Thermoelectric Harvesting and Seasonal Routing in Wireless Sensor Networks

by

Aristotelis Kollias

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

Department of Computing Science
University of Alberta

Abstract

We propose the use of embedded in-wall thermoelectric harvester to power the nodes of a wireless sensor network. We exploit significant temperature differences of indoor and outdoor environments in cold climates. We use real measurements to evaluate the use of the temperature difference as a proxy to heat flow through walls, and the feasibility of creating a network of in-wall thermoelectric harvesting sensors. We also discuss the seasonal availability of thermoelectric energy and get in depth to the issues and advantages this presents. We formulate the corresponding multi-commodity routing flow problem, where each commodity is data from every node, and using the indoor/outdoor temperature as a proxy we investigate the effect of fairness in routing. We also present a low overhead method of constructing a seasonally-aware routing scheme and study its performance. Finally we talk about techniques for predicting the incoming available energy in the future.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Wireless sensor networks are used to assist in a number of applications, such as structure monitoring and maintenance, to monitor human activity, to help prevent disasters like forest fires, and, in general, to collect data for scientific and business purposes, [1]. In the construction industry, sensors can be used to alert of dangerous situations, in cases where the infrastructure is critically compromised and to monitor wear and the "health" of buildings in general [2, 3, 4]. For example, humidity sensors inside wall structures can provide advance warning of excess humidity which could lead to toxic mold growth, and strain dynamometers sensors can be used to determine the response of the building during earthquakes.

We consider the standard model for wireless sensor network (WSN) nodes, i.e., as consisting of a transceiver, the sensor, a microcontroller and an energy source, usually a battery. Generally, when the energy source for the WSN node is a battery, there exists a limit to how long the node can function without servicing, i.e., changing of batteries. With current technology, the node modules generally exhibit low energy consumption, therefore the lifetime limit that a battery imposes can be long enough that, depending on the situation, replacing batteries will not impose a significant cost. However, in particular situations, the sensors are inaccessible, i.e., embedded within structures, like walls, and the cost of replacing the energy source might be prohibitive.

There are three possible solutions for difficult–to-access sensor nodes. They can be either abandoned after they have performed their function for some time (with

the hope that other, nearby, sensors will take over the task of measuring the same phenomenon), or they can be all connected to a wired power distribution subsystem. Finally, energy harvesting could be employed. Abandoning the sensors is an expensive option and possibly unacceptable. The wiring option is expensive in terms of labor cost and materials, exacerbated by the price of copper necessary for wiring, and wiring defeats the advantage of using wireless, since communication can be accomplished over the same wires that provide power. Wiring also results in a structure which is more complex and could be prone to faults (accidental puncture/cutting of the wires, problems with single points of failure, collection of RF energy by acting as low frequency antennas, etc.). In short, the ability of each node to power itself from energy harvesting leads to better node autonomy, and overall system resilience.

The solution we adopt for this work, is to use thermoelectric energy harvesters to power each sensor node. Energy harvesters exploit the ambient energy of the environment to replenish the energy stored in the battery or the super capacitor. Photovoltaic harvesters have been studied extensively in previous works e.g., [5]. Buildings with good illumination can use photovoltaic energy to power sensor modules. Instead, in this paper we consider thermoelectric energy harvesting because: (a) the potential for photovoltaic energy can be limited due to sub-optimal placement, and because of long nights, as is the case at latitudes of northern continental climates, while, (b), during winter the indoor to outdoor temperature difference can reach as much as 60 degrees Celsius creating unparalleled opportunities for thermoelectric harvesting. More specifically, we take advantage of the difference between indoor and outdoor temperature in buildings in northern climates. To this end, we use data collected from an actual inhabited apartment complex in Fort McMurray, Alberta, Canada. The particular apartment complex has been constructed using modular construction techniques.

Placing thermoelectric harvesters inside walls serves the application of powering co-located sensors that monitor the wall and building behaviour. In climates like the one considered in our study, extreme weather conditions can cause events important to the integrity of a building, e.g., breakage of water pipes, more frequently occurring

than in moderate climates. The ability to embed sensors in inaccessible locations that autonomously operate for several decades (i.e., as long as the building lasts), monitoring for such events, can currently only be supported using energy harvesting.

1.1 Motivation

Our research is primary motivated by the need for almost zero maintenance sensor networks, that will be able to operate independently and autonomously. Networks like these will be able to be installed inside buildings, in remote infrastructure and generally in any difficult to reach and maintain areas. Their primary job will be to monitor and transmit information, dealing with data collection needed for applications such as: monitoring the operational efficiency of a system, collecting usage statistics, reporting critical information e.g., structure fatigue, etc.

For the above to be achieved the following goals have to be met:

1. energy autonomy of the nodes throughput the year,
2. dynamic routing of data based on energy availability,
3. node placement to maximize their application-specific utility,
4. considerations of cost vs. robustness tradeoffs.

In this thesis we are not considering item (3) as it is specific to the applications at hand. For illustration purposes, when necessary, we consider that the application is simply data collection to a sink node from all nodes, involving possibly multi-hop communication (routing) to deliver the data to the sink.

Energy harvesting on wireless networks was introduced as a way to prolong the lifetime of the nodes. By harvesting, you can replenish the energy in the nodes battery and ideally extend the sensors lifetime forever. Unfortunately, batteries have a limited number of possible recharges, and they slowly degrade, especially in adverse environments [6]. Also due to the, generally, variable availability of energy to harvest, in most cases we cannot rely on the energy to be sufficient for the continuous operation

of the nodes. A solution to those two conflicting requirements is to use (super) capacitors as energy storage for the harvesting. While they cannot store as much energy as a battery, they are ideal for harvesting applications, as they don't have a limited number of charge/discharge cycles. As seen in Simjee and Chou [7], while a node that operates with a battery source has a 1-2 years of lifetime, due to the amount of charge/discharge cycles, their system, that operates with an ultracapacitor, has a theoretical lifetime of 20 years. To keep the size of the nodes small, we assume that the capacitors used provide typical storage in order of a few Farads.

Another interesting question is that of determining the most cost-effective solution: one can use bigger energy harvesters, which allows harvesting more energy, but they increase the overall cost and could even create secondary problems. For example, a secondary problem with photovoltaic harvesters is that a bigger harvester needs more exposed surface area, which can range from inconvenient to impossible. Equally, to make good use of a bigger harvester, additional labor costs could be involved to get the exact positioning right. In other cases, e.g., in-wall thermoelectric harvesters, the harvested energy comes from an increased energy loss through the wall area where the harvester is installed, i.e., reducing the insulation of the wall from the outside environment which defeats the design purpose of a wall.

Lastly the nodes need to dynamically manage their power, due to the variable nature of the energy. Routing and sensing of the data should adapt to the current energy constraints. In periods of time with superfluous energy the nodes are not required to conserve anything, and they can work with no constraints. On periods with really constrained energy, the nodes need to optimize their energy usage, by sensing according to their energy reserves, routing their data through possibly different paths, according to the reserves of the other nodes of the network, and generally manage their operation according to the current energy situation of the network.

It could be argued that making the network into a wired one would solve most of the problems, but, even discarding the cases where cables are impossible to use, cables impose an extra cost, create extra problems on their own, like an increased weakness to corrosion etc.. Wireless networks add more flexibility to the system, freeing us to

perform node placement guided by the application needs. The wireless sensor network deployment is constrained by node energy and the achievable transmission range. In this thesis we focus on the node energy component.

1.2 Research Questions

The intention of this thesis is to act as, primarily, a feasibility study for the use of inexpensive in-wall thermoelectric harvesting to power commonplace wireless sensor nodes. To this end, we consider specific technologies for the harvesters and the nodes and provide absolute numbers of their energy harvesting and corresponding data volume transmission capabilities. However, the question of designing such networks is not yet tackled and to this end we have to determine what kinds of information a network designer would need to study, e.g., to simulate, a WSN powered this way before deciding the deploy the nodes. In particular, in-wall deployment means that the burden of sensor placement has to be addressed at the earliest stages of designing a building, conceivably using tools that are the same, or similar to those used by Civil Engineers. The simplest form of information a designer might have access to are temperature measurements; outdoor temperatures based on historical data, indoor temperatures based on resident behavior models. We therefore study into how good are indoor/outdoor temperature differences in providing a basis for thermoelectric energy harvesting estimates. The advantage we exploit is that we have access to highly accurate heat flow measurements from an actual apartment complex. The heatflow measurements are the most directly representative examples of energy transfer measurements (a fraction of which can be harvested) and depend on the wall structure and the choice of insulating and other materials.

Beyond the feasibility and design considerations, another aspect of the long-term operation for a WSN utilizing in-wall thermoelectric harvesting is the anticipation that it will, on occasion, fail to send the data needed due to the scarcity of harvested energy. Hence, at any point in time a certain total volume of data transfer is feasible given the current energy storage. We attempt to characterize the volume of data that

can be delivered from the nodes in the maximal sense and in a fair sense. The maximal approach looks into maximizing the amount of data delivered to the sink regardless of the fact that some nodes could deliver less (even zero!) of their data volume to the sink. The fair approach allows for a balanced, equal across all nodes, volume of delivered data albeit typically the sum of data volume delivered in total is less than that of the maximal case. We attempt an optimization formulation that can reflect the reality of such network. We take advantage of the, usually, static topology of such networks and the general approach of duty cycling the operation of the network which means that all transmission/forwarding action takes place at regular, periodic, instants. Based on a formulation of the underlying optimization problem, we come up with realistic routing schemes that exploit the strongly seasonal behavior that characterizes the amount of thermoelectrically harvested energy.

This thesis is strongly data-driven, utilizing the data collected from a project involving the instrumentation of a multi-unit apartment complex in the area of Fort MacMurray, Alberta, Canada [8]. Data sets of temperature and heat flow measurements used throughout the thesis were collected from that particular building. Layouts and network topologies used also correspond to the same building.

1.2.1 Contributions

The contributions of the present thesis are:

- A study demonstrating that the use of indoor and outdoor temperature readings are good proxies based on which the thermoelectric harvesting potential for in-wall WSNs can be assessed.
- A feasibility study demonstrating the adequacy of thermoelectrically harvested energy to power WSN nodes using off-the-shelf technologies for both harvester and nodes alike.
- A modeling approach of the data delivery process at "snapshot" instants (once in every collection "round") based on multicommodity flow and two varieties of objectives.

- A routing algorithm driven by simple lookup tables that takes into consideration the seasonality of the available energy and uses past measurements to construct the routing alternatives.

Towards the end of the thesis, we also make a first step towards the prediction of how thermoelectric energy is available in the near future, with the ultimate goal of optimizing the current set of routing decisions based on several steps of predicted harvesting into the future. However, this enters into the area of a combined prediction model that involves, on one hand weather patterns, and on the other, resident behavior, whose combined effect is left for a future study.

Most of the work on this thesis has been reported in conference papers [8, 9]

1.3 Thesis Outline

We consider the problem of multi-hop routing in WSNs of static topology composed of nodes that exploit thermoelectric energy harvesting. Our objective is the multi-decade autonomous operation of the WSN. The specific application domain considered is sensors embedded in exterior walls in buildings, and more specifically in Northern climates where, especially in the colder months, the temperature difference between indoor and outdoor provides abundant opportunities for thermoelectric harvesting. The reasons for embedding wireless rather than wired sensors is primarily due the increased costs of wiring and labor required to install such wiring. Moreover, the sensing we would like to perform is generally taking place in hard to reach locations, not necessarily conducive to other forms of energy harvesting, e.g., photovoltaic harvesting, due to lack of light and constraints stemming from orientation/placement. On the contrary, heat transfer is a universal phenomenon evidenced everywhere, albeit not always at levels that would provide for effective harvesting.

Chapter 2 provides a review of basic concepts and work carried out by other researchers in the area of energy harvesting for WSN regarding both the practical and theoretical aspects. Chapter 3 provides the first step to addressing the issue of determining, at the design stage, what kinds of information is needed to derive

relatively accurate estimates of the thermoelectrically harvested energy. Our access to a unique data set from an apartment complex allows us to determine whether the differences of indoor / outdoor temperatures is a good proxy for the actual heat flow – and indirectly for the harvested energy. It is envisioned that the network designer, knowing the use of a building (and possibly having a resident behavior model in mind) and the weather in the particular area, can then make an informed decision on the amount of data that such nodes could collect and forward. We make a number of observations on the time-varying nature of thermoelectric harvesting and conjecture on the links of its performance to factors such as the occupant behavior.

Chapter 4 presents the main contributions of the thesis, (a) an optimization formulation and the solution of time-dependent routing problem given energy levels and the corresponding fairness formulation of the same, (b) a study of the impact of the time-varying nature of the thermoelectric harvesting on the cycle-by-cycle routing decisions, leading to, (c), a seasonally–aware routing scheme which is possible to deploy using routing decision made in previous years and by building simple look-up tables that can be downloaded to the sensor nodes. Techniques to reduce the risk of complete energy outage are also introduced, by reserving some of the energy of the nodes for subsequent cycles.

Concluding, Chapter 5, presents the possible extensions that could lead, ultimately, to routing strategies developed over a finite horizon of future energy harvesting. The concluding chapter describes the roadblocks that still exist in the deployment of thermoelectrically–powered WSNs and suggests future directions to address them.

Chapter 2

Related Work

The field of Wireless Sensor Networks (WSNs) is a well established field of research. WSNs typically rely on low power electronic systems (like micro-controllers, communication modules etc.) [10]. They consist of autonomous nodes, that provide sensed information from the natural environment, such as temperature and humidity, have some (usually limited) processing capabilities and some (again, limited) wireless communication capability. The nodes are often connected to a finite power source (battery), or as is less often the case, to a renewable energy module, which draws power from the environment in the form of energy harvesting [11, 5, 12, 1, 13]. The nodes usually possess limited or even no knowledge about their remaining energy reserves. To the extent that they are aware of their energy reserves, they can adjust their operation accordingly to make optimal usage of this energy [14]. The described research touches on both the energy harvesting and the energy–aware operation of WSNs.

2.1 ENERGY HARVESTING

Energy harvesting for powering WSN nodes, with example applications such as smart buildings and predictive maintenance of structures, is not new [1]. Some of the most common types of harvesting, used for WSNs, are photovoltaic harvesting, piezoelectric harvesting and thermoelectric harvesting. Photovoltaic harvesting is the most commonly explored type, which draws energy from light and transforms it into elec-

trical power. Piezoelectric energy is the creation of electrical power from specific materials, when they get subjected to vibrations and or mechanical stress. Lastly thermoelectric harvesting, which is the focus of this thesis, involves the transformation of the heat transferred from an area of higher to an area of lower temperature, into electrical energy.

The reason we select the thermoelectric harvesting can be understood when we consider the following facts:

- In photovoltaic harvesting, the energy is dependent on the amount of light that exists in the environment. That means that a well-illuminated environment is necessary to harvest sufficient amounts of energy. Some indoor environments, such as fairly well-lit office buildings, provide remarkably stable energy availability for photovoltaic harvesting [5]. Yet other, and in particular shielded, environments such as inside a wall or within pipes and structures are by definition unsuitable for photovoltaic harvesting.

- Piezoelectric harvesting is more suited to environments where mechanical moving parts are involved, e.g., automobiles, doors, etc. , and on occasion it is the very vibrations (and resulting mechanical stress) that they both exploit for powering themselves and about which they report sensed data. Clearly, not all environments involve moving parts or opportunities to attach a harvester to a moving part.

- Thermoelectric harvesting is ideal for situations where we are in close proximity to surfaces that are consistently at different temperatures with the environment around them. That includes server rooms, where the surface of microprocessors is always warmer than the environment [13], or, as in our case taking advantage of the fact that in Northern climates the inside surface of a wall in residential buildings has a significant temperature difference to that of the walls outer surface. (The basis of thermoelectric harvesting is the *Seebeck effect* which is described in Chapter 3.)

Notice that we are not concerned with the task of increasing the energy harvesters efficiency, a topic of intense activity anyway, e.g., [15, 16]. We consider off-the-shelf components and we do not even use a sensor platform optimized for energy efficiency. In other use words, the results we present here are very close to representing a "worst case" scenario with respect to the devices employed.

Integrated nodes built around a specific type of harvesting have been presented in the past, e.g., the one reported by Raghunathan et al.[17] built around solar harvesting. In addition to designing and building such a node, they evaluate its energy consumption and its rechargeability capability. However, more notable work in this field is that of Gorlatova et al. [11, 5] which focuses on how to use photovoltaic harvesting under diverse use scenarios, and proposes suitable optimization models. Another distinct feature of their work is the study of photovoltaic harvesting in environments under the control of the users (depending on indoor illumination, or in the pocket of users), i.e., with idiosyncratic and sometimes unpredictable behaviour. We follow on their example to the extent that we observe the thermoelectric energy harvesting over time in realistic scenarios.

Where our work differs significantly from that of Gorlatova et al. is on our focus on the energy performance of the whole network given real measurements that allow us to incorporate the location-dependent effect of energy harvesting. Circumstances differ for the various nodes of a WSN relying on energy harvesting, hence while some might be more energy "rich" other are not. Such an imbalance is unavoidable in real systems. Hence, while Gorlatova et al. provide algorithms that optimize the energy usage of a single node over time, we are trying to optimize[1] the energy usage through the whole network. Effectively, Gorlatova et al., have an application model in mind where nodes broadcast some type of data (measurements, unique ID etc.) that gets received directly, without multi-hop forwarding, by the sink node. On the contrary, we adopt multi-hop forwarding (even if limited to a few hops) as a necessary feature to allow communication with hard-to-reach (in the RF transmission sense) nodes.

Towards this end we use a device model loosely based on that of Mateu et al.

[1]The exact nature of this optimization is described later in this thesis.

