Delft University of Technology
M.Sc. Thesis in Embedded Systems

Do we really need *A Priori* Link Quality Estimation?

Vasileios S. Vasilopoulos

embedded software

TU Delft Delft University of Technology

Do we really need *A Priori* Link Quality Estimation?

M.Sc. Thesis in Embedded Systems

Embedded Software Group
Faculty of Electrical Engineering, Mathematics and Computer Science
Delft University of Technology
Mekelweg 4, 2628 CD Delft, The Netherlands

Vasileios S. Vasilopoulos
vvasilop@gmail.com

August 5^{th}, 2013

Author
Vasileios S. Vasilopoulos (vvasilop@gmail.com)
Title
Do we really need *A Priori* Link Quality Estimation?
M.Sc. Presentation
August 13^{th}, 2013

Graduation Committee

Prof. Dr. K. G. Langendoen (Chair)	Delft University of Technology
M. A. Zúñiga Zamalloa, Ph.D. (Daily Supervisor)	Delft University of Technology
Dr. Z. Al-Ars (Core Member)	Delft University of Technology

Abstract

Traditionally, link quality estimation (LQE) has been viewed as an *a priori* step in sensor network routing protocols because it filters out unreliable links before data transmission. Recent results, however, show that protocols can perform well without *a priori* LQE. Because getting rid of LQE seems rather counter-intuitive, the aim of this work is to look deeper into the behavior of LQE-free protocols. Our results, based on one of the state-of-the-art LQE-free protocols, show two interesting insights. First, LQE-free protocols manage to choose links that are slightly better than the ones obtained with *a priori* LQE methods. Second, in traditional protocols, the effort needed to identify good links accounts for 40% to 60% of the energy consumption of nodes. By eliminating this overhead, LQE-free protocols can save a significant amount of energy compared to standard approaches. On the other hand, clear downsides of LQE-free operation are the longer paths and the worse load balance. The latter poses a higher transmission burden on some nodes.

Keywords: Link Quality Estimation, Data Collection, Routing, Collection Tree Protocol, Broadcast-Free Collection Protocol, Wireless Sensor Networks

"Knowledge is power." – Sir Francis Bacon

Preface

This Master of Science thesis brings my student life in Delft and in general to an end. I started my graduation project as an independent research endeavor at the Embedded Software Group roughly nine months ago with a great enthusiasm and a passion for achieving something novel in terms of research value. I decided to work on the promising topic of collection protocols in the area of Wireless Sensor Networks (WSNs). By the time I began working on this project, my interaction with WSNs was not that new. In fact, I was initiated into the area of WSNs in late 2007 in the context of my undergraduate work at the University of Patras in Greece. In the end, I consider myself extremely lucky because I had the unique chance to gain meaningful insight into how state-of-the-art research is conducted in an inspiring academic setting including all the stages from formulating a research proposal to validating it through concrete empirical results.

Each and every achievement in life is the outcome of a joint effort, either explicitly or implicitly. This thesis, along with the preceding work, is no exception. I am particularly grateful to Marco Zúñiga and Daniele Puccinelli whose truly positive attitude and trust helped me make the most of my graduation project. Through my interaction with Marco specifically, I have started learning not only what critical thinking is all about but also a few abstract lessons which will certainly help me improve myself throughout my life. I also want to acknowledge Koen Langendoen for hosting me at his group. What is more, his insightful comments and suggestions helped me improve the quality of this thesis. Furthermore, I want to acknowledge Zaid Al-Ars for joining my graduation committee. A special thank you goes to all the friends in the Netherlands and in Greece. Finally, I am so grateful to my beloved family consisting of my father Spyros, my mother Alexandra and my sisters Konstantina and Eleanna for their limitless support, both financial and emotional. Needless to say, this thesis is dedicated to them.

Vasilis Vasilopoulos

Delft, The Netherlands
August 5th, 2013

Contents

Preface **vii**

1 Introduction **1**
- 1.1 Problem Statement 2
- 1.2 Contributions 2
- 1.3 Novelty 3
- 1.4 Thesis Organization 3

2 Related Work **5**
- 2.1 Link Quality Estimation in WSNs 5
- 2.2 The Collection Tree Protocol 6
- 2.3 State-of-the-art Collection Protocols 7
- 2.4 The Broadcast-Free Collection Protocol 9

3 Evaluation Framework **11**
- 3.1 Methodology 11
- 3.2 Implementation Details 12

4 Empirical Results and Analysis **15**
- 4.1 Routing Structure Analysis 15
  - 4.1.1 Link Quality 15
  - 4.1.2 Path Length 18
  - 4.1.3 Link Quality vs. Path Length 18
- 4.2 Operational Analysis 19
  - 4.2.1 Energy Efficiency 20
  - 4.2.2 Load Balancing 22
  - 4.2.3 Energy Efficiency vs. Load Balancing 24

5 Conclusions **29**

Bibliography **31**

Chapter 1

Introduction

Wireless sensor networks (WSNs [1]) emerged roughly fifteen years ago, representing a unique class of computing with large numbers of resource-constrained nodes necessarily cooperating on a single application. Several design goals are inherently linked with WSNs, notably including, but not limited to, reliability and energy efficiency. Data collection is a common application in the field of WSNs. More specifically, a network deployment formed by a number of sensor nodes performing a task has to forward its data to a central station, most commonly referred to as the sink. The goal for the nodes is to discover the routes to the sink. The aforementioned requirement for energy efficiency (following from the limited energy resources of the nodes) has substantiated the need for low-power sensor networks. As a natural consequence, energy-saving techniques and appropriate protocol designs have been an important topic of research.

Traditional collection protocols in sensor networks, such as CTP [11], Arbutus [19] and MintRoute [21], employ a two step approach to build their routing structure. First, nodes perform link quality estimation (LQE) to get a sense of all the possible usable links. Then, each node forms a routing path on top of these usable links (according to a given metric). Performing link quality estimation makes sense because unreliable links are commonplace in sensor networks (due to the use of low-cost radios) and it is important to filter them out to increase reliability.

In the last couple of years, a new generation of collection protocols [9, 12, 15, 20] has challenged the LQE-based approach, and they have achieved better energy efficiency than CTP-like approaches. While all these new protocols have their own distinctive features, they share a common characteristic: they do not perform *a priori* link quality estimation. This observation motivated us to look deeper into the (in)significance of link quality estimation for sensor networks.

1.1 Problem Statement

The most important task of sensor networks is the ability to deliver data. Link quality and path length play a central role in this respect: the better the link and the shorter the path, the higher the delivery rate. But, if no *a priori* link quality estimation is performed, several important questions arise. Without LQE, how good (or bad) are the selected links? What is the impact on the path length? The new generation of collection protocols reports delivery rates that are comparable or higher than CTP; therefore, they must be utilizing good links. Does this imply that *a priori* LQE is not necessary?

Furthermore, since link quality estimation typically comes with a sizable control traffic overhead and, therefore, a significant energy footprint, it is important to gain insight into the energy demands of LQE methods (since the elimination of these extra-energy demands may be the core reason why the new generation of protocols outperforms the older one).

1.2 Contributions

This work is part of a wider research effort aiming at understanding LQE-free protocols and has two overarching goals. First, to understand the impact of eliminating *a priori* LQE on the underlying routing structure (i.e. analyzing the resulting link quality and path length). Second, to quantify the impact of LQE methods on the network's energy consumption (by analyzing the nodes' duty cycle and load balance).

We selected the Broadcast-Free Collection Protocol (BFC [20]) as an initial step to delve into the questions posed in Section 1.1 and to meet the aforementioned goals. Overall, our empirical work captures several interesting insights:

- ⋆ LQE-free protocols seem to select links that are as good as, or even slightly better than those chosen by *a priori* LQE methods. While our results are based on BFC, we conjecture that this behavior also holds for other protocols (at least in some scenarios).

- ⋆ The paths of LQE-free protocols tend to be longer than those found in LQE-based protocols.

- ⋆ We analyze the energy budget of CTP, dividing it between the parts used for control traffic (mainly LQE) and data transmissions. Our results show that the control traffic part uses between 40% and 60% of the energy resources of nodes. Quantifying this overhead is central to understand the (in)efficiency of methods aiming at identifying good links.

