not directly considered in combination effect analysis, but indirectly involved in evaluation of other parameters combination effects as context parameters. They are Channel_States (Bit Eerror Rate) and Number_of_Users.

As shown in Table 4, the minimum contention window size is increased exponentially from 16 to 256, the frame size is chosen from 500 bytes to 1500 bytes with the step size of 250 bytes, the data rates are selected from 1, 2 and

11 Mbps, and the retransmission times are selected from 2 to 10 with the step size of 2.

Besides crafty consideration on the choice of parameters, another critical issue raised from Table 4 is the data normalization. As mentioned in [10], direct reduction on raw data from the Choquet model to linear multi_regression model may cause “bad” solutions, where non-additive measures on some subsets are often not able to be determined due to different underlying scale of measurements**.** Recall how we construct the parameter matrix, a measurement of parameter Data_Rate can be either Mbps as unit or 10^6 bps as unit. There is no difference in essence as we interpret these data, but for matrix construction algorithm we described before, it’s a serious problem**.** To avoid observed data of one parameter is doomed bigger or smaller than another parameter due to underlying scale difference, a normalization process is scheduled before matrix construction step such that,

$$f_{ji}^{'} \leftarrow {}^{f_{ji}}\!\big/_{median(x_i)}.$$

Using non-additive measure theory, we follow the process step by step, each possible combination out of $(2^4 - 1)$ will generates a value of significance measure for each context setup, as shown in Table 4. We get conditional conclusion as:

- When the network is lightly loaded, whether the channel quality is good or bad, μ[14] = 3.0609 or 1.9730 has the largest absolute value with positive sign indicates that the interdependency between frame size, data rate and retransmission times have the most significant effect on the throughput performance.

- When the network is heavily loaded, whether the channel quality is good or bad, $\mu[7]$ = 2.4404 or 1.3275 has the largest absolute value with positive sign indicates that the interdependency between frame size, minimum contention window size and data rate have the most significant effect on the throughput performance.

μ	Parameter subset	Number of users=10		Number of users=50	
		BER=10^{-5}	BER=10^{-3}	BER=10^{-5}	BER=10^{-3}
$\mu[1]$	$\{cw\}$	-0.0229	-0.0272	0.0331	-0.0012
$\mu[2]$	$\{fm\}$	0.3437	0.2358	-0.3390	-0.1278
$\mu[3]$	$\{cw, fm\}$	0.3721	0.2647	0.0193	0.0929
$\mu[4]$	$\{dr\}$	0.7644	0.4961	0.2424	0.1547
$\mu[5]$	$\{dr, cw\}$	0.5862	0.3756	0.6372	0.3326
$\mu[6]$	$\{dr, fm\}$	3.0575	1.9643	1.6536	1.0212
$\mu[7]$	$\{dr, fm, cw\}$	3.0334	1.8538	**2.4404**	**1.3275**
$\mu[8]$	$\{r\}$	0.0120	-0.0216	-0.4497	-0.1973
$\mu[9]$	$\{r, cw\}$	0.0209	0.0178	-0.0285	-0.0085
$\mu[10]$	$\{fm, r\}$	0.2998	0.2231	-0.1841	-0.0461
$\mu[11]$	$\{r, fm, cw\}$	0.3538	0.2345	0.1100	0.1736
$\mu[12]$	$\{dr, r\}$	0.7358	0.4811	-0.1272	-0.2542
$\mu[13]$	$\{dr, r, cw\}$	0.5715	0.3359	0.6385	0.2952
$\mu[14]$	$\{dr, r, fm\}$	**3.0609**	**1.9730**	1.2545	1.1846
$\mu[15]$	$\{dr, r ,fm, cw\}$	3.0217	1.8546	2.3174	1.2678

Table 5.resluts from [19]

6.6 Comparison of the two models: Linear regression model and Choquet integral model.

The rationale behind the two model has one common point, both of them apply linear regression and least square methodology at the last step to find the parameter significance relevant to performance metric. The advantage of linear regression model is computation simplicity. N parameters only have n-dimensional linear relationship to solve. But the problem is that, we have to assume the parameters involved are orthogonal between each other, which is intolerable for wireless network with many issues are still in exploring stage with uncertainty. On the other hand, although Choquet integral model succeeds in parameters interaction analysis, practically, the design space has to be reduced before applying the model, due to n parameters nonlinear aggression problem has to transform to 2^n-dimensional linear regression problem. The computation cost is too high to ignore.