[12], whereby thermoelectric energy harvesting takes place by exploiting the fact that the human body maintains a roughly constant temperature regardless of the environment, and hence the heat flow from/to the body to the environment can be captured using some "wearable" device. The body–to–environment heat flow capture is an idea explored by a number of other researchers as well [18, 19]. Whereas this previous work exploits the temperature homeostasis of the human body, we exploit the less predictable but consistent behavior of keeping indoor spaces at reasonable, for human habitation, temperatures compared to the outdoors environments.

In the rest of the thesis, it is tacitly assumed that the outdoor environment is usually at a lower (to much lower temperature) than the indoor environment because our source of data is in Northern climates and close to the Arctic Circle. The reverse relation (colder indoors than outdoors) can also hold with, presumably, similar results e.g., in environments closer to the tropics. Indeed, as the results will show in the thesis, the most challenging time for thermoelectric harvesting is when the indoor and outdoor temperature are the same, which is a characteristic in milder climate regions. Even there, temperature differences between night and day and the resulting heat flow might be sufficient to produce the requisite energy. In short, the heat flow is a universal characteristic and the essence of the laws of thermodynamics. It is impossible to find environments where it does not take place. The question is therefore whether it is sufficient to be harvested, and if today's off–the–shelf harvester can do so effectively.

2.2 ENERGY AWARE OPERATION

Apart from harvesting, or increasingly together with it, there is an effort to make the nodes individually and as a group more energy efficient. In the case of no harvesting, both researchers and the industry are trying to lower the energy consumption of the nodes, from either the hardware or the software point of view.

From the hardware point of view, electronics suppliers are making lower power electronics. One of the first commercial–grade off–the–shelf low power microprocessor

was the Texas Instruments MSP 430 microcontroller [20]. The energy consumption of this type of processors can be as low as a few microWatts on both active and standby modes. With the microcontroller energy demands being that minuscule, the real attention is, justifiably, turned to the RF transceiver whose energy consumption for unit of data (received or transmitted) is the significant source of energy cost. For this reason, our results indicate the energy capacity indirectly, by presenting the number of bits that could have been transmitted.

On the software side of this problem, researchers have tried to minimize energy consumption, while still maximizing throughput. The main way this is done, is by duty cycling nodes in the network, changing the node from high consumption to low consumption states and back, according to a certain prescribed standard of operation. For long–term energy independence, the nodes have to stay in standby mode (low consumption) for most of the time, and only awake when they need to do something (high consumption) which in our approach is when they need to receive/transmit. In other words we consider the energy for the actual sensing to be negligible and leave outside the scope of the thesis the inclusion of an energy model to describe the part of the energy that goes to sensing and/or computation. There clearly exist other approaches as well. For example, according to one, nodes adapt to incoming available energy, executing more computation operations only when there is a surplus of energy, e.g., Vigorito et al. [21].

Previous works, such as that in [22], have assumed that the energy harvesting nodes are the leaf nodes of the network. This assumption simplifies matters as energy harvesting nodes are not burdened by traffic forwarding tasks. Our work assumes that all nodes (bar the sink) are exploiting energy harvesting. Clearly, if a leaf node runs out of energy it impacts only its own ability to collect and send data, but such depletion can have far more severe results if a node is part of the routing of data. What we point out in the thesis is that an interior node with limited energy harvesting output, depending on its location to the overall topology, can become the bottleneck of the entire system. Consequently, it is through the "balanced" use of interior nodes that we might be able to maximize the ability to acquire (or acquire

in a fair manner) data from the entire network. This balance can be expressed via a multi-path forwarding whereby a node's data can be split to follow multiple paths towards the sink.

In summary, the problem we study involves the combined effects of (a) the data collection as such and the well-known fact that nodes closer to the sink are depleted of their energy first as well as, (b), the effect of the variable time-dependent behavior of the energy harvested across different nodes. The added insights provided in this thesis stem from the fact that previous work, in attempting to treat the problem in a more abstract manner (e.g., the work carried out in energy efficient routing such as in [14, 23] and similar works) in that not all nodes start with the same energy reserves or are able to replenish to the same level their energy reserves.

A simplification we adopt is to consider that the network periodically (once per "cycle") collects and transmits the data – an assumption in line with the duty-cycling principle of operation. We consider various options for the period of this cycle. Clearly, our attention is not on delay. Data can be stored and not be forwarded for some time, waiting for the next "cycle" of communication. As it turns out, multi–hop communication aggravates the situation since the network may not have enough energy to forward all the data collected during a period within the next period's active ("ON" interval). We view delay as a secondary issue which does not enter our model. We note also that the per-cycle operation we propose has strong similarities to [24] in that, given the energy budget (or predictions thereof) a multi-commodity flow problem needs to be solved.

The solution of the optimization problem can be seen as instructions to each node to send a certain amount of data to a certain set of next hop destinations towards the sink, for each "flow". A flow is the data originating from one sensor that need to be delivered to the sink. Clearly, given the energy situation in the network, delivering the data of all flows may be infeasible and the optimization problems, depending on its exact formulation, favors one of many possible ways to allocate the available energy to the forwarding of flows. In this thesis we consider both a maximization of the total volume of traffic delivered to the sink formulation, as well as a formulation

which was not considered in [24], i.e., a fair (equal volume delivered from each source) formulation.

Additionally, the attention of [24] revolves around the distributed implementation of the algorithm which we forego for two reasons: (a) as in almost all WSN research, a supporting infrastructure is assumed, such as a capable sink and/or an additional backbone of computation resources outside the WSN – we use this infrastructure to solve the flow problems we have formulated and, (b), rather than spend energy sending extra control messages between nodes in the interest of an iterative distributed solution (shown in [24] to be $O(N^4)$ or worse), we send a short status update (and new energy level) message from each node to the sink (typically requiring $O(N^2)$ messages), thus simplifying the communication needs and leaving to the sink the burden of the computational problem (in our case, the solution of a, possibly large, LP problem).

Finally, a notable work about efficient operation of a network using energy harvesting appeared recently by Marasevic et al. [25]. There, the authors find the optimal routing of the data by using a water-filling algorithmic framework. Their solution maximizes the amount of data sent (which because it is in reference to a fixed length time "slot" – akin to the period we assume in our model – is often described as a "rate"), while ensuring fairness across nodes and across time, by making certain that no single node sends more data at the expense of others. A major modeling assumption they adopt is that the future levels of harvested energy availability for each node in the network are known, up to a certain time T. The solutions they present are, correspondingly, over a time duration of T. The adoption of a water–filling framework means that the rate of data transmitted by each node is gradually increased, and the rates of the nodes that cannot be increased any further without hurting the rates of the nodes that received lower rate are fixed and cannot further increase. It should be emphasized that the fairness expressed in this framework is across time T. They point out that time-dependent routing combined with splittable flows, which is the basis of this thesis, is the most challenging scenario. Hence, while solutions for fixed routing or for unsplittable flows are of manageable complexity, the suggested

time-dependent routing / splittable flows technique is both an approximation (albeit of a fully polynomial-time approximation nature) and it requires a massive amount of computation and variables. For illustration purposes, a step inside each water–filling iteration is the so-called "rate fixing" which in all varieties of the problem involves a time complexity proportional to the number of edges in the network or proportional to the number of edges, m multiplied by T. In the case of time-dependent routing and splittable flows, the corresponding step requires the solution of an LP of mT variables with nT constraints (where n is the number of nodes). Through our communication with the authors of [25] we have confirmed that indeed even the most trivial examples of networks and time durations are exceedingly computationally demanding to calculate and this is the reason that no such attempt was made to include computation results in their publication, but instead focused on the complexity analysis.

Clearly, the conclusions in [25] regarding the complexity analysis of this family of problems do not diminish the need to determine solutions of *practical* value where the future energy harvesting is not known and where the flows can be split and the routing can dynamically change from round to round. It is only in very controlled and predictable environments that the future energy harvesting potential can be assumed to be known. Examples exist where this (accurate) predictability is possible, e.g., harvesting from server machines in a server room [13]. There is a plethora environments that the harvested energy changes proportionally to phenomena that exhibit random or even chaotic behavior, like the weather (cloudy skies affect photovoltaic energy, temperature can affect thermoelectric, etc.). If anything, the work in [25] could be a good benchmark though against which to compare other routing schemes, but first its computation needs would have to be significantly reduced.

Chapter 3

Thermoelectric Harvesting Performance

A common problem in energy harvesting research is the use of synthetic data sets. The basis of our work are actual measurements of heat flow, based on which, two things can be determined: (a) whether the difference between indoor and outdoor temperature is a good proxy for the potential energy harvested and (b) whether off–the–shelf harvesters and wireless nodes can indeed convey an appreciable amount of data (or if the current technology has to improve by orders of magnitude before this is possible).

3.1 THE DATA SET

We use data collected over a period of a year, from eleven different apartments within a single apartment complex in Fort McMurray, Alberta, Canada. The data collected is comprehensive, including such aspects as water flow and temperature for the water used by radiators for heating, water flow and temperature for residential water, CO2 concentration, etc. For the purposes of this study we consider only the heat flow through exterior (outside-facing) walls, the indoor air temperature, and the outdoor air temperature. The data collection was conducted in real–time and is still taking place, but we extract a one year period (8th of September of 2012 to the 8th of September of 2013) which is sufficient for the purposes of capturing seasonal variations.

In our data set, we obtain a separate indoor air temperature for each apartment but have a single outdoor air temperature, as acquired by the Building Automation System (BAS). It has to be noted that a single outdoor air temperature is, again, only an approximation of the locally specific outdoor wall temperature of each wall unit, since phenomena like convection can, depending on airspeed, result in different temperatures at different spots and orientations.

The heat flow measurements are obtained at two locations (one on a stud, and one on the insulation) of the exterior–facing wall of each apartment. Of the two heat flow measurements the one most relevant to our study, which we subsequently use, is the heat flow via the stud. Studs are necessary for the structure and proper framing of the walls but at the same time they are responsible for loss of heat as they represent a "bridge" of smaller thermal resistance (compared to the insulated area of the wall) between interior and exterior.

3.1.1 The Methodological Approach

The heat flow through a wall reflects the combined results of wall construction qualities (stud spacing, insulation, etc.), of human activity (e.g. the choice of thermostat setpoints, opening of windows, etc.), of weather phenomena (temperature, wind direction, etc.), the particular orientation and location of the apartment, and of course the exact location of the sensor in the wall. As we will see, the combination of the factors leads to a highly variable heat flow which nevertheless exhibits distinct seasonal characteristics.

We use heat flow data to describe the extent to which thermoelectric energy harvesting through the exterior walls is adequate to power (and to what degree) WSN nodes. Heat flow (measured in W/m^2 units) through exterior walls is directly related to the temperature difference between the two sides of the wall. Heat flow through an infinitesimally narrow slice of surface is defined as, $\vec{q} = -k\nabla T$ where $\vec{q}$ is the local heat flow, k is the materials heat conductivity and ∇T is the temperature gradient. We note that heat flow is a vector. We use the convention that a positive heat flow represents loss of heat (i.e., radiating from the interior side of the wall to

the exterior side) while a negative indicates the reverse direction. Naturally, due to the climate characteristics in Fort McMurray, the latter case is infrequent, and occurs almost exclusively during the warmest summer months.

More precisely, if the wall is to be treated as a single homogeneous material of infinite area and thickness L the heat flow is inversely proportional to the thickness, that is, $\vec{q} = -kL^{-1}\nabla T$. In reality, the walls are more complex non-homogenous structures containing cavities, insulation, studs for wall support, studs for window framing, etc. impacting collectively on the heat flow magnitude and direction.

The thermoelectric harvesters are also dependent on the temperature differences to produce electricity and their output is governed by $-\nabla V = S\nabla T$ where S is the Seebeck coefficient of the material and ∇V is the electric potential between the two harvester terminals. We will examine the correlation between the heat flow data (q_{hf}) with the air temperature difference between indoor and outdoor (ΔT_{air}). We will subsequently use ΔT_{air} as a proxy of the actual temperatures of the two (inside and outside) wall surfaces. This is primarily because the air temperature represents averages whereas q_{hf} is specific to the location of the wall where the heat flow sensor is mounted. However, our reasoning is that any thermal harvester installed in the wall will experience similar heat transfer behaviour as the heat flow sensors measure at the same location. Hence, whereas ΔT_{air} is a good basis for an overall estimate of energy harvesting potential, regardless of where the harvester is placed on the exterior wall, q_{hf} allows us to examine the highly idiosyncratic behaviour (captured by the standard deviation) due to the factors we listed at the beginning of this section.

Our interest in ΔT_{air} and q_{hf} and their relation is also motivated by the intention to use, in a future design, a thermoelectric harvester inside the wall whose two sides are in contact to the two wall surfaces via materials of high thermal conductivity, e.g., metal. This would allow the thermal harvester to be a low thermal resistance "bridge" and hence receive the full benefit of the temperature difference of the two surfaces. However, such a design is future work because it needs to satisfy several other structural, mechanical, and safety constraints for in-wall embedding.

3.2 THE DEVICE MODEL

3.2.1 The Harvester Model

Our model is based on the performance of the TEC1-12703 Peltier module. The module has a surface area of 16 cm^2. The TEC1-12703 is designated as a thermoelectric cooler. Thermoelectric coolers are build on the same operating principles as thermoelectric harvesters. They use the same or similar materials and can be used both ways (as a harvester or a cooler). The main difference between the two is their optimal operating temperature. Modules that are made for harvesting tend to operate better at temperature ranges between 50 and 200 Celsius. The efficiency of the modules depends on their thermal and electrical conductivity which changes based on the ambient temperature. The reason that the harvesting modules are optimized for higher temperatures, is that common applications try to take advantage of the heat lost from machines operating at high temperatures. Another advantage of that module is in cost, with the price of these modules being to around 4 USD, compared harvesting modules that are normally magnitudes more expensive. Because the module is designated for cooling, its datasheet is really not that helpful for harvesting purposes.

We carried out characterization experiments to determine the relation between ΔT and energy harvested. The setup was a small "refrigerator", using TEC1-12703 modules (see Figure 3.1), arranged such that, a constant ΔT was created and the resulting energy harvested measured. This refrigerator is made by using (top to bottom): (a) a heat sink, which is connected to a TEC1-12703 (c), also connected to a power supply, to provide the hot side temperature for the harvester, (d) the harvester, with its other side connected to, (e) another TEC1-12703 module, again connected to the power supply to provide the cold side temperature. Lastly, (f)another heat sink along with a small fan was used to help regulate the heat and stabilize ΔT. This setup was used to maintain constant ΔT ranging up to 49 degrees Celsius. The temperatures were sensed using (b) thermistors (US Sensor USP12397), which have a resistance to temperature function, and where placed between the thermoelectric

modules in the middle of the surfaces, to get accurate temperature readings, and a separate sensor module to record the temperatures. Our findings indicate that the power output, W (in mW), of the particular harvester relates to ΔT (in Celsius) as $W = 0.0096\ \Delta T^2 + 0.2292\ \Delta T + 0.11$ (Least Squares fit with $R^2 = 0.9949$[1]). The power output was measured by calculating the current on a 120 Ohm resistive load.

The power output model is sufficient for determining the output current/power of a harvester but is insufficient for calculating the actual charge of an energy storage device, i.e., how fast could a (super) capacitor actually be charged, the leakage that a charging module can exhibit, and the general efficiency of said charging circuitry needs to be accounted. For this reason we connected the harvester to a BQ25504 Battery management module from Texas Instruments responsible to charge a 1 Farad super-capacitor. After running a series of experiments we determined that the energy harvested this way was better captured by $W = (2.57\ \Delta T^2 + 5.88\ \Delta T + 0.11) \cdot 10^{-3}$ (also in mW). Basically, we operated the refrigerator at stable temperature differences and slowly started charging the capacitor, measuring the voltage it held every few minutes/seconds (according to the operating ΔT). This gave us charge over time for different ΔT, and we inferred the energy and power of the system.

3.2.2 The Sensor Node Model

In order to develop a WSN node energy consumption model, we use the NanoZ-CC2530 device which employs the Texas Instruments CC2530 microcontroller. We operate it with another node acting as data sink, communicating using the z-stack tool [26] (Zigbee compliant). We carried out energy-exhaustion tests to determine for how long (how many bytes) of *payload* can be transmitted for the amount of energy accumulated to sufficiently charged 1 Farad capacitor (rated at $5\ V$) up to $3.6\ V$. The packets transmitted followed the standard Zigbee data frame structure. We first conducted experiments using the TEC1-12703 harvester connected to a Texas Instruments BQ25504 Evaluation Board to regulate the voltage and to determine the ability of the harvester to charge the capacitor. After the success of the first

[1]Notice that the fit is not (cannot be) perfect as at $\Delta T = 0$ we should be getting $W = 0$.

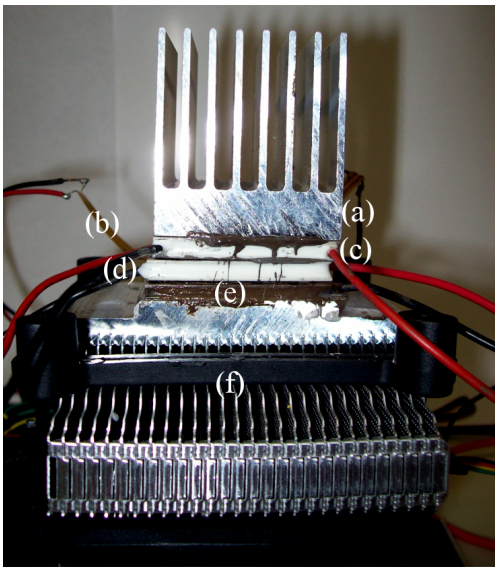

Figure 3.1: The harvester experimental setup. Shown in the figure are: (a) heat sink for temperature control of the "hot side", (b) thermistor for measuring the "hot side" temperature, (c) thermoelectric heater for the hot side, (d) the thermoelectric harvester, (e) the same as (c) but for generating the "cold side" temperature, (f) fan with heat sink for thermal control of the "cold side".

step, and in the interest of accelerating the experiments, we used a standard power supply (GPC-3030) to charge the capacitor to the same 3.6 V. The capacitor was then connected to the NanoZ module and used to send data until exhaustion. From the measurements we concluded that it was possible to transmit an average of 4103.7 bytes of payload per Joule using messages of maximum payload size (90 bytes per packet). The sensor modules are not energy efficient, as we measured that they consume 0.49 mW in their sleep state when the processor is set to the sleep mode PM2 (only low-frequency oscillator operating).