⋆ Depending on the method, LQE-free protocols may lead to a better or worse load balance. In the case of BFC, the load balance and, therefore, the energy balance is worse.

1.3 Novelty

This study is the first to identify that many of the newly proposed collection protocols have a common characteristic: the lack of LQE. Considering that we use BFC as a stepping stone to understand LQE-free protocols, it is important to highlight how our work differs from the study presented by Puccinelli et al. [20]. In that work, the authors propose BFC, provide an analytical model for the duty cycle of different types of nodes and compare BFC with CTP with respect to energy consumption, delivery rate and latency. However, their evaluation results do not look into the routing structure resulting from the lack of LQE methods. Also, while the authors do provide a thorough evaluation of the energy consumption, they do not compare it with a dissected energy budget from CTP. This dissection is important because, as it will be explained later (see Section 4.2.1), it allows us to compare LQE-free protocols with the "LQE-free" version of CTP which is obtained if the period of the Trickle timer in CTP increases to infinity (see Section 2.2 for further information on CTP). What is more, the above-mentioned dissection allowed us to trace a bug in the BFC code that was penalizing BFC's energy performance substantially.

Finally, while the authors of [20] do mention that BFC may result in unbalanced networks, we provide data that supports their hypothesis. Overall, our goal is to better understand the properties of LQE-free operations using BFC as a case study of LQE-free operation and CTP as a case study of LQE-based operation.

1.4 Thesis Organization

The remainder of this thesis is organized as follows. First, we discuss the related work in the field in Chapter 2. Next, we outline our evaluation methodology and highlight a few important implementation details in Chapter 3. We then present the empirical results and convey the insights that can be derived from them in Chapter 4. Finally, we provide our concluding remarks in Chapter 5.

Chapter 2

Related Work

This chapter discusses the related work. We begin with a discussion on the importance of LQE in the field of WSNs and some background information on LQE (see Section 2.1). Next, we report a number of important design features with respect to the de facto standard LQE-based protocol for WSNs (see Section 2.2). We then provide an overview of the recent literature on LQE-free collection protocols and argue over the protocols we chose to validate our research proposal (see Section 2.3). Finally, we present a number of important design features with respect to the LQE-free protocol we chose for comparison (see Section 2.4).

2.1 Link Quality Estimation in WSNs

For roughly a decade now, it has been widely accepted that low-power links, such as those found in WSNs, are unreliable, mainly owing to the following three factors:

- ⋆ The *surrounding environment* leading to multipath propagation effects and contributing to background noise.
- ⋆ The *interference* resulting from concurrent transmissions within a wireless network or between cohabiting wireless networks and other electromagnetic sources.
- ⋆ The *hardware transceivers*, especially those employed in WSNs (low-cost), make radiated signals more susceptible to noise, interference, and multipath distortion.

As a natural consequence, low-power links are inherently governed by dynamics in time and space. In time, links go up and down, while, in space, they do not conform to the ideal unit disk (transmission radii) model.

The aforementioned facts have motivated the need for assessing link quality prior to establishing the routing paths. The fundamental role that link quality plays has rendered LQE an active area of research within the WSN community, subsequently leading to a significant body of work in the related field. Baccour et al. have conducted a comprehensive survey on link quality estimators targeting WSNs [3]. More specifically, they have presented LQE as a three-step procedure, including link monitoring, link measurements, and metric evaluation. The way the first step is performed is what really distinguishes one estimator from another. Bacour et al. identify three types of link monitoring, namely active, passive, and hybrid.

In the context of active monitoring, a node evaluates the links to its neighbors by sending probe packets, either in a broadcast fashion or in a unicast one. Probe packets are generally sent at a certain rate, thus yielding a trade-off between energy efficiency (low rates) and accuracy (high rates). Estimators with higher sampling rates tend to be more accurate (employing a large number of samples) and, therefore, less energy-efficient than estimators with lower sampling rates.

On the other hand, estimators using passive monitoring exploit existing traffic without incurring additional communication overhead. In accordance with this scheme, nodes heavily rely upon overhearing transmitted packets that are not addressed to them. The estimators of this category tend to be more energy-efficient when used in duty-cycled WSNs. However, when the network operates at low data rate or unbalanced traffic, which is, more often than not, the case in WSNs, passive link monitoring may potentially lead to a lack of up-to-date link measurements, thus posing some inaccuracy.

Finally, hybrid link monitoring is, in principle, a combination of active and passive monitoring ultimately leading to a better balance between up-to-date link quality measurements and routing (network layer) advertisements. CTP can be regarded as a collection protocol employing hybrid link monitoring, as we will elucidate in Section 2.2.

Overall, a link quality estimator is an efficient one when it is energy-efficient, accurate, reactive and stable [3]. However, following from the above-mentioned discussion, these four criteria cause emerging trade-offs. In fact, energy efficiency is at odds with accuracy and the same holds for reactivity and stability.

We will now proceed to outline the key mechanisms behind the CTP design.

2.2 The Collection Tree Protocol

CTP [11] employs *a priori* LQE and a mechanism to identify short paths with the aim of discovering and subsequently maintaining the routes to the sink. The following three mechanisms are at the heart of CTP:

- ⋆ LQE: The four-bit LQE scheme [10] is responsible for assessing the quality of the links.

- ⋆ Adaptive Beaconing: The nodes in the network broadcast control beacons at non-fixed intervals in order to discover their neighbors. To this end, CTP designers use a modified version of the Trickle Timer [13] to fit their purposes. More specifically, the nodes send beacons every 64 ms and the inter-beacon interval grows exponentially until it reaches its maximum value in the steady state, roughly 8.6 minutes. With this approach, each and every node is given the chance to join a short-path route to the sink. In the event of any network change, the Trickle Timer is reset and any node with erratic connectivity or even faulty behavior is forced to identify a route to the sink.
- ⋆ Data Path Validation: The nodes keep track of the route cost by making use of Expected Number of Transmissions (ETX [6]) as the cost metric. In a nutshell, the nodes use data packets in order to discover and fix routing inconsistencies, i.e. erroneous paths and/or routing loops, rapidly.

The main drawback of the overall CTP design is the high overhead posed by the mechanisms to maintain short paths to the sink. We will now proceed to elaborate on a new generation of collection protocols that foregoes *a priori* LQE in spite of its, so far, proven significance for low-power WSNs.

2.3 State-of-the-art Collection Protocols

In recent years, several collection protocols have been proposed that do not rely on LQE as heavily as in previous research efforts, which treated it as an indispensable building block. Before describing the related work in detail, it is important to highlight that this new generation of LQE-free protocols provides the same delivery rate as CTP and most of these new protocols do so by consuming less energy.

The summary of the state-of-the-art collection protocols is captured in Table 2.1. For the duty cycle we do not provide the actual numbers, but rather a relative metric (since most of these papers used different evaluation scenarios). For example, the value of 0.4 for ORW indicates that the duty cycle of this protocol is on average 60% less than in CTP (as reported by the authors).

	BCP	LWB	ORW	BFC	CTP
relative duty cycle (%)	>5	0.34	0.4	0.53	1
delivery rate (%)	>97	>99	>99	>99	>99

Table 2.1: Performance of recent collection protocols.

The Backpressure Collection Protocol (BCP) [15] employs a data-driven approach and forwards packets to the neighbor with the lowest queue level.

It is a notable example of structure-free routing, but it is not a great fit for duty-cycled sensor networks because it requires a significant level of offered load to be effective, and this increases the energy consumption of nodes (see Table 2.1).

The Low-Power Wireless Bus (LWB) [9] describes a way to perform data dissemination without utilizing LQE. LWB effectively turns a multi-hop low-power wireless network into a shared bus, where all nodes are potential receivers of all data. LWB leverages constructive interference to perform ultra-cheap and ultra-light network wide floods. The main limitation of this protocol with respect to CTP is that it is centralized; one node in the network is responsible for synchronizing the fast floods of all the other nodes.

The aforementioned protocols are great examples of new radical ways to disseminate and collect data without *a priori* LQE, but their mechanisms are fundamentally different from CTP. BCP requires an artificially high offered load to bootstrap the system and LWB is centralized. To have a good (and fair) insight into the benefits of LQE-free approaches, we believe that it is better to first use protocols that, in spirit, are closer to CTP. Below we describe two of such protocols.