For a complex system, we believe a preliminary investigation into certain parameters individual contribution to performance metric with only linear regression solution is beneficial. Then we can utilize Choquet integral model with reduce parameters design space to further investigate the interdependency between parameters and combining effects to final performance. We are here to show the feasibility of our evaluation methodology, i.e., how to reduce the parameters to an effective parameter set effectively and efficiently. This is an important and serious ongoing research topic in wireless network in general and particularly in WSN. As more and more efficient mathematical models are developed and applied to such problem space, the evaluation methodology will become more powerful tool to tackle the WSN evaluation problem.

CHAPTER 7

CONCLUSION AND FUTURE PROSPECT

From historic viewpoint, WSNs related technologies will keep evolving as past history indicated. 1). More efficient CPU and memory chip will be available with low power requirements in a small size according Moore's Law. 2). Power source improvements in batteries, as well as passive power sources such as solar or vibration energy, are expanding application options. 3). Research in Materials Science has resulted in novel sensing materials with lower energy cost and higher reliability. 4). Transceivers for wireless devices are becoming smaller, less expensive, and less power hungry. Nonetheless, the enabling technologies are evolving at different paces as indicated in [43]. Comparing the progress of computing with battery and wireless transceiver technologies, computing resources (CPU and memory) are becoming easily available (more powerful, cheaper, smaller), while battery and transceiver capacity improves slower. This phenomenon has big impact on futuristic architecture design and evaluation from triangular constraint tradeoffs perspective. We can change the importance of subcomponents (CPU, Memory, and Battery) of the cost indicator to reflect the nature of the contribution of each of these subcomponents as: relaxed, constrained, or bottleneck. Hence, we can use more relaxed resources to gain more performance improvement and degree of generality while avoiding the energy hungry or bottleneck resources like transmitting power. In this context, we can foresee more complex algorithms can be handled by WSNs computing facilities in near future for specific application domain. As enabling technologies and commercialization progress, generality eventually will regain its importance. Architecture evaluation should follow the technical trends to reflect the dynamic nature of technology development. Our evaluation framework is open enough to tolerate enabling technology evolvement.

In this work, we analyze the elements contributing to stochastic natured performance at chapter 3. We further introduce new metrics to reflect the randomness in system evaluation. To combat the dilemma of multi objective decision making in application specific design context, a single performance index algorithm is established to balance the importance of application specific requirement and performance measurement result. In chapter 4, we develop a qualitative evaluation framework to serve as a master key to unlock the myth of multifaceted WSNs applications performance. This framework do not evaluate performance only from application QoS performance perspective, more importantly, also from SDLC system development perspective. In chapter 5, a practical performance benchmarking workflow and relevant algorithm are presented as case study of effort to quantify chapter 4 qualitative concepts. Parameters interaction and design space problem are identified in this chapter. To make experiment operation practical, in chapter 6, we explore the ways to reduce the evaluation space by parameter significance analysis. Linear and nonlinear models are introduced and compared to each other.

WSNs are unique in terms of its node volume, wireless channel, resource constraint and vast application potentials. Uncertainty attribute of WSNs performance still remains a difficult problem to address in future. Especially quantitative mapping of decisions to performance and cost metrics is difficult to handle. There needs further theoretical advance to give more insight into the fundamental collective properties of stochastic wireless network. As Gnedenko and Kolmogorov in their classical work on limit distributions [49] described, "the epistemological value of the theory of probability is based on this: that large-scale random phenomena in their collective action create strict, nonrandom regularity".

We have illustrated how to get the single evaluation index for QoS performance. The other two elements in the triangular constraints model, cost and generality indexing needs to be carried out as our future research. As a qualitative model it is efficient. But if we can precisely model the cost relating to specific attributes, like generality cost and development efficiency, it will be helpful to system designer for evaluation and comparison of alternative solutions.

References

[1]. S. N. Diggavi, N. Al-Dhahir, A. Stamoulis, and A. R. Calderbank, “Great expectations: The value of spatial diversity in wireless networks,” *Proc. IEEE*, vol. 92, no. 2, pp. 219–270, Feb. 2004.