3.3 NUMERICAL RESULTS

We first consider the correlation of the ΔT_{air} and q_{hf} time series over the entire year and separately for each apartment ($Correlation(X,Y) = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\sum(x-\bar{x})^2 \sum(y-\bar{y})^2}}$). The time series represent hourly averages of the respective values. Table 3.1 demonstrates the overall strong correlation of the two time series but there exist differences (e.g.

Table 3.1: Correlation of ΔT_{air} and q_{hf}.

Apartment	$corr(\Delta T_{air}, q_{hf})$
1	0.91
2	0.95
3	0.91
4	0.89
5	0.89
6	0.91
7	0.87
8	0.93
9	0.87
10	0.88
11	0.93

0.95 for apartment 2 and 0.87 for apartment 4) that are ultimately related to the occupant(s) actions and behaviour. To illustrate the nature of differences consider the scatterplots of ΔT_{air} and q_{hf} in Figure 3.3 and 3.4 for the entire first year, showing an overall strong correlation but with significant variance and occasional outlier points.

Without precisely knowing the inhabitants behaviour, one can only conjecture on several reasons for certain outliers (the resident could have been using an electrical radiator close to the location of the heat flow sensor, or touching the wall at the sensor location, or had the windows open, etc.). To illustrate the differences that might show up, consider Figure 3.2 showing the heat flow during a specific day (in mid-March of 2013) during which one of the two presented apartments had almost constant heat flow (possibly the apartment was vacant that day) while the other one had highly variable heat flow (the oscillations are probably due to the heating cycling around the thermostat setpoint, while higher setpoints and possible resident activity are evident from approximately 6am to 12pm and from approximately 8pm to 10pm).

Let us define $s^{(d)}_{avg}$, $s^{(d)}_{max}$, and $s^{(d)}_{min}$ as, respectively, the average, maximum, and minimum of daily q_{hf} standard deviation, calculated based on the hourly measurements of each day. Essentially we try to capture the statistics of variability of q_{hf} within the same day as they behave across the year. As can be seen from Table 3.2 the remarkable fact is that there exist days that almost every apartment shows a dras-

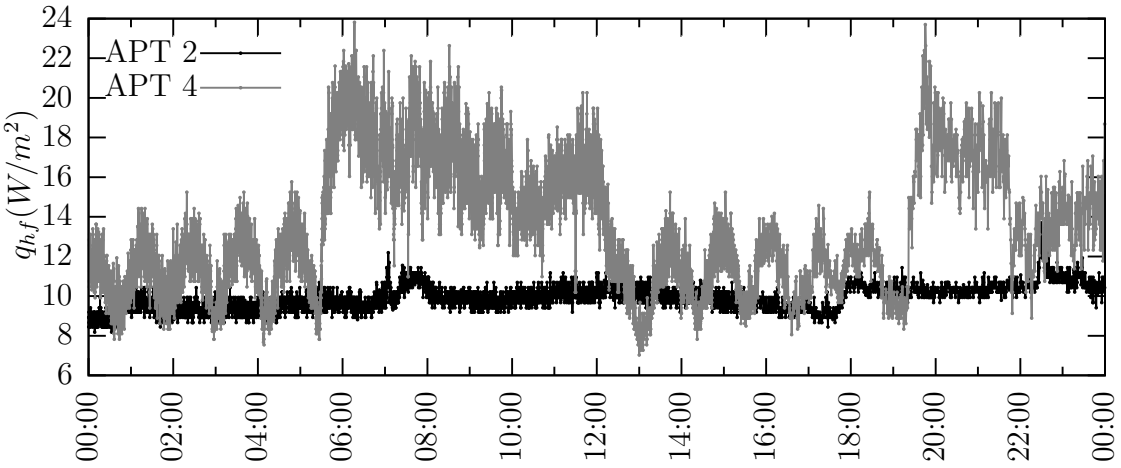

Figure 3.2: Intra-day q_{hf} in apartments 2 and 4.

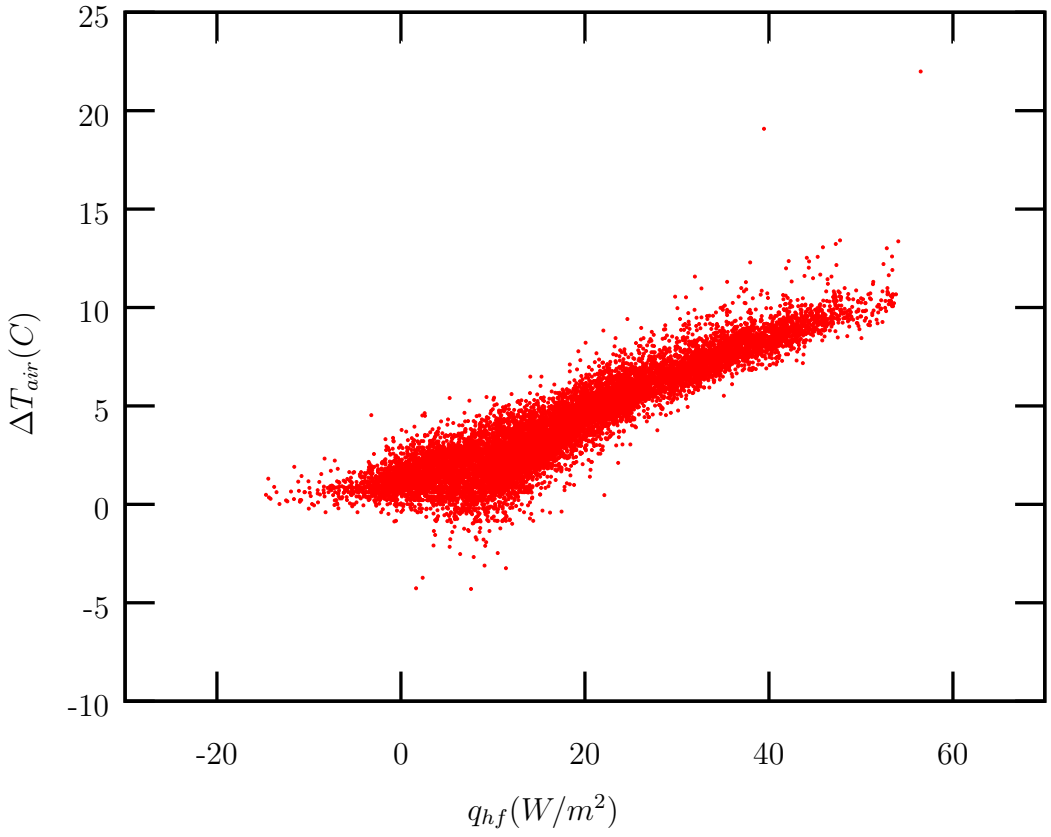

Figure 3.3: Apartment 2 ΔT_{air} vs. q_{hf}.

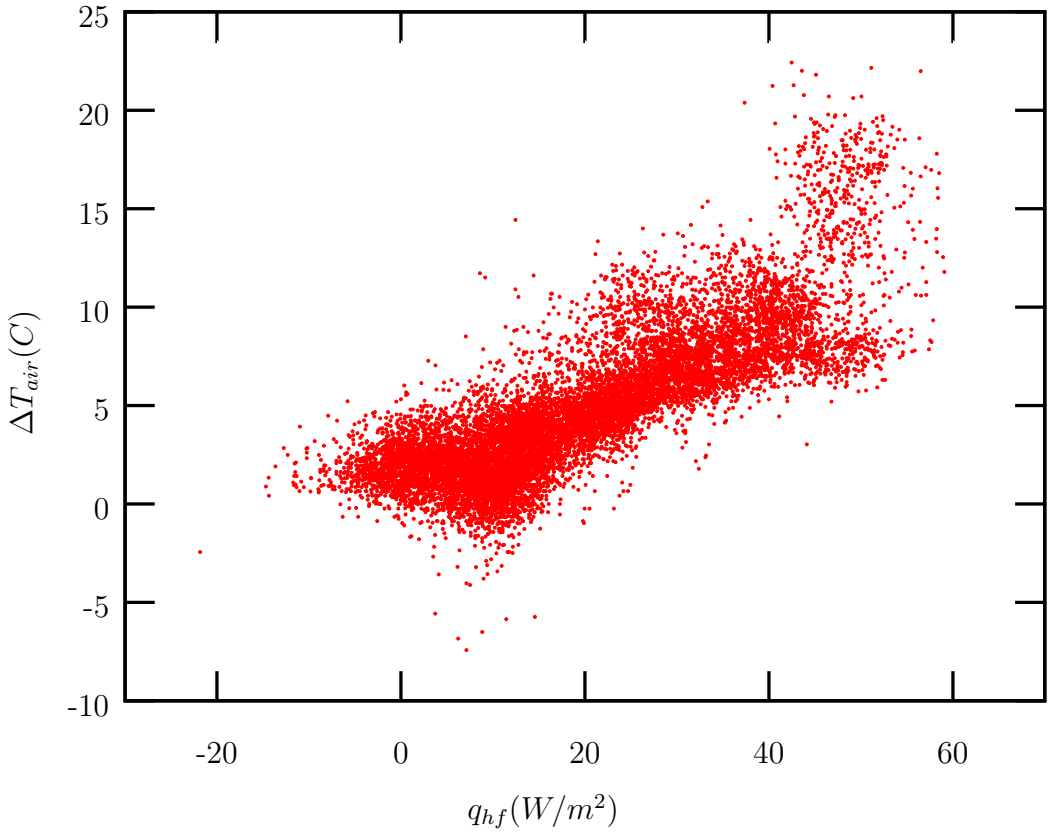

Figure 3.4: Apartment 4 ΔT_{air} vs. q_{hf}.

Table 3.2: q_{hf} and its standard deviation for various apartments, and potential harvesting output.

Apt.	$avg.\ q_{hf}$	$s^{(d)}_{avg}$	$s^{(d)}_{max}$	$s^{(d)}_{min}$	harvested (mW)	bytes/day
1	7.50	1.32	4.24	0.21	2.40	850943(677209)
2	6.02	0.99	2.95	0.21	1.87	663027(489292)
3	8.28	1.65	5.31	0.17	2.68	950220(776486)
4	6.85	1.47	5.01	0.26	2.2	780031(606297)
5	6.44	1.30	4.90	0.26	2.32	822579(648844)
6	8.67	2.41	7.34	0.27	2.47	875762(702028)
7	5.71	2.25	7.15	0.20	2.26	801305(627571)
8	8.41	1.42	4.96	0.16	2.58	914764(741030)
9	7.41	2.48	6.79	0.34	2.41	854489(680755)
10	7.40	3.01	10.0	0.42	2.36	836761(663027)
11	7.58	1.48	5.70	0.13	2.11	748121(574387)

tically small standard deviation. Apartment 11 exhibit a $s^{(d)}_{min}$ of 0.13 and even the apartment 11 which has the largest intra-day standard deviation of 10.0 still has low variability days as its $s^{(d)}_{min}$ of 0.42 illustrates this. Our current conjecture is that the days of low variability represent days that the apartments were possibly vacant, and hence no resident–related influence was introduced apart from leaving the thermostat at a particular (and possibly low) setpoint.

Next, we try to capture the differences across groups of apartments based on the floor and their orientation. To this end, we define $\sigma^{(y)}_{avg}$, $\sigma^{(y)}_{max}$ and $\sigma^{(y)}_{min}$, representing, respectively, the average, maximum and minimum of the standard deviation of the daily averages of particular groupings of apartments across the entire year. Table 3.3 (ϵ stands for a quantity less than 0.005) provides some interesting results. As expected, by comparing to Table 3.2, the overall variability is less pronounced at the larger time scale of a year as it dilutes the effects of intra-day variance seen in Table 3.2. While the maximum variability, i.e., $\sigma^{(y)}_{max}$, can still reach significant levels, it is less than the one observed intra-day.

The minimum, $\sigma^{(y)}_{min}$, reaches a small value which occurs during summer days when the outdoor and indoor temperatures are almost equal and the temperature across all apartments is similar as well. Unfortunately, less variable days across apartments are

Table 3.3: Average q_{hf} vs. orientation/elevation.

	$avg.q_{hf}$	$\sigma^{(y)}_{avg}$	$\sigma^{(y)}_{max}$	$\sigma^{(y)}_{min}$
Orientation				
North	7.30	1.05	3.53	0.07
South	7.30	0.84	3.13	0.15
Elevation				
1st floor	7.13	0.28	4.30	ϵ
2nd floor	7.37	0.75	3.64	0.03
3rd floor	6.93	0.86	3.42	0.03
4th floor	7.70	1.19	4.78	0.06

also days of small harvesting potential because ΔT is small. Furthermore, due to the occupant behaviour, we encounter cases such as the average q_{hf} at floor 3 and 4 being quite different (6.93 vs. 7.70). The implication of this observation is that, should multi-hop forwarding be used in the sensor nodes, the bottleneck (in terms of nodes with least harvested energy) could assume undesirable topological characteristics, by restricting the paths that could be followed to collect the data to sink nodes. In this example, an entire floor may not have enough energy to forward traffic between adjacent floors, towards a sink node placed at the bottom floor.

Despite the variability, seasonal patterns are evident across all apartments. Consider for example Figure 3.5 which shows the differences of two apartments (number 2 and 4) over the entire year. Indeed, even though the differences on certain days can be significant, they follow identical seasonal trends. The trends are also similar to all the other apartments (not shown here for the sake of brevity).

Finally, we note that there exist relatively important differences in the heat flow, and hence on harvesting potential, between night and day. To put it differently, the amount of energy collected in the morning hours is generally different from what could be collected overnight. Therefore, the intra-day variability could be handled by collecting energy over an entire (and possibly more than one) day, modulo of course the seasonal variations. For the night vs. day differences, we present the, rather arbitrarily chosen, intervals of day (6am to 6pm) and night (6pm to 6am). Hence, each day represents two (average) measurements. When considered in groupings of

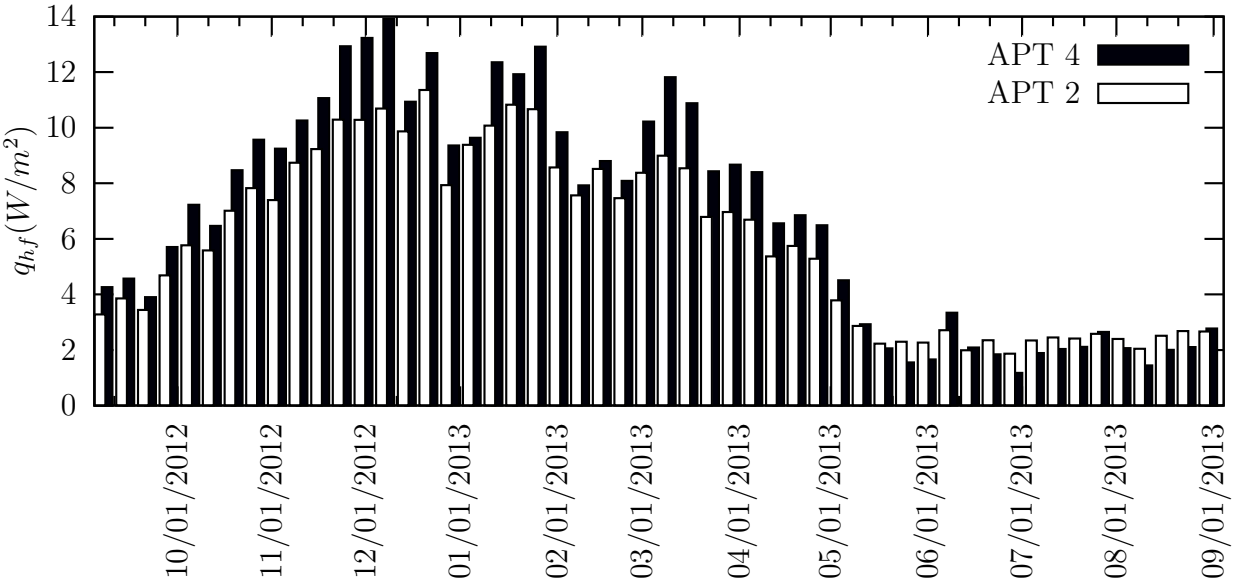

Figure 3.5: Daily average q_{hf} in apartments 2 and 4.

weeks, and across all weeks of the year, we produce Table 3.4 where we can readily see the ratio between maximum and minimum standard deviation can be significant and it is not uncommon to find a 7-times difference (apartment 6 is a good example with maximum over a week of 4.66 and minimum over another week of 0.67). Extending the averaging over a longer time scale will naturally smooth the variability extremes. For example, in Table 3.5, each entire day is represented by its average value, and the standard deviation of the daily measurements over a month (defined as a sequence of 30 days) are reported, across the entire year. The maximum variability is tamed but the difference between minimum and maximum standard deviation of the same apartment can still be surprisingly large (apartment 6 has a maximum of 3.55 and a minimum of only 0.74).

3.3.1 Data Transfer Capabilities

Let us return to Table 3.2 and notice the harvested power potential using ΔT_{air} and the model of the harvester noted earlier. The average power that can be harvested in the apartments exterior wall is 2.33 mW with a standard deviation of 0.21 mW. The highest average is 2.68 mW, and the apartment producing the lowest could

Table 3.4: Average, maximum and minimum weekly standard deviation of q_{hf}.