In Opportunistic Routing in Wireless Sensor Networks (ORW) [12], the sink first performs floods to form a gradient centered around itself. The nodes then use opportunistic unicast transmissions to forward data to the first neighbor that wakes up and that provides progress towards the sink (i.e. a node with a higher gradient). ORW does not rely on *a-priori* LQE methods except for one instance, when nodes can not find a parent. The authors describe their approach as a "coarse-grained flavor of LQE".

The Broadcast-Free Collection protocol (BFC) [20] completely avoids LQE while building a data collection tree. BFC relies on eavesdropping nodes that already have a path towards the sink. The protocol is bootstrapped by the sink neighbors, which upon transmitting their first data packets get direct acks from the sink. The eavesdropping process is repeated by nodes further down the tree until all nodes get a path. If a node is unable to eavesdrop on a potential parent, it remains disconnected until a potential parent appears in its vicinity.

Selecting the best candidate for comparison. As mentioned in Section 1.2, this work describes our first effort aimed at understanding LQE-free protocols. As a starting point we required a protocol that resembles CTP as much as possible but that requires no *a-priori* LQE. ORW and BFC are similar to CTP, but ORW still performs some level of LQE (albeit minimal). To avoid obtaining wrong insights, such as getting LQE-help from ORW when a node cannot find a parent, we decided to use BFC as a stepping stone.

We will now proceed to outline the key mechanisms behind the BFC design.

2.4 The Broadcast-Free Collection Protocol

As mentioned in Section 2.3, BFC does not make use of *a priori* LQE in the process of building the data collection tree. Instead, the following three mechanisms are inherently linked with BFC, rendering the protocol itself purely LQE-free:

- ⋆ On Demand Snooping: At start-up, the goal is for the sink's neighbors, i.e. the nodes that are one hop away from the sink, to discover the sink. The rest of the nodes enter a de facto hibernation state until they snoop on unicast transmissions from a potential parent, i.e. when the on-time of a hibernating node matches (part of) the on-time of a potential parent. It should be highlighted that we focus on duty-cycled sensor networks whereby the nodes, and therefore their radio, are not always on. Low Power Listening (LPL [17]) is a very popular (energy-saving) mechanism for link layer duty-cycling. LPL aims at reducing the on-time of the radio transceivers on the sensor nodes (since radio communication is the primary source of energy consumption in sensor networks). With LPL, all nodes are duty-cycled in an asynchronous manner and each transmitter must ensure that the transmission duration of each of its packets overlaps with the on-time of the receiver. BFC works on top of LPL due to LPL's inefficiency not only when it comes to broadcast, but also when it comes to unicast: LPL's unicast transmissions last longer than needed because receivers do not wake up as soon as unicast transmissions begin. In the parentless state, LPL continues to check the medium every active period (on-time period) but takes no action if radio activity is not detected. This approach is employed to fully identify all the sink's neighbors while maintaining the protocol 100% broadcast-free.

- ⋆ Adaptive Low Power Listening: Since BFC discards broadcast control traffic completely, the tree constructed by BFC is way more unbalanced, in terms of load (i.e. number of forwarded packets), compared to the one formed by CTP. To ameliorate this load-balancing problem, BFC employs a built-in adaptive LPL mechanism. Every node i monitors its forwarding load and adapts its LPL active time accordingly. If i has a heavy load, its parent j is bound to have a heavier load and adapts its LPL active time (halves it). As a result, the cost of i's unicast transmissions is also halved. If i also halves its active time, the cost of the unicast transmissions of its children also gets halved.

- ⋆ Data Path Validation: The nodes keep track of the route cost by counting the Required Number of Transmissions (RNP [5]) as the cost metric.

The main downside of the overall BFC design is the substantial latency introduced at start-up until all the nodes have joined the tree, as reported in [20].

Chapter 3

Evaluation Framework

This chapter outlines the evaluation framework upon which our experimental work is based. We present the employed methodology in Section 3.1 and then we shed light on a number of implementation details related to our work in Section 3.2.

3.1 Methodology

Our experimental evaluation is completely based on a publicly available testbed. We selected the Indriya testbed [7] because it offers reasonably challenging testbed conditions as well as a relatively large scale network (138 active nodes as of spring 2013).

We used existing TinyOS implementations of CTP and BFC, and ran both protocols on top of BoX-MAC [16]. BoX-MAC is the standard MAC layer employed in TinyOS and the most popular flavor of LPL. In all of our experiments nodes inject one data packet every 5 minutes (a reasonable interval value for several low-power data collection applications). The sink was located at one edge of the network (node 1 in the top right corner) and it was always on (a typical choice since the sink is normally connected to a station with unlimited access to energy resources). Given that sensor network testbeds can have different connectivity patterns at different times [18], we always ran BFC and CTP back-to-back so that both protocols could operate under similar conditions. We used a transmit power of 0 dBm on 802.15.4 channel 26 in all experiments.

Even though we ran multiple experiments over several LPL wake-up intervals during weekdays in the mornings and early afternoons (so as to ascertain that protocols' behavior is examined under different times), in this thesis we show the results for a wake-up interval of 1 sec because it is in the optimal operating range of CTP – with respect to energy consumption – for the assumed offered load of 1 packet per node every 5 minutes. For all the other LPL intervals (>1 sec), BFC performed even better than CTP. In line with most experimental studies in this area, we use the duty cycle as a proxy for energy consumption [8], since measuring the energy consumption of each and every node on a remote-access testbed is, in principle, not possible.

In Chapter 4, we report our results based on the different roles that nodes play in the network. This is done because the sink's neighbors, relays and leaves can have vastly different performances. The sink's neighbors are the nodes that can reach the sink in one hop, relays are the nodes that are not sink's neighbors but that need to forward packets for other nodes (besides their own packets), and leaves are the nodes that only have to send their own packets.

It should be noted that we focus on the trend of the results and not on the exact numbers, although we do report numbers in Chapter 4. While it is true that link dynamics can vary from experiment to experiment, we do not average the results among experiments for the same LPL wake-up interval because, as it was mentioned above, the trend of the results aptly fits our purposes (subsequently allowing us to convey our key insights). The aforementioned trend was the same among all the experiments for the same LPL wake-up interval.

3.2 Implementation Details

Although this thesis focuses on evaluating existing protocols rather than proposing a novel one, we had to implement a number of features in order to realize our research. Our implementation procedure can be divided into two main parts, namely running code in nesC and data processing code in MATLAB.

The first part was about modifying existing TinyOS implementations of CTP and BFC. That would not be feasible without a very good understanding of what both implementations do. It should be noted that each implementation includes a number of files all written in nesC, a C variant used to build sensor network applications in TinyOS. In order to separate the control traffic from the data transmissions of CTP (see Sections 4.2.1 and 4.2.3), we added code (most notably) in the following two modules: *CtpForwardingEngineP*, *CtpRoutingEngineP*. The former handles the data traffic while the latter is responsible for the control traffic part and, consequently, the tree formation. Other modifications required to enable the above-mentioned separation included changes in the configuration *CtpP* and the introduction of new interfaces that allowed for passing variables from one module to the other. We also had to add/change a number of other lines of code to assess the link quality of CTP and to get the correct information with respect to the path length of BFC (among other things). It should be highlighted that, by the time this M.Sc. project began, the implementation of BFC was quite raw. For instance, the main part of the BFC code ($>1{,}000$ lines) was heavily uncommented and unindented (unstructured). The fact that we found a critical bug in the BFC code merely proves the aforementioned point (regarding the raw implementation).

The second part consisted of a heavily iterative processing procedure in order to assure that the collected data was manipulated properly without overestimating/underestimating the performance of the two protocols under scrutiny. That part was implemented in MATLAB ($> 1{,}000$ lines) allowing for a faster verification due to a number of handy features available in MATLAB.

Chapter 4

Empirical Results and Analysis

This chapter is the essence of our work. It presents our results and conveys the key insights. It is important to note that we divide the analysis of the results into two main categories. First, we delve into the routing structure (see Section 4.1). Then, we focus on the operational part (see Section 4.2). Dividing the analysis into the two aforementioned categories helps us in conveying our insights more effectively.

4.1 Routing Structure Analysis

Routing has two main components, namely link quality and path length. Those two components are jointly responsible for the formation of the data gathering tree. Therefore, we will proceed to examine both in Sections 4.1.1 and 4.1.2, respectively.