[2]. Zuniga, M. and Krishnamachari, B.. 2004. “Analyzing the transitional region in low power wireless links”. In Proceedings of the IEEE 1st Annual Conference on Sensor and Ad Hoc Communications and Networks (SECON). 517–526.

[3]. Zuniga, M. and Krishnamachari, B. 2007. “An analysis of unreliability and asymmetry in low-power wireless links”. ACM Trans. Sen. Netw. 3, 2, 7.

[4]. M. Mamun, T. Hasan-Al-Mahmud, S. Debnath, and M. Islam, “Analyzing the Low Power Wireless Links for Wireless Sensor Networks,” Journal of Telecommunications, vol.1, issue.1, pp.123-127, 2010

[5]. Baccour, N.; Aa, A.K.; Mottola, L.; Youssef, H.; Boano, C.A.; Ario, M. “Radio Link Quality Estimation in Wireless Sensor Networks: A Survey”. ACM Trans. Sens. Netw. 2012, submitted.

[6]. The Ruckus white paper, available at :

http://www.ruckuswireless.com/whitepapers/preview/wireless-network-performance

[7]. J. Haapola, F. Martelli, and C. Pomalaza-R´aez, “Application-driven analytic toolbox for wsns,” in 8th International Conference on Ad-Hoc Networks and Wireless (ADHOC-NOW), LNCS 5793, pp. pp. 112 – 125, September 2009.

[8]. V. Czitrom, “One-factor-at-a-time versus designed experiments,” *The American Statistician*, vol. 53, pp. 126–131, May 1999.

[9]. L. Wang and Y. Xiao, “A survey of energy-efficient scheduling mechanisms in sensor networks," Mobile Network Applications, vol. 11, no. 5, pp. 723-740, 2006.

[10]. D. Yu, P. Nanda, & R. Braun. 2011, ‘Credibility Problems and Tradeoffs between Realistic and Abstraction in WANET and WSN Simulation’, The 7th International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM2011), 23-25 September 2011, Wuhan, China

[11]. N. Kamyabpour and D. B. Hoang, "A study on Modeling of Dependency between Configuration Parameters and Overall Energy Consumption in Wireless Sensor Network (WSN)," CoRR, 2011.

[12]. N. B. Rizvandi, et al., "On Modeling Dependency between MapReduce Configuration Parameters and Total Execution Time," CoRR, 2011.

[13]. N. B. Rizvandi, et al., "An Accurate Fir Approximation of Ideal Fractional Delay Filter with Complex Coefficients in Hilbert Space," Journal of Circuits, Systems, and Computers, vol. 14, pp. 497-506, 2005.

[14]. N. B. Rizvandi, et al., "Preliminary Results on Modeling CPU Utilization of MapReduce Programs," School of Information Technologies, University of Sydney, Sydney2010.

[15]. Z. Wang, “A new model of nonlinear multiregressions by projection pursuit based on generalized choquet integrals,” in IEEE International Conference on Fuzzy Systems, 2002, pp. 1240–1244.

[16]. D. Denneberg, Non-Additive Measure and Integral. Kluwer Academic Publisher, 1994.

[17]. S. Ci and H. Guo, “Quantitative Dynamic Interdependency Measure and Significance Analysis for Cross-Layer Design under Uncertainty,” IEEE ICCCN, 2007.

[18]. S. Ci and H.-F Guo, “Significance measure with nonlinear and incommensurable observations,” in Proceedings of IEEE Global

Communications Conference (GLOBECOM '07), Washington, DC, USA, November 2007.

[19]. R. Chen, W. Wang, Z. Mi, "Interdependency Measure and Analysis for Cross-Layer Design in WLAN", IJACT: International Journal of Advancements in Computing Technology, Vol. 3, No. 11, pp. 238 ~ 245, 2011

[20]. H. Guo, W. Zheng, S. Ci, “Precise Determination of Non-additive Measures”, In Proceeding(s) of IEEE International Conference on Cybernetics and Intelligent Systems (CIS), Chengdu, China, pp.911-916, June 2008.