Apt.	$s_{avg}^{(w)}$	$s_{max}^{(w)}$	$s_{min}^{(w)}$
1	1.25	2.59	0.54
2	1.06	2.24	0.35
3	1.48	3.03	0.71
4	1.38	3.32	0.46
5	1.10	2.81	0.40
6	2.43	4.66	0.67
7	1.91	3.38	0.55
8	1.50	3.54	0.54
9	2.13	4.20	0.75
10	2.66	4.46	0.86
11	1.42	3.77	0.66

Table 3.5: Average, maximum and minimum monthly standard deviation of q_{hf}.

Apt.	$s_{avg}^{(m)}$	$s_{max}^{(m)}$	$s_{min}^{(m)}$
1	1.49	2.58	0.58
2	1.12	2.06	0.38
3	1.58	2.83	0.44
4	1.52	2.51	0.59
5	1.34	2.70	0.46
6	2.07	3.55	0.74
7	1.57	2.79	0.74
8	1.71	3.03	0.56
9	1.88	3.19	0.77
10	2.12	3.59	0.92
11	1.67	3.73	0.45

harvest 1.87 mW. Assuming we harvest energy to transmit once a day, each node could transmit an average of approximately 770 kbytes of payload per day. The per-apartment potential daily transfer volume can be seen in Table 3.2. The numbers in parentheses are the payload that could be transmitted per day assuming at all other times the sensor node idles as described in the node model, i.e., consuming 0.49 mW in its sleep state. Admittedly, much better performance is possible with better designed nodes. However, even at 600 kbytes of payload per day, a sensor can adequately send samples of its own sensing (e.g.. slowly changing humidity values or accelerometer activity compressed to the interesting events only) and still have some energy capacity to perform multi-hop routing.

However, as Figure 3.5 indicates, during the summer there might not be enough power to send data without risking an outage. Specifically, apartment 11 during the week of 6/29 to 7/5 was able to harvest only an average power of 0.113 mW, that is not sufficient to even power the sensor module. The exact sensor node design is also important. For example, assuming that an external circuit duty-cycles the operation of the entire node, then the restarting (equivalent to a cold boot) of the particular nodes we employed takes around 2 seconds to complete during which time it consumes the same power as when transmitting, 84.51 mW. After those 2 seconds the device enters sleep mode where it consumes $0.49 mW$. Hence, power-up is a costly overhead of $\sim 169.02 mJoule$, and a strategy of repeatedly powering up on-demand to gather data (not even transmitting) is probably unacceptable for the particular node design. Or, equivalently, the particular sensor would have to harvest energy for an average of 1497 seconds, to just cope with the 2 seconds startup energy cost before it performed any useful sampling, computation, and transmission (energy permitting), thus limiting the rate of sampling/sensing.

If the energy storage capacity is small to allow the longer term harvesting, outage is almost certain during summer months. The particular apartments do not have air-conditioning units for cooling, as they are rarely used in such northern climates. Thermoelectric harvesting during the summer occurs mostly at night when the outside temperature drops.

The good news is that due to the significant variability across apartments and throughout the day, a protocol to determine, at least locally, which node has (or has had in the recent past) the good fortune of harvesting more energy could be suitable for routing. In other words, there indeed exists diversity of opportunities to spend energy of another neighboring node because of the corresponding diversity in inhabitant behaviour. As a rule though, such routing strategies must become more conservative during the summer when the harvesting potential is reduced in both absolute numbers and in terms of variability across apartments. Our recommendation would therefore be in favour of "seasonally-aware" routing algorithms discussed in the next chapter.

Concluding, the potential of photovoltaic output from a cell of the same surface area ($16\,cm^2$) as the thermoelectric harvester used, at the same geographical location, and using off-the-shelf solar cells with efficiency 17% on a south-facing vertical wall is an average daily power of $88.739\,mW$ (according to data in Natural Resources Canada website [27]). The average power from thermoelectric harvesting appears low by comparison. However, one has to consider that (1) the solar power favors the south facing side of the building, over the north facing ones, and (2) to harvest solar energy the placement of the photovoltaic cells is crucial and one has to consider problems of occlusion of the light source, compared to a fairly flexible placement of the thermoelectric harvesters. Finally, as a matter of aesthetics, a photovoltaic harvester requires that it be exposed to outside view, whereas a thermoelectric harvester can be embedded "out of sight" within the wall structures.

Chapter 4

Seasonal Routing

In this chapter we develop a model and a subsequent time-varying routing strategy. The routing approach is based on the observation that the thermoelectric energy harvesting exhibits strong seasonality despite its overall variability. Hence, a previous year's routing decisions can be used as a routing guideline in subsequent years. We recommend routing strategies based on these observations, after we first examine the underlying abstract routing problem.

4.1 The Network Routing Model

Each sensor collects data independently of the rest. The data of each node are treated as a separate commodity. The goal of every node is to send as much data as possible with its current energy to the sink using, possibly, multiple paths. Instead of stipulating which paths are to be used and which ones are not, we pose the question as determining what fraction of traffic for each commodity should flow across each link with the purpose of either maximizing the total, or, (in the second version) the total concurrent volume of data delivered to the sink. By adopting a multi-commodity model, we do not force a particular routing, but rather we anticipate to observe that in optimal routing, when flows are split to traverse multiple paths, and we anticipate that such splitting will exhibit seasonal characteristics, *i.e.*, a particular link will be used certain times of the year and not at others.

4.1.1 Maximum Multi-commodity Flow Model

We consider the following multi-commodity maximization formulation of the routing problem on the n sensor nodes (with t denoting the sink):

$$\max \quad \sum_{i=1}^{n}(f_i(s_i,t)) \tag{4.1}$$

s. t.

$$f_i(s_i,t) = \sum_{w\in\mathcal{N}_{s_i}} f_i(s_i,w) = \sum_{w\in\mathcal{N}_t} f_i(w,t) \tag{4.2}$$

$$\sum_{w\in\mathcal{N}_u} f_i(u,w) = \sum_{v\in\mathcal{N}_u} f_i(v,u) \tag{4.3}$$

$$q\left(\sum_{i=1}^{n}\sum_{v\in\mathcal{N}_u} f_i(u,v)\right) + p\left(\sum_{i=1}^{n}\sum_{w\in\mathcal{N}_u} f_i(w,u)\right) \leq c(u) \tag{4.4}$$

Where $f_i(v,w)$ is the flow of commodity i from node v to node w. Exceptionally, the auxiliary notation $f_i(s_i,t)$ indicates the total flow from the origin of commodity i (node s_i) towards the sink t over possibly multiple hops and paths. $\mathcal{N}_u$ indicates the neighboring (adjacent) nodes to node u. We assume the number of nodes, minus the sink, is n. Equation 4.2 applies to all commodities i and indicates that the total flow out of the source and into the sink must be the same and is equal to $f_i(s_i,t)$. Equation 4.3 holds for each node u (other than the sink) and commodity i and represents the flow balance equation into and out of node u. Finally, equation 4.4 represents the constraint that the energy expended at node u cannot be more than $c(u)$ (the available energy). Here, q and p represent the ratios of energy spent per unit of flow for transmitting and receiving respectively.

By solving the above problem to determine $f_i(v,w)$ in each cycle, using an LP solver, we derive the maximum amount of data that can be sent with the current energy levels in the network. Using the computed solution, we can determine how much energy each sensor has to spend. We subtract that from the current energy of the nodes and then we move to the next cycle, where the amount of energy harvested is added to each node, and the multi-commodity flow is solved again. We note that,

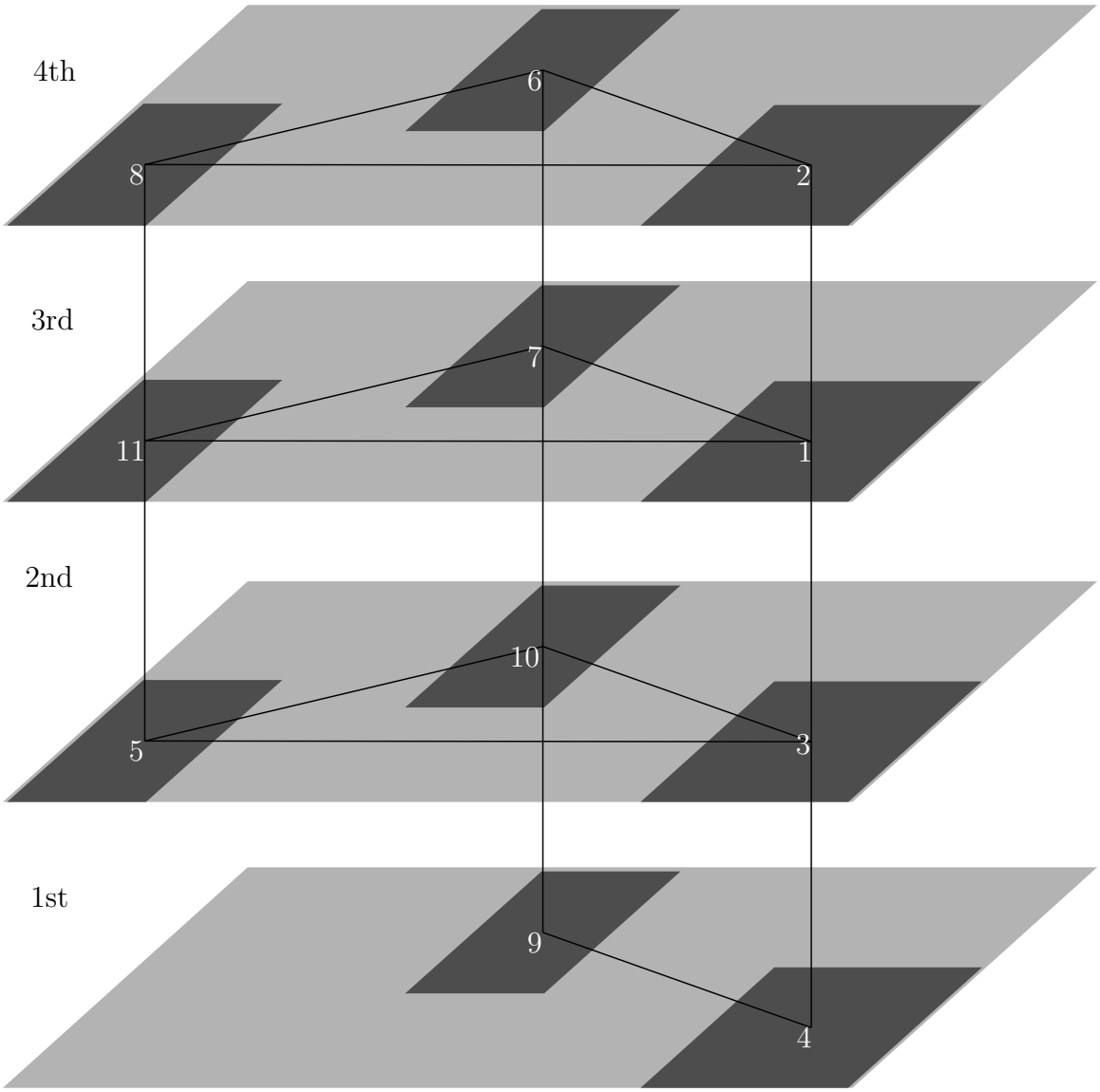

Figure 4.1: The example network topology across building floors.

as we noted in the introduction, the LP can be centrally solved at the sink and the results communicated to the nodes with negligible message overhead.

4.1.2 Maximum Concurrent Multi-commodity Flow Model

Since the maximization of the total flow leads to unfair results (demonstrated in the next section), we also consider a version whereby the objective is changed to the following one:

$$\max \quad f_1(s_1, t) \tag{4.5}$$

and we additionally introduce the constraint:

$$f_i(s_i, t) = f_1(s_1, t) \quad \forall i \tag{4.6}$$

The goal here is to maximize the total flow as long as all the commodities can deliver the same amount of data. Thus, the additional constraint ensures that no commodity is going to receive any worse service than any of the rest. Nevertheless, we expect this to happen at the detriment of the total flow. Furthermore, it should be clear that solving the concurrent flow problem does not necessarily result in the optimal use of the energy of the nodes. In the concurrent version there will always be sensors that have an excess of energy, which can lead to many different maximum solutions, some of them having wasteful energy expenditure, i.e., leaving drastically different (and possibly low) residual energy at the nodes. Consider the following example, taken from our sample network in Figure 4.1, that illustrates the problem: Node 3 routes flow 3 to node 5. Node 5 proceeds to route this flow to node 11 which then routes it along the path to the sink. Node 5 also routes flow 5 to node 3, which in turn routes it to node 1 to be routed along the path to the sink. If instead of this, node 3 routed flow 3 to node 2, and node 5 routed flow 5 to node 8, there would be less energy spent. Nodes 11 and 1, which are closer to the sink, would still need to route one complete commodity flow each, which for them would cost the same, but node 5 and 3, instead of receiving one flow and sending two out, will now just send one flow each. The reason that this is allowed is that since the nodes 5 and 3 are closer to the

edges of the network, they have a lot more residual energy, which allows them some flexibility on how to spend their energy. The problem is that unnecessary usage of energy like this, can lead to depletion of the energy of those nodes that could have been useful in future cycles.

To address this shortcoming, we optimize with respect to a secondary objective whose purpose is to maximize the residual energy of nodes in anticipation that it could be used in subsequent cycles. We remove wasteful solutions produced by the maximum concurrent formulation, by creating a second LP problem which, using the solution to the concurrent version (let's denote it by $f^\star$) explicitly minimizes the sum across all nodes of the consumed energy, as captured by equation 4.4. That is,

$$\min \sum_{u=1}^{n} \left(q \left(\sum_{i=1}^{n} \sum_{v \in \mathcal{N}_u} f_i(u,v) \right) + p \left(\sum_{i=1}^{n} \sum_{w \in \mathcal{N}_u} f_i(w,u) \right) \right) \tag{4.7}$$

s. t.

$$f^\star = f_1(s_1, t) \tag{4.8}$$

plus the additional constraints for flow conservation and source/sink flow summation we already presented. The reader should note however that this minimization takes place over the sum of energy expended across all nodes, with no specific attention to any single node.

A few technical remarks are in order: (a) we take the approach that solving off-line the optimization problem(s) and informing the nodes of the way to route data is acceptable because the topology is static and the information about the energy levels is relatively short and could be communicated to the sink (and from there to any optimization solution facility) at the beginning of the duty cycle and the nodes can be informed about the solution (delay is not a concern as the operation of the network is duty cycled anyway), and, (b) it is possible to generalize the fairness captured by the concurrent formulation to a weighted fairness by setting $f_i(s_i, t) = w_i f_1(s_1, t)$ where

w_i's are fixed weights, in particular when it is known that certain nodes produce a constant factor more data than others by virtue of the sensing they perform.

4.2 Evaluation

We use data collected over the period from 25th of June 2012 to 25th of June 2014 (the evaluation was done at a later stage, which gave us access to more data). The structure of our network is based on the structure of the apartment building that we have instrumented and from which the measurements are collected. There are 4 floors in the building, and 11 apartments are monitored. On every floor except the bottom level there are 3 apartments, two of which face South, while the third one faces to the North. For our network we assume that there is one node in each apartment attached to the exterior wall, and each node can communicate with all the nodes on the same floor, as well as with the nodes at the same location on the floor plan on adjacent floors. The topology of the network (without the sink) can be seen in Figure 4.1. We also consider two different sink node placements: one at the fourth floor and one at the second floor. A sink at a specific floor can communicate in one hop with all the sensors at the same floor. The reason for choosing these two floors for the sink placement is to have a location close to one extreme end of the building as well as one closer to the "middle".

For our experiments in this paper, q and p have been chosen based on ratios characterizing actual RF transceivers. In particular we adopt a model consistent with the Silicon Labs Si106x [28]. The q/p ratios are: $1.31, 2.12, 5.11, 5.47, 6.20$. For our duty-cycle, we have different timescales (per hour, twice per day, per day, per week and per month), which has given us a multitude of results. For this reason, we choose in each figure to present the most informative and general experiments, for each specific situation.

We started by trying to solve the routing using the maximum multi-commodity flow model. Initially, we assumed an unlimited energy storage capacity at the nodes, for the purpose of seeing how the harvested energy scales across time. Results are

shown in Figure 4.2. The results correspond to a daily duty cycling over one year (starting on the 25th of June 2012 and ending on the 25th of June 2013) and the solution of the multi-commodity flow on a daily basis. After finding the solution for each flow, we subtract the energy used from the energy already in the sensors. With this we move to the next cycle, we add the new energy harvested and run the problem again. We can see how the different q/p ratios bring different magnitude of results. Clearly the range of values is vast. More informative is Figure 4.3 which shows the residual energy at the nodes on the daily timeframe if the duty cycling was performed on a daily basis. We can see that the nodes 3, 5 and 10, which are next to the sink and represent a bottleneck, are normally out of energy, while all remaining sensors appear to have significant unused energy reserves. Indeed, if the objective is to maximize the delivered data to the sink, regardless of which sensor sends them, it is usually enough that all the nodes near the sink spend all of their energy in each cycle, trying to send only their commodity. This leads to the maximum amount of data, since there is no cost incurred by the sink's neighbors for receiving, consuming it exclusively to send data.