4.1.1 Link Quality

For LQE-free protocols to be competitive, the quality of the selected links should be high. To utilize a common metric, we define link quality as the average number of transmissions required to send successfully a packet from any node to its immediate parent. Figures 4.1 and 4.2 show the cumulative distribution function (*cdf*) for all child-parent links for relay and leaf nodes, respectively, and Table 4.1 reports the mean and standard deviation.

We observe that performing *a priori* LQE does not help in selecting better links. In fact, BFC tends to select better links. CTP's links are roughly 9% costlier than BFC's for relays and 5.5% costlier for leaves. Below we provide a conjecture as to why this happens. We do not show the results for the sink's neighbors because all of these nodes had perfect links with CTP and BFC.

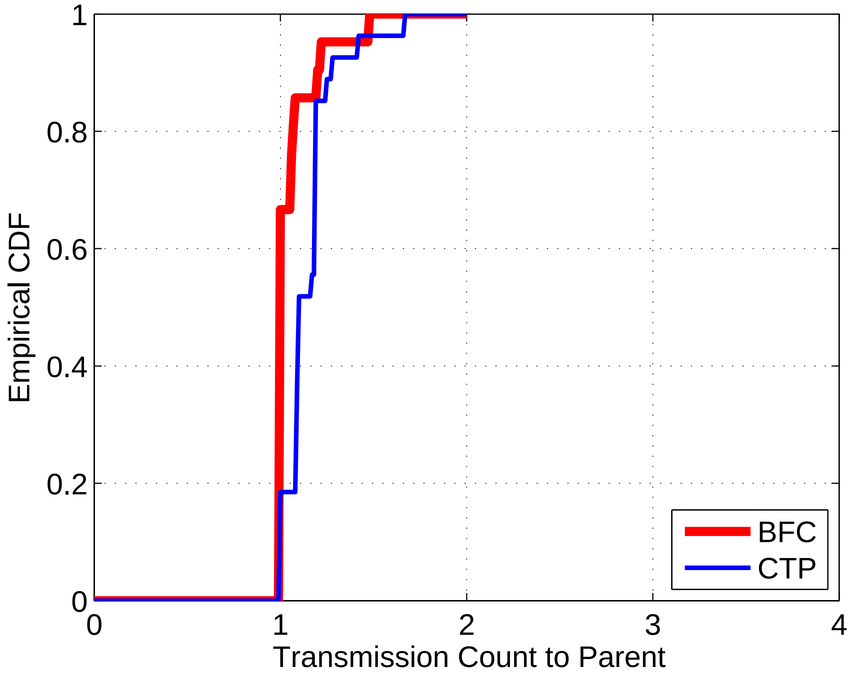

Figure 4.1: CDF of the average number of transmissions to the selected parent for the relays.

	relays		leaves	
	CTP	BFC	CTP	BFC
average (μ)	1.15	1.05	1.27	1.20
std dev (σ)	0.14	0.11	0.42	0.36

Table 4.1: Link quality (# transmissions).

Why does *a priori* LQE not seem to matter? The key limitation of LQE methods is the presence of "intermittent" links. Based on the various LQE studies in the literature [3], let us assume the following simplified classification of sensor networks links: (i) good, when they rarely drop a packet (very high signal-to-noise ratio, way above the radio sensitivity), (ii) intermittent, when they have periods of good quality but also periods of bad quality (the signal-to-noise ratio oscillates around the radio sensitivity), (iii) unreliable, when most of the time their quality is below good (the signal-to-noise ratio is usually below the radio sensitivity). LQE methods are very good at discarding unreliable links, but not so good at discarding intermittent links (because there is a chance that the quality is measured during a good period). In fact, Alizai et al. identified this limitation as a weakness of CTP but from a different perspective [2]. That study claimed

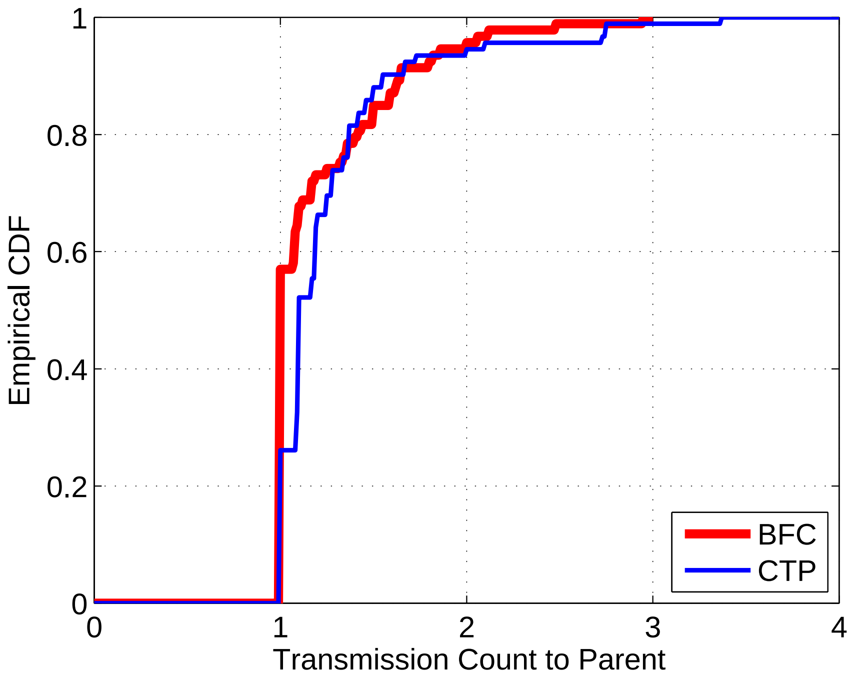

Figure 4.2: CDF of the average number of transmissions to the selected parent for the leaves.

that CTP was too conservative. By discarding some of these intermittent links, CTP was not taking advantage of those links' good periods, which decreased its performance. In this study, we move in a different direction. We conjecture that LQE-free protocols overcome the lack of LQE methods by being even more conservative on the utilization of intermittent links.

In BFC, nodes have a pre-determined parent (as in CTP), but the network relies solely on good links. In BFC, a node is considered a parent if three consecutive data packets arrive from the (potential) child node. Notice that BFC does not use a single beacon to find a parent, BFC simply relies on sniffing and latching on nodes that already found a path towards the sink. Given that (i) in most data collection applications the period between consecutive data packets is of a few minutes, and that (ii) intermittent links are bursty in nature [3]; by validating a link every few minutes, the chances of picking an intermittent link are significantly reduced. BFC is conservative in the sense that, by default, its data path validation mechanisms discard unreliable and intermittent links. This conservative approach, letting BFC assume symmetric links, may explain why BFC selects slightly better links than CTP.

In ORW [12], nodes do not have a pre-determined parent; they transmit information opportunistically using a unicast paradigm. In this way, nodes do not assume that a link exists (which is what protocols with *a priori* LQE methods do); a node uses what is available at the moment. Given that link quality is known to be highly correlated in time, ORW usually ends up selecting either good links or intermittent links during good periods. ORW is conservative in the sense that unless a link is currently good, it assumes that it does not exist.

4.1.2 Path Length

In CTP, the *a priori* LQE phase not only allows the network to identify good links, but this preliminary information is also used to identify short paths. On the other hand, in LQE-free protocols nodes do not have this extra information, and hence, they may end up using longer routes. Figures 4.3 and 4.4 show the *cdf* of path lengths for relays and leaves, and Table 4.2 reports the mean and standard deviation. The absence of LQE clearly affects the depth of the data collection tree. Thanks to LQE, CTP's routes are roughly 8% shorter for relays and 13% shorter for leaves in this particular experiment. Even though we did not perform experiments with ORW, we expect this trend to be the same. Due to the opportunistic nature of ORW transmissions, packets can follow paths that deviate significantly from the shorter routes identified by CTP.

	relays		leaves	
	CTP	BFC	CTP	BFC
average (μ)	2.72	2.95	3.36	3.85
std dev (σ)	0.77	0.80	0.99	0.87

Table 4.2: Path length (# hops).