[21]. R. Hoes, T. Basten, C.-K. Tham, M. Geilen, H. Corporaal, “Analysing QoS trade-offs in wireless sensor networks”, in: 10th ACM Symposium on Modeling, Analysis, and Simulation of Wireless and Mobile Systems (MSWiM), Proc., ACM Press, New York, NY, USA, 2007, pp. 60.

[22]. T. Saaty, *The Analytic Hierarchy Process*. New York: McGraw-Hill, 1980.

[23]. Rehman S., Turletti T., Dabbous W. “A Roadmap for Benchmarking in Wireless Networks”(2011*).* INRIA Technical Report [inria-00614167, version 1], available: http://hal.inria.fr/inria-00614167

[24]. J. Gerwen, S. Bouckaert, I. Moerman, P. Demeester. " Benchmarking for Wireless Sensor Networks," *SENSORCOMM '11. The Fifth International*

Conference on Sensor Technologies and Applications, vol., no., pp. 134-139, August 21-27, 2011

[25]. Rehman S., Turletti T., Dabbous W. “Benchmarking in Wireless Networks” (2010). INRIA Technical Report [inria-00530329 - version 1] available: http://hal.inria.fr/inria-00530329

[26]. CREW project, “Home page,” http://www.crew-project.eu/

[27]. BonFIRE project, “Home page,” http://www.bonfireproject.eu/

[28]. OneLab, “Home page,” http://www.onelab.eu/

[29]. M. Hempstead, M. Welsh, and D. Brooks, “Tinybench: The case for a standardized benchmark suite for tinyos based wireless sensor network devices.” LCN, pp. 585–586, 2004.

[30]. T. Kim, J. Kim, S. Lee, I. Ahn, M. Song, and K. Won, “An automatic protocol verification framework for the development of wireless sensor networks,” in TridentCom, 2008.

[31]. P. Gunningberg, M. Bjorkman, E. Nordmark, S. Pink, P. Sjodin, and J.-E. Stromquist, “Application protocols and performance benchmarks,” IEEE Communications Magazine, vol. 27, no. 6, pp. 30 –36, jun. 1989.

[32]. G MARTINOVIC, J BALEN, D ZAGAR “Performance Evaluation Model of Optimization Methods for Wireless Sensor Networks”, *WSEAS Transactions On Communications,* Issue 10, Volume 8, October 2009 ISSN 1109-2742

[33]. Maximilian Ott, Ivan Seskar, Robert Siraccusa, Manpreet Singh, "Orbit Testbed Software Architecture: Supporting Experiments as a service," Testbeds and Research Infrastructures for the DEvelopment of NeTworks and COMmunities (Tridentcom), Feb 23-25 2005, Trento, Italy.

[34]. Wolfgang Kiess, "On Real-world Experiments With Wireless Multihop Networks," PhD dissertation, 2008.

[35]. University of Utah Flux Research Group, "Emulab: The Utah Network Emulation Testbed", http://www.emulab.net/

[36]. H. Lundgren, D. Lundberg, J. Nielsen, E. Nordström, and C. Tschudin, "A large-scale testbed for reproducible Ad Hoc protocol evaluations," In Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC), pages 337-343, March 2002.

[37]. P. De, A. Raniwala, S. Sharma, and T. Chiueh, "Mint: a miniaturized network testbed for mobile wireless research," INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE, vol. 4, pp. 2731-2742 vol. 4, March 2005.

[38]. J. Vanhie-Van Gerwen, E. De Poorter, B. Latre, I. Moerman and P. Demeester, “Real-life performance of protocol combinations for Wireless Sensor Networks”, The Third IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing (SUTC2010), Newport Beach, California, USA, June 7-9, 2010

[39]. B. V. Gnedenko, A. N. Kolmogorov, Limit distributions for sums of independent random variables. Trans. and annotated by K. L. Chung. Cambridge, Addison-Wesley, 1954. 264+9 pp.

[40]. M. Ali, T. Voigt, U. Saif, K. Rmer, A. Dunkels, K. Langendoen, J. Polastre, and Z. A. Uzmi, “Medium access control issues in sensor networks,” SIGCOMM 2006, pp. 33–36, 2006.

[41]. G. Barrenetxea, F. Ingelrest, G. Schaefer, and M. Vetterli, “The Hitchhiker’s Guide to Successful Wireless Sensor Network Deployments.” in SenSys, 2008, pp. 43–56.