The apparent unfairness caused by maximizing the delivered data is rectified by considering the concurrent version of the multi-commodity flow problem. In Figures 4.4 and 4.5 we can see the difference in results between the maximum multi-commodity problem and the maximum concurrent flow. For ease of comparison we have added all the flows in the concurrent flow (essentially multiplying the flow by 11). We can see a more significant difference when the ratio is smaller. This happens because in both cases the sensors next to the sink (there are three of them) are the bottleneck of the routing. In the case of small q/p ratio, in the first version of the multi-commodity problem they can just use up all the energy transmitting, normalized by the ratio to the sink, ($energyofthenode = datasent * q$) while in the concurrent version the data that arrives to the sink has a total cost to the three nodes around it equal to $(p+q)*datasent*8/3+q*datasent$ (where 8 is the number of nodes that are not neighbors of the sink, and hence rely on those neighbors for routing). In the case of small ratio, q is smaller, therefore the amount of data sent scales better

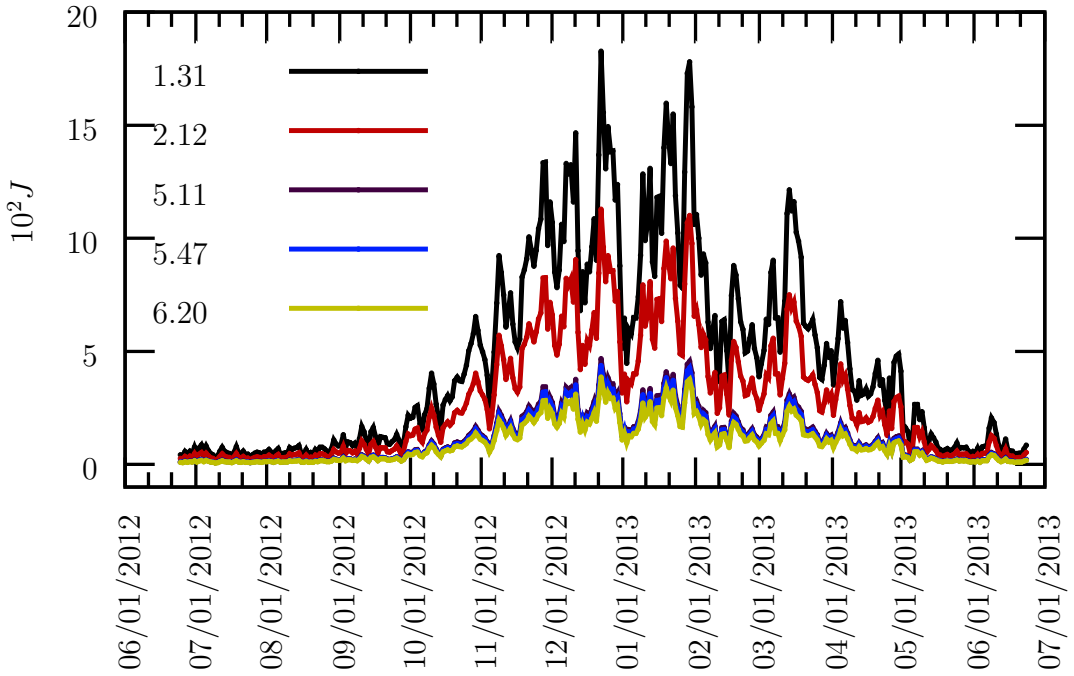

Figure 4.2: Maximum multi-commodity flow solutions (daily duty cycling, second floor sink, various q/p)).

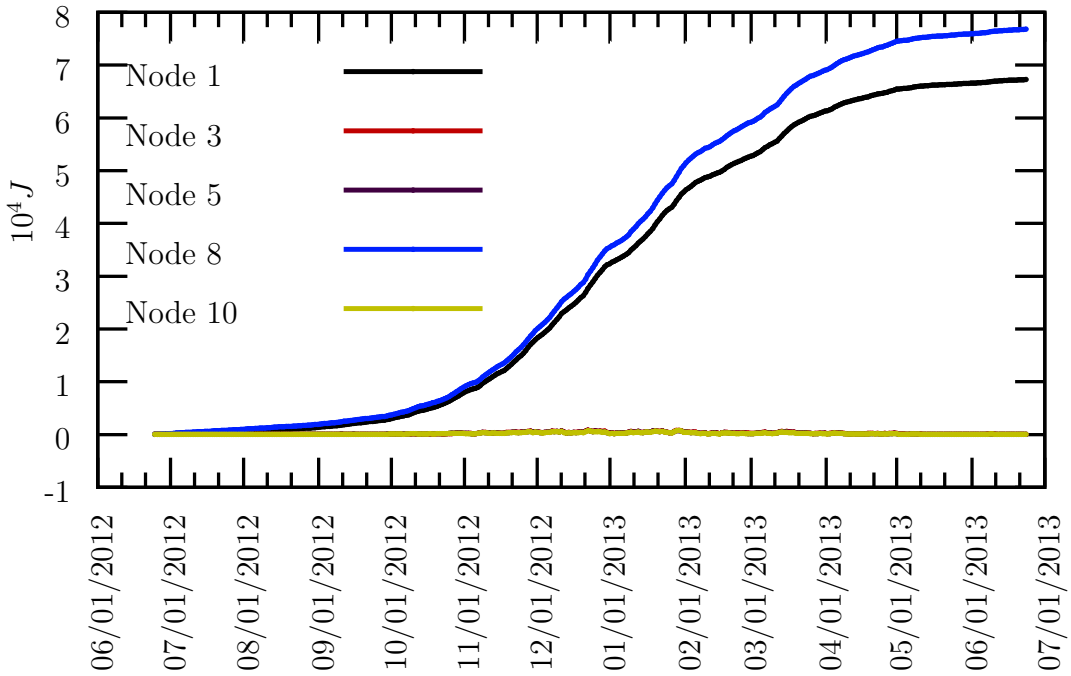

Figure 4.3: Residual energy at certain nodes (second floor sink, $q/p = 1.31$).

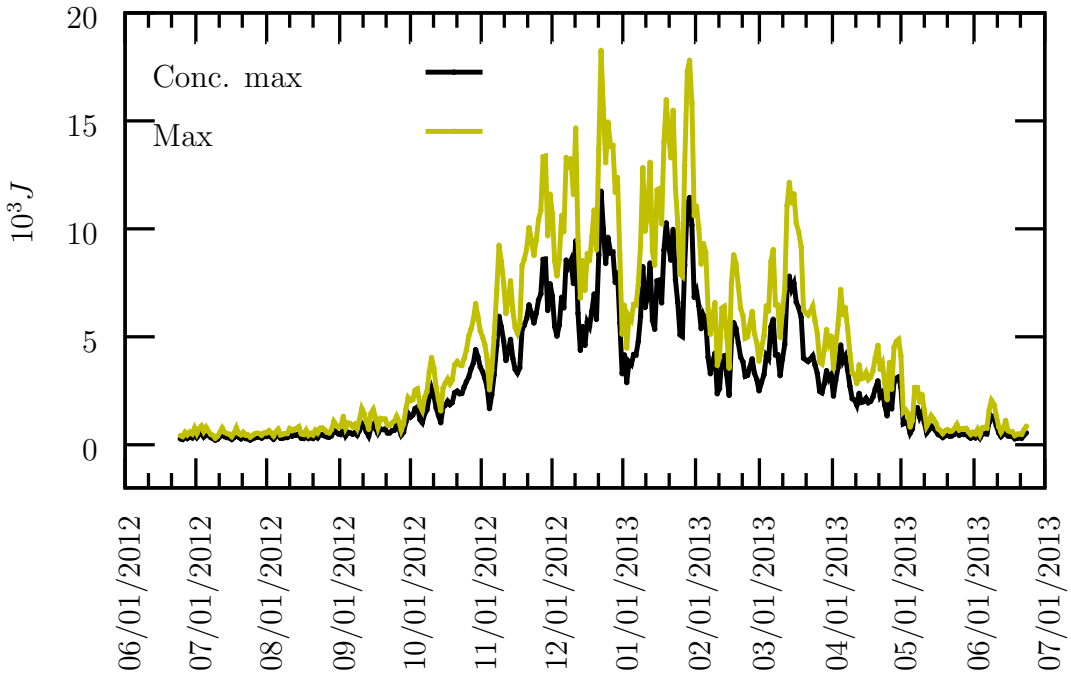

Figure 4.4: Comparison of maximum vs. concurrent maximum ($q/p = 1.31$).

for the first version. Even though the maximum multi-commodity problem provides a better total throughput, the concurrent flow is more useful due to the fact that the flows from all nodes are equal, hence fair.

In the third set of experiments we limited the energy harvesting storage capacity. We assumed a super capacitor of $5Farad$ and $5V$ that can store energy up to 62.5 *Joules*. The relevant results are shown in Figures 4.6 to 4.9. We compare the residual energy at certain nodes, when the sink is on the second floor. It is obvious that the capacitor adds a ceiling to the energy gathered. We can see that when the duty cycling is hourly, the limited capacity only affects the nodes farther away from the sink, since those nodes generally do not route much traffic through them, and since the maximum amount of data they send is a small portion to their overall en-

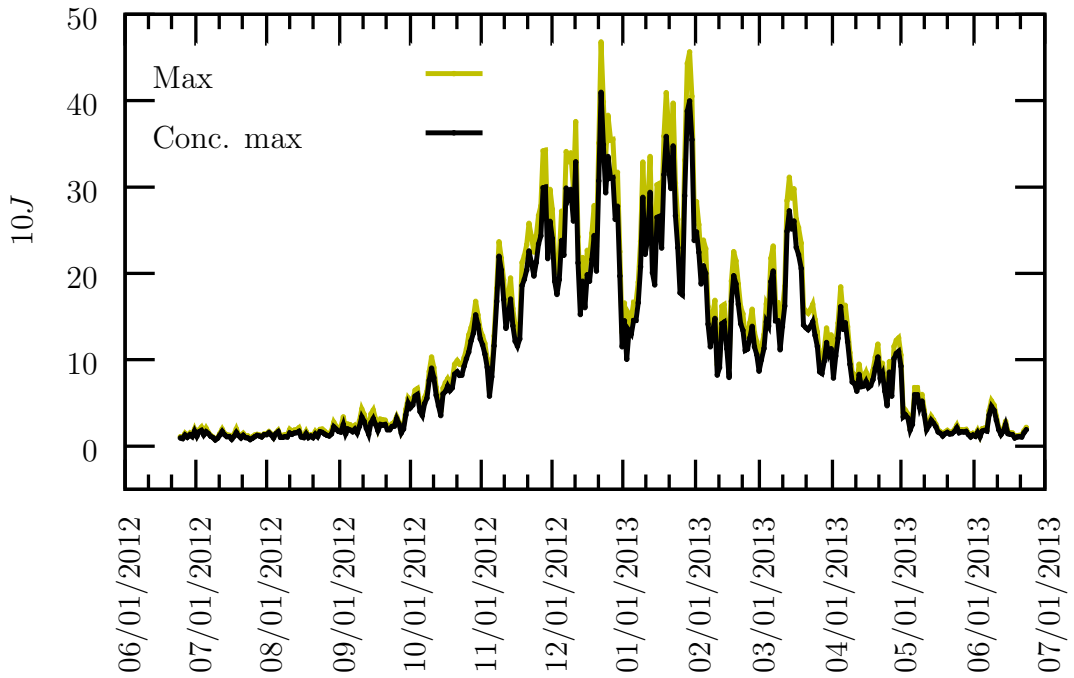

Figure 4.5: Comparison of maximum vs. concurrent maximum ($q/p = 5.11$).

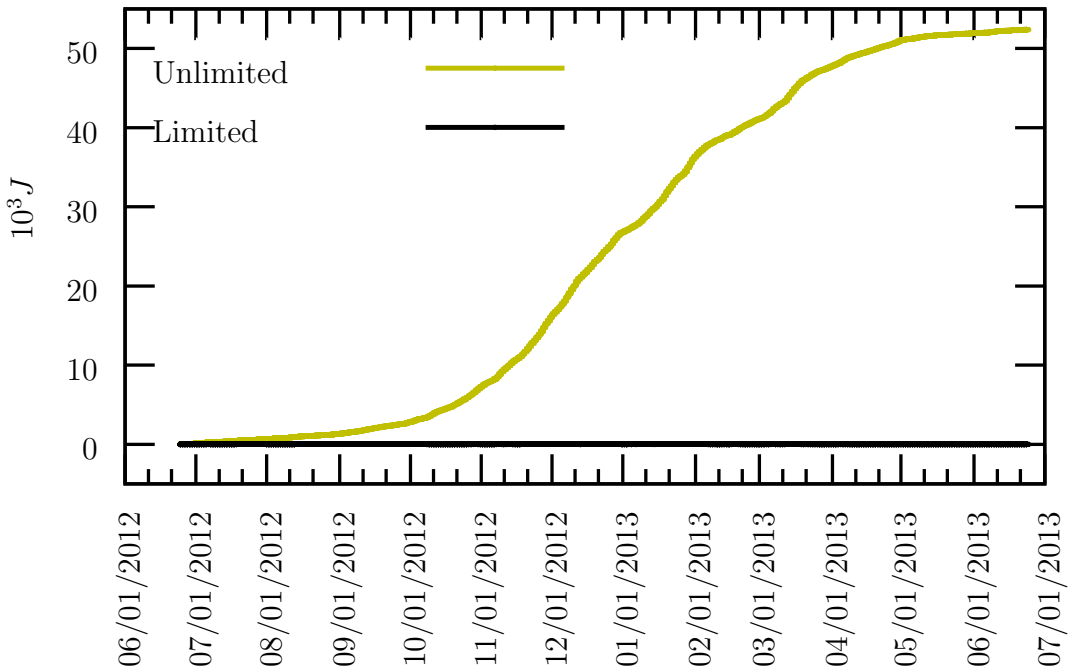

Figure 4.6: Energy at node 4 (1st floor) with and without limited capacity (the limited capacity line is imperceptible at, almost, 0).

ergy levels, they tend to build up excess energy (only limited by the finite capacitor). A good example of this behavior is node 4 at the first floor (Figure 4.6), which needs to transmit only its own data, by virtue of being at the outskirts of the topology, which means that it spends only a portion of the energy other nodes spend – leading it to accumulate a lot of energy over the course of a year.

Node 3, on the other hand, (Figure 4.7) is adjacent to the sink, which leads it to use all of its energy in every cycle. When the duty-cycle is per hour, we cannot even notice a difference to the amount of energy the node has in the beginning of each cycle. In Figures 4.8 and 4.9 we can see how much the ceiling of limited capacity affects the nodes that are not next to the sink.

Figures 4.10 and 4.11 demonstrate how the capacity limit affects the results at

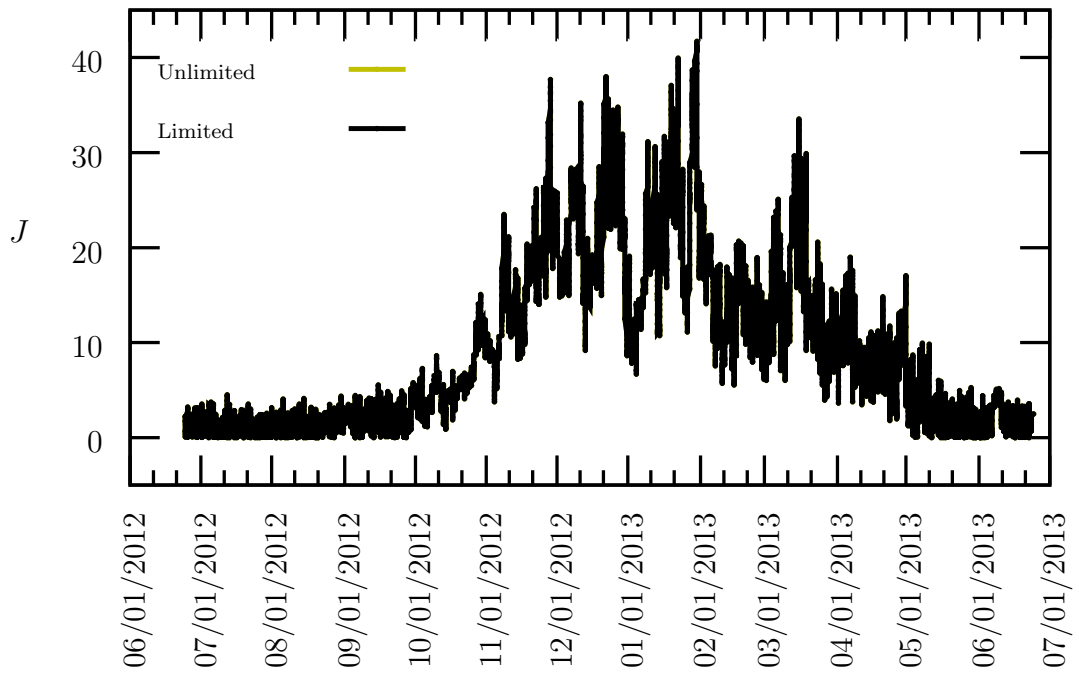

Figure 4.7: Energy at node 3 with and without limited capacity (no evident difference).