With BFC, every node tends to stick to a workable parent, i.e. a parent that offers a viable route to the sink and it switches parents only after six successive unacknowledged transmissions. The new parent, however, does not necessarily offer a better (either shorter or longer) route to the sink. On the other hand, the LQE-based operation of CTP allows nodes to decide more drastically on picking better routes to the sink thus rendering the changes in the number of hops more often.

4.1.3 Link Quality vs. Path Length

To provide a comprehensive view of the relation between link quality and path length, Figures 4.5 and 4.6 show the routing cost (number of transmissions required to deliver a packet to the sink) versus the path length for relays and leaves. If all paths were error-free, the plot would show a line with slope 1. We observe that, in this particular experiment, the path

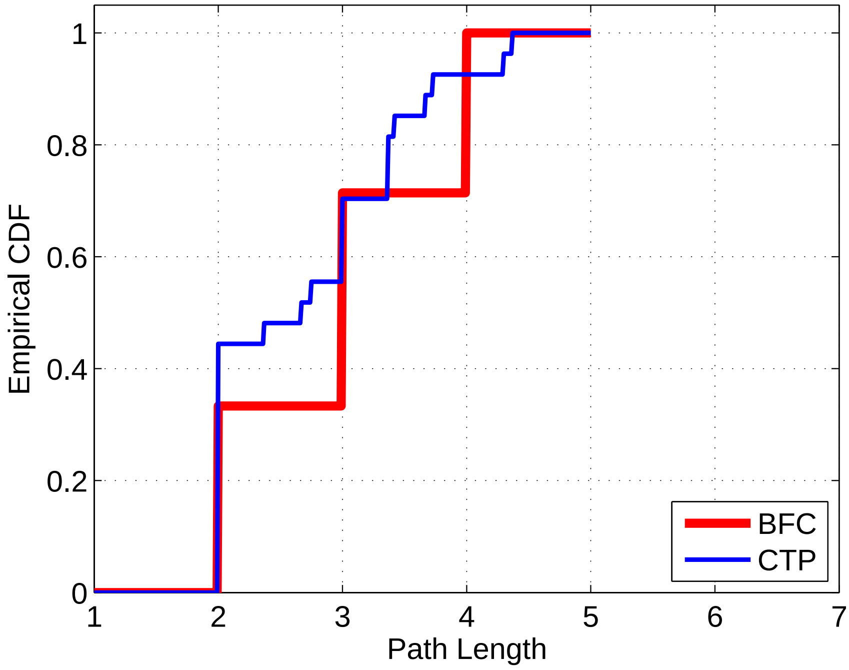

Figure 4.3: CDF of the average number of hops to the sink for the relays.

lengths are much noisier for the leaves than they are for the relays, mainly for CTP. We hypothesize that, because the number of leaves is way higher than the number of relays for both protocols (see Table 4.3), the probability of picking an intermittent link, mainly with CTP, increases significantly.

At first glance, we could state that LQE-methods are more efficient because they provide shorter paths with similar link qualities, that is, they require fewer transmissions to deliver the same amount of data. But, as we will observe next, these savings are minimal compared to the energy expenses of CTP control traffic (which is required to find these good short routes in the first place).

4.2 Operational Analysis

We will proceed to look into energy efficiency (see Section 4.2.1) and load balancing (see Section 4.2.2), and subsequently associate both with the first part of our analysis where we investigated link quality and path length.

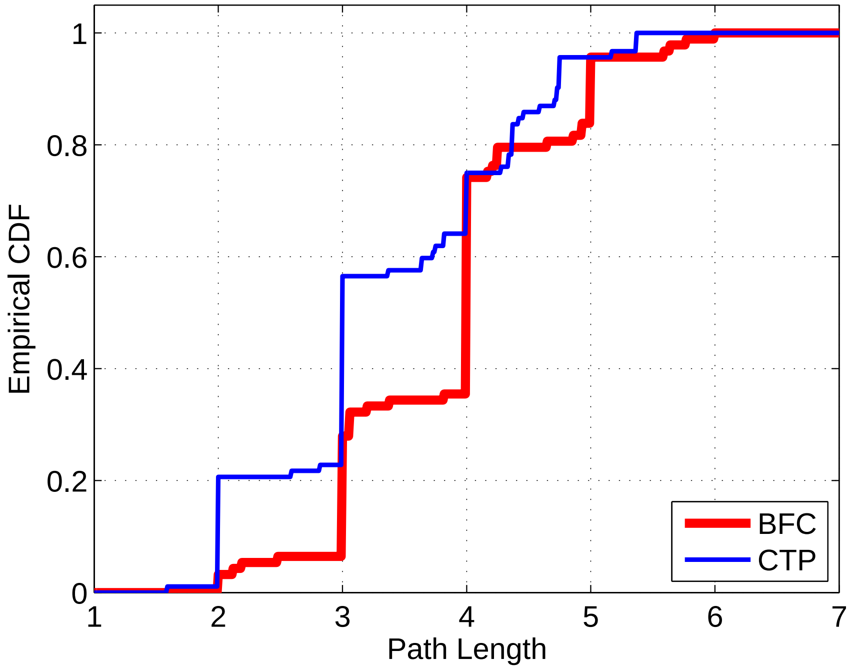

Figure 4.4: CDF of the average number of hops to the sink for the leaves.

4.2.1 Energy Efficiency

We have seen that without employing LQE underneath, BFC still opts for good links. Now, we will proceed to understand why the overall operation of BFC, and LQE-free protocols in general, can be so energy-efficient. To this end, we changed slightly the CTP code to monitor the fractions of the duty cycle related to control traffic and to data traffic. We also found a bug in the BFC code that penalized its duty cycle (i.e. BFC's energy performance is better than the measurements reported in [20]). Figures 4.7 and 4.8 show the mean and median duty cycle for the three types of nodes. The median plot allows us to filter out connectivity outliers. Furthermore, Figures 4.11 and 4.12 allow for a complete view of the energy consumption on a node basis (see Section 4.2.3).

Figure 4.8 shows that the cost of CTP control traffic is similar for all node classes, in the order of 0.5%. This occurs because CTP periodically sends probes to test the connectivity of nodes. It is important to highlight that these measurements were performed once CTP and BFC reached a stable state, that is, after the initial probing required to bootstrap the network (otherwise the duty cycle of CTP would be even higher). The important lesson from this result is that it gives us an "LQE-free" version of CTP (captured by the white bars). Assuming an ideal static environment, CTP

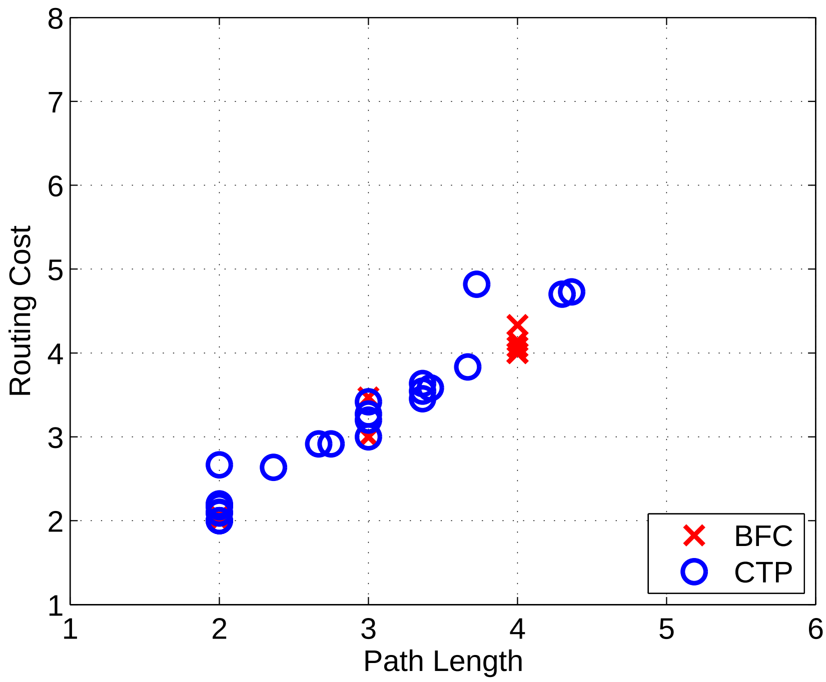

Figure 4.5: Routing cost vs. path length for the relays.

would need to run the Trickle timer only once (at the beginning). Then, after the routes are formed, the LQE phase would not need to be run again because links would not change. This result allows us to benchmark an LQE-free protocol, such as BFC, with an ideal "LQE-free" version of CTP.