[42]. Polastre, J., Hui, J., Levis, P., Zhao, J., Culler, D., Shenker, S., and Stoica, I. 2005. A unifying link abstraction for wireless sensor networks. 3rd international Conference on Embedded Networked Sensor Systems, USA, November 02 - 04, 2005

[43]. Paradiso, J.A. and Starner, T.,Energy Scavenging for Mobile and Wireless Electronics, IEEE Pervasive Computing, Vol. 4, No. 1, February 2005, pp. 18-27.

[44]. G. Zakharia, “Architectural Efficiency” Proceeding of the Computer Management Group CMG’10, Orlando, FL, 2010

[45]. Saaty, T.: The Analytic Hierarchy Process. McGraw-Hill, New York (1980)

[46]. L. Zhu, A. Aurum, I. Gorton, and R. Jeffery. Tradeoff and sensitivity analysis in software architecture evaluation using analytic hierarchy process. Software Quality Control, 13(4):357–375, 2005.

[47]. McCall, J. A., Richards, P. K., and Walters, G. F., "Factors in Software Quality", Nat'l Tech.Information Service, no. Vol. 1, 2 and 3, 1977.

[48]. J. Yick, B. Mukherjee, and D. Ghosal. Wireless sensor network survey. Computer Networks, 52(12), 2008.

[49]. B. V. Gnedenko and A. N. Kolmogorov *Limit distributions f or sums of independent random variables.* . Trans, and annotated by K. L. Chung. Cambridge, Addison-Wesley, 1954. 264+9 pp.

[50]. R. Prasad, “A perspective of layerless communications,” *Wireless Personal Communications,* vol. 44, no. 1, pp. 95–100, 2008

[51]. Vikas Kawadia and P. R. Kumar, “A Cautionary Perspective on Cross-Layer Design”, IEEE wireless communication, 2005

[52]. Y. Gu and T. He, “Data forwarding in extremely low dutycycle sensor networks with unreliable communication links.” in SenSys ’07, 2007, pp. 321–334.

[53]. M. Buettner, G. V. Yee, E. Anderson, and R. Han, “X-mac: a short preamble mac protocol for duty-cycled wireless sensor networks.” in SenSys ’06, 2006, pp. 307–320.

[54]. R. Fonseca, O. Gnawali, K. Jamieson, S. Kim, P. Levis, and A. Woo, “The collection tree protocol (CTP).”
http://www.tinyos.net/tinyos-2.x/doc/html/tep123.html.

[55]. R. Thouvenin, “Implementing and evaluating the dynamic manet on-demand protocol in wireless sensor networks.” Master’s thesis, University of Aarhus, 2007.

[56]. E. Osipov, “tinylunar: One-byte multihop communications through hybrid routing in wireless sensor networks.” In NEW2AN, vol. 4712. Springer, 2007, pp. 379–392.

[57]. C. J. Sreenan, S. Nawaz, T. D. Le, and S. Jha, “On the sensitivity of sensor network simulations.” in VTC Spring. IEEE, 2006, pp. 1043–1047.

[58]. B.N. Bhandari, R.V.R. Kumar, R. Banjari, and S.L. Maskara, “Sensitivity of the IEEE 802.11b MAC Protocol Performance to the Various Protocol Parameters”, *ICCCAS 2004, International Conference on Communications, Circuits and Systems*, 27-29 June 2004, Volume 1, Pages: 359-363.

[59]. F. Tirkawi, S. Fischer, “Generality Challenges and Approaches in WSNs”. IJCNS 2(1): 58-63 , 2009.

[60]. W. Dargie, "Qualitative Evaluation of Cross-Layer Approaches in Wireless Sensor Networks," *Proceedings of 19th International Conference on Computer Communications and Networks (ICCCN)*, pp.1-6, 2-5 Aug. 2010.

[61]. S. Tong**,** **“**An Evaluation Framework for middleware approaches on Wireless Sensor Networks”. In: Seminar on Internetworking, Helsinki University of Technology, Finland, April 27th 2009.

[62]. L. Yuan, X. Wang, "Study on performance evaluation method based on measurement for wireless sensor network,". *ICCTA '09.* pp.201-206, 16-18 Oct. 2009.