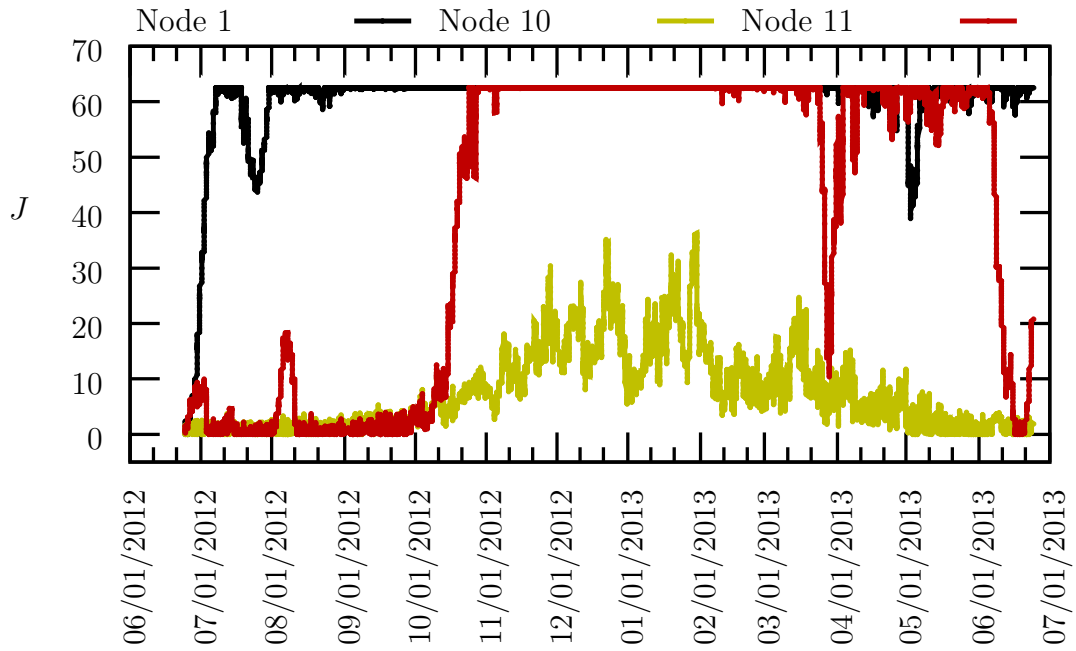

Figure 4.8: Energy at nodes 1, 10 (adjacent to sink), and 11 with limited capacity (node 10 is a bottleneck for routing).

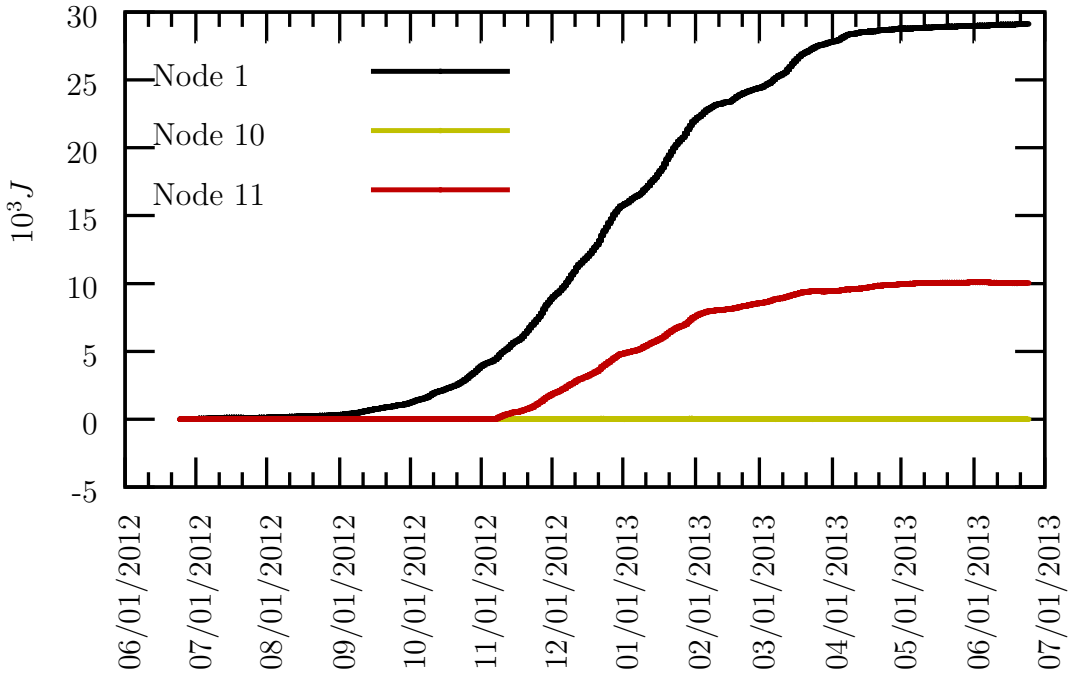

Figure 4.9: Energy at nodes 1, 10, and 11 without limited capacity (node 10 is the bottom line).

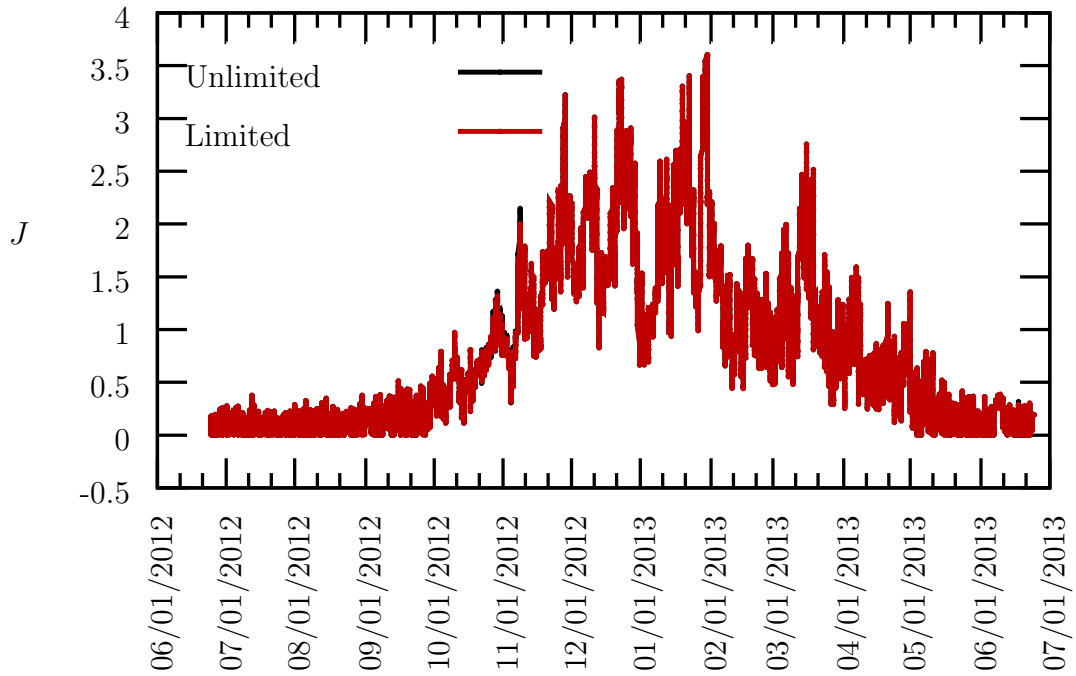

Figure 4.10: Hourly duty cycling with and without capacity limit (second floor sink).

hourly and twice per day duty cycling. In Figure 4.10, the timescale is hourly and the result of the multi-commodity routing does not change significantly. This happens because the bottleneck sensors (adjacent to the sink) cannot gather energy fast enough to reach the capacity limit. It is more interesting to see what happens when the timeframe is twice per day (Figure 4.11). We can see that for most part in the middle of the studied period (corresponding to cold months where indoor/outdoor temperature difference is significant, and hence harvesting is most productive), there is a certain limit to what the sensors can send. Ideally at that point we would use the excess energy for other functions, or even storing it to a long term power storage like a battery [13]. However, what is noticeable overall from the results so far, is that the extreme variance of energy harvesting and usage at different times.

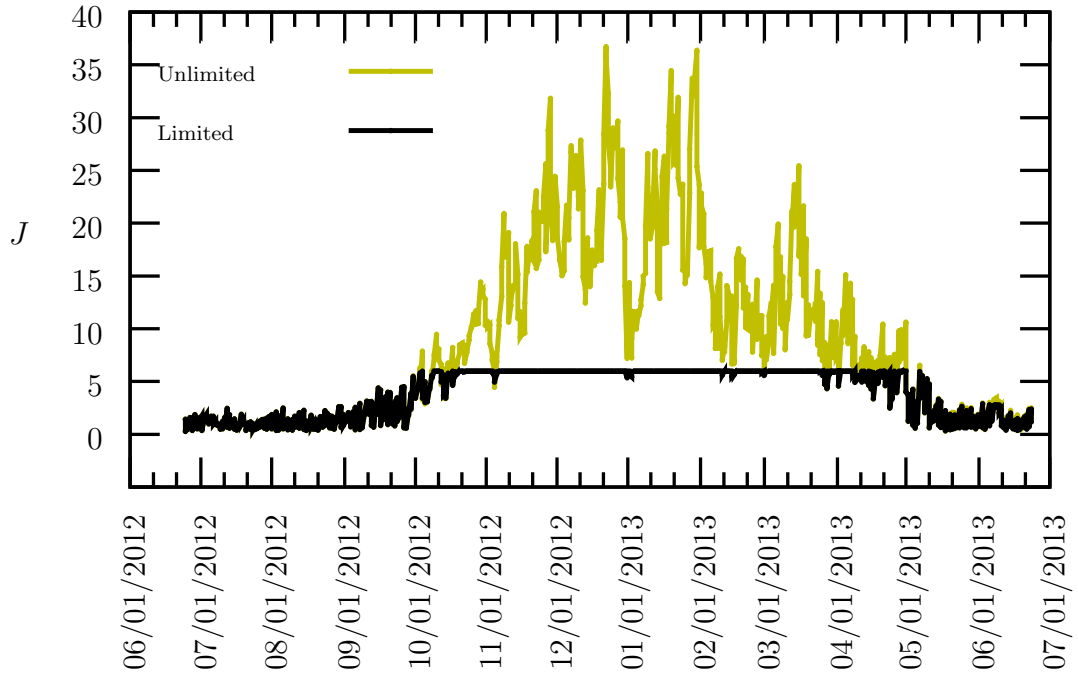

Figure 4.11: Flow difference with and without capacity limit ($q/p = 2.11$, twice per day duty cycling, fourth floor sink).

4.2.1 Dealing with Variability

In trying to solve the significant variance we witnessed in empirical results, we imposed the ad-hoc limit of using only 80% of the energy stipulated by the flow problem solutions. That is, we reserve 20% of each solution as backup that can then be part of the residual energy surviving into the next cycle. The simplicity of such a scheme was intentional, since one could easily program such a "safety margin". This led to some interesting results, which can be seen in Figure 4.12. The amount of data sent was very similar, as a whole, to the standard concurrent maximum problem – Figure 4.13 shows the results for duty cycling twice per day (every 12 hours). The 20% reserve means that there are not really completely depleted sensors in this case.

Tables 4.1 and 4.2 present data about what happens when we use the 20% reserve. In table 4.2 we can see that the flows in the system do not really change if we impose reserve if the duty cycling is on an hourly basis. The "bad days" continue to be very bad and the "good days" do not change towards the better. When the duty cycle is twice per day we find bigger differences, as is also shown in Figure 4.13 – the flows are slightly less than without the 20% reserve, which is the case because for a significant part of the year the capacity of the network is reached on every duty cycle. The interesting part is that the standard deviation is less, which means the network behavior is more predictable, but even more importantly we can see that the minimum flow is proportionally bigger than the one without the 20% reserve. This can be explained by the data in Table 4.1, where we can see the residual energy for the three nodes adjacent to the sink: on average they hold more energy in storage and the minimum energy left on them in the beginning of a duty cycle is significantly larger. It is this behavior that provides more consistency to the network, allowing for less chances that the nodes are depleted.

4.3 A Seasonal Routing Algorithm

Until this point we have been examining the potential of a centralized optimization execution to inform the nodes as to their routing decisions. We remark now on a

Table 4.1: Residual energy at nodes adjacent to the sink with and without the 20% reserve (twice per day duty cycle).

	node 3 w/ reserve	node 3 w/out reserve
Average	$46.58J$	$44.18J$
Std Dev.	20.23	22.21
Min	$7.55J$	$3.37J$
	node 5 w/ reserve	node 5 w/out reserve
Average	$50.69J$	$48.81J$
Std Dev.	17.96	19.58
Min	$5.02J$	$2.35J$
	node 10 w/ reserve	node 10 w/out reserve
Average	$44.44J$	$42.02J$
Std Dev.	21.84	23.33
Min	$4.97J$	$2.56J$

Table 4.2: Results for the concurrent maximum with and without the 20% reserve (hourly and twice per day duty cycle).

	Hourly	
	w/ reserve	w/out reserve
Average	$0.79J$	$0.79J$
Std Dev.	0.760	0.761
Min	$0.0009J$	$0.0001J$
Max	$3.59J$	$3.60J$
	Twice/day	
	w/ reserve	w/out reserve
Average	$3.45J$	$4.09J$
Std Dev.	1.643	2.20
Min	$0.502J$	$0.306J$
Max	$4.79J$	$5.99J$

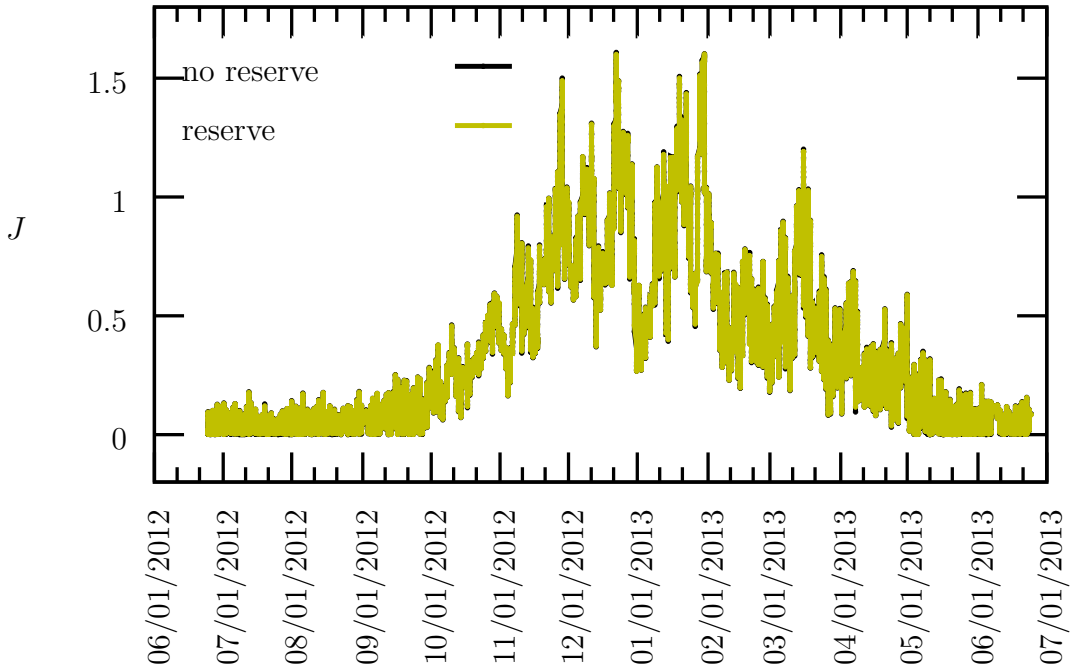

Figure 4.12: Hourly duty cycle, $q/p = 5.11$, with and without 20% reserve, sink at the fourth floor.

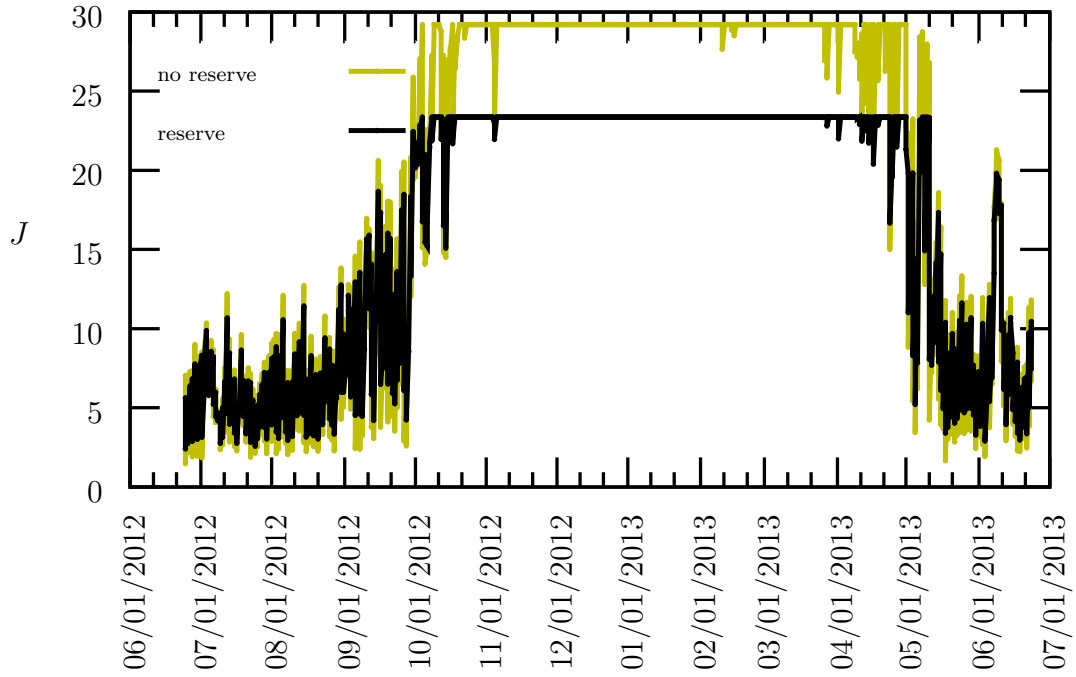

Figure 4.13: Twice daily duty cycle, with limited capacity, $q/p = 5.11$, sink at the second floor.

strategy whereby the routing can be guided by purely seasonal adjustments, independently by each node without the continuous assistance of an optimization to be run on the side. As stated in the work where we first commented on the collected data [8] and presented in the 3rd chapter of this thesis, thermoelectric energy harvesting follows a strongly seasonal pattern. We tried to take advantage of this quality of the data, by creating a very low overhead seasonal routing scheme. The idea here is that we "train" the sensors by using the data of the previous year(s), so that they can be prepared for the second year. The process can be extended to multiple years. In essence the multi-commodity flow problem is solved for the first year, and with it we create simple lookup tables (to guide routing) at the granularity of week or month that are downloaded to the sensors and based on which routing will be performed in the subsequent year(s).