Figure 4.8 also allows us to convey a few other noteworthy insights. As mentioned in Section 3.1, the sink was always on in the course of our experimental campaign. This choice favorably affects the duty cycles of the sink's neighbors for both protocols because these nodes forward their packets to the sink immediately without waiting for the sink to wake up. As for the relays, the fact that the BFC duty cycle is less than the CTP data duty cycle can be attributed to the built-in adaptive LPL mechanism of BFC. The adaptive LPL mechanism of BFC allows for additional energy savings for increasing LPL wake-up intervals. Regarding the leaves, BFC is more expensive than CTP when comparing the duty cycle of the former to the data duty cycle of the latter. We hypothesize that BFC's adaptive LPL mechanism has minimal effect on the leaves for a wake-up interval of 1 sec.

It is important to note that BFC has some other disadvantages besides the long latency mentioned in [20]. *A priori* LQE also allows a better responsiveness. CTP would react much faster to changes in the environment than BFC. For instance, if a malicious interfering node is first turned on

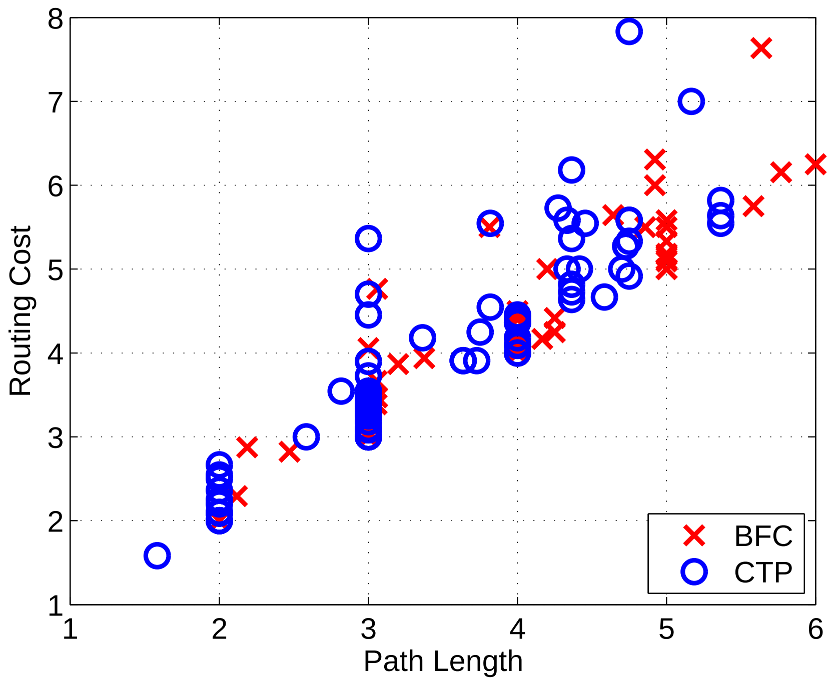

Figure 4.6: Routing cost vs. path length for the leaves.

and then off. CTP would be able to rapidly (re)identify the shorter paths "erased" by the interfering node. On the other hand, BFC, would not be able to recover (or at least not fast enough) because no periodic LQE measurements are done. As part of our future work, we plan to use the JamLab tool [4] to stress LQE-free protocols under different interference patterns.

4.2.2 Load Balancing

In the original BFC paper [20], the authors mention that BFC seemed to have a worse load balance than CTP. We validate this hypothesis. Figures 4.9 and 4.10 illustrate the *cdf* of the forwarding load for the sink's neighbors and relays. The forwarding load is defined as the ratio of the total number of forwarded packets (which includes locally generated and relayed packets) per locally generated packet. We do not present the *cdf* for the leaves because the vast majority of them had a forwarding load of 1. Strictly speaking, all the leaves should have a forwarding load of 1, but, in a real network, leaves may occasionally act as relays for short periods of time. We, therefore, use a threshold of 2 to filter out such nodes from the set of real relays.

Recalling that neither CTP nor BFC employ an explicit load balancing scheme, we see that BFC performs clearly worse than CTP with respect

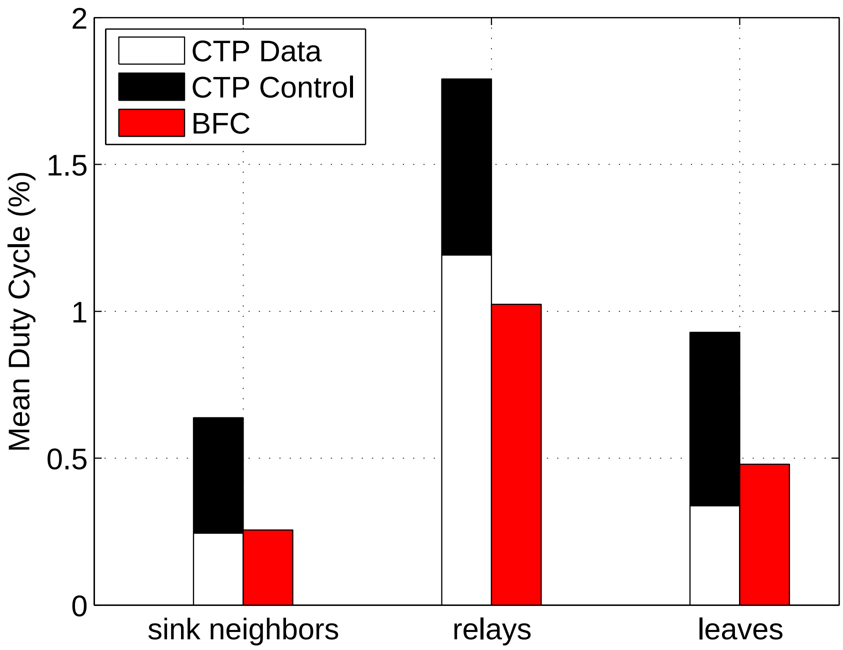

Figure 4.7: Mean duty cycle for sink's neighbors, relays and leaves.

to sink's neighbors (see Figure 4.9). A few sink's neighbors in BFC are responsible for forwarding 80% of the load. The load balancing is better for the relays (see Figure 4.10). To quantify the unfair share of the forwarding load and consequently of the energy resources, we compute the Jain's fairness index for both CTP and BFC for all node classes in Table 4.3. A Jain's index of 1 denotes perfect balance (best-case value), while $1/n$ (where n denotes the number of nodes) indicates clear unfairness (worst-case value). If we consider all nodes, the Jain's fairness index is 0.29 for CTP and 0.18 for BFC.

	# sink's neighbors	index	# relays	index	# leaves	index
CTP	17	0.44	27	0.66	92	0.95
BFC	22	0.18	21	0.44	93	0.96

Table 4.3: Jain's fairness index.

The reason for the worse load balance in BFC seems to be the same as for the longer paths (see Section 4.1.2). Without *a priori* LQE, a network not only looses the ability to identify good links but also to have a broader view of the underlying communication graph. This broader view could facilitate the identification of shorter paths and better load balancing.

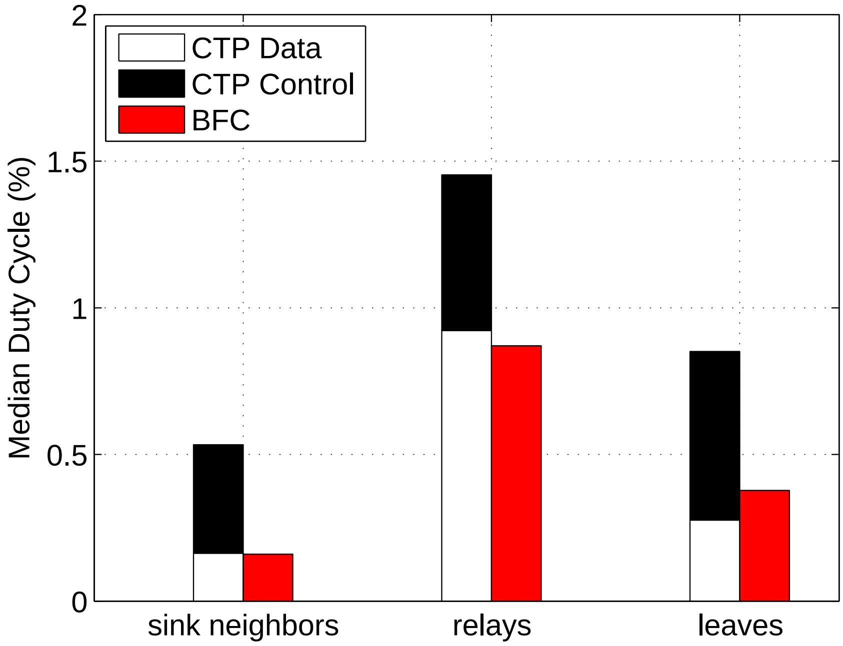

Figure 4.8: Median duty cycle for sink's neighbors, relays and leaves.