In Figure 4.14 one can see an example of how a node can split its flows based on specific needs. This example has the sink at the second floor, so node 7 is 2 hops away from it (flow to 10 and then to the sink). The boxes on the figure represent how node 7 would split the total outbound flow from itself to the nodes 1,10 and 11 on the monthly timescale. Even though the shortest path to the sink lies with node 10 (See Figure 4.1) we can see that while node 7 sends the majority of its flow every month to node 10, to be forwarded straight to the sink, it also sends significant amounts to nodes 11 and 1 which are on the same distance to the sink as it is. This happens because node 10 does not have enough energy to route all the flows that pass through 7. For this reason the node sends some of the flows to 1 and 11 to be forwarded to nodes 5 and 3, who have some energy to spare, with an extra hop. This leads to more expenditure of energy but ends up with a greater amount of flow for the network.

Our approach consists of the following: we run the multi-commodity flow problem for our network based on the energy harvesting data from the previous year, for weekly and monthly timeframes, keeping a record of the split of the flows from each sensor (*e.g.* the previous example of node 7). Subsequently, in every timeframe of the second year (our validation set) each node splits its flow of data according to a lookup table created by the solutions from the previous year. In essence each

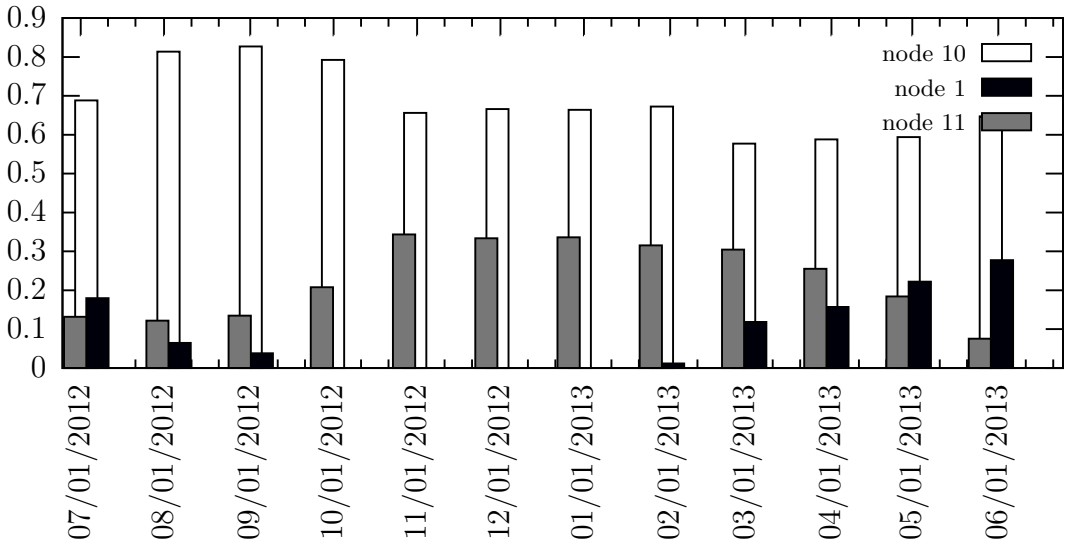

Figure 4.14: The split of flows from sensor 7 on a monthly basis to nodes 1, 10 (adjacent to the sink) and 11.

node holds a table that directs it to send data according to calendar information[1]. Additionally, for the amount of data that the node actually sends, the decision is taken according to how much concurrent flow the whole network can send, then reduced by 20% reserve, for the same reason of reducing the impact of variability as shown earlier. Figures 4.15 and 4.16 show the results from seasonal aware routing (for $q/p = 2.11$) where the season is a week and a month, respectively. The results are shown against the corresponding hourly duty cycling (with 20% reserve) concurrent maximum solutions. Our results indicate that the seasonal aware routing was capable of routing 86.6% of the optimal flow allocation as would have been determined by the concurrent maximum multi-commodity flow over the same year. The percentage was not noticeably different regardless of whether weekly or monthly season was used.

[1]The rows are the different seasons (we considered weeks and months as two alternatives), and the columns are the neighboring sensors. Each entry in the table is a ratio that expresses the fraction of the data the node has to send to that particular neighbor.

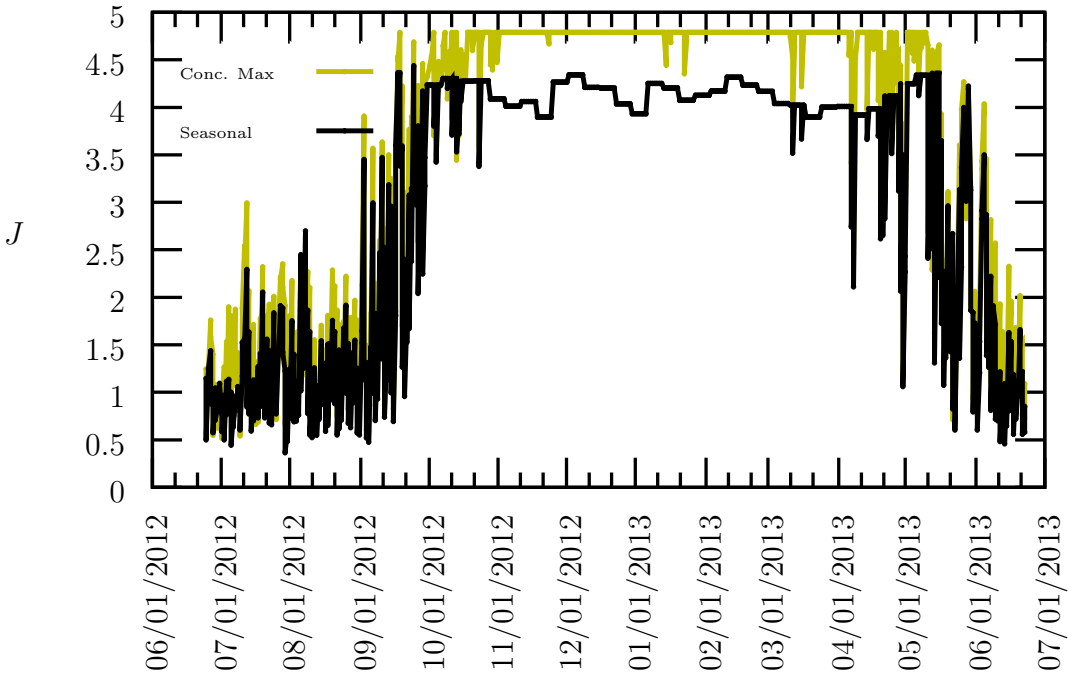

Figure 4.15: Twice per day duty cycling, weekly season, $q/p = 2.11$.

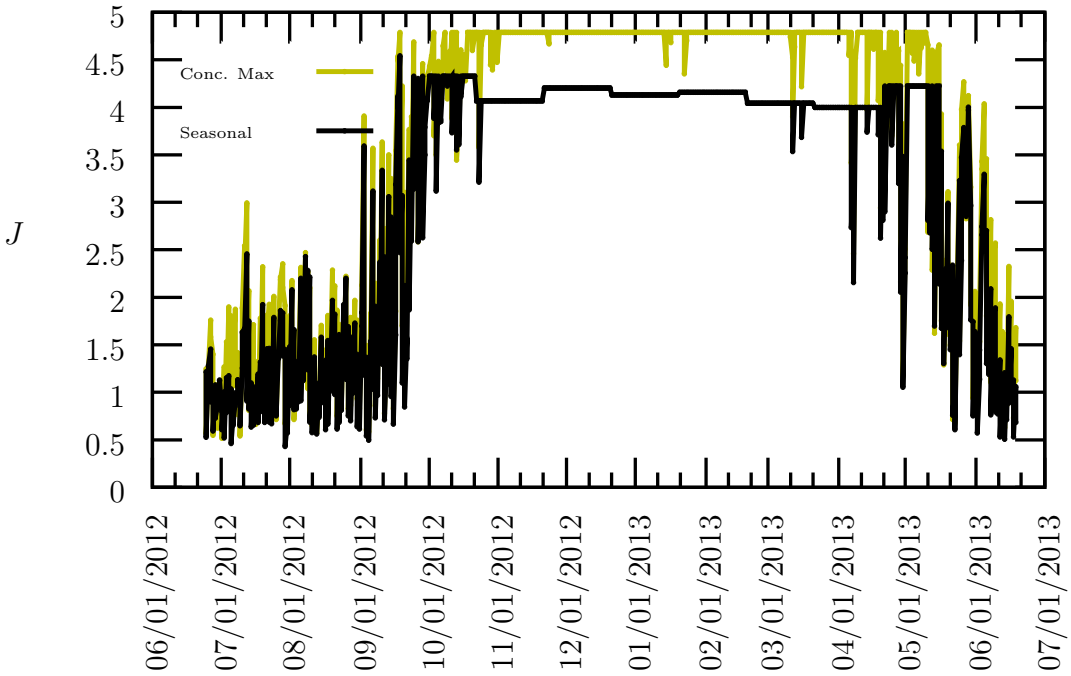

Figure 4.16: Twice per day duty cycling, monthly season, $q/p = 2.11$.

4.4 Discussion

4.4.1 Maximum vs. concurrent maximum

The solution to the multi-commodity problem we formulated has confirmed a well-known fact, that is, that the biggest bottleneck is at the nodes adjacent to the sink, as they are the ones who have to shoulder the routing of all the data to their final destination. As can be seen in Figure 4.9, other nodes might also struggle in the beginning with initially small deposits of energy, but after the winter season (where most harvesting happens) begins we can clearly see that the sensors start accumulating significant energy deposits (e.g., see nodes 11 and 1).

If the solution to the maximum concurrent multi-commodity flow problem is to be used as the centralized routing scheme, it should be noted that the sink should be placed at a location where its adjacent nodes are the ones mostly benefiting from energy harvesting. In Figure 4.17 we can see that if the sink is at the extreme side of the network (top floor of the apartment building), nodes before that part can also impose bottlenecks (here node 11 is the only node that feeds node 8, which up to a point of the year does not deplete its energy, because node 11 cannot forward enough data to do so).

4.4.2 Infinite vs. limited capacity

In a realistic setting the nodes possess a finite energy storage capacity, but as shown with Figure 4.6 they do not really need it. The limited capacity impact depends on the duty cycling period. If the duty cycle is small enough (hourly in our case) it does not result in any significant difference, and as it increases, changes become evident, especially in the high energy availability seasons (Canadian winter). In Figure 4.11 we can see an actual difference for twice a day duty cycling. There, fewer data are able to be transferred to the sink from every node throughout most of the winter period.

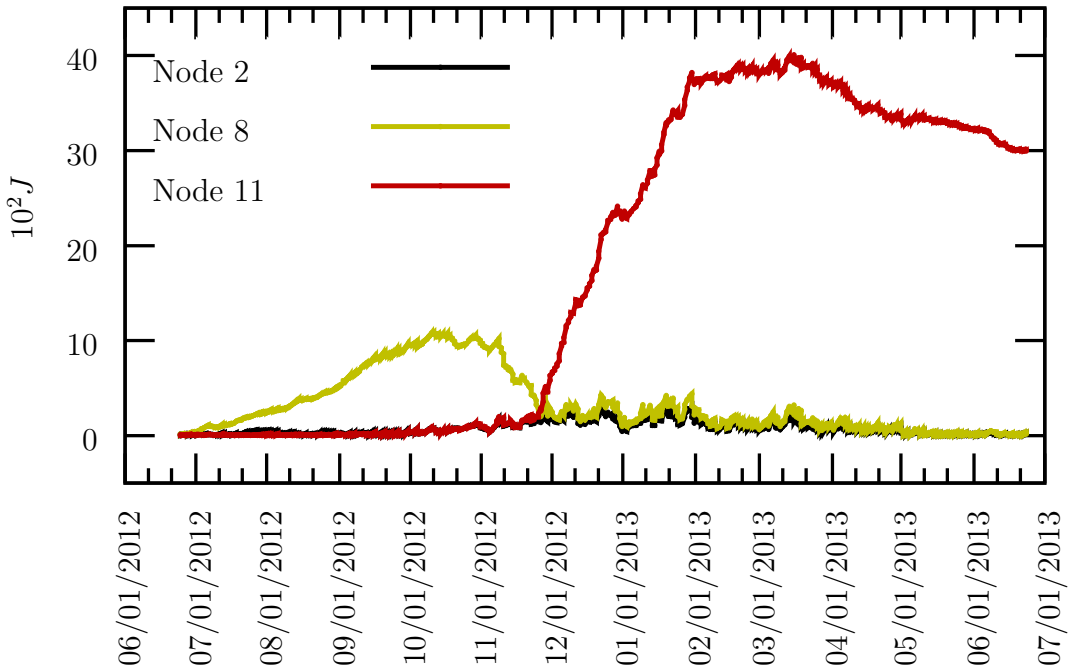

Figure 4.17: Residual energy at nodes 2 and 8 (both adjacent to the sink) and node 11 (3rd floor), with sink at the fourth floor.

4.4.3 20% reserve and limited capacity

We tried to even out the variance of the flows across different timeframes. Our solution to this was to engineer a 20% reserve of the energy indicated by the concurrent maximum solution, thus ensuring that there is always residual energy at all the nodes. Figure 4.18 illustrates this technique by showing how much energy exists in node 2 with and without the reserve. We can see that in almost every cycle the node maintains more energy in it, which results in the network being able to route more flow per cycle, so in the end we do not notice any significant difference in total flows routed compared to not having a 20% reserve. However, if nodes in the network reach the maximum capacity often, as in Figure 4.13, the scheme results in a decrease of the total amount of data sent, because, systematically, 20% less data are being transferred.

4.4.4 Seasonal routing vs. 20% reserve

The idea of seasonally–aware routing is that in every timeframe the node consults a lookup table to determine what "season" it is in, and blindly sends the data, spit accordingly, to the directions the table entry points to. According to our experiments, (see Figure 4.15) this seasonal scheme, was successful at routing the 86% of the optimal concurrent maximum flow even in the presence of a 20% reserve (to handle unanticipated variability). Without being difficult to compute and to install in the nodes, this routing scheme seems very promising, and eliminates the need to derive solutions to the multi–commodity flow problems in every duty cycle.

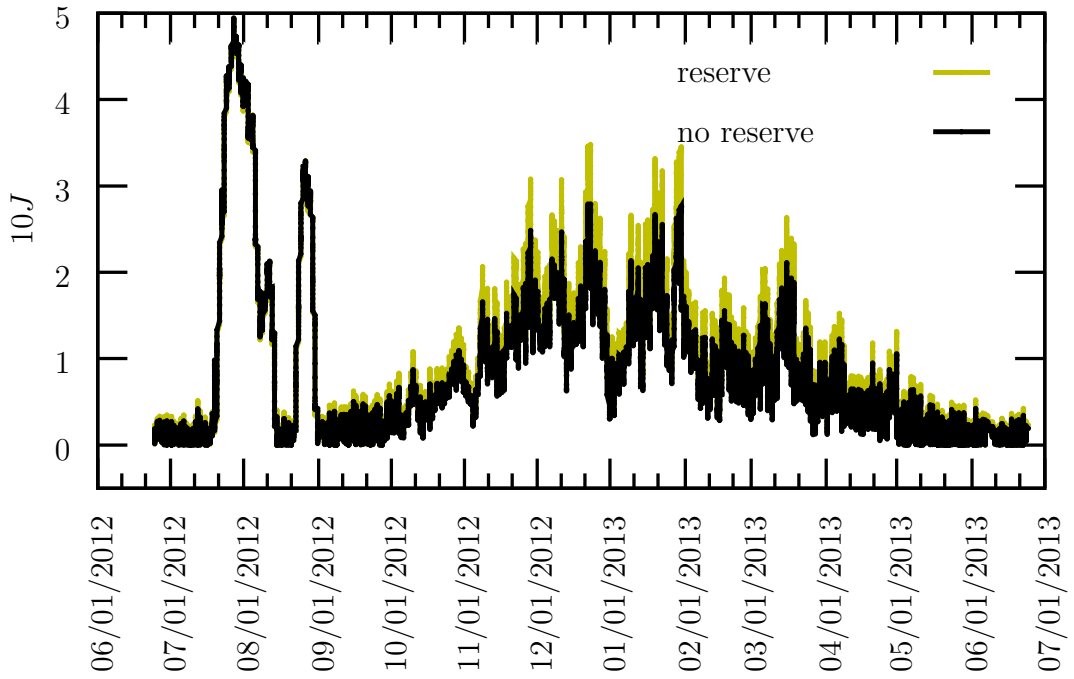

Figure 4.18: Node 2 with limited capacity, with and without the 20% reserve.

Chapter 5

Extensions and Future Work

5.1 Prediction

In the previous chapters we emphasized the need to produce solutions to the routing problem either (a) without any information about the future energy harvesting outcomes, or, (b) under the assumption that entire seasons will repeat, at least in a proportional sense to influence the routing choices in the same way each year. In this section we briefly examine an intermediate step, that of producing estimates for the near future based on an Autoregressive model. An accurate prediction model would, at least technically, allow us to see the pursue the solution for future routing decisions using the time-varying splittable flow model presented in [25]. If indeed those predictions are accurate, one could then attempt to develop more efficient computational techniques that those outlined in [25]. A secondary benefit is that, the generation of an autoregressive model would allow us to generate synthetic but realistic looking series of energy harvesting values.

We tried two methods for predicting the future harvested energy values: an Autoregressive Integrated Moving Average (ARIMA) model and a trivial method which predicts the next cycle to have the exact same harvested energy values as the previous one. The ARIMA model is an attempt to capture a more refined behavior of the physical process underlying the harvesting. Yet, it is also intuitively compelling that the trivial approach could be able to produce good results often enough. Both methods were chosen for their simplicity and ease of implementation and lack of complexity

(on the prediction part). This is useful because it lets us build nodes with the capability of predicting their own future energy availability and that of their neighboring nodes, which are the ones relevant to their routing decisions.

5.1.1 Predicting energy values

The idea behind forecasting by using the value of the previous time step comes from the fact that temperature changes gradually, and if the cycle period is small enough, the temperature and the temperature difference will not have changed a lot from one cycle to the next. With this idea in mind we calculate the mean square error (MSE) for this method and compare it to the ARIMA model in table 5.2. Specifically, the ARIMA model is fitted to the known part of the time series and it has the following form $(1-\sum_{i=1}^{p}\phi_i L^i)(1-L)^d X_t = c+(1+\sum_{i=1}^{q}\theta_i L^i)\epsilon_t$ where L is the lag operator and ϕ_i are the parameters of the autoregressive part and θ_i of the moving average part. c is a constant and ϵ_t is the error. The lag operator L is defined by $X_t = LX_{t+1}$. Table 5.1 summarizes the fitting MSE for each apartment, using a model with autoregressive coefficients at 1, 2, 4, 12, 36, 48, 60 and moving average coefficients at 1 through five, 12, 24, 48, 72. This model was chosen based on its fitness to the data. It was fitted separately to each apartment. The coefficients are determined by fitting the model over an entire year's collected data.

When the cycle is hourly and twice per day, we calculate the next value using hourly values – in the case of hourly by predicting the next hour, and in the case of twice per day by predicting the next 12 hourly values and averaging them. In the case of a daily cycle averaged daily data are used to make the predictions. In the table 5.1 we present the MSE when using an ARIMA model built based on the fit of a year's data to make the prediction across the cycles of the second year. The furthest back lags go in the hourly model is 72 hours in the past, relative to current time. The MSE for the hourly and the twice per day cycle can be seen for each apartment. While it can be seen that the error is bigger in the case of the longer cycle, it has to be noted that this happens mainly for the following reason: the model used for the hourly cycle needs to predict the temperature difference of the next hour, while the

twice per day needs to calculate it for the next 12, using the same data. On figure 5.1 we can see how close the predictions of the temperature differences are to the real data in the case of the hourly cycle in apartment 3.

From the tables we can see that predicting the values using ARIMA is close to the real data, but for the hourly cycle it is less accurate then the trivial prediction. Hence, it appears that, at least this particular ARIMA model is needlessly complicated for any benefit it might result in. In the case of daily and the twice per day duty cycle the ARIMA is more accurate than the trivial method, but the daily duty cycle exhibits a big enough error no matter the prediction. Clearly, the trivial method works well with the hourly cycle, because the temperature differences from hour to hour are minor most of the time.

Overall, we expect that the worsening of the prediction the further in the future we construct predictions for (regardless of the prediction technique) will seriously jeopardize the application of the water–filling method in [25], even if its computational complexity could be somehow reduced. A possible alternative is to consider cycles of length longer than a day, in the hope that the averaging of the energy harvested will be smoothed out and become more predictable. This however introduces two technical issues, first that the energy storage supported by the nodes might need to be significant, requiring appropriately sized super capacitors possibly in the order of tens of Farads, and secondly the traffic will be delayed by a significant amount of time before it can be delivered to the sink. It is debatable whether fairness across time is an absolute requirement when it comes at the cost of more expensive devices and delayed traffic collection.

5.2 Conclusions and Future Work

In anticipation of deploying in-wall wireless sensors for structure monitoring this thesis reports on measurements of temperature difference and heat flow, to evaluate the feasibility of using thermoelectric energy harvesting in Northern Climates. We conclude that the use of such harvesters for wireless sensor modules is possible with

Table 5.1: MSE of the ARIMA model for each apartment with hourly cycle (columns 1 and 2) or on twice per day cycle (column 3).

Apartment	fit Error (for hourly)	Hourly	Twice per day
1	2.8471	3.6421	14.807
2	2.6861	3.1263	11.996
3	3.1106	3.0894	13.020
4	2.7658	3.3727	13.743
5	3.7787	4.4287	12.449
6	2.7888	3.2392	13.221
7	4.4242	5.4527	14.486
8	2.6012	2.7467	13.637
9	2.8556	3.0305	11.722
10	2.8666	3.0076	10.137
11	3.7000	3.7872	20.825

Table 5.2: MSE for ARIMA versus trivial (prediction equal to current) for apartment 3.

Cycle	ARIMA	Trivial
Hourly	3.089	2.273
Twice per day	13.020	34.089
Once per day	20.8756	21.689

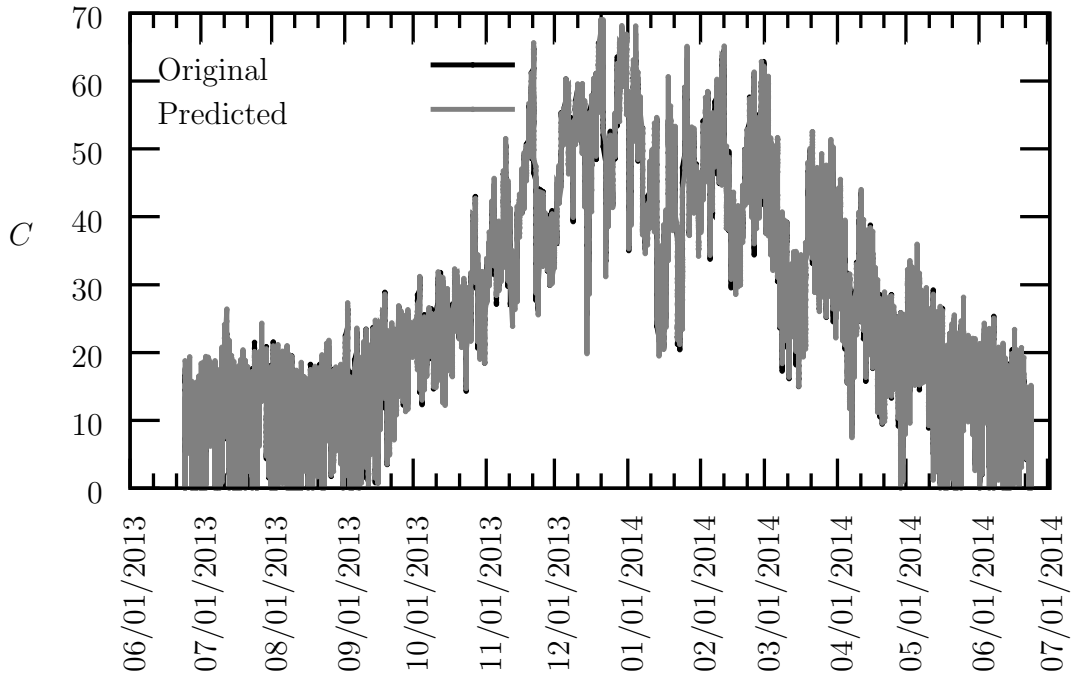

Figure 5.1: Hourly cycle predicted values vs. actual for the second year.

current technology albeit not without some challenges.

First, even though the difference between indoor and outdoor air temperature is a good guide for the potential energy harvesting, the exact behaviour of heat flow that ultimately governs the thermoelectric harvesting potential at particular points of the wall structure are subject to factors that are dependent on weather phenomena (e.g., convection phenomena on the wall surfaces) but, more importantly, on the behaviour patterns of the residents who are in control of not only the thermostat setpoints but also of objects attached to or close to the walls, extra forms of directional heat radiators (e.g. electric heaters), and so on.

Second, even though the heat flow follows, as expected, a seasonal pattern, its variability from one outside wall to another (one apartment to another) can be significant when observed in small time scales, e.g., intra-day. Hence, despite the regularity of the harvesting potential, it is to be expected even within the same day, that the wall on some apartments can sustain higher volume of sensor data transfers than others. The implications of this behaviour to multi-hop routing are obvious. Short-term alternative routing paths would need to be considered.

In the particular environments we consider, the harvesting potential is reduced in the summer months moving almost exclusively to later in the day. Additionally, in the summer months, the variability of harvesting becomes smaller. Hence, when sizing the energy storage (e.g., capacitance of a supercapacitor) necessary to sustain the sensor node operations, we should size based on the worst case summer harvesting potential. One could argue that we need to consider complementing thermoelectric harvesting with photovoltaic harvesting which reaches its peak output during the summer months. Nevertheless, such a decision is dependent on a decision to expose the photovoltaic element which we try to completely avoid as we wish to deploy as inconspicuous sensors as possible.

We also contributed a model and a solution to the routing through an energy harvesting network, by formulating multi-commodity flow problems. The multi-commodity flow model can be used, along with its slight variations, as a centralized routing scheme, where the sink/central controller of the network decides how the flow

of data is going to be routed inside the network, after getting all the information about the residual energy at each of the nodes of the network. We additionally proposed a decentralized (to the extent that each node needs to be equipped with a simple lookup table) seasonally aware scheme based on the concurrent multi-commodity flow problem, which can run individually at each node.

Predicting the energy that will be harvested in the future can further help the basic idea of seasonal routing and hopefully tap the benefits of time-varying, flow splitting routing that, at least in theory, have been noted by other researchers. Early indications are not very encouraging, at least to the extent that they rely on a simple ARIMA model. Further research in this direction is required. It is not far fetched to also assume that the nodes could be informed (from listening occasionally to a "weather channel") of the upcoming outdoor temperature short-term predictions and tune their energy consumption accordingly.

A full implementation The idea would be for the node to predict its neighbours energy. The lookup table idea at each node can be extended with some cutoff ratios, where the node stops sending traffic to some of the adjacent nodes, because their energy is small enough, that, had we solved the multi-commodity problem, its solution would have chosen to ignore those node for routing particular flows.

Finally, in the future we would like to modify the multi-commodity flow formulation so that it tries to optimize the residual energy in the nodes for use in the next cycle, but with the added knowledge that it distinguishes where this residual energy would bring more benefit (i.e., at nodes adjacent to the sink). A useful concept is the so–called "energy-neutral" operation, from previous research works [21, 29], which was proposed to achieve perpetual operation of the nodes. However any such extensions will have to go hand–in–hand with the integration of a new self–contained thermoelectric harvesting node for in-wall use, because by the time this research was completed, new devices with even lower energy requirements were being made widely available.

Bibliography

[1] R. J. Vullers, R. Schaijk, H. J. Visser, J. Penders, and C. V. Hoof, "Energy harvesting for autonomous wireless sensor networks," *Solid-State Circuits Magazine, IEEE*, vol. 2, no. 2, pp. 29–38, 2010.

[2] N. Xu, S. Rangwala, K. K. Chintalapudi, D. Ganesan, A. Broad, R. Govindan, and D. Estrin, "A wireless sensor network for structural monitoring," in *Proceedings of the 2nd international conference on Embedded networked sensor systems*. ACM, 2004, pp. 13–24.

[3] M. Ceriotti, L. Mottola, G. P. Picco, A. L. Murphy, S. Guna, M. Corra, M. Pozzi, D. Zonta, and P. Zanon, "Monitoring heritage buildings with wireless sensor networks: The torre aquila deployment," in *Proceedings of the 2009 International Conference on Information Processing in Sensor Networks*. IEEE Computer Society, 2009, pp. 277–288.

[4] S. Kim, S. Pakzad, D. Culler, J. Demmel, G. Fenves, S. Glaser, and M. Turon, "Health monitoring of civil infrastructures using wireless sensor networks," in *Information Processing in Sensor Networks, 2007. IPSN 2007. 6th International Symposium on*. IEEE, 2007, pp. 254–263.

[5] M. Gorlatova, A. Wallwater, and G. Zussman, "Networking low-power energy harvesting devices: Measurements and algorithms," in *INFOCOM, 2011 Proceedings IEEE*, 2011, pp. 1602–1610.

[6] R. Rao, S. Vrudhula, and D. N. Rakhmatov, "Battery modeling for energy aware system design," *Computer*, vol. 36, no. 12, pp. 77–87, 2003.

[7] F. Simjee and P. H. Chou, "Everlast: long-life, supercapacitor-operated wireless sensor node," in *Low Power Electronics and Design, 2006. ISLPED'06. Proceedings of the 2006 International Symposium on.* IEEE, 2006, pp. 197–202.

[8] A. Kollias and I. Nikolaidis, "In-wall thermoelectric harvesting for wireless sensor networks," in *Proceedings of the 3rd International Conference on Smart Grids and Green IT Systems*, 2014, pp. 213–221. [Online]. Available: http://www.scitepress.org/DigitalLibrary/Link.aspx?doi=10.5220/0004864102130221

[9] ——, "Seasonally aware routing for thermoelectric energy harvesting wireless sensor networks," in *Smart Cities and Green ICT Systems (SMARTGREENS), 2015 International Conference on.* IEEE, 2015, pp. 1–11.

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

[11] M. Gorlatova, P. Kinget, I. Kymissis, D. Rubenstein, X. Wang, and G. Zussman, "Challenge: ultra-low-power energy-harvesting active networked tags (enhants)," in *Proceedings of the 15th annual international conference on Mobile computing and networking.* ACM, 2009, pp. 253–260.

[12] L. Mateu, C. Codrea, N. Lucas, M. Pollak, and P. Spies, "Energy harvesting for wireless communication systems using thermogenerators," in *Proceeding of the XXI Conference on Design of Circuits and Integrated Systems (DCIS), Barcelona, Spain*, 2006.

[13] L. Rizzon, M. Rossi, R. Passerone, and D. Brunelli, "Wireless sensor networks for environmental monitoring powered by microprocessors heat dissipation," in *Proceedings of the 1st International Workshop on Energy Neutral Sensing Systems*, ser. ENSSys '13. New York, NY, USA: ACM, 2013, p. 8:1178:6. [Online]. Available: http://doi.acm.org/10.1145/2534208.2534216

[14] S. Singh, M. Woo, and C. S. Raghavendra, "Power-aware routing in mobile ad hoc networks," in *Proceedings of the 4th annual ACM/IEEE international conference on Mobile computing and networking*. ACM, 1998, pp. 181–190.

[15] N. S. Hudak and G. G. Amatucci, "Small-scale energy harvesting through thermoelectric, vibration, and radiofrequency power conversion," *Journal of Applied Physics*, vol. 103, no. 10, pp. 101 301–101 301, 2008.

[16] E. J. Luber, M. H. Mobarok, and J. M. Buriak, "Solution-processed zinc phosphide (α-zn3p2) colloidal semiconducting nanocrystals for thin film photovoltaic applications," *ACS nano*, vol. 7, no. 9, pp. 8136–8146, 2013.

[17] V. Raghunathan, A. Kansal, J. Hsu, J. Friedman, and M. Srivastava, "Design considerations for solar energy harvesting wireless embedded systems," in *Proceedings of the 4th international symposium on Information processing in sensor networks*. IEEE Press, 2005, p. 64.

[18] Y. Ramadass and A. Chandrakasan, "A battery-less thermoelectric energy harvesting interface circuit with 35 mv startup voltage," *Solid-State Circuits, IEEE Journal of*, vol. 46, no. 1, pp. 333–341, 2011.

[19] Z. Wang, V. Leonov, P. Fiorini, and C. Van Hoof, "Realization of a wearable miniaturized thermoelectric generator for human body applications," *Sensors and Actuators A: Physical*, vol. 156, no. 1, pp. 95–102, 2009.

[20] ti.com, "Low-power microcontroller," http://www.ti.com/lsds/ti/microcontrollers_16-bit_32-bit/msp/ultra-low_power/overview.page, 2015, [Online; accessed October 2015].

[21] C. M. Vigorito, D. Ganesan, and A. G. Barto, "Adaptive control of duty cycling in energy-harvesting wireless sensor networks," in *Sensor, Mesh and Ad Hoc Communications and Networks, 2007. SECON'07. 4th Annual IEEE Communications Society Conference on*. IEEE, 2007, pp. 21–30.

[22] L. Yerva, B. Campbell, A. Bansal, T. Schmid, and P. Dutta, "Grafting energy-harvesting leaves onto the sensornet tree," in *Proceedings of the 11th international conference on Information Processing in Sensor Networks*. ACM, 2012, pp. 197–208.

[23] W. Li, F. C. Delicato, and A. Y. Zomaya, "Adaptive energy-efficient scheduling for hierarchical wireless sensor networks," *ACM Trans. Sen. Netw.*, vol. 9, no. 3, p. 33:11733:34, Jun. 2013. [Online]. Available: http://doi.acm.org/10.1145/2480730.2480736

[24] N. Sadagopan and B. Krishnamachari, "Maximizing data extraction in energy-limited sensor networks," *International Journal of Distributed Sensor Networks*, vol. 1, no. 1, pp. 123–147, 2005.

[25] J. Marašević, C. Stein, and G. Zussman, "Max-min fair rate allocation and routing in energy harvesting networks: Algorithmic analysis," in *Proceedings of the 15th ACM international symposium on Mobile ad hoc networking and computing*. ACM, 2014, pp. 367–376.

[26] ti.com, "A fully compliant zigbee 2012 solution: Z-stack," http://www.ti.com/tool/z-stack, 2012, [Online; accessed January 2014].

[27] pv.nrcan.gc.ca, "Photovoltaic potential and solar resource maps of canada," http://pv.nrcan.gc.ca/index.php?n=794&m=u&lang=e, 2013, [Online; accessed January 2014].

[28] silabs.com, "Si106x-8x ultra-low power mcu with integrated high-performance sub-1 ghz transceiver," http://www.silabs.com/Support%20Documents/TechnicalDocs/Si106x-8x.pdf, 2014, [Online; accessed July 2014].

[29] A. Kansal, J. Hsu, S. Zahedi, and M. B. Srivastava, "Power management in energy harvesting sensor networks," *ACM Transactions on Embedded Computing Systems (TECS)*, vol. 6, no. 4, p. 32, 2007.