Load balancing has been closely related to the hotspot problem, namely a well-known problem in the area of wireless sensor networks that cuts a deployment short due to an emerging load imbalance. Therefore, a number of efforts in the field have focused on investigating load balancing schemes to mitigate the problem and subsequently increase the network lifetime. However, a recent research study in the field has shown that load balancing schemes might have been overrated [14].

4.2.3 Energy Efficiency vs. Load Balancing

To provide a complete view of the effect of LQE on the energy consumption of individual nodes, Figures 4.11 and 4.12 show the duty cycle versus the forwarding load for the sink's neighbors and relays. Leaves can be regarded as relays that are responsible for forwarding only their own load (normally). Therefore, Figure 4.12 depicts both node classes jointly. We observe that, in general, BFC penalizes some individual nodes. As hypothesized in [20], this may be due to a "rich gets richer effect". Nodes that are picked first as parents are eavesdropped first, which in turn increases their number of children and their forwarding load. This load balancing problem is not an intrinsic problem of LQE-free protocols. We believe that the opportunistic nature of ORW will naturally balance the load in the network because it can choose vastly different paths at each time (due to the presence of intermittent links). Analyzing the load balancing properties of ORW (among its many other LQE-free characteristics) is part of our future work.

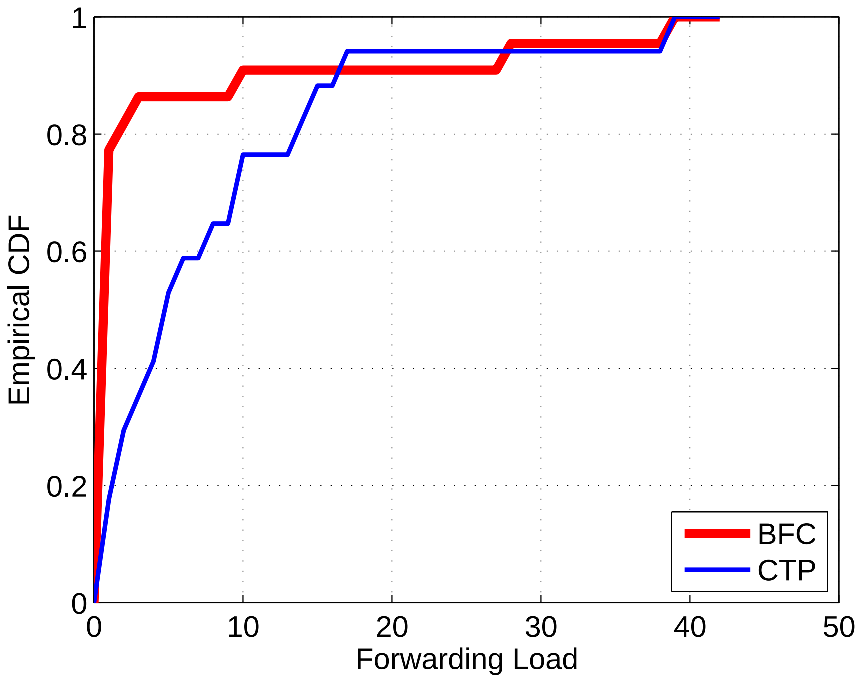

Figure 4.9: CDF of the forwarding load for the sink's neighbors.

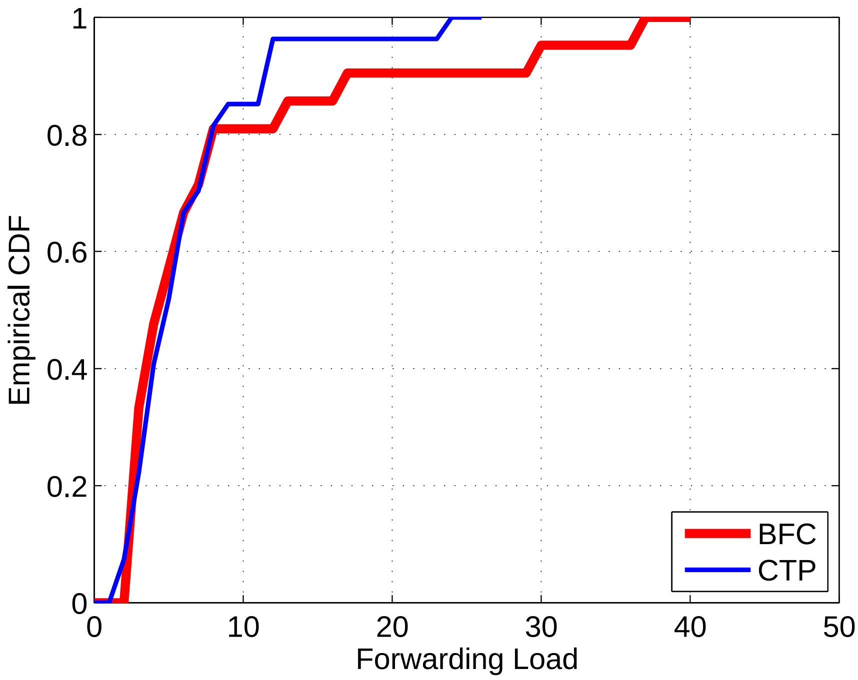

Figure 4.10: CDF of the forwarding load for the relays.

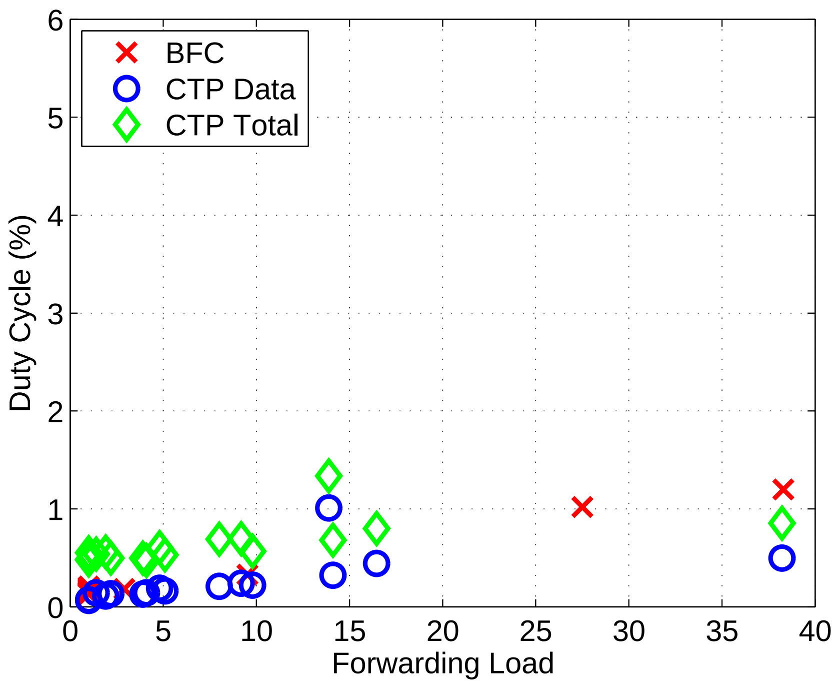

Figure 4.11: Duty cycle vs. forwarding load for the sink's neighbors.

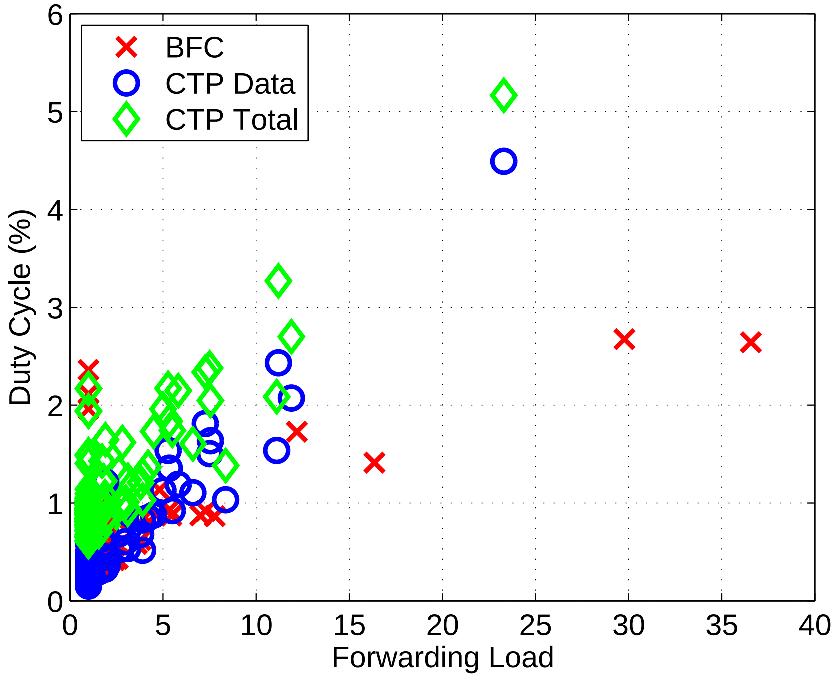

Figure 4.12: Duty cycle vs. forwarding load for the relays and the leaves.

Chapter 5

Conclusions

Hitherto most data collection protocols in sensor networks typically construct the routing structure by relying on LQE methods and on mechanisms to identify short(est) paths. The motivation for our research stems from a number of efforts in the field that aim to contradict this notion either partially or completely. Our goal was to shed some light on why network protocols employing an LQE-free operation offer a few substantial benefits compared to conventional approaches. To this end, we have focused on two protocols at the extremes, namely CTP (LQE-based) and BFC (LQE-free), and subsequently evaluated their performance on Indriya, a publicly available large-scale testbed (138 active nodes as of spring 2013).

Our evaluation indicates that LQE might have been overrated. Not using LQE results in longer routes with extra hops, but it also results in more conservative routing decisions, i.e. choosing individual links with a lower cost. The path length inefficiency caused by the lack of LQE is therefore amortized at the root; on top of that, no LQE means no control overhead, which translates into a reduction of the energy footprint of collection protocols between 40% and 60%. Lack of fairness posed by load imbalance is a clear drawback of LQE-free operation, though it can be avoided by abstracting away from the network topology and getting every node to receive every packet as in [9]. Our results, however, clearly show that load imbalance does not have sizable consequences in terms of energy consumption.

Bibliography

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci. A Survey on Sensor Networks. In *IEEE Communications Magazine*, pages 102–114, Aug. 2002.

[2] M. H. Alizai, O. Landsiedel, J. Á. B. Link, S. Götz, and K. Wehrle. Bursty traffic over bursty links. In *Proceedings of the 7th ACM Conference on Embedded Networked Sensor Systems*, pages 71–84. ACM, 2009.

[3] N. Baccour, A. Koubaa, L. Mottola, M. A. Zuniga, H. Youssef, C. A. Boano, and M. Alves. Radio link quality estimation in wireless sensor networks: A survey. *ACM Transactions on Sensor Networks (TOSN)*, 8(4):34, 2012.

[4] C. Boano, T. Voigt, C. Noda, K. Römer, and M. Zúñiga. JamLab: Augmenting sensornet testbeds with realistic and controlled interference generation. In *Proceedings of the* 10^{th} *IEEE International Conference on Information Processing in Sensor Networks (IPSN)*, pages 175–186, apr 2011.

[5] A. Cerpa, J. L. Wong, M. Potkonjak, and D. Estrin. Temporal properties of low power wireless links: modeling and implications on multi-hop routing. In *Proceedings of the 6th ACM international symposium on Mobile ad hoc networking and computing*, pages 414–425. ACM, 2005.

[6] D. S. De Couto, D. Aguayo, J. Bicket, and R. Morris. A high-throughput path metric for multi-hop wireless routing. *Wireless Networks*, 11(4):419–434, 2005.

[7] M. Doddavenkatappa, M. Chan, and A. A. L. Indriya: A low-cost, 3d wireless sensor network testbed. In *Proc. of the* 8^{th} *Int. ICST Conf. on Testbeds and Research Infrastructures for the Development of Networks and Communities*, 2011.

[8] A. Dunkels, F. Osterlind, N. Tsiftes, and Z. He. Software-based on-line energy estimation for sensor nodes. In *Proceedings of the 4th workshop on Embedded networked sensors (EmNets)*, Cork, Ireland, 2007.

[9] F. Ferrari, M. Zimmerling, L. Mottola, and L. Thiele. Low-Power Wireless Bus. In *Proceedings of the 10th ACM Conference on Embedded Networked Sensor Systems (SenSys)*, Toronto, Canada, November 2012.

[10] R. Fonseca, O. Gnawali, K. Jamieson, and P. Levis. Four-Bit Wireless Link Estimation. In *Sixth Workshop on Hot Topics in Networks (HotNets-VI)*, Atlanta, GA, Nov. 2007.

[11] O. Gnawali, R. Fonseca, K. Jamieson, D. Moss, and P. Levis. Collection Tree Protocol. In *7th ACM Conference on Embedded Networked Sensor Systems (SenSys'09)*, Berkeley, CA, Nov. 2009.

[12] O. Landsiedel, E. Ghadimi, S. Duquennoy, and M. Johansson. Low power, low delay: Opportunistic routing meets duty cycling. In *IPSN'12: Proceedings of the 11th ACM/IEEE International Conference on Information Processing in Sensor Networks*, April 2012.

[13] P. Levis, N. Patel, D. Culler, and S. Shenker. Trickle: A Self-Regulating Algorithm for Code Propagation and Maintenance in Wireless Sensor Networks. In *First USENIX/ACM Symposium on Networked Systems Design and Implementation (NSDI'04)*, San Francisco, CA, Mar. 2004.

[14] Y. Liu, Y. He, M. Li, J. Wang, K. Liu, L. Mo, W. Dong, Z. Yang, M. Xi, J. Zhao, et al. Does wireless sensor network scale? a measurement study on greenorbs. In *INFOCOM, 2011 Proceedings IEEE*, pages 873–881. IEEE, 2011.

[15] S. Moeller, A. Sridharan, B. Krishnamachari, and O. Gnawali. Routing without routes: the backpressure collection protocol. In *Proceedings of the 9th International Conference on Information Processing in Sensor Networks (IPSN)*, pages 279–290, apr 2010.

[16] D. Moss and P. Levis. BoX-MACs: Exploiting Physical and Link Layer. Technical Report 08-00, Stanford University, 2008.

[17] J. Polastre, J. Hill, and D. Culler. Versatile Low Power Media Access for Wireless Sensor Networks. In *2nd ACM Conference on Embedded Networked Sensor Systems (SenSys'04)*, Baltimore, MD, USA, Nov. 2004.

[18] D. Puccinelli, O. Gnawali, S. Yoon, S. Santini, U. Colesanti, S. Giordano, and L. Guibas. The Impact of Network Topology on Collection Performance. In *European Conference on Wireless Sensor Networks (EWSN'11)*, Bonn, Germany, Feb. 2011.

[19] D. Puccinelli and M. Haenggi. Reliable data delivery in large-scale low-power sensor networks. *ACM Transactions on Sensor Networks (TOSN)*, 6(4):28, 2010.

[20] D. Puccinelli, M. Zuniga, S. Giordano, and P. J. Marrón. Broadcast-Free Collection. In *10th ACM Conference on Embedded Networked Sensor Systems (SenSys '12)*, Toronto, Canada, Nov. 2012.

[21] A. Woo, T. Tong, and D. Culler. Taming the underlying challenges of reliable multihop routing in sensor networks. In *Proceedings of the 1st international conference on Embedded networked sensor systems*, pages 14–27. ACM, 2